

**Interaktionsanalysen zur spezifischen Integration des  
Retrotransposons TRE5-A.1 in *Dictyostelium discoideum***

Dissertation  
zur Erlangung des Doktorgrades  
der Naturwissenschaften

vorgelegt beim Fachbereich  
Chemische und Pharmazeutische Wissenschaften  
der Johann Wolfgang Goethe Universität  
in Frankfurt am Main

von  
Thanh Chung  
aus Quang Dong, Vietnam

Frankfurt am Main, 2007  
(D30)

Vom Fachbereich  
Chemische und Pharmazeutische Wissenschaften  
der Johann Wolfgang Goethe-Universität in Frankfurt am Main  
als Dissertation angenommen.

Dekan : Prof. Dr. H. Schwalbe

1. Gutachter : Prof. Dr. T. Winckler

2. Gutachter : Prof. Dr. T. Dingermann

Datum der Disputation:

*Für meine Eltern*

DON'T **EVER** GIVE UP

# INHALTSVERZEICHNIS

Inhaltsverzeichnis .....	I
Abbildungsverzeichnis .....	V
Tabellenverzeichnis.....	VII
Abkürzungsverzeichnis .....	VII
<b>1 Einleitung .....</b>	<b>1</b>
1.1 Transposable Elemente .....	1
1.1.1 Allgemeines.....	1
1.1.2 DNA-Transposons .....	2
1.1.3 Retrotransposons .....	3
1.1.3.1 LTR-Retrotransposons.....	3
1.1.3.2 Non-LTR-Retrotransposons .....	7
1.1.4 Insertionspräferenzen .....	12
1.2 Expression von tRNA-Genen in Eukaryonten.....	13
1.2.1 tRNA-Gene.....	13
1.2.2 Pol III-Expression.....	13
1.2.2.1 TBP .....	15
1.2.2.2 Brf1 .....	17
1.2.2.3 Bdp1.....	18
1.3 Dictyostelium discoideum .....	19
1.3.1 Anmerkungen zur Systematik und Entwicklung .....	19
1.3.2 Lebenszyklus .....	20
1.3.3 Das Genom – Aufbau und Struktur .....	22
1.3.4 tRNA-Gene in <i>Dictyostelium discoideum</i> .....	23
1.4 Transposable Elemente in <i>Dictyostelium discoideum</i> .....	25
1.4.1 DNA-Transposons .....	26
1.4.2 LTR-Retrotransposons.....	26
1.4.2.1 DIRS-1 .....	26
1.4.2.2 Skipper.....	26
1.4.2.3 DGLT-A und B .....	27
1.4.3 tRNA-spezifische Non-LTR-Retrotransposons (TREs) .....	27
1.4.3.1 TRE5-Gruppe .....	27
1.4.3.2 TRE3-Gruppe .....	32
1.5 Transposons in der Gentherapie .....	32

1.6	Zielsetzung-----	34
<b>2</b>	<b>Material und Methoden -----</b>	<b>35</b>
2.1	Material -----	35
2.1.1	Chemikalien, Geräte und ihre Bezugsquellen-----	35
2.1.2	Allgemeine Puffer und Lösungen-----	37
2.1.3	Nährmedien-----	38
2.1.4	Antikörper -----	39
2.1.5	Verwendete Organismen -----	39
2.1.6	Vektoren und Plasmide -----	40
2.1.7	Synthetische Oligonukleotide -----	44
2.1.7.1	Oligonukleotide, die im Vektor binden -----	44
2.1.7.2	Oligonukleotide, die in TRE-Elementen binden-----	45
2.1.7.3	Oligonukleotide, die in DdTFIIIB binden -----	45
2.1.7.4	Oligonukleotide, die in HsTBP binden -----	46
2.2	Molekularbiologische Methoden -----	47
2.2.1	Standardmethoden-----	47
2.2.2	Sequenzanalysen -----	47
2.3	Bakterielle <i>Zweihybrid</i> -Versuche -----	48
2.3.1	Herstellung der Konstrukte-----	50
2.3.2	Doppeltransformation und Selektion-----	50
2.4	Interaktionsanalyse <i>in vitro</i> -----	51
2.4.1	Expression des rekombinanten Fusionsproteins MBP-ORF1 -----	51
2.4.2	<i>In vitro</i> Transkription und Translation-----	52
2.4.3	MBP- <i>Pulldown</i> -Assay -----	52
2.5	Herstellung eines monoklonalen Antikörpers gegen <i>Dictyostelium</i> -TBP -----	53
2.5.1	Expression des Antigens GST-TBP im präparativen Maßstab-----	53
2.5.2	Aufreinigung des rekombinanten Proteins über Ni <sup>2+</sup> -NTA-Säule ----	54
2.5.3	Elektroelution -----	55
2.5.4	Immunisierung der Mäuse -----	56
2.5.5	ELISA-Screening -----	56
2.5.6	Western-Blot-Analyse -----	57
2.6	Arbeiten mit <i>Dictyostelium discoideum</i> -----	59
2.6.1	Allgemeine Kulturbedingungen -----	59

2.6.2	Herstellung von Zellpellets	59
<b>3</b>	<b>Ergebnisse</b>	<b>60</b>
3.1	Arbeitshypothese zur Erkennung der Integrationsorte durch TRE5-A.1	60
3.2	TFIIIB in <i>Dictyostelium discoideum</i>	62
3.2.1	DdTBP	62
3.2.2	DdBrf1	63
3.2.3	DdBdp1	64
3.3	TRE5-A.1 Proteine	66
3.3.1	ORF1-Protein	66
3.3.2	ORF2-Protein	68
3.4	Interaktionsanalyse im bakteriellen Zweihybrid-System	69
3.4.1	Proteininteraktionen zwischen TFIIIB-Untereinheiten	73
3.4.2	Dimerisierung der TRE5A.1-Proteine ORF1p und ENp	74
3.4.3	Interaktionen zwischen TRE5-A.1-ORF1p und DdTFIIIB	75
3.4.4	Keine Interaktion zwischen TRE5-A.1-ORF1p und humanem TBP	76
3.4.5	Kartierung der Interaktionsflächen von DdTBP	76
3.4.6	Kartierung der Interaktionsflächen von TRE5-A.1-ORF1p	81
3.4.7	Interaktion von TRE5-B-ORF1p mit DdTFIIIB	82
3.5	Interaktionen von ORF1p mit TBP und Brf 1 <i>in vitro</i>	84
3.5.1	Rekombinante Expression von ORF1p	84
3.5.2	MBP-Pulldown-Versuche	84
3.6	Herstellung eines monoklonalen Antikörpers gegen TBP	86
3.6.1	Expression und Aufreinigung des Antigens	86
3.6.2	Isolierung des Antikörpers	87
3.6.3	Charakterisierung eines potentiellen „Epitop-Tags“	89
<b>4</b>	<b>Diskussion</b>	<b>92</b>
4.1	Das bakterielle Zweihybrid-System	92
4.2	Definition des Integrationsortes von TRE5-A.1 durch ORF1p-DdTFIIIB-Interaktionen	95
4.3	ORF1p ist Teil des Präintegrationskomplexes von TRE5-A.1	99
4.4	RTp und ENp	100
4.5	Konvergente Entwicklung eines zielgerichteten Integrationsmechanismus	101
4.6	Ausblick	104

<b>5</b>	<b>Zusammenfassung</b>	<b>106</b>
<b>6</b>	<b>Literatur</b>	<b>107</b>
	<b>Lebenslauf</b>	<b>115</b>
	<b>Publikationen</b>	<b>116</b>
	<b>Danksagung</b>	<b>117</b>

## ABBILDUNGSVERZEICHNIS

Abb. 1.1: Integrationsmechanismen von DNA- und Retrotransposons.....	2
Abb. 1.2: Aufbau der LTR-Retrotransposons Ty1 und Ty3.....	4
Abb. 1.3: Struktur eines Non-LTR-Retrotransposons .....	8
Abb. 1.4: Modell des TPRT-Mechanismus .....	10
Abb. 1.5: Struktur eines Alu-Elementes .....	11
Abb. 1.6: Transkription eines tRNA-Gens.....	14
Abb. 1.7: Dreidimensionale Struktur von HsTBP .....	16
Abb. 1.8: Struktureller Aufbau TFII-B verwandter Transkriptionsfaktoren.....	18
Abb. 1.9: Strukturen von ScBdp1 und HsBdp1 .....	19
Abb. 1.10: Phylogenetische Einordnung von <i>Dictyostelium</i> .....	20
Abb. 1.11: Schematische Darstellung des Entwicklungszyklus von <i>D. discoideum</i> ..	22
Abb. 1.12: Konsensussequenz der B-Boxen in <i>D. discoideum</i> .....	24
Abb. 1.13: Transposons in <i>Dictyostelium discoideum</i> .....	25
Abb. 1.14: Aufbau der TRE5-A-Elemente.....	30
Abb. 1.15: TRE5-B .....	31
Abb. 1.16: TRE5-C .....	31
Abb. 1.17: TRE3-Element.....	32
Abb. 3.1: Mögliche Zusammensetzung des Pol III-Initiationskomplexes .....	61
Abb. 3.2: Sequenz des <i>tbpA</i> -Gens .....	62
Abb. 3.3: DdTBP – Alignment.....	63
Abb. 3.4: DdBrf1-Alignment .....	64
Abb. 3.5: Identifizierung von DdBdp1 .....	65
Abb. 3.6: Schematische Darstellung der klonierten TRE5-A.1-kodierten Proteine ...	66
Abb. 3.7: Alignment der ORF1-Sequenzen .....	67
Abb. 3.8: ORF1p-Sequenz aus DDB0218337 .....	68
Abb. 3.9: Das <i>BacterioMatch<sup>TM</sup> Two Hybrid</i> System.....	69
Abb. 3.10: DNA-Sequenz des in pBT klonierten <i>tbpA</i> -Gens.....	70
Abb. 3.11: Expression der Fusionsproteine im B2H II-Reporterstamm .....	72
Abb. 3.12: Interaktion zwischen DdTBP und DdBrf1 im B2H I-System .....	73
Abb. 3.13: Interaktionen zwischen DdTFIIIB und ORF1p und ORF2p.....	75
Abb. 3.14: Struktur von DdTBP .....	77
Abb. 3.15: Modelle der DdTBP-Mutanten .....	78
Abb. 3.16: Helix H2' in DdTBP und HsTBP .....	79



Abb. 3.17: Interaktionsanalysen der DdTBP-Mutanten.....	80
Abb. 3.18: <i>Two Hybrid</i> Versuche mit der Chimäre HsTBP(dH2').....	81
Abb. 3.19: Bestimmung der Interaktionsflächen von ORF1p.....	82
Abb. 3.20: Clustal X-Alignment der ORF1-Proteine aus TRE5-A und TRE5-B.....	83
Abb. 3.21: <i>Two Hybrid</i> -Versuche mit TRE5-B ORF1p.....	83
Abb. 3.22: MBP- <i>Pulldown</i> Experimente.....	85
Abb. 3.23: Expression des Antigens GST-TBP in W3110 .....	86
Abb. 3.24: Aufgereinigtes GST-TBP .....	87
Abb. 3.25: Western Blot Anti-TBP 7A3 .....	88
Abb. 3.26: Identifikation von DdTBP mit Anti-TBP 7A3 .....	89
Abb. 3.27: Sequenz von GST-TBP.....	89
Abb. 3.28: Western Blot Anti-TBP(ep) 6B3.....	90
Abb. 3.29: Bestimmung des antigenen Epitops .....	91
Abb. 4.1: Schematische Darstellung verschiedener ORF1-Proteine .....	97

## TABELLENVERZEICHNIS

Tab. 3.1: Zusammenfassung der <i>Two Hybrid</i> -Versuche .....	72
Tab. 3.2: Vergleich des Methionin-Gehalts.....	85
Tab. 3.3: Spezifität und Sensitivität verschiedener Antikörper gegen TBP .....	88

## ABKÜRZUNGSVERZEICHNIS

3-AT	3-Amino-1,2,4-Triazol
A	Adenin
Abb.	Abbildung
ABTS	2,2'-Azino-bis(3-ethylbenzthiazolin-6-sulfonsäure)- diammonium Salz
Amp	Ampicillin
APOD	Ammoniumperoxodisulfat
AS	Aminosäure
ATP	Adenosintriphosphat
Bdp1	<i>B double prime</i> (Untereinheit von TFIIIB)
β-ME	β-Mercaptoethanol
bp	Basenpaar(e)
Brf	<i>TFII related factor</i> (Untereinheit von TFIIIB)
BSA	<i>Bovine Serum Albumin</i>
°C	Grad Celsius
C	Cytosin
Cam	Chloramphenicol
CbfA	C-Modul bindender Faktor
DDT	<i>Dictyostelium</i> DNA Transposon
DGLT	<i>Dictyostelium</i> gypsy like transposon
DIRS	<i>Dictyostelium</i> intermediate repeat sequence
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
DNase	Desoxyribonuklease
dNTPs	Desoxyribonukleosidtriphosphate (A, T, G, C)
ds	<i>double stranded</i> (doppelsträngige DNA)
DTT	Dithiothreitol
ECL	<i>Enhanced Chemiluminescence</i>
EDTA	Ethylendiamintetraessigsäure
ELISA	<i>Enzyme Linked Immuno Sorbent Assay</i>
EN	Endonuklease
g	Erdbeschleunigung oder Gramm
G	Guanin
h	Stunde
HC	Histidin-Cystein-Motiv (Zinkfinger-Sequenz)
HisDO	<i>Histidin Dropout Amino acid supplement</i>

IN	Integrase
IPTG	Isopropyl- $\beta$ -D-thiogalaktopyranosid
ITR	<i>inverted terminal repeat</i>
kb	Kilobase(n)
kDa	Kilodalton
LINE	<i>long interspersed nuclear element</i>
LTR	<i>long terminal repeat</i>
m/v	Masse pro Volumen
mA	Milliampere
mbar	Millibar
MCS	<i>Multiple Cloning Site</i>
$\mu$ g	Mikrogramm
$\mu$ M	Mikromolar
min	Minuten
ml	Milliliter
mM	millimolar
mm	Millimeter
MW	Molekulargewicht
MQ Wasser	Milli-Q Wasser
ng	Nanogramm
Ni <sup>2+</sup> -NTA	Nickel-Nitrilo-tri-essigsäure
nm	Nanometer
OD <sub>600</sub>	optische Dichte, gemessen bei 600 nm
ORF	<i>open reading frame</i> (offener Leserahmen)
PAA	Polyacrylamid
PAGE	Polyacrylamid-Gelelektrophorese
PBS	<i>Phosphate buffered saline</i>
PCR	Polymerase chain reaction (Polymerase Kettenreaktion)
PEG	Polyethylenglykol
PETG	Phenylethyl- $\beta$ , D-Thiogalactosid
PIC	Präintiationskomplex
Pol	RNA-Polymerase
Pos.	Position
q. s.	<i>quantum satis</i>
RNA	Ribonukleinsäure
RNAP-alpha	$\alpha$ -Untereinheit der RNA-Polymerase
RNase	Ribonuklease
RNP	Ribonukleinpartikel
rRNA	ribosomale RNA
RT	Raumtemperatur oder Reverse Transkriptase
SDS	Natriumdodecylsulfat
sec	Sekunden
SINE	<i>short interspersed nuclear element</i>
snRNA	<i>small nuclear RNA</i>

Strep	Streptomycin
SV	Säulenvolumen
T	Thymin
TBP	TATA-Box bindendes Protein (Untereinheit von TFIIB)
Tdd	Transposon in <i>Dictyostelium discoideum</i>
TE	transposables Element
Tet	Tetracyclin
TF	Transkriptionsfaktor
TEMED	N, N, N', N'-Tetramethylethyldiamin
TPRT	<i>target primed reverse transcription</i>
TR	Transposase
TRE	tRNA-Gen assoziiertes Retroelement
Tris	Tris(hydroxymethyl)-aminomethan
TSD	<i>target site duplication</i>
Ty	<i>transposon yeast</i>
U	Unit (Einheit) oder Uracil
UTR	<i>untranslated region</i> (nicht translatierter Bereich)
ÜS	Überstand
UpM	Umdrehungen pro Minute
UV	Ultraviolett
V	Volt
VLP	<i>virus like particle</i> (virusähnlicher Partikel)
v/v	Volumen pro Volumen
vs.	<i>versus</i>
X-Gal	5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galaktopyranosid

# 1 Einleitung

## 1.1 Transposable Elemente

### 1.1.1 Allgemeines

Transposable Elemente (TEs) sind DNA-Abschnitte im Genom, die in der Lage sind, innerhalb einer Zelle ihre genetische Position zu verändern. Synonym werden die Begriffe Transposons oder mobile genetische Elemente verwendet [CRAIG *et al.*, 2002].

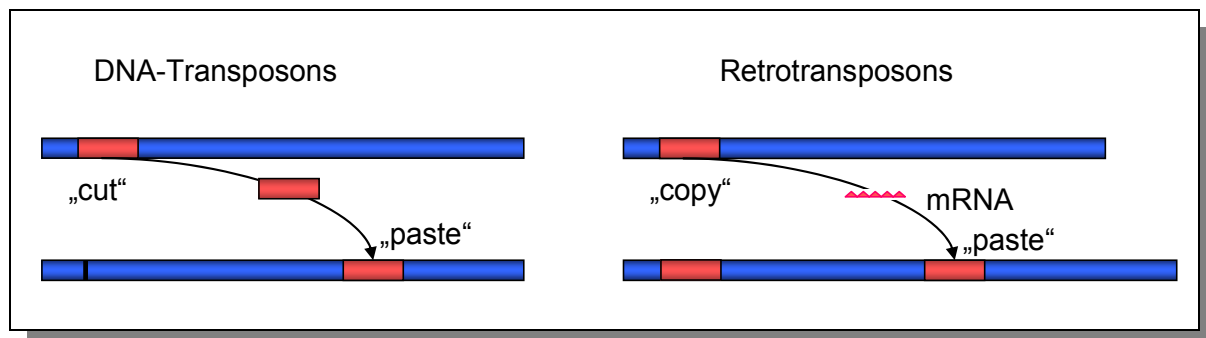
Bereits in den 1940er Jahren wurden mobile genetische Elemente durch genetische Experimente in *Zea mays* entdeckt [MCCLINTOCK 1948]. Dies ist umso bemerkenswerter, als dass man zu dieser Zeit weder genaue Vorstellungen über die molekulare Natur der Erbinformation hatte noch die heutigen molekularbiologischen Techniken zur Charakterisierung mobiler DNA zur Verfügung standen.

Mittlerweile sind Transposons in allen untersuchten Organismen gefunden worden, sowohl in Eukaryonten als auch in Prokaryonten. Sie können einen beträchtlichen Teil des Genoms ausmachen und dieses somit auch maßgeblich beeinflussen. Im Menschen sind fast 50 schwere Erkrankungen bekannt, deren Entstehung auf die Mobilisierung von Retrotransposons zurückgeführt werden kann [KAZAZIAN 1998]. Transposons können aber auch das Genom vergrößern und flexibler und komplexer gestalten, genomische DNA rearrangieren, Wirtsgene mutieren und die Expressionsstärke nahe gelegener Gene modulieren [KAZAZIAN 2004]. Mehrere entscheidende Unterschiede zwischen den Genomen der Affen und des Menschen beruhen auf Alu-vermittelten Rekombinationen [MOURIER 2005].

Je nach Transpositionsmechanismus teilt man die TEs in zwei große Klassen ein: Retrotransposons (Klasse I-Elemente) und DNA-Transposons (Klasse II-Elemente).

Die strukturell sehr heterogene Klasse der DNA-Transposons integriert nach einem „*cut-and-paste*“-Mechanismus. Das Transposon wird mit Hilfe einer Transposase als Fragment aus seinem Ursprungsort herausgeschnitten und an anderer Stelle integriert. Die Anzahl mobiler genetischer Elemente bleibt dabei konstant. Im Gegensatz zu DNA-Transposons vermehren sich Retrotransposons bei jedem Transpositionereignis, da sie nach einem „*copy-and-paste*“-Mechanismus integrieren [DEININGER & BATZER 2002]. Von dem ursprünglichen Element wird eine

RNA-Kopie hergestellt und nach reverser Transkription an einer neuen Stelle im Genom integriert (Abb. 1.1).



**Abb. 1.1: Integrationsmechanismen von DNA- und Retrotransposons**

Links: DNA-Transposons werden von einer Stelle des Genoms entfernt („cut“) und an anderer Stelle wieder inseriert („paste“). Die Kopienzahl bleibt unverändert. Rechts: Retrotransposons verbleiben an ihrem Ursprungsort. Sie werden in ein RNA-Intermediat kopiert („copy“), das dann durch Reverse Transkription in DNA überführt und an einer neuen Stelle des Genoms integriert („paste“) wird. Die Kopienzahl erhöht sich bei jedem Transpositionereignis.

Man bezeichnet TEs als autonom, wenn sie die für die Transposition notwendigen Proteine selbst kodieren und nicht (oder nur in sehr begrenztem Rahmen) auf fremde Hilfsproteine angewiesen sind. Die Mehrheit der DNA-Transposons scheint autonom zu sein, bei den Retrotransposons dagegen gibt es eine große Gruppe nicht-autonomer Elemente.

Bisher wurden Retrotransposons im engeren Sinn nur in Eukaryonten gefunden, während DNA-Transposons auch in prokaryontischen Genomen verbreitet sind [MALIK & EICKBUSH 2001].

### 1.1.2 DNA-Transposons

DNA-Transposons sind etwa 2,5 kb groß und besitzen an ihren Enden entgegengesetzt orientierte terminale Wiederholungseinheiten (*inverted terminal repeats*, ITRs) als Bindemotiv für die Transposase. Die 5'-ITR besitzt Promotoraktivität und kontrolliert die Transkription des offenen Leserahmens (*open reading frame*, ORF), der für eine Transposase kodiert. Während des Transpositionsprozesses erzeugt das Enzym durch einen versetzten Schnitt der Ziel-DNA Zielsequenzverdoppelungen (*target site duplication*, TSD) von 2-14 bp. TSDs verbleiben als Fußabdruck im Genom, auch wenn das Transposon den genetischen Ort verlassen hat. Die Länge einer TSD ist häufig charakteristisch für ein bestimmtes Transposon [CRAIG *et al.*, 2002; KAZAZIAN 2004]. DNA-Transposons integrieren sequenzspezifisch. Die von ihnen erkannte Zielsequenz ist allerdings sehr

kurz, so dass das Genom an zahlreichen Stellen von DNA-Transposons besiedelt ist. Viele DNA-Transposons inserieren bevorzugt in der unmittelbaren Nähe des Ursprungsortes (*local hopping*) [KAZAZIAN 2004].

### 1.1.3 Retrotransposons

Retrotransposons transponieren über ein RNA-Intermediat. Eine Reverse Transkriptase erzeugt aus dem Intermediat eine DNA-Kopie, die dann in den Zielort integriert wird. Durch „*copy-and-paste*“ erhöht sich bei jedem Transpositionereignis die Anzahl der Elemente in einer Zelle.

Nach diesem Mechanismus integrieren Retrotransposons, echte und endogene Retroviren, Retrons und Gruppe-II-Introns. Die Reversen Transkriptasen aller Elemente zeigen einen hohen Konservierungsgrad, so dass man von einem gemeinsamen Ursprung der Enzyme ausgeht [MALIK *et al.*, 1999].

Die meisten Retrotransposons lassen sich nach ihrem Transpositionsmechanismus und ihrer Struktur in LTR- (*long terminal repeat*) und Non-LTR-Retrotransposons einteilen. Non-LTR-Retrotransposons wiederum werden in autonome (LINEs) und nichtautonome (SINEs) Elemente aufgeteilt [CAPY 2005]. Nicht-autonome Elemente imitieren die 3'-Enden autonomer Retrotransposons und werden so von der Reversen Transkriptase und der Integrase autonomer Elemente erkannt und mobilisiert [OKADA *et al.*, 1997].

#### 1.1.3.1 LTR-Retrotransposons

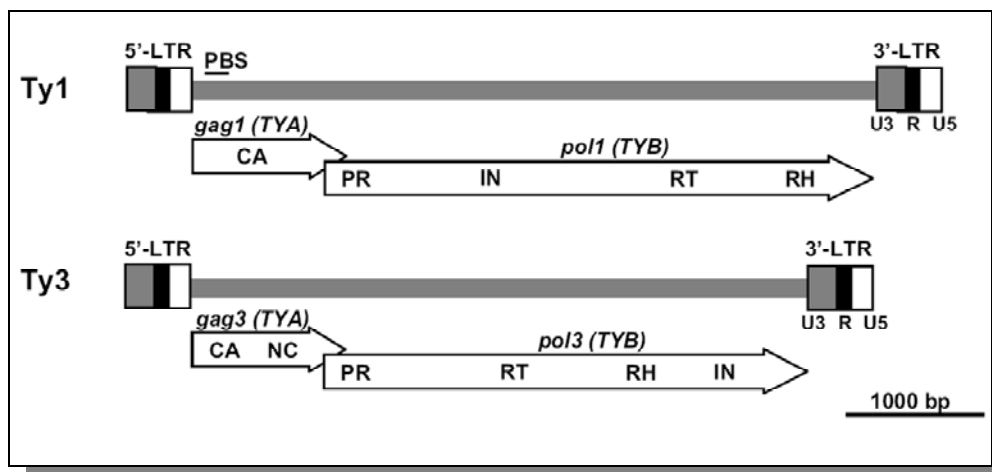
LTR-Retrotransposons kommen ausschließlich in Eukaryonten vor. Es handelt sich um eine bezüglich Struktur und Transpositionsmechanismus sehr einheitliche Gruppe, deren Eigenschaften dank der engen Verwandtschaft zu den gut untersuchten Retroviren ebenfalls weitgehend bekannt sind [BOEKE & CORCES 1989; KATZ & SKALKA 1994]. Die reverse Transkription dieser Elemente findet in *virus like particles* im Zytoplasma statt [CRAIG *et al.*, 2002].

Die 4-10 kb großen Elemente werden beiderseits von 50-100 bp langen LTRs flankiert. Sie haben meistens die gleiche Orientierung (*direct repeats*, DR), nur selten findet man invers orientierte Wiederholungseinheiten (*inverted repeats*, IR) (Abb. 1.2).

Sie werden in die drei Regionen U3, R und U5 unterteilt. In den LTRs sind die notwendigen Motive sowohl für Promotor- als auch Terminator-Aktivitäten zu finden. Die Transkription beginnt im 5'-LTR an der U3/R-Grenzregion und endet im 3'-LTR

an der R/U5-Grenzregion. Somit sind die U3- bzw. U5-Regionen (*unique to the 3'/5' end*) jeweils nur einmal im Transkript vertreten, die R-Region (*repeated region*) dagegen zweimal.

Zwischen den beiden LTRs befinden sich meist zwei zueinander versetzte, sich um 10-50 bp überlappende offene Leserahmen (*open reading frame*, ORF). Der erste ORF wird mit *gag* bezeichnet. Er kodiert für strukturelle Capsid-(CA) und Nukleocapsid (NA)-Proteine, die in Analogie zu den Retroviren für die Ausbildung der virusähnlichen Partikel (*virus like particles*, VLP) zuständig sind. Der zweite ORF, der mit *pol* bezeichnet wird, ist die am besten konservierte Sequenz der LTR-Retrotransposons und weist alle für den Replikationszyklus essentiellen Motive auf. Er kodiert für die Enzyme Protease (PR), Reverse Transkriptase (RT), Ribonuklease H (RH) und Integrase (IN). Je nach Anordnung dieser Domänen auf dem zweiten ORF können die meisten LTR-Retrotransposons der *Ty3/gypsy* oder der *Ty3/copia*-Familie zugeordnet werden. Den LTR-Retrotransposons fehlt das Gen für das ENV-Protein, welches den Retroviren ihre Infektiösität verleiht. Deshalb können LTR-Retrotransposons nur in das Genom re-inserieren, aus dem sie stammen.



**Abb. 1.2: Aufbau der LTR-Retrotransposons Ty1 und Ty3**

Dargestellt sind die Elemente Ty1 (5,9 kb) und Ty3 (5,4 kb) aus *S.cerevisiae*. CA: Capsid-Domäne (Ausbildung der VLPs), HC: Histidin-Cystein-reiche Domäne (Ausbildung von Zinkfingerstrukturen), IN: Integrase-Domäne (Integration der Retrotransposon-DNA), NC: Nukleocapsid-Domäne (Ausbildung der VLPs), PBS: Primer-tRNA-binding Site (Bindestelle für die Primer-tRNA zur Initiation der reversen Transkription), PR: Protease-Domäne (Spaltung des Poly-Proteins in die einzelnen funktionellen Proteine), RT: Reverse Transkriptase-Domäne (reverse Transkription der Retrotransposon-RNA in den VLPs). Die LTRs bestehen aus den drei Regionen U3, R und U5.



Bei LTR-Retrotransposons, die zwei ORFs (*gag* und *pol*) besitzen, wird nur eine gemeinsame bicistronische mRNA transkribiert. Das erzeugte Transkript dient einerseits in den VLPs als Matrize für die reverse Transkription, andererseits wird es an den Ribosomen zur Translation der *gag*- und *pol*-Proteine benötigt.

Die Translation der zumeist um +1 zueinander versetzten ORFs erfolgt gemeinsam durch einen ribosomalen *frameshift*, so dass zunächst Fusionsproteine entstehen, die später prozessiert werden müssen. Da ein ribosomaler *frameshift* nur in ca. 5-10% aller Fälle eintritt, gewährleistet die Nutzung dieses Mechanismus, dass das Strukturprotein GAG, das für die Ausbildung der VLPs in größeren Mengen benötigt wird, im Vergleich zu den enzymatischen Proteinen des POL-Proteins im Überschuss gebildet wird [WILHELM *et al.*, 2005]. Nach spezifischer Spaltung des Polyproteins durch eine Protease lagern sich die einzelnen Proteine im Zytoplasma zu den VLPs zusammen. Diese enthalten einige Retrotransposontranskripte, tRNAs sowie Desoxynukleotide für die Initiation und Durchführung der reversen Transkription [HAAG *et al.*, 2000]. In den VLPs findet die vollständige Synthese eines DNA-Doppelstrangs des Retrotransposons statt.

Durch Auflösen der VLPs wird die doppelsträngige Retrotransposon-DNA freigesetzt. Sie bildet vermutlich zumindest mit der Integrase einen Präintegrationskomplex, der allerdings klein genug sein muss um in den Zellkern zu gelangen. Bei einigen LTR-Transposons wie den Ty-Elementen aus *S. cerevisiae* konnte eine Kernlokalisationssequenz (*nuclear localization signal*, NLS) identifiziert werden, so dass sie wahrscheinlich aktiv in den Zellkern transportiert werden können. Ohne eine NLS können die Elemente nur während der Mitose, also wenn die Zellmembran aufgelöst ist, an die genomische DNA gelangen. Dort leitet die Integrase einen Doppelstrangbruch in der Zielsequenz ein und integriert die Retrotransposon-DNA [KENNA *et al.*, 1998]. DNA-Reparaturmechanismen füllen den versetzten Schnitt wieder auf und erzeugen so die TSDs.

Im Genom von *S. cerevisiae* gibt es fünf LTR-Retrotransposon-Familien (Ty1-5). Alle Mitglieder dieser Familien weisen die typische Struktur eines LTR-Retrotransposons auf, Ty3 als *gypsy-like element* zeigt im zweiten ORF eine etwas abweichende Anordnung der Domänen. Etwa 3,1% des 12 Mb großen Hefe-Genoms bestehen aus Ty-Sequenzen. Alle Ty-Elemente zeigen Insertionspräferenzen, wenn auch unterschiedliche. Bemerkenswert ist aber die Tatsache, dass sie stets Regionen

besiedeln, die (i) arm an kodierenden Sequenzen oder (ii) transkriptionell inaktiv sind [LESAGE & TODESCHINI 2005].

### Ty1

Ty1 integriert bevorzugt in Regionen 75-750 bp oberhalb von Genen, die von der RNA-Polymerase III (Pol III) transkribiert werden [KIM *et al.*, 1998b; BACHMAN *et al.*, 2004]. Die TFIIIB-Untereinheit Bdp1 und der Chromatin modellierende Faktor Isw2 stellen notwendige Proteine für die periodische Integration von Ty1 dar. Nach Rekrutierung durch Bdp1 arrangiert Isw2 die Nukleosomen im Abstand von 80 bp. In den gleichen Abständen wurden Ty1-Insertionen beobachtet [BACHMAN *et al.*, 2004; BACHMAN 2005]. Fehlen Chromatin-modellierende Faktoren wie die Histondeacetylase Hos3 oder das Trithoraxprotein Set3, finden Transpositionen deutlich seltener statt. Das weist darauf hin, dass die Chromatin-Struktur des transkriptionell aktiven tRNA-Gens eine wichtige Rolle für die Integration von Ty1 spielt [MOU *et al.*, 2006].

Ty2 und Ty4 transponieren ebenfalls in Bereiche, die bis zu 1 kb oberhalb von tRNA-Genen liegen.

### Ty3

Das 5,4 kb große Ty3-Element integriert positionsspezifisch 1-4 bp oberhalb der Transkriptionsinitiationsstelle von Pol III-Genen [CHALKER & SANDMEYER 1992].

Die Bildung des Präintegrationskomplexes von Ty3 ist sehr eng mit der Rekrutierung der Pol III verknüpft. Versuche *in vitro* und *in vivo* bestätigen die wichtige Funktion der Transkriptionsfaktoren TFIIIC und TFIIIB während des Integrationsprozesses, zeigen aber auch, dass die Pol III selbst nicht notwendig ist, sondern dass diese vielmehr mit der Ty3-kodierten Integrase um Bindestellen konkurriert [CHALKER & SANDMEYER 1992; KIRCHNER *et al.*, 1995; CONNOLLY & SANDMEYER 1997; YIEH *et al.*, 2002].

Der zentrale Faktor für die Integration von Ty3 ist TFIIIB. Während die Untereinheiten TBP und Brf1 für die Integration notwendig sind und Ty3 durch Wechselwirkung mit der IN an seinen Zielort dirigieren, erhöht die dritte Untereinheit Bdp1 lediglich die Effektivität der Integration. TFIIIC beeinflusst das Geschehen in zweierlei Hinsicht: Zum einen trägt er zur Erkennung des Zielortes bei, indem er die Orientierung beeinflusst, mit der TFIIIB an die TATA-Box bindet [YIEH *et al.*, 2000; 2002]. TFIIIC

bindet sequenzspezifisch mit zwei seiner Untereinheiten an die A- und die B-Box und rekrutiert anschließend TFIIIB. Mutationen in der A- und B-Box von tRNA-Genen, die die Transkription behindern, hemmen auch die Integration von Ty3-Elementen. Zum anderen wurde eine schwache Interaktion zwischen dem C-Terminus von TFIIIC95 und einem N-terminalen Bereich der Ty3-kodierten Integrase gefunden. Dies lässt darauf schließen, dass TFIIIC direkt mit dem Präintegrationskomplex von Ty3 interagiert und die Wahl des Integrationsortes lenkt [AYE *et al.*, 2001]. Die Strukturen von Ty1 und Ty3 sind in Abb. 1.2 dargestellt.

## Ty5

Ty5 besiedelt spezifisch Regionen im Heterochromatin der Telomere sowie „silent chromatin“-Bereiche [ZOU & VOYTAS 1997]. Es konnten sechs Aminosäuren im C-Terminus der Integrase (LXSSXP) identifiziert werden, die durch Protein-Protein-Interaktion mit einer C-terminalen Domäne von Sir4p, einer Struktur des „*silent chromatin*“, die Integration vermitteln [XIE *et al.*, 2001; ZHU *et al.*, 2003].

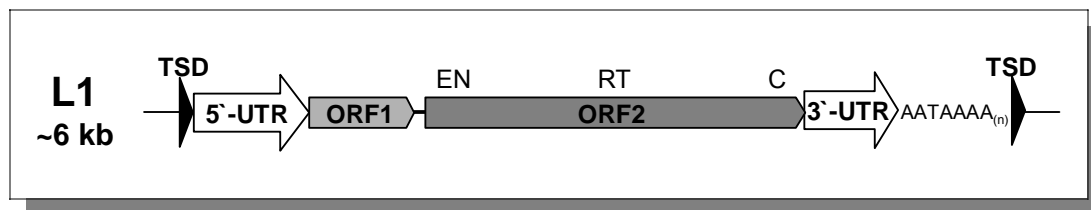
### 1.1.3.2 Non-LTR-Retrotransposons

Von allen mobilen Elementen wurden Non-LTR-Retrotransposons aufgrund ihrer heterogenen Struktur zuletzt entdeckt, obwohl sie in praktisch allen eukaryontischen Genomen in oft hohen Kopienzahlen vorkommen [WEINER 1986].

Non-LTR-Retrotransposons kommen ausschließlich in Eukaryonten vor und stellen die größte Klasse der Transposons dar. Evolutionär stammen sie vermutlich von den prokaryontischen Gruppe II-Introns ab [MALIK *et al.*, 1999]. Aufgrund phylogenetischer Vergleiche ihrer RT-Domänen - der einzigen Domäne, die allen Elementen gemeinsam ist - werden sie in zwölf Klassen eingeteilt [MALIK *et al.*, 2000]. Die Non-LTR-Retrotransposons der Säuger bezeichnet man als LINEs (*long interspersed nuclear elements*). Der bekannteste Vertreter dieser Gruppe ist das ~6 kb große LINE-1 oder L1-Retrotransposon, das ca. 17% des humanen und 19% des murinen Genoms ausmacht [LANDER *et al.*, 2001]. Allerdings sind aufgrund von 5'-Verkürzungen oder Mutationen nur noch sehr wenige Kopien aktiv [BOISSINOT & FURANO 2005].

Ein vollständiges Non-LTR-Retrotransposon ist 4-6 kb groß und hat einen oder meist zwei ORFs, die teilweise überlappend und im Leserahmen versetzt (Drosophila, Maus) oder durch eine intergenische Region (*intergenic region*, IGR) voneinander

getrennt sind (Mensch) [LI *et al.*, 2006]. Für die Retrotransposition von L1 sind beide Proteine *in cis* notwendig [MORAN *et al.*, 1996].



**Abb. 1.3: Struktur eines Non-LTR-Retrotransposons**

Struktur eines Non-LTR-Retrotransposons am Beispiel des humanen L1-Elementes (modifiziert nach Kazazian 2004).

ORF1 kodiert für ein 40-50 kDa großes Nukleinsäure-bindendes Protein (ORF1p), dessen genaue Funktion noch nicht bekannt ist und das keine Sequenzhomologien zu bekannten Proteinen aufweist. Es wurde gezeigt, dass es als Nukleinsäure-Chaperon dienen kann [MARTIN & BUSHMAN 2001; MARTIN *et al.*, 2005] und mit hoher Affinität mit dem C-Terminus an RNA bindet [KOLOSHA & MARTIN 2003]. Umstritten ist noch, ob die Bindung allgemein an RNA erfolgt oder ob ORF1p seine eigene RNA bevorzugt [KOLOSHA & MARTIN 2003; KULPA & MORAN 2005; MATSUMOTO *et al.*, 2006]. Aufgereinigtes ORF1p aus dem humanen L1-Element formt durch Protein-Protein-Interaktionen seines N-terminalen Drittels stabile Homotrimere und bedeckt in dieser Form etwa 50 bp einer RNA. Die Bindung durch ORF1p soll die RNA möglicherweise vor dem Abbau schützen, zusätzlich ist es denkbar, dass ORF1p die Sekundärstruktur der RNA auflockert und dadurch die reverse Transkription erleichtert [MARTIN *et al.*, 2000; BASAME *et al.*, 2006]. Sedimentationsanalysen zeigen, dass es ORF1p/ L1-RNA-Komplexe mit unterschiedlichen Stöchiometrien gibt [KULPA & MORAN 2005].

Aus Mutationsversuchen mit murinem und humanem ORF1p weiß man, dass das Protein für die Formation von Ribonukleoproteinkomplexen (RNPs) essentiell ist, dass es aber noch weitere Funktionen in späteren Stadien der Retrotransposition haben muss. Denkbar sind z. B. eine mittelbare Beteiligung am RNP-Transport, am Import in den Zellkern oder an den initialen Schritten der kombinierten reversen Transkription und Integration (siehe unten) [KULPA & MORAN 2005; MARTIN *et al.*, 2005].

Eine genauere Strukturvorstellung hat man von ORF1p des telomer-spezifischen Non-LTR-Retrotransposons SART1 aus *B. mori*. Dort konnte im N-Terminus eine Kernlokalisierungssequenz identifiziert werden. Der C-Terminus enthält drei Zinkfinger, die (i) für die RNP-Formation essentiell sind und (ii) an der Interaktion von

ORF1p mit sich selbst als auch zwischen ORF1p/ORF2p beteiligt sind. Dies gilt aber nicht für alle Non-LTR-Retrotransposons, da L1H auch ohne ein Zinkfingermotiv multimerisiert. Zumindest der erste der drei Zinkfinger ist relativ gut konserviert [MATSUMOTO *et al.*, 2006].

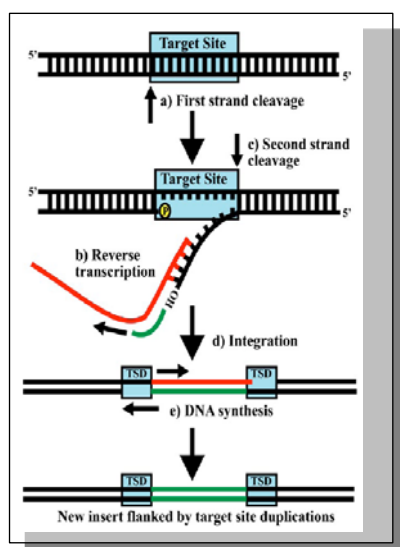
ORF1p telomerspezifischer Non-LTR-Retrotransposons wie SART1 aus *B. mori* sowie TART und HeT-A aus *D. melanogaster* konnten im Nukleus nachgewiesen werden [RASHKOVA *et al.*, 2002; MATSUMOTO *et al.*, 2004], ORF1p aus murinem L1 oder *Drosophila factor* dagegen nicht [MATSUMOTO *et al.*, 2004].

Der zweite ORF kodiert für die enzymatischen Proteine ENp und RTp, die für die Retrotransposition notwendig sind. Zusätzlich können Domänen für eine RNase H und Nukleinsäure-bindende Motive vorhanden sein. Mutationen in den katalytischen Zentren beider Enzyme haben keinen Einfluss auf die Bildung von RNPs, da sie nur während des TPRT (siehe unten) aktiv sind [KULPA & MORAN 2005]. Die RT des humanen L1-Elements bindet bevorzugt ihre eigene RNA am polyA-Schwanz und konnte in RNPs nachgewiesen werden [KULPA & MORAN 2006].

In Non-LTR-Retrotransposons werden die beiden ORFs von unterschiedlich langen nicht-translatierten Regionen (*untranslated regions*, UTRs) flankiert. Dabei besitzt die 5'-UTR häufig Pol-II-Promotoraktivität, während die 3'-UTR neben einem Terminationssignal ein Polyadenylierungssignal sowie einen adenosinreichen Anhang trägt [OSTERTAG & KAZAZIAN 2001a; DEININGER & BATZER 2002]. Promotoren der Non-LTR-Retrotransposons sind relativ schwach und liegen im Gegensatz zu üblichen Pol II-Promotoren innerhalb des transkribierten Bereiches, sind in dieser Hinsicht also den Pol III-Promotoren ähnlich. Möglicherweise bedeutet das, dass Non-LTR-Retrotransposons nach einem hybriden Mechanismus transkribiert werden, in dem sowohl Eigenschaften des Pol II- als auch des Pol III-Aparates vertreten sind [DEININGER & BATZER 2002].

Die Translation der beiden ORFs von einem bicistronischen Transkript erfolgt meist durch interne Initiation unabhängig voneinander. In murinem L1 wurde für jeden ORF eine interne ribosomale Eintrittsstelle (*internal ribosomal entry site*, IRES) identifiziert. Da eine Überexpression von ORF2p für die Zelle toxisch ist, diskutiert man die Möglichkeit, dass die Translation von ORF1p und ORF2p zunächst gleich stark

beginnt, gebildetes ORF1p dann aber die IRES von ORF2 blockiert und somit die Menge an ORF2p limitiert. IRES-Strukturen sind kaum konserviert [LI *et al.*, 2006]. In Mensch und Maus bilden ORF1p und ORF2p zusammen mit der RNA des Retrotransposons den Ribonukleinkomplex (RNP), bei dem aus *Drosophila* stammenden Element *I factor* dagegen ist ORF2p für die Ausbildung des RNP nicht notwendig [SELEME *et al.*, 2005]. Aufgrund der enzymatischen Natur der ORF2-kodierten Proteine findet man deutlich mehr ORF1p als ORF2p in solchen Komplexen [KULPA & MORAN 2006]. Eine Kernlokalisierungssequenz in der EN-Domäne vieler Non-LTR-Retrotransposons ermöglicht den Eintritt der RNPs in den Zellkern. Dort binden sie an die Ziel-DNA, deren (-)-Strang anschließend von der Endonuklease geschnitten wird. Die frei gewordene 3'-OH-Gruppe der Ziel-DNA wird von der RT für die Initiation der reversen Transkription genutzt. Dieser Mechanismus, bei dem chromosomale Insertion mit reverser Transkription kombiniert wird, wird *target primed reverse transcription* (TPRT) genannt, da die Initiation der reversen Transkription durch die Ziel-DNA selbst erfolgt (Abb. 1.4). Die typischen TSDs von bis zu 100 bp entstehen dadurch, dass der (+)-Strang der Ziel-DNA meist deutlich 3'-versetzt zum (-)-Strang geschnitten wird [GILBERT *et al.*, 2005]. Häufig wird die reverse Transkription vorzeitig beendet, so dass 5'-verkürzte, inaktive da promotorlose Non-LTR-Retrotransposons integriert werden [DEWANNIEUX & HEIDMANN 2005]. Der genaue Mechanismus, nach dem das 5'-Ende des transkribierten Retroelementes mit der Ziel-DNA verknüpft wird, ist noch nicht verstanden [GILBERT *et al.*, 2005].

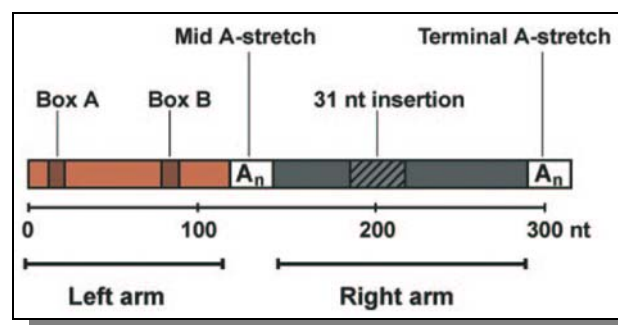


**Abb. 1.4: Modell des TPRT-Mechanismus**

Rot: RNA, Grün: DNA. **a.** Die Endonuklease trennt den Minusstrang der Ziel-DNA auf und erzeugt dabei ein freies 3'-OH-Ende. **b.** Das 3'-Ende der Retrotransposon-RNA lagert sich an an die freie DNA an und die RT nutzt das freie 3'-OH-Ende der DNA als Primer für die reverse Transkription. **c.** Der Schnitt des Plusstrangs der Ziel-DNA erfolgt leicht versetzt zum ersten Schnitt. **d.** Die cDNA integriert nach einem noch nicht aufgeklärten Mechanismus. **e.** Die RNA des Retrotransposons wird entfernt und durch DNA komplementär ersetzt. Durch das komplementäre Auffüllen des versetzten Schnittes der EN entsteht die TSD [OSTERTAG & KAZAZIAN 2001b].

Non-LTR-Retrotransposons lassen sich in autonome und nicht-autonome Elemente einteilen. Im Gegensatz zu den autonomen Retrotransposons besitzen die nicht-autonomen Elemente keine bzw. nicht alle für die Retrotransposition notwendigen kodierenden Sequenzen. Die meisten der nicht-autonomen Retrotransposons gehören zur Familie der SINEs (*short interspersed nuclear elements*), deren bekanntester und erfolgreichster Vertreter das humane Alu-Element ist. Etwa eine Million Kopien besiedeln das menschliche Genom, allerdings ist wie bei den LINEs auch hier nur ein kleiner Bruchteil von 0,5% tatsächlich aktiv [DEININGER *et al.*, 2003].

Das ursprüngliche Alu entstand vermutlich durch eine 5'→3'-Fusion des 7SL-RNA-Gens. Die heute bekannten Alus mit zwei ähnlichen Domänen und einer adenosinreichen Verbindungssequenz gehen aus einer Kopf-an-Schwanz-Fusion zweier fossiler Alu-Monomere hervor. Im linken Arm findet man einen schwachen, aber funktionellen Pol-III-Promotorbereich (A- und B-Box). Am 3'-Ende sitzt ein kurzer Poly-A-Schwanz (Abb. 1.5):



**Abb. 1.5: Struktur eines Alu-Elementes**

Alu-Elemente sind etwa 300 bp lang und bestehen aus zwei ähnlichen Monomeren (linker und rechter Arm), die über einen Adenosin-reichen Stretch verbunden sind. Am 3'-Ende befindet sich ein weiterer, kurzer A-Stretch. Der linke Arm trägt schwache, aber funktionelle Promotoren (A- und B-Box) [HÄSLER. J. 2006].

Alus können von LINE-Elementen mobilisiert werden. Dabei ist es nicht erforderlich, dass es sich um ein vollständiges LINE handelt, so dass Alus also auch von den zahlreichen 5'-verkürzten Kopien *in trans* mobilisiert werden können [DEWANNIEUX *et al.*, 2003]. Das menschliche Genom wird aktiv von Alus gestaltet. Sie modulieren die Genexpression auf posttranskriptionaler Ebene. Sie beeinflussen alternative Spleißprozesse und die Regulation der Translation, induzieren nicht-homologe Rekombinationen und können durch ihre Mobilisierung regulatorische Einheiten und Gene zusammenbringen.

### 1.1.4 Insertionspräferenzen

Einige Retrotransposons haben die Fähigkeit erlangt, nicht zufällig, sondern sequenz-, positions- oder strukturspezifisch in unkritische, d.h. nicht-kodierende oder -regulierende Regionen des Wirtsgenoms zu inserieren. In Gen-dichten Genomen ist daher der Anteil an transposablen Elementen allgemein geringer und die Integrationsspezifität ausgeprägter [BOEKE & DEVINE 1998; EICHINGER *et al.*, 2005].

Viele mobile Elemente integrieren bevorzugt an Stellen, die bereits von Retroelementen besiedelt sind und bilden so Transposoncluster [KAZAZIAN 2004; EICHINGER *et al.*, 2005]. Die Integration in ein anderes TE findet man besonders häufig in den dichten Genomen niederer Eukaryonten [KIM *et al.*, 1998a]. Es wurde gezeigt, dass für den Erkennungs- und/ oder Integrationsprozess häufig Proteine der Wirtszelle benötigt werden. So wird die Positionsspezifität von Ty3 durch Interaktionen mit den ScTFIIIB-Untereinheiten ScTBP und ScBrf1 vermittelt [YIEH *et al.*, 2002] und Ty5 benötigt für seine gerichtete Integration das Protein Sir4 um seine cDNA an die telomerische DNA anzuheften [XIE *et al.*, 2001; ZHU *et al.*, 2003].

Humane L1-Elemente erkennen bestimmte „Zielsequenzen“. Da diese allerdings relativ unspezifisch und im humanen Genom weit verbreitet sind, wird die Wirtszelle nur unzureichend geschützt [KAZAZIAN 1998].

In *D. discoideum* haben Retrotransposons zwei verschiedene Strategien entwickelt um eine Integration in kodierende Sequenzen zu vermeiden. Elemente der TRE- und der DGLT-A-Familien integrieren ausschließlich in der Nähe von tRNA-Genen, während die restlichen Retroelemente bereits vorhandene Cluster besiedeln. Die Integrationsspezifität der TREs wird wahrscheinlich ähnlich wie bei Ty3 aus *S. cerevisiae* durch Proteininteraktionen zwischen Pol III-Transkriptionsfaktoren und transposon-kodierten Proteinen vermittelt [WINCKLER *et al.*, 2005].

tRNA-Gene als Integrationsort haben zwei Vorteile. Zum einen sind sie ubiquitär in vielen Kopien im Genom vertreten. Zum anderen gibt es in der unmittelbaren Umgebung von tRNA-Genen in der Regel keine Pol II-Gene, so dass diese Regionen als sichere Integrationsorte angesehen werden können [BOLTON & BOEKE 2003; WINCKLER *et al.*, 2005].



## 1.2 Expression von tRNA-Genen in Eukaryonten

### 1.2.1 tRNA-Gene

tRNA-Gene kommen ubiquitär im Genom vor und werden von der RNA-Polymerase III (Pol III) transkribiert. Dabei werden die 76-97 bp kleinen Gene zunächst als Vorläufer-tRNAs synthetisiert und im Anschluss zu funktionsfähigen tRNAs prozessiert. Durch Paarung einiger Basen nehmen sie die typische Kleeblattform an, bei der sowohl das 3'- als auch das 5'-Ende im Akzeptorarm liegen. Spezifische Aminoacyl-tRNA-Synthetasen erkennen charakteristische Stellen der tRNA und beladen die 3'-OH-Gruppe mit den zugehörigen Aminosäuren. Die wichtigste Erkennungsstelle ist die Position 73 der tRNA-Gene, viele Synthetasen erfassen aber zusätzlich den Akzeptorarm des tRNA-Gens oder überprüfen das Anticodon [RAMESH 2001; FUKAI 2003].

In *S. cerevisiae* wurde gezeigt, dass die Aktivität von tRNA-Genen die transkriptionelle Aktivität benachbarter Pol II-Genen negativ beeinflusst. Dies könnte eine Erklärung dafür sein, dass selbst in kompakten Genomen proteinkodierende Gene meist einen Abstand von mehreren 100 bp zu tRNA-Genen halten [BOLTON & BOEKE 2003].

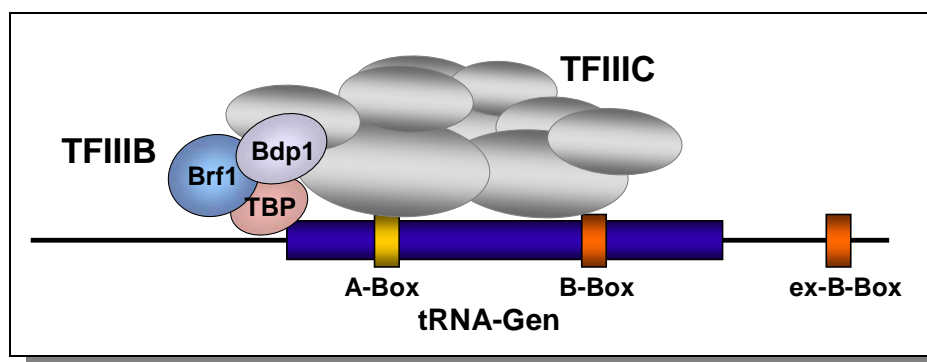
### 1.2.2 Pol III-Expression

Eukaryonten besitzen drei Formen der RNA-Polymerase, die die Transkription nukleärer Gene katalysieren. Trotz ihrer komplexen, aus zahlreichen Untereinheiten bestehenden Strukturen benötigen sie Hilfsproteine oder allgemeine Transkriptionsfaktoren, um die Transkription zu initiieren.

Die RNA-Polymerase III transkribiert im Komplex mit ihren Transkriptionsfaktoren TFIIIA, B und C kleine untranslatierte RNA-Arten, zu denen neben tRNA-Genen auch das ribosomale 5S-Gen und das U6 snRNA-Gen gehören [GEIDUSCHEK & TOCCHINI-VALENTINI 1988]. Aufgrund von Promotoreigenschaften sowie der (Nicht-)Beteiligung der Transkriptionsfaktoren TFIIIA und TFIIIC können diese Gene in drei Typen unterteilt werden [WILLIS 1993; KASSAVETIS 1994; WHITE 1994]. Das ribosomale 5S-Gen (Typ1) und tRNA-Gene (Typ2) besitzen einen geninternen Promotor, der von den Transkriptionsfaktoren erkannt wird. Typ3-Gene wie das U6 snRNA-Gen besitzen dagegen einen Promotor oberhalb der kodierenden Sequenz und *Enhancer*-Elemente wie z.B. eine TATA-Box. Zusammen mit der DNA bilden die

Transkriptionsfaktoren den Präinitiationskomplex und stellen die korrekte Positionierung der Pol III am Transkriptionsstart sicher.

Zu Beginn der Transkription von tRNA-Genen erkennt und bindet TFIIIC die geninternen Promotorsequenzen (A- und B-Box) und rekrutiert gleichzeitig TFIIIB. Dieser bindet durch Proteininteraktionen ca. 30 bp oberhalb des Transkriptionsstarts an die DNA (Abb. 1.6). Eine Rekrutierung durch TFIIIC ist dann nicht notwendig, wenn das Gen eine starke TATA-Box besitzt und somit TFIIIB durch die Bindung von TBP an die TATA-Box an die DNA herangeführt werden kann [GEIDUSCHEK & KASSAVETIS 2001; SCHRAMM & HERNANDEZ 2002].



**Abb. 1.6: Transkription eines tRNA-Gens**

Schematische Darstellung eines tRNA-Gens aus *D. discoideum*. Zu Beginn der Transkription bindet TFIIIC (grau) an die geninterne Promotorsequenz (B-Box) und rekrutiert im Anschluss TFIIIB, der in *D. discoideum* aus den drei Untereinheiten TBP, Brf1 und Bdp1 besteht. TFIIIB führt die Pol III an die DNA heran (nicht dargestellt). Unterhalb des Gens befindet sich häufig eine externe B-Box.

Für die Transkription von tRNA-Genen sowie des U6 snRNA-Gen ist der zinkfingerreiche TFIIIA nicht notwendig, der Faktor TFIIIB dagegen wird für die Transkription aller drei Klassen von Pol III-Genen benötigt [JOAZEIRO *et al.*, 1994]. Die zentrale Rolle von TFIIIB liegt dabei in der Rekrutierung der Pol III an die DNA. Kassavetis *et al.* zeigten bereits 1990, dass zumindest in *S. cerevisiae* sowohl TFIIIA als auch TFIIIC aus dem Präinitiationskomplex entfernt werden konnten, die Pol III aber dennoch auf der DNA verblieb und aktiv war. Somit dienen diese TFs in erster Linie dazu, TFIIIB an die DNA zu führen [KASSAVETIS *et al.*, 1990; 1992].

Neben der Rolle bei der Heranführung der RNA-Polymerase an ihre Promotoren ist TFIIIB außerdem maßgeblich an der Öffnung des Promotors und damit der Initiation der Transkription beteiligt [KASSAVETIS *et al.*, 1998b].

TFIIIB besteht aus den drei Untereinheiten TBP (TATA-Box bindendes Protein), Brf1 und dem Pol III-spezifischen Bdp1. Zumindest in *S. cerevisiae* sind alle drei Untereinheiten für die Transkription notwendig [KASSAVETIS *et al.*, 1999]. TBP und

Brf1 bilden in *S. cerevisiae* und im Menschen den stabilen Komplex B' und können gemeinsam aufgereinigt werden, während Bdp1 leicht dissoziiert [KASSAVETIS *et al.*, 2006].

### 1.2.2.1 TBP

Einer der Kernvorgänge bei der Transkriptionsinitiation ist die Erkennung der TATA-Box durch TBP [ZHAO & HERR 2003].

TBP ist einzigartig unter den Transkriptionsfaktoren, da es das einzige bekannte Protein ist, das von allen drei eukaryontischen RNA-Polymerasen benötigt wird [HERNANDEZ 1993]. Die Mechanismen, nach denen TBP an der Transkription beteiligt sind, wurden für den Pol II-Apparat am besten untersucht. Entsprechend seiner zentralen Rolle im Transkriptionsprozess handelt es sich um ein Protein mit einem hochkonservierten, etwa 180 bp großen C-terminalen Kernstück (*core*). Hier flankieren zwei sich wiederholende Einheiten (*repeats*) eine stark basische Domäne [NIKOLOV *et al.*, 1992]. Der *core*-Bereich bindet durch Interaktion mit der kleinen Furche mit hoher Affinität an die TATA-Box von Pol II und Pol III-Promotoren. An dieser Stelle wird die DNA entwunden und scharf geknickt, um die Bildung des TBP-TATA-Box-Komplexes zu ermöglichen. Der TBP-induzierte Knick in der DNA spielt eine wichtige Rolle für den korrekten Aufbau des Präinitiationskomplexes. Etwa 70% der 180 AS des *Core*-Bereiches sind in allen untersuchten Organismen identisch. Bei den Unterschieden handelt es sich häufig um konservative Austausche, die den hydrophoben Kern von TBP nicht destabilisieren [KIM *et al.*, 1993; JUO *et al.*, 1996; NIKOLOV *et al.*, 1996]. Im humanen Protein wurde eine inhibitorische DNA-Bindestelle (IDB) auf der Oberfläche des *Core*-Bereiches identifiziert, die die zunächst langsame und instabile Bindung von TBP an die DNA erklärt. Der primäre Komplex wird durch TFIIB gefestigt [ZHAO & HERR 2003].

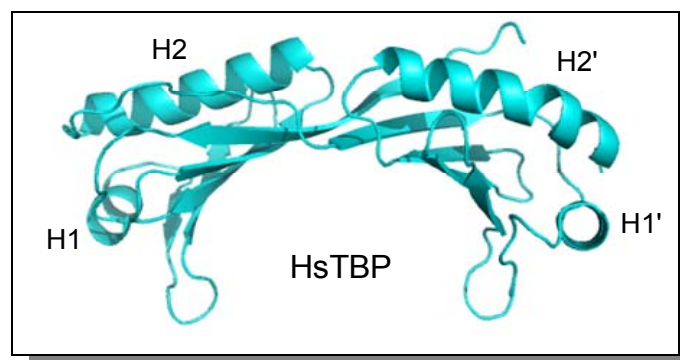
TBP besitzt außerdem eine N-terminale Region, die sowohl bezüglich der Länge als auch der Sequenz zwischen den Spezies deutlich variieren kann. In der Hefe wurde nachgewiesen, dass dieser Teil des Proteins kaum an der Transkription beteiligt ist [CORMACK & STRUHL 1992]. Bei Vertebraten wird diese Region durch einen 28-42 AS langen Glutaminsäure-Stretch charakterisiert. Längere Stretche sind pathologisch und führen zu zerebellären Ataxien und anderen neurodegenerativen Erkrankungen [FUJIGASAKI *et al.*, 2001; NAKAMURA *et al.*, 2001].

Stets mit TBP assoziiert sind die TAFs (TBP assoziierte Faktoren), die die Polymerase-Spezifität vermitteln und gemeinsam mit TBP einen Multiproteinkomplex

bilden, der als Transkriptionsfaktor TFIID (Pol II) oder TFIIB (Pol III) bezeichnet wird. Zusammen mit weiteren generellen Transkriptionsfaktoren koordiniert der TBP-haltige Komplex die Bildung eines funktionellen PIC. Die Pol III-spezifischen TAFs in *S. cerevisiae* sind Brf1 und Bdp1. [SHEN *et al.*, 1998].

Aus Kristallstrukturanalysen geht hervor, dass TBP eine beinahe spiegelsymmetrische sattelförmige Struktur mit einem hydrophoben Kern besitzt (Abb. 1.7). Jede Teildomäne umfasst zwei  $\alpha$ -Helices (H1, H2 bzw. H1' und H2') und fünf antiparallele  $\beta$ -Faltblätter aus 88-89 AS (S1-S5 bzw. S1'-S5'). Die Verbindung zwischen beiden Teildomänen wird durch ein kurzes Linkerpeptid (*basic repeat*) hergestellt [NIKOLOV *et al.*, 1992; 1996]. Die konkave Unterseite wird von acht antiparallelen  $\beta$ -Faltblättern geformt. An dieser Unterseite findet der Kontakt mit der DNA statt. Auf der Oberseite des Sattels bilden zwei große  $\alpha$ -Helices und Teile von S1 und S1' eine konvexe, weitgehend positiv geladene Oberfläche. Die Homologie-Region II von Brf1 tritt über fast die gesamte Helix H2 und zusätzlich Teile von S1 mit TBP in Wechselwirkung, die Hauptkontaktstellen auf HsTBP sind aber drei Arginine (R<sup>231</sup>, R<sup>235</sup>, R<sup>239</sup>) auf H2 [SHEN *et al.*, 1998]. Die Kristallstruktur des Brf1 (Homologie-Region II)-TBP-DNA-Komplexes wurde kürzlich aufgeklärt [JUO *et al.*, 2003]. Ein schwächerer Kontakt findet außerdem zwischen dem N-Terminus von Brf1 und dem C-terminalen Steigbügel von TBP statt. Das zum N-terminalen Teil von Brf1 homologe Protein der Pol II-Maschinerie, TFIIB, bindet prinzipiell an die gleichen Oberflächen am C-Terminus von TBP [SCHRÖDER *et al.*, 2003].

Trotz seiner annähernd symmetrischen Struktur bindet TBP in einer definierten Orientierung an die TATA-Box [NIKOLOV *et al.*, 1996].



**Abb. 1.7: Dreidimensionale Struktur von HsTBP**

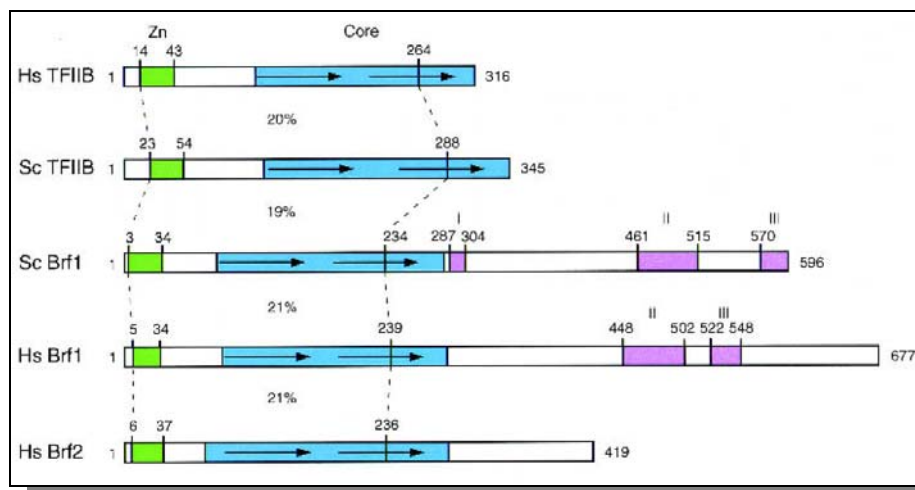
Kristallstruktur von HsTBP nach Nikolov 1996 (SwissProt P20226). Die großen Helices H2 und H2' bilden eine konvexe Oberseite, die als Plattform für Protein-Protein-Interaktionen dient. An der konkaven Unterseite des Moleküls bindet die DNA. Die Grafik wurde mit PyMOL erstellt.

### 1.2.2.2 Brf1

Für das humane Protein Brf1 sowie sein Ortholog in *S. cerevisiae* sind die Gene bekannt [COLBERT & HAHN 1992; WANG & ROEDER 1995]. Sie sind gut konserviert (30% Identität, 53% Homologie) und haben beide eine etwa 310 bp lange N-terminale Domäne, die über die gesamte Länge homolog zu TFIIIB ist. Sie ist hauptsächlich für die Rekrutierung der Pol III sowie die Öffnung des Promotors verantwortlich. Der C-Terminus, der die Pol III-Spezifität vermittelt, ist dagegen deutlich schlechter konserviert. Die Kontaktflächen dieser Domäne sind allerdings maßgeblich für den Zusammenhalt der drei Untereinheiten von TFIIIB. Die gut konservierte Homologie-Region II im C-Terminus stellt die wesentliche Bindestelle für TBP dar, eine geringere Affinität wird durch die N-terminale Domäne vermittelt. Das gleiche gilt für die Bindung von Bdp1 an Brf1 [KASSAVETIS *et al.*, 2006]. Bei der Assemblierung von TFIIIB bindet Brf1 an den TBP-DNA-Komplex und wird auf den gegenüberliegenden Seiten der geknickten DNA an den Steigbügelstrukturen von TBP fixiert.

*In vitro* Versuche haben gezeigt, dass die C-terminale Hälfte von Brf1 einen stabilen aber transkriptionell beinahe inaktiven Komplex mit der DNA bildet, während der Komplex mit der N-terminalen Hälfte zwar instabil ist, aber eine ähnliche Transkriptionsrate wie der Wildtyp-Komplex erreicht [KASSAVETIS *et al.*, 1998a]. Ein artifizielles Fusionsprotein, das die Topologie von TFIIIB imitiert, indem der N-Terminus von Brf1 über die Kernstruktur von TBP mit einem großen C-terminalen Fragment von Brf1 verbunden wird, kann diese beiden TFIIIB-Untereinheiten im Transkriptionsprozess vollständig ersetzen [JUO *et al.*, 2003].

TFIIIB, Brf1 sowie das humane Brf2, das im menschlichen TFIIIB an Typ3-Promotoren die Funktion von Brf1 übernimmt, bilden eine Familie von strukturell und funktionell verwandten Transkriptionsfaktoren (Abb. 1.8). Im C-Terminus sind drei Homologie-Regionen I-III in den Hefen *C. albicans*, *K. lactis*, *S. pombe* und *S. cerevisiae* erhalten, die Regionen II und III finden man auch in HsBrf1 [KHOO *et al.*, 1994; ANDRAU *et al.*, 1999]. Zinkfingerstrukturen bilden den N-Terminus.



**Abb. 1.8: Struktureller Aufbau TFIIB-verbundener Transkriptionsfaktoren**

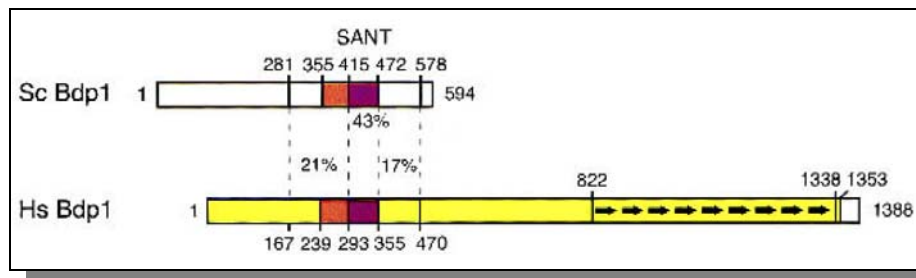
Grün: Zinkfinger, blau: Kernbereich von TFIIB, violett: Homologie-Regionen I-III des C-Terminus. In Prozent sind die Identitäten der einzelnen Proteine im Bereich der maximalen Konservierung angegeben [SCHRAMM & HERNANDEZ 2002].

### 1.2.2.3 Bdp1

Bdp1 ist ein Pol III-spezifischer Faktor, für den es in anderen Transkriptionssystemen keine paraloge Proteine gibt. Bdp1 vermittelt TFIIB seine außergewöhnliche Stabilität gegenüber Hochsalzpuffern oder polyanionischen Stressbedingungen [KASSAVETIS *et al.*, 1990]. Bei Eintritt dieser letzten TFIIB-Untereinheit in den TBP-Brf1-DNA-Komplex verstärkt Bdp1 die Beugung der DNA, so dass die DNA-Bereiche oberhalb und unterhalb der TFIIB-Bindestelle näher zusammenrücken [BRAUN *et al.*, 1992]. Versuche mit ScBdp1 zeigen, dass durch Bindung von Bdp1 TFIIC – mittelbar oder unmittelbar - von seiner Kontaktstelle verdrängt wird [KUMAR *et al.*, 1997].

Bdp1 ist nicht nur für die Rekrutierung der Pol III essentiell, sondern wird ebenfalls für die initiale Öffnung des Promotors benötigt. Dennoch können vor allem N-terminal große Segmente entfernt werden ohne die Funktion des Proteins zu beeinträchtigen [KUMAR *et al.*, 1997]. Die stärkste Konservierung findet man im Bereich der SANT-Domäne im C-terminalen Drittel [KASSAVETIS *et al.*, 2001]. Diese Domäne, die aus drei Helices besteht, findet man typischerweise in Proteinen, die an der transkriptionellen Aktivierung beteiligt sind. Ihre strukturelle Ähnlichkeit zur DNA-bindenden Domäne von Myb-Proteinen lassen eine DNA-Bindefähigkeit vermuten, die es Bdp1 ermöglicht, nach Rekrutierung durch TFIIC auch ohne TFIIC auf der DNA zu verbleiben [AASLAND *et al.*, 1996]. Außerdem wurden Protein-Protein-Interaktionen beobachtet, an denen die SANT-Domäne beteiligt war

[BOYER *et al.*, 2004; DING *et al.*, 2004; MO *et al.*, 2005]. In *S. cerevisiae* ist für die Interaktion zwischen Bdp1 und dem C-Terminus von Brf1 ebenfalls die SANT-Domäne verantwortlich, aber auch über die N-Termini der beiden Proteine findet eine schwache Wechselwirkung statt. Es wird vermutet, dass die Kontaktdomänen der beiden TFIIIB-Untereinheiten ko-konserviert wurden [KASSAVETIS *et al.*, 2001].



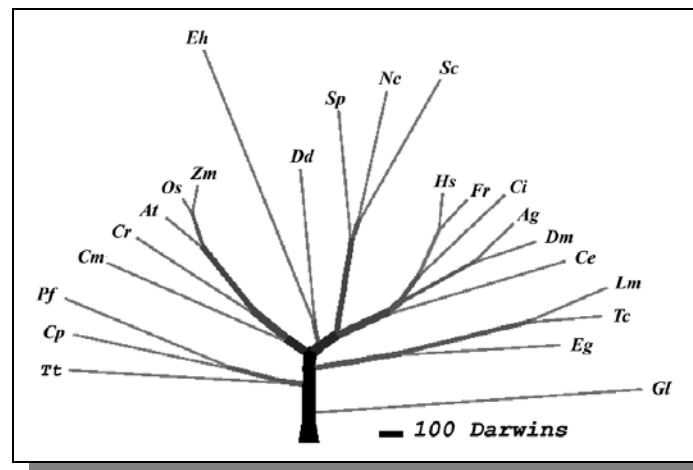
**Abb. 1.9: Strukturen von ScBdp1 und HsBdp1**

Bdp1 ist die am schlechtesten konservierte Einheit von TFIIIB. Am besten konserviert ist die SANT-Domäne (violett), aber auch eine kleine Region oberhalb (orange) weist erkennbare Homologien auf. Die Prozentzahlen geben das Maß der Identitäten für die jeweils gestrichelten Bereiche wider [SCHRAMM & HERNANDEZ 2002].

## 1.3 Dictyostelium discoideum

### 1.3.1 Anmerkungen zur Systematik und Entwicklung

Die erste entdeckte *Dictyostelium*-Art, *D. mucoroides*, wurde von O. Brefeld 1869 irrtümlich als Pilz charakterisiert [BREFELD 1869]. Seitdem hält sich hartnäckig der Begriff „zellulärer Schleimpilz“ (*Dictyostelidae*), auch wenn *Dictyostelium* phylogenetisch nicht mit den Pilzen verwandt ist. Statt dessen legen phylogenetische Analysen mit Gensequenzen ribosomaler RNA, verschiedenen Proteinen und ganzen Genomen den Schluss nahe, dass *Dictyostelium* mit anderen Vertretern der *Amoebozoa* eine monophyletische Gruppe bildet, die als Schwesterstamm zu den Opisthokonten (Tiere und Pilze) anzusehen ist [BAPTESTE *et al.*, 2002; EICHINGER *et al.*, 2005]. Aufgrund ihrer funktionellen und sequentiellen Ähnlichkeit wurden etwa 1850 Proteine zusammengefasst, die spezifisch in Eukaryonten auftreten. Lediglich 746 dieser Proteine konnten auch in *D. discoideum* identifiziert werden. Daraus wird die phylogenetische Distanz zwischen *Dictyostelium discoideum* und Säugern ersichtlich, die eine Identifizierung von orthologen Proteinen erschweren [EICHINGER *et al.*, 2005].



**Abb. 1.10: Phylogenetische Einordnung von *Dictyostelium***

Der Vergleich von Proteomen vollständig sequenzierter Genome ordnet *Dictyostelium* (Dd) als Schwesterstamm der Opisthokonten ein [SONG 2005].

*Dictyostelium discoideum*, der mit Abstand am besten untersuchte Vertreter der *Dictyostelidae*, wurde 1935 in den Wäldern North Carolinas entdeckt und beschrieben [RAPER 1935]. Zwischenzeitlich sind über 70 verschiedene Arten bekannt [MUTZEL 1991]. Vertreter der „sozialen Amöben“ sind in humusreichen Waldböden der gemäßigten Zonen heimisch und leben bei ausreichender Nahrung als einzellige Amöben, während sie bei Nahrungsmangel einen Entwicklungszyklus durchlaufen und einen vielzelligen Organismus bilden.

### 1.3.2 Lebenszyklus

Als Amöbe ernährt sich *Dictyostelium discoideum* durch Phagozytose von Bodenbakterien. Die Bakterien werden von einem unspezifischen Rezeptor der Zellmembran erkannt, von Pseudopodien umschlossen und nach endozytotischer Aufnahme in den Lysosomen zersetzt [BOZZARO & PONTE 1995]. Für eine einzige Zellteilung phagozytiert eine Amöbe etwa 1000 Bakterien.

Während dieser Phase des vegetativen Wachstums liegt *D. discoideum* als 8-12  $\mu\text{m}$  große Amöbe vor, die sich durch Aktin- und Myosinfilamente fortbewegen kann.

Sobald nicht mehr ausreichend Nahrung zur Verfügung steht, die Zelldichte zu groß wird oder eine bestimmte Grenztemperatur überschritten ist, geht *Dictyostelium* in eine 24-30 Stunden andauernde Entwicklungsphase über, in deren Verlauf aus einer Population individueller Amöben ein multizellulärer Organismus entsteht. Die dabei stattfindenden morphologischen Veränderungen und Zelldifferenzierung enden schließlich in der Bildung eines Fruchtkörpers [MAEDA *et al.*, 1989]. Angelockt durch sogenannte Schrittmacherzellen, die periodisch cAMP sezernieren, strömen einzelne



Amöben durch Chemotaxis innerhalb von 6-7 Stunden zu Aggregationszentren (*Mound*) zusammen, die aus etwa  $10^5$  Zellen bestehen. Dabei bewegen sie sich stets in Richtung der höheren cAMP-Konzentration, so dass es zur Bildung charakteristischer, astförmiger Bahnen kommt (*Streaming*) [KONIJN *et al.*, 1967; LOOMIS 1993]. Die Größe dieser Aggregate wird durch den *Counting Factor*, einem Multiproteinkomplex, reguliert [GOMER 1999]. Im Aggregat fusionieren die einzelnen Amöben im Gegensatz zu den echten plasmodischen Schleimpilzen allerdings nicht, sondern bilden unter Aufrechterhaltung der individuellen Zellstruktur ein Pseudoplasmodium (*Slug*), das in der Lage ist, sich aufgrund thermischer oder chemotaktischer Reize zu bewegen (wandernder *Slug*) [BONNER 1994]. Nach etwa 10 Stunden wird die Differenzierung der Zellen in Prästiel- und Präsporenzellen eingeleitet. Ca. 75% der Zellen eines *Slugs* entwickeln sich zu Präsporenzellen und 15% zu Prästielzellen, die unter Ausbildung einer Stielröhre (*Mexican Hat*) auf den Boden des Pseudoplasmodiums wandern, stark anschwellen und dort schließlich absterben. Gleichzeitig werden die Präsporenzellen angehoben (Kulmination). Im Gegensatz zu den Stielzellen, die absterben, gehen Sporenzellen in einen Ruhezustand über. Sie keimen bei günstigen Bedingungen wieder aus um anschließend erneut in die vegetative Phase einzutreten [COTTER *et al.*, 1992]. Ein schematischer Überblick der Entwicklungsphasen ist in Abb. 1.11 dargestellt. Durch diese Form des Lebenszyklus ist *Dictyostelium* in der Lage, auch schwierige Umweltbedingungen zu überdauern.

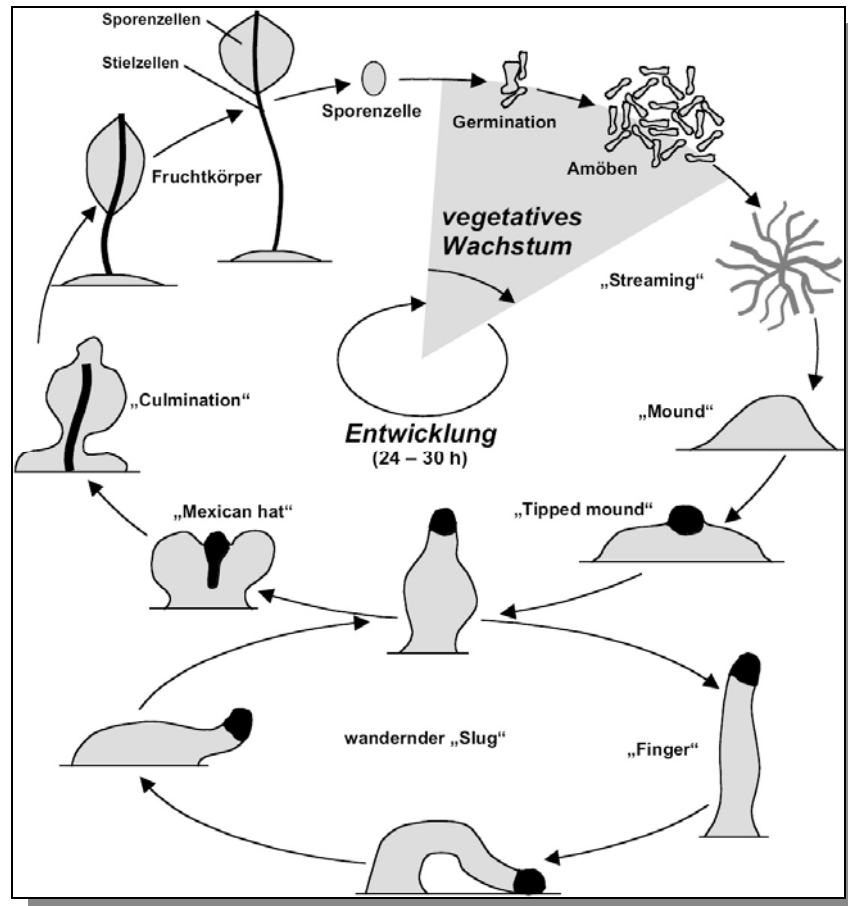


Abb. 1.11: Schematische Darstellung des Entwicklungszyklus von *D. discoideum*

### 1.3.3 Das Genom – Aufbau und Struktur

*Dictyostelium discoideum* besitzt ein etwa 34 Mb großes haploides Genom, das sich auf sechs 4-8 Mb große Chromosomen verteilt. Neben den Chromosomen befinden sich im Zellkern etwa 100 Kopien eines 88 kb großen, extrachromosomalen Palindroms, das die Information für die ribosomalen RNAs enthält [SUCGANG *et al.*, 2003; EICHINGER *et al.*, 2005].

2005 wurde die Sequenzierung des Genoms des *Dictyostelium*-Stammes AX4 durch das Internationale *Dictyostelium*-Genom-Projekt abgeschlossen [EICHINGER *et al.*, 2005]. Bemerkenswert ist der überdurchschnittlich hohe A/T-Gehalt (78%). Während er in Exons 73% beträgt, liegt der A/T-Anteil in nichtkodierenden Regionen, in denen homopolymere Sequenzen bis zu 180 bp lang sein können, sogar bei bis zu 96%. Für die Kodierung von Aminosäuren werden A/T-reiche Kodons gegenüber den G/C-reichen Triplets bevorzugt. Ebenso sind A/T-reiche Aminosäuren wie Asn, Lys, Ile, Tyr und Phe in *Dictyostelium*-Proteinen deutlich häufiger anzutreffen als GC-reiche Aminosäuren [EICHINGER *et al.*, 2005].

Etwa 71% des Erbguts kodieren für 125000 gleichmäßig im Genom verteilte Gene, deren Exons 63% des Genoms ausmachen. Es gibt nur wenige, meist kurze Introns. Verglichen mit dem humanen Erbgut, das mit einer Größe von 2900 Mb nur 2-3mal so viele Gene kodiert wie das *Dictyostelium*-Genom, handelt es sich bei letzterem also um ein sehr kompaktes Genom, in dem die kodierenden Bereiche der einzelnen Gene kaum 1000 bp auseinander liegen [LANDER *et al.*, 2001; EICHINGER *et al.*, 2005]. In diesen intergenischen Abschnitten befinden sich allerdings noch regulatorische Einheiten wie Promotoren oder Terminatoren. Es wird deutlich, dass mobile Elemente in *Dictyostelium discoideum* nur sehr begrenzte Möglichkeiten zur sicheren Integration haben, ohne Gene oder deren regulatorische Einheiten durch Insertionsmutation zu zerstören. Dennoch bestehen rund 10% des Genoms von *D. discoideum* aus mobilen Elementen, was auf zuverlässige positionsspezifische Integrationsmechanismen schließen lässt [GLÖCKNER *et al.*, 2002].

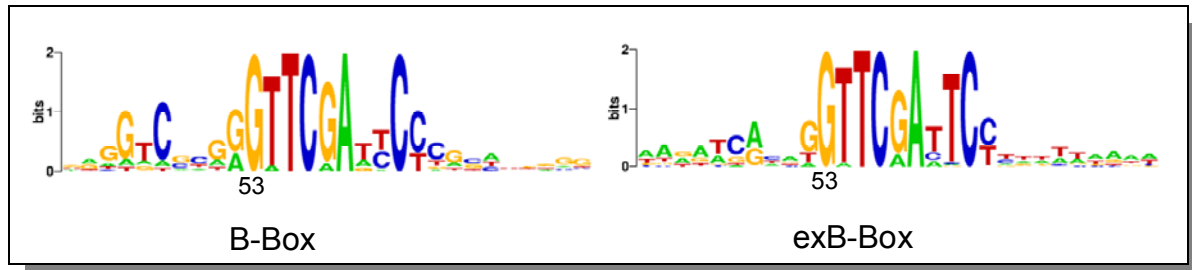
#### 1.3.4 tRNA-Gene in *Dictyostelium discoideum*

In *D. discoideum* liegen insgesamt rund 390 tRNA-Gen-Kopien vor, die homogen im Genom verteilt sind. Überdurchschnittlich viele tRNA-Gene treten als tandemartige Verdoppelungen auf. Etwa 54% sind mit Retroelementen assoziiert (siehe unten) [EICHINGER *et al.*, 2005; WINCKLER *et al.*, 2005].

Entsprechend des hohen A/T-Gehaltes des Erbguts zeigt *D. discoideum* einen starken Bias zur Nutzung von tRNAs mit A/T-reichen Kodons [EICHINGER *et al.*, 2005]. Von den drei Stopkodons *ochre* (UAA), *amber* (UAG) und *opal* (UGA) wird in 90% aller Fälle das *ochre*-Stopkodon verwendet, die beiden anderen Kodons terminieren zu gleichen Teilen die Translation der übrigen Gene [WINCKLER *et al.*, 2005].

tRNA-Gene in *Dictyostelium* besitzen zwei interne Promotorregionen: die A-Box (Position 8-19) und die B-Box (Position 53-61). Zwar kontaktiert TFIIIC in anderen Organismen auch die A-Box, allerdings konnte dies in *Dictyostelium* bisher noch nicht nachgewiesen werden [BUKENBERGER *et al.*, 1994]. Ein *Alignment* der B-Boxen praktisch aller tRNA-Gene zeigt, dass die konservierten Domänen der A- und B-Boxen der Konsensus-Sequenz anderer eukaryontischer Promotorsequenzen folgt (O. Siol, persönliche Mitteilung). Eine genaue Analyse der B-Boxen in *D. discoideum*-tRNA-Genen hat eine zusätzliche konservierte Trinukleotidsequenz (GTC) 4 bp oberhalb der B-Box-Konsensussequenz identifiziert (Abb. 1.12). Für das Guanin an Position 53 wurde gezeigt, dass es für die Bindung von TFIIIC

wahrscheinlich essentiell ist. tRNA-Gene mit einer mutierten B-Box, die nicht mehr von TFIIC erkannt wird, werden auch von TRE5-A nicht mehr als möglicher Integrationsort erfasst [SIOL *et al.*, 2006].



**Abb. 1.12: Konsensussequenz der B-Boxen in *D. discoideum***

Dargestellt ist ein Alignment der Konsensussequenz der B-Boxen (links) und der exB-Boxen (rechts) von *D. discoideum*-tRNA-Genen inkl. 10 flankierender Basenpaare (T.Winckler, nicht publiziert). Guanin 53 ist markiert. Die Sequenzlogos wurden mit dem Programm WebLogo erzeugt (<http://weblogo.berkeley.edu>).

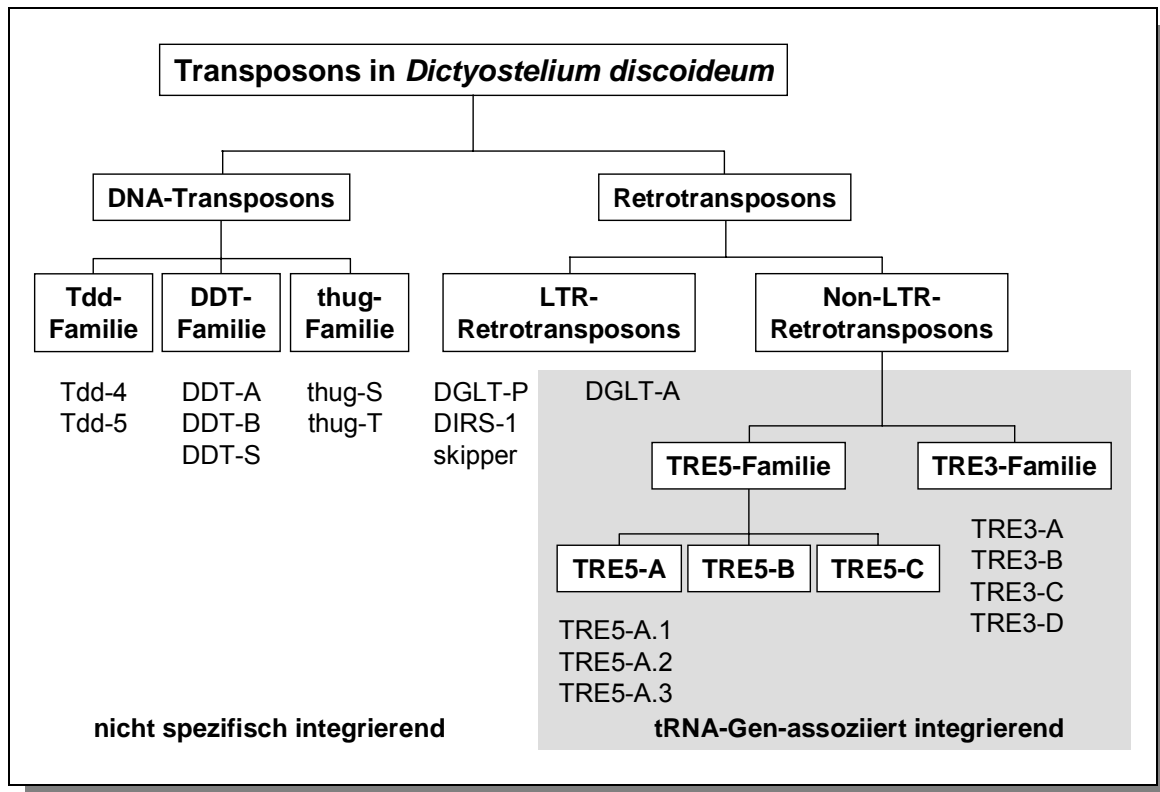
Ca. 66% aller tRNA-Gene in *Dictyostelium discoideum* besitzen 19-35 bp unterhalb des Gens eine zusätzliche externe B-Box, die möglicherweise als „Parkposition“ für TFIIC während der Transkription genutzt wird [SIOL *et al.*, 2006]. Die Mehrzahl der externen B-Boxen zeigt eine stark ausgeprägte Homologie zur Konsensus-Sequenz der internen B-Box. Obwohl die genaue Funktion der externen B-Box noch nicht bekannt ist, deutet die starke Konservierung aber darauf hin, dass sie nicht unbedeutend ist. Erste Versuche lassen vermuten, dass eine externe B-Box die Funktion einer inaktiven internen B-Box teilweise ersetzen kann [HOFMANN *et al.*, 1991; SIOL *et al.*, 2006].

Die genaue Struktur von *Dictyostelium*-TFIIC ist noch unbekannt. In der Bäckerhefe besteht TFIIC aus den sechs Untereinheiten TFIIC55, TFIIC60, TFIIC91, TFIIC95, TFIIC131 und TFIIC138. Obwohl die Aktivität von *Dictyostelium*-TFIIC biochemisch eindeutig gezeigt wurde [BUKENBERGER *et al.*, 1994], konnten bisher nur zwei Proteine mit den orthologen Untereinheiten aus *S. cerevisiae* sicher identifiziert werden (TFIIC95 = DDB0231091, TFIIC131 = DDB0231090).

## 1.4 Transposable Elemente in *Dictyostelium discoideum*

Das Genom von *D. discoideum* besteht zu ca. 10% aus transposablen Elementen [GLÖCKNER *et al.*, 2001b]. Der Großteil dieser Elemente gehört zur Gruppe der Retrotransposons, lediglich 1,5% sind DNA-Transposons [WINCKLER 1998; SZAFRANSKI *et al.*, 1999a]. Non-LTR-Retrotransposon-Sequenzen machen 3,7% des Genoms aus [GLÖCKNER *et al.*, 2001b].

Ein Überblick der Transposablen Elemente ist in Abb. 1.13 dargestellt. Demzufolge gibt es 18 verschiedene Transposons, die in 7 Familien gegliedert werden können. Die Verteilung im Genom ist nicht homogen. Vielmehr drängen sich die Elemente in „Clustern“ in der Nähe von tRNA-Genen oder den Zentromeren der jeweiligen Chromosomen [WINCKLER *et al.*, 2005].



**Abb. 1.13: Transposons in *Dictyostelium discoideum***  
 Überblick über die transposablen Elemente in *Dictyostelium discoideum*

### 1.4.1 DNA-Transposons

Zur Gruppe putativer DNA-Transposons gehören die drei Familien Tdd-4 (Transposons in *Dictyostelium discoideum*), DDT und *thug*. Analysen der Aminosäuresequenzen der ORFs (Transposasen) geben keine Hinweise auf Verwandtschaften zu anderen Elementen. Das Vorhandensein von Introns sowie ITRs und die Länge der TSDs lassen vermuten, dass Tdd-4-Elemente kanonische DNA-Transposons sind, die im Laufe der Evolution die IN-Domäne eines LTR-Transposons adaptiert haben und somit den Übergang zwischen DNA- und Retrotransposon darstellen [WINCKLER *et al.*, 2002a; 2005].

### 1.4.2 LTR-Retrotransposons

#### 1.4.2.1 DIRS-1

DIRS-1 (*Dictyostelium intermediate repeat sequence*) war das erste mobile Element, das in *D. discoideum* entdeckt wurde. Es ist mit Abstand das am häufigsten vertretene TE, allerdings sind die meisten Kopien aufgrund interner Rekombination defekt. Vollständige DIRS1-Elemente sind 4,8 kb groß und besitzen zwei sich partiell überlappende ORFs sowie nicht-identische LTRs, die ungewöhnlicherweise invers angeordnet sind. Der zweite ORF kodiert u.a. für eine Integrase und überdeckt vollständig einen dritten ORF einer RT/RNaseH-Domäne, die eine entfernte Verwandtschaft zu korrespondierenden Strukturen in Ty3/*gypsy*-like LTR-retrotransposons aufweist [MALIK *et al.*, 2000]. DIRS-1 integriert bevorzugt in andere DIRS-1-Elemente und bildet so große Cluster. Jedes *D. discoideum*-Chromosom trägt nahe einem Ende ein solches Cluster, so dass man annimmt, dass diese Cluster als Centromere dienen könnten [LOOMIS *et al.*, 1995; EICHINGER *et al.*, 2005].

#### 1.4.2.2 Skipper

Durch vergleichende Analysen der RT-Sequenzen konnten Skipper-Elemente der Ty3/*gypsy*-Familie zugeordnet werden [MALIK & EICKBUSH 2001]. Die Elemente sind ca. 7 kb lang, besitzen zwei 390 bp große LTRs und zwei ORFs, deren Leserahmen versetzt sind. Ungewöhnlich ist die Anordnung der Gene, bei der im ersten ORF die Gene *pro* und *gag* fusioniert vorliegen [WINCKLER *et al.*, 2005].

### 1.4.2.3 DGLT-A und B

DGLT-Retrotransposons (*Dictyostelium gypsy-like transposon*) wurden erst durch das *Dictyostelium* Genom Projekt entdeckt. Deutlich früher bekannt war das „H3R-Element“, das heute als „Solo-LTR“ von DGLT-A aufzufassen ist [HOFMANN *et al.*, 1991]. DGLT-A-Elemente integrieren 10-30 bp oberhalb von tRNA-Genen. Sie sind etwa 5 kb groß, besitzen einen ORF mit einer RT- und IN-Domäne und werden von zwei 260 bp langen LTRs flankiert. Die fünf Kopien, die sich im Genom befinden, sind aufgrund von Mutationen alle inaktiv.

DGLT-B-Elemente sind rund 6,2 kb groß und werden von zwei LTRs eingerahmt. Sie besitzen wahrscheinlich vier ORFs, was aber wegen gravierender Degeneration der Elemente während der Evolution nicht mehr genau bestimmbar ist. Im Gegensatz zu den DGLT-A-Transposons zeigen sie keine Integrationsspezifität für tRNA-Gene [GLÖCKNER *et al.*, 2001b; WINCKLER *et al.*, 2005].

Ein weiteres *gypsy-like* TE wurde DGLT-P (Pseudotransposon) genannt. Es handelt sich hierbei möglicherweise nicht um ein vollständiges Transposon, da kein RT-kodierender ORF nachgewiesen werden konnte. Man geht davon aus, dass sich dieses Element im Laufe der Evolution nicht in *Dictyostelium discoideum* durchsetzen konnte [GLÖCKNER *et al.*, 2001b].

### 1.4.3 tRNA-spezifische Non-LTR-Retrotransposons (TREs)

Aufgrund ihrer bemerkenswerten Integrationsspezifität für tRNA-Gene werden die sieben Non-LTR-Retrotransposons aus *D. discoideum* als *tRNA-gene targeted retroelements* (TREs) bezeichnet [SZAFRANSKI *et al.*, 1999a]. Je nach Position der Integration - entweder am 5'- oder 3'-Ende eines tRNA-Gens - unterscheidet man zwischen TRE5- und TRE3-Elementen. Alle TREs integrieren orientierungsspezifisch mit ihrem 5'-Ende in Richtung des tRNA-Gens.

#### 1.4.3.1 TRE5-Gruppe

Elemente der TRE5-Gruppe integrieren mit eindrucksvoller Präzision orientierungsspezifisch mit ihrem 5'-Ende in Richtung des tRNA-Gens  $48 \pm 2$  bp oberhalb von tRNA-Genen [WINCKLER *et al.*, 2005]. Nach ihrer Struktur unterscheidet man zwischen TRE5-A, -B und -C-Elementen.

## TRE5-A

Retrotransposons dieser Gruppe wurden in den späten 1980er Jahren als DREs (*Dictyostelium repetitive element*) entdeckt und intensiv untersucht [MARSCHALEK *et al.*, 1989; HOFMANN *et al.*, 1991; MARSCHALEK *et al.*, 1992b; 1992a]. Sie zeigen zahlreiche strukturelle Variationen, so dass man nach der neuen Namensgebung zwischen TRE5-A.1, -A.2 und -A.3 unterscheidet [SZAFRANSKI *et al.*, 1999b; BECK *et al.*, 2002]. Zusammengenommen befinden sich etwa 70 Kopien im Genom [GLÖCKNER *et al.*, 2001b].

TRE5-A-Elemente können bei der Integration zusätzliche Nukleotide einführen und erzeugen 12-16 bp lange TSDs [WINCKLER *et al.*, 2002a].

Vollständige TRE5-A.1-Elemente sind 5,7 kb groß, und besitzen zwei ORFs, die sich auf rund 70 bp in unterschiedlichen Leserahmen überlappen. Die ORFs von TRE5-A werden beidseitig von modularen nicht translatierten Regionen (UTRs) eingefasst (siehe unten).

Der erste ORF kodiert für ein 51 kDa großes Protein, das keine Homologie zu bekannten Proteinen aufweist, also auch nicht zu den ORF1-Proteinen anderer Non-LTR-Retrotransposons. Im Zytosol von *D. discoideum* bildet rekombinant exprimiertes ORF1p große Komplexe [HENTSCHEL *et al.*, 2001]. In Analogie zu den ORF1-Proteinen anderer Non-LTR-Retrotransposons ist es möglich, dass mehrere Moleküle ORF1p mit TRE5-A-Transkripten Ribonukleoproteinpartikel (RNPs) bilden, die anschließend zur reversen Transkription und Integration durch den TPRT-Mechanismus an die genomische DNA gebracht werden [WINCKLER *et al.*, 2002a]. Im Gegensatz zu ORF2p kann endogenes ORF1p in Western Blot Analysen detektiert werden. Ein Überschuss an ORF1p gegenüber ORF2p wird wahrscheinlich durch eine ineffiziente Termination und Reinitiation der Translation erzeugt.

ORF2p besteht aus drei distinkten Domänen, die Homologien zu Reversen Transkriptasen (RT), Endonukleasen (EN) und cysteinreichen Domänen aus anderen Non-LTR-Retrotransposons aufweisen [MALIK *et al.*, 1999; YANG *et al.*, 1999; OSTERTAG & KAZAZIAN 2001a]. Bakteriell exprimiertes ENp bildet ebenfalls unlösliche Aggregate [HENTSCHEL *et al.*, 2001]. Weder in ORF1p noch in ORF2p konnte ein Kernlokalisierungssignal gefunden werden. Da es sich bei *D. discoideum* allerdings um einen sich oft teilenden Organismus mit kurzer Generationszeit handelt, können



die RNPs während der Mitose die partiell aufgelöste nukleäre Membran passieren und in den Zellkern gelangen [WINCKLER *et al.*, 2002a].

Die 5'-UTR beginnt mit einem A-Modul. Dabei handelt es sich um eine 200 bp lange Sequenz, die als interner Promotor die Synthese der Plusstrang-RNA kontrolliert [SCHUMANN *et al.*, 1994a]. Das Transkript beginnt mit der ersten Base des A-Moduls um zu gewährleisten, dass ein mobilisiertes TRE5-A nach wie vor einen vollständigen Promotor besitzt. Bei der vollständigen Sequenzierung des *D. discoideum*-Genoms wurden TRE5-A-Elemente mit bis zu sieben aufeinander folgenden A-Modulen gefunden. Einen ähnlichen Modul-Block in der 5'-UTR findet man im murinen L1-Element. Dort hat ein monomeres Modul allerdings nur sehr schwache Promotoraktivität, während in TRE5-A bereits ein einzelnes A-Modul als starker Promotor funktioniert [SCHUMANN *et al.*, 1994b; DEBERARDINIS & KAZAZIAN 1999].

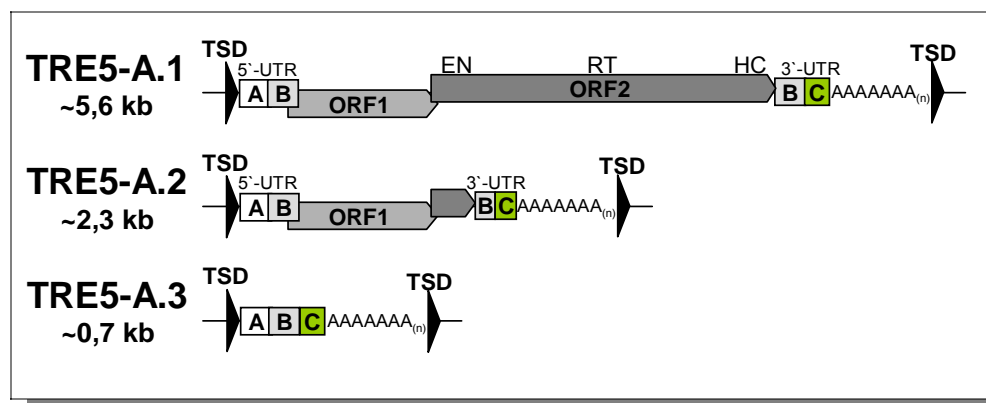
Durch eine Kopie der ersten 70 bp des A-Moduls sowie einen 39 bp langen homopolymeren A-Stretch vom A-Modul getrennt folgt das 290 bp große B-Modul. Hier liegt das Startkodon des ersten ORF. Es ist möglich, dass das B-Modul der ursprüngliche Promotor der TRE5-Elemente war, die genaue Funktion ist allerdings nicht bekannt [MARSCHALEK *et al.*, 1992a; WINCKLER *et al.*, 2002a]. Im Laufe der Evolution der TRE5-A-Elemente wurde das B-Modul dupliziert und liegt nun als Kopie auch in der 3'-UTR vor.

Die 3'-UTR besteht aus der Kopie des B-Moduls sowie einem C-Modul und endet in einem Poly-Adenosin-Schwanz. Das knapp 140 bp große C-Modul ist ein schwacher Promotor, der die Synthese einer Population von Minusstrang-RNA steuert [SCHUMANN *et al.*, 1994a]. Die Minusstrang-RNA reguliert möglicherweise die Menge an Plusstrang-RNA, welche als mRNA fungiert und für die Retrotransposition notwendig ist. Das Protein CbfA aus *D. discoideum* bindet *in vitro* an das C-Modul und könnte so der Wirtszelle ermöglichen, regulierend in Retrotranspositionseignisse einzugreifen [HORN *et al.*, 1999; WINCKLER *et al.*, 2002a]. Die meisten TRE5-A.1-Elemente sind vermutlich inaktiv, da ihnen aufgrund von 5'-Verkürzungen ein funktionsfähiger Promotor fehlt [BECK *et al.*, 2002]. Diese 5'-Verkürzungen werden häufig bei Non-LTR-Retrotransposons gefunden und gehen wahrscheinlich auf eine unvollständige reverse Transkription während des Integrationsprozesses zurück [DEWANNIEUX & HEIDMANN 2005]. Jedoch wurde kürzlich gezeigt, dass TRE5-A-Elemente in *D. discoideum*-Zellen aktiv sind. Für den

Nachweis wurde ein *in vivo* Retrotranspositions-Testsystem entwickelt, das die Isolierung neuer Integrationen ermöglicht. Die „TRE-Falle“ trägt das Gen für die UMP-Synthase (*pyr5-6*), in dessen Intron ein Köder-tRNA-Gen eingefügt wurde. Stabil in *D. discoideum*-Zellen transformiert, wird das *pyr5-6*-Gen nach Herausspleißen des Introns und damit des Ködergens exprimiert. Diese Zellen (Phänotyp: *ura*<sup>+</sup>) sind 5-Fluoro-Orotat-sensitiv (5-FO), da die UMP-Synthase das Zytostatikum in das giftige 5-Fluoro-Uridin-Monophosphat umwandelt. *Ura*<sup>+</sup>-Zellen sterben somit nach 2-3 Wochen in 5-FO-haltigem Medium ab. Wird die TRE-Falle allerdings von einem TRE-Element angesprungen, wird das UMP-Synthase-Gen durch Insertionsmutation inaktiviert (Phänotyp: *ura*<sup>-</sup>). Solche Zellen sind resistent gegen 5-FO und können isoliert und charakterisiert werden [BECK *et al.*, 2002; SIOL *et al.*, 2006].

TRE5-A-Elemente gibt es in zahlreichen strukturellen Varianten. Sehr geläufig ist die Deletion von 3,1 kb des ORF2 in Kombination mit drei kleineren Deletionen im 3'-UTR sowie verschiedenen Punktmutationen im A-Modul [MARSCHALEK *et al.*, 1992a]. Diese Variante, TRE5-A.2, ist nicht mehr autonom und muss von TRE5-A.1-Elementen in *trans* mobilisiert werden. Dennoch sind diese Elemente im heutigen *D. discoideum*-Genom mindestens genauso erfolgreich wie die autonomen TRE5-A.1-Elemente [BECK *et al.*, 2002; SIOL *et al.*, 2006].

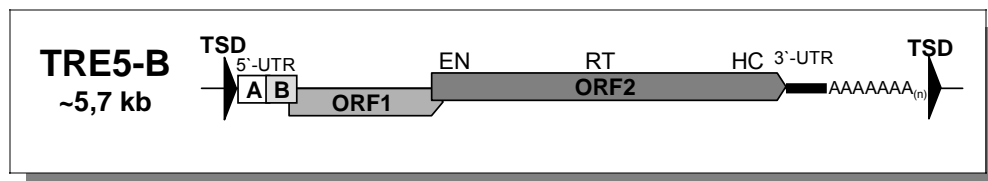
Durch interne Rekombination zwischen den zwei B-Modulen eines Elementes kann es zur vollständigen Deletion beider ORFs kommen. TRE5-A.3-Elemente bestehen also nur aus einem A-, B- und C-Modul. Ob sie aktiv sind oder erst nach der Mobilisierung entstanden ist unbekannt.



**Abb. 1.14: Aufbau der TRE5-A-Elemente**  
Modifiziert nach Winckler 2005.

## TRE5-B

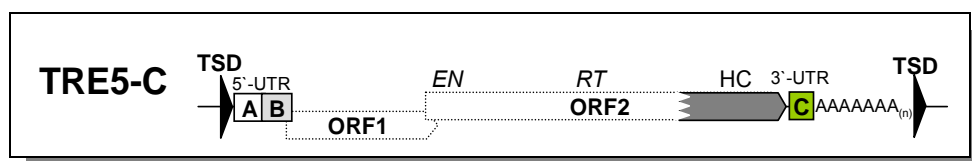
TRE5-B ist ebenso wie TRE5-A modular aufgebaut, allerdings fehlt ihm in der 3'-UTR sowohl das B- als auch das C-Modul. Elemente dieser Gruppe sind 5,9 kb groß, integrieren im Abstand von 45-50 bp ebenfalls oberhalb von tRNA-Genen und hinterlassen ähnliche TSDs wie TRE5-A-Retrotransposons. Bisher wurde zumindest eine Kopie gefunden, die statt des typischen Abstands von 45-50 bp 96 bp zum tRNA-Gen aufwies. Insgesamt gibt es ca. 15 Kopien von TRE5-B im Genom.



**Abb. 1.15: TRE5-B**  
Schematischer Aufbau eines TRE5-B-Elementes. Modifiziert nach Winckler 2005.

## TRE5-C

Wie auch die TRE5-B-Elemente wurde diese Gruppe erst im Laufe des *Dictyostelium* Genom-Projektes entdeckt. TRE5-C-Retrotransposons enthalten vermutlich ebenfalls wie die vorangegangenen beiden Gruppen in der 5'-UTR ein A- und B-Modul sowie zwei sich überlappende ORFs. Allerdings konnte kein komplettes TRE5-C durch Genomprojekt-Daten assembliert werden. Es ist somit unklar, ob das *Dictyostelium discoideum*-Genom ein vollständiges, aktives TRE5-C enthält. In der 3'-UTR besitzt TRE5-C ein C-Modul, das dem aus TRE5-A recht ähnlich ist. Vier der im Genom gefundenen Kopien befinden sich in einem Abstand von 44-46 bp oberhalb eines tRNA-Gens, eine Kopie ist durch eine nachträgliche Insertion eines DGLT-A weiter vom assoziierten tRNA-Gen entfernt [WINCKLER *et al.*, 2002a].



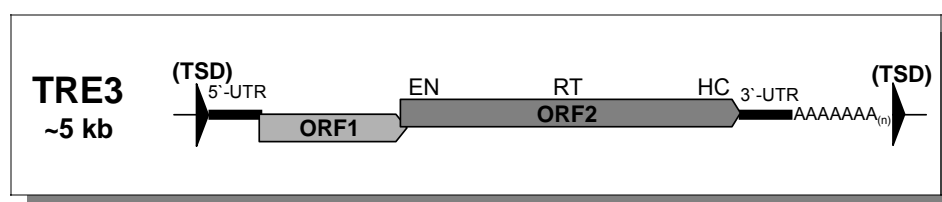
**Abb. 1.16: TRE5-C**  
Schematische Darstellung eines TRE5-C-Elementes. Modifiziert nach Winckler 2005.

### 1.4.3.2 TRE3-Gruppe

Die vier TRE3-Elemente weisen eine ähnliche Struktur wie die TRE5-Retrotransposons auf, die Gruppe ist allerdings wesentlich einheitlicher. Vor allem die ORFs sind gut konserviert, Variationen finden sich in der 5'- und 3'-UTR. In TRE3-A konnte im Bereich der ersten 350 bp der 5'-UTR eine schwache Promotoreigenschaft festgestellt werden [WINCKLER *et al.*, 2002a].

Das Genom von *D. discoideum* enthält ca. 60 Kopien von TRE3-A, 50 von TRE3-B, 30 von TRE3-C und weitere 5 von TRE3-D. Die meisten sind 5'-verkürzt und somit inaktiv. Sie haben eine Länge von 4,7-5,3 kb und enthalten zwei entweder sequentiell angeordnete (TRE3-A und -C) oder sich teilweise überlappende ORFs (TRE3-B), die von UTRs flankiert werden. Ein bis zu 70 bp langer homopolymerer A-Stretch bildet das Ende der 3'-UTR. Der Abstand zum tRNA-Gen ist bei TRE3-Elementen deutlich unterschiedlicher als bei den TRE5-Elementen [WINCKLER *et al.*, 1998; SZAFRANSKI *et al.*, 1999a].

Etwa 66% der tRNA-Gene besitzen eine zusätzliche externe B-Box unterhalb der kodierenden Bereiche. Es scheint, dass TRE3-Retrotransposons sowohl die Gen-interne als auch die externe B-Box als Zielstruktur erkennen. Von beiden B-Boxen ist bekannt, dass sie Bindungsstellen für den Pol III Transkriptionsfaktor TFIIIC darstellen [BUKENBERGER *et al.*, 1994], so dass angenommen wird, dass die Integrationsspezifität der TRE3-Elemente durch Wechselwirkungen zwischen TRE3-Proteinen und TFIIIC vermittelt wird [WINCKLER *et al.*, 2005].



**Abb. 1.17: TRE3-Element**

Schematischer Aufbau eines TRE3-Elementes. Modifiziert nach Winckler 2005.

## 1.5 Transposons in der Gentherapie

Die Gentherapie ist eine experimentelle Strategie der Medizin, bei der Fremd-DNA in die Zellen des Patienten eingebracht wird. In ihrer engsten Definition bedeutet Gentherapie die Reparatur eines defekten Genabschnittes (Genkorrektur), sie umfasst im weiteren Sinne aber auch den Ersatz defekter Gene durch das

Einbringen funktioneller, zusätzlicher Kopien (Genaddition), die Inaktivierung pathogener Genprodukte (Anti-Genherapie) oder auch die Immunisierung gegen das Genprodukt.

Für die dauerhafte Expression des therapeutischen Gens ist es erforderlich, dass es stabil in das Genom des Patienten integriert wird. Die Verwendung viraler Vektoren ermöglicht neben der stabilen Integration des Gens eine effektive Transfektion und ausgeprägte Expression des Proteins, daher werden bei den meisten Genherapie-Ansätzen Retroviren, Adenoviren oder Adenoassoziierte Viren verwendet.

Daneben gibt es verschiedene Ansätze, die stabile und effektive Integration von Transposon-DNA in das Wirtsgenom für die Lösung dieser Problematik zu nutzen. So wurde „*Sleeping Beauty*“, ein Transposon aus dem Lachs, synthetisch wiederhergestellt, nachdem es seine Insertionsfähigkeit im Lauf der Evolution durch mehrere Mutationen verloren hatte. Die Transposase dieses Elements mobilisiert jede DNA-Sequenz *in trans*, die von „*Sleeping Beauty*“-ITRs flankiert wird. In ersten Versuchen wurde ein Expressionsvektor für das Enzym sowie ein Plasmid, das das von ITRs flankierte therapeutische Gen enthält, in Mäuse injiziert [CARLSON *et al.*, 2005]. Die persistierende Expression der Transposase beeinträchtigte allerdings die Stabilität des Transposons. Da zudem eine potentielle Genotoxizität beobachtet wurde, wird zur Zeit die Möglichkeit untersucht, das Enzym in Form seiner mRNA in die Zellen einzubringen [WILBER *et al.*, 2006; WILBER *et al.*, 2007].

In einem anderen Ansatz wurde ein L1-Element mit einem Adenovirus fusioniert [SOIFER & KASAHARA 2004]. Der chimäre Virus kodierte damit unter anderem für das L1-Element, in welches das therapeutische Gen eingebracht werden sollte. Ein wichtiger Vorteil dieses Ansatzes war, dass der Adenovirus sowohl ruhende als auch sich teilende Zellen sehr effektiv infizierte [KUBO *et al.*, 2006]. Es konnte beobachtet werden, dass in Abhängigkeit der vorgeschalteten Pol II-Promotoren Retrotranspositionsfrequenz und Expressionsrate des L1-Elementes in verschiedenen Zelltypen unterschiedlich ausgeprägt waren, so dass sich hier eine weitere Möglichkeit zur gezielten Integration des therapeutischen Gens in definierte Gewebe eröffnet. Als limitierender Faktor dieses Ansatzes stellte sich die Größe des therapeutischen Gens heraus, da die Retrotranspositionsfrequenz des Vektors ab einer Insertgröße von etwa 4 kb deutlich herabgesetzt war [SOIFER & KASAHARA 2004].

## 1.6 Zielsetzung

Trotz zahlreicher Bemühungen um eine sichere Gentherapie sind bisher noch viele Probleme ungelöst. Darunter fallen eine geringe Transfektionsrate, eine häufig nur transiente Expression des therapeutischen Gens durch die fehlende oder instabile Integration in das Wirtsgenom und die ungerichtete Insertion.

Retrovirale Systeme sind in der Lage, die beiden erstgenannten Probleme weitgehend zu lösen. Allerdings findet die Integration ungerichtet statt, so dass durch die Insertion des therapeutischen Gens funktionelle Gene zerstört werden und weitere Krankheiten resultieren können.

Das Non-LTR-Retrotransposon TRE5-A integriert ausnahmslos positions- und orientierungsspezifisch in nicht-kodierende Bereiche ca. 50 bp oberhalb von tRNA-Genen im Genom von *Dictyostelium discoideum*. Dieses Element könnte also als Modell dienen, um einem retroviral basierten System Integrationsspezifität zu verleihen und die Gentherapie deutlich sicherer zu gestalten.

Als Arbeitshypothese wurde formuliert, dass die restriktive Erkennung der möglichen Integrationsorte nicht aufgrund einer spezifischen Bindung von TRE5-A-kodierten Proteinen an flankierende genomische DNA-Sequenzen erfolgt. Vielmehr wird angenommen, dass der Zielort indirekt durch Interaktionen zwischen dem tRNA-Gen-spezifischen Transkriptionsfaktor TFIIIB und TRE5-A-kodierten Proteinen (ORF1p, ORF2p) erkannt wird.

Ziel dieser Arbeit war es daher, durch die Analyse potentieller Proteininteraktionen zwischen TRE5-A und tRNA-Gen-spezifischen Transkriptionsfaktoren den Mechanismus der Zielorterkennung durch TRE5-A genauer zu analysieren. Mit Hilfe eines bakteriellen Zweihybrid-Assays sollten alle TRE5-A.1-kodierten Proteine und die drei DdTFIIIB-Untereinheiten DdTBP, DdBrf1 und DdBdp1 systematisch auf mögliche Interaktionen untersucht werden. Zusätzlich sollte das interne Zusammenspiel der drei Untereinheiten von DdTFIIIB charakterisiert werden. Nach der positiven Identifizierung möglicher Interaktionen zwischen TRE5-A und DdTFIIIB-Untereinheiten sollten die jeweils beteiligten Kontaktdomänen möglichst genau eingegrenzt werden.

## 2 Material und Methoden

### 2.1 Material

#### 2.1.1 Chemikalien, Geräte und ihre Bezugsquellen

Substanzen	Hersteller/Lieferant
3-1,2,4-Aminotriazol	Sigma
Acrylamid	Roth
Adeninhydrochlorid	Sigma
Agar	Sigma
Agarose	Invitrogen
Ampicillin	Roth
APOD	Roth
ATP	Roche Molecular Biochemicals
Bacto-Trypton	Difco
BSA	NEB New England Biolabs
Coomassie Brilliant Blue G250	Sigma
DMSO	Sigma
dNTPs	AbiPrism
DTT	Roth
EDTA	Roth
Ethidiumbromid	Sigma
Glutathion-Sepharose-Beads	Roth
Glycerin 86%	Roth
GSH	Fluka
Guanidin HCl	Roth
Harnstoff	Serva
IPTG	Saxon
L-[ <sup>35</sup> S]-Methionin (1000Ci/ mMol)	Amersham
Lysozym	Sigma
Natriumazid	Roth
Phenylethyl-β-D-Thiogalactosid	Sigma
Phenol	Roth
Ponceau S	Roth
ProteinA-Agarose/ Sepharose	Sigma
Proteingrößenmarker	New England Biolabs
Röntgen-Fixierlösung	Tenetal
Röntgen-Rapid-Entwickler	Tenetal
SDS	Biomol
β-Mercaptoethanol	Roth
TEMED	Sigma
Tetracyclin	Sigma
Triton-X 100	Sigma
Trockenmilch fettarm	Frema
Tween 20	Sigma
X-Gal	Roth
Yeast Extract	Difco

Kits	Hersteller/Lieferant
BacterioMatch <sup>®</sup> II TwoHybridSystem Vector Kit	Stratagene
BacterioMatch <sup>®</sup> TwoHybridSystem Vector Kit	Stratagene
BigDye <sup>™</sup> Cycle Sequencing Kit	PE Applied Biosystems
ECL <sup>™</sup> Kit	Qiagen
Long Range DNA Polymerase Kit	Eppendorf
pGEM <sup>®</sup> -T-Vector System	Promega
Plasmid Midi Kit	Qiagen
Qiaquick <sup>®</sup> Gel Extraction Kit	Qiagen
RNEasy <sup>®</sup> mini Kit	Qiagen
Taq Polymerase Kit	Eppendorf
TNT <sup>®</sup> Coupled Reticulocyte Lysate System	Promega

Verbrauchsmaterial	Hersteller/Lieferant
Elektroporationsküvetten 4mm	Molecular Bioproducts
Eppendorf Protein LoBind Caps 1,5 ml	Eppendorf
Filterpapier 3MM und 6MM	Whatmann
Microtiterplatten	Greiner
Ni <sup>2+</sup> -NTA-Agarose	Qiagen
Nitrozellulose-Membran Protran BA 85	Schleicher & Schüll
Penta-His Antibody Mouse IgG	Qiagen
Petrischalen 10 cm	Greiner
Polaroidfilme Typ 667	Polaroid
Prestained Protein Marker, Broad Range	New England BioLabs
Reaktionsgefäße (1,5/15/50 ml)	Greiner
Röntgenfilme X-Omat AR <sup>™</sup>	Kodak
Sterilfilter 0,2 µm	Schleicher & Schüll

Enzyme	Hersteller/Lieferant
Calf Intestinal Alkaline Phosphatase (CIP)	New England BioLabs (NEB)
<i>E. coli</i> DNA-Ligase	New England BioLabs (NEB)
Restriktionsendonukleasen	New England BioLabs (NEB)
RNase A	New England BioLabs (NEB)
T4-DNA-Ligase	New England BioLabs (NEB)
T4-Polynukleotidkinase	New England BioLabs (NEB)

Geräte	Hersteller/Lieferant
Brutschrank	Heraeus
ColdLab	Revco Scientific
DNA Sequencer AbiPrism <sup>™</sup> 310	PE Applied Biosystems
Elektroblotting Apparatur	Fröbel
ELISA-Reader MR 5000	Dynatech
French <sup>®</sup> Pressure Cell Press FA-078	SLM Aminco



Geltrockner	Fröbel
Gene Pulser™	BioRad
Gene Quant RNA/ DNA Calculator	Pharmacia
Lyophilisator ALPHA 1-2	Christ
PCR-Gerät	Hybaid
Rotationsschüttler HT	Infors
Schlauchpumpe Cycl I	Roth
Speed Vac Concentrator	Heraeus
Spektralphotometer Hitachi U 2000	Colora
Tischzentrifuge 5415 C	Eppendorf
UV-Schirm	Fröbel
Vortexer REAX 2000	Heidolph
Waage LC 4200 S	Sartorius
Wasserbad	Julabo
Ultraschall-Sonde UW 70	Bandelin
Zentrifuge J2-HS	Beckman
Zentrifuge J6-HC	Beckman

### 2.1.2 Allgemeine Puffer und Lösungen

Lösung/Puffer	Zusammensetzung	
Auftragspuffer, nativ	Sucrose	50 % (m/v)
	EDTA (pH8)	100 mM
	Bromphenolblau	0,1 % (m/v)
	Xylencyanol FF	0,1 % (v/v)
LowTE (pH 8,0)	Tris	10 mM
	EDTA	1 mM
Phenol, verflüssigt	Phenol	1 kg
	H <sub>2</sub> O	250 ml
Phenol, gesättigt	Phenol, verfl.	100 ml
	LowTE	3 ml
Phosphat-Puffer 50 x (pH 7,0)	KH <sub>2</sub> PO <sub>4</sub>	735 mM
	Na <sub>2</sub> HPO <sub>4</sub> x 2 H <sub>2</sub> O	100 mM

### 2.1.3 Nährmedien

#### Flüssige Nährmedien

#### Zusammensetzung

HL5-Medium	Bacto-Pepton	14,3 g/l
	Hefe-Extrakt	7,15 g/l
	Na <sub>2</sub> HPO <sub>4</sub> x 2H <sub>2</sub> O	0,64 g/l
	KH <sub>2</sub> PO <sub>4</sub>	0,49 g/l
	Autoklavieren	
	Glucose	1 %(w/v)
LB-Medium	Bacto-Trypton	10 g/l
	Hefe-Extrakt	5 g/l
	NaCl	10 g/l
	Autoklavieren	
SOC-Medium	Bacto-Trypton	5 g/l
	Hefe-Extrakt	5 g/l
	Glucose	0,1 %(w/v)
YT-Medium	Bacto-Pepton	10 g/l
	Hefe-Extrakt	5 g/l
	NaCl	5 g/l
	Autoklavieren	
YT <sub>Amp</sub>	YT-Medium mit 60 µg/ ml Ampicillin	
LB <sub>Cam</sub>	LB-Medium mit 20 µg/ ml Chloramphenicol	
LB <sub>Tet</sub>	LB-Medium mit 6,25 µg/ ml Tetracyclin	

#### Agar-Nährmedien

#### Zusammensetzung

YT-Agar	Agar	12 g/l
	YT-Medium	
	Autoklavieren	
LB-Agar	Agar	20 g/l
	LB-Medium	
	Autoklavieren	

## 2.1.4 Antikörper

### Primäre Antikörper

Name	Verdünnung	Beschreibung
Anti-Lambda cl	1:2000	polyklonaler Antikörper aus dem Kaninchen gegen eine N-terminale Region (aa 21-36) des Lambda cl-Repressors (Stratagene)
Anti-GST	1:2000	Muriner monoklonaler Antikörper gegen GST (Sigma)
Anti-ORF1 10F11	Überstand	Muriner monoklonaler Antikörper gegen eine C-terminale Region von TRE5A.1-ORF1
Anti-TBP(ep) 6B3	Überstand	Muriner monoklonaler Antikörper gegen das GST-AAA-TBP-Mischepitop („Epitop-Tag“)
Anti-TBP 7A3	Überstand	Muriner monoklonaler Antikörper gegen eine N-terminale Region von DdTBP

### Sekundäre Antikörper

Name	Verdünnung	Beschreibung
Anti-Maus	1:10000	Monoklonaler Anti-Maus IgG (Sigma)
Anti-Kaninchen	1:10000	Polyklonaler Anti-Rabbit IgG (Sigma)

## 2.1.5 Verwendete Organismen

### *Escherichia coli*-Stämme (*E. coli* K12)

Bezeichnung	Genotyp
BacterioMatch <sup>®</sup> Reporterstamm	$\Delta(\text{mcrA})$ 183 $\Delta(\text{mcrCB-hsdSMR-mrr})$ 173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac[F' laqI bla lacZ Kanr]]
BacterioMatch <sup>®</sup> II Reporterstamm	$\Delta(\text{mcrA})$ 183 $\Delta(\text{mcrCB-hsdSMR-mrr})$ 173 endA1 hisB suppE44 thi-1 recA1 gyrA96 relA1 lac[F' laqI q HIS3 aadA Kanr]
BL21 (DE3)	<i>E. coli</i> B, F <sup>-</sup> , dcm, ompT, hsdS(rb <sup>-</sup> mb <sup>-</sup> ), gal $\lambda$ (DE3) [STUDIER <i>et al.</i> , 1990]
DH5 $\alpha$	supE44, $\Delta$ lacU169, ( $\phi$ 80dlacZ $\Delta$ M15), hsdR17, recA1, endA1 gyrA96, thi-1, relA1 [HANAHAN 1983]

RZ1032	F', cat, (pCJ105; m13s, CMR) dut, ung1, thi-1, relA1, spoT1, mcrA (Sambrook 1989)
W3110	F <sup>-</sup> IN (rrnD-rrnE)1 lambda <sup>-</sup> ; (ATCC 27325)
XL1 blue MRF Kan	Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr) 173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac[F' proAB lac <sup>q</sup> ZΔM15 Tn(Kan <sup>R</sup> )]

### ***Dictyostelium discoideum*-Stämme**

<b>Bezeichnung</b>	<b>Genotyp</b>
AX2	Axenisch wachsende Mutante [Sussman & Sussman 1967; Loomis 1982] des Wildtyps NC4 [Raper 1935].

### **2.1.6 Vektoren und Plasmide**

<b>Kommerzielle Vektoren</b>	<b>Beschreibung</b>
pBT (1.1670)	<i>Bait</i> -Vektor des B2H-Systems (3,2 kb), Cam <sup>R</sup> , IPTG-induzierbarer lac-UV5-Promotor, p15A ori, trägt die Sequenz für das λ Repressorprotein λcl inkl. DNA-Bindedomäne (Fa. Stratagene)
pGEM <sup>®</sup> -T	<i>E. coli</i> -Klonierungsvektor (3001 bp), Amp <sup>R</sup> , f1 ori, lacZ, MCS, T7-Promotor, SP6-Promotor (Fa. Promega)
pGEX-5T (1.1893)	Bakterieller, IPTG-induzierbarer Expressionsvektor, ermöglicht die Herstellung N-terminaler 6xHis-GST-markierter rekombinanter Proteine [Berthold et al., 1992]
pMALc2X	IPTG-induzierbarer <i>E. coli</i> -Expressionsvektor (6,6 kb) zur Herstellung von MBP-Fusionsproteinen, die in das Periplasma der Bakterien sezerniert werden. Der Fusionsanteil kann durch Faktor X von MBP getrennt werden (Fa. New England Biolabs)
pTNT (1.1788)	Vektor für <i>in vitro</i> -Expressionen (2871 bp), Amp <sup>R</sup> , f1 ori, verfügt sowohl über einen T7- als auch SP6-Promotor und T7-Terminator (Fa. Promega)
pTRG (1.1671)	<i>Target</i> -Vektor des B2H-Systems (4,4 kb), Tet <sup>R</sup> , lpp/ lac-UV5-Promotor, ColE1 ori, trägt die Sequenz für den N-Terminus von RNAP-alpha (Fa. Stratagene)

Wo nicht ausdrücklich anders beschrieben wurden alle nun aufgelisteten Konstrukte als *NotI*-Fragmente in die einzelnen Vektoren eingeführt. In der Regel wurden die *Inserts* mittels PCR amplifiziert, in pGEM-T zwischenkloniert und anschließend in die *Two Hybrid*-Vektoren pBT bzw. pTRG eingefügt. pTNT-Konstrukte dienen der

Proteinexpression *in vivo*, pMALc2X-Konstrukte der bakteriellen Expression. Die laborinterne Nummer der Plasmide ist in Klammern angegeben.

Vektoren/ Plasmide	Beschreibung
pGEMT-Bdp1	(1.1801) PCR der TFIIIB-Untereinheit Bdp1 an genom. AX2-DNA über Bdp1 1/5 und Bdp1 2/6
pGEMT-BrfC	(1.1674) PCR des C-Terminus der TFIIIB-Untereinheit Brf an genom. AX2-DNA mit Brf-10 u. -11
pGEMT-BrfN	(1.1694) Klonierung des N-Term. von Brf als PCR-Produkt an genom. AX2-DNA mit Brf-12/13 in pGEM T
pGEMT-BrfNC	(1.1707) Klonierung der TFIIIB-Untereinheit BrfNC als PCR-Produkt an genom. AX2-DNA mit Brf-11/12 in pGEMT
pGEMT-EN	(1.1751) PCR der ORF2-Untereinheit EN an pB3 mit DRE EN-01/02, als NotI-Fragment in pGEMT kloniert
pGEMT-EN(mut)	(1.1786) Einzelstrangmut. an pGEMT-EN mit Rep81; Inaktivierung d. Enzyms
pGEMT-GST	(1.1860) PCR an pGEX-5T mit GSTfor/ back; GST als EcoRI-Fragment in pGEMT
pGEMT-HC	(1.1752) PCR der ORF2-Untereinheit HC an pB3 mit DRE HC-01/02
pGEMT-hTBP	(1.1658) PCR von hTBP an cDNA RS4;11 (JZ) mit hTBP-01/02
pGEMT-hTBP(dH2') 09	(1.1942) Einführung v. dH1' in hTBP(dH2') mit Primer hTBP-09
pGEMT-hTBP(dH2') 10	(1.1943) Einführung d. authentischen C-Terminus v. DdTBP in hTBP(dH2') mit Primer hTBP-10
pGEMT-hTBP(dH2')	(1.1924) Einführung v. dH2' in hTBP02/03 mit Mutagenese-Primer hTBP-08
pGEMT-hTBP02/03	(1.1676) PCR von hTBP an cDNA RS4;11 (JZ) mit hTBP-02/03, entfernt N-terminalen Poly-Glu-Stretch
pGEMT-hTBP06/07	(1.1909) PCR an pGEMT-hTBP02/03 mit Primer hTBP-06/07; hH2' als NotI-Fragment in pGEMT
pGEMT-ORF1(Not)	(1.1757) PCR an pB3 mit Primer DRE ORF1-01/02
pGEMT-RT(ATG, tag)	(1.1808) PCR an pB3 mit Primer Rep66/65 ergibt EcoRI/Xho-Fragment (1.1036), darin Einführung einer EcoRI-Site + Stop-Codon an 3' mit Primer pGEMT-RT (tag) sowie Einführung eines ATG an 5' mit Primer RT-mut2, dadurch Zerstörung EcoRI-Site an 5'-Ende. RT(ATG, tag) als Eag-Fragment in pGEMT
pGEMT-TBP	(1.1673) PCR der vollständigen kodierenden Sequenz von TBP an cDNA von AX2 als NotI-Fragment mit Primer TBP10/11
pGEMT-TBP(Not108)	(1.1930) Einführung einer NotI-Site an Pos 108 in pGEMT-TBP mittels ss-Mutagenese mit Primer TBP-15
pGEMT-TBP(Not128)	(1.1931) Einführung einer NotI-Site an Pos 108 in pGEMT-TBP mittels ss-Mutagenese mit Primer TBP-14
pGEMT-TBP(S195A)	(1.1929) Austausch Ser gg. Ala an Pos 195 im vollständigen TBP mittels ss-Mutagenese mit Primer TBP-13
pGEMT-TBP09/11 (S195A)	(1.1938) Austausch Ser gg. Ala an Pos. 195 in Helix H2' mit ss-Mutagenese-Primer TBP-13
pGEMT-TBP09/11 mut12	(1.1913) Austausch Ser gg. Lys an Pos 195 in Helix H2' von TBP mittels ss-Mutagenese mit Primer TBP-12
pGEMT-TBP10/11 mut12	(1.1912) Austausch Ser gg. Lys an Pos 195 im vollständigen TBP mittels ss-Mutagenese mit Primer TBP-12
pBT-Bdp1	(1.1822) kodierende Sequenz v. Bdp1 als Not-Fragment aus pGEMT-Bdp1 in pBT

pBT-BrfC(439-706)	(1.1672) Isolierung d. Fragments aus pTNT-BrfC(Not435) mit NotI und Klonierung in pBT
pBT-BrfN	(1.1722) N-terminales Fragment von Brf1 (1-270) in pTRG kloniert
pBT-BrfNC	(1.1734) vollst. kodierende Sequenz von Brf1 aus pGEMT in pBT kloniert
pBT-EN(mut)	(1.1794) EN(mut) aus pGEMT in pBT kloniert
pBT-HC	(1.1799) HC aus pGEMT in pBT kloniert
pBT-hTBP(dH2')	(1.1928) Klon. der Chimäre hTBP(dH2') als NotI-Fragment aus pGEMT in pBT
pBT-hTBP02/03	(1.1696) N-terminal trunkiertes humanes TBP aus pGEM-T in pBT eingefügt
pBT-ORF1	(1.1779) vollst. kodierende Sequenz von TRE5-A.1 ORF1 aus pGEMT in pBT kloniert
pBT-ORF1(K277)	(1.1910) Klon. von ORF1(1-276) als NotI-Fragment aus pTNT in pBT
pBT-ORF1(L159)	(1.1905) Klon. von ORF1(1-158) als NotI-Fragment aus pTNT in pBT
pBT-RT(ATG tag)	(1.1875) Klon. v. RT(ATG, tag) als Eag-Fragment aus pGEMT in pBT
pBT-TBP	(1.1821) vollst. kodierende Sequenz von TBP aus pGEMT in pBT kloniert
pBT-TBP(1-108)	(1.1933) Isolierung d. Fragments aus pGEMT-TBP (Not 108) mit NotI u. Klonierung d. linken Teildomäne v. TBP in pBT
pBT-TBP(1-128)	(1.1937) Isolierung d. Fragments aus pGEMT-TBP (Not 128) mit NotI u. Klonierung d. N-terminalen Fragments (Helices H1 u. H2 bis direkt vor H1') in pBT
pBT-TBP(113-205)	(1.1934) Isolierung d. Fragments aus pGEMT-TBP (Not 108) mit NotI u. Klonierung d. C-terminalen Fragments (Helices H1' u. H2', d.h. rechten Teildomäne v. TBP) in pBT
pBT-TBP(1-136)	(1.1806) Isolierung d. Fragments aus pGEMT-TBP (Not 128) mit NotI u. Klonierung d. N-terminalen Fragments (Helices H1, H2 u. H1') in pBT
pBT-TBP(137-205)	(1.1710) Isolierung d. Fragments aus pGEMT-TBP (Not 128) mit NotI u. Klonierung d. C-terminalen Fragments (H2'+C-Terminus) in pBT
pBT-TBP(m3)	(1.1883) TBP(m3) aus pTNT in pBT kloniert
pBT-TBP(S195A)	(1.1935) vollst. TBP mit Ser gg. Ala-Austausch in Pos 195 aus pGEMT in pBT kloniert
pBT-TBP09/11 mut12	(1.1918) isolierte Helix dH2'TBP mit Ser gg. Lys-Austausch in Pos 195 aus pGEMT in pBT kloniert
pBT-TBP09/11	(1.1789) isolierte Helix dH2'TBP aus pGEMT in pBT kloniert
pBT-TBP09/11(S195A)	(1.1939) isolierte Helix dH2'TBP mit Ser gg. Ala-Austausch in Pos 195 aus pGEMT in pBT kloniert
pBT-TBP10/11 mut12	(1.1917) vollst. TBP mit Ser gg. Lys-Austausch in Pos 195 aus pGEMT in pBT kloniert
pGEX 5T(NotI)	(1.1900) Einführung einer NotI-Site in Originalvektor mittels Linkerligation mit Primer Bam-Not-Bam, ergibt GST-Fusionsproteine mit einem zusätzlichen Linker aus 3 Alaninen zw. GST und Fusionsanteil
pGEX(NotI)-Bdp1	(1.1807) Klonierung d. Fragments Bdp1 in Expressionsvektor pGEX-5T(Not) aus pGEMT-Bdp1
pGEX(NotI)-BrfC	(1.1736) Klonierung d. Fragments BrfC in Expressionsvektor pGEX-5T(Not) aus pGEMT-BrfC
pGEX(NotI)-BrfNC	(1.1746) Klonierung v. Brf (volle Länge) in Expressionsvektor aus pGEMT-BrfNC
pGEX(NotI)-ORF1	(1.1741) Klonierung v. ORF1 in Expressionsvektor aus pGEMT-ORF1
pGEX(NotI)-TBP	(1.1660) Klonierung d. vollständigen TBP in Expressionsvektor aus pGEMT-TBP

pMALc2x(NotI)	(1.1904) Einführung einer NotI-Site in MCS des Vektors mit Primer pMAL-Not(neu), ergibt MBP-Fusionsproteine mit einem zusätzlichen Linker aus 3 Alaninen zw. MBP und Fusionsanteil
pMALc2x(NotI)-ORF1	(1.1908) ORF1 aus pGEMT in pMALc2X für bakterielle Expressionen
pMALc2x(NotI)-ORF1(K277)	(1.1926) ORF1(K277) aus pTNT in pMALc2X für bakterielle Expressionen
pMALc2x(NotI)-ORF1(L159)	(1.1925) ORF1(L159) aus pTNT in pMALc2X für bakterielle Expressionen
pMALc2x(NotI)-ORF1(S405)	(1.1927) ORF1(S405) aus pTNT in pMALc2X für bakterielle Expressionen
pMALc2x(NotI)-TBP	(1.1907) vollständiger TBP aus pGEMT in pMALc2X für bakterielle Expressionen
pMALc2x-TBP(ep)	(1.1932) ss-Mut. mit Primer pMAL-TBP(ep) an pMALc2X(NotI)-TBP zur Einführung einer GST-Sequenz in C-Terminus v. MBP
pMAL-TBP(P384I)	(1.1941) Austausch Pro gg. Ile an Pos. 384 in pMAL-TBP(ep) mit ss-Mutagenese-Primer pMAL-TBP(P384I); N-terminale Begrenzung d. Epitops
pMAL-TBP(T396P)	(nur Minilysat) Austausch Thr gg. Pro an Pos. 396 in pMAL-TBP(ep) mit Primer pMAL-TBP(T396P); C-terminale Begrenzung d. Epitops
pTNT-Bdp1	(1.1816) Klonierung der TFIIIB-Untereinheit Bdp1 als Not-Fragment in pTNT
pTNT-BrfC	(1.1766) Brf(271-706) aus pGEM-T in pTNT für in vitro Expressionen
pTNT-BrfC(439-535)	(1.1903) Einführung einer NotI-Site in pTNT-BrfC(Not435) mit Primer Brf(439-535) ergibt bei NotI-Verdau das Fragment BrfC(439-535)
pTNT-BrfC(439-635)	(1.1902) Einführung einer NotI-Site in pTNT-BrfC(Not435) mit Primer Brf(439-635) ergibt bei NotI-Verdau das Fragment BrfC(439-635)
pTNT-BrfC(Not435)	(1.1662) Einführung einer NotI-Site an AS439 in BrfC zur Fragmentierung, Primer BrfC(Not435)
pTNT-BrfN	(1.1767) Brf(1-270) aus pGEM-T in pTNT für in vitro Expressionen
pTNT-BrfNC	(1.1758) Klonierung der vollständigen TFIIIB-Untereinheit Brf1 als NotI-Fragment in pTNT
pTNT-EN(mut)	(1.1803) Klonierung v. EN(mut) als Not-Fragment in pTNT
pTNT-GST	(1.1866) Einführung eines GST-Tags als EcoRI-Fragment in pTNT, PCR an pGEX-5T mit Primer GSTfor/ back
pTNT-GST-ORF1	(1.1878) Klonierung v. ORF1 in pTNT-GST zur Gewinnung d. Fusionsproteins GST-ORF1 für Pulldowns
pTNT-GST-TBP	(1.1890) Klonierung v. TBP in pTNT-GST zur Gewinnung d. Fusionsproteins GST-ORF1 für Pulldowns
pTNT-ORF1	(1.1802) vollständiger ORF1 aus pGEM-T-Konstrukt in pTNT für in vitro Expressionen
pTNT-ORF1(K277)	(1.1892) Einführung einer NotI-Site+Stop-Codon mittels Einzelstrangmutagenese und Primer ORF1(K277) an pTNT-ORF1
pTNT-ORF1(S405)	(1.1695) Einführung der AS Ser als Stopcodon an Pos. 405 mittels Einzelstrangmutagenese und Primer ORF1(S405) an pTNT-ORF1
pTNT-ORF1(W108)	(1.1922) Einführung einer NotI-Site+Stop-Codon mittels Einzelstrangmutagenese und Primer ORF1(W108) an pTNT-ORF1(L159)
pTNT-RT(ATG, tag)	(1.1854) RT(ATG, tag) als Eag-Fragment aus pGEMT in pTNT für in vitro Expressionen
pTNT-TBP	(1.1768) vollständiger TBP aus pGEM-T-Konstrukt in pTNT für in vitro Expressionen

pTNT-TBP(m3)	(1.1675) ss-Mutagenese an pTNT-TBP mit Primer TBPm3, Austausch von drei für die WW mit BrfC essentiellen Arginine
pTNT-TBP(Not137)	(1.1446) Einführung einer NotI-Site zur Fragmentierung v. TBP mit Primer TBP(Not137) an pTNT-TBP
pTNT-TBP10/11 mut12	(1.1921) TBP10/11 mit AS-Austausch S195K aus pGEMT als NotI-Fragment in pTNT für in vitro Expressionen
pTGR-ORF1(B)	(1.1940) TRE5-B ORF1 als NotI-Fragment aus pGEMT-ORF1(B) isoliert und in pTRG eingefügt
pTRG-Bdp1	(1.1617) kodierende Sequenz v. Bdp1 als Not-Fragment aus pGEMT-Bdp1 aus pGEMT in pTRG
pTRG-BrfC	(1.1730) C-Terminus v. Brf1 als NotI-Fragment aus pGEMT in pTRG kloniert
pTRG-BrfN	(1.1731) N-terminales Fragment von Brf1 (1-270) in pTRG kloniert
pTRG-BrfNC	(1.1759) vollständige kodierende Sequenz v. BrfNC als NotI-Fragment aus pGEMT in pTRG
pTRG-EN(mut)	(1.1795) Klon. von EN(mut) als NotI-Fragment in pTRG aus pGEMT
pTRG-HC	(1.1798) Klon. von HC als NotI-Fragment in pTRG aus pGEMT
pTRG-ORF1	(1.1796) vollst. kodierende Sequenz v. TRE5-A.1 ORF1 als NotI-Fragment aus pGEMT in pTRG
pTRG-ORF1(112-158)	(1.1944) Fragment mit NotI aus pTNT-ORF1(W108) isoliert und in pTRG eingefügt
pTRG-ORF1(402-444)	(1.19xx) C-terminale Coil-Struktur von ORF1 aus pGEMT
pTRG-ORF1(K277)	(1.1920) ORF1(K277) als NotI-Fragment aus pTNT in pTRG kloniert
pTRG-ORF1(L159)	(1.1906) ORF1(K277) als NotI-Fragment aus pTNT in pTRG kloniert
pTRG-ORF1(S405)	(1.1805) ORF1(S405) als NotI-Fragment aus pTNT in pTRG kloniert
pTRG-ORF1(W108)	(1.1936) ORF1(W108) als NotI-Fragment aus pTNT in pTRG kloniert
pTRG-RT(ATG tag)	(1.1874) RT(ATG, tag) als EagI-Fragment aus pGEMT in pTRG kloniert
pTRG-TBP	(1.1703) vollst. kodierende Sequenz v. TBP in pTRG aus pGEMT kloniert

## 2.1.7 Synthetische Oligonukleotide

Alle für Klonierungen, Einzelstrangmutagenesen und Sequenzierungen verwendeten Oligonukleotide wurden von den Firmen MWG oder Metabion synthetisiert.

### 2.1.7.1 Oligonukleotide, die im Vektor binden

Name	Sequenz (5' → 3')
Bam Not Linker	GATCCGCGGCCGCG
GST back	GGAATTCATGTCACGATGAATTGGTACCTCGAGTC
GST for	GAATTCGCAATGGATTCCCCTATACATCACCATCACC
pBT_3' seq	CAGCATCCTGCGATGCAGATCC
pBT-seq	CCACAGTACCCAATGATCCCATG
pGEM-Not	GCATGCTCGCGGCCGCCATGG
pGEX Seq	GTATTGAAGCTATCCCACAAATTGATAAG
pMAL-Not(neu)	GATCGAGGGAAGGGCGGCCGCGCATTTCGGATCCTCTAGAGTCGACC



pMAL-seq	GGTCGTCAGACTGTCGATGAAGCCC
pMAL-TBP(ep)	CAACAACCTCGGGCCCCGGGGAAGTGCTGCCGCAATGAGCACC
pMAL-TBP(P384I)	CAACAACCTCGGGATCCGGGGAAGTGCTGC
pMAL-TBP(T396P)	CACCGCAACCCCTACATCTACACCAGCAC
pTNT-seq	CGACCACTATAGGCTAGCATTTAG
pTNT-seq	CGACTCACTATAGGCTAGCATTTAG
pTRG-seq	GCAACCATTCTGGCTGAACAACCTG
RSP	CGGATAACAATTCACACAGG
SP	CGCCAGGGTTTTCCAGTCAC

### 2.1.7.2 Oligonukleotide, die in TRE-Elementen binden

Name	Sequenz (5' → 3')
DRE EN-01	GGCGGCCGCAATGGATTCTACCCCAAATAAACTATAAAG
DRE EN-02	GGCGGCCGCTTAAAGGAAGATTGCTCTTGAATCTTTGGC
DRE HC-01	GGCGGCCGCAATGGAAAGGTACCTCGGCTTTGACTTCAAC
DRE HC-02	GGCGGCCGCTTATTTATAGAAAGTACTGAGTGAATTATTG
DRE ORF1-01	GGCGGCCGCGAGCCGCCATGAAAATCACAAACAACACCACAAATATTAAG
DRE ORF1-02	GGCGGCCGCCTTATAGTTTTATTTGGGGTAGATTTTTATCTACC
DRE RT-01	GGCGGCCGCAATGGAAACAGCCAAAGATTCAAGAGCAATCTTCC
DRE RT-02	GGCGGCCGCTTAAACCCTTATTGTTGAAGTCAAAGCCGAGG
DRE-RT-01Eco	GATGATTGAATTCGGCGGCCGCAATGGAAACAGCCAAAGATTCAAGA GCAATCTTCC
DRE-RT-02Eco	GGTTGAATTCGGCGGCCGCTTAAACCCTTATTGTTGAAGTCAAAGC CGAGG
ORF1 (S405*)	GTGGGGTTTTTTCATTTACTGATTTGTTTTAG
ORF1(K277*)	CTCTTTGTTTCGCGGCCGCGTCATATTGGTGTATTGTATGG
ORF1(L159*)	CCTTTGTTGGAGCGGCCGCTTCATGTTAGATTGTCTAGTTC
ORF1(W108*)	CCATTCTATTTCTGTGAGTGCGGCCGCTAATCGAAATCCGTATAC
pGEM-RT(tag)	GATACCCTTAGGCTAGAATTCAAAGCCGAGG
RT-mut2	CTTTGGCTGTTTCCATTTCCCGAATCA

### 2.1.7.3 Oligonukleotide, die in DdTFIIIB binden

Name	Sequenz (5' → 3')
Brf-10	GGCGGCCGCAATGGATTTAGAGGATCAATGTGATCCACC
Brf-11	GGCGGCCGCTTAAATAATCATCTTCTTCATAGTTATTAATAATTACGAC
Brf-12	GGCGGCCGCAATGGATTCAAAAAATTGTGGAAATTG
Brf-13	GGCGGCCGCTCAACCATTGAATTCATCGAATCTAAG
Brf-14	CCTTAATAGATTGAACATTTTGTGAACC
BrfC(439-535)	CTTCATCACTTAAAGCGGCCGCTCAATCATTTGGTGCAAAAACCTG
BrfC(439-635)	CCCCACCAAATGCGGCCGCTCAAACGAAATTTGGTTTGAAG
dBdp1-01	GGCGGCCGCAATGGATAAAGAAAAATCAGATAATGA

dBdp1-01Eco	GATGATTGAATTCGGCGGCCGCAATGGATAAGAAAAATCAGAT AATGAAGTTAATG
dBdp1-02	GGCGGCCGCTTAATATTCATAATCATTGTATTATA
dBdp1-02Eco	GGTTGAATTCGGCGGCCGCTTAATATTCATAATCATTGTATTATAA TTATCATTTTCC
dBdp1-03	GGTGATGCACCCGTATTAGAAGATGG
dBdp1-05	GGTCTGAAGCTGAAACTCAAGACTTTTTTTATTCAATTA AAAAGATATGG TACAGATTTTTTC",
dBdp1-06	GAAAAATCTGTACCATATCTTTTTAATTGAATAAAAAAGTCTTGAGTTTCAG CTTCAGACC
dBdp1-07	GTTTTTATTGTTATCGTCATCATCATCATC
dBdp1-08	GTGGTATTAGTGGTGCTTATCGTCCTGCAG
dBrf (Not 435)	CTTTTCAATATTCGCGGCCGCTGTTGTTGGAATTGAATCTACATTTG
dTBP(Not 137)	CATAATTGGTGAAAGATGCGGCCGCATTATGTAAGAGTTC
dTBP-07	GGCGGCCGCAATGGTTTTGCACTGGAGCAAAATCTG
dTBP-08	GGCGGCCGCTATGCAGGGAAATCCAATTTTTGAATG
dTBP-09	GGCGGCCGCTGGTGCTAAAGTTCGTGAAGA
dTBPm3	CCAATTTTTGAATGATACTTGCATATTTTTCTGCAGCAAATTCACTTGCA TCTTCAG
TBP-10	GGCGGCCGCAATGAGCACCGCAACCACTACATC
TBP-11	GGCGGCCGCTTATTGAGTAATTGCATTAACTTTTT
TBP-12	TTTATCCTGTACTTAAGGCCTTTAAAAAAGTTAATG
TBP-13	TTTATCCTGTACTTGCGGCCTTTAAAAAAGTTAATG
TBP-14	GTGATGTCAAATCTAGCGGCCGCTAGAACTCTTACATAATG
TBP-15	CAAAAATTGGATTTCTAGCGGCCGCTACAGATTTTAAGATTC
TBP-16	CAAATTAGAACTCTTAGCAGCAGCACATACATCTTTCACC

#### 2.1.7.4 Oligonukleotide, die in HsTBP binden

Name	Sequenz (5' → 3')
hTBP-01	GGCGGCCGCAATGGATCAGAACAACAGCCTGCCACCTTAC
hTBP-02	GGCGGCCGCTTACGTCGTCTTCCTGAATCCCTTTAGAATAG
hTBP-03	GGCGGCCGCAATGGGCACCACTCCACTGTATCCCTC
hTBP-04	GTTGGTGGGTGAGCACAAGGCCTTC
hTBP-05	GATGCCTTATGGCACTGGACTGACC
hTBP-06	GGCGGCCGCTGGTGCTAAAGTCAGAGCAG
hTBP-07	GGCGGCCGCTTACGTCGTCTTCCTGAATCC
hTBP-08	GAAGCATTTGAAAACATCTACCCTGTACTTTCCGCCTTTAAAAAAGTTAAT TAAGCTTGCGGCCGCCATCACTAGTG
hTBP-09	CTATAAGGTTAGAACTCCTTCACAACGCCCAACAATTTAGTAGTTAT GAGC
hTBP-10	CTTTCGCCTTTAAAAAAGTTAATGCAATTA CTCAATAACATCACTAGTGC

## 2.2 Molekularbiologische Methoden

### 2.2.1 Standardmethoden

Molekularbiologische Standardmethoden wurden, wenn nicht anders erwähnt, nach Sambrook *et al.* [SAMBROOK *et al.*, 1989] durchgeführt. Molekularbiologische Kits wurden entsprechend den Herstellerangaben verwendet.

### 2.2.2 Sequenzanalysen

Sequenzierungen wurden unter Verwendung des ABI Prism<sup>®</sup> BigDye<sup>™</sup> Cycle Sequencing Kits nach der Kettenabbruchmethode durchgeführt [SANGER *et al.*, 1977]. Dabei wurde die zu analysierende DNA in einer PCR mit unterschiedlichen Fluoreszenzfarbstoffen markiert:

Primer	10 pmol
DNA	200-500 ng
Ready Reaction Mix	2 $\mu$ l
MilliQ Wasser	ad 5 $\mu$ l

Das Temperatur-Zeit-Schema der PCR:

Initiale Denaturierungsphase	96°C	2 min	} x25
Denaturierungsphase	96°C	20 s	
Hybridisierungsphase	56°C	10 s	
Synthesephase	60°C	4 min	

Die Probe wurde mit Wasser auf 50  $\mu$ l aufgefüllt und mittels EtOH-Fällung gereinigt und aufkonzentriert. Nach Aufnahme in 15  $\mu$ l *Template Suppression Reagent* (TSR) wurden die DNA-Fragmente im ABI Prism<sup>®</sup> 310 Genetic Analyzer mittels Kapillarelektrophorese aufgetrennt und die Sequenz automatisch ausgelesen.

## 2.3 Bakterielle *Zwei*hybrid-Versuche

Alle verwendeten Lösungen wurden, sofern kein Autoklavieren der fertigen Zubereitung möglich war, sterilfiltriert.

<b>Lösung</b>	<b>Zusammensetzung</b>	
10x HisDO	HisDO	7,7 g/l
	Autoklavieren 15 min, 121°C	
3-AT 1 M	3-AT	841 mg
	DMSO	10 ml
10x M9-Salze	NH <sub>4</sub> Cl	10 g/l
	Na <sub>2</sub> HPO <sub>4</sub> x2H <sub>2</sub> O	75 g/l
	KH <sub>2</sub> PO <sub>4</sub>	30 g/l
	NaCl	50 g/l
	Autoklavieren	
M9 Additivum (für 500 ml-Ansatz)	Glukose 20%	10 ml
	Adenin 20 mM	5 ml
	10x HisDO	50 ml
	MgSO <sub>4</sub> 1 M	0,5 ml
	Thiamin-HCl 1 M	0,5 ml
	ZnSO <sub>4</sub> 10 mM	0,5 ml
	CaCl <sub>2</sub> 100 mM	0,5 ml
	IPTG 50 mM	0,5 ml
<b>Nähr- und Selektionsmedien</b>	<b>Zusammensetzung</b>	
HisDO-Agar	Agar	15 g/l
	VE-Wasser	760 ml
	Autoklavieren	
	10xM9-Salze	100 ml
	M9 Additivum	135 ml
	Cam 25 mg/ ml	1 ml
	Tet 5 mg/ ml	2,5 ml
3-AT-5mM-Agar	HisDO-Agar	1000 ml
	3-AT 1 M	5 ml
Strep12-Agar	HisDO-Agar	1000 ml
	Streptomycin 10 mg/ml	1,2 ml
LB <sub>Cam</sub> 34	LB-Agar	1000 ml

	Cam 25 mg/ml	1,36 ml
LB <sub>Cam</sub> 20	LB-Medium	250 ml
	Cam 25 mg/ml	200 µl
LB <sub>Tet</sub> 12,5	LB-Agar	1000 ml
	Tet 5 mg/ml	2,5 ml
LB <sub>Tet</sub> 6,25	LB-Medium	250 ml
	Tet 5mg/ ml	0,313 ml

In dem *BacterioMatch*<sup>TM</sup> *Two Hybrid*-System (Stratagene) erfolgt ähnlich wie bei Hefe-Zweihybrid-Systemen die Identifizierung von Interaktionen durch die transkriptionelle Aktivierung von Reporter genen.

Der Reporterstamm trägt die Reporter gene *aadA*, welches die Streptomycin-Resistent vermittelt, und *HIS3* unter der Kontrolle des  $\lambda$ -Operators. Expression des *HIS3*-Gens ermöglicht Bakterien das Wachstum auf Histidin-Mangelmedium. Um die Expression der Reporter gene zu aktivieren, muss das  $\lambda$ cl-Repressorprotein des Bakteriophagen  $\lambda$  an den Operator und die RNA-Polymerase an den Promotor binden. Die gleichzeitige Bindung von Repressorprotein und Polymerase erfolgt jedoch nur dann, wenn sie mit zwei interagierenden Proteinen (*Bait* und *Target*) fusioniert sind und im Reporterstamm koexprimiert werden.

Zur Herstellung dieser Fusionsproteine wurde ein Protein (*Bait*) durch Klonierung in den pBT-Vektor mit dem  $\lambda$ cl-Repressorprotein fusioniert, das N-terminal eine DNA-Bindedomäne und C-terminal eine Dimerisierungsdomäne enthielt. Das andere Protein (*Target*) wurde in den pTRG-Vektor kloniert und auf diese Weise an die N-terminale Domäne der  $\alpha$ -Untereinheit der RNA-Polymerase fusioniert.

Für die Dimerisierungsdomäne des Transkriptionsfaktors Gal4 aus Hefen und die Gal11<sup>P</sup>-Domäne, einer Mutante des Gal11-Proteins, konnte eine Interaktion bereits nachgewiesen werden. Diese Polypeptide werden von den beiden Kontrollplasmiden pBT-Gal4 und pTRG- Gal11<sup>P</sup> kodiert.

### 2.3.1 Herstellung der Konstrukte

Zielvektor und Insert wurden im Verhältnis 1:10 einer Restriktionshydrolyse mit *NotI* unterworfen und nach Aufreinigung aus dem Gel zwei Tage bei 12°C ligiert. Anschließend wurde die DNA mittels EtOH-Fällung gereinigt und in 10 µl MilliQ-Wasser aufgenommen. Von diesem Ansatz wurde ein Mikroliter in 50 µl elektrokompetente XL-1 blue MRF Kan-Zellen transformiert (2,5 kV, 25 µF). Die Transformationsreaktion wurde 30 min bei 37°C inkubiert. pBT-Konstrukte wurden auf LB<sub>Cam34</sub>-, pTRG-Konstrukte auf LB<sub>Tet</sub>12,5-Platten ausgebracht und zwei Tage bei 30°C bebrütet. Zur Überprüfung der korrekten Orientierung des Fragments, also *in frame* mit dem  $\lambda$ cl- (pBT) bzw. dem RNAP $\alpha$ -Protein (pTRG), wurden erst eine Restriktionshydrolyse und anschließend eine Sequenzierung durchgeführt. Dafür verwendete Minikulturen enthielten 20 µg/ml Chloramphenicol bzw. 6,25 µg/ml Tetracyclin und wurden bei 30°C inkubiert. Da es sich sowohl bei pBT als auch bei pTRG um *low copy*-Vektoren handelt, wurden die anschließenden Plasmidpräparationen nur in der Hälfte des üblichen MilliQ-Volumens aufgenommen.

### 2.3.2 Doppeltransformation und Selektion

Je ein Mikrogramm des pBT- und pTRG-Konstruktes wurden gemeinsam in 50 µl elektrokompetente BacterioMatch II Reporterzellen (2,5 kV und 25 µF) transformiert. Der Reaktionsansatz wurde 30 min bei 37°C inkubiert, auf nichtselektive Histidin-*dropout*-Platten (HisDO) ausgebracht und 24 Stunden bei 37°C bebrütet. Am folgenden Tag wurden pro Doppeltransformation zehn Klone mittels Impföse auf eine HisDO-Masterplatte übertragen. Von der Masterplatte wurden die Klone anschließend auf HisDO-Selektionsplatten mit 3-Amino-1,2,4-Triazol (3-AT, Endkonzentration 4-6 mM) überimpft und nach 24 Stunden bei 37°C auf Wachstum untersucht. Jede Doppeltransformation wurde mindestens drei Mal durchgeführt. Aus positiv getesteten Klonen wurden die Plasmide wieder isoliert und mittels Restriktionshydrolyse und Sequenzierung eine Verwechslung der Konstrukte ausgeschlossen. Zusätzlich wurden die Ergebnisse auf Streptomycinplatten (12 µg/ml) verifiziert. Um Eigenaktivierungen auszuschließen sowie als optischer „Leerwert“ wurde jeweils eine Negativkontrolle mitgeführt. Dazu wurde das zu testende Konstrukt entweder mit dem Leervektor pBT oder pTRG in die

Reporterzellen eingebracht. Von ausgesuchten Doppeltransformatanten wurden Glycerin-Stocks hergestellt und bei  $-80^{\circ}\text{C}$  gelagert.

## 2.4 Interaktionsanalyse *in vitro*

### 2.4.1 Expression des rekombinanten Fusionsproteins MBP-ORF1

Säulenpuffer	Tris (pH 7,5)	20 mM
	NaCl	300 mM
	DTT	2 mM
	EDTA	1 mM
	$\beta$ -ME	10 mM

Für die Expression des Fusionsproteins wurde TRE5-A ORF1 als *NotI*-Fragment in den Vektor pMALc2x (New England Biolabs) kloniert. Dazu musste die entsprechende Schnittstelle durch Einzelstrangmutagenese in den Vektor eingeführt werden, allerdings wurde dadurch die vektorkodierte Schnittstelle für Fusionsproteine (Faktor X) zerstört. Zur Unterscheidung vom Originalvektor wurde der modifizierte Vektor pMALc2x(Not) genannt.

Das Konstrukt pMALc2x(Not)-ORF1 wurde in den Bakterienstamm W3110 transformiert. Eine 50 ml  $\text{YT}_{\text{amp}}$ -Kultur, die zusätzlich 0,2% Glukose enthielt, wurde mit 1 ml einer Übernachtskultur angeimpft und bei  $37^{\circ}\text{C}$  bis zu einer  $\text{OD}_{600}$  von 0,6 geschüttelt. Durch Zugabe von 0,3 mM IPTG wurde die Expression des Fusionsproteins induziert. Sowohl vor der Zugabe von IPTG als auch nach der Induktion wurde jeweils ein kleines Aliquot der Kultur entnommen und die Expression mittels SDS-PAGE überprüft. Nach zweistündiger Inkubation bei  $37^{\circ}\text{C}$  wurden die Zellen pelletiert ( $4^{\circ}\text{C}$ , 5000 UpM, JA10 Rotor Beckman), zweimal mit kaltem PBS gewaschen und anschließend erneut zentrifugiert. Pro 50 ml Kultur wurden die Pellets in einem Milliliter kalten Säulenpuffer aufgenommen und bis zu ihrer weiteren Verwendung bei  $-20^{\circ}\text{C}$  gelagert.

Für Kontrollversuche wurden Bakterienzellen mit pMALc2x(Not) transformiert und MBP unter den gleichen Bedingungen exprimiert.

### 2.4.2 *In vitro* Transkription und Translation

Zur Expression  $^{35}\text{S}$ -markierter Proteine wurde der *TNT<sup>®</sup> T7 Coupled Reticulocyte Lysate Kit* (Promega) nach den Angaben des Herstellers verwendet. Die radioaktive Markierung erfolgte durch Zusatz von  $2\ \mu\text{l}$  [ $^{35}\text{S}$ ]-Methionin (Amersham) entsprechend  $20\ \mu\text{Ci}$ . Vom  $50\ \mu\text{l}$ -Ansatz wurden  $5\ \mu\text{l}$  als *Input*-Kontrolle abgenommen, der Rest wurde komplett für einen *Pulldown*-Versuch verwendet. Es wurden ausschließlich *Protein-LoBind*-Reaktionsgefäße von Eppendorf verwendet.

### 2.4.3 MBP-*Pulldown*-Assay

Für die *Pulldown*-Versuche wurden *Amylose-Beads* jeweils frisch mit MBP-ORF1p bzw. MBP beladen. Dazu wurden die gefrorenen Bakteriensuspensionen auf Eis aufgetaut, anschließend mit  $5 \times 10$  Pulsen sonifiziert und unlösliche Zellbestandteile durch Zentrifugation ( $20\ \text{min}$ ,  $14000\ \text{UpM}$ ,  $4^\circ\text{C}$ , Tischzentrifuge Eppendorf) abgetrennt. Da die Expressionsstärke von MBP-ORF1p deutlich schwächer war als die von MBP und MBP-ORF1p zusätzlich mit einer geringeren Affinität an die Säule band, wurde für das Fusionsprotein der gesamte lösliche Überstand verwendet, während für MBP lediglich  $10\ \mu\text{l}$  genommen und mit Säulenpuffer auf das ursprüngliche Volumen ergänzt wurden. Der Überstand wurde in ein frisches Cap überführt und am Überkopffrotor  $30\ \text{min}$  bei  $4^\circ\text{C}$  mit  $100\ \mu\text{l}$  äquilibrierten *Amylose-Beads* inkubiert. Ungebundenes Protein wurde durch dreimaliges Waschen mit je  $500\ \mu\text{l}$  kaltem Säulenpuffer entfernt. Aufgenommen wurden die beladenen Säulen in  $200\ \mu\text{l}$  Säulenpuffer.  $10\ \mu\text{l}$  dieser Suspension wurden mittels SDS-PAGE und Coomassie-Färbung auf ausreichende Beladung kontrolliert.

Präzipitiertes MBP-ORF1 bzw. MBP wurde  $2\ \text{h}$  zusammen mit *in vitro* translatiertem, radioaktiv markiertem Protein auf der niedrigsten Stufe des Vortexers bei  $4^\circ\text{C}$  inkubiert. Nach dreimaligem Waschen mit je  $1\ \text{ml}$  Säulenpuffer, wobei die *Beads* bei jedem Waschdurchgang in ein frisches Cap überführt wurden, wurde die *Amylose*-Säule abzentrifugiert und in  $30\ \mu\text{l}$  LowTE und  $30\ \mu\text{l}$  2x Laemmli-Puffer aufgekocht. Die Analyse der ko-präzipitierten Proteine erfolgte über SDS-PAGE und anschließender Autoradiographie.

Es wurden ausschließlich *Protein-LoBind*-Reaktionsgefäße von Eppendorf verwendet.



## 2.5 Herstellung eines monoklonalen Antikörpers gegen *Dictyostelium*-TBP

Injiziert man einer Maus ein Antigen, produziert sie im Rahmen der spezifischen Abwehr polyklonale Antikörper gegen das Fremdprotein. Um einen geeigneten, d.h. ausreichend spezifischen und affinen monoklonalen Antikörper gegen das eingesetzte Antigen zu isolieren werden Milzzellen mit Myelomzellen fusioniert und die Klone anschließend einem ELISA- sowie einem *Western-Screening* unterzogen [Brennand *et al.*, 1986]

### 2.5.1 Expression des Antigens GST-TBP im präparativen Maßstab

Resuspensionspuffer	Tris (pH 8,0)	20 mM
	NaCl	200 mM
	EDTA	1 mM
	β-ME	10 mM
Wasch-Puffer	Tris (pH 8,0)	20 mM
	NaCl	200 mM
	EDTA	1 mM
	β-ME	10 mM
	Triton X 100	1 % (v/v)
Lösungspuffer	Tris (pH 8,0)	10 mM
	NaH <sub>2</sub> PO <sub>4</sub>	100 mM
	Guanidin HCl	6 M

Das Plasmid pGEX(Not)-TBP wurde in W3310-Zellen transformiert. Zur Überprüfung der Expressionsrate und der Löslichkeit des Fusionsproteins GST-TBP wurden Vorversuche analytischen Maßstab durchgeführt. Aliquots der nichtinduzierten, induzierten, löslichen und unlöslichen Fraktion wurden mit SDS-PAGE analysiert und zeigten, dass GST-TBP in hohem Maß unlöslich war.

Eine 1000 ml YT<sub>amp</sub>-Kultur wurde mit 20 ml einer Übernachtskultur angeimpft und bei 37°C bis zu einer OD<sub>600</sub> von 0,5 geschüttelt. Durch Zugabe von 0,5 mM IPTG wurde die Expression des Fusionsproteins induziert. Nach zweistündiger Inkubation bei 37°C wurden die Zellen pelletiert (4°C, 5000 UpM, JA10 Rotor Beckman), mit Resuspensionspuffer gewaschen und anschließend erneut zentrifugiert. Das Pellet wurde anschließend in 50 ml kaltem Resuspensionspuffer aufgenommen und zweimal mit der vorgekühlten *French Press* bei 1200 PSI lysiert. Durch Zentrifugation

bei 4°C und 5000 UpM wurden die *Inclusion bodies* abgetrennt, zweimal mit eiskaltem Waschpuffer gewaschen und in sechs Aliquots aufgeteilt. Ein Aliquot wurde in 4 ml Lösungspuffer aufgenommen und das Protein über Nacht bei RT herausgelöst. Unlösliche Bestandteile wurden durch Zentrifugation (20 min, 4°C, 5000 UpM, JA-20 Rotor Beckman) abgetrennt. Die übrigen Pellets wurden bis zur weiteren Verwendung bei -20°C gelagert.

### **2.5.2 Aufreinigung des rekombinanten Proteins über Ni<sup>2+</sup>-NTA-Säule**

Aus pGEX-Vektoren exprimierte GST-Fusionsproteine besitzen N-terminal einen His<sub>6</sub>-Tag, so dass sie an Ni<sup>2+</sup>-NTA-Säulen immobilisiert werden können [Hochuli 1990].

Waschpuffer A (pH 8,0)	NaH <sub>2</sub> PO <sub>4</sub>	100 mM
	Tris	10 mM
	Guanidin-HCl	6 M
Waschpuffer B (pH 8,0)	NaH <sub>2</sub> PO <sub>4</sub>	100 mM
	Tris	10 mM
	Harnstoff	8 M
Elutionspuffer C, D, E	Waschpuffer B mit pH 6,3 (C), pH 5,9 (D), pH 4,5 (E)	
Bradford-Reagenz	Coomassie Brilliant Blue G250	0,01 % (m/v)
	Ethanol	5 % (v/v)
	Phosphorsäure	10 % (v/v)

4 ml Säulenmaterial wurden mit 3 Säulenvolumen Waschpuffer B äquilibriert und anschließend im *Batch*-Verfahren über Nacht bei RT mit der Proteinlösung beladen. Die Säule wurde mit 3-4 Säulenvolumen Waschpuffer B gewaschen. Unspezifisch gebundene Proteine wurden durch je 2 Säulenvolumen Waschpuffer C und D entfernt. Für die Elution von GST-TBP wurden 2 Säulenvolumen Elutionspuffer auf die Säule gegeben und nach 30 min langsam eluiert. Dabei wurde der Proteingehalt des Eluats mittels der Bradford-Methode engmaschig überprüft, indem 2 µl Probe mit

100 µl Bradford-Reagenz versetzt wurden [BRADFORD 1976]. In Anwesenheit von Proteinen zeigte das Reagenz eine Blaufärbung.

Für die weitere Verwendung musste das Eluat über Nacht, gegen 5 mM Tris dialysiert und dann lyophilisiert werden. Schließlich wurde das Protein in 2 mM Tris, 4 M Harnstoff aufgenommen.

### 2.5.3 Elektroelution

Elutionspuffer	Tris (pH 7,4)	200 mM
	DTT	100 mM
	SDS	0,1 % (m/v)
	pH mit Essigsäure einstellen	
Laufpuffer	Tris (pH 7,4)	50 mM
	DTT	1 mM
	SDS	0,1 % (m/v)
	pH mit Essigsäure einstellen	

Bei der Elektroelution werden Proteine gelelektrophoretisch aufgetrennt und anschließend durch Anlegen einer Spannung aus dem Gel isoliert. Sie kann zusätzlich zu einer Säulenaufreinigung oder als alleinige Aufreinigung durchgeführt werden. Lösliche Proteine können direkt aus dem Überstand des Bakterienlysates verwendet werden, unlösliche müssen mit Lösungspuffer und 4 M Harnstoff aus den *Inclusion bodies* isoliert werden.

Das Fusionsprotein GST-TBP wurde aus *Inclusion bodies* aufgearbeitet, gelelektrophoretisch aufgetrennt und das Gel zwei Stunden mit 4 M Natriumacetatpuffer gefärbt. Mit dem Skalpell wurde die weiß angefärbte Bande ausgeschnitten, kurz in Elektroelutionspuffer gewaschen und dann mit 5-10 ml dieses Puffers in einen Dialyseschlauch gegeben. Das Protein wurde in einer Agarosegelkammer mit Laufpuffer 2 h bei 100 mA aus dem Gel eluiert. Zur weiteren Verwendung wurde es 3 h gegen 5 mM Tris dialysiert und lyophilisiert.

### 2.5.4 Immunisierung der Mäuse

Jeweils 100 µg des GST-TBP-Fusionsproteins wurden mit Freund'schem Adjuvans versetzt und der Maus fünf Mal im Abstand von zwei Wochen in die Bauchdecke injiziert. Der Antikörpertiter wurde nach der dritten Applikation im ELISA getestet. War der Titer ausreichend hoch, wurde die Maus nach zwei weiteren Antigen-Gaben getötet und die Milz präpariert.

Die Lymphozyten wurden isoliert und anschließend im Verhältnis 1:5 mit NS1-Myelomzellen fusioniert. Die Reaktion wurde dabei durch Polyethylenglykol induziert.

Zusätzlich wurde das polyklonale Serum für erste *Screenings* aufgefangen.

### 2.5.5 ELISA-Screening

Coating-Puffer 10x	NaHCO <sub>3</sub>	100 mM
	Na <sub>2</sub> CO <sub>3</sub>	100 mM
	im Verhältnis 3:1, einstellen auf pH 9,6	
PBS 10x (pH 7,2)	Na <sub>2</sub> HPO <sub>4</sub>	80 mM
	KH <sub>2</sub> PO <sub>4</sub>	15 mM
	KCl	25 mM
	NaCl	1,4 M
PBS-T	Tween <sup>®</sup> 20	0,1 % (v/v)
	PBS 1x	
Gelatine-Blockierlösung	Gelatine	0,25 % (m/v)
	PBS 1x	
Zweiter Antikörper	Anti-Maus-AK- POD-Konjugat	3 µl
	Gelatine-Blockierlösung	30 ml

Pro Vertiefung einer Mikrotiterplatte wurden 0,5 µg GST-TBP als Antigen eingesetzt. Das Fusionsprotein aus den gereinigten *Inclusion bodies* wurde in *Coating* Puffer mit 4 M Harnstoff aufgenommen und über Nacht bei 4°C an den Kunststoffboden der Mikrotiterplatte adsorbiert. Überschüssiges Protein wurde anschließend durch Waschen mit 200 µl PBS-T entfernt, freie Valenzen eine Stunde lang mit Gelatine-Lösung bei RT blockiert und eine weitere Stunde der Erstantikörper aus dem Überstand der Hybridomklone gebunden. Nach dreimaligem Waschen mit je 200 µl Gelatinelösung wurde dann 60 min lang mit dem Zweitantikörper (Verdünnung

1:10000) inkubiert. Durch dreimaliges Waschen mit je 200 µl PBS-T wurde ungebundener Antikörper entfernt und anschließend 200 µl ABTS zugegeben. Nach 30 minütiger Inkubation bei RT wurden die Absorptionen im ELISA-Reader bestimmt. Positive Ergebnisse mit dem Antigen GST-TBP wurden mit einem weiteren Fusionsprotein, MBP-TBP, verifiziert, um Antikörperreaktionen gegen den Fusionsteil des rekombinanten Proteins auszuschließen.

Nur Klone mit deutlich positiver Reaktion sowohl mit GST-TBP als auch mit MBP-TBP wurden in der Western-Blot-Analyse näher untersucht.

### 2.5.6 Western-Blot-Analyse

Anoden-Puffer I	Tris (pH 10,4) Methanol	300 mM 20 % (v/v)
Anoden-Puffer II	Tris (pH 10,4) Methanol	25 mM 20 % (v/v)
Kathoden-Puffer	Tris (pH 9,4) 6-Aminohexansäure Methanol	25 mM 40 mM 20 % (v/v)
TBS 10x	Tris (pH 7,5) NaCl	100 mM 1,5 M
TBS-T	TBS 10x Tween <sup>®</sup> 20 H <sub>2</sub> O ad	100 ml 500 µl 1000 ml
BSA-Blockierlösung	Albumin Fraktion V NaN <sub>3</sub> , 10% (m/v) TBS-T ad	3 g 500 µl 100 ml
Milch-Blockierlösung	Milchpulver TBS-T ad	5 g 100 ml
Ponceau-Färbelösung	Ponceau S TCA	0,1 % (m/v) 2,0 % (m/v)

Bei der Western-Blot-Analyse werden elektrophoretisch aufgetrennte Proteine durch ein senkrecht zum Gel angelegtes elektrisches Feld aus der Polyacrylamidmatrix auf eine Nitrozellulosemembran (Protran<sup>®</sup> BA 85) übertragen [KYHSE-ANDERSEN 1984].

Nach der *Semidry*-Methode wurde dazu das *Blotsandwich* wie folgt aufgebaut:

6 Lagen 3MM-Papier, getränkt mit Anodenpuffer I

3 Lagen 3MM-Papier, getränkt mit Anodenpuffer II

Nitrozellulosemembran, etwa 30 min gewässert

PAA-Gel

6 Lagen 3MM-Papier, getränkt mit Kathodenpuffer

Luftblasen zwischen den einzelnen Schichten und überschüssiger Puffer wurden entfernt, um eine gleichmäßige und vollständige Übertragung der Proteine zu gewährleisten. Der Transfer erfolgte entsprechend der Größe des PAA-Gels bei 0,8-1,2 mA/cm<sup>2</sup> für 60 - 90 Minuten. Die übertragenen Proteine wurden zur Kontrolle reversibel mit Ponceau S angefärbt und anschließend mit reichlich Wasser wieder entfärbt. Zur Blockade unspezifischer Valenzen wurde die Membran 30 min bei RT in BSA-Lösung geschwenkt, dann inkubierte man mit dem entsprechend verdünnten Erstantikörper. Nach 60 min bei RT wurde ungebundener Antikörper durch dreimaliges Waschen mit TBS-T entfernt. Die Membran wurde dreimal 5 min mit Milch blockiert, eine Stunde bei RT mit dem Peroxidase-gekoppelten Zweitantikörper (Verdünnung 1:10000 in Milch-Blockierlösung) inkubiert und schließlich wieder je dreimal mit Milch und TBS-T gewaschen. Die Antikörper-markierten Proteine wurden nun durch eine Chemilumineszenz-Reaktion und Auflegen eines Röntgenfilms detektiert. Für bakteriell exprimierte Proteine betrug die Expositionszeit 5-15 min, für Proteine aus *Dictyostelium discoideum* war eine Über-Nacht-Exposition notwendig.

## 2.6 Arbeiten mit *Dictyostelium discoideum*

### 2.6.1 Allgemeine Kulturbedingungen

Untransformierte, axenisch wachsende Zellen des *D.-discoideum*-Stammes AX2 wurden bei 22°C in HL5-Medium in Schüttelkultur (150 rpm) bis zu einer Dichte von  $5 \times 10^6$  Zellen/ml kultiviert [Ashworth & Watts 1970]. Die Bestimmung der Zellzahl erfolgte mittels einer Zählkammer oder einem CASY-Zellzähler. Unter den genannten Bedingungen betrug die Generationszeit 8 bis 10 Stunden.

### 2.6.2 Herstellung von Zellpellets

Ein Volumen Zellsuspension, das  $2 \times 10^7$  Zellen entsprach, wurde pelletiert (10 min, 4°C, Zentrifuge Rotixa/P). Die Zellen wurden zweimal in kaltem Phosphatpuffer gewaschen, dabei fand der letzte Waschschrift in einem 1,5 ml-Reaktionsgefäß statt. Bis zur weiteren Verwendung wurden die Zellen als Pellets ohne Pufferüberstand bei -80°C gelagert.

Für Western-Blot-Analysen von *D. discoideum*-Gesamtextrakten wurde ein Pellet in 200 µl LowTE aufgenommen und durch Vortexen lysiert. Anschließend wurde ein Volumen Laemmli-Puffer zugegeben und aufgeköcht. Von der Probe wurden 10-20 µl entsprechend 50-100 µg Gesamtprotein aufgetragen.

## 3 Ergebnisse

### 3.1 Arbeitshypothese zur Erkennung der Integrationsorte durch TRE5-A.1

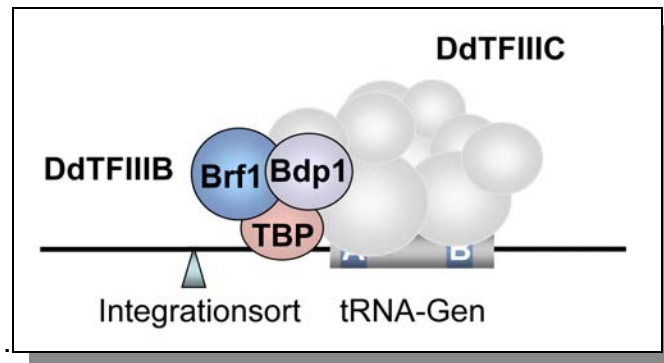
TRE5-Elemente integrieren mit einer bemerkenswerten Präzision positions- und orientierungsspezifisch ca. 48 bp vor tRNA-Gene.

Da tRNA-Gene in ihrer unmittelbaren Nähe keine konservierten DNA-Sequenzen aufweisen, kann die Erkennung von tRNA-Genen durch TRE5-A-Elemente nicht aufgrund einer spezifischen Bindung TRE5-kodierter Proteine an genomische DNA-Sequenzen in der Nähe von tRNA-Genen erfolgen. Vielmehr wird angenommen, dass TRE5-Elemente tRNA-Gene indirekt erkennen, nämlich durch die Interaktion mit tRNA-Gen-spezifischen Transkriptionsfaktoren (siehe Abschnitt 1.2.2).

Das gemeinsame Sequenzmerkmal aller tRNA-Gene ist die hoch konservierte B-Box (siehe Abschnitt 1.3.4). Prinzipiell ist es denkbar, dass TREs mit Hilfe konservierter, flankierender Motive wie der B-Box ihr Integrationsziel direkt erkennen. Allerdings wird das von Pol III transkribierte ribosomale 5S-Gen, das keine B-Box besitzt, ebenfalls als orthodoxer Integrationsort identifiziert [SIOL *et al.*, 2006]. Ein weiteres Argument gegen diese Hypothese ist die Tatsache, dass die B-Box rund 100 bp von der Integrationsstelle entfernt liegt. Nach dem TPRT-Mechanismus wird der Integrationsort von Non-LTR-Retrotransposons durch die Position bestimmt, an der die Endonuklease die Ziel-DNA schneidet. Bei den sogenannten Apurinischen/ Apyrimidinischen Endonukleasen (APE), mit denen die EN-Domäne von TRE5-A phylogenetisch verwandt ist, findet der Schnitt in unmittelbarer Nähe derjenigen Stelle statt, an der das Enzym an die DNA bindet [COST & BOEKE 1998; FENG *et al.*, 1998b; MALIK & EICKBUSH 1999; CHRISTENSEN & EICKBUSH 2005]. Würde das Retrotransposon also unmittelbar die B-Box erkennen, müsste es auch dort in die Ziel-DNA integrieren und nicht 100 bp oberhalb davon.

Bei Bildung eines Präinitiationskomplexes bindet TFIIIC an die interne B-Box der tRNA-Gene und führt TFIIIB an das 5'-Ende des Gens heran (Abb. 3.1):





**Abb. 3.1: Mögliche Zusammensetzung des Pol III-Initiationskomplexes**

Das tRNA-Gen trägt eine geninterne A- und B-Box. Einer der ersten Schritte während der Pol III-Transkription ist die Bindung von TFIIIC an die B-Box. TFIIIC rekrutiert TFIIIB, der in *Dictyostelium discoideum* aus den drei Untereinheiten DdTBP, DdBrf1 und DdBdp1 besteht. Die dreidimensionale Abschirmung der DNA durch TFIIIB endet etwa 50 bp oberhalb des tRNA-Gens. Dort wurden bisher alle Insertionen von TRE5-A-Elementen beobachtet. DdTFIIIC ist schematisch in grau dargestellt. Seine genaue Zusammensetzung ist nicht bekannt.

Während der Transkription kann TFIIIC durch Pol III von seiner DNA-Bindestelle in eine „Parkposition“ verdrängt werden, TFIIIB dagegen bleibt permanent gebunden [ROBERTS *et al.*, 2003]. Die dreidimensionale Abschirmung der DNA durch TFIIIB endet ca. 50 bp oberhalb des tRNA-Gens. Dieser Abstand entspricht ziemlich genau den rund 48 bp zwischen dem 5'-Ende eines integrierten TRE5-Elementes und dem 5'-Ende des tRNA-Gens. Die übereinstimmenden Abstände legen die Möglichkeit nahe, dass TFIIIB und nicht TFIIIC die Zielstruktur der TRE-Integration ist. Demnach ist TFIIIC für die gerichtete Integration von TRE5 nicht unmittelbar essentiell, sondern wird in erster Linie für die Rekrutierung und korrekte Positionierung von TFIIIB benötigt. Durch Protein-Protein-Interaktionen zwischen TRE5-A-kodierten Proteinen und DdTFIIIB findet das Retrotransposon dann sein Integrationsziel.

Diese Arbeitshypothese zur Erkennung seiner Integrationsorte durch TRE5-A sollte durch die nachfolgend beschriebenen Arbeiten direkt überprüft werden. Dazu war es zunächst notwendig, potentielle TFIIIB-Untereinheiten im Genom von *Dictyostelium discoideum* zu identifizieren, da zu Beginn dieser Arbeit praktisch nichts über *D. discoideum*-TFIIIB (DdTFIIIB) bekannt war.

### 3.2 TFIIIB in *Dictyostelium discoideum*

Durch die kürzlich abgeschlossene Sequenzierung des Genoms von *D. discoideum* wurde es möglich, DdTFIIIB *in silico* zu identifizieren. [EICHINGER *et al.*, 2005]. Dazu wurden die bekannten Sequenzen der drei TFIIIB-Untereinheiten TBP, Brf1 und Bdp1 des Menschen (Hs), *Saccharomyces cerevisiae* (Sc) und *Schizosaccharomyces pombe* (Sp) mit dem fertig annotierten *D. discoideum*-Genom verglichen [CHISHOLM *et al.*, 2006]. Die Auswertung der vorhandenen Daten sprechen für die Existenz eines TFIIIB in *D. discoideum*, der sich wie TFIIIB der Hefen aus drei Untereinheiten zusammensetzt (vgl. Abb. 3.1).

Für jedes der drei Proteine DdTBP, DdBrf1 und DdBdp1 wurde nur ein Gen gefunden. Ein orthologes Protein zum humanen Brf2 wurde nicht identifiziert.

#### 3.2.1 DdTBP

DdTBP wurde durch Blastp-Suchen gegen orthologe Proteinsequenzen aus *S. cerevisiae* und dem Menschen identifiziert. Das *tbpA*-Gen (Dictybase Eintrag DDB0185066) befindet sich auf Chromosom 2 und umfasst 819 bp inklusive eines Introns von 200 bp Länge. Entsprechend der kodierenden Sequenz besteht DdTBP aus 205 Aminosäuren und hat ein errechnetes Molekulargewicht von 22,8 kDa:

```

1 ATG AGC ACC GCA ACC ACT ACA TCT ACA CCA GCA CAA AAT GTT GAT CTT TCA AAA CAT CCA TCT GGT ATT ATT CCA 75
1 M S T A T T T S T P A Q N V D L S K H P S G I I P 25

76 ACC CTC CA gtaagtggtgtgtgtttttatatttttttttttttttttttttttttttttttttttttttttttttttttttttttttcaatcac 172
26 T L Q 28

173 tgcaaccattgaagttatataaataaaaaaaaaaaaattttttacttttgatttttttaccttttttttttttttttttttttttttttttttttcaaaaaaa 272

273 aaaaaaaaaatag A AAT ATA GTT TCA ACA GTA AAT ATG GCA ACA GAA TTA GAT TTA AAA GCA ATT GCA TTG GGA GCA 348
29 N I V S T V N M A T E L D L K A I A L G A 49

49 AGA AAT GCA GAA TAT AAT CCA AAA CGT TTT GCA GCA GTT ATT ATG AGA ATT AGA GAA CCA AAG ACA ACA GCA TTA 423
50 R N A E Y N P K R F A A V I M R I R E P K T T A L 74

424 ATT TTC AAG AGT GGT AAA ATG GTT TGC ACT GGA GCA AAA TCT GAA GAT GCA AGT AGA TTT GCT GCA AGA AAA TAT 498
75 I F K S G K M V C T G A K S E D A S R F A A R K Y 99

499 GCA AGA ATC ATT CAA AAA TTG GAT TTC CCT GCA AGA TTT ACA GAT TTT AAG ATT CAA AAT ATC GTT GGA TCT TGT 573
100 A R I I Q K L D F P A R F T D F K I Q N I V G S C 124

574 GAT GTC AAA TTC CCA ATC AAA TTA GAA CTC TTA CAT AAT GCA CAT ACA TCT TTC ACC AAT TAT GAA CCT GAA ATC 648
125 D V K F P I K L E L L H N A H T S F T N Y E P E I 149

649 TTT CCT GGT TTA ATC TAT AAA ATG ATT CAA CCT AAA GTA GTT TTA TTG ATT TTT GTC TCT GGC AAA ATC GTC TTG 723
150 F P G L I Y K M I Q P K V V L L I F V S G K I V L 174

724 ACT GGT GCT AAA GTT CGT GAA GAG ATT TAT GAA GCT TTT GAA AAT ATT TAT CCT GTA CTT TCC GCC TTT AAA AAA 798
175 T G A K V R E E I Y E A F E N I Y P V L S A F K K 199

799 GTT AAT GCA ATT ACT CAA TAA 819
200 V N A I T Q * 206

```

**Abb. 3.2: Sequenz des *tbpA*-Gens**

Vollständige Gensequenz von DdTBP (DDB0185066). Das Intron ist mit Kleinbuchstaben angegeben.

Humaner TBP (Hs) und DdTBP sind im C-terminalen Bereich zu 68% identisch, die Identität mit ScTBP beträgt über das gesamte Protein 70% (Abb. 3.3). Im Gegensatz zum humanen Protein hat *D. discoideum*-TBP keine aminoterminaler Verlängerung.

DdTBP	1	-----
HsTBP	1	MDQNNLPPYAQGLASPOGAMTPGIPFISPMMPYGTGLTPQPIQNTNLSLSILEEQQRQQQ
ScTBP	1	-----
DdTBP	1	-----
HsTBP	61	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQAVAAAAVQQSTSQQATQGTSGQAPQ
ScTBP	1	-----MADEERLKEFKEANKIVFDPNTR
DdTBP	1	-----MSTATTSTPAQNVDSLKHPSGIPTLQNIIVSTVNMATELLE
HsTBP	121	LFHSQTLTAPLPGETTPLYSPMPMPPTIPATPASESSGIIVPOLQNIIVSTVNLGCKLL
ScTBP	24	QVWENQNRDGTTPATTFQSEEDIKRAAPESEKDT--ATSGIIVPTLQNIIVATVTLGCRLL
		H1 <span style="float: right;">H2</span>
DdTBP	43	KATALGARNAEYNPKRFAAVIMRIREPKTTALIFSGKMVCTGAKSEEDASRFARVKYARI
HsTBP	181	KTIALRARNAEYNPKRFAAVIMRIREPRTTALIFSGKMVCTGAKSEEQSRLAARKYARV
ScTBP	83	KTVALRARNAEYNPKRFAAVIMRIREPKTTALIFASGKMVCTGAKSEDDSKLASRKYARI
		H1'
DdTBP	103	IQKLDGPARFDFKIQNIIVGSCDVKFPILLELLHNAHTSFTNYEPELFPGLIYKMIQPKV
HsTBP	241	VQKLGFAKFLDFKIQNIIVGSCDVKFPILLEGLVLT HQQFSSYEPELFPGLIYRMKPRI
ScTBP	143	IQKIGFAAKFDFKIQNIIVGSCDVKFPILLEGLAFSGHTFSSYEPELFPGLIYRMVQPKI
		H2'
DdTBP	163	VLLIFVSGKIVLTGAKVREBIYBAFENIYPVLSAFKRVNAITQ
HsTBP	301	VLLIFVSGKVVLTGAKVRAEIIYBAFENIYPIIKGFFKTT----
ScTBP	203	VLLIFVSGKIVLTGAKVREBIYQAFEAIIYPVLSAFKRM----

**Abb. 3.3: DdTBP – Alignment**

Clustal X-Alignment von DdTBP, HsTBP und ScTBP. Die exponierten  $\alpha$ -Helices auf der konvexen Oberseite sind blau markiert. Schwarz unterlegte Aminosäuren sind in den Proteinen identisch, bei den grau hinterlegten Aminosäuren handelt es sich um konservative Austausche. Die Proteine sind zu knapp 70% identisch.

### 3.2.2 DdBrf1

DdBrf1 (TFIIB-*related factor*) wird von dem Gen *brf1* (Dictybase Eintrag DDB0204087) auf Chromosom 3 kodiert. Es ist 706 Aminosäuren groß und hat ein errechnetes Molekulargewicht von 79,9 kDa.

Der N-Terminus von DdBrf1 (AS 1-270, nachfolgend „DdBrfN“) ist im Vergleich mit orthologen Brf1-Proteinen gut konserviert und weist die erwartete Homologie zu *D. discoideum*-TFIIB auf. Sie ist allerdings scheinbar geringer als zu orthologen TFIIB-ähnlichen Domänen orthologer Brf1-Proteine. In dieser Region ist DdBrf1 zu etwa 42% mit ScBrf1 identisch (Abb. 3.4). Eine Zinkfingerdomäne des Typs CXXC(H)-15/17-CXXC liegt zwischen den Aminosäuren 4-32 [HAHN & ROBERTS 2000]. Dagegen zeigt die C-terminale Region (AS 271-706; im weiteren „DdBrfC“), die wahrscheinlich die Spezifität für die Pol III vermittelt, nur sehr schwache Homologien zu Brf1-Proteinen anderer Organismen. Keine der drei Homologie-Regionen I-III, die in den Hefen *C. albicans*, *K. lactis*, *S. pombe* und *S.*

*cerevisiae* erhalten sind [KHOO *et al.*, 1994; ANDRAU *et al.*, 1999], konnten in DdBrfC identifiziert werden. In den Hefen stellen diese Regionen durch Wechselwirkung mit der C34-Untereinheit der Pol III die wesentliche Kontaktdomäne zwischen Brf1 und dem Enzym dar [KHOO *et al.*, 1994].

DdBrf1	1	MS-KNCGNCG-NDSFASN--SVVCTACGTVKESANIVSEVQF---GGDSSIVGTFV
HsBrf1	1	MTGRVCRGCGGTDIELDAARG--DAVCTACGSVLEDNIIIVSEVQFVSSGGSSAVGQEV
ScBrf1	1	MP--VCKNCHGTEFERDLSNANNDLVCKACGVVSEDNPIVSEVTFGETSAGAAVVQGSFI
DdBrf1	53	S---ATRKPS---YRSLGGRDSRAMSVENARKKLDQIGNSDRIRSHLIDSQRTFELAM
HsBrf1	59	SLDGAGKTPTLGGGFHVNLGKESRAQTLONGRRRIHHLGNQLQLNQHCLDTAFNFFKMAV
ScBrf1	59	G---AGQSHAA---FGGSSALESREATLNNARKKLRAVSYAHIPEYITDAAFQWYKLA
DdBrf1	107	ENNFTKGRQTRLVAAACLYIVCRRETRTPHLLIDFSENLOVNVFIVAGTFLLQIKLNI-Q
HsBrf1	119	SRHLTRGRKMAHVIAACLYIVCRTEGTPHLLDLDLLOVNVVVLGKTFLLARELCT-N
ScBrf1	113	ANFVQGRRSQNVIASCLYVACRKEKTHMLIDFSSRLQVSVYSIGATFLKMVKKLHTE
DdBrf1	166	LPIVDPSFFIQRFASALEFGDQTKDVIATANKLVARMKRDWMSIGRKFSGICGASLYIAS
HsBrf1	178	APAIDPCLYIPRFAHLLFEGKNHEVSMALRLLORMKRDWMHTGRRPSGLCGAALLVAA
ScBrf1	173	LPALDPSLFIQHFAEKLDLADKKIKVVKDAVKLAORMSKDWMFEGRRPAGIAGACILAC
DdBrf1	226	KIHGFKRSMKEIVHVVKIGEQTLINRLDEFKKTPTATLRFDEFNG
HsBrf1	238	RMHDFRRVVKVVISVVKVCSLTKRRLTEPEDTPTSQLTIDEFMK
ScBrf1	233	RMNLRRTHTTEIVAUSHVAEPLQORLNEFKNPKAAKLSVQKFR

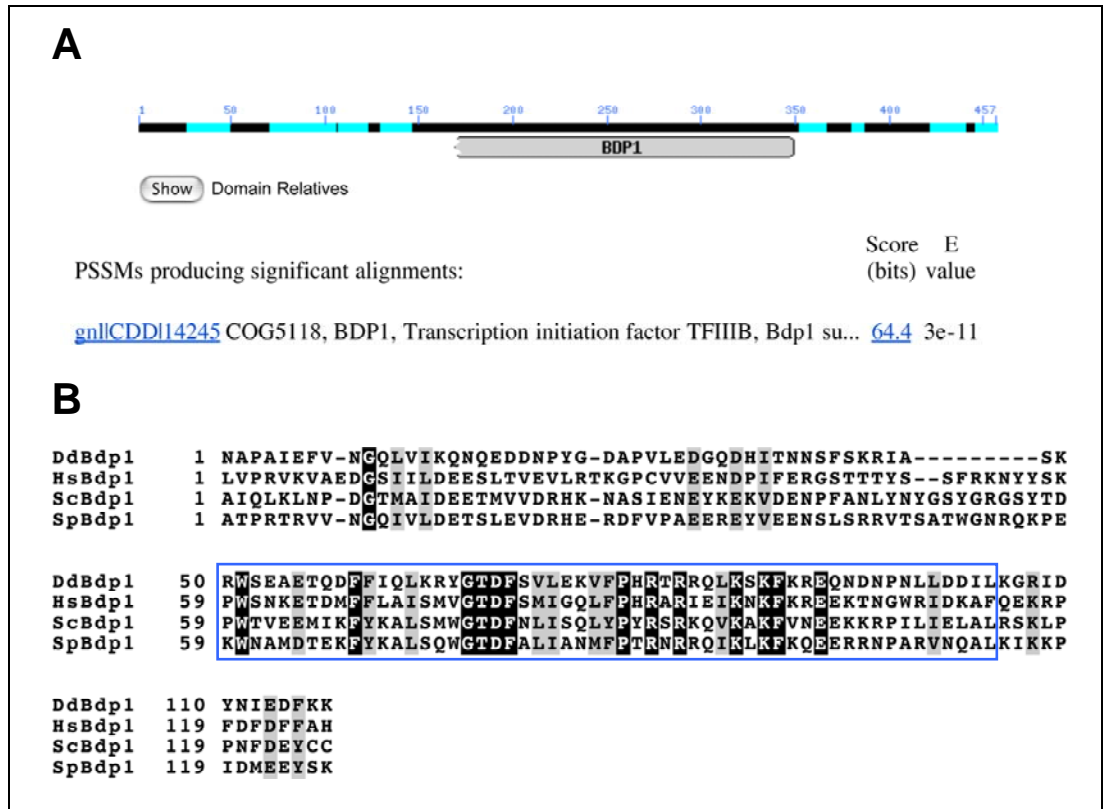
**Abb. 3.4: DdBrf1-Alignment**

Clustal X-Alignment der konservierten TFII-homologen Domäne von DdBrf1 (DdBrfN) und den entsprechenden Domänen aus HsBrf1 und ScBrf1. Schwarz unterlegte Aminosäuren sind in beiden Proteinen identisch, bei grau unterlegte Positionen handelt es sich um konservative Austausche.

### 3.2.3 DdBdp1

2002 einigte man sich auf eine einheitliche Nomenklatur für die Untereinheiten des Transkriptionsfaktors TFIIIB [WILLIS 2002]. Bdp1, vorher B", wurde nach der biochemischen Fraktion, aus der es ursprünglich isoliert wurde, *TFIIIB double prime* benannt [KASSAVETIS *et al.*, 1991]. Bdp1 ist die am geringsten konservierte Untereinheit von TFIIIB, deshalb erwiesen sich Blastp-Recherchen in den *Dictyostelium*-Genomdaten anhand der Sequenzen orthologer Proteine aus *S. cerevisiae* und *H. sapiens* als schwierig (Abb. 3.5, B). Am besten konserviert ist die SANT-Domäne im C-terminalen Drittel von Bdp1. Sie konnte in dem putativen DdBdp1 eindeutig identifiziert werden (Abb. 3.5, A). SANT-Domänen sind meist in Proteinen vertreten, die an der transkriptionellen Regulation beteiligt sind. Es wird außerdem vermutet, dass die persistierende Bindung von TFIIIB an die DNA auch ohne TFIIIC auf diese Domäne zurückzuführen ist [AASLAND *et al.*, 1996]. Eine Interaktion der SANT-Domäne aus ScBdp1 und der Homologie-Region II aus ScBrf1 konnte kürzlich nachgewiesen werden [KASSAVETIS *et al.*, 2006].

Nach den Recherchen in *Dictybase* kodiert das Gen *bdp1* (Dictybase Eintrag DDB0220508) für ein 457 AS großes Protein mit einem Molekulargewicht von 53 kDa. Das Gen umfasst 1474 bp inklusive einer 100 bp langen Intronsequenz und liegt auf Chromosom 4.

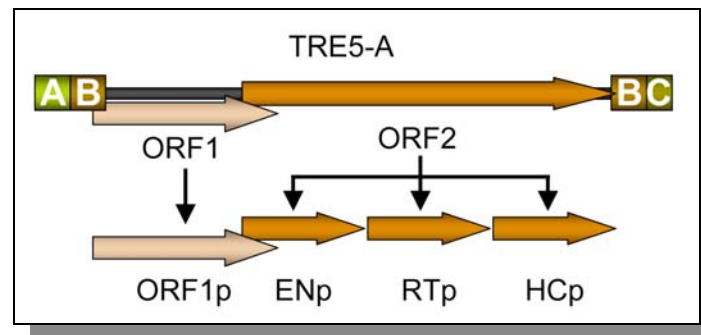


### Abb. 3.5: Identifizierung von DdBdp1

**A.** Mit Hilfe der *Conserved Domain Database* ([www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi](http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi)) wurde in der Sequenz des möglichen DdBdp1 eine konservierte Domäne gefunden, die mit guter Signifikanz als Bdp1 identifiziert wird [MARCHLER-BAUER *et al.*, 2005]. **B.** Alignment des Kernbereichs von DdBdp1 mit den entsprechenden Bdp1-Regionen aus *H. sapiens*, *S. cerevisiae* und *S. pombe*. Die SANT-Domäne ist blau markiert.

### 3.3 TRE5-A.1 Proteine

Ein vollständiges TRE5-A.1-Element enthält zwei offene Leserahmen (ORFs). Sie kodieren für die Proteine ORF1p und ORF2p, die beide als potentielle Interaktionspartner für DdTFIIIB während der Identifizierung des Integrationsziels in Frage kommen (Abb. 3.6).



**Abb. 3.6: Schematische Darstellung der klonierten TRE5-A.1-kodierten Proteine**

Die beiden offenen Leserahmen von TRE5-A.1 kodieren für die Proteine ORF1p und ORF2p. ORF2p wurde in Form seiner funktionellen Domänen Endonuklease (ENp), Reverse Transkriptase (RTp) und Histidin/ Cystein-Domäne separat kloniert und exprimiert. A, B und C stellen die Module der 5'- bzw. 3'-UTR dar.

#### 3.3.1 ORF1-Protein

TRE5-A.1 ORF1 umfasst 1335 bp, die ein 444 Aminosäuren großes Protein kodieren. Das Startkodon liegt im B-Modul des Elements.

Nach der Klonierung der ORF1-Sequenz anhand des Plasmids pB3 (Genbank X57034 [MARSCHALEK *et al.*, 1992a]) wurden im Vergleich zu der vom *Dictyostelium*-Genomprojekt etablierten Konsensus-Sequenz für TRE5-A.1 sechs Aminosäureaustausche festgestellt (vgl. Abb. 3.7) [MARSCHALEK *et al.*, 1992a; GLÖCKNER *et al.*, 2001b]. Das auf dem Plasmid pB3 enthaltene TRE5-A.1-Element wurde aus einer genomischen Genbank isoliert [MARSCHALEK *et al.*, 1992a]. Aufgrund der festgestellten Abweichungen der neu klonierten ORF1-Sequenz von dem mit bioinformatischen Methoden festgelegten Konsensus wurde zunächst die ORF1-Sequenz des pB3-kodierten TRE5-A-Elements resequenziert. Es stellte sich heraus, dass die DNA-Sequenz sowohl mit der in der Genbank publizierte Sequenz für TRE5-A.1 (ehemals DREa, X57034) als auch mit der neu klonierten Sequenz übereinstimmte (Abb. 3.7).

ORF1p_cons	MKITNNTTNIKQHKCLQKISEIVDKTQLKKKQFKYVININHEPDDKIKKDLEKSLDKKDV
ORF1p_X57034	MKITNNTTNIKQHKCLQKISEIVDKTQLKKKQFKYVININHEPDDKIKKDLEKSLDKKDV
ORF1p_cons	LIKSNNTFKILTDSISTVIYFVNMFDAELKGQFITTLRPKNMYTDFDWYKSLTEIEWVIE
ORF1p_X57034	LIKSNNTFKILTDSISTVIYFVNMFDAELKGQFITTLRPKNMYTDFDWYKSLTEIEWVIE
ORF1p_cons	NEGCKLIKSEIKGETLIKTVKRVVEEYDTIIELDNLTLIHGSNKGWRDQTSLPENQEEEN
ORF1p_X57034	NEGCKLIKSEIKGETLIKTVKRVVEEYDTIIELDNLTLIHGSNKGWRDQTSLPENQEEEN
ORF1p_cons	KNKKQSDEQQPQSTSEQIPKQGKRTIIIPPITSSNDKTFMALDDTAQKVIASLENVDE
ORF1p_X57034	KNKKQSDEQQPQSTSEQIPKQGKRTIIIPPITSSNDKTFMALDDTAQKVIASLENVDE
ORF1p_cons	LTRIENEINTQSIINQEIKTSEKPIEPPSPYNTPIKQHGETKGLTTEEKREANDIFDF
ORF1p_X57034	LTRIENEINTQSIKNQEIKTSEKPIEPPSPYNTPIKQHGETKRLTTEEKREANDIFDF
ORF1p_cons	TSPNTDAIRKIEFPLSIPSNQWFFSSPIKHIMSFSPTYGTPSKIIDSPSKLLNTPKNKQQ
ORF1p_X57034	TSPNTDAIRKIEFPLSIPSNQWFFSSPIKHIMSFSPTYGTPSKIIDSPSKLLNTPKNKQQ
ORF1p_cons	DKTNRTNKTPVKPKLGEYGGQKSIKTKMDEDIDNFKELKQINKSKNPTNTQDQQHNEQS
ORF1p_X57034	DKTNRTNKTPVKPKLGEYGGQKSIKTKMDEDIDNFKELKQISKSKNPTNTQDQQHNEQS
ORF1p_cons	KDQLEDKINQKLQKVDKNLPQIKL
ORF1p_X57034	KDQLEDKINQKLQKVDKNLPQIKL

### Abb. 3.7: Alignment der ORF1-Sequenzen

Der Vergleich der bioinformatisch erstellten Konsensus-Sequenz für TRE5-A.1-Elemente (cons) und der pB3-kodierten Sequenz von ORF1p (Genbank X57034) zeigt sechs Abweichungen in der pB3-kodierten ORF1p-Sequenz, die blau markiert sind.

Um zu überprüfen, ob es sich bei den festgestellten Aminosäure-Abweichungen um PCR- bzw. Klonierartefakte während der Neuklonierung von ORF1p für die nachfolgend beschriebenen Versuche handelte oder ob die Mutationen in authentischen, retrotranspositionskompetenten Elementen vorhanden sind, wurde die klonierte ORF1p-Sequenz mit genomischen TRE5-A-Kopien verglichen. Blastp-Suchen ergaben mindestens eine ORF1p-Sequenz, die mit der klonierten ORF1p-Sequenz vollständig übereinstimmte (DDB0218337) (Abb. 3.8). Diese Sequenz stammt von einem intakten TRE5-A.1-Element, das 48 bp vor einem Thr(GCU)-tRNA-Gen auf Chromosom 3 integriert ist. Es wurden mindestens neun weitere Elemente identifiziert, die die ORF1p-Variante E232D, G284R, A294V, N403S enthalten (DDB0118161, DDB0217959, DDB0217293, DDB0217122, DDB0216652, DDB0216617, DDB0218400, DDB0215551, DDB0217749). Ebenso wurden fünf Elemente im Genom identifiziert, die die seltenen Mutationen Q192K und N254K aufweisen (DDB0218400, DDB0217749, DDB0001853, DDB0219771, DDB0219554). Diese Beobachtungen legen nahe, dass sich die erwähnten Elemente innerhalb des Genoms trotz der Aminosäure-Abweichungen amplifiziert haben und nicht erst nach erfolgter Retrotransposition entstanden sind. Die Mutationen sind demnach keine Klonierartefakte, sondern repräsentieren wahrscheinlich funktionell aktive ORF1p-Derivate.

```

>DDB0218337 |Protein| gene: JC3V2_0_02027 on chromosome: 3 position 5068971 to 5070305
Length = 444
Score = 889 bits (2296), Expect = 0.0 Identities = 444/444 (100%), Positives = 444/444 (100%)

Query: 1 MKITNNTTNIKQHKCLQKI SEIVDKTQLKKKQFKYVININHEPDDKIKKDLEKSLDKKDV 60
Sbjct: 1 MKITNNTTNIKQHKCLQKI SEIVDKTQLKKKQFKYVININHEPDDKIKKDLEKSLDKKDV 60

Query: 61 LIKSNNTPKILTDSISTVIYFVNMFD AELKGQFITTLRPNMYTDFDWYKSLTEIEWVIE 120
Sbjct: 61 LIKSNNTPKILTDSISTVIYFVNMFD AELKGQFITTLRPNMYTDFDWYKSLTEIEWVIE 120

Query: 121 NEGCKLIKSEIKGETLI IKTVKRVVEEYDTII ELDNLT LIGHSNKGWRDQTSLPENQEEEN 180
Sbjct: 121 NEGCKLIKSEIKGETLI IKTVKRVVEEYDTII ELDNLT LIGHSNKGWRDQTSLPENQEEEN 180

Query: 181 KKKQSDEQQPKSTSEQIPKQGKRTII IPPITSSNDKTFMALDDTAQKVIDASLENVDE 240
Sbjct: 181 KKKQSDEQQPKSTSEQIPKQGKRTII IPPITSSNDKTFMALDDTAQKVIDASLENVDE 240

Query: 241 LTRIENEINTQSIKNQBIKTKSEKPIBPPSPYNTPIKQHGETKRLTTEEKLREVNDFDF 300
Sbjct: 241 LTRIENEINTQSIKNQBIKTKSEKPIBPPSPYNTPIKQHGETKRLTTEEKLREVNDFDF 300

Query: 301 TSPNTDAIRKIEFPLSIPSNQWFFSSPIKHIMSFSPYGTSPSKI IDSPSKLLNTPKNKQQ 360
Sbjct: 301 TSPNTDAIRKIEFPLSIPSNQWFFSSPIKHIMSFSPYGTSPSKI IDSPSKLLNTPKNKQQ 360

Query: 361 DKTNRNTNKT PVKPKLGEYGQKSIKTKMDEDIDNFNKEILKQISKSKNPTNTQDQQHNEQS 420
Sbjct: 361 DKTNRNTNKT PVKPKLGEYGQKSIKTKMDEDIDNFNKEILKQISKSKNPTNTQDQQHNEQS 420

Query: 421 KDQLEDKINQKLQKVDK NLPQIKL 444
Sbjct: 421 KDQLEDKINQKLQKVDK NLPQIKL 444

```

### Abb. 3.8: ORF1p-Sequenz aus DDB0218337

Eine Blastp-Suche mit der ORF1p-Sequenz aus dem Plasmid pB3 (Query) im *D. discoideum*-Genom ergab u. a. die Genvorhersage DDB0218337 (Sbjct). Die Sequenzen des neu klonierten ORF1p und des intakten TRE5-A.1-Elementes stimmen zu 100% überein.

### 3.3.2 ORF2-Protein

Mit 3456 bp, die für ein 135 kDa großes Protein kodieren, war ORF2 einerseits nicht als ganzes klonierbar, andererseits auch zu groß, um im B2H-System getestet werden zu können. Potentielle Interaktionspartner sollten nicht größer als 50 kDa sein, da ansonsten die Induktion der Reportergene auch bei ausreichender Expression der Fusionsproteine möglicherweise fehlschlägt. Aus diesem Grund wurde ORF2p in Form seiner drei funktionellen Domänen Endonuklease (ENp), Reverse Transkriptase (RTp) und der Histidin/Cystein-reichen Domäne (HCp) separat kloniert (vgl. Abb. 3.6).

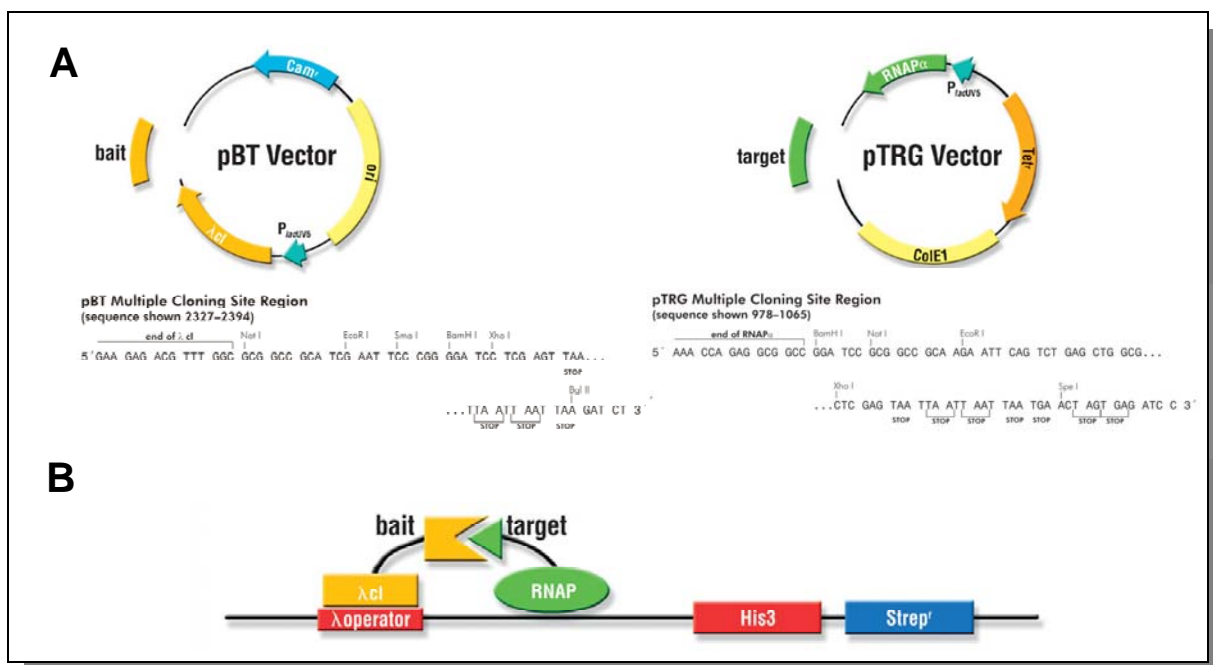


### 3.4 Interaktionsanalyse im bakteriellen Zweihybrid-System

Zweihybrid-Experimente dienen dazu, potentielle Wechselwirkungen zwischen zwei Proteinen *in vivo* nachzuweisen bzw. Interaktionspartner zu identifizieren und die Kontaktdomänen näher zu charakterisieren. Im Gegensatz zu biochemischen Methoden sind hierfür keine aufwendig gereinigten Proteine notwendig.

Durch Interaktion zweier Testproteine wird eine bakterielle DNA-Bindedomäne (DNA-BD) mit einer Transkriptionsaktivierenden Domäne (AD) zusammengeführt und wieder zu einem aktiven Transkriptionsfaktor rekonstituiert, so dass die Expression von Reportergenen induziert wird (Abb. 3.9).

Im bakteriellen Zweihybrid-System *BacterioMatch*<sup>TM</sup> II der Firma Stratagene (im folgenden B2H-II abgekürzt) besteht die DNA-Bindedomäne aus dem Repressorprotein des Phagen Lambda ( $\lambda$ cl), das an spezifische Operatorsequenzen im starken lacUV5-Promotor bindet. Das  $\lambda$ cl-Gen befindet sich auf dem pBT-Vektor (Abb. 3.9). Die Aktivierungsdomäne besteht aus der  $\alpha$ -Untereinheit der bakteriellen RNA-Polymerase und wird von dem Zielplasmid pTRG mit dem Tandempromotor lpp/lacUV5 exprimiert (Abb. 3.9).



**Abb. 3.9: Das *BacterioMatch*<sup>TM</sup> Two Hybrid System**

**A.** Vektorkarte des verwendeten Köderplasmids pBT bzw. des Zielplasmids pTRG. P: Promotor, ori: Replikationsursprung **B.** Prinzip des bakteriellen Zweihybrid- Systems. Interaktionen zwischen den Fusionsproteinen führen zu einer transkriptionellen Aktivierung der Reportergenkassette, die aus dem *HIS3*- und dem *aadA*-Gen besteht.

Protein-Protein-Interaktionen werden durch transkriptionelle Aktivierung des *HIS3*-Gens und des *aadA*-Gens, das eine Streptomycin-Resistenz vermittelt, detektiert. Starke Interaktionen konnten noch in Gegenwart von 6 mM 3-Amino-1,2,4-Triazol (3-AT) detektiert werden, schwache nur bei 4 mM 3-AT. Nur Klone mit einem interagierenden Proteinpaar sind Histidin-prototroph und streptomycinresistent.

Zu Beginn der Arbeit wurde die erste Version dieses bakteriellen Zweihybrid-Systems nach Dove und Joung verwendet (B2H-I). Diese unterschied sich von der späteren Version lediglich durch die Verwendung einer anderen Reporter-gen-Kassette. Statt des *HIS3*- und des *aadA*-Gens wurden Gene für eine  $\beta$ -Lactamase (*amp<sup>R</sup>*) und eine  $\beta$ -Galactosidase (*lacZ*) verwendet. Positive Klone konnten so auf X-Gal-haltigen Selektionsplatten aufgrund einer Blaufärbung der Kolonien identifiziert werden. Zum anderen vermittelte die Interaktion zwischen den zu testenden Proteinen eine Ampicillin-Resistenz. Allerdings war die basale Transkriptionsrate so hoch, dass mit diesem System viele falsch positive Klone identifiziert wurden. Aus diesem Grund werden an dieser Stelle nur wenige beispielhafte Daten gezeigt, die aus Experimenten mit dem B2H-I stammen. Stattdessen wurden die Versuche auf das wesentlich zuverlässigere B2H-II-System umgestellt, in dem auf Wachstum auf Histidin-Mangelmedium bzw. Streptomycinresistenz getestet wird (siehe unten).

Mit Hilfe des bakteriellen Zweihybrid-Assays B2H-II sollten systematisch alle TRE5-A.1-kodierten Proteine (ORF1p und ORF2p) und die drei DdTFIIIB-Untereinheiten DdTBP, DdBrf1 und DdBdp1 auf mögliche Interaktionen untersucht werden. Potentielle Interaktionspartner aus DdTFIIIB und TRE5-A.1 wurden jeweils in beide Vektoren kloniert. Für die Klonierungen wurden die Primer allgemein so gewählt, dass sie an beiden Enden des Amplifikats jeweils eine *NotI*-Schnittstelle generierten:

```

           $\lambda$ cI          --Not I--          tbpA
CAGTGGCCTGAAGAGACGTTTGGCgggcgccaATGAGCACCGCAACCACTACATCT ...
Q W P E E T F G A A A M S T A T T T S

          tbpA          --Not I--
... GCCTTAAAAAAGTTAATGCAATTACTCAATAA gggcgccc
A F K K V N A I T Q *

```

**Abb. 3.10: DNA-Sequenz des in pBT klonierten *tbpA*-Gens**  
Gezeigt sind die Übergänge zwischen Vektorsequenz und der DNA-Sequenz von DdTBP.

Am 5'-Ende des PCR-Fragments wurde durch die Primer außerdem eine Konsensus-Sequenz für einen Translationsstart in Säugerzellen eingeführt (5'-GCCGCC(A/G)CCATGG-3', Translations-Startkodon unterstrichen) [KOZAK 1989]. Dadurch konnten die PCR-Fragmente nach Umklonierung in den Vektor pTNT in einem säugerzellbasierten *in vitro* Translationssystem exprimiert werden (siehe Abschnitt 3.5). Die kodierenden Sequenzen von DdTBP und DdBdp1 werden jeweils durch ein Intron unterbrochen. Deshalb wurde zunächst die RNA aus *D. discoideum* AX2-Zellen isoliert, daran die cDNA synthetisiert und diese dann als *Template* für die PCR eingesetzt. Für die Amplifizierung von DdBrf1 wurde genomische AX2-DNA verwendet. Für die Experimente im Zweihybridsystem wurde DdBrf1 zwischen seinen beiden funktionellen Domänen in eine TFIIIB-homologe (DdBrfN) und eine C-terminale Hälfte (DdBrfC) geteilt.

TRE5-A.1-ORF2p wurde entsprechend seinen funktionellen Domänen in drei Proteine geteilt: die Endonuklease (ENp), die reverse Transkriptase (RTp) und ein Histidin/ Cystein-reiches Protein (HCp) (Abb. 3.6).

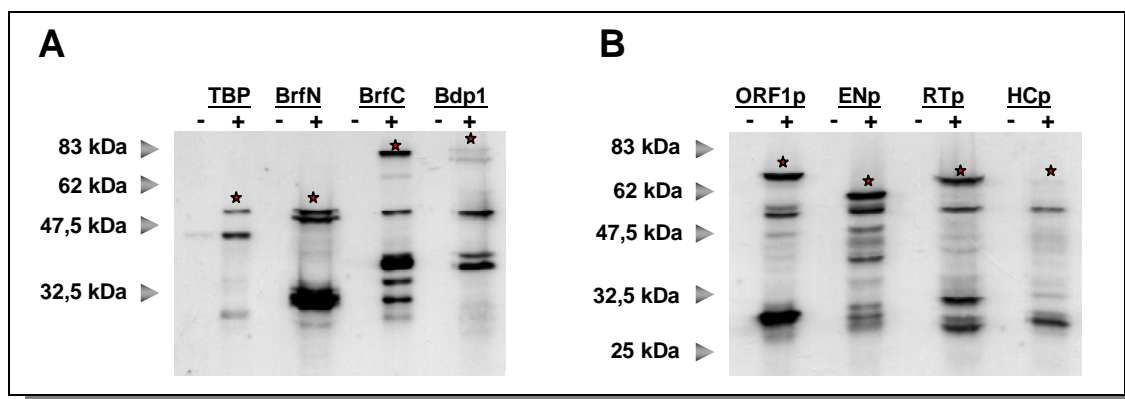
Alle Fragmente wurden mittels PCR hergestellt. Die einzelnen PCR-Fragmente wurden zunächst in pGEM-T zwischenkloniert und anschließend als *NotI*-Fragmente in die B2H-Vektoren pBT und pTRG kloniert. Jedes Fragment wurde sowohl in pBT als auch pTRG getestet um auszuschließen, dass Interaktionen wegen möglicher sterischer Hinderungen mit einem der Konstrukte nicht detektiert werden konnten. Die zu testenden Plasmide wurden in Reporterzellen kotransformiert und die Klone anschließend auf Platten mit Histidin-Mangelmedium, das mit unterschiedlichen Konzentrationen an 3-AT supplementiert wurde, ausgestrichen. Standardmäßig wurden Platten mit 5 mM 3-AT verwendet. Eine Übersicht der prinzipiellen Interaktionen zwischen TRE5-A.1 und DdTFIIIB wird in Tab. 3.1 gegeben.

pTRG \ pBT	ORF1p	ENp	RTp	HCp	DdTBP	DdBrf1 1-270	Brf1 271-706	DdBdp1
ORF1p	++++	-	-	-	++++	-	+++	++
ENp	-	+++	-	-	-	-	-	-
RTp	-	-	-	-	-	-	-	-
HCp	-	-	-	-	-	-	-	-
DdTBP	-	-	-	-	-	-	+	-
DdBrf1 <sup>1-270</sup>	-	-	-	-	-	-	-	-
DdBrf1 <sup>271-706</sup>	+	-	-	-	++++	-	-	-
DdBdp1	-	-	-	-	-	-	+	-

**Tab. 3.1: Zusammenfassung der *Two Hybrid*-Versuche**

++++: stabile Interaktion bei bis zu 6 mM 3-AT, +++: stabile Interaktion bei bis zu 5 mM 3-AT, ++: schwache Interaktion, nur bei 4 mM 3-AT detektierbar, +: sehr schwache Interaktion bei 2-3 mM 3-AT, -: keine Interaktion detektierbar

Die Expression der in pBT-kodierten  $\lambda$ cl-Fusionsproteine wurde mittels Western-Blot-Analyse verifiziert (Abb. 3.11). Eine entsprechende Überprüfung der pTRG-kodierten Fusionsproteine war nicht möglich, da kein Antiserum gegen den RNA-Polymerase-Fusionsanteil zur Verfügung stand. Für jede Doppeltransformation wurde eine Negativkontrolle als optischer „Leerwert“ mitgeführt. Eine Autoaktivierung der Reportergene wurde in Vorversuchen ausgeschlossen, indem die jeweiligen rekombinanten Plasmide mit einem Leervektor kotransformiert wurden. Mit 3-AT wurde der Selektionsdruck auf die Zellen so eingestellt, dass die „Leerwert“-Transformanden unter den gewählten Bedingungen kein Wachstum zeigten.



**Abb. 3.11: Expression der Fusionsproteine im B2H II-Reporterstamm**

Für die Western Blot Analysen wurden TBP (50 kDa), BrfN (57 kDa), BrfC (77 kDa), Bdp1 (73 kDa), ORF1 (78 kDa), EN (65 kDa), RT (78 kDa) und HC (76 kDa) als pBT-Konstrukte in den Reporterstamm transformiert und die Expression durch Zugabe von 0,5 mM IPTG 3h bei 30°C induziert. In Klammern sind die errechneten Molekulargewichte der Fusionsproteine angegeben. Gesamtzell-Lysat von nichtinduzierten (-) und induzierten (+) Bakterien wurde aufgetragen und die Fusionsproteine mit einem polyklonalen, gegen  $\lambda$ cl gerichteten Antiserum nachgewiesen. Die entsprechenden Banden der vollständigen Fusionsproteine sind markiert.

### 3.4.1 Proteininteraktionen zwischen TFIIIB-Untereinheiten

Um die Zusammensetzung von TFIIIB aus *D. discoideum* näher zu untersuchen, wurden die drei Untereinheiten von TFIIIB sowohl in den *Bait*-Vektor pBT als auch in den *Target*-Vektor pTRG kloniert. DdTBP und DdBdp1 wurden hierbei als vollständige Proteine untersucht, während die dritte Untereinheit DdBrf1, zwischen ihren zwei funktionellen Domänen geteilt und anschließend separat als DdBrfN (AS 1-270) und DdBrfC (AS 271-706) getestet wurde. Die Konstrukte wurden durch eine Doppeltransformation in Kanamycin-resistente kompetente Reporterzellen eingebracht und auf Platten mit dem notwendigen Selektionsdruck ausgebracht.

Sowohl im B2H-I-System als auch im B2H-II-System konnte eine starke Interaktion zwischen DdBrfC und DdTBP nachgewiesen werden (Abb. 3.12). Im B2H II-System blieb die Interaktion auch bei 6 mM 3-AT noch nachweisbar.



**Abb. 3.12: Interaktion zwischen DdTBP und DdBrf1 im B2H I-System**

Die Aktivierung der Reportergene *amp<sup>R</sup>* und *lacZ* im B2H I-System ermöglicht auf X-Gal-haltigen Platten eine Detektion positiver Klone durch Blaufärbung (linke u. rechte Platte). Das zweite Reportergen vermittelt zusätzlich eine Carbenicillin-Resistenz (rechte Platte). 1: Positivkontrolle pBT-Gal/ pTRG-LGF, 2: Negativkontrolle pBT/ DdTBP, 3: DdBrfC/ DdTBP, 4: Negativkontrolle DdBrfC/ pTRG

Da in *S. cerevisiae* die C-terminale Homologie-Region II von Brf1 mit TBP den stabilen B'-Komplex bildet, wurde eine ähnliche Interaktion auch in *D. discoideum* erwartet und konnte mit diesem Versuch bestätigt werden [KASSAVETIS *et al.*, 1998a]. Eine Wechselwirkung zwischen DdTBP und DdBrfN wurde ebenso wenig beobachtet wie mit DdBdp1.

Zwischen DdBrf1 und DdBdp1 konnte eine sehr schwache Interaktion detektiert werden, allerdings war das Wachstum der Transformanden kaum von dem der „Leerwert“-Transformanden zu unterscheiden (siehe Tab. 3.1).

### 3.4.2 Dimerisierung der TRE5A.1-Proteine ORF1p und ENp

Aus vorhergehenden Arbeiten war bereits bekannt, dass ORF1p große Aggregate bildet, die mittels GFP-Markierung im Zytosol von *D. discoideum*-Zellen nachweisbar waren [HENTSCHEL *et al.*, 2001].

Im B2H-II-System konnte eine starke Interaktion von ORF1p mit sich selbst gezeigt werden, die auch in nachfolgenden *Pulldown*-Versuchen bestätigt wurde (Abb. 3.13). Computeranalysen sagen eine C-terminal etwa zwischen den Aminosäuren 410-444 von ORF1p gelegene *Coil*-Domäne voraus, die möglicherweise an einer Dimerisierung bzw. Oligomerisierung beteiligt ist. Eine Deletionsmutante ORF1p<sup>1-404</sup> war im B2H-II-System dann tatsächlich nicht mehr zu Dimerisierung fähig (Abb. 3.19). Die direkte positive Testung des Fragments ORF1p<sup>402-444</sup> bestätigte, dass der C-Terminus von ORF1p die Dimerisierung des Proteins vermittelt (Daten nicht gezeigt).

Die Endonuklease-kodierende Sequenz des ORF2 konnte trotz großer Bemühungen zunächst nicht in die Vektoren des B2H-II-Systems kloniert werden. Durch gerichtete Einzelstrangmutagenese wurde die Aminosäure Asp 188 (D<sup>188</sup>) im aktiven Zentrum der EN gegen ein Glycin ausgetauscht und das Enzym so vermutlich inaktiviert. Die D188G-Mutation wurde in Analogie zur D205G-Mutation im humanen Retrotransposon L1 gewählt, dessen Endonuklease dadurch inaktiv wurde [FENG *et al.*, 1996]. Die EN(D188G)-Mutante konnte nun ohne Weiteres in die B2H-II-Vektoren kloniert werden.

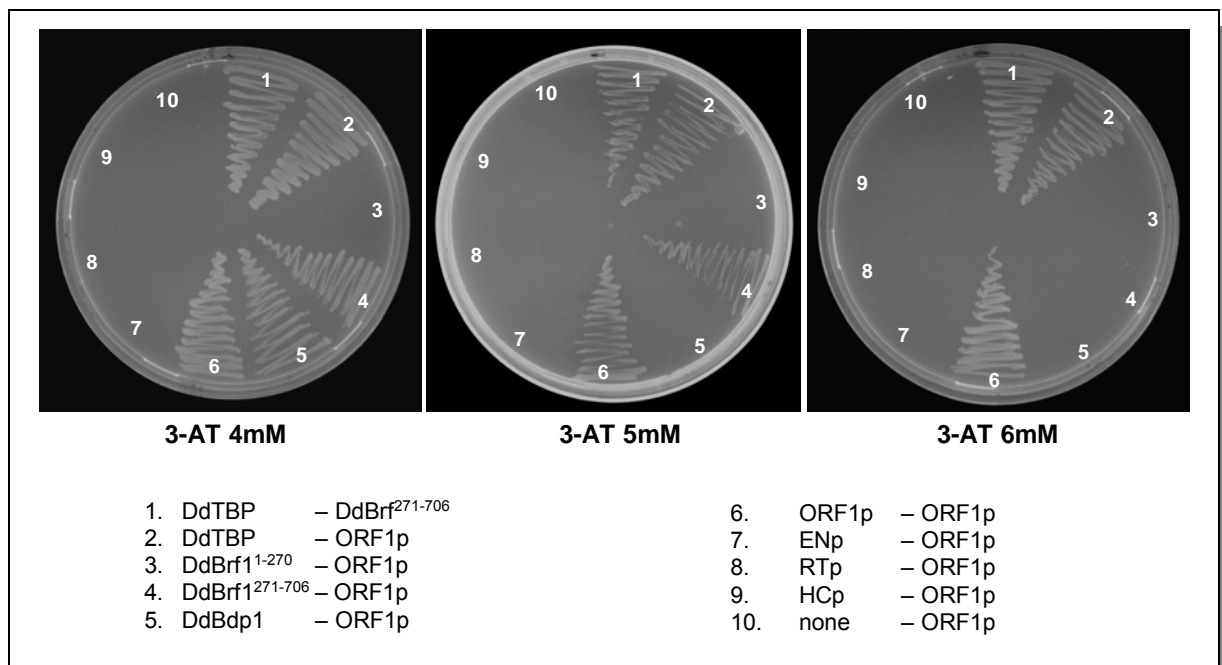
Hentschel *et al.* zeigten 2001, dass ENp sowohl *in vitro* (in Niedrigsalzpuffern) als auch *in vivo* (im Zytosol von *D. discoideum*-Zellen) zur Aggregation neigen. Dies konnte reproduzierbar auch im B2H-System gezeigt werden und weist auf ein bisher noch nicht identifiziertes Dimerisierungsmotiv in der EN-Domäne von TRE5-A.1 ORF2p hin (siehe Abb. 3.19, A).

Außer der Oligomerisierung von ENp wurden keine anderen Interaktionen der Proteine ENp, RTp und HCp des ORF2 mit sich selbst oder Teilen von DdTFIIIB gefunden. Ebenso waren keine Wechselwirkungen zwischen ORF1p und ORF2p detektierbar (vgl. Tab. 3.1). Die ausreichende Expression der Proteine wurde durch eine Western Blot Analyse kontrolliert (siehe Abb. 3.11).

### 3.4.3 Interaktionen zwischen TRE5-A.1-ORF1p und DdTFIIIB

Über die Struktur oder Funktion von ORF1p während der Retrotransposition von TRE5-A ist wenig bekannt. ORF1p zeigt keine signifikanten Homologien zu bekannten Proteinen, auch nicht zu ORF1-Proteinen anderer TREs oder Non-LTR-Retrotransposons aus anderen Organismen. Die relative Position von ORF1 im TRE5-A-Element im Vergleich mit anderen Non-LTR-Retrotransposons lässt die Vermutung zu, dass es sich um ein RNA-bindendes Protein handelt, das mit Transkripten des eigenen Elementes Ribonukleoproteinpartikel bilden kann und somit ein Teil des Retrotranspositionsintermediats darstellt [WINCKLER *et al.*, 2002a].

TRE5-A.1-ORF1p interagiert im B2H-II-System mit allen drei Untereinheiten von DdTFIIIB. Als am stärksten erwies sich die Wechselwirkung DdTBP-ORF1p, die auch hochstringenten Selektionsbedingungen (6 mM 3-AT) standhielt (Abb. 3.13). Schwächer, aber ebenfalls signifikant waren die Interaktionen mit DdBrfC sowie mit DdBdp1, die aber vermutlich zumindest teilweise durch eine geringere Expressionsrate der beiden DdTFIIIB-Proteine bedingt sind (vgl. Abb. 3.11 und Abb. 3.13). Eine Wechselwirkung zwischen DdBrfN und ORF1p wurde nicht beobachtet.



**Abb. 3.13: Interaktionen zwischen DdTFIIIB und ORF1p und ORF2p**

DdTBP und DdBrfC bilden einen stabilen Komplex. ORF1p tritt mit jeder der drei DdTFIIIB-Untereinheiten in Wechselwirkung. Interaktionen zwischen DdTFIIIB und ORF2p konnten auch bei der geringsten 3-AT-Menge nicht beobachtet werden. ORF1p bildet stabile Homodimere. In den Plasmidkombinationen sind die pBT-Konstrukte auf der linken Seite und die pTRG-Konstrukte auf der rechten Seite gelistet.

#### 3.4.4 Keine Interaktion zwischen TRE5-A.1-ORF1p und humanem TBP

Aufgrund der großen Ähnlichkeit von HsTBP und DdTBP war es naheliegend direkt zu testen, ob HsTBP in der Lage ist mit ORF1p zu interagieren.

Das Gen für HsTBP liegt im humanen Genom auf Chromosom 6q27. Das Protein besitzt N-terminal einen Polyglutaminsäurestretch, der je nach Allel aus 29-42 Glutaminsäuremolekülen besteht [PURRELLO *et al.*, 1994]. Eine Verlängerung dieses Glutaminsäure-Abschnitts ist mit verschiedenen zerebellären Erkrankungen assoziiert, unter anderem mit SCA 17 [FUJIGASAKI *et al.*, 2001; NAKAMURA *et al.*, 2001].

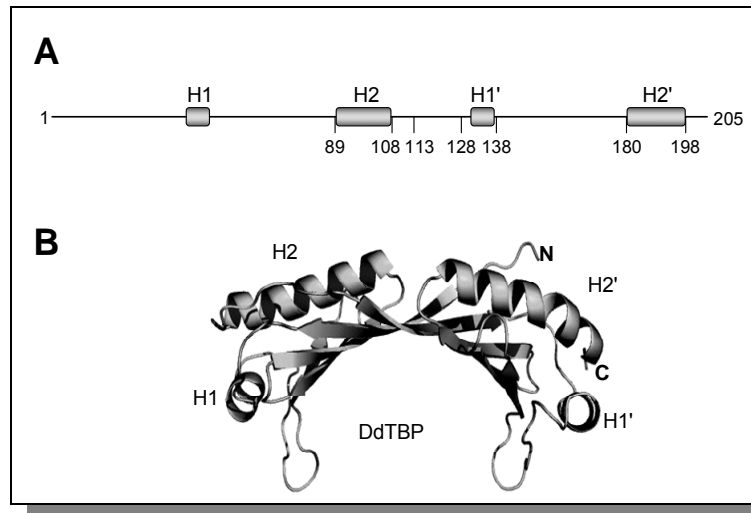
Zunächst wurde versucht, die vollständige kodierende Sequenz von HsTBP in den Zweihybrid-Vektor pBT zu klonieren. Dies gelang jedoch nicht, da das  $\lambda$ cl/HsTBP-Fusionsprotein wahrscheinlich für die Bakterienzellen toxisch war und es dadurch während der Klonierung zu einer Gegenselektion kam. Deshalb wurde mit einem zweiten Primerpaar ein N-terminal verkürztes Fragment ohne den Poly-Glutaminsäure-Stretch amplifiziert. Dieses Protein entsprach in seiner Länge in etwa dem *Dictyostelium discoideum*-TBP (nicht gezeigt) und konnte ohne weitere Schwierigkeiten kloniert werden. Da das exprimierte Polypeptid offenbar keine toxische Wirkung auf die Bakterien hatte und sich die Sequenzhomologie zwischen HsTBP und DdTBP vor allem im mittleren und C-terminalen Bereich befand (vgl. Abb. 3.2), wurde für alle weiteren Versuche das verkürzte Protein HsTBP02/03 verwendet.

HsTBP wurde im bakteriellen Zweihybrid-System auf Interaktion mit ORF1p getestet. Wie Abb. 3.18 zeigt, war allerdings keine Wechselwirkung zwischen den beiden Proteinen nachweisbar.

#### 3.4.5 Kartierung der Interaktionsflächen von DdTBP

Aufgrund der relativ hohen Homologie zwischen HsTBP und DdTBP war es möglich, anhand der Kristallstruktur von HsTBP (Model SwissProt P20226) ein Modell für DdTBP zu errechnen [NIKOLOV *et al.*, 1996]. Demnach besitzt auch DdTBP eine beinahe spiegelsymmetrische sattelförmige Struktur, deren konvexe Oberseite durch vier  $\alpha$ -Helices (H1, H2 bzw. H1', H2') geformt wird und deren Unterseite aus zehn antiparallelen  $\beta$ -Faltblättern (S1 - S5 bzw. S1' - S5') besteht:





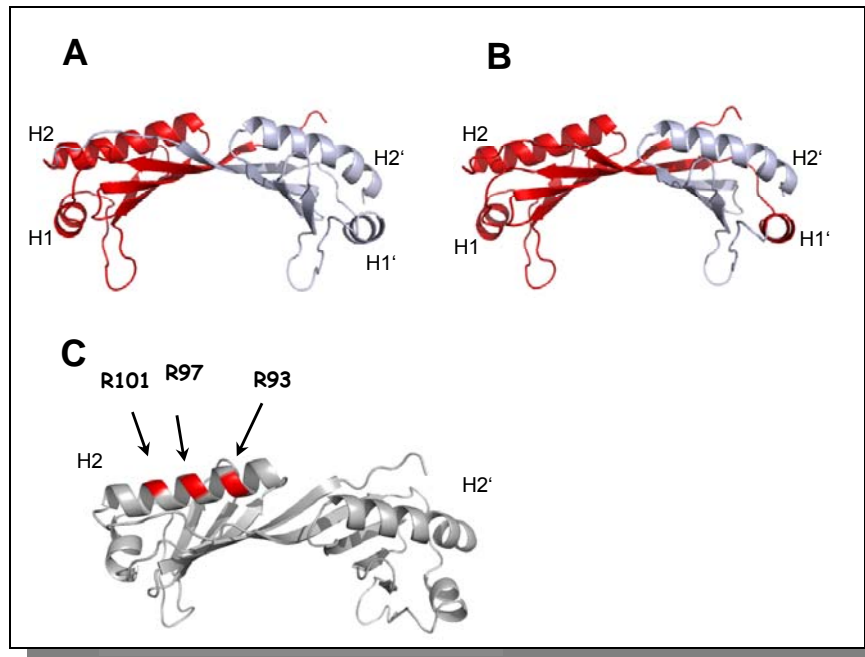
**Abb. 3.14: Struktur von DdTBP**

**A:** Lage der  $\alpha$ -Helices in der Primärstruktur und **B:** in der dreidimensionalen Struktur von DdTBP. Die Struktur des Modells wurde mit Hilfe des Programms SWISS MODEL [SCHWEDE *et al.*, 2003] und *Template File* 1CDW.png des humanen TBP erstellt.

Die Sequenzen des *Core*-Bereichs von DdTBP und HsTBP sowie TBP aus *S. cerevisiae* sind zu rund 70% identisch, so dass es nahe lag, zunächst bekannte Interaktionsflächen aus Hefe und Mensch zu untersuchen.

DdTBP wurde an Position 108 in zwei symmetrische Fragmente  $\text{DdTBP}^{1-108}$  und  $\text{DdTBP}^{113-205}$  geteilt und anschließend im B2H-II-Assay gegen DdBrfC und ORF1p getestet. Es zeigte sich, dass die eingesetzten Proteine unterschiedliche Bereiche von DdTBP kontaktierten. Während DdBrfC mit dem N-Terminus von DdTBP in Wechselwirkung trat und keine Interaktion mit  $\text{DdTBP}^{113-205}$  aufwies, verhielt sich ORF1p komplementär dazu (Abb. 3.17).

Die Affinität von DdBrfC zu  $\text{DdTBP}^{1-108}$  war allerdings deutlich geringer als zum vollständigen Protein. Offensichtlich gab es auf DdTBP weitere Kontaktflächen. Deshalb wurde  $\text{DdTBP}^{1-108}$  um die  $\alpha$ -Helix H2 sowie einige direkt angrenzende Aminosäuren verlängert und die resultierende Mutante  $\text{DdTBP}^{1-136}$  im B2H-II-System getestet. Mit dieser Mutante konnte die ursprüngliche Interaktionsstärke wieder vollständig hergestellt werden, während mit der separaten  $\alpha$ -Helix H2 ( $\text{DdTBP}^{82-110}$ ) zwar reproduzierbare, aber nur schwache Wechselwirkungen nachgewiesen werden konnten.



**Abb. 3.15: Modelle der DdTBP-Mutanten**

**A.** DdTBP<sup>1-108</sup> umfasst die linke Teildomäne von DdTBP, **B.** DdTBP<sup>1-136</sup> enthält zusätzlich Helix H1' der rechten Teildomäne, **C.** Die drei Arginine R93, R97 und R101 auf Helix H2 stellen die Hauptkontakte für DdBrfC auf DdTBP dar.

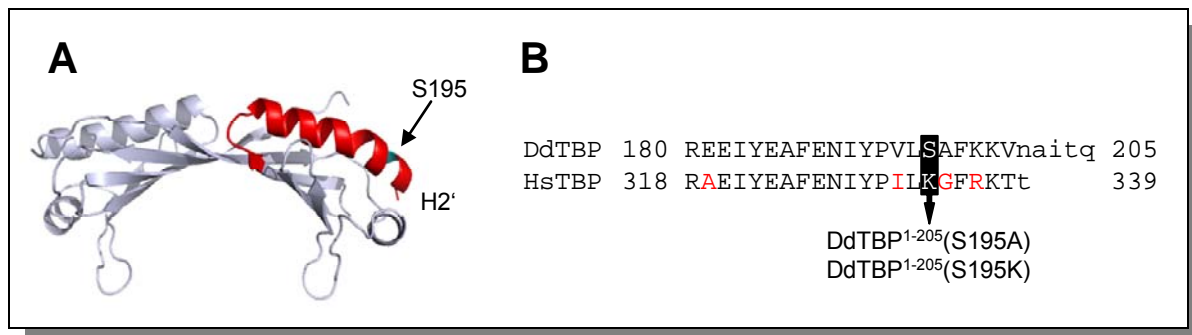
In HsTBP unterbindet eine dreifache Punktmutation (R<sup>231</sup>E, R<sup>235</sup>E, R<sup>239</sup>S) auf der konvexen Seite der  $\alpha$ -Helix H2 die Interaktion mit HsBrf1 [SHEN *et al.*, 1998]. Durch Einzelstrangmutagenese wurden an den korrespondierenden Stellen von DdTBP die gleichen Mutationen eingeführt (R<sup>93</sup>E, R<sup>97</sup>E, R<sup>101</sup>S) und die korrekte Sequenz durch Sequenzierung verifiziert.

Diese Triple-Mutante DdTBP<sup>1-205</sup>(EES) war wie erwartet nicht mehr zu einer Wechselwirkung mit DdBrf1 in der Lage. Eine mangelhafte Expression des Proteins wurde in Westernblot-Analysen ausgeschlossen (Daten nicht gezeigt).

DdTBP<sup>1-205</sup>(EES) interagiert weiterhin mit ORF1p (Abb. 3.17). Dies stimmt mit der ersten groben Kartierung überein, nach der die Kontaktfläche für ORF1p im C-terminalen Bereich von DdTBP liegt.

Trotz der großen strukturellen Ähnlichkeit zwischen HsTBP und DdTBP konnte keine Wechselwirkung zwischen dem humanen Protein und ORF1p detektiert werden (vgl. Abschnitt 3.4.4). Bei genauerer Betrachtung der beiden TBP-Sequenzen wurde deutlich, dass sich Unterschiede vor allem in den beiden  $\alpha$ -Helices H1' und H2' des C-terminalen Kernbereichs von TBP befinden. Die Sequenzen der jeweils 18 AS langen Helix H2' in HsTBP und DdTBP unterscheiden sich lediglich in fünf Positionen (Abb. 3.16, B). Eine sehr auffällige, nicht-konservative Sequenzabweichung stellt Lysin<sup>333</sup> aus HsTBP dar, die Serin<sup>195</sup> in DdTBP entspricht. Wie in Abb. 3.16

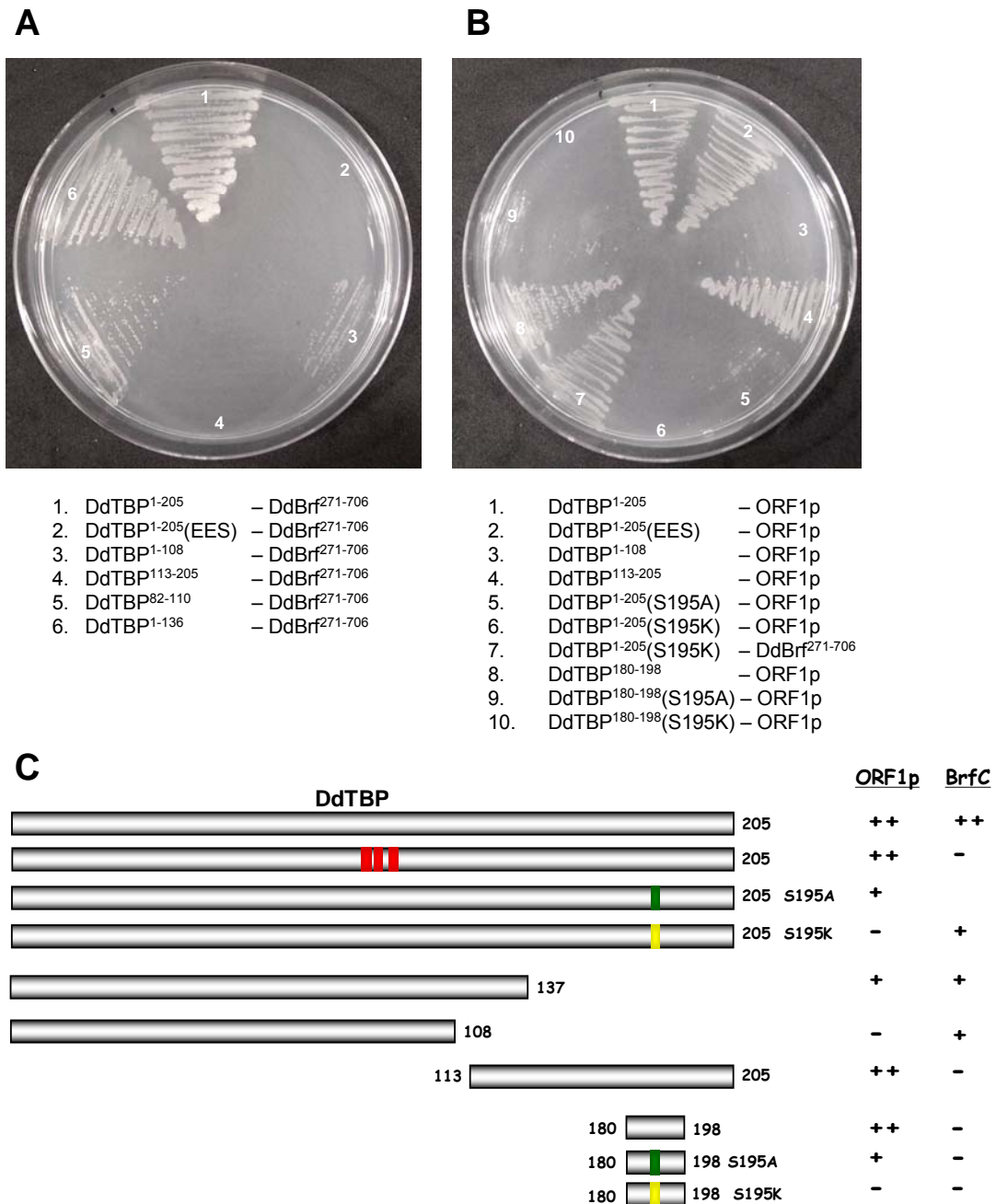
angedeutet sitzen beide Aminosäuren exponiert auf der konvexen Oberseite der jeweiligen Helix.



**Abb. 3.16: Helix H2' in DdTBP und HsTBP**

**A.** Die Mutante DdTBP<sup>180-198</sup> umfasst die rot markierte Helix H2'. S<sup>195</sup> ist blau markiert. **B.** Aminosäuresequenzen der C-Termini von DdTBP und HsTBP. Die Mutationen S195A und S195K wurden sowohl im vollständigen Protein als auch in der separaten Helix experimentell überprüft.

Die Mutation von Serin<sup>195</sup> zu Lysin in DdTBP–Helix H2' (DdTBP<sup>180-198</sup> (S195K)) unterband die Interaktion von DdTBP und ORF1p vollständig. Tauschte man dagegen Serin<sup>195</sup> gegen die neutrale Aminosäure Alanin aus (DdTBP<sup>180-198</sup> (S195A)), blieb eine schwache Interaktion detektierbar (Abb. 3.17). Das gleiche Bild ergab sich, wenn die Mutationen nicht nur in der separaten Helix getestet wurden, sondern auch im vollständigen Protein (DdTBP(S195K) bzw. DdTBP(S195A)) (Abb. 3.17). Serin<sup>195</sup> ist demnach eine wichtige Aminosäure für die Interaktion von DdTBP mit ORF1p. Austausch gegen neutrale Aminosäuren wie Alanin werden im B2H-II-System toleriert, extreme Austausch gegen geladene Aminosäuren dagegen nicht. Die Aktivität oder Expression von DdTBP(S195K) wurde durch die Punktmutation nicht beeinträchtigt, da das Protein noch in der Lage war mit DdBrfC zu interagieren.

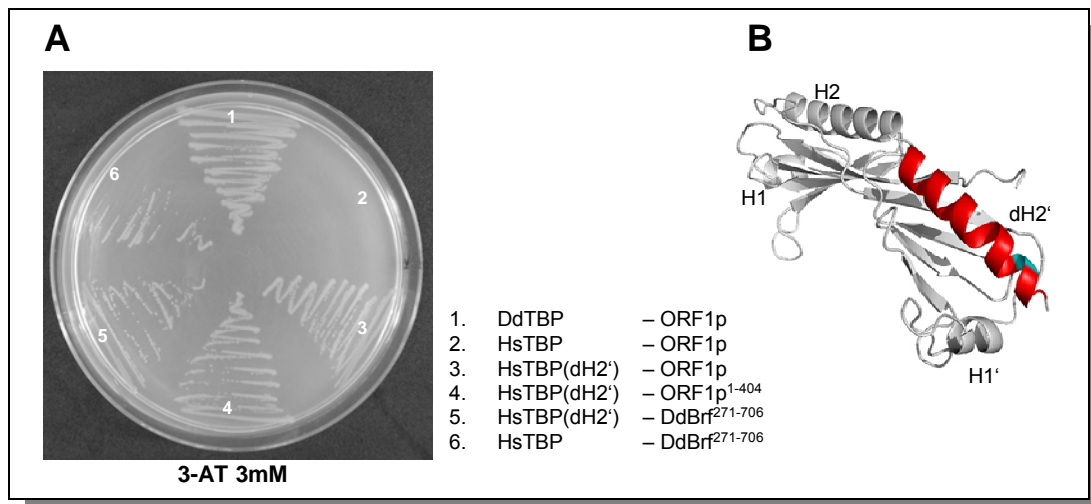


### Abb. 3.17: Interaktionsanalysen der DdTBP-Mutanten

**A.** Zweihibrid-Assays zeigen die Interaktionen der DdTBP-Mutanten mit DdBrfC und **B.** mit ORF1p. Die Selektionsplatten enthielten 5 mM 3-AT. **C.** Zusammenfassung der Interaktionen der getesteten DdTBP-Mutanten.

Die vorangegangenen Versuche zeigten, dass vor allem die  $\alpha$ -Helix H2' für die Interaktion zwischen ORF1p und DdTBP verantwortlich ist. Deshalb wurde durch Einzelstrangmutagenese ein chimäres HsTBP-Protein mit der  $\alpha$ -Helix H2' aus DdTBP generiert (HsTBP(dH2')) und ebenfalls im Zweihibrid-Assay gegen ORF1p und DdBrfC getestet. Wie in Abb. 3.18 dargestellt konnte zwar eine Interaktion zwischen der Chimäre und ORF1p detektiert werden, allerdings nur bei einer 3-AT-Konzentration von 3 mM. Demnach scheint ORF1p außer Kontakten in Helix

H2' noch andere Kontaktflächen auf DdTBP zu haben. Positive Wechselwirkungen des chimären Proteins mit ORF1p<sup>1-404</sup> und DdBrfC bestätigten noch einmal die vorangegangenen Versuchen gefundenen Interaktionsflächen für ORF1p und DdBrfC auf DdTBP.



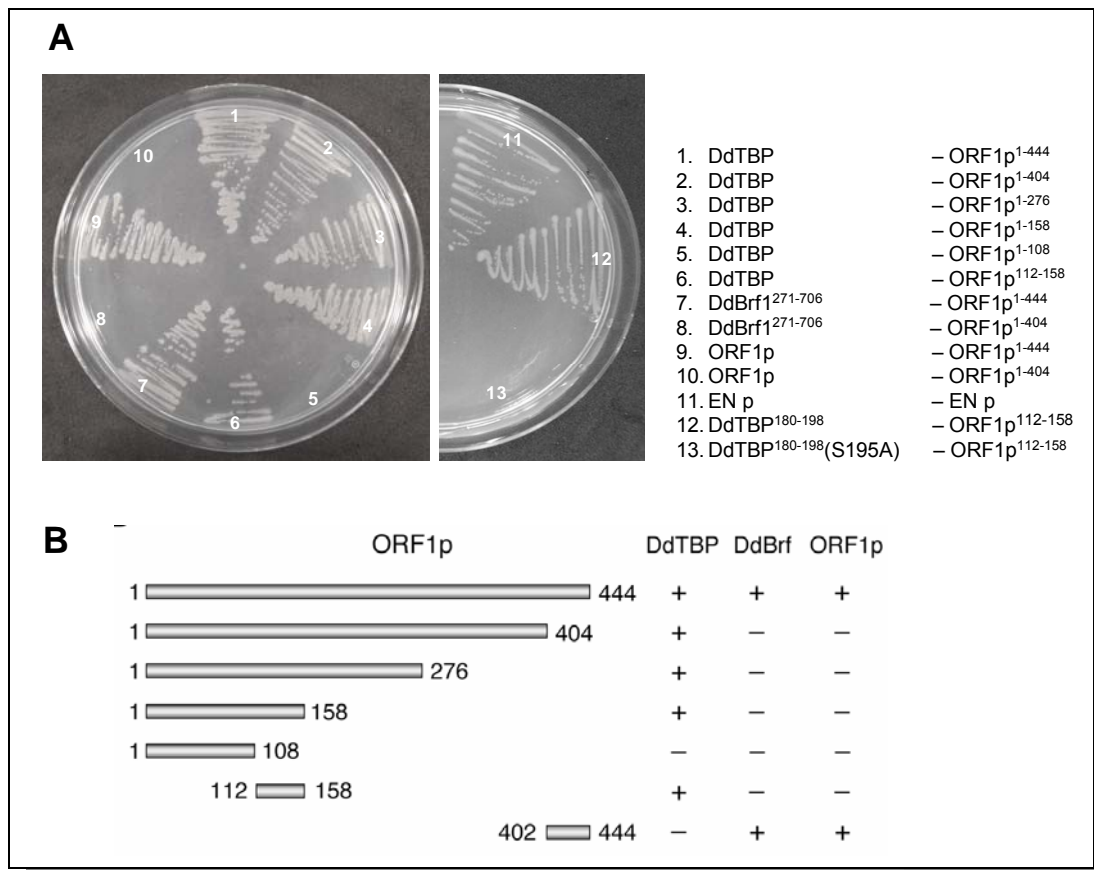
**Abb. 3.18: Two Hybrid Versuche mit der Chimäre HsTBP(dH2')**

**A.** Das chimäre Protein HsTBP(dH2') ist im Gegensatz zu HsTBP in der Lage, mit ORF1p zu interagieren. Eine Wechselwirkung mit DdBrfC konnte ebenso nachgewiesen werden. **B.** Die Helix dH2' ist im Modell von DdTBP rot markiert, S<sup>195</sup> blau. Humane Sequenzen sind grau dargestellt. Generiert mit PyMOL.

### 3.4.6 Kartierung der Interaktionsflächen von TRE5-A.1-ORF1p

Zur näheren Bestimmung der Kontaktflächen auf ORF1p wurden durch Einzelstrangmutagenesen die Deletionsmutanten ORF1p<sup>1-404</sup>, ORF1p<sup>1-276</sup>, ORF1p<sup>1-158</sup> und ORF1p<sup>1-108</sup> generiert. Die Mutagenesen wurden im pTNT-Vektor durchgeführt, das Plasmid einem *NotI*-Verdau unterzogen und das erhaltene Fragment dann in pTRG kloniert.

Wie in Abb. 3.19 dargestellt, interagierte DdTBP mit allen Deletionsvarianten von ORF1p außer ORF1p<sup>1-108</sup>. Folglich muss die Interaktionsfläche von ORF1p mit DdTBP zwischen den AS 108-158 liegen. Dies konnte durch positive Testung des Fragments ORF1p<sup>112-158</sup> sowohl mit DdTBP als auch mit der isolierten Helix H2' (DdTBP<sup>180-198</sup>) belegt werden. DdTBP<sup>180-198</sup>(S195K) (siehe Abschnitt 3.4.5) dagegen trat wie erwartet nicht mit ORF1p<sup>112-158</sup> in Wechselwirkung:



**Abb. 3.19: Bestimmung der Interaktionsflächen von ORF1p**

Zusammenfassende Darstellung der im Zweihybrid-System getesteten ORF1p-Deletionen. (+) bezeichnet Wachstum der Transformanten bei 5 mM 3-AT.

Analoge B2H-II-Versuche mit DdBrfC gegen die Deletionsreihe von ORF1p zeigten, dass die Interaktion zwischen diesen beiden Proteinen bereits durch Deletion der letzten 40 AS von ORF1p aufgehoben werden konnte (vgl. Abb. 3.19). Die separate *Coil*-Domäne von ORF1p (ORF1p<sup>402-444</sup>) dagegen zeigte eine Wechselwirkung mit DdBrfC (Daten nicht gezeigt). D.h. die *Coil*-Struktur, die für die Dimerisierung von ORF1p verantwortlich ist, stellt auch die Kontaktfläche für DdBrf1 dar, so dass beide Proteine möglicherweise um diese Bindungsstelle konkurrieren.

### 3.4.7 Interaktion von TRE5-B-ORF1p mit DdTFIIIB

Die Retrotransposons TRE5-A und TRE5-B sind entfernt miteinander verwandt und stammen vermutlich von einem gemeinsamen Vorläufer ab [GLÖCKNER *et al.*, 2001a]. Allerdings sind die Sequenzen von TRE5-A.1-ORF1p und TRE5-B-ORF1p lediglich zu 25% identisch (Abb. 3.20). Beide Retrotransposons integrieren etwa 45-50 bp vor tRNA-Genen, so dass die Vermutung nahe lag, dass TRE5-A und TRE5-B ähnliche Mechanismen zur Erkennung des Zielortes verwenden [WINCKLER *et al.*, 2002b]. Um diesen Gedanken zu überprüfen, wurde die kodierende Sequenz von ORF1p aus

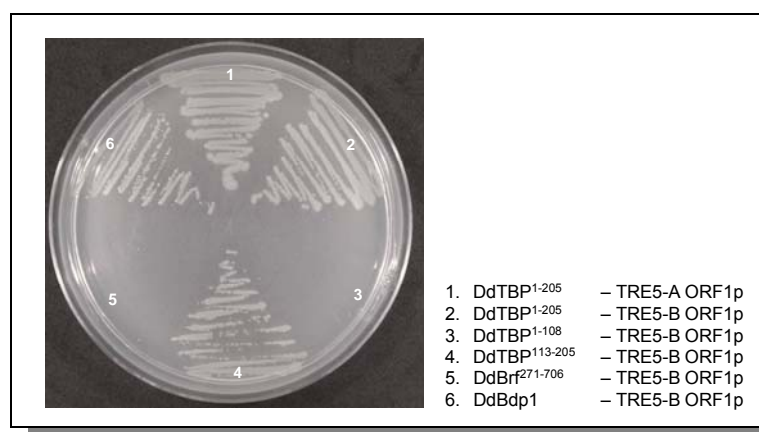
TRE5-B als *NotI*-Fragment in den Vektor pTRG kloniert und anschließend im B2H-II-System gegen alle drei Untereinheiten von DdTFIIIB getestet.

TRE5-A	1	MKI	TNNT	NIKQ	HKCLO	KISB	IVDK	TOLKK	QFRVY
TRE5-B	1	MNKLANTIMTESTETLADKIKKSNELRTRKFKYLKNTYRTLDRKISBAIYHTQIELKQVRFM							
TRE5-A	37	ININHEPDDKIKKDLKSLDKKDVLTKSNNTFKILTDSISTVIYFVNMPDAELKQOFIT							
TRE5-B	61	INIPDEISNNLKAELRHLQGGDKIISKNNKFKILTD CRTQVIKYSNVFNGVLKGVHRAI							
TRE5-A	97	LRFKNMYTDLDWYKSLTEIEVVIENEGCKLIKSEIKGETLTIKTVKRVVEEYDTIIELDN							
TRE5-B	121	IKPKKIFKTFNWKSMDIVKTLHEEGLEIKKLDNLGENLIHHSFNSIISP---EHKQD							
TRE5-A	157	FTLIGHSNKGRDQTSLEPNOEENKNNKQSDSQPKS--TSEQIPKQGRRTIIPPIITSS							
TRE5-B	177	HIHCKHITILFLTTAETKNQRKT-NKEIEDEISENSKITHELPNPEIKENARINIKSKEE							
TRE5-A	215	SNDKTFMALDDTAQKVTDASLENVDELTRIEINEINTOSIKNOEIKTKSEKPIEPPSPYNT							
TRE5-B	236	SNHNTSESKNNNNKLENRRGNRKEIN---NDNIKDKSKNEKVENENEIENKKTITIRAP							
TRE5-A	275	PIKQHGERTKRLTTEEKLREVNDIFDFTSPNTDAIRKIEFPLSIPSNQWFFSSPIRHIMSF							
TRE5-B	293	TTTTTETNTINNIERKQVLRVPPITVTVYSVDS-REKRKMELTEDEEAEVVRKMPFENE							
TRE5-A	335	SPTYGTPSKTIDSPSKLLNTPKKNQDQKTNRTNKTPVKPKLGEYGQKSIKTKMDEDIDNF							
TRE5-B	352	EENRIFEEKINNQKKKE--TPRLASIKK--RQTMSPIQNPESPATPETVRRKRSNSSPNP							
TRE5-A	395	NKEILKQISKSKNPTNTQDQQHNEQSKDQLEDKINOKLQKVDKN-LPQIKL-							
TRE5-B	408	TTNHQEISTPYKTPTKITK---EES-DLMEAAALRDKLQKVLQDILPSVSVFV							

**Abb. 3.20: Clustal X-Alignment der ORF1-Proteine aus TRE5-A und TRE5-B**

Schwarz unterlegte Aminosäuren sind in beiden Proteinen identisch, bei den grau hinterlegten Aminosäuren handelt es sich um konservative Austausche. Die Proteine ORF1p aus jeweils TRE5-A und TRE5-B sind schwach konserviert. Die Sequenz in ORF1p aus TRE5-A.1, die mit DdTBP interagiert, ist blau markiert.

Wie in Abb. 3.21 dargestellt zeigte TRE5-B-ORF1p mit DdTBP eine deutliche Interaktion, die auf den C-terminalen Bereich DdTBP<sup>113-205</sup> eingegrenzt werden konnte. Außerdem band TRE5-B ORF1p schwach an DdBdp1. Eine Wechselwirkung mit DdBrf1 konnte allerdings auch geringer Stringenz des Assays (3 bzw. 4 mM 3-AT) nicht nachgewiesen werden. Ebenso wurde keine Bindung zwischen TRE5-A-ORF1p und TRE5-B-ORF1p beobachtet (Daten nicht gezeigt).



**Abb. 3.21: Two Hybrid-Versuche mit TRE5-B ORF1p**

TRE5-B ORF1p zeigte stabile Interaktionen mit DdTBP und DdBdp1 bei 5 mM 3-AT

### 3.5 Interaktionen von ORF1p mit TBP und Brf 1 *in vitro*

Die im Zweihybrid-System identifizierten Wechselwirkungen zwischen ORF1p und den DdTFIIIB-Untereinheiten sollten mit einer biochemischen Methode *in vitro* bestätigt werden. Dazu wurde die Sequenz von TRE5-A.1 ORF1 in pMALc2X(Not) kloniert und das Plasmid als lösliches Fusionsprotein MBP-TBP exprimiert. Die potentiellen Interaktionspartner DdTBP sowie DdBrf1 wurden unter Verwendung von radioaktivem Methionin aus dem Vektor pTNT *in vitro* transkribiert und translatiert. Mit Hilfe der [<sup>35</sup>S]-markierten Proteine wurden in *Pulldown*-Versuchen Interaktionen mit ORF1p analysiert.

#### 3.5.1 Rekombinante Expression von ORF1p

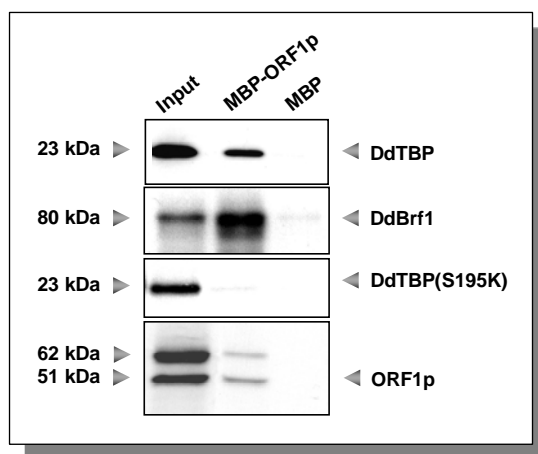
Die DNA-Sequenz von TRE5-A.1 ORF1 wurde als *NotI*-Fragment in den bakteriellen Expressionsvektor pMALc2X(Not) kloniert und als lösliches Fusionsprotein mit dem Maltose-bindenden Protein (MBP) exprimiert. Vorangegangene Arbeiten zeigten, dass die Löslichkeit von Proteinen durch die Fusionierung mit MBP deutlich erhöht werden kann [KAPUST & WAUGH 1999]. MBP fungiert dabei wahrscheinlich ähnlich wie ein Chaperon und verhindert durch die Bindung an Aggregationsintermediate die Selbstaggregation des Fusionspartners [Fox *et al.*, 2001]. Der Vektor pMALc2X besitzt keine Signalsequenz zur Sekretion, so dass gebildetes Protein im Zytoplasma akkumuliert. Um die kodierende Sequenz von ORF1 als *NotI*-Fragment in den IPTG-induzierbaren Expressionsvektor pMALc2X klonieren zu können, wurde zunächst eine *NotI*-Schnittstelle mittels Einzelstrangmutagenese in den Vektor eingefügt. Zur Unterscheidung wurde der neue Vektor in pMAL-c2X(Not) umbenannt. Die Expression des MBP-ORF1-Proteins erfolgte nach Transformation des Plasmids pMALc2X-ORF1 in W3110 bei 37 °C über 2 Stunden. Die Zugabe von 0,2% Glukose zum Medium unterdrückte die Expression einer Amylase, die die Bindung von MBP-Fusionsproteinen an die Säule reprimiert.

#### 3.5.2 MBP-*Pulldown*-Versuche

Um die Bindung der TFIIIB-Untereinheiten DdTBP, DdBrf1 und DdBdp1 an ORF1p biochemisch nachzuweisen, wurden MBP-*Pulldown*-Versuche durchgeführt. Dazu wurden Amylose-Säulen jeweils frisch mit etwa der gleichen Menge MBP-ORF1 bzw. MBP als Negativexperiment beladen und anschließend mit <sup>35</sup>S-markierten TFIIIB-Proteinen inkubiert und präzipitiert. Da vor allem radioaktiv markiertes DdTBP,



aber auch die anderen *in vitro* translatierten Proteine dazu neigten an die Innenwand des Reaktionsgefäßes zu binden, wurden ausschließlich *Protein-LoBind-Caps* für die Präzipitationsversuche verwendet. Durch SDS-Gelelektrophorese wurden die erhaltenen Proteinkomplexe aufgetrennt und radiologisch analysiert. Die Interaktionen von DdTBP und DdBrf1 mit ORF1p konnten dabei ebenso bestätigt werden wie die Homodimerisierung von ORF1p (Abb. 3.22). Die Mutante DdTBP(S195K) interagierte wie bereits im B2H-II-System nicht mehr mit ORF1p. Es ließ sich außerdem zeigen, dass die C-terminal deletierte Mutante ORF1p<sup>1-404</sup> nicht mehr in der Lage war zu dimerisieren (Daten nicht gezeigt).



**Abb. 3.22: MBP-Pulldown Experimente**

Die radioaktiv markierten Proteine DdTBP, DdBrf1 und ORF1p banden auch *in vitro* an MBP-ORF1. Keine Interaktion fand dagegen zwischen DdTBP(S195K) und MBP-ORF1 statt. ORF1p läuft vermutlich sowohl als Monomer als auch als Dimer. 10% des Markierungsreaktionsansatzes (*Input*) wurden als Mengenvergleich aufgetragen.

Die Interaktion zwischen ORF1p und DdBrfC konnte wahrscheinlich aufgrund einer unzureichenden Markierung der Proteine durch [<sup>35</sup>S]-Methionin nicht im *Pulldown*-Versuch verifiziert werden. Für ein positives Signal mit ORF1p mußte der Versuch mit dem vollständigen DdBrf1 durchgeführt werden. DdBdp1, das prozentual genauso wenige Methionine in seiner Aminosäuresequenz aufweist wie BrfC, konnte ebenfalls nicht detektiert werden (Tab. 3.2). Eine Interaktion der Proteine mit dem MBP-Fusionsanteil wurde in Vorversuchen ausgeschlossen.

	AS	n(Met)	%
DdTBP	205	5	2,4
DdBrf1	706	11	1,6
DdBdp1	458	5	1,1
DdBrfC	437	5	1,1

**Tab. 3.2: Vergleich des Methionin-Gehalts**

Die Interaktionen zwischen ORF1p und DdBrf(271-706) bzw. DdBdp1 konnten wahrscheinlich aufgrund einer unzureichenden Markierung durch [<sup>35</sup>S]-Methionin nicht überprüft werden.

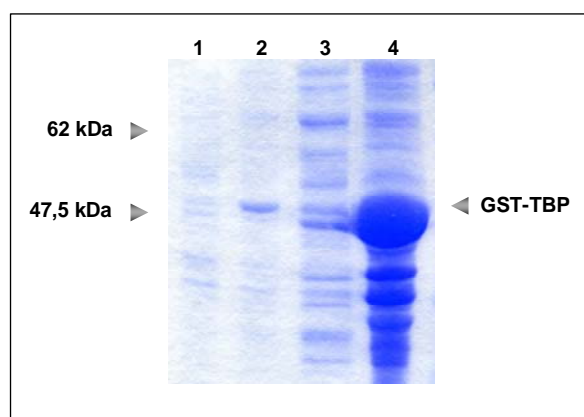
### 3.6 Herstellung eines monoklonalen Antikörpers gegen TBP

Da sich im Lauf der Arbeit das TATA-Box-bindende Protein TBP als ein wichtiger Faktor sowohl für die Assemblierung des Transkriptionsfaktors TFIIB als auch für die Integration des untersuchten Retrotransposons herausstellte, war es von Interesse, für weitergehende Analysen einen Antikörper gegen DdTBP zu gewinnen. Dieser sollte endogenes TBP aus *D. discoideum* ebenso erkennen wie bakteriell exprimiertes, rekombinantes Protein. Kommerzielle Antikörper waren meist gegen eine N-terminale Region des humanen TBP gerichtet. In diesem Bereich weisen die beiden ansonsten stark konservierten Proteine allerdings die geringste Homologie auf, da der humane Faktor N-terminal einen ausgedehnten Poly-Glutaminsäure-Stretch besitzt.

Zur bakteriellen Expression von TBP wurde ein modifizierter pGEX-5T-Vektor verwendet, indem eine *NotI*-Schnittstelle an das 3'-Ende der GST-Sequenz eingefügt wurde. Dadurch wurden nach bakterieller Expression des Plasmids drei zusätzliche Alanine zwischen GST und dem Fusionsprotein eingeführt. Auch der Vektor pMAIc2x-TBP, der für die Expression des Fusionsproteins MBP-TBP verwendet wurde, trug diese Modifikation (vgl. Abb. 3.29, A).

#### 3.6.1 Expression und Aufreinigung des Antigens

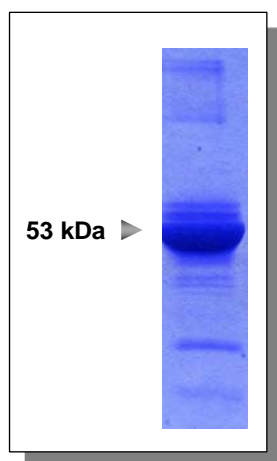
Die DNA-Sequenz von DdTBP wurde als *NotI*-Fragment in den bakteriellen Expressionsvektor pGEX-5T kloniert und als GST-Fusionsprotein exprimiert. Allerdings war es in diesem Vektorsystem nicht möglich, ausreichende Mengen an löslichem Protein zu erhalten, da GST-TBP weitgehend unabhängig von äußeren Expressionsbedingungen und Lyseprotokollen (errechnetes Molekulargewicht 53 kDa) als unlösliches Protein in *Inclusion bodies* vorlag:



**Abb. 3.23: Expression des Antigens GST-TBP in W3110**

Spur 1: Gesamtextrakt (nichtinduziert), Spur 2: Gesamtextrakt (2h induziert mit 0,5 mM IPTG), Spur 3: lösliche Proteine, Spur 4: unlösliche Proteine aus *Inclusion bodies*. Das Fusionsprotein GST-TBP (errechnetes Molekulargewicht 53 kDa) liegt ausschließlich in unlöslicher Form vor.

Aufgrund der Unlöslichkeit des GST-TBP-Fusionsproteins wurde das Antigen unter denaturierenden Bedingungen aufgereinigt. Es wurde ein präparativer Expressionsansatz durchgeführt, ein Aliquot davon über Nacht bei Raumtemperatur in 6 M Guanidinium-HCl gelöst und am folgenden Tag über eine Ni<sup>2+</sup>-NTA-Säule aufgereinigt. Zwar konnte auf diese Weise ein Großteil der unspezifisch gebundenen Proteine entfernt und das GST-TBP mit großer Reinheit gewonnen werden, allerdings war die Ausbeute des Antigens für eine Immunisierung nicht ausreichend. In einem zweiten Ansatz wurde deshalb auf die säulenchromatographische Aufreinigung verzichtet. Es wurde ein weiteres Aliquot der *Inclusion bodies* über Nacht bei Raumtemperatur in 300 µl LowTE und 300 µl Laemmli-Puffer gelöst und über ein präparatives 10%iges PAA-Gel aufgereinigt. Das Gel wurde mit 4 M Natriumacetat-Lösung gefärbt, das Protein durch Elektroelution aus der ausgeschnittenen Bande herausgelöst und das Eluat abschließend gefriergetrocknet. Die Überprüfung des aufgereinigten Proteins zeigte nur geringe Verunreinigungen durch *E. coli*-Proteine und konnte anschließend für die Immunisierung der Mäuse verwendet werden (Abb. 3.24). Die gewählte Methode stellte einen Kompromiss zwischen hinreichender Reinheit und notwendiger Proteinmenge dar.



**Abb. 3.24: Aufgereinigtes GST-TBP**

Die Aufreinigung über ein PAA-Gel und anschließende Elektroelution erreicht eine hinreichende Reinheit des gewünschten Antigens.

### 3.6.2 Isolierung des Antikörpers

Die präparierten Milzzellen einer gegen GST-TBP immunisierten Maus wurden mit Myelomzellen fusioniert und die Antikörper-produzierenden Hybridomklone anschließend im ELISA-Screening näher untersucht. Um Antikörperreaktionen gegen den Fusionsteil GST auszuschließen, wurden die positiven Klone zusätzlich gegen das Fusionsprotein MBP-TBP getestet (siehe Abschnitt 3.6.3). Die Klone, die im ELISA ein positives Signal gaben, wurden anschließend in der Western-Blot-Analyse

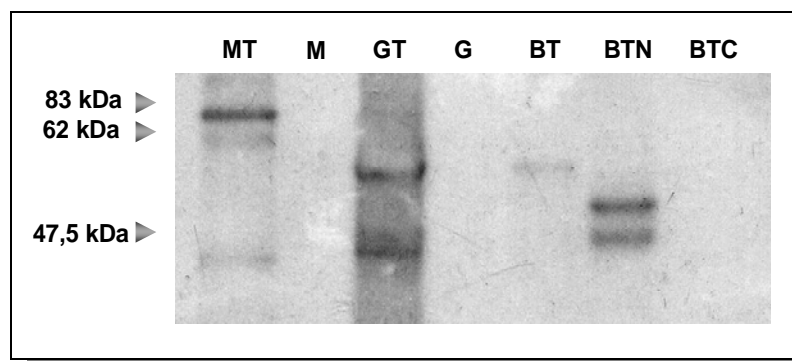
gegen MBP-TBP, MBP, GST-TBP, GST sowie Gesamtzell-Lysat aus *D. discoideum*-AX2 getestet. Klone, die Antikörper gegen *E. coli*-Proteine produzierten, konnten teilweise mit Hilfe von bakteriell exprimiertem MBP bzw. GST eliminiert werden. Aufgrund der Aufreinigungsstrategie (siehe Abschnitt 3.6.1) enthielt das gereinigte Protein noch einen relativ hohen Anteil an *E. coli*-Proteinen, so dass nicht alle Antikörper bezüglich der Spezifität gegenüber TBP oder bakteriellen Proteinen eingeordnet werden konnten.

	7A3	7A7	8B5	8B6	8B12	9B4	9G5	6B3
<b>MBP-TBP</b>	+++	+	+++	+++	-	++	+	-
<b>GST-TBP</b>	+++	++	+++	+++	+	+	+	+++
<b>Spezifität</b>	TBP	TBP	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	?

**Tab. 3.3: Spezifität und Sensitivität verschiedener Antikörper gegen TBP**

Dargestellt sind die Kurzbezeichnungen der monoklonalen Antikörper. Die Empfindlichkeit wurde nach den gemessenen Absorptionen im ELISA-Screening in drei Stufen eingeteilt. Für die Bestimmung der Spezifität wurden außerdem Western Blot Analysen durchgeführt. Antikörpern, die *E. coli*-Proteine ohne Zusammenhang mit den getesteten Fusionsproteinen erkannten, wurde die Spezifität *E. coli* zugeordnet.

Der Hybridomaklon 7A3 sezernierte Antikörper vom Typ IgG<sub>2Ak</sub>. Mit den eingesetzten Fusionsproteinen ergab er eine gute Signalstärke und zeigte nur sehr wenige unspezifische Banden an. Das antigene Epitop von TBP lag im N-terminalen Bereich, da nur das mit pBT-TBP<sup>1-108</sup> exprimierte Protein erkannt wurde. Das C-terminale Fragment von TBP aus pBT-TBP<sup>113-205</sup> rief dagegen keine Antikörperreaktion hervor (Abb. 3.25).

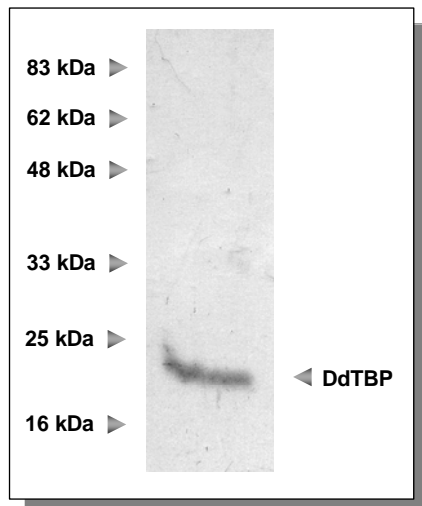


**Abb. 3.25: Western Blot Anti-TBP 7A3**

MT, M: MBP-TBP (67 kDa) bzw. MBP (42,5 kDa), GT, G: GST-TBP (53 kDa) bzw. GST (30 kDa), BT, BTN, BTC: vollständiges Fusionsprotein aus pBT-TBP bzw. der N-/C-Terminus (50/39/37 kDa). Angegeben sind die errechneten Molekulargewichte.

Die Western-Blot-Analysen gegen Gesamtextrakte von *D. discoideum*-Zellen ergaben eine spezifische Bande bei 23 kDa (Abb. 3.26). Dies entsprach dem Molekulargewicht von DdTBP, so dass davon ausgegangen werden konnte, dass der Hybridomaklon 7A3 einen spezifischen Antikörper gegen DdTBP sezernierte.

Um ein vergleichbares Signal zum bakteriell exprimierten Protein zu erhalten war allerdings eine deutlich längere Expositionszeit notwendig.



**Abb. 3.26: Identifikation von DdTBP mit Anti-TBP 7A3**  
Western Blot Analyse. Aufgetragen wurde Gesamt-Zellysate von *D. discoideum* AX2 ( $t_0$ ) und der Film über Nacht exponiert.

### 3.6.3 Charakterisierung eines potentiellen „Epitop-Tags“

Für die Expression von TBP als GST-markiertes Protein war es notwendig, den originalen Vektor pGEX-5T zu modifizieren. Dazu wurde eine *NotI*-Schnittstelle in die MCS eingefügt, so dass nach der Translation zwischen GST und dem Fusionsanteil ein Linker aus drei zusätzlichen Alanin-Molekülen entstand.

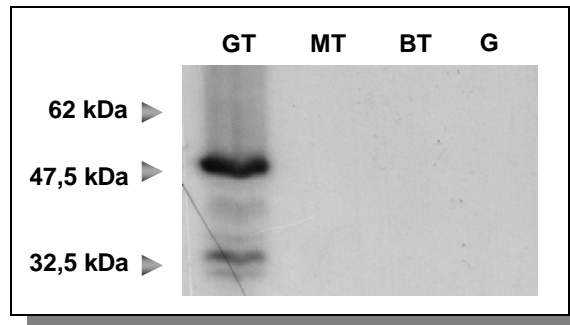
LVPRGS<sup>AAA</sup>MSTATTST

#### Abb. 3.27: Sequenz von GST-TBP

Gezeigt ist die Aminosäure-Sequenz von GST-TBP aus dem Plasmid pGEX(*NotI*)-TBP am Übergang zwischen GST und TBP. Rot: GST-Sequenz, grün: Alanin-Linker, schwarz: TBP-Sequenz

Im *ELISA-Screening* nach DdTBP-spezifischen Antikörpern fielen einige Hybridoma-Klone auf, die mit GST-TBP eine Antikörperreaktion zeigten. Mit anderen Fusionsproteinen wie MBP-TBP fiel sie jedoch aus, ebenso wie mit GST oder MBP alleine (Abb. 3.28). Endogenes TBP aus *D. discoideum* konnte mit diesen Antikörpern nicht detektiert werden. Aufgrund der stärksten und spezifischsten

Antikörperreaktion gegen GST-TBP wurde der Klon 6B3 für die weiteren Analysen ausgesucht.



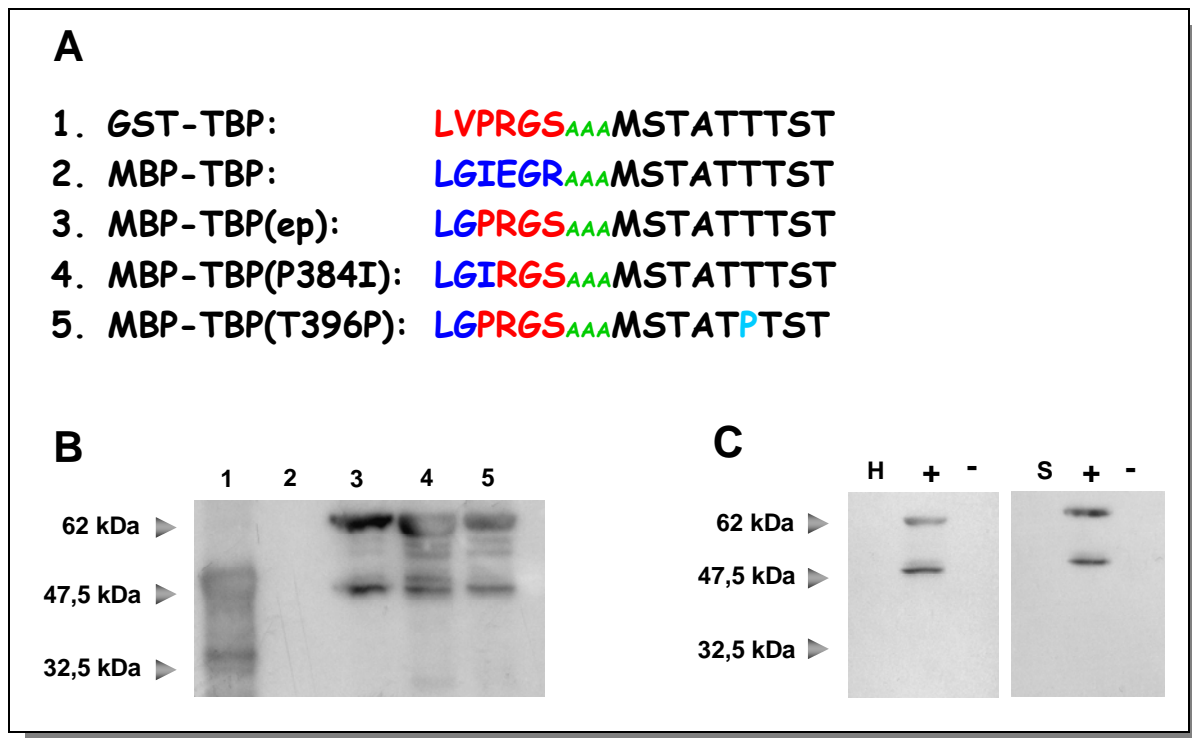
**Abb. 3.28: Western Blot Anti-TBP(ep) 6B3**

GT, G: GST-TBP (53 kDa) bzw. GST (30 kDa), MT: MBP-TBP (67 kDa), BT: vollständiges Fusionsprotein aus pBT-TBP (50 kDa). Angegeben sind die errechneten Molekulargewichte. Der Antikörper 6B3 erkennt nur das Antigen GST-TBP, rekombinante TBP-Proteine mit einem anderen Fusionsteil oder GST alleine werden nicht detektiert.

Dieser Antikörper erkannte offenbar ein Epitop, das sich über einen C-terminalen Teil von GST, die drei zusätzlichen Alanin-Moleküle und einer N-terminale Region von TBP erstreckte.

Um dies nachzuweisen, wurden in MBP-TBP im Bereich des C-Terminus von MBP vier Aminosäuren gegen die entsprechenden vier Aminosäuren aus GST ausgetauscht (siehe Abb. 3.29, A). Wie erwartet wurde dieses Protein MBP-TBP(ep) von den Antikörpern des Hybridomaklons 6B3 mit der gleichen Intensität erkannt wie das Antigen GST-TBP (siehe Abb. 3.29, B).

Zur näheren Charakterisierung des vom Klon 6B3 gebundenen Epitops wurden Einzelstrangmutagenesen an der DNA-Sequenz von MBP-TBP(ep) durchgeführt. Durch die Mutation P384I wurde auf N-terminaler Seite Pro<sup>384</sup> aus GST wieder gegen die ursprüngliche Aminosäure Isoleucin aus MBP ausgetauscht. Der Austausch von Threonin gegen die strukturbrechende Aminosäure Prolin an Position 396 sollte den C-Terminus des Epitops in TBP markieren.



### Abb. 3.29: Bestimmung des antigenen Epitops

Anti-TBP(ep) 6B3 erkennt eine Mischsequenz, die sich aus den letzten drei AS von GST, dem Alanin-Linker und einer N-terminalen TBP-Sequenz zusammensetzt. **A.** Aminosäuresequenzen der (mutierten) Proteine: rot: GST-Sequenz, grün: Alanin-Linker, der durch Einführung der NotI-Schnittstelle in die Vektoren pGEX-5T und pMALc2x entsteht, blau: MBP-Sequenz, schwarz: TBP-Sequenz, hellblau: T396P-Mutation in TBP. **B.** Western Blot. Aufgetragen wurde Gesamtzelllysate. Spur 1: GST-TBP (53 kDa), Spur 2: MBP-TBP (67 kDa), Spur 3: MBP-TBP(ep) (67kDa), Spur 4: MBP-TBP(P384I) (67 kDa), Spur 5: MBP-TBP(T396P) (67 kDa). Alle drei MBP-TBP-Mutanten werden vom Antikörper Anti-TBP(ep) 6B3 erkannt. **C.** Western Blot. Aufgetragen wurde Gesamtzelllysate. H: AH109 Hefezellen, S: 293 wt Säugerzellen. Als Positivkontrolle (+) wurde MBP-TBP(ep) aufgetragen, als Negativkontrolle (-) MBP. Die Analysen zeigen keine Kreuzreaktionen mit Proteinen aus Hefe- oder Säugerzellen.

Es konnte gezeigt werden, dass beide Mutantenproteine MBP-TBP(P384I) und MBP-TBP(T396P) eine genau so starke Antikörperreaktion hervorriefen wie das Antigen GST-TBP, so dass bei Abschluss der Arbeit das antigene Epitop auf folgende Aminosäure-Sequenz reduziert werden konnte (Abb. 3.29, A, B):

**RGS<sup>AAA</sup>MSTAT**

Eine Kreuzreaktion des Antikörpers fand weder mit Proteinen aus Säuger- noch aus Hefezellen statt (Abb. 3.29, C). Der charakterisierte „Tag“ könnte somit zum Beispiel für die Markierung von säugerzell-exprimierten Proteinen verwendet werden.

## 4 Diskussion

TRE5-A.1-Retrotransposons aus *D. discoideum* erkennen ihren Integrationsort mit Hilfe des Pol III-Transkriptionsapparates. In der vorliegenden Arbeit werden Daten präsentiert, die zeigen, dass DdTFIIB ein essentieller Faktor für die Retrotransposition von TRE5-A.1 ist und dass das Retrotransposon durch direkte Interaktion seines ORF1-Proteins mit DdTFIIB an seinen Integrationsort herangeführt wird.

### 4.1 Das bakterielle Zweihybrid-System

Zweihybrid-Systeme sind sehr leistungsfähige Verfahren zur Identifizierung von Protein-Interaktionspartnern. Da das System nur auf physikalischen Bindungen beruht, kann eine breite Variation verschiedener Protein-Protein-Interaktionen durch die Verwendung eines einzigen allgemeinen Protokolls detektiert und charakterisiert werden.

In dieser Arbeit wurde anfangs das *BacterioMatch*<sup>TM</sup> *Two-Hybrid*-System I (B2H-I) verwendet, das auf einer Arbeit von Dove, Joung und Hochschild basiert [DOVE *et al.*, 1997]. Positive Interaktionen führen zu Carbenicillin-resistenten Klonen. Eine zusätzliche Kontrolle erfolgt durch eine Blaufärbung positiver Klone auf X-Gal-Platten, die mit dem  $\beta$ -Galaktosidase-Inhibitor Phenylethyl- $\beta$ -D-thiogalaktosid (PETG) versetzt werden. Zur Detektion schwächerer Interaktionen ist der Assay ungeeignet, da die Reportergene – zwar schwächer – aber auch ohne die transkriptionelle Aktivierung durch ein interagierendes Proteinpaaar exprimiert werden. Die basale Transkription des *amp*<sup>R</sup>-Gens konnte auch durch eine deutlich höhere Carbenicillin- sowie eine höhere PETG-Konzentration nicht hinreichend kompensiert werden.

Die zweite Version des bakteriellen Zweihybrid-Systems B2H-II verwendet eine *HIS3-aadA*-Reportergenkassette. Proteininteraktionen werden durch transkriptionelle Aktivierung des *HIS3*-Gens detektiert, das den Klonen das Wachstum in Gegenwart von 3-Amino-1,2,4-Triazol (3-AT) ermöglicht. 3-AT hemmt kompetitiv das *HIS3*-Enzym, das in geringer Menge auch ohne transkriptionelle Aktivierung exprimiert wird. Klone wachsen erst dann auf Histidin-Mangelmedium, wenn ausreichend *HIS3* gebildet wird, um die kompetitive Inhibition durch 3-AT zu überwinden. Das



verbesserte System nach Joung und Pabo zeigt einen deutlich verminderten *Background*, so dass auch schwächere Interaktionen identifiziert werden können.

Dennoch macht ein Zweihybrid-Experiment nur qualitative Aussagen. Interaktionsstärken und daraus abgeleitete Relevanzen der Wechselwirkungen können nur bedingt miteinander verglichen werden.

Demnach können die robusten Interaktionen zwischen DdTBP und DdBrfC bzw. DdTBP und ORF1p darauf zurückzuführen sein, dass sie (i) besonders stark sind und/ oder (ii) die interagierenden Proteine besonders gut exprimiert wurden. Stratagene bietet zwar ein polyklonales Antiserum gegen den in pBT klonierten  $\lambda$ cl-Fusionsanteil des ersten Interaktionspartners an, die Expression des pTRG-kodierten Fusionsproteins mit dem zweiten Interaktionspartner kann allerdings nicht überprüft werden. Für eine bessere quantitative Auswertung der Versuche wäre ein zusätzliches Antiserum gegen die pTRG-kodierte RNA-Polymerase-Untereinheit wünschenswert. Die Konzentration an 3-AT, die eine Interaktion aushält, stellt zwar ein gewisses Maß für die Interaktionsstärke dar, muss aber immer in Zusammenhang mit der Expressionsrate gesehen werden. Die Relevanz der Interaktion zwischen DdTBP und ORF1p sollte deshalb zusätzlich *in vivo* überprüft werden.

Dies kann z. B. durch die Herstellung von *D. discoideum Knock-in*-Mutanten mit DdTBP-Proteinen geschehen, die theoretisch nicht mehr in der Lage sind, mit ORF1p zu interagieren. Um die Expression mutierter DdTBP vom Wildtyp-Protein in *D. discoideum*-Zellen unterscheiden zu können, wurde in dieser Arbeit erfolgreich ein monoklonaler Antikörper gegen Wildtyp-DdTBP entwickelt. Die zusätzliche Isolierung eines hochspezifischen „Epitop-Tag“-Antikörpers könnte genutzt werden, um DdTBP-Mutanten nachzuweisen, die durch den *Knock-in* in *Dictyostelium discoideum*-Zellen eingebracht wurden. Die fehlende Interaktion zwischen ORF1p und DdTBP-Mutanten und damit auch der Integration von TRE5-A.1 könnte durch genetische Retrotranspositionstests in *D. discoideum*-Zellen erfolgen. Dazu sind in unserem Arbeitskreis bereits zwei Testsysteme etabliert: Die TRE-Falle sowie ein artifizielles Retrotransposon, das die Bestimmung des Integrationsortes ermöglicht [BECK *et al.*, 2002; SIOL *et al.*, 2006]. Die TRE-Falle erlaubt die Isolation von *de novo*-Integrationen, während mit dem artifizialen Element ein Testsystem zur Verfügung steht, das es ermöglicht, Retrotransposons zu modifizieren, ihre Mobilität zu testen und anschließend inklusive der neuen flankierenden Sequenzen aus dem

Genom zu isolieren. Auf diese Weise ergänzen sich beide Systeme bei der Untersuchung des Retotranspositionsmechanismus.

Zusätzlich könnten CoIP-Versuche durchgeführt und das Integrationsverhalten sowie der Phänotyp der Mutante bestimmt werden. Da es sich bei TBP allerdings um einen generellen Transkriptionsfaktor handelt, ist es schwierig abzuschätzen, wie gravierend ein Knock-in den Zellzyklus von *Dictyostelium discoideum* beeinträchtigt.

Trotz des generell sehr erfolgreichen Einsatzes von Hefe-basierten Zweihybrid-Systemen und trotz eines Mangels an potentiell für die Interaktion wichtiger posttranslationaler Modifikationen der Fusionsproteine in einem bakteriellen Assay stellt letzterer eine interessante Alternative zum Hefe-System dar. Die Transformationsrate in Bakterien ist größer, es sind keine aufwendigen Aufarbeitungsschritte zur Isolierung der DNA erforderlich und die kurze Generationszeit von Bakterien ermöglicht schneller sichtbare Ergebnisse. Insbesondere eukaryontische Testproteine, die an universellen Prozessen des Zellzyklus wie der Transkription beteiligt sind, können die transkriptionellen und translatorischen Vorgänge in Hefezellen empfindlich stören. Daneben ist es unwahrscheinlicher, dass eine bakterielle Wirtszelle eukaryontische Homologe der zu testenden Proteine in sich trägt und das Ergebnis der Interaktionsanalyse auf diese Weise verfälscht wird. Es ist denkbar, dass frühere Zweihybrid-Experimente der in dieser Arbeit untersuchten Proteine in einem Hefe-basierten System aus diesen Gründen fehlschlugen.

Als Fazit kann festgehalten werden, dass das eingesetzte bakterielle Zweihybrid-System bei sorgfältiger Einstellung der experimentellen Bedingungen zur Unterdrückung falsch positiver Ergebnisse eine sehr attraktive Alternative zum Hefe-basierten System darstellt. Dies gilt insbesondere dann, wenn bestimmte Proteine direkt auf mögliche Interaktionen getestet werden sollen. Es scheint nach den in dieser Arbeit gemachten Erfahrungen, dass das bakterielle System aufgrund des hohen „Hintergrund-Wachstums“ einem Hefe-basierten System dann unterlegen ist, wenn ein *Screening* einer cDNA-Bank nach unbekanntem Interaktionspartnern eines Proteins erforderlich ist.

## 4.2 Definition des Integrationsortes von TRE5-A.1 durch ORF1p-DdTFIIIB-Interaktionen

Die Wahl des Integrationsortes ist für den Transpositions-/ Lebenszyklus eines Retrotransposons von entscheidender Bedeutung. Sie kann sowohl die Auswirkungen der Insertion auf das Wirtsgenom als auch die Expression des Retrotransposons beeinflussen.

Die Integrationspräferenz von TRE5-A für tRNA-Gene ist nicht sequenz-bedingt sondern hängt vielmehr von den Transkriptionsfaktoren der Pol III ab. Vorangegangene Arbeiten haben gezeigt, dass die gerichtete und positionsspezifische Integration von TRE5-A.1 oberhalb von tRNA-Genen DdTFIIIC benötigen [SIOL *et al.*, 2006]. Die Ergebnisse dieser Arbeiten ließen vermuten, dass der Faktor allerdings nicht unmittelbar an der Heranführung des Retrotransposons an seine Ziel-DNA beteiligt ist, sondern dass DdTFIIIC in erster Linie für die Rekrutierung und korrekte Positionierung von DdTFIIIB benötigt wird. Konform mit dieser Hypothese konnte gezeigt werden, dass ein 5S-Gen ebenfalls als orthodoxes Integrationsziel erkannt wird [SIOL *et al.*, 2006]. Die Transkription eines 5S-Gens erfolgt unabhängig von der geninternen B-Box. Stattdessen bindet ein zinkfingerreicher Transkriptionsfaktor (TFIIIA) an die Promotorregion und rekrutiert im Anschluss TFIIIC, der wiederum TFIIIB an die DNA heranzuführt. Die Topologie des Präinitiationskomplexes eines 5S-Gens unterscheidet sich im Bereich des Gens also wesentlich von dem eines tRNA-Gens und schließt somit direkte Interaktionen zwischen TFIIIC und TRE5-A.1 aus. In der vorliegenden Arbeit konnte nun direkt gezeigt werden, dass TRE5-A-kodiertes ORF1p unmittelbar DdTFIIIB kontaktiert. Im Kontext der dreidimensionalen Abdeckung der DNA durch DdTFIIIB, die etwa 50 bp oberhalb der transkriptionsaktiven tRNA-Gene endet, erscheint es auch plausibel, dass TRE5-A.1 DdTFIIIB erkennt und in direkter Nachbarschaft oberhalb des Transkriptionsfaktors in die Ziel-DNA integriert. Dadurch ergeben sich die beobachteten Abstände von  $48 \pm 2$  bp zwischen TRE5-A.1 und tRNA-Gen.

Die kürzlich abgeschlossene Sequenzierung des Genoms von *Dictyostelium discoideum* AX4 ermöglichte die *In-Silico*-Identifizierung von DdTFIIIB, der wie in *S. cerevisiae* aus den drei Untereinheiten DdTBP, DdBrf1 und DdBdp1 besteht [EICHINGER *et al.*, 2005].

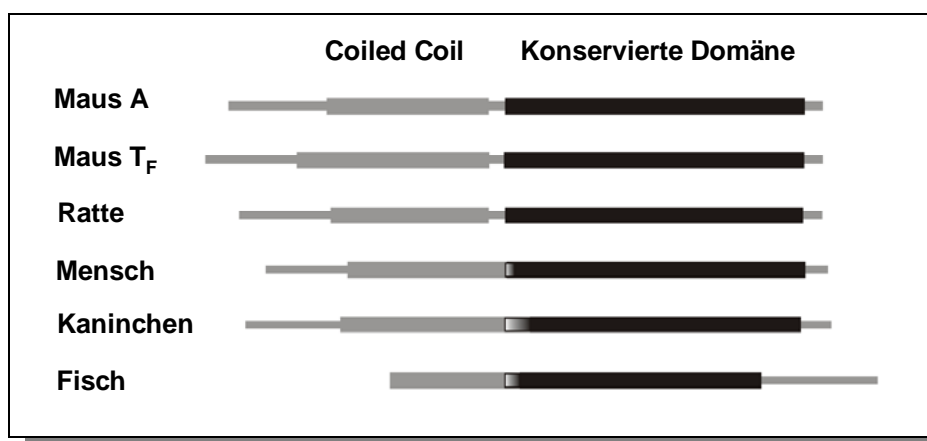
Die robuste Interaktion zwischen DdTBP und DdBrfC lässt die Vermutung zu, dass diese beiden Untereinheiten wie in der Bäckerhefe einen stabilen Komplex (B') bilden [KASSAVETIS *et al.*, 2006]. Mutationen und Deletionsversuche zeigen, dass DdTBP mit seiner N-terminal gelegenen  $\alpha$ -Helix H2 mit DdBrfC in Wechselwirkung tritt, während die C-terminale Helix H2' mit ORF1p interagiert. Da der Core-Bereich von TBP verschiedener Spezies sehr gut konserviert ist [NIKOLOV *et al.*, 1996], führten Mutationen, die analog zu bekannten, die Interaktion zwischen HsTBP und HsBrf1 unterbindenden Mutationen gewählt wurden, auch in DdTFIIIB dazu, dass eine Interaktion nicht mehr stattfinden konnte.

Als letzte Untereinheit tritt in *S. cerevisiae* ScBdp1 dem TBP-Brf1-DNA-Komplex bei und bindet mit Hilfe der SANT-Domäne an den C-Terminus von ScBrf1 [KASSAVETIS *et al.*, 2001; 2006]. In *Dictyostelium discoideum* konnte eine solche Interaktion zwar ebenfalls beobachtet werden, allerdings war sie so schwach, dass diese Daten durch zusätzliche Versuche verifiziert werden müssen. Eine Interaktion zwischen Bdp1 und den anderen Untereinheiten von TFIIIB darf mit großer Wahrscheinlichkeit angenommen werden. Neben der nachgewiesenen Wechselwirkung in *S. cerevisiae* gibt es außerdem in HsTFIIIB Hinweise auf eine Interaktion zwischen TBP und Bdp1, möglicherweise auf der DNA [SAXENA *et al.*, 2005]. Der Eintritt von Bdp1 in den B'-Komplex verstärkt die Beugung der DNA und schafft so eine für die Transkription optimale Topologie [BRAUN *et al.*, 1992]. Bei den durchgeführten Versuchen stand möglicherweise durch eine zu geringe Expression des  $\lambda$ cl-Bdp1-Fusionsproteins nicht genügend Protein für eine Interaktion zur Verfügung (siehe Abb. 3.11). Alternativ ist es denkbar, dass DdBdp1 *in vivo* an eine zusammengesetzte Interaktionsoberfläche aus DdBrf1 und DdTBP bindet.

Um diese Theorie zu überprüfen, könnten Dreihybrid-Versuche durchgeführt werden. Durch sensiblere Verfahren (photochemisches Cross-Linking, andere Zweihybrid-Systeme) müssen die ersten Hinweise aus den Zweihybrid-Experimenten überprüft werden.

TRE5-A.1 ist ein Retrotransposon aus *D. discoideum*, das nach seiner Mobilisierung mit hoher Spezifität oberhalb von tRNA-Genen integriert [WINCKLER *et al.*, 2005]. Das Element besitzt zwei ORFs, die sich auf ca. 70 bp in unterschiedlichen Leserahmen überlappen. Aufgrund von Sequenzhomologien zu anderen Non-LTR-Retrotransposons konnten in ORF2 Domänen für eine Endonuklease (EN) und eine

Reverse Transkriptase (RT) sowie eine Domäne mit einem typischen Histidin-Cystein-Motiv identifiziert werden. Die RT und die EN werden während der gekoppelten reversen Transkription und Integration, die nach dem TPRT-Mechanismus verläuft, benötigt. Im Gegensatz dazu kodiert ORF1 für ein Protein, dessen Funktion bisher praktisch unbekannt ist. TRE5-A.1-ORF1p zeigt keine signifikanten Homologien zu den Sequenzen anderer untersuchter ORF1-Proteine. Orthologe Proteine aus L1-Elementen von Mensch, Maus, Ratte und Kaninchen besitzen N-terminal eine ausgedehnte *Coiled-coil*-Domäne, die sich über etwa 100-120 Aminosäuren erstreckt und die für Protein-Interaktionen genutzt wird (Abb. 4.1). Man vermutet aufgrund der schwachen Konservierung dieser Sequenzen, dass diese *Coils* entweder durch nichthomologe Rekombination erworben wurden oder dass sie sich von einem gemeinsamen Vorläufer abstammend rapide auseinander entwickelt haben [MARTIN 2006].



**Abb. 4.1: Schematische Darstellung verschiedener ORF1-Proteine**

Die dünnen Balken repräsentieren die Länge des vollständigen Proteins. Dicke graue Balken stellen die *Coiled-coil*-Domäne dar, dicke schwarze Balken die stark basische, konservierte Domäne. Die Domänen überlappen sich bei Mensch, Kaninchen und Fisch (modifiziert nach Martin 2006). Für ORF1p aus TRE5-A.1 konnten keine Sequenzhomologien zu den anderen untersuchten ORF1-Proteinen gefunden werden. Eine putative *Coil*-Domäne sitzt im C-Terminus (nicht dargestellt).

Umso interessanter und aufschlussreicher sind daher die Ergebnisse der Zweihybrid-Experimente. Sie belegen Interaktionen zwischen TRE5-A.1-ORF1p und allen drei Untereinheiten von DdTFIIIB, die für DdTBP und DdBrf1 zusätzlich in *Pulldown*-Versuchen bestätigt werden konnten.

Durch Deletions- und Mutationsversuche konnte die minimale Kontaktdomäne des ORF1-Proteins für die Interaktion mit DdTBP auf 46 Aminosäuren (AS 112-158) eingegrenzt werden. Da dieser Bereich keine bekannten proteinbindenden Motive

wie *Coil*-Domänen oder Leucin-Zipper enthält, sollen weiterführende Arbeiten die dreidimensionale Struktur von ORF1p mittels NMR-Analysen aufklären. In einem ersten Schritt in diese Richtung wird zur Zeit an der Expression eines löslichen ORF1p-Fragmentes gearbeitet.

Die für die DdTBP-ORF1p-Interaktion notwendige DdTBP-Kontaktdomäne konnte im Bereich der  $\alpha$ -Helix H2' lokalisiert werden. Verpflanzt man diese Helix aus DdTBP in HsTBP, so ist die resultierende Chimäre HsTBP(dH2') in der Lage, mit ORF1p zu interagieren. HsTBP selbst bindet trotz der großen Strukturähnlichkeit zu DdTBP nicht an ORF1p. Die Affinität der Chimäre zu TRE5-A.1-ORF1p ist allerdings deutlich geringer als die von DdTBP, da sie lediglich bis zu einer 3-AT-Konzentration von 3 mM nachweisbar war. Demnach tritt ORF1p wahrscheinlich noch mit weiteren Regionen von DdTBP in Kontakt. Die vollständige Wiederherstellung der Interaktionsstärke könnte durch die zusätzliche Einführung weiterer Epitope aus DdTBP wie der Helix H1' und des genuinen C-Terminus gelingen.

Nach Programmen zur Strukturvorhersage können etwa in der Mitte und am C-Terminus von TRE5-A.1-ORF1p potentielle, kleine *Coil*-Domänen lokalisiert sein. Unabhängig von der unterschiedlichen Lage und Größe der *Coiled-coils* im Vergleich zu den orthologen ORF1-Proteinen anderer Organismen (vgl. Abb. 4.1) scheinen diese auch in TRE5-A.1 für die Dimerisierung verantwortlich zu sein. Die Deletion der terminalen *Coil*-Struktur in ORF1p<sup>1-405</sup> unterbindet wie erwartet die Dimerisierung von ORF1p. Umgekehrt konnte für das isolierte ORF1p<sup>402-444</sup>-Fragment eine positive Wechselwirkung mit ORF1p und mit DdBrfC nachgewiesen werden.

Die Interaktion von DdBrfC findet ebenfalls über den C-Terminus von ORF1p statt. Demnach konkurrieren DdBrfC und ORF1p um die Bindung der terminalen 40 Aminosäuren. Alternativ ist es denkbar, dass DdBrfC nur mit ORF1p-Dimeren interagiert und die Deletionsmutante ORF1<sup>1-404</sup> erst in zweiter Konsequenz nicht mehr in der Lage ist, DdBrfC zu binden. Recherchen in den gängigen Datenbanken sagen für DdBrf1<sup>443-463</sup> eine *Coiled-coil*-Domäne voraus, die möglicherweise mit der entsprechenden Struktur in ORF1p interagieren kann. Erste Zweihybrid-Experimente mit Deletionsmutanten von DdBrf1 (DdBrf1<sup>439-706</sup>) sind mit dieser Computervorhersage konform (Daten nicht gezeigt). Die genauere Lage der Kontaktdomäne zwischen DdBrf1 und ORF1p wird im Rahmen einer weiteren Dissertation untersucht.

### 4.3 ORF1p ist Teil des Präintegrationskomplexes von TRE5-A.1

Im Gegensatz zu TRE5-A.1-ORF2p ist die genaue Rolle von ORF1p während der Retrotransposition von TRE5-A.1 noch ungeklärt. Sequenzanalysen ergaben keine signifikanten Homologien zu Proteinen bekannter Funktion. Der Nachweis, dass es sich bei ORF1p aus dem phylogenetisch verwandten humanen L1-Element um ein RNA-bindendes Protein handelt, das mit hoher Affinität an RNA bindet und somit Ribonukleopartikel bildet, stellte deshalb einen großen Fortschritt dar. Es konnte auch gezeigt werden, dass ORF1-Proteine als Nukleinsäure-Chaperone dienen und dass diese Aktivität während der Retrotransposition von L1 notwendig ist [HOHJOH & SINGER 1997; KULPA & MORAN 2005; MARTIN 2006]. Beide Funktionen werden von distinkten Regionen des Proteins wahrgenommen [MARTIN *et al.*, 2000]. In Analogie zu ORF1-Proteinen anderer Non-LTR-Retrotransposons wird vermutet, dass TRE5-A.1-ORF1p ebenfalls an RNA bindet und dadurch Ribonukleopartikel bildet, die möglicherweise während der Mitose zum Zellkern gelangen [WINCKLER *et al.*, 2002a]. Es wurde beobachtet, dass rekombinant exprimiertes ORF1p im Zytosol von *Dictyostelium discoideum*-Zellen große Aggregate bildet [HENTSCHEL *et al.*, 2001]. Ob diese Komplexe RNA enthalten, ist bisher noch spekulativ, da Sequenzanalysen bisher keine RNA-bindenden Motive ergeben haben.

Versuche mit ORF1p aus murinen und humanen L1-Elementen zeigten überraschenderweise, dass Mutationen in ORF1p die Retrotransposition mindestens genauso stark, wenn nicht eher stärker beeinträchtigen als Mutationen der enzymatischen Proteine EN und RT und eine Retrotransposition vollständig unterbinden können [MORAN *et al.*, 1996; COST *et al.*, 2002; KULPA & MORAN 2005; MARTIN *et al.*, 2005]. Die direkte Interaktion von TRE5-A.1-ORF1p mit DdTFIIIB unterstützt die Vermutung, dass ORF1p allgemein eine weitaus größere Rolle vor und während der Integration spielt als bisher angenommen.

Nach den Daten von Siol *et al.* 2006 (unveröffentlichte Daten) zeigt TRE5-A-ORF1p *in vivo* eine ausgeprägte *cis*-Präferenz. Die Bindung von ORF1p an seine eigene RNA würde zum einen die eigene RNA als Intermediat der Retrotransposition vor dem Abbau schützen, zum anderen würde die Bindung von RNA an das Protein, das durch Interaktion mit DdTFIIIB direkt an die DNA geführt wird, auch die RNA direkt an den Zielort bringen. Da weder TRE5-A.1-ORF1p noch -ORF2p eine Kernlokalisierungssequenz tragen und ORF1p die einzige Retrotransposon-kodierte

Struktur ist, für die bisher eine Interaktion mit DdTFIIIB nachgewiesen wurde, erscheint es plausibel, dass die RNA des Elementes mit ORF1p an die DNA gelangt. Die in dieser Arbeit erhobenen Daten zeigen eindrucksvoll, dass ORF1p ein Teil des Präintegrationskomplexes eines Non-LTR-Retrotransposons (TRE5-A) sein muss und im Falle der TRE5-A-Retrotransposition mindestens eine Aufgabe hat – nämlich die Festlegung des Integrationsortes.

#### 4.4 RTp und ENp

TRE5-A.1 ist ein aktives Retrotransposon aus *Dictyostelium discoideum* [SIOL *et al.*, 2006], das durch die reverse Transkription eines RNA-Intermediates immer wieder neue Kopien von sich selbst erstellt. Die dafür notwendigen Enzyme Endonuklease (EN) und Reverse Transkriptase (RT) sind im zweiten ORF von TRE5-A kodiert. Das Element gehört zu der großen Gruppe der Non-LTR-Retrotransposons, die sich mit Hilfe eines Mechanismus replizieren, bei dem die chromosomale Insertion eng mit der reversen Transkription verknüpft ist (TPRT-Mechanismus) [OSTERTAG & KAZAZIAN 2001a]. Während der TPRT schneidet die Element-kodierte EN den Minusstrang der Ziel-DNA. Da die EN-Domäne der TRE5-A.1-Elemente mit den Apurinischen/Apyrimidinischen Endonukleasen verwandt ist, findet dieser Schnitt in unmittelbarer Nähe derjenigen Stelle statt, an der das Enzym an die DNA bindet [FENG *et al.*, 1998a; MALIK & EICKBUSH 1999; CHRISTENSEN & EICKBUSH 2005]. Kulpa *et al.* konnten die Zusammensetzung der Ribonukleopartikel des humanen L1-Elements aufklären und zeigten, dass diese aus L1-RNA, ORF1p und ORF2p bestehen [KULPA & MORAN 2006].

In den Zweihybrid-Experimenten konnten auch bei der niedrigsten 3-AT-Konzentration keine Interaktionen zwischen TRE5-A.1-kodiertem ORF2p und ORF1p bzw. DdTFIIIB detektiert werden. Eine ausreichende Expression der pBT-kodierten  $\lambda$ cl-Fusionsproteine wurde in Western-Blot-Analysen überprüft (siehe Abb. 3.11). Das entsprechende Kontrollexperiment für die pTRG-kodierten Fusionsproteine konnte allerdings nicht durchgeführt werden, da kein Antiserum gegen den RNA-Polymerase-Fusionsanteil zur Verfügung stand.

Zahlreiche Kotransformationsversuche haben gezeigt, dass die Nachweisbarkeit einer Interaktion unter anderem davon abhing, aus welchem Vektor die Fusionsproteine exprimiert wurden. So wuchsen z. B. Doppeltransformatanten in Anwesenheit von 3-AT, die die Plasmide pBT-TBP und pTRG-ORF1 trugen. Die



Transformanden der inversen Plasmidkombination pBT-ORF1/ pTRG-TBP dagegen zeigten unter den gleichen Bedingungen entweder gar kein oder ein nur sehr schwaches Wachstum. Demnach könnten die sterischen Verhältnisse in Kombination mit einer möglicherweise unzureichenden Expression dazu führen, dass vorhandene Interaktionen zwischen ORF2p und ORF1p bzw. DdTFIIIB nicht detektiert wurden.

#### **4.5 Konvergente Entwicklung eines zielgerichteten Integrationsmechanismus**

Retrotransposons integrieren nicht an willkürlichen Stellen des Wirtsgenoms, sondern bevorzugen Regionen mit einer geringen Gendichte. Diese Verteilung kann auf eine Selektion gegen schädigende Integrationen zurückzuführen sein, oder auf eine aktive Selektion von chromosomalen Bereichen außerhalb des Gen-reichen Euchromatins, wo Integrationen die zelluläre Integrität des Wirts nicht beeinträchtigen [BOEKE & DEVINE 1998]. Gerade in haploiden Genomen mit einer so hohen Gendichte wie bei *D. discoideum* oder *S. cerevisiae* herrscht ein großer Selektionsdruck auf mobilen Elementen. Nur ein Mechanismus, der zuverlässig die Insertion in „sichere“, also nichtkodierende Bereiche der DNA vermittelt, ermöglicht das evolutionäre Überleben der Transposons. Ein faszinierendes Beispiel für einen solchen Mechanismus stellt die Integration in die unmittelbare Nachbarschaft von tRNA-Genen dar. Für die Bäckerhefe wurde gezeigt, dass aktive tRNA-Gene die Aktivität benachbarter Pol II-Gene reprimieren und sich aus diesem Grund kaum proteinkodierende Gene in direkter Nachbarschaft befinden [BOLTON & BOEKE 2003]. Auch wenn die Abstände von tRNA-Genen und Strukturgenen im Genom von *D. discoideum* bislang nicht systematisch untersucht wurden, kann angenommen werden, dass gewisse Abstände zwischen Pol II- und Pol III-Genen in allen Lebewesen existieren und deshalb tRNA-Gene auch in *Dictyostelium discoideum* zu einem sicheren Ort für die Integration mobiler Elemente werden. Pol II-Gene bieten auch deshalb keine guten Ziele für eine nahe Integration, weil sich in den flankierenden Bereichen dieser Gene häufig noch regulatorische Bereiche wie Promotoren oder Terminatoren befinden.

Mobile Elemente, die spezifisch in die Nähe von tRNA-Genen integrieren, findet man sowohl in der Klasse der Non-LTR- als auch der LTR-Retrotransposons. Diese Integrationsspezifität kann man allein in *D. discoideum* mehrfach beobachten. So integrieren TRE5-Elemente etwa 48 bp und DGLT-A-Elemente ca. 30 bp oberhalb

von tRNA-Genen [GLÖCKNER *et al.*, 2001b], während TRE3-Elemente ca. 100 bp unterhalb von tRNA-Genen integrieren [SZAFRANSKI *et al.*, 1999a; GLÖCKNER *et al.*, 2001b]. Der Abstand zwischen TRE3 und tRNA-Gen wird sehr wahrscheinlich von der Lage entweder der geninternen oder extern gelegenen B-Box bestimmt und lässt darauf schließen, dass tRNA-Gene durch eine Wechselwirkung zwischen TRE3-kodierten Proteinen und DdTFIIIC erkannt werden [SZAFRANSKI *et al.*, 1999a]. TRE5-B-Elemente scheinen ebenfalls Interaktionen mit Transkriptionsfaktoren der Pol III-Maschinerie für die Definition ihres Integrationsortes zu nutzen (vgl. Abschnitt 3.4.7). Die Daten der Zweihybrid-Versuche legen den Schluss nahe, dass TRE5-A und TRE5-B sich im wesentlichen der gleichen Wechselwirkungen mit DdTBP bedienen. Da die Aminosäuresequenzen der beiden ORF1-Proteine auch im Bereich der Kontaktdomäne von ORF1p für DdTBP nur schwach konserviert sind, muss die eventuelle Identifizierung anderer gemeinsamer Motive Gegenstand weiterer Untersuchungen sein.

Der Integrationsmechanismus von DGLT-A ist zur Zeit noch nicht beschrieben [GLÖCKNER *et al.*, 2001b]. Es liegt nahe anzunehmen, dass die Integrationsspezifität dieser LTR-Retrotransposons für einen Bereich etwa 30 bp oberhalb von tRNA-Genen ebenfalls durch Interaktionen mit DdTFIIIB vermittelt wird. Durch die Bindung von DdTFIIIB an die transkriptionell aktive DNA wird diese durch Beugung stark deformiert und potentielle Integrationsstellen exponiert. Alternativ ist es denkbar, dass wie im Falle der Integration des LTR-Retrotransposons Ty5 aus *S. cerevisiae* durch Bindung des Transkriptionsfaktors an verschiedenen Stellen oberhalb des tRNA-Gens Bedingungen geschaffen werden, die eine Retrotransposition erst ermöglichen. DdTFIIIB wäre somit auch ein potentieller Bindungspartner für DGLT-A. Vor allem vor dem Hintergrund, dass es sich bei DGLT-A mit hoher Wahrscheinlichkeit um ein LTR-Retrotransposon handelt, das phylogenetisch enger mit den Retroviren verwandt ist als ein Non-LTR-Retrotransposon, wäre die Aufklärung des Integrationsmechanismus dieses Elementes von großem Interesse und möglichem Nutzen für die Gentherapie. Der Nachteil der heutigen retroviralen Gentransfer-Vektoren, nämlich die bevorzugte Integration in transkriptionsaktive Bereiche des Genoms, könnte durch eine Adaption der tRNA-Gen-Erkennung an einen solchen Gentransfer-Vektor aufgehoben und die Sicherheit der Gentherapie wesentlich verbessert werden [Wu *et al.*, 2003].

Auch in *S. cerevisiae* spielen tRNA-Gene eine wichtige Rolle bei der positionsspezifischen Integration der LTR-Elemente Ty3 und Ty1. Hier sind die Retrotranspositionsprozesse ebenfalls eng mit der Transkription von tRNA-Genen verknüpft. Proteininteraktionen am Integrationsort positionieren die Transposons in unterschiedlichen Abständen zum tRNA-Gen [KIM *et al.*, 1998b].

Für Ty3 wurde gezeigt, dass diese Spezifität durch Wechselwirkungen zwischen Ty3-kodierten Proteinen und Transkriptionsfaktoren der Pol III vermittelt wird. Die RNA-Polymerase III selbst nicht notwendig, sondern konkurriert vielmehr mit der Ty3-kodierten Integrase um Bindestellen auf der Ziel-DNA [CONNOLLY & SANDMEYER 1997]. Mutationen der A-oder B-Box von tRNA-Genen, die die weitere Transkription beeinträchtigen, führen zu einer deutlich verminderten Transpositionsrate durch Ty3-Elemente [CHALKER & SANDMEYER 1992]. Allerdings können Integrationsereignisse an DNA-TBP-Brf1-Komplexen (B'-Komplex) beobachtet werden, so dass wie bei *Dictyostelium discoideum* angenommen werden kann, dass der Ty3-Präintegrationskomplex mit Teilen von ScTFIIIB in Wechselwirkung tritt. Obwohl beide ScBrf1-Hälften die Integration von Ty3 fördern, scheint die wesentliche Interaktion durch die TFIIIB-homologe N-terminale Domäne zustande zu kommen [YIEH *et al.*, 2002]. Die Integrationsrate von Ty3 wird in Anwesenheit von ScBdp1 deutlich gesteigert, was zeigt, dass das Element für eine effiziente Transposition entweder zusätzliche Kontakte mit ScTFIIIB benötigt oder dass ScBdp1 den B'-Komplex in eine für die Transposition günstigere Konformation überführt [YIEH *et al.*, 2000]. Im Unterschied zu Ty3 wird die Affinität von TRE5-A.1 zu DdBrf1 vor allem durch den kaum konservierten C-Terminus vermittelt. Aber auch TRE5-A.1 benötigt zusätzliche Kontakte mit DdBdp1.

Im Gegensatz dazu ist die Integrationsspezifität von Ty1 nicht unmittelbar mit den Transkriptionsfaktoren der Pol III verknüpft. Vielmehr führt die Bindung Chromatin modellierender Proteine wie Isw2 aber auch Bdp1 dazu, dass die Topologie von Nukleosomen oberhalb des tRNA-Gens eine periodische Integration von Ty1 ermöglicht [BACHMAN *et al.*, 2004; 2005].

Die erfolgreiche Verbreitung eines mobilen Elementes ist zu einem wesentlichen Teil von der Fähigkeit, sichere Integrationsorte zu finden, abhängig. Die Assoziation so unterschiedlicher Retrotransposons wie Ty1, Ty3, DGLT-A, TRE5 und TRE3 mit tRNA-Genen in Genomen mit einer extrem hohen Gendichte zeigt, dass tRNA-Gene

solche sicheren Orte darstellen und mobile Elemente durch Insertionen an diesen Orten das Genom signifikant mitgestaltet haben.

## 4.6 Ausblick

Ziel der Gentherapie ist die spezifische Einführung therapeutischer Gene in die Zellen des Patienten. Die Übertragung des genetischen Materials gelingt zur Zeit mit Hilfe retroviraler Vektoren, die die Fremd-DNA mit guter Effizienz stabil in das Erbgut des Patienten integrieren. Allerdings erfolgen diese Integrationen bevorzugt in der Nähe des Transkriptionsstarts stark transkribierter Gene, so dass ein hohes Risiko der Insertionsmutagenese besteht [Wu *et al.*, 2003].

In Hinblick auf die Entwicklung sicherer Gentransfer-Vektoren wurden in dieser Arbeit daher einige wichtige Erkenntnisse zum Mechanismus der positionsspezifischen Integration von TRE5-A.1 gewonnen. Auf diesen Ergebnissen aufbauend bieten sich weitere Untersuchungen an:

- Die Relevanz der Interaktion zwischen DdTBP und ORF1p *in vivo* kann durch *Knock-in*-Versuche mit mutierten DdTBP-Varianten in *D. discoideum*-Zellen untersucht werden. Dazu würde endogenes TBP durch Mutanten ersetzt, die *in vitro* nicht mit ORF1p von TRE5-A.1 interagieren können. In solchen Zellen könnte dann untersucht werden, ob markierte TRE5-A-Retrotransposons noch in der Lage sind, tRNA-Gene als Integrationsorte zu identifizieren. Genetisch markierte TRE5-A-Derivate werden zur Zeit im Arbeitskreis entwickelt.
- Sowohl in Retrotransposons der Bäckerhefe als auch des Menschen konnten Interaktionen der beiden TFIIIB-Untereinheiten Bdp1 und Brf1 nachgewiesen werden. Da es bisher bereits erste Hinweise auf Interaktionen zwischen DdBdp1 und DdBrf1 gibt, sollten weitere Experimente diese Ergebnisse verifizieren. Auch das mit ORF1p interagierende Epitop auf DdBrfC sollte weiter eingegrenzt werden.
- Das mit ORF1p interagierende Epitop auf DdTBP kann durch Deletionen und Mutationen weiterhin eingegrenzt werden. Das gleiche gilt für ORF1p.
- Die schrittweise Anpassung des chimären Proteins HsTBP(dH2') an die Kontaktoberfläche von ORF1p würde der direkten Überprüfung der in DdTBP gefundenen Epitope dienen.

- Interessanterweise gab es in der vorliegenden Arbeit keine Hinweise auf eine Interaktion der enzymatischen Proteine EN und RT mit DdTFIIIB. Da beide Enzyme für die Retrotransposition allerdings notwendig sind, sollte dieses Ergebnis nochmals überprüft werden (vgl. Abschnitt 4.4)
- Für eine mögliche Anwendung in der Gentherapie wäre es wichtig, ORF1p strukturbasiert an HsTBP anzupassen, d.h. durch die gezielte Veränderung der ORF1p-Struktur seine Spezifität für DdTBP in eine Spezifität für HsTBP umzuwandeln. Dazu ist es notwendig, die dreidimensionale Struktur von ORF1p z.B. mittels NMR-Analysen aufzuklären. Als ersten Schritt in diese Richtung wird zur Zeit an der Expression eines löslichen ORF1p-Fragmentes gearbeitet.
- Die Entwicklung eines aktiven Hybridelementes aus einem humanen L1-Element und einem TRE5-A.1, das das Prinzip der tRNA-gerichteten Integration eines Non-LTR-Retrotransposons in Säugerzellen nachahmt, wäre ein erster Schritt zu retroviralen Gentransfer-Vektoren, die durch die spezifische Integration nahe von tRNA-Genen die Gentherapie in Zukunft wesentlich sicherer gestalten könnten.

## 5 Zusammenfassung

Das Non-LTR-Retrotransposon TRE5-A.1 aus *Dictyostelium discoideum* integriert positionsspezifisch  $48 \pm 2$  bp oberhalb von tRNA-Genen. Es konnte gezeigt werden, dass TRE5-A.1 Wechselwirkungen zwischen ORF1p und dem RNA-Polymerase III-spezifischen Transkriptionsfaktor DdTFIIIB nutzt, um seinen Integrationsort zu identifizieren. Damit wurden in dieser Arbeit erstmals direkte Proteininteraktionen zwischen einem Non-LTR-Retrotransposon und Komponenten des Chromatins zur Definition des Integrationsortes nachgewiesen.

DdTFIIIB wurde durch Blastp-Suchen *in silico* identifiziert. Wie auch in *S. cerevisiae* wird innerhalb des *D. discoideum* TFIIIB-Komplexes eine stabile Interaktion zwischen der N-terminal gelegenen Helix H2 von DdTBP und dem C-Terminus von DdBrf1 gebildet.

Es konnte sowohl mittels bakterieller Zweihybrid-Versuche als auch durch biochemische *Pulldown*-Experimente gezeigt werden, dass TRE5-A-kodiertes ORF1-Protein (ORF1p) mit allen drei Untereinheiten von DdTFIIIB in Wechselwirkung tritt. Am ausgeprägtesten war diese mit DdTBP. Durch Mutations- und Deletionsanalysen wurden die Kontaktflächen auf DdTBP und ORF1p näher charakterisiert. Demnach interagiert ORF1p über seinen N-Terminus (AS 112-158) mit DdTBP, während die C-terminalen 40 Aminosäuren sowohl von DdBrf1 als auch von ORF1p beansprucht werden und beide Proteine vermutlich um diese Bindestelle konkurrieren. DdTBP bindet hauptsächlich mit der C-terminalen  $\alpha$ -Helix H2' an ORF1p. Eine wichtige Position für diese Interaktion ist Ser<sup>195</sup> auf DdTBP. Obwohl HsTBP selbst nicht mit ORF1p interagiert, kann die Einführung der Helix H2' aus DdTBP in das humane Protein die Interaktion mit ORF1p wiederherstellen. Interaktionen zwischen DdTFIIIB und TRE5-A-ORF2p wurden nicht detektiert, allerdings konnte ein vermutliches Dimerisationsmotiv in ENp entdeckt werden. Für weiterführende Versuche, die die Relevanz der hier gefundenen Proteininteraktionen für die Target-Identifizierung in *D. discoideum*-Zellen untersuchen soll, wurden in dieser Arbeit zwei monoklonale Antikörper gegen DdTBP und einen zufällig entdeckten „Epitop-Tag“ isoliert und charakterisiert.

Die genaue Rolle von ORF1p während der Retrotransposition von TRE5-A.1 ist noch ungeklärt. Erste grundlegende Einblicke weisen darauf hin, dass ORF1p Teil des Präintegrationskomplexes von TRE5-A sein muss.

## 6 Literatur

- AASLAND, R., STEWART, A.F. & GIBSON, T. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends Biochem Sci* **21**, 87-8 (1996).
- ANDRAU, J.C., SENTENAC, A. & WERNER, M. Mutagenesis of yeast TFIIIB70 reveals C-terminal residues critical for interaction with TBP and C34. *J Mol Biol* **288**, 511-20 (1999).
- ASHWORTH, J.M. & WATTS, D.J. Metabolism of the cellular slime mould *Dictyostelium discoideum* grown in axenic culture. *Biochem. J.* **119**, 175-182 (1970).
- AYE, M., DILDINE, S.L., CLAYPOOL, J.A., JOURDAIN, S. & SANDMEYER, S.B. A truncation mutant of the 95-kilodalton subunit of transcription factor IIIC reveals asymmetry in Ty3 integration. *Mol. Cell. Biol.* **21**, 7839-7851 (2001).
- BACHMAN, N., EBY, Y. & BOEKE, J.D. Local definition of Ty1 target preference by long terminal repeats and clustered tRNA genes. *Genome Res* **14**, 1232-47 (2004).
- BACHMAN, N., GELBART, M., TSUKIYAMA, T., BOEKE, J. TFIIIB subunit Bdp1 is required for periodic integration of the TY1 retrotransposon and targeting of *Isw2p* to *S. cerevisiae* tDNAs. *Genes & Dev* **19**, 955-964 (2005).
- BAPTESTE, E., BRINKMANN, H., LEE, J.A. *et al.* The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. *Proc. Natl. Acad. Sci. USA* **99**, 1414-1419 (2002).
- BASAME, S., WAI-LUN LI, P., HOWARD, G., BRANCIFORTE, D., KELLER, D. & MARTIN, S.L. Spatial assembly and RNA binding stoichiometry of a LINE-1 protein essential for retrotransposition. *J Mol Biol* **357**, 351-7 (2006).
- BECK, P., DINGERMANN, T. & WINCKLER, T. Transfer RNA gene-targeted retrotransposition of *Dictyostelium* TRE5-A into a chromosomal UMP synthase gene trap. *J. Mol. Biol.* **318**, 273-285 (2002).
- BERTHOLD, H., FRORATH, B., SCANARINI, M., ABNEY, C.C., ERNST, B. & NORTHEMANN, W. Plasmid pGEX-5T - An Alternative System for Expression and Purification of Recombinant Proteins. *Biotechnology Letters* **14**, 245-250 (1992).
- BOEKE, J. & CORCES, V. Transcription and reverse transcription of retrotransposons. *Ann. Rev. Microbiol.* **43**, 403-434 (1989).
- BOEKE, J.D. & DEVINE, S.E. Yeast retrotransposons: finding a nice quiet neighborhood. *Cell* **93**, 1087-1089 (1998).
- BOISSINOT, S. & FURANO, A.V. The recent evolution of human L1 retrotransposons. *Cytogenet Genome Res* **110**, 402-6 (2005).
- BOLTON, E.C. & BOEKE, J.D. Transcriptional interactions between yeast tRNA genes, flanking genes and Ty elements: a genomic point of view. *Genome Res.* **13**, 254-263 (2003).
- BONNER, J.T. The migration stage of *Dictyostelium*: behavior without muscles or nerves. (minireview). *FEMS Microbiol. Lett.* **120**, 1-8 (1994).
- BOYER, L.A., LATEK, R.R. & PETERSON, C.L. The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol* **5**, 158-63 (2004).
- BOZZARO, S. & PONTE, E. Cell adhesion in the life cycle of *Dictyostelium*. *Experientia* **51**, 1175-1188 (1995).
- BRADFORD, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).
- BRAUN, B.R., KASSAVETIS, G.A. & GEIDUSCHEK, E.P. Bending of the *Saccharomyces cerevisiae* 5S rRNA gene in transcription factor complexes. *J Biol Chem* **267**, 22562-9 (1992).
- BREFELD, O. *Dictyostelium mucoroides*. Ein neuer Organismus aus der Verwandtschaft der Myxomyceten. *Abhandlungen der Senckenbergischen Naturforschenden Gesellschaft Frankfurt* **7**, 85-107 (1869).

- BRENNAND, D.M., DANSON, M.J. & HOUGH, D.W.** A comparison of ELISA screening methods for the production of monoclonal antibodies against soluble protein antigens. *J Immunol Methods* **93**, 9-14 (1986).
- BUKENBERGER, M., DINGERMANN, T., MEISSNER, W., SEIFART, K.H. & WINCKLER, T.** Isolation of transcription factor Illc from Dictyostelium discoideum. *Eur. J. Biochem.* **220**, 839-846 (1994).
- CAPY, P.** Classification and nomenclature of retrotransposable elements. *Cytogenet Genome Res* **110**, 457-61 (2005).
- CARLSON, C.M., FRANSEN, J.L., KIRCHHOF, N., MCLIVOR, R.S. & LARGAESPADA, D.A.** Somatic integration of an oncogene-harboring Sleeping Beauty transposon models liver tumor development in the mouse. *Proc Natl Acad Sci U S A* **102**, 17059-64 (2005).
- CHALKER, D.L. & SANDMEYER, S.B.** Ty3 integrates within the region of RNA polymerase III transcription initiation. *Genes & Dev* **6**, 117-128 (1992).
- CHISHOLM, R.L., GAUDET, P., JUST, E.M. et al.** dictyBase, the model organism database for Dictyostelium discoideum. *Nucleic Acids Res* **34**, D423-7 (2006).
- CHRISTENSEN, S.M. & EICKBUSH, T.H.** R2 target-primed reverse transcription: ordered cleavage and polymerization steps by protein subunits asymmetrically bound to the target DNA. *Mol Cell Biol* **25**, 6617-28 (2005).
- COLBERT, T. & HAHN, S.** A yeast TFIIIB-related factor involved in RNA polymerase III transcription. *Genes Dev* **6**, 1940-9 (1992).
- CONNOLLY, C. & SANDMEYER, S.** RNA polymerase III interferes with Ty3 integration. *FEBS Lett.* **405**, 305-311 (1997).
- CORMACK, B.P. & STRUHL, K.** The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell* **69**, 685-96 (1992).
- COST, G.J. & BOEKE, J.D.** Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* **37**, 18081-18093 (1998).
- COST, G.J., FENG, Q., JACQUIER, A. & BOEKE, J.D.** Human L1 element target-primed reverse transcription in vitro. *EMBO J.* **21**, 5899-5910 (2002).
- COTTER, D.A., SANDS, T.W., VIRDY, K.J., NORTH, M.J., KLEIN, G. & SATRE, M.** Patterning of development in Dictyostelium discoideum: factors regulating growth, differentiation, spore dormancy, and germination. *Biochem. Cell Biol.* **70**, 892-919 (1992).
- CRAIG, N.L., CRAIGIE, R., GELLERT, M. & LAMBOWITZ, A.M.** (eds.). *Mobile DNA II*, (ASM Press, Washington DC, 2002).
- DEBERARDINIS, R.J. & KAZAZIAN, H.H.** Analysis of the promoter from an expanding mouse retrotransposon subfamily. *Genomics* **56**, 317-323 (1999).
- DEININGER, P.L. & BATZER, M.A.** Mammalian retroelements. *Genome Res* **12**, 1455-65 (2002).
- DEININGER, P.L., MORAN, J.V., BATZER, M.A. & KAZAZIAN, H.H., JR.** Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev* **13**, 651-8 (2003).
- DEWANNIEUX, M., ESNAULT, C. & HEIDMANN, T.** LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* **35**, 41-8 (2003).
- DEWANNIEUX, M. & HEIDMANN, T.** LINEs, SINEs and processed pseudogenes: parasitic strategies for genome modeling. *Cytogenet Genome Res* **110**, 35-48 (2005).
- DING, Z., GILLESPIE, L.L., MERCER, F.C. & PATERNO, G.D.** The SANT domain of human MI-ER1 interacts with Sp1 to interfere with GC box recognition and repress transcription from its own promoter. *J Biol Chem* **279**, 28009-16 (2004).
- DOVE, S.L., JOUNG, J.K. & HOCHSCHILD, A.** Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386**, 627-30 (1997).
- EICHINGER, L., PACHEBAT, J.A., GLÖCKNER, G. et al.** The genome of the social amoeba Dictyostelium discoideum. *Nature* **435**, 43-57 (2005).
- FENG, Q., SCHUMANN, G. & BOEKE, J.D.** Retrotransposon R1Bm endonuclease cleaves the target sequence. *Proc Natl Acad Sci U S A* **95**, 2083-8 (1998a).



- FENG, Q.H., MORAN, J.V., KAZAZIAN, H.H. & BOEKE, J.D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905-916 (1996).
- FENG, Q.H., SCHUMANN, G. & BOEKE, J.D. Retrotransposon R1Bm endonuclease cleaves the target sequence. *Proc Natl Acad Sci USA* **95**, 2083-2088 (1998b).
- FOX, J.D., KAPUST, R.B. & WAUGH, D.S. Single amino acid substitutions on the surface of Escherichia coli maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci* **10**, 622-30 (2001).
- FUJIGASAKI, H., MARTIN, J.J., DE DEYN, P.P. *et al.* CAG repeat expansion in the TATA box-binding protein gene causes autosomal dominant cerebellar ataxia. *Brain* **124**, 1939-47 (2001).
- FUKAI, S., NUREKI, O., SEKINE, S., SHIMADA, A., VASSYLYEV, D.G., YOKOYAMA, S. Mechanisms of molecular interactions for tRNA(Val) recognition by valyl-tRNA synthetase. *RNA* **9**, 100-111 (2003).
- GEIDUSCHEK, E.P. & TOCCHINI-VALENTINI, G.P. Transcription by RNA polymerase III. *Annu Rev Biochem* **57**, 873-914 (1988).
- GEIDUSCHEK, E.P. & KASSAVETIS, G.A. The RNA polymerase III transcription apparatus. *J Mol Biol* **310**, 1-26 (2001).
- GILBERT, N., LUTZ, S., MORRISH, T.A. & MORAN, J.V. Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol Cell Biol* **25**, 7780-95 (2005).
- GLÖCKNER, G., SZAFRANSKI, K., WINCKLER, T. *et al.* The complex repeats of Dictyostelium discoideum. *Genome Res.* **11**, 585-594 (2001a).
- GLÖCKNER, G., SZAFRANSKI, K., WINCKLER, T. *et al.* The complex repeats of Dictyostelium discoideum. *Genome Res.* **11**, 585-594 (2001b).
- GLÖCKNER, G., EICHINGER, L., SZAFRANSKI, K. *et al.* Sequence and analysis of chromosome 2 of Dictyostelium discoideum. *Nature* **418**, 79-85 (2002).
- GOMER, R.H. Cell density sensing in a eukaryote. *ASM News* **65**, 23-29 (1999).
- HAAG, A.L., LIN, J.H. & LEVIN, H.L. Evidence for the packaging of multiple copies of Tf1 mRNA into particles and the trans priming of reverse transcription. *J Virol* **74**, 7164-70 (2000).
- HAHN, S. & ROBERTS, S. The zinc ribbon domains of the general transcription factors TFIIB and Brf: conserved functional surfaces but different roles in transcription initiation. *Genes Dev* **14**, 719-30 (2000).
- HANAHAH, D. Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166**, 557-80 (1983).
- HÄSLER, J., S., K. Alu elements as regulators of gene expression. *Nuc. Acids Res.* **34**, 5491-5497 (2006).
- HENTSCHEL, U., ZÜNDORF, I., DINGERMAN, T. & WINCKLER, T. On the problem of establishing the subcellular localization of Dictyostelium retrotransposon TRE5-A proteins by biochemical analysis of nuclear extracts. *Anal. Biochem* **296**, 83-91 (2001).
- HERNANDEZ, N. TBP, a universal eukaryotic transcription factor? *Genes Dev* **7**, 1291-308 (1993).
- HOCHULI, E. Purification of recombinant proteins with metal chelate adsorbent. *Genet Eng (NY)* **12**, 87-98 (1990).
- HOFMANN, J., SCHUMANN, G., BORSCHET, G. *et al.* Transfer RNA genes from Dictyostelium discoideum are frequently associated with repetitive elements and contain consensus boxes in their 5'-flanking and 3'-flanking regions. *J. Mol. Biol.* **222**, 537-552 (1991).
- HOHJOH, H. & SINGER, M.F. Sequence specific single-stranded RNA-binding protein encoded by the human LINE-1 retrotransposon. *EMBO J.* **16**, 6034-6043 (1997).
- HORN, J., DIETZ-SCHMIDT, A., ZUNDORF, I., GARIN, L., DINGERMAN, T. & WINCKLER, T. A Dictyostelium protein binds to distinct oligo(dA).oligo(dT) DNA sequences in the C-module of the retrotransposable element DRE. *Eur. J. Biochem.* **265**, 441-448 (1999).
- JOAZEIRO, C.A., KASSAVETIS, G.A. & GEIDUSCHEK, E.P. Identical components of yeast transcription factor IIIB are required and sufficient for transcription of TATA box-containing and TATA-less genes. *Mol Cell Biol* **14**, 2798-808 (1994).

- JUO, Z.S., CHIU, T.K., LEIBERMAN, P.M., BAIKALOV, I., BERK, A.J. & DICKERSON, R.E. How proteins recognize the TATA box. *J Mol Biol* **261**, 239-54 (1996).
- JUO, Z.S., KASSAVETIS, G.A., WANG, J., GEIDUSCHEK, E.P. & SIGLER, P.B. Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* **422**, 534-9 (2003).
- KAPUST, R.B. & WAUGH, D.S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668-74 (1999).
- KASSAVETIS, G., BARDELEBEN, BRAUN, JOAZEIRO. Transcription by RNA Polymerase III. *Transcription mechanisms and regulation* **3**, 107-126 (1994).
- KASSAVETIS, G.A., BRAUN, B.R., NGUYEN, L.H. & GEIDUSCHEK, E.P. S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* **60**, 235-45 (1990).
- KASSAVETIS, G.A., BARTHOLOMEW, B., BLANCO, J.A., JOHNSON, T.E. & GEIDUSCHEK, E.P. Two essential components of the Saccharomyces cerevisiae transcription factor TFIIIB: transcription and DNA-binding properties. *Proc Natl Acad Sci U S A* **88**, 7308-12 (1991).
- KASSAVETIS, G.A., JOAZEIRO, C.A., PISANO, M. *et al.* The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIIB. *Cell* **71**, 1055-64 (1992).
- KASSAVETIS, G.A., KUMAR, A., RAMIREZ, E. & GEIDUSCHEK, E.P. Functional and structural organization of Brf, the TFIIIB-related component of the RNA polymerase III transcription initiation complex. *Mol Cell Biol* **18**, 5587-99 (1998a).
- KASSAVETIS, G.A., KUMAR, A., LETTS, G.A. & GEIDUSCHEK, E.P. A post-recruitment function for the RNA polymerase III transcription-initiation factor IIIB. *Proc Natl Acad Sci U S A* **95**, 9196-201 (1998b).
- KASSAVETIS, G.A., LETTS, G.A. & GEIDUSCHEK, E.P. A minimal RNA polymerase III transcription system. *Embo J* **18**, 5042-51 (1999).
- KASSAVETIS, G.A., LETTS, G.A. & GEIDUSCHEK, E.P. The RNA polymerase III transcription initiation factor TFIIIB participates in two steps of promoter opening. *Embo J* **20**, 2823-34 (2001).
- KASSAVETIS, G.A., DRISCOLL, R. & GEIDUSCHEK, E.P. Mapping the principal interaction site of the Brf1 and Bdp1 subunits of Saccharomyces cerevisiae TFIIIB. *J Biol Chem* **281**, 14321-9 (2006).
- KATZ, R. & SKALKA, A. The retroviral enzymes. *Ann. Rev. Biochem.* **63**, 133-173 (1994).
- KAZAZIAN, H.H. Mobile elements and disease. *Curr. Opin. Genet. Dev.* **8**, 343-350 (1998).
- KAZAZIAN, H.H. Mobile Elements: Drivers of Genome Evolution. *Science* **303**, 1626-1632 (2004).
- KENNA, M.A., BRACHMANN, C.B., DEVINE, S.E. & BOEKE, J.D. Invading the yeast nucleus: a nuclear localization signal at the C terminus of Ty1 integrase is required for transposition in vivo. *Mol Cell Biol* **18**, 1115-24 (1998).
- KHOO, B., BROPHY, B. & JACKSON, S.P. Conserved functional domains of the RNA polymerase III general transcription factor BRF. *Genes Dev* **8**, 2879-90 (1994).
- KIM, J.L., NIKOLOV, D.B. & BURLEY, S.K. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**, 520-7 (1993).
- KIM, J.M., VANGURI, S., BOEKE, J.D., GABRIEL, A. & VOYTAS, D.F. Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence. *Genome Res* **8**, 464-78 (1998a).
- KIM, J.M., VANGURI, S., BOEKE, J.D., GABRIEL, A. & VOYTAS, D.F. Transposable elements and genome organization: A comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence. *Genome Res* **8**, 464-478 (1998b).
- KIRCHNER, J., CONOLLY, C.M. & SANDMEYER, S.B. Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. *Science* **267**, 1488-1491 (1995).
- KOLOSHA, V.O. & MARTIN, S.L. High-affinity, non-sequence-specific RNA binding by the open reading frame 1 (ORF1) protein from long interspersed nuclear element 1 (LINE-1). *J Biol Chem* **278**, 8112-7 (2003).

- KONIJN, T.M., VAN DE MEENE, J.G.C., BONNER, J.T. & BARKLEY, D.S. The acrasin activity of adenosine-3',5'-cyclic phosphate. *Proc. Natl. Acad. Sci. USA* **58**, 1152-1154 (1967).
- KOZAK, M. The scanning model for translation: an update. *J. Cell Biol.* **108**, 229-241 (1989).
- KUBO, S., SELEME MDEL, C., SOIFER, H.S. *et al.* L1 retrotransposition in nondividing and primary human somatic cells. *Proc Natl Acad Sci U S A* **103**, 8036-41 (2006).
- KULPA, D.A. & MORAN, J.V. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum Mol Genet* **14**, 3237-48 (2005).
- KULPA, D.A. & MORAN, J.V. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat Struct Mol Biol* **13**, 655-60 (2006).
- KUMAR, A., KASSAVETIS, G.A., GEIDUSCHEK, E.P., HAMBALCO, M. & BRENT, C.J. Functional dissection of the B" component of RNA polymerase III transcription factor IIIB: a scaffolding protein with multiple roles in assembly and initiation of transcription. *Mol Cell Biol* **17**, 1868-80 (1997).
- KYHSE-ANDERSEN, J. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J Biochem Biophys Methods* **10**, 203-9 (1984).
- LANDER, E.S., LINTON, L.M., BIRREN, B. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
- LESAGE, P. & TODESCHINI, A.L. Happy together: the life and times of Ty retrotransposons and their hosts. *Cytogenet Genome Res* **110**, 70-90 (2005).
- LI, P.W., LI, J., TIMMERMAN, S.L., KRUSHEL, L.A. & MARTIN, S.L. The dicistronic RNA from the mouse LINE-1 retrotransposon contains an internal ribosome entry site upstream of each ORF: implications for retrotransposition. *Nucleic Acids Res* **34**, 853-64 (2006).
- LOOMIS, W.F. *The development of Dictyostelium discoideum.*, 551 (Ac. Press, New York, 1982).
- LOOMIS, W.F. Lateral inhibition and pattern formation in Dictyostelium. *Curr. Topics Dev. Biol.* **28**, 1-46 (1993).
- LOOMIS, W.F., WELKER, D., HUGHES, J., MAGHAKIAN, D. & KUSPA, A. Integrated maps of the chromosomes in Dictyostelium discoideum. *Genetics* **141**, 147-157 (1995).
- MAEDA, Y., OHMORI, T., ABE, T., ABE, F. & AMAGAI, A. Transition of starving Dictyostelium cells to differentiation phase at a particular position of the cell cycle. *Differentiation* **41**, 169-175 (1989).
- MALIK, H., HENIKOFF, S. & EICKBUSH, T.H. Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res.* **10**, 1307-1318 (2000).
- MALIK, H.S., BURKE, W.D. & EICKBUSH, T.H. The age and evolution of non-LTR retrotransposable elements. *Mol. Biol. Evol.* **16**, 793-805 (1999).
- MALIK, H.S. & EICKBUSH, T.H. Modular evolution of the integrase domain in the Ty3/gypsy class of LTR retrotransposons. *J. Virol.* **73**, 5186-5190 (1999).
- MALIK, H.S. & EICKBUSH, T.H. Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. *Genome Res.* **11**, 1187-1197 (2001).
- MARCHLER-BAUER, A., ANDERSON, J.B., CHERUKURI, P.F. *et al.* CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res* **33**, D192-6 (2005).
- MARSCHALEK, R., BRECHNER, T., AMON-BOHM, E. & DINGERMANN, T. Transfer RNA genes: landmarks for integration of mobile genetic elements in Dictyostelium discoideum. *Science* **244**, 1493-1496 (1989).
- MARSCHALEK, R., HOFMANN, J., SCHUMANN, G., GOSSERINGER, R. & DINGERMANN, T. Structure of DRE, a retrotransposable element which integrates with position specificity upstream of Dictyostelium discoideum tRNA genes. *Mol. Cell. Biol.* **12**, 229-239 (1992a).
- MARSCHALEK, R., HOFMANN, J., SCHUMANN, G. & DINGERMANN, T. Two distinct subforms of the retrotransposable DRE element in NC4 strains of Dictyostelium discoideum. *Nucl. Acids Res.* **20**, 6247-6252 (1992b).

- MARTIN, S.L., LI, J. & WEISZ, J.A. Deletion analysis defines distinct functional domains for protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *J Mol Biol* **304**, 11-20 (2000).
- MARTIN, S.L. & BUSHMAN, F.D. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol Cell Biol* **21**, 467-75 (2001).
- MARTIN, S.L., CRUCEANU, M., BRANCIFORTE, D. *et al.* LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J Mol Biol* **348**, 549-61 (2005).
- MARTIN, S.L. The ORF1 Protein Encoded by LINE-1: Structure and Function During L1 Retrotransposition. *J Biomed Biotechnol* **2006**, 45621 (2006).
- MATSUMOTO, T., TAKAHASHI, H. & FUJIWARA, H. Targeted nuclear import of open reading frame 1 protein is required for in vivo retrotransposition of a telomere-specific non-long terminal repeat retrotransposon, SART1. *Mol Cell Biol* **24**, 105-22 (2004).
- MATSUMOTO, T., HAMADA, M., OSANAI, M. & FUJIWARA, H. Essential domains for ribonucleoprotein complex formation required for retrotransposition of telomere-specific non-long terminal repeat retrotransposon SART1. *Mol Cell Biol* **26**, 5168-79 (2006).
- MCCCLINTOCK, B. Mutable loci in maize. *Carnegie Inst. Wash. Year Book* **47**, 155-169 (1948).
- MO, X., KOWENZ-LEUTZ, E., LAUMONNIER, Y., XU, H. & LEUTZ, A. Histone H3 tail positioning and acetylation by the c-Myb but not the v-Myb DNA-binding SANT domain. *Genes Dev* **19**, 2447-57 (2005).
- MORAN, J.V., HOLMES, S.E., NAAS, T.P., DEBERARDINIS, R.J., BOEKE, J.D. & KAZAZIAN, H.H. High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917-927 (1996).
- MOU, Z., KENNY, A.E. & CURCIO, M.J. Hos2 and Set3 promote integration of Ty1 retrotransposons at tRNA genes in *Saccharomyces cerevisiae*. *Genetics* **172**, 2157-67 (2006).
- MOURIER, T. Reverse transcription in genome evolution. *Cytogenet Genome Res* **110**, 56-62 (2005).
- MUTZEL, R. Cellular slime molds: why and how to become pluricellular. *Bull. Inst. Pasteur* **89**, 51-58 (1991).
- NAKAMURA, K., JEONG, S.Y., UCHIHARA, T. *et al.* SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet* **10**, 1441-8 (2001).
- NIKOLOV, D.B., HU, S.H., LIN, J. *et al.* Crystal structure of TFIID TATA-box binding protein. *Nature* **360**, 40-6 (1992).
- NIKOLOV, D.B., CHEN, H., HALAY, E.D., HOFFMAN, A., ROEDER, R.G. & BURLEY, S.K. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A* **93**, 4862-7 (1996).
- OKADA, N., HAMADA, M., OGIWARA, I. & OHSHIMA, K. SINEs and LINEs share common 3' sequences: a review. *Gene* **205**, 229-243 (1997).
- OSTERTAG, E.M. & KAZAZIAN, H.H. Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* **35**, 501-538 (2001a).
- OSTERTAG, E.M. & KAZAZIAN, H.H., JR. Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. *Genome Res* **11**, 2059-65 (2001b).
- PURRELLO, M., DI PIETRO, C., RAPI SARDA, A. *et al.* Localization of the human genes encoding the two subunits of general transcription factor TFIIE. *Genomics* **23**, 253-5 (1994).
- RAMESH, V., RAJBHANDARY, U.L. Importance of the anticodon sequence in the aminoacylation of tRNAs by methionyl-tRNA synthetase and by valyl-tRNA synthetase in an Archaeobacterium. *J Biol Chem* **276**, 3660-3665 (2001).
- RAPER, K.B. *Dictyostelium discoideum*, a new species of slime mold from decaying forest leaves. *J. Agr. Res.* **50**, 135-147 (1935).
- RASHKOVA, S., KARAM, S.E. & PARDUE, M.L. Element-specific localization of *Drosophila* retrotransposon Gag proteins occurs in both nucleus and cytoplasm. *Proc Natl Acad Sci U S A* **99**, 3621-6 (2002).

- ROBERTS, D.N., STEWART, A.J., HUFF, J.T. & CAIRNS, B.R. The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proc Natl Acad Sci U S A* **100**, 14695-700 (2003).
- SAMBROOK, J., MANIATIS, T. & FRIETSCH, E.F. *Molecular cloning : a laboratory manual*, 3 v. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989).
- SANGER, F., NICKLEN, S. & COULSON, A.R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-7 (1977).
- SAXENA, A., MA, B., SCHRAMM, L. & HERNANDEZ, N. Structure-function analysis of the human TFIIB-related factor II protein reveals an essential role for the C-terminal domain in RNA polymerase III transcription. *Mol Cell Biol* **25**, 9406-18 (2005).
- SCHRAMM, L. & HERNANDEZ, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev* **16**, 2593-620 (2002).
- SCHRÖDER, O., BRYANT, G.O., GEIDUSCHEK, E.P., BERK, A.J. & KASSAVETIS, G.A. A common site on TBP for transcription by RNA polymerases II and III. *Embo J* **22**, 5115-24 (2003).
- SCHUMANN, G., ZÜNDORF, I., HOFMANN, J., MARSCHALEK, R. & DINGERMANN, T. Internally located and oppositely oriented polymerase II promoters direct convergent transcription of a LINE-like retroelement, the Dictyostelium Repetitive Element, from Dictyostelium discoideum. *Mol. Cell. Biol.* **14**, 3074-3084 (1994a).
- SCHUMANN, G., ZUNDORF, I., HOFMANN, J., MARSCHALEK, R. & DINGERMANN, T. Internally located and oppositely oriented polymerase II promoters direct convergent transcription of a line-like retroelement, the Dictyostelium repetitive element, from Dictyostelium discoideum. *Mol. Cell. Biol.* **14**, 3074-3084 (1994b).
- SCHWEDE, T., KOPP, J., GUEX, N. & PEITSCH, M.C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* **31**, 3381-5 (2003).
- SELEME, M.C., DISSON, O., ROBIN, S., BRUN, C., TENINGES, D. & BUCHETON, A. In vivo RNA localization of I factor, a non-LTR retrotransposon, requires a cis-acting signal in ORF2 and ORF1 protein. *Nucleic Acids Res* **33**, 776-85 (2005).
- SHEN, Y., KASSAVETIS, G.A., BRYANT, G.O. & BERK, A.J. Polymerase (Pol) III TATA box-binding protein (TBP)-associated factor Brf binds to a surface on TBP also required for activated Pol II transcription. *Mol Cell Biol* **18**, 1692-700 (1998).
- SIOL, O., BOUTLILISS, M., CHUNG, T., GLÖCKNER, G., DINGERMANN, T. & WINCKLER, T. Role of RNA polymerase III transcription factors in the selection of integration sites by the dictyostelium non-long terminal repeat retrotransposon TRE5-A. *Mol Cell Biol* **26**, 8242-51 (2006).
- SOIFER, H.S. & KASAHARA, N. Retrotransposon-adenovirus hybrid vectors: efficient delivery and stable integration of transgenes via a two-stage mechanism. *Curr Gene Ther* **4**, 373-84 (2004).
- SONG, J., XU, Q., OLSEN, R., LOOMIS, W., SHAULSKY, G., KUSPA, A., SUCGANG, R. Comparing the Dictyostelium and Entamoeba genomes reveals an ancient split in the conosa lineage. *PLoS Computational Biology* **1**, 579-584 (2005).
- STUDIER, F.W., ROSENBERG, A.H., DUNN, J.J. & DUBENDORFF, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60-89 (1990).
- SUCGANG, R., CHEN, G., LIU, W. *et al.* Sequence and structure of the extrachromosomal palindrome encoding the ribosomal RNA genes in Dictyostelium. *Nucleic Acids Res* **31**, 2361-8 (2003).
- SUSSMAN, R. & SUSSMAN, M. Cultivation of Dictyostelium discoideum in axenic culture. *Biochem. Biophys. Res. Commun.* **29**, 53-55 (1967).
- SZAFRANSKI, K., GLÖCKNER, G., DINGERMANN, T. *et al.* Non-LTR retrotransposons with unique integration preferences downstream of Dictyostelium discoideum transfer RNA genes. *Mol. Gen. Genet.* **262**, 772-780 (1999a).
- SZAFRANSKI, K., GLOCKNER, G., DINGERMANN, T. *et al.* Non-LTR retrotransposons with unique integration preferences downstream of Dictyostelium discoideum tRNA genes. *Mol Gen Genet* **262**, 772-80 (1999b).
- WANG, Z. & ROEDER, R.G. Structure and function of a human transcription factor TFIIB subunit that is evolutionarily conserved and contains both TFIIB- and high-mobility-group protein 2-related domains. *Proc Natl Acad Sci U S A* **92**, 7026-30 (1995).

- WEINER, A.M.** Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu Rev Biochem* **55**, 631-61 (1986).
- WHITE, R.J.** RNA Polymerase III transcription. *R.G. Landes Co., Austin, Texas* (1994).
- WILBER, A., FRANDSEN, J.L., GEURTS, J.L., LARGAESPADA, D.A., HACKETT, P.B. & McIVOR, R.S.** RNA as a source of transposase for sleeping beauty-mediated gene insertion and expression in somatic cells and tissues. *Mol Ther* **13**, 625-30 (2006).
- WILBER, A., WANGENSTEEN, K.J., CHEN, Y. et al.** Messenger RNA as a Source of Transposase for Sleeping Beauty Transposon-mediated Correction of Hereditary Tyrosinemia Type I. *Mol Ther* (2007).
- WILHELM, F.X., WILHELM, M. & GABRIEL, A.** Reverse transcriptase and integrase of the *Saccharomyces cerevisiae* Ty1 element. *Cytogenet Genome Res* **110**, 269-87 (2005).
- WILLIS, I.** RNA polymerase III: Genes, factors and transcriptional specificity. *Eur. J. Biochem.* **212**, 1-11 (1993).
- WILLIS, I.M.** A universal nomenclature for subunits of the RNA polymerase III transcription initiation factor TFIIIB. *Genes Dev* **16**, 1337-8 (2002).
- WINCKLER, T.** Retrotransposable elements in the *Dictyostelium discoideum* genome. *Cell. Mol. Life Sci.* **54**, 383-393 (1998).
- WINCKLER, T., TSCHEPKE, C., DEHOSTOS, E.L., JENDRETZKE, A. & DINGERMANN, T.** Tdd-3, a tRNA gene-associated poly(A) retrotransposon from *Dictyostelium discoideum*. *Mol Gen Genet* **257**, 655-661 (1998).
- WINCKLER, T., DINGERMANN, T. & GLÖCKNER, G.** *Dictyostelium* mobile elements: strategies to amplify in a compact genome. *Cell. Mol. Life Sci.* **59**, 2097-2111 (2002a).
- WINCKLER, T., DINGERMANN, T. & GLOCKNER, G.** *Dictyostelium* mobile elements: strategies to amplify in a compact genome. *Cell. Mol. Life Sci.* **59**, 2097-2111 (2002b).
- WINCKLER, T., SZAFRANSKI, K. & GLÖCKNER, G.** Transfer RNA gene-targeted integration: an adaptation of retrotransposable elements to survive in the compact *Dictyostelium discoideum* genome. *Cytogenet Genome Res* **110**, 288-98 (2005).
- WU, X., LI, Y., CRISE, B. & BURGESS, S.M.** Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300**, 1749-51 (2003).
- XIE, W., GAI, X., ZHU, Y., ZAPPULLA, D.C., STERNGLANZ, R. & VOYTAS, D.F.** Targeting of the yeast Ty5 retrotransposon to silent chromatin is mediated by interactions between integrase and Sir4p. *Mol. Cell. Biol.* **21**, 6606-6614 (2001).
- YANG, J., MALIK, H.S. & EICKBUSH, T.H.** Identification of the endonuclease domain encoded by R2 and other site-specific, non-long terminal repeat retrotransposable elements. *Proc. Natl. Acad. Sci. USA* **96**, 7847-7852 (1999).
- YIEH, L., KASSAVETIS, G., GEIDUSCHEK, E.P. & SANDMEYER, S.B.** The Brf and TBP subunits of the RNA polymerase III transcription factor IIIB mediate position-specific integration of the gypsy-like element Ty3. *J. Biol. Chem.* **275**, 29800-29807 (2000).
- YIEH, L., HATZIS, H., KASSAVETIS, G. & SANDMEYER, S.B.** Mutational analysis of the transcription factor IIIB-DNA target of Ty3 retroelement integration. *J Biol Chem* **277**, 25920-8 (2002).
- ZHAO, X. & HERR, W.** Role of the inhibitory DNA-binding surface of human TATA-binding protein in recruitment of human TFIIIB family members. *Mol Cell Biol* **23**, 8152-60 (2003).
- ZHU, Y., DAI, J., FUERST, P.G. & VOYTAS, D.F.** Controlling integration specificity of a yeast retrotransposon. *Proc. Natl. Acad. Sci. USA* **100**, 5891-5895 (2003).
- ZOU, S. & VOYTAS, D.F.** Silent chromatin determines target preference of the retrotransposon Ty5. *Proc. Natl. Acad. Sci. USA* **94**, 7412-7416 (1997).

# Lebenslauf

## Persönliche Daten:

Name: Tieu Thanh Chung  
Geboren am 01.03.1978  
Geburtsort Quang Dong, Vietnam  
Staatsangehörigkeit deutsch

## Schulbildung:

1984 – 1988 Grundschule St. Nikolaus, Konz  
1988 – 1994 Gymnasium Konz  
1994 – 1997 Max-Planck-Gymnasium Trier, Abschluss: Abitur

## Studium:

10/1997 – 04/2002 Pharmaziestudium an der Johannes Gutenberg  
Universität, Mainz  
04/2002 2. Staatsexamen in Pharmazie  
05/2003 3. Staatsexamen in Pharmazie und Erlangung der  
Approbation

## Promotion

05/2003 – 05/2007 Promotion an der Johann Wolfgang Goethe Universität  
Frankfurt am Institut für Pharmazeutische Biologie unter  
der Leitung von Prof. Dr. T. Winckler

## Akademische Lehrer:

Prof. Dr. G. Dannhardt, Prof. Dr. T. Dingermann, Prof. Dr. B. Epe, Prof. Dr. P.  
Langguth, Prof. Dr. J. Stöckigt, Prof. Dr. U. Pindur und Prof. Dr. T. Winckler

## Publikationen

***Molecular and Cellular Biology, Nov. 2006, Vol. 26, No.22 (8242 – 8251)  
Role of RNA Polymerase III Transcription Factors in the Selection of  
Integrations Sites by the Dictyostelium Non-Long Terminal Repeat  
Retrotransposon TRE5-A***

Oliver Siol, Moustapha Boutiliss, Thanh Chung, Gernot Glöckner, Theodor Dingermann, Thomas Winckler

***Convergent evolution of tRNA gene targeting by retrotransposons in compact  
genomes***

Thanh Chung, Oliver Siol, Theodor Dingermann, Thomas Winckler  
Molecular Cell 2007, submitted

## Posterpräsentationen

***How TRE5-A Retrotransposons select tRNA-genes as Integration Targets – an  
Update***

Oliver Siol, Thanh Chung, Theodor Dingermann und Thomas Winckler  
International *Dictyostelium* Conference 2004, Sainte Adèle, Québec (Kanada)

***Targeted Integration of Dictyostelium Retrotransposons: Concept for Safer  
Gene Transfer Vectors***

Thomas Winckler, Oliver Siol, Thanh Chung, Theodor Dingermann  
ZAFES Meeting 2005, Frankfurt am Main

***RNA Polymerase III Transcription Factor Complex TFIIIB determines  
Chromosomal Integrations Sites of Retrotransposon TRE5-A***

Thanh Chung, Oliver Siol, Theodor Dingermann, Thomas Winckler  
International *Dictyostelium* Conference 2005, Autrans (Frankreich)



# Danksagung

Allen voran danke ich Prof. Dr. Thomas Winckler. Für die Überlassung des sehr interessanten Themas, seine hervorragende wissenschaftliche Betreuung, geduldige Anleitung und Unterstützung, seinen steten Überblick, seine Wertschätzung und für die Aufnahme in einen tollen Arbeitskreis.

Prof. Dr. Theodor Dingermann für die freundliche Aufnahme in seinem Institut, die zahlreichen Diskussionsbeiträge und Ideen während der Seminare sowie für die Übernahme des Zweitgutachtens.

Prof. Dr. Rolf Marschalek für seine kritischen und kreativen Diskussionsbeiträge während der Arbeitsseminare.

Dr. Ilse Zündorf für die engagierte und kompetente Herstellung der monoklonalen Antikörper (in memoriam Sir Peter und Ustinov).

Meinen unmittelbaren Kollaborateuren für die lustige Zeit in und außerhalb des S1-Bereichs, dafür, dass sie mich in ihren Kreis aufgenommen, molekularbiologischen Input-overflow in verständliche Häppchen übersetzt und mich immer wieder aufgefangen haben:

Dr. Moni Baik für wertvolle Alltagstipps (beschrifte alles, sonst landet bei Oli in der Schublade) und dafür, dass sie mir jedes Fettnäpfchen vergibt;

Dr. Oli Siol alias linke Magenhälfte für die praxisorientierte Einführung ins effiziente Klonieren;

Dr. Ingo Jennes für das Besorgen von einfach allem, angefangen bei Verbrauchsmaterial über Frühstück bis hin zu Schultern. Mein Respekt und meine Freundschaft gelten einem Mann, der ein Wort niemals zweimal identisch schreibt;

Moustapha Boutliliss, auch wenn wir uns bis zuletzt nicht einigen konnten, ob die asiatische oder die marokkanische Höflichkeit sich durchsetzen wird.

Den Rolfis Dr. Adelheid Bursen, Dr. Sven Moritz, Dr. Sebastian Scharf, Dr. Claus Meyer, Dr. Anne Gaußmann, Dr. Björn Schneider, Erik Kowarz, Irina Eberle, Miriam Jacobi, Clarissa Oehm, Anne Benedikt, Karen Schley und der *Bioll next generation* Johanna Blees und Nadine Spilger möchte ich für so manchen wertvollen Labortip und die angenehme Zeit abseits der Bench danken – und natürlich für das Heruntergeben ungezählter Flaschen etc., die sich irgendwie immer außerhalb meiner Reichweite befanden.

Jens Rabenstein für seine Freundschaft, zahllose Mitnahmen im BMW und für die Erkenntnis, dass man morgen auch noch mit der Enttäuschung leben kann. Meinem Bioll-Ziehvater Zech, Dr. Jürgen für die Einführung in die Dunkle Seite der Macht und Ute Steiner für die angenehme Zusammenarbeit während (viel zu vieler!) Bio III-Praktika.

Den TAs Silvia Bracharz und Jenny Merkens, besonders aber Freia Pfeil für ihre liebe und kompetente Unterstützung, egal wie kurzfristig sie erbeten wurde.

Frau Lemmer für ihre Hilfe bei der „Reise nach Jerusalem – Hörsaal-Edition“ und ihre stets offene Tür.

Dr. Anke Lühe für die Paper-„Axis of Swiss“ und Not-Formatierungen.

Meinen Freunden für ihr Verständnis und dafür, dass sie mich trotz seltener Kontakte noch nicht verstoßen haben.

Leif dafür, dass er alles geduldig erträgt.

Meinen Geschwistern für ihren technical support zu jeder Tages- und Nachtzeit – Ihr habt Euer Bestes gegeben...

Meiner Familie für ihre Unterstützung und ihren Rückhalt auf dem steinigen Weg zur Weltherrschaft.