

**Die Flexibilität von TAL-Effektoren in der Natur und als
molekulare Werkzeuge zur Genaktivierung und
Genommodifikation**

Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität
Halle-Wittenberg,

vorgelegt

von Frau Annekatriin Richter

geb. am 31.01.1987 in Wippra

Gutachter:

1. Prof. Dr. Jens Boch
2. Prof. Dr. Sven-Erik Behrens
3. Prof. Dr. Alain Tissier

Tag der Verteidigung: 11.10.2019

Zusammenfassung

Transcription activator like effectors (TALEs) sind bakterielle Effektorproteine, die von phytopathogenen *Xanthomonas spp.* über das Typ-III-Sekretionssystem in die Zellen des pflanzlichen Wirts transloziert werden. Dort gelangen sie in den Nukleus, binden spezifisch an Promotorsequenzen und induzieren die Transkription von Genen. Dafür verfügen TALEs über eine DNA-Bindedomäne aus 33-35 Aminosäuren (AS)-Wiederholungen, den *repeats*, die sich vor allem an den AS-Positionen 12 und 13 (*repeat variable diresidue*, RVD) voneinander unterscheiden. Das *repeat*-Array bildet eine hochregelmäßige superhelikale Struktur, wobei jeder *repeat* ein Nukleotid des codierenden DNA-Stranges erkennt. Die Basenspezifität wird dabei durch die 13. AS eines jeden *repeats* vermittelt. Auf Grund dieser einfachen, modularen DNA-Erkennung werden TALEs zur Erstellung molekularer Werkzeuge zur spezifischen Geninduktion und Genommodifikation verwendet. In dieser Arbeit wurde zum einen der Einfluss seltener *repeats* einer aberranten Länge von 30, 39, 40 oder 42 AS auf die DNA-Erkennung durch TALEs untersucht. Es wurde festgestellt, dass TALEs mit einem aberranten *repeat* neben einer optimalen auch eine Zielsequenz mit einer Leserasterverschiebung um -1 Nukleotid in der Nähe des aberranten *repeats* erkennen. Zum anderen wurde die Geninduktion durch TALE-Aktivatoren am Modell des HIV-1-Promotors untersucht. Es war bereits bekannt, dass TALEs zwar nicht auf, aber stromaufwärts von den drei SP1-Bindestellen des Kernpromotors Transkription induzieren können. In dieser Arbeit wurde mit Hilfe einer Promotormutante gezeigt, dass eine effiziente Geninduktion mit artifiziellen TALEs durch das kooperative Zusammenwirken mit endogenen Transkriptionsfaktoren wie SP1 möglich ist. Darüber hinaus konnte eine effiziente Geninduktion am mutierten HIV-1-Promotor auch durch die synergistische Aktivierung mit mehreren künstlichen TALEs, sowie durch einen TALE mit multiplen Aktivierungsdomänen erzielt werden. Unter Verwendung eines aberranten *repeats* und RVDs multipler Basenspezifität wurden außerdem TALEs erstellt, die Transkription an variablen Promotorsequenzen verschiedener HIV-1 Subtypen induzieren konnten. TALE-Nukleasen (TALEN) zur Genommodifikation werden durch Fusion einer FokI-Endonukleasedomäne erstellt. Da FokI DNA als Dimer schneidet, werden TALEN als Paare eingesetzt. In dieser Arbeit wurden zur Vereinfachung dieser Technologie *single chain* TALEN mit einem einkettigen FokI-Endonukleasedomänendimer erstellt, die spezifisch und effizient DNA *in vitro* schneiden, jedoch keine Mutationen *in vivo* verursachen. Weiterhin wurden durch zusätzliche RVD-Base-Fehlpaarungen hochspezifische TALEN-Paare erstellt, die zwei Sequenzen anhand nur einer Base unterscheiden.

Summary

Transcription activator like effectors (TALEs) are effector proteins of plant pathogenic *Xanthomonas* species. The bacteria inject them via the type III-secretion system into the host plant cells, where they enter the nucleus, specifically bind to promoter sequences and thereby induce genes. Therefore, TALEs contain a unique DNA-binding domain consisting of 33 to 35 amino acid tandem repeats, which mostly differ at positions 12 and 13, the so called repeat variable diresidue (RVD). The repeat array forms a highly regular superhelical structure that wraps around the DNA double helix exposing the 13th residue of each repeat to the bases of the DNA coding strand. Thus, each repeat recognizes one nucleotide with its 13th amino acid specifying which base is bound. Due to this simple mode of DNA recognition, TALEs are applied as molecular tools for gene induction and genome modification. In this work, the influence of rare aberrant repeats of 30, 39, 40 or 42 amino acids on TALE-DNA recognition was investigated. It was found that TALEs with aberrant repeats not only recognize their optimal target sequence but also a sequence with a -1-nucleotide frameshift in proximity to the aberrant repeat. Furthermore, the induction of transcription at the HIV-1 promoter by TALE-activators was studied. It was known that TALEs do not activate transcription at the three SP1-binding sites of the HIV-1 core promoter, but do so when targeting upstream sequences. Using HIV-1 promoter mutants, it was shown in this work that artificial TALEs are able to cooperatively induce transcription together with endogenous transcription factors like SP1. On the other hand, an efficient transcriptional induction at the mutant promoter was possible by synergistic activation with multiple TALE-activators or a TALE with multiple activation domains. Exploiting an aberrant repeat and RVDs with multiple base specificities, TALE-activators were generated that efficiently induced transcription at variable promoter sequences of different HIV-1 subtypes. TALE nucleases (TALEN) for genome editing are fusion proteins with a FokI endonuclease domain, which only cleaves DNA as a dimer. Therefore, TALEN are applied pairwise. To simplify this method single chain TALEN (scTALEN) with a monomeric FokI endonuclease domain dimer were used in this work. Although specifically cleaving DNA *in vitro* efficiently, scTALEN did not induce mutation in living cells. Finally, TALEN-pairs with single nucleotide specificity were generated employing additional RVD-base mismatches.

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis	V
Tabellenverzeichnis	VII
Abkürzungsverzeichnis	VIII
1. Einleitung	1
1.1 TAL-Effektoren	1
1.1.1 Zielgene	2
1.1.2 TALE-basierte Pflanzenresistenzmechanismen	3
1.1.3 Der strukturelle Aufbau und die DNA-Erkennung von TALEs	5
1.1.4 Die Variabilität der repeats	6
1.2 Die Anwendung von TALEs	7
1.2.1 TALE-Aktivatoren	8
1.2.1.1 Die Transkription	8
1.2.1.2 Die transkriptionelle Aktivierung durch TALEs	9
1.2.2 TALEN	10
1.3 Das CRISPR/Cas-System	11
1.4 Das humane Immundefizienz-Virus	13
1.5 Vorarbeiten	14
1.6 Zielstellung	16
2. Material und Methoden	17
2.1 Material	17
2.1.1 Verwendete Plasmide	17
2.1.2 Verwendete Oligonukleotide	20
2.1.3 Biologisches Material	21
2.1.3.1 Verwendeter Bakterianstamm	21
2.1.3.2 Verwendete Zelllinien, Kultivierung und Passage	21
2.1.3.3 Verwendetes Pflanzenmaterial	22
2.1.4 Chemikalien, Geräte und Verbrauchsmaterialien	22
2.1.5 Verwendete Medien und Zusätze	23
2.1.6 Verwendete Lösungen und Puffer	24
2.2 Molekulargenetische Methoden	25

2.2.1	Präparation von Nukleinsäuren.....	25
2.2.1.1	Plasmidpräparation	25
2.2.1.2	Isolation von RNA aus humanen Zellen.....	25
2.2.1.3	Isolation von genomischer DNA aus humanen Zellen.....	25
2.2.2	Spektrophotometrische Nukleinsäurebestimmung	26
2.2.3	PCR-Amplifikation von DNA-Fragmenten.....	26
2.2.4	Restriktionsanalysen	26
2.2.5	Agarosegelelektrophorese.....	27
2.2.5.1	Auftrennung von DNA-Fragmenten.....	27
2.2.5.2	Auftrennung von Gesamt-RNA.....	27
2.2.6	Reinigung von DNA-Fragmenten aus Reaktionsansätzen oder Agarosegelen	27
2.2.7	Phosphorylierung von PCR-Produkten	27
2.2.8	Klonierung von DNA-Fragmenten	28
2.2.9	Die Erstellung von Luciferase-Reporterkonstrukten	29
2.2.10	Die Erstellung von β -Glucuronidase-Reporterkonstrukten	29
2.2.11	Die Erstellung artifizierlicher TALE-Derivate mit der Golden TAL Technologie ...	29
2.2.12	Transformation chemisch kompetenter <i>E. coli</i>	31
2.2.13	Transformation elektrisch kompetenter Bakterienzellen.....	31
2.2.14	Sequenzierung von DNA.....	32
2.2.15	Transfektion humaner Zellen mit Plasmiden.....	32
2.2.16	Transkriptionsstartbestimmung mittels 5'-RACE	33
2.2.17	T7 Endonuklease I-Verdau zur Detektion von Mutationen.....	33
2.3	Biochemische Methoden	34
2.3.1	Quantitative Analyse der Luciferaseaktivität	34
2.3.2	Untersuchung der subzellulären Lokalisation	35
2.3.3	Expressionsanalysen mittels Immunoblot	36
2.3.3.1	SDS-PAGE	36
2.3.3.2	Immunoblot	37
2.3.4	Proteinbestimmung nach Bradford	38
2.3.5	Proteinreinigung aus <i>E. coli</i>	38
2.3.6	EMSA	39
2.3.7	Analyse der β -Glucuronidase Aktivität	39
2.3.7.1	Quantitative GUS-Bestimmung.....	40
2.3.7.2	Qualitative GUS-Bestimmung.....	40

2.3.8	TALEN-Aktivitätsanalysen <i>in vitro</i>	41
3.	Ergebnisse	42
3.1	TALE-repeats mit abweichender Aminosäuren-Anzahl	42
3.1.1	Ein aberranter repeat erlaubt die Erkennung einer -1 Nukleotid- Leserasterverschiebung in der Ziel-DNA-Sequenz <i>in planta</i> und <i>in vitro</i>	42
3.1.2	TALEs und TALEN mit aberrantem <i>repeat</i> tolerieren eine -1 Nukleotid- Leserasterverschiebung an verschiedenen Positionen der Ziel-DNA-Sequenz.....	44
3.1.3	Ein aberranter <i>repeat</i> erlaubt die Erkennung einer Ziel-DNA-Sequenz bei Deletion des korrespondierenden oder eines benachbarten Nukleotides	46
3.1.4	TALEs mit geringer <i>repeat</i> -Anzahl werden durch einen 40 <i>As-repeat</i> geschwächt.....	48
3.1.5	Aberrante <i>repeats</i> nehmen an der Erkennung der Ziel-DNA-Sequenz teil.....	50
3.1.6	Zwei aberrante <i>repeats</i> im Tandem verhalten sich anders als ein Einzelner.....	52
3.1.7	PthXo3 und AvrXa7 tolerieren Indels in ihrer Ziel-DNA-Sequenz	53
3.1.8	Künstliche AvrXa7-Derivate erlauben eine Beurteilung des Einflusses von aberrantem <i>repeat</i> und RVD-Komposition auf die Erkennung verschiedener Ziel- DNA-Sequenzen	58
3.2	Untersuchung der TALE-vermittelten Transkriptionsinduktion am HIV-1-Promotor	62
3.2.1	Bestimmung des ArHIV6-induzierten Transkriptionsstarts am HIV-1-Promotor	62
3.2.2	Cas9-Aktivatoren zeigen ähnliches Aktivitätsmuster am HIV-1-Promotor wie TALE-Aktivatoren.....	64
3.2.3	Analyse der Positionsabhängigkeit der TALE-vermittelten Aktivierung am HIV-1-Promotor mit Hilfe einer Promotormutante	67
3.2.3.1	Nachweise der Bindung des HIV-1-Promotors mit mutierten SP1- Bindestellen durch Derivate von ArHIV1 und 2.....	68
3.2.3.2	Synergistische Aktivierung an der SP1-Bindestellenmutante des HIV-1- Promotors durch TALEs.....	71
3.2.4	Verstärkte Geninduktion durch artifizielle TALEs mit mehreren Aktivierungsdomänen.....	75
3.2.5	TALE-Aktivatoren für HIV-1 Subtypen	78
3.3	Die Konstruktion von <i>single chain</i> TALEN (scTALEN)-Varianten und hochspezifischen TALEN-Paaren	83
3.3.1	Studien zu Design und Aktivität von scTALEN	84

3.3.1.1	Die Auswahl einer geeigneten FokI-Variante	85
3.3.1.2	Restriktionsstudien mit verschiedenen scTALEN-Varianten.....	87
3.3.1.2.1	TALEN und scTALEN mit der FokI-inaktivierenden D450A-Mutation erzeugen DSBs	88
3.3.1.2.2	Beide scFokI-Domänen einer dscTALEN können DSBs induzieren.....	90
3.3.1.3	Test von scTALEN- und Doppel-scTALEN-Varianten in zwei Modellsystemen.....	93
3.3.1.3.1	scTALEN- und Doppel-scTALEN für zwei Gene des Zebrafischblinds <i>Danio rerio</i>	93
3.3.1.3.2	Für <i>CCR5</i> entworfene scTALEN- und Doppel-scTALEN zeigen keine Aktivität in humanen Zellen	96
3.3.2	Hochspezifische TALEN-Paare.....	101
4.	Diskussion.....	108
4.1	Aberrante TALE- <i>repeats</i>	108
4.1.1	Wie aberrante <i>repeats</i> die DNA-Erkennung durch TALEs beeinflussen.....	108
4.1.2	Die Flexibilität natürlicher TALEs	113
4.1.3	Aberrante <i>repeats</i> erweitern das Anwendungsspektrum von TALEs	115
4.2	Die TALE-vermittelte Transkriptionsinduktion am HIV-Promotor.....	117
4.2.1	Die Aktivierung von HIV-1 durch künstliche Transkriptionsfaktoren.....	117
4.2.2	Strategien zur effizienten Anwendung künstlicher Transkriptionsaktivatoren ...	119
4.2.3	Die Aktivierung verschiedener HIV-1-Varianten	122
4.3	Betrachtungen zur nicht nachweisbaren Aktivität von scTALEN <i>in vivo</i>	124
4.4	Die Erstellung hochspezifischer TALEN	127
5.	Literatur	131
6.	Publikationen	145

Abbildungsverzeichnis

Abbildung 1.1: Funktionsweise und struktureller Aufbau natürlicher TALEs	4
Abbildung 1.2: Schematische Darstellung eines DNA-gebundenen TALEN-Paares.....	11
Abbildung 1.3: Schematische Darstellung der Bindung von <i>Streptococcus pyogenes</i> Cas9 (SpCas9) an eine Ziel-DNA-Sequenz mittels <i>single guide</i> RNA (sgRNA). 12	
Abbildung 1.4: Künstliche TALE-Aktivatoren am HIV-1-Promotor	15
Abbildung 3.1.1: TALE-Derivate mit aberrantem <i>repeat</i> tolerieren -1 Nukleotid- Leserasterverschiebungen	43
Abbildung 3.1.2: Aberrante <i>repeats</i> funktionieren an verschiedenen Positionen der <i>repeat</i> - Folge.....	45
Abbildung 3.1.3: Ein aberranter <i>repeat</i> kann die <i>repeat</i> -Folge verlassen.....	47
Abbildung 3.1.4: Ein 40 <i>As-repeat</i> schwächt TALEs mit geringer <i>repeat</i> -Anzahl.....	49
Abbildung 3.1.5: Aberrante <i>repeats</i> nehmen an der Erkennung der Ziel-DNA-Sequenz teil .	51
Abbildung 3.1.6: Einer von zwei 40 <i>As-repeats</i> im Tandem verlässt die <i>repeat</i> -Region bei DNA-Bindung	52
Abbildung 3.1.7: PthXo3 erkennt eine Vielzahl von Ziel-DNA-Sequenzen	54
Abbildung 3.1.8: AvrXa7 erkennt eine Vielzahl von Ziel-DNA-Sequenzen.....	56
Abbildung 3.1.9: Artificielle AvrXa7-Derivate zeigen den Einfluss von aberrantem <i>repeat</i> und RVD-Komposition auf das Spektrum akzeptierter Boxen.....	59
Abbildung 3.2.1: Der durch ArtHIV6 induzierte Transkriptionsstart entspricht dem Natürlichen	63
Abbildung 3.2.2: Cas9-Aktivatoren zeigen mit sgRNAs für den HIV-1-Promotor ähnliche Positionseffekte wie die entsprechenden TALE-Aktivatoren, wobei die induzierte Aktivität von der Menge der sgRNA abhängt.....	66
Abbildung 3.2.3: TALEs am originalen HIV-1-Promotor und einer SP1-Bindestellen- Promotormutante	68
Abbildung 3.2.4: Die DNA-Bindedomänen von ArtHIV1 und 2 binden ein Fragment des originalen HIV-1-Promotors mit höherer Affinität als ein Fragment der Promotormutante	70
Abbildung 3.2.5: Die TALE-Aktivatoren ArtHIV1, 2, 5, 6 und TALE5 induzieren Luciferasereporteraktivität am HIV-1-Promotor und einer SP1- Bindestellen-Promotormutante auf unterschiedliche Weise	74
Abbildung 3.2.6: ArtHIV6-Derivate mit multiplen ADs aktivieren unterschiedliche Luciferasereporterkonstrukte mit erhöhter Effizienz	77

Abbildung 3.2.7: RVDs multipler Spezifität und aberrante <i>repeats</i> ermöglichen künstlichen TALEs an Promotorsequenzen verschiedener HIV-1 Subtypen Reporteraktivität zu induzieren	81
Abbildung 3.3.1: : Ziel-DNA-Sequenzen aus dem humanen <i>CCR5</i> nach Mussolino <i>et al.</i> , 2012, mit den verwendeten TALEN-Grundgerüsten	85
Abbildung 3.3.2: Vergleich von <i>Sharkey</i> und Wildtyp (wt)-scTALEN und TALEN-Paaren .	86
Abbildung 3.3.3: TALEN mit einer D450A-Mutation in einer <i>Sharkey</i> -FokI- Endonukleasedomäne bleiben aktiv	89
Abbildung 3.3.4: Beide scFokI-Domänen einer Doppel-scTALEN (dscTALEN) können DNA-DSBs erzeugen	92
Abbildung 3.3.5: Die für Gene des Zebrafisches <i>Danio rerio</i> erstellten TALEN-Varianten erzeugen DSB in <i>in vitro</i> -Restriktionsanalysen	95
Abbildung 3.3.6: Weder mit scTALEN noch mit dscTALEN sind Mutationen in humanen Zellen nachweisbar.....	99
Abbildung 3.3.7: TALEN mit scFokI zeigen eine unvollständige Kernlokalisierung.....	101
Abbildung 3.3.8: Aufbau und Positionierung von TALEN-Paaren zur Unterscheidung des gesunden und des mutierten <i>FANCI</i> -Genes anhand eines Nukleotides	103
Abbildung 3.3.9: TALEN-Paare können zwei Sequenzen anhand einer Fehlpaarung an Position T ₀ unterscheiden.....	105
Abbildung 3.3.10: TALEN-Paare können zwei Sequenzen anhand einer Fehlpaarung an einer <i>repeat</i> -Position unterscheiden	107
Abbildung 4.1.1: Modell der DNA-Erkennung durch ein Standard- <i>repeat</i> -Array aus 34 AS <i>repeats</i> (A) ohne oder (B) mit aberrantem <i>repeat</i>	109
Abbildung 4.1.2: PthXo3 könnte den <i>OsSWEET14</i> -Promotor in <i>Oryza brachyantha</i> in anderer Konformation als in <i>O. sativa</i> erkennen	114
Abbildung 4.2: Verschiedene Strategien zur Genaktivierung mit künstlichen Aktivatoren.....	122

Tabellenverzeichnis

Tabelle 2.1:	Verwendete Plasmide.....	17
Tabelle 2.2:	Verwendete Oligonukleotide	20
Tabelle 2.3:	Eigenschaften der verwendeten Bakterienstämme	21
Tabelle 2.4:	Zusammensetzung des zur Bakterienanzucht verwendeten Nährmediums	23
Tabelle 2.5:	Für bakterielle Kulturen verwendete Antibiotika	23
Tabelle 2.6:	Zusammensetzung des für humane Zellen verwendeten Kulturmediums	24
Tabelle 2.7:	Verwendete Lösungen und Puffer	24
Tabelle 2.8:	Zusammensetzung der Transfektionsansätze für unterschiedliche Kulturdimensionen.....	32
Tabelle 2.9:	Zusammensetzung der SDS-Polyacrylamidgele.....	37
Tabelle 2.10:	Für Immunoblots verwendete Antikörper.....	37

Abkürzungsverzeichnis

A	Adenin
Abb.	Abbildung
AAD	<i>acidic activation domain</i> , saure Transkriptionsaktivierungsdomäne
AD	Aktivierungsdomäne
AIDS	<i>acquired immune deficiency syndrome</i>
AIM	<i>Agrobacterium infiltration medium</i>
Amp	Ampicillin
APS	Ammoniumpersulfat
ART	<i>antiretroviral rherapy</i>
ArtHIV	Artifizieller TAL-Effektor mit HIV-1-Genom-spezifischer Zielsequenz
Avr	Avirulenz
BSA	Bovines Serumalbumin
bp	<i>base pair</i> , Basenpaar
bzw.	beziehungsweise
C	Cytosin
CRISPR	<i>clustered regularly interspaced short palindromic repeats</i>
CRISPR/Cas	<i>CRISPR/CRISPR-associated protein</i>
CTR	C-terminale Region
D	Aspartat
DMEM	Dulbecco's Modified Eagle's Medium
DMF	Dimethylformamid
DMSO	Dimethylsulfoxid
DNA	<i>desoxyribonucleic acid</i> , Desoxyribonucleinsäure
ECL	<i>enhanced chemiluminescence</i> , Detektionsverfahren für Immunoblot
ECW	Kulturvarietät <i>Early Cal Wonder</i> von <i>Capsicum annuum</i>
EMSA	<i>enzyme mobility shift assay</i>
G	Glycin (Aminosäure)
G	Guanin (Nukleotid)
GFP	<i>green fluorescent protein</i>
GGC	<i>Golden Gate Cloning</i>
GTT	<i>Golden TAL Technologie</i>
H	Histidin

HDR	<i>homology directed repair</i>
HIV-1	Humanes Immundefizienz-Virus Typ 1
HR	Hypersensitive Reaktion
HTH	Helix- <i>turn</i> -Helix-Motiv
I	Isoleucin
i.d.R.	in der Regel
K	Lysin
Kan	Kanamycin
kDa	kilo-Dalton
LB	<i>lysogenic broth</i>
LRR	<i>leucin rich repeat</i> , Leucin-reiche Sequenzwiederholung
LTR	<i>long terminal repeat</i>
LV	Leervektor
M	molare Masse
min	Minute(n)
N	Asparagin
NHEJ	<i>non-homologous end joining</i>
NLS	<i>nuclear localisation signal</i> , Kernlokalisierungssignal
NTR	N-terminale Region
PAMP	<i>Pathogen associated molecular pattern</i>
PCR	<i>polymerase chain reaction</i> , Polymerase-Kettenreaktion
pv.	Pathovar
R, R	Resistenz
RBF	RVD-Base-Fehlpaarung
RT	Reverse Transkriptase
RVD	<i>repeat variable diresidue</i>
rpm	<i>revolutions per minute</i> , Umdrehungen pro Minute
S	Serin
SDS	Natriumdodecylsulfat
s	Sekunde
ssp.	Subspezies
T	Thymin
Tab.	Tabelle
TAE	Tris-Acetat-EDTA

TALE	<i>transcriptionactivator-like effector</i>
TALEN	TALE-Nuklease
TBST	Tris-buffered saline + Tween20
Tris	Tris-(hydroxymethyl)-aminomethan
U	<i>unit(s)</i> , Einheit(en)
u.a.	unter anderem/n
X-Gal	5-Brom-4-Chlor-3-indoxyl- β -D-galactosid
<i>Xac</i>	<i>Xanthomonas axonopodis</i> pv. <i>citri</i>
<i>Xcc</i>	<i>Xanthomonas campestris</i> pv. <i>campestris</i>
<i>Xcm</i>	<i>Xanthomonas campestris</i> pv. <i>malvacearum</i>
<i>Xcv</i>	<i>Xanthomonas campestris</i> pv. <i>vesicatoria</i>
<i>Xoc</i>	<i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>
<i>Xoo</i>	<i>Xanthomonas. oryzae</i> pv. <i>oryzae</i>
UPA	<i>upregulated by AvrBs3</i>
z.B.	zum Beispiel
ZFN	Zinkfinger-Nuklease
ZFP	Zinkfingerprotein
% (v/v)	<i>volume per volume</i> , Volumenprozent
% (w/v)	<i>weight per volume</i> , Gewichtsprozent

1. Einleitung

Als Erbinformation aller Lebewesen verschlüsselt die DNA die Bausteine der grundlegenden Lebensprozesse, sowie artspezifischer und individueller Merkmale. Diese Bausteine sind in als Gene bezeichneten DNA-Abschnitten codiert. Wie einzelne Gene und deren räumlich-zeitliche Regulation zur Gesamtheit eines Organismus beitragen, ist jedoch in vielen Fällen noch nicht geklärt. Die denkbar zuverlässigste und eleganteste Methode, um dies zu untersuchen, ist die direkte Manipulation des Gens von Interesse im Genom des Organismus. Diesen Ansatz ermöglicht das *Genome Engineering*. Dessen Grundlage bilden Proteine, die zur Erkennung einer spezifischen DNA-Sequenz programmiert werden können. Dadurch kann gezielt eine bestimmte Stelle in einem Genom angesteuert werden, um eine beliebige Aktion an der DNA durchzuführen. Die Modifikation solcher Werkzeuge, z.B. durch die Fusion funktionaler Domänen, erlaubt einen vielseitigen Einsatz. Neben der Zerstörung oder gezielten Veränderung eines Gens – dem *Genome Engineering* im engeren Sinne – sind unter anderem auch Genaktivierung oder epigenetische Veränderungen möglich. Die ersten Werkzeuge des *Genome Engineerings* waren Zinkfinger- und Meganukleasen. Die sequenzspezifische Programmierung dieser Genschere ist jedoch aufwendig, was eine weit verbreitete Anwendung verhinderte. Dies änderte sich mit der Entschlüsselung des TALE (*transcription activator-like effector*)-Codes. Der einfache, modulare Mechanismus zur DNA-Erkennung und Bindung von TALEs vereinfachte die Konstruktion von *Genome Engineering*-Werkzeugen drastisch, sodass TALE-Nukleasen (TALEN) und ihre Vorgänger zu Beginn dieser Arbeit von Nature Methods zur Methode des Jahres 2011 gewählt wurden (Baker, 2012). Seither hat das *Genome Engineering* eine rasante Entwicklung erlebt, zu der insbesondere die jüngeren CRISPR/Cas (*clustered regularly interspaced short palindromic repeats/CRISPR associated protein*)-basierten Werkzeuge beigetragen haben.

1.1 TAL-Effektoren

Transkriptionsaktivator-ähnliche Effektoren (*transcription activator-like effectors*), kurz TAL-Effektoren oder TALEs, wurden als Effektorproteine bakterieller Phytopathogene der Gattung *Xanthomonas* entdeckt (Bonas *et al.*, 1989). Sie werden vom Bakterium mit Hilfe des Typ-III-Sekretionssystems (T3S) in die Zellen der Wirtspflanze injiziert (Abb. 1.1 A; Van den Ackerveken *et al.*, 1996; Szurek *et al.*, 2002). Im Zytoplasma der Pflanzenzelle werden die TALEs von Importin α gebunden und in den Nukleus transportiert (Szurek *et al.*, 2001). Dort binden sie spezifisch an DNA-Sequenzen in Promotoren von Zielgenen, deren Transkription

die TALEs induzieren und so die Besiedlung der Wirtspflanze durch das Bakterium unterstützen (Abb. 1.1 A; Boch und Bonas, 2010).

1.1.1 Zielgene

TALEs sind in *Xanthomonas spp.* zu finden, die eine Vielzahl verschiedener ein- und zweikeimblättriger Pflanzen, darunter bedeutende Kulturpflanzen wie z. B. Reis, befallen. Um ihr Wirtsspektrum zu kennzeichnen, sind *Xanthomonas spp.* in Pathovare (pv.) eingeteilt. So besiedeln beispielsweise *X. oryzae* pv. *oryzae* (*Xoo*) und pv. *oryzicola* (*Xoc*) Reispflanzen (Gattung *Oryza*), auf denen sie die schwerwiegenden Krankheiten bakterielle Weißblättrigkeit bzw. bakterielle Streifenkrankheit auslösen (Nino-Liu *et al.*, 2006; Verdier *et al.*, 2012). Entsprechend ihrer weiten phylogenetischen Verbreitung sind mehr als 100 verschiedene TALEs bekannt, wobei die Anzahl der TALE-Gene in den verschiedenen *Xanthomonas*-Stämmen sehr unterschiedlich ist (Scholze und Boch, 2011). Besonders viele TALE-Gene (7-28) treten z.B. in *Xoo* und *Xoc* Stämmen auf (Scholze und Boch, 2011). Trotz dieser Vielzahl bekannter TALEs ist nur ein Teil ihrer pflanzlichen Zielgene identifiziert. Zielgene, deren Aktivierung zur Besiedlung und Erkrankung der Wirtspflanze durch Xanthomonaden beitragen, werden als Suszeptibilitätsgene bezeichnet. In vielen Fällen ist jedoch nicht aufgeklärt, ob und wie die Aktivierung identifizierter Zielgene die bakterielle Kolonisation der Pflanze beeinflussen. Eine mögliche Strategie stellt die Manipulation des pflanzlichen Transkriptoms durch die Aktivierung von Transkriptionsfaktoren dar (Boch *et al.*, 2014). Beispiele für TALE-aktivierte Transkriptionsfaktoren sind das durch PthA aus *X. citri* bei der Auslösung von Zitruskrebs induzierte Suszeptibilitätsgen *CsLOB* (Hu *et al.*, 2014; Li *et al.*, 2014; Pereira *et al.*, 2014), das durch AvrBs3 aus *X. campestris* pv. *vesicatoria* (*Xcv*) induzierte, in Paprika Hypertrophie-verursachende *UPA20* (Kay *et al.*, 2007), sowie die durch PthXo6 und PthXo7 aus *Xoo* induzierten Gene *OsTFXI* bzw. *OsTFIIAγ1* (Sugio *et al.*, 2007).

Eine andere Strategie stellt die TALE-vermittelte Induktion von Transportergenen dar (Boch *et al.*, 2014). Dies könnte zur Anpassung des Apoplastenmilieus an die Bedürfnisse des Pathogens beitragen, indem das Bakterienwachstum hemmende Stoffe ins Pflanzenzellinnere und förderliche Stoffe in den Apoplasten transportiert werden (Boch *et al.*, 2014). Beispielsweise trägt Talg2 durch die Aktivierung von *OsSULTR3;6* zur Besiedlung von Reispflanzen durch *Xoc* bei (Cernadas *et al.* 2014). Dabei codiert *OsSULTR3;6* einen Sulfattransporter, der mutmaßlich Sulfatimport in die Pflanzenzelle vermittelt (Cernadas *et al.* 2014). Auch die prominenteste Gruppe TALE-induzierter Suszeptibilitätsgene, die Familie der *SWEET*-Gene in Reis, gehört zu den Transportergenen. Die Aktivierung dieser

Zuckertransporter kodierenden Gene trägt wahrscheinlich zur Bereitstellung von Saccharose und Glukose im Apoplasten und Xylem bei und ist essentiell für die Virulenz von *Xoo*. Gleich mehrere TALEs aus verschiedenen *Xoo*-Stämmen induzieren drei Gene dieser Familie. Während *OsSWEET11* und *OsSWEET13* von PthXo1 bzw. PthXo2 aktiviert werden, wird *OsSWEET14* von vier verschiedenen TALEs, TalC, AvrXa7, PthXo3, sowie TAL5, angesteuert (Yang *et al.* 2006; Antony *et al.* 2010; Römer *et al.*, 2010; Liu *et al.*, 2011; Yu *et al.* 2011; Zhou *et al.* 2015). Darüber hinaus wurde für die übrigen *OsSWEETs* aus dieser Verwandtschaftsgruppe (Klade III), *OsSWEET12* und *OsSWEET15*, gezeigt, dass auch sie bei Induktion durch künstliche TALEs zur Besiedlung von Reis durch *Xoo* beitragen (Li, T. *et al.*, 2013; Streubel *et al.* 2013).

1.1.2 TALE-basierte Pflanzenresistenzmechanismen

Wirtspflanzen haben verschiedene Resistenzbarrieren gegen Phytopathogene entwickelt, auch basierend auf Effektorproteinen. In Bezug auf die größte Effektorfamilie der TALEs sind mehrere Resistenzmechanismen bekannt. So stellen Gene, die einen lokalen Zelltod als Resistenzreaktion auslösen, einen weiteren Typ von TALE-induzierten Genen dar. Zu diesen Resistenzgenen gehören *Bs3* und *Bs4C* aus Paprika, die durch die *Xcv*-TALEs AvrBs3 bzw. AvrBs4 induziert werden (Römer *et al.*, 2007; Strauß *et al.*, 2012), sowie *Xa10* und *Xa27* aus Reis, die von den *Xoo*-TALEs AvrXa10 bzw. AvrXa27 aktiviert werden (Gu *et al.*, 2005; Tian *et al.*, 2014). Neben der transkriptionellen Aktivierung von Resistenzgenen durch TALEs kann eine Resistenzreaktion durch die Erkennung der Effektoren auf Proteinebene ausgelöst werden. Dies ist im Zusammenhang mit TALEs für das Paprika-Resistenzprotein Bs4, das u.a. AvrBs4 aus *Xcv* erkennt, beschrieben (Schornack *et al.*, 2004). Andere TALE-basierte Pflanzenresistenzmechanismen, die nicht auf der Auslösung einer Resistenzreaktion beruhen, werden durch Mutationen vermittelt, die die Aktivierung von Suszeptibilitätsgenen verhindern. Ein mögliches Beispiel ist eine Mutation in der γ -Untereinheit des generellen Transkriptionsfaktors (TF) IIA. In Reis existieren zwei Paraloge des TFIIA γ -Gens, *OsTFIIA γ 1* und *OsTFIIA γ 5* (*Xa5*), von denen *Xa5* konstitutiv exprimiert wird. Das rezessive Allel *xa5* verringert die Virulenz von *Xoo* (Iyer and McCouch, 2004; Iyer-Pascuzzi *et al.*, 2008; Gu *et al.*, 2009), was durch die PthXo7-vermittelte Induktion von *OsTFIIA γ 1* z.T. aufgehoben wird (Sugio *et al.*, 2007). Eine andere Möglichkeit stellt die Mutation der DNA-Zielsequenzen von TALEs in Zielgenpromotoren dar (Chu *et al.*, 2006a, b; Yang *et al.*, 2006; Liu *et al.*, 2011). Dadurch wird die TALE-abhängige Geninduktion und ein entsprechender Beitrag zur Virulenz des Pathogens verhindert. Betrifft dies wichtige Suszeptibilitätsgene wie die *OsSWEETs*,

vermitteln solche biallelischen Promotormutationen Resistenz der Pflanze gegen das Pathogen. Dementsprechend sind Reisvarianten, die Insertionen, Deletionen oder Nukleotidsubstitutionen in den Promotoren von *OsSWEET11* und *OsSWEET13* tragen, resistent gegenüber *Xoo*-Stämmen mit den TALEs PthXo1 bzw. PthXo2 (Chu *et al.*, 2006a, b; Antony *et al.*, 2010; Liu *et al.*, 2011; Yuan *et al.*, 2011; Zhou *et al.* 2015).

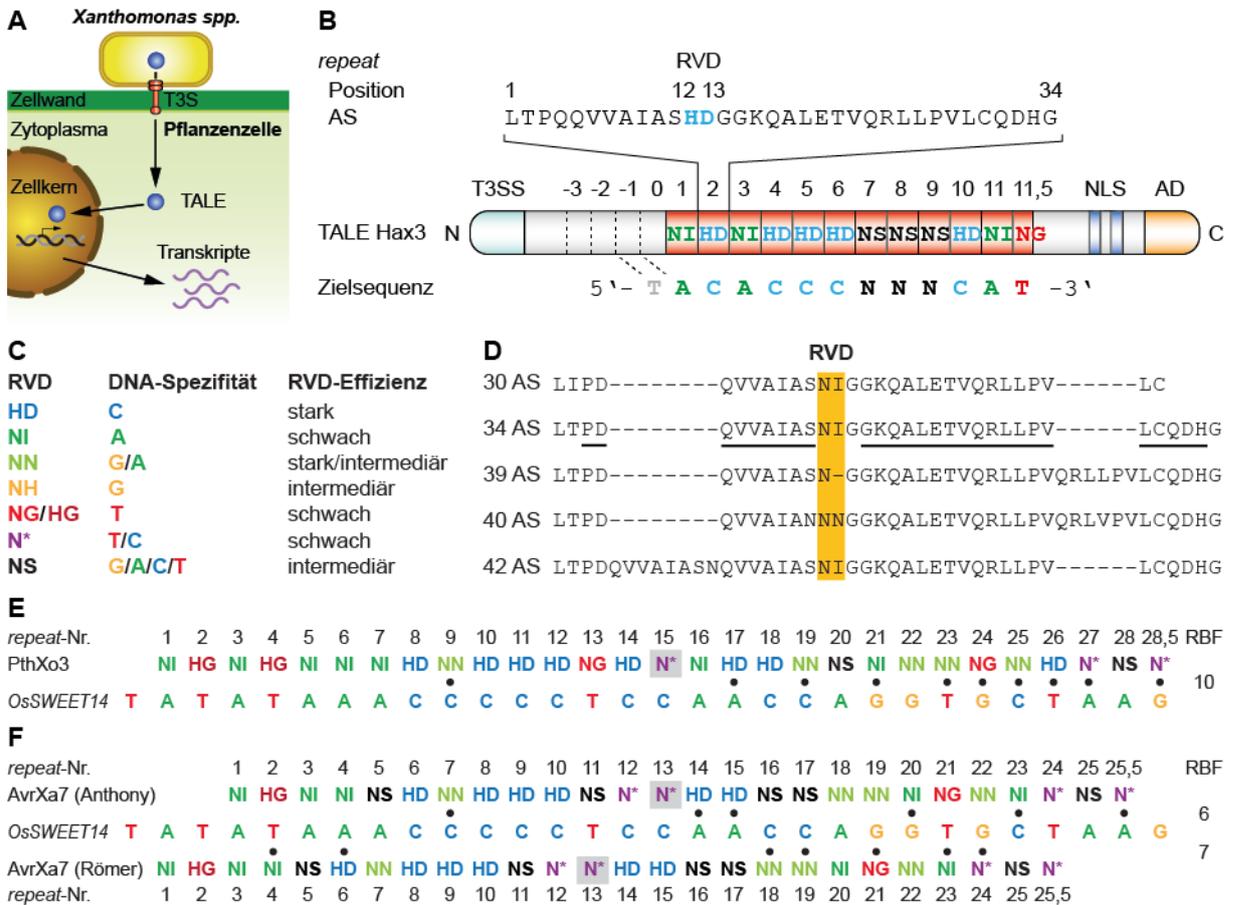


Abb. 1.1: Funktionsweise und struktureller Aufbau natürlicher TALEs. (A) TALEs werden über das Typ-III-Sekretionssystem (T3S) in die Pflanzenzelle übertragen und aktivieren dort im Zellkern Genexpression. Abb. verändert nach Streubel *et al.*, 2013. (B) Aufbau und DNA-Erkennung des natürlichen TALEs Hax3. TALEs enthalten ein T3S-Signal (T3SS, türkis), nukleare Lokalisierungssignale (NLS, blau), eine saure Aktivierungsdomäne (AD, orange) und eine DNA-Bindedomäne aus hoch konservierten Aminosäuren (AS)-Wiederholungen (*repeats*, rot). Jedes *repeat* erkennt ein Nukleotid der Ziel-DNA-Sequenz. Dabei bestimmen die AS an Position 12 und 13, genannt RVD (*repeat variable diresidue*), die Basenspezifität. Der degenerierte *repeat* -1 im TALE N-Terminus erkennt zusätzlich ein 5'-vorgelagertes Thymin. Abb. verändert nach Richter *et al.*, 2016. (C) RVD-Spezifitäten und Effizienzen. Ein * kennzeichnet das Fehlen der AS an dieser Position. (D) Alignment der AS-Sequenzen natürlicher TALE-*repeats* unterschiedlicher AS-Anzahl. Die AS, die die beiden α -Helices formen, sind in einem typischen 34 AS *repeat* unterstrichen. (E) Alignment der PthXo3 RVDs mit der *OsSWEET14*-Zielsequenz nach Anthony *et al.*, 2010. Es treten mehr RVD-Base-Fehlpaarungen (RBF, schwarze Punkte) im 3'-Bereich der Zielsequenz auf. (F) Alignment der AvrXa7 RVDs mit der *OsSWEET14*-Zielsequenz in zwei um zwei Nukleotide verschobenen Varianten nach Anthony *et al.*, 2010, oder Römer *et al.*, 2010. (E, F) 39 AS *repeat* in grau. (D-F) Abb. verändert nach Richter *et al.*, 2014.

1.1.3 Der strukturelle Aufbau und die DNA-Erkennung von TALEs

Um ihre Funktion als Transkriptionsaktivatoren in Pflanzen zu erfüllen, besitzen TALEs bestimmte strukturelle Merkmale. Neben einem N-terminalen Typ-III-Sekretionssignal (Abb. 1.1 B, T3SS; Szurek *et al.*, 2002) zur Translokation aus der Bakterien- in die Pflanzenzelle verfügen sie über Domänen eukaryotischer Transkriptionsaktivatoren. Dabei handelt es sich um zwei funktionale Kernlokalisierungssignale (nukleare Lokalisierungssignale, NLS) und eine saure Aktivierungsdomäne (AD) im C-Terminus (Abb. 1.1 B; Van den Ackerveken *et al.*, 1996; Zhu *et al.*, 1998; Szurek *et al.*, 2001; Marois *et al.*, 2002).

Entscheidend für die spezifische Erkennung von Zielgenpromotoren ist jedoch die einzigartige, zentrale DNA-Bindedomäne von TALEs. Dabei handelt es sich um ein Array von hoch konservierten Aminosäuresequenz-Wiederholungen, den sogenannten *repeats* (Abb. 1.1 B; Boch und Bonas, 2010). Die i. d. R. 34 Aminosäuren eines jeden *repeats* bilden eine Helix-Loop-Helix-Struktur, wodurch die variablen Aminosäuren an *repeat*-Position 12 und 13, genannt *repeat variable diresidue* (Abb. 1.1 B, RVD), in der großen Furche der DNA-Doppelhelix exponiert werden (Deng *et al.*, 2012; Mak *et al.*, 2012; Gao *et al.*, 2012). Dort vermittelt der RVD jeweils eines *repeats* die Erkennung jeweils eines Nukleotids des DNA-Leitstranges in konsekutiver Folge. Somit bestimmt die Anzahl der *repeats* auch die Länge der vom TALE erkannten DNA-Sequenz (Boch *et al.*, 2009; Moscou und Bogdanove, 2009). Der C-terminal letzte, sogenannte halbe *repeat*, der nur 19 oder 20 der 34 konservierten Aminosäuren enthält, begrenzt dabei die DNA-Zielsequenz am 3'-Ende (Abb. 1.1 B). Zusätzlich zu der vom *repeat*-Array erkannten DNA-Sequenz gehört zur Zielsequenz von TALEs ein unmittelbar 5'-vorgelagertes Thymin, das als initiales T (T₀) bezeichnet wird (Abb. 1.1 B; Boch *et al.*, 2009; Moscou und Bogdanove, 2009). TALE-Kristallstrukturen haben bis zu vier degenerierte *repeats* (-3 bis 0) im N-Terminus gezeigt, wobei ein Tryptophan in *repeat* -1 für die Koordination des initialen T notwendig ist (Abb. 1.1 B; Deng *et al.*, 2012; Mak *et al.*, 2012; Gao *et al.*, 2012; Stella *et al.*, 2013; Lamb *et al.*, 2013; Tsuji *et al.*, 2013; Schreiber *et al.*, 2014). Ähnlich der Spezifizierung des initialen T durch das Tryptophan im degenerierten *repeat* -1 interagiert auch die Aminosäure an Position 13 eines jeden kanonischen *repeats* mit einer spezifischen Nukleobase (Deng *et al.*, 2012; Mak *et al.*, 2012). Die Aminosäure an Position 12 hingegen vermittelt intra-*repeat*-Interaktionen, die zur strukturellen Stabilisierung des *repeats* und der Positionierung von Aminosäure Nr. 13 zur DNA-Base beitragen (Deng *et al.*, 2012; Mak *et al.*, 2012). Die so vermittelte DNA-Basenspezifität der RVDs wurde bereits entschlüsselt, wobei einige RVDs die Erkennung mehrerer Nukleobasen erlauben (Abb. 1.1 C;

Boch *et al.*, 2009; Moscou und Bogdanove, 2009; Streubel *et al.*, 2012; Le Cong *et al.*, 2012). So erkennen z. B. die häufigen RVDs HD, NI und NG spezifisch die DNA-Basen C (Cytosin), A (Adenin) bzw. T (Thymin), wohingegen NN und NS mit G (Guanin) und A bzw. A, C, G und T jeweils mehrere Basen erkennen (Abb. 1.1 C; Boch *et al.*, 2009; Moscou und Bogdanove, 2009). Darüber hinaus wurde festgestellt, dass die verschiedenen RVDs unterschiedlich stark zur Aktivität von TALEs beitragen (Abb. 1.1 C; Streubel *et al.*, 2012). Dabei wurde einem RVD eine starke Effizienz in Kombination mit einer bestimmten Base zugeordnet, wenn ein TALE, der fast ausschließlich aus *repeats* mit diesem RVD bestand, Aktivität an einer aus dem entsprechenden Nukleotid bestehenden DNA-Sequenz zeigte. Als starke RVDs wurden nur HD und NN an G identifiziert. Eine intermediäre Effizienz wurde zugeordnet, wenn ein TALE mit einer Folge des entsprechenden RVDs noch Genaktivierung vermittelte. Dies traf u. a. auf den G-spezifischen RVD NH aus einem TALE-Homolog von *Ralstonia solanacearum* zu. Als schwache RVDs wurden solche bezeichnet, die nur in einem *repeat*-Array mit starken oder zumindest intermediären RVDs TALE-Aktivität vermitteln konnten, was z. B. für NI und NG der Fall war.

Zusätzlich zu T_0 legt also die RVD-Abfolge die Zielsequenz eines TALEs fest (Abb. 1.1 B). Das kann genutzt werden, um durch die computergestützte Vorhersage von Zielsequenzen in Promotorregionen neue Zielgene von TALEs zu identifizieren (Doyle *et al.*, 2012; Grau *et al.*, 2013; Pérez-Quintero *et al.*, 2013). Abgesehen von der Erkennung mehrerer Basen durch einige RVDs binden und aktivieren natürliche TALEs jedoch oft an Zielsequenzen die Fehlpaarungen enthalten (Kay *et al.*, 2007; Römer *et al.*, 2007 und 2010; Boch *et al.*, 2009; Antony *et al.*, 2010; Moscou und Bogdanove, 2009; Strauß *et al.*, 2012), was Vorhersagen erschwert. Diese Fehlpaarungstoleranz hängt von Länge und Zusammensetzung des *repeat*-Arrays ab (Boch *et al.*, 2009; Mali *et al.*, 2013 b). Längere *repeat*-Arrays tolerieren dabei i. d. R. mehr Fehlpaarungen als Kürzere.

1.1.4 Die Variabilität der *repeats*

Die höchste Variabilität in einem *repeat* zeigen die beiden Aminosäuren des RVD an Position 12 und 13 (Boch *et al.*, 2009; Moscou und Bogdanove, 2009). Außerhalb des RVDs variieren nur wenige Aminosäuren. Dabei handelt es sich vor allem um die Aminosäuren an *repeat*-Position 4 und 24 (Streubel *et al.*, 2012), deren Einfluss auf die Effizienz eines *repeats* nicht näher bekannt ist.

Neben Aminosäuerevariationen treten *repeats* auf, deren Länge von den üblichen 34 Aminosäuren abweicht. Zwei relativ häufige Ausnahmen vom 34 Aminosäuren-*repeat* sind 33 und 35 Aminosäuren. 33 Aminosäuren-*repeats* fehlt die Aminosäure an Position 13 des RVD (RVD: N*; Abb. 1.1 C), wobei in 35 Aminosäuren-*repeats* ein zusätzliches Prolin auf die Aminosäure an Position 32 folgt. Diese Veränderungen des *repeats* haben jedoch keinen Einfluss auf die DNA-Bindesepezifität von TALEs (Boch *et al.*, 2009; Streubel *et al.*, 2012; Cong *et al.*, 2012). Außer in der Gattung *Xanthomonas* treten auch TALE-Homologe mit ausschließlich 35 Aminosäuren-*repeats* in *Ralstonia solanacearum* auf. Auch diese verfügen über eine RVD-basierte DNA-Bindeaktivität (de Lange *et al.*, 2013; Li, L. *et al.*, 2013). Allerdings gibt es auch verschiedene *repeats*, die mehr als eine Aminosäure vom üblichen 34 Aminosäuren-*repeat* aus *Xanthomonas spp.* abweichen. Bei diesen aberranten *repeats* handelt es sich um 30, 39 oder 40 (abhängig vom RVD), und 42 Aminosäuren-*repeats*, wobei die Längenunterschiede durch Deletion oder Duplikation von Aminosäuren in der ersten oder zweiten *repeat*-Helix zustande kommen (Abb. 1.1 D; Boch und Bonas, 2010). Sie treten nur einzeln eingestreut in regulären 33-35 Aminosäuren-*repeat*-Arrays auf (Yang *et al.*, 2000; Boch und Bonas, 2010). Obwohl zu erwarten wäre, dass ein aberranter *repeat* die regelmäßige Struktur des *repeat*-Arrays stören sollte, enthalten die essentiellen *Xoo*-Virulenzfaktoren PthXo3 und AvrXa7 beide einen 39 Aminosäuren-*repeat*, sind funktional und in der Lage *OsSWEET14* zu induzieren (Yang *et al.* 2000; Yang and White, 2004; Antony *et al.* 2010; Römer *et al.*, 2010). Alignments der beiden TALEs mit den Zielsequenzen zeigen jedoch viele Fehlpaarungen zwischen den RVDs und der DNA (Abb. 1.1 E, F; Antony *et al.* 2010; Römer *et al.*, 2010). Fast alle Fehlpaarungen von PthXo3 mit der *OsSWEET14*-Zielsequenz befinden sich dabei im Bereich C-terminal des aberranten *repeats* (Abb. 1.1 E; Antony *et al.* 2010). Im Falle von AvrXa7 sind hingegen zwei zueinander verschobene Zielsequenzen im *OsSWEET14*-Promotor publiziert (Abb. 1.1 F; Antony *et al.* 2010; Römer *et al.*, 2010). Ob der 39 Aminosäuren-*repeat* im Falle von PthXo3 und AvrXa7 oder aberrante *repeats* im Allgemeinen das DNA-Bindevverhalten von TALEs beeinflussen, war zu Beginn dieser Arbeit jedoch noch nicht geklärt.

1.2 Die Anwendung von TALEs

Auf Grund der modularen DNA-Bindedomäne, deren Spezifität einem einfachen Code folgt, sind TALEs sehr gut zur Erstellung molekularer Werkzeuge geeignet. Durch entsprechende Modifikation der Proteine ist es möglich jede denkbare Aktion an oder mit der gebundenen DNA durchzuführen. Dafür ist es notwendig die *repeat*-Domäne für die gewünschte DNA-

Zielsequenz zu programmieren. Um das zu realisieren, wurden verschiedene Protokolle zur Assemblierung von Designer *repeat*-Arrays entwickelt (Geißler *et al.*, 2011; Zhang *et al.*, 2011; Reyon *et al.*, 2012; Schmid-Burgk *et al.*, 2013). Insbesondere die Methode der „Golden Gate“-Klonierung wurde auf TALEs angewendet (Engler *et al.*, 2008; Geißler *et al.*, 2011; Zhang *et al.*, 2011; Cermak *et al.*, 2011; Morbitzer *et al.*, 2011). Dabei werden Typ IIS-Restriktionsenzyme eingesetzt, die ein Sequenz-unabhängiges, schnittstellenfreies Klonieren ermöglichen (Engler *et al.*, 2008). Dadurch können z. B. bei der „Golden TAL Technology“-*repeat*-Module mit dem gewünschten RVD anhand der gewählten Überhänge gezielt im Array positioniert werden (Geißler *et al.*, 2011). Entsprechend der natürlichen Funktion von TALEs war zunächst eine Anwendung als transkriptionelle Aktivatoren naheliegend (Morbitzer *et al.*, 2010; Miller *et al.*, 2011; Geißler *et al.*, 2011; Zhang *et al.*, 2011). Andere bekannte Anwendungen von TALEs sind TALE-Repressoren, TALEN und TALE-Rekombinasen (Miller *et al.*, 2011; Mahfouz *et al.*, 2012; Blount *et al.*, 2012; Garg *et al.*, 2012; Mercer *et al.*, 2012). Solche TALE-basierten Werkzeuge, vor allem aber TALEN, wurden in einer Vielzahl von Organismen und Modellsystemen verwendet, neben humaner Zellkultur auch in Hefe, Nematoden, Insekten, Fischen, Mäusen, Schweinen, Rindern und natürlich in Pflanzen (Cermak *et al.*, 2011; Wood *et al.*, 2011; Huang *et al.*, 2011; Ma *et al.*, 2012; Carlson *et al.*, 2012; Wang *et al.*, 2013; Christian *et al.*, 2013).

1.2.1 TALE-Aktivatoren

1.2.1.1 Die Transkription

Um die Transkription proteincodierender Gene in Eukaryoten zu initiieren, muss die RNA-Polymerase II (Pol II) an den Promotor des Gens rekrutiert werden (Alberts *et al.*, 2004). Dies geschieht im Rahmen des Präinitiationskomplexes (PIC), der sich aus der Pol II und den generellen Transkriptionsfaktoren (GTFs) zusammensetzt, die in folgender Reihenfolge binden: TFIID, TFIIA, TFIIB, Pol II, TFIIF, TFIIE und TFIIH (Smale und Kadonaga, 2003; Alberts *et al.*, 2004; Kornberg, 2007). Die Bindung von TFIID an die DNA sorgt dabei für eine starke Verformung dieser, was die weitere Proteinanlagerung ermöglicht. Der TFIIH verfügt sowohl über eine Helikase-Untereinheit, die die DNA entwindet, als auch über eine Proteinkinase-Untereinheit, die die C-terminale Domäne der Pol II phosphoryliert und damit für deren Freisetzung aus dem PIC sorgt (Alberts *et al.*, 2004). Damit ist die Transkriptionsinitiation beendet, die GTFs werden wieder frei und die Pol II geht in die Elongation der mRNA über. Der Ort der Transkriptionsinitiation wird dabei durch *cis*-Elemente bestimmt, die direkt von GTFs gebunden werden, wie z. B. der TATA-Box und dem Initiator

(INR), die vom TATA-Bindeprotein (TBP) bzw. einer anderen TFIID-Untereinheit erkannt werden (Smale und Kadonaga, 2003; Alberts *et al.*, 2004). Darüber hinaus sind Transkriptionsaktivatoren, ein Mediator genannter Proteinkomplex und Chromatin-modifizierende Enzyme für die Geninduktion notwendig (Alberts *et al.*, 2004). Transkriptionsaktivatoren sind verschiedene Transkriptionsfaktoren die an Enhancer genannte *cis*-Elemente, die bisweilen mehrere 1000 Basenpaare (bp) vom Gen entfernt sein können, binden und die Anlagerung des PIC an den Promotor fördern. Die Präsenz oder Abwesenheit einzelner Transkriptionsaktivatoren in der Umgebung eines Gens trägt somit zu dessen Regulation bei. Die Kommunikation der Transkriptionsaktivatoren mit dem PIC wird dabei durch den Mediator vermittelt (Alberts *et al.*, 2004; Krishnamurthy und Hampsey, 2009). Dieser Multiproteinkomplex interagiert auch mit Chromatin-modifizierenden Enzymen, die die Zugänglichkeit der Nukleosomen verpackten DNA für die Transkription ermöglichen (Alberts *et al.*, 2004).

1.2.1.2 Die transkriptionelle Aktivierung durch TALEs

Da TALEs bereits von Natur aus als Transkriptionsaktivatoren in Pflanzen fungieren, reicht es aus ihre DNA-Spezifität durch Modifikation der *repeat*-Region anzupassen, um ein beliebiges Pflanzengen zu induzieren (Morbitzer *et al.*, 2010). Dabei sind eine funktionale NLS und die saure AD von TALEs für die Geninduktion in Pflanze notwendig (Zhu *et al.*, 1998; Szurek *et al.*, 2001; Marois *et al.*, 2002). Der Funktionsverlust eines TALEs durch Deletion der eigenen AD wird jedoch teilweise durch Ersatz mit der sauren VP16-AD kompensiert (Zhu *et al.*, 1999; Szurek *et al.*, 2001; Geißler *et al.*, 2011). Daher ist anzunehmen, dass die TALE-AD ähnlich anderen sauren ADs wie der VP16-AD mit Proteinen, die an der Transkription beteiligt sind, interagiert. Kandidaten sind GTFs, sowie Bestandteile des Mediators und des Chromatin-Rekonstruktionskomplexes (Triezenberg, 1995; Hall und Struhl, 2002; Carpenter *et al.*, 2005; Hirai *et al.*, 2010). Ein Beispiel für einen GTF, der mit sauren ADs interagiert, ist TFIIA (Ozer *et al.*, 1994; Kobayashi *et al.*, 1995; Stargell und Struhl, 1995). Kürzlich wurde außerdem eine Interaktion des C-terminalen Bereiches hinter den *repeats* von *Xoo*-TALEs mit TFIIA γ 5 aus Reis modelliert und nachgewiesen (Dehury *et al.*, 2015; Yuan *et al.*, 2016). Darüber hinaus verschieben einige natürliche TALEs den Transkriptionsstart ihrer Zielgene auf 40-60 bp stromabwärts ihrer DNA-Bindestelle (Kay *et al.*, 2007; Römer *et al.*, 2009a und b; Antony *et al.*, 2010). Diese Erkenntnisse sprechen für eine TALE-vermittelte Geninduktion in Pflanze durch direkte Interaktion mit dem PIC. Für die Anwendung von TALEs als Genaktivatoren in humaner Zellkultur war die natürliche TALE-AD allerdings nicht effizient (Geißler *et al.*,

2011). Eine effektive Aktivierung war jedoch durch den Austausch der TALE-AD gegen eine VP16- oder VP64-AD möglich (Miller *et al.*, 2011; Geißler *et al.*, 2011; Zhang *et al.*, 2011).

1.2.2 TALEN

Um TALEs als natürliche Transkriptionsaktivatoren in Nukleasen zur Genommodifikation in Eukaryotenzellen umzuwandeln, ist die Fusion einer Endonukleasedomäne notwendig. Wie zuvor bei der Konstruktion von Zinkfinger-nukleasen (ZFN; Kim *et al.*, 1996; Smith *et al.*, 1999; Urnov *et al.*, 2005) wurde zur Erstellung von TALEN die Endonukleasedomäne von FokI verwendet (Miller *et al.*, 2011; Li *et al.*, 2011; Mahfouz *et al.*, 2011; Mussolino *et al.*, 2011). Die Typ IIS Endonuklease FokI erzeugt einen vier Nukleotide langen 5'-Überhang (Sugisaki und Kanazawa, 1981) und hat den Vorteil, dass sich die Endonukleasedomäne von der sequenzspezifischen DNA-Bindedomäne isolieren lässt (Li *et al.*, 1992 und 1993; Kim *et al.*, 1994). Durch Fusion der FokI-Endonukleasedomäne ohne Sequenzspezifität mit einem programmierbaren DNA-Bindeprotein, wie einem Zinkfingerprotein (ZFP) oder TALE, ist es daher also möglich beliebige DNA-Sequenzen zu schneiden (Miller *et al.*, 2011; Mussolino *et al.*, 2011). Im Falle von TALEs wurden dabei verschiedene Verkürzungen von N- und C-Terminus vorgenommen, um die DNA-Bindedomäne so kompakt wie möglich zu gestalten und nicht zweckdienliche Domänen wie T3SS und AD zu entfernen. Da die FokI-Endonukleasedomäne zum schneiden der DNA dimerisieren muss (Bitinaite *et al.*, 1998), werden TALEN darüber hinaus als Paare eingesetzt (Abb. 1.2; Miller *et al.*, 2011; Mahfouz *et al.*, 2011; Mussolino *et al.*, 2011). Es wurde festgestellt, dass die Distanz des von der FokI-Endonukleasedomäne erzeugten Doppelstrangbruchs (DSB) zu der von den TALE-*repeats* erkannten Zielsequenz von der Länge des mit der Endonukleasedomäne fusionierten TALE N- oder C-Terminus abhängt (Miller *et al.*, 2011; Li *et al.*, 2011; Mahfouz *et al.*, 2011; Mussolino *et al.*, 2011). Dementsprechend muss auch der Abstand der beiden Zielsequenzen der TALEN eines Paares, der sogenannte *spacer*, der Länge des FokI-fusionierten TALE-Terminus angepasst werden (Miller *et al.*, 2011; Mahfouz *et al.*, 2011; Mussolino *et al.*, 2011). Häufig wird z.B. ein TALEN-Aufbau mit 136 N- und 63 C-terminalen AS und C-terminal fusionierter FokI-Endonukleasedomäne gewählt (Abb. 1.2; Miller *et al.*, 2011). Dabei ist ein *spacer* von 14-16 bp optimal. Epitope und eine NLS werden bei diesem Aufbau N-terminal fusioniert. Die eigentliche Genommodifikation wird jedoch nicht durch die TALEN eingeführt, die lediglich einen gezielten DSB verursachen, sondern durch die Reparatur dieses DSB durch die eukaryotische Zelle. Die Reparatur eines DSB kann dabei durch zwei alternative Mechanismen zusammengefasst werden, die Verknüpfung nicht-homologer Enden (*non-homologous end-*

RNA-Komplex so sein Ziel in der DNA gefunden, erfolgt durch die voneinander unabhängigen, HNH- und RuvC-ähnlichen Endonukleasedomänen ein DNA-Doppelstrangbruch mit glatten Enden. Im Rahmen biotechnologischer Anwendungen wurde CRISPR/Cas gemäß seiner Natur zuerst als Nuklease im *Genome Editing* eingesetzt (Le Cong *et al.*, 2013; Mali *et al.*, 2013 a). Dabei wurden crRNA-tracrRNA-Chimären, genannt *single guide* oder nur *guide* RNA (sgRNA/gRNA), und das bis heute meistverwendete Cas9 aus *Streptococcus pyogenes* (SpCas9) genutzt (Abb. 1.3). Geeignete SpCas9-Zielsequenzen umfassen 20 bp und müssen unmittelbar stromaufwärts einer NGG- oder mindestens NRG-Sequenz als PAM liegen (Mali *et al.*, 2013 b). Für die Anwendung dieses programmierbaren, DNA-bindenden Molekülkomplexes als Genaktivator oder -repressor wurden Nuklease-negative Cas9-Derivate durch Mutation der beiden Endonukleasedomänen erstellt und NLS, sowie entsprechende Aktivator- oder Repressordomänen wie VP64, bzw. KRAB fusioniert (Gilbert *et al.*, 2013). Eine Alternative zur Fusion funktionaler Domänen an das Cas9-Derivat stellt die Fusion an ein Protein dar, das ein entsprechendes, in die sgRNA-integriertes Aptamer binden kann und so zum Wirkort des künstlichen Genregulators rekrutiert wird (Mali *et al.*, 2013 b).

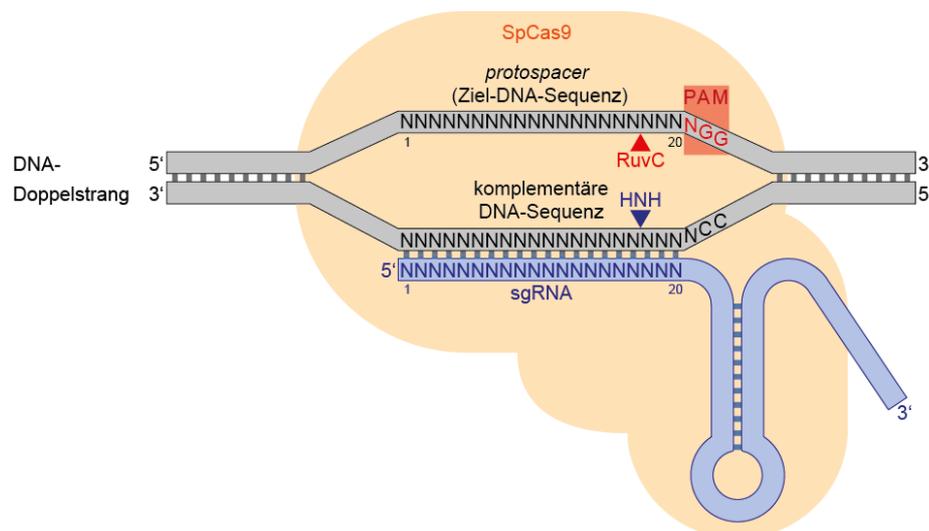


Abb. 1.3: Schematische Darstellung der Bindung von *Streptococcus pyogenes* Cas9 (SpCas9) an eine Ziel-DNA-Sequenz mittels *single guide* RNA (sgRNA). SpCas9 (orange) entwindet den DNA-Doppelstrang (grau), sodass das *protospacer adjacent motif* (PAM, rot) vom Protein und die zum *protospacer* (Ziel-DNA-Sequenz, N₁₋₂₀, schwarz) komplementäre DNA-Sequenz von der sgRNA (N₁₋₂₀, blau) gebunden werden kann. Die HNH-ähnliche Endonukleasedomäne (blau) schneidet den komplementären und die RuvC-ähnliche Endonukleasedomäne (rot) den Ziel-DNA-Strang. Die Schnittstellen sind durch farbige Dreiecke gekennzeichnet.

1.4 Das humane Immundefizienz-Virus

Bei dem humanen Immundefizienz-Virus Typ 1 (HIV-1) handelt es sich um eines der erfolgreichsten humanpathogenen Viren der jüngeren Zeit, das zur Ausprägung des letal verlaufenden Erworbenen Immundefizienzsyndroms (*acquired immune deficiency syndrome*, AIDS) führt (Kayser *et al.*, 2010). HIV-1 ist ein komplexes Retrovirus aus der Gattung der Lentiviren, das Zellen des Immunsystems, wie Makrophagen und CD4-positive (CD4⁺) T-Zellen, befällt (Flint *et al.*, 2009; Kayser *et al.*, 2010). Das HIV-Genom ist wie bei allen Retroviren eine (+)-Strang-RNA, die nach Eintritt in die Wirtszelle durch die virale reverse Transkriptase in eine DNA-Kopie umgeschrieben wird. Diese wird in den Zellkern transportiert und dort von einem weiteren, viralen Enzym, der Integrase, in das Wirtszellgenom integriert. So entsteht das Provirus, welches von Sequenzwiederholungen, den sogenannten *long terminal repeats* (LTR), flankiert ist. Das Provirus dient der zellulären Transkriptionsmaschinerie als Vorlage für die viralen Transkripte, die sowohl mRNAs für die Synthese der Virusproteine als auch neue Virusgenome darstellen. Dabei dient der 5'-LTR als Promotor. An der Wirtszellmembran werden Virusgenome und -Proteine dann zu neuen Viruspartikeln assembliert. Reverse Transkription, Provirus-Integration, sowie Partikelassemblierung sind typische Angriffspunkte der antiretroviralen Therapie (ART), durch die eine Reduktion der Viruslast im Blutplasma auf ein mit den üblichen diagnostischen Mitteln nicht detektierbares Level erreicht werden kann (Kayser *et al.*, 2010). Dennoch führt ART nicht zur Auslöschung des Virus im Organismus, sodass bei Unterbrechung ein neuer Virusausbruch erfolgen kann (Finzi *et al.*, 1997; Wong *et al.*, 1997).

Eine mögliche Strategie zur Heilung von HIV-1 besteht darin eine natürliche Resistenz gegen HIV-1 nachzuahmen. CCR5 ist ein HIV-1 Co-Rezeptor, der dem Virus zum Zelleintritt dient. Personen, die das *CCR5*Δ32-Allel homozygot tragen, sind daher weitestgehend gegen HIV-1 resistent (Liu *et al.*, 1996). Auch eines der ersten mit TALEN ausgeschalteten Gene war *CCR5* (Miller *et al.*, 2011; Mussolino *et al.*, 2011), das zuvor bereits als Zielgen für die Mutation durch ZFNs diente (Perez *et al.*, 2008; Holt *et al.*, 2010). Die Inaktivierung von *CCR5* durch ZFNs wird bereits als hämatopoetische Stamm- und T-Zelltherapie in klinischen Studien (Phase 1/2) durch Sangamo Therapeutics getestet (ClinicalTrials.gov Identifier: NCT00842634, NCT01044654, NCT01252641, NCT01543152, NCT02225665, NCT02500849). Das unterstreicht das hohe therapeutische Potential der gezielten Genommodifikation durch TALEN und vergleichbare Werkzeuge.

Eine andere, potentielle Strategie zur Heilung von HIV-1 ist die Induktion der viralen Genexpression in latent infizierten Zellreservoir, wie den CD4⁺ T-Gedächtniszellen (Williams und Greene, 2007; Geeraert *et al.*, 2008; Choudhary und Margolis, 2011; Katlama *et al.*, 2013; Xing und Siliciano, 2013). Dadurch sollen diese Zellen für die Zerstörung durch das Immunsystem markiert werden oder auf Grund der Virusreproduktion zugrunde gehen, während ART eine weitere Verbreitung des Virus verhindert (*shock and kill*-Strategie; Geeraert *et al.*, 2008). Pharmakologische Ansätze zur transkriptionellen Induktion des Provirus verfolgten eine generelle Stimulation der Transkription durch Auflockerung des Chromatins oder die Aktivierung von Transkriptionsfaktoren, wie z. B. NFκB oder NFAT, die ähnlich wie bei der T-Zellaktivierung den viralen Promotor stimulieren (Williams und Greene, 2007; Choudhary und Margolis, 2011; Katlama *et al.*, 2013; Xing und Siliciano, 2013). Ein Nachteil von Pharmazeutika ist, dass sie nicht nur auf die infizierten Zellen, sondern auf den gesamten Organismus wirken und dementsprechend Nebenwirkungen hervorrufen können (Choudhary und Margolis, 2011; Katlama *et al.*, 2013; Xing und Siliciano, 2013).

1.5 Vorarbeiten

TALE-basierte Transaktivatoren, die auf eine spezifische DNA-Sequenz programmiert werden können, stellten eine Alternative zur HIV-1-Induktion ohne diesen Nachteil dar. Daher wurden in Vorarbeiten artifizielle TALE-Aktivatoren (ArtHIV1, 2, 5 und 6) mit VP16-AD zur Induktion der Transkription am HIV-1-Promotor erstellt (Abb. 1.4). Sie binden in einem Bereich stromaufwärts des natürlichen HIV-1-Transkriptionsstarts, der die TATA-Box, drei SP1-Boxen und zwei κB-Motive enthält (Abb. 1.4; van Opijnen *et al.*, 2004; Flint *et al.*, 2009). Jede der SP1-Boxen ist dabei eine Bindestelle des monomeren Zinkfingerproteins SP1, das ein für die HIV-1-Transkription grundlegender Transkriptionsfaktor ist (Rohr *et al.*, 2003; Flint *et al.*, 2009). Die κB-Motive werden von dem Transkriptionsaktivator NFκB erkannt, der ein Heterodimer aus den Proteinen p50 und p65 (RelA) darstellt. Neben NFκB kann dort auch NFAT binden. Beide sind vor allem zur Induktion der Transkription des Virus in T-Zellen wichtig (Rohr *et al.*, 2003; Flint *et al.*, 2009; Karn und Stoltzfus, 2012).

Um zu untersuchen, ob die artifiziellen TALE-Aktivatoren am HIV-1-Promotor Transkription induzieren können, wurden Luciferasereporterstudien mit Fragmenten des 5'-LTR des HIV-1 Subtyp B-Provirus aus der Zelllinie 8E5 durchgeführt (Folks *et al.*, 1986). Dabei wurde eine Abhängigkeit der Transkriptionsaktivierung durch die ArtHIVs von der Position im HIV-1-Promotor festgestellt (Abb. 1.4). So konnte nur ArtHIV6, der über den κB-Motiven bindet, eine Luciferaseaktivität, die deutlich über das Basalniveau des Reporterkonstruktes hinausging,

induzieren. Als Ursachen für diese Positionsabhängigkeit der TALE-vermittelten Transkriptionsaktivierung am HIV-1-Promotor wurde eine Unzugänglichkeit der DNA durch Chromatin oder andere *trans*-Elemente, insbesondere SP1, angenommen.

Ein ähnliches Ergebnis lieferte eine parallele Studie an dem gleichen, genomisch integrierten HIV-1-Promotor in 8E5-Zellen (Geißler *et al.*, 2015). Dabei wurde ein größerer Bereich des HIV-1-Promotors mit artifiziellen TALE-Aktivatoren abgedeckt. Die stärkste Induktion lieferte ein TALE-Aktivator (TALE5), dessen Zielsequenz sich unmittelbar stromaufwärts der κ B-Motive befand.

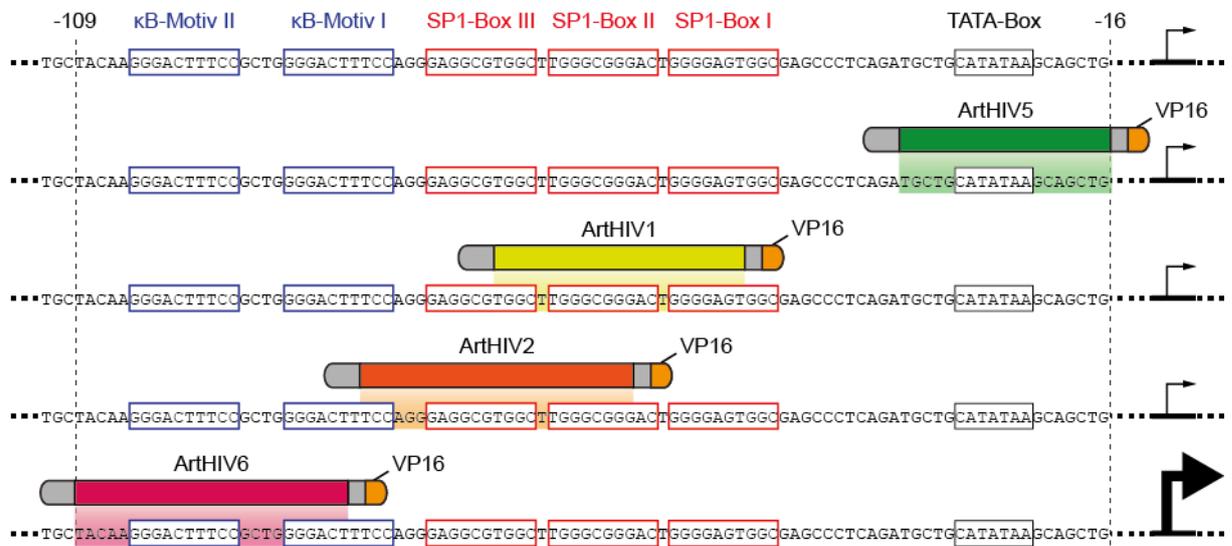


Abb. 1.4: Künstliche TALE-Aktivatoren am HIV-1-Promotor. In Vorarbeiten wurden vier künstliche TALE-Aktivatoren für den HIV-1-Promotor erstellt (ArtHIV1, 2, 5 und 6), die zwischen Position -109 und -16 in Bezug auf den natürlichen HIV-1-Transkriptionsstart binden (Annekatriin Richter, Diplomarbeit 2011). Zur effizienten Aktivierung in humanen Zellen wurde die TALE-AD durch eine VP16-AD ausgetauscht. Die TATA-Box (nach van Opijnen *et al.*, 2004), SP1-Bindestellen (Box I-III) und κ B-Motive im HIV-1-Promotorausschnitt sind markiert. Luciferasereporterstudien haben gezeigt, dass nur ArtHIV6 die Promotoraktivität im Vergleich zum Basalniveau ohne TALE erhöht (siehe Pfeilstärke). Bindestellen von Transkriptionsfaktoren im HIV-1-Promotor sind umrahmt.

1.6 Zielstellung

Die folgende Arbeit beschäftigt sich zum einen mit dem Einfluss seltener, aberranter *repeats* auf die DNA-Erkennung durch TALEs, insbesondere im Hinblick auf die beiden Virulenzfaktoren AvrXa7 und PthXo3.

Zum anderen soll die Anwendbarkeit von TALE-Aktivatoren zur Geninduktion im Hintergrund des HIV-1-Promotors untersucht werden. Außerdem sollten die Ursachen für die unterschiedliche Aktivität der verschiedenen HIV-TALEs auch im Vergleich zu künstlichen Cas9-basierten Aktivatoren analysiert werden. Darüber hinaus wurde eine Verbesserung der TALE-Aktivatoren für eine effizientere Aktivierung und im Hinblick auf die Erkennung verschiedener Virusvarianten angestrebt.

Schließlich war es ein Ziel dieser Arbeit neue TALEN-Varianten unter zwei verschiedenen Gesichtspunkten der Anwendung zu konstruieren. Ein Schwerpunkt bei der Verwendung molekularer Werkzeuge ist dabei die Effizienz ihrer Erstellung. So muss für die Nutzung von CRISPR/Cas9 nur eine sgRNA zur Erzeugung eines DSBs erstellt werden, während TALEN die Generierung zweier Moleküle je DSB erfordern. Hier sollten daher Untersuchungen zur Entwicklung und Anwendung monomerer TALEN-Varianten durchgeführt werden. Des Weiteren ist es von Interesse genommodifizierende Werkzeuge so spezifisch für eine Zielsequenz wie möglich zu gestalten. Dafür sollten hochspezifische TALEN, die zwei Sequenzen anhand nur einer Base unterscheiden können, erstellt werden.

2. Material und Methoden

2.1 Material

2.1.1 Verwendete Plasmide

Tab. 2.1: Verwendete Plasmide

Plasmid	Beschreibung ^(a)	Verwendung ^(b)	Resistenz	Referenz
pMP:GFP	GFP, CMV-Promotor (pcDNA3-Derivat)	Expression in humanen Zellen	Amp	Geißler <i>et al.</i> , 2011
pMP:H3VP16AD	Hax3 mit VP16AD, N-terminales GFP, CMV-Promotor (pcDNA3-Derivat)	Expression in humanen Zellen	Amp	Geißler <i>et al.</i> , 2011
pGGC	pUC57 ohne BsaI-Schnittstelle	Klonierung DNA-Fragmente	Amp	Geißler <i>et al.</i> , 2011
pGTLA	Zwischenvektor L-A, <i>BpiI</i> - und <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i>	Erstellung Sechser-repeat-Modul GTT	Kan	Geißler <i>et al.</i> , 2011
pGTAB	Zwischenvektor A-B, <i>BpiI</i> - und <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i>	Erstellung Sechser-repeat-Modul GTT	Kan	Geißler <i>et al.</i> , 2011
pGTBR	Zwischenvektor B-R, <i>BpiI</i> - und <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i>	Erstellung Sechser-repeat-Modul GTT	Kan	Geißler <i>et al.</i> , 2011
pGTBC	Zwischenvektor B-C, <i>BpiI</i> - und <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i>	Erstellung Sechser-repeat-Modul GTT	Kan	Geißler <i>et al.</i> , 2011
pGTCR	Zwischenvektor C-R, <i>BpiI</i> - und <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i>	Erstellung Sechser-repeat-Modul GTT	Kan	Geißler <i>et al.</i> , 2011
pGTCD	Zwischenvektor C-D, <i>BpiI</i> - und <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i>	Erstellung Sechser-repeat-Modul GTT	Kan	Richter <i>et al.</i> , 2014
pGTDR	Zwischenvektor D-R, <i>BpiI</i> - und <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i>	Erstellung Sechser-repeat-Modul GTT	Kan	Richter <i>et al.</i> , 2014
pcDHN	CMV-Promotor, <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i> (pcDNA3-Derivat)	Erstellung TAL-Effektoren mit N-terminalem Epitop zur Expression in humanen Zellen GTT	Amp	Geißler <i>et al.</i> , 2011
pcDHNS	CMV-Promotor, <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i> (pcDHN-Derivat)	Erstellung TAL-Effektoren mit N-terminalem Epitop zur Expression in humanen Zellen GTT	Spec	Ivonne Renner
pcDHO	CMV-Promotor, <i>BsaI</i> -Schnittstellen, <i>E. coli lacZ</i> (pcDNA3-Derivat)	Erstellung TAL-Effektoren zur Expression in humanen Zellen GTT	Amp	Geißler <i>et al.</i> , 2011
pNTH3	N-terminale Hax3-Region in pJet1.2, <i>BsaI</i> -Schnittstellen	Modul für N-terminale Hax3-Region GTT	Amp	Geißler <i>et al.</i> , 2011
pCTH3VP16	C-terminale Hax3-Region mit VP16-AD, <i>BsaI</i> -Schnittstellen, in pGGC (pUC57-Derivat für GGC)	Modul für C-terminale Hax3-Region GTT	Amp	Geißler <i>et al.</i> , 2011
pJetNtGFP	GFP in pJet1.2, <i>BsaI</i> -Schnittstellen	Modul für N-terminales Epitop GTT	Amp	Sabine Thieme
pcAH1	pcDHN-Derivat GFP:: <i>ArtHIV1</i>	mit Expression in humanen Zellen	Amp	Vorarbeiten
pcAH2	pcDHN-Derivat GFP:: <i>ArtHIV2</i>	mit Expression in humanen Zellen	Amp	Vorarbeiten
pcAH5	pcDHN-Derivat FLAG:: <i>ArtHIV5</i>	mit Expression in humanen Zellen	Amp	Vorarbeiten
pcAH6	pcDHN-Derivat FLAG:: <i>ArtHIV6</i>	mit Expression in humanen Zellen	Amp	Vorarbeiten
pFABHIV67	Box- <i>ArtHIV6/7</i> , Minimal-CMV-Promotor, pFAX-Derivat	Reporterkonstrukt	Amp	Vorarbeiten
pFAHPs	HIV-1 Promotor	Reporterkonstrukt	Amp	Vorarbeiten

pFAHPI	HIV-1 Promotor mit <i>gag leader sequence</i>	Reporterkonstrukt	Amp	Vorarbeiten
pSAR6	TALE34	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR7	sTALE30p8	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR8	TALE40p3	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR9	TALE40p8	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR10	TALE40p89	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR11	TALE40p14	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR12	TALE42p8	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR13	TALE34s (11.5 repeats)	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR14	TALE40p8s (11.5 repeats)	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR1	TALE40p3s (11.5 repeats)	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR2	AvrXa7	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR4	PthXo3	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR15	ArtXa72	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pSAR16	ArtXa73	Expression in <i>planta</i>	Spec	Richter <i>et al.</i> , 2014
pEBAR29	TALE-Box	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pEBAR30	TALE-Box -1p9	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pEBAR31	TALE-Box -2p9	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pEBAR32	TALE-Box +1p9	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pEBAR33	TALE-Box +2p9	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pEBAR35	TALE-Box -1p4	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pEBAR36	TALE-Box -1p15	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pE-AGT2	AGT2-Box	TALEN-Restriktionanalyse	Kan	Richter <i>et al.</i> , 2014
pGAR30	TALE-Box +1p9	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR31	TALE-Box +2p9	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR32	TALE-Box -1p4	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR28	TALE-Box 9.1	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR27	TALE-Box 9.3	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR30	TALE-Box 8.1	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR31	TALE-Box 8.3	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR17	TALE-Box	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR18	TALE-Box-1p3	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR19	TALE-Box-1p7	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR20	TALE-Box-1p8	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR21	TALE-Box-1p9	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR22	TALE-Box-1p15	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR24	TALE-Box-2p9	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR25	TALE-Box 8.2	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR26	TALE-Box 9.2	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR32	PthXo3-Box-2	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR1	PthXo3-Box-1	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR2	PthXo3-Box	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR3	PthXo3-Box+1	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR4	PthXo3-Box+2	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR5	AvrXa7 <i>Os11N3</i> -Box	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR6	AvrXa7-Box	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR7	AvrXa7-Box-1	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR8	AvrXa7-Box+1	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR9	AvrXa7-Box+2	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pGAR28	AvrXa7-Box-2	GUS-Reporterkonstrukt	Kan, Hyg	Richter <i>et al.</i> , 2014
pDAR9	AGT2-TALEN	<i>in vitro</i> Expression	Amp	Richter <i>et al.</i> , 2014
pDAR11	TALEN34	<i>in vitro</i> Expression	Amp	Richter <i>et al.</i> , 2014
pDAR12	TALEN40p3	<i>in vitro</i> Expression	Amp	Richter <i>et al.</i> , 2014
pDAR13	TALEN40p8	<i>in vitro</i> Expression	Amp	Richter <i>et al.</i> , 2014
pDAR14	TALEN40p14	<i>in vitro</i> Expression	Amp	Richter <i>et al.</i> , 2014
pDAR15	Bs4R-TALEN	<i>in vitro</i> Expression	Amp	Richter <i>et al.</i> , 2014
pAKR34	His ₆ -ArtHIV1C17	Expression in <i>E. coli</i>	Kan	diese Arbeit
pAKR37	His ₆ -ArtHIV2C17	Expression in <i>E. coli</i>	Kan	diese Arbeit

pcTAR26	FLAG-Hax3-VP16	Expression in hum. Zellen	Amp	diese Arbeit
pAKR71	FLAG-Cas9N0-VP16	Expression in hum. Zellen	Spec	diese Arbeit
pAKR72	GFP-Cas9N0-VP16	Expression in hum. Zellen	Spec	diese Arbeit
MLM3636	sgRNA	Expression in hum. Zellen	Amp	Addgene plasmid # 43860
JDS246	Cas9	Expression in hum. Zellen	Amp	Addgene plasmid # 43861
pcTAR17	TALE5	Expression in hum. Zellen	Amp	diese Arbeit
pAKR43	ArtHIV6-mVP16	Expression in hum. Zellen	Spec	diese Arbeit
pAKR44	ArtHIV6-VP64	Expression in hum. Zellen	Spec	diese Arbeit
pAKR45	ArtHIV6-p65	Expression in hum. Zellen	Spec	diese Arbeit
pAKR46	ArtHIV6-VpH	Expression in hum. Zellen	Spec	diese Arbeit
pAKR47	ArtHIV6-Vp	Expression in hum. Zellen	Spec	diese Arbeit
pAKR48	ArtHIV6-HSA1-AD	Expression in hum. Zellen	Spec	diese Arbeit
pAKR68	ArtHIV6-W7	Expression in hum. Zellen	Spec	diese Arbeit
pAKR69	ArtABC6	Expression in hum. Zellen	Spec	diese Arbeit
pAKR70	TALEABC5	Expression in hum. Zellen	Spec	diese Arbeit
pFAR3	HIV-1 MP	Luciferasereporterkonstrukt	Amp	diese Arbeit
pHIVA1	HIV-1 Promotor Subtyp A	Luciferasereporterkonstrukt	Amp	diese Arbeit
pHIVC1	HIV-1 Promotor Subtyp C	Luciferasereporterkonstrukt	Amp	diese Arbeit
pHIVD1	HIV-1 Promotor Subtyp D	Luciferasereporterkonstrukt	Amp	diese Arbeit
pHIVE1	HIV-1 Promotor Subtyp E	Luciferasereporterkonstrukt	Amp	diese Arbeit
pHIVF1	HIV-1 Promotor Subtyp F	Luciferasereporterkonstrukt	Amp	diese Arbeit
pCCR54	<i>hCCR5Δ32</i>	TALEN-Restriktionanalyse	Amp	diese Arbeit
pCCR51	<i>hCCR5</i>	TALEN-Restriktionanalyse	Amp	diese Arbeit
pUAR1	HIV-1 <i>gagpol</i>	TALEN-Restriktionanalyse	Amp	diese Arbeit
pTUBA1A	<i>D. rerio tuba1a</i>	TALEN-Restriktionanalyse	Amp	diese Arbeit
ptardbp	<i>D. rerio tardbp</i>	TALEN-Restriktionanalyse	Amp	diese Arbeit
pFANCI1	<i>Hs FANCI</i>	TALEN-Restriktionanalyse	Amp	diese Arbeit
pFANCI M1	<i>Hs FANCI</i> mutiertes Allel	TALEN-Restriktionanalyse	Amp	diese Arbeit
pAKR4-31	FTA-TALEN	<i>in vitro</i> Expression	Spec	diese Arbeit
pDAR22-33	Zef-TALEN	<i>in vitro</i> Expression	Amp	diese Arbeit
pQEAR1	TALEN L1	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR2	TALEN R1	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR3	TALEN scR1	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR4	TALEN NscR1	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR5	TALEN NscR2	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR6	TALEN dscR1	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR7	TALEN dscR2	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR8	TALEN Lds	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR9	TALEN Rrr	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR10	TALEN scRds	<i>in vitro</i> Expression	Kan	diese Arbeit
pQEAR11	TALEN scRrr	<i>in vitro</i> Expression	Kan	diese Arbeit
pcNAR4	TALEN wtR	<i>in vitro</i> Expression	Spec	diese Arbeit
pAKR66	TALEN wtL	<i>in vitro</i> Expression	Spec	diese Arbeit
pAKR67	TALEN wtscR	<i>in vitro</i> Expression	Spec	diese Arbeit
pcNAR23	TALEN R	<i>in vitro</i> Expression	Amp	diese Arbeit
pcNAR7	TALEN L	<i>in vitro</i> Expression	Spec	diese Arbeit
pcNAR13	TALEN scR	<i>in vitro</i> Expression	Spec	diese Arbeit
pAKR50	TALEN L1c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR51	TALEN R1c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR54	TALEN scR1c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR55	TALEN NscL1c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR56	TALEN NscL2c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR57	TALEN NscR1c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR58	TALEN NscR2c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR60	TALEN dscR1c	Expression in hum. Zellen	Spec	diese Arbeit
pAKR61	TALEN dscR2c	Expression in hum. Zellen	Spec	diese Arbeit

(a) GGC = *Golden Gate Cloning* (b) Die für Klonierungen mit der *Golden TAL* Technologie verwendeten Plasmide sind mit GTT gekennzeichnet.

2.1.2 Verwendete Oligonukleotide

Die verwendeten Oligonukleotide (Tab. 2.2) wurden von der Firma Metabion GmbH (Martinsried) oder Eurofins MWG Operon (Ebersberg) bezogen.

Tab. 2.2: Verwendete Oligonukleotide

Name ^(a)	Sequenz	Verwendung
NTM_FLAG_fw	ATCCATGGACTACAAGGATGACGATGACAAGCA	N-terminales FLAG-Epitop-Modul für TALEs/ TALEN
NTM_FLAG_rev	CATATGCTTGTGCATCGTCATCCTTGTAGTCCAT	
HA_F/AR	ATCCATGTACCCCTACGACGTGCCCGACTACGCCCA	N-terminales HA-Epitop-Modul für TALEs/ TALEN
HA_R/AR	CATATGGGCGTAGTCGGGCACGTCGTAGGGGTACAT	
HIVfwXbalPro (E)	GATTCTAGATGGAAGGGCTAATTCCTCCC	Erstellung von
HIVrvHindIIIPro (E)	CCCAAGCTTTCCAGGCTCAGATCTGGTC	Luciferasereporter-
HIVrvPro (E)	CTCTCTCCTTCTAGCCTCCGC	Konstrukten mit HIV-1
LuciferaseATG (E)	ATGGAAGACGCCAAAACATAAAGAAAGGC	Promotoren
pF12Arv1 (E)	ATCTTTCTAGAGTCGACTATCAC	
HIV-A-F	CTGACACAGAAGTTGCTGACAGGGACTTTCCGCTGG GGACTTTC	Erstellung von
HIV-A-R	CCCTGTCAGCAACTTCTGTGTCAGCAGTTCTGAAGT ACTCCGG	Luciferasereporter-
HIV-C-F	CCGCTGGGACTTTCCACTGGGGCGTTCCAGGGAGGC GTGGCTTGG	Konstrukten mit HIV-1
HIV-C-R	CCCAGTGGAAAGTCCCAGCGGAAAGTCCCTTCTGTG TCAGCAGTTCTTGAAGTACTCCGG	Subtyp Promotoren
HIV-D-F	CTACAAGGGACTTTCCGCTGGGACTTTCCAGGGAGG CGTG	
HIV-D-R	GTCCAGCGGAAAGTCCCTTGTAGAAACCTCGGTGT CAGCAGTTCTTGAAGTACTCCGG	
HIV-E-F	GGACTTCCGCTGGGGACTTTCCAGGG	
HIV-E-R	GGAAAGTCCCAGCGGAAAGTCTAGTTAGAAACTTC TTTGTGTCAGCAGTTCTTGAAGTACTCCGG	
HIV-F-F	GATTGCTGACACAGAAGATTCTAAAGGGACTTTCCG CTGGGGACTTTC	
HIV-F-R	GTCCCTTTAGAATCTTCTGTGTCAGCAATCTCTGTGT CATCAGTTCTTGAAGTACTCCGGATGCAG	
HP-SPI_F	GGACTGAAAAGTGGCGAGCCCTCAG	Erstellung von Luciferase-
HP-SPI+ME_R	CTCGCCACTTTTTCAGTCCCGTTTAAGCCACGTTTCCC	reporter-Konstrukt mit SPI-
Luc_36_R_AR	CGCCGGGCTTTTCTTTATG	mutiertem HIV-1 Promotor
Luc973R_AR	CCCCGACTTCTTAGAGAGGGGAGCGCC	5'-RACE
Luc321R_AR	GTCGTTCCGCGGGCGCAACTGCAACTCCG	
HIV_wt-P_F/AR	CCGCTGGGGACTTTCCAGGGAGGCGTGGCTTGGGCG GGACTGGGGAGTGGCGAGCCCTCAG	EMSA
HIV_wt-P_R/AR	CTGAGGGCTCGCCACTCCCCAGTCCCGCCCAAGCCA CGCTCCCTGGAAAGTCCCCAGCGG	
HIV_mu-P_F/AR	CCGCTGGGGACTTTCCAGGGAAACGTGGCTTAAACG GGACTGAAAAGTGGCGAGCCCTCAG	
HIV_mu-P_R/AR	CTGAGGGCTCGCCACTTTTTCAGTCCCGTTTAAGCCAC GTTTCCCTGGAAAGTCCCCAGCGG	
B-HIV_wt-P_F/AR	Biotin-CCGCTGGGGACTTTCCAGGGAGGCGTGGCTT GGGCGGGACTGGGGAGTGGCGAGCCCTCAG	
B-HIV_wt-P_R/AR	Biotin-CTGAGGGCTCGCCACTCCCCAGTCCCGCCCA AGCCACGCCTCCCTGGAAAGTCCCCAGCGG	

B-HIV_mu-P_F/AR	Biotin-CCGCTGGGGACTTTCCAGGGAAACGTGGCTT AAACGGGACTGAAAAGTGGCGAGCCCTCAG	EMSA
B-HIV_mu-P_R/AR	Biotin-CTGAGGGCTCGCCACTTTTCAGTCCCCTTTAA GCCACGTTTCCCTGGAAAGTCCCCAGCGG	
B-HIV_neg-P_F/AR	Biotin-CCAGTGGAGCCAGATAAAGTAGAAGAAGCC AATAAAGGAGAGAACACCCGCTTGTTACACC	
B-HIV_neg-P_R/AR	Biotin-GGTGTAACAAGCGGGTGTCTCTCCTTTATTG GCTTCTTCTACTTTATCTGGCTCCACTGG	
HIV_neg-P_F/AR	CCAGTGGAGCCAGATAAAGTAGAAGAAGCCAATAA AGGAGAGAACACCCGCTTGTTACACC	
HIV_neg-P_R/AR	GGTGTAACAAGCGGGTGTCTCTCCTTTATTGGCTTC TTCTACTTTATCTGGCTCCACTGG	
B-HIV_GC2-P_F/AR	Biotin-CCGCTGGGGACTTTCCAGGGAAACGTGGCTT GGGCGGGACTGAAAAGTGGCGAGCCCTCAG	
B-HIV_GC2-P_R/AR	Biotin-CTGAGGGCTCGCCACTTTTCAGTCCCGCCCA AGCCACGTTTCCCTGGAAAGTCCCCAGCGG	
HIV_GC2-P_F/AR	CCGCTGGGGACTTTCCAGGGAAACGTGGCTTGGGCG GGACTGAAAAGTGGCGAGCCCTCAG	
HIV_GC2-P_R/AR	CTGAGGGCTCGCCACTTTTCAGTCCCGCCCAAGCCAC GTTTCCCTGGAAAGTCCCCAGCGG	
MussCCR5-F/AR	AAGATGGATTATCAAGTGTCAAGTCC	PCR für T7 Endonuklease I- Verdau
MussCCR5-R/AR	CAAAGTCCCCTGGGGCG	

(a) Mit (E) gekennzeichnete Oligonukleotide wurden von Eurofins MWG Operon (Ebersberg) bezogen. Die Übrigen von Metabion GmbH (Martinsried).

2.1.3 Biologisches Material

2.1.3.1 Verwendete Bakterienstämme

Zur Vermehrung von Plasmiden wurden in dieser Arbeit *Escherichia coli* OneShot® Top10 (Invitrogen, Groningen) verwendet. Für die Proteinexpression wurden *E. coli* BL2-AI verwendet (Invitrogen, Groningen). Die transiente Expression von Genen in *Nicotiana benthamiana* erfolgte mittels *Agrobacterium tumefaciens* (Larebeke *et al.*, 1974). Die Eigenschaften der Stämme sind Tabelle 2.3 zu entnehmen.

Tab. 2.3: Eigenschaften der verwendeten Bakterienstämme

Bakterienstamm	Eigenschaften
<i>E. coli</i> OneShot Top10	F- <i>mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15\Delta lacX74$ <i>recA1</i> <i>araD139</i> $\Delta(araleu)$ 7697 <i>galU</i> <i>galK</i> <i>rpsL</i> (StrR) <i>endA1</i> <i>nupG</i>
<i>E. coli</i> OneShot BL21-AI	F <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻</i> <i>m_B⁻</i>) <i>gal</i> <i>dcm</i> <i>araB::T7RNAP-tetA</i>
<i>A. tumefaciens</i> GV3101	C58, RifampicinR; Ti-Plasmid: pMP90 (pTIC58_T-DNA) GentamycinR, Gene für Nopalinsynthese

2.1.3.2 Verwendete Zelllinien, Kultivierung und Passage

Reportergenaktivierungsstudien, Protein- und RNA-Analysen wurden unter Verwendung der *human embryonic kidney* Zelllinie Hek293T-REx™ (Invitrogen, Karlsruhe) durchgeführt. Diese aus einer embryonalen Niere stammende humane Zelllinie wurde durch Transformation mit der DNA-Sequenz des E1 Proteins des Adenovirus 5 generiert (Graham *et al.*, 1977) und

exprimiert zusätzlich das SV40 *large T-Antigen*. Darüber hinaus ist ein durch Zugabe von Blastocidin selektierbares Tetracyclin-abhängiges Expressionssystem (pcDNA6/TR) enthalten.

Für immunhistochemische Untersuchungen wurde eine andere humane Zelllinie, *human hepato cellular carcinoma* Huh7-Zellen (Ralf Bartenschlager, Heidelberg, Deutschland), verwendet. Sie wurde aus Lebertumorgewebe eines 57-jährigen Japaners gewonnen (Nakabayashi *et al.*, 1984).

Die Kultivierung beider adhärent wachsender Zelllinien erfolgte in 10 cm Kulturschalen (Corning, Amsterdam, Niederlande) mit 10 ml Medium bei 37°C und 5% CO₂. Bei einem Bedeckungsgrad von 90% vor Erreichen der Konfluenz wurden die Zellen passagiert. Dazu wurde das überstehende Medium entfernt und die Zellen mit 10 ml 1xPBS kurzzeitig überschichtet um Mediumreste zu entfernen (Waschschritt). Im Anschluss wurden die Zellen bis zur Ablösung vom Boden der Kulturschale mit 500 µl 0,25% Trypsin-EDTA (GIBCO, Invitrogen) inkubiert. Nach Abnahme der Lösung wurden die Zellen zusätzlich durch Abklopfen mechanisch gelöst, um dann in 10 ml Medium aufgenommen zu werden. Die hergestellte, homogenisierte Zellsuspension wurde zur Aussaat in neue 10 cm Kulturschalen mit Medium so verdünnt, dass ein Fünftel der Hek293T-REx- und ein Zehntel der Huh7-Zellen in die Folgekultur umgesetzt wurden.

2.1.3.3 Verwendetes Pflanzenmaterial

In dieser Arbeit wurde das Pflanzenmodell *Nicotiana benthamiana* verwendet. *N. benthamiana* wurde unter Gewächshausbedingungen angezogen (26-28 °C; 60-70 % relative Luftfeuchte; Langtag-Lichtperiode 6 bis 22 Uhr).

2.1.4 Chemikalien, Geräte und Verbrauchsmaterialien

Chemikalien, Geräte und Verbrauchsmaterialien wurden, sofern nicht anders angegeben, von folgenden Firmen bezogen:

Amersham Pharmacia Biotech (Freiburg i. Breisgau), AppliChem GmbH (Darmstadt), Carl Roth GmbH & Co. (Karlsruhe), Clontech Laboratories GmbH (Heidelberg), Corning (Amsterdam, Niederlande) Difco (Augsburg), Eppendorf (Hamburg), GE Healthcare (Freiburg), Greiner Labortechnik GmbH (Solingen), Invitrogen (Groningen, Niederlande), Fermentas GmbH (St. Leon Rot), Merck (Darmstadt), Metabion GmbH (Martinsried), New England Biolabs GmbH (Schwalbach/Taunus), peqlab Biotechnologie GmbH (Erlangen), Promega Corporation (Madison USA), Qiagen GmbH (Hilden), Roche Molecular

Biochemicals GmbH (Mannheim), Sigma Aldrich Chemie GmbH (Taufkirchen), VWR (Darmstadt)

2.1.5 Verwendete Medien und Zusätze

Zur Anzucht der verwendeten *E. coli*-Stämme bei 37°C wurde LB (*lysogeny broth*) Medium genutzt. Zur Anzucht von *A. tumefaciens* bei 30°C wurde YEB (*yeast extract broth*) Medium verwendet. Die Zusammensetzungen der Medien ist Tabelle 2.4 zu entnehmen. Zur Herstellung von Nährplatten wurden 1,5% (w/v) Bacto-Agar vor dem Sterilisieren (bei 121°C und 2 bar für 20 min im Autoklaven) zum Flüssigmedium hinzugefügt.

Tab. 2.4: Zusammensetzung der zur Bakterienanzucht verwendeten Nährmedien

Bezeichnung	Zusammensetzung
LB Medium	1% (w/v) Bakto-Trypton 0,5% (w/v) Hefe Extrakt 1% (w/v) NaCl pH 7,5
YEB Medium	0,5 % (w/v) Fleischextrakt 0,5 % (w/v) Trypton 0,5 % (w/v) Saccharose 2 mM MgSO ₄ 0,1 % (w/v) Hefeextrakt pH 7,2

Zur selektiven Kultivierung wurden autoklavierten Medien folgende Antibiotika zugesetzt:

Tab. 2.5: Für bakterielle Kulturen verwendete Antibiotika

Antibiotikum	Endkonzentration
Ampicillin (Amp)	100 µg/ml (festes Medium) 50 µg/ml (flüssiges Medium)
Kanamycin (Kan)	25 µg/ml
Hygromycin (Hyg)	50 µg/ml
Spectinomycin (Spec)	100 µg/ml
Rifampicin (Rif)	100 µg/ml

Zur Blau-Weiß-Selektion wurden Festmedien vor dem Erstarren mit 0,4% (w/v) X-Gal als Indikator versetzt. Dazu diente eine 2%ige (w/v) X-Gal-Lösung in DMF (**D**imethylformamid).

Nährplatten und Flüssigmedien wurden bei 4 °C gelagert. Für die dauerhafte Lagerung von Bakterien bei -80 °C wurde eine Bakteriensuspension mit 20% Glycerin versetzt.

Das für die Kultivierung beider verwendeter, humaner Zelllinien eingesetzte Medium basierte auf einem **D**ulbecco's **M**odified **E**agle's **M**edium (DMEM; GIBCO®, Invitrogen, Darmstadt). Zum Schutz vor mikrobiellem Wachstum wurde Penicillin-Streptomycin, liquid (10000

units/ml Penicillin, 10000 µg/ml Streptomycin in 0,85% Salzlösung; GIBCO®, Invitrogen) hinzugefügt. Die Zusammensetzung des Mediums ist Tabelle 2.6 zu entnehmen.

Tab. 2.6: Zusammensetzung des für humane Zellen verwendeten Kulturmediums

Bestandteil	Anteil/Endkonzentration
Grundmedium	
DMEM	88,8% (v/v)
[+] D-Glukose	4 g/l
[+] Nicht essentielle Aminosäuren	
[-] L-Glutamin	
[-] Pyruvat	
Andere Bestandteile	
FCS (<i>foetal calf serum</i> ; PAN Biotech)	10% (v/v)
Biotin (Sigma-Aldrich)	0,1 µg/ml
Hypoxanthin (Sigma-Aldrich)	0,25 µg/ml
Penicillin-Streptomycin, liquid	1% (v/v)
Penicillin	100 units/ml
Streptomycin	100 µg/ml

2.1.6 Verwendete Lösungen und Puffer

Tab. 2.7: Verwendete Lösungen und Puffer

Bezeichnung	Zusammensetzung
Puffer für DNA-Analysen	
10xPCR-Puffer	100 mM Tris-HCl, pH 8,5 500 mM KCl 15 mM MgCl ₂ 1 % (v/v) Triton X100 0,1 % (w/v) Gelatine
1xTAE	40 mM Tris-Acetat, pH 8,0 1 mM EDTA
5xDNA-Ladepuffer	15 % Ficoll 400 Farbstoff: Orange G
Puffer für Proteinanalysen	
5x Lämmli	250 mM Tris-HCl, pH 6,8 50 % (v/v) Glycerin 10 % (w/v) SDS 25 % (v/v) 2-Mercaptoethanol 0,005 % (w/v) Bromphenolblau
1x TANK	25 mM Tris 250 mM Glycin 0,1% (w/v) SDS
1x TBST	50 mM Tris-HCl, pH 8,0 150 mM NaCl 0,05% (v/v) Tween20
Blockierungslösung	1xTBST 5% (w/v) Magermilchpulver 3% (w/v) BSA (B ovines S erumalbumin F raktion V)
Transferpuffer	20% Methanol 20 mM Tris 150 mM Glycin

Trenngelpuffer	1,5 M Tris-HCl pH 8,8
Sammelgelpuffer	1 M Tris-HCl pH 6,8
Detektionslösung	100 mM Tris-HCl, pH 8,5 225 nM p-Cumarinsäure 1,25 mM 3-Aminophthalhydrazid 0,3 % (v/v) H ₂ O ₂
Puffer für RNA-Analysen	
1xMOPS	20 mM MOPS (3-(N-Morpholino)propansulfonsäure) 5 mM Natriumacetat 1 mM EDTA, pH 7
10xLadepuffer	50% (v/v) Glycerin 0,1 M EDTA, pH 8 0,25% (w/v) Bromphenolblau 0,25% (w/v) Xylencyanol 100 µg/ml Ethidiumbromid
Puffer und Lösungen für Zellkulturanwendungen und Immunhistochemie	
1xPBS	137 mM NaCl 2,7 mM KCl 1,5 mM KH ₂ PO ₄ 6,5 mM Na ₂ HPO ₄
Paraformaldehydlösung	4% (w/v) Paraformaldehyd (<i>reagent grade, crystalline</i> ; Sigma-Aldrich) gelöst in 1xPBS

2.2 Molekulargenetische Methoden

2.2.1 Präparation von Nukleinsäuren

2.2.1.1 Plasmidpräparation

Die Isolierung von Plasmiden aus *E. coli* wurde dem Prinzip der alkalischen Lyse (Sambrook *et al.*, 1989) folgend mittels QIAprep Spin Miniprep Kit oder QIAGEN Plasmid *Plus* Midi Kit (Qiagen GmbH, Hilden) nach Herstellerprotokoll durchgeführt. Eluiert wurde die Plasmid-DNA mit 50 µl oder 100 µl (Qiagen Plasmid *Plus* Midi Kit) destilliertem Wasser.

2.2.1.2 Isolation von RNA aus humanen Zellen

Zur Isolation von Gesamt-RNA inklusive eines DNase-Verdaus aus Hek293T-REx-Zellen wurde das RNeasy Mini Kit (Qiagen GmbH, Hilden) in Kombination mit dem RNase-Free DNase Set (Qiagen GmbH, Hilden) nach Herstellerprotokoll verwendet. Die RNA wurde in 30 µl RNase-freiem Wasser eluiert.

2.2.1.3 Isolation von genomischer DNA aus humanen Zellen

Zur Isolation genomischer DNA aus Hek293T-REx-Zellen wurde das DNeasy Blood & Tissue Kit (Qiagen GmbH, Hilden) nach Herstellerprotokoll verwendet. Die DNA wurde in 200 µl nukleasefreiem Wasser eluiert.

2.2.2 Spektrophotometrische Nukleinsäurebestimmung

Die Konzentration von Nukleinsäuren in ng/μl (c_{NA}) wurde nach der Formel $c_{NA}=A_{260} \cdot X_i$ mit dem Umrechnungsfaktor $X_1=50$ für DNA und $X_2=40$ für RNA aus der Absorption bei 260 nm (A_{260}) bestimmt. Die Absorption bei 280 nm (A_{280}) diente der Beurteilung der Reinheit nach dem Quotienten A_{260}/A_{280} . Die Absorptionsmessung erfolgte direkt mit 1 μl Lösung an einem Thermo Scientific NanoDrop 1000 Spektrophotometer gegen destilliertes Wasser.

2.2.3 PCR-Amplifikation von DNA-Fragmenten

Die PCR (*polymerase chain reaction*, Saiki *et al.*, 1985) zur Amplifikation von DNA-Fragmenten erfolgte mit der Phusion-Polymerase (Thermo Fisher Scientific, Vantaa) nach Herstellerangaben. Diese amplifiziert DNA auf Grund einer Korrekturlese- (*proof-reading*) Funktion mit geringer Fehlerrate. Dauer der Elongationszeit (30 s pro 1 kb) und Anzahl der Zyklen wurden jeweils angepasst. Die Anlagerungstemperatur wurde abhängig von den jeweiligen als Primer eingesetzten Oligonukleotiden gewählt.

Die Kolonie-PCR an *E. coli* wurde unter Verwendung von 1U *Taq*-Polymerase in 1xPCR-Puffer-gepufferten 20 μl Ansätzen mit 4 nmol je **Desoxyribonukleosidtriphosphat** (dNTP) und mindestens 1 pmol beider Oligonukleotide ebenfalls unter entsprechender Anpassung von Elongationszeit (60 s pro 1 kb) und Anlagerungstemperatur analog einer hier für *Taq*-PCR generalisierten Prozedur durchgeführt. Dieser entsprechend erfolgte die Reaktion:

Initiale Denaturierung	94°C	5 min
35 Zyklen:	Denaturierung	94°C 30 s
	Anlagerung	50-70°C 30 s
	Elongation	72°C 20 s -3 min
Finale Elongation	72°C	5 min
	4-8°C	∞

Die Durchführung aller Reaktionen erfolgte in Temperaturzyklus-Steuergeräten (PTC-225Peltier Thermal Cycler, MJ Research FlexCycler, Analytik Jena) mit Heizdeckel. Erhaltene DNA-Fragmente wurden mittels Agarose-Gelelektrophorese analysiert.

2.2.4 Restriktionsanalysen

Um isolierte und amplifizierte DNA sequenzspezifisch zu fragmentieren wurden Restriktionsendonukleasen verwendet. Bei Puffereinsatz, Reaktionstemperatur sowie -zeit und

eingesetzten Enzymmengen wurden die Herstellerangaben (Fermentas GmbH, St. Leon Rot bzw. New England Biolabs, Schwalbach/Taunus) befolgt.

2.2.5 Agarosegelelektrophorese

2.2.5.1 Auftrennung von DNA-Fragmenten

DNA-Fragmente der mit DNA-Ladepuffer versetzten Proben wurden in einer horizontalen Gelelektrophorese aufgetrennt. Für Fragmente kleiner als 300 bp und für Fragmente größer als 300 bp wurden 2%ige (w/v) bzw. 1%ige (w/v) 1xTAE-gepufferte Agarosegele verwendet. 400 ml Agaroselösung wurden vor Beginn des Erstarrungsprozesses mit 12 µl einer 10 mg/ml Stammlösung des in DNA interkalierenden Ethidiumbromids zur Sichtbarmachung der DNA unter UV-Licht im Transilluminator versetzt. Der Laufpuffer war 1x TAE. Als Größenstandards wurden der GeneRuler™ 1 kb DNA Ladder, der GeneRuler™ 100 bp DNA Ladder oder der O'Range™ Ruler 100bp+500bp DNA Ladder (Fermentas GmbH, St. Leon Rot) verwendet.

2.2.5.2 Auftrennung von Gesamt-RNA

Die Auftrennung von RNA erfolgte nach analogem Prinzip zur DNA-Auftrennung und sollte der Untersuchung der Isolate auf möglicherweise stattgefundene Degradation dienen. Dazu wurde zum Erhalt der aufgetragenen RNA-Probe ein denaturierendes 1,2%iges (w/v) Agarosegel im Laufpuffer 1xMOPS mit 2% (v/v) Formaldehyd hergestellt. Ethidiumbromid war im verwendeten Ladepuffer enthalten.

2.2.6 Reinigung von DNA-Fragmenten aus Reaktionsansätzen oder Agarosegelen

Auf geeigneten Agarosegelen aufgetrennte DNA-Fragmente wurden zur Reinigung derer mit gewünschter Größe auf einem UV-Tisch sichtbar gemacht und mit einem Skalpell ausgeschnitten. Die Elution wurde mit dem *MinElute Gel Extraction Kit* (Qiagen GmbH, Hilden) nach Herstellerangaben durchgeführt. Zur direkten Reinigung von DNA-Fragmenten aus Reaktionsgemischen wurde das *MinElute PCR Purification Kit* (Qiagen GmbH, Hilden) oder das *GeneJET PCR Purification Kit* (Thermo Scientific) nach Herstellerangaben verwendet. Die DNA wurde anschließend in 10 bis 20 µl destilliertem Wasser wieder aufgenommen.

2.2.7 Phosphorylierung von PCR-Produkten

Aus Amplifikation mittels PCR stammende DNA-Fragmente haben keine für die enzymatische Ligation mit anderen DNA-Fragmenten notwendige Phosphat-Gruppe an den freien 5'-Enden

von Plus- und Minusstrang. Daher wurde für Klonierungen, die die Ligation zweier PCR-Produkte erforderten, ein gereinigtes DNA-Fragment in einem 20 µl Ansatz mit 10 U T4 Polynukleotid-Kinase (T4 PNK, Fermentas GmbH, St. Leon Rot) 45 min in 1x T4-DNA-Ligasepuffer phosphoryliert. Der Ansatz wurde im Anschluss zur Klonierung eingesetzt.

2.2.8 Klonierung von DNA-Fragmenten

Die Ligation von DNA-Molekülen wurde in 20 µl Ansätzen mit 5 U T4-DNA-Ligase (Fermentas GmbH, St. Leon Rot) in 1x T4-DNA-Ligasepuffer 10 min bei 22°C durchgeführt. Das zu klonierende DNA-Fragment wurde dabei in fünffachem Überschuss zum Vektor eingesetzt.

Die Klonierung von PCR-Fragmenten in pGGC (Geißler *et al.*, 2011) erfolgte in einer *restriction ligation* durch Nutzung von 10 U des *blunt end-cutters* SmaI (Fermentas GmbH, St. Leon Rot) und 1 U T4-DNA-Ligase in einem 1x Tango (Fermentas GmbH, St. Leon Rot)-gepufferten 10 µl Ansatz mit 10 pmol ATP. Bei Einsatz von 1 µl des Vektors verblieben 5 µl Volumen für das gereinigte PCR-Produkt. So wurden auch neue Module für die *Golden TAL* Technologie erstellt.

Die nicht gerichtete Klonierung von *Taq*-PCR-Fragmenten mittels TOPO TA Cloning[®] Kit (Invitrogen, Groningen, Niederlande) in den T-Überhangvektor pCR2.1[®]-TOPO[®] erfolgte ebenso nach Herstellerangaben wie die gerichtete Klonierung von PCR-Produkten mit stumpfen Enden in pENTR/D-TOPO (Invitrogen, Groningen, Niederlande). Zur Ausrichtung des DNA-Fragments wurde das 5'-Ende bei der PCR-Amplifikation durch eine CACC-Sequenz im *forward*-Primer festgelegt.

In pENTR/D-TOPO klonierte DNA-Fragmente können über LR-Rekombination gerichtet in Zielvektoren übertragen werden (*GATEWAY*-System; Life Technologies). Dafür flankieren attL-Sequenzen das zu rekombinierende DNA-Fragment in pENTR/D-TOPO. Die LR-Rekombination zwischen attL- und attR-Sequenzen bewirkt den Austausch einer attR-flankierten Selektionskassette im Zielvektor durch das zu rekombinierende DNA-Fragment in pENTR/D-TOPO. Dafür wurde der *Gateway[®] LR Clonase[®] Enzyme mix* (Life Technologies) nach Herstellerangaben verwendet. Der Ansatz wurde 30 min bis 1 h bei Raumtemperatur inkubiert.

Die jeweils erhaltenen Plasmide wurden chemisch in *E. coli* transformiert.

2.2.9 Die Erstellung von Luciferase-Reporterkonstrukten

Die Luciferase-Reporterkonstrukte mit mutiertem oder lokal subtypischem HIV-1-Promotor wurden mit Hilfe von SOE-PCRs (*splicing by overlapping extension PCR*) erstellt. Dabei wurden durch Primer-Überhänge die gewünschten Veränderungen in einer ersten PCR an den Luciferase-Reporterkonstrukten pFAHPs und pFAHPI, die LTR-Fragmente des HIV-1 Subtyp B-Provirus aus 8E5-Zellen (Folks *et al.*, 1986) enthalten, in die beiden Halbseiten des HIV-1-Promotors (stromauf- und stromabwärts der Veränderung) eingeführt. Die im Bereich der Veränderung überlappenden, gereinigten PCR-Produkte dienten in der zweiten PCR als Matrize. Die Luciferase-Reporterplasmide pFAHPs und pFAHPI dienten außerdem als Klonierungsvektoren. Die SOE-PCR-Produkte mit den subtypisch angepassten HIV-1-Promotoren wurden über *XbaI* und *HindIII* in pFAHPs kloniert, sodass der Subtyp B Promotor ausgetauscht wurde. Bei der Amplifikation erfolgte dabei auch das Anhängen der entsprechenden Schnittstellen. Bei der Klonierung des mutierten HIV-1-Promotors wurde nur die *XbaI*-Schnittstelle 5' angehängt. Der benötigte Vektoranteil von pFAHPI wurde so in einer PCR amplifiziert (mit den Oligonukleotiden LuciferaseATG und pF12Arv1), dass eine gerichtete Klonierung über den *XbaI*-Überhäng und ein stumpfes Ende nach Phosphorylierung eines Fragmentes möglich wurde.

2.2.10 Die Erstellung von β -Glucuronidase-Reporterkonstrukten

Natürliche und artifizielle TALE-Boxen wurden zusammen mit einem Bs4-Minimalpromotor gerichtet in pENTR/D-TOPO (Invitrogen) kloniert und mittels LR-Rekombination in pGWB3 (Nakagawa *et al.*, 2007) rekombiniert, sodass das Promotorderivat vor dem β -Glucuronidasegen *uidA* inseriert wird.

2.2.11 Die Erstellung artifizieller TALE-Derivate mit der *Golden TAL Technologie*

Die Erstellung von TALE-Derivaten erfolgte mit Hilfe der *Golden TAL Technologie*. Zur Verfügung standen die in Geißler *et al.* (2011), Streubel *et al.* (2012) und Richter *et al.* (2014) publizierten *repeat*-Module. In einem ersten Schritt wurden Multi-*repeat*-Module aus Einzel-*repeat*-Modulen in einem Zwischenvektor assembliert. Dazu wurden DNA-Lösungen mit 50 ng/ μ l verwendet. Die Reaktion wurde wie folgt zusammengesetzt:

- 2 μ l Zwischenvektor (pGTLA/pGTAB/pGTBR/pGTBC/pGTCR/ pGTCD/ pGTDR)
- 2 μ l je Einzel-*repeat*-Modul (maximal 6)
- 2 μ l 10 mM ATP (Fermentas)

2 μ l NEB Puffer 4 (New England Biolabs)

1 μ l T4-DNA-Ligase (30 U/ μ l, Fermentas)

1 μ l *Bpi*I (10 U/ μ l, New England Biolabs)

ad 20 μ l (H₂O)

Die *cut ligation* erfolgte in einem Temperaturzyklus-Steuergerät mit Heizdeckel (PTC-225 Peltier Thermal Cycler, MJ Research FlexCycler, Analytik Jena) nach folgendem Programm:

10 Zyklen: 37°C 2 min

16°C 5 min

16°C 20 min

80°C 20 min

4°C ∞

Ein Teil des Ansatzes wurde dann in chemisch kompetente *E. coli* TOP10 transformiert. Nach Anzucht auf Nährplatten erfolgte die Blau-Weiß-Selektion zur Auswahl geeigneter Klone, die durch Restriktionsanalyse und auf Sequenzebene überprüft wurden.

Die so erstellten Multi-*repeat*-Module in pGT-Vektoren wurden zusammen mit bis zu zwei N- und bis zu zwei C-terminalen Modulen zur Erstellung künstlicher TALE-Derivate im Zielvektor verwendet. Das eingesetzte Volumen für jedes N- und C-terminale Modul wurde dessen Größe angepasst und schwankte zwischen 0,5 bis 3 μ l. Die Reaktion wurde wie folgt angesetzt:

2 μ l Zielvektor

x μ l je N-terminales Modul

x μ l je C-terminales Modul

1 μ l je Multi-*repeat*-Modul

2 μ l 10 mM ATP (Fermentas)

2 μ l NEB Puffer 4 (New England Biolabs)

1 μ l T4-DNA-Ligase (30 U/ μ l, Fermentas)

1 μ l *Bsa*I (10 U/ μ l, New England Biolabs)

ad 20 μ l (H₂O)

Die Reaktion erfolgte wie bei der Erstellung der Multi-repeat-Module mit folgendem Programm:

15 Zyklen: 37°C 5 min

16°C 5 min

16°C 20 min

50°C 30 min

80°C 20 min

4°C ∞

Der 30 min 50°C Schritt wurde eingeführt um abschließend die Restriktion durch *BsaI* zu fördern, sodass die eingesetzten Module, die die gleiche Resistenz wie der Zielvektor trugen, möglichst nicht religieren konnten.

Ein Teil des Ansatzes wurde dann in chemisch kompetente *E. coli* TOP10 transformiert. Die Selektion und Kontrolle positiver Klone erfolgte je nach Konstrukt mit Blau-Weiß-Selektion, Kolonie-PCR mit geeigneten Primern und Restriktionsverdau, gefolgt von der Sequenzierung der Plasmide.

2.2.12 Transformation chemisch kompetenter *E. coli*

Chemisch kompetente *E. coli*-Zellen (Top10) wurden zur Transformation auf Eis aufgetaut. 5 bis 10 µl des jeweiligen Reaktionsansatzes wurden mit 50 µl der Zellen gemischt und 30 min lang auf Eis belassen. Der anschließende Hitzeschock bei 42°C wurde für 45 s durchgeführt und die Bakterien zur einstündigen Erholung bei 37°C mit 250 µl LB-Medium verdünnt. Je 200 µl der Suspension wurden dann auf einem LB-Festmedium mit entsprechendem Antibiotikum zur Selektion ausplattiert.

2.2.13 Transformation elektrisch kompetenter Bakterienzellen

Die Transformation von 2 µl Plasmid-DNA in 50 µl elektrisch kompetenter *E. coli* (B121 AI) oder *A. tumefaciens* (GV3101) erfolgte in 1 mm Küvetten bei einer Spannung von $U=2500$ V und einem Widerstand von $R=200$ Ω. Die Bakterienzellen wurden anschließend in 250 µl LB-Medium (*E. coli*) oder 400 µl YEB-Medium (*A. tumefaciens*) aufgenommen und 1h bei 37°C (*E. coli*) bzw. 30°C (*A. tumefaciens*) inkubiert. Danach wurden die Bakteriensuspensionen auf selektives Festmedium ausplattiert und über Nacht bei 37°C bzw. zwei Tage bei 30°C angezogen.

2.2.14 Sequenzierung von DNA

Sequenzdaten wurden von einem entsprechenden Dienstleister (Microsynth Seqlab Göttingen GmbH, Göttingen) bezogen. Dazu die DNA mit dem Entsprechenden Oligonukleotid nach Angaben des Dienstleisters eingeschickt und nach dem Protokoll *Extended HotShot* vom Dienstleister bearbeitet.

2.2.15 Transfektion humaner Zellen mit Plasmiden

Die Transfektion von Hek293T-REx- und Huh7 Zellen erfolgte mit Turbofect *in vitro Transfection Reagent* (Fermentas GmbH, St. Leon-Rot, Deutschland). Dabei handelt es sich um eine sterile Lösung eines kationischen Polymers, das kompakte, stabile, positiv geladene Komplexe mit DNA formt, diese vor Degradation schützt und die Aufnahme in eukaryotische Zellen vermittelt.

Für Experimente wurde zur Zellaussaat in Multiwell-Platten (Corning, Amsterdam, Niederlande) $\frac{1}{4}$ des jeweiligen Kulturvolumens an Zellsuspension eingesetzt. Transfiziert wurden die Zellen 24 h nach Passage. Dazu wurden die zu transfizierenden Plasmide mit dem hier verwendeten Grundmedium DMEM in hundertfachem Volumen des eingesetzten Transfektionsreagenz verdünnt. Nach Vermischung mit dem Transfektionsreagenz inkubierte das Gemisch 20 min bei Raumtemperatur und wurde dann der jeweiligen Kultur zugefügt, die dann mindestens weitere 24 h bis zur Verwendung in Experimenten im Brutschrank verbrachte. Für die Transfektion von Zellen in verschiedengroßen Kulturen wurden unterschiedliche Mengen Reporter- und/oder Expressionskonstrukt, DMEM-Grundmedium und dem entsprechend Transfektionsreagenz, i.d.R. wie in Tabelle 2.8 aufgeführt, eingesetzt.

Tab. 2.8: Zusammensetzung der Transfektionsansätze für unterschiedliche Kulturdimensionen

Kultur ^(a)				Transfektion			
Gefäß	d	A	V	DMEM	Turbofect	DNA R. ^(b)	DNA gesamt
12 Well-Platte	2,2 cm	3,8 cm ²	2 ml	400 µl	4 µl	200 ng	2 µg
24 Well-Platte	1,56 cm	1,9 cm ²	1 ml	200 µl	2 µl	100 ng	1 µg
48 Well-Platte	1,1 cm	0,95 cm ²	0,5 ml	100 µl	1 µl	50 ng	0,5 µg

(a) d und A sind Durchmesser und Wachstumsfläche der eingesetzten Kulturgefäße, in denen Überstandskulturen mit dem Kulturmediumvolumen V angelegt wurden. Die zur Transfektion eingesetzten Mengen Plasmid-DNA und ggf. der (b) Anteil an Reporterkonstrukt sind angegeben.

2.2.16 Transkriptionsstartbestimmung mittels 5'-RACE

Das 5'-RACE (*rapid amplification of cDNA ends*) zur Transkriptionsstartpunktbestimmung wurde mit dem BD SMARTer™RACE cDNA Amplification Kit (Clontech Laboratories GmbH, Heidelberg) dem Herstellerprotokoll folgend durchgeführt. Geeignete cDNA wurde unter Verwendung der zugehörigen Oligonukleotide 5'-RACE CDS Primer A (ein dT-stretch Oligonukleotid) und SMART II A nach Herstellerprotokoll synthetisiert. Die eigentliche 5'-RACE-PCR wurde mit dem sogenannten Universal Primer A Mix (UPM) und dem genspezifischen Oligonukleotid Luc_973_R_AR nach Herstellerprotokoll für Anlagerungstemperaturen von über 70°C durchgeführt. Dabei entstammten einige Komponenten nicht dem Kit. Dies waren der 10x PCR-Puffer, die verwendete 2 mM dNTP-Lösung, sowie die rekombinante *Taq*-DNA-Polymerase (Fermentas, St. Leon Rot) von der 1 U eingesetzt wurde. Als Positivkontrolle diente eine PCR-Reaktion mit den genspezifischen Oligonukleotiden LuciferaseATG und Luc_321_R_AR. Die Anlagerungstemperatur in dieser Reaktion betrug 68°C. In 25 Zyklen wurde je 1 min amplifiziert. Eine zweite geschachtelte (*nested*) PCR wurde analog der 5'-RACE-PCR mit 2,5 µl deren Produktes als Matrize sowie 10 pmol des Nested Universal Primer A (NUP) und des genspezifischen Oligonukleotides Luc_321_R_AR angesetzt. Die Reaktion erfolgte bei einer Anlagerungstemperatur von 65°C und 1 min Elongationszeit in 25 Zyklen. Das Produkt wurde über ein Agarosegel gereinigt und in pCR2.1®-TOPO® kloniert. Die erhaltenen Klone wurden anschließend sequenziert.

2.2.17 T7 Endonuklease I-Verdau zur Detektion von Mutationen

Zur Detektion von TALEN-vermittelten Mutationen wurde zunächst der von den TALEN angesteuerte genomische Lokus an 100 ng genomischer DNA von TALEN-transfizierten Hek293T-Rex-Zellen mit der Phusion-Polymerase (Thermo Scientific) in einem Gesamtvolumen von 50 µl amplifiziert. 5 µl der PCR-Reaktion wurden mittels Agarose-Gelelektrophorese auf die spezifische Amplifikation eines PCR-Produktes überprüft, bevor der verbleibende PCR-Ansatz mit dem *GeneJET PCR Purification Kit* (Thermo Scientific) gereinigt wurde. 200 ng des gereinigten PCR-Produktes wurden mit 1 µl NEB Buffer 2 (New England Biolabs) in einem 9,5 µl Ansatz wie folgt in einem Thermocycler denaturiert und rehybridisiert:

5 min	95°C
	-2°C/s (Abkühlrate)
1 s	85°C

$$\begin{array}{l} -0.1^{\circ}\text{C/s (Abkühlrate)} \\ \infty \quad 25^{\circ}\text{C} \end{array}$$

Danach wurden 0,5 μl T7 Endonuklease I (T7EI; New England Biolabs) hinzugefügt und der Reaktionsansatz für 15 min bei 37°C im Thermocycler inkubiert. War das PCR-Produkt ein Gemisch aus mutierten und Wildtyp-Sequenzen, so haben sich bei der Denaturierung und Rehybridisierung DNA-Heteroduplexe gebildet, deren lokal einzelsträngige Bereiche nun von der T7EI erkannt und geschnitten werden können. Nach der Inkubation wird die Reaktion zum Abstoppen direkt auf Eis transferiert und/oder mit EDTA-haltigem Ladepuffer versetzt. Die Reaktion wird zur Analyse über ein 1,5%-iges Agarosegel elektrophoretisch aufgetrennt. Der Anteil geschnittenen PCR-Produktes (x_{cut}) wird ggf. mit ImageJ ermittelt und der Prozentsatz der Modifikation ($y_{mod} [\%]$) nach Guschin *et al.* (2010) berechnet:

$$y_{mod} [\%] = 100 \times (1 - \sqrt{1 - x_{cut}})$$

2.3 Biochemische Methoden

2.3.1 Quantitative Analyse der Luciferaseaktivität

Für Reporter-genaktivierungsstudien wurden Hek293T-REx-Zellen verwendet, die 24 h nach Passage mit einem Luciferasereporterkonstrukt und Expressionskonstrukten entsprechend Tabelle 2.8 co-transfiziert wurden und weitere 24 h zur transienten Genexpression inkubierten.

Für die Bestimmung der Luciferaseaktivität wurde das *Luciferase Assay System* (E1500, Promega) verwendet. Die Zellen wurden nach Herstellerangaben mit CCLR (*cell culture lysis reagent*) direkt im Kulturgefäß lysiert. Dabei wurden zur Bedeckung der Zellen 200 μl CCLR auf 1,9 cm^2 eingesetzt. Die erhaltenen, auf Eis gelagerten Lysate wurden nach der vom Hersteller empfohlenen Klärung durch Zentrifugation (2 min, 12000 \times g, 4°C) zur Messung der Lumineszenz am Sirius Single Tube Luminometer (Berthold Detection Systems) eingesetzt. Dazu wurden 20 μl Luciferase Assay Reagent (LAR), das Luciferin als Enzymsubstrat enthält, in konischen 11x70 mm Rotilabo-Reagenzröhrchen aus Polysterol (Carl Roth GmbH & Co., Karlsruhe) vorgelegt und 4 μl Lysat hinzugefügt. Kurzes Vortexen diente der Durchmischung. Das Reagenzröhrchen mit Reaktionsansatz wurde nun zügig in den Mess-Schacht eingeführt und eine Messung über eine Dauer von 10 s nach 2 s Verzögerung entsprechend der Angaben zum *Luciferase Assay System* durchgeführt. Die erhaltenen Messwerte wurden direkt vom Luminometer ausgedruckt und konnten zur Berechnung der Luciferaseaktivität verwendet

werden. Dazu wurde das Gesamtprotein im Lysat nach Bradford bestimmt und in folgende Formel eingesetzt:

$$A_{Luc} = \frac{P_{Lum}}{m_{GP}}$$

P_{Lum} = Lumineszenz als Leistung in relativen Lichteinheiten pro Sekunde (RLE/s)

m_{GP} = Masse Gesamtprotein im eingesetzten Lysat in μg

A_{Luc} = Luciferaseaktivität in RLE/s/ μg Gesamtprotein

Die hier bestimmte Luciferaseaktivität stellt die mit der Substratumsetzung korrelierende Lichtleistung des Gesamtproteins dar, die von der enthaltenen Luciferase erbracht wird, und kann somit als Maß für den Anteil der Luciferase am Gesamtprotein betrachtet werden. Die Produktion der Luciferase in den Zellen ist dabei von der Transkription abhängig, die durch die TAL-Effektoren beeinflusst werden soll. Die Auswertung der Daten erfolgte mit Microsoft Excel.

2.3.2 Untersuchung der subzellularen Lokalisation

Um zu überprüfen, ob die zur Expression in humanen Zellen erstellten TALEN entsprechend der beabsichtigten Funktion im Zellkern lokalisieren, wurde ein immunhistochemisches Verfahren angewandt. Dazu wurden Huh7-Zellen in 48 Well-Platten überführt, die zuvor unter sterilen Bedingungen mit Deckgläschen bestückt wurden. 24 h nach Zellaussaat wurde nur ein TALEN-Expressionsvektor entsprechend Tabelle 2.8 transfiziert. Weitere 24 h später wurden die Deckgläschen den Kulturgefäßen entnommen und in eine 48 Well-Platte (Corning, Amsterdam, Niederlande) mit 1x PBS zum Entfernen von Mediumresten überführt. Nach Abnahme des Waschpuffers wurden die Zellen auf dem Deckgläschen mit 4% Paraformaldehydlösung für 20 min fixiert. Nach ein- bis zweimaligem Waschen mit PBS wurden die Zellen mit 0,5% Triton x 100 1x PBS für 5 min permeabilisiert. Ein weiterer Waschschriff mit 1x PBS folgte. Die Zellpräparate, an denen die Lokalisation von FLAG-Fusionsproteinen untersucht werden sollte, wurden einer Antikörperbehandlung unterzogen. Das 1x PBS des letzten Waschschriffes wurde abgenommen und zum einstündigen Blockieren gegen 10% FCS 1x PBS ersetzt. Es wurde unter leichtem Schwenken bei Raumtemperatur inkubiert. Im Anschluss erfolgte die Behandlung mit dem ersten Antikörper gegen das Epitop. Es wurde monoklonaler, muriner anti-FLAG® M2 Antikörper (Stratagene) 1:1000 in 3% FCS 1x PBS verdünnt verwendet. Die Präparate wurden über 1 h mit der Antikörperlösung bei 37°C geschwenkt. Es folgte zweimalig 5 min Waschen mit 3% FCS 1x PBS bei Raumtemperatur auf

einem auf niedrige Frequenz eingestellten Schüttler. Die Inkubation mit dem zweiten Antikörper erfolgte 1 h in analoger Weise wie mit dem Ersten. Da es sich um einen mit dem Fluoreszenzfarbstoff Carbocyanin 3 (Cy3) konjugierten anti-Maus IgG (aus Ziege, Dianova), der ebenfalls 1:1000 verdünnt wurde, handelte, wurde dunkel inkubiert. Die Präparate wurden dann zweimal 5 min mit 1x PBS gewaschen (schwenkend bei Raumtemperatur) und auf Objektträger aufgebracht. Dazu wurden die Deckgläschen dem Gefäß entnommen, zweimal kurz in Wasser und zweimal kurz in vergällten Ethanol getaucht, getrocknet und die Zellseite auf Objektträgern mit ProLong® Gold *antifade reagent with DAPI* (4',6-Diamidin-2-phenylindol, Invitrogen, Karlsruhe, Deutschland) eingebettet. Alle Präparate wurden über Nacht bei Raumtemperatur getrocknet. Die Prozedur kann nach der Fixierung unterbrochen und das Material in 1x PBS im Kühlschrank gelagert werden. Die Präparate wurden am Fluoreszenzmikroskop untersucht.

2.3.3 Expressionsanalysen mittels Immunoblot

Zur Untersuchung der Expression der TALE-Derivate in humanen Zellen oder *N. benthamiana* wurden Immunoblots durchgeführt. Dafür wurden mit CCLR (Promega *Luciferase Assay System*) hergestellte Lysate von Hek293T-REx-Zellen oder mit GUS-Extraktionspuffer hergestellte Pflanzenextrakte verwendet. Das auf Eis gehaltene Lysat wurde nach Proteinbestimmung nach Bradford für die SDS-PAGE (Sodiumdodecylsulfat-Polyacrylamidgelelektrophorese) mit Lämmli gepuffert.

2.3.3.1 SDS-PAGE

Mit der SDS-PAGE können Proteine nach Molekulargewicht aufgetrennt werden. Dies wird durch das in Probenpuffer (Lämmli) und Gel enthaltene SDS ermöglicht, dass den Proteinen durch Anlagerung eine negative Gesamtladung verleiht. Für eine höhere Trennleistung wurden diskontinuierliche Gele aus 5%igem Sammel- und 8%igem Trenngel hergestellt (Tab. 2.9) und verwendet. Die Gele wurden in vertikalen Apparaturen für Midi- oder Maxi-Gele (Amersham Biosciences, Freiburg) eingespannt. Als Größenstandard diente die PageRuler™ Prestained Protein Ladder (Fermentas). Die aufgetragenen Probenvolumina wurden anhand der bestimmten Gesamtproteinmenge angeglichen und dem Versuch angepasst. Die Elektrophorese fand bei 200 V und 20 mA pro Midi- und bei 30 mA pro Maxi-Gel in 1 x TANK-Puffer statt.

Tab. 2.9: Zusammensetzung der SDS-Polyacrylamidgele

Bestandteil ^(a)	Endkonzentration/Anteil	
	5%iges Sammelgel	8%iges Trenngel
30% Acrylamid-Mix	17% (v/v)	27% (v/v)
Tris HCl (Gel-Puffer)	130 mM (pH 6,8)	375 mM (pH 8,8)
SDS	0,1% (w/v)	0,1% (w/v)
APS	0,1% (w/v)	0,1% (w/v)
TEMED	0,1% (v/v)	0,06% (v/v)

(a) Gel-Puffer siehe 2.1.7; APS = Ammoniumpersulfat; TEMED = N,N,N',N'-Tetramethyldiamin

2.3.3.2 Immunoblot

Rekombinante Proteine wurden im Immunoblot nachgewiesen. Die elektrophoretisch aufgetrennten Proteine wurden im Tank-Blot Verfahren auf eine Nitrocellulose Membran („Whatman“ Protran[®], Schleicher und Schuell GmbH) bei 4°C für 90 min und 95V übertragen. Dazu wurde Transferpuffer verwendet. Danach inkubierte die Membran für 1 Stunde unter leichtem Schwenken in Blockierungslösung. Die anschließende Inkubation der Membran mit dem Primärantikörper erfolgte über Nacht bei 4°C bzw. 2 h bei Raumtemperatur. Benutzt wurde für das jeweilige Epitop ein entsprechend in TBST verdünnter Antikörper (Tab. 2.10). Die Membran wurde nach Inkubation mit dem primären Antikörper dreimal für 5 bis 10 min mit 1x TBST gewaschen und im Anschluss 1 h bei Raumtemperatur mit dem sekundären Antikörper inkubiert. Es wurde das jeweils entsprechende *hrp*-Konjugat verdünnt in TBST verwendet. Folgend wurde die Membran erneut 3x für 10 min mit 1xTBST gewaschen. Detektiert wurde mittels der *enhanced chemiluminescence* (ECL; Amersham Biosciences, Freiburg) Methode durch 1 min Inkubation der Membran mit der Detektionslösung und anschließender Belichtung und Entwicklung eines Biomax Light 1-Film (Eastman Kodak Company, Rochester, USA).

Tab. 2.10: Für Immunoblots verwendete Antikörper

Antikörper	Konzentration/Verdünnung der Gebrauchslösung
Primärantikörper	
Anti-GFP (Kaninchen; Life Technologies)	1 µg/ml
Anti-FLAG (M2, Maus; Sigma-Aldrich)	1 µg/ml
Anti-c-Myc (A-14, Kaninchen; Santa Cruz)	1 µg/ml
Anti-HA (3F10, Ratte; Sigma-Aldrich)	100 ng/ml
Sekundäre Antikörper	
Anti-Kaninchen IgG (GE Healthcare)	1:10000
Anti-Maus IgG (GE Healthcare)	1:10000
Anti-Ratte IgG (Sigma-Aldrich)	1:10000

2.3.4 Proteinbestimmung nach Bradford

Zur Bestimmung des Gesamtproteins in Zelllysaten und Pflanzenextrakten wurde mit *Roti*[®]-*Quant* (Carl Roth GmbH & Co.; Karlsruhe) eine Proteinbestimmung nach Bradford (Bradford, 1976) durchgeführt. In dem 1 ml Ansatz wurden dabei 20% (v/v) des Reagenz eingesetzt. 4 µl Zelllysat oder 10 µl Pflanzenextrakt in aqua dest. wurden in der Messküvette vorgelegt. Nach Zugabe und Vermischung des Reagenzes inkubierten die Ansätze für 20 min bei Raumtemperatur. Bei einer Wellenlänge von 595 nm wurde die Absorption gegen einen Nullstandard (20% (v/v) *Roti*[®]-*Quant* in aqua dest.) gemessen und aufgezeichnet. BSA (Molecular Biology Grade, New England Biolabs, Schwalbach/Taunus)-Standards in aqua dest. dienten zur Erstellung einer Standardkurve der Proteinkonzentration gegen die Absorption. In Microsoft Excel wurde die Proteinkonzentration aus der Absorption anhand der Standardkurve berechnet.

2.3.5 Proteinreinigung aus *E. coli*

Zur Synthese und Reinigung von His₆-markierten TALE-Derivaten wurden die Expressionsplasmide in *E. coli* BL21-AI transformiert. Transformanten wurden über Nacht in LB-Flüssigmedium mit dem entsprechenden Antibiotikum bei 37 °C schüttelnd angezogen (20 ml Vorkultur). Am folgenden Tag wurden Hauptkulturen (500 ml LB-Flüssigmedium mit Antibiotikum) mit einer optischen Dichte von OD₆₀₀ = 0,1 angeimpft und schüttelnd bei 37 °C bis zu einer OD₆₀₀ von 0,6 - 0,8 vermehrt. Die TALE-Expression wurde durch Zugabe von 0,2% (w/v) L-Arabinose und 3% (v/v) Ethanol (Endkonzentrationen) induziert und erfolgte für zwei Stunden bei Raumtemperatur. Die *E. coli* Zellen wurden durch Zentrifugation (8000 rpm; 10 min; 4 °C) pelletiert und in 20 ml Lysispuffer (50 mM Tris; 300 mM NaCl, 10 mM Imidazol; pH 8,0; 8 °C) mit Proteinase-Inhibitor (1 Tablette/10 ml complete, EDTA-free; Roche) resuspendiert. Im Anschluss wurden die Zellen für 10 Minuten auf Eis mit Lysozym (Spatelspitze, Sigma Aldrich) inkubiert und durch drei Gefrier-Tau-Zyklen lysiert. Zur Zerstörung der Nukleinsäuren erfolgte eine einstündige Inkubation mit RNase A und DNase I (Endkonzentration 20 µg/ml) auf Eis, gefolgt von einem einstündigen Zentrifugationsschritt (20000 x g; 4°C). Die im Zelllysat befindlichen His₆-markierten Proteine wurden über eine Ni-NTA-Matrix (QIAGEN) immobilisiert. Nach dreimaligem Waschen der Ni-NTA-Matrix mit Waschpuffer (50 mM Tris; 300 mM NaCl, 20 mM Imidazol; pH 8,0; 8 °C) wurden die immobilisierten Proteine mit 2 ml Elutionspuffer (50 mM Tris; 300 mM NaCl, 250 mM Imidazol; pH 8,0; 8 °C) eluiert. Der gereinigten Proteine wurden mittels Amicon Ultra 15 ml Zentrifugenfiltern (Merk Millipore, Darmstadt) in Proteinlagerungs-Puffer (50 mM

Natriumphosphat; 300 mM NaCl; 10 % (v/v) Glycerin; pH 7,0) überführt, aliquotiert, in Flüssigstickstoff eingefroren und bei -20 °C gelagert. Die Reinheit der Proteine wurde mittels SDS-PAGE und anschließender Coomassie-Färbung analysiert. Die Proteinkonzentrationen wurden nach Bradford bestimmt.

2.3.6 EMSA

Für den *electrophoretic mobility shift assay* (EMSA) wurden die gereinigten His₆-markierten TALE-Derivate zunächst mit 10 mM DTT (Dithiothreitol) über Nacht bei 8 °C reduziert. Die Biotin-markierten bzw. unmarkierten, einzelsträngigen Oligonukleotide wurden um doppelsträngige DNA zu erhalten zu gleichen Teilen gemischt, für 5 Minuten bei 95 °C denaturiert und über Nacht langsam im Heizblock auf Raumtemperatur abgekühlt. Am Folgetag wurden die Verdünnungen der reduzierten Proteine sowie der nicht markierten doppelsträngigen DNA-Fragmente für Wettbewerbsanalysen mittels Proteinlagerungs-Puffer mit 10 mM DTT bzw. aqua dest. erstellt. Zunächst wurden für Wettbewerbsanalysen die TALEs mit einem Überschuss an unmarkierter DNA im EMSA-Puffer (100 mM KCl; 10 mM HEPES pH 7,5; 5 mM MgSO₄; 5 % (v/v) Glycerin; 10 mM DTT) für 10 min bei Raumtemperatur inkubiert. Danach wurden die Biotin-markierten DNA-Fragmente hinzugefügt und die Inkubationszeit um weitere 30 min verlängert. Die gebildeten DNA-Protein-Komplexe wurden im Anschluss über 5 %-ige Polyacrylamidgele im EMSA-Laufpuffer (0,5 x TBE; 1 % (v/v) Glycerin) bei 8 °C aufgetrennt. Danach wurden die DNA-Fragmente mittels Tank-Blot (0,5 x TBE, 100 V, max. Ampere, 30 Minuten, 8 °C) auf eine positiv geladene Nylonmembran überführt. Die Membran wurde daraufhin für eine Stunde bei 99 °C gebacken. Die Detektion der biotinylierten DNA erfolgte mittels *LightShift Chemiluminescent EMSA Kit-Chemiluminescent Nucleic Acid Detection Module* (Thermo Scientific) entsprechend den Herstellerangaben.

2.3.7 Analyse der β -Glucuronidase Aktivität

Für β -Glucuronidase (GUS)-Reporterstudien wurden *A. tumefaciens* GV3101-Stämme mit den entsprechenden GUS-Reporter- und Proteinexpressions-Konstrukten in AIM (*Agrobacterium*-Infiltrationsmedium: 10 mM MgCl₂; 5 mM MES pH 5,3; 150 μ M Acetosyringon) resuspendiert und auf eine OD₆₀₀ von 0,8 eingestellt. Jeweils ein *Agrobacterium*-Stamm mit einem GUS-Reporterkonstrukt wurde dann 1:1 mit einem *Agrobacterium*-Stamm mit einem Expressionskonstrukt gemischt und in die Blätter von drei 5-7 Wochen alten *N. benthamiana*-Pflanzen inokuliert. Die GUS-Reporter- und Proteinexpressions-Konstrukte werden dann als

Transfer-DNAs von *A. tumefaciens* in die Pflanzenzellen übertragen. Zwei Tage nach Inokulation erfolgte die Ernte der Blattscheiben mit einem Korkbohrer (Durchmesser 0,9 cm).

2.3.7.1 Quantitative GUS-Bestimmung

Zwei Blattscheiben wurden zusammen mit zwei Metallkugeln in *Safelock* Eppendorf-Reaktionsgefäße überführt und in flüssigem Stickstoff eingefroren. Der Aufschluss des tiefgefrorenen Blattmaterials erfolgte mit Hilfe einer Kugelmühle (Retsch GmbH; Haan) für 40 s bei einer Frequenz von 30 Hz. Das Pflanzenmaterial wurde im Anschluss mit 300 µl GUS-Extraktionspuffer (50 mM Na₂HPO₄ (pH 7,0); 10 mM EDTA; 10 mM β-Mercaptoethanol; 0,1 % Triton-X100; 0,1 % SDS) versetzt und 5 min bei 14.000 rpm und 4°C zentrifugiert. 200µl Überstand wurden in eine 96 Well-Platte übertragen, nochmals zentrifugiert (4.000 rpm; 5min; 4°C) und 100 µl in eine neue PCR-Platte überführt. In zwei technischen Replikaten wurden je 10 µl Probe in neue PCR-Platten pipettiert und mit 90µl GUS-Extraktionspuffer mit 5 mM MUG (4-Methylumbelliferyl-β-D-glucuronid; AppliChem GmbH, Darmstadt) gemischt. Es folgte eine Inkubation bei 37°C für 60-90 min. Zum Abstoppen der Reaktion wurden 10 µl von dieser mit 90 µl 0,2 M Natriumcarbonat versetzt. Die Fluoreszenz (Anregung: 360 nm; Emission: 465 nm) des Reaktionsproduktes 4-Methylumbelliferon (4-MU) wurde anschließend in schwarzen Platten mit durchsichtigem Boden mit dem Spectrafluor (Tecan Group Ltd.; Mannedorf Schweiz) gemessen. Als Standard diente 10 µM 4-MU (Sigma-Aldrich Chemie GmbH, München). Abschließend Die Proteinbestimmung wurde nach Bradford durchgeführt. Für die Berechnung der GUS-Aktivität wurde folgende Formel in Microsoft Excel verwendet:

$$A_{GUS} = \frac{\Delta F \times 10 \text{ pmol MU}}{t[\text{min}] \times \text{Protein} [\text{mg}] \times F_{10 \text{ pmol MU}}}$$

A_{GUS} = GUS-Aktivität in pmol 4-MU/min/µg Protein

ΔF = Fluoreszenzwert der Probe abzüglich der Hintergrundfluoreszenz

t = Reaktionszeit in Minuten

$F_{10 \text{ pmol MU}}$ = Fluoreszenzwert des Standards (10 µM MU)

2.3.7.1 Qualitative GUS-Bestimmung

Bei der quantitativen Analyse der GUS-Aktivität wird das Substrats X-Gluc (5-bromo-4-chloro-3-indolyl-β-D-Glucuronsäure; Carl Roth GmbH & Co., Karlsruhe) durch die β-Glucuronidase hydrolysiert, was durch Reaktion mit dem Luftsauerstoff zur Blaufärbung führt. Die Intensität der Färbung ist demnach ein Maß für die GUS-Aktivität. Dazu wurde je eine

Blattscheibe je Inokulationsstelle geerntet (insgesamt 3 Blattscheiben von 3 Pflanzen). Das Blattgewebe wurde mit 1 ml GUS-Färbelösung (10 mM Natriumphosphat, pH 7,0; 10 mM EDTA, pH 8,0; 1 mM Kalium-ferricyanat; 1 mM Kalium-ferrocyanat; 0,1 % Triton X-100; 0,1% X-Gluc [5-bromo-4-chloro-3-indolyl-b-Dglucuronide]) bedeckt und mit einer kanülenloser Spritze vakuumfiltriert. Anschließend wurden die Blattscheiben über Nacht bei 37°C inkubiert. Danach wurden die Blattscheiben mit Ethanol entfärbt und zwischen Acetat-Folie getrocknet.

2.3.8 TALEN-Aktivitätsanalysen *in vitro*

Um die Aktivität von TALEN zu testen wird zunächst die Zielsequenz mittels PCR amplifiziert und in ein Plasmid kloniert. In dieser Arbeit wurden dafür die Vektoren pENTR/D-TOPO und pGGC verwendet. Anschließend wurde das jeweilige Plasmid mit der Zielsequenz linearisiert. Für pENTR-Derivate wurde die Restriktionsendonuklease Alw44I (Thermo Scientific) und für pGGC-Derivate ScaI (Thermo Scientific) verwendet. Die linearisierten Plasmide wurden dann mit Hilfe des *GeneJET PCR Purification Kits* (Thermo Scientific) aus der Reaktion gereinigt. Die TALEN wurden einzeln oder gemeinsam in einem 1:1 Gemisch von insgesamt 500 ng Expressionsplasmid mittels *TnT T7 Quick Coupled Transcription/Translation System* (Promega) *in vitro* exprimiert. Von der TnT-Reaktion, die die exprimierten TALEN enthält, wurden 4 µl direkt für den *in vitro*-Aktivitätstest verwendet. Jeder Restriktionsansatz enthielt weiterhin 200 ng des linearisierten Plasmides, 1 x NEBuffer 3 (New England Biolabs) und 2,5 µg/µl BSA (New England Biolabs) in einem Gesamtvolumen von 20 µl. Nach einer Inkubation für 1h bei 37°C wurden die Reaktionsansätze 20 min bei 65°C inaktiviert und anschließend für 3 min bei 16.000 x g zentrifugiert. 16 µl des entstandenen Überstandes wurden dann mittels Elektrophorese auf einem 1% Agarose-Gel analysiert.

3. Ergebnisse

3.1 TALE-*repeats* mit abweichender Aminosäuren-Anzahl

Im Folgenden wird untersucht, ob und welchen Einfluss aberrante *repeats* auf die DNA-Erkennung durch TALEs haben. Da für den natürlichen TALE AvrXa7, der einen aberranten *repeat* enthält, zwei zueinander verschobene Zielsequenzen im *OsSWEET14*-Promotor publiziert sind (Abb. 1.1 F; Antony *et al.* 2010; Römer *et al.*, 2010), wurden TALEs mit aberranten *repeats* auf ihre Fähigkeit untersucht, Leserasterverschiebungen zu erkennen. Zum Test wurden verschiedene Ziel-DNA-Sequenzen abgeleitet und 5' mit einem *Bs4*-Minimalpromotor fusioniert. Diese Zielpromotoren wurden entweder in linearisierten, pENTR/D-TOPO-basierten Plasmiden TALEN *in vitro* angeboten oder 5' eines GUS-Reportergens in einen T-DNA Vektor inseriert.

3.1.1 Ein aberranter *repeat* erlaubt die Erkennung einer -1 Nukleotid-Leserasterverschiebung in der Ziel-DNA-Sequenz *in planta* und *in vitro*

Zur Untersuchung aberranter *repeats* auf eine mögliche Erkennung von Ziel-DNA-Sequenzen (Boxen) mit Leserasterverschiebung wurden artifizielle TALEs und TALEN konstruiert, die neben einem aberranten *repeat* an Position 8 (p8) nur *repeats* mit eindeutiger Nukleobasen-Spezifität enthielten (Abb. 3.1.1 A). Dadurch sollten im Falle einer Leserasterverschiebung in der Ziel-DNA-Sequenz möglichst viele RVD-Base-Fehlpaarungen die Aktivität des jeweiligen TALE-Derivates inhibieren, sodass nur eine Veränderung des DNA-Erkennungsvermögens durch den jeweiligen aberranten *repeat* Aktivität wiederherstellen könnte. Zum Vergleich wurde stets ein entsprechendes TALE-Derivat (TALE/TALEN34) ohne aberranten *repeat* mitgeführt (Abb. 3.1.1 A). Als Ziel-DNA-Sequenzen wurden neben der optimalen Box DNA-Sequenzen angeboten, die -1, -2, +1 und +2 Nukleotide-Leserasterverschiebungen nahe der dem aberranten *repeat* korrespondierenden Position enthielten (Abb. 3.1.1 A). In GUS-Reporterstudien in *N. benthamiana* wurden artifizielle TALEs mit jeweils einem natürlich vorkommenden 30, 40 oder 42 *As-repeat* auf ihr Vermögen untersucht, GUS-Aktivität an den verschiedenen Boxen zu induzieren (Abb. 3.1.1 B). Darüber hinaus wurde der 40 *As-repeat* exemplarisch in TALEN-Aktivitätsanalysen *in vitro* getestet (Abb. 3.1.1 D, E). Dazu wurden die beschriebenen TALE-Derivate als *forward*-bindende TALEN eingesetzt und mit einer einheitlichen *reverse*-bindenden TALEN ohne aberranten *repeat* kombiniert. Die Plasmide mit den entsprechenden Boxen wurden vor der Inkubation mit den TALEN linearisiert, sodass TALEN-Aktivität durch zwei Restriktionsfragmente detektierbar ist. Als Spezifitätskontrolle

diente eine *forward* TALEN, die eine AGT2-Box genannte Sequenz erkennen und gemeinsam mit der *reverse*-bindenden TALEN schneiden kann (Abb. 3.1.1 D; Streubel *et al.*, 2012).

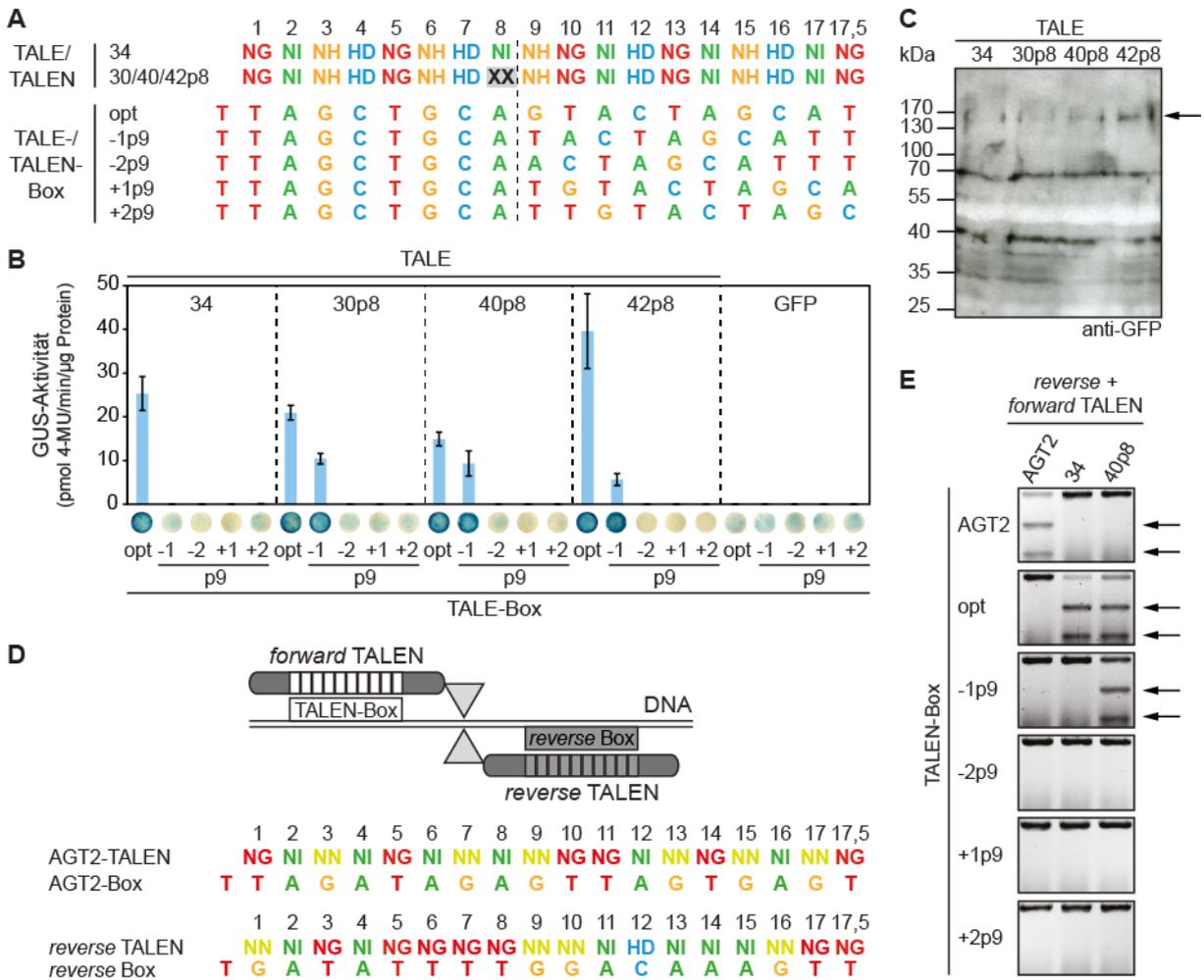


Abb. 3.1.1: TALE-Derivate mit aberrantem repeat tolerieren -1 Nukleotid-Leserasterverschiebungen.

(A) RVD-Abfolgen der verwendeten TALEs und TALEN, sowie DNA-Sequenzen der getesteten Boxen. Die TALE-Derivate enthalten entweder nur 34 As-repeats oder einen natürlichen aberranten repeat mit 30, 40 oder 42 As an Position 8 (p8, grau hinterlegt) und dem RVD XX=NI für den 30 und 42 As-repeat oder XX=NN im Falle des 40 As-repeat. Neben der optimalen Box (opt) wurden verschiedene Leserasterverschiebungen (-1, -2, +1, +2) an Position 9 (p9) angeboten, gekennzeichnet durch eine gestrichelte Linie. (B) Quantitative und qualitative GUS-Reporterstudien im Vergleich zu GFP in *N. benthamiana* wurden 48 h nach Inokulation durchgeführt ($n=3$). Fehlerindikatoren kennzeichnen Standardabweichungen im quantitativen Ansatz. Eine repräsentative Blattscheibe des qualitativen Ansatzes wird gezeigt. (C) Immunologischer Nachweis der GFP-gekoppelten TALEs 48 h nach *Agrobacterium*-vermittelter Expression in *N. benthamiana* (Richter *et al.*, 2014). Ein Pfeil markiert die detektierten TALEs. (D) Versuchsaufbau der TALEN-Aktivitätsanalysen *in vitro*, sowie RVD-Abfolgen und Boxen der AGT2- und der in allen Experimenten konstanten *reverse*-bindenden TALEN. Das TALEN-Paar schneidet die linearisierte DNA, wenn die Boxen von beiden TALEN erkannt werden. Dreiecke symbolisieren die C-terminal fusionierten *FokI*-Endonukleasedomänen. (E) TALEN-Aktivitätsanalysen *in vitro*. Nach einstündiger Restriktion wurden die Reaktionen inaktiviert und über ein 0,8%-iges Agarosegel aufgetrennt. Pfeile markieren die geschnittenen Fragmente.

Sowohl in *N. benthamiana* als auch *in vitro* zeigten die unterschiedlichen TALE-Derivate, artifizielle TALEs und TALEN, ein ähnliches Aktivitätsmuster. Dabei wird deutlich, dass im Vergleich zu TALE-Derivaten mit nur 34 *As-repeats* solche mit einem aberranten *repeat* neben der optimalen Box auch eine Ziel-DNA-Sequenz erkennen, die eine -1 Nukleotid-Leserasterverschiebung nahe der dem aberranten *repeat* korrespondierenden Position enthält (Abb. 3.1.1 A, B, E). Andere Leserasterverschiebungen wurden weder von artifiziellen TALEs noch TALEN toleriert (Abb. 3.1.1 A, B, E). Die vollständige Synthese der artifiziellen TALEs in *N. benthamiana* wurde mittels Immunoblot überprüft (Abb. 3.1.1 C). Wurde die AGT2-TALEN als *forward* TALEN verwendet, konnte nur Restriktion der DNA mit AGT2-Box nachgewiesen werden. Umgekehrt konnte keine andere *forward* TALEN die Restriktion dieser DNA vermitteln (Abb. 3.1.1 E). Die verwendeten TALEN zeigten also keine unspezifische Restriktion.

3.1.2 TALEs und TALEN mit aberrantem *repeat* tolerieren eine -1 Nukleotid-Leserasterverschiebung an verschiedenen Positionen der Ziel-DNA-Sequenz

Da sich einzelne aberrante *repeats* in natürlichen TALEs an zentralen Positionen in der *repeat*-Folge befinden, wurde in GUS-Reporterstudien und TALEN-Aktivitätsanalysen *in vitro* untersucht, ob auch mit einem aberranten *repeat* an einer mehr N- oder C-terminalen Position der *repeat*-Domäne ein funktionales TALE-Derivat erstellt werden kann, das -1 Nukleotid-Leserasterverschiebungen an entsprechenden Positionen akzeptiert. Dazu wurden zwei weitere Derivate vom Typ der in 3.1.1 vorgestellten TALEs und TALEN erstellt, die einen 40 *As-repeat* mit dem RVD NN, der Guanin und Adenin erkennt, an Position 3 (40p3) oder 14 (40p14) der *repeat*-Folge enthielten (Abb. 3.1.2 A). Diese wurden mit den TALE-Derivaten ohne und mit 40 *As-repeat* an Position 8 der *repeat*-Folge an der optimalen Box und -1 Nukleotid-Leserasterverschiebungen an den Positionen 4 (-1p4), 9 (-1p9) und 15 (-1p15) verglichen (Abb. 3.1.2 A). Artifizielle TALEs dieser Bauart wurden in GUS-Reporterstudien in *N. benthamiana* getestet (Abb. 3.1.2 B) und im Immunoblot nachgewiesen (Abb. 3.1.2 C). Für *in vitro*-Studien wurden analoge TALEN wie unter 3.1.1 beschrieben mit linearer DNA mit den entsprechenden Boxen getestet (Abb. 3.1.2 D). AGT2-Box und -TALEN wurden ebenfalls als Kontrollen für eine spezifische Restriktion mitgeführt (Abb. 3.1.1 D, 3.1.2 D).

Quantitative und qualitative GUS-Reporterstudien zeigten, dass auch artifizielle TALEs mit 40 *As-repeat* an verschiedenen nicht zentralen Positionen der *repeat*-Folge sowohl an der optimalen als auch einer korrespondierenden -1 Nukleotid-Leserasterverschiebungs-Box das Reportergen aktivieren konnten (Abb. 3.1.2 B). Dabei konnte jedoch eine deutlich geringere

Aktivität des TALE40p14 im Vergleich zu artifiziellen TALEs mit 40 *As-repeat* an anderer Position an der optimalen Box beobachtet werden (Abb. 3.1.2 B). Befand sich die -1 Nukleotid-Leserasterverschiebung nicht nahe der dem 40 *As-repeat* korrespondierenden Box-Position konnte in keinem Fall Reporteraktivität nachgewiesen werden (Abb. 3.1.2 B). Dieses positionsgebundene Verhalten lässt darauf schließen, dass der Mechanismus, der die -1 Nukleotid-Leserasterverschiebungstoleranz zulässt, unmittelbar an der Position des aberranten *repeats* im Array ausgelöst wird.

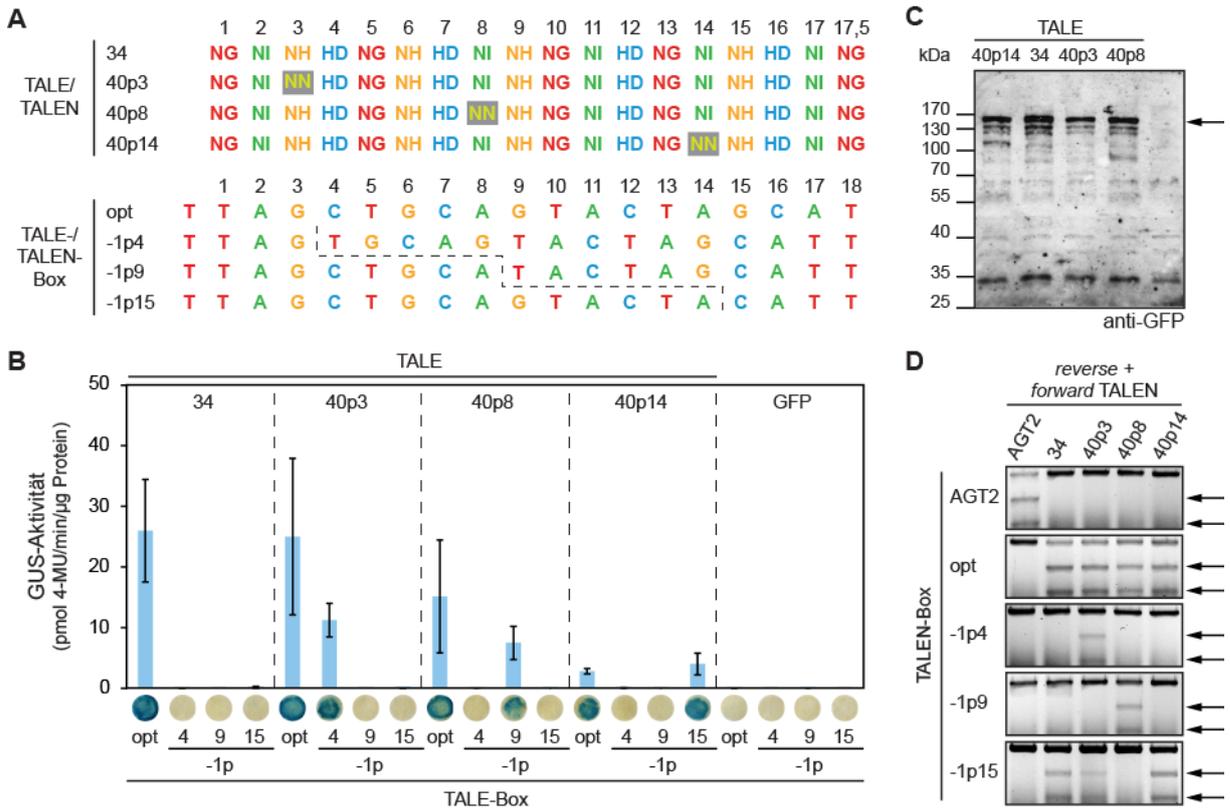


Abb. 3.1.2: Aberrante *repeats* funktionieren an verschiedenen Positionen der *repeat*-Folge.

(A) RVD-Abfolgen der verwendeten TALEs und TALEN, sowie DNA-Sequenzen der getesteten Boxen. Die TALE-Derivate enthalten entweder nur 34 *As-repeats* oder einen natürlichen aberranten *repeat* mit 40 *As* an Position 3, 8 oder 14 (40p3/p8/p14, grau hinterlegt) und dem RVD NN. Neben der optimalen Box (opt) wurden -1 Nukleotid-Leserasterverschiebungen an den Positionen 4, 9 und 15 (-1p4/p9/p15) angeboten, gekennzeichnet durch eine gestrichelte Linie. (B) Quantitative und qualitative GUS-Reporterstudien im Vergleich zu GFP in *N. benthamiana* wurden 48 h nach Inokulation durchgeführt ($n=3$). Fehlerindikatoren kennzeichnen Standardabweichungen im quantitativen Ansatz. Eine repräsentative Blattscheibe des qualitativen Ansatzes wird gezeigt. (C) Immunologischer Nachweis der GFP-gekoppelten TALEs 48 h nach *Agrobacterium*-vermittelter Expression in *N. benthamiana* (Richter *et al.*, 2014). Ein Pfeil markiert die detektierten TALEs. (D) TALEN-Aktivitätsanalysen *in vitro*. Nach einstündiger Restriktion wurden die Reaktionen inaktiviert und über ein 0,8%-iges Agarosegel aufgetrennt. Pfeile markieren die geschnittenen Fragmente.

Das in GUS-Reporterstudien ermittelte Erkennungsmuster ließ sich auch in TALEN-Aktivitätsanalysen *in vitro* nachweisen. Neben einer DNA, die die optimale Box enthielt wurden von den TALEN mit 40 *As-repeat* auch solche Boxen erkannt und geschnitten, die eine -1 Nukleotid-Leserasterverschiebung nahe der dem 40 *As-repeat* korrespondierenden Position enthielten (Abb. 3.1.2 D). Im Unterschied zum Aktivitätsmuster der artifiziellen TALEs in *N. benthamiana* (Abb. 3.1.2 B) zeigten TALEN34 und 40p3 auch Restriktionsaktivität an einer Box mit -1 Nukleotid-Leserasterverschiebung an Position 15 (Abb. 3.1.2 D). Da TALEN34 keinen aberranten *repeat* enthält und der 40 *As-repeat* in TALEN40p3 zu einer 12 Nukleotide stromaufwärts gelegenen Position korrespondiert, scheint diese beobachtete Aktivität nicht auf den Mechanismus für die -1 Nukleotid-Leserasterverschiebungserkennung zurückzuführen zu sein, sondern die Toleranz von RVD-Base-Fehlpaarungen widerzuspiegeln. Für beide TALEN würden drei Fehlpaarungen von RVD 15 bis 17 mit den Basen 15 bis 17 der Box -1p15 auftreten (Abb. 3.1.2 A). Diese Fehlpaarungen verhinderten eine GUS-Reportergenaktivierung durch entsprechende, artifizielle TALEs (Abb. 3.1.2 B TALE34, 40p3). Demnach tolerieren TALEN offenbar RVD-Base-Fehlpaarungen besser als TALEs. Dies könnte in der für ihre Aktivität notwendigen Dimerisierung der fusionierten *FokI*-Endonukleasedomänen liegen. Durch diese könnte die zweite hier perfekt zur ihrer Box passende TALEN des Paares wie ein Anker wirken und die *FokI*-Endonukleasedomänen für die Funktion ausreichend zur DNA bringen, obwohl die C-terminalen *repeats* der Partner-TALEN nicht in der großen Furche der DNA binden. Für die Funktion der TALE-AD scheint dies jedoch notwendig zu sein.

3.1.3 Ein aberranter *repeat* erlaubt die Erkennung einer Ziel-DNA-Sequenz bei Deletion des korrespondierenden oder eines benachbarten Nukleotides

Nachdem nun gezeigt wurde, dass TALEs mit einem aberranten *repeat* -1 Nukleotid-Leserasterverschiebungen, bzw. die Deletion eines Nukleotides, an verschiedenen Positionen der Ziel-DNA-Sequenz tolerieren können, wurde im Folgenden untersucht, an welcher Box-Position in Bezug auf den aberranten *repeat* die Toleranz einer Leserasterverschiebung am höchsten ist. Aus 3.1.2 ist bekannt, dass -1-Nukleotid-Leserasterverschiebungen nur in räumlicher Nähe zum 40 *As-repeat* akzeptiert werden, so wurde z.B. das GUS-Reportergen von TALE40p8 nicht aktiviert, wenn das Nukleotid an Box-Position 4 oder 15 deletiert ist (Abb. 3.1.2 A, B). Sowohl in diesem Experiment als auch in 3.1.1 wurden TALE-Derivate mit aberranten *repeats* mit -1 Nukleotid-Leserasterverschiebungsboxen getestet, in denen nicht das dem aberranten *repeat* korrespondierende, sondern das 3'-folgende Nukleotid deletiert ist. Hier wurden die artifiziellen TALEs ohne oder mit 30, 40 oder 42 *As-repeat* an Position 8 aus 3.1.1

an Boxen mit einer -1 Nukleotid-Leserasterverschiebung an Position 7, 8 oder 9 getestet (Abb. 3.1.3 A).

Quantitative und qualitative GUS-Reporterstudien in *N. benthamiana* ergaben, dass artifizielle TALEs mit aberrantem *repeat* die höchste GUS-Reporteraktivität an -1 Nukleotid-Leserasterverschiebungsboxen dann auslösen, wenn das dem aberranten *repeat* gegenüberliegende bzw. korrespondierende Nukleotid deletiert ist (Abb. 3.1.3 B). Für alle drei getesteten aberranten *repeats* an Position 8 konnte eine geringere GUS-Reporteraktivierung mit Boxen gezeigt werden, in denen das Nukleotid an Position 7 deletiert ist (Abb. 3.1.3 B). Für die beiden verlängerten 40 und 42 As-*repeats* konnte dies auch für die 3'-verschobene -1 Nukleotid-Leserasterverschiebung gezeigt werden (Abb. 3.1.3 B). Diese Ergebnisse legen nahe, dass der aberrante *repeat* beim Erkennen einer -1 Nukleotid-Leserasterverschiebung möglicherweise das kontinuierliche *repeat*-Array verlässt (aus-loop) und nicht mehr an der DNA-Erkennung teilnimmt.

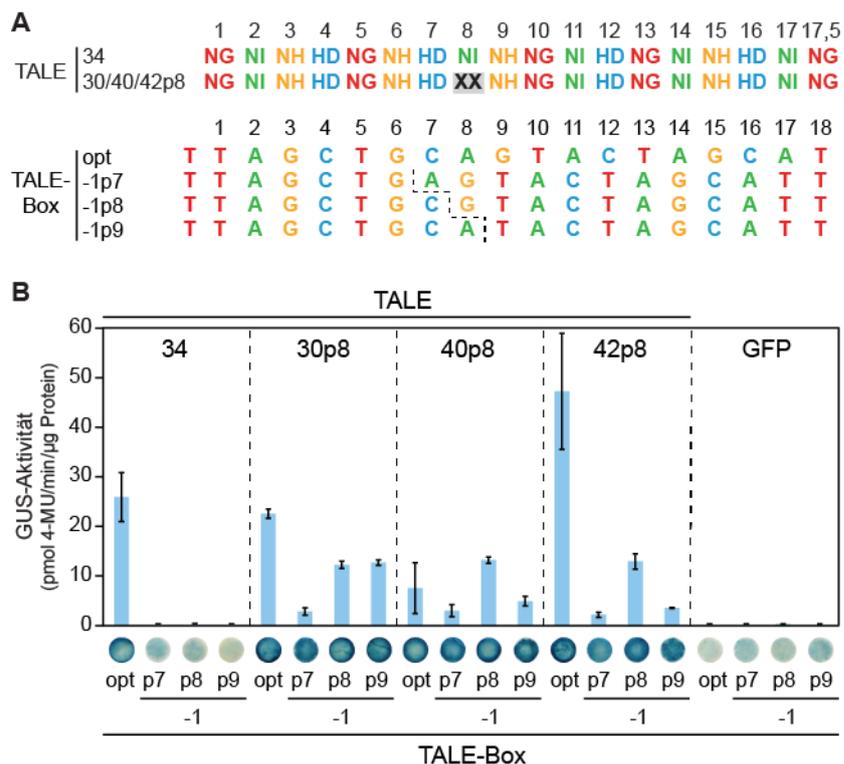


Abb. 3.1.3: Ein aberranter *repeat* kann die *repeat*-Folge verlassen. (A) RVD-Abfolgen verwendeter TALEs und DNA-Sequenzen getesteter Boxen. Die TALEs enthalten entweder nur 34 As-*repeats* oder einen aberranten *repeat* mit 30, 40 oder 42 As an Position 8 (grau) und dem RVD XX (NI bei 30 und 42 As, NN bei 40 As). Neben der optimalen Box (opt) wurden -1 Nukleotid-Lese-rasterverschiebungen an Position 7, 8 oder 9 angeboten (gestrichelte Linie). (B) Quantitative und qualitative GUS-Reporterstudien *N. benthamiana* im Vergleich zu GFP 48 h nach Inokulation ($n=3$). Fehlerindikatoren kennzeichnen Standardabweichungen. Eine repräsentative Blattscheibe wird gezeigt.

3.1.4 TALEs mit geringer *repeat*-Anzahl werden durch einen 40 *As-repeat* geschwächt

Da die bekannten, natürlichen TALEs mit einem aberranten *repeat* mindestens 15,5 *repeats* enthalten, wurde untersucht, ob TALEs mit weniger *repeats* und einem aberranten *repeat* funktional sind und eine -1 Nukleotid-Leserasterverschiebung tolerieren können. Um einen Vergleich zu ermöglichen wurden artifizielle TALEs erstellt, die innerhalb ihrer 11,5 *repeats* die gleichen RVDs enthalten wie die in 3.1.1 bis 3.1.3 verwendeten TALEs, und auf ihre GUS-Reporteraktivierung getestet (Abb. 3.1.4 A, B, D, E). Exemplarisch wurde ein aberranter 40 *As-repeat* entweder an Position 3 oder 8 der 11,5 *repeats* positioniert und mit einem entsprechenden 17,5 *repeat*-TALE und TALEs ohne aberranten *repeat* verglichen. Dazu wurden für den 40 *As-repeat* an Position 8 die Boxen mit lokal verschobener -1 Nukleotid-Leserasterverschiebung aus 3.1.3 verwendet (Abb. 3.1.4 A, B). Im Falle des 40 *As-repeat* an Position 3 wurde die TALE-Box -1p4 aus 3.1.2 und eine -1 Nukleotid-Leserasterverschiebungsbox an der dem 40 *As-repeat* korrespondierenden Position 3 genutzt (Abb. 3.1.4 D, E).

GUS-Reporterstudien haben gezeigt, dass bereits der 11,5 *repeat*-TALE (TALE34s) mit den hier verwendeten RVDs eine deutlich geringere Aktivität an der optimalen Box hat als ein vergleichbarer 17,5 *repeat*-TALE (TALE34) und 17,5 *repeat*-TALEs mit 40 *As-repeat* (TALE40p8/p3, Abb. 3.1.4 B, E). Ein 40 *As-repeat* an Position 8 eines 11,5 *repeat*-TALEs (TALE40p8s) führte sogar zu gänzlichem Aktivitätsverlust an optimaler und -1 Nukleotid-Leserasterverschiebungsboxen, an denen ein analoger 17,5 *repeat*-TALE noch Aktivität zeigt (Abb. 3.1.4 B). Die vollständige Synthese von TALE40p8s in *N. benthamiana* konnte allerdings im Immunoblot nachgewiesen werden, sodass von der Intaktheit des Proteins ausgegangen werden kann (Abb. 3.1.4 C). Im Falle eines 40 *As-repeats* an Position 3 in einem 11,5 *repeat*-TALE (TALE40p3s) konnte ein ähnliches Verhalten wie für den entsprechenden 17,5 *repeat*-TALE (TALE40p3) beobachtet werden, allerdings bei viel geringerer GUS-Reporteraktivierung (Abb. 3.1.4 E). Auch die Aktivität von TALE40p3s an der optimalen Box war deutlich schwächer als die von dem 11,5 *repeat*-TALE34s ohne aberranten *repeat*, aber vergleichbar der Aktivierung an der Leserasterverschiebungsbox -1p3 (Abb. 3.1.4 E). Auch der TALE40p3 mit 17,5 *repeats* aktiviert an der optimalen und der -1p3 Box mit gleicher Intensität. Im Gegensatz zu der Variante mit nur 11,5 *repeats* (TALE40p3s) ist TALE40p3 jedoch auch in der Lage GUS-Reporteraktivität an einer Box mit -1 Nukleotid-Leserasterverschiebung an Position 4 (-1p4) zu induzieren (Abb. 3.1.4). Unter der Annahme, dass der aberrante *repeat* die *repeat*-Folge bei Bindung einer Leserasterverschiebungsbox verlässt, bedeutet dies eine RVD-Base-Fehlpaarung (HD mit G, Abb. 3.1.4 D), die TALE40p3s

mit seinen 11,5 *repeats* offenbar nicht kompensieren kann (Abb. 3.1.4 E). Die korrekte Synthese von TALE40p3s *in planta* konnte im Vergleich zu dem ebenfalls 11,5 *repeats* enthaltenden TALE34s gezeigt werden (Abb. 3.1.4 F).

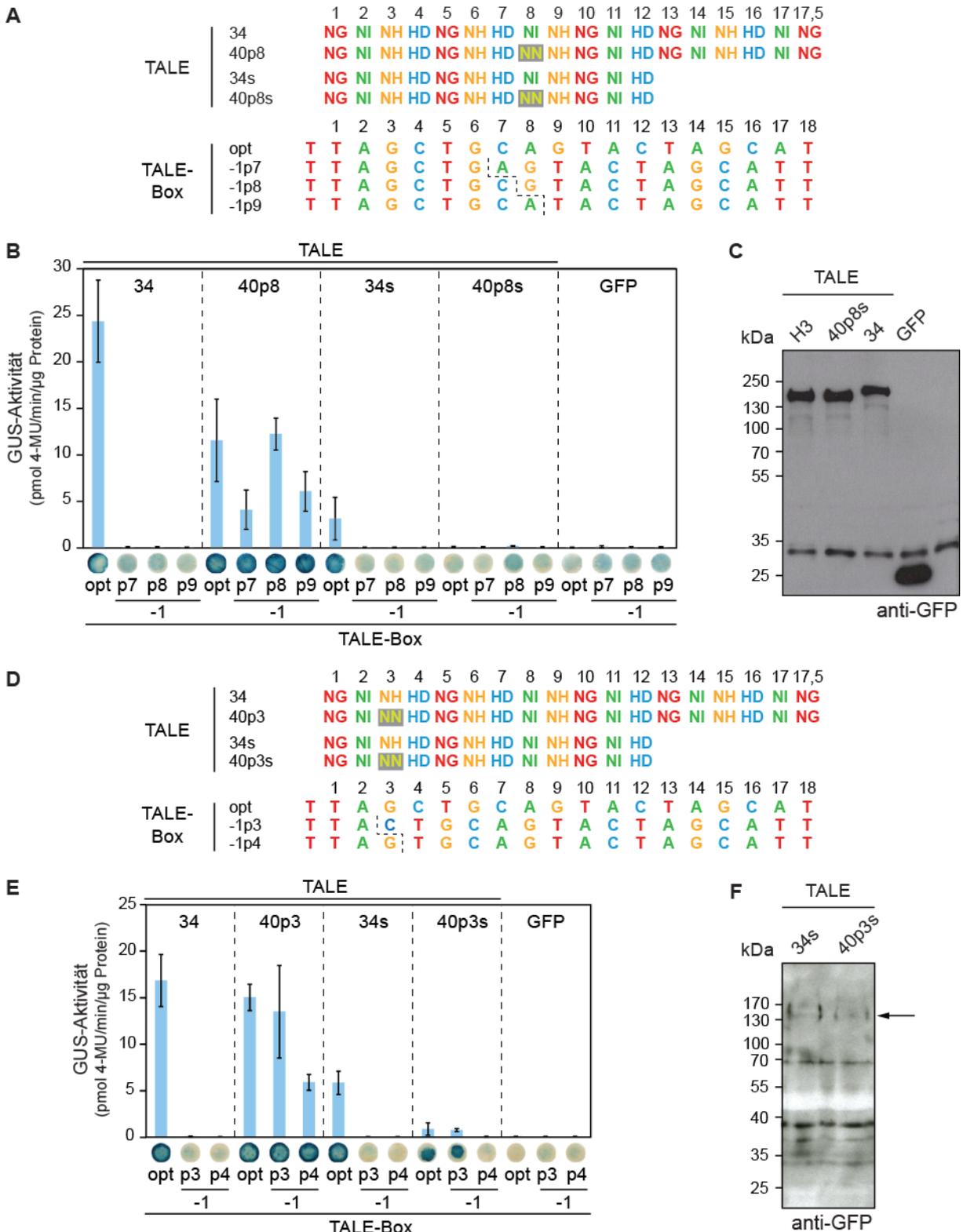


Abb. 3.1.4: Ein 40 *As-repeat* schwächt TALEs mit geringer *repeat*-Anzahl. (A) RVD-Abfolgen der in (B) verwendeten TALEs, sowie DNA-Sequenzen der getesteten Boxen. Die TALEs

enthalten entweder 17,5 oder 11,5 nur 34 *As-repeats* oder einen natürlichen aberranten *repeat* mit 40 *As* an Position 8 (grau hinterlegt) und dem RVD NN. Neben der optimalen Box (opt) wurden -1 Nukleotid-Leserasterverschiebungen an den Positionen 7, 8 und 9 (-1p7/p8/p9) angeboten, gekennzeichnet durch eine gestrichelte Linie. (B, E) Quantitative und qualitative GUS-Reporterstudien im Vergleich zu GFP in *N. benthamiana* wurden 48 h nach Inokulation durchgeführt ($n=3$). Fehlerindikatoren kennzeichnen Standardabweichungen im quantitativen Ansatz. Eine repräsentative Blattscheibe des qualitativen Ansatzes wird gezeigt. (C, F) Immunologischer Nachweis GFP-gekoppelter TALEs 48 h nach *Agrobacterium*-vermittelter Expression in *N. benthamiana*. (C) TALE40p8s wurde im Vergleich zu TALE34 mit 17,5 *repeats* und dem natürlichen 11,5 *repeat* TALE Hax3 (H3) mit dem Richtigen Molekulargewicht detektiert. Als Kontrollen diente GFP und nicht inokuliertes Pflanzenmaterial. (D) RVD-Abfolgen der in (E) verwendeten TALEs, sowie DNA-Sequenzen der getesteten Boxen. Die TALEs enthalten entweder 17,5 oder 11,5 nur 34 *As-repeats* oder einen natürlichen aberranten *repeat* mit 40 *As* an Position 3 (grau hinterlegt) und dem RVD NN. Neben der optimalen Box (opt) wurden -1 Nukleotid-Leserasterverschiebungen an den Positionen 3 und 4 (-1p3/p4) angeboten, gekennzeichnet durch eine gestrichelte Linie. (F) Immunoblot (Richter *et al.*, 2014). Ein Pfeil markiert die detektierten, GFP-gekoppelten TALEs.

Zusammenfassend konnte gezeigt werden, dass aberrante *repeats* TALEs mit geringer *repeat*-Anzahl stark in ihrer Aktivität beeinträchtigen, was mutmaßlich an einer durch den aberranten *repeat* verringerten DNA-Affinität von TALE-Derivaten liegen könnte (Abb.3.1.4 A, B, D, E).

3.1.5 Aberrante *repeats* nehmen an der Erkennung der Ziel-DNA-Sequenz teil

Um die Frage zu beantworten, ob aberrante *repeats* überhaupt an der DNA-Bindung teilnehmen, wurden die artifiziellen TALEs mit den verschiedenen aberranten *repeats* und ohne aus 3.1.1 und 3.1.3 im Vergleich zur optimalen Box an Boxen mit RVD-Base-Fehlpaarungen getestet (Abb. 3.1.5). Dazu wurden ausgehend von dem Nukleotid an Position 8 oder 9 eine, zwei oder drei Nukleobasen ausgetauscht, sodass Fehlpaarungen mit den RVDs der TALEs entstehen (Abb. 3.1.5 A). Diese Austausche sollten zum einen eine Beurteilung darüber erlauben, ob der aberrante *repeat* bei Bindung der optimalen Box einen Kontakt mit dem gegenüberliegenden Nukleotid herstellt, und zum anderem, ob die RVDs unmittelbar nachfolgender *repeats* in der Lage sind, an die korrespondierenden Nukleotide zu binden. Sollten Basenaustausche und somit RVD-Base-Fehlpaarungen keinen Einfluss auf die GUS-Reportergenaktivierung durch die artifiziellen TALEs im Vergleich zur optimalen Box haben, wäre dies ein Indiz für eine lokal unterbrochene DNA-Bindung durch die *repeat*-Domäne.

GUS-Reporterstudien zeigten für den artifiziellen TALE ohne aberranten *repeat* (TALE34) eine Verringerung der Induktion durch alle getesteten Fehlpaarungen (Abb. 3.1.5 B), sodass davon ausgegangen werden kann, dass der Aufbau des Experimentes durchaus geeignet ist um den beschriebenen Sachverhalt zu untersuchen. Offenbar erlaubt die Auswahl hochspezifischer

3.1.6 Zwei aberrante *repeats* im Tandem verhalten sich anders als ein Einzelner

Bisher wurde von der natürlichen Situation ausgegangen, in der aberrante *repeats* nur einzeln und eingestreut in Abfolgen von 34 *As-repeats* auftreten. Im Folgenden wurde nun überprüft, ob ein artifiziereller TALE mit zwei aberranten *repeats* in unmittelbarer Folge an einer DNA-Sequenz GUS-Reporteraktivität induzieren kann (Abb. 3.1.6). Dazu wurden zwei 40 *As-repeats* an Position 8 und 9 der *repeat*-Abfolge platziert und im Vergleich zu TALE34 und TALE40p8, d.h. einem entsprechenden TALE ohne und mit nur einem 40 *As-repeat*, an den Boxen mit verschiedenen Leserasterverschiebungen aus 3.1.1 und 3.1.3 getestet (Abb. 3.1.6 A).

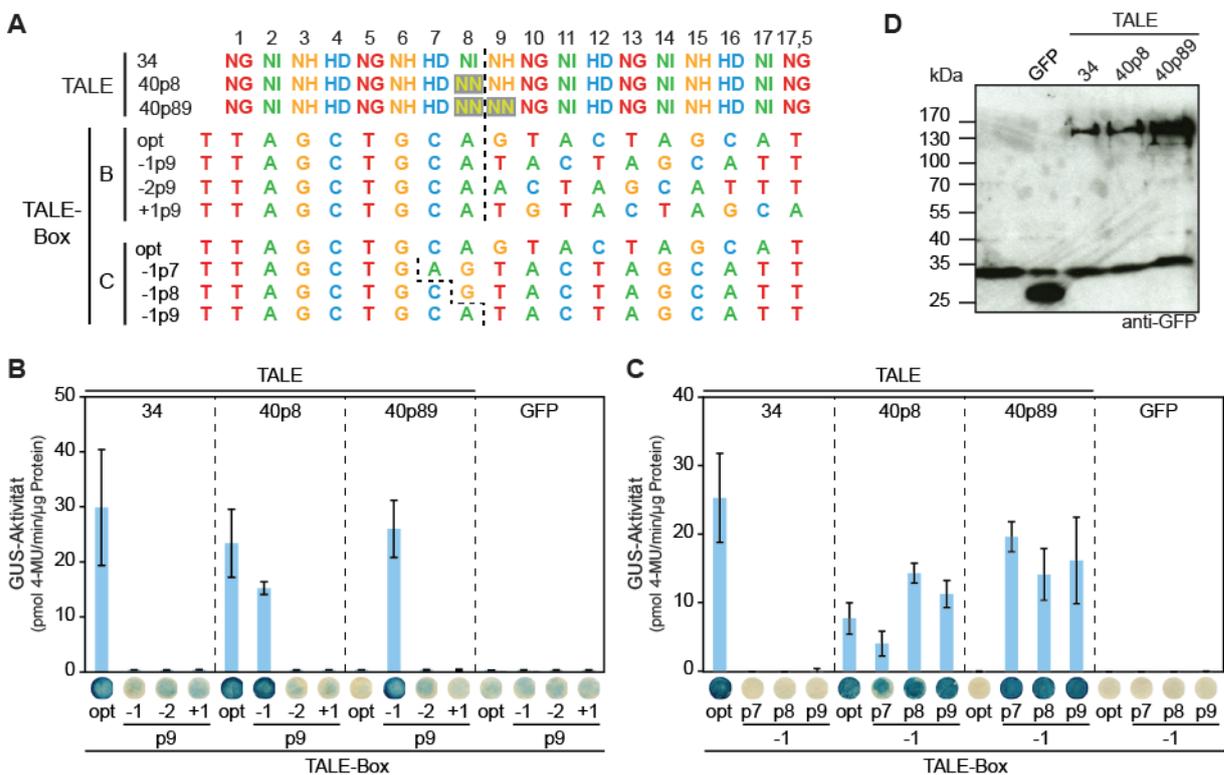


Abb. 3.1.6: Einer von zwei 40 *As-repeats* im Tandem verlässt die *repeat*-Region bei DNA-Bindung. (A) RVD-Abfolgen der artifizierellen TALEs, sowie DNA-Sequenzen der getesteten Boxen. Die verwendeten TALEs enthalten entweder nur 34 *As-repeats* oder einen, bzw. zwei natürliche aberrante *repeats* mit 40 *As* an Position 8 oder 8 und 9 (40p8/p89, grau hinterlegt) und dem RVD NN. Neben der optimalen Box (opt) wurden in verschiedenen Leserasterverschiebungen, gekennzeichnet durch eine gestrichelte Linie, angeboten. In (B) waren dies Deletionen von ein oder zwei Nukleotiden (-1, -2), bzw. die Insertion eines Thyminidins (+1) an Position 9 (p9). In (C) wurden -1 Nukleotid-Leserasterverschiebungen an den Positionen 7, 8 und 9 (-1p7/p8/p9) getestet. (B, C) Quantitative und qualitative GUS-Reporterstudien im Vergleich zu GFP in *N. benthamiana* wurden 48 h nach Inokulation durchgeführt ($n=3$). Fehlerindikatoren kennzeichnen Standardabweichungen im quantitativen Ansatz. Eine repräsentative Blattscheibe des qualitativen Ansatzes wird gezeigt. (D) Immunologischer Nachweis der GFP-gekoppelten TALEs 48 h nach *Agrobacterium*-vermittelter Expression in *N. benthamiana* (Richter *et al.*, 2014).

Dabei konnte beobachtet werden, dass der artifizielle TALE mit zwei 40 *As-repeats* (TALE 40p89) im Gegensatz zu dem artifiziellen TALE mit einem 40 *As-repeat* (TALE40p8) nicht in der Lage war zwei Boxen mit unterschiedlichem Leseraster zu erkennen (Abb. 3.1.6 B). Anders als TALE34 erkennt TALE40p89 jedoch nicht die optimale Box, sondern ausschließlich eine -1 Nukleotid-Leserasterverschiebungsbox (Abb. 3.1.6 B, -1p9). Dies lässt die Vermutung zu, dass im Falle zweier 40 *As-repeats* in Tandem einer die *repeat*-Folge verlassen muss. Dabei ist zu beachten, dass 40 *As-repeats* den RVD NN haben, welcher sowohl die Erkennung von Guanin als starker und von Adenin als intermediärer RVD zulässt. Deshalb kann keine Aussage darüber getroffen werden, welcher der beiden 40 *As-repeats* die *repeat*-Abfolge verlässt (Abb. 3.1.6 A). Um nachzuweisen, dass TALE40p89 trotz der beiden aberranten *repeats* in *N. benthamiana* korrekt synthetisiert wird, wurde ein Immunoblot erstellt, der das Protein neben den anderen, hier getesteten artifiziellen TALEs mit dem korrekten Molekulargewicht zeigt (Abb. 3.1.6 D).

3.1.7 PthXo3 und AvrXa7 tolerieren Indels in ihrer Ziel-DNA-Sequenz

Da mit Hilfe artifizieller TALEs gezeigt werden konnte, dass aberrante *repeats* die Erkennung einer optimalen und einer -1 Nukleotid-Leserasterverschiebungsbox ermöglichen, wurde nun untersucht, ob natürliche TALEs mit einer 39 *As-repeat* Leserasterverschiebungsbox erkennen. Dazu wurden TALE-Expressionskonstrukte für *N. benthamiana* erstellt, die die originale *repeat*-Region der natürlichen TALEs PthXo3 und AvrXa7 enthalten. Zur Vereinfachung werden diese im Folgenden PthXo3, bzw. *pthXo3*, und AvrXa7, bzw. *avrXa7*, genannt. Um PthXo3 und AvrXa7 nicht nur auf die Erkennung ihrer *OsSWEET14* (*Os11N3*)-Zielsequenzen (Römer *et al.*, 2010; Antony *et al.*, 2010; Abb. 3.1.8 A, 3.1.8 A), sondern auch auf die Erkennung gezielt veränderter Zielsequenzen zu testen, wurde für jeden der beiden TALEs eine optimale Box abgeleitet (Abb. 3.1.7 A, 3.1.8 A). Dabei wurden die DNA-Spezifitäten der RVDs in eine DNA-Sequenz übersetzt (Abb. 3.1.7 B), wobei für RVDs mit multipler Spezifität jeweils nur eine mögliche DNA-Base verwendet wurde, Guanin für NN, Thymin für N* und Adenin für NS (Abb. 3.1.7 A, 3.1.8 A). Auf der Grundlage dieser optimalen Boxen wurden für beide TALEs verschiedene Leserasterverschiebungsboxen erstellt, in denen entweder ein oder zwei Nukleotide nahe des aberranten *repeats* deletiert oder inseriert wurden (Abb. 3.1.7 A, 3.1.8 A, optimale Box -1, -2, +1, +2). Die optimalen Boxen von PthXo3 und AvrXa7 und ihre Derivate wurden ebenfalls stromaufwärts des *Bs4*-Minimalpromotors vor ein GUS-Reportergen fusioniert und in transienten GUS-Reportergenaktivierungsstudien mit dem jeweiligen TALE getestet (Abb. 3.1.7 B, 3.1.8 B).

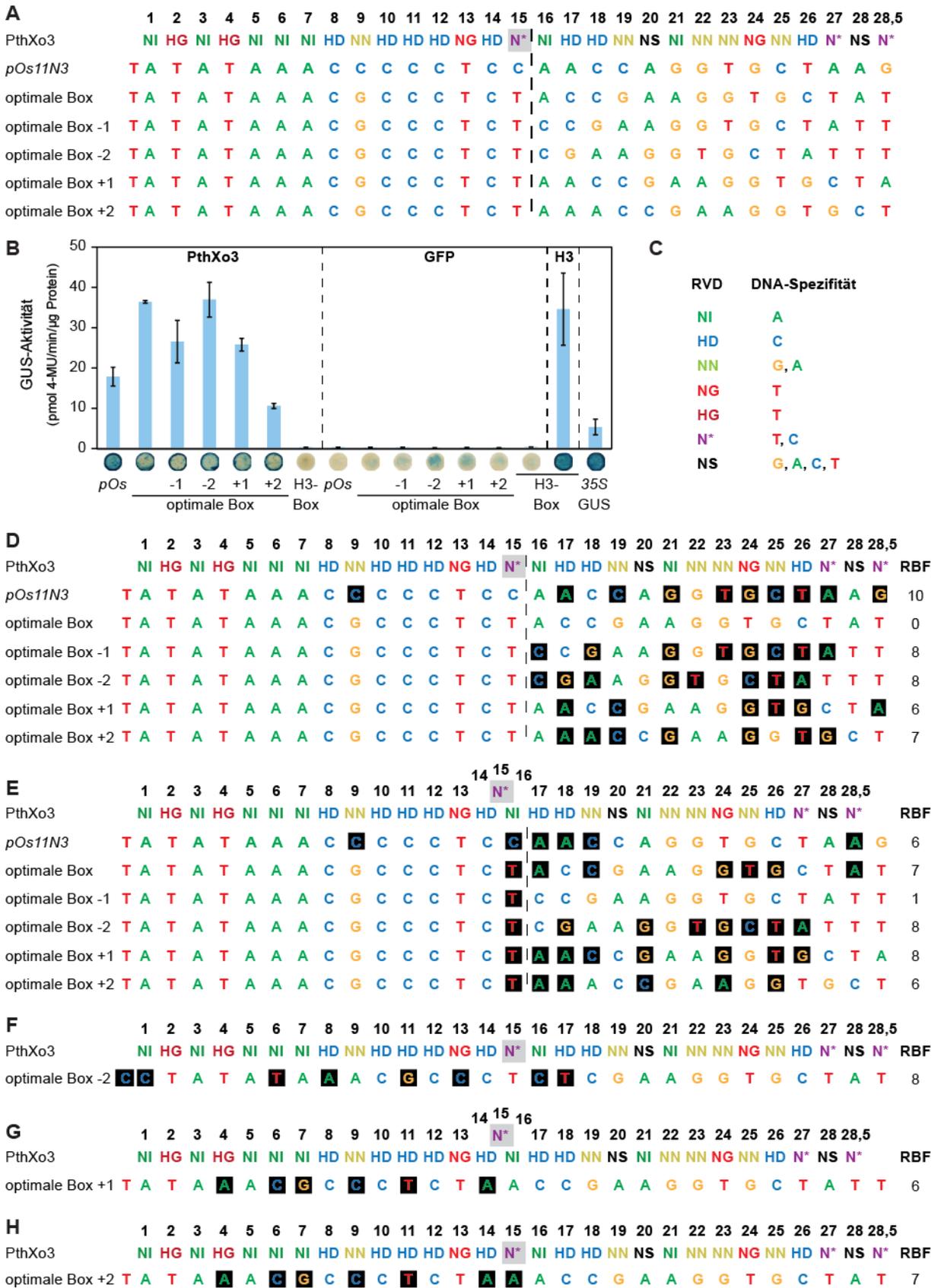


Abb. 3.1.7: PthXo3 erkennt eine Vielzahl von Ziel-DNA-Sequenzen. (A) PthXo3 RVDs und Ziel-DNA-Sequenzen. Der natürliche TALE PthXo3 enthält einen aberranten N*-repeat von 39 As an Position 15 (grau hinterlegt). Die natürliche PthXo3-Box im Reis-*Os11N3* (*Os11g31190*, *OsSWEET14*)-Promotor (Antony *et al.*, 2010; *pOs11N3*, *pOs*) oder eine optimale Box, bzw. optimale Boxen mit Deletionen oder Insertionen von ein oder

zwei Basenpaaren an Position 16 (optimale Box -1, -2, +1, +2; gestrichelte Linie), stromabwärts eines *Bs4*-Minimalpromotors wurden an ein promotorloses GUS-Reporter gen fusioniert. (B) GUS-Reporter genanalyse von PthXo3 mit den verschiedenen Boxen ($n=3$). GUS-Reporter konstrukte wurden zusammen mit konstitutiv *35S*-getriebenem *pthXo3*, bzw. *GFP*, mittels *A. tumefaciens* in *N. benthamiana*-Blattzellen übertragen. *35S::uidA* (*35S* GUS) und der natürliche TALE Hax3 (H3) mit seiner perfekten Hax3-Box (H3-Box) dienten als Kontrollen. Fehlerindikatoren kennzeichnen Standardabweichungen im quantitativen Ansatz. Eine repräsentative Blattscheibe des qualitativen Ansatzes wird gezeigt. (C) RVD-Spezifitäten. (D-H) Fehlpaarungsanalyse von PthXo3 an den verschiedenen Boxen. Unpassende RVD-Base-Paarungen sind schwarz hinterlegt. Die Anzahl der RVD-Base-Fehlpaarungen (RBF) ist hinter den Boxen indiziert. (D) Der aberrante *repeat* von PthXo3 in regulärer RVD-Base-Paarung. (E) Der aberrante *repeat* von PthXo3 nimmt nicht an der DNA-Erkennung teil. (F-H) Alternative *alignments* von PthXo3 mit (F) der optimalen Box -2, (G) der optimalen Box +1 und (H) der optimalen Box +2.

PthXo3 induzierte erwartungsgemäß GUS-Reporter genaktivität am *Os11N3*-Promotor (Abb. 3.1.7 B). Fehlpaarungsanalysen anhand der gegebenen RVD-DNA-Spezifitäten (3.1.7 C) zeigen eine Reduktion der RVD-Base-Fehlpaarungen (RBF) mit dem *Os11N3*-Promotor von 10 auf 6, wenn der 39 *As-repeat* nicht an der DNA-Sequenzerkennung teilnimmt (3.1.7 D, E, *pOs11N3*). Auch bei einer entsprechenden Zielsequenzvorhersage mit dem Programm TALgetter konnte ohne den aberranten *repeat* eine Verbesserung des Ranges der *Os11N3*-Zielsequenz unter allen vorhergesagten PthXo3-Zielsequenzen in Reis von 119 mit auf 1 erreicht werden (Richter *et al.*, 2014). Das bedeutet, dass PthXo3 *OsSWEET14* vermutlich ohne Teilnahme des 39 *As-repeats* an der Sequenzerkennung induziert. Neben dem *Os11N3*-Promotor konnte PthXo3 GUS-Reporter genaktivität an seiner optimalen Box, sowie an allen Leserasterverschiebungsboxen induzieren (Abb. 3.1.7 B). Da PthXo3 keine Aktivierung des GUS-Reporter konstruktes mit der H3-Box als Negativkontrolle zeigt, scheint dies jedoch auf spezifischer Sequenzerkennung zu beruhen. Auch für die optimale Box -1 ergab die Fehlpaarungsanalyse erwartungsgemäß eine Verringerung der RBF von 8 auf 1 bei Abwesenheit des 39 *As-repeat* von der DNA-Bindung (3.1.7 D, E). Die übrigen Leserasterverschiebungsboxen ergaben jedoch auch bei Berücksichtigung der Bindung eines alternativen Leserasters durch die N-terminalen *repeats* RBF von 6 bis 8 (3.1.7 D-H). Eine eindeutige Aussage darüber wie PthXo3 diese Sequenzen erkennt ist hier nicht möglich. Es zeigt aber das PthXo3 neben dem aberranten *repeat* über Mechanismen verfügt, die eine hohe Fehlpaarungstoleranz erlauben.

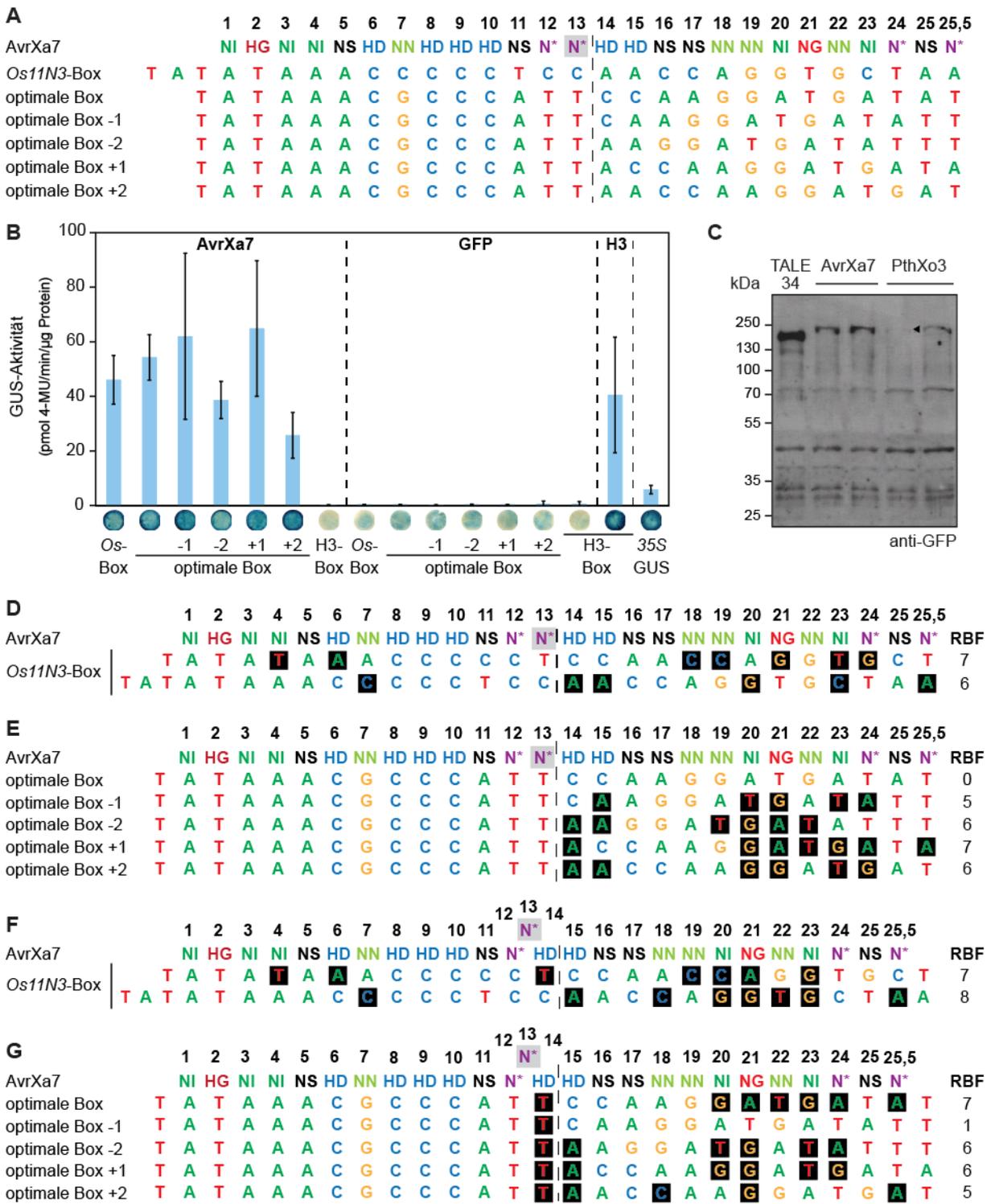


Abb. 3.1.8: AvrXa7 erkennt eine Vielzahl von Ziel-DNA-Sequenzen. (A) AvrXa7 RVDs und Ziel-DNA-Sequenzen. Der natürliche TALE AvrXa7 enthält einen aberranten N*-repeat von 39 As an Position 13 (grau hinterlegt). Die natürliche AvrXa7-Box aus dem Reis-Os11N3 (*Os11g31190*, *OsSWEET14*)-Promotor (*Os11N3*-Box, *Os*-Box), eine optimale Box, bzw. optimale Boxen mit Deletionen oder Insertionen von ein oder zwei Basenpaaren an Position 14 (optimale Box -1, -2, +1, +2; gestrichelte Linie), stromabwärts eines *Bs4*-Minimalpromotors wurden an ein promotorloses GUS-Reportergen fusioniert. (B) GUS-Reportergenanalyse von AvrXa7 mit den verschiedenen Boxen ($n=3$). GUS-Reportergenkonstrukte wurden zusammen mit konstitutiv 35S-getriebenem *avrXa7*, bzw. *GFP*, mittels *A. tumefaciens* in *N. benthamiana*-Blattzellen übertragen. 35S::uidA (35S

GUS) und der natürliche TALE Hax3 (H3) mit seiner perfekten Hax3-Box (H3-Box) dienen als Kontrollen. Fehlerindikatoren kennzeichnen Standardabweichungen im quantitativen Ansatz. Eine repräsentative Blattscheibe des qualitativen Ansatzes wird gezeigt. (C) Immunoblot der natürlichen TALEs AvrXa7 und PthXo3 mit N-terminalem GFP-Epitop im Vergleich zu dem artifiziellen TALE34 48 h nach *Agrobacterium*-vermittelter Expression in *N. benthamiana*. AvrXa7 und PthXo3 in jeweils zwei Spuren. Eine schwarze Pfeilspitze kennzeichnet ein PthXo3-Signal. (D-G) Fehlpaarungsanalyse von AvrXa7 an den verschiedenen Boxen. Unpassende RVD-Base-Paarungen sind schwarz hinterlegt. Die Anzahl der RVD-Base-Fehlpaarungen (RBF) ist hinter den Boxen indiziert. (D, E) Der aberrante *repeat* von AvrXa7 in regulärer RVD-Base-Paarung. (F, G) Der aberrante *repeat* von AvrXa7 nimmt nicht an der DNA-Erkennung teil. (D, F) AvrXa7 mit der natürlichen Box aus dem Reis-*Os11N3*-Promotor entweder aligniert an der linksmöglichsten Sequenz beschrieben von Römer *et al.*, 2010, oder der zweiten möglichen Sequenz beschrieben von Antony *et al.*, 2010. (E, G) AvrXa7 aligniert an den optimalen Boxen.

Um sicherzustellen, dass die hochmolekularen TALEs PthXo3 als auch AvrXa7 in *N. benthamiana* vollständig synthetisiert werden, erfolgte der Nachweis mittels Immunoblot (Abb. 3.1.8 C). Wie PthXo3 aktivierte auch AvrXa7 das GUS-Reportergen an der *Os11N3*- und optimalen Box, sowie an allen ihren Derivaten mit Leserasterverschiebung, jedoch nicht an der Negativkontrolle H3-Box (Abb. 3.1.8 B). Fehlpaarungsanalysen zeigten, dass auch AvrXa7 eine optimale Box mit einer -1 Nukleotid-Leserasterverschiebung vermutlich ohne Teilnahme des aberranten *repeat* erkennt (Abb. 3.1.8 E, G). Für alle anderen angebotenen Zielsequenzen kann keine Aussage getroffen werden, ob der aberrante *repeat* von AvrXa7 an der DNA-Sequenzerkennung teilnimmt oder nicht (Abb. 3.1.8 E, D-G). Für die Erkennung der Vielzahl von Zielsequenzen durch die natürlichen TALEs AvrXa7 und PthXo3 und somit ihrer hohen Fehlpaarungstoleranz scheinen noch andere Parameter außer dem aberranten *repeat* eine Rolle zu spielen. Beide TALEs enthalten viele RVDs mit multipler Basen-Spezifität, namentlich NN, NS und N* (Abb. 3.1.7 A, 3.1.8 A), die dazu beitragen, dass die Leserasterverschiebungen zu weniger RBF führen als theoretisch zu erwarten wären. So ergeben sich für PthXo3 und AvrXa7 nur maximal 8 bzw. 6 RBF bei einer Verschiebung von insgesamt 14 bzw. 13 Nukleotiden stromauf- oder stromabwärts relativ zum *repeat*-Array (Abb. 3.1.7 D-H, 3.1.8 D-G). Darüber hinaus verfügen beide TALEs über eine Vielzahl der starken RVDs HD und NN, sowie eine generell hohe Anzahl an *repeats*. Diese Faktoren könnten die maximale Anzahl tolerierter RBF beider TALEs womöglich steigern. Um den Einfluss von aberrantem *repeat*, RVD-Komposition und *repeat*-Anzahl zu beurteilen bedarf es jedoch weiterer Analysen.

3.1.8 Künstliche AvrXa7-Derivate erlauben eine Beurteilung des Einflusses von aberrantem *repeat* und RVD-Komposition auf die Erkennung verschiedener Ziel-DNA-Sequenzen

Um den Einfluss von aberrantem 39 *As-repeat* und der RVD-Zusammensetzung von AvrXa7 auf die Erkennung von Leserasterverschiebungsboxen zu untersuchen, wurden artifizielle AvrXa7-Derivate erstellt, deren *repeat*-Regionen in ihrer Zusammensetzung verändert wurden (Abb. 3.1.9 A, ArtXa70-73; Katrin Schlien). Zunächst wurde der 39 *As-repeat* gegen einen 34 *As-repeat* mit gleichem RVD ausgetauscht (ArtXa71). Da AvrXa7 zahlreiche RBF toleriert (Abb. 3.1.8), wurden separat oder zusätzlich zum Austausch des aberranten *repeats* auch die RVDs mit multipler Spezifität zu RVDs, die nur ein Nukleotid erkennen, verändert (Abb. 3.1.9 A). Die Auswahl des entsprechenden RVDs erfolgte dabei anhand der optimalen AvrXa7-Box (Abb. 3.1.9 A). Dabei wurde das RVD NS stets zu NI für Adenin, NN zu NH für Guanin und N* zu NG für Thymin ausgetauscht. Das RVD N* an Position 13 der *repeat*-Region wurde in den beiden Derivaten, ArtXa72 und ArtXa73, beibehalten um den natürlich vorkommenden 39 *As-repeat* nicht zu verändern und den TALE ohne aberranten *repeat*, ArtXa73, bzgl. der RVDs vergleichbar zu belassen.

Zunächst wurden GUS-Reporteranalysen durchgeführt, um AvrXa7 mit dem nachgebauten ArtXa70 an den verschiedenen Boxen zu vergleichen (Abb. 3.1.9 B). Dabei konnte festgestellt werden, dass beide TALEs an der natürlichen und allen Varianten der optimalen Box das Reporter gen induzieren. Ein Unterschied konnte jedoch bei der Aktivierung an den Boxen mit Nukleotid-Insertionen festgestellt werden. An der optimalen Box +1 und +2 aktivierte AvrXa7 das GUS-Reporter gen stärker als der artifizielle TALE. Zwar war bei beiden TALEs zu beobachten, dass die optimale Box +2 im Vergleich zu allen übrigen Boxen zur geringsten Aktivität führte, an der optimalen Box +1 induzierte AvrXa7 jedoch ähnlich hohe Reporter genaktivität wie an der optimalen Box selbst, wohingegen die durch ArtXa70 vermittelte Aktivierung deutlich geringer ausfiel (Abb. 3.1.9 B). Dies weist darauf hin, dass neben der RVD-Komposition und dem aberranten *repeat* noch andere Bereiche des Proteins AvrXa7 dessen großes Spektrum an Zielsequenzen beeinflussen. Dabei kann es sich um unterschiedliche Aminosäuren an nicht RVD-Positionen der *repeats* oder im N- oder C-Terminus handeln. Sie könnten die Syntheserate oder Stabilität des TALEs und somit die Verfügbarkeit in *N. benthamiana* beeinflussen. Ein Indiz dafür lieferte ein vergleichender Immunoblot von in *N. benthamiana* synthetisiertem AvrXa7 mit ArtXa70, der eine höhere Expression von *avrXa7* zeigte (Abb. 3.1.9 C).

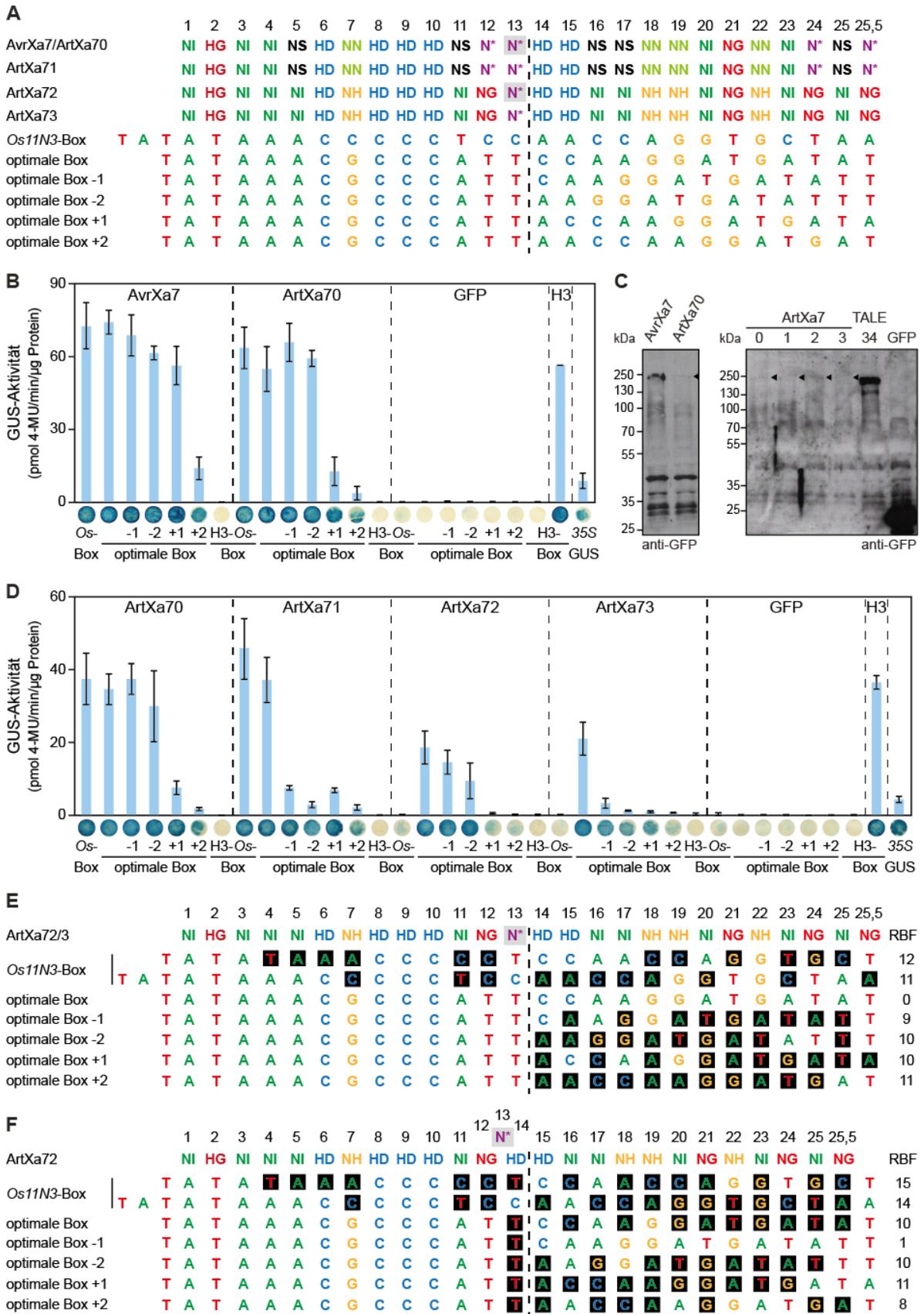


Abb. 3.1.9: Artificielle AvrXa7-Derivate zeigen den Einfluss von aberrantem repeat und RVD-Komposition auf das Spektrum akzeptierter Boxen. (A) TALE RVDs und Ziel-DNA-Sequenzen. Der natürliche TALE AvrXa7 und die künstlichen TALEs ArtXa70 und ArtXa72 enthalten einen aberranten N*-repeat von 39 As an Position 13 (grau hinterlegt). ArtXa70 und ArtXa71 (Katrin Schlien) haben mit AvrXa7 identische RVDs. In

ArtXa72 und ArtXa73 wurden anhand der optimalen Box mit Ausnahme des N* im 13. *repeat* RVDs multipler Spezifität zu RVDs ausgetauscht, die nur ein Nukleotid spezifizieren (NS zu NI, NN zu NH, N* zu NG). Die natürliche AvrXa7-Box aus dem Reis-*Os11N3* (*Os11g31190*, *OsSWEET14*)-Promotor (*Os11N3*-Box, *Os*-Box), eine optimale Box, bzw. optimale Boxen mit Deletionen oder Insertionen von ein oder zwei Basenpaaren an Position 14 (optimale Box -1, -2, +1, +2; gestrichelte Linie), stromabwärts eines *Bs4*-Minimalpromotors wurden an ein promotorloses GUS-Reportergen fusioniert. (B, D) GUS-Reportergenanalysen der TALEs mit den verschiedenen Boxen ($n=3$). GUS-Reporterkonstrukte wurden zusammen mit konstitutiv *35S*-getriebenem *TALE*, bzw. *GFP*, mittels *A. tumefaciens* in *N. benthamiana*-Blattzellen übertragen. *35S::uidA* (*35S* GUS) und der natürliche TALE Hax3 (H3) mit seiner perfekten Hax3-Box (H3-Box) dienten als Kontrollen. Fehlerindikatoren kennzeichnen Standardabweichungen im quantitativen Ansatz. Eine repräsentative Blattscheibe des qualitativen Ansatzes wird gezeigt. (B) Vergleich des natürlichen AvrXa7 mit dem künstlichen ArtXa70. (C) Immunoblots der artifiziellen TALEs ArtXa70-3 mit N-terminalem GFP-Epitop im Vergleich zu AvrXa7 oder dem artifiziellen TALE34 48 h nach *Agrobacterium*-vermittelter Expression in *N. benthamiana*. Schwarze Pfeilspitzen kennzeichnen die Signale von ArtXa70-3. (D) GUS-Reporterstudien zum Vergleich der artifiziellen AvrXa7-Derivate ArtXa70-3 mit und ohne aberranten *repeat* oder RVDs multipler Spezifität. (E, F) Fehlpaarungsanalyse von ArtXa72-3 an den verschiedenen Boxen. Unpassende RVD-Base-Paarungen sind schwarz hinterlegt. Die Anzahl der RVD-Base-Fehlpaarungen (RBF) ist hinter den Boxen indiziert. (E) Der aberrante *repeat* von ArtXa72 in regulärer RVD-Base-Paarung. (F) Der aberrante *repeat* von ArtXa72 nimmt nicht an der DNA-Erkennung teil.

Der immunologische Nachweis der übrigen artifiziellen AvrXa7-Derivate, ArtXa71 bis ArtXa73, in *N. benthamiana* zeigte eine mit ArtXa70 vergleichbare, schwache Synthese der Proteine mit entsprechendem Molekulargewicht (Abb. 3.1.9 C). Als Kontrollen dienten TALE34 mit geringerem Molekulargewicht und GFP, die ein deutlich stärkere Signale lieferten. GUS-Reporterstudien der artifiziellen AvrXa7-Derivate mit der optimalen Box und den Leserasterverschiebungsboxen zeigten sowohl einen Einfluss des 39 *As-repeat* als auch der RVDs mit multipler Spezifität auf das Sequenzerkennungsspektrum der TALEs (Abb. 3.1.9 D). Der Vergleich des künstlichen AvrXa7, ArtXa70, mit ArtXa71 ohne aberranten *repeat* zeigt, dass der aberrante *repeat* in AvrXa7 für die Erkennung von Leserasterverschiebungen um -1 oder -2 Nukleotide notwendig ist (Abb. 3.1.9 A, D). Fehlpaarungsanalysen konnten bereits die Erkennung der optimalen Box -1 durch AvrXa7 erklären, indem der 39 *As-repeat* nicht an der DNA-Bindung teilnimmt (3.1.7, Abb. 3.1.8 G). Der Erkennung der optimalen Box -2 durch AvrXa7 hingegen scheint ein anderer Mechanismus zugrunde zu liegen (3.1.7, Abb. 3.1.8 G). Da die künstlichen AvrXa7-Derivate mit und ohne 39 *As-repeat*, ArtXa70 und ArtXa71, beide in der Lage sind an der *Os11N3*-Box das Reportergen zu induzieren (Abb. 3.1.9 D), ist der aberrante *repeat* während der Bindung der *Os11N3*-Box durch AvrXa7 vermutlich in das *repeat*-Array integriert (Abb. 3.1.8 D). Beide 2010 publizierten Ziel-DNA-Sequenzen nach Römer *et al.* und Antony *et al.* stellen dabei mögliche Boxen dar. Im Falle der, wenn auch vergleichsweise schwachen, Aktivierung an den optimalen Boxen +1 und +2 scheint die

Besonderheit des aberranten *repeats*, nicht an der DNA-Bindung teilnehmen zu müssen, ebenfalls keine Rolle zu spielen (Abb. 3.1.9 D).

Bei Betrachtung der Aktivität der artifiziellen AvrXa7-Derivate mit spezifischeren RVDs, ArtXa72 und ArtXa73, an den verschiedenen GUS-Reporterkonstrukten werden Gemeinsamkeiten und Unterschiede zu ArtXa70 und ArtXa71 bei der Boxerkennung deutlich. Wie für ArtXa70 und ArtXa71 gezeigt, konnte mit ArtXa72 und ArtXa73 eine mit der optimalen Box vergleichbare Aktivierung der Boxen mit ein oder zwei Nukleotiddeletionen nur dann erreicht werden, wenn der aberrante *repeat*, hier in ArtXa72, vorhanden war (Abb. 3.1.9 D). Auch die Erkennung der optimalen Box -1 durch ArtXa72 kann dabei durch aus-loopen des 39 *As-repeat* erklärt werden, was jedoch nicht die Erkennung der optimalen Box -2 erklärt (Abb. 3.1.9 E, F). Erwartungsgemäß wird das Reporterkonstrukt mit der *OsIIN3*-Box nicht durch die für die optimale Box designten TALEs ArtXa72 und ArtXa73 aktiviert (Abb. 3.1.9 D). Fehlpaarungsanalysen ergaben für diese Paarungen 11 oder mehr RBF unter Berücksichtigung der verschiedenen Bindungsmodi (Abb. 3.1.9 E, F). Allerdings lässt sich auch bei der Aktivierung der optimalen Box durch ArtXa72 und ArtXa73 eine geringere GUS-Reportergenaktivität messen als bei ArtXa70 und ArtXa71, obwohl keine RBF vorliegen (Abb. 3.1.9 D, E). Dies liegt im Austausch der RVDs mit multipler Spezifität zu RVDs, die nur ein Nukleotid spezifizieren, begründet. Bei den in ArtXa70 und ArtXa71 enthaltenen RVDs NN und NS handelt es sich um einen starken und einen intermediären RVD, die in ArtXa72 und ArtXa73 zu dem intermediären RVD NH und dem schwachen RVD NI ausgetauscht wurden (Streubel *et al.*, 2012). Diese Veränderung der RVD-Effizienzen schlägt sich entsprechend auf die Aktivität der TALEs nieder. An den Insertionsboxen induzierten sowohl ArtXa72 mit, als auch ArtXa73 ohne 39 *As-repeat* im Unterschied zu ArtXa70 und 71 kaum noch Reportergenaktivität (Abb. 3.1.9 D). Dies zeigt den Einfluss der RVD-Komposition auf das Spektrum akzeptierter Ziel-DNA-Sequenzen. Obwohl die Anzahl an RBF von ArtXa72 und ArtXa73 mit den optimalen Boxen +1 und +2 im Vergleich zu ArtXa70 und ArtXa71, bzw. AvrXa7 unabhängig davon, ob der aberrante *repeat* an der DNA-Erkennung teilnimmt oder nicht, in jedem Falle gestiegen ist (Abb. 3.1.8 E, G; Abb. 3.1.9 E, F), muss auch die verringerte RVD-Effizienz für die Reduktion der Reportergeninduktion in Betracht gezogen werden.

3.2 Untersuchung der TALE-vermittelten Transkriptionsinduktion am HIV-1-Promotor

In Vorarbeiten wurde mit Luciferase-Reporterstudien bereits eine Positionsabhängigkeit der Aktivität von TALEs am HIV-1-Promotor gezeigt. Im Folgenden wurde dieses System zur Untersuchung dieser Positionsabhängigkeit und zur Optimierung von HIV-1-Aktivatoren verwendet.

3.2.1 Bestimmung des ArtHIV6-induzierten Transkriptionsstarts am HIV-1-Promotor

In Vorarbeiten konnte ein TALE, ArtHIV6, identifiziert werden, der in der Lage ist an dem HIV-1-Promotor aus 8E5-Zellen Luciferase-Reporteraktivität zu induzieren. Für eine Anwendung im ursprünglichen Sinne zur Aktivierung des Virus‘ in ruhenden, latent infizierten Immunzellen eines Patienten oder als Werkzeug zur HIV-1-Induktion für Forschungszwecke wäre es von Interesse näher zu untersuchen, in wie fern die Aktivierung durch ArtHIV6 der natürlichen HIV-1-Aktivierung durch NF- κ B oder NFAT ähnelt (Karn und Stoltzfus, 2012). Hierbei ist vor allem die Lage des Transkriptionsstarts von Bedeutung, da die ersten 59 Nukleotide der HIV-1-mRNA das vollständige TAR-Element umfassen und somit für den regulatorischen Rückkopplungsmechanismus durch Tat wichtig sind (Karn und Stoltzfus, 2012). Um den durch ArtHIV6 initiierten Transkriptionsstartpunkt zu bestimmen wurden 5‘-RACE-Analysen in Hek293T-RexTM-Zellen durchgeführt (Abb. 3.2.1). Dazu wurde das Luciferase-Reporterkonstrukt pFAHPs mit dem HIV-1-Promotorfragment *pHIV-1** verwendet (Abb. 3.2.1 A). In diesem Konstrukt befindet sich das Startcodon der Luciferase nur 44 bp stromabwärts des natürlichen Transkriptionsstarts, wobei nur die ersten 35 bp nach dem Startnukleotid aus dem HIV-1-Promotor stammen. Dementsprechend ist das TAR-Element nicht vollständig enthalten und die typische Sekundärstruktur des 5‘-Bereiches der HIV-1-mRNA kann sich nicht ausbilden. Da jedoch kein Tat-Protein in der verwendeten Zelllinie vorhanden ist und gezeigt wurde, dass weder Tat noch die Sekundärstruktur des TAR-Elements Einfluss auf den Transkriptionsstart nehmen (Rittner *et al.*, 1995), sollte dieser Umstand weder die induzierbare Reporteraktivität noch den Transkriptionsstart an pFAHPs beeinflussen. Zur Kontrolle wurde der Transkriptionsstart an diesem Reporterkonstrukt in Abwesenheit eines TALEs bestimmt (Abb. 3.2.1 B). Dabei konnte der annotierte, natürliche Transkriptionsstart von HIV-1 auch am Reporterkonstrukt anhand von 13 aus 17 sequenzierten Klonen, die ein Fragment der 5‘-RACE-PCR enthielten, bestätigt werden. Am Reporterkonstrukt entspricht dies einer 5‘-UTR von 44 bp. Die übrigen vier Klone zeigten nach 3‘-verschobene Transkriptionsstarts, wobei dreimal eine 5‘-UTR von 40 bp nachgewiesen werden konnte. In Anwesenheit von ArtHIV6 wurde ein ähnliches Ergebnis erzielt (Abb. 3.2.1 C). In acht von

zwölf Klonen wurde ebenfalls der natürliche Transkriptionsstart von HIV-1 gefunden, der einen Abstand von 84 bp zum 3'-Ende der ArtHIV6-Box in *pHIV** hat. Auch der um 4 bp stromabwärts verschobene Startpunkt mit einer 5'-UTR von 40 bp konnte dreimal wiedergefunden werden. Eine weitere Sequenz eines Klonen hatte eine deutlich längere 5'-UTR und startete 138 bp stromaufwärts des 3'-Endes der ArtHIV6-Ziel-DNA-Sequenz.

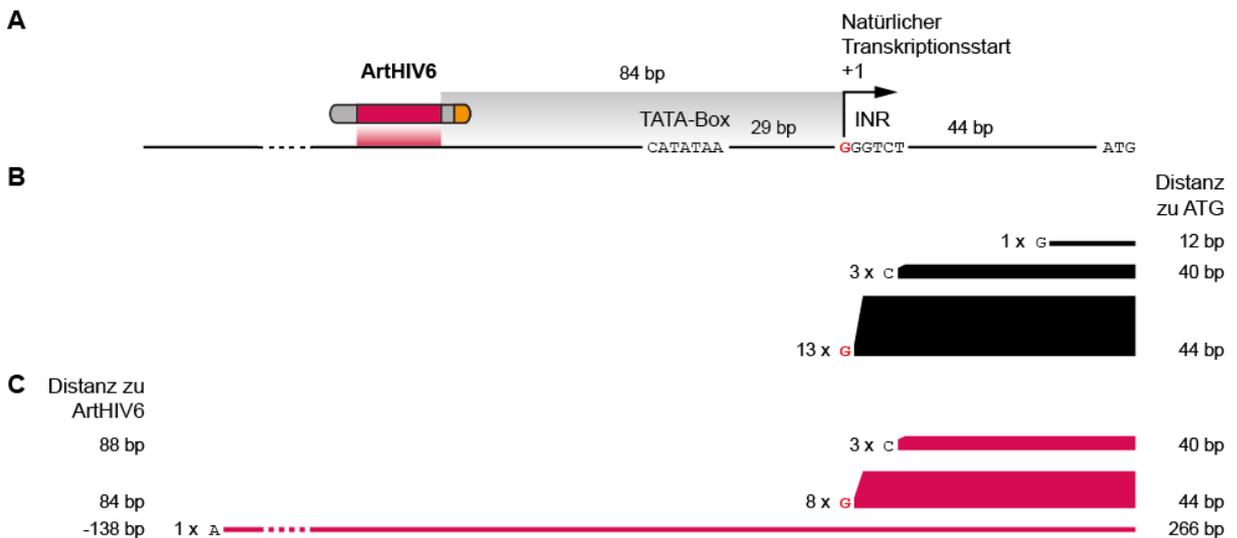


Abb. 3.2.1: Der durch ArtHIV6 induzierte Transkriptionsstart entspricht dem Natürlichen.

(A) Schematische Darstellung der Verhältnisse am HIV-1-Promotor *pHIV-1** im Luciferase-Reporterkonstrukt pFAHPs. Der artifizielle TALE ArtHIV6 (N- und C-Terminus grau, *repeats* pink, C-terminale VP16-AD orange) erkennt eine Ziel-DNA-Sequenz, deren 3'-Ende 84 bp stromaufwärts des bekannten, natürlichen HIV-1-Transkriptionsstart (+1, initiale Base rot) am Initiator-Element (INR, Rittner *et al.*, 1995) entfernt ist. Die Sequenz der TATA-Box (Van Opijnen *et al.*, 2004) 29 bp stromaufwärts und des Startcodons der Luciferase 44 bp stromabwärts des natürlichen Transkriptionsstartes sind aufgeführt. (B, C) Bestimmung des Transkriptionsstarts mittels 5'-RACE in Hek293T-RexTM-Zellen 24 h nach Transfektion. Die Insertionen der erhaltenen Klone wurden ausgehend von der Vektorsequenz sequenziert. Die Anzahl der jeweiligen Sequenzen, die identifizierte Startbase und die Distanz zum Luciferase-ATG sind indiziert. (B) Verifizierung des natürlichen Transkriptionsstartes am Luciferase-Reporterkonstrukt mit *pHIV-1**. Das Reporterkonstrukt pFAHPs wurde im Verhältnis 1:9 mit dem Plasmid pcDHN co-transfiziert. (C) Bestimmung des von ArtHIV6 induzierten Transkriptionsstartes am Luciferase-Reporterkonstrukt mit *pHIV-1**. Das Reporterkonstrukt pFAHPs wurde im Verhältnis 1:9 mit dem ArtHIV6-Expressionsplasmid co-transfiziert.

Dieses Ergebnis der 5'-RACE-Analysen lässt die Annahme zu, dass ArtHIV6 am HIV-1-Promotor ein Transkript mit natürlichem Transkriptionsstart induziert. Diese Annahme wird durch die in Vorarbeiten beobachtete, mindestens fünffache Induktion der Luciferaseaktivität des hier verwendeten Reporterkonstruktes durch ArtHIV6 gestützt. Dadurch kann eine entsprechend mindestens fünffache Erhöhung des Transkriptlevels vorausgesetzt werden. Daraus ergibt sich die Wahrscheinlichkeit, dass alle acht gefundenen Sequenzen mit

natürlichem Transkriptionsstart von der Basalaktivität des Reporterkonstruktes stammen, mit $(1/5)^8 = 2,56 \cdot 10^{-6}$ und ist damit sehr gering.

3.2.2 Cas9-Aktivatoren zeigen ähnliches Aktivitätsmuster am HIV-1-Promotor wie TALE-Aktivatoren

Auf Grund der mit TALE-Aktivatoren festgestellten Positionsabhängigkeit der Aktivierung am HIV-1-Promotor war es von Interesse zu untersuchen, ob die Art der DNA-Bindung dabei eine Rolle spielt. Da Cas9 durch die Wahl der sgRNA leicht zu einer beliebigen Ziel-DNA-Sequenz dirigiert werden kann, bot dieses molekulare Werkzeug eine einfache Alternative zur TALE-DNA-Bindedomäne für die Erstellung von HIV-1-Aktivatoren. Um Cas9-basierte Aktivatoren zu erstellen wurde ein für *H. sapiens* codonoptimiertes *cas9* mit den Mutationen D10A (Asparaginsäure an Position 10 zu Alanin, Jinek *et al.*, 2012) und N863A (Asparagin an Position 863 zu Alanin, Mali *et al.*, 2013 b) in der RuvC-, bzw. der HNH-ähnlichen Endonukleasedomäne genutzt (Ordon *et al.*, 2017). Diese Cas9-Nuklease-Null-Variante (Cas9N0) wurde zur Verwendung als Aktivator mit N- und C-terminalen NLS, sowie einer C-terminalen VP16-AD analog zu den am HIV-1-Promotor eingesetzten TALEs versehen. Für die sgRNAs wurden die Ziel-DNA-Sequenzen im HIV-1-Promotor so gewählt, dass jeweils eine einem der TALEs ArtHIV1, 2, 5, 6 und TALE5 aus Geißler *et al.* (2015) entspricht (Abb. 3.2.2 A). Zwecks Vergleichbarkeit wurde TALE5, der in Geißler *et al.* (2015) mit einer für *H. sapiens* Codon-optimierten Modul-Bibliothek erstellt wurde, mit der Golden TALE Technologie nachgebaut. Dabei ist zu beachten, dass auf Grund der festgelegten Länge von 20 bp und der Notwendigkeit einer PAM die Ziel-DNA-Sequenzen der sgRNAs 1, 2, 5, 6 und T5 nicht vollständig mit denen der entsprechenden TALEs übereinstimmen. Um jedoch eine maximale Übereinstimmung der Position zu ermöglichen wurde neben NGG auch NAG als PAM verwendet. Zur Analyse des erstellten Cas9-Aktivators mit den verschiedenen sgRNAs am HIV-1-Promotor wurden Luciferasereporterstudien in Hek293T-RexTM-Zellen durchgeführt. Um einen Vergleich der TALEs mit einem Cas9-Aktivator und den verschiedenen sgRNAs zu ermöglichen wurden zunächst exemplarisch zwei verschiedene Verhältnisse von Cas9-Aktivator- zu sgRNA-Plasmid getestet (Abb. 3.2.2 B). Dazu wurde bei konstanter Gesamtmenge der transfizierten Plasmide ein Cas9-Aktivator (Cas9N0-VP16) entweder im Verhältnis 1:1 oder 1:3 mit verschiedenen sgRNAs am HIV-1-Promotor getestet. Dabei ließ sich eine Verdopplung der induzierten Luciferaseaktivität am verwendeten Reporterkonstrukt bei Überschuss an sgRNA-Plasmid feststellen (Abb. 3.2.2 B). Das Molekulargewicht und die Synthese des Cas9-Aktivators im Beisein der verschiedenen

sgRNAs wurde im Immunoblot bestätigt (Abb. 3.2.2 C). Es konnte im Vergleich der Synthese des Proteins bei Co-Transfektion mit sgRNAs in unterschiedlichen Verhältnissen kein schwächeres Cas9N0-VP16-Signal nachgewiesen werden, wenn das Plasmid zu dessen Expression im Überschuss transfiziert wurde. Dies deutet auf eine Sättigung der Synthesekapazitäten für Cas9N0-VP16 der hier verwendeten Zellen hin. Ein Verhältnis zwischen Cas9-Aktivator und sgRNA von 1:3 erscheint demnach für den Vergleich mit den TALE-Aktivatoren geeignet. Die sgRNAs 1, 2 und 5 wurden in diesem Experiment ebenfalls, jedoch nur im Verhältnis 1:1 zu Cas9N0-VP16 am HIV-1-Promotor getestet (Abb. 3.2.2 B). Dadurch lassen sich bereits Rückschlüsse auf eine mögliche Positionsabhängigkeit auch von Cas9-Aktivatoren ziehen. Es wird deutlich, dass auch Cas9N0-VP16 mit den sgRNAs 1, 2 und 5 geringere Luciferaseaktivität induziert als mit den sgRNAs 6 und T5, was dem aus Vorarbeiten bekannten Aktivitätsmuster der korrespondierenden TALEs entspricht. An den beiden am weitesten stromaufwärts zum Transkriptionsstart gelegenen Positionen lässt sich die stärkste Aktivierung durch ArtHIV6 und TALE5 als auch GFP-Cas9N0-VP16 mit den sgRNAs 6 und T5 nachweisen. Dabei ist das Aktivitätsniveau der verschiedenen Aktivatoren ähnlich. Im Immunoblot konnte die Synthese von GFP-Cas9N0-VP16 mit allen sgRNAs und ohne mit dem korrekten Molekulargewicht gezeigt werden (Abb. 3.2.2 E). Mit ArtHIV1, 2 und 5 ließ sich jedoch in Vorarbeiten keine reproduzierbare Induktion des Reporters nachweisen. Die sgRNAs 1, 2 und 5 hingegen scheinen die Luciferaseaktivität des Reporterkonstruktes im Vergleich zu den Kontrollen H3-VP16 und Cas9N0-VP16 ohne sgRNA zu verdoppeln. Beim Vergleich der Luciferaseaktivitäten eines N-terminal mit GFP fusionierten Cas9-Aktivators (GFP-Cas9N0-VP16) mit den verschiedenen sgRNAs im Verhältnis 1:3 und den TALE-Aktivatoren am HIV-1-Promotor ist die Induktion des Reporterkonstruktes durch die sgRNAs 1, 2 und 5 noch stärker (2,5-3,5-fach; Abb. 3.2.2 D). Hier wird der Kontrast zu den TALE-Aktivatoren an diesen Positionen deutlich, die keine (ArtHIV1, 2) oder geringe Luciferaseaktivität (ArtHIV5) am Reporterkonstrukt induzieren. Auffällig war eine auch bei geringerer Menge eingesetzten sgRNA-Plasmids nachweisbare Aktivierung des Reportergens durch Cas9-Aktivatoren mit den sgRNAs 1, 2 und 5, die so weder hier noch in Vorarbeiten nicht für die entsprechenden TALE-Aktivatoren gezeigt werden konnte. Dabei hat möglicherweise die unterschiedliche Art der DNA-Bindung eine Bedeutung. Die starke Entwindung der DNA durch das Cas9-Derivat könnte die für die fehlende TALE-Aktivierung ursächlichen Faktoren, evtl. SP1, verdrängen. Da der Vergleich von TALE- und Cas9-Aktivatoren mit VP16-AD am HIV-1-Promotor ein ähnliches Aktivitätsmuster für beide Gruppen zeigte, kann angenommen werden, dass für die effiziente Aktivierung durch ArtHIV6

und TALE5, bzw. sgRNA6 und T5 bestimmte Elemente im hier untersuchten Promotorabschnitt eine Rolle spielen. Die Ziel-DNA-Sequenzen dieser Aktivatoren befinden sich alle stromaufwärts der SP1-Bindestellen und der TATA-Box. Aktivatoren, die auf diesen Elementen platziert waren (ArtHIV1, 2, 5, sgRNA1, 2, 5), erreichten geringere oder keine Aktivierung des Luciferasereporterkonstruktes.

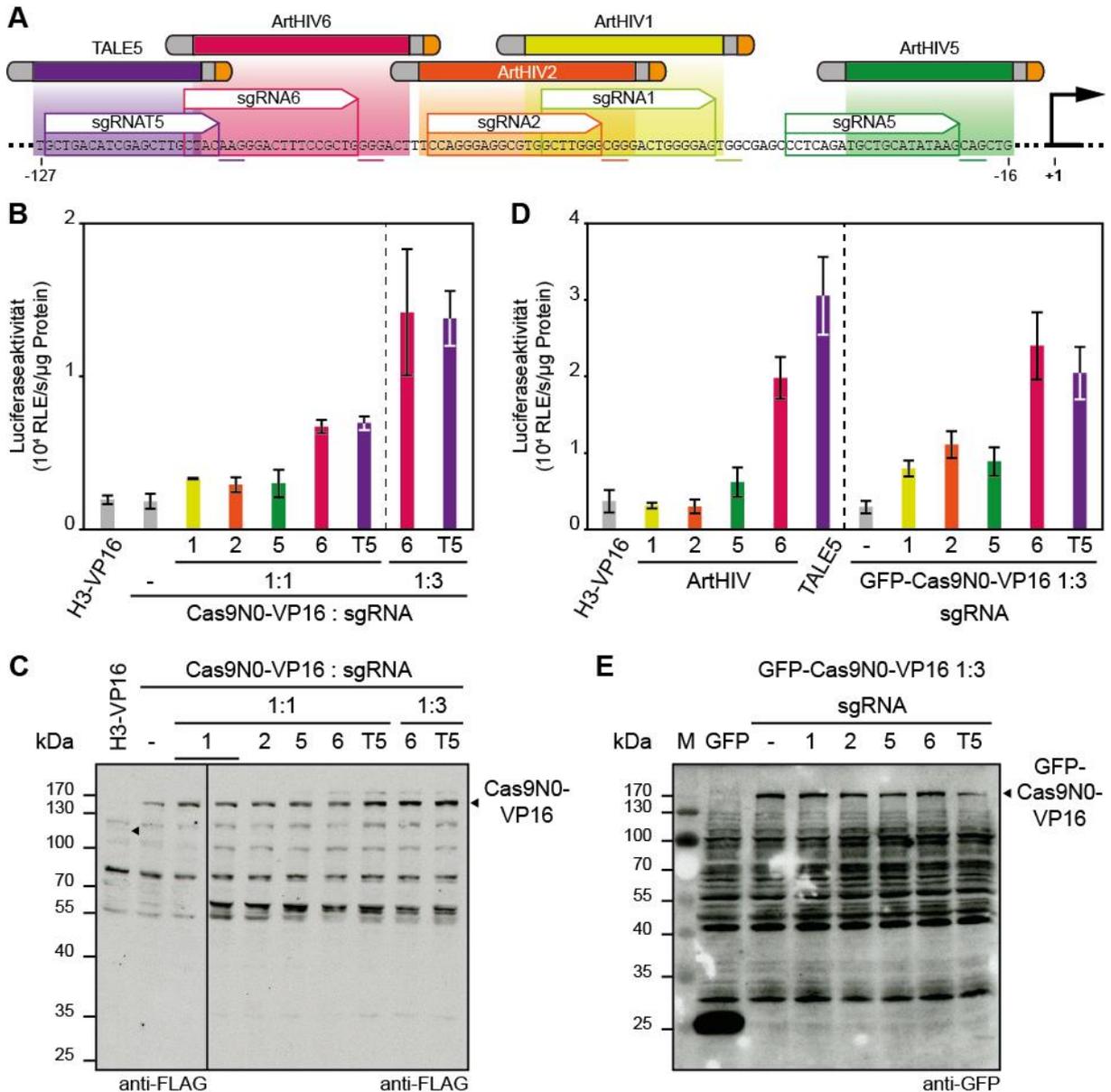


Abb. 3.2.2: Cas9-Aktivatoren zeigen mit sgRNAs für den HIV-1-Promotor ähnliche Positionseffekte wie die entsprechenden TALE-Aktivatoren, wobei die induzierte Aktivität von der Menge der sgRNA abhängt. (A) Lage der Zielsequenzen von ArtHIV1, 2, 5 und 6, sowie TALE5 (Geißler *et al.*, 2015) als auch der entsprechenden sgRNAs 1, 2, 5 und 6, sowie T5 zum Transkriptionsstart (+1) am HIV-1-Promotor aus 8E5. Die korrespondierenden Positionen sind farbcodiert, die PAMs der einzelnen Cas9-Aktivatorbindestellen sind in der entsprechenden Farbe unterstrichen und kennzeichnen neben der Pfeildarstellung der sgRNAs die Ausrichtung der Bindestelle. N- und C-Terminus (grau) der TALEs sind dargestellt. Die natürliche AD am C-Terminus wurde durch eine VP16-AD (orange) substituiert. (B, D) Luciferase-Reporteranalysen der Aktivatoren mit dem

Reporterkonstrukt pFAHPI ($n=3$). Das Luciferase-Reporterkonstrukt wurde im Verhältnis 1:9 zusammen mit konstitutiv CMV-getriebenem *TALE* oder Cas9-Aktivatorgen und hU6-getriebener *sgRNA* in verschiedenen Verhältnissen in Hek293T-RexTM-Zellen co-transfiziert. Als Kontrollen dienten das Reporterkonstrukt mit H3-VP16 oder dem jeweiligen Cas9-Aktivator ohne *sgRNA* (-). Fehlerindikatoren kennzeichnen Standardabweichungen. (B) Ein N-terminal mit FLAG-Epitop fusionierter Cas9-Aktivator (Cas9N0-VP16) wurde entweder im Verhältnis 1:1 oder 1:3 mit verschiedenen *sgRNAs* am HIV-1-Promotor getestet. (D) Ein N-terminal mit GFP fusionierter Cas9-Aktivator (GFP-Cas9N0-VP16) wurde im Verhältnis 1:3 mit den verschiedenen *sgRNAs* am HIV-1-Promotor mit den TALE-Aktivatoren verglichen. Die TALE-Aktivatorplasmide wurden im Verhältnis 1:4 mit dem H3-VP16-codierenden Plasmid als Gemisch transfiziert. (C, E) Immunoblots der verwendeten Cas9-Aktivatoren mit den verschiedenen *sgRNAs*. Die Zelllysate der Dreifachbestimmung des jeweiligen Proteins mit der jeweiligen *sgRNA* oder ohne am HIV-1-Reporterkonstrukt wurden zu gleichen Teilen gemischt und ca. 10 µg Totalprotein aufgetrennt. Die Signale des jeweiligen Cas9-Aktivators sind markiert. (C) Immunoblot zu (B). Die gezeigten Ausschnitte stammen von zwei Bereichen derselben Membran. Die Spur Cas9N0-VP16 1:1 *sgRNA1* ist zum Vergleich in beiden Ausschnitten enthalten. Eine schwarze Pfeilspitze kennzeichnet das sehr schwache Signal von H3-VP16. (E) Immunoblot zu (D). Als Kontrolle wurde ein gleich behandeltes Lysat von *gfp*-exprimierenden Hek293T-RexTM-Zellen verwendet. M: Marker (PageRulerTM Prestained Protein Ladder).

3.2.3 Analyse der Positionsabhängigkeit der TALE-vermittelten Aktivierung am HIV-1-Promotor mit Hilfe einer Promotormutante

In Vorarbeiten konnte eine starke Positionsabhängigkeit der TALE-abhängigen Aktivierung am HIV-1-Promotor festgestellt werden. So wurde eine effiziente Aktivierung nur von Positionen stromaufwärts der drei Bindestellen des Transkriptionsfaktors SP1 festgestellt (Abb. 3.2.3). Ein ähnliches Ergebnis lieferten Studien an dem gleichen, genomisch integrierten HIV-1-Promotor in 8E5-Zellen (Geißler *et al.*, 2015). Um den Einfluss der drei SP1-Boxen auf die TALE-abhängige Aktivierung am HIV-1-Promotor näher zu untersuchen, wurden diese durch Mutationen inaktiviert (Abb. 3.2.3). Dazu wurden ähnlich wie in anderen Arbeiten Positionen in der ersten Hälfte der SP1-Bindestellen verändert (Jones *et al.*, 1986; Burnett *et al.*, 2009; Abb. 3.2.3). Es wurden Guanin zu Adenin-Transitionen durchgeführt, sodass in allen drei SP1-Boxen die zweite bis einschließlich vierte Position mit einem Adenin besetzt wurde. Diese Mutationen wurden gewählt, um neben einer Zerstörung der SP1-Bindestellen den Erhalt funktionaler Bindestellen für ArtHIV1 und 2 zu erreichen (Abb. 3.2.3). Diese mit dem RVD NN für Guanin erstellten TALEs sollten auch in der Lage sein, Sequenzen mit Adeninen an den beschriebenen Positionen zu erkennen. Der ursprüngliche HIV-1-Promotor und die SP1-Bindestellen-Promotormutante werden fortan als originaler (OP) bzw. mutierter Promotor (MP) bezeichnet.

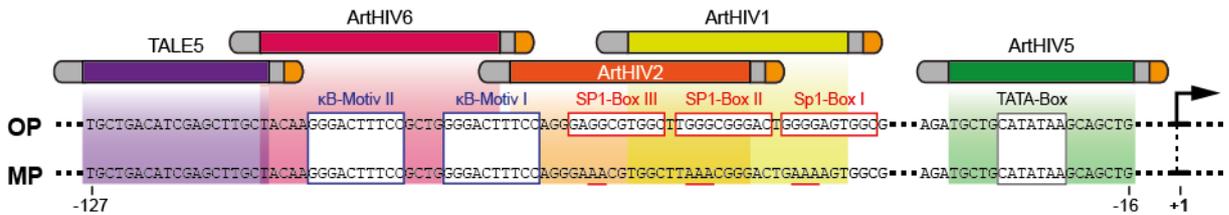


Abb. 3.2.3: TALEs am originalen HIV-1-Promotor und einer SP1-Bindestellen-Promotormutante. Die TALEs ArtHIV1, 2, 5, 6 und TALE5 (Geißler *et al.*, 2015), gezeigt mit N- und C-Terminus (grau), sowie C-terminaler VP16-AD (orange), erkennen mit ihrer DNA-Bindedomäne die farblich gleich markierte Zielsequenz im originalen HIV-1-Promotor (OP). Die im OP enthaltenen zwei κ B-Motive, drei SP1-Bindestellen (SP1-Boxen), TATA-Box (Van Opijnen *et al.*, 2004) und Transkriptionsstart (+1) sind gezeigt. Im Vergleich dazu sind die Verhältnisse an einer in den SP1-Boxen mutierten Promotorvariante (MP) gezeigt. Die G zu A-Mutationen, die zur Zerstörung der SP1-Boxen führen sollten, sind rot unterstrichen.

3.2.3.1 Nachweise der Bindung des HIV-1-Promotors mit mutierten SP1-Bindestellen durch Derivate von ArtHIV1 und 2

Um zu überprüfen, ob ArtHIV1 und 2 tatsächlich in der Lage sind auch den mutierten Promotor zu erkennen, wurden DNA-Bindestudien durchgeführt. Zum vergleichenden Test der DNA-Bindung von ArtHIV1 und 2 am OP und MP *in vitro* wurden 6x His-markierte, N- und C-terminal verkürzte Derivate (H_6 -ArtHIV1/2) dieser Proteine in *E. coli* produziert und gereinigt. Die gewählten Verkürzungen auf N-terminal 137 und C-terminal 17 verbleibende Aminosäuren sollten zur Stabilisierung der gereinigten Proteine beitragen und dabei das spezifische DNA-Bindevermögen der TALEs erhalten (Szurek *et al.*, 2002; Miller *et al.*, 2011; Mussolino *et al.*, 2011). Dazu enthält der gewählte N-Terminus die in TALE-Kristallstrukturen beobachteten, bis zu vier degenerierten *repeats* -3 bis 0, von denen angenommen wird, dass sie für die DNA-bindende Aktivität des N-Terminus wichtig sind (Deng *et al.*, 2012; Mak *et al.*, 2012; Gao *et al.*, 2012; Stella *et al.*, 2013). Des Weiteren ist bekannt, dass ein Tryptophan in *repeat* -1 ähnlich einem RVD für die Spezifizierung und Koordination des initialen T notwendig ist (Mak *et al.*, 2012; Lamb *et al.*, 2013; Tsuji *et al.*, 2013; Schreiber *et al.*, 2014). Unter Verwendung dieser Proteinpräparate in unterschiedlicher Konzentration und markierten 61 bp-Fragmenten des OP und MP, die die Bindestellen beider TALEs enthielten, wurden EMSA (*electrophoretic mobility shift assay*) durchgeführt (Abb. 3.2.4). Dabei wird das Laufverhalten einer markierten DNA-Spezies im elektrischen Feld durch Komplexbildung in der vorausgegangenen Bindungsreaktion mit dem Protein verändert und das Signal im Verhältnis zur freien DNA verschoben. Diese Verschiebung visualisiert somit die DNA-Bindung des Proteins. Für H_6 -ArtHIV1 und 2 konnte Komplexbildung ab einer Konzentration von 16,6 nM nach einer Bindungsreaktion sowohl mit 16,6 nM markierten OP- als auch MP-Fragmenten beobachtet

werden (Abb. 3.2.4). Um die Spezifität dieser Bindungen zu überprüfen wurden Kompetitionsanalysen mit verschiedenen nicht markierten DNA-Spezies durchgeführt. Zum einen wurden hierzu die von den Proteinen gebundenen DNA-Fragmente des OP und MP verwendet, zum anderen DNA-Fragmente, die nicht erkannt werden sollten (Abb. 3.2.4).

H₆-ArtHIV1 wurde für die Kompetitionsanalysen 33,2 nM und damit im 2x Überschuss zu den verwendeten markierten DNA-Spezies eingesetzt (Abb. 3.2.4 A). Bei der Konkurrenz der Bindung von H₆-ArtHIV1 sowohl an das markierte OP- als auch MP-Fragment mit der nicht markierten OP-DNA wurde ein Ausbleiben des Signals erreicht (Abb. 3.2.4 A). Ein vollständiges Ausbleiben der Signale für die H₆-ArtHIV1-Komplexe mit den markierten OP- und MP-Fragmenten wurde nur mit 100x Überschuss des nicht markierten MP-Fragmentes erreicht, wobei das bei 25x Überschuss verbleibende Signal des H₆-ArtHIV1-OP-Komplexes deutlich stärker als das des H₆-ArtHIV1-MP-Komplexes ausfiel. Die Kompetitionsanalyse mit dem nicht markierten DNA-Fragment ohne ArtHIV1-Bindestelle zeigte im Vergleich kaum eine Abschwächung des Komplexsignals mit dem markierten OP-Fragment (Abb. 3.2.4 A). Die Konkurrenz des H₆-ArtHIV1-MP-Komplexes fiel hingegen stärker aus, sodass bei 100x Überschuss des nicht markierten DP-Fragmentes kein Signal mehr nachzuweisen war (Abb. 3.2.4 A). Im Falle von 16,6 nM H₆-ArtHIV2 wurde sowohl die Bindung des markierten OP- als auch des MP-Fragmentes durch Konkurrenz mit eben diesen spezifischen Ziel-DNA-Sequenzen in nicht markierter Form in jeder eingesetzten Menge aufgehoben (Abb. 3.2.4 B). Die Bindung von H₆-ArtHIV2 an das OP-Fragment konnte im Vergleich dazu durch ein 47 bp-Fragment mit randomisierter DNA-Sequenz nur teilweise und nur im 100x Überschuss kompetiert werden (Abb. 3.2.4 B). Der Komplex von H₆-ArtHIV2 mit dem MP-Fragment wurde hingegen durch die randomisierte DNA im 100x Überschuss vollständig aufgelöst und geringere Mengen dieses DNA-Fragmentes (25x, 50x) führten bereits zu einer Schwächung des Komplexsignals (Abb. 3.2.4 B).

Da die Konkurrenz mit den nicht markierten OP- und MP-Fragmenten im Vergleich zu den DNA-Spezies ohne bekannte ArtHIV1- oder 2-Bindestelle deutlich stärker ausfiel, kann eine spezifische Bindung des OP und MP-Fragmentes durch die TALEs angenommen werden. Für beide Proteine ließ sich auch eine schwächere Bindung an das MP-Fragment im Vergleich zum bevorzugt gebundenen OP-Fragment nachweisen (Abb. 3.2.4). Dies ist auf die bekannte, geringere Effizienz der Bindung des RVDs NN an Adenin im Vergleich zu Guanin zurückzuführen (Streubel *et al.*, 2012; Le Cong *et al.*, 2012; Christian *et al.*, 2012; Meckler *et al.*, 2013).

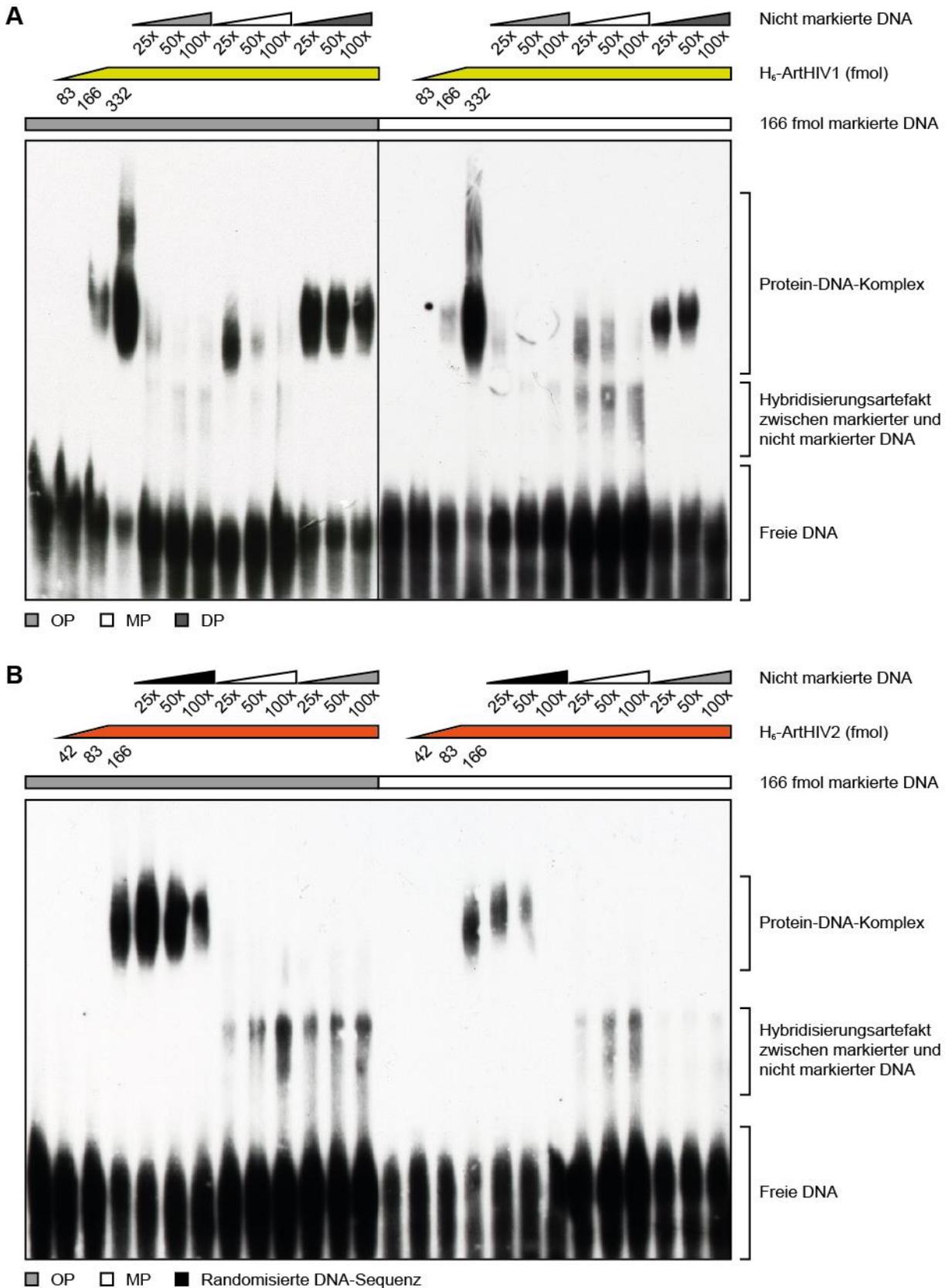


Abb. 3.2.4: Die DNA-Bindedomänen von ArthHIV1 und 2 binden ein Fragment des originalen HIV-1-Promotors mit höherer Affinität als ein Fragment der Promotormutante. (A, B) EMSA (*electrophoretic mobility shift assay*) mit (A) ArthHIV1- und (B) ArthHIV2-Derivaten an Fragmenten des originalen HIV-1-Promotors (OP) aus 8E5-Zellen und des in den SP1-Bindestellen mutierten HIV-1-Promotors (MP). Die in *E. coli* produzierten, gereinigten, 6x His-markierten Proteine mit N-terminalen 137 und C-terminalen 17 Aminosäuren

(H₆-ArtHIV1/2) verlagern bei Bildung eines Protein-DNA-Komplexes in der 30-minütigen Bindungsreaktion mit der markierten DNA deren Signal als freie DNA (*Shift*). Die Proteinlösungen und die verschiedenen DNA-Spezies wurden so eingesetzt, dass in dem verwendeten Volumen von 10 µl die angegebenen Stoffmengen enthalten waren. Als markierte DNA wurden 61 bp OP- und MP-Fragmente von -96 bis -35 gemessen am natürlichen Transkriptionsstart am OP in konstanter Menge dargeboten. Durch Titration wurden geeignete Proteinmengen für Kompetitionsanalysen bestimmt. Dabei wurden die nicht markierten Kompetitor-DNA-Spezies im 25x (4,15 pmol), 50x (8,3 pmol) oder 100x (16,6 pmol) Überschuss zugegebenen. Außer den beschriebenen OP- und MP-Fragmenten wurden dafür auch DNA Fragmente ohne die Bindestellen beider TALEs verwendet. Diese waren ein 61 bp-Fragment des zum Transkriptionsstart distalen Promotorbereiches (DP) von -304 bis -243 in (A) und ein 47 bp-Fragment mit einer randomisierten DNA-Sequenz in (B).

3.2.3.2 Synergistische Aktivierung an der SP1-Bindestellenmutante des HIV-1-Promotors durch TALEs

Zur Untersuchung des Einflusses der SP1-Bindestellen im HIV-1-Promotor auf die Aktivierbarkeit durch die artifiziiellen TALEs ArtHIV1, 2, 5, 6 und TALE5 (Geißler *et al.*, 2015) wurden Luciferasereporterstudien mit dem OP und MP in Hek293T-RExTM durchgeführt (Abb. 3.2.5 A). Als Kontrolle diente Hax3-VP16, ein TALE, der weder im OP noch im MP binden kann. Am OP waren nur die beiden am weitesten stromaufwärts liegenden TALEs, ArtHIV6 auf den NFκB-Bindestellen und TALE5, zu einer mindestens siebenfachen Aktivierung dieses Luciferasereporters in der Lage (Abb. 3.2.5 A). Weiterhin induzierte ArtHIV5 zumindest geringe Reporteraktivität am OP, wohingegen durch ArtHIV1, sowie ArtHIV2 gar keine Aktivierung an diesem Promotor nachzuweisen war (Abb. 3.2.5 A). Dabei wird deutlich, dass die Basalaktivität des MP im Vergleich zum OP stark reduziert ist (Abb. 3.2.5 A). Dies zeigt zum einen, dass die Mutation der SP1-Bindestellen erfolgreich war, und zum anderen bestätigt es die wichtige Rolle von SP1 bei Aktivierung der HIV-1-Transkription (Jones *et al.*, 1986; Ross *et al.*, 1991; Perkins *et al.*, 1993; Burnett *et al.*, 2009; Flint *et al.*, 2009). Beim Test der einzelnen für HIV-1 erstellten, artifiziiellen TALEs mit dem MP lässt sich im Gegensatz zum OP für jeden der TALEs eine Aktivierung des Luciferasereporters nachweisen (Abb. 3.2.5 A). Obwohl die dabei erreichten Luciferaseaktivitäten die Basalaktivität des OP nicht übersteigen, entspricht dies jedoch für alle fünf TALEs einer mindestens 7,5-fachen Induktion des MP. Dabei ist die von ArtHIV5 an der Position der TATA-Box induzierte Aktivität geringer als die durch die vier anderen, stromaufwärts bindenden TALEs (Abb. 3.2.5 A). Diese Unterschiede in der Aktivierung des OP und MP durch die einzelnen TALEs zeigen, dass die SP1-Bindestellen einen Einfluss auf die TALE-abhängige Aktivierung am HIV-1-Promotor haben. Dies ist insbesondere für ArtHIV1 und 2, die Sequenzen im Bereich der SP1-Bindestellen erkennen, deutlich. Die fehlende Aktivierung am

OP lässt sich dabei vermutlich auf eine Konkurrenz mit SP1 um die Bindestelle zurückführen. Mögliche Erklärungen wären eine fehlende DNA-Bindung der TALEs auf Grund der Präsenz von SP1 oder ein Verdrängen der Zinkfingerproteine von den Bindestellen, wobei die TALE-induzierte Aktivität die SP1-induzierte Aktivität nur substituiert und nicht weiter erhöht. Starke Aktivierung des HIV-1-OP ist hingegen nur von Positionen stromaufwärts der SP1-Bindestellen möglich (Abb. 3.2.5 A: ArtHIV6, TALE5). Kann SP1 im Falle des MP jedoch nicht binden, ist die durch diese TALEs induzierte, absolute Aktivität ebenfalls bezogen auf den OP gering. SP1 scheint also auch für eine effiziente Aktivierung des HIV-1-Promotors durch TALEs notwendig zu sein. Die mit ArtHIV6 und TALE5 am OP erzielten Luciferaseaktivitäten sind dabei höher als die Summe der Basalaktivität des OP, die als Entsprechung der Induktion durch SP1 betrachtet werden kann, und der durch den jeweiligen TALE am MP ohne SP1 induzierten Aktivität (Abb. 3.2.5). Dieser mehr als additive, kooperative Effekt auf die Transkriptionsinduktion durch mehrere Aktivatoren wird als transkriptionelle Synergie bezeichnet (Lin *et al.*, 1990; Joung *et al.*, 1994; Chi *et al.*, 1995). Um auch durch artifizielle TALEs eine möglichst effiziente transkriptionelle Aktivierung zu erreichen wurde bereits in mehreren Studien ein synergistischer Ansatz gewählt (Perez-Pinera *et al.*, 2013 a; Maeder *et al.*, 2013 a). Es konnte dabei beobachtet werden, dass der Einsatz mehrerer synthetischer TALEs für einen Zielpromotor in jedem Falle zu einer um ein Vielfaches höheren Induktion führte als mit den einzelnen TALEs (Perez-Pinera *et al.*, 2013 a; Maeder *et al.*, 2013 a). Eine mehr als additive, also synergistische Induktion konnte nicht für alle, aber für die meisten Zielpromotoren festgestellt werden (Perez-Pinera *et al.*, 2013 a; Maeder *et al.*, 2013 a). Auf Grund dieser Beobachtungen war es auch von Interesse die kooperative Induzierbarkeit des originalen HIV-1-Promotors im Vergleich zu der SP1-Bindestellenmutante, für die bereits fünf verschiedene artifizielle TALEs zur Verfügung standen, zu testen (Abb. 3.2.5 A). Dazu wurde von jedem TALE-Plasmid in Kombinationen oder einzeln die gleiche Menge eingesetzt. Die fehlende DNA zur konstanten Transfektionsmenge wurde durch das Hax3-VP16-codierende Kontrollplasmid ergänzt. Auf diese Weise konnte am OP weder mit Zweier- oder Dreierkombinationen noch mit allen TALEs eine Steigerung der Aktivität des jeweils stärksten eingesetzten TALEs und dementsprechend keine Synergie der TALEs festgestellt werden (Abb. 3.2.5 A). Bei einigen TALE-Kombinationen am OP lässt sich jedoch eine etwas geringere Luciferaseaktivität als die des stärksten TALE beobachten (Abb. 3.2.5 A). Hierbei ist zu beachten, dass die Ziel-DNA-Sequenzen mehrerer TALEs im HIV-1-Promotor überlappen, so von ArtHIV6 und TALE5, sowie ArtHIV1 und ArtHIV2, oder im Falle von ArtHIV2 und ArtHIV6 sehr dicht

nebeneinander liegen (Abb. 3.2.3). Bei der Kombination dieser TALEs könnte es zu einer sterischen Verhinderung einer simultanen Bindung an die DNA kommen, was eine weniger effiziente Aktivierung verursachen könnte. An dem MP zeigt sich mit den verschiedenen TALE-Kombinationen ein anderes Bild (Abb. 3.2.5 A). Bereits die meisten Paarungen artifizierender TALEs für den HIV-1-Promotor induzierten am MP eine im Vergleich zu den von den einzelnen TALEs induzierten Aktivitäten gesteigerte Luciferaseaktivität. Dies wurde nicht für die Kombination von ArtHIV6 und TALE5, deren Bindestellen überlappen, beobachtet. Wurde jedoch einer dieser beiden TALEs mit ArtHIV1 oder 2, deren Ziel-DNA-Sequenzen über den mutierten SP1-Bindestellen liegen, kombiniert, war mit Ausnahme von ArtHIV2 und 6, deren Bindestellen dicht nebeneinander liegen, eine synergistische Aktivierung zu beobachten. Darüber hinaus wurde für alle Kombinationen von drei oder mehr TALEs eine bis zu 118-fache synergistische Aktivierung nachgewiesen. Die dabei durch einige Kombinationen am MP erreichten absoluten Luciferaseaktivitäten entsprachen sogar den durch diese Kombinationen am OP induzierten Aktivitäten, wobei dies in Bezug auf die Basalaktivität des OPs nur einer ca. vierfachen Induktion entsprach. Die vollständige Synthese der einzelnen, artifizierender TALEs in Hek293T-RExTM-Zellen wurde im Immunoblot verifiziert (Abb. 3.2.5 B). Von den aufgeführten Ergebnissen lassen sich folgende Annahmen ableiten. Zum einen scheint die Höhe der mit artifizierender TALEs erreichbaren Induktionsrate an einem Promotor stark von dessen basaler Aktivität abzuhängen. Zum anderen kann angenommen werden, dass synergistische Effekte nicht nur zwischen mehreren synthetischen Transkriptionsfaktoren sondern auch zwischen diesen und endogenen Transkriptionsfaktoren wie SP1 zu einer effizienten Aktivierung eines Zielpromotors führen können. Ein einzelner günstig positionierter, künstlicher Aktivator, wie hier ArtHIV6 oder TALE5, könnte demnach für die Induktion eines Zielgens bereits maximal effizient sein.

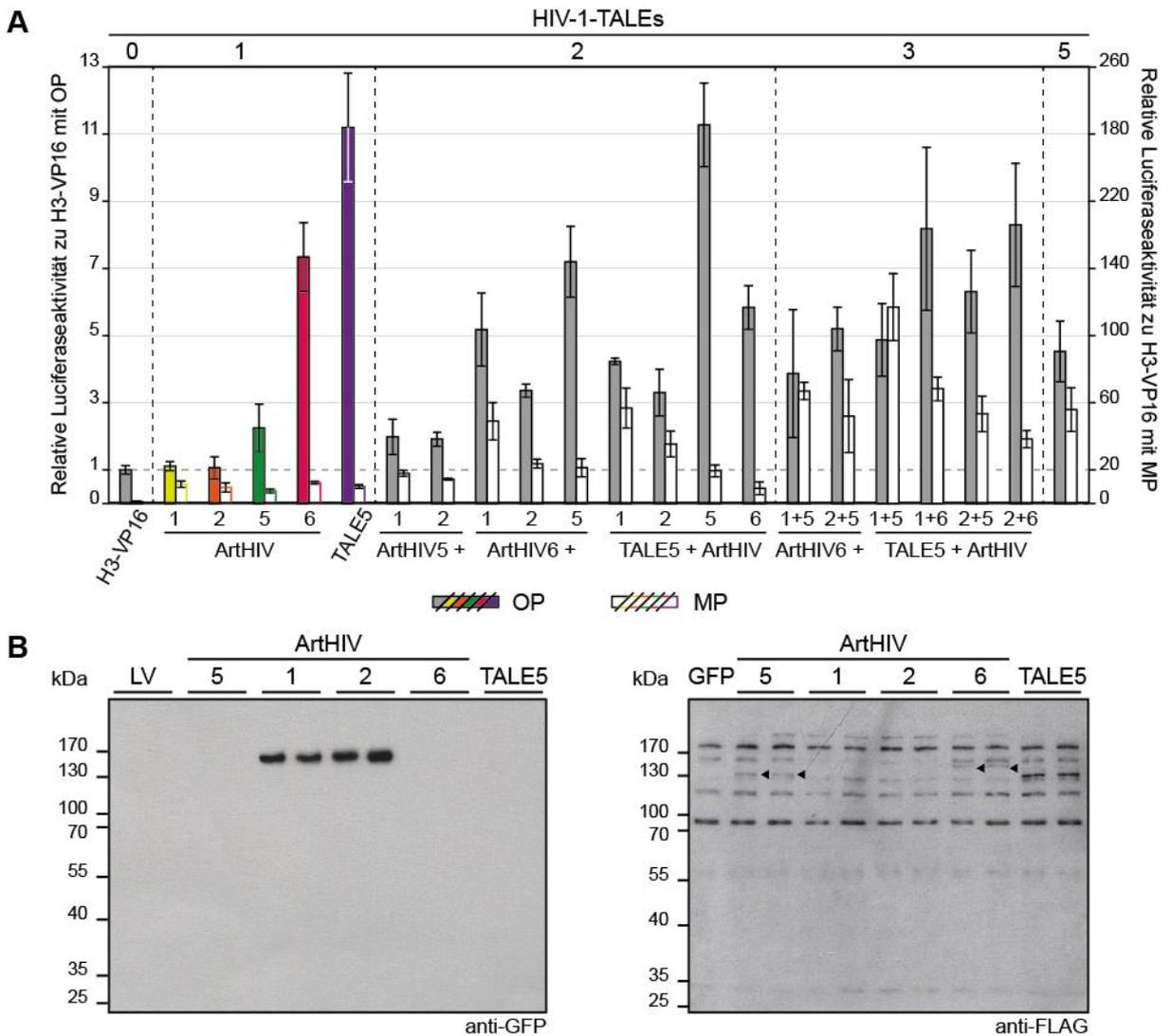


Abb. 3.2.5: Die TALE-Aktivatoren ArtHIV1, 2, 5, 6 und TALE5 induzieren Luciferasereporterogenaktivität am HIV-1-Promotor und einer SP1-Bindestellen-Promotormutante auf unterschiedliche Weise.

(A) Luciferase-Reportergenanalyse der HIV-1-TALE-Aktivatoren mit dem Originalpromotor (OP, pFAHPI) und dem in den SP1-Bindestellen mutierten Äquivalent (MP, $n=3$). Die Luciferase-Reporterkonstrukte wurden im Verhältnis 1:9 zusammen mit konstitutiv CMV-getriebenen TALE-Expressionsvektoren in Hek293T-RExTM-Zellen co-transfiziert. Dabei wurde von jedem TALE-Plasmid in jeder Kombination als auch einzeln die gleiche Menge eingesetzt. Die fehlende DNA zur konstanten Transfektionsmenge wurde durch ein Hax3-VP16 (H3-VP16)-codierendes Plasmid ergänzt. Als Kontrolle diente das jeweilige Reporterkonstrukt mit H3-VP16. Fehlerindikatoren kennzeichnen Standardabweichungen. (B) Exemplarische Immunoblots der HIV-1-TALEs ArtHIV1, 2, 5, 6 und TALE5. ArtHIV1 und 2 tragen ein GFP-, ArtHIV5, 6 und TALE5 ein FLAG-Epitop. Zwei Lysate von Hek293T-RExTM-Zellen, die mit dem jeweiligen TALE-Expressionsvektor und entweder mit dem Leervektor pcDHN (LV) und dem OP-Reporterkonstrukt (links) oder einem *gfp*-Expressionsvektor und dem MP-Reporterkonstrukt (rechts) co-transfiziert waren, wurden aufgearbeitet, sodass jeweils ca. 36 μ g Totalprotein pro Spur aufgetrennt werden konnten. Schwarze Pfeilspitzen kennzeichnen die Signale von ArtHIV5 und 6. Als Kontrollen wurden gleich behandelte Lysate von nur mit LV und OP-Reporterkonstrukt transfizierten (links) oder in Anwesenheit des MP-Reporterkonstruktes nur *gfp*-exprimierenden (rechts) Hek293T-RexTM-Zellen verwendet.

3.2.4 Verstärkte Geninduktion durch artifizielle TALEs mit mehreren Aktivierungsdomänen

Neben der Strategie effiziente Geninduktion mittels synergistischer Aktivierung durch mehrere synthetische Aktivatoren zu erreichen, konnten Aktivatoren auch durch multiple ADs verstärkt werden (Perez-Pinera *et al.*, 2013 a; Maeder *et al.*, 2013 a; Chavez *et al.*, 2014; Konermann *et al.*, 2015). Dabei wurden mehrere, unterschiedliche ADs entweder durch direkte Fusion an den C-Terminus oder durch proteinvermittelte Bindung an ein Aptamer in der sgRNA in räumliche Nähe zu einem Cas9-Aktivator (dCas9) gebracht (Chavez *et al.*, 2014; Konermann *et al.*, 2015). Dies führte im Vergleich zu den einzelnen ADs zu einer um ein vielfaches höheren Genaktivierung, wobei Kombinationen von drei verschiedenen ADs effizienter waren als Paare (Chavez *et al.*, 2014; Konermann *et al.*, 2015). In den vorgestellten Dreierkombinationen wurden stets die beiden für synthetische Aktivatoren üblichsten ADs, das VP16-Derivat VP64 und eine p65-AD, verwendet und mit einer humanen HSF1-, MYOD1- oder Epstein-Barr viralen Rta-AD kombiniert (Chavez *et al.*, 2014; Konermann *et al.*, 2015). Davon wurde die Kombination VP64-p65-Rta als C-terminal fusionierte AD, genannt VPR, eingesetzt und konnte erfolgreich auf artifizielle ZNF- und TALE-Aktivatoren übertragen werden (Chavez *et al.*, 2014).

Um die hocheffizient induzierende AD-Kombination einer dCas9-fusionierten VP64 mit einer sgRNA-Aptamer-bindenden p65-HSF1-Fusion (Konermann *et al.*, 2015) ähnlich der VPR auf TALEs zu übertragen, wurden schrittweise eine *Mus musculus* p65-AD aus pReg neo (Promega) und eine von Hek293T-RExTM-cDNA stammende HSF1-AD sequenzidentisch zu Konermann *et al.*, 2015, und daher mit einer vorrausgehenden NLS an ArHIV6 mit einer VP64-AD fusioniert (Abb. 3.2.6 A). Die entstandenen ADs werden im Folgenden als Vp (VP64-p65) und VpH (VP64-p65-HSF1) bezeichnet.

Da Kombinationen multipler ADs zu stark vergrößerten Konstrukten führen, z.B. VpH mit 1164 bp, war es auch von Interesse kleine ADs im Vergleich zu testen. Hierzu wurde zum einen die zweite Subdomäne der VP16-AD, genannt mVP16, mit autonomer transkriptioneller Aktivität benutzt (Abb. 3.2.6 A; Sullivan *et al.*, 1998). Diese Domäne entspricht nicht der in VP64 vervierfachen AS-Sequenz aus der VP16-AD, obwohl strukturelle Ähnlichkeiten bestehen (Abb. 3.2.6 A). In beiden ADs, sowie in den korrespondierenden Bereichen von VP16, ist ein Gemisch saurer, bzw. negativ geladener, und unpolarer, hydrophober AS wiederzufinden. Arbeiten in *Saccharomyces cerevisiae* haben gezeigt, dass hydrophobe AS allein bereits ausreichend für eine effiziente Genaktivierung waren, sodass hier zum anderen

ein in Hefe vielversprechendes Peptid aus sieben Tryptophanen, genannt W₇, C-terminal an ArtHIV6 fusioniert wurde (Lin *et al.*, 2010).

Zum Test des Aktivierungspotentials der ArtHIV6-Derivate mit den sieben verschiedenen AD-Varianten im Vergleich zu ArtHIV6 mit VP16-AD wurden Luciferasereporteranalysen in Hek293T-RExTM-Zellen mit zwei Reporterkonstrukten mit geringer Basalaktivität durchgeführt (Abb. 3.2.6 B). Dabei handelte es sich zum einen um das MP-Reporterkonstrukt und zum anderen um ein Minimal-CMV-Promotor-Reporterkonstrukt mit einer 5'-vorgelagerten, Box6/7 genannten, 29 bp umfassenden HIV-1-Promotorsequenz, die die ArtHIV6-Bindestelle enthält (Abb. 3.2.6 B). Um die mit den ArtHIV6-AD-Varianten am MP erreichten Luciferaseaktivitäten mit synergistischer Aktivierung vergleichen zu können, wurde ArtHIV6 zusammen mit ArtHIV1 mitgeführt (Abb. 3.2.6 B). Als Kontrolle diente H3-VP16, der in beiden Reporterkonstrukten keine Ziel-DNA-Sequenz besitzt.

Die Luciferasereporteranalysen zeigten für beide Reporterkonstrukte ein ähnliches Bild (Abb. 3.2.6 B). Die kleinen ADs mVP16 und W₇ induzierten keine oder nur sehr geringe Luciferaseaktivität (Abb. 3.2.6 B). Im Gegensatz dazu war die mit multiplen ADs beobachtete Induktion um ein Vielfaches höher als dies mit den einzelnen ADs VP64, p65 und HSF1 der Fall war (Abb. 3.2.6 B). Eine stärkere Induktion durch VpH im Vergleich zu Vp, sowie eine definitiv synergistische und nicht additive Aktivierung bezogen auf die Aktivitäten der einzelnen ADs, konnte jedoch nur am Box6/7-Reporterkonstrukt beobachtet werden (Abb. 3.2.6 B). Am MP entsprach die Aktivierung durch Vp und VpH allerdings zumindest den Wert, den auch die Kombination von ArtHIV1 und 6 mit VP16-AD erreichte (Abb. 3.2.6 B). Beim Vergleich der einzelnen ADs wurde an beiden Promotoren deutlich, dass die bereits selbst ein Tetramer einer VP16-Subdomäne darstellende VP64-AD eine effizientere Aktivierung als VP16, p65 und HSF1 vermittelte (Abb. 3.2.6 B). Die durch den ArtHIV6-VP64 induzierte Aktivität entsprach jeweils ungefähr einer Verdopplung der durch VP16 induzierten Aktivität, die wiederum zumindest am Box6/7-Reporterkonstrukt stärker ausfiel als bei p65 und HSF1 und auch am MP diesen nicht unterlegen war (Abb. 3.2.6 B). Die Expression, bzw. co-Expression, aller erstellten, hier verwendeten TALE-Gene konnte mit dem korrekten Molekulargewicht im Immunoblot nachgewiesen werden (Abb. 3.2.6 C).

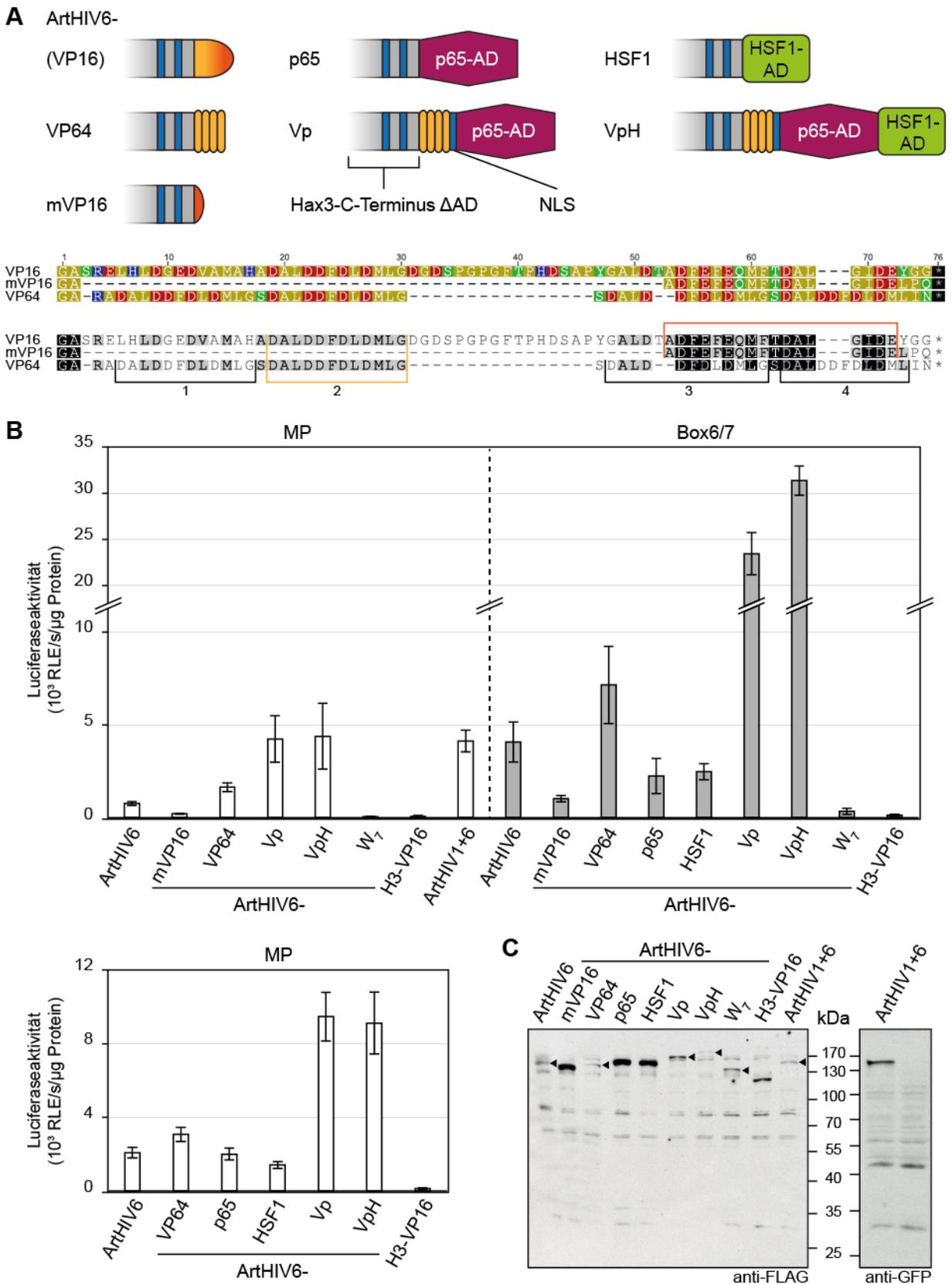


Abb. 3.2.6: ArtHIV6-Derivate mit multiplen ADs aktivieren unterschiedliche Luciferasereporterkonstrukte mit erhöhter Effizienz. (A) Schematische Darstellung der verschiedenen AD-Konstrukte (oben) außer W₇ und ein in Geneious Pro 5.3.6 erstelltes AS-alignment der verwendeten VP16-AD-Derivate im Ein-Buchstabencode, eingefärbt nach Polarität (ocker – unpolar, grün – polar, blau – positiv geladen, rot – negativ geladen; Mitte) oder

AS-Ähnlichkeit untereinander (schwarz – ähnliche AS in drei Sequenzen, grau - ähnliche AS in zwei Sequenzen; unten). Die an den Hax3-C-Terminus ohne natürliche AD von ArtHIV6 fusionierte VP16-AD wurde durch die jeweilige AD ersetzt. Die obere, dunkelorange Klammer am unteren *alignment* markiert den in mVP16 enthaltenen Bereich der VP16-AD. Die untere, hellorange Klammer am unteren *alignment* kennzeichnet die in VP64 viermal enthaltene VP16-AS-Sequenz, deren Wiederholungen durch schwarze Klammern verdeutlicht und nummeriert sind. (B) Luciferase-Reporteranalysen der ArtHIV6-Derivate mit dem MP-Reporterkonstrukt und einem Minimal-CMV-Promotorreporterkonstrukt mit einem Box6/7 genannten Abschnitt des HIV-1-Promotors, der die ArtHIV6-Bindestelle enthält ($n=3$). Das jeweilige Luciferase-Reporterkonstrukt wurde im Verhältnis 1:9 zusammen mit konstitutiv CMV-getriebenem TALE oder TALEs im Verhältnis 1:1 (ArtHIV1+6) in Hek293T-REx™-Zellen co-transfiziert. Als Kontrolle diente das jeweilige Reporterkonstrukt mit H3-VP16. Fehlerindikatoren kennzeichnen Standardabweichungen (C) Immunoblots der verwendeten TALEs. Die Zelllysate einer Dreifachbestimmung des jeweiligen Proteins oder der jeweiligen Proteine am MP-Reporterkonstrukt wurden zu gleichen Teilen gemischt und wie unter Material und Methoden beschrieben aufgearbeitet, sodass jeweils ca. 7,5 µg Totalprotein aufgetrennt wurden. Mit Ausnahme des GFP-markierten ArtHIV1 sind alle TALEs N-terminal mit einem FLAG-Epitop fusioniert. Schwache TALE-Signale im linken Bild sind durch einen schwarzen Pfeil markiert.

Die vorliegenden Ergebnisse bestätigen die in der Literatur beobachteten Phänomene (Chavez *et al.*, 2014; Konermann *et al.*, 2015) und erweitern das Repertoire für TALEs etablierter, multipler ADs um eine VP64-p65-HSF1-Fusion (Abb. 3.2.6 A, B). Darüber hinaus konnte gezeigt werden, dass synthetische Aktivatoren mit multiplen ADs so effizient wie mehrere, synergistisch agierende Aktivatoren mit einfacher AD induzieren können (Abb. 3.2.6 B). Ein artifizieller TALE mit einem in Hefe effizient Gene induzierenden, hydrophoben Peptid (Lin *et al.*, 2010), ArtHIV6-W₇, war hingegen kaum zur Geninduktion eines Luciferasereporters in der Lage (Abb. 3.2.6 B), was auf Unterschiede in der benachbarten AS-Sequenz der fusionierten Aktivatoren zurückzuführen sein könnte oder auf Abweichungen im Mechanismus der Geninduktion zwischen Hefe und Säugern hinweisen könnte.

3.2.5 TALE-Aktivatoren für HIV-1 Subtypen

Neben ArtHIV6 als geeignetem TALE-Aktivator für HIV-1 konnte in einer Studie der synthetische TALE5, der unmittelbar stromaufwärts von ArtHIV6 im LTR bindet, als effizientester Aktivator in 8E5-Zellen identifiziert werden (Abb. 3.2.7 A; Geißler *et al.*, 2015). Diese beiden TALEs wurden für den Promotor im LTR des Provirus von 8E5-Zellen entworfen und sind somit nur für ein bestimmtes HIV-1-Isolat spezifisch. Da HIV-1 als Retrovirus einer hohen Mutationsrate unterliegt und daher in hoher Diversität auftritt, werden die Zielsequenzen der gefundenen, funktionalen TALEs in vielen HIV-1-Isolaten so verändert sein, dass keine Erkennung mehr stattfinden kann. In Geißler *et al.*, 2015, wurde dieses Problem bei der Anwendung von TALE5 zur HIV-1-Induktion in Patientenzellen durch Sequenzierung des in

den isolierten Zellen vorherrschenden Provirus adressiert. Für veränderte Zielsequenzen wurde ein neuer, für diese spezifischer TALE erstellt und verwendet. Obgleich eine individualisierte Anwendung dank effizienter Klonierungsmethoden für TALEs möglich ist, kann nicht davon ausgegangen werden, dass alle vorhandenen Proviren in einem infizierten Individuum im Bereich der TALE-Bindestelle identisch sind. Kalkulationen ergaben, dass allein in einem Patienten täglich mehrere Millionen Virusvarianten entstehen (Perelson *et al.*, 1996; besprochen in Taylor *et al.*, 2008). Global betrachtet, existieren allerdings zahlreiche Viruspopulationen der genetisch distinkten HIV-1 Subtypen A-H, J und K in unterschiedlichen Verbreitungsgebieten (besprochen in Taylor *et al.*, 2008). Beim Vergleich mit den LTR-Sequenzen von 60 verschiedenen Isolaten aus den Subtypen A-G wird der HIV-1 LTR aus 8E5 phylogenetisch in Subtyp B eingeordnet (De Baar *et al.*, 2000). Dabei weisen die Sequenzen eine hohe Diversität auch im Bereich der Zielsequenzen von TALE5 und ArtHIV6 auf (De Baar *et al.*, 2000). Per Auge waren vor allem Substitutionen im 3'-Bereich der TALE5- und dementsprechend im 5'-Bereich der ArtHIV6-Bindestelle festzustellen. Dies entspricht der Region unmittelbar stromaufwärts der beiden, für die Aktivierung der viralen Transkription wichtigen NFκB-Bindestellen. Zwischen diesen stark konservierten *cis*-Elementen, sowie direkt stromauf- und stromabwärts waren in einigen Isolaten Indels an verschiedenen Positionen der ArtHIV6-Zielsequenz zu finden. Darüber hinaus gab es größere Sequenzumstellungen, die für bestimmte Subtypen charakteristisch waren. Dabei handelte es sich zum einen um eine zumindest hier für Subtyp C spezifische dritte NFκB-Bindestelle, durch die die Sequenz im 3'-Bereich der TALE5-Zielsequenz stark verändert und die ArtHIV6-Zielsequenz stromaufwärts in das Zentrum der TALE5-Zielsequenz verschoben wird. Zum anderen waren Sequenzverdopplungen zu beobachten, die nicht ausschließlich aber zumindest in allen LTR-Sequenzen vom Subtyp F dieses Datensets zu finden waren. Diese ließen in der Regel eine zweite, versetzte TALE5-Bindestelle entstehen. Es war davon auszugehen, dass diese hohe Sequenzvielfalt die Induzierbarkeit der Transkription durch TALE5 oder ArtHIV6 in vielen Fällen eingeschränkt oder verhindert hätte. Daher erschien dieser Sequenzvergleich als geeigneter Hintergrund für das Design von TALE-Aktivatoren mit multipler Sequenzspezifität.

Um TALE-Aktivatoren zu erstellen und zu testen, die die Bindestellen von TALE5 und ArtHIV6 in verschiedenen HIV-1 Subtypen erkennen, wurden aus den beschriebenen Isolaten subtyp-spezifische Sequenzen ausgewählt (Abb. 3.2.7 A; De Baar *et al.*, 2000). Für Subtyp B wurde die bereits aus 8E5-Zellen in Reporterkonstrukte übertragene LTR-Sequenz verwendet (pFAHPs). Aus Subtyp G wurde kein Isolat ausgewählt, da dadurch auf Grund lokaler

Ähnlichkeiten mit Isolaten anderer Subtypen, vor allem A, keine erhöhte Variabilität in beiden TALE-Zielsequenzen erreicht werden konnte (De Baar *et al.*, 2000). Zur Erstellung der Reporterkonstrukte für die Subtypen A, C-F wurde das Subtyp B-Reporterkonstrukt pFAHPs mit dem Promotor aus dem 8E5-Provirus verwendet. Der Sequenzabschnitt von Position -127 bis -85 relativ zum natürlichen und ArtHIV6-induzierten Transkriptionsstart wurde gegen die Sequenz der ausgewählten Isolate ausgetauscht, die die Zielsequenzen beider TALEs enthalten (Abb. 3.2.7 A). Dabei wurden alle drei NFκB-Bindestellen aus Subtyp C, sowie die beiden, potentiellen TALE5-Zielsequenzen des ausgewählten Subtyp F-Isolates mit übertragen.

Im *alignment* der ausgewählten Sequenzen mit den RVD-Abfolgen beider TALEs werden die RVD-Base-Fehlpaarungen (RBF) deutlich (Abb. 3.2.7 A). Anhand dieser RBF wurden neue TALEs, TALEABC5 und ArtABC6, für die Zielsequenzen von TALE5 und ArtHIV6 zur Erkennung aller ausgewählten Isolatsequenzen gestaltet (Abb. 3.2.7 B). Dabei wurden auch die beiden möglichen TALE5-Bindestellen in der Subtyp F-Sequenz berücksichtigt. Um TALEABC5 an die vornehmlich im Zentrum und in der 3'-Hälfte der TALE5-Bindestellen vorhandenen RBF anzupassen wurden die RVDs der *repeats* Nr. 8 und 15 zu N* und der *repeats* Nr. 9, 13, 14 und 16 zu NS ausgetauscht (Abb. 3.2.7 B). Diese RVDs ermöglichen die Erkennung mehrerer DNA-Basen (Abb. 3.2.7 C). Während N* Pyrimidine erkennt, ist NS in der Lage alle vier DNA-Basen zu erkennen (Streubel *et al.*, 2012). Diese RVDs wurden auch zur Erstellung von ArtABC6 mitverwendet (Abb. 3.2.7 B). So wurden N* *repeats* an Position 10 und 23 und NS *repeats* an den Positionen 2 bis 4, sowie an Position 22 eingesetzt. Außerdem wurde NI an Position 1 gegen ein NN ausgetauscht, das neben Guanin auch Adenin erkennt (Streubel *et al.*, 2012). Neben Änderungen der RVDs wurde in ArtABC6 auch ein künstlicher aberranter *repeat* mit 40 As und dem RVD NI für Adenin eingesetzt (Abb. 3.2.7 B, C). Dieser sollte eine verbesserte Erkennung der Subtyp E- und F-Sequenzen durch den TALE erlauben, indem er die Bindung einer um -1 Nukleotid verschobenen Zielsequenz ermöglicht. Für die Erstellung der neuen TALEs wurde die bereits für ArtHIV6 verwendete Golden TALE Technologie – erweitert um die verschiedenen, neuen *repeat*-Typen – und eine VP16-AD verwendet (Geißler *et al.*, 2011; Streubel *et al.*, 2012; Richter *et al.*, 2014).

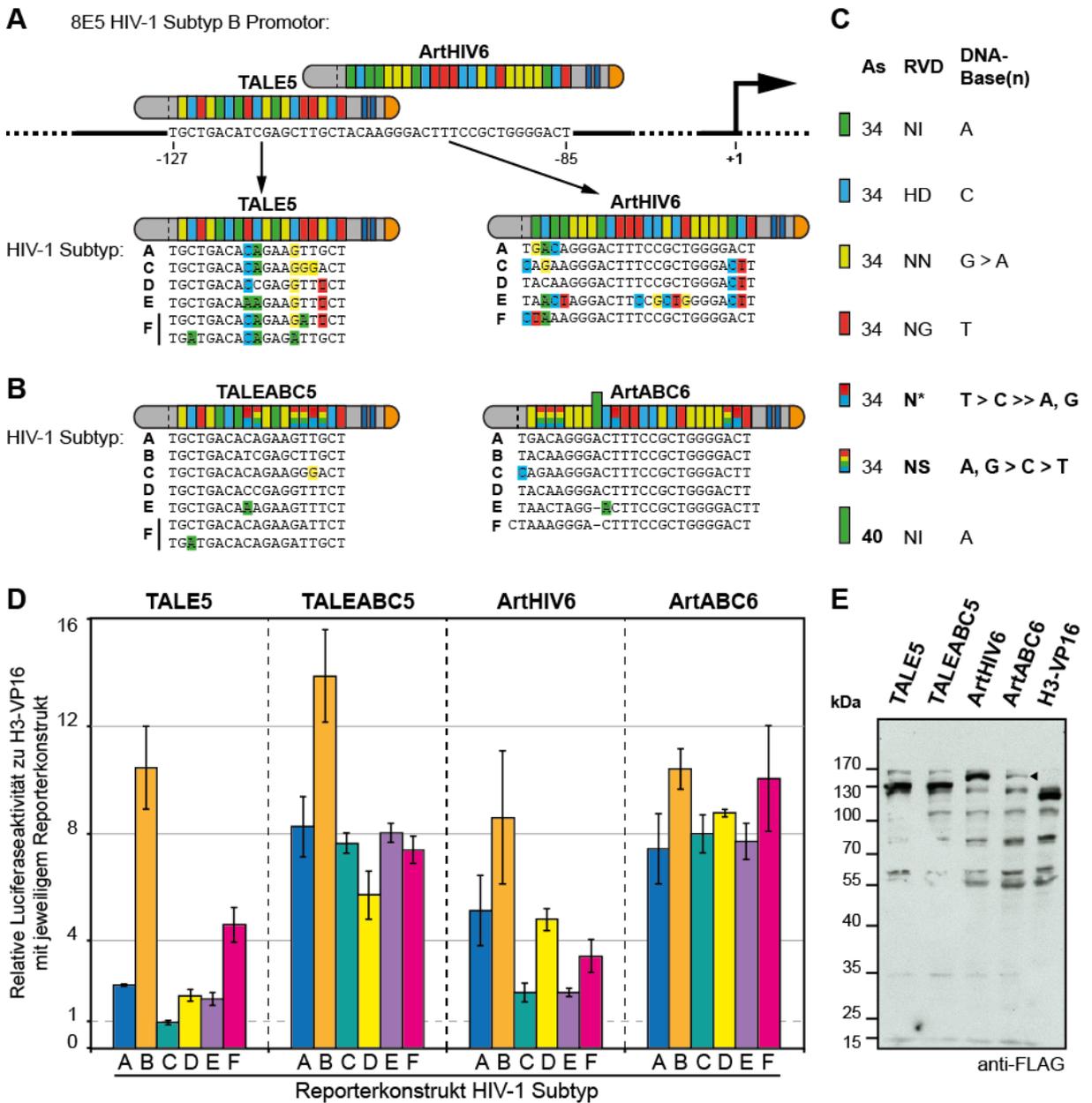


Abb. 3.2.7: RVDs multipler Spezifität und aberrante repeats ermöglichen künstlichen TALEs an Promotorsequenzen verschiedener HIV-1 Subtypen Reporteraktivität zu induzieren. (A) Lage der Zielsequenzen von TALE5 und ArtHIV6 im HIV-1 Subtyp B-Promotor aus 8E5, sowie *alignment* der TALEs mit möglichen Zielsequenzen aus anderen HIV-1 Subtyp Isolaten. An den N-Terminus (grau) beider TALEs wurde ein FLAG-Epitop (nicht gezeigt) fusioniert. Ihr C-Terminus (grau) enthält die natürlichen funktionalen NLS (blau), die natürliche AD wurde durch eine VP16-AD (orange) substituiert. Der Sequenzabschnitt von Position -127 bis -85 relativ zum natürlichen und ArtHIV6-induzierten Transkriptionsstart am HIV-1 Subtyp B-Promotor aus 8E5 (hier: *pHIV-1** in pFAHPs) wurde zur Erstellung der Reporterkonstrukte für die Subtypen A, C-F gegen die Sequenz der entsprechenden, im Jahr 2000 in De Baar *et al.* publizierten Isolate ausgetauscht, die die im *alignment* gezeigten Zielsequenzen enthalten (*Accession*: A, AF196743; C, AF196731; D, AF196726; E, AF196746; F, AF196762). Von den beiden im Reporterkonstrukt Subtyp F enthaltenen TALE5-Bindestellen beginnt die Untere stromaufwärts der Oberen. (B) *Alignment* der unter (A) beschriebenen HIV-1 Subtypsequenzen mit den rekonstruierten TALEs TALEABC5 und ArtABC6, die zur Adressierung all dieser Sequenzen RVDs multipler Spezifität und einen aberranten *repeat* enthalten. (A, B) Farbige hinterlegte DNA-Basen zeigen RVD-

Base-Fehlpaarungen. (C) As, RVDs und DNA-Basenspezifitäten der in (A) und (B) gezeigten *repeats*. RVDs und *repeats*, die den TALEs neue Eigenschaften verleihen, sind fett dargestellt. (D) Luciferase-Reporteranalysen der TALEs mit den Reporterkonstrukten verschiedener HIV-1 Subtypen ($n=3$). Luciferase-Reporterkonstrukte wurden im Verhältnis 1:9 zusammen mit konstitutiv CMV-getriebenem TALE in Hek293T-RexTM-Zellen co-transfiziert. Die Luciferaseaktivität des jeweiligen Reporterkonstruktes mit einem Derivat des natürlichen TALEs Hax3 (H3) mit N-terminalem FLAG-Epitop und zur VP16-AD substituierten TALE-AD, H3-VP16, diente als Bezugsgröße für die Berechnung der relativen Luciferaseaktivität. Fehlerindikatoren kennzeichnen Standardabweichungen. (E) Immunoblot der verwendeten TALEs mit N-terminalem FLAG-Epitop. Die Zellysate der Dreifachbestimmung des jeweiligen TALEs mit Reporterkonstrukt HIV-1 Subtyp E wurden zu gleichen Teilen gemischt und wie unter Material und Methoden beschrieben aufgearbeitet, sodass jeweils ca. 8 µg Totalprotein aufgetrennt wurden. Eine schwarze Pfeilspitze kennzeichnet das Signal von ArtABC6.

Zum Test der für die ausgewählten HIV-1 Subtyp-Isolate neu erstellten TALEs, TALEABC5 und ArtABC6, wurde die Luciferasereporteraktivität der verschiedenen HIV-1 Subtyp-Reporterkonstrukte in Anwesenheit dieser TALEs in Hek293T-RExTM-Zellen ermittelt und mit der Luciferasereporteraktivität des jeweiligen Reporterkonstruktes in Anwesenheit von TALE5, bzw. ArtHIV6 verglichen (Abb. 3.2.7 D). Zum Vergleich wurde die Luciferasereporteraktivität der Reporterkonstrukte mit Hax3-VP16, das keine Bindestelle in den Reporterkonstrukten besitzt, verwendet. Diese Luciferasereporteranalyse ergab eine verbesserte, um die achtfache Induktion der neuen Reporterkonstrukte für die HIV-1 Subtypen A und C bis F durch die neuen TALEs im Vergleich zu TALE5 und ArtHIV6 (Abb. 3.2.7 D). Am HIV-1 Subtyp B-Reporterkonstrukt aus dem 8E5-Provirus hingegen konnte eine annähernd gleiche Aktivität beobachtet werden (Abb. 3.2.7 D). Es konnte also durch Veränderungen in der *repeat*-Region eine größere Bandbreite an induzierbaren Sequenzen bei gleichbleibender Effizienz erreicht werden. Im Immunoblot mit Lysaten aus Hek293T-RExTM-Zellen konnten darüber hinaus alle TALEs mit dem korrekten Molekulargewicht nachgewiesen werden (Abb. 3.2.7 E). Dabei scheint ArtABC6 im Vergleich zu den übrigen TALE-VP16-Fusionen schlechter synthetisiert zu werden oder weniger stabil zu sein.

3.3 Die Konstruktion von *single chain* TALEN (scTALEN)-Varianten und hochspezifischen TALEN-Paaren

Gegenüber anderen Anwendungen von TALEs, z. B. als Aktivatoren, ist die Genomeditierung am populärsten, da sie das einfache, aber hochpräzise Verändern von DNA-Sequenzen *in situ* erlaubt. Dies bietet enorme Möglichkeiten bei der Untersuchung von Genfunktionen in der Forschung und zur Adressierung genetisch bedingter Krankheiten. Dennoch gibt es Nachteile bei der Anwendung von TALEN, insbesondere im Vergleich zu dem jüngst für die Genomeditierung entdeckten CRISPR/Cas-System. Dies liegt vor allem in der paarigen Anwendung von TALEN begründet, der die vergleichsweise aufwändige und teure Erstellung zweier *repeat*-Regionen erfordert, während ein sgRNA-Molekül für einen DSB mittels Cas9 ausreicht. Außerdem erschwert die Notwendigkeit zwei große Proteine in die Zielzelle zu verbringen die Applikation von TALEN gegenüber dem ebenfalls großen Protein Cas9 und der allerdings sehr kleinen sgRNA. Darüber hinaus verlangt die paarige Anwendung die Zugänglichkeit eines relativ großen Bereichs, 50 bis 60 bp (zweimal ca. 20 bp TALEN-Bindestellen, 14 bis 16 bp *spacer*), im Chromatin des Zielgenoms. Cas9, beladen mit sgRNA, erkennt insgesamt nur 23 bp, was diesem Komplex vermutlich den Schnitt einer Ziel-DNA-Sequenz in mehr heterochromatischen Bereichen vereinfacht. Für den Genetiker stellt CRISPR/Cas9 daher das einfachere Werkzeug dar. Die paarige Anwendung von TALEN hat aber auch Vorteile. Durch die große Anzahl spezifischer Sequenzpositionen der beiden TALEN-Ziel-DNA-Sequenzen ist die Anzahl möglicher *off-target*-Sequenzen vor allem im Vergleich zu den Monomeren CRISPR/Cas9-Komplexen gering. Dennoch weisen TALEN eine gewisse Fehlpaarungstoleranz auf (Abb. 3.1.2 D; Tesson *et al.*, 2011; Osborn *et al.*, 2013). Und selbst TALE-Aktivatoren sind abhängig von der Anzahl an *repeats* in der Lage ein bis drei RBF zu überwinden (Boch *et al.*, 2009; Mali *et al.*, 2013 b). Dies stellt vor allem bei der Anwendung von TALEN für gentherapeutische Zwecke ein Problem dar. Zwar können *off-targets* durch entsprechendes Design reduziert werden, doch eine Differenzierung zwischen gesundem und krankem Allel anhand eines *single nucleotide polymorphism* (SNP) ist nicht möglich (Garg *et al.*, 2012). Dabei stellen SNPs den Großteil der genetischen Variation beim Menschen dar und sind somit auch ursächlich für viele genetisch bedingte Erkrankungen (Wang und Moulton, 2001; Lee und Shatky, 2008).

Im Folgenden wurden zwei Ansätze gewählt, um zum einen die effizientere Erstellung von TALEN und zum anderen die Diskriminierung eines Nukleotides zu realisieren.

3.3.1 Studien zu Design und Aktivität von scTALEN

Zu Beginn dieser Arbeit waren noch keine scTALEN, jedoch scZFN publiziert, deren scFokI-Variante den zu designenden scTALEN zugrunde gelegt wurde (Minczuk *et al.*, 2008; Mino *et al.*, 2009). Dazu wurden zwei FokI-Endonukleasedomänenmonomere durch ein Glycin (G)- und Serin (S)-haltiges Verbindungsstück, *linker*, miteinander verbunden. In Minczuk *et al.*, 2008, wurde ein 35 AS-*linker* in *in vitro*-Studien als ausreichend für eine volle Nukleaseaktivität von 95% identifiziert und entsprechende scZFNs erfolgreich zur Depletion mitochondrialer DNA in Zellkultur eingesetzt. Dabei enthielt die Peptidkette zwischen den Zinkfingern und der scFokI ca. 46 AS, bestehend aus einem HA-Epitop, einer NES und einem weiteren GS-*linker* (Minczuk *et al.*, 2008). Im Vergleich dazu wurde in Mino *et al.*, 2009, ein längerer GS-*linker* von 60 AS zwischen den Endonukleasedomänen in Restriktionsanalysen als optimal geeignet ermittelt. Ein ähnlicher, aber nicht zu Minczuk *et al.*, 2008, identischer 35 AS-*linker* zwischen den FokI-Monomeren erlaubte nur eine deutlich schwächere Aktivität, wobei die Verbindung der Zinkfinger mit der scFokI nur aus einem 15 AS GS-*linker* bestand (Mino *et al.*, 2009). Es ist anzunehmen, dass eine scFokI eine gewisse Flexibilität zum Schneiden einer DNA benötigt. Diese erscheint in den beiden Publikationen unterschiedlich realisiert worden zu sein, durch erhöhte Beweglichkeit zwischen entweder der DNA-Bindedomäne und der scFokI (Minczuk *et al.*, 2008) oder den beiden monomeren FokI-Endonukleasedomänen (Mino *et al.*, 2009). Da hier ein TALEN-Grundgerüst nach Miller *et al.*, 2011, mit 63 TALE-C-terminalen AS (Abb. 3.3.1) als relativ langem Verbindungsstück zwischen *repeats* und FokI-Endonukleasedomäne gewählt wurde, fiel die Entscheidung für den *linker* in der scFokI auf die 35 AS aus Minczuk *et al.*, 2008, der darüber hinaus im Vergleich zu Mino *et al.*, 2009, auch die erfolgreiche Applikation in Zellen gezeigt hatte (Maik Reschke, Dissertation 2018).

Für die Entwicklung der scTALEN wurden mit Rücksicht auf die für die *in vivo*-Studien angedachte Zelllinie, Hek293T-RexTM, die in Mussolino *et al.*, 2011, in solchen Zellen mit einem TALEN-Paar adressierten Ziel-DNA-Sequenzen aus dem humanen Gen *CCR5* verwendet (Abb. 3.3.1). Die Bindestellen für die linke (*Left*) L-Typ- und die rechte (*Right*) R-Typ-TALEN mit jeweils 17,5 *repeats* sind dabei durch einen 15 bp-*spacer* separiert, der eine Hpy188III-Restriktionsschnittstelle enthält, die bei der Analyse der Nukleaseaktivität in Zellen zur Anreicherung mutierter Sequenzen genutzt werden kann. Die mit der RVD-Abfolge aus Mussolino *et al.*, 2011, erstellten *repeat*-Module, sowie die neu kreierten N- und C-terminalen TALEN-Module basierten auf einer für die Golden TALE Technologie gestalteten, für *Homo sapiens* Codon-optimierten, Hax3-sequenzbasierten Modulbibliothek (Geißler *et al.*, 2015).

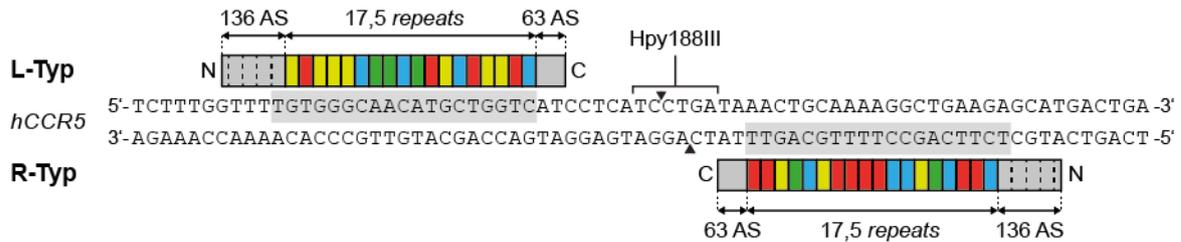


Abb. 3.3.1: Ziel-DNA-Sequenzen aus dem humanen *CCR5* nach Mussolino *et al.*, 2012, mit den verwendeten TALEN-Grundgerüsten. L (*Left*)- und R (*Right*)-Typ-TALEN enthielten jeweils einen 136 AS N- und einen 63 AS C-Terminus wie in Miller *et al.*, 2011, beschrieben. Die RVD-Abfolge zur Erkennung der Ziel-DNA-Sequenzen wurde von Mussolino *et al.*, 2011, übernommen mit NI (grün) für Adenin, HD (hellblau) für Cytosin, NN (gelb) für Guanin und NG (rot) für Thymin.

3.3.1.1 Die Auswahl einer geeigneten FokI-Variante

Im Rahmen der Entwicklung von FokI-basierten ZFNs gab es Bemühungen zur Verbesserung der Endonukleasedomäne für eine erhöhte Spezifität durch Heterodimerisierung der Paare und verbesserte Effizienz (Miller *et al.*, 2007; Szczepek *et al.*, 2007; Guo *et al.*, 2010). Um eine effiziente scFokI-Domäne zu erstellen wurde daher neben einer Wildtyp (wt)-FokI auch eine als *Sharkey* bezeichnete Variante aus Guo *et al.*, 2010, getestet. Dabei handelt es sich um zwei, in einem gerichteten Evolutionsansatz identifizierte Mutationen, die eine 15-fach höhere Nukleaseaktivität vermittelten (Abb. 3.3.2 A; Guo *et al.*, 2010). Darüber hinaus war die Kompatibilität solcher Mutationen mit obligat heterodimerisierenden FokI-Varianten bekannt, wobei die Paarung einer sogenannten DS- mit einer RR-Variante als am besten geeignet erschien (Abb. 3.3.2 A; Guo *et al.*, 2010; Maik Reschke, Dissertation 2018). Daher wurden *Sharkey* DS- und RR-Endonukleasedomänen für die L- und R-Typ-TALEN eines Kontrollpaares, sowie aus praktischen Gründen auch für eine scFokI, verwendet, sodass jeweils ein TALEN-Paar und eine scTALEN einmal mit Wildtyp- und einmal mit *Sharkey*-FokI-Endonukleasedomäne erstellt wurden (Abb. 3.3.2 A, B). Zum Test der Nukleaseaktivität wurden *in vitro*-Restriktionsanalysen mit linearisierten Plasmid-DNAs mit den Ziel-DNA-Sequenzen aus *hCCR5* durchgeführt (Abb. 3.3.2 C). Obwohl alle TALEN-Varianten an der Ziel-DNA aktiv waren, konnte sowohl für die Paare als auch für die scTALEN eine höhere Aktivität mit *Sharkey* FokI-Endonukleasedomänen festgestellt werden, wobei dies bei den scTALEN noch deutlicher auffiel. An der Nicht-Ziel-DNA (*gagpol*) konnte hingegen keine Restriktion beobachtet werden (Abb. 3.3.2 C). Die vollständige Transkription und Translation der verschiedenen TALEN *in vitro* wurde mittels Immunoblot nachgewiesen (Abb. 3.3.2 D).

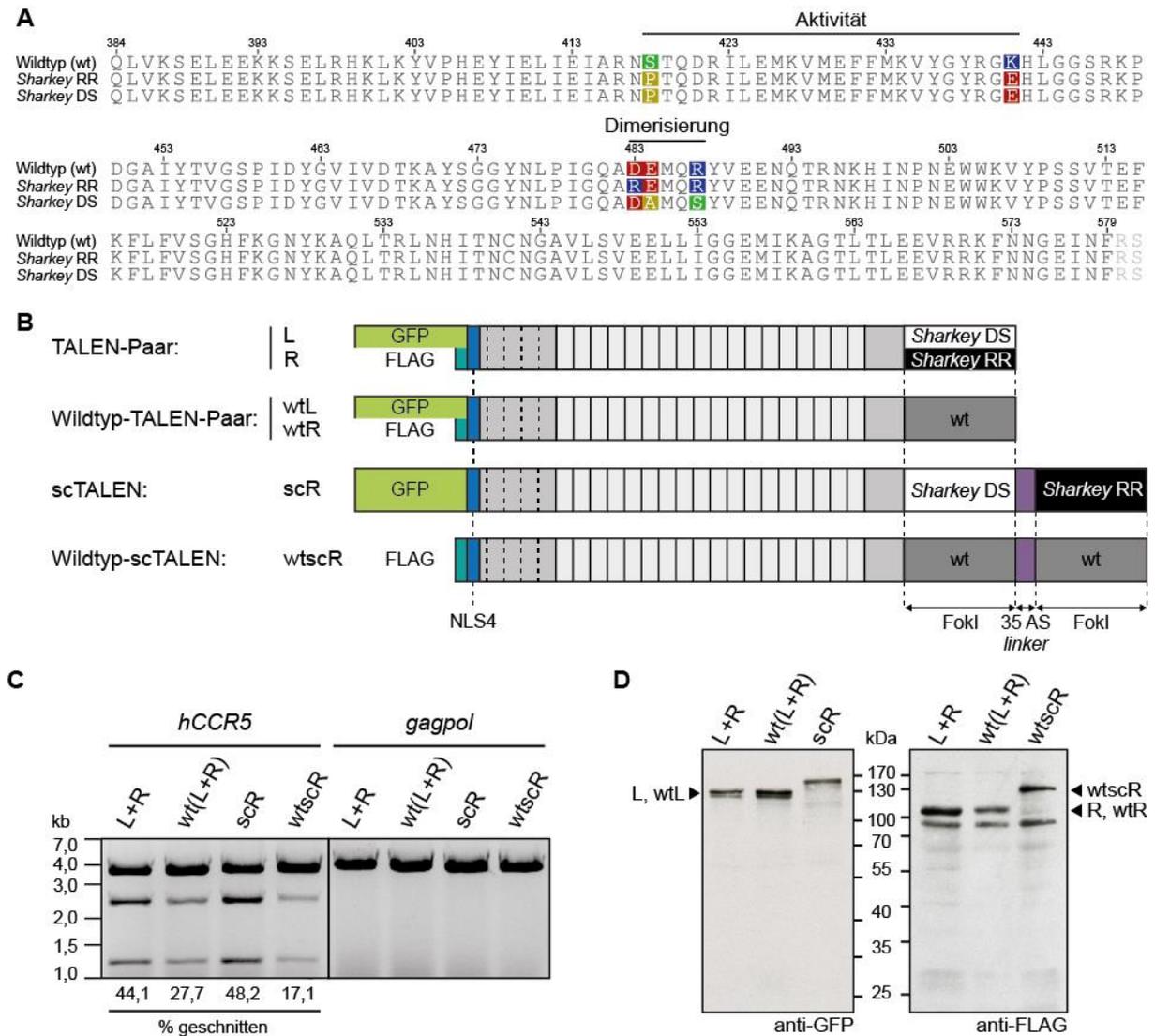


Abb. 3.3.2: Vergleich von Sharkey und Wildtyp (wt)-scTALEN und TALEN-Paaren. (A) AS-Sequenzalignment mit Geneious Pro 5.3.6 der verwendeten FokI-Endonukleasedomänen. Die zugunsten der Aktivität der Sharkey-Varianten oder der Heterodimerisierung der DS- und RR-Varianten vom Wildtyp (wt) abweichenden AS sind abhängig von ihrer Polarität hervorgehoben (ocker – unpolar, grün – polar, blau – positiv geladen, rot – negativ geladen). Die Nummerierung entspricht der Position im vollständigen FokI-Protein. Hellgraue AS entstammen nicht FokI. (B) Schematische Darstellung vom Aufbau der getesteten TALEN. Das TALE-Grundgerüst und die repeats der L- und R-Typ-TALEN wurden bereits detailliert in Abb. 3.3.1 beschrieben. Die zwischen das N-terminale Epitop, GFP (hellgrün) oder FLAG (cyan), und den TALE-Anteil fusionierte NLS4 (blau) wurde in Kapitel 3.2.6 ermittelt (Abb. 3.2.9). Die verschiedenen FokI-Endonukleasedomänen und der 35 AS-linker (violett) aus Minczuk *et al.*, 2008, sind gekennzeichnet. (C) Restriktionsanalyse mit den Sharkey- und Wildtyp-TALEN- und scTALEN-Varianten wie in (B) beschrieben. *In vitro* transkribierte und translatierte TALEN wurden für eine Stunde bei 37°C mit einem linearisierten Plasmid, das entweder ein die Ziel-DNA-Sequenzen enthaltendes Fragment aus *hCCR5* oder ein Fragment aus dem HIV-Gen *gagpol* ohne TALEN-Bindestellen enthielt, inkubiert. Die inaktivierten Reaktionen wurden über ein 0,8%-iges Agarosegel aufgetrennt und der Anteil TALEN-geschnittenen Plasmids mit ImageJ ermittelt. (D) Immunoblots der verwendeten TALEN-Varianten. Die *in vitro*-Transkriptions- und Translationsreaktionen der TALEN-Paare oder scTALEN wurden so aufgearbeitet, dass jeweils die in der Restriktionsanalyse verwendete Menge aufgetrennt wurde.

Der Vergleich hat deutlich gezeigt, dass die verwendete *Sharkey*-FokI-Endonukleasedomäne auch für scTALEN eine höhere Aktivität vermittelt (Abb. 3.3.2 C). Daher wurde für alle weiteren Experimenten diese scFokI-Variante und ein entsprechendes TALEN-Paar als Kontrolle gewählt. In der nach Beginn dieser Arbeit veröffentlichten, bisher einzigen Publikation mit FokI-basierten scTALEN wurde die Wildtyp-FokI- Endonukleasedomäne verwendet (Sun und Zhao, 2014). Die so erstellten scTALEN hatten eine nur geringe *in vivo*-Aktivität im Vergleich zu TALEN-Paaren für die gleichen Ziel-DNA-Sequenzen. Es gab jedoch weitere Unterschiede im Aufbau der in Sun und Zhao, 2014, genutzten scTALEN. Obwohl der verwendete TALE-C-Terminus ebenfalls 63 AS enthielt, wurde ein 95 AS-*linker* mit vermehrt unpolaren und nur wenigen polaren AS, der in einem *linker*-Bibliotheksscreen identifiziert wurde, zwischen den einzelnen FokI- Endonukleasedomänen eingesetzt. Ein vergleichbarer 100 AS GS-*linker* nach Mino *et al.*, 2009, hatte jedoch keine Nukleaseaktivität *in vitro* im Vergleich zum entsprechenden TALEN-Paar (Sun und Zhao, 2014). Darüber hinaus bestanden ein Drittel der zwölf in dem Bibliotheksscreen identifizierten *linker* aus 35 und ein Viertel aus 33 AS. Somit stellten kürzere, zu dem hier gewählten vergleichbare *linker* die Mehrheit der identifizierten Sequenzen dar. Nukleaseaktivität konnte für entsprechende scTALEN, sowie für scTALEN mit vier anderen *linker*-Regionen mit 57 bis 104 AS, jedoch nicht nachgewiesen werden (Sun und Zhao, 2014). Die Auswahl eines geeigneten *linker* für eine scFokI zur Erstellung einer scTALEN ist demnach schwierig. Der hier gewählte 35 AS GS-*linker* erlaubte zu entsprechenden TALEN-Paaren vergleichbare Nukleaseaktivitäten *in vitro* und wurde deshalb beibehalten.

3.3.1.2 Restriktionsstudien mit verschiedenen scTALEN-Varianten

Um die möglichen Vorteile von scTALEN weiter auszubauen wurden weitere Varianten erstellt, die andere Nachteile vorhandener, programmierbaren Nukleasen adressieren sollten. Zum einen wurde eine deaktivierende Mutation in jeweils eine der beiden FokI-Endonukleasedomänen eines Paares oder der scTALEN zur Kreation von DNA-Einzelstrangbrüchen verursachenden Nickasen eingeführt. Studien zunächst mit ZFN und in jüngerer Zeit auch mit TALEN haben für FokI-basierte Nickasen eine niedrigere Zytotoxizität bei relativ zum Anteil an NHEJ höherer HDR gezeigt, weshalb eine Übertragung dieser Vorteile auf eine einzelne Nuklease von Interesse war (Wang *et al.*, 2012; Ramirez *et al.*, 2012; Kim *et al.*, 2012; Wu *et al.*, 2014). Zum anderen wurden TALEN mit N- und C-terminaler scFokI-Endonukleasedomäne erstellt. Diese sollten die Anzahl benötigter Nukleasemoleküle zur gezielten Vermittlung kleiner Deletionen von entweder vier TALEN zweier Paare oder zwei

scTALEN auf nur eine Doppel-scTALEN (*double* scTALEN, dscTALEN) vermindern, was den Konstruktionsaufwand verringern und die Applikationseffizienz erhöhen würde. Darüber hinaus könnte so auch die scTALEN-Ziel-DNA-Sequenz direkt ausgeschnitten werden, wodurch ein mögliches wiederholtes Schneiden verhindert würde. Eine höhere Effizienz und geringere Zytotoxizität könnten somit durch diesen Ansatz erreicht werden.

3.3.1.2.1 TALEN und scTALEN mit der FokI-inaktivierenden D450A-Mutation erzeugen DSBs

Zur Erstellung potentieller Nickasen wurde die Mutation der Asparaginsäure an Position 450 bezogen auf das native Restriktionsenzym FokI zu Alanin (D450A) zur Deaktivierung der Endonukleasedomäne gewählt (vgl. Abb. 3.3.2 A; Waugh und Sauer, 1993). Diese wurde auch bereits in anderen Studien zum Design von Nickasen verwendet (Ramirez *et al.*, 2012; Kim *et al.*, 2012; Wu *et al.*, 2014). Im Vergleich mit zwei anderen, für diesen Zweck benutzten Mutationen, D450N und D467A, konnte für D450A in Wu *et al.*, 2014, die niedrigste Zytotoxizität und hohe HDR-Effizienz festgestellt werden. Um zu testen, ob D450A auch zur Deaktivierung der hier erstellten *Sharkey*-TALEN und scTALEN geeignet ist, wurden L- und R-Typ-TALEN mit dieser Mutation erstellt, die in verschiedenen Kombinationen mit dem vorhandenen TALEN-Paar in *in vitro*-Restriktionsanalysen eingesetzt werden konnten (Abb. 3.3.3 A, B). Außerdem wurden zusätzlich zu der vorhandenen scTALEN zwei weitere konstruiert, die die Mutation entweder in dem mehr N-terminal gelegenen DS (scRds)- oder in dem mehr C-terminal gelegenen RR (scRrr)-Teil der scFokI trugen (Abb. 3.3.3 A). Auch deren Aktivität wurde mittels *in vitro*-Restriktion untersucht (Abb. 3.3.3 B). Dabei konnte nur bei der Paarung einer L- mit einer R-Typ-TALEN, die jeweils D450A enthielten (Lds+Rrr), keine Restriktion der hCCR5-Ziel-DNA festgestellt werden. Die übrigen, als Nickasen gedachten TALEN-Paare (L+Rrr, Lds+R) und scTALEN (scRds, scRrr) vermittelten Doppelstrangbrüche mit unterschiedlicher Effizienz. Bei D450A-Mutation der *Sharkey*-FokI-RR-Endonukleasedomäne wurde eine Reduktion der Nukleaseaktivität des TALEN-Paares (L+Rrr) auf ca. 50% und der scTALEN (scRrr) auf ca. 20% beobachtet (Abb. 3.3.3 B). Die Mutation der *Sharkey*-FokI-DS-Endonukleasedomäne führte hingegen in beiden Fällen (Lds+r, scRds) zu keiner Verringerung der Restriktionseffizienz an der Ziel-DNA (Abb. 3.3.3 B). Eine Nicht-Ziel-DNA (*gagpol*) wurde jedoch von keiner TALEN-Kombination oder scTALEN geschnitten (Abb. 3.3.3 B), wobei für alle die vollständige *in vitro*-Transkription und Translation im Immunoblot nachgewiesen werden konnte (Abb. 3.3.3 C).

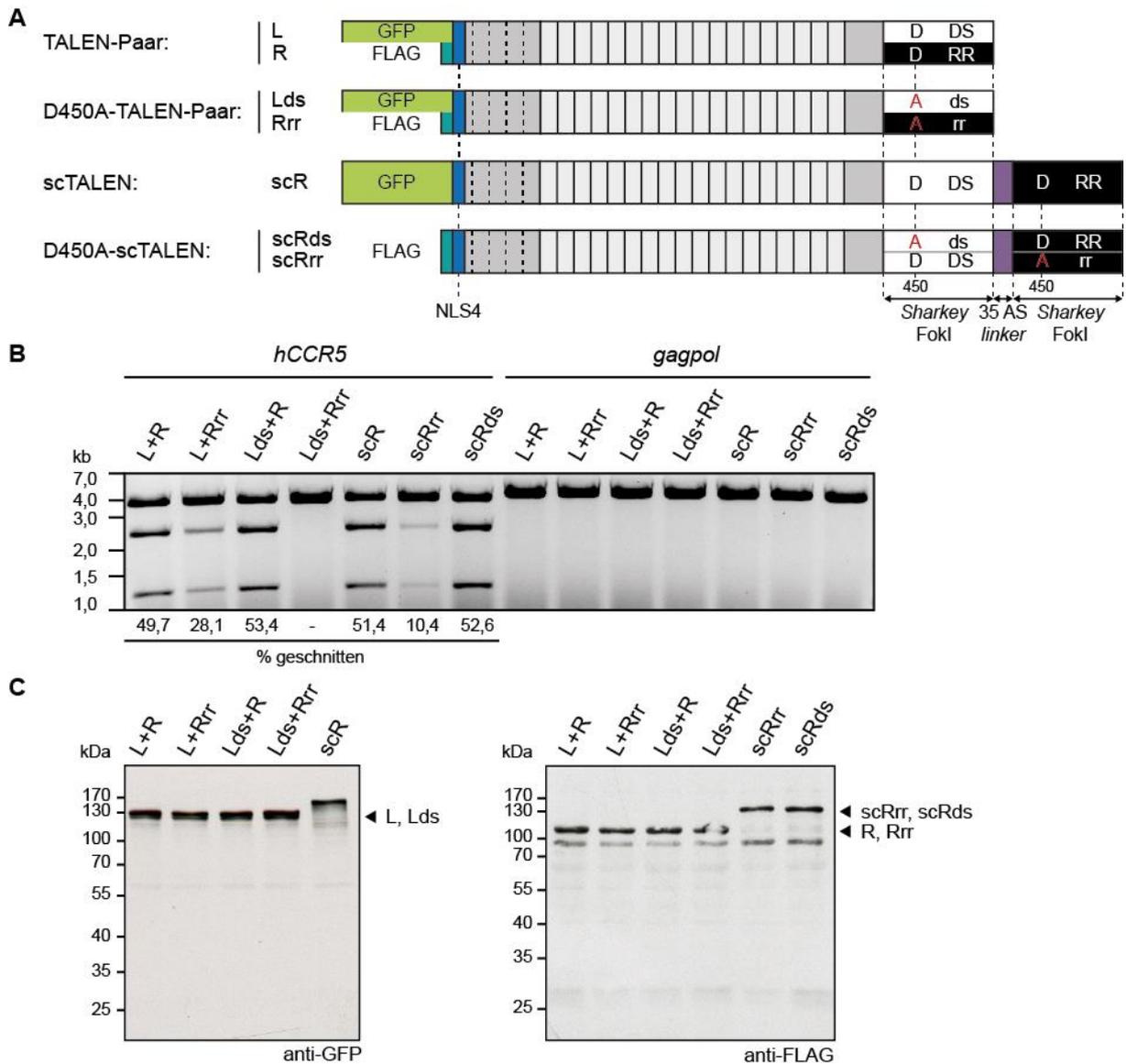


Abb. 3.3.3: TALEN mit einer D450A-Mutation in einer Sharkey-FokI-Endonukleasedomäne bleiben aktiv.

(A) Schematische Darstellung vom Aufbau der getesteten TALEN. Das TALE-Grundgerüst und die *repeats* der L- und R-Typ-TALEN wurden bereits detailliert in Abb. 3.3.1 beschrieben. Die zwischen das N-terminale Epitop, GFP (hellgrün) oder FLAG (cyan), und den TALE-Anteil fusionierte NLS4 (blau) wurde in Kapitel 3.2.6 ermittelt (Abb. 3.2.9). Die verschiedenen Sharkey-FokI-Endonukleasedomänen und der 35 AS-linker (violett) aus Minczuk *et al.*, 2008, sind gekennzeichnet. Die heterodimerisierenden FokI-Varianten sind anhängig von der AS an Position 450 der vollständigen FokI-AS-Sequenz als DS und RR im Falle der ursprünglichen Asparaginsäure oder ds und rr bei Mutation zu Alanin bezeichnet. (B) Restriktionsanalyse mit den Sharkey- und D450A-TALEN- und scTALEN-Varianten wie in (A) beschrieben. *In vitro* transkribierte und translatierte TALEN wurden für eine Stunde bei 37°C mit einem linearisierten Plasmid, das entweder ein die Ziel-DNA-Sequenzen enthaltendes Fragment aus *hCCR5* oder ein Fragment aus dem HIV-Gen *gagpol* ohne TALEN-Bindestellen enthielt, inkubiert. Die inaktivierten Reaktionen wurden über ein 0,8%-iges Agarosegel aufgetrennt und der Anteil TALEN-geschnittenen Plasmids mit ImageJ ermittelt. (C) Immunoblots der verwendeten TALEN-Varianten. Die *in vitro*-Transkriptions- und Translationsreaktionen der TALEN-Paare oder scTALEN wurden so aufgearbeitet, dass jeweils die in der Restriktionsanalyse verwendete Menge aufgetrennt wurde.

Anhand dieser Ergebnisse kann festgehalten werden, dass die erstellten „Nickasen“ mit D450A-Mutation *in vitro* beide DNA-Stränge schneiden (Abb. 3.3.3 B). Eine mögliche Ursache wäre der Schnitt beider DNA-Einzelstränge durch die aktive FokI-Endonukleasedomäne eines Nickase-Paares oder einer scNickase. Voraussetzung dafür wäre eine hohe Beweglichkeit der Endonukleasedomänen am gewählten TALE-C-Terminus. Die unterschiedliche Schnitteffizienz der putativen Nickasen mit D450A im DS- und RR-Monomer (Abb. 3.3.3 B) legt weiterhin nahe, dass die DS- und RR-Varianten der FokI-Endonukleasedomäne eine unterschiedliche Schnitteffizienz besitzen.

3.3.1.2.2 Beide scFokI-Domänen einer dscTALEN können DSBs induzieren

Zur Konstruktion von dscTALEN wurden Varianten N-terminaler scTALEN erstellt (Abb. 3.3.4 A). Dazu wurde neben dem bereits vorgestellten TALE-Grundgerüst mit N-terminalen 136 und C-terminalen 63 AS (NscR1) eine zweite Variante mit vollständigem TALE-N-Terminus (NscR2) gewählt, da die Funktionalität einer FokI-Endonukleasedomäne in diesem Kontext bereits gezeigt wurde (Li *et al.*, 2011). In diesem Falle wurde NLS4 zwischen das N-terminale Epitop und die scFokI verschoben. Entsprechend wurden zwei dscTALEN-Varianten (dscR1, dscR2) erstellt, die neben der jeweiligen N-terminalen scFokI eine C-terminale scFokI analog zur C-terminalen scTALEN trugen (Abb. 3.3.4 A). Dabei wurde die vollständige *in vitro*-Synthese der verschiedenen Konstrukte im Immunoblot bestätigt (Abb. 3.3.4 C). Zum Test der Nukleaseaktivitäten dieser TALEN-Varianten wurden *in vitro*-Restriktionsstudien durchgeführt, in denen die Funktionalität aller erstellten Nukleasen bei Vorhandensein der spezifischen Ziel-DNA-Sequenz nachgewiesen werden konnte (Abb. 3.3.4 B). Darüber hinaus wurde, um die Größe der mit einer dscTALEN deletierbaren Bereiche abschätzen zu können, die durch NscR1 induzierte Schnittstelle mittels Sequenzierung bestimmt (Abb. 3.3.4 D). Anhand einer Sequenz konnte unter Berücksichtigung des von FokI erzeugten 4 nt langen 5'-Überhangs ein 16 nt von der scTALEN-Bindestelle entfernter Schnittpunkt im NscR1-gebundenen und -ungebundenen DNA-Strang angenommen werden. Im Vergleich dazu war von TALEN mit FokI-Endonukleasedomäne am vollständigen TALE-N-Terminus ein mit 14 nt im ungebundenen Strang erwartungsgemäß etwas weiter von der Bindestelle entfernter Schnittpunkt beobachtet worden (Li *et al.*, 2011). Da die optimale *spacer*-Länge für TALEN-Paare mit FokI-Endonukleasedomäne an 63 TALE-C-terminalen AS auf 14 bis 16 bp eingegrenzt wurde (Miller *et al.*, 2011), wurde für die hier verwendete C-terminale scFokI ein 5 oder 6 nt entfernter Schnittpunkt im gebundenen, bzw. ein 9 oder 10 nt entfernter Schnittpunkt im ungebundenen DNA-Strang geschätzt. Damit waren für dcsTALEN mit 19 bp Ziel-DNA-

Sequenz Deletionen von mindestens 40 bp anzunehmen. In Restriktionsanalysen war dieser Größenunterschied zwischen N- und C-terminaler scTALEN für dieselbe Ziel-DNA-Sequenz am kleineren Restriktionsfragment zu erkennen (Abb. 3.3.4 E). Für die dscTALEN-Varianten waren Fragmente beider Größenordnungen nachzuweisen, was bedeutet, dass beide scFokI-Endonukleasedomänen der Moleküle allerdings in Abhängigkeit vom gewählten TALE-N-Terminus mit unterschiedlicher Effizienz schneiden können (Abb. 3.3.4 B, E). Das gemeinsame Auftreten beider Fragmente zeigte jedoch auch, dass die N- und die C-terminale scFokI zeitlich unabhängig voneinander schneiden (Abb. 3.3.4 E). Würden beide Endonukleasedomänen immer zeitgleich schneiden, wäre jeweils nur das kleinere Fragment nachzuweisen gewesen. Die vorliegenden Ergebnisse schließen jedoch nicht aus, dass ein synchroner Schnitt von N- und C-terminaler scFokI möglich ist. Die beobachtete, mit den getesteten scTALEN vergleichbare Nukleaseaktivität der dscTALEN könnte auf synchronen DNA-Schnitt hinweisen (Abb. 3.3.4 B). Aufschluss darüber, ob Deletionen mit dscTALEN möglich sind, könnte der Test dieser neben den scTALEN in *in vivo*-Systemen geben.

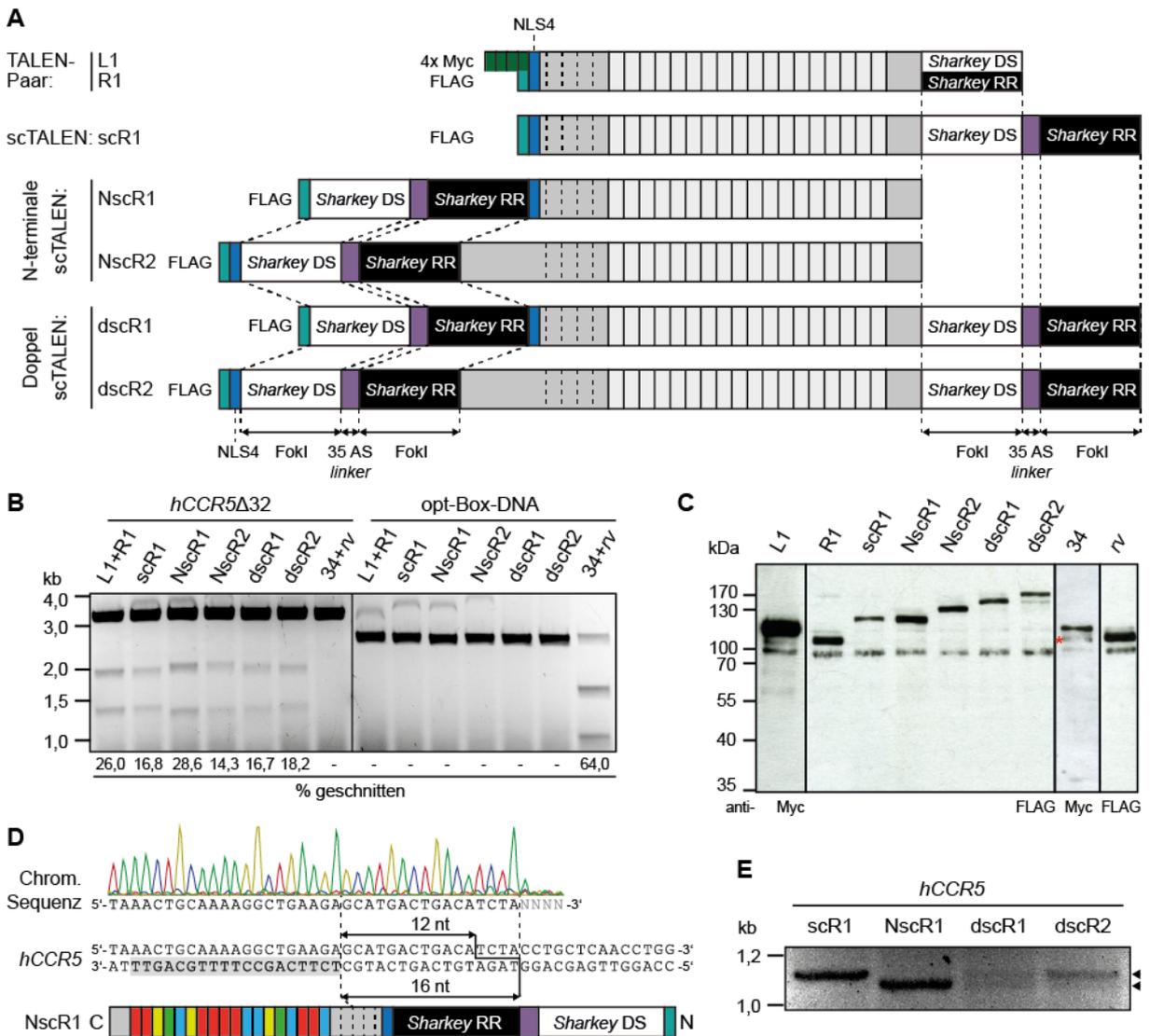


Abb. 3.3.4: Beide scFokI-Domänen einer Doppel-scTALEN (dscTALEN) können DNA-DSBs erzeugen.

(A) Schematische Darstellung vom Aufbau der getesteten TALEN. Das TALE-Grundgerüst der ersten TALEN-Variante (1) und die *repeats* der L- und R-Typ-TALEN wurden bereits detailliert in Abb. 3.3.1 beschrieben. Die TALEN der zweiten Variante (2) enthalten einen vollständigen TALE-N-Terminus. Im Unterschied zu den R-Typ-TALEN wurde der L-Typ-TALEN L1 mit einem 4x Myc-Epitop (dunkelgrün) versehen. Die in Kapitel 3.2.6 ermittelte NLS4 (blau, Abb. 3.2.9) wurde entweder unmittelbar N-terminal des verkürzten TALE-N-Terminus platziert oder zwischen N-terminales FLAG-Epitop (cyan) und scFokI fusioniert, wenn ein vollständiger TALE-N-Terminus verwendet wurde. Die Sharkey-FokI-Endonukleasedomänen und der 35 AS-linker (violett) aus Minczuk *et al.*, 2008, sind gekennzeichnet. (B) Restriktionsanalyse mit dem TALEN-Paar L1 und R1, den N- und C-terminalen scTALEN, sowie den dscTALEN-Varianten wie in (A) beschrieben. *In vitro* transkribierte und translatierte TALEN wurden für eine Stunde bei 37°C mit linearisierten Plasmiden inkubiert, die entweder ein die Ziel-DNA-Sequenzen enthaltendes Fragment aus *hCCR5Δ32* oder die opt-Box aus Kapitel 3.1.1 (Abb. 3.1.1 A, D, E) mit TALEN34- und *reverse (rv)* TALEN-Bindestellen enthielten. Die inaktivierten Reaktionen wurden über ein 0,8%-iges Agarosegel aufgetrennt und der Anteil TALEN-geschnittenen Plasmids mit ImageJ ermittelt. (C) Immunoblots der verwendeten TALEN-Varianten. Die *in vitro*-Transkriptions- und Translationsreaktionen der verschiedenen TALEN-Varianten wurden so aufgearbeitet, dass jeweils die in der Restriktionsanalyse verwendete

Menge aufgetrennt wurde. Das Signal der 1x Myc-markierten TALEN34 ist durch einen roten Asterisk gekennzeichnet. (D) Schnittstellenbestimmung von NscR1 mittels Sanger-Sequenzierung eines Restriktionsproduktes. Die erhaltene Sequenz und das Chromatogramm (Chrom.), sowie die NscR1-*repeats* sind im *alignment* an hCCR5 gezeigt. (E) Detailausschnitt einer wie unter (B) beschriebenen Restriktionsanalyse an einem Plasmid mit *hCCR5*-Fragment. Fragmente unterschiedlicher Größe sind durch schwarze Pfeile markiert.

3.3.1.3 Test von scTALEN- und Doppel-scTALEN-Varianten in zwei Modellsystemen

Bisher konnte die Nukleaseaktivität von scTALEN und dscTALEN in *in vitro*-Restriktionsanalysen nachgewiesen werden (siehe 3.3.1.1, 3.3.1.2.2). Um die Anwendbarkeit zur Genom-Editierung beurteilen zu können, soll im Folgenden überprüft werden, ob scTALEN Mutationen und dscTALEN Deletionen *in vivo* erzeugen können. Dazu wurden zum einen TALEN, scTALEN und dscTALEN für Gene des Zebrafisch *Danio rerio* erstellt und teilweise im Rahmen einer Kollaboration in diesem Modellorganismus getestet (3.3.1.3.1). Dabei ist *Danio rerio* durch seine gut observierbare Embryonalentwicklung sowie die hohe Regenerationsfähigkeit für die Entwicklungsbiologie, aber auch die Medizinische Forschung bedeutsam und auch die erste Vertebraten-Spezies in der Genom-Editierung mittels TALEN gezeigt wurde (Huang *et al.*, 2011). Zum anderen wurde die Aktivität von TALEN, scTALEN und dscTALEN nach Mussolini *et al.*, 2012, nun auch in Hek293T-RexTM-Zellen untersucht (3.3.1.3.2). Für beide Ansätze wurden die verschiedenen TALEN-Varianten C-terminal mit einer zusätzlichen, zweiteiligen (bipartiten) NLS versehen, mit deren Hilfe bereits in anderen Studien die Kernlokalisierung von SpCas9 vermittelt werden konnte (Le Cong *et al.*, 2013).

3.3.1.3.1 scTALEN- und Doppel-scTALEN für zwei Gene des Zebrafisch *Danio rerio*

Zum Test von scTALEN und dscTALEN in *Danio rerio* wurden Loci in bereits durch andere Nukleasen erfolgreich editierten Genen gewählt (Abb. 3.3.5 A). Dabei handelte es sich um das α -Tubulin-Gen *tuba1a* und das für das TAR-DNA-bindende Protein codierende Gen *tardbp* (http://www.ensembl.org/Danio_rerio/Gene/Summary). Der gewählte Locus des letzteren konnte bereits sowohl mit ZFN (Schmid *et al.*, 2013), als auch mit CRISPR/Cas9 (Hruscha *et al.*, 2013) modifiziert werden. Zur Detektion der herbeigeführten Veränderung in *Danio rerio* wurde für beide Gen-Loci jeweils eine zur Verfügung stehende Restriktionsschnittstelle verwendet (Abb. 3.3.5 A). Diese sollte im Falle einer stattfindenden Modifikation zerstört werden, was in einer PCR-basierten Analyse als Restriktionslängenpolymorphismus nachweisbar ist. Die hier erstellten TALENs wurden so entworfen, dass diese Analyse im Zebrafisch (*zebrafish*, Zef) möglich sein sollte (Abb. 3.3.5 A). Dazu wurde für jedes Gen ein heterodimerisierendes TALEN-Paar gewählt, das flankierend zur Restriktionsschnittstelle

bindet (Abb. 3.3.5 A, B). Dieses TALEN-Paar wurde mit Zef11 und Zef21 bezeichnet. Für die Ziel-DNA-Sequenz einer TALEN eines jeden Paares wurde jeweils eine scTALEN, ZefX2, und eine dscTALEN, ZefX3, erstellt. Zusätzlich wurde je Gen eine weitere dscTALEN für eine dritte Zielsequenz, Zef3 in *tubala* und Zef6 in *tardbp*, konstruiert. Um die generelle Funktionalität dieser TALEN-Varianten zu überprüfen, wurden Restriktionsanalysen *in vitro* durchgeführt (Abb. 3.3.5 C). Als zu schneidende DNA wurden linearisierte Plasmide mit Fragmenten der beiden Zielgene, *tubala* und *tardbp*, verwendet. Dabei konnte die Aktivität aller getesteten Nukleasen mit zusätzlicher NLS an einer jeweils geeigneten Ziel-DNA-Sequenz bestätigt werden. Auch die *in vitro*-Synthese der verschiedenen TALEN mit dem entsprechenden Molekulargewicht wurde im Immunoblot nachgewiesen (Abb. 3.3.5 D). Um zu untersuchen, ob die erstellten Nukleasevarianten zur Genommodifikation im Modellorganismus in der Lage sind, wurde das TALEN-Set für *tubala* im Rahmen einer Kooperation von Alexander Hruscha vom Deutschen Zentrum für Neurodegenerative Erkrankungen (DZNE), Schillerstraße 44 in 80336 München, getestet. Dazu wurden die verschiedenen TALEN *in vitro* transkribiert und die gewonnene mRNA in frisch befruchtete Eier des Zebraäbrblings injiziert. Für die co-injizierten TALEN eines Paares wurde jeweils eine mRNA-Konzentration von ca. 0,6 µg/µl eingesetzt. Für die einzeln injizierten scTALEN und dscTALEN wurde hingegen die doppelte mRNA-Konzentration von ca. 1,2 µg/µl verwendet. Zur Detektion der Modifikation von *tubala* wurden Restriktionsresistenzanalysen mit dem Enzym EaeI an PCR-Produkten des Lokus von genomischer DNA von jeweils zehn injizierten und zum Vergleich zwei nicht-injizierten Embryonen einen Tag nach der Befruchtung durchgeführt. Dabei konnte für das TALEN-Paar Zef11/21 eine Modifikationseffizienz von ca. 50% im Vergleich zu den Kontrollembryonen nachgewiesen werden (Alexander Hruscha, persönliche Kommunikation). Für die scTALEN Zef12 und Zef22, sowie die dscTALEN Zef23 und Zef3 konnte jedoch keine Modifikation von *tubala* mit der verwendeten Methode nachgewiesen werden (Alexander Hruscha, persönliche Kommunikation).

Die *in vitro* gezeigten Aktivitäten lassen darauf schließen, dass die zusätzliche, C-terminale NLS die Funktion der FokI-Endonukleasedomänen nicht stört. Darüber hinaus konnte die modifizierende Aktivität eines TALEN-Paares in einem Organismus gezeigt werden. Wider Erwarten haben die *in vitro* sehr aktiven scTALEN-Varianten *in vivo* nicht zu einer Mutation geführt. Ob dies auch für den zweiten Lokus der Fall wäre, müsste durch ein entsprechendes Experiment untersucht werden. Zu beachten ist, dass keine Aussage über die Integrität der scTALEN- und dscTALEN-Proteine in den Zielzellen getroffen werden kann, da kein Western-Blot durchgeführt wurde.

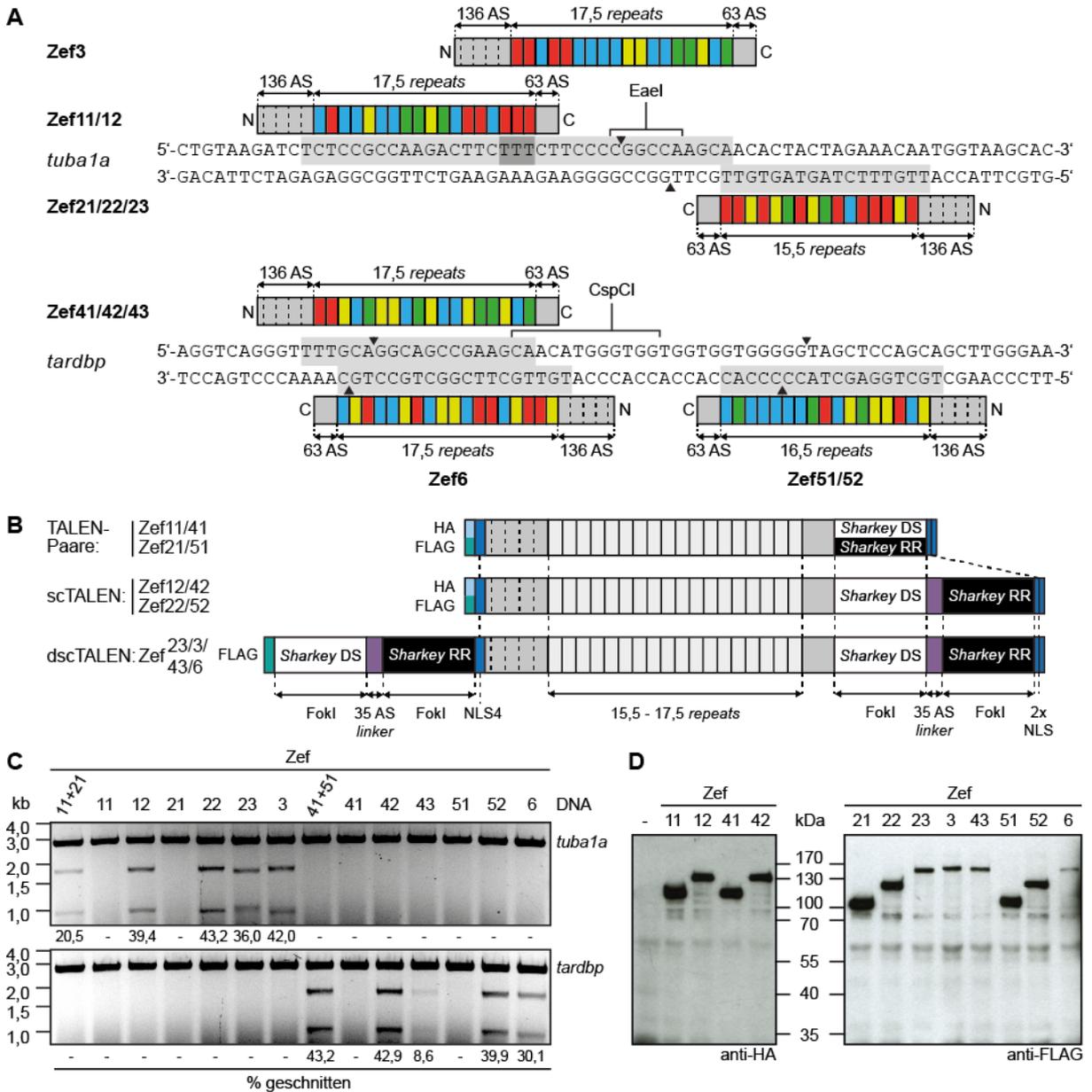


Abb. 3.3.5: Die für Gene des Zebrafisches *Danio rerio* erstellten TALEN-Varianten erzeugen DSB in *in vitro*-Restriktionsanalysen. (A) TALEN-Zielsequenzen und -Grundgerüste an den *D. rerio*-Genen *tuba1a* und *tardbp*. Die gewählten Loci in *tuba1a* und *tardbp* wurden bereits erfolgreich mit CRISPR/Cas9 (Alexander Hruscha, nicht publiziert) oder mit ZFN (Schmid *et al.*, 2013) und CRISPR/Cas9 (Hruscha *et al.*, 2013) editiert. Zur Detektion der Mutationen in *D. rerio* diente die Zerstörung von Endonukleaseschnittstellen (EaeI, CspCI) in PCR-basierten Restriktionsresistenzanalysen. Das TALE-Grundgerüst enthielt jeweils einen 136 AS N- und einen 63 AS C-Terminus wie in Miller *et al.*, 2010, beschrieben. Als RVDs zur Erkennung der Ziel-DNA-Sequenzen wurden NI (grün) für Adenin, HD (hellblau) für Cytosin, NN (gelb) für Guanin und NG (rot) für Thymin verwendet. Die Benennung erfolgte anhand des Zielorganismus (*zebrafish*, Zef) mit Indexziffern für Zielsequenz (X, 1-6) und TALEN-Variante (Y, 1-3) in der Form ZefXY, wobei Y entfällt, wenn nur eine TALEN-Variante für eine Zielsequenz erstellt wurde. (B) Schematische Darstellung vom Aufbau der erstellten Zef-TALEN. ZefX1-TALEN enthielten entweder eine Sharkey-FokI-Endonukleasedomäne vom heterodimerisierenden Typ DS oder RR, sodass Zef11 und 21, sowie Zef41 und 51 als TALEN-Paare für *tuba1a*, bzw. *tardbp* eingesetzt werden konnten. Für dieselben Ziel-DNA-Sequenzen wurden scTALEN, ZefX2, und dscTALEN, ZefX3, neben den zusätzlichen

dscTALEN Zef3 und 6 erstellt. Um ggf. einen unabhängigen immunologischen Nachweis aller potentiell paarig einsetzbarer TALEN zu ermöglichen wurde an Zef11, 12, 41 und 42 ein HA-Epitop (hellblau) und an Zef21, 22, 51 und 52 ein FLAG-Epitop (cyan) N-terminal fusioniert. Darüber hinaus wurden alle dscTALEN mit einem FLAG-Epitop versehen. Die in Kapitel 3.2.6 ermittelte NLS4 (blau, Abb. 3.2.9) wurde unmittelbar N-terminal des verkürzten TALE-N-Terminus platziert. Für eine effiziente Kernlokalisierung wurde zusätzlich eine zweiteilige (bipartite) NLS (2x NLS, blau) an den C-Terminus jedes Konstruktes fusioniert (Le Cong *et al.*, 2013). Die *Sharkey-FokI-Endonukleasedomänen* und der *35 AS-linker* (violett) aus Minczuk *et al.*, 2008, sind gekennzeichnet. (C) Restriktionsanalyse mit den in (A) und (B) beschriebenen TALEN-Varianten. *In vitro* transkribierte und translatierte Zef-TALEN wurden für eine Stunde bei 37°C mit einem linearisierten Plasmid, das entweder das Ziel-DNA-Sequenzen tragende *tuba1a*- oder *tardbp*-Fragment enthielt, inkubiert. Die inaktivierten Reaktionen wurden über ein 0,8%-iges Agarosegel aufgetrennt und der Anteil TALEN-geschnittenen Plasmids mit ImageJ ermittelt. (D) Immunoblots der verwendeten TALEN-Varianten. Die *in vitro*-Transkriptions- und Translationsreaktionen der TALEN-Paare, scTALEN oder dscTALEN wurden wie unter Material und Methoden beschrieben aufgearbeitet, sodass jeweils die in der Restriktionsanalyse verwendete Menge aufgetrennt wurde.

3.3.1.3.2 Für *CCR5* entworfene scTALEN- und Doppel-scTALEN zeigen keine Aktivität in humanen Zellen

Für den Test von scTALEN und dscTALEN in der humanen Zelllinie Hek293T-RExTM wurden neue CMV-Promotor-getriebene *TALEN*-Expressionskonstrukte erstellt, die im Vergleich zu den in 3.3.1.2.2 vorgestellten TALEN-Varianten (Abb. 3.3.4 A) über die C-terminale, bipartite NLS verfügten um einen möglichst effizienten Kerntransport zu ermöglichen (Abb. 3.3.5 A). Darüber hinaus wurden die verschiedenen TALEN-Varianten über ein sogenanntes selbstspaltendes 2A-Peptid mit GFP zur Überprüfung einer erfolgreichen Transfektion verknüpft (Abb. 3.3.5 A; Ran *et al.*, 2013). Als Kontrollen dienten neben nicht transfizierten Zellen, das bereits in Restriktionsanalysen getestete TALEN-Paar L+R (Abb. 3.3.2, 3.3.3) sowie das TALEN-Paar (L+R)1c mit zusätzlicher NLS und Transfektionsmarker (Abb. 3.3.6 A). Die C-terminale scTALEN scR1c, sowie die beiden unterschiedlichen dscTALEN dscR1c und dscR2c wurden einzeln getestet, wohingegen die N-terminalen scTALEN-Varianten zur Erzeugung von Deletionen paarig, Nsc(L+R)1c und Nsc(L+R)2c, eingesetzt wurden. Zur Detektion möglicher Modifikationen wurden verschiedene Analysen durchgeführt. Zunächst wurde 72 Stunden nach Transfektion eine PCR des Lokus mit beschriebenen Oligonukleotiden zum Nachweis von Deletionen durchgeführt (Abb. 3.3.4 B; Mussolino *et al.*, 2011). Die Funktionalität dieses Ansatzes zur Deletionsdetektion konnte mit Cas9 und drei *CCR5*-gerichteten sgRNAs gezeigt werden (Abb. 3.3.4 B). So verursachte Cas9 mit zwei von drei möglichen Zweierkombinationen der sgRNAs Deletionen von erwartungsgemäß ca. 50 und 100 bp. Bei Präsenz nur einer sgRNA konnten hingegen keine Deletionen nachgewiesen werden. Bei der entsprechenden Analyse TALEN-transfizierter Zellen konnten jedoch keine Deletionen

detektiert werden, obwohl dies für die paarigen N-terminalen scTALEN und die dscTALEN angenommen wurde (Abb. 3.3.4 B). Auch bei Wiederholung der Analyse nach Hpy188III-Verdau der genomischen DNA zur Anreicherung modifizierter Sequenzen waren keine Deletionen nachweisbar. Um kleinere Modifikationen, die bei Aktivität eines einzelnen FokI-Endonukleasedomänen-Heterodimers zu erwarten wären, sichtbar zu machen, wurden die PCR-Produkte des angesteuerten Locus nach Denaturierung und Rehybridisierung einem T7 Endonuklease I (T7EI)-Verdau unterzogen (Abb. 3.3.4 B). Dabei konnten ähnliche Editierungseffizienzen durch die beiden getesteten TALEN-Paare nachgewiesen werden, jedoch durch keine der scTALEN-Varianten. Die korrekte Synthese aller getesteten TALEN in Hek293T-RExTM wurde mittels Immunoblot untersucht (Abb. 3.3.4 B). Es wurde festgestellt, dass alle Proteine mit dem richtigen Molekulargewicht produziert wurden und das GFP-Marker-Modul vollständig von den TALEN getrennt vorlag. Anhand dieser Beobachtung lässt sich also nicht erklären, warum von den getesteten TALEN-Varianten nur die TALEN-Paare Modifikationen in zellulären Systemen herbeiführen (siehe auch Kapitel 3.3.1.3.1), obwohl *in vitro* Nukleaseaktivität für alle getesteten Typen gezeigt wurde (Abb. 3.3.4, 3.3.5).

Die zusätzliche NLS schien im Vergleich der TALEN-Paare keinen Einfluss auf die Aktivität zu haben (Abb. 3.3.6 B). Es wäre daher anzunehmen, dass beide eine effiziente Kernlokalisierung zeigen. Um dies vor allem auch für die strukturell andersartigen scTALEN-Varianten zu überprüfen, wurden Lokalisierungsstudien durchgeführt (Abb. 3.3.7). Dazu wurden die R-Typ-TALEN, -scTALEN und -dscTALEN in Huh7-Zellen transfiziert und die subzelluläre Lokalisierung mittels Immunfluoreszenzfärbung untersucht (Abb. 3.3.6 A, 3.3.7). Dabei zeigte sich, dass sich die verschiedenen Konstrukte sehr unterschiedlich verhalten. Für die R-Typ-TALEN des Paares, R1c, konnte eine vollständige Kernlokalisierung in allen Zellen nachgewiesen werden (Abb. 3.3.7). Im Vergleich dazu wurden zu R baugleiche TALEN für andere Zielsequenzen, die entsprechend nur über die N-terminale NLS verfügten, sowohl im Nukleus einiger Huh7-Zellen als auch ausschließlich im Zytoplasma anderer Huh7-Zellen der gleichen Präparation lokalisiert. Das Verhältnis zwischen den beiden beobachteten Zuständen war ca. 1:1. Demnach erlaubte nur das Vorhandensein beider NLS eine vollständige Kernlokalisierung der einfachen TALEN eines Paares. Dennoch waren die Editierungseffizienzen beider TALEN-Paare mit ein oder zwei NLS vergleichbar (Abb. 3.3.6 B). Bei Betrachtung der Lokalisierung der scTALEN-Varianten, war für die verschiedenen N- oder C-terminalen scTALEN mit beiden NLS ein ähnliches Verhalten wie für die TALEN mit nur einer NLS zu beobachten (Abb. 3.3.7). Da letztere Aktivität in humanen Zellen zeigte, kann nicht angenommen werden, dass diese unvollständige Kernlokalisierung die Ursache für die

nicht nachweisbare Aktivität der scTALEN scR1c, NscR1c und NscR2c war (Abb. 3.3.6 B, 3.3.7). Im Kontrast dazu war die dscTALEN dscR1c ausschließlich im Zytoplasma nachzuweisen (Abb. 3.3.7). Ähnlich verhielt es sich für die andere, untersuchte dscTALEN dscR2c, die jedoch in vereinzelt Zellen im Nukleus zu beobachten war. Diese dscTALEN hatte einen vollständigen TALE N-Terminus. Dies traf auch auf die scTALEN NscR2c zu, die ebenfalls etwas häufiger im Kern zu detektieren war als das entsprechende Konstrukt mit verkürztem TALE N-Terminus, NscR1c.

Ursachen für die stark unterschiedliche subzelluläre Lokalisierung der verschiedenen TALEN-Varianten könnten in der unterschiedlichen Proteingröße, -struktur oder in der Positionierung der NLS liegen. Allerdings liegt auch die NLS in den Konstrukten mit unterschiedlichen N-Termini an anderer Position zur scFokI. Obgleich die C-terminale, bipartite NLS in einer einfachen TALEN die Kernlokalisierung maximal verbessert, scheint dieser Effekt C-terminal einer scFokI nicht aufzutreten. Auch in Konstrukten mit einer scFokI an einem Verkürzten N-Terminus befindet sich die N-terminale NLS stromabwärts der dimeren Endonukleasedomäne. Es wäre denkbar, dass eine einfache oder bipartite NLS, die einer scFokI folgt, möglicherweise der Kernimportmaschinerie strukturell nicht zugänglich ist. Demnach hätte die dscTALEN dscR1c keine zugängliche NLS, was die fehlende Kernlokalisierung begründen würde. Für diesen auch in Zebrafischembryonen getesteten Nukleasetyp wäre also anzunehmen, dass die fehlende Aktivität auf die nicht stattfindende Kernlokalisierung zurückzuführen ist. Allerdings konnte auch Aktivierung durch TALEs nachgewiesen werden, die vornehmlich nicht im Nukleus nachgewiesen wurden (Kapitel 3.2.6, Abb. 3.2.9). Auch wenn die Lokalisierung als Ursache für die ausbleibende Aktivität der dscTALEN angenommen werden könnte, ist nicht geklärt, warum die einfachen scTALEN nicht in zellulären Systemen aktiv waren. Diese waren mit vergleichbarer Häufigkeit zur TALEN R des funktionalen Paares L+R vermutlich ausreichend in Zellkernen vorhanden. Es wäre allerdings möglich, dass die Nukleaseaktivität der scTALEN ihren eigenen Nachweis maskiert. Schneiden scTALEN, führt die zelluläre DNA-Reparatur meist zu eher kleinen Veränderungen, die wahrscheinlich nicht weitreichend genug sind, um die Erkennungssequenz der scTALEN zu zerstören. Bei TALEN-Paaren führen die induzierten Veränderungen im Zielgen vermutlich häufig zur Verringerung oder Vergrößerung des *spacers* zwischen den einzelnen Zielsequenzen, wodurch eine effiziente Dimerisierung der FokI-Endonukleasedomänen verhindert würde. Eine scTALEN könnte bei Erhalt ihrer Zielsequenz jedoch erneut schneiden und zwar solange bis die DNA-Reparatur die Erkennungssequenz soweit verkürzt hat, dass der TALEN nicht mehr binden kann. PCR-basierte Detektionmethoden versagen daher beim Nachweis dieses Ereignisses. Mit einer

dscTALEN, die ihre Zielsequenz theoretisch ausschneidet, könnte dies umgangen werden. Leider ist nicht klar, ob die beiden scFokI-Domänen auch zeitgleich schneiden können. Die Optimierung der Kernlokalisierung von dscTALEN, z.B. durch NLS stromaufwärts der beiden scFokI, oder längere Wirkzeiten der scTALEN in den Zellen, um eine Zerstörung der Zielsequenz zu erreichen, könnten den Nachweis einer möglichen Aktivität eventuell erlauben.

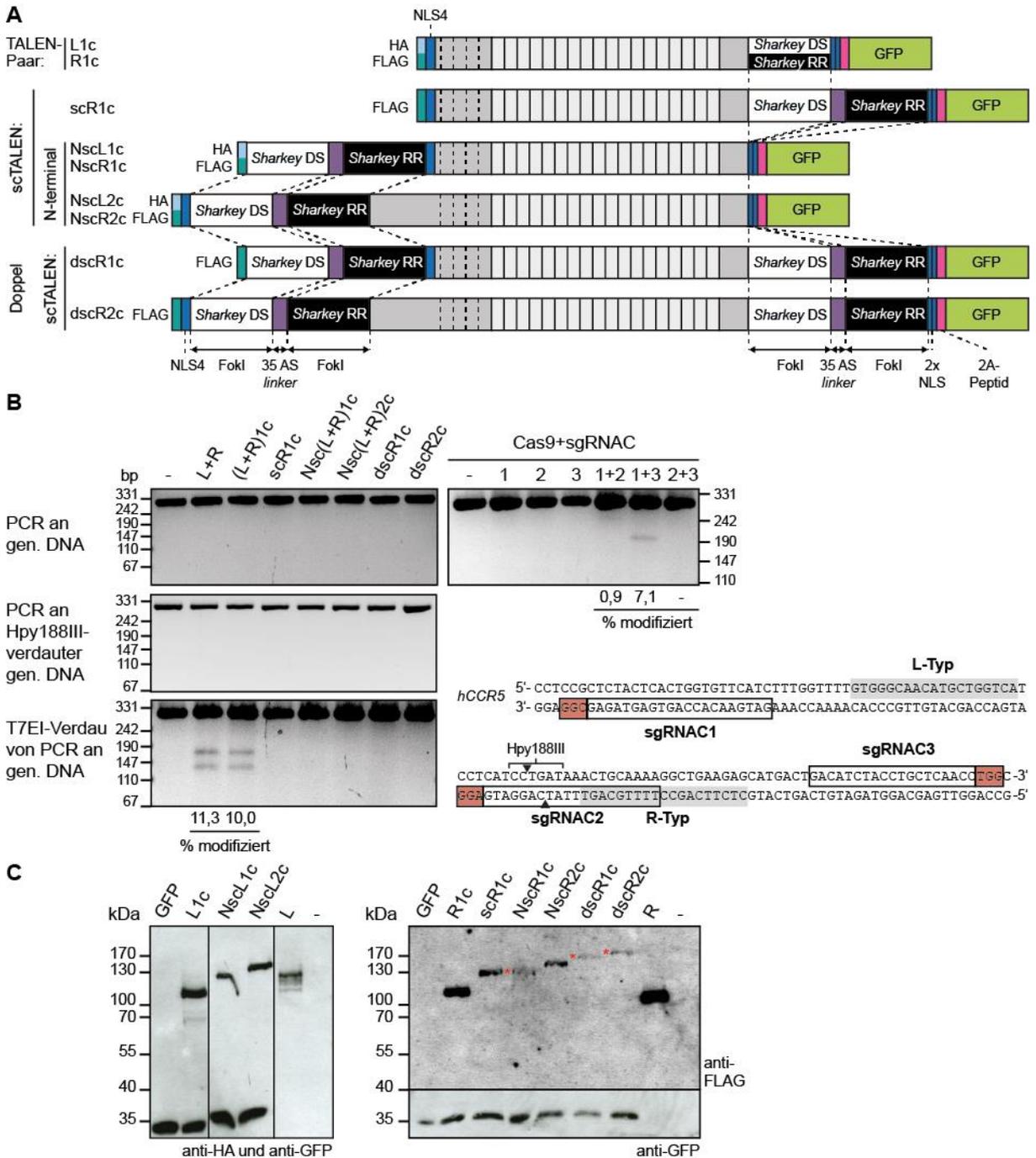


Abb. 3.3.6: Weder mit scTALEN noch mit dscTALEN sind Mutationen in humanen Zellen nachweisbar. (A) Schematische Darstellung vom Aufbau der getesteten TALEN. Das TALE-Grundgerüst der ersten TALEN-Variante (1c) und die *repeats* der L- und R-Typ-TALEN wurden bereits detailliert in Abb. 3.3.1 beschrieben. Die TALEN der zweiten Variante (2c) enthalten einen vollständigen TALE-N-Terminus. Im Unterschied zu den R-

Typ-TALEN wurden die L-Typ-TALEN mit einem HA-Epitop (hellblau) versehen. Die in Kapitel 3.2.6 ermittelte NLS4 (blau, Abb. 3.2.9) wurde entweder unmittelbar N-terminal des verkürzten TALE-N-Terminus platziert oder zwischen N-terminales FLAG-Epitop (cyan) und scFokI fusioniert, wenn ein vollständiger TALE-N-Terminus verwendet wurde. Für den Test in humanen Zellen wurde zusätzlich eine zweiteilige (bipartite) NLS (2x NLS, blau) an den C-Terminus jeden Konstruktes fusioniert (Le Cong *et al.*, 2013). Die Sharkey-FokI-Endonukleasedomänen und der 35 AS-linker (violett) aus Minczuk *et al.*, 2008, sind gekennzeichnet. Zur Überprüfung erfolgreicher Transfektion wurde C-terminal GFP über ein selbst-spaltendes 2A-Peptid fusioniert (Ran *et al.*, 2013). (B) Genotypisierung von *CCR5* in Hek293T-RexTM-Zellen 72 h nach Transfektion der unter (A) beschriebenen TALEN. Dabei wurde zur Transfektion einzelner TALEN-Expressionskonstrukte die Gesamtmenge zu applizierender DNA ausgeschöpft. Zur paarigen Anwendungen bestimmter TALEN wurde die transfizierte DNA-Menge im Verhältnis 1:1 auf die beiden Plasmide aufgeteilt. Als generelle Kontrollen dienten nicht transfizierte Zellen (-), sowie das in 3.3.1.1 beschriebene TALEN-Paar L und R (Abb. 3.3.2 A). Zur Überprüfung der durchgeführten Deletionsanalyse wurden jeweils zwei *CCR5*-gerichtete sgRNAs (sgRNAC1-3) in MLM3636 (Keith Joung) zu gleichen Teilen im Verhältnis 1:1 mit dem *cas9*-Expressionsvektor JDS246 (Keith Joung) transfiziert. Als Negativkontrollen hierfür dienten der Expressionsvektor MLM3636 (Keith Joung) ohne (-) und mit jeweils einem der Gene für sgRNAC1-3 im Verhältnis 1:1 mit dem *cas9*-Expressionsvektor JDS246. Zur Analyse möglicher TALEN-induzierter Genommodifikationen wurde zunächst eine PCR mit den in Mussolino *et al.*, 2012, für *CCR5* beschriebenen Oligonukleotiden an isolierter genomischer DNA vor und nach einem Restriktionsverdau mit Hpy188III zur Detektion größerer (ab ca. 50 bp) Deletionen durchgeführt. Um kleinere Deletionen, Insertionen und Punktmutationen festzustellen, wurde das PCR-Produkt der nicht Hpy188III-verdauten genomischen DNA denaturiert und unter Bedingungen, die die Bildung von Heteroduplex-DNA erlauben, rehybridisiert und im Anschluss mit der T7 Endonuklease I (T7EI) behandelt. Dabei erlaubt die Restriktion der Heteroduplex-DNA durch T7EI den Nachweis von Mutationen. Die T7EI-verdauten oder unverdauten PCR-Produkte wurden über ein 1,5%-iges Agarosegel aufgetrennt und der Anteil des geschnittenen oder kleineren PCR-Produktes mit ImageJ ermittelt. Im Anschluss wurde aus diesem Wert für die T7EI-Genotypisierung der Prozentsatz der Modifikation nach Guschin *et al.*, 2010, berechnet. (C) Immunoblots der verwendeten TALEN-Varianten. Lysate von Hek293T-RExTM-Zellen, die 24 h zuvor mit dem jeweiligen TALEN- oder einem *gfp*-Expressionsvektor transfiziert worden waren, wurden so aufgearbeitet, dass für die verschiedenen TALEN-Proben, sowie die nicht transfizierte Leerkontrolle (-), 4,4 µg und für GFP-Proben 3,3 µg Totalprotein pro Spur aufgetrennt werden konnten. Die Detektion erfolgte mit Hilfe von Antikörpern gegen die TALEN-fusionierten Epitope HA und FLAG, sowie jeweils gegen das mitsynthetisierte GFP. Schwache Signale im rechten Immunoblot sind durch einen roten Asterisk gekennzeichnet.

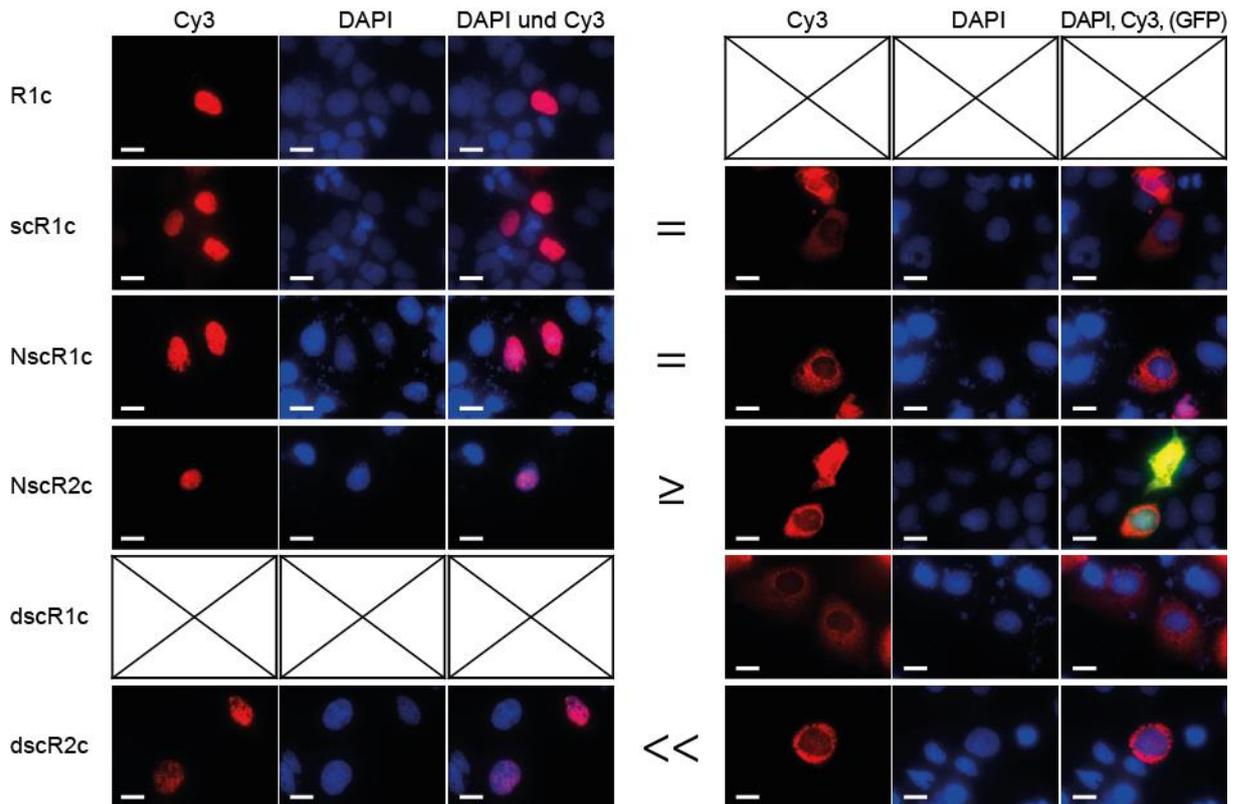


Abb. 3.3.7: TALEN mit scFokI zeigen eine unvollständige Kernlokalisierung. Lokalisierungsstudien der verwendeten TALEN in 24 h nach Transfektion fixierten Huh7-Zellen. Wie in Material und Methoden beschrieben wurde zur Detektion FLAG-markierter TALEN zunächst eine Immunfluoreszenzfärbung mit Cy3 durchgeführt und bei der Einbettung eine DAPI-Färbung der Nuklei vorgenommen. Es wurde zwischen Lokalisierung im Nukleus (links) und im Zytoplasma (rechts) kategorisiert. Falls einzelne TALEN-Varianten nicht in einem der Kompartimente nachgewiesen werden konnten, ist dies durch Fehlen der Bilder gekennzeichnet. Gleichheits- und Vergleichszeichen geben die geschätzte Häufigkeit des beobachteten Auftretens der Lokalisierungstypen wieder. Für NscR2c ist rechts auch das Signal des mitsynthetisierten GFP gezeigt. Der gezeigte Maßstab (weiß) entspricht 20 μm .

3.3.2 Hochspezifische TALEN-Paare

Um TALEN-Paare zu entwickeln, die zwei DNA-Sequenzen anhand eines Basenunterschiedes differenzieren können, wurde ein SNP in dem Gen *FANCI* (**F**anconi **a**nemia **c**omplementation **g**roup **I**, ehemals *KIAA1794*, Chromosom 15) als klinisch relevantes Beispiel gewählt. *FANCI* wurde 2007 als ein kausales Gen der Fanconi-Anemie identifiziert, wobei Defekte in *FANCI* in autosomal rezessiver Form zur Ausprägung der Erbkrankheit führen (Dorsman *et al.*, 2007; Sims *et al.*, 2007; Smogorzewska *et al.*, 2007). Die Krankheit ist gekennzeichnet durch Knochenmarksversagen, Myelodysplastische Syndrome (MDS) und eine Prädisposition für akute myeloische Leukämie (AML), sowie Kopf-Hals-Karzinome (Gueiderikh *et al.*, 2017). Ihre Ursache liegt in der Störung eines DNA-Reparaturmechanismus, dem sogenannten *interstrand cross-link* (ICL) *repair*, der für die Auflösung von DNA-Vernetzungen benötigt

wird. Dabei fungiert das Phosphoprotein FANCI im Heterodimer mit FANCD2 als Vermittler zwischen dem FANC-Kernkomplex, der den DNA-Schaden detektiert und beide Proteine monoubiquitiniert, und den eigentlichen DNA-Reparaturproteinen (Smogorzewska *et al.*, 2007; Gueiderikh *et al.*, 2017). Insgesamt sind neben *FANCI* 20 weitere Gene bekannt, die an diesem DNA-Reparaturmechanismus beteiligt sind und mit der Krankheit assoziiert werden (Gueiderikh *et al.*, 2017). Dennoch tritt die Fanconi-Anemie nur bei einem von ca. 250.000-350.000 Neugeborenen auf, wobei defektes FANCI wiederum eine seltene Ursache der Krankheit darstellt. Dementsprechend sind in solchen Fällen z. T. zwei verschiedene *FANCI*-Allele mit Funktionsverlust in gemischt heterozygoter Form beteiligt (Dorsman *et al.*, 2007). Die Mutationen in den defekten Allelen können dabei mehrere 1000 bp voneinander entfernt liegen. Ein gezielter Schnitt eines der beiden Allele anhand des pathogenen SNPs ohne Zerstörung des lokal wildtypischen zweiten Allels könnte durch Homologie-gerichtete Reparatur zur Wiederherstellung der Wildtypsequenz des ersten Allels führen. Bei einer rezessiven Erbkrankheit reicht die Wiederherstellung der Wildtypsequenz eines Allels bereits zur Beseitigung der Krankheitsursache aus. In dem hier als Zielsequenz gewählten Beispiel handelt es sich um ein pathogenes *FANCI*-Allel (NM_001113378.1(FANCI):c.1461T>A (p.Tyr487Ter)), bei dem eine T zu A Transversion im 15. Exon zur vorzeitigen Termination des Proteins führt. In einem klinischen Fall trat dieses Allel in gemischt heterozygoter Form neben einem noch unbeschriebenen Allel auf, das eine Mutation 22671 bp stromabwärts des Transversionslokus aufwies (Dr. Mark J. Osborn, University of Minnesota; persönliche Kommunikation). Ziel dieser Untersuchung war es nun, TALEN-Paare zu konstruieren, die bevorzugt die Zielsequenz des beschriebenen, pathogenen *FANCI*-Allels schneiden und die korrespondierende Wildtypsequenz nicht mehr oder kaum noch erkennen (Abb. 3.3.8). Dabei wurde ein heterodimerisierender Aufbau gewählt (Abb. 3.3.8 A). Zur Differenzierung der mutierten DNA-Sequenz des pathogenen Allels von der Wildtypsequenz wurden folgende Strategien angewendet:

- 1) Da bekannt ist, dass längere TALEs mehr Fehlpaarungen und kürzere TALEs weniger Fehlpaarungen tolerieren (Mali *et al.*, 2013 b), wurde zunächst die **repeat-Anzahl** der über dem SNP-bindenden TALEN variiert (Abb. 3.3.8 B). Um dabei die Gesamtzahl an *repeats* und den Abstand der TALEN eines Paares zueinander konstant zu halten, wurde die *repeat*-Anzahl der jeweils korrespondierenden TALEN angepasst (Abb. 3.3.8 B).
- 2) Um eine auf die Base genaue Spezifität der *repeat*-Region zu erreichen, wurde grundlegend für jede der vier DNA-Basen ein spezifischer RVD – NI, HD, NH und NG für A, C, G und T – gewählt (Abb. 3.3.8 C). Für G-bindende *repeats* ist bekannt, dass

sie unterschiedlich stark zur Aktivität künstlicher TALE und damit vorrausichtlich zur DNA-Bindung beitragen (Streubel *et al.*, 2012). Daher wurde jede nicht SNP-bindende TALEN zusätzlich in einer Variante mit dem stärkeren RVD NN für G konstruiert (Abb. 3.3.8 B, C). Diese TALEN sollten im Paar als eine Art **Anker** für die durch Fehlpaarung schwächer bindenden, über dem SNP lokalisierten TALEN wirken und die Aktivität des Paares ggf. gewährleisten.

- 3) Darüber hinaus sollte die Stringenz der SNP-bindenden TALEN durch zusätzliche **Fehlpaarungen** erhöht werden. Diese wurden durch bewusst falsch gepaarte RVD in der *repeat*-Region erzeugt (Abb. 3.3.9 A, 3.3.10 A). In Anhängigkeit von der *repeat*-Anzahl wurden dabei eine bis drei RBF platziert.

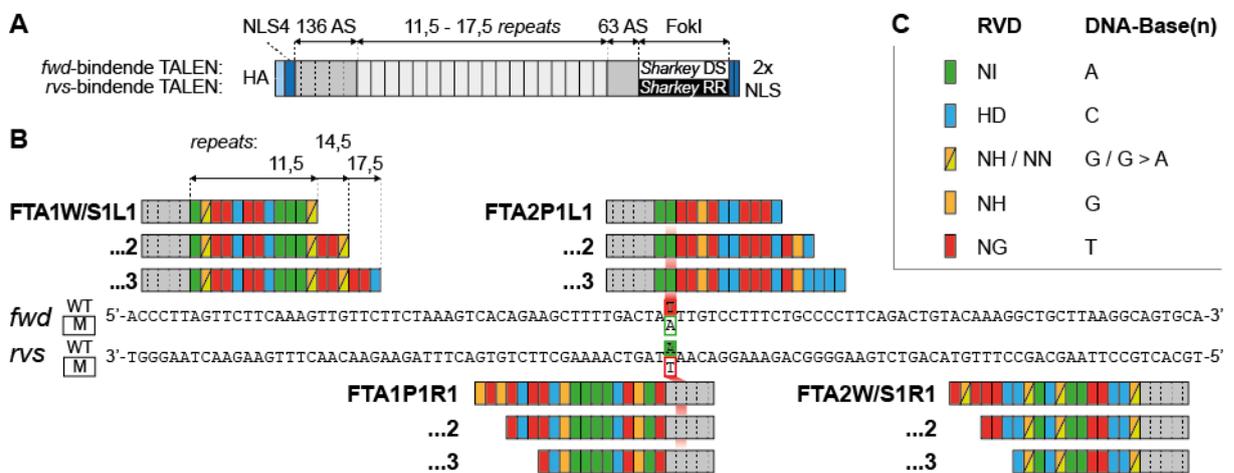


Abb. 3.3.8: Aufbau und Positionierung von TALEN-Paaren zur Unterscheidung des gesunden und des mutierten *FANCI*-Genes anhand eines Nucleotides. (A) Aufbau *forward* (*fwd*)- und *reverse* (*rvs*)-bindender TALEN für heterodimerisierende Paare, wie in 3.3.1.1 beschrieben. HA-Epitop (HA), fusionierte NLS-Varianten, die Anzahl N- und C-terminaler Aminosäuren (AS), sowie die Anzahl der *repeats* sind gekennzeichnet. (B) Positionierung der TALEN-Paare 1-3 beider Sets, *FANCI*-TALEN 1 und 2 (FTA1 und 2), am *fwd*- und *rvs*-Strang des wildtypischen (WT) oder mutierten (M) Allels des Gens *FANCI*. Die drei Paare jeden Sets unterscheiden sich in der *repeat*-Anzahl der L- und R-Typ-TALEN. Ein Paar besteht jeweils aus einer TALEN mit 11,5 und einer mit 17,5 *repeats* oder aus zwei TALEN mit 14,5 *repeats*. Die über dem SNP bindenden TALEN sind zur perfekten (P1) Erkennung der mutierten Sequenz entworfen, wodurch eine RBF mit der Wildtypsequenz entsteht. Die nicht über dem SNP bindenden TALEN wurden in zwei Versionen konstruiert. In Version W1 (*weak* 1) wurde das schwächere RVD NH und in Version S1 (*strong* 1) das stärkere RVD NN für G verwendet. (C) RVD-Legende zu (B).

Die entworfenen Sets von TALEN-Paaren binden dabei so, dass sich entweder die linke oder die rechte TALEN eines Paares über der Transversion befindet und so eine Fehlpaarung mit der wildtypischen *FANCI*-Sequenz entsteht (Fig. 338 B). Dabei ist der SNP unterschiedlich zu der jeweiligen *repeat*-Region positioniert. Beim ersten Set befindet sich der SNP an Position T₀ der

rechten TALEN (Abb. 3.3.9). Im zweiten TALEN-Set hingegen, erkennt der zweite *repeat* der linken TALEN den SNP (Abb. 3.3.10). Um die Aktivität und Spezifität der erstellten TALEN-Paare zu analysieren wurden *in vitro*-Restriktionsanalysen an DNA mit der Wildtyp- und mutierten *FANCI*-Sequenz durchgeführt (Abb. 3.3.9 B, 3.3.10 B).

Für das erste TALEN-Set, bei dem sich der SNP an Position T_0 befindet, konnte ein Einfluss aller drei angewandten Strategien auf die Spezifität beobachtet werden. Betrachtet man zunächst die TALEN-Paare, die keine zusätzlichen RBF enthalten, wird deutlich, dass die einzelne Fehlpaarung mit A statt T an Position 0 der Wildtypsequenz bereits die Aktivität der TALEN-Paare schwächt. Dieser Effekt nimmt mit zunehmender *repeat*-Anzahl der rechten TALEN ab, wobei stärkere RVD in der linken TALEN nur in einzelnen Fällen die Restriktionsaktivität fördern (vgl. Abb. 3.3.9 B Paar FTA1P1R1/S1L1 mit FTA1P1R1/W1L1). Für die TALEN mit zusätzlichen RBF fiel der Einfluss von *repeat*-Anzahl und Stärke des Ankers ähnlich aus. Die Diskriminierung zwischen Wildtyp- und mutierter Sequenz wurde jedoch viel deutlicher. Während die Schnitteffizienz aller TALEN mit zusätzlichen RBF auf der Wildtypsequenz mit der Fehlpaarung an Position 0 auf unter 2% sank, blieb Aktivität, wenn auch stark reduziert, an der mutierten *FANCI*-Sequenz erhalten. Darüber hinaus konnte anhand der rechten TALEN mit 11,5 *repeats* ein Einfluss der RBF-Position und Anzahl auf die Aktivität festgestellt werden. Die mehr C-terminal gelegene RBF an Position 11 führte zu einer stärkeren Absenkung der Schnitteffizienz als die an Position 2. Der stärkste Aktivitätsverlust war jedoch zu beobachten, wenn an beiden Positionen RBF platziert waren.

Für das zweite TALEN-Set, bei dem der SNP durch den zweiten *repeat* der linken TALEN erkannt wird, zeigte sich ein anderes Bild. Bei den TALEN-Paaren, die keine zusätzlichen RBF enthalten, war nur im Falle der linken TALEN mit 11,5 *repeats* ein negativer Einfluss der Fehlpaarung mit der Wildtypsequenz auf die Nukleaseaktivität festzustellen (Abb. 3.3.10 B). Durch zusätzliche RBF wurde jedoch auch für die linken TALEN mit 14,5 und 17,5 *repeats* eine Diskriminierung zwischen Wildtyp- und mutierter Sequenz möglich, wobei die Schnitteffizienz für die mutierte *FANCI*-Sequenz eher zu als abnahm. Die größten Aktivitätsunterschiede zwischen Wildtyp- und mutiertem *FANCI*-Allel waren für linke TALEN mit 11,5 *repeats* nachzuweisen. Dabei nahm auch hier die Nukleaseaktivität auf der Wildtypsequenz mit zunehmend C-terminaler Position und Anzahl der RBF ab, wohingegen die Aktivität auf der mutierten Sequenz konstant hoch blieb. Ein positiver Effekt einer stärkeren Anker-TALEN (hier: rechte TALEN) auf die Aktivität war in geringem Maße nur für Paare mit linken 11,5-er TALEN nachzuweisen. Für Paare, bei denen die linken TALEN über 11,5

repeats enthielten, beeinflusste ein Anker-TALEN mit stärkeren RVDs die Aktivität eher negativ.

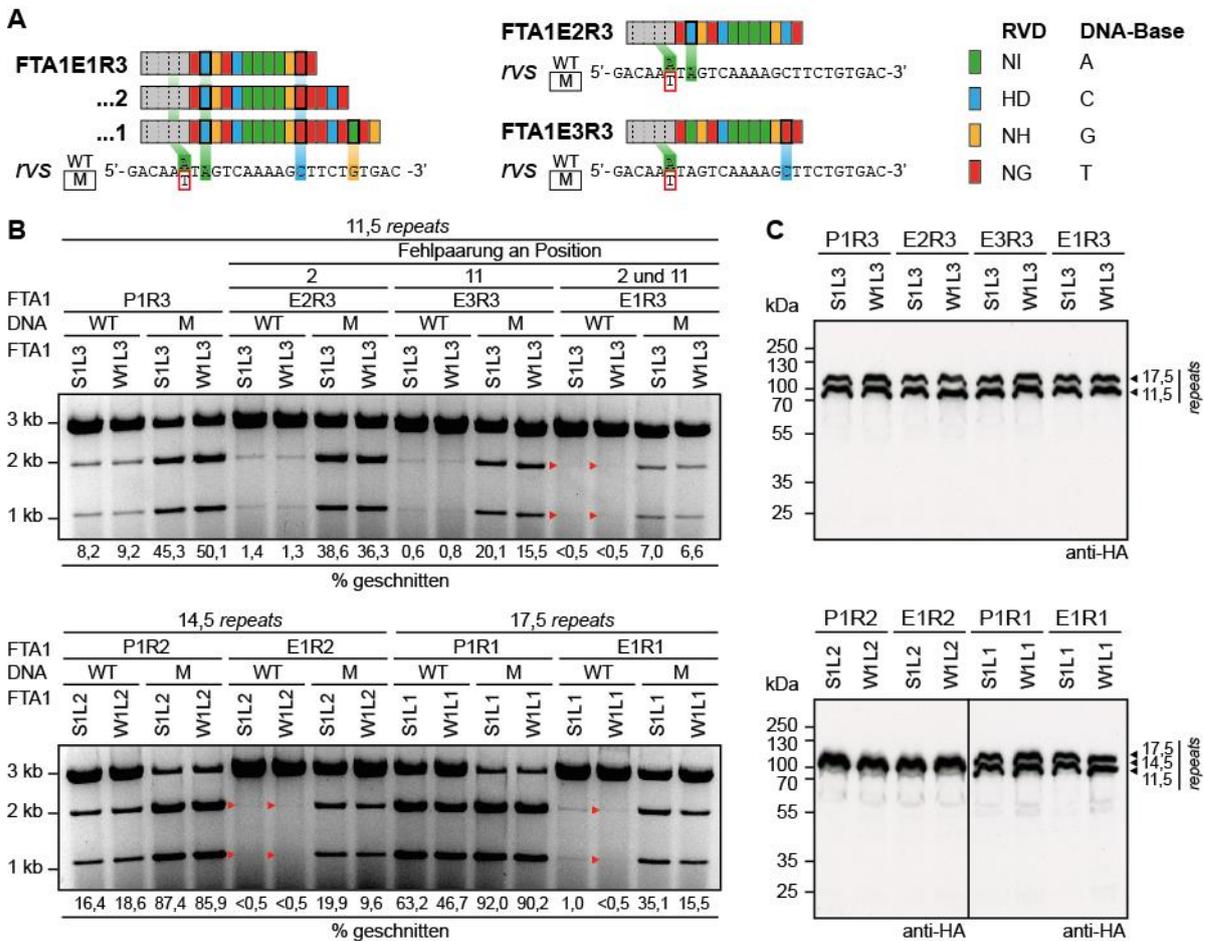


Abb. 3.3.9: TALEN-Paare können zwei Sequenzen anhand einer Fehlpaarung an Position T₀ unterscheiden.

(A) *Alignment* von fünf zusätzlichen R-Typ-TALEN des FTA1-Sets mit der wildtypischen (WT) und mutierten (M) *FANCI*-Sequenz. RVD-Austausche (fett schwarz umrahmt) führen zu Fehlpaarungen (*errors*, E1-3) mit beiden Zielsequenzen (farbig markiert). Nur in der M-Sequenz ist ein Thymin an Position 0 vorhanden. Im WT erzeugt ein Adenin eine RBF. (B) Restriktionsanalyse mit den TALEN-Paaren des FTA1-Sets. *In vitro* transkribierte und translatierte TALEN wurden für eine Stunde bei 37°C mit einem linearisierten Plasmid, das entweder *FANCI* WT- oder M-DNA enthielt, inkubiert. Die inaktivierten Reaktionen wurden über ein 0,8%-iges Agarosegel aufgetrennt und der Anteil TALEN-geschnittenen Plasmids mit ImageJ ermittelt. (C) Immunoblots der verwendeten TALEN. Die *in vitro*-Transkriptions- und Translationsreaktionen der TALEN-Paare wurden wie unter Material und Methoden beschrieben aufgearbeitet, sodass jeweils die in der Restriktionsanalyse verwendete Menge aufgetrennt wurde.

Zur Kontrolle der korrekten *in vitro* Transkription und Translation, sowie vergleichbarer Mengen der verwendeten TALEN wurden Immunoblots angefertigt (Abb. 3.3.9 C, 3.3.10 C). Dabei konnten alle Proteine mit dem richtigen Molekulargewicht und ähnlicher Signalintensität nachgewiesen werden. Auf Grund der Proteinsynthese der Paare in einem Reaktionsansatz

konnten die beiden 14,5-er TALEN in einem Ansatz nicht unterschieden werden. Da jedoch für alle Paare Nukleaseaktivität detektiert wurde, kann eine effektive Synthese beider 14,5-er TALEN angenommen werden.

Zusammenfassend ist festzustellen, dass zusätzliche RBF gut geeignet sind, um die Unterscheidung zweier DNA-Sequenzen anhand eines Basenpaares durch TALEN zu ermöglichen. Dabei könnten RBF an mehr C-terminalen Positionen zu besseren Ergebnissen beitragen. Eine Fehlpaarung an Position 0 hat zudem einen sehr starken Einfluss. Darüber hinaus zeigen TALEN mit weniger *repeats* erwartungsgemäß eine höhere Stringenz auf Grund ihrer höheren Anfälligkeit für Fehlpaarungen. Eine verstärkte Aktivität durch eine Ankerfunktion der zweiten, nicht über dem SNP-bindenden TALEN konnte nur in einigen Fällen gezeigt werden. Hier scheinen noch andere Faktoren, wie das Verhältnis der *repeat*-Anzahl oder der DNA-Bindungsstärke zwischen den beiden TALEN eines Paares, eine Rolle zu spielen. Als geeignete Kandidaten für eine mögliche Anwendung an entsprechenden Patientenzellen wären hier die Paare FTA1E1R2/S1R2 und FTA2E1L1/S1R1 zu nennen. Paar FTA1E1R2/S1R2 verfügt zwar nur noch über ca. 20% Aktivität an dem mutierten *FANCI*-Allel, hat dabei jedoch eine kaum nachweisbare Aktivität auf der Wildtypsequenz. FTA2E1L1/S1R1 hingegen zeigt den größten Aktivitätsunterschied zwischen Wildtyp- und mutierter Sequenz, wobei das mutierte Allel mit 70%iger und die Wildtypsequenz mit ca. 4%iger Effizienz geschnitten wird. Ob eine Restaktivität von 4% auf der Wildtypsequenz in Zellen tatsächlich noch mit hoher Frequenz das lokal wildtypische Allele schneidet, müsste überprüft werden.

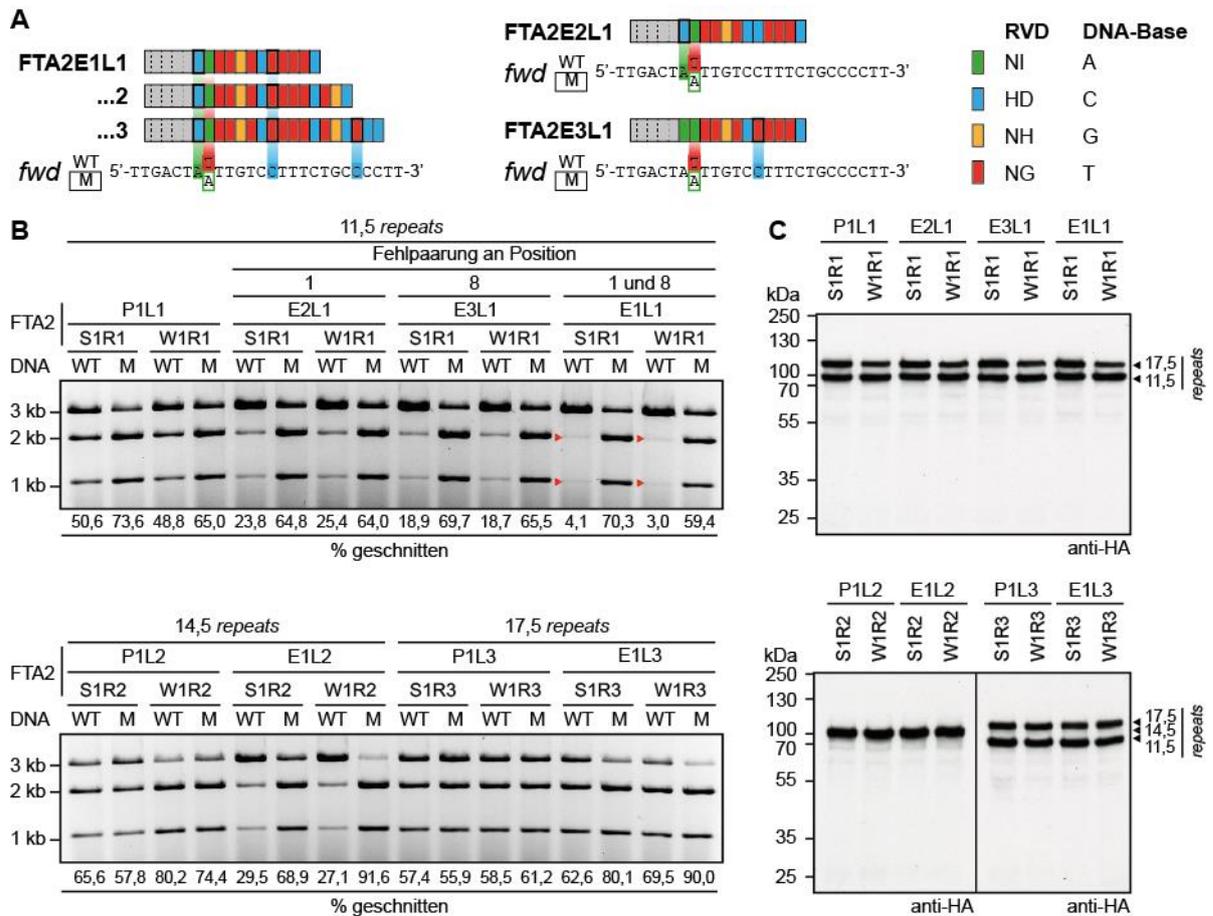


Abb. 3.3.10: TALEN-Paare können zwei Sequenzen anhand einer Fehlpaarung an einer *repeat*-Position unterscheiden. (A) *Alignment* von fünf zusätzlichen L-Typ-TALEN des FTA2-Sets mit der wildtypischen (WT) und mutierten (M) *FANCI*-Sequenz. RVD-Austausche (fett schwarz umrahmt) führen zu Fehlpaarungen (*errors*, E1-3) mit beiden Zielsequenzen (farbig markiert). Nur in der WT-Sequenz führt der fehlende SNP zu einer weiteren RBF. (B) Restriktionsanalyse mit den TALEN-Paaren des FTA1-Sets. *In vitro* transkribierte und translatierte TALEN wurden für eine Stunde bei 37°C mit einem linearisierten Plasmid, das entweder *FANCI* WT- oder M-DNA enthielt, inkubiert. Die inaktivierten Reaktionen wurden über ein 0,8%-iges Agarosegel aufgetrennt und der Anteil TALEN-geschnittenen Plasmids mit ImageJ ermittelt. (C) Immunoblots der verwendeten TALEN. Die *in vitro*-Transkriptions- und Translationsreaktionen der TALEN-Paare wurden wie unter Material und Methoden beschrieben aufgearbeitet, sodass jeweils die in der Restriktionsanalyse verwendete Menge aufgetrennt wurde.

4. Diskussion

4.1 Aberrante TALE-repeats

In dieser Arbeit wurden seltene *repeats* mit einer von 34 abweichenden AS-Anzahl von 30, 39, 40 oder 42 AS (Abb. 1.1 D), die in der Natur nur vereinzelt in *repeat*-Arrays auftreten, auf ihren Einfluss auf die DNA-Sequenzerkennung durch TALEs untersucht. Es war anzunehmen, dass sie die regelmäßige Struktur der *repeat*-Region stören, was sich auf die Aktivität der TALEs auswirken könnte. Interessanterweise tritt ein solcher aberranter *repeat* (39 AS) auch in den bedeutenden Virulenzfaktoren AvrXa7 und PthXo3 des Reispathogens *Xoo* auf, weshalb anzunehmen war, dass aberrante *repeats* die Geninduktion durch TALEs nicht generell verhindern.

4.1.1 Wie aberrante *repeats* die DNA-Erkennung durch TALEs beeinflussen

Mit Hilfe von GUS-Reporterstudien in *N. benthamiana* wurde festgestellt, dass künstliche TALEs mit jeweils einem aberranten *repeat* von 30, 40 oder 42 AS sowohl die von den RVDs abgeleitete optimale Zielsequenz (Box) als auch eine Box mit einer Leserasterverschiebung um -1 Nukleotid (Nt) in der Nähe des vom aberranten *repeat* erkannten Nt erkennen können (Abb. 3.1.1). Für den 40 AS-*repeat* konnte darüber hinaus gezeigt werden, dass die Erkennung einer -1 Nt-Leserasterverschiebung an verschiedenen Box-Positionen möglich ist, wenn der aberrante *repeat* entsprechend im Array verschoben wird (Abb. 3.1.2). Diese neue Form der DNA-Erkennung durch TALE-*repeats* konnte auch mit TALEN, die einen 40 AS-*repeat* enthielten, in Restriktionsstudien bestätigt werden (Abb. 3.1.1, 3.1.2). In weiteren GUS-Reporterstudien wurde die Position der von aberranten *repeats* tolerierten Nukleotid-Deletion auf vorzugsweise das von ihnen erkannte oder ein benachbartes Nukleotid eingegrenzt (Abb. 3.1.3, 3.1.4). Eine mögliche Erklärung für die Erkennung einer -1 Nt-Leserasterverschiebung durch aberrante *repeats* wäre, dass der aberrante *repeat* das Array verlassen kann („aus-loopt“) und so die Bindung der folgenden *repeats* um eine Basenposition in 5'-Richtung verschiebt (Abb. 4.1.1 B). Das ist in einem reinen 34 AS-*repeat* Array nicht möglich (Abb. 4.1.1 A). Analysen mit Fehlpaarungsboxen (Abb. 3.1.5) ergaben darüber hinaus, dass sowohl der aberrante als auch die ihm nachfolgenden *repeats* an der Bindung einer optimalen Box teilnehmen (Abb. 4.1.1). Dabei führten alle getesteten RVD-Base-Fehlpaarungen zu einem ähnlichen Aktivitätsverlust bei TALEs mit aberrantem *repeat* wie bei einem vergleichbaren TALE ohne aberrantem *repeat* (Abb. 3.1.5).

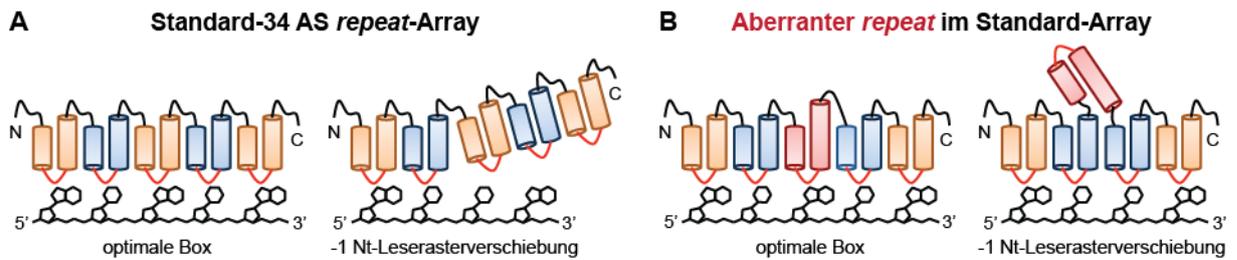


Abb. 4.1.1: Modell der DNA-Erkennung durch ein Standard-repeat-Array aus 34 AS repeats (A) ohne oder (B) mit aberrantem repeat. Die schematische Darstellung zeigt jeweils ein Aligment der repeats an den Leitstrang der DNA einer optimalen Zielsequenz (Box) oder einer Box mit einer -1 Nukleotid (Nt)-Leserasterverschiebung. An der Box mit Leserasterverschiebung loopt der aberrante repeat (dunkelrot) aus. Orange = Purin-bindend, Blau = Pyrimidin-bindend. Abb. verändert nach Richter *et al.*, 2014.

Dass alle untersuchten, aberranten repeats in ähnlicher Weise dazu in der Lage waren, die Erkennung einer optimalen und einer -1 Nt-Leserasterverschiebungsbox durch einen TALE zu vermitteln, könnte durch die Sequenzverwandtschaft mit dem üblichen 34 AS repeat erklärt werden. Dabei kommt die von 34 abweichende AS-Anzahl aberranter repeats durch Deletion oder Duplikation von Aminosäuren in der ersten oder zweiten α -Helix des repeats zustande (Abb. 1.1 D; Boch und Bonas, 2010). Vermutlich lassen diese Veränderungen der repeat-Helices noch immer die Ausbildung von inter-repeat-Interaktionen zu, die das repeat-array stabilisieren (Deng *et al.*, 2012; Mak *et al.*, 2012), sodass eine optimale Zielsequenz erkannt wird. Andererseits passen die aberranten repeats wahrscheinlich nicht perfekt in das repeat-Array, wodurch lokale Flexibilität und strukturelle Spannung erzeugt werden könnten. Das könnte dem aberranten repeat erlauben bei Bindung einer -1 Nt-Leserasterverschiebungsbox aus dem repeat-Array zu loopen. Außerdem haben Computersimulationen eines TALEs ohne aberranten repeat bereits eine hohe Flexibilität des Moleküls berechnet (Wan *et al.*, 2013), was die Annahme unterstützt, dass der aberrante repeat unterschiedliche Konformationen im repeat-Array annehmen könnte.

Es ist weiterhin aus Kristallstrukturanalysen bekannt, dass die superhelikale Struktur von TALEs in DNA-gebundener Form viel dichter gepackt ist als in freier Form (Deng *et al.*, 2012). Jüngere Studien, bei denen das Verhalten einzelner TALE-Moleküle an der DNA beobachtet wurde, haben darüber hinaus gezeigt, dass TALEs bei der Suche nach einer passenden Zielsequenz verschiedene Modi durchlaufen (Cuculis *et al.*, 2015 und 2016). Im abgeleiteten Modell initiiert der N-Terminus zunächst die unspezifische Bindung an die DNA, gefolgt von einem eindimensionalen, rotationsunabhängigen Gleiten entlang der Doppelhelix (Cuculis *et al.*, 2015 und 2016; Becker und Boch, 2016). Es wurde vermutet, dass die repeat-Region während dieses Suchmodus eine ähnlich lockere Konformation annimmt wie in einem DNA-

freien TALE (Cuculis *et al.*, 2015). Dadurch wäre die Superhelix der TALE-*repeats* asynchron mit der DNA-Doppelhelix, was die aus Kristallstrukturen bekannten Interaktionen der *repeats* mit dem Zuckerphosphat-Rückgrat der DNA (Deng *et al.*, 2012) minimieren und eine sehr schnelle Suche des TALEs entlang der DNA ermöglichen würde (Cuculis *et al.*, 2015). Beim Übergang in den Erkennungsmodus würden die TALE-*repeats* nun versuchen die für eine „feste“ DNA-Bindung benötigten, molekularen Kontakte mit der DNA herzustellen (Cuculis *et al.*, 2015). Das sind neben den Interaktionen mit dem Rückgrat der DNA vor allem die Kontakte der jeweiligen 13. AS eines jeden *repeats* mit den DNA-Basen (Deng *et al.*, 2012; Mak *et al.*, 2012). Dadurch würde die Kondensation der TALE-*repeats* in die DNA-gebundene Form nur bei Erkennung einer geeigneten Zielsequenz erfolgen (Cuculis *et al.*, 2015).

Wie könnte nun ein aberranter *repeat* den vorgeschlagenen Such- und Erkennungsvorgang beeinflussen? Da anzunehmen ist, dass ein aberranter *repeat* zu einer lokalen Störung der sonst hochregelmäßigen *repeat*-Arraystruktur führt, könnte dies eine Separierung der *repeat*-Region in zwei DNA-Bindedomänen hervorrufen, die evtl. unabhängig voneinander vom Such- in den Erkennungsmodus übergehen, bzw. an der DNA kondensieren können. Dabei müssten die Wechselwirkungen zwischen den *repeats* beider Bindedomänen mit der DNA zu einem energetisch günstigen Zustand führen, sodass der aberrante *repeat* entweder bei Erkennung einer optimalen Box mit in das Array eingefügt werden oder bei Erkennung einer -1 Nt-Leserasterverschiebungsbox aus-loopen kann. Falls der aberrante *repeat* wirklich zu einer Teilung der DNA-Bindedomäne von TALEs führt, könnte dies eine generelle Schwächung der TALE-DNA-Bindung verursachen und eine mögliche Erklärung liefern, warum einige TALEs mit aberranten *repeats* nur eine schwache oder gar keine Aktivität zeigten (Abb. 3.1.2, 3.1.4). Dabei handelt es sich zum einen um einen 17,5 *repeat*-TALE mit einem 40 AS-*repeat* an 14. Position, der sowohl mit der optimalen als auch einer -1 Nt-Leserasterverschiebungsbox eine nur schwache GUS-Reporteraktivierung hervorrief (Abb. 3.1.2). Es könnte sein, dass die DNA-Interaktionen der vier dem Aberranten folgenden *repeats* zu schwach sind, um eine stabile Bindung aufrechtzuerhalten. Dementsprechend könnte der C-Terminus des TALEs zeitweise relativ zur DNA fehlplatziert sein, wodurch keine kontinuierliche Reporteraktivierung erfolgen würde. Zum anderen zeigten TALEs mit 11,5 *repeats* und einem 40 AS *repeat* an achter oder dritter Position keine bzw. nur eine geringe Aktivierung von Reporterkonstrukten mit den entsprechenden Boxen (Abb. 3.1.4), was mit der geringen *repeat*-Anzahl dieser TALEs zusammenhängen könnte. So wurde für TALEs mit 15,5 *repeats* eine relativ stabile und dauerhafte DNA-Assoziation im Suchmodus gezeigt, während es zur gelegentlichen Dissoziation kürzerer TALEs mit nur 11,5 *repeats* von der DNA kam (Cuculis *et al.*, 2016).

Außerdem wurde gezeigt, dass TALEs erst ab 10,5 *repeats* effizient aktivieren (Boch *et al.*, 2009). Daher ist anzunehmen, dass bereits die unspezifische DNA-Bindung im Suchmodus bei kürzeren schwächer als bei längeren TALEs ist. Eine zusätzliche Schwächung der unspezifischen DNA-Interaktion durch eine möglicherweise vom 40 AS-*repeat* verursachte Diskontinuität des *repeat*-Arrays, könnte die Häufigkeit mit der ein kurzer TALE seine Zielsequenz findet weiter verringern und die Reporterogenaktivierung schwächen. Das Ausbleiben jeglicher Reporterogenaktivierung durch einen 11,5 *repeat*-TALE mit einem 40 AS *repeat* an achter Position (Abb. 3.1.4 A, TALE40p8s) lässt sich so jedoch nicht erklären und könnte mit einer zusätzlichen Störung des Erkennungsmodus zusammenhängen. So sind kürzere TALEs auch generell anfälliger gegenüber Fehlpaarungen zwischen den RVDs und DNA-Basen (Mali *et al.*, 2013 b), was wahrscheinlich auf eine Verhinderung der spezifischen DNA-Bindung durch den TALE zurückzuführen ist. Falls ein aberranter *repeat* die DNA-Bindung schwächt, ist es vorstellbar, dass zumindest eine der beiden potentiell separierten DNA-Bindedomänen eines TALEs mit aberrantem *repeat* wie eine Art „Kondensationskeim“ allein in der Lage sein muss, die DNA-gebundene Konformation anzunehmen. Im Falle von TALE40p8s bestehen die potentiell separierten Arrays aus sieben bzw. vier *repeats* und sind möglicherweise zu kurz, um allein energieeffizient an die DNA zu binden (Abb. 3.1.4 A).

Interessanterweise scheint ein 40 AS-*repeat* im N-terminalen Bereich der *repeat*-Region gut toleriert zu werden, was durch die mit einem Standard-TALE vergleichbare GUS-Reporteraktivierung durch einen TALE mit diesem *repeat* an dritter Position gezeigt werden konnte (Abb. 3.1.2, 3.1.4). Das ist überraschend, da zum einen die verbliebene N-terminale Array-Hälfte aus nur zwei *repeats* besteht und zum anderen bekannt ist, dass den N-terminalen *repeats* eine größere Wichtung bei der DNA-Erkennung zukommt als den mehr C-terminal gelegenen *repeats* (Meckler *et al.*, 2013; Schreiber *et al.*, 2015). Dass ein aberranter *repeat* im N-terminalen Bereich des Arrays einen TALE nicht zu beeinträchtigen scheint, könnte an den besonderen Eigenschaften des TALE-N-Terminus liegen. So wurde gezeigt, dass der N-Terminus mit den vier degenerierten *repeats* selbst eine hohe DNA-Affinität hat (Gao *et al.*, 2012) und für die initiale unspezifische DNA-Bindung, sowie den schnellen Suchmodus benötigt wird (Cuculis *et al.*, 2015). Darüber hinaus ist der *repeat* -1 für die Erkennung von T₀ verantwortlich (Mak *et al.*, 2012; Lamb *et al.*, 2013; Tsuji *et al.*, 2013; Schreiber *et al.*, 2014), wobei Kristallstrukturdaten darauf hinweisen, dass diese Interaktion evtl. nicht bei jedem TALE in der DNA-gebundenen Konformation erhalten bleibt (Stella *et al.*, 2013). Weiterhin wurde eine Kondensation der degenerierten, N-terminalen *repeats* nicht gezeigt. Der 40 AS-*repeat* an dritter Array-Position könnte also ein Ausnahmefall sein, bei dem der C-terminale Bereich der

repeat-Region die Zielsequenz erkennt und entsprechend an der DNA kondensiert, während der N-Terminus und die beiden verbliebenen N-terminalen *repeats* die DNA evtl. unkomprimiert binden. Das könnte auf Grund der geringen Anzahl von nur zwei auch für die regulären *repeats* möglich sein (vgl. DNA-gebundene und freie TALE-Struktur in Deng *et al.*, 2012). Dabei scheint dieser Prozess unabhängig von der Teilnahme des 40 AS-*repeat* an der DNA-Erkennung spezifisch zu sein, was ein Vergleich mit einem entsprechenden Standard-TALE zeigt, der -1 Nt-Leserasterverschiebungsboxen nicht erkennt (Abb. 3.1.2, 3.1.4).

Ob sich alle aberranten *repeats* in kurzen TALEs oder an verschiedenen Array-Positionen wie der 40 AS-*repeat* verhalten, müsste noch untersucht werden. Darüber hinaus kann aus den vorhandenen Daten nicht abgeschätzt werden, ob aberrante *repeats* erst bei Bindung einer -1 Nt-Leserasterverschiebungsbox das Array verlassen oder auch im DNA-freien bzw. im unspezifisch assoziierten Zustand plastisch ihre Konformation wechseln und zeitweilig ins Array inseriert oder aus-geloopt vorliegen können. Um das vorgeschlagene „aus-loop“-Modell zu überprüfen, wäre die Kristallstruktur eines TALEs mit aberrantem *repeat* gebunden an eine -1 Nt-Leserasterverschiebungsbox notwendig. Kristallstrukturanalysen und Simulationen der molekularen Dynamik eines TALEs mit aberrantem *repeat* in DNA-freier Form könnten auch Aufschluss darüber geben, wie plastisch sich ein aberranter *repeat* im Array verhält. Einen ersten Hinweis auf das dynamische Verhalten von aberranten *repeats* könnte der jüngst untersuchte 28 AS-*repeat* (Deletion der sechs AS von Position 23 bis 28) liefern, der ebenfalls einzeln im Standard-Array natürlich verkürzter TALEs von *X. oryzae* spp. auftritt (Read *et al.*, 2016). Ähnlich den hier untersuchten 40 AS-*repeats* im Tandem (Abb. 3.1.6) ermöglicht dieser einem TALE zwar die Erkennung einer entsprechenden -1 Nt-Leserasterverschiebungsbox, aber keiner optimalen Box (Read *et al.*, 2016). Es wäre also anzunehmen, dass der 28 AS-*repeat* ebenso wie einer der beiden 40 AS-*repeats* bei der DNA-Bindung aus-loopen muss. Die verkürzten TALEs, die einen 28 AS-*repeat* enthalten, sind allerdings keine Transkriptionsaktivatoren, sondern inhibieren durch die Reigene *Xal* und *Xol*-vermittelte Resistenzen gegen *Xo*-Stämme (Ji, Z. *et al.*, 2016; Read *et al.*, 2016). Es wird angenommen, dass die vermittelte Resistenz durch die direkte Proteinbindung von TALEs im Zytoplasma verursacht wird und die verkürzten TALEs die Resistenzproteine zwar binden, aber die Auslösung der Resistenzreaktion verhindern (Read *et al.*, 2016). Dafür wird der 28 AS-*repeat* jedoch nicht benötigt (Read *et al.*, 2016), weshalb angenommen werden könnte, dass er die vom Resistenzprotein erkannte TALE-Struktur nicht verändert. Dementsprechend könnten auch aberrante *repeats* im DNA-freien Zustand ins *repeat*-Array integriert sein.

4.1.2 Die Flexibilität natürlicher TALEs

Als Pflanzenpathogene befinden sich *Xanthomonas spp.* im ständigen evolutiven Wettrüsten mit ihren Wirtspflanzen (Jones und Dangl, 2006). So finden sich in *Xoo* bedeutende Virulenzfaktoren, wie *OsSWEET*-induzierende TALEs, die für die Besiedlung der Wirtspflanze Reis essentiell sind. Im Gegenzug haben sich in Reis z. B. die von *Xal* und *Xol* codierten Resistenzproteine entwickelt, die TALEs erkennen und eine Resistenzreaktion auslösen (Ji, Z. *et al.*, 2016; Read *et al.*, 2016). Verkürzte *Xo*-TALEs, die als Antagonisten der Resistenzgene *Xal* und *Xol* agieren, können wiederum als evolutive Antwort auf diese TALE-vermittelte Pflanzenresistenz betrachtet werden (Ji, Z. *et al.*, 2016; Read *et al.*, 2016). Die in dieser Arbeit untersuchten, aberranten *repeats* erlauben einem TALE die Erkennung einer optimalen Zielsequenz, sowie einer -1 Nt-Leserasterverschiebungsbox. Diese Eigenschaft könnte für TALEs, die als Transkriptionsaktivatoren agieren, einen evolutiven Vorteil darstellen. Während TALEs häufig Zielsequenzen mit Substitutionen immer noch erkennen können (Boch *et al.*, 2009; Moscou und Bogdanove, 2009; Römer *et al.*, 2010; Antony *et al.*, 2010; Strauß *et al.*, 2012), führen Leserasterverschiebungen durch Deletionen oder Insertionen meist dazu, dass TALEs ihr Zielgen nicht mehr aktivieren können (Gu *et al.*, 2005; Yang *et al.*, 2006; Chu *et al.*, 2006a, b; Liu *et al.*, 2011; Yuan *et al.*, 2011). So findet sich eine Leserasterverschiebung um 1 Nt in der PthXo2-Zielsequenz im Promotor des *OsSWEET13*-Allels *xa25*, das eine Resistenz gegen entsprechende *Xoo*-Stämme vermittelt (Zhou *et al.*, 2015). Die *xa25*-Resistenz konnte mit einem künstlichen TALE, der ein 40 AS-*repeat* enthielt und die veränderte Zielsequenz von PthXo2 ansteuert, überwunden werden, wobei die *OsSWEET13*-Aktivierung auch am nicht-Resistenz-Allel erhalten blieb (Richter *et al.*, 2014). Mit diesem Experiment konnte nachgewiesen werden, dass die hier untersuchten aberranten *repeats* prinzipiell die Überwindung von Resistenzen erlauben, die auf Leserasterverschiebungen um 1 Nt beruhen. Um zu untersuchen, ob *OsSWEET14*-Allele mit einer solchen Leserasterverschiebung zur evolutiven Konservierung des 39 AS-*repeat* (40 AS-*repeat* RVD N*) in den bedeutenden Virulenzfaktoren AvrXa7 und PthXo3 beigetragen haben könnten, wurden entsprechende Zielsequenzen in den verfügbaren *OsSWEET14*-Promotoren verschiedener *Oryza spp.* gesucht (Richter *et al.*, 2014). Dabei wurde eine 1 bp-Insertion in der PthXo3-Box von *O. brachyantha* entdeckt, die im Unterschied zur *O. sativa*-Box mit Beteiligung des aberranten *repeat* gebunden werden könnte (Abb. 4.1.2; Richter *et al.*, 2014). Ob dadurch das Wirtsspektrum von *Xoo* tatsächlich erweitert wird, müsste untersucht werden.

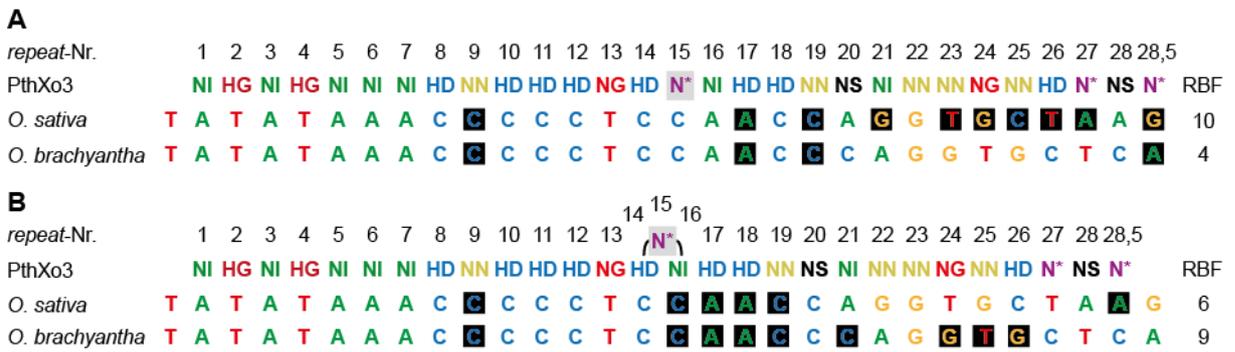


Abb. 4.1.2: PthXo3 könnte den *OsSWEET14*-Promotor in *Oryza brachyantha* in anderer Konformation als in *O. sativa* erkennen. Alignment der PthXo3 RVDs mit der *OsSWEET14*-Zielsequenz aus *O. sativa* und *O. brachyantha* (A) mit und (B) ohne Teilnahme des aberranten repeat an der DNA-Bindung. Eine Reduktion der RVD-Base-Fehlpaarungen (RBF, schwarz hinterlegt) zeigt, dass der aberrante repeat in *O. sativa* vermutlich aus-loop und in *O. brachyantha* nicht. 39 AS repeat in grau. Abb. verändert nach Richter *et al.*, 2014.

In dieser Arbeit wurde außerdem gezeigt, dass die natürlichen TALEs PthXo3 und AvrXa7 eine extrem hohe Toleranz gegen bis zu acht bzw. sechs RBF in ihrer jeweiligen Erkennungssequenz haben (Abb. 3.1.7, 3.1.8). So konnten sie neben Reporterkonstrukten einer optimalen oder einer -1 Nt-Leserasterverschiebungsbox auch solche mit Leserasterverschiebungen um -2, +1 und +2 Nt aktivieren. Mit Hilfe künstlicher AvrXa7-Derivate, in denen entweder der 39 AS- gegen einen 34 AS-repeat oder die RVDs multipler gegen solche singulärer Basenspezifität ausgetauscht wurden, konnte festgestellt werden, dass die Erkennung der Deletionsboxen (-1 und -2 Nt) hauptsächlich durch den aberranten repeat vermittelt wird, während die Insertionsboxen (+1 und +2 Nt) durch RVDs multipler Spezifität toleriert werden können (Abb. 3.1.9). Dabei wurde beobachtet, dass der Austausch stärkerer, unspezifischer RVDs (NS, NN) durch schwächere, spezifische RVDs (NI, NH) zu einer reduzierten Aktivierung an der optimalen Box führte (Abb. 3.1.9). Diese Ergebnisse zeigen zum einen, dass die generell hohe Toleranz natürlicher TALEs gegenüber RBF in ihren Zielsequenzen in diesem Fall vor allem durch starke RVDs, die mehrere Nukleobasen erkennen können, verursacht wird. Zum anderen deuten sie darauf hin, dass aberrante repeats unter Umständen auch die Erkennung von Leserasterverschiebungen von -2 Nt vermitteln können, was in vorrausgegangenen Analysen mit künstlichen TALEs nicht der Fall war (Abb. 3.1.1). Neben dem aberranten repeat unterschied sich das AvrXa7-Derivat mit spezifischen RVDs von den künstlichen TALEs mit einem 30, 40 oder 42 AS-repeat vor allem in der repeat-Anzahl. Mit 25,5 repeats war das artifizielle AvrXa7 deutlich länger als die TALEs mit 17,5 repeats (vgl. Abb. 3.1.1 und 3.1.9). Möglicherweise könnte also die durch einen aberranten repeat gewonnene TALE-Flexibilität bei ausreichender repeat-Anzahl der umgebenden Array-Abschnitte weiter ausgenutzt werden. Ob dies durch das Aus-looping mehrerer repeats zustande kommen könnte oder ein anderer

Mechanismus greift, kann nur spekuliert werden. Längere repeat-Arrays haben darüber hinaus auch den Vorteil, dass sie eine höhere Fehlpaarungstoleranz vermitteln (Mali *et al.*, 2013 b; Juillerat *et al.*, 2014).

Insgesamt lassen diese Ergebnisse die Formulierung einer Hypothese zu drei möglichen, evolutiven Anpassungen von TALEs zur Überwindung von Wirtsresistenzen zu. Demnach entwickeln sich in TALEs unter selektivem Druck, wie bei den wichtigen Virulenzfaktoren AvrXa7 und PthXo3, *repeat*-Arrays mit erstens möglichst vielen, zweitens starken, sowie unspezifischen und drittens besonderen *repeats*, die dem TALE neue Eigenschaften verleihen können. Diese Faktoren sollten auch bei der computergestützten Vorhersage von TALE-Zielsequenzen und somit Zielgenen berücksichtigt werden. So könnte auch die Rolle der übrigen natürlichen TALEs mit aberranten *repeats* für das jeweilige Pathogen untersucht werden. Alle bisher identifizierten aberranten *repeats* wurden dabei in den *X. oryzae* Pathovaren *oryzae* und *oryzicola* gefunden (Boch und Bonas, 2010; Richter *et al.*, 2014; Read *et al.*, 2016). Es wäre interessant, ob weitere aberrante *repeats* in diesen oder anderen Spezies auftreten und welche neuen Eigenschaften diese einem *repeat*-Array verleihen könnten, nicht zuletzt im Hinblick auf die biologische Funktion der entsprechenden TALEs.

4.1.3 Aberrante *repeats* erweitern das Anwendungsspektrum von TALEs

Der in dieser Arbeit entdeckte neuartige Erkennungsmodus von TALEs mit aberranten *repeats*, der es erlaubt neben optimalen Zielsequenzen auch solche mit -1 Nt-Leserasterverschiebungen zu nutzen, bietet auch interessante neue Möglichkeiten für die Anwendung von TALEs in der Biotechnologie. So könnten sowohl TALEs als auch TALEN zur Erkennung allelischer Genvarianten verwendet werden, die sich durch eine 1 Nt-Leserasterverschiebung unterscheiden. Beim Design solcher TALEs oder TALEN müsste jedoch die Positionierung des aberranten *repeat* im Array als auch dessen Länge beachtet werden, da TALEs mit aberrantem *repeat* und geringer *repeat*-Anzahl (11,5) eine stark reduzierte oder keine Reporteraktivität induzierten (Abb. 3.1.4). Für TALEs, die als Genaktivatoren eingesetzt werden sollen, ist für eine effiziente Aktivierung eine Platzierung des aberranten *repeats* am Anfang oder im Zentrum des *repeat*-Arrays zu empfehlen, da ein aberranter *repeat* im C-terminalen Bereich der *repeat*-Region die induzierende Aktivität schwächte (Abb. 3.1.2). Dies ist aus anderen Gründen auch für TALEN zu empfehlen. Einerseits wurde ihre Restriktionsaktivität nicht durch einen aberranten *repeat* im C-terminalen Bereich des Arrays beeinträchtigt (Abb. 3.1.2). Andererseits war unter anderem auch eine TALEN ohne aberrantem *repeat* dazu in der Lage, eine DNA zu schneiden, die eine entsprechende -1 Nt-Leserasterverschiebung im 3'-Bereich der Zielsequenz

enthielt (Abb. 3.1.2). Das zeigt, dass TALEN möglicherweise auf Grund der paarigen Anwendung eine höhere Toleranz gegenüber Fehlpaarungen im C-terminalen Bereich des *repeat*-Arrays haben als TALEs. Diese beobachtete Eigenschaft von TALEN-Paaren sollte auch bei deren Design beachtet werden, sodass ein Schneiden einer der Zielsequenz ähnlichen Sequenz an falscher Stelle (*off-target*) vermieden werden kann.

In dieser Arbeit wurden TALEs mit aberranten *repeats* zur Induktion der Transkription an Promotoren verschiedener HIV-1 Subtypen verwendet (Abb. 3.2.7). Dabei konnte auch eine effiziente Aktivierung der Subtypen E und F erreicht werden, die sich u. a. durch lokale 1 Nt-Deletionen von den Subtypen A-D unterscheiden. Neben dem aberranten *repeat* wurden in diesem (ArtABC6) und einem weiteren TALE (TALEABC5), der an anderer Stelle ohne Leserasterverschiebung bindet, erfolgreich RVDs multipler Spezifität zur Erkennung der verschiedenen Subtypen eingesetzt (Abb. 3.2.7). Die weitreichende, in dieser Arbeit um die aberranten *repeats* ergänzte Kenntnis über die individuellen Eigenschaften einzelner *repeat*-Typen erlauben also ein Design von DNA-Bindeproteinen mit einem definierbaren Spektrum verschiedener/alleler Zielsequenzen.

Im Unterschied zu TALEs erlaubt die auf Basenpaarung beruhende DNA-Bindung durch CRISPR/Cas9 per Definition keinen gezielten Einbau von *wobble*-Positionen, an denen die Erkennung mehrerer Nukleobasen möglich ist. Aber könnten mit CRISPR/Cas9 gezielt Leserasterverschiebungen wie mit aberranten *repeats* adressiert werden? Tatsächlich ist Cas9 in der Lage Zielsequenzen mit Indels durch Auswölbung der ungepaarten Nukleotide der sgRNA oder des komplementären DNA-Einzelstranges zu binden und zu schneiden (Lin *et al.*, 2014). Dabei wird die Auswölbung eines Nukleotides in der einzelsträngigen DNA (entspricht einer Leserasterverschiebung von +1 Nt) und in Ausnahmefällen von bis zu vier Nukleotiden in der sgRNA (entspricht Leserasterverschiebungen von bis zu -4 Nt) toleriert, wobei nur 1 Nt-Leserasterverschiebungen in beide Richtungen systematisch untersucht wurden. Weiterhin wurde gezeigt, dass 1 Nt-Leserasterverschiebungen im äußersten 5'-Bereich der sgRNA keine oder kaum Aktivitätseinbußen verursachen (Lin *et al.*, 2014), was vermutlich auf die generell hohe Fehlpaarungstoleranz des CRISPR/Cas9-Systems in diesem Bereich zurückzuführen ist (Hsu *et al.*, 2013). Außerhalb dieses Bereiches war die Erkennung von 1 Nt-Leserasterverschiebungen extrem sequenzabhängig (Lin *et al.*, 2014) und ist dementsprechend im Unterschied zu TALEs mit aberranten *repeats* nicht vorhersagbar. Darüber hinaus konnten Cas9 *off-targets* identifiziert werden, die neben einer 1 Nt-Leserasterverschiebung zusätzliche Fehlpaarungen enthielten (Lin *et al.*, 2014). Im Unterschied dazu scheint ein aberranter *repeat*

einen TALE eher zu schwächen und dadurch gegenüber Fehlpaarungen spezifischer zu machen (Abb. 3.1.5). Die Leserasterverschiebungstoleranz von Cas9 erscheint also auf Grund ihrer Unvorhersagbarkeit wenig für Anwendungen geeignet, muss aber beim sgRNA-Design bzgl. möglicher *off-targets* berücksichtigt werden. TALEs oder TALEN mit aberrantem *repeat* haben hingegen abhängig vom übrigen Array-Design (*repeat*-Anzahl, RVDs) eine hohe Spezifität für ihre optimale Box und Boxen mit -1 Nt-Leserasterverschiebung im Bereich des aberranten *repeat*.

4.2 Die TALE-vermittelte Transkriptionsinduktion am HIV-Promotor

Ein möglicher Ansatz, der zur Heilung von HIV-1 verfolgt wird, ist die sogenannte *shock and kill*-Strategie (Geeraert *et al.*, 2008). Dabei soll die virale Genexpression in latent infizierten Zellreservoirs induziert werden, um diese durch die Virusreproduktion oder das Immunsystem zu zerstören, wobei ART eine weitere Verbreitung des Virus verhindert. Eine spezifischere Alternative zur pharmazeutischen Aktivierung von HIV-1 stellen künstliche TALE-Aktivatoren dar, da diese nicht in andere zelluläre Prozesse eingreifen und somit auch auf nicht infizierte Zellen keinen Effekt haben sollten. Beim Test verschiedener künstlicher TALEs wurde festgestellt, dass deren Fähigkeit zur Transkriptionsinduktion stark von ihrer Position im HIV-1-Promotor (LTR) abhing. So erfolgte eine Reporteraktivierung am LTR nur, wenn der entsprechende TALE eine Sequenz stromaufwärts der drei SP1-Bindestellen ansteuerte.

4.2.1 Die Aktivierung von HIV-1 durch künstliche Transkriptionsfaktoren

In dieser Arbeit wurde mit Hilfe eines Reporterkonstruktes in HEK293T-Zellen gezeigt, dass TALEs die Expression am LTR-Promotor induzieren können, sodass die induzierte mRNA den natürlichen HIV-1 Transkriptionsstart besitzt (Abb. 3.2.1). Damit ist das vollständige TAR-Element enthalten, wodurch eine regulatorische Rückkopplung wie bei der natürlichen viralen RNA durch das virale Protein Tat möglich wäre. Kooperationspartner und andere haben darüber hinaus die Aktivierung eines Provirus in latent infizierten Zellen mit künstlichen ZFP-, TALE- oder Cas9-basierten Aktivatoren gezeigt (Wang *et al.*, 2014 und 2015; Geißler *et al.*, 2015; Perdigao *et al.*, 2016; Ji, H. *et al.*, 2016; Bialek *et al.*, 2016; Limsirichai *et al.*, 2016; Saayman *et al.*, 2016). Auch in einigen dieser Studien wurden die künstlichen Aktivatoren zunächst auf die Induktion eines Reporterkonstruktes mit HIV-1-Promotor in einer HEK293T-Zelllinie getestet. Wie in dieser Arbeit waren auch dabei sowohl TALE- als auch Cas9-Aktivatoren effizient, wenn sie Sequenzen im Bereich der beiden κ B-Motive oder unmittelbar stromaufwärts dieser erkannten (Abb. 3.2.2; Perdigao *et al.*, 2016; Ji, H. *et al.*, 2016; Saayman

et al., 2016). Bei der Induktion von Proviren oder endogener Reporter wurden allerdings neben Ähnlichkeiten auch Unterschiede bezüglich der geeigneten Positionierung künstlicher Aktivatoren im LTR in den verschiedenen Zellmodellen beobachtet. So waren die beiden κ B-Motive stromaufwärts der SP1-Bindestellen nicht immer für eine effiziente Aktivierung durch künstliche TALEs oder Cas9-Aktivatoren geeignet, wohingegen die weiter 5'-vorgelagerte Region eine maximale Induktion im jeweiligen Testsystem erlaubte (Geißler *et al.*, 2015; Bialek *et al.*, 2016). Dies könnte an einer guten Zugänglichkeit dieser Region liegen, die zum einen durch die Abwesenheit von Nukleosomen im aktiven LTR (Van Lint *et al.*, 1996) und zum anderen der geringen Dichte bekannter Transkriptionsfaktor-Bindestellen begründet ist. Die SP1-Bindestellen des LTR erschienen hingegen mit einer Ausnahme als Zielsequenzen für künstliche Aktivatoren ungeeignet. So wurden in Limsirichai *et al.* (2016) Cas9-Aktivatoren u. a. mit sgRNAs für Zielsequenzen eingesetzt, die mit den Bindestellen von NF κ B oder SP1 überlappen, und in den zwei verschiedenen klonalen Zelllinien J-Lat 9.2 und 10.6 getestet. J-Lat sind latent mit einem defekten HIV-1, der einen GFP-Reporter enthält, infizierte T-Lymphozyten der Jurkat Zelllinie (Jordan *et al.*, 2003). Während in J-Lat 9.2 die gegen die κ B-Motive gerichtete sgRNA die bei weitem stärkste und nahezu einzige GFP-Aktivierung von allen getesteten sgRNAs hervorrief, zeigte sich in J-Lat 10.6 ein vollkommen gegensätzliches Bild (Limsirichai *et al.*, 2016). Alle getesteten sgRNAs, inklusive der gegen die SP1-Boxen gerichteten, führten zu einer GFP-Aktivierung, wobei die gegen die κ B-Motive gerichtete sgRNA mit am schwächsten war. Für die J-Lat Zelllinien ist bekannt, dass ihre Aktivierbarkeit durch NF κ B stark von der Chromatin-Zugänglichkeit des proviralen LTR beeinflusst wird, die auch vom Integrationsort der Proviren abhängt (Miller-Jensen *et al.*, 2012). So ist die Aktivierung des HIV-Reporters durch sowohl eine TNF (Tumornekrosefaktor)- α vermittelte NF κ B-Antwort als auch einen Histon-Deacetylase-Inhibitor (HDACi) in J-Lat 10.6 viel stärker als in J-Lat 9.2 (Miller-Jensen *et al.*, 2012). Es könnte also angenommen werden, dass der HIV-1-Promotor in J-Lat 10.6 auch für andere Transkriptionsfaktoren, die im LTR binden können, sowie die Cas9-Aktivatoren aus Limsirichai *et al.* (2016) besser zugänglich ist als in J-Lat 9.2, wobei die NF κ B-Bindestellen womöglich besetzt sind. In Limsirichai *et al.* (2016) wurden jedoch keine sgRNAs für die Region unmittelbar stromaufwärts der κ B-Motive getestet, weshalb offen blieb, ob diese Region ein besseres Ziel für künstliche Aktivatoren im schwer induzierbaren Provirus von J-Lat 9.2 darstellt als die κ B-Motive selbst. In anderen Studien waren künstliche Aktivatoren, die die 5' der κ B-Motive gelegene Region ansteuern, allerdings in der Lage in schwer aktivierbaren J-Lat Klonen den proviralen Reporter zu induzieren

(Perdigao *et al.*, 2016; Bialek *et al.*, 2016). Daher ist diese Region als relativ „sicherer Hafen“ für die Aktivierung von latent HIV-1 infizierten Zellen anzunehmen.

Insgesamt scheint die Aktivierbarkeit von HIV-1-Proviren vor allem von der Zugänglichkeit des Chromatins und den vorhandenen Transkriptionsfaktoren in der individuellen Zelle abzuhängen. Dies sind auch die Stellschrauben, die durch Pharmazeutika im Rahmen der *shock and kill*-Strategie angesteuert werden. Beispiele für solche Präparate sind HDACi wie Vorinostat und Proteinkinase C-Aktivatoren wie Prostratin, die die Aktivierung zellulärer Transkriptionsfaktoren, die am LTR agieren wie z. B. NF κ B, fördern (Xing und Siliciano, 2013). Im Vergleich zu diesen Präparaten waren künstliche Aktivatoren mindestens genauso effizient (Ji, H. *et al.*, 2016; Saayman *et al.*, 2016) oder besser (Geißler *et al.*, 2015; Perdigao *et al.*, 2016; Bialek *et al.*, 2016; Limsirichai *et al.*, 2016) darin, die Expression von HIV-1 zu aktivieren. Im kombinierten Einsatz künstlicher Aktivatoren mit Pharmazeutika konnte darüber hinaus eine verstärkte Aktivierung proviraler Reporter beobachtet werden (Perdigao *et al.*, 2016; Limsirichai *et al.*, 2016). Bei der Verwendung von Cas9-Aktivatoren mit effizienten LTR-spezifischen sgRNAs in nichtinfizierten Zellen wurde jedoch keine Zytotoxizität oder T-Zellaktivierung, die bei der Anwendung von Pharmazeutika auftreten könnten, festgestellt (Ji, H. *et al.*, 2016; Saayman *et al.*, 2016). Die hohe Effizienz bei der Reaktivierung von HIV-1 in Kombination mit der hohen Spezifität ohne unerwünschte Nebeneffekte machen künstliche HIV-Aktivatoren zu einer attraktiven Alternative zur Realisierung der *shock and kill*-Strategie zur Heilung von HIV-1.

4.2.2 Strategien zur effizienten Anwendung künstlicher Transkriptionsaktivatoren

Die am LTR von HIV-1 beobachtete Positionsabhängigkeit der Transkriptionsinduktion durch künstliche Aktivatoren wurde auch bereits an anderen Promotoren beobachtet (Zhang *et al.*, 2011; Bultmann *et al.*, 2012; Renner *et al.*, 2014; Barbon *et al.*, 2016). Die gezielte Genaktivierung mit einem künstlichen Transkriptionsfaktor ist dementsprechend nicht uneingeschränkt möglich und bedarf oftmals der Optimierung. So wurde in dieser Arbeit und anderen Studien untersucht, wodurch die Geninduktion durch künstliche Aktivatoren beeinflusst wird und welche Strategien zur effizienten Genaktivierung führen können.

Bereits aus Vorarbeiten war bekannt, dass ein Luciferasereporter mit dem HIV-1-Promotor nicht effizient durch artifizielle TALEs, deren Zielsequenzen mit den SP1-Boxen überlappen (ArtHIV1 und 2), aktiviert werden kann. Dies konnte in dieser Arbeit auch für einen Cas9-Aktivator mit entsprechenden sgRNAs nachgewiesen werden (Abb. 3.2.2). Auch in anderen

Studien wurde festgestellt, dass einige TALE-Aktivatoren, deren Zielsequenzen mit bestimmten *cis*-Elementen überlappen, keine effiziente Aktivierung vermitteln können (Bultmann *et al.*, 2012; Geißler *et al.*, 2015; Barbon *et al.*, 2016). Dabei ist davon auszugehen, dass die künstlichen Aktivatoren mit entsprechenden *trans*-Elementen, wie z. B. SP1, um die Bindestellen konkurrieren. So waren bei Veränderung der SP1-Boxen im HIV-1-Promotor, sodass nur die artifiziellen TALEs aber voraussichtlich nicht Sp1 noch binden konnten (Abb. 3.2.3), die TALEs in der Lage den Luciferasereporter zu aktivieren (Abb. 3.2.4). Dabei war festzustellen, dass die Aktivierung des mutierten LTR geringer war als die Aktivität des intakten Promotors in An- oder Abwesenheit von ArtHIV1 oder 2. Das zeigt zum einen, welche hohe Basalaktivität SP1 dem HIV-1-Promotor im Reporterkonstrukt verleiht, und zum anderen, dass SP1 vermutlich die Bindung von ArtHIV1 und 2 blockiert. Interessanterweise war auch die Reporteraktivierung durch die stromaufwärts der SP1-Boxen bindenden TALE-Aktivatoren (ArtHIV6 und TALE5) durch die Mutation der SP1-Boxen auf unter das Basalniveau des Reporterkonstruktes mit dem originalen LTR reduziert, obwohl ArtHIV6 und TALE5 am originalen LTR effizient Luciferaseaktivität induzieren (Abb. 3.2.4). Auch in anderen Studien führte die Deletion oder Mutation benachbarter *cis*-Elemente zum Verlust bzw. zur Reduktion ursprünglich starker Aktivierung durch einen künstlichen TALE am jeweiligen Promotor (Bultmann *et al.*, 2012; Barbon *et al.*, 2016). Diese Beobachtungen könnten auf ein kooperatives oder sogar synergistisches Zusammenwirken der künstlichen Aktivatoren mit endogenen Transkriptionsfaktoren zurückzuführen sein. Dabei bedeutet Synergie, dass die Aktivierung bei Anwesenheit mehrerer Transkriptionsaktivatoren höher ausfällt als die Summe der Aktivitäten, die durch die einzelnen Transkriptionsaktivatoren vermittelt werden (Carey *et al.*, 1990). Da die sauren ADs von VP16 und p65 funktionale Ähnlichkeiten aufweisen (Blair *et al.*, 1994), könnten ArtHIV6 und TALE5 mit VP16 demnach ähnlich wie NF κ B durch eine kooperative Interaktion mit SP1 die Transkription am HIV-1-Promotor befördern (Perkins *et al.*, 1993; Burnett *et al.*, 2009). Eine mögliche Strategie, um mit künstlichen Aktivatoren effiziente Geninduktion zu erreichen, könnte also die geeignete Positionierung zu *cis*-Elementen sein, sodass eine synergistische Aktivierung zusammen mit endogenen Transkriptionsfaktoren ermöglicht wird (Abb. 4.2 A, B).

Um einen künstlichen Aktivator geeignet platzieren zu können, ist eine gute Kenntnis der Transkriptionsregulation am Promotor des Zielgenes oder vorausgehende Tests mehrerer Aktivatoren notwendig. Da dies nicht immer realisiert werden kann, ist eine andere Möglichkeit starke Effekte durch künstliche Aktivatoren zu erreichen, der Einsatz multipler ADs (Abb. 4.2 C). Dabei werden mehrfach die gleichen oder unterschiedliche ADs entweder direkt an das

DNA-bindende Protein fusioniert (VPR; Chavez *et al.*, 2014) oder über ein repetitives Epitop (SunTag; Tanenbaum *et al.*, 2014) oder im Falle von Cas9-Aktivatoren ein sgRNA-Aptamer (*synergistic activation mediator*, SAM; Konermann *et al.*, 2015) mit dem DNA-bindenden Protein assoziiert. Dadurch wird eine stärkere Aktivierung als durch einzelne ADs ermöglicht. VPR, SunTag, SAM, sowie eine Kombination von VPR und SAM wurden auch erfolgreich zur gesteigerten HIV-1-Aktivierung durch Cas9-Aktivatoren eingesetzt (Ji, H. *et al.*, 2016; Bialek *et al.*, 2016; Limsirichai *et al.*, 2016; Saayman *et al.*, 2016). In dieser Arbeit wurden ebenfalls multiple ADs an ArtHIV6 fusioniert (Abb. 3.2.5). Dabei wurde eine Kombination aus VP64-, p65- und HSF1-AD (VpH) gewählt, die bisher nur im SAM-System verwendet wurde (Konermann *et al.*, 2015). Im Vergleich mit den einzelnen ADs zeigte ArtHIV6 mit der neuen AD VpH am SP1-defizienten LTR und an einem minimalen CMV-Promotor eine mehr als additive Aktivierung (Abb. 3.2.5). Die an ein Protein fusionierten ADs können also gewissermaßen synergistisch zusammenwirken.

Eine weitere Möglichkeit um synergistische Aktivierung zu erreichen ist die Kombination mehrerer künstlicher Aktivatoren, die im weiteren Bereich des Genpromotors sowohl den codierenden als auch den codogenen DNA-Strang ansteuern können (Abb. 4.2 D; Richter und Boch, 2013). Diese Methode wurde sowohl für ZFP (Liu *et al.*, 2001) als auch TALE- (Maeder *et al.*, 2013 a; Perez-Pinera *et al.*, 2013 a) und Cas9-Aktivatoren (Maeder *et al.*, 2013 b; Perez-Pinera *et al.*, 2013 b) etabliert. Da Cas9 durch die Applikation mehrerer, kleiner sgRNAs leicht mehrere Zielsequenzen ansteuern kann, ohne dass mehrere Proteine in eine Zelle gebracht werden müssen (Le Cong *et al.*, 2013), erscheint diese Form der transkriptionellen Synergie für Cas9-Aktivatoren besonders geeignet. Für die HIV-1-Aktivierung wurden sowohl mehrere sgRNAs mit einem Cas9-Aktivator als auch mehrere TALE-Aktivatoren eingesetzt (Perdigao *et al.*, 2016; Bialek *et al.*, 2016). In beiden Fällen konnte eine leicht synergistische Aktivierung beobachtet werden. In dieser Arbeit war mit den verwendeten HIV-TALEs keine Synergie am originalen HIV-1-Promotor nachweisbar (Abb. 3.2.4), wobei ein deutlich kleinerer Promotorbereich als in den anderen Studien abgedeckt wurde (Perdigao *et al.*, 2016; Bialek *et al.*, 2016). Im Unterschied dazu war am SP1-defizienten LTR starke synergistische Aktivierung mit mehreren TALE-Kombinationen, insbesondere mit TALE5, ArtHIV1 und 5, nachweisbar (Abb. 3.2.4). Dies legt nahe, dass diese Form der synergistischen Aktivierung vor allen für die Induktion von Genen mit geringer Basalaktivität geeignet ist. Darüber hinaus wurde gezeigt, dass die Verwendung mehrerer künstlicher TALE-Aktivatoren unabhängig von der Chromatin-Zugänglichkeit der einzelnen Zielsequenzen zur effizienten Aktivierung führen kann (Perez-

Pinera *et al.*, 2013 a). Dies könnte eine einfache Induktion ohne Vorwissen über die Regulation eines Gens in der jeweiligen Zelle erlauben.

Neben den verbreiteten transkriptionellen ADs wie VP16 und p65, die in der Regel von Transkriptionsfaktoren stammen, können auch Domänen an DNA-Bindeproteine fusioniert werden, die durch direkte Beeinflussung der Chromatin-Zugänglichkeit Transkription aktivieren. So wurden durch Fusion der katalytischen Kerndomäne von p300 an ZFP, TALEs oder Cas9N0 spezifische Histon-Acetyltransferasen geschaffen, die zur effizienten Genaktivierung führten (Hilton *et al.*, 2015). Cas9-basierte Histon-Acetyltransferasen wurden auch erfolgreich zur HIV-Induktion eingesetzt (Limsirichai *et al.*, 2016) und stellen eine interessante Alternative, insbesondere im Hinblick auf womöglich epigenetisch stillgelegte Proviren in latenten Zellreservoirs, dar.

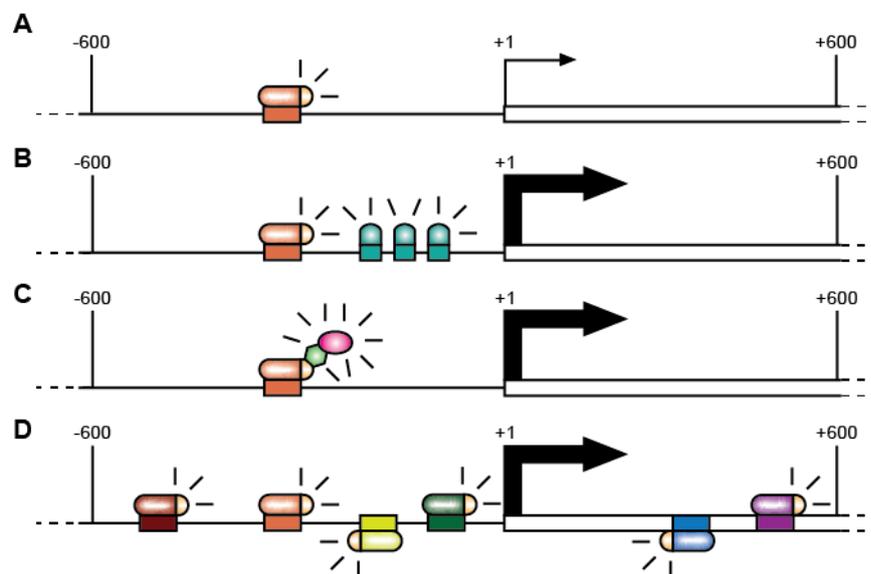


Abb. 4.2: Verschiedene Strategien zur Genaktivierung mit künstlichen Aktivatoren. Ein artifizielles Aktivator (orange) kann ein Zielgen (A) nicht oder nur geringfügig oder (B) durch synergistisches Zusammenwirken mit endogenen Transkriptionsfaktoren (türkis) effizient aktivieren. Effiziente Aktivierung kann auch durch (C) multiple Aktivierungsdomänen (hellgrün, pink) oder (D) synergistisches Zusammenwirken mit anderen künstlichen Aktivatoren (rot, gelb, grün, blau, violett) für dasselbe Zielgen erreicht werden. Abb. verändert nach Richter und Boch, 2013.

4.2.3 Die Aktivierung verschiedener HIV-1-Varianten

Auf Grund der hohen retroviralen Mutationsrate tritt HIV-1 mit einer großen Variabilität auf. Dies wird vor allem in der starken Diversität der Virus-Subtypen deutlich (de Baar *et al.*, 2000; Taylor *et al.*, 2008; de Arellano *et al.*, 2010). Selbst innerhalb eines einzelnen Patienten sind zahlreiche verschiedene provirale Klone zu finden (Geißler *et al.*, 2015). Sogar im Bereich der wichtigsten regulatorischen Promotorelemente des LTR, den κ B-Motiven und SP1-Boxen,

lassen sich Veränderungen mit Einfluss auf die virale Replikation in einzelnen Viruspopulationen finden. So führen z. B. kleine Mutationen in den SP1-Bindestellen chinesischer HIV-1 Isolate zu einer verstärkten viralen Transkription (Qu *et al.*, 2016) und die evolutive Umwandlung eines κ B-Motives zur Bindestelle eines anderen Transkriptionsfaktors zur verbesserten Fitness des Virus in Zellen mit wenig NF κ B (Verhoef *et al.*, 1999). Vor diesem Hintergrund müssen auch künstliche Aktivatoren für die *shock and kill*-Strategie betrachtet werden. In Bialek *et al.* (2016) wurde daher die Sequenz der effektivsten sgRNAs mit verfügbaren Sequenzen verschiedener Subtypen verglichen, wobei sich eine geringe Konservierung der Bindestellen herausstellte. In anderen Studien wurden zur Aktivierung einzelner Subtypen oder von Patientenisolaten individuelle sgRNAs oder TALE-Aktivatoren für die gleiche Position im Promotor eingesetzt (Geißler *et al.*, 2015; Saayman *et al.*, 2016). In Geißler *et al.* (2015) wurde dabei ein TALE für die Konsensus-Sequenz der im Patienten auftretenden proviralen Klone mit den üblichen *repeats* (NN, HD, NI, NG) erstellt, der das Virus in den jeweiligen Patientenzellen aktivieren konnte. Ob jedoch alle proviralen Klone des Patienten mit dem individuell erstellten TALE aktivierbar sind, wurde nicht geklärt. Das wäre jedoch im Hinblick auf die *shock and kill*-Strategie sehr interessant, da für deren Umsetzung das Virus in allen latent infizierten Zellen aktiviert werden müsste. In einem der in Geißler *et al.* (2015) untersuchten Patienten wurde auch ein proviraler Klon mit Indels in der Zielsequenz des verwendeten TALEs gefunden. Solche Sequenzen können i. d. R. nicht von herkömmlichen TALEs erkannt werden.

In dieser Arbeit wurden unter Verwendung von RVDs mit multipler Spezifität (NN, NS, N*) und eines aberranten *repeats* ArtHIV6- und TALE5-Varianten (ArtABC6 und TALEABC5) erstellt, die in der Lage waren, gleichzeitig an Promotorsequenzen der HIV-1 Subtypen A bis F Transkription zu aktivieren (Abb. 3.2.6). Dabei wurden auch kleine Indels überwunden. Die Möglichkeit derart flexible HIV-Aktivatoren zu erstellen ist nur mit TALEs und nicht mit ZFPs oder Cas9-Aktivatoren gegeben. Ein optimierter HIV-TALE könnte den Bereich unmittelbar stromaufwärts der κ B-Motive ansteuern und so wie TALEABC5 aus dieser Arbeit für ein Spektrum möglicher Zielsequenzen angepasst sein. Darüber hinaus könnte die Effizienz der Aktivierung noch durch Fusion multipler ADs oder eine Kombination mehrerer flexibler TALEs erhöht werden.

4.3 Betrachtungen zur nicht nachweisbaren Aktivität von scTALEN *in vivo*

TALEN werden üblicherweise als Paare eingesetzt, wobei jedes Einzelmolekül aus einem TALE-Grundgerüst zur Sequenzerkennung und einer FokI-Endonukleasedomäne ohne eigene DNA-Spezifität besteht. Dies hat Nachteile gegenüber monomeren Genomeditierungs-Werkzeugen wie z. B. CRISPR/Cas9. So sind paarige Nukleasen im Vergleich zu entsprechenden monomolekularen Nukleasen zum einen in der Erstellung doppelt so aufwendig und teuer und zum anderen auf Grund der längeren DNA-Zielsequenz vermutlich stärker von der Chromatin-Zugänglichkeit abhängig. Daher sind monomere TALEN-Varianten für die biotechnologische Anwendung von generellem Interesse. Eine Möglichkeit, monomere TALEN zu erstellen, war dabei die Fusion anderer Nukleasedomänen als der von FokI. So wurden monomere TALEN mit Hilfe eines fusionierten PvuII-Dimers (Yanik *et al.*, 2013) oder der katalytisch aktiven Domäne von I-TevI erstellt (Beurdeley *et al.*, 2013; Kleinstiver *et al.*, 2014). Diese Varianten hatten jedoch den Nachteil, dass durch die DNA-Spezifität der verwendeten Nukleasen zusätzliche Anforderungen an die Zielsequenz anfielen. Eine Alternative ohne eigene Sequenzspezifität stellte eine scFokI-Endonukleasedomäne dar, wie sie bereits für monomere ZFN verwendet wurde (Minczuk *et al.*, 2008; Mino *et al.*, 2009). So wurden scTALEN von Sun und Zhao (2014) entwickelt, die im Vergleich zu TALEN-Paaren jedoch eine deutlich geringere Mutationseffizienz in humanen Zellen aufwiesen.

Auch in dieser Arbeit wurden verschiedene scTALEN-Varianten erstellt, von denen alle DNA spezifisch und effizient *in vitro* schnitten (Abb. 3.3.2, 3.3.4, 3.3.5). Beim Test der mit *Sharkey*-FokI-Endonukleasedomänen optimierten scTALEN-Varianten in *D. rerio*-Embryonen (Alexander Hruscha, persönliche Kommunikation) und humanen Zellen (Abb. 3.3.6) wurde jedoch keine Aktivität in Form von Indel-Mutationen *in vivo* nachgewiesen. Eine mögliche Erklärung für die hier fehlende und in Sun und Zhao (2014) geringe Mutationseffizienz wäre, dass scTALEN immer wieder die DNA binden und schneiden können, weil ihre Zielsequenz nur mit geringer Effizienz durch die zellulären Reparaturmechanismen zerstört wird. Während bei TALEN-Paaren NHEJ vermutlich über die Zeit die *spacer*-Länge so stark verändert, dass das Paar nicht mehr schneiden kann, müssten sich die Mutationen bei scTALEN bis in die Zielsequenz erstrecken. Ein immerzu wiederholtes Schneiden der DNA könnte zum einen zum Tod der mit scTALEN-transfizierten Zelle führen oder verhindern, dass sich in der Inkubationszeit (72 h) eine nachweisbare Mutation etabliert, da der DSB per se nicht mit PCR-basierten Methoden nachweisbar ist. Insbesondere bei der konstanten Expression der scTALEN von Plasmiden (Abb. 3.3.6; Sun und Zhao, 2014) könnte ein solcher Effekt eine Rolle spielen,

da ein Endzustand auch nicht durch Abwesenheit der Nuklease erreicht werden kann. Im Unterschied dazu wurden die Zebrafischzygoten mit scTALEN-mRNA injiziert (Alexander Hruscha, persönliche Kommunikation), die mit jeder Zellteilung des Embryos auf die Tochterzellen verteilt und somit ausgedünnt wird. Darüber hinaus unterliegen sowohl die mRNA als auch das Protein dem Abbau, sodass im Laufe der Zeit keine scTALEN-Moleküle im Zebrafischembryo nachproduziert werden, während die Vorhandenen abnehmen. Daher sollte sich in diesem System zumindest auf Grund der Abwesenheit der scTALEN eine Mutation etablieren können. Ob dies bereits nach dem gewählten Untersuchungszeitraum von 24 h der Fall wäre, müsste z. B. mittels Immunoblot überprüft werden. Da der Embryo zu diesem Zeitpunkt jedoch bereits recht weit entwickelt ist (Segmentierung abgeschlossen, Schwanz und Organe angelegt, Herzschlag, rote Blutzellen, Pigmentierung der Retina nach Kimmel *et al.*, 1995), kann angenommen werden, dass die scTALEN, deren mRNA im Einzellstadium injiziert wurde, in den meisten Zellen abgebaut wurde. Dennoch war auch in *D. rerio* kein eindeutiger Nachweis der scTALEN-Aktivität *in vivo* möglich. Um zu untersuchen, ob kontinuierlich wiederholtes Schneiden der scTALEN wirklich der Grund für die nicht nachweisbare Aktivität *in vivo* ist und um diese potentielle Ursache zu umgehen, wurden auch dscTALEN erstellt, die ein wiederholtes Schneiden durch Ausschneiden ihrer eigenen Zielsequenz verhindern sollten (Abb. 3.3.4, 3.3.5, 3.3.6). Obwohl gezeigt werden konnte, dass sowohl die N- als auch die C-terminale scFokI-Endonukleasedomäne der dscTALEN DNA schneidet (Abb. 3.3.4), bleibt offen, ob dies auch zeitgleich stattfinden kann, was für ein Ausschneiden der Zielsequenz *in vivo* nötig wäre. Darüber hinaus könnte die fehlende Aktivität der dscTALEN in beiden *in vivo*-Systemen, *D. rerio* und humane Zellen (Alexander Hruscha, persönliche Kommunikation; Abb. 3.3.6), auch auf die geringe bis fehlende Kernlokalisierung der dscTALEN zurückzuführen sein (Abb. 3.3.7).

Eine andere mögliche Erklärung für die nicht nachweisbare *in vivo*-Aktivität der scTALEN-Varianten könnte sein, dass scTALEN womöglich gar keinen DSB auslösen und somit nicht zu Indel-Mutationen in Zellen führen. So wurde in dieser Arbeit in Restriktionsanalysen gezeigt, dass TALEN und scTALEN mit einem aktiven und einem inaktiven FokI-Monomer DSB erzeugen können (Abb. 3.3.3), obwohl sie nur einen DNA-Strang schneiden sollten. Es wurde angenommen, dass der gewählte TALE-C-Terminus eventuell genügend Flexibilität erlaubt, um der aktiven FokI-Endonukleasedomäne das Schneiden beider DNA-Stränge zu ermöglichen. Dies würde auch bedeuten, dass mit dem in dieser Arbeit gewählten Aufbau scFokI-Nukleasen nicht von scFokI-Nickasen im Restriktionsverdau zu unterscheiden sind. Demnach könnten im Umkehrschluss alle erstellten scTALEN-Varianten womöglich scTALE-

Nickasen sein. Falls die scTALEN tatsächlich scTALE-Nickasen wären, könnte es sein, dass ihre Aktivität mit den gewählten Methoden (PCR, Restriktions- oder T7EI-Verdau) *in vivo* nicht nachzuweisen ist, weil die Zellen einen Einzelstrangbruch (ESB) nicht durch NHEJ sondern durch Basenexzision reparieren (de Laat *et al.*, 1999), ohne dass Mutationen erzeugt werden. Dieser Prozess würde sich dann immer wiederholen. Im Unterschied dazu könnte eine ESBs erzeugende scTALEN *in vitro* erneut binden und den zweiten noch intakten Strang schneiden, wodurch der DSB gewissermaßen in zwei Schritten erzeugt werden würde. Dass mit einem *linker* zu einer scFokI verknüpfte Endonukleasedomänen tatsächlich nicht nur DSBs sondern auch ESBs verursachen können, wurde bereits *in vitro* für bestimmte scZFN gezeigt (Mino *et al.*, 2009). Dabei war der Anteil an ESBs und DSBs von der *linker*-Länge abhängig. Bei einem 35 AS-*linker*, der dem in dieser Arbeit verwendeten ähnelte, war das Verhältnis von DSB zu ESB ca. eins zu zwei. Ein 60 AS-*linker* erlaubte hingegen die eingesetzte DNA vollständig zu schneiden. In einer geringeren Effizienz DSBs zu erzeugen, könnte also die Ursache dafür liegen, dass Sun und Zhao (2014) zumindest eine geringe Aktivität ihrer scTALEN mit einem 95 AS-*linker* in humanen Zellen gesehen haben, während in dieser Arbeit keine nachweisbar war. In dem scZFN-Experiment war die scFokI jedoch nur durch 15 AS mit dem ZFP verbunden (Mino *et al.*, 2009), was ihr weniger Flexibilität erlauben sollte als der 63 AS TALE-C-Terminus der scTALEN. Dies könnte erklären, warum für die scZFN nicht wie in dieser Arbeit nur DSBs sondern auch ESBs *in vitro* beobachtet wurden (Mino *et al.*, 2009). Dennoch ist anhand dieser Ergebnisse anzunehmen, dass die in dieser Arbeit erstellten scTALEN auch DSBs erzeugen können. Darüber hinaus ist nicht auszuschließen, dass der hier postulierte DSB durch scTALE-Nickasen in zwei Schritten nicht auch erfolgt, ohne dass sich die scTALE-Nickase von der DNA löst, und die scFokI einfach nur flexibel den DNA-Strang wechselt, was dann auch einen DSB *in vivo* erlauben sollte.

Eine weitere mögliche Erklärung für die fehlende *in vivo*-Aktivität der in dieser Arbeit entwickelten scTALEN könnte mit ihrer Löslichkeit in der Zelle zusammenhängen. So konnte für keine der getesteten scTALEN-Varianten eine vollständige Kernlokalisierung nachgewiesen werden (Abb. 3.3.7). Darüber hinaus zeigten die Lokalisierungsstudien auch, dass die scTALEN sowohl im Zytoplasma als auch im Nukleus nicht gleichmäßig verteilt vorlagen, was auf die mögliche Bildung von Proteinaggregaten schließen lässt. In Ren *et al.* (2017) wurden TALE-Derivate als Sonden zur Sichtbarmachung repetitiver DNA-Abschnitte wie Telomere oder Zentromere in lebenden Zellen angewendet. Dabei wurde ebenfalls beobachtet, dass die TALE-Fusionsproteine Aggregate im Nukleus bildeten und nicht mit ihren Zielsequenzen co-lokalisierten. TALE-Derivate scheinen also dazu zu tendieren, in lebenden

Zellen Aggregate zu bilden, was vermutlich durch eine starke Expression wie durch die in dieser Arbeit gewählten Konstrukte begünstigt wird. Neben einer möglichen Aggregation von TALE-Derivaten in Zellen könnte auch die Fusion von FokI-Endonukleasedomänen zu einer schlechten zellulären Löslichkeit von scTALEN beitragen. Ein möglicher Ansatz zur Erhöhung der Löslichkeit von TALEN-Varianten könnte die Fusion eines Löslichkeit-erhöhenden Peptides sein. In Ren *et al.* (2017) wurden verschiedene solcher Peptide zur Optimierung der subnuklearen Lokalisierung von TALE-Derivaten getestet, wobei nur die Fusion von Thioredoxin zur korrekten Lokalisierung der TALE-Derivate führte. Daher wäre ein möglicher Ansatz scTALEN so zu optimieren, dass sie in lebenden Zellen Mutationen verursachen können, die Fusion von Löslichkeit-erhöhenden Peptiden wie Thioredoxin. Die so optimierten scTALEN müssten dann *in vivo* sowohl auf die Erzeugung von DSBs als auch ESBs untersucht werden, um festzustellen, ob sie möglicherweise als Nickasen agieren. Neben der Untersuchung auf Indels durch NHEJ, müsste also auch ein Ansatz zur Feststellung von HDR als ein Reparaturmechanismus von ESBs gewählt werden. Da für TALE-Nickase-Paare eine bessere HDR-Effizienz im Vergleich zu TALEN-Paaren erreicht wurde (Wu *et al.*, 2014), könnten scTALE-Nickasen nützliche neue Werkzeuge für die Genomeditierung darstellen.

4.4 Die Erstellung hochspezifischer TALEN

Für die gentherapeutische Anwendung ist es notwendig, programmierbare Nukleasen so spezifisch wie möglich zu gestalten, um *off-target*-Effekte zu vermeiden oder idealerweise auszuschließen. Dazu wurde in dieser Arbeit beispielhaft ein defektes *FANCI*-Allel (NM_001113378.1(*FANCI*):c.1461T>A (p.Tyr487Ter)), das rezessiv zur Ausprägung der Fanconi-Anämie führt, mit TALEN angesteuert (Abb. 3.3.8). Da defekte *FANCI*-Allele relativ selten sind, treten sie häufig in gemischt heterozygoter Form auf. Das hier gewählte Allel, das sich nur durch einen SNP vom WT unterscheidet, stammt z. B. aus einem Patienten, bei dem der Defekt im zweiten Allel über 20000 bp entfernt liegt. Das eröffnet die theoretische Möglichkeit eines eleganten Therapieansatzes, bei dem spezifisch nur das lokal defekte Allel zerstört wird und die Zelle das zweite Allel, das in diesem Bereich eine wildtyp-Sequenz aufweist, als homologe Reparaturvorlage verwenden könnte. Dadurch würde eines der beiden defekten Allele repariert, was zur Heilung der rezessiven Erbkrankheit ausreichend wäre. In einer anderen Studie wurden bereits Fibroblasten des Patienten mit dem hier adressierten Allel zu induzierten pluripotenten Stammzellen (iPSC) reprogrammiert, das defekte Allel repariert und die geheilten iPSCs zu Zellen hämatopoetischen Charakters differenziert (Osborn *et al.*, 2016). Die homologe Reparatur des defekten Allels erfolgte dabei mit nicht Allel-spezifischen

Cas9-Nickasen und einer Reparaturvorlage mit einer Marker-Kassette, die in einem zweiten Schritt entfernt werden musste. Dabei bleiben Veränderungen in *FANCI* zurück, die mit dem hier vorgeschlagenen Ansatz nicht auftreten würden. Die Erstellung geheilter, potentiell therapeutisch einsetzbarer Zellen in Osborn *et al.* (2016) zeigt, dass solche Ansätze realistische Möglichkeiten zur Behandlung von Krankheiten darstellen.

Die in dieser Arbeit verwendeten TALEN, die nur die defekte, aber nicht die Wildtyp-*FANCI*-Sequenz schneiden sollten, wurden zu diesem Zweck nach verschiedenen Kriterien modifiziert um möglichst spezifisch zwischen beiden Allelen unterscheiden zu können. Dabei wurde die Länge der *repeat*-Arrays verändert, die Bindestärke der TALEN über das Verhältnis von schwachen zu starken RVDs variiert und zusätzliche RBF in die über dem SNP bindende TALEN eingeführt (Abb. 3.3.8, 3.3.9, 3.3.10). Auf die Differenzierung der beiden Sequenzen anhand des 1 bp-Unterschiedes hatten in Restriktionsanalysen *in vitro* vor allem die Veränderung der *repeat*-Anzahl und die zusätzlichen RBF im SNP-bindenden TALEN einen Einfluss, während die Wirkung starker oder schwacher RVDs vernachlässigbar war (Abb. 3.3.9, 3.3.10). Es war zu beobachten, dass die Aktivität kürzerer TALEN stärker durch den SNP und die zusätzlichen RBF betroffen war als die Aktivität längerer TALEN, wobei auch deren Sensitivität für den SNP durch zusätzliche RBF wuchs. Diese geringere Fehlpaarungstoleranz kürzerer *repeat*-Arrays wurde bereits für TALEs und TALEN gezeigt (Mali *et al.*, 2013 b; Juillerat *et al.*, 2014).

Einen starken Einfluss auf die Restriktionsaktivität der TALEN hatte darüber hinaus die Position des SNPs (vgl. Abb. 3.3.9 und 3.3.10). Wurde der SNP als initiales T verwendet, führte das bereits ohne zusätzliche RBF und unabhängig von der *repeat*-Anzahl zu einem starken Aktivitätsunterschied der TALEN zwischen der mutierten und wildtypischen *FANCI*-Sequenz, in der sich ein Adenin an dieser Position befindet (Abb. 3.3.9). Bei einer SNP-bindenden TALEN mit nur 11,5 *repeats* reichte diese Fehlpaarung aus, um die Aktivität an der Wildtypsequenz bereits auf unter 10% zu senken (Abb. 3.3.9). Mit zusätzlichen RBF war es sogar möglich, die Restriktionsaktivität der TALEN aller Längen an der Wildtypsequenz auf 1% oder weniger zu senken, wobei eine Aktivität von ca. 20% oder mehr an der mutierten Sequenz erhalten blieb (Abb. 3.3.9). Es ist bekannt, dass die Position T₀ einen größeren Einfluss auf die Gesamtbindung von TALEs besitzt, als einzelne RVDs (Boch *et al.*, 2009; Schreiber *et al.*, 2014). Dementsprechend wurden TALEN-Varianten, die Sequenzen anhand einer Fehlpaarung an Position 0 diskriminieren, auch in anderen Studien erfolgreich eingesetzt (Bacman *et al.*, 2013 und 2018; Pereira *et al.*, 2018). Sie wurden dabei zur Depletion pathogener

mitochondrialer DNA (mtDNA), die zusammen mit wildtypischer mtDNA im selben Organell auftritt, verwendet.

Im Vergleich zur Platzierung des SNP an Position T_0 hatte der SNP an zweiter Position per se nur einen geringen und mit zunehmender *repeat*-Anzahl gar keinen Einfluss auf die Restriktionsaktivität der TALEN (Abb. 3.3.10). Für eine SNP-bindende TALEN mit 11,5 *repeats* konnte jedoch durch zwei zusätzliche RBF ein Aktivitätsunterschied zwischen mutierter und WT-Sequenz von ca. 70% zu nur ca. 4% erreicht werden (Abb. 3.3.10). Obwohl die verbleibende Restriktionsaktivität an der WT-Sequenz höher war als bei Positionierung des SNP an T_0 , war dieser Aktivitätsunterschied zwischen den beiden *FANCI*-Allelsequenzen der Höchste der beobachtet wurde. Da die TALEN-Konzentration bei zellulärer Anwendung vermutlich geringer ist als im Restriktionsverdau, kann angenommen werden, dass dieser beobachtete Unterschied für die Diskriminierung des wildtypischen *FANCI*-Allels *in vivo* ausreicht, ohne dass die Aktivität am mutierten Allel darunter leidet. Im Unterschied dazu führten die drei zusätzlichen RBF in TALEN mit 17,5 *repeats* und dem SNP an zweiter Position nur zu Aktivitätsunterschieden von ca. 20% zwischen mutierter und WT-Sequenz, wobei die verbleibende Restriktionsaktivität an der WT-Sequenz immer noch 60 bis 70% betrug (Abb. 3.3.10). Dass die vier Fehlpaarungen relativ gut toleriert wurden, könnte an den für die zusätzlichen RBF gewählten RVD-Base-Kombinationen liegen. So ist bekannt, dass die Spezifität von TALEs eher durch negative Diskriminierung von Nukleobasen als durch positive Interaktion mit der passenden Base durch die RVD-AS 13 vermittelt wird (Wicky *et al.*, 2013). Dementsprechend hat jeder individuelle RVD ein gewisses Toleranzspektrum gegenüber den nicht „aktiv“ spezifizierten DNA-Basen. Für die den SNP an zweiter Position bindenden TALEN wurden die zusätzlichen RBF HD zu Adenin und NG zu Cytosin gewählt (Abb. 3.3.10), von denen gezeigt wurde, dass sie in TALEN relativ gut tolerierte Fehlpaarungen sind (Juillerat *et al.*, 2014). Möglicherweise hätten Kombinationen von RVDs mit von ihnen stärker diskriminierten Basen, wie z. B. NI mit Thymin oder HD mit Guanin (Juillerat *et al.*, 2014), zu einer besseren Unterscheidung der beiden *FANCI*-Allele auch durch längere TALEN geführt. Vor diesem Hintergrund stellen alternative RVDs eine weitere Möglichkeit dar, die Stringenz von TALEN gegenüber Fehlpaarungen zu erhöhen. So wurden in Yang *et al.* (2014) alle durch Kombination der 20 proteinogenen AS möglichen RVDs auf ihre Basenspezifität getestet und einige neue RVDs identifiziert. Solche wurden erfolgreich eingesetzt um die Spezifität von TALEN durch Austausch aller (Miller *et al.*, 2015) oder einiger RVDs an bestimmten Positionen (Juillerat *et al.*, 2015; Gautron *et al.*, 2017) zu erhöhen und *off-target*-Effekte zu reduzieren, was insbesondere im Hinblick auf eine mögliche therapeutische Anwendung von

TALEN interessant ist. Ob TALEN allerdings mit Hilfe eines alternativen RVDs, der eine geringe Fehlpaarungstoleranz aufweist, zwei Sequenzen anhand eines SNP differenziert schneiden können, ist bisher nicht getestet worden. Es ist jedoch anzunehmen, dass eine spezifische Unterscheidung zweier Sequenzen anhand eines 1 bp-Unterschiedes mit TALEN nicht in jedem Falle ohne zusätzliche RBF möglich ist. Bei dieser Herangehensweise muss berücksichtigt werden, dass durch die zusätzlichen RBF neue *off-targets* kreiert werden können. Generell sollten daher möglichst spezifische natürliche oder alternative RVDs verwendet werden. So wurde z. B. auch in dieser Arbeit NH (erkennt G) statt NN (erkennt G und A) zur Bindung von Guanin eingesetzt (Abb. 3.3.8, 3.3.9, 3.3.10). Darüber hinaus ist es mit geeigneten Algorithmen bioinformatisch möglich *repeat*-Arrays so zu gestalten, dass sie möglichst wenige *off-targets* im Genom aufweisen (Garg *et al.*, 2012).

Die in dieser Arbeit vorgestellte Differenzierung zweier Sequenzen anhand eines 1 bp-Unterschiedes durch TALEN ist nur bedingt mit dem CRISPR/Cas-System umsetzbar. So wurde gezeigt, dass Cas9 i. d. R. sogar mehr als eine Fehlpaarung zwischen der sgRNA und der Zielsequenz toleriert (Mali *et al.*, 2013 b; Hsu *et al.*, 2013). Eine Ausnahme stellt dabei die PAM-Sequenz dar, die entsprechend bereits für Anwendungen mit 1 bp Auflösung ausgenutzt wurde (Li *et al.*, 2016). Allerdings ist dieser Ansatz auf Grund der minimalen PAM-Sequenz NRG von Cas9 aus *S. pyogenes* (Mali *et al.*, 2013 b) nur auf einen Teil aller möglichen SNPs anwendbar, so z. B. nicht auf das in dieser Arbeit adressierte *FANCI*-Allel. Darüber hinaus haben auch optimierte Cas9-Varianten, die deutlich weniger *off-targets* erkennen, noch keine Einzelnukleotidspezifität außerhalb der PAM erreicht (Kleinstiver *et al.*, 2016; Slaymaker *et al.*, 2016; Nelson und Gersbach, 2016).

5. Literatur

Van Den Ackerveken, G., Marois, E. and Bonas, U. (1996) 'Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell', *Cell*. doi: 10.1016/S0092-8674(00)81825-5.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. 2004. *Molekularbiologie der Zelle*, Vierte Auflage

Ann Ran, F. *et al.* (2013) 'Genome engineering using the CRIPR-Cas9 system', *Nature*. doi: 10.1038/nprot.2013.143.

Antony, G. *et al.* (2010) 'Rice xa13 Recessive Resistance to Bacterial Blight Is Defeated by Induction of the Disease Susceptibility Gene Os-11N3', *the Plant Cell Online*, 22(11), pp. 3864–3876. doi: 10.1105/tpc.110.078964.

De Arellano, E. R. *et al.* (2010) 'Drastic decrease of transcription activity due to hypermutated long terminal repeat (LTR) region in different HIV-1 subtypes and recombinants', *Antiviral Research*. doi: 10.1016/j.antiviral.2010.08.007.

De Baar, M. P. *et al.* (2000) 'Subtype-specific sequence variation of the HIV type 1 long terminal repeat and primer-binding site', *AIDS Res Hum Retroviruses*. 16(5):499-504.

Bacman, S. R. *et al.* (2013) 'Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs', *Nature Medicine*, 19(9), pp. 1111–1113. doi: 10.1038/nm.3261.

Bacman, S. R. *et al.* (2018) 'MitoTALEN reduces mutant mtDNA load and restores tRNA^{Ala} levels in a mouse model of heteroplasmic mtDNA mutation', *Nature Medicine*, 24(11), pp. 1696–1700. doi: 10.1038/s41591-018-0166-8.

Baker, M. (2012) 'Gene-editing nucleases', *Nature Methods*. Nature Publishing Group, 9(1), pp. 23–26. doi: 10.1038/nmeth.1807.

Barbon, E. *et al.* (2016) 'An engineered tale-transcription factor rescues transcription of factor VII impaired by promoter mutations and enhances its endogenous expression in hepatocytes', *Scientific Reports*. doi: 10.1038/srep28304.

Becker, S. and Boch, J. (2016) 'Single-molecule biophysics: TALEs spin along, but not around', *Nature Chemical Biology*. doi: 10.1038/nchembio.2182.

Beurdeley, M. *et al.* (2013) 'Compact designer TALENs for efficient genome engineering', *Nature Communications*. doi: 10.1038/ncomms2782.

Bialek, J. K. *et al.* (2016) 'Targeted HIV-1 latency reversal using CRISPR/Cas9-derived transcriptional activator systems', *PLoS ONE*, 11(6), pp. 1–19. doi: 10.1371/journal.pone.0158294.

Bitinaite, J. *et al.* (1998) 'FokI dimerization is required for DNA cleavage', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.95.18.10570.

Blair, W. S. *et al.* (1994) 'Mutational analysis of the transcription activation domain of RelA: identification of a highly synergistic minimal acidic activation module.', *Molecular and Cellular Biology*, 14(11), pp. 7226–7234. doi: 10.1128/mcb.14.11.7226.

- Blount, B. A. *et al.* (2012) 'Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology', *PLoS ONE*, 7(3), pp. 1–11. doi: 10.1371/journal.pone.0033279.
- Boch, J. *et al.* (2009) 'Breaking the code of DNA binding specificity of TAL-type III effectors', *Science*. 11;326(5959):1509-12. doi: 10.1126/science.1178811.
- Boch, J. and Bonas, U. (2010) 'Xanthomonas AvrBs3 Family-Type III Effectors: Discovery and Function', *Annual Review of Phytopathology*. 48:419-36. doi: 10.1146/annurev-phyto-080508-081936.
- Boch, J., Bonas, U. and Lahaye, T. (2014) 'TAL effectors - pathogen strategies and plant resistance engineering', *New Phytologist*. doi: 10.1111/nph.13015.
- Bonas, U., Stall, R. E. and Staskawicz, B. (1989) 'Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*', *MGG Molecular & General Genetics*. doi: 10.1007/BF00330575.
- Bultmann, S. *et al.* (2012) 'Targeted transcriptional activation of silent *oct4* pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers', *Nucleic Acids Research*. doi: 10.1093/nar/gks199.
- Burnett, J. C. *et al.* (2009) 'Control of stochastic gene expression by host factors at the HIV promoter', *PLoS Pathogens*, 5(1). doi: 10.1371/journal.ppat.1000260.
- Carey, M. *et al.* (1990) 'A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives', *Nature*. doi: 10.1038/345361a0.
- Carlson, D. F. *et al.* (2012) 'Efficient TALEN-mediated gene knockout in livestock', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1211446109.
- Carpenter, A. E. *et al.* (2005) 'Common Effects of Acidic Activators on Large-Scale Chromatin Structure and Transcription', *Molecular and Cellular Biology*, 25(3), pp. 958–968. doi: 10.1128/mcb.25.3.958-968.2005.
- Cermak, T. *et al.* (2011) 'Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting', *Nucleic Acids Research*. doi: 10.1093/nar/gkr218.
- Cernadas, R. A. *et al.* (2014) 'Code-Assisted Discovery of TAL Effector Targets in Bacterial Leaf Streak of Rice Reveals Contrast with Bacterial Blight and a Novel Susceptibility Gene', *PLoS Pathogens*. doi: 10.1371/journal.ppat.1003972.
- Chavez, A. *et al.* (2015) 'Highly efficient Cas9-mediated transcriptional programming', *Nature Methods*. doi: 10.1038/nmeth.3312.
- Chi, T. *et al.* (1995) 'A general mechanism for transcriptional synergy by eukaryotic activators', *Nature*. doi: 10.1038/377254a0.
- Choudhary, S. K. and Margolis, D. M. (2011) 'Curing HIV: Pharmacologic Approaches to Target HIV-1 Latency', *Annual Review of Pharmacology and Toxicology*. doi: 10.1146/annurev-pharmtox-010510-100237.
- Christian, M. L. *et al.* (2012) 'Targeting G with TAL Effectors: A Comparison of Activities of TALENs Constructed with NN and NK Repeat Variable Di-Residues', *PLoS ONE*. doi: 10.1371/journal.pone.0045383.

- Christian, M. *et al.* (2013) 'Targeted Mutagenesis of *Arabidopsis thaliana* Using Engineered TAL Effector Nucleases', *Genes/Genomes/Genetics*. 3(10):1697-705. doi: 10.1534/g3.113.007104.
- Chu, Z., Yuan, M., *et al.* (2006 a) 'Promoter mutations of an essential gene for pollen development result in disease resistance in rice', *Genes and Development*, 20(10), pp. 1250–1255. doi: 10.1101/gad.1416306.
- Chu, Z., Fu, B., *et al.* (2006 b) 'Targeting xa13, a recessive gene for bacterial blight resistance in rice', *Theoretical and Applied Genetics*, 112(3), pp. 455–461. doi: 10.1007/s00122-005-0145-6.
- Cong, L. *et al.* (2012) 'Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains', *Nature Communications*. Nature Publishing Group, 3, pp. 966–968. doi: 10.1038/ncomms1962.
- Cong, L. *et al.* (2013) 'Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121), pp. 819–823. doi: 10.1126/science.1231143.Multiplex.
- Cuculis, L. *et al.* (2015) 'Direct observation of TALE protein dynamics reveals a two-state search mechanism', *Nature Communications*. Nature Publishing Group, 6, pp. 1–11. doi: 10.1038/ncomms8277.
- Cuculis, L. *et al.* (2016) 'TALE proteins search DNA using a rotationally decoupled mechanism', *Nature Chemical Biology*. doi: 10.1038/nchembio.2152.
- Dehury, B. *et al.* (2015) 'Molecular recognition of avirulence protein (avrxa5) by eukaryotic transcription factor xa5 of rice (*Oryza sativa* L.): Insights from molecular dynamics simulations', *Journal of Molecular Graphics and Modelling*. Elsevier Inc., 57, pp. 49–61. doi: 10.1016/j.jmglm.2015.01.005.
- Deng, D. *et al.* (2012) 'Structural basis for sequence-specific recognition of DNA by TAL effectors', *Science*. doi: 10.1126/science.1215670.
- Dorsman, J. C. *et al.* (2007) 'Identification of the Fanconi anemia complementation group I gene, FANCI.', *Cellular oncology : the official journal of the International Society for Cellular Oncology*.
- Doyle, E. L. *et al.* (2012) 'TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction', *Nucleic Acids Research*. doi: 10.1093/nar/gks608.
- Engler, C., Kandzia, R. and Marillonnet, S. (2008) 'A one pot, one step, precision cloning method with high throughput capability', *PLoS ONE*, 3(11). doi: 10.1371/journal.pone.0003647.
- Finzi, D. *et al.* (1997) 'Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy', *Science*. doi: 10.1126/science.278.5341.1295.
- Flint, S. J., Enquist, L. W., Racaniello, V. R., Skalka, A. M. 2009. Principles of Virology (Vol. I - Molecular Biology), Third Edition
- Folks, T. M. *et al.* (1986) 'Biological and biochemical characterization of a cloned Leu-3- cell surviving infection with the acquired immune deficiency syndrome retrovirus.', *The Journal of experimental medicine*. 164(1):280-90.
- Gao, H. *et al.* (2012) 'Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region', *Cell Research*. doi: 10.1038/cr.2012.156.

- Garg, A. *et al.* (2012) 'Engineering synthetic TAL effectors with orthogonal target sites', *Nucleic Acids Research*, 40(15), pp. 7584–7595. doi: 10.1093/nar/gks404.
- Gautron, A. S. *et al.* (2017) 'Fine and Predictable Tuning of TALEN Gene Editing Targeting for Improved T Cell Adoptive Immunotherapy', *Molecular Therapy - Nucleic Acids*. doi: 10.1016/j.omtn.2017.10.005.
- Geeraert, L., Kraus, G. and Pomerantz, R. J. (2007) 'Hide-and-Seek: The Challenge of Viral Persistence in HIV-1 Infection', *Annual Review of Medicine*. doi: 10.1146/annurev.med.59.062806.123001.
- Geißler, R. *et al.* (2011) 'Transcriptional activators of human genes with programmable DNA-specificity', *PLoS ONE*, 6(5), pp. 1–7. doi: 10.1371/journal.pone.0019509.
- Geißler, R. *et al.* (2015) 'Patient-adapted, specific activation of HIV-1 by customized TAL effectors (TALEs), a proof of principle study', *Virology*. Elsevier, 486, pp. 248–254. doi: 10.1016/j.virol.2015.09.018.
- Gilbert, L. A. *et al.* (2013) 'CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes', *Cell*. doi: 10.1016/j.cell.2013.06.044.
- Graham, F. L. *et al.* (1977) 'Characteristics of a human cell line transformed by DNA from human adenovirus type 5', *Journal of General Virology*. doi: 10.1099/0022-1317-36-1-59.
- Grau, J. *et al.* (2013) 'Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites', *PLoS Computational Biology*. doi: 10.1371/journal.pcbi.1002962.
- Gu, K. *et al.* (2005) 'R gene expression induced by a type-III effector triggers disease resistance in rice', *Nature*. doi: 10.1038/nature03630.
- Gu, K. *et al.* (2009) 'Transcription activator-like type III effector avrxa27 depends on ostfiiay5 for the activation of Xa27 transcription in rice that triggers disease resistance to *Xanthomonas oryzae* pv. *oryzae*: Short communication', *Molecular Plant Pathology*. doi: 10.1111/j.1364-3703.2009.00567.x.
- Gueiderikh, A., Rosselli, F. and Neto, J. B. C. (2017) 'A never-ending story: The steadily growing family of the FA and FA-like genes', *Genetics and Molecular Biology*. doi: 10.1590/1678-4685-GMB-2016-0213.
- Guo, J., Gaj, T. and Barbas, C. F. (2010) 'Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases', *Journal of Molecular Biology*. Elsevier Ltd, 400(1), pp. 96–107. doi: 10.1016/j.jmb.2010.04.060.
- Guschin, D. Y. *et al.* (2010) 'A rapid and general assay for monitoring endogenous gene modification', *Methods in Molecular Biology*. doi: 10.1007/978-1-60761-753-2_15.
- Hall, D. B. and Struhl, K. (2002) 'The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo', *Journal of Biological Chemistry*, 277(48), pp. 46043–46050. doi: 10.1074/jbc.M208911200.
- Hilton, I. B. *et al.* (2015) 'Activates Genes From Promoters and Enhancers', *Nature Biotechnology*, 33(5), pp. 510–517. doi: 10.1038/nbt.3199.Epigenome.
- Hirai, H., Tani, T. and Kikyo, N. (2011) 'Structure and functions of powerful transactivators: VP16, MyoD and FoxA', *International Journal of Developmental Biology*. doi: 10.1387/ijdb.103194hh.

- Holt, N. *et al.* (2010) 'Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo', *Nature Biotechnology*. doi: 10.1038/nbt.1663.
- Hruscha, A. *et al.* (2013) 'Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish', *Development*. doi: 10.1242/dev.099085.
- Hsu, P. D. *et al.* (2013) 'DNA targeting specificity of RNA-guided Cas9 nucleases', *Nature Biotechnology*. doi: 10.1038/nbt.2647.
- Hu, Y. *et al.* (2014) 'Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1313271111.
- Huang, P. *et al.* (2011) 'Heritable gene targeting in zebrafish using customized TALENs', *Nature Biotechnology*. doi: 10.1038/nbt.1939.
- Iyer-Pascuzzi, A. S. *et al.* (2008) 'Genetic and Functional Characterization of the Rice Bacterial Blight Disease Resistance Gene xa5', *Phytopathology*. doi: 10.1094/phyto-98-3-0289.
- Iyer, A. S. and McCouch, S. R. (2004) 'The Rice Bacterial Blight Resistance Gene xa5 Encodes a Novel Form of Disease Resistance', *Molecular Plant-Microbe Interactions*, 17(12), pp. 1348–1354. doi: 10.1094/mpmi.2004.17.12.1348.
- Ji, H. *et al.* (2016) 'Specific reactivation of latent HIV-1 by dCas9-SunTag-VP64-mediated guide RNA targeting the HIV-1 promoter', *Molecular Therapy*. doi: 10.1038/mt.2016.7.
- Ji, Z. *et al.* (2016) 'Interfering TAL effectors of *Xanthomonas oryzae* neutralize R-gene-mediated plant disease resistance', *Nature Communications*. doi: 10.1038/ncomms13435.
- Jinek, M. *et al.* (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science*. doi: 10.1126/science.1225829.
- Jones, K. A. *et al.* (1986) 'Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1', *Science*. doi: 10.1126/science.3008338.
- Jones, J. D. G. and Dangl, J. L. (2006) 'The plant immune system', *Nature*. doi: 10.1038/nature05286.
- Jordan, A., Bisgrove, D. and Verdin, E. (2003) 'HIV reproducibly establishes a latent infection after acute infection of T cells in vitro', *EMBO Journal*. doi: 10.1093/emboj/cdg188.
- Joung, J. K., Koepf, D. M. and Hochschild, A. (1994) 'Synergistic activation of transcription by bacteriophage λ cI protein and E. coli cAMP receptor protein', *Science*. doi: 10.1126/science.8091212.
- Juillerat, A. *et al.* (2014) 'Comprehensive analysis of the specificity of transcription activator-like effector nucleases', *Nucleic Acids Research*, 42(8), pp. 5390–5402. doi: 10.1093/nar/gku155.
- Juillerat, A. *et al.* (2015) 'Optimized tuning of TALEN specificity using non-conventional RVDs', *Scientific Reports*, 5, pp. 1–7. doi: 10.1038/srep08150.
- Karn, J. and Stoltzfus, C. M. (2012) 'Transcriptional and posttranscriptional regulation of HIV-1 gene expression', *Cold Spring Harbor Perspectives in Medicine*, 2(2), pp. 1–17. doi: 10.1101/cshperspect.a006916.

- Katlama, C. *et al.* (2013) 'Barriers to a cure for HIV: New ways to target and eradicate HIV-1 reservoirs', *The Lancet*. doi: 10.1016/S0140-6736(13)60104-X.
- Kay, S. *et al.* (2007) 'A bacterial effector acts as a plant transcription factor and induces a cell size regulator', *Science*. doi: 10.1126/science.1144956.
- Kayser, F. H., Böttger, E. C., Zinkernagel, R. M., Haller, O., Eckert, J., Deplazes, P. 2010. Taschenlehrbuch Medizinische Mikrobiologie, Zwölfte Auflage
- Kim, Y. G. and Chandrasegaran, S. (1994) 'Chimeric restriction endonuclease.', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.91.3.883.
- Kim, Y. G., Cha, J. and Chandrasegaran, S. (1996) 'Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.93.3.1156.
- Kim, E. *et al.* (2012) 'Precision genome engineering with programmable DNA-nicking enzymes', *Genome Research*. doi: 10.1101/gr.138792.112.
- Kimmel, C. B. *et al.* (1995) 'Stages of embryonic development of the zebrafish', *Developmental Dynamics*. doi: 10.1002/aja.1002030302.
- Kleinstiver, B. P. *et al.* (2014) 'The I-TevI Nuclease and Linker Domains Contribute to the Specificity of Monomeric TALENs', *Genes/Genomes/Genetics*. 16;4(6):1155-65. doi: 10.1534/g3.114.011445.
- Kleinstiver, B. P. *et al.* (2016) 'High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects', *Nature*. 529(7587):490-5. doi: 10.1038/nature16526.
- Kobayashi, N., Boyer, T. G. and Berk, A. J. (1995) 'A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly.', *Molecular and Cellular Biology*, 15(11), pp. 6465–6473. doi: 10.1128/mcb.15.11.6465.
- Konermann, S. *et al.* (2015) 'Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex', *Nature*. doi: 10.1038/nature14136.
- Kornberg, R. (2007) 'The molecular basis of eukaryotic transcription (Nobel Lecture)', *Angewandte Chemie - International Edition*, 46(37), pp. 6956–6965. doi: 10.1002/anie.200701832.
- Krishnamurthy, S. and Hampsey, M. (2009) 'Eukaryotic transcription initiation', *Current Biology*, 19(4), pp. R153–R156. doi: 10.1016/j.cub.2008.11.052.
- De Laat, W. L., Jaspers, N. G. J. and Hoeijmakers, J. H. J. (1999) 'Molecular mechanism of nucleotide excision repair', *Genes and Development*. doi: 10.1101/gad.13.7.768.
- Lamb, B. M., Mercer, A. C. and Barbas, C. F. (2013) 'Directed evolution of the TALE N-terminal domain for recognition of all 50 bases', *Nucleic Acids Research*. doi: 10.1093/nar/gkt754.
- De Lange, O. *et al.* (2013) 'Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease', *New Phytologist*. doi: 10.1111/nph.12324.
- Van Larebeke, N. *et al.* (1974) 'Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability', *Nature*. doi: 10.1038/252169a0.

- Lee, P. H. and Shatkay, H. (2008) 'F-SNP: Computationally predicted functional SNPs for disease association studies', *Nucleic Acids Research*. doi: 10.1093/nar/gkm904.
- Li, L., Wv, L. P. and Chandrasegaran, S. (1992) 'Functional domains in *Fok* I restriction endonuclease', *Biochemistry*, 89(May), pp. 4275–4279.
- Li, L. and Chandrasegaran, S. (1993) 'Alteration of the cleavage distance of *Fok* I restriction endonuclease by insertion mutagenesis.', *Proceedings of the National Academy of Sciences*, 90(7), pp. 2764–2768. doi: 10.1073/pnas.90.7.2764.
- Li, T. *et al.* (2011) 'TAL nucleases (TALNs): Hybrid proteins composed of TAL effectors and *Fok*I DNA-cleavage domain', *Nucleic Acids Research*. doi: 10.1093/nar/gkq704.
- Li, L. *et al.* (2013) 'Characterization and DNA-binding specificities of *Ralstonia* TAL-like effectors', *Molecular Plant*. doi: 10.1093/mp/sst006.
- Li, T. *et al.* (2013) 'Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice', *Molecular Plant*. © The Authors. All rights reserved., 6(3), pp. 781–789. doi: 10.1093/mp/sst034.
- Li, Z. *et al.* (2014) 'A potential disease susceptibility gene *CsLOB* of citrus is targeted by a major virulence effector *PthA* of *Xanthomonas citri* subsp. *citri*', *Molecular Plant*. doi: 10.1093/mp/sst176.
- Li, Y. *et al.* (2016) 'Exploiting the CRISPR/Cas9 PAM constraint for single-nucleotide resolution interventions', *PLoS ONE*. doi: 10.1371/journal.pone.0144970.
- Limsirichai, P., Gaj, T. and Schaffer, D. V. (2016) 'CRISPR-mediated activation of latent HIV-1 expression', *Molecular Therapy*. American Society of Gene & Cell Therapy, 24(3), pp. 499–507. doi: 10.1038/mt.2015.213.
- Lin, Y. S. *et al.* (1990) 'How different eukaryotic transcriptional activators can cooperate promiscuously', *Nature*. doi: 10.1038/345359a0.
- Lin, C. H. *et al.* (2010) 'A tryptophan-rich peptide acts as a transcription activation domain', *BMC Molecular Biology*. doi: 10.1186/1471-2199-11-85.
- Lin, Y. *et al.* (2014) 'CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences', *Nucleic Acids Research*. doi: 10.1093/nar/gku402.
- Van Lint, C. *et al.* (1996) 'Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation.', *The EMBO Journal*. doi: 10.1002/j.1460-2075.1996.tb00449.x.
- Liu, R. *et al.* (1996) 'Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection', *Cell*. doi: 10.1016/S0092-8674(00)80110-5.
- Liu, P. Q. *et al.* (2001) 'Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions: Activation of vascular endothelial growth factor A', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M011172200.
- Liu, Q. *et al.* (2011) 'A paralog of the *MtN3*/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice', *Plant, Cell and Environment*, 34(11), pp. 1958–1969. doi: 10.1111/j.1365-3040.2011.02391.x.
- Ma, S. *et al.* (2012) 'Highly Efficient and Specific Genome Editing in Silkworm Using Custom TALENs', *PLoS ONE*, 7(9), pp. 1–7. doi: 10.1371/journal.pone.0045035.

- Maeder, M. L., Linder, S. J., Reyon, D., *et al.* (2013 a) 'Robust, synergistic regulation of human gene expression using TALE activators', *Nature Methods*. doi: 10.1038/nmeth.2366.
- Maeder, M. L., Linder, S. J., Cascio, V. M., *et al.* (2013 b) 'CRISPR RNA-guided activation of endogenous human genes', *Nature Methods*. doi: 10.1038/nmeth.2598.
- Mahfouz, M. M. *et al.* (2011) 'De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1019533108.
- Mahfouz, M. M. *et al.* (2012) 'Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein', *Plant Molecular Biology*. doi: 10.1007/s11103-011-9866-x.
- Mak, A. N. S. *et al.* (2012) 'The crystal structure of TAL effector PthXo1 bound to its DNA target', *Science*. doi: 10.1126/science.1216211.
- Mali, P., Yang, L., *et al.* (2013 a) 'RNA-guided human genome engineering via Cas9', *Science*. doi: 10.1126/science.1232033.
- Mali, P., Aach, J., *et al.* (2013 b) 'CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering', *Nature Biotechnology*. doi: 10.1038/nbt.2675.
- Marois, E., Van den Ackerveken, G. and Bonas, U. (2002) 'The Xanthomonas Type III Effector Protein AvrBs3 Modulates Plant Gene Expression and Induces Cell Hypertrophy in the Susceptible Host', *Molecular Plant-Microbe Interactions*, 15(7), pp. 637–646. doi: 10.1094/mpmi.2002.15.7.637.
- Meckler, J. F. *et al.* (2013) 'Quantitative analysis of TALE-DNA interactions suggests polarity effects', *Nucleic Acids Research*. doi: 10.1093/nar/gkt085.
- Mercer, A. C. *et al.* (2012) 'Chimeric TALE recombinases with programmable DNA sequence specificity', *Nucleic Acids Research*, 40(21), pp. 11163–11172. doi: 10.1093/nar/gks875.
- Miller-Jensen, K. *et al.* (2012) 'Chromatin accessibility at the HIV LTR promoter sets a threshold for NF- κ B mediated viral gene expression', *Integrative Biology (United Kingdom)*. doi: 10.1039/c2ib20009k.
- Miller, J. C. *et al.* (2007) 'An improved zinc-finger nuclease architecture for highly specific genome editing', *Nature Biotechnology*. doi: 10.1038/nbt1319.
- Miller, J. C. *et al.* (2011) 'A TALE nuclease architecture for efficient genome editing', *Nature Biotechnology*. doi: 10.1038/nbt.1755.
- Miller, J. C. *et al.* (2015) 'Improved specificity of TALE-based genome editing using an expanded RVD repertoire', *Nature Methods*. Nature Publishing Group, 12(5), pp. 465–471. doi: 10.1038/nmeth.3330.
- Minczuk, M. *et al.* (2008) 'Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA', *Nucleic Acids Research*. doi: 10.1093/nar/gkn313.
- Mino, T., Aoyama, Y. and Sera, T. (2009) 'Efficient double-stranded DNA cleavage by artificial zinc-finger nucleases composed of one zinc-finger protein and a single-chain FokI dimer', *Journal of Biotechnology*, 140(3–4), pp. 156–161. doi: 10.1016/j.jbiotec.2009.02.004.

- Mladenov, E. and Iliakis, G. (2011) 'Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways', *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. doi: 10.1016/j.mrfmmm.2011.02.005.
- Morbitzer, R. *et al.* (2010) 'Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors', *Proceedings of the National Academy of Sciences*, 107(50), pp. 21617–21622. doi: 10.1073/pnas.1013133107.
- Morbitzer, R. *et al.* (2011) 'Assembly of custom TALE-type DNA binding domains by modular cloning', *Nucleic Acids Research*, 39(13), pp. 5790–5799. doi: 10.1093/nar/gkr151.
- Moscou, M. J. and Bogdanove, A. J. (2009) 'A simple cipher governs DNA recognition by TAL effectors', *Science*. doi: 10.1126/science.1178817.
- Mussolino, C. *et al.* (2011) 'A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity', *Nucleic Acids Research*. doi: 10.1093/nar/gkr597.
- Nakabayashi, H. *et al.* (1984) 'Phenotypical Stability of a Human Hepatoma Cell Line, HuH-7, in Long-Term Culture with Chemically Defined Medium', *GANN Japanese Journal of Cancer Research*. doi: 10.20772/cancersci1959.75.2_151.
- Nakagawa, T. *et al.* (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* 104, 34–41.
- Nelson, C. E. and Gersbach, C. A. (2016) 'Cas9 loosens its grip on off-target sites', *Nature Biotechnology*. doi: 10.1038/nbt.3501.
- Niño-Liu, D. O., Ronald, P. C. and Bogdanove, A. J. (2006) 'Xanthomonas oryzae pathovars: Model pathogens of a model crop', *Molecular Plant Pathology*. doi: 10.1111/j.1364-3703.2006.00344.x.
- van Opijnen, T. *et al.* (2004) 'The Human Immunodeficiency Virus Type 1 Promoter Contains a CATA Box Instead of a TATA Box for Optimal Transcription and Replication', *Journal of Virology*. doi: 10.1128/jvi.78.13.6883-6890.2004.
- Ordon, J. *et al.* (2017) 'Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit', *Plant Journal*, 89(1), pp. 155–168. doi: 10.1111/tpj.13319.
- Osborn, M. J. *et al.* (2013) 'TALEN-based gene correction for epidermolysis bullosa', *Molecular Therapy*. The American Society of Gene & Cell Therapy, 21(6), pp. 1151–1159. doi: 10.1038/mt.2013.56.
- Osborn, M. J. *et al.* (2016) 'CRISPR/Cas9 Targeted Gene Editing and Cellular Engineering in Fanconi Anemia', *Stem Cells and Development*. doi: 10.1089/scd.2016.0149.
- Ozer, J. *et al.* (1994) 'Molecular cloning of the small (gamma) subunit of human TFIIA reveals functions critical for activated transcription.', *Genes & Development*, 8(19), pp. 2324–2335. doi: 10.1101/gad.8.19.2324.
- Perdigão, P. *et al.* (2016) 'Reactivation of latent HIV-1 expression by engineered TALE transcription factors', *PLoS ONE*, 11(3), pp. 1–18. doi: 10.1371/journal.pone.0150037.

- Pereira, A. L. A. *et al.* (2014) 'Identification of putative TAL effector targets of the citrus canker pathogens shows functional convergence underlying disease development and defense response', *BMC Genomics*. doi: 10.1186/1471-2164-15-157.
- Pereira, C. V *et al.* (2018) 'mitoTev-TALE: a monomeric DNA editing enzyme to reduce mutant mitochondrial DNA levels', *EMBO Molecular Medicine*. doi: 10.15252/emmm.201708084.
- Perelson, A. S. *et al.* (1996) 'HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time', *Science*. doi: 10.1126/science.271.5255.1582.
- Perez-Pinera, P., Ousterout, D. G., *et al.* (2013 a) 'Synergistic and tunable human gene activation by combinations of synthetic transcription factors', *Nature Methods*. doi: 10.1038/nmeth.2361.
- Perez-Pinera, P., Kocak, D. D., *et al.* (2013 b) 'RNA-guided gene activation by CRISPR-Cas9-based transcription factors', *Nature Methods*. doi: 10.1038/nmeth.2600.
- Pérez-Quintero, A. L. *et al.* (2013) 'An Improved Method for TAL Effectors DNA-Binding Sites Prediction Reveals Functional Convergence in TAL Repertoires of *Xanthomonas oryzae* Strains', *PLoS ONE*. doi: 10.1371/journal.pone.0068464.
- Perez, E. E. *et al.* (2008) 'Establishment of HIV-1 resistance in CD4+T cells by genome editing using zinc-finger nucleases', *Nature Biotechnology*. doi: 10.1038/nbt1410.
- Perkins, N. D. *et al.* (1993) 'A cooperative interaction between NF- κ B and Sp1 is required for HIV-1 enhancer activation', *The EMBO journal*, 12(9), pp. 3551–3558.
- Qu, D. *et al.* (2016) 'The variances of Sp1 and NF- κ B elements correlate with the greater capacity of Chinese HIV-1 B'-LTR for driving gene expression', *Scientific Reports*, 6(September), pp. 1–11. doi: 10.1038/srep34532.
- Ramirez, C. L. *et al.* (2012) 'Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects', *Nucleic Acids Research*. doi: 10.1093/nar/gks179.
- Read, A. C. *et al.* (2016) 'Suppression of Xo1 -Mediated Disease Resistance in Rice by a Truncated , Non-DNA-Binding TAL Effector of *Xanthomonas oryzae*', *Frontiers in Plant Science*. doi: 10.3389/fpls.2016.01516.
- Ren, R. *et al.* (2017) 'Visualization of aging-associated chromatin alterations with an engineered TALE system', *Cell Research*. doi: 10.1038/cr.2017.18.
- Renner, I. *et al.* (2014) 'Antiviral interferon-beta signaling induced by designed transcription activator-like effectors (TALE)', *PLoS ONE*, 9(12), pp. 3–10. doi: 10.1371/journal.pone.0114288.
- Reschke, M. (2018) 'Funktionelle Analyse von TALEs in *Xanthomonas oryzae* und Entwicklung von TALENs zum Genome Editing', Dissertation
- Reyon, D. *et al.* (2012) 'FLASH assembly of TALENs for high-throughput genome editing', *Nature Biotechnology*. doi: 10.1038/nbt.2170.
- Rittner, K. *et al.* (1995) 'The human immunodeficiency virus long terminal repeat includes a specialised initiator element which is required for tat-responsive transcription', *Journal of Molecular Biology*, 248(3), pp. 562–580. doi: 10.1006/jmbi.1995.0243.
- Rohr, O. *et al.* (2003) 'Regulation of HIV-1 gene transcription: from lymphocytes to microglial cells', *Journal of Leukocyte Biology*, 74(5), pp. 736–749. doi: 10.1189/jlb.0403180.

- Römer, P. *et al.* (2007) 'Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene', *Science*. doi: 10.1126/science.1144958.
- Römer, P. *et al.* (2009 a) 'Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper *Bs3* alleles.', *Plant physiology*. doi: 10.1104/pp.109.139931.
- Römer, P., Recht, S. and Lahaye, T. (2009 b) 'A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0908812106.
- Römer, P. *et al.* (2010) 'Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*', *New Phytologist*, 187(4), pp. 1048–1057. doi: 10.1111/j.1469-8137.2010.03217.x.
- Ross, E. K. *et al.* (1991) 'Contribution of NF-kappa B and Sp1 binding motifs to the replicative capacity of human immunodeficiency virus type 1: distinct patterns of viral growth are determined by T-cell types.', *Journal of virology*. 65(8):4350-8.
- Saayman, S. M. *et al.* (2016) 'Potent and targeted activation of latent HIV-1 using the CRISPR/dCas9 activator complex', *Molecular Therapy*. doi: 10.1038/mt.2015.202.
- Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., Arnheim, N. 1985. Enzymatic Amplification of β -Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. *Science* 230:1350-1354
- San Filippo, J., Sung, P. and Klein, H. (2008) 'Mechanism of Eukaryotic Homologous Recombination', *Annual Review of Biochemistry*. 77:229-57. doi: 10.1146/annurev.biochem.77.061306.125255.
- Schmid-Burgk, J. L. *et al.* (2013) 'A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes', *Nature Biotechnology*. doi: 10.1038/nbt.2460.
- Schmid, B. *et al.* (2013) 'Loss of ALS-associated TDP-43 in zebrafish causes muscle degeneration, vascular dysfunction, and reduced motor neuron axon outgrowth', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1218311110.
- Scholze, H. and Boch, J. (2011) 'TAL effectors are remote controls for gene activation', *Current Opinion in Microbiology*. doi: 10.1016/j.mib.2010.12.001.
- Schornack, S. *et al.* (2004) 'The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3', *Plant Journal*. doi: 10.1046/j.1365-313X.2003.01937.x.
- Schreiber, T. and Bonas, U. (2014) 'Repeat 1 of TAL effectors affects target specificity for the base at position zero', *Nucleic Acids Research*. doi: 10.1093/nar/gku341.
- Schreiber, T. *et al.* (2015) 'Refined requirements for protein regions important for activity of the TALE AvrBs3', *PLoS ONE*. doi: 10.1371/journal.pone.0120214.
- Sims, A. E. *et al.* (2007) 'FANCI is a second monoubiquitinated member of the Fanconi anemia pathway', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb1252.
- Slaymaker, I. M. *et al.* (2016) 'Rationally engineered Cas9 nucleases with improved specificity', *Science*. doi: 10.1126/science.aad5227.

- Smale, S. T. and Kadonaga, J. T. (2003) 'The RNA Polymerase II Core Promoter', *Annual Review of Biochemistry*, 72(1), pp. 449–479. doi: 10.1146/annurev.biochem.72.121801.161520.
- Smith, J., Berg, J. M. and Chandrasegaran, S. (1999) 'A detailed study of the substrate specificity of a chimeric restriction enzyme', *Nucleic Acids Research*. doi: 10.1093/nar/27.2.674.
- Smogorzewska, A. *et al.* (2007) 'Identification of the FANCI Protein, a Monoubiquitinated FANCD2 Paralog Required for DNA Repair', *Cell*. doi: 10.1016/j.cell.2007.03.009.
- Stargell, L. A. and Struhl, K. (1995) 'The TBP-TFIIA interaction in the response to acidic activators in vivo', *Science*. doi: 10.1126/science.7604282.
- Stella, S. *et al.* (2013) 'Structure of the AvrBs3-DNA complex provides new insights into the initial thymine-recognition mechanism', *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 69(9), pp. 1707–1716. doi: 10.1107/S0907444913016429.
- Strauß, T. *et al.* (2012) 'RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1212415109.
- Streubel, J. *et al.* (2012) 'TAL effector RVD specificities and efficiencies', *Nature Biotechnology*. doi: 10.1038/nbt.2304.
- Streubel, J. *et al.* (2013) 'Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*', *New Phytologist*. doi: 10.1111/nph.12411.
- Sugio, A. *et al.* (2007) 'Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OsTFIIAgamma1* and *OsTFXI* during bacterial blight of rice.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(25), pp. 10720–5. doi: 10.1073/pnas.0701742104.
- Sugisaki, H. and Kanazawa, S. (1981) 'New restriction endonucleases from *Flavobacterium okeanoikoites* (FokI) and *Micrococcus luteus* (MluI)', *Gene*. 16(1-3):73-8.
- Sullivan, S. M. *et al.* (1998) 'Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus', *Nucleic Acids Research*. doi: 10.1093/nar/26.19.4487.
- Sun, N. and Zhao, H. (2014) 'A single-chain TALEN architecture for genome engineering', *Molecular BioSystems*, 10(3), pp. 446–453. doi: 10.1039/c3mb70412b.
- Szcepek, M. *et al.* (2007) 'Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases', *Nature Biotechnology*. doi: 10.1038/nbt1317.
- Szurek, B. *et al.* (2001) 'Eukaryotic features of the *Xanthomonas* type III effector AvrBs3: Protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper', *Plant Journal*, 26(5), pp. 523–534. doi: 10.1046/j.0960-7412.2001.01046.x.
- Szurek, B. *et al.* (2002) 'Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell', *Molecular Microbiology*. doi: 10.1046/j.1365-2958.2002.03139.x.
- Tanenbaum, M. E. *et al.* (2014) 'A protein-tagging system for signal amplification in gene expression and fluorescence imaging.', *Cell*. doi: 10.1016/j.cell.2014.09.039.

- Taylor, B. S. *et al.* (2008) 'The Challenge of HIV-1 Subtype Diversity', *New England Journal of Medicine*. doi: 10.1056/nejmra0706737.
- Tesson, L. *et al.* (2011) 'Knockout rats generated by embryo microinjection of TALENs', *Nature Biotechnology*. doi: 10.1038/nbt.1940.
- Tian, D. *et al.* (2014) 'The rice TAL effector-dependent resistance protein XA10 triggers cell death and calcium depletion in the endoplasmic reticulum.', *The Plant cell*. doi: 10.1105/tpc.113.119255.
- Triezenberg, S. J. (1995) 'Structure and function of transcriptional activation domains', *Current Opinion in Genetics and Development*, 5(2), pp. 190–196. doi: 10.1016/0959-437X(95)80007-7.
- Tsuji, S., Futaki, S. and Imanishi, M. (2013) 'Creating a TALE protein with unbiased 5'-T binding', *Biochemical and Biophysical Research Communications*. doi: 10.1016/j.bbrc.2013.10.060.
- Urnov, F. D. *et al.* (2005) 'Highly efficient endogenous human gene correction using designed zinc-finger nucleases', *Nature*. doi: 10.1038/nature03556.
- Verdier, V., Vera Cruz, C. and Leach, J. E. (2012) 'Controlling rice bacterial blight in Africa: Needs and prospects', *Journal of Biotechnology*. doi: 10.1016/j.jbiotec.2011.09.020.
- Verhoef, K. *et al.* (1999) 'Evolution of the Human Immunodeficiency Virus Type 1 Long Terminal Repeat Promoter by Conversion of an NF- κ B Enhancer Element into a GABP Binding Site', *Journal of Virology*, 73(2), pp. 1331–1340. Available at: <http://jvi.asm.org/content/73/2/1331.full.pdf>.
- Wan, H. *et al.* (2013) 'Molecular Dynamics Simulations of DNA-Free and DNA-Bound TAL Effectors', *PLoS ONE*, 8(10). doi: 10.1371/journal.pone.0076045.
- Wang, Z. and Moulton, J. (2001) 'SNPs, protein structure, and disease', *Human Mutation*. doi: 10.1002/humu.22.
- Wang, J. *et al.* (2012) 'Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme', *Genome Research*. doi: 10.1101/gr.122879.111.
- Wang, H. *et al.* (2013) 'TALEN-mediated editing of the mouse y chromosome', *Nature Biotechnology*. doi: 10.1038/nbt.2595.
- Wang, P. *et al.* (2014) 'Specific reactivation of latent HIV-1 with designer zinc-finger transcription factors targeting the HIV-1 5'-LTR promoter', *Gene Therapy*, 21(5), pp. 490–495. doi: 10.1038/gt.2014.21.
- Wang, X. *et al.* (2015) 'Designed Transcription Activator-Like Effector Proteins Efficiently Induced the Expression of Latent HIV-1 in Latently Infected Cells', *AIDS Research and Human Retroviruses*. doi: 10.1089/aid.2014.0121.
- Waugh, D. S. and Sauer, R. T. (1993) 'Single amino acid substitutions uncouple the DNA binding and strand scission activities of Fok I endonuclease.', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.90.20.9596.
- Wicky, B. I. M., Stenta, M. and Dal Peraro, M. (2013) 'TAL effectors specificity stems from negative discrimination', *PLoS ONE*, 8(11). doi: 10.1371/journal.pone.0080261.
- Williams, S. A. and Greene, W. C. (2007) 'Regulation of HIV-1 latency by T-cell activation', *Cytokine*. doi: 10.1016/j.cyto.2007.05.017.

- Wong, J. K. *et al.* (1997) 'Recovery of replication-competent HIV despite prolonged suppression of plasma viremia', *Science*. doi: 10.1126/science.278.5341.1291.
- Wood, A. J. *et al.* (2011) 'Targeted genome editing across species using ZFNs and TALENs', *Science*. doi: 10.1126/science.1207773.
- Wu, Y. *et al.* (2014) 'TALE nickase mediates high efficient targeted transgene integration at the human multi-copy ribosomal DNA locus', *Biochemical and Biophysical Research Communications*. Elsevier Inc., 446(1), pp. 261–266. doi: 10.1016/j.bbrc.2014.02.099.
- Xing, S. and Siliciano, R. F. (2013) 'Targeting HIV latency: Pharmacologic strategies toward eradication', *Drug Discovery Today*. doi: 10.1016/j.drudis.2012.12.008.
- Yang, B. *et al.* (2000) 'The virulence factor AvrXa7 of *Xanthomonas oryzae* pv. *oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.170286897.
- Yang, B. and White, F. F. (2004) 'Diverse Members of the AvrBs3/PthA Family of Type III Effectors Are Major Virulence Determinants in Bacterial Blight Disease of Rice', *Molecular Plant-Microbe Interactions*, 17(11), pp. 1192–1200. doi: 10.1094/mpmi.2004.17.11.1192.
- Yang, B., Sugio, A. and White, F. F. (2006) 'Os8N3 is a host disease-susceptibility gene for bacterial blight of rice', *Proceedings of the National Academy of Sciences*, 103(27), pp. 10503–10508. doi: 10.1073/pnas.0604088103.
- Yang, J. *et al.* (2014) 'Complete decoding of TAL effectors for DNA recognition', *Cell Research*, 24(5), pp. 628–631. doi: 10.1038/cr.2014.19.
- Yanik, M. *et al.* (2013) 'TALE-PvuII fusion proteins - Novel tools for gene targeting', *PLoS ONE*, 8(12), pp. 1–13. doi: 10.1371/journal.pone.0082539.
- Yu, Y. *et al.* (2011) 'Colonization of Rice Leaf Blades by an African Strain of *Xanthomonas oryzae* pv. *oryzae* Depends on a New TAL Effector That Induces the Rice Nodulin-3 *Os11N3* Gene', *Molecular Plant-Microbe Interactions*. doi: 10.1094/MPMI-11-10-0254.
- Yuan, M. *et al.* (2016) 'A host basal transcription factor is a key component for infection of rice by TALE-carrying bacteria', *eLife*, 5(JULY), pp. 1–17. doi: 10.7554/eLife.19605.
- Yuan, T. *et al.* (2011) 'Characterization of *xanthomonas oryzae*-responsive cis-acting element in the promoter of rice race-specific susceptibility gene *Xa13*', *Molecular Plant*. doi: 10.1093/mp/ssq076.
- Zhang, F. *et al.* (2011) 'Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription', *Nature Biotechnology*. doi: 10.1038/nbt.1775.
- Zhou, J. *et al.* (2015) 'Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice', *Plant Journal*, 82(4), pp. 632–643. doi: 10.1111/tpj.12838.
- Zhu, W. *et al.* (1998) 'AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus', *Mol Plant Microbe Interact.* doi: 10.1094/MPMI.1998.11.8.824.
- Zhu, W. *et al.* (1999) 'The C Terminus of AvrXa10 Can Be Replaced by the Transcriptional Activation Domain of VP16 from the Herpes Simplex Virus', *The Plant Cell Online*, 11(9), pp. 1665–1674. doi: 10.1105/tpc.11.9.1665.

6. Publikationen:

Richter A, Streubel J, Boch J (2016). TAL Effector DNA-Binding Principles and Specificity. *Methods Mol Biol.* 1338:9-25.

Geissler R, Hauber I, Funk N, **Richter A**, Behrens M, Renner I, Chemnitz J, Hofmann-Sieber H, Baum H, van Lunzen J, Boch J, Hauber J, Behrens SE (2015). Patient-adapted, specific activation of HIV-1 by customized TAL effectors (TALEs), a proof of principle study. *Virology* 486:248-54.

Richter A, Streubel J, Blücher C, Szurek B, Reschke M, Grau J, Boch J (2014). A TAL effector repeat architecture for frameshift binding. *Nature communications* 5: 3447.

Streubel J, **Richter A**, Reschke M, Boch J (2013). TALEs - Proteine mit programmierbarer DNA-Bindesepezifität. *BIOspektrum* 19: 370-372.

Richter A, Boch J (2013). Designer TALEs team up for highly efficient gene induction. *Nat Methods.* 10(3):207-8.

Danksagung

Hiermit möchte ich mich recht herzlich bei Prof. Dr. Jens Boch und Prof. Dr. Sven-Erik Behrens für die Betreuung und Erstellung der Gutachten bedanken.

Weiterhin bedanke ich mich bei Prof. Ulla Bonas für die Möglichkeit in ihrem Labor an meiner Promotion arbeiten zu können.

Allen Mitgliedern der Arbeitsgruppen Boch, Bonas und Behrens möchte ich für das angenehme Arbeitsklima, die entgegengebrachte Unterstützung im Laboralltag und den fachlichen Austausch danken. Mein besonderer Dank gilt meinen Laborkollegen und Mitstreitern Jana, Maik und Sebastian.

Außerdem danke ich meinen neuen Kollegen in der Arbeitsgruppe von Prof. Christoph Englert für das angenehme Arbeitsklima und die Möglichkeit ein interessantes neues Forschungsfeld kennenzulernen.

Zum Schluss möchte ich meiner Familie danken, die mich unterstützt und an mich geglaubt hat.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende wissenschaftliche Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Ich erkläre weiterhin, dass andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht wurden.

Mit dieser Arbeit bewerbe ich mich erstmals um die Erlangung des Doktorgrades.

Halle/Saale, den 28.02.2019

Lebenslauf

Name: Annekatriin Richter
Anschrift: Katharinenstr. 16, 07743 Jena
Geburtsdatum: 31.01.1987
Geburtsort: Wippra
Staatsangehörigkeit: Deutsch

Ausbildung

1993 – 1997 Grundschole, Gerbstedt
1997 – 1999 Sekundarschole, Gerbstedt
1999 – 2006 Gymnasium am Markt, Hettstedt, Abschluss: Abitur
2006 – 2011 Studium der Biologie an der Martin-Luther-Universität Halle-Wittenberg
2010 – 2011 Diplomarbeit im Fachbereich Genetik am Institut für Biologie der Naturwissenschaftlichen Fakultät I der Martin-Luther-Universität HalleWittenberg mit dem ThemaThema: „Untersuchungen zur TAL-Effektor-induzierbaren Transkription an proviralen Sequenzen des Humanen Immundefizienz-Virus“
Abschluss: Diplom
2011 – 2015 Promotionsarbeit im Fachbereich Genetik am Institut für Biologie der Naturwissenschaftlichen Fakultät I der Martin-Luther-Universität HalleWittenberg
2016 – 2019 Wissenschaftliche Mitarbeiterin am Leibniz-Institut für Alternsforschung – Fritz-Lipmann-Institut e. V. in Jena

Halle (Saale), den 28.02.2019