

**Generierung hochspezifischer Nukleasen für die
zielgerichtete Spaltung genomischer
DNA-Sequenzen**

Inauguraldissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften

Dr. rer. nat.

vorgelegt von
Mert Yanik, M. Sc.

angefertigt am
Institut für Biochemie
Fachbereich Biologie und Chemie
Justus-Liebig-Universität Gießen

Gießen 2013

Die vorliegende Doktorarbeit wurde im Zeitraum von November 2009 bis Oktober 2013 im Institut für Biochemie des Fachbereichs 08 an der Justus-Liebig-Universität Gießen unter der Leitung von Prof. Dr. A. Pingoud angefertigt und durch die von der DFG finanzierte IRTG (International Research Training Group) „*Enzymes and Multienzyme complexes acting on nucleic acids*“ (GRK 1384) unterstützt.

Erstgutachter:

Prof. Dr. Alfred Pingoud
Institut für Biochemie
Fachbereich 08; Biologie und Chemie
Justus-Liebig-Universität Gießen
Heinrich-Buff-Ring 58
35392 Gießen

Zweitgutachter:

Prof. Dr. Michael U. Martin
Institut für Immunologie
Fachbereich 08; Biologie und Chemie
Justus-Liebig-Universität Gießen
Schubertstraße 81
35394 Gießen

Erklärung

Ich erkläre: „Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.“

(Mert Yanik)

Gießen, den

Publikationen

Yanik, M., Alzubi, J., Lahaye, T., Cathomen, T., Pingoud, A. & Wende, W. *PvuII fusion proteins - novel tools for gene targeting*
PLoS ONE 2013 8(12): e82539.

Eingereichte Patente

Yanik, M., Wende, W., & Pingoud, A. (2011) *Chimeric nucleases for gene targeting*:
EP11182725.9.

“ Die Wissenschaft fängt eigentlich erst da an, interessant zu werden, wo sie aufhört.”

Justus von Liebig

Inhaltsverzeichnis

1	Einleitung	9
1.1	Zunahme genomischer Sequenzen in Datenbanken	9
1.2	Gentherapie / <i>Gene Targeting</i>	10
1.3	Hochspezifische Nukleasen	12
1.3.1	Zink-Finger-Nukleasen.....	13
1.3.2	Homing Endonukleasen.....	15
1.3.3	TFO-verknüpfte Nukleasen	17
1.3.4	TALE-Nukleasen.....	18
1.3.5	RNA vermittelte Nukleasen.....	20
1.3.6	Proteinfusionen.....	21
1.4	Spaltmodule	22
1.4.1	Restriktionsenzym BfiI.....	22
1.4.2	Restriktionsenzym PvuII	25
1.5	Zielgerichteter Transport der Nukleasen in die Zelle	26
1.6	Epigenetische Aspekte für die Genommodifikation	27
1.7	Zielsetzung.....	28
2	Material und Methoden.....	29
2.1	Materialien	29
2.1.1	Enzyme & Marker	29
2.1.2	Kits zur Aufreinigung von DNA und Protein.....	29
2.1.3	Chemikalien.....	30
2.1.4	Radiochemikalien	31
2.1.5	<i>E.coli</i> Stämme.....	31
2.1.6	Puffer & Medien.....	32
2.1.7	Chromatographie Säulen	33
2.1.8	Plasmide	33
2.1.9	Primer & Oligonukleotide	34
2.1.10	Substrate	36
2.2	Methoden	37
2.2.1	Molekularbiologische Methoden.....	37
2.2.1.1	Klonierungen	37
2.2.1.2	Klonierung des BfiI-I-SceI Konstruktes	38
2.2.1.3	Zielgerichtete Mutagenese der BfiI Bindedomäne	39
2.2.1.4	Klonierung des TALE-PvuII Konstruktes	40
2.2.1.5	Zielgerichtete Mutagenese des TALE-PvuII Konstruktes.....	41
2.2.1.6	Zufall basierte Mutagenese des TALE-PvuII Konstruktes.....	41
2.2.1.7	Klonierung der Substratplasmide.....	41
2.2.2	Biochemische Methoden	42
2.2.2.1	Denaturierende-Gelelektrophorese	42
2.2.2.2	Expression und Aufreinigung von BfiI-I-SceI Varianten.....	42
2.2.2.3	Expression und Aufreinigung von TALE-PvuII Varianten.....	43

2.2.2.4	Chromatographische Trennverfahren.....	44
2.2.2.5	Bestimmung der Proteinkonzentration	44
2.2.2.6	Aktivitätstest der Fusionskonstrukte.....	45
3	Ergebnisse	46
3.1	Hochspezifische Nuklease BfiI-I-SceI.....	46
3.1.1	Aktivitätstest des Fusionsproteins BfiI-I-SceI.....	47
3.1.2	Aktivierung der Spaltung	48
3.1.3	Mutagenese der BfiI Bindedomäne	51
3.1.4	Aktivitätstest der BfiI Einzelmutanten	52
3.1.5	Spaltung fluoreszenz-markierter Oligos.....	54
3.2	Hochspezifische Nuklease TALE-PvuII Homodimer	55
3.2.1	Aktivitätstest des Fusionsproteins TALE-PvuII.....	57
3.2.2	Spaltungskinetiken mit TALE-28-L-wtPvuII HF.....	59
3.2.3	Linkervarianten und präferierte Substrate	63
3.2.4	Spaltungskinetiken mit bipartiten Substraten	64
3.2.5	Stabilitätstest des Enzyms TALE-28-L-wtPvuII HF.....	65
3.2.6	Vergleich physiologischer Puffer	66
3.2.7	Viabilitätsexperiment des TALE-PvuII Konstrukts in <i>E.coli</i>	67
3.2.8	Neue PvuII Varianten im Fusionskonstrukt	67
3.2.9	Viabilitätstests der PvuII Varianten in <i>E. coli</i>	70
3.2.10	Toxizitätstest der TALE-PvuII Varianten in HEK293 Zellen.....	70
3.2.11	Aktivitätstest der homodimeren TALE-PvuII Varianten in HEK293 Zellen	72
3.3	TALE-PvuII Heterodimere	73
3.3.1	Heterodimere mit PvuII und TALE AvrBs3 & AvrBs4.....	74
3.3.2	Heterodimere mit unterschiedlichen PvuII Untereinheiten.....	75
3.3.3	Aktivitätstest der heterodimeren Varianten in HEK293 Zellen	77
3.3.4	Toxizitätstest der heterodimeren Varianten in HEK293 Zellen	78
3.3.5	Adressierung des humanen Interleukin 2 Rezeptor γ Gens.....	79
4	Diskussion	82
4.1	Hochspezifische Nukleasen als Werkzeuge für das <i>genome engineering</i>	82
4.2	Generierung des BfiI-I-SceI Fusionsenzym.....	84
4.3	Generierung des AvrBs3-PvuII Fusionsenzym.....	87
4.3.1	Spezifität der TALE-PvuII Konstrukte.....	88
4.3.2	TALE-PvuII <i>multi-turnover</i> Eigenschaften.....	90
4.3.3	Spaltungskinetiken auf bipartiten Substraten	91
4.3.4	Physiologische Puffer und Stabilität des TALE-PvuII Enzyms.....	92
4.3.5	Viabilitätsexperimente / Selektion neuer Varianten.....	92
4.3.6	TALE-PvuII <i>in vivo</i> Experimente in HEK293 Zellen.....	93
4.3.6.1	TALE repeat 0 Limitierung	96
4.4	Generierung der heterodimeren AvrBs3-PvuII Fusionsenzyme.....	96

5	Fazit	98
6	Zusammenfassung.....	100
7	Abkürzungsverzeichnis.....	101
8	Abbildungsverzeichnis	102
9	Tabellenverzeichnis.....	103
10	Danksagung	104
11	Literaturverzeichnis.....	107

1 Einleitung

1.1 Zunahme genomischer Sequenzen in Datenbanken

Seit der Entdeckung des strukturellen Aufbaus der DNA durch James Watson und Francis Crick im Jahre 1953 (Watson and Crick 1953) und der Entwicklung der Sequenzierungstechnik der DNA durch Fred Sanger im Jahre 1977 (Sanger, Nicklen and Coulson 1977) wurden eine Fülle von genomischen Daten verschiedener Organismen zusammengetragen. Einen Meilenstein erreichte dabei das humane Genomprojekt im Jahre 2001 mit der Entschlüsselung des menschlichen Genoms (Lander *et al.* 2001, Venter *et al.* 2001, McPherson *et al.* 2001).

In der *GenBank*-Datenbank sind die öffentlich zugänglichen Sequenzen der bis jetzt sequenzierten Organismen abgelegt und abrufbar (Benson *et al.* 2013). Die Datenbank wurde 1982 vom *National Center for Biotechnology Information* (NCBI, USA) gegründet und wird täglich mit der europäischen (*European Nucleotide Archive*, ENA) und der japanischen (*DNA Data Bank of Japan*, DDBJ) Datenbank synchronisiert, um eine weltweite Abdeckung zu gewährleisten (Benson *et al.* 2011). Über 164 Millionen Eintragungen sind zum 01.05.2013 erfasst (Abb. 1.1).

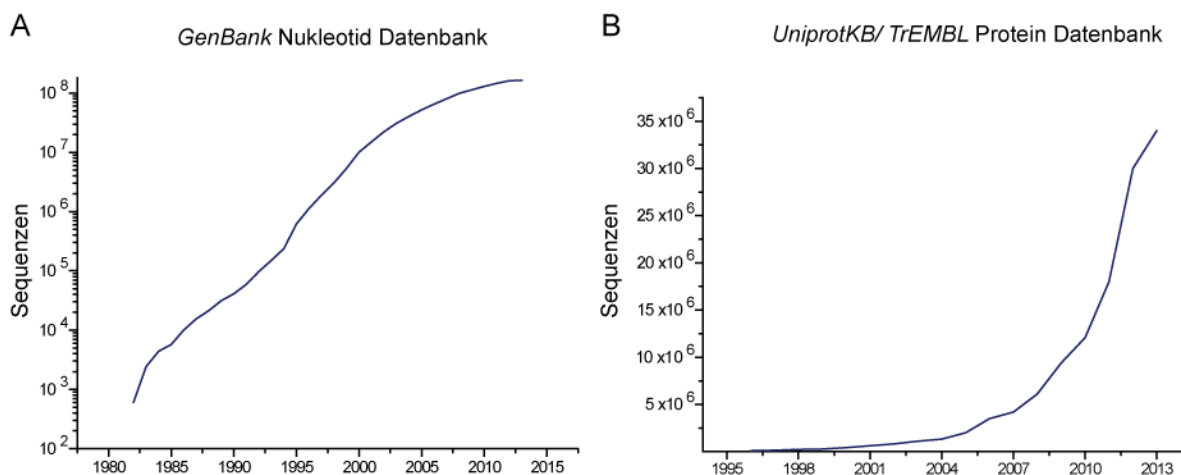


Abb. 1.1: Nukleotid- und Protein-Statistik

A zeigt die Nukleotidsequenzen sämtlicher Organismen, die seit 1982 erfasst wurden. Europa, USA und Japan synchronisieren zusammen die Datenbank. Zum 1.5.2013 sind 164136731 Sequenzen erfasst (www.ncbi.nlm.nih.gov/genbank/statistics). **B** die Uniprot/TrEMBL ProteinDatenbank verwaltet Anfang Mai 2013 knapp 34 Millionen Proteinsequenzen (www.ebi.ac.uk/uniprot/TrEMBLstats/).

Analog zu den Gensequenzen sind Proteinsequenzen ebenfalls in Datenbanken erfasst und der Öffentlichkeit zugänglich. Die Uniprot-Datenbank (*universal protein resource*) fasst und kombiniert die Daten der Proteindatenbanken Swiss-Prot, TrEMBL und PIR-PSD zusammen und ist selbst in vier Kerndatenbanken unterteilt. Davon ist die UniProtKB (Knowledgebase)/TrEMBL-Datenbank diejenige, die am 01.05.2013 knapp 34 Millionen Proteinsequenzen aufweist. Uniprot wird von *European Bioinformatics Institute* (EBI, England), *Swiss Institute of Bioinformatics* (SIB, Schweiz) und *Protein Information Resource* (PIR, USA) verwaltet und zählt zu den größten Proteindatenbanken (Apweiler, Bairoch and Wu 2004, Magrane and Consortium 2011, UniProt 2012, UniProt 2010). Durch die immer besser werdenden Sequenziermethoden und deren Automatisierung können größere Datenmengen in einer kürzeren Zeit erfasst werden (Bonetta 2006, Schmitt *et al.* 2012). Die neue Generation der Sequenziermethoden, auch bekannt als *third generation sequencing*, beschreibt Hochdurchsatzverfahren, mit denen im Vergleich zu den Kapillarelektrophoresen nach der Sangermethode, eine über 100 mal schnellere Generierung von Daten möglich ist (Pareek, Smoczynski and Tretyn 2011, Liu *et al.* 2012).

1.2 Gentherapie / *Gene Targeting*

Mit den Datenmengen und den darauf basierenden Analysen und Forschungen steigt das Verständnis über die Funktionen verschiedener Gene und Proteine. Im komplexen menschlichen Genom mit über 3 Milliarden DNA-Basen konnten Mutationen für viele monogenetische Erkrankungen identifiziert werden (Chial 2008). Durch die spezifische Adressierung von Gensequenzen, dem sogenannten *gene targeting*, ist im Prinzip die Möglichkeit gegeben, monogenetische Erkrankungen zu therapieren (Gentherapie) (Lander 2011). Im Folgenden werden beispielhaft drei Krankheitsbilder gentherapeutisch relevanter Erkrankungen kurz erläutert:

1. Die vererbare Immunschwächekrankheit SCID (*severe combined immunodeficiency*) führt zu einer fehlerhaften Reifung der B- und T-Lymphozyten, wodurch das Immunsystem nicht funktionsfähig ist. Die Lebenserwartung erkrankter Kinder liegt bei wenigen Jahren. Eine Reihe von Mutationen kann zu SCID führen. Die häufigste Form ist die sogenannte X-SCID Immunschwäche mit einer Mutation im humanen Interleukin 2 Rezeptor- γ (*hIL2RG*)-Gen, das sich auf dem X-Chromosom befindet (Fischer, Hacein-Bey-Abina and Cavazzana-Calvo 2010, Rivat *et al.* 2012). Als mögliche

Therapiemaßnahmen wurden Knochenmarktransplantationen (Antoine *et al.* 2003, Buckley *et al.* 1999), Genaddition über virale Systeme (Gaspar *et al.* 2004, Huston *et al.* 2011) und Gentergeting mit spezifischen Zink-Finger-Nukleasen und Reparatur mit Hilfe von homologer Rekombination (Kapitel 1.3) (Moehle *et al.* 2007, Urnov *et al.* 2005) durchgeführt.

2. Als wichtiges gentherapeutisches *target* bei der Bekämpfung von HIV-Viren hat sich das *CCR5*-Gen erwiesen. Es bildet einen Co-Rezeptor auf den T-Helferzellen aus, welcher von HIV-Viren genutzt wird, um die Zelle zu infizieren. Im Jahr 1996 wurde entdeckt, dass eine Deletion von 32 Nukleotiden im *CCR5*-Gen zur Immunität gegen HIV-Viren führt (Deng *et al.* 1996, Liu *et al.* 1996, Samson *et al.* 1996). Durch Modifikation des *CCR5*-Rezeptors konnte die Infektion der Zellen durch HIV-Viren gemindert bzw. vermieden werden (Perez *et al.* 2008).

3. Die Augenkrankheit *Retinitis pigmentosa* führt zur Erblindung und kann ebenfalls auf eine einzelne Mutationen zurückgeführt werden, wie z.B. im *Retinitis pigmentosa* GTPase-Regulator (*RPGR*)-Gen. Sie ist daher auch im Fokus von gentherapeutischen Maßnahmen (Beltran *et al.* 2012, Stieger and Lorenz 2010).

Das Potential des *gene targetings* erkannte der Nobelpreisträger David Baltimore bereits im Jahre 1978.

“. . . it would be a triumph of medicine if the effects of such mutant genes could be countered... A patient could, for instance, be treated in this way for a blood disease caused by an abnormal protein made by a mutant gene. A normal gene would be inserted into the precursor cells - immature bone marrow cells that ultimately develop into functioning blood cells. In this way, a normal protein could be made in place of, or along with, the aberrant protein. The genetically altered blood cell precursor could then cure the patient's disease... It is likely to be the first type of genetic engineering tried on human beings, and might be tried within the next five years" (Baltimore 1978).

Baltimore „hoffte“, dass die Modifikation komplexer Genome bereits Anfang der 80er möglich wäre. Selbst im Jahre 2013 konnte sich seine Vision noch nicht erfüllen, aber an den Perspektiven und Möglichkeiten des *gene targetings* hat sich bis heute nichts geändert.

Eine Integration von DNA-Sequenzen ins Genom kann auch mit Hilfe von Viren durchgeführt werden. Sie haben den Vorteil nahezu jede Zelle *in vivo* adressieren zu können, DNA einzuschleusen und die Zelle zu „reparieren“. Jedoch ist diese Art von Integration nicht dirigierbar. Der zufällige Einbau von DNA-Sequenzen ins Genom erhöht

das Risiko, Onkogene zu aktivieren und Krebs auszulösen oder essentiell wichtige Gene auszuschalten (Cavazzana-Calvo *et al.* 2010, Howe *et al.* 2008, Stein *et al.* 2010). Diese Problematik kann durch Verwendung hochspezifischer Nukleasen zur zielgerichteten Integration gelöst werden.

Aus heutiger Sicht stellen solche hochspezifischen Nukleasen, die spezifische DNA-Spaltungen in komplexen Genomen ausführen, essentielle molekularbiologische Werkzeuge dar.

1.3 Hochspezifische Nukleasen

Die zielgerichtete Adressierung von Genen in einem komplexen Genom (> 1 Milliarde Basenpaare) setzt spezifische und sehr präzise arbeitende Nukleasen voraus. Im Idealfall sollte solch eine Nuklease nur einmal, u.zw. im gewünschten Genlokus, schneiden und keine zytotoxische Aktivität aufweisen. Mit den Nukleasen und der resultierenden Spaltung aktiviert man die zelleigenen Reparatursysteme. Bei einem DNA-Doppelstrangbruch können beide Enden durch Deletion oder Insertion einzelner Basen wieder ligiert werden. Diese Sequenzänderung im Gen führt in den meisten Fällen zu einem Ausfall des Gens (*knock out*). Diese Reparatur wird als *non-homologous end-joining* (NHEJ) bezeichnet und wird verwendet, um Gene auszuschalten (Lieber 2010). Das Gegenstück zu NHEJ ist die homologe Rekombination (HR). HR tritt in Säugerzellen selten auf, ein Ereignis pro 10^5 bis 10^7 Zellen (Vasquez *et al.* 2001), aber durch einen Einzel- oder Doppelstrangbruch in der Nähe einer Mutation kann die Frequenz um mehrere Zehnerpotenzen gesteigert werden (Paques and Duchateau 2007, Smih *et al.* 1995, Pingoud and Silva 2007). Für den Erfolg einer HR ist das Vorhandensein eines DNA-Fragments mit der gewünschten Sequenz und flankierenden homologen Bereichen zum genomischen Abschnitt wichtig. Dieses wird dann anstelle des defekten Gens eingesetzt (Abb. 1.2). Das korrigierte Gen wird z.B. mit Hilfe eines Plasmids in die Zelle eingeschleust, oder das intakte Gen auf dem zweiten Allel dient als Matrize. Die Reparatur des DNA-Strangbruchs via HR wird zur Aktivierung, Insertion und Wiederherstellung von Genen verwendet. Bei der Gentherapie wäre eine Nuklease, die nur einen Einzelstrangbruch ausführt (*nicking*), wünschenswert, da die Reparatur durch NHEJ ausgeschlossen ist und somit auch die Toxizität gemindert wird (Metzger *et al.* 2011, Davis and Maizels 2011).

Es wurden bereits artifizielle Nukleasen mit unterschiedlichen Architekturen generiert, um diesem Ziel näher zu kommen (Pingoud and Wende 2011). Die Genommodifikation mit

Hilfe von hochspezifischen Nukleasen wurde zur Methode des Jahres 2011 in Nature Methods (Methods 2012) gekürt. Im Prinzip wird eine Nuklease als Spaltmodul mit einem DNA-bindenden Protein (Bindungsmodul) kombiniert, um mit dem Bindungsmodul die gewünschte („programmierbare“) Spezifität zu erreichen. Eine Ausnahme bilden die Homing-Endonukleasen (HE), die nur durch aufwendiges *protein engineering* in ihrer Spezifität zu verändern sind. Abb. 1.2 zeigt eine Übersicht der verwendeten Architekturen:

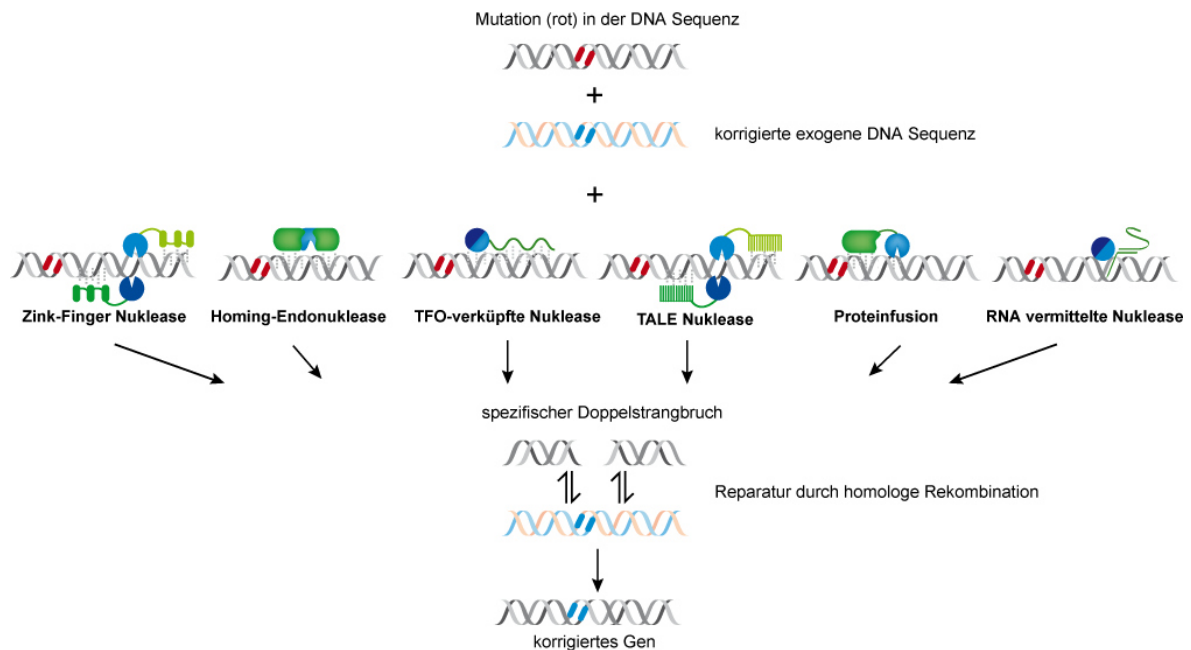


Abb. 1.2: Übersicht hochspezifischer Nukleasen

Doppelstrangbrüche in der Nähe einer Mutation können durch unterschiedliche Nukleasekonstrukte eingefügt werden, um durch homologe Rekombination (HR) das mutierte durch das korrigierte Gen zu ersetzen. Zu den Nukleasearchitekturen zählen die Zink-Finger-Nukleasen, Homing-Endonukleasen, *triple-helix forming oligo* (TFO)-verknüpfte Nukleasen, TALE-Nukleasen, Proteinfusionen, die beispielsweise als monomere Nukleasen agieren und RNA-vermittelte Nukleasen, wie z.B. die Cas9-Endonuklease. Modifiziert nach Pingoud and Wende 2011.

1.3.1 Zink-Finger-Nukleasen

Zink-Finger-Nukleasen (ZFN) zählen zu den bisher meistgenutzten und untersuchten Fusionsnukleasen. 1996 wurde erstmals die unspezifische katalytische Domäne des FokI-Restriktionsenzym an Zink-Finger-Module fusioniert und untersucht (Kim, Cha and Chandrasegaran 1996). Ein Zink-Finger-Modul erkennt und bindet drei DNA-Basen und besteht aus 30 Aminosäuren (Klug 2010). Über verschiedene Kombinationen der Aminosäuren ist es gelungen, nahezu jedes beliebige DNA-Triplett mit Zink-Finger-Modulen zu adressieren (Segal *et al.* 1999, Choo and Klug 1994, Greisman and Pabo 1997, Hurt *et al.* 2003).

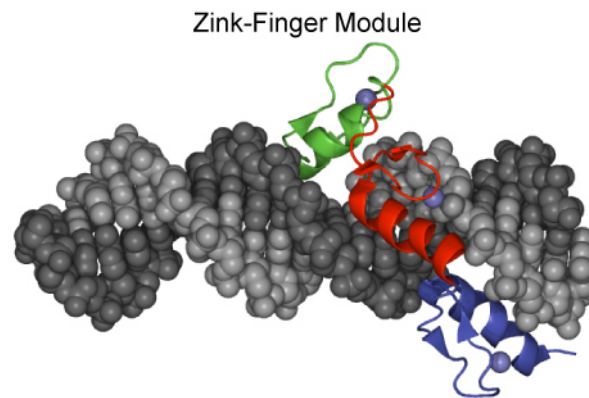


Abb. 1.3: Zink-Finger-Bindungsmodule Co-Röntgenkristallstruktur

Zink-Finger-Proteine (grün, rot und blau) wickeln sich um die große Furche der DNA (grau) und binden an jeweils drei Nucleotide. Zinkionen (hellblau) werden von Zink-Finger-Modulen komplexiert, um die Proteinstruktur zu stabilisieren (pdb: 1P47).

Durch die Aneinanderreihung von drei bis vier ZF-Modulen ist es möglich, längere DNA-Sequenzen zu adressieren. Daher zählen die ZFN zu den ersten „programmierbaren“ Nucleasen und wurden in vielen Organismen wie Taufliegen, Zebrafischen, Tabakpflanzen und humanen Zellen in Kulturen angewendet (Camenisch, Brilliant and Segal 2008, Urnov *et al.* 2010, Handel and Cathomen 2011, Rahman *et al.* 2011).

Die Strategie, ZF-Module als unabhängige Einheiten zu betrachten und sie aneinander zu reihen, ist nicht immer erfolgreich, da die ZF-Module nicht kontext-unabhängig mit der DNA interagieren. Des Weiteren können sich einzelne ZF-Module von der DNA ablösen und somit bei sehr langen ZF-Ketten (> 5) an Spezifität verlieren. Durch diesen Umstand kann eine 100 %ige Bindung nicht gewährleistet werden, und das ist einer der Gründe, warum die ZF-Module sehr gut, entsprechend ihrer Zielsequenz, selektiert werden müssen, um Gentoxizität zu vermeiden (Isalan, Choo and Klug 1997, Ramirez *et al.* 2008, Wright *et al.* 2006). Neue Methoden wie OPEN (*oligomerized pool engineering*) (Maeder *et al.* 2008) und CoDA (*context-dependent assembly*) (Sander *et al.* 2011) sollen dabei helfen, die kontextspezifischen Interaktionen zu berücksichtigen. Dies verdeutlicht die Limitierung der „Programmierbarkeit“ (Kim, Lee and Carroll 2010, Ramirez *et al.* 2008, Sander *et al.* 2011, Pingoud and Wende 2011). Dennoch lassen sich sehr gute Fortschritte erzielen, wenn ein System für einen bestimmten Locus optimiert wird (Gupta *et al.* 2012, Zhu *et al.* 2011), wie beispielsweise die ZFN der Firma Sangamo Biosciences (USA). Diese werden in einer zweiten klinischen Erprobungsphase untersucht, um das *CCR5*-Gen zu adressieren (Tang 2013a, Tang 2013b, Tebas P 2013).

1.3.2 Homing-Endonukleasen

Homing-Endonukleasen (HE) nehmen in dieser Auflistung eine ganz besondere Position ein, denn sie sind bereits von Natur aus sehr spezifisch. Ein Fusionskonstrukt mit einem Bindungsmodul und einem Spaltmodul ist nicht vonnöten, um die Spezifität weiter zu erhöhen. HEs binden zwischen 14 und 40 bp und spalten innerhalb oder in der Nähe ihrer DNA-Erkennungssequenz (Stoddard 2005, Stoddard 2011, Paques and Duchateau 2007, Silva *et al.* 2011). Aufgrund des Aufbaus ihres aktiven Zentrums können die HEs in fünf Familien untergegliedert werden: LAGLIDADG, His-Cys Box, HNH, GIY-YIG und PD-(D/E)_xK. Die LAGLIDADG-Familie ist die bisher größte unter den HEs und ihre Mitglieder werden auch als „Meganukleasen“ bezeichnet, um die HEs von den übrigen hochspezifischen Nukleasekonstrukten zu differenzieren (Stoddard 2005). HEs können als dimere (I-CreI) oder als monomere Nukleasen (I-SceI) aktiv sein. Die Besonderheit der HEs liegt in ihrer Kodierung. Entweder befinden sich ihre Gensequenzen innerhalb von Introns und werden durch Spleißen freigesetzt oder sie existieren in Inteinen und werden aus Vorläuferproteinen durch autokatalytisches Proteinspleißen aktiviert. Ihre biologische Aufgabe ist bis heute unbekannt. Die einzige bekannte Funktion ist die Verbreitung ihrer Gene („*selfish gene*“), welche durch Spaltung des Allels ohne die HE-Gensequenz und der anschließenden HR mit dem Allel, welches die Gensequenz trägt, erfolgt (Chevalier and Stoddard 2001, Jurica and Stoddard 1999).

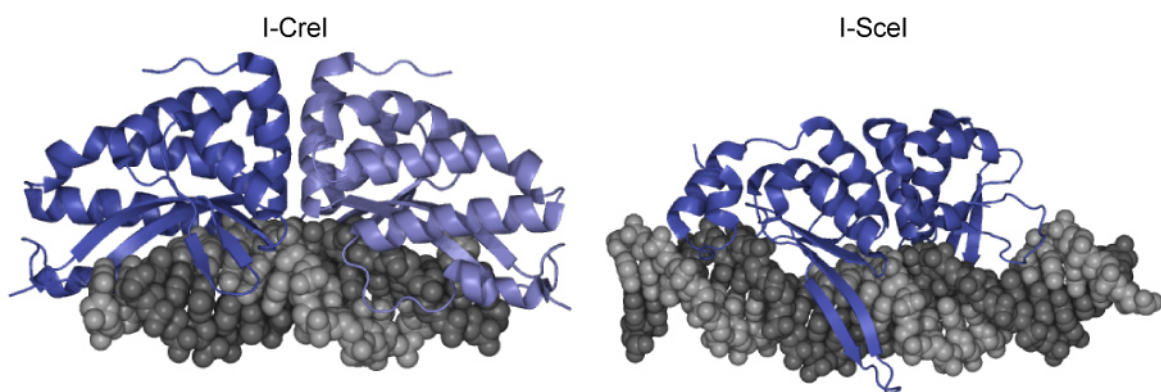


Abb. 1.4: I-CreI & I-SceI – Homing-Endonukleasen aus der LAGLIDADG-Familie

Die Abbildung zeigt die Röntgenkristallstrukturen der beiden Homing-Endonukleasen I-CreI und I-SceI gebunden an ihre Erkennungssequenz. Beide Enzyme sind jeweils intron-kodiert und sind Mitglieder der LAGLIDADG-Familie, die größte innerhalb der HEs, I-CreI (pdb: 1G9Z), ist als Dimer aktiv (blaue und hellblaue Untereinheit), und I-SceI (pdb: 1R7M) ist als Monomer in der Lage, einen DNA-Doppelstrangbruch zu erzielen.

Die von Natur aus gegebene hohe Spezifität war Anlass für zahlreiche Anstrengungen, die Erkennungssequenzen der HEs zu verändern und an neue Sequenzen anzupassen.

Verschiedene Strategien wurden verfolgt, um die Spezifität zu verändern und die katalytische Aktivität beizubehalten. Mit computergestützten Methoden wurden einzelne AS innerhalb der Bindedomänen verändert (Ashworth *et al.* 2006, Ashworth *et al.* 2010) und diverse *in vivo*-Selektion und *screening*-Methoden entwickelt, die auf der Bindung und Spaltung veränderter DNA-Erkennungssequenzen beruhen (Gimble, Moure and Posey 2003, Gruen *et al.* 2002, Rosen *et al.* 2006, Seligman *et al.* 2002, Sussman *et al.* 2004). Hervorzuheben ist dabei die Arbeit des französischen Unternehmens Cellectis (Romainville), die mit Hilfe eines *high-throughput*-Verfahrens eine Bibliothek mit über 30.000 I-CreI Varianten hergestellt hat, die sich in ihren AS für die DNA-Bindungsregion unterscheiden und unterschiedliche DNA-Sequenzen binden. Mit dieser Bibliothek ist man in der Lage, für fast jede beliebige Sequenz eine passende I-CreI-Variante einzusetzen (Arnould *et al.* 2006, Chames *et al.* 2005, Smith *et al.* 2006). Zwei I-CreI-Varianten wurden im humanen Genom eingesetzt. Zum einen im *XPC*-Genlokus, welches in der Krankheit *Xeroderma Pigmentosum* (Arnould *et al.* 2007) involviert ist, und zum anderen im *RAG1*-Gen, das zur Immunkrankheit SCID führt (Grizot *et al.* 2009, Munoz *et al.* 2011). Des Weiteren setzte man Meganukleasen beim sogenannten *virus clipping* ein. Hier nutzt man das Prinzip des NHEJ, indem man virale Sequenzen im Genom adressiert. Diese verlieren nach der Spaltung durch NHEJ ihre Funktion. Am Beispiel des Herpes-simplex-Virus‘ konnte eine um 50 % verringerte Anzahl an Infektionen ermittelt werden (Grosse *et al.* 2011).

HEs auf eine neue Gensequenz zu „programmieren“ ist mit einem enormen Aufwand verbunden (Takeuchi *et al.* 2011). Der Fokus der Forschung liegt nunmehr auf der Entwicklung alternativer Strategien (siehe Kapitel 1.3.4).

Abgesehen von der technisch aufwendigen gerichteten Evolution von HEs, könnten HEs trotzdem auch in Zukunft eine wichtige Rolle beim *gene targeting* einnehmen, denn sie kommen in allen drei Bereichen des Lebens vor. Sie sind in Bakterien, Archaeen und einzelligen Eukaryoten vorhanden. Mit dieser Vielzahl von Organismen und den stetig wachsenden Informationen in Datenbanken (siehe Abb. 1.1: Nukleotid- und Protein-Statistik) steigt auch die Wahrscheinlichkeit, eine passende Meganuklease für eine bestimmte humane Gensequenz zu finden und diese spezifisch zu spalten. Mit neuen Datenbankalgorithmen konnten aus bereits sequenzierten Genomen bisher nicht entdeckte Gene für HEs detektiert werden. Hochrechnungen zufolge wäre es möglich, für jedes humane Transkript eine passende Meganuklease zu finden (Barzel *et al.* 2011).

1.3.3 TFO-verknüpfte Nukleasen

In dieser Kategorie von spezifischen Nukleasen werden DNA-Oligonukleotide als Bindungsmodule verwendet und mit Spaltmodulen verknüpft. Das Oligonukleotid windet sich um die DNA und bildet mit den entsprechenden Basen Hoogsteen-Basenpaare aus, so dass eine Triple-Helix entsteht (TFO, *triple-helix forming oligo*) (Rogers, Lloyd and Glazer 2005, Duca *et al.* 2008, Giovannangeli *et al.* 1992).

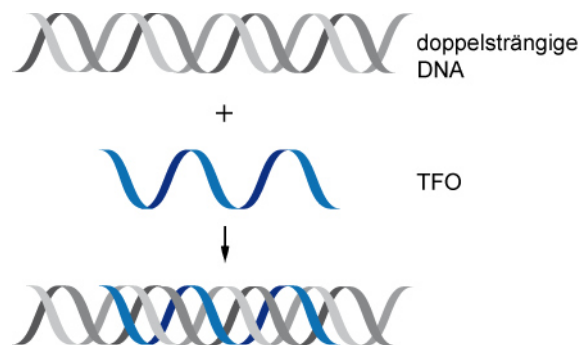


Abb. 1.5: *Triple-helix forming Oligo (TFO)*

Mit TFOs (blau) ist es nahezu möglich, jede beliebige DNA-Sequenz (grau) zu adressieren. Die TFOs können sich über die große Furche der DNA winden und bilden dabei über Hoogsten-Basenpaare eine *triple helix* aus.

Mit modifizierten Oligonukleotiden ist es nahezu möglich, jede beliebige Sequenz zu adressieren und in Fusion mit einer Nuklease zu spalten (Eisenschmidt *et al.* 2005). Aufgrund der Tatsache, dass ein Hybrid-Enzym aus einer Kombination von DNA und Protein sich nicht in Zellen exprimieren lässt, sondern aufwendig hergestellt und in Zellen über „Profektion“ (Chiang *et al.* 2010, Geel *et al.* 2009) transfiziert werden muss, ist diese Klasse der Nukleasen für weitere Forschung und Anwendung im Bereich hochspezifischer Nukleasen uninteressant geworden (Pingoud and Wende 2011). Hinzu kommt die Eigenschaft, dass TFOs sehr langsam an ihre Zielsequenz binden, wodurch die Gefahr einer toxischen Spaltung über das Spaltmodul erhöht wird (Eisenschmidt *et al.* 2005).

1.3.4 TALE-Nukleasen

TALE-Nukleasen, auch TALEN genannt, sind artifizielle Fusionsproteine mit einem TALE (*transcription-activator-like effector*)-Protein als Bindungsmodul und einer Nuklease als Spaltmodul.

Die TAL-Effektoren wurden in phytopathogenen Bakterien der Gattung *Xanthomonas* entdeckt und sind in der Lage, über ein Typ-III-Sekretionssystem (Bonas *et al.* 1991) in Pflanzenzellen einzudringen (Boch and Bonas 2010). Im Wirtszellkern induzieren sie die Transkription von Wirtsgenen (Kay *et al.* 2007, Romer *et al.* 2007), wodurch die Vermehrung und Verbreitung der Bakterien begünstigt wird (Yang, Sugio and White 2006). Die befallenen Pflanzen entwickeln Nekrosen an Blättern und Stängel, können Blatt- und Stängelfäule aufweisen, und die Blätter welken rapide (Boch and Bonas 2010). Zu den Wirtspflanzen zählen viele Nutz- und Kulturpflanzen wie z.B. Reis, Paprika, Baumwolle, Soja, Banane, Zwiebel, Zitrone und Kohl (Scholze, Streubel and Boch 2011). Die ersten Beobachtungen infizierter Tomaten und Paprika wurden 1921 protokolliert (Jones *et al.* 2000). Das erste Effektorprotein AvrBs3 konnte 1989 aus *Xanthomonas campestris pv. vesicatoria* (Xcv) aufgrund einer induzierten Resistenzreaktion in Paprikapflanzen isoliert werden (Bonas, Stall and Staskawicz 1989). Im Laufe der Evolution entwickelten die Pflanzen Resistenzgene als Maßnahme gegen die eindringenden Effektorproteine. Die TAL-Effektoren können neben den Wirtsgenen auch Resistenzgene induzieren und somit eine hypersensitive Reaktion auslösen, die mit lokalem Zelltod verbunden ist. Das korrespondierende Resistenzgen in der Paprika wird als Bs3 bezeichnet und das Zielgen für den AvrBs3-Effektor als UPA (*upregulated by AvrBs3*)-Gen (Bonas *et al.* 1989, Buttner and Bonas 2010, Romer *et al.* 2007, Romer *et al.* 2009). Die meisten Namen der Effektorproteine wurden ursprünglich basierend auf ihre Eigenschaften im Wirt gegeben. Avr steht beispielsweise für Avirulenz. Der DNA-Bindungscode der TAL-Effektoren wurde erstmals im Jahr 2009 (Boch *et al.* 2009, Moscou and Bogdanove 2009) in *Xanthomonas* und später im Jahr 2013 in der Bakteriengattung *Ralstonia solanacearum* entschlüsselt (de Lange *et al.* 2013).

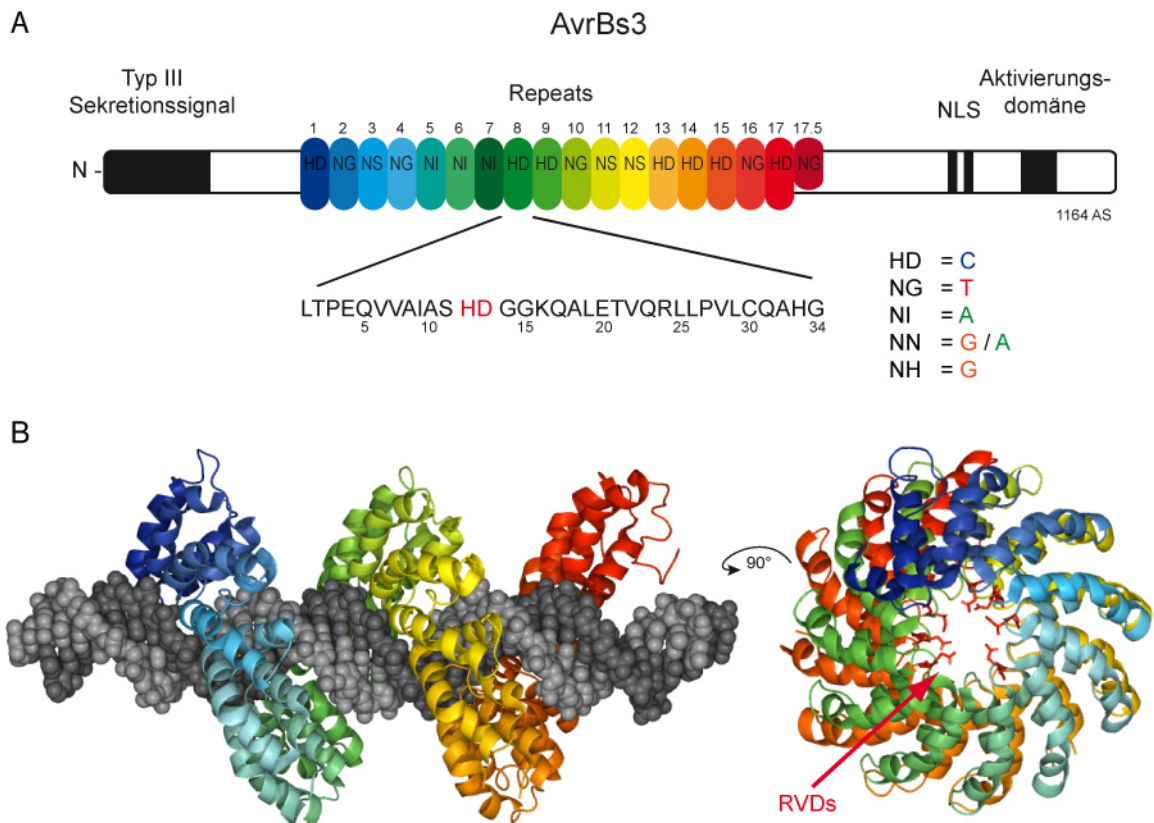


Abb. 1.6: Aufbau und Co-Kristallstruktur der TAL-Effektoren

A zeigt den schematischen Aufbau eines TALE-Proteins. Am N- und C-Terminus sind die wichtigen und konservierten Steuersignale (schwarz) für das Protein. Die *repeats* (bunt) in der Mitte zeigen die jeweiligen RVDs des AvrBs3-Proteins, die spezifisch an die DNA-Basen binden. **B** die Co-Kristallstruktur zeigt nur die *repeat*-Region des TALE-Proteins, wie es sich über die große Furche der DNA (grau) windet. Im rechten Bild sind die die Basen ererkennenden Aminosäurereste rot markiert. Die Struktur stammt vom PthXo1 TAL-Effektor (pdb: 3UGM).

TALE-Proteine besitzen konservierte Strukturen in ihren N- und C-terminalen Bereichen. Die für das Einschleusen in die Wirtszelle verantwortlichen Signale, wie das Typ-III-Sekretions- und Translokationssignal, befinden sich im N-Terminus. Der Transport in den Wirtszellkern erfolgt über das Kernlokalisierungssignal (*nuclear localisation signal*, NLS), und die Induktion der Wirtsgene wird durch die Aktivierungsdomäne eingeleitet. Sowohl das NLS-Signal als auch die Aktivierungsdomäne sind am C-Terminus lokalisiert (Abb. 1.6A). Für die Erkennung und Bindung der DNA-Sequenzen ist die zentrale Region im TALE-Protein verantwortlich. Diese ist durch konservierte, (fast) identische Sequenzwiederholungen, den sogenannten *repeats*, aufgebaut. Ein *repeat* besteht aus 34 Aminosäuren. Das Ende der Wiederholungen bildet ein sogenannter halber *repeat* mit 20 AS. Die *repeats* wiederholen sich zwischen 1,5 bis 33,5 Mal, wobei eine *repeat*-Anzahl von 15,5, 17,5 und 19,5 in der Natur dominiert (Boch and Bonas 2010). Die zwei AS-Positionen 12 und 13 in jedem *repeat* sind hypervariabel, bekannt auch als RVD (*repeat*

variable diresidue), und sind verantwortlich für die Erkennung und Bindung einzelner Nukleobasen eines Stranges der DNA-Sequenz. Die Anzahl und Kombination der *repeats* bestimmt die Bindungsspezifität der TALE-Proteine. Die einzige Limitierung stellt der *repeat* an der Position 0 dar, diese setzt immer eine Thymin-Base voraus (Boch *et al.* 2009, Doyle *et al.* 2013, Mak *et al.* 2012, Streubel *et al.* 2012, Moscou and Bogdanove 2009). Mit der Entschlüsselung des TALE-Bindungscode erkannte man das Potential der TALE-Effektoren als programmierbare Bindungsmodule für die Generierung von hochspezifischen Nukleasen. Anstelle von ZFN wurden TALE-FokI-Nukleasen hergestellt (Christian *et al.* 2010, Li *et al.* 2011, Miller *et al.* 2011, Mussolino *et al.* 2011). Die Signale im N- und C-Terminus wurden komplett deletiert, so dass nur die eigentliche Kern-*repeat*-Sequenz des Effektors genutzt wird. Durch den simplen Bindungscode, ihre Spezifität und die einfache Synthese der TALE für jede beliebige DNA-Sequenz (Briggs *et al.* 2012, Reyon *et al.* 2012, Wang *et al.* 2012, Zhang *et al.* 2011) haben sich die TALE-Nukleasen innerhalb kürzester Zeit zu den Favoriten innerhalb der Klasse der hochspezifischen Nukleasen entwickelt und stellen heute die 17 Jahre lange Forschung mit ZFN in den Schatten (DeFrancesco 2011, Bogdanove and Voytas 2011, Clark, Voytas and Ekker 2011, Mak *et al.* 2013, Perez-Pinera, Ousterout and Gersbach 2012).

1.3.5 RNA-vermittelte Nukleasen

1987 entdeckte man im *E. coli*- Genom „ungewöhnliche“ 29 Nukleotid lange Wiederholungseinheiten, die wiederum durch 32 Nukleotide separiert sind (Ishino *et al.* 1987). 2002 erkannte und analysierte man diese Muster in vielen Archaeen und Bakterien und gab ihnen erstmalig das Akronym SPIDR (*spacer interspersed direct repeat*) (Jansen *et al.* 2002). Heute sind diese Muster als das CRISPR (*clustered regularly interspaced short palindromic repeat*)-System bekannt. Die „ungewöhnlichen“ *repeats* stammen von infektiösen Viren bzw. von Fremd-DNA, die die Prokaryoten erfolgreich abwehren konnten. Mit Hilfe dieser Sequenzen wird eine adaptive Immunantwort aufgebaut und als RNA-Matrize gegen invasive DNA-Fragmente verwendet. Die kurzen RNA-Sequenzen (tracrRNA & crRNA) binden spezifisch an die Fremd-DNA und markieren diese für den Abbau durch z.B. die Cas9-Endonuklease (Horvath and Barrangou 2010, Wiedenheft, Sternberg and Doudna 2012).

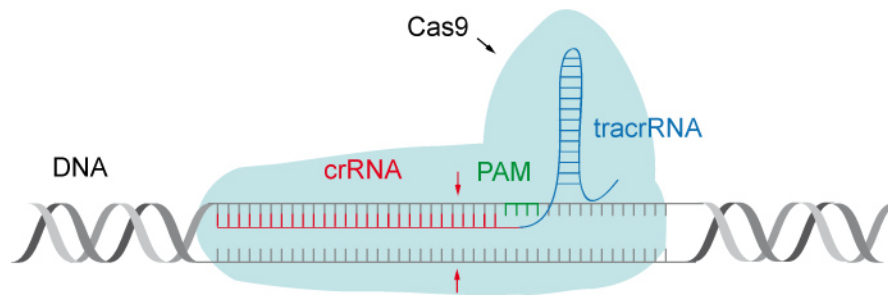


Abb. 1.7: Adressierte Spaltung mit der Cas9-Nuklease aus dem CRISPR-System

Zwei kurze RNA-Sequenzen sind zusammen mit der Cas9-Endonuklease verantwortlich für eine spezifische DNA-Spaltung (rote Pfeile). Die crRNA (rot) markiert die zu spaltende DNA-Sequenz und sorgt zusammen mit der tracrRNA (blau) für die Bindung der Cas9-Nuklease und der anschließenden Spaltung. Zum Schutz des eigenen Genoms bindet die crRNA neben der kurzen PAM-Sequenz (grün), die zur Identifizierung der infektiösen DNA beiträgt. Diese ist meist vier Nucleotide lang und stellt eine Limitierung bei der Programmierung dar.

2012 konnten mit artifiziell hergestellten chimären RNA-Matrizen DNA-Sequenzen *in vitro* gezielt adressiert und gespalten werden (Jinek *et al.* 2012). Anfang 2013 wurden erstmals Versuche in zellulären Systemen durchgeführt und die Ergebnisse mit denen von TALEN verglichen. Dabei wurde mit dem CRISPR-System zunächst eine minimal höhere Effizienz bei der Adressierung von spezifischen DNA-Sequenzen ermittelt (Cong *et al.* 2013a, Mali *et al.* 2013). Einige Monate später folgten *in vivo*-Experimente in Mausmodellen (Wang *et al.* 2013). Der Vorteil dieses neuen Ansatzes ist, dass die Adressierung auf Watson-Crick-Basenpaarungen beruht und die komplementäre Nucleinsäure unabhängig von der DNA-Modifikation immer gut an die DNA binden kann. Zudem ist die Adressierung multipler Sequenzen in einem Ansatz möglich (Cong *et al.* 2013b). Im Vergleich zur Generierung von hochspezifischen Fusionsnukleasen ist die „Programmierung“ einer Nuklease alleine durch RNA-Sequenzen einfach. Die Frage, ob die evolutiv entwickelte Spezifität des CRISPR-Systems in Prokaryoten dieselbe Effizienz in Eukaryoten aufweist, ist im Detail noch zu klären. Erste Veröffentlichungen Mitte 2013 zeigen hohe unspezifische Aktivitäten in humanen Zellen (Fu *et al.* 2013, Cradick *et al.* 2013).

1.3.6 Proteinfusionen

Die in Abb. 1.2 gezeigten Nukleasearchitekturen sind größtenteils Fusionsenzyme / Proteinfusionen, dennoch wurde eine Nuklesekategorie „Proteinfusion“ integriert, um auf abweichende Architekturen bzw. Strategien aufmerksam zu machen. So wurden beispielsweise in Kombination von MutH mit TALE-Proteinen (Gabsalilow *et al.* 2013), hochspezifische monomere Nucleasen oder in Kombination von PvuII (Fonfara *et al.*

2012) und FokI (Lippow *et al.* 2009) mit inaktivierten HEs als Bindungsmodule hochspezifische dimere Nukleasen generiert. Als Alternative zu FokI als Spaltmodul wurde die HE TevI in Kombination mit Zink-Fingern verwendet (Kleinstiver *et al.* 2012).

1.4 Spaltmodule

Die unspezifische Nukleasedomäne des Restriktionsenzym FokI ist das am meisten benutzte Spaltmodul bei der Generierung von hochspezifischen Nukleasen (Gaj, Gersbach and Barbas 2013). Die Idee einer unspezifischen Nuklease mit einem Bindungsmodul Spezifität zu verleihen, um jede beliebige DNA-Sequenz zu adressieren, stellt die bestmögliche Architektur einer hochspezifischen Fusionsnuklease dar. Aber bisher, unabhängig von den Bindungsmodulen, wurde diese Idee nicht zufriedenstellend umgesetzt. Konstrukte mit FokI als Spaltmodul neigen zu gentoxischer Aktivität (Gabriel *et al.* 2011, Pattanayak *et al.* 2011, Paques and Duchateau 2007). In der vorliegenden Arbeit wurden zwei Restriktionsenzyme als alternative Nukleasen verwendet. Durch ihre kurze Erkennungssequenz besitzen Restriktionsenzyme schon eine Spezifität, dadurch sind die Fusionsproteine nicht zu 100 % programmierbar, bzw. sind auf die Anwesenheit einer Restriktionsspaltstelle angewiesen, aber eine erhöhte Spezifität durch die Fusion mit einem Bindungsmodul könnte bei der Minderung der toxischen Aktivität eine Rolle spielen. Zum anderen kommt eine 6 bp lange Erkennungssequenz (typische Länge bei Restriktionsenzymen) alle 4096 bp im Genom vor. Für die Induktion der HR wird eine Spaltung in der Nähe der Mutation mit einem tolerierbaren Abstand von einigen 100 bis wenigen 1000 bp, abhängig vom Locus, benötigt (Paques and Duchateau 2007). Aufgrund dieser Tatsachen eignen sich Restriktionsenzyme für die Verwendung als Spaltmodul in hochspezifischen Nukleasen. Die verwendeten Restriktionsenzyme BfiI und PvuII werden in den nächsten beiden Kapiteln näher beschrieben.

1.4.1 Restriktionsenzym BfiI

Restriktionsenzyme sind unentbehrliche Werkzeuge in der Molekularbiologie und sind Bestandteil von Restriktions-Modifikationssystemen (RM), die hauptsächlich in Bakterien vorkommen. Ihre Hauptaufgabe ist die Abwehr eindringender DNA. RM-Systeme zeichnen sich durch das Zusammenspiel eines Restriktionsenzym und der dazugehörigen Methyltransferase aus. Die Methyltransferase besitzt dieselbe Erkennungssequenz und methyliert die bakterielle DNA an Cytosin (N4- oder C5-Methylierung) oder an Adenin

(N6-Methylierung), um sie, in Gegenwart des Restriktionsenzym, vor der Spaltung zu bewahren. Mit diesem System ist das eigene bakterielle Genom vor der Spaltung geschützt, wohingegen eingedrungene Fremd-DNA, welche das entsprechende Methylierungsmuster nicht aufweist, gespalten werden kann, sofern sie vorher nicht methyliert wird und ihre schädliche Wirkung entfalten kann. Restriktionsenzyme werden in vier Gruppen aufgeteilt, abhängig von Kofaktoren, Reaktionsmechanismen und der Zusammensetzung ihrer Untereinheiten. Die Typ II-Restriktionsenzyme sind im Reich der Bakterien am stärksten vertreten und sehr gut charakterisiert. Sie spalten innerhalb oder unmittelbar neben ihrer 4-8 bp langen Erkennungssequenz (Pingoud *et al.* 2005, Roberts *et al.* 2003).

Das Restriktionsenzym BfiI wurde 1998 aus dem Organismus *Bacillus firmus* S8120 isoliert und zählt zu den Typ II S (*shifted*)-Restriktionsenzymen. Es erkennt die 6 bp lange, nicht-palindrome Sequenz 5'ACTGGG(N)5/4 und spaltet diese 5 bp außerhalb im oberen und 4 bp im unteren DNA-Strang (Vitkute *et al.* 1998). Im Gegensatz zu den bis heute bekannten Restriktionsenzymen benötigt es für die Katalyse keine divalenten Metallionen (Sapranaukas *et al.* 2000). Eine weitere Besonderheit stellt der strukturelle Aufbau dar. BfiI ist ein Homodimer mit einer Größe von ca. 38 kDa pro Untereinheit. Jede Untereinheit besitzt eine DNA-bindende Domäne und bildet als Dimer eine unspezifische Nukleasedomäne mit nur einem aktiven Zentrum aus, die verantwortlich für die Spaltung der doppelsträngigen DNA ist (Lagunavicius *et al.* 2003, Sasnauskas, Halford and Siksnys 2003). Durch das HxK-Motiv im aktiven Zentrum zählt BfiI evolutionär zu den Mitgliedern der Phospholipase D-Familie (Zaremba *et al.* 2004). Mit der strukturellen Aufklärung des Enzyms im Jahre 2005 erkannte man die Aktivitätskontrolle der unspezifischen Nukleasedomäne durch einen internen Peptidlinker. Dieser ist 28 AS lang und verbindet die Bindedomäne mit der Nukleasedomäne. Erst durch die spezifische Bindung der Bindedomäne an die DNA ermöglicht der interne Peptidlinker durch eine Konformationsänderung, dass das aktive Zentrum zugänglich wird (Grazulis *et al.* 2005).

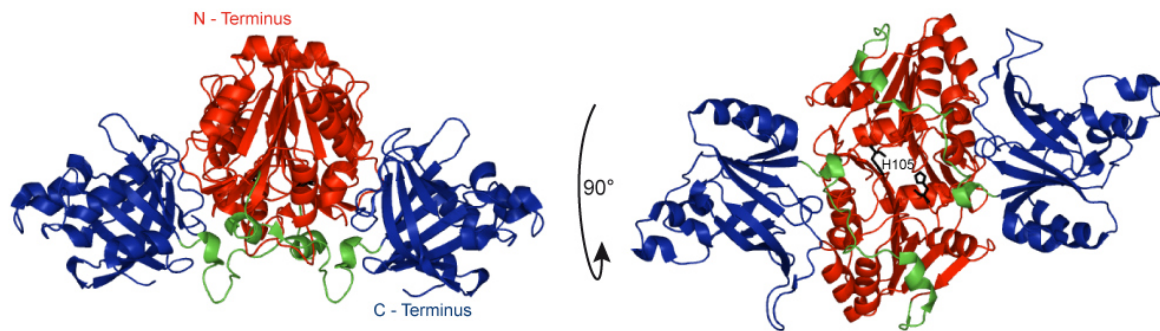


Abb. 1.8: Restriktionsenzym BfiI

In der Kristallstruktur des homodimeren BfiI-Restriktionsenzym (pdb: 2C1L) sind die DNA-Bindedomänen (AS 199-358) in blau, die Nukleasedomäne (1-170) in rot und der peptidinterne Linker (171-198) in grün abgebildet. Jede Untereinheit besitzt eine Bindedomäne und beide Untereinheiten zusammen bilden eine katalytisch aktive Nukleasedomäne aus. Im rechten Bild sind die beiden Histidin-Reste (schwarz) hervorgehoben; sie bilden das aktive Zentrum des Enzyms. Bei spezifischer Bindung der Bindedomänen ermöglichen die flexiblen Linker den Kontakt der katalytischen Domäne mit der DNA.

Der Reaktionsmechanismus der BfiI-Restriktionsendonuklease ist aufgrund des Vorhandenseins nur eines aktiven Zentrums sehr speziell. Die Spaltung der DNA erfolgt immer sequentiell mit dem Beginn der Hydrolyse der Phosphodiesterbindung im unteren DNA-Strang. Durch den Wechsel zum oberen Strang und einer Reorientierung der Nukleasedomäne wird der obere Strang gespalten (Abb. 1.9).

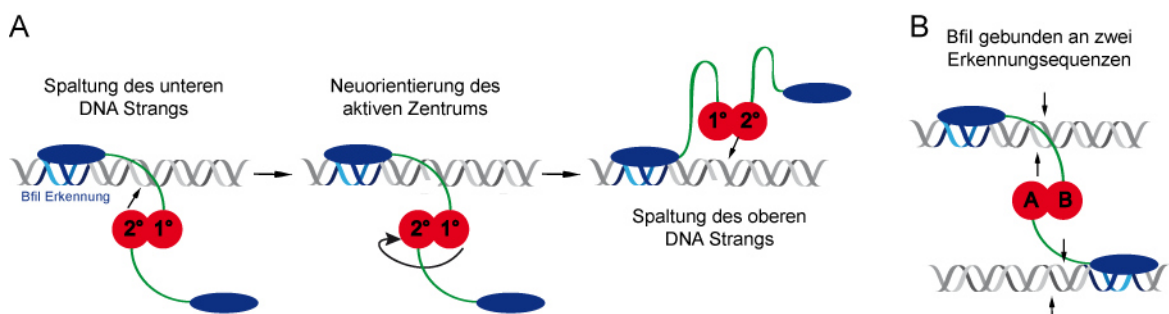


Abb. 1.9: BfiI-Reaktionsschema

A Die Bindung einer DNA-Bindedomäne (blau) reicht für einen DNA-Doppelstrangbruch aus. Zuerst wird der untere DNA-Strang gespalten und durch Neuorientierung der Nukleasedomäne (rot) erfolgt der nukleophile Angriff auf die Phosphodiesterbindung der DNA, erneut mit der gleichen katalytischen Untereinheit. **B** In *cis* oder *trans* gebundenen BfiI-Enzym ist ebenfalls in der Lage, eine DNA-Doppelstrangspaltung auszuführen (Sasnauskas *et al.* 2010).

Neben der außergewöhnlichen „Akrobatik“ auf der DNA ist die katalytische Reaktion des BfiI-Enzyms ohne Kofaktoren einzigartig. Der Histidinrest an der Position 105 in jeder Nukleaseuntereinheit bildet das aktive Zentrum. Der nukleophile Angriff des Histidins auf die Phosphatgruppe des Zuckers erfolgt immer von der gleichen Untereinheit, unabhängig davon, um welchen Strang es sich handelt (Abb. 1.10). Dadurch wird bei einem

Strangwechsel eine Neuorientierung des aktiven Zentrums erzwungen, und es entsteht ein Enzym-DNA-Intermediat. Das Histidin in der zweiten Untereinheit fungiert als Base und aktiviert ein Wassermolekül, um mit Hilfe des Hydroxidions das kovalente Enzym-DNA-Intermediat aufzulösen und den Einzelstrangbruch abzuschließen (Sasnauskas *et al.* 2010).

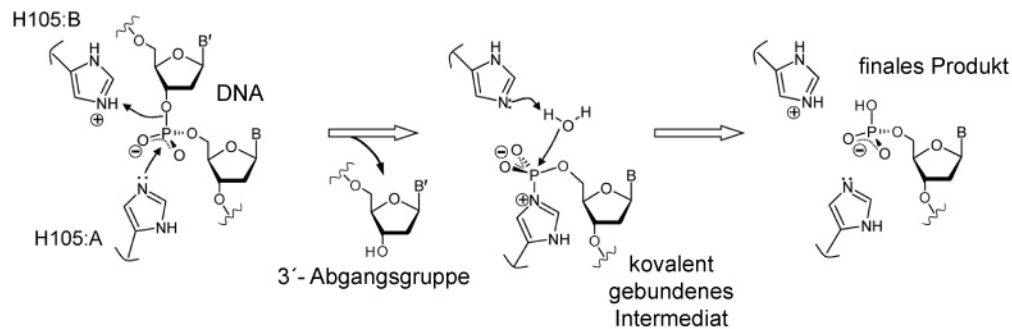


Abb. 1.10: BfiI-Reaktionsmechanismus

H105A / H105B kennzeichnen die beiden Histidin-Reste im aktiven Zentrum in der jeweiligen katalytischen Untereinheit. Durch den nukleophilen Angriff von H105A auf das Phosphat der DNA entsteht ein kovalent gebundenes Enzym-DNA-Intermediat. Mit Hilfe von H105B löst ein Hydroxidion das Intermediat auf und führt zu einem erfolgreichen Abschluss des Einzelstrangbruchs (Sasnauskas *et al.* 2010).

1.4.2 Restriktionsenzym PvuII

Das Restriktionsenzym PvuII zählt zu der Gruppe der Typ II P (*palindromic*)-Restriktionsenzyme. PvuII stammt aus dem Bakterium *Proteus vulgaris* und zählt zu den bisher kleinsten entdeckten Restriktionsenzymen mit einer monomeren Größe von 18 kDa. Als Homodimer spaltet es die palindrome Erkennungssequenz 5' CAG↓CTG 3' zwischen den Basen Guanin und Cytosin, welches zu einem Doppelstrangbruch mit glatten Enden führt. Es besitzt das konservierte „PD...(D/E)xK“-Motiv (in jeder Untereinheit) mit den entscheidenden AS D58, E68 und K70 im aktiven Zentrum (Aggarwal 1995, Pingoud and Jeltsch 1997, Cheng *et al.* 1994). Für die Katalyse werden Mg^{2+} Ionen als Kofaktoren benötigt (Dupureur and Conlan 2000, Horton and Cheng 2000, Jose, Conlan and Dupureur 1999, Spyridaki *et al.* 2003, Conlan and Dupureur 2002). Eine besondere Variante dieses Enzyms stellt die *single-chain*-PvuII (scPvuII)-Variante dar. Hier sind die beiden identischen Untereinheiten durch einen vier AS-Linker (Gly-Ser-Gly-Gly) zu einer Polypeptidkette miteinander verbunden, mit kaum merklichen Einbußen auf die Aktivität (Simoncsits *et al.* 2001, Meramveliotaki *et al.* 2007).

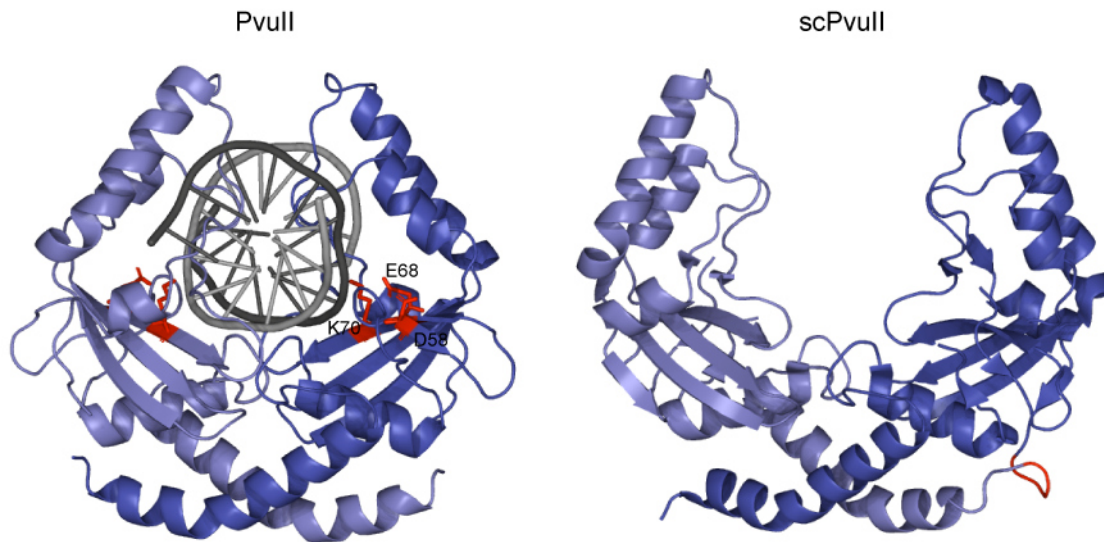


Abb. 1.11: Kristallstrukturen von PvuII und scPvuII

Die beiden identischen Untereinheiten sind in blau und hellblau hervorgehoben. Durch die Bindung an die DNA verändert sich der Winkel der Untereinheiten, so dass die katalytisch aktiven AS wirken können (rote AS; pdb: 1PVI). Bei der scPvuII-Variante (pdb: 3KSK) sind beide Untereinheiten durch einen 4 AS-Linker (rot) miteinander verbunden.

1.5 Zielgerichteter Transport der Nukleasen in die Zelle

Die Wahl der Methoden, die Nukleasen für eine Genommodifikation in die Zellen zu transportieren, hängt von mehreren Faktoren ab. Zum einen kann es von Bedeutung sein, um welche Zellen es sich handelt und ob diese leicht kultivierbar sind. Beispielsweise hätte eine *ex vivo*-Behandlung enorme Vorteile bezüglich der Selektion der „korrigierten“ Zellen, und somit die Kontrolle über unspezifische, toxische Resultate. Des Weiteren ist es wichtig, wie gut das Zielgewebe/ -organ zugänglich ist. Die Wahl hängt von der Effizienz, der Realisierbarkeit und der Frage ab, welche Konzentrationen der Nukleasen erforderlich sind. Letzteres hat eine große Bedeutung in Bezug auf zytotoxische Effekte, die im Zusammenhang mit der Nukleasekonzentration stehen (Pattanayak *et al.* 2011). Eine transiente Expression kann mit Hilfe von Plasmid-Vektoren (Hockemeyer *et al.* 2009, Holt *et al.* 2010, Zou *et al.* 2009) oder viralen Vektoren (Integrase-defiziente Lentivirale Vektoren (Cornu *et al.* 2008, Lombardo *et al.* 2007), wie Adenovirale Vektoren (Holkers *et al.* 2013, Perez *et al.* 2008), Adeno-assoziierten Vektoren (Gellhaus *et al.* 2010, Metzger *et al.* 2011, Porteus *et al.* 2003), welche nicht im Genom integrieren, erreicht werden. Durch das Einschleusen von für Nukleasen kodierende mRNA können die Expressionszeiten und Konzentrationen nochmal verkürzt und die Integration von Genen im Genom verhindert werden (Beumer *et al.* 2006, Foley *et al.* 2009, Zou *et al.* 2011). Bei der sogenannten Profektion können die Nukleasen sogar direkt als Protein verabreicht

werden, aber hierfür müssen die Proteine sehr aufwendig hergestellt und aufgereinigt werden (Gaj *et al.* 2012). Unabhängig von der Transportmethode können Nukleasen mit spezifischen Sequenzen versehen werden, wodurch sich deren Halbwertszeiten in der Zelle verringern. (Pruett-Miller *et al.* 2009).

1.6 Epigenetische Aspekte für die Genommodifikation

Variationen der genomischen Zugänglichkeit aufgrund der Chromatinstruktur ist ein weiterer Aspekt, der einen Einfluss auf die Erkennung der adressierten Sequenzen und die Effektivität der Nukleasen haben könnte.

Verschiedene Erkenntnisse in der Wissenschaft führen zu der Annahme, dass die Effizienz der Integration eines Gens über HR durch die Suppression oder Überexpression bestimmter Gene erhöht werden kann. Beispielsweise wird die HR durch die Überexpression des RAD51-Gens oder durch die Suppression der Gene EP300 oder ATF71P verbessert, welche in der Modellierung der Chromatinstruktur involviert sind (Delacote *et al.* 2011). Die Effektivität der Adressierung mit Meganukleasen konnte in Zellen, die mit Komponenten zur Lockerung der Chromatinstruktur versehen waren, signifikant gesteigert werden (Delacote *et al.* 2011). Die verminderte TALEN-Affinität durch eine 5-Methylcytosin-Methylierung in der DNA-Erkennungssequenz konnte durch eine Demethylierung aufgehoben werden (Bultmann *et al.* 2012, Valton *et al.* 2012). Als mögliche Strategie, methylierte Sequenzen mit TAL-Effektoren zu adressieren, können modifizierte *repeats* eingesetzt werden (Valton *et al.* 2012).

Trotz des erheblichen Einflusses der Chromatinstruktur sind die meisten Zellen während einer Anwendung der Nukleasen dynamischen, vom Zellzyklus abhängigen Prozessen unterworfen; somit besteht eine hohe Wahrscheinlichkeit, dass die Nukleasen an ihre spezifische Erkennungssequenz binden können. Beispielsweise wird die Zugänglichkeit der DNA während der S-Phase, aufgrund der weniger dicht gepackten Chromatinstruktur, nahezu im gesamten Genom erleichtert (Bell *et al.* 2011).

1.7 Zielsetzung

Das Ziel dieser Arbeit ist die Generierung hochspezifischer Nukleasen im Kontext eines Fusionsenzym. Um die unspezifische Aktivität der FokI-Nuklease als Spaltmodul zu vermeiden, werden zwei Strategien verfolgt. Als Alternative zu FokI werden zwei Restriktionsendonukleasen, BfiI bzw. PvuII, als Spaltmodule eingesetzt, um mit Hilfe von DNA-bindenden Proteinen ihre Spezifität zu erweitern.

Zum einen liegt der Fokus auf der Nuklease. Es soll untersucht werden, ob die allosterische Aktivierung des BfiI-Enzyms sich mit Hilfe von Bindungsproteinen kontrollieren lässt, um mit der Autoinhibition Zytotoxizität zu vermeiden. Als Bindungsmodul wird die sehr spezifisch bindende und inaktivierte Homing-Endonuklease I-SceI verwendet, um die Erkennungssequenzen des BfiI-Enzyms zu erweitern. Zudem besitzt der außergewöhnliche Reaktionsmechanismus des BfiI-Enzyms das Potenzial, durch die Inaktivierung der nativen Bindedomäne, eine 100 % programmierbare Nuklease zu generieren.

Zum anderen sollen programmierbare, hochspezifische Nukleasen in Fusion mit TALE-Proteinen generiert werden. Als Nuklease wird in diesem Zusammenhang das gut charakterisierte, spezifische Restriktionsenzym PvuII in Zusammenhang mit adressierten bi- bzw. tripartiten Substraten verwendet. Es soll gezeigt werden, ob neben der homodimeren auch eine heterodimere Architektur des Fusionsenzym möglich wäre. Des Weiteren besteht die Möglichkeit, eine monomere hochspezifische Nuklease mit *single-chain*-PvuII zu generieren. Dieses Konstrukt hätte große Vorteile im Vergleich zu den bisherigen heterodimeren FokI-Architekturen, bezüglich ihrer Größe und Spaltungsraten, sowie ihres Transports mit Hilfe von viralen Transportsystemen (Lu 2004).

2 Material und Methoden

2.1 Materialien

2.1.1 Enzyme & Marker

Die verwendeten Restriktionsenzyme zählen zu den Standardenzymen der Hersteller *New England Biolabs* (NEB) oder *Fermentas* und werden bei den jeweiligen Klonierungen aufgelistet. Sie wurden in den mitgelieferten Puffern nach Herstellerangaben verwendet.

Alle Größenstandards, die als Marker bei der gelelektrophoretischen Auftrennung von Proteinen und DNA (SDS-PAGE, Agarose-, Polyacrylamidgele) verwendet wurden, sind von der Firma *Fermentas*. Dazu zählen *PageRulerTM Unstained Protein ladder*, *pUC 8 Mix* und *GeneRulerTM 1kb DNA ladder*.

2.1.2 Kits zur Aufreinigung von DNA und Protein

Die im Folgenden aufgelisteten Kits wurden nach den Herstellerangaben verwendet.

Tabelle 1: Verwendete Kits

Kit	Hersteller	Verwendung
Wizard Plus SV Minipreps DNA Purification System	Promega (Mannheim)	Plasmidpräparation im Mini-Maßstab (1,5-7,5 µg)
Wizard Plus SV Midipreps DNA Purification System	Promega (Mannheim)	Plasmidpräparation im Midi-Maßstab (100-200 µg)
Wizard SV Gel and PCR Clean-Up System	Promega (Mannheim)	Aufreinigung von PCR-Ansätzen / Restriktionsspaltungen und DNA Aufreinigungen aus präparativen Agarosegelen
Protino Ni-TED 1000 Packed Columns	Macherey- Nagel (Düren)	Aufreinigung von His-tag-Proteinen
Protino Ni-IDA 1000 Packed Columns	Macherey- Nagel (Düren)	Aufreinigung von His-tag-Proteinen

2.1.3 Chemikalien

Tabelle 2 gibt die in dieser Arbeit verwendeten Chemikalien mit Angabe der jeweiligen Hersteller wieder. Für Verdünnungen und Lösungen wurde Reinstwasser aus dem Milli-Q Synthesis Wasseraufbereitungssystem (Millipore, Schwallbach) verwendet.

Tabelle 2: Verwendete Chemikalien

Chemikalie	Hersteller	Chemikalie	Hersteller
2-Mercaptoethanol	Merck	Kaliumacetat	Appllichem
40 % Acrylamid/Bisacrylamid 29:1	Appllichem	Kaliumchlorid	Roth
Agar	Appllichem	Kaliumhydrogenphosphat	Merck
Agarose	Invitrogen	Kanamycinsulfat	Roth
Aluminiumsulfat	Appllichem	Lubrol (Polyethylene-9-laurylether)	Sigma
Ammoniumperoxodisulfat	Merck	Magnesiumchlorid	Merck
Ammoniumsulfat	Appllichem	Methanol	Merck
Ampicillin-Natriumsalz	Appllichem	Natriumacetat	Appllichem
Bromphenolblau	Merck	Natriumcarbonat	Merck
Coomassie Brilliant Blue G 250	Appllichem	Natriumchlorid	Appllichem
DMSO	Merck	Natriumhydroxid	Merck
dNTPs	Fermentas	Natriumthiosulfat	Merck
Dithiothreitol (DTT)	Appllichem	Ni ²⁺ -NTA-Agarose	Qiagen
Ethylendiamintetraessigsäure (EDTA)	Appllichem	Phosphorsäure	Roth
Essigsäure	Roth	Phenylmethylsulfonylfluorid (PMSF)	Appllichem
Ethanol 96 %, 99,8 %	Appllichem	Saccharose	Merck
Ethidiumbromid	Roth	Salzsäure	Merck
Formaldehyd	Merck	Sodiumdodecylsulfate (SDS)	Roth
Glycerin	Appllichem	Silbernitrat	Merck
Glycin	Appllichem	Tetramethylethylendiamin (TEMED)	Appllichem
Hefeextrakt	Appllichem	Tetracyclin-Hydrochlorid	Appllichem
Imidazol	Appllichem	Tris-(hydroxymethyl)-aminomethan	Appllichem
Isopropyl-β-D-thiogalactopyranosid (IPTG)	Appllichem	Trypton	Appllichem
Isopropanol	Roth		

2.1.4 Radiochemikalien

[α - ^{32}P]-dATP wurde von der Firma Hartmann Analytic (Braunschweig) bezogen.

2.1.5 *E.coli*-Stämme

ER2566

Für die Expression von BfiI-Konstrukten wurden ER2566-Zellen zusammen mit dem pET-Vektorsystem verwendet. Genotyp: *F*- λ - *fhuA2 [lon] ompT lacZ::T7 gene 1 gal sulA11 Δ (mcrC-mrr)114::IS10 R(mcr-73::miniTn10-TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm]*.

XL10 Gold

XL10 Gold-Zellen wurden für die Expression von TALE-PvuII-Proteinen zusammen mit dem pQE30-Vektorsystem eingesetzt. Genotyp: *endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ (mcrA)183 Δ (mcrCB-hsdSMR-mrr)173 tet^R F'[proAB lacI^qZ Δ M15 Tn10(Tet^R Amy Cm^R)]*.

DH5 α

Dieser Stamm wurde für Substratklonierungen herangezogen. Genotyp: *F*- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ 80dlacZ Δ M15 Δ (lacZYA-argF)U169, hsdR17(*r_K⁻m_K⁺*), λ -.*

JM109

Der JM109-Stamm diente ebenfalls für Klonierungsarbeiten für Substratherstellungen. Genotyp: *endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB⁺ Δ (lac-proAB) e14- [F' traD36 proAB⁺ lacI^q lacZ Δ M15] hsdR17(*r_K⁻m_K⁺*)*.

2.1.6 Puffer & Medien

Die Zusammensetzung der Puffer und Medien sind den folgenden Tabellen zu entnehmen.

Tabelle 3: Puffer & Medien

Puffer / Lösungen	Zusammensetzung
1x Physiologischer Puffer	20 mM Tris-Ac, 120 mM K-Ac, 1 mM Mg ²⁺ -Ac, pH 7.5
10x TPE	0.9 M Tris, 20 mM EDTA, pH 8.2 (eingestellt mit Phosphorsäure)
10x TBE	1 M Tris, 1 M Borsäure, 25 mM EDTA, pH 8.2 (eingestellt mit NaOH)
1x STE	10 mM Tris/ HCl, 100 mM NaCl, 0.1 mM EDTA, pH 8.0
5x DNA Auftragspuffer	25 % (w/v) Saccharose, 1.2 % (w/v) SDS, 250 mM EDTA, 0.1 % (w/v) Bromphenolblau, pH 8.0 (eingestellt mit NaOH)
1x Laemli-SDS-Elektrophoresepuffer	25 mM Tris, 0.19 M Glycin, 0.1 % (w/v) SDS, pH 8.3
SDS 8%-Trenngel (8 ml)	1.6 ml 40 % Acrylamid (29:1), 2.24 ml 1.5 M Tris/HCl pH 8.8, 80 µl 10 % SDS, 4.08 ml H ₂ O
SDS 6 % Sammelgel (4 ml)	500 µl 40 % Acrylamid (29:1), 520 µl 1 M Tris/HCl pH 6.8, 40 µl 10 % SDS, 2.94 ml H ₂ O
5x Laemmli-gelaufragspuffer (5x LAP)	350 mM Tris/ HCl, 0.1 % (w/v) SDS, 6 % (v/v) 2-Mercaptoethanol, 30 % (v/v) Glycerin, 0.05 % (w/v) Bromphenolblau, pH 6.8
Kolloidales Coomassie	0.1 % (w/v) Coomassie Blue-G 250, 2 % (w/v) Phosphorsäure, 5 % (w/v) Aluminiumsulfat, 10 % (v/v) Ethanol
Aufschluss- / Waschpuffer I	30 mM KPi, 1 M NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, 0,01 % Lubrol, pH 7,8
Waschpuffer II	30 mM KPi, 500 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.01 % Lubrol, 15 mM Imidazol, 1 mM PMSF, pH 7.8
Elutionspuffer	50 mM KPi, 300 mM NaCl, 250 mM imidazole, pH 8
Dialysepuffer Bfil	30 mM KPi, 300 mM NaCl, 1 mM EDTA, 1 mM DTT, 60 % Glycerin, pH 7.8
Dialysepuffer PvuII	30 mM KPi, 550 mM KCl, 1 mM DTT, 1 mM EDTA, 60 % v/v Glycerin, pH 7.8
Heparin Niedrigsalzpuffer	30 mM KPi, 300 mM KCl, 1 mM DTT, 1 mM EDTA, 1 mM PMSF, 0.01 % Lubrol, pH 7.8
Heparin Hochsalzpuffer	Wie Heparin Niedrigsalzpuffer, aber mit 1 M KCl
Medium	Zusammensetzung
Luria Bertani	1 % (w/v) Trypton, 0.5 % (w/v) Hefeextrakt, 0.5 % (w/v) NaCl, pH 7.5

2.1.7 Chromatographie Säulen

Für die Aufreinigung der TALE-PvuII-Fusionskonstrukte wurde die Anionenaustauschersäule HiTrapTM Heparin 5 ml (GE Healthcare) verwendet.

2.1.8 Plasmide

pET-21b

Der pET-21b-Expressionsvektor dient zur Expression der BfiI-Nukleasevarianten. Die Expression wird über einen T7-Promotor gesteuert, welcher unter der Kontrolle des *Lac-Repressors* steht. BfiI besitzt einen N-terminalen His-tag und eine Ampicillinresistenz.

pACYC184 M.BfiI

Das *low copy*-Plasmid pACYC184 trägt neben der Chloramphenicolresistenz auch eine Tetracyclinresistenz und kodiert für die BfiI-Methyltransferase.

pASK45-X

Das Expressionsplasmid pASK45 wird als Substratplasmid für die BfiI-Nukleasevarianten verwendet. Der Vektor hat eine Ampicillinresistenz und einen TetracyclinPromotor. Rekombinante Proteine können von einem terminalen Strep-tag und einem terminalen His-tag flankiert werden.

pAT153

Das Plasmid pAT153 dient als Vektor für verschiedene Nukleasesubstrate in Spaltungskinetiken. Es trägt eine Ampicillinresistenz.

pLGM

Das Plasmid pLGM kodiert für die PvuII-Methyltransferase. Es handelt sich um ein *low copy*-Plasmid mit einer Kanamycinresistenz (Simoncsits *et al.* 2001).

pQE30

Dieser Expressionsvektor kodiert für die verschiedenen TALE-PvuII-Varianten und ist mit einer Ampicillinresistenz versehen. Er besitzt einen modifizierten T5-Promotor, der unter der Kontrolle des *Lac-Repressors* steht. Der auf dem Plasmid vorhandene N-terminale His-tag wurde an den C-Terminus der kodierenden Sequenz umklont.

pVAX

Expressionsvektor für eukaryotische Zellen mit einer Kanamycinresistenz zur Selektion in *E.coli*-Zellen. Die Regulation der Expression der rekombinanten Proteine erfolgt über einen CMV-Promotor.

Die verschiedenen verwendeten Antibiotika und deren eingesetzten Konzentrationen werden in der folgenden Tabelle aufgelistet.

Tabelle 4: Verwendete Antibiotika

Antibiotika	LB- Medium ($\mu\text{g} / \text{ml}$)	Agar Platte ($\mu\text{g} / \text{ml}$)
Ampicillin	75	100
Chloramphenicol	20	30
Kanamycin	25	25
Tetracyclin	10	10

2.1.9 Primer & Oligonukleotide

Tabelle 5: Fluorophor-gelabelte Oligos

Primer	Sequenz (5' \rightarrow 3')	Anmerkung
Oligo-175	CACTTAGGGATAACAGGGTAATGGTACTCAGCTGATTCATATTACCCTGTTATCCCT ACGCT	up-3 FAM-6 IPI18
Oligo-176	AGCGTAGGGATAACAGGGTAATATGAATCAGCTGAGTACCATTACCCTGTTATCCC TAAGTG	down IPI18
Oligo-223	AGCGTAGGGATAACAGGGTAATATGAATCAGCTGAGTACCATTACCCTGTTATCCC TAAGTG	down 5'Cy3 IPI18
Oligo-224	CACTTAGGGATAACAGGGTAATGGTACTCAGCTGATTCATATTACCCTGTTATCCCT ACGCT	up IPI18

Tabelle 6: Primer & Oligonukleotide

Primer	Sequenz (5' → 3')	Anmerkung
Oligo-41	TAATACGACTCACTATAGGG	T7 Prom Primer
Oligo-57	GCATGCATGCTAGCAACATCAAAAAAACCAGGTA	NheI, BfiI-I-SceI
Oligo-58	GCGCAGATCTCTCGAGTTTCAGGAAAGTTTCGGAG	XhoI, BfiI- I-SceI
Oligo-67	AATTCGGATCCTTACTCGAGGCTAGCCCAATAATCCTTCCGAATTCTTC	NheI Marker Fusion BfiI
Oligo-85	TACTGGGAGATCGTAC	BfiI Kasette up
Oligo-86	GATCTCCCAGTAGTAC	BfiI Kasette down
Oligo-100	TGTATTCTTTCAGCAGTTTAGAGTTCGGAC	BfiI Mutanten fwd
Oligo-103	TCATTGCAGCGCTAGGACCAGATGG	Entfernt BfiI, pASK
Oligo-133	TTTACCGGTTCTGCTAGCAACATCAAAAAAACCAG	BfiI-5AS-Linker
Oligo-134	TTTACCGGTCCAATAATCCTTCCGAATTC	BfiI-5AS-Linker
Oligo-183	CCGGCTCTGGTTCGGCAGCGGTT	BfiI 10AS <i>up</i>
Oligo-184	CTAGAACCGCTGCCGGAACCAGAG	BfiI 10AS <i>down</i>
Oligo-189	AGTTCGAGAAGTCGACCGGGGTTGATCTTAGAACATTGGGATA	Sall <i>Fwd</i> AvrBs3
Oligo-190	GATAACGGAGCTCGCTACTCTATGGCTGGTGCGC	AvrBs3-63AS
Oligo-191	GGATAACGGAGCTCCCTAGGCATGCTAAAGCAAC	AvrBs3-28AS
Oligo-194	ATTATTAGAGCTCTGGGGTCATATACAGG	PvuII P14G
Oligo-195	CCTGGTAGTACTGTTAATCCTGTAATTAGAAG	<i>Rev</i> zu Oligo-194
Oligo-245	AGTTACCTTCGAAGCTGAAGCCAACCTTGATTTTA	BfiI -N279A
Oligo-247	GGTCACGCTGATACTGCAGCCATAACAAGCTGCGC	BfiI-R212A
Oligo-309	ATTAACATATGCATGTTGATCTTAGAACATTGGGATA	Entfernt Strep-tag
Oligo-310	GATAACGGAGCTCCCTCAAGTGCTGGCCGTCCTCCT	Entfernt 28AS in AvrBs3
Oligo-311	CTGGGTGGGTACCTAGGCATGCTAAAGCAACGAGGT	Entfernt L in AvrBS3
Oligo-343	CATGGAATAAATAAAATATTTCAAGATAATGG	PvuII D30K + SspI
Oligo-346	CAAGATAATGGTGGAGATCTCCTTCAAGTCCTTC	PvuII K38D + BglIII
Oligo-401	ACGGGATCCCACCTCGTTGCTTTAGCATGCCTAGG	BamHI, C-Term AvrBs3
Oligo-402	TTGTACAGGGCCCTTATGATCCGTAATCTTTGTCC	ApaI, C-Term PvuII
Oligo-460	CTCCACTGCAGTTAGATACA	PstI, N-Term AvrBs3

2.1.10 Substrate

Erkennungssequenzen sind fett dargestellt und die Sequenzen für die Bindungsmodule sind zusätzlich unterstrichen.

Tabelle 7: Erkennungssequenzen der Plasmidsubstrate

Plasmidsubstrate	Erkennungssequenzen (5' → 3')
-Pvull-	CAGCTG
T3-2-P-2-T3*	TTTAT <u>TCTATAAACCTAACCCCTCT</u> GCAGCTGCG <u>AGAGGGTTAGGTTTATAGA</u> TAAACATG
T3-4-P-4-T3	TTTAT <u>TCTATAAACCTAACCCCTCT</u> CCTGCAGCTGCGGG <u>AGAGGGTTAGGTTTATAGA</u> TAAACATG
T3-6-P-6-T3	TTTAT <u>TCTATAAACCTAACCCCTCT</u> TTCCCTGCAGCTGCGGGAA <u>AGAGGGTTAGGTTTATAGA</u> TAAACATG
T3-8-P-8-T3	TTTAT <u>TCTATAAACCTAACCCCTCT</u> TATTCCTGCAGCTGCGGGAATA <u>AGAGGGTTAGGTTTATAGA</u> TAAACATG
T3-6-P	TTTAT <u>TCTATAAACCTAACCCCTCT</u> GGATCCCAGCTGTCTAGACATG
T4-6-P-6-T4	<u>TATAATTAATAATCCACTT</u> GGATCCCAGCTGTCTAGAA <u>AAGTGGATTATTAATTATACATG</u>
T3-18-T3	<u>TCTATAAACCTAACCCCTCT</u> GGATCCGATATCTCTAGAA <u>AGAGGGTTAGGTTTATAGACATG</u>
T-8-P-6-T hIL2RG	ATC <u>ATTACCATTACATCCCTCT</u> TATTCCTGCAGCTGCCCTG <u>CTGGGAGTGGGGCTGAACACGACCATG</u>
T-4-P-4-T hIL2RG	ATC <u>ATTACCATTACATCCCTCT</u> CCTGCAGCTGCCCC <u>CTGGGAGTGGGGCTGAACACGACCATG</u>
-Bfil-	TACTGGGAGATCGTAC
I-SceI-18-I-SceI	<u>TAGGGATAACAGGGTAAT</u> GGTACTCAGCTGATTCAT <u>ATTACCCTGTTATCCCTA</u>
I-SceI-24-I-SceI	<u>TAGGGATAACAGGGTAAT</u> GGTACTGAATTCAGCTGATTCAT <u>ATTACCCTGTTATCCCTA</u>
I-SceI-30-I-SceI	<u>TAGGGATAACAGGGTAAT</u> GGTACAGGATCCGGTACCCAGCTGATTCAT <u>ATTACCCTGTTATCCCTA</u>
I-SceI-0-Bfil	GCT <u>TAGGGATAACAGGGTAATACTGGG</u> AGAT
I-SceI-3-Bfil	GCT <u>TAGGGATAACAGGGTAAT</u> ACTACTGGGAGAT
I-SceI-6-Bfil	GCT <u>TAGGGATAACAGGGTAAT</u> GGTACTACTGGGAGAT
I-SceI-8-Bfil	GCT <u>TAGGGATAACAGGGTAAT</u> ATGGTACTACTGGGAGAT
I-SceI-10-Bfil	GCT <u>TAGGGATAACAGGGTAAT</u> ATAGGGTACTACTGGGAGAT
I-SceI-12-Bfil	GCT <u>TAGGGATAACAGGGTAAT</u> ATAGCTGGTACTACTGGGAGAT

Tabelle 8: Verwendete PCR-Substrate und deren zur Herstellung verwendeten Primerpaare

PCR-Substrate	verwendete Primerpaare	
T3-n-P-n-T3 / T3-6-P (~390 bp)	Primer Oligo 382 Primer Oligo 396	Fwd: 5' CCGCCCAGTCCTGCTCGCTTC 3' Rev: 5' ACCCAGAGCGCTGCCGGCAC 3'
-P- 562 bp	Primer Oligo 126 Primer Oligo 117	Fwd: 5' GTATCACGAGGCCCTTTTCGTCT 3' Rev: 5' TTAGGAAGCAGCCCAGTAGTAG 3'
-P- 900 bp	Primer Oligo 032 Primer Oligo 117	Fwd: 5' AAAACAGGAAGGCAAAATGC 3' Rev: 5' TTAGGAAGCAGCCCAGTAGTAG 3'

2.2 Methoden

Mikrobiologische Standardmethoden (Nährmedien zur Kultivierung von *E.coli*, Elektrotransformation, Herstellung kompetenter Zellen), molekularbiologische Standardmethoden (Agarose- und Polyacrylamidgelelektrophorese, Plasmid-DNA Isolierung, Bestimmung der DNA-Konzentration, Restriktionsspaltung, Ligation von DNA-Fragmenten, Polymerase-Kettenreaktion (PCR)) und biochemische Standardmethoden (SDS-Polyacrylamid-Gelelektrophorese, Coomassie- und Silberfärbung) wurden entweder nach den Herstellerangaben oder nach den *Current Protocols in Molecular Biology* und *Protein Science* durchgeführt (Ausubel *et al.* 2012, Coligan *et al.* 2012).

2.2.1 Molekularbiologische Methoden

2.2.1.1 Klonierungen

Der Transfer von Genen von einem Plasmid in einen anderen Vektor erfolgt über eine Amplifizierung der Gene per PCR, wobei gleichzeitig die gewünschten Restriktionsschnittstellen eingefügt werden oder über eine Doppel-Restriktion, bei der das gewünschte DNA-Fragment über ein präparatives Agarosegel isoliert wird. Die kompatiblen Enden der DNA-Fragmente werden über Nacht bei 16°C mit der T4-DNA-Ligase (NEB) nach Herstellerangaben ligiert. Dabei werden 100 ng Vektor mit einem 3-5 fachen molaren Überschuss an *Insert* eingesetzt. Von je 20 µl Ligationsansatz werden die *E. coli*-Zellen mit 10 µl transformiert.

2.2.1.2 Klonierung des BfiI-I-SceI-Konstruktes

Das Restriktionsenzym BfiI befindet sich innerhalb des Proteinfusionkonstruktes am N-Terminus von I-SceI. Die Gene von BfiI (Primer Oligo-41 und -67) und I-SceI (Primer Oligo-57 und -58) wurden per PCR amplifiziert und über die eingebrachte NheI-Schnittstelle miteinander fusioniert. Die für die drei letzten Aminosäuren (AS), Valin, Threonin und Phenylalanin, von BfiI kodierenden Sequenzen wurden dabei deletiert. Die ersten beiden AS (Met und Lys) von I-SceI wurden durch die NheI-Sequenzschnittstelle ersetzt, welche für Ala und Ser kodiert. Am C-Terminus von I-SceI befindet sich ein His-tag zur Aufreinigung mit Hilfe der Metall-Chelat-Affinitätschromatographie. Dieses Konstrukt bildet die Grundlage weiterer Glycin-Serin-Linkervarianten. Die NheI-Schnittstelle repräsentiert einen Linker von 2 AS. Durch die Oligos-183 und -184 wurde in die NheI-Erkennungssequenz eine Kasette von zusätzlichen 8 AS eingefügt, so dass das Konstrukt BfiI-10AS-I-SceI entsteht. Die Klonierung der BfiI-5AS-I-SceI-Variante erfolgte über eine Kreislauf-PCR. Die Oligos-133 und -134 besitzen eine AgeI-Schnittstelle an ihrem 5'-Ende. Die Ligation nach der PCR über die AgeI-Erkennung führt zu einem zusätzlichen Einbau von 3 AS zwischen den Genen von BfiI und I-SceI. Die entsprechende PCR wurde mit der Pfu-Polymerase und dem folgendem Programm durchgeführt:

1.	Initialisierung	95°C	3 min		
2.	Denaturierung	95°C	1 min	}	3 Zyklen
3.	Annealing	52°C	50 sek		
4.	Elongation	68°C	10 min		
5.	Denaturierung	95°C	1 min	}	25 Zyklen
6.	Annealing	60°C	50 sek		
7.	Elongation	68°C	10 min		
8.	Finalisierung	68°C	20 min		
9.	Pause	4°C			

10 µl des aufgereinigten PCR-Produkts wurden für den AgeI-Verdau verwendet.

2.2.1.3 Zielgerichtete Mutagenese der BfiI-Bindedomäne

Der zielgerichtete Austausch von Aminosäuren innerhalb von Enzymen resultiert aus der entsprechenden Substitution auf DNA-Ebene (Kirsch and Joly 1998). Hierfür wird ein Mutagenese-Primer gewählt, der die gewünschte Mutation und eine zusätzliche Restriktionsschnittstelle trägt, welche später als Marker dient. Eine Standard-PCR wird mit dem zu veränderndem Gen durchgeführt, so dass das PCR-Produkt eine Länge von 150-500 bp aufweist. Dieser sogenannte Megaprimer wird aufgereinigt (*Wizard SV Gel and PCR Clean-up system Kit – Promega*), per Polyacrylamid-Gelelektrophorese wird seine Größe überprüft und eine Konzentrationsbestimmung durchgeführt (*Nanodrop 1000, Peqlab*). Im zweiten Schritt wird eine PCR zusammen mit dem Megaprimer und dem BfiI-Plasmid ausgeführt. Die beiden Stränge des Megaprimers dienen als Primer und der Polymerisierungsstartpunkt im Plasmid geht in beide Richtungen, wodurch das ganze Plasmid amplifiziert wird. In einem 100 µl-Ansatz werden 40 µl (200-300 ng/µl) Megaprimer zusammen mit 2 µl Pfu-Polymerase (Eigenpräparation 5 *units*/ µl) eingesetzt. Das Programm für die entsprechende PCR mit der Pfu-Polymerase sieht wie folgt aus:

1.	Initialisierung	95°C	3 min		
2.	Denaturierung	95°C	1 min	}	9 Zyklen
3.	Annealing	60°C	50 sek		
4.	Elongation	68°C	15 min		
5.	Denaturierung	95°C	1 min	}	9 Zyklen
6.	Annealing	60°C	50 sek		
7.	Elongation	68°C	15 min		
8.	Finalisierung	68°C	20 min		
9.	Pause	4°C			

Im Anschluss wird der Ansatz der PCR mit 30 *units* DpnI für 3 h bei 37°C verdaut. Dieses Restriktionsenzym spaltet spezifisch methylierte G^mATC-Sequenzen, die nur auf dem Template-Plasmid vorkommen, da diese zuvor in *dam*⁺-*E. coli*-Zellen durch die Dam-Methylase methyliert wurden. Diese Spaltung verhindert, dass nach der Transformation die *E. coli*-Zellen einen Selektionsvorteil durch die Template-DNA ohne Mutation bekommen. *E. coli* ER2566 werden mit einem Zehntel des Ansatzes mit der BfiI-Methyltransferase transformiert, auf Chloramphenicol- und Ampicillin-haltigen Agarplatten ausgestrichen

und über Nacht bei 37°C inkubiert. Die Plasmid-DNA der gewachsenen Klone wird mit Hilfe einer PCR (Taq-Polymerase) überprüft und zusätzlich über die Spaltung der durch die Megaprimer eingeführten Restriktionsschnittstellen kontrolliert. Klone mit den eingebrachten Marker-Erkennungssequenzen werden durch die Sequenzierung anschließend verifiziert. Die Oligos 245 (N279A, BstBI) und 247 (R212A, PstI) dienen als *reverse* Primer zusammen mit dem Oligo-100 als *forward* Primer für die Megaprimer-PCR, um die Mutationen in der BfiI-Bindedomäne einzuführen.

2.2.1.4 Klonierung des TALE-PvuII-Konstruktes

Sowohl das PvuII- als auch das scPvuII-Gen, befinden sich im pQE30-Vektor C-terminal zu einem Zink-Finger (ZF)-Modul (Schierling *et al.* 2012). Dieser Vektor wird als Ausgangsplasmid für die Klonierung des TALE-PvuII-Konstruktes verwendet. Mit den Enzymen SalI und SacI wird das ZF-Modul herausgeschnitten und durch die Gene der TALE-Proteine AvrBs3 oder AvrBs4 ersetzt. Die Schnittstellen von SalI und SacI werden zuvor durch eine PCR in die TALE-Sequenz eingebracht. Das Oligo-189 bindet am 5'-Terminus und fügt die SalI-Erkennungsstelle ein. Die Klonierung über SalI führt zu einer Deletion der ersten 150 AS. Die SacI-Schnittstelle wurde am 3'-Terminus über die *reverse* Primer 190 bzw. 191 eingeführt. Mit den Primern erhält man unterschiedlich lange C-Termini der TALE-Proteine. Der Primer 190 führt zu einem 63 AS langen und der Primer 191 zu einem 28 AS langen C-Terminus (gezählt vom letzten *repeat*). Nach der Ligation befindet sich zwischen der TALE- und PvuII-Sequenz zusätzlich eine kodierende Sequenz für einen 16 AS langen Linker (L = SSVIPNRGVTKQLVKG), welcher von der SacI- und KpnI-Erkennungsstelle flankiert ist. Über diese Restriktionsschnittstellen werden weitere TALE-PvuII-Linkervarianten generiert. Das Konstrukt TALE-28-PvuII wird mit Hilfe des Oligos-311 über die KpnI-Erkennung generiert und die TALE-L-PvuII-Variante mit dem kürzesten Linker wird mit Hilfe des Oligos-310 über die SacI-Sequenz kloniert. Das Oligo-309 dient bei beiden Klonierungen als *forward* Primer und entfernt zusätzlich den Strep-tag mit Hilfe der NsiI-Erkennungssequenz, welcher sich am 5'-Ende befindet (Abb. 2.1).

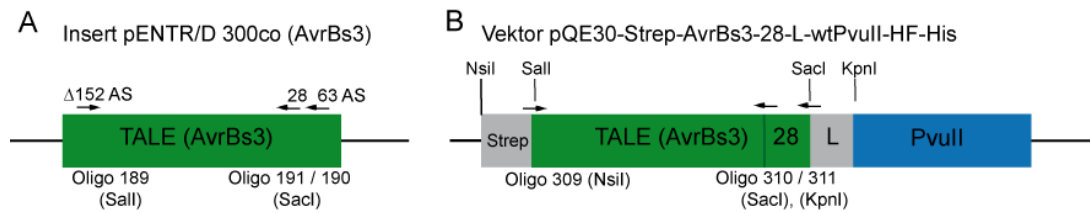


Abb. 2.1: TALE-PvuII-Klonierungsstrategie

A Das AvrBs3-Gen wird per PCR amplifiziert und mit den entsprechend eingeführten Schnittstellen in den pQE30-Vektor eingebracht. Die verwendeten Oligos verkürzen dabei das Gen am 5'- sowohl als auch am 3'-Ende. **B** Das TALE-28-L-PvuII-Fusionsenzym wird mit den Oligos-310 und -311 am Linker variiert. Das Oligo-310 generiert dabei die TALE-L-PvuII- und das Oligo-311 die TALE-28-PvuII-Variante. Mit dem *forward* Primer 309 wird mit Hilfe der NsiI-Schnittstelle der Strep-tag entfernt. L = Linker

2.2.1.5 Zielgerichtete Mutagenese des TALE-PvuII-Konstruktes

Im TALE-PvuII-Konstrukt wird eine zusätzliche, zielgerichtete Mutation (P14G) im PvuII-Gen eingefügt. Dies erfolgt nach der Anleitung in Kapitel 2.2.1.3 mit den Primern 194 (Mutagenese-Primer) und 195 unter Verwendung von *E. coli*-XL10Gold-pLGM-Zellen.

2.2.1.6 Zufall-basierte Mutagenese des TALE-PvuII-Konstruktes

Das Gen für das TALE-PvuII-Konstrukt (AvrBs3-28-L-PvuII T46G) wird in das pVAX-Vektorsystem kloniert, und steht dabei unter die Kontrolle des eukaryotischen CMV-Promotors. Ohne den Schutz der PvuII-Methyltransferase (pLGM) wird der Vektor in JM109- oder DH5 α -Zellen transformiert. Durch den CMV-Promotor oder einem anderen Promotor der sich auf dem Plasmid befindet (verantwortlich für die Kanamycinresistenz), findet eine geringe basale Expression der TALE-PvuII-Nuklease statt. Die Konzentration ist so gering, dass sie sich nicht letal auswirkt, sondern eine Anpassung der *E. coli*-Zellen an die geringe Toxizität durch Mutagenese ermöglicht (Oelgeschlager *et al.* 1990).

2.2.1.7 Klonierung der Substratplasmide

Als Grundlage für die Substratherstellung dient das Plasmid pAT153. In einer vorangegangenen Arbeit wurde bereits die Erkennungssequenz des I-SceI-Enzyms mit einem variablen Abstand am 3'-Ende von 0, 3, 6, 8, 10 und 12 bp bis zur KpnI-Erkennung eingebracht (Fonfara 2011). In die gespaltene KpnI-Sequenz wird eine Kasette mit der BfiI-Erkennungsstelle integriert (Oligos-85 und -86). Der Vektor enthält insgesamt vier BfiI-Erkennungssequenzen. Weitere Substratplasmide mit wenigen oder keinen nativen BfiI-Erkennungssequenzen werden mit dem pASK-Vektor generiert. Die bipartite Erkennungssequenz von I-SceI und BfiI mit den jeweiligen Abständen werden über die

Restriktionsenzyme SallI und NheI in den pASK-Vektor umgesetzt. Die zusätzliche BfiI-Erkennungssequenz in pASK wird mit dem Oligo-103 durch zielgerichtete Mutagenese (2.2.1.3) entfernt.

Für die TALE-PvuII-Substrate werden ebenfalls Kassetten (Kapitel 2.1.10) mit den jeweiligen Sequenzen in ein Derivat des pAT153-Plasmids über die Restriktionsschnittstellen KpnI und SphI eingeführt. Bei der Integration wird die KpnI-Schnittstelle deletiert; das erleichtert somit die Selektion positiver Klone.

2.2.2 Biochemische Methoden

2.2.2.1 Denaturierende Gelelektrophorese

Bei DNA-Einzelstrangbrüchen ist es wichtig zu unterscheiden, welcher der beiden Stränge gespalten wurde. Dazu wird eine Standard-PCR mit Fluorophor-markierten Primern mit dem jeweiligen Substratplasmid durchgeführt. Aufgrund der unterschiedlichen Laufeigenschaften in einem denaturierenden Gel (7 M Harnstoff), abhängig von der Fragmentgröße der markierten Einzelstränge, kann darauf zurückgeschlossen werden, welcher Einzelstrang gespalten wurde. Als Substrat für das BfiI-I-SceI-Konstrukt wird die I-18-I (2.1.10)-Kassette mit den Fluorophoren FAM-6 (3'-Modifikation) und Cy3 (5'-Modifikation) verwendet. Vor dem Probenauftrag wird mit den Gelen (5 % Acrylamid (19:1), 7 M Harnstoff, 0,08 % APS, 0,2 % TEMED und 1x TBE-Puffer) ein Vorlauf bei 8 Watt für 30 min durchgeführt. Die Gelelektrophorese erfolgt nach Probenauftrag bei 2-3 Watt für ca. 1 h.

2.2.2.2 Expression und Aufreinigung von BfiI-I-SceI-Varianten

Das BfiI-Restriktionsenzym würde bei einer Überexpression in *E. coli*-Zellen die bakterielle genomische DNA schneiden. Die Toxizität wird durch die Anwesenheit der BfiI-Methyltransferase (BfiI-M) verhindert, da sie die BfiI-Erkennungssequenzen methyliert und sie vor Restriktion schützt. Daher wird für das pET-Expressionssystem der *E. coli*-ER2566-Stamm, mit der bereits enthaltenen BfiI-M, verwendet. Der Vektor pACYC mit dem Gen für BfiI-M trägt eine Chloramphenicolresistenz. Durch das Expressionsplasmid haben die *E. coli*-Zellen eine zusätzliche Ampicillinresistenz.

500 ml LB-Medium mit den entsprechenden Antibiotika werden mit 10 ml einer ü.N.-Kultur angeimpft und bei 37°C inkubiert. Nach Erreichen einer Zelldichte von etwa OD₆₀₀ 0,7, erfolgt die Induktion der Proteinexpression durch Zugabe von 1 mM IPTG. Nach

weiteren 4 h bei 37°C werden die Zellen abzentrifugiert (3600 x g, 4°C, 15 min), mit 40 ml 1 x STE-Puffer gewaschen, erneut abzentrifugiert und das Zellpellet bei -20°C gelagert. Die Proteinaufreinigung erfolgt über Metall-Chelat-Affinitätschromatographie mit den Protino Ni-TED bzw. -IDA 1000 Packed Columns (Macherey & Nagel). Zunächst wird das eingefrorene Zellpellet mit Aufschlußpuffer resuspendiert, per Ultraschall (12-mal 15 sek. Sonifier 250, Branson) bei Eiskühlung aufgeschlossen, und anschließend werden die Zelltrümmer abzentrifugiert (31000 x g, 4°C, 30 min). Der Überstand wird mit sterilen Vorsatzfiltern (0.45 µm Celluloseacetat Membran, VWR) filtriert und nach Herstellerangaben über die Protino-Säulen aufgereinigt.

2.2.2.3 Expression und Aufreinigung von TALE-PvuII-Varianten

Analog zu der Aufreinigung in Kapitel 2.2.2.2 werden auch die TALE-PvuII-Konstrukte aufgereinigt. Mit dem Unterschied, dass die PvuII-Methyltransferase (PvuII-M), kodiert auf dem pLGM-Plasmid mit einer Kanamycinresistenz, in XL10Gold-Zellen zusammen mit dem Expressionsvektor der PvuII-Varianten (pQE-Vektor, Ampicillinresistenz) cotransformiert wird. Nach der Induktion mit IPTG werden die TALE-scPvuII-Konstrukte nach 4 h bei 37°C und die Zellen der dimeren Varianten nach ü.N.-Inkubation bei 23°C abzentrifugiert. Die Proteinaufreinigung erfolgt über Affinitätschromatographie mit Ni²⁺-NTA-Agarose und anschließender Ionenaustauschchromatographie über eine Heparin-Säule, Kapitel 2.2.2.4. Der Überstand nach Ultraschallaufschluss und Zentrifugation, welcher das lösliche TALE-PvuII-Protein enthält, wird für 30 min mit 5 mM Mg²⁺ und 10 *units* DNaseI (Fermentas) bei RT inkubiert. Anschließend erfolgt die Bindung an 1 ml Ni-NTA-Agarose für 1 h bei 4°C. Der Ansatz wird bei niedriger Geschwindigkeit in der Zentrifuge pelletiert (130 x g, 4°C, 5 min), und das Pellet wird je einmal mit 40 ml Waschpuffer I und II (2.1.6) gewaschen (5 min Inkubation bei 4°C) und erneut zentrifugiert. Nach dem letzten Waschritt wird der Ansatz auf eine *Poly-Prep Chromatography-Säule* (Bio-Rad) überführt. Die Elution des gebundenen Proteins erfolgt durch viermalige Zugabe von jeweils 1 ml Elutionspuffer. Die vereinten Elutionsfraktionen werden über eine HPLC aufgereinigt oder ü.N. gegen Dialysepuffer dialysiert.

2.2.2.4 Chromatographische Trennverfahren

Im Gegensatz zur Aufreinigung des BfiI-Konstrukts wurden für die PvuII-Konstrukte zwei Aufreinigungsverfahren gewählt. Neben der Metall-Chelat-Affinitätschromatographie über den His-tag wurde im anschließenden Reinigungsverfahren die Ionenaustauschchromatographie mit einer Heparin-Säule durchgeführt. Durch die reineren Proteine lassen sich die Enzym-DNA-Verhältnisse genauer berechnen und somit besser charakterisieren. Die Ionenaustauschchromatographie wurde mit einem Äktapurifier 10-FPLC Chromatographie-System (GE Healthcare) bei 4°C durchgeführt.

Mit dem Heparin-Niedrig- und Hochsalzpuffer (2.1.6) wird ein Gradient gebildet, bei dem die Elution des TALE-PvuII-Proteins bei einer KCl-Konzentration von ca. 550 mM begünstigt wird.

2.2.2.5 Bestimmung der Proteinkonzentration

Die Proteinkonzentration wird durch photometrische Messung (NanoDrop, Peqlab Biotechnologie GmbH) bestimmt. Die Extinktion bei 280 nm wird gegen den verwendeten Puffer gemessen. Die aromatischen AS Tryptophan und Tyrosin sowie die Disulfidbrücken der Cysteinreste absorbieren bei dieser Wellenlänge (Pace *et al.* 1995). Der molare Extinktionskoeffizient kann mit Hilfe der Beziehung

$$\epsilon_{\text{Protein}}^{280\text{nm}} = n \times \text{Trp} \times 5500 + n \times \text{Tyr} \times 1490 + n \times \text{Cys-S-S-Cys} \times 125 \text{ [M}^{-1} \text{ cm}^{-1}\text{]}$$

berechnet werden, wobei n die Anzahl der AS darstellt. Über das Lambert-Beersche Gesetz wird die Proteinkonzentration aus $A^{280\text{nm}}$ und $\epsilon^{280\text{nm}}$ bestimmt: $c = \frac{A^{280\text{nm}}}{\epsilon^{280\text{nm}} \cdot d}$ (d = 1 cm Schichtdicke).

2.2.2.6 Aktivitätstest der Fusionskonstrukte

Die Aktivität der verschiedenen Fusionsenzyme wird durch eine Spaltungskinetik auf den jeweiligen adressierten und nicht-adressierten Substraten durchgeführt (siehe 2.1.10). Bei Plasmid-DNA wird die Aktivität vom überspiralisierten Plasmid (*supercoil*) über das Zwischenprodukt des offenen Rings (*open circle*), bei dem ein Einzelstrang gespalten wird, zum linearen Zustand der DNA bestimmt. Das Fortschreiten der Linearisierung (oder Bildung des offenen Rings bei Nickasen) und die Abnahme der *supercoil*-DNA ist ein Maß für die Aktivität. Dazu werden Ansätze mit 8 nM DNA in einem Volumen von 15 µl und 1 x Aktivitätspuffer (Kapitel 2.1.6; physiologischer Puffer) vorbereitet. Die Reaktion wird durch die Zugabe verschiedener Enzymkonzentrationen bei 37°C gestartet. Die eingesetzten Konzentrationen sind bei den jeweiligen Experimenten im Einzelnen aufgeführt. Zum Abstoppen der Reaktion wird die Probe in ein Reaktionsgefäß mit bereits vorgelegtem DNA-Auftragspuffer (2.1.6) pipettiert und anschließend durch Agarosegelelektrophorese analysiert.

Bei Spaltungskinetiken mit kurzen PCR-Fragmenten und radioaktiv markierten Substraten werden analog 20 nM Substrat mit entsprechenden Enzymkonzentrationen eingesetzt und auf PAA-Gelen analysiert. Das relative Spaltprodukt, welches über die Quantifizierung der Bandenintensitäten erfolgt, wird anschließend ermittelt (rel. Anteil des Spaltprodukts = Intensität Spaltprodukt / Summe der Intensität aller Banden). Die Daten werden anschließend mit Hilfe der linearen Regressionsanalyse ausgewertet.

3 Ergebnisse

3.1 Hochspezifische Nuklease Bfil-I-SceI

Hochspezifische Nukleasen zeichnen sich durch ihre Genauigkeit aus, in einem komplexen Genom nur einmal zu schneiden. Dies wird durch eine lange Erkennungssequenz von mindestens 16 bp gewährleistet, denn statistisch betrachtet kommt eine Nukleotidabfolge mit 16 bp im humanen Genom nur einmal vor. Als zuverlässiges Bindungsmodul hat sich die inaktive Variante der Homing-Endonuklease I-SceI (D44N, D145A) erwiesen (Fonfara *et al.* 2012). Die Abb. 3.1.1 zeigt schematisch die Orientierung des Fusionsenzym mit den drei unterschiedlichen Linkerlängen. Der Linker besteht jeweils aus einer abwechselnden Abfolge der Aminosäuren Glycin und Serin. Die Klonierung ist unter Methoden 2.2.1.2 beschrieben.

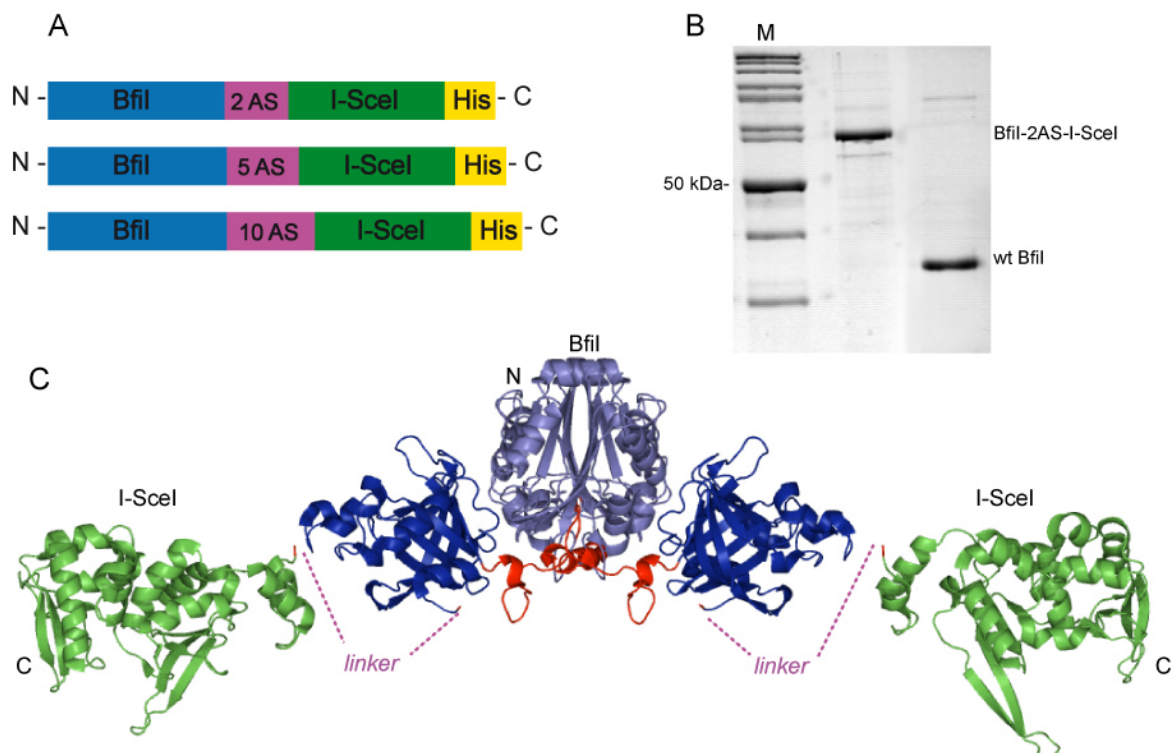


Abb. 3.1.1: Fusionsenzym Bfil-I-SceI

A I-SceI wurde mit dem C-Terminus von Bfil fusioniert. Dabei wurden drei Fusionsproteine mit drei unterschiedlich langen Glycin-Serin-Linkern hergestellt. **B** 8% SDS-PAGE nach Coomassie-Färbung, nur die Aufreinigung der Bfil-10AS-I-SceI-Variante (~68 kDa als Monomer) im Vergleich zum Wildtyp Bfil (~40 kDa als Monomer) wird gezeigt (Marker; PageRuler von Fermentas). Vom aufgereinigten Protein wurden ca. 2 µg aufgetragen. **C** Eine Modelldarstellung des Bfil (PDB code:2c11) - I-SceI (1r7m)-Fusionsenzym, generiert mit Pymol.

Die Proteinaufreinigung der Variante BfiI-10AS-I-SceI im Vergleich zum BfiI-Wildtyp (wt) wird exemplarisch für alle Konstrukte in Abb. 3.1.1 gezeigt. Von beiden Enzymen sind Proteinkristallstrukturen vorhanden (BfiI pdb code:2C1L; I-SceI: pdb code: 1R7M). Die Struktur des Fusionsproteins wurde mit Pymol modelliert (Abb. 3.1C). Die native Größe beträgt ca. 140 kDa.

3.1.1 Aktivitätstest des Fusionsproteins BfiI-I-SceI

BfiI nimmt unter den Restriktionsenzymen eine Sonderstellung ein (Kapitel 1.4.1), da es nur ein katalytisches Zentrum besitzt, mit der ein DNA-Doppelstrangbruch erzielt wird. Um zu überprüfen, ob BfiI auch im Fusionskonstrukt in der Lage ist, DNA zu spalten, wurden Aktivitätstests durchgeführt. Neben Substratplasmiden mit nur einer BfiI-Erkennungssequenz wurden zusätzlich Plasmide mit Erkennungssequenzen für I-SceI und BfiI verwendet, um die erweiterten Bindungseigenschaften des Fusionsproteins zu analysieren. Die in Abb. 3.1.1 gezeigten Kinetiken wurden mit den den Substraten in Abb. 3.1.1A durchgeführt. Diese Substrate besitzen neben der bipartiten Erkennungssequenz (I-SceI-BfiI) eine zusätzliche, 1554 bp entfernte BfiI-Erkennungstelle. Die Frage, die sich hier stellt, ist, ob das Fusionsprotein eine Präferenz gegenüber der erweiterten oder einfachen BfiI-Erkennungssequenz zeigt. Zwischen der bipartiten Zielsequenz wurden zusätzliche DNA-Basenpaare eingefügt, die als Abstandshalter (*spacer*) dienen, um den bestmöglichen *spacer* für die Bindung und Spaltung zu ermitteln. Die in Abb. 3.1.1B und C gezeigten Spaltungskinetiken mit dem Fusionsprotein und der BfiI-Wildtypvariante zeigen das Produkt der adressierten Bande bei 2300 bp (roter Pfeil). Auch die Spaltung der unspezifischen bzw. nicht-adressierten DNA ist sichtbar. Das Bandenmuster der Enzyme ähnelt sich sehr. Die übrigen Fusionsproteine BfiI-2AS-I-SceI und BfiI-5AS-I-SceI zeigten die gleichen Ergebnisse auf den jeweiligen Substraten. Diese Kinetiken stellen stellvertretend die Ergebnisse aller Varianten mit den unterschiedlichen Substraten dar.

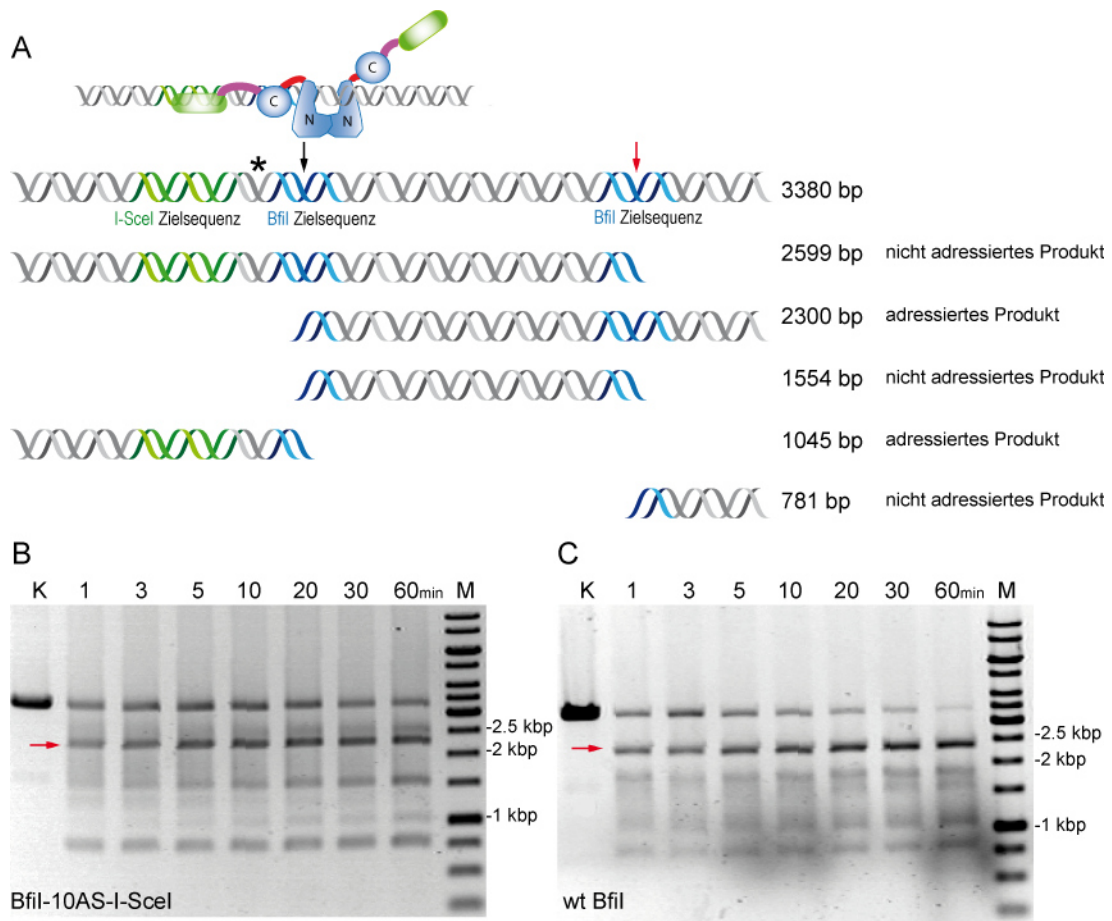


Abb. 3.1.1: Aktivitätstest BfiI-I-SceI

A Ein linearisiertes DNA-Substrat mit einer adressierten bipartiten Erkennungssequenz (I-SceI-BfiI) und einer einzelnen, nicht-adressierten, 1554 bp entfernten BfiI-Sequenz wurden verwendet, um die Präferenz des Enzyms für eine der beiden Zielsequenzen zu ermitteln. Der Stern (*) markiert die Position der unterschiedlich langen Basenpaare, die zwischen 0, 3, 8, 10, 12, 14 bp variieren. **B** zeigt die Spaltungskinetik der Variante BfiI-10AS-I-SceI unter Verwendung des linearisierten Substrates mit 8 bp Abstand (9 nM DNA; 90 nM Enzym) Der rote Pfeil zeigt das Produkt der adressierten Spaltung. **C** Im Vergleich zu **B** wird die Spaltungskinetik des BfiI-Wildtypenzym (0.7 units auf 9 nM DNA, 0 bp Abstand) gezeigt.

3.1.2 Aktivierung der Spaltung

Die BfiI-Nukleasedomäne ist eine unspezifische Nuklease. Die Spezifität wird durch die Bindungsdomänen gegeben. Um zu klären, ob die Aktivierung auch über andere Bindungsmodule stattfinden kann und nicht nur über die BfiI-Bindungsdomäne, wurde ein Plasmid als Substrat generiert, welches keine BfiI-Erkennungssequenz, aber dafür zwei I-SceI-Erkennungssequenzen besitzt, die *head-to-head* angeordnet sind, mit einem Abstand von 18 bp zueinander. Die Abb. 3.1.2 zeigt das Modell der Bindung und Spaltung des BfiI-I-SceI-Fusionsproteins auf einem Substrat ohne BfiI-Erkennungssequenz.

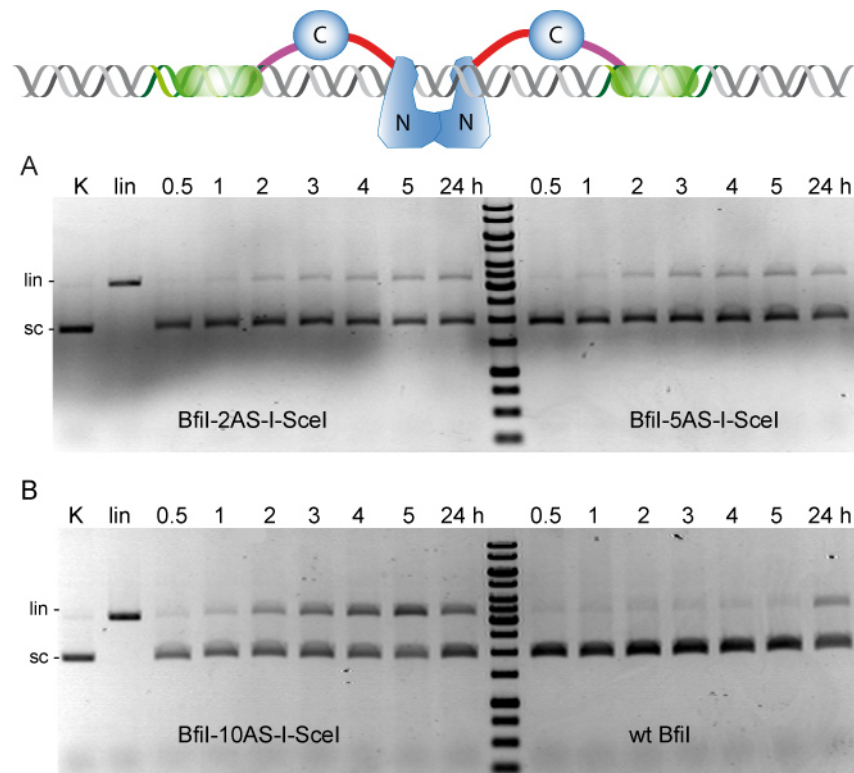


Abb. 3.1.2 BfiI-I-SceI-Spaltungskinetiken auf dem *head-to-head*-Substrat

Bei den Kinetiken wurde ein zehnfacher Enzymüberschuss gegenüber der DNA (12 nM) im Fermentas-Tango-Puffer eingesetzt (0,8% Agarosegel, Et-Br-Färbung). Als Kontrolle dient das ungeschnittene Plasmid mit der Sequenz I-SceI-18 bp-I-SceI (K). Gezeigt sind die Bandenmuster der überspiralisierte DNA (sc) und die Linearisierung (lin) mit dem Restriktionsenzym SpeI. **A** zeigt die Spaltungskinetiken auf dem adressierten *head-to-head*-Plasmidsubstrat, beginnend bei 30 min bis 24 h, für die Variante BfiI-2AS-I-SceI (links) und BfiI-5AS-I-SceI (rechts). In **B** ist links die Spaltungskinetik mit dem *head-to-head*-Substrat für die Variante BfiI-10AS-I-SceI im Vergleich rechts mit dem BfiI-Wildtyp zu sehen.

Die Varianten BfiI-2AS-I-SceI und BfiI-5AS-I-SceI zeigen eine sehr schwache Aktivität auf dem *head-to-head* orientierten Substrat mit 18 bp Abstand zwischen den beiden I-SceI-Erkennungssequenzen (Abb. 3.1.2A). BfiI-10AS-I-SceI hingegen spaltet das Plasmid deutlich schneller und wandelt mehr überspiralisierte DNA (sc) in lineare DNA (lin) um (Abb. 3.1.2B). Das Substrat wird zu ca. 50 % innerhalb von 5 h umgesetzt. Die Kontrolle mit dem BfiI-Wildtypenzym zeigt keine Aktivität, außer bei dem 24 h-Zeitwert. Dies ist auf die lange Inkubationszeit zurückzuführen, bei der unspezifische Reaktionen stattfinden können. Native Restriktionsenzyme spalten ihre Erkennungssequenz innerhalb von Minuten. Die Ergebnisse bestätigen die Strategie, dass eine Spaltung mit BfiI als Spaltmodul auch ohne BfiI-Erkennungssequenz induziert werden kann. Um mehr über die Aktivierung zu erfahren, wurden weitere Kontroll-Plasmide hergestellt. Zunächst wurde überprüft, ob eine Initiierung der Spaltung durch BfiI im Fusionsansatz auch mit nur einer I-SceI-Erkennungssequenz möglich ist (Abb. 3.1.3).

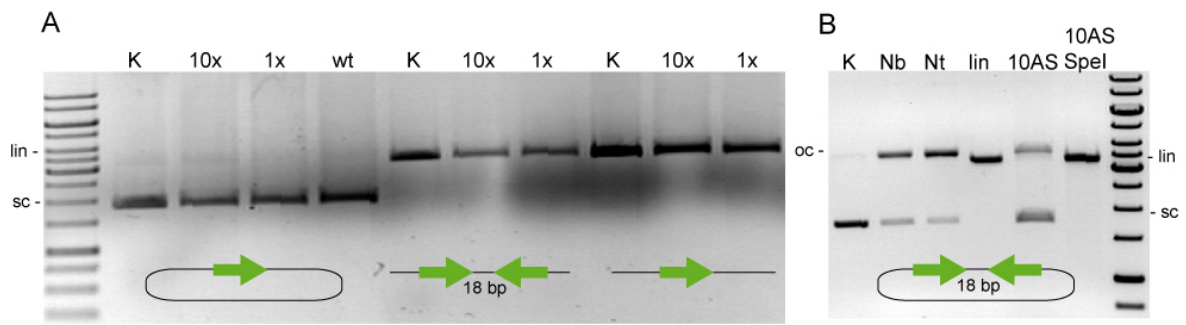


Abb. 3.1.3: BfiI-10AS-I-SceI-Aktivität auf Kontroll-Substraten

A Das Enzym BfiI-10AS-I-SceI wurde auf Substraten (9 nM) mit nur einer oder zwei I-SceI-Erkennungssequenzen (grüner Pfeil) mit einem einfachen oder zehnfachen Überschuss an Enzym bei 37°C inkubiert (0,8 % Agarosegel, Et-Br-Färbung). Die Substrate wurden entweder als überspiralisierte DNA- *supercoil* (sc)-Plasmide eingesetzt oder als linearisierte (lin) DNA. Der BfiI-Wildtyp (wt) wurde mit einer *unit* an Enzym (Fermentas) eingesetzt. **B** Zur Unterscheidung zwischen Einzel- und Doppelstrangbruch der Plasmid-DNA wurden die Nickasen Nb BpU10I und Nt BpU10I mit je einer *unit* auf der Plasmid-DNA mit zwei I-SceI-Sequenzen inkubiert. Die DNA weist ein Bandenmuster auf der Höhe von offenem Zirkel (oc) auf und liegt über der linearisierten DNA, die mit dem Restriktionsenzym SpeI (1 *unit*) linearisiert wurde. Die vorletzte Spalte zeigt die Inkubation mit dem Fusionsenzym BfiI-10AS-I-SceI (10AS) und anschließender Spaltung mit SpeI.

Ein DNA-Substrat mit nur einer I-SceI-Erkennung wird weder als scPlasmid, noch als lineares Substrat von der BfiI-10AS-I-SceI-Variante umgesetzt. Dies zeigt, dass eine *cis*-Erkennung bevorzugt wird, aber selbst das lineare Substrat mit zwei I-SceI-Erkennungssequenzen in *cis* wird nicht durchgespalten (Abb. 3.1.3A). Um zu kontrollieren, ob es sich um einen Einzelstrangbruch handelt, wurde das Plasmid I-SceI-18 bp-I-SceI neben der ungeschnittenen und linearisierten Kontrolle zusätzlich mit den Enzymen NtBp10I und NbBp10I inkubiert, siehe Abb. 3.1.2. Diese führen einen DNA-Einzelstrangbruch ein, welcher auf dem Agarosegel zu einem langsameren DNA-Laufverhalten führt als bei linearisierter DNA (Doppelstrangbruch). Die Abb. 3.1.3 zeigt, dass das Fusionsprotein einen Einzelstrangbruch (*nick*) einführt. Eine Nachspaltung mit SpeI führte zu keinem „DNA-Schmier“, welches auf einen spezifischen DNA-Einzelstrangbruch hindeutet.

Weitere I-SceI *cis*-Plasmidsubstrate wurden generiert, um die Präferenz des bevorzugten Abstands zu ermitteln. Ein Abstand von 18 bp wird gegenüber 24 bp und 30 bp bevorzugt (Abb. 3.1.4).

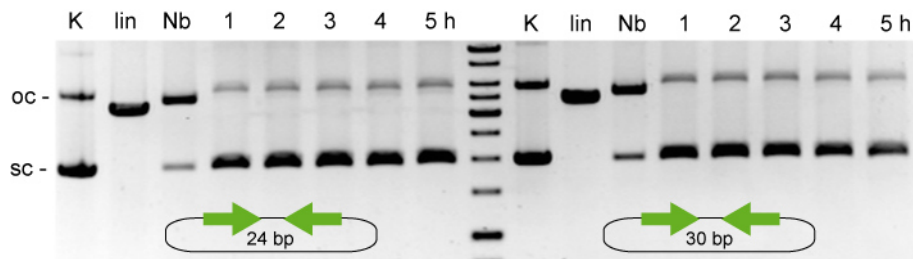


Abb. 3.1.4: BfiI-10AS-I-SceI-Kinetiken auf größeren I-SceI-cis-Abstandsplasmiden

BfiI-10AS-I-SceI wurde auf den *head-to-head* orientierten I-SceI-Substraten (9 nM) in einem zehnfachen Überschuss inkubiert. Der Abstand zwischen den beiden I-SceI-Sequenzen (grüner Pfeil) beträgt 24 bzw. 30 bp. Als Kontrolle dient das ungeschnittene (K), das linearisierte (lin) und das genickte (Nb) Plasmid. Die Kinetiken wurden im Fermentas-Tango-Puffer von 30 min bis 5 h bei 37°C durchgeführt (0.8% Agarosegel, Et-Br-Färbung).

3.1.3 Mutagenese der BfiI-Bindedomäne

Das dimere Restriktionsenzym BfiI erkennt die Erkennungssequenz 5'-ACTGGG(N)4/5-3' und spaltet 4 bzw. 5 bp vom 3'-Ende der Erkennungssequenz die DNA-Stränge. Die Bindung an die DNA erfolgt über die Bindedomäne. Das Ziel der Mutagenese innerhalb der BfiI-Domäne ist, die Bindung nur über das Bindungsmodul I-SceI zu dirigieren und DNA-Spaltung von BfiI-Erkennungssequenzen zu unterdrücken. Hierzu wurden zwei Aminosäuren innerhalb der Bindungsdomäne ausgesucht und durch Alanin ersetzt. Informationen bezüglich dieser Positionen erhielten wir von unserem Kooperationspartner Prof. Dr. Virginijus Siksnys aus Vilnius. Bei einem Ansatz, die Erkennung von BfiI zu verändern, was nicht zum gewünschten Ergebnis führte, wurde man auf diese Mutationen aufmerksam. Die Ergebnisse wurden nicht veröffentlicht.

Abb. 3.1.5 zeigt die Positionen der Aminosäuren, an denen die Mutationen R212A und N279A in der jeweiligen BfiI-Bindedomäne eingeführt wurden.

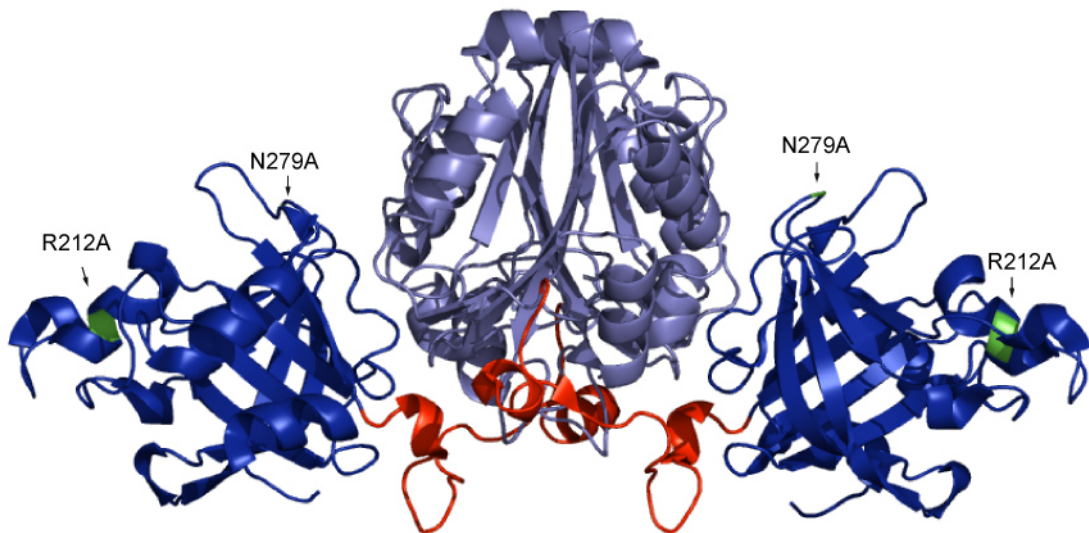


Abb. 3.1.5: Mutagenese der BfiI-Bindedomäne

Die BfiI-Nukleasedomäne ist in hellblau und die Bindedomäne in dunkelblau dargestellt. Der Peptidlinker ist in rot gezeigt. Die Mutationen R212A und N279A in der Bindedomäne sind grün hervorgehoben.

3.1.4 Aktivitätstest der BfiI-Einzelmutanten

Die Kontrolle der Bindung und Spaltung der BfiI-Einzelmutanten R212A und N279A erfolgte in einem Aktivitätstest unter Verwendung eines Substratplasmids mit nur einer BfiI-Erkennungssequenz. Beide Einzelmutanten zeigen im Vergleich zu den Kontrollen keine Aktivität (Abb. 3.1.6).

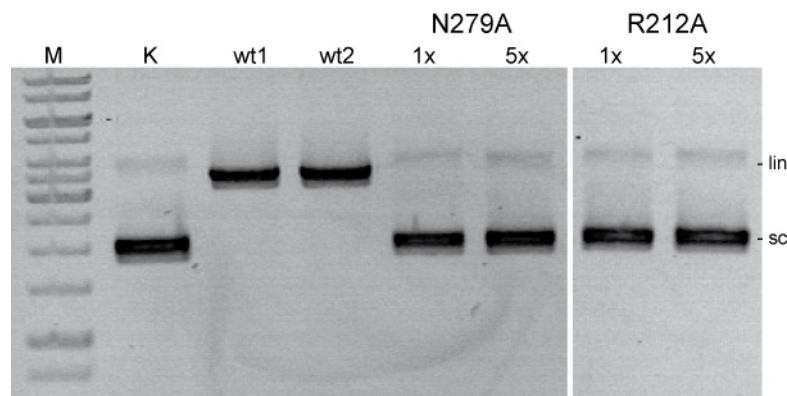


Abb. 3.1.6: Aktivitätstest der BfiI-Einzelmutanten N279A & R212A¹

Der Aktivitätstest erfolgte für 1 h bei 37°C im Fermentas-Tango-Puffer auf dem pASK 45+-Plasmid mit nur einer BfiI-Erkennungssequenz. Es wurden 7 nM DNA eingesetzt bei einem equimolaren (1 x) Verhältnis oder fünffachen (5 x) Enzymüberschuss. Als Kontrolle dient das Plasmid im ungeschnittenen Zustand (K) und gespalten mit kommerziell erworbenen BfiI-Restriktionsenzym von Fermentas (wt1) und der selbst aufgereinigten Wildtyp-BfiI-Variante (wt2). Aufgetragen ist der 1kb Marker von Fermentas.

¹ Die Mutagenese und der Aktivitätstest der BfiI-Einzelmutanten wurden von Tim Schneider während seiner Bachelorarbeit unter meiner Betreuung, durchgeführt.

Die Mutagenese wurde auch am Fusionsprotein durchgeführt, um zu überprüfen, ob die Bindung und Spaltung nur über das Bindungsmodul (I-SceI) gesteuert werden kann. Die Ergebnisse aus Abb. 3.1.2 konnten bestätigt und wiederholt werden. Die BfiI-Mutanten sind in der Lage an dem adressierten Substrat (keine BfiI-Zielsequenz) nur mit Hilfe des Bindungsmoduls I-SceI einen DNA-Einzelstrangbruch einzuführen. Die Spaltungskinetiken sind in der Abb. 3.1.7 abgebildet.

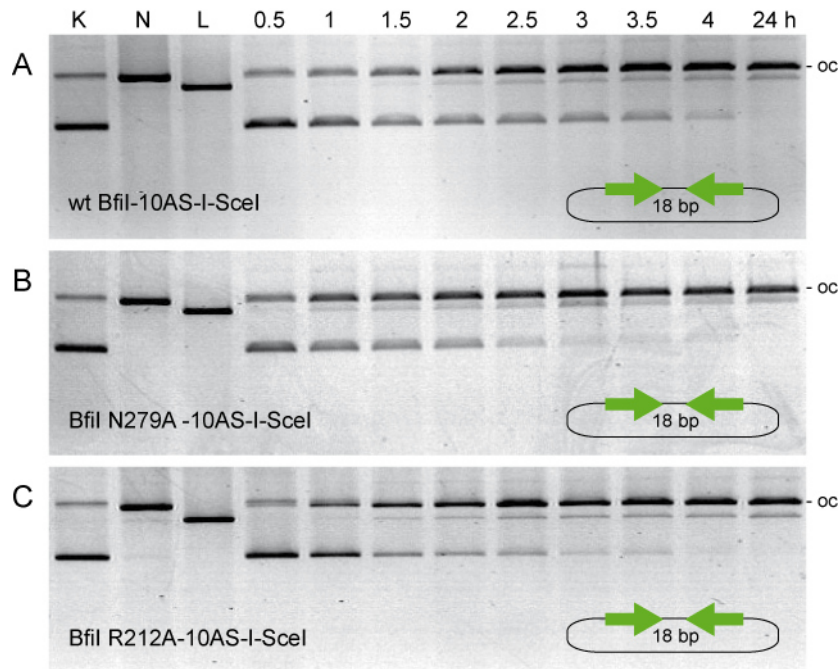


Abb. 3.1.7: Kinetiken der Einzelmutanten im Fusionsansatz²

Die Kinetiken wurden mit dem Substratplasmid I-SceI18 bp-I-SceI im Fermentas-Tango-Puffer durchgeführt. Die Konzentration der DNA beträgt 9 nM. Das jeweilige Enzym wurde mit einem zehnfachen Überschuss eingesetzt. K steht für die ungeschnittene DNA-Kontrolle, N für die *nicking*-Kontrolle (Bande läuft auf der Höhe des offenen Zirkels (oc)) mit dem Enzym Bpu10I (eine *unit*) und L für das linearisierte Plasmid mit einer *unit* des Restriktionsenzym SpeI. **A** Die Kinetik für das Fusionsenzym mit dem Wildtyp-BfiI, **B** der N279A-Variante und **C** der R212A-Variante wurde für einen Zeitraum von 30 min bis über Nacht durchgeführt.

² Die Mutagenese und der Aktivitätstest der BfiI-Einzelmutanten wurden von Tim Schneider, während seiner Bachelorarbeit unter meiner Betreuung, durchgeführt.

3.1.5 Spaltung fluoreszenz-markierter Oligos

Das Substrat mit den *head-to-head* I-SceI-Erkennungssequenzen wird von dem BfiI-SceI-Konstrukt nur in einem DNA-Strang gespalten. Um zu klären, welcher der beiden DNA-Stränge gespalten wird, wurden fluoreszenz-markierte Oligos verwendet. Der obere Strang wurde mit Cy3 (rot) und der untere Strang mit Fam6 (grün) markiert. Beide Oligos wurden einzeln mit einem nicht markierten Gegenstrang verwendet und auf einem PAA-Harnstoffgel aufgetragen. Abb. 3.1.8 zeigt, dass das Fusionsenzym den unteren Strang mehrmals spaltet.

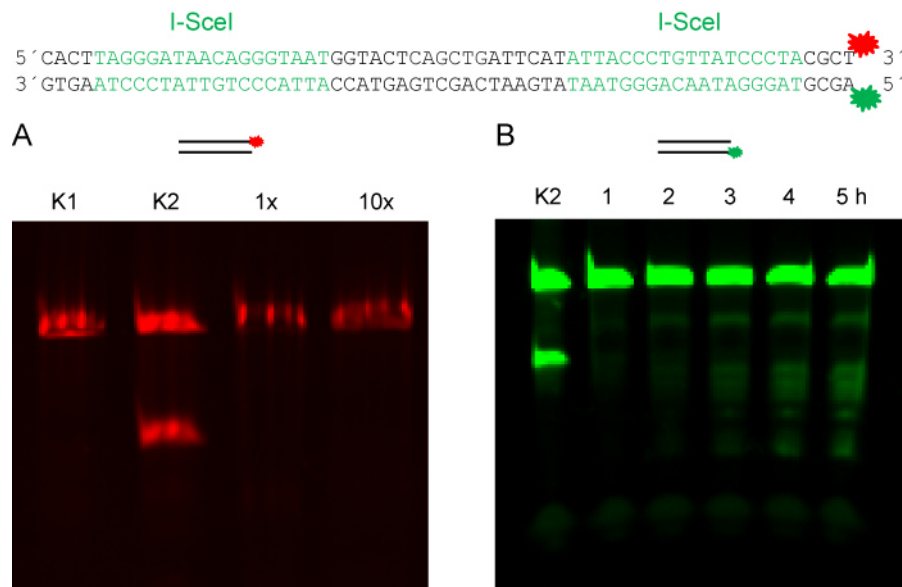


Abb. 3.1.8: Spaltung der fluoreszenz-markierten Oligos

Die 64 bp langen adressierten Oligos mit *head-to-head* orientierten I-SceI-Erkennungssequenzen mit 18 bp Abstand zueinander wurden als Substrat verwendet. Aufgetragen wurden die Proben auf ein 17 % PAA-Harnstoffgel (7 M). **A** Der obere Strang wurde mit FAM-6 (rot) am 3'-Ende markiert und zusammen mit nicht markiertem Gegenstrang bei 37°C für 5 h inkubiert (Oligos-175 und -176). K1 entspricht der Kontrolle nur mit dem Substrat, K2 ist das Substrat ungeschnitten und geschnitten mit PvuII (PvuII-Erkennungssequenz (CAGCTG) liegt zentral auf den Oligos). Das Fusionsenzym wurde in einem ein- bzw. zehnfachen (1x, 10x) Überschuss gegenüber dem Substrat eingesetzt. **B** Der untere markierte DNA-Strang am 5'-Ende mit Cy3 (grün) wurde zusammen mit nicht-markiertem Gegenstrang inkubiert (Oligos-223 und -224). Das Fusionsenzym wurde bei der Spaltungskinetik in einem zehnfachen Überschuss eingesetzt.

3.2 Hochspezifische Nuklease TALE-PvuII-Homodimer

Im vorigen Kapitel wurde die inaktive Homing-Endonuklease I-SceI als Bindungsmodul an der Nuklease BfiI verwendet. Der entscheidende Nachteil von I-SceI ist, dass es sich nicht programmieren lässt, d.h. die Zielsequenz ist nicht frei wählbar und somit limitierend bei der Adressierung genomischer Sequenzen. Eine hochspezifische Nuklease sollte jedoch neben ihrer hohen Spezifität auch vielseitig anwendbar sein. Eine interessante Alternative zu programmierbaren Zink-Finger-Modulen (Kapitel 1.3.1), bieten die *transcription activator like effector* (TALE)-Proteine (Kapitel 1.3.4). In ihrer repetitiven Kernstruktur, die 1,5 bis 30,5 *repeats* enthalten kann, sind die AS an den Positionen 12 und 13 eines jeden *repeats* entscheidend an der Erkennung der DNA-Sequenz beteiligt (Mak *et al.* 2012; Boch *et al.* 2009). Eine Substitution des I-SceI-Enzyms durch TALE im BfiI-I-SceI-Konstrukt ist nicht möglich, denn aufgrund der TALE-Struktur schien es von Vorteil zu sein, eine C-terminale Fusion am TAL-Effektorprotein durchzuführen und die ursprüngliche Aktivierungsdomäne durch eine Nuklease zu ersetzen. In dieser Arbeit wurde hierfür, aufgrund der positiven Erfahrungen mit PvuII innerhalb des Instituts für Biochemie, das Restriktionsenzym PvuII gewählt. PvuII wurde bereits erfolgreich mit den Bindungsmodulen TFOs (*triple-helix-forming oligos*) (Eisenschmidt *et al.* 2005), Zink-Finger (Schierling *et al.* 2012) und inaktiven Homing-Endonukleasen (Fonfara *et al.* 2012) fusioniert. Auf diesem Hintergrund lassen sich die Bindungs- und Spaltungseigenschaften genauer vergleichen und studieren.

Für die Fusion von PvuII und TALE wurde eine verkürzte Variante des AvrBs3-TALE-Proteins verwendet, an dessen N-Terminus die ersten 152 AS deletiert sind. Bezüglich des C-Terminus wurden zwei unterschiedlich lange Varianten hergestellt. Angelehnt an die Veröffentlichungen von Miller (Miller *et al.* 2011) (Abb. 3.2.1) ist das Protein ausgehend vom letzten *repeat* ab der 28. AS bzw. 63. AS verkürzt, wodurch lediglich die DNA-Bindungseigenschaften des Proteins zu Nutze gemacht werden. Diese wurden an das native PvuII-Enzym und an die *single-chain*-Variante fusioniert (Meramveliotaki *et al.* 2007). Mit den PvuII-Varianten erfolgte eine Integration des 16 AS langen FokI-Linkers (L). In beiden Fällen ist in PvuII die T46G-Mutation vorhanden, die auch als PvuII HF (*high fidelity*) von der Firma New England Biolabs (NEB) bekannt ist (Zhu 2009). Diese Mutation bewirkt eine geringere Affinität zur DNA und reduziert die Star-Aktivität. Die Abb. 3.2.1B zeigt, dass die *single-chain*-Variante nur ein TALE-Bindungsmodul besitzt und als monomere Nuklease fungiert. Ein Modell der homodimeren Variante ist in Abb.

3.2.1C zeigt. Hierfür wurden die Kristallstrukturen beider Enzyme mit Hilfe von Pymol zu einem Fusionskonstrukt samt DNA modelliert.

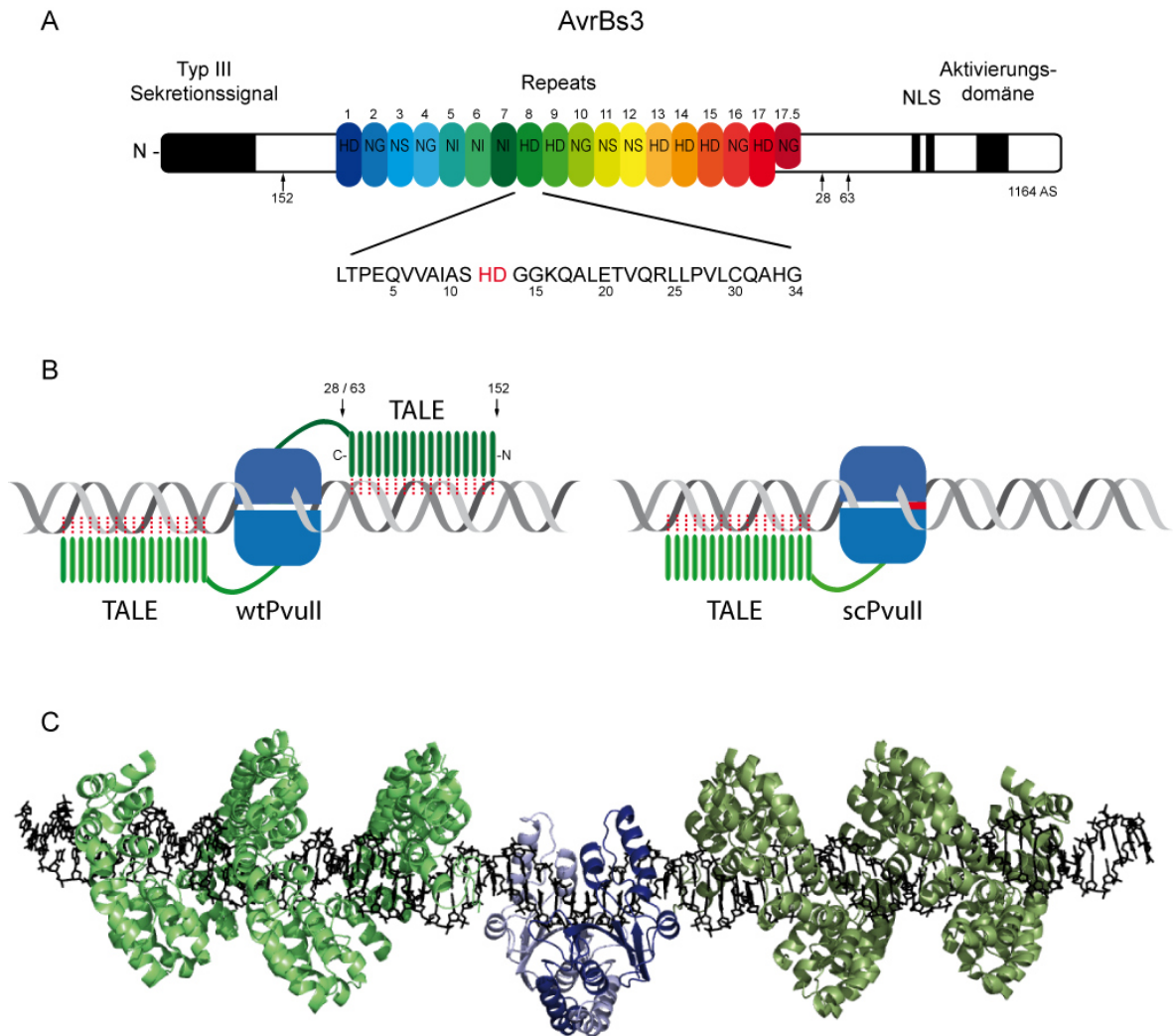


Abb. 3.2.1: AvrBs3-PvuII-Fusionsprotein

A Schema des nativen AvrBs3-TALE-Proteins. Für die Fusion mit PvuII wurden am N-Terminus die ersten 152 AS und am C-Terminus vom letzten *repeat* ausgehend die restlichen AS nach der 28. bzw. 63. AS deletiert. Jeder *repeat* (in Farbe dargestellt) besteht aus 34 AS und erkennt spezifisch eine DNA-Base mit den AS an den Positionen 12 und 13. **B** Schematische Darstellung der Fusionsproteine mit dem dimeren PvuII-Enzym (links, wtPvuII) und der PvuII *single-chain*-Variante (rechts, scPvuII). Die single-chain-Variante ist in der Lage als monomeres Enzym zu binden und zu spalten. **C** Mit Hilfe der Kristallstrukturen beider Enzyme wurde das dimere Fusionsprotein modelliert. Es ist ca. 203 kDa groß.

3.2.1 Aktivitätstest des Fusionsproteins TALE-PvuII

Als Ausgangsvektor für die Klonierung wurde das Zink-Finger-PvuII-Fusionskonstrukt aus der Arbeit von Benno Schierling verwendet (Schierling *et al.* 2012). Die Zink-Finger-Module wurden von der Studentin Martina Saling durch das TALE-Gen ersetzt. Die Aufreinigung der TALE-PvuII-Fusionsproteine erfolgte in zwei Schritten. Nach einer Affinitätschromatographie mit Hilfe des *His-tags* am C-Terminus des PvuII-Enzyms wurden die Proteine im zweiten Schritt über eine Ionenaustauschchromatographie isoliert. Abb. 3.2.2 zeigt das Resultat der Aufreinigungen.

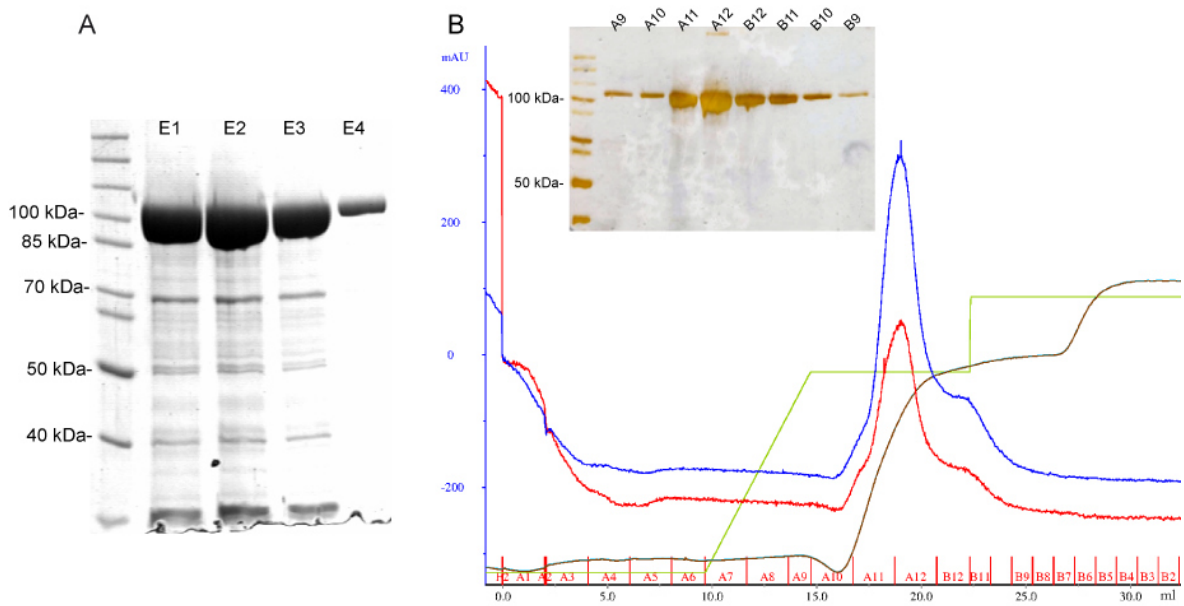


Abb. 3.2.2: Proteinaufreinigung des TALE-PvuII-Konstrukts

A 8 % SDS-PAGE mit Fraktionen der affinitätschromatischen Aufreinigung über den *His-tag*. Von den Eluaten eins bis vier wurden folgende Mengen aufgetragen: 11, 13, 6, 2 μg Protein. **B** Chromatogramm der Ionenaustauschchromatographie (Heparin). Jeweils 10 μl der Fraktionen A9 bis B9 wurden auf ein 8 % SDS-PAGE aufgetragen und mit anschließender Silberfärbung nachgewiesen (Proteinmengen A11 bis B11; 2, 4,4, 1,7, 0,9 μg). Absorptionen bei 280 (blaue Linie) und 260 nm (rot) wurden gemessen. Der eingestellte Salzgradient (grün) und der gemessene Gradient (braun) führen zur Elution des Proteins bei ca. 500 mM KCl.

Die Aktivität der Fusionsenzyme wurde mit Hilfe einer Spaltungskinetik auf einem tripartiten adressierten Substrat durchgeführt. Das Substratplasmid enthält die Erkennungssequenz für das AvrBs3-TALE-Protein und für das PvuII-Restriktionsenzym mit je 6 bp Abstand zwischen den Erkennungssequenzen (TALE-6 bp-PvuII-6 bp-TALE). Zusätzlich zum adressierten Substrat wurde ein nicht-adressiertes Plasmid eingesetzt, welches nur eine PvuII-Erkennungssequenz besitzt. Im Idealfall sollte also nur die adressierte Sequenz gespalten werden.

Die Abb. 3.2.3A zeigt Spaltungskinetiken mit TALE-*single-chain* (sc)-PvuII-Varianten im Niedrigsalzpuffer (Ionenstärke ca. 76 mM). Die oben dargestellten Gele zeigen die Spaltung auf dem adressierten Substrat. Innerhalb von Minuten wird das Substrat vollständig, sowohl von dem TALE-scPvuII-Konstrukt mit dem langen Linker (63 AS) als auch von der Variante mit dem kürzeren Linker (28 AS) umgesetzt.

Der Unterschied beider Linkervarianten wird in den unten dargestellten Gelen mit dem nicht adressierten Substrat deutlich. Während die lange Linkervariante (63 AS) das Plasmid innerhalb einer Stunde vollständig umsetzt, ist bei der kürzeren Variante (28 AS) nur eine geringfügige Zunahme des Doppelstrangbruchs (lineare Bande) zu beobachten. Im Vergleich dazu zeigt Abb. 3.2.3B die Spaltungskinetik mit der homodimeren TALE-28-L-*wt*PvuII-Variante. Wie bereits bei dem TALE-28-L-scPvuII-Konstrukt wird das adressierte Substrat innerhalb von Minuten gespalten, während auf dem nicht-adressierten Substrat keine Aktivität zu beobachten ist.

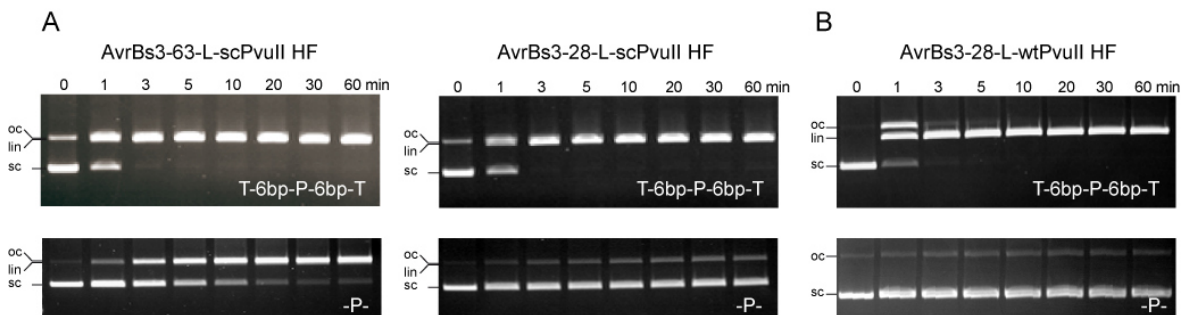


Abb. 3.2.3: Vergleich der Linkervarianten

A Vergleich des AvrBs3-63-L-scPvuII-Konstrukts mit der 28 AS scPvuII-Variante im äquimolaren Verhältnis (8 nM) auf einem adressierten, tripartiten Plasmidsubstrat (oben) und nicht-adressiertem Plasmid mit nur einer PvuII-Erkennungssequenz (unten) auf einem 0.8 % Agarose-Et-Br-Gel. **B** zeigt die Spaltungskinetik mit der dimeren Fusionsvariante und dem 28er Linker in einem 1.25 % Agarose-Et-Br-Gel. Die Spaltungskinetiken wurden im Niedrigsalzpuffer durchgeführt, mit einer Ionenstärke von ca. 76 mM (20 mM Tris-Ac, 50 mM K-Ac, 2 mM Mg-Ac, pH 7.5). *Open circle* (oc), *linear* (lin), *supercoiled* (sc). Adressiertes Substratplasmid (T-6bp-P-6bp-T), nicht-adressiertes Plasmid (P) mit einer PvuII-Erkennungssequenz.

Der kürzere Linker (28 AS) hat sich bei den Spaltungskinetiken in Zusammenhang mit TALE-PvuII als spezifischer erwiesen. Um zu überprüfen, ob es einen Unterschied zwischen der dimeren (wtPvuII) und der monomeren (scPvuII) Variante bezüglich der Spezifität gibt, wurden diese im Überschuss bei Spaltungskinetiken nur auf dem nicht-adressierten Substrat eingesetzt (siehe Abb. 3.2.4). Für dieses Experiment wurde ein Puffer mit annähernd physiologischer Ionenstärke gewählt. Dazu wurde der Standardpuffer (Niedrigsalzpuffer) für PvuII (20 mM Tris-Ac, 50 mM K-Ac, 2 mM Mg²⁺-Ac, pH 7.5) den physiologischen Bedingungen mit einer Ionenstärke von ca. 150 mM und 1 mM Mg²⁺

angepasst (20 mM Tris-Ac, 120 mM K-Ac, 1 mM Mg-Ac und pH 7.5; mit Einberechnung der Ionen über die Enzymstammlösung nach der Dialyse wird eine Ionenstärke von ca. 150 mM erreicht) (Mano, Mao and Heller 2002, Rubelt *et al.* 2012). Die Kinetik mit der scPvuII-Variante wurde bis zu einer Stunde erfasst, und die der wtPvuII-Variante dauerte zwei Stunden. Im Gegensatz zu TALE-28-L-scPvuII, ist bei TALE-28-L-wtPvuII keine Aktivität zu beobachten. Die scPvuII-Variante *nicht* hauptsächlich die DNA (Einzelstrangspaltung) innerhalb der ersten Minuten mit einer Zunahme von Doppelstrangbrüchen (lineare Bande) gegen Ende der Kinetik.

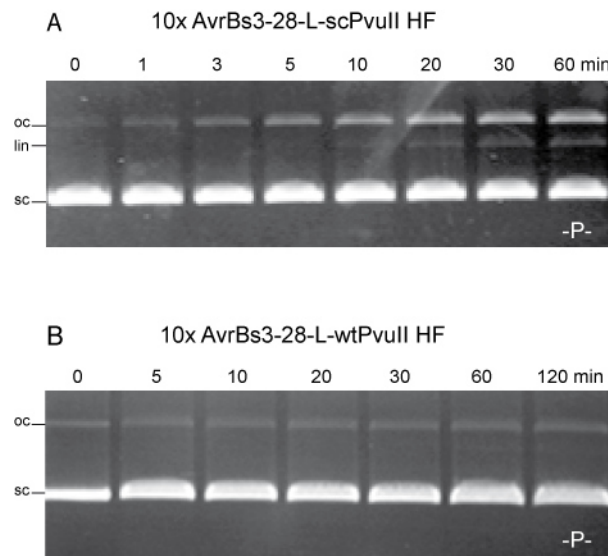


Abb. 3.2.4: Spezifitätsvergleich des Fusionsproteins mit wt- und scPvuII

Spaltungskinetik auf dem nicht-adressiertem Plasmid (P) (8 nM, eine PvuII-Erkennung) mit einem zehnfachen Überschuss an Enzym. Die Kinetiken wurden unter physiologischen Bedingungen durchgeführt (Ionenstärke 150 mM, 1 mM Mg²⁺, pH 7.5). **A** Spaltungskinetik bis zu 1 h des Fusionsproteins mit der scPvuII-Variante und **B** die Kinetik mit der wtPvuII-Variante bis zu 2 h. (*open circle* (oc), *linear* (lin), *supercoiled* (sc))

3.2.2 Spaltungskinetiken mit TALE-28-L-wtPvuII-HF

Die dimere Architektur des Fusionsenzym in Kombination mit den 28 AS hat sich als die spezifischere Variante erwiesen. Eine Spaltungskinetik in Kombination mit einem nicht-adressierten 900 bp langen PCR-Fragment sollte diese Feststellung bekräftigen. Als adressiertes Substrat wurden 8 nM Plasmid-DNA mit der Erkennungssequenz TALE-6bp-PvuII-6bp-TALE verwendet, in Kombination mit 32 nM nicht-adressiertes PCR-Fragment mit nur einer PvuII-Erkennungssequenz. Das Enzym wurde zum einen im Unterschuss, und zum anderen im Überschuss, bezogen auf die Plasmid-DNA, eingesetzt.

Innerhalb von Minuten ist eine vollständige Spaltung des adressierten Plasmids zu beobachten, und trotz der 24 h Inkubationszeit ist keine Aktivität an dem PCR-Produkt sichtbar (Abb. 3.2.5A).

In der Arbeit von Fonfara 2012 (Fonfara *et al.* 2012) wurde in das Fusionsprotein PvuII-I-SceI eine P14G-Mutation innerhalb des PvuII-Enzyms eingeführt. Die Mutation schwächt die Bindung zwischen den beiden PvuII-Untereinheiten, wodurch die Dimerisierung auf der gebundenen DNA vermutlich nur mit Hilfe der Bindungsmodule (I-SceI) stattfinden kann. Diese Variante zeigte die höchste Spezifität innerhalb dieser Architektur und wurde auch in dem TALE-PvuII-Konstrukt dieser Arbeit eingebracht. In der Abb. 3.2.5B ist die kompetitive Spaltungskinetik mit beiden Enzymarchitekturen dargestellt. Die TALE-28-L-wt PvuII-HF-Variante ist deutlich spezifischer.

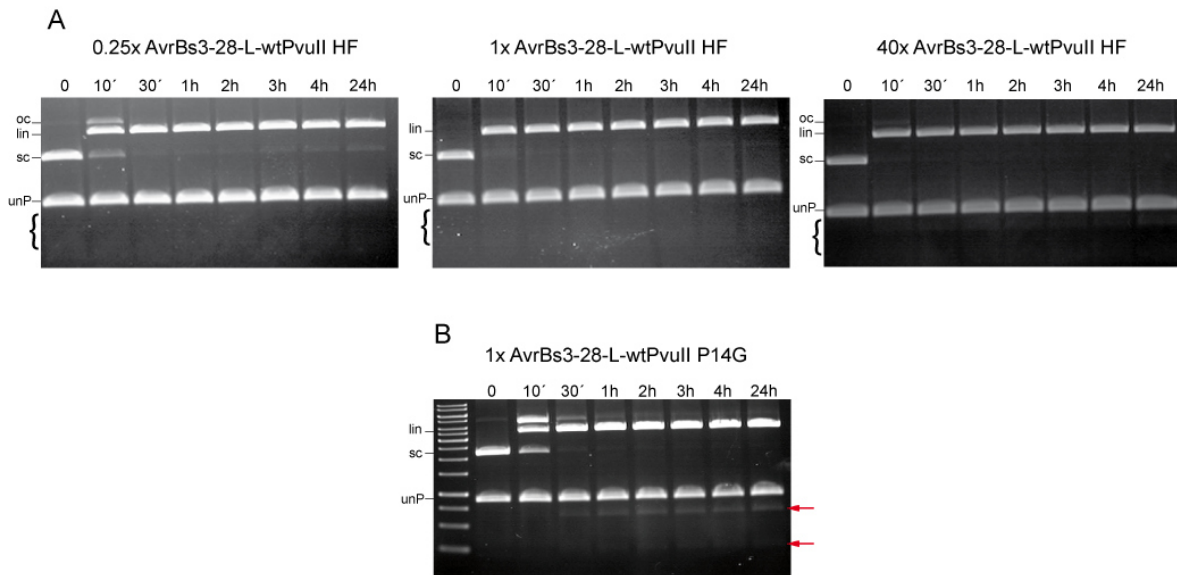


Abb. 3.2.5: Kompetitive Spaltungskinetiken

A Spaltungskinetik auf dem adressierten, tripartiten Plasmid (8 nM) unter physiologische Bedingungen (Ionenstärke 150 mM, 1 mM Mg^{2+} , pH 7.5) in Konkurrenz mit einem nicht-adressiertem PCR-Substrat (32nM). Das Fusionsprotein AvrBs3-28-L-wtPvuII-HF wird im Unterschuss (2 nM), äquimolar und im Überschuss (320 nM) bezogen auf die Plasmidkonzentration eingesetzt. **B** zeigt die Spaltungskinetik des Fusionsproteins mit einer zusätzlichen P14G-Mutation innerhalb des PvuII-HF-Enzyms. Die roten Pfeile zeigen die Spaltungsprodukte des nicht-adressierten PCR-Fragments (unP). (*open circle* (oc), *linear* (lin), *supercoiled* (sc)).

Um eine genauere Aussage über die Spezifität treffen zu können, wurden quantitative Messungen mit radioaktiv markierten Substraten durchgeführt. Bei der Spaltungskinetik wurde neben dem adressierten, tripartiten Substrat mit einer Länge von 393 bp ein nicht-adressiertes 1138 bp langes Substrat mit nur einer einzelnen PvuII-Erkennungssequenz verwendet. Im äquimolaren Verhältnis und einer Inkubationszeit von 24 h wird das adressierte Substrat innerhalb von Minuten gespalten, während auf dem nicht-adressierten

Substrat keine Aktivität zu erkennen ist. Validiert wurde dieses Ergebnis mit einer Spaltungskinetik mit beiden Substraten in Konkurrenz (Abb. 3.2.6). Unter Einbezug des Detektionslimits des *Instant Imagers*, wurde eine 34.000-fache Präferenz für das adressierte Substrat berechnet.

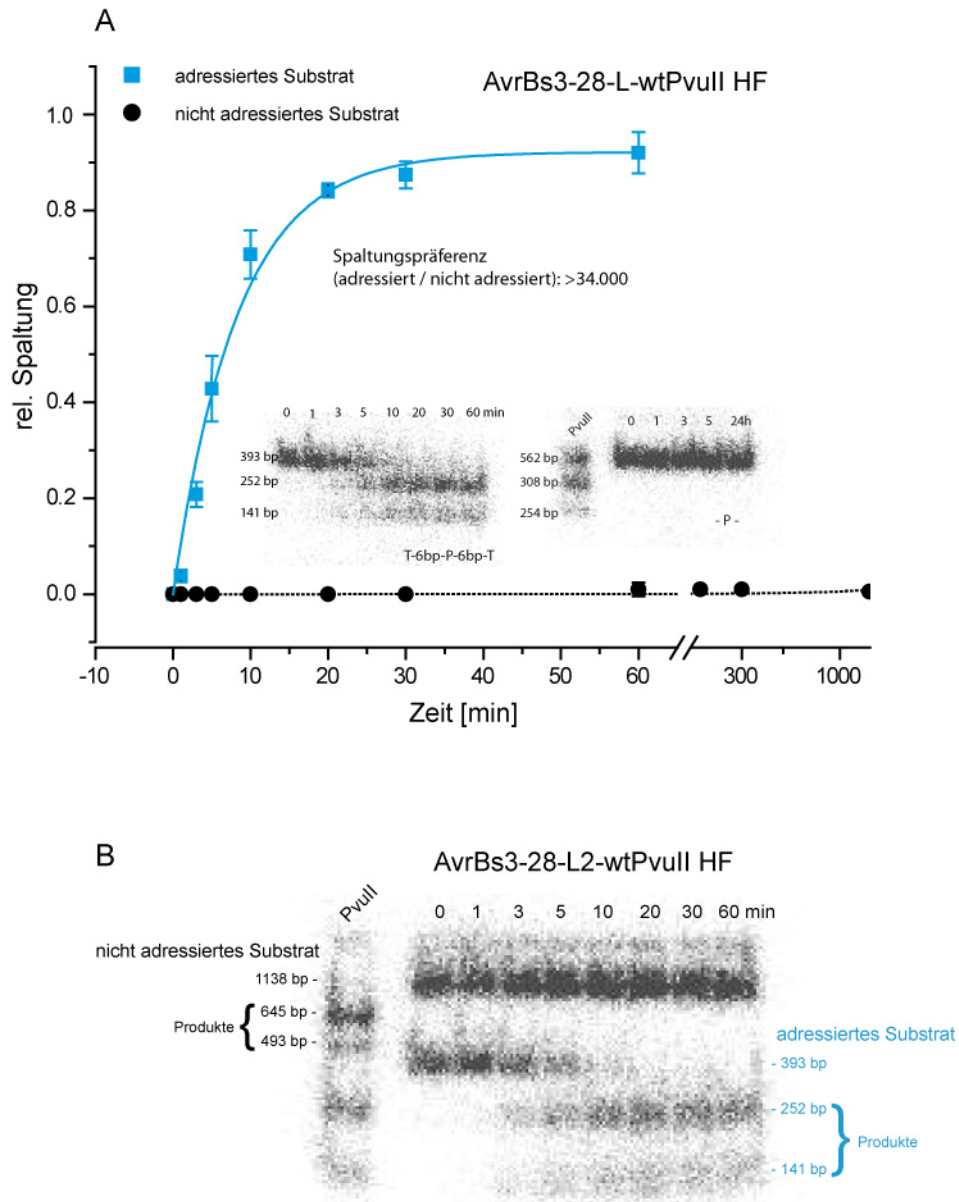


Abb. 3.2.6: Spezifität des AvrB3-28-L-wtPvuII-HF-Fusionsenzym

Die Spaltungskinetiken wurden unter physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg²⁺, pH 7.5) mit jeweils 20 nM Substrat und 20 nM Enzym durchgeführt. **A** dreifache Ausführung der Spaltungskinetiken auf dem adressierten Substrat (blaue Kurve), mit der tripartiten Erkennungssequenz (TALE-6bp-PvuII-6bp-TALE), und dem nicht-adressierten Substrat mit nur einer PvuII-Erkennungssequenz (schwarze Kurve). **B** Autoradiographie von kompetitiver Spaltungskinetik mit adressierten und nicht-adressierten Substrat (je 20 nM Substrat mit 20 nM Enzym).

Die Ergebnisse in Abb. 3.2.5 machen nicht nur die hohe Spezifität des Enzyms deutlich, sondern auch die *multi-turnover*-Eigenschaften. Trotz der langen Erkennungssequenz, vergleichbar der von Homing-Endonukleasen bindet und spaltet AvrBs3-28-L-wtPvuII-HF jedoch mehrfach. Um einen Beweis für die *multi-turnover*-Eigenschaft des Enzyms zu erbringen, wurde ein Zehntel der Menge an Enzym im Vergleich zum Substrat eingesetzt. Dabei wurde anstatt der bei *single turnover* erwarteten 10 % mehr als 60 % des Substrates gespalten (Abb. 3.2.7).

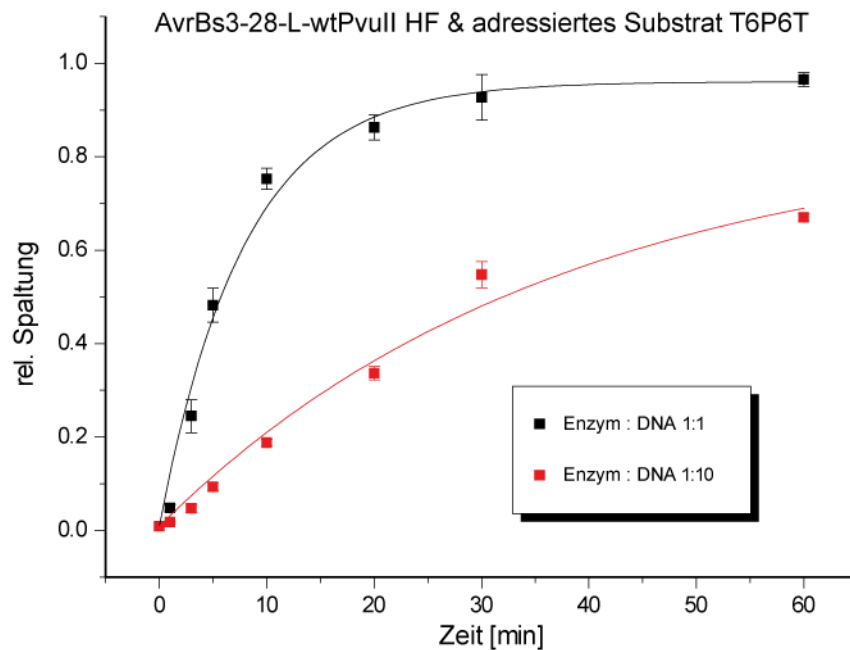


Abb. 3.2.7: Spaltungskinetik unter Verwendung radioaktiv markierter DNA

Spaltungskinetiken mit P^{32} radioaktiv markierten, adressierten Substraten. 20 nM des tripartiten Substrats (TALE-6bp-PvuII-6bp-TALE) wurde im äquimolaren Verhältnis und im Unterschuss (2 nM) mit dem Konstrukt AvrBs3-28-L-wtPvuII-HF unter annähernd physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg^{2+} , pH 7.5) inkubiert.

3.2.3 Linkervarianten und präferierte Substrate

Die Spezifität des AvrBs3-28-L-wtPvuII-Enzyms wurde bisher nur auf dem adressierten Substrat mit 6 bp Abstand getestet (TALE-6bp-PvuII-6bp-TALE). Um zu überprüfen, ob weitere Abstände bevorzugt werden oder ob weitere Linkervariationen innerhalb des Fusionskonstrukts zu einer anderen Präferenz führen, wurden nicht nur die Linker im Fusionsprotein gekürzt, sondern auch neu-adressierte Substrate generiert. Abb. 3.2.8 stellt die Spaltungsraten der Enzyme auf den verschiedenen Substraten dar. Die Daten resultieren aus den Spaltungskinetiken mit radioaktiv markierten Substraten im äquimolaren Verhältnis. Das Enzym AvrBs3-28-L-wtPvuII-HF weist die höchste Rate mit den Substraten mit 4 bp und 6 bp *spacer* auf. Die nächst kürzere Linkervariante AvrBs3-28-wtPvuII-HF weist die geringste Aktivität auf. Der kürzeste Linker (AvrBs3-L-wtPvuII), mit nur 16 AS zwischen AvrBs3 und wtPvuII, zeigt die höchste Spaltungsrate unter Verwendung des adressierten Substrates mit dem kürzesten *spacer* von nur 2 bp.

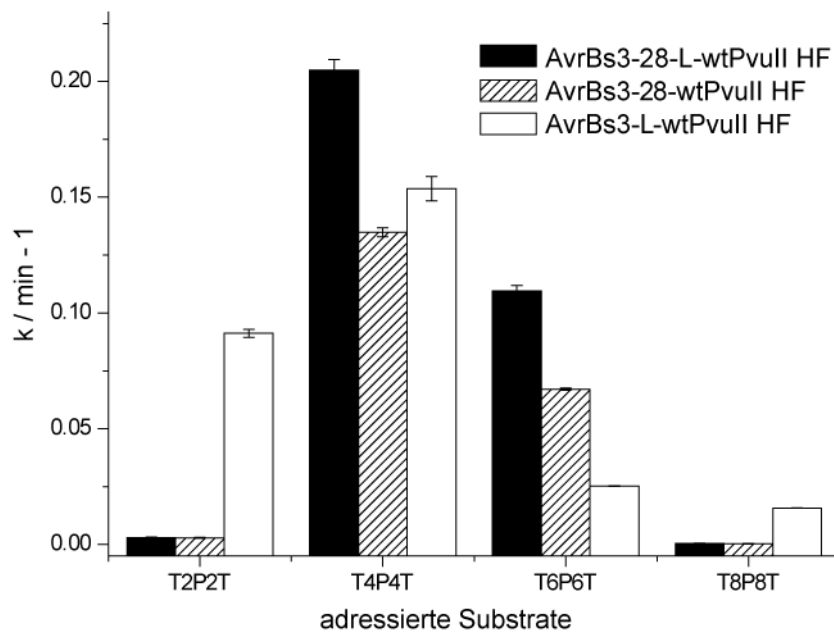


Abb. 3.2.8: Spaltungsraten unter Verwendung adressierter Substrate unterschiedlicher Länge

20 nM radioaktiv markierte Substrate mit 2, 4, 6 und 8 bp zwischen der AvrBs3 und der PvuII-Erkennungssequenz, wurden für 60 min mit 20 nM AvrBs3-28-L-wtPvuII-HF, AvrBs3-28-wtPvuII-HF und AvrBs3-L-wtPvuII-HF unter annähernd physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg²⁺, pH 7.5) inkubiert.

3.2.4 Spaltungskinetiken mit bipartiten Substraten

Die bisherigen Experimente haben gezeigt, dass das dimere Enzym AvrBs3-28-L-wtPvuII-HF eine tripartite Sequenz adressieren kann. Als nächstes sollten die Spaltungseigenschaften unter Verwendung eines bipartiten (TALE-6bp-PvuII) Substrates überprüft werden. Abb. 3.2.9A zeigt die Spaltungskinetik auf einem adressierten Plasmid in Konkurrenz mit einem nicht-adressierten PCR-Fragment. Dabei erfolgt lediglich eine Spaltung des adressierten Plasmids. Das PCR-Fragment hingegen bleibt intakt. Zur quantitativen Bestimmung wurde die Spaltungskinetik mit bipartiten, radioaktiv markierten Substraten durchgeführt. Innerhalb einer Stunde wurden knapp 25 % des Substrats umgesetzt (Abb. 3.2.9B).

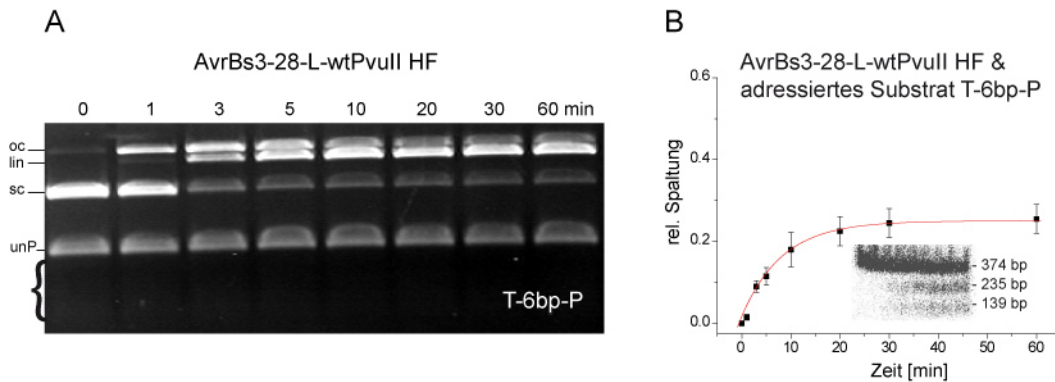


Abb. 3.2.9: Spaltungskinetik des bipartiten Substrats

A 8 nM des adressierten Plasmids (bipartite Sequenz: TALE-6bp-PvuII) wurde im äquimolaren Verhältnis unter annähernd physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg^{2+} , pH 7.5) mit dem Enzym AvrBs3-28-L-PvuII-HF (8 nM) eingesetzt. Zusätzlich wurde ein nicht-adressiertes PCR-Fragment (unP) mit nur einer PvuII-Erkennung im Überschuss (32 nM) mit inkubiert. **B** Spaltungskinetik mit radioaktiv markiertem bipartitem Substrat. (*open circle* (oc), *linear* (lin), *supercoiled* (sc))

3.2.5 Stabilitätstest des Enzyms TALE-28-L-wtPvuII-HF

Mit Hilfe des Stabilitätstests kann überprüft werden, wie lange ein Enzym aktiv ist und Substrate umsetzen kann. Dazu wurde das Enzym zusammen mit dem tripartiten, adressierten Plasmidsubstrat im äquimolaren Verhältnis ohne Mg^{2+} inkubiert. Nach festgelegten Zeitpunkten wurden den Ansätzen 1 mM Mg^{2+} zugegeben. Die Reaktion wurde nach 30 min gestoppt und über ein Agarosegel überprüft. Der Stabilitätstest zeigt, dass das Enzym nach einer vierstündigen Inkubationszeit noch 50 % Aktivität und nach 24 h nahezu keine Aktivität mehr besitzt (Abb. 3.2.10).

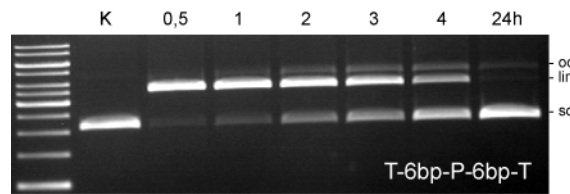


Abb. 3.2.10: Stabilitätstest des AvrBs3-28-L-wtPvuII-HF-Enzyms

Die Stabilität des Fusionsenzym wurde unter annähernd physiologischen Bedingungen (Ionenstärke 150 mM, pH 7.5) durch einen Aktivitätstest mit anschließender Elektrophorese in einem 1.25 % igem Agarosegel überprüft. 8 nM des adressierten Substrats (Plasmid-DNA mit der tripartiten Erkennungssequenz TALE-6bp-PvuII-6bp-TALE) wurden mit 8 nM des Fusionsenzym für 0.5, 1, 2, 3, 4 und 24 h ohne Mg^{2+} inkubiert. Nach der Inkubationszeit wurde die Reaktion mit 1 mM Mg^{2+} für 30 min gestartet. (*open circle* (oc), *linear* (lin), *supercoiled* (sc))

3.2.6 Vergleich physiologischer Puffer

Der bisher benutzte annähernd physiologische Puffer basierte auf der Grundlage des empfohlenen PvuII-Puffers der Firma NEB. Der Tris-/Kalium-Acetat-Puffer wurde auf eine Ionenstärke von ca. 150 mM eingestellt (20 mM Tris-Ac, 120 mM K-Ac, 1 mM Mg^{2+} -Ac, pH 7.5), unter Berücksichtigung der Ionen im Enzym-Dialysepuffer. Um sicher zu stellen, dass die ermittelten Ergebnisse mit diesem „physiologischer Puffer“ auch vergleichbar sind mit weiteren oder ähnlichen Puffern wurde eine Spaltungskinetik mit radioaktiven Substraten im KGB-Puffer (115 mM Kaliumglutamat, 25 mM Tris-Acetat, 1 mM Magnesium-Acetat, 0,5 mM 2-Mercaptoethanol, 50 μ g/ml BSA, pH 7.6) durchgeführt (McClelland *et al.* 1988). Dieser Puffer weist mit Hinblick auf die Ionenstärke mit ca. 150 mM und 1 mM Mg^{2+} Konzentration physiologische Eigenschaften auf. Die Ergebnisse der Spaltungskinetiken in Abb. 3.2.11 zeigen für beide Reaktionsbedingungen eine vollständige Spaltung innerhalb von Minuten.

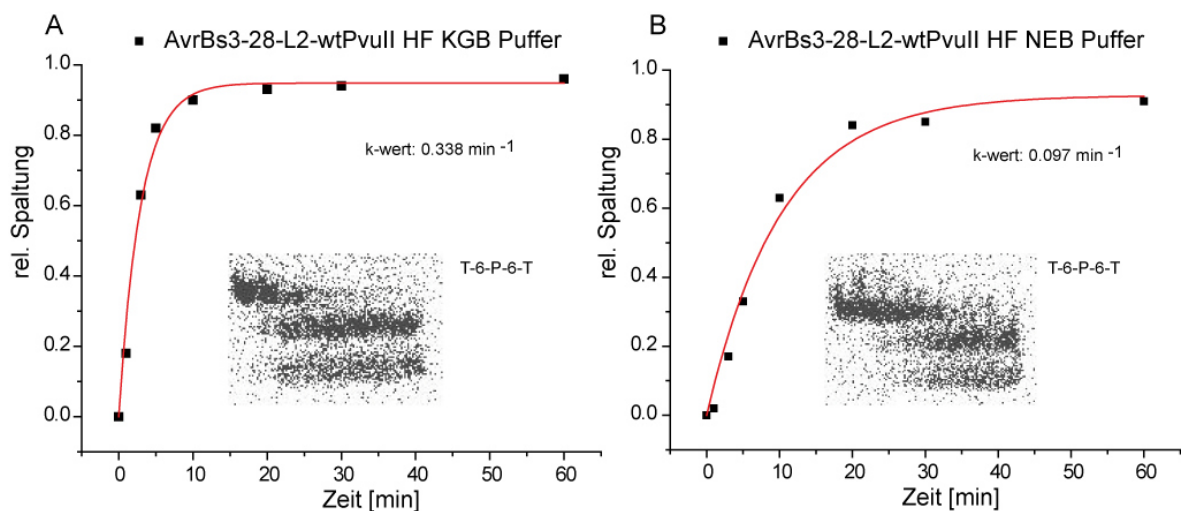


Abb. 3.2.11: Vergleich der Spaltungskinetik in physiologisch angepassten Puffern

Die Spaltungskinetiken wurden im äquimolaren Verhältnis (radioaktiv markiertes Substrat und Enzym je 20 nM) durchgeführt. Das Substrat hat die tripartite, adressierte Erkennungssequenz TALE-6bp-PvuII-6bp-TALE. **A** Spaltungskinetik im KGB-Puffer (115 mM Kalium-Glutamat, 25 mM Tris-Ac, 1 mM Mg^{2+} , 50 μ g/ml BSA, 0,5 mM 2-Mercaptoethanol, pH 7.6). **B** Spaltungskinetik mit dem von NEB angeglichenen physiologischen Puffer (20 mM Tris-Ac, 120 mM K-Ac, 1 mM Mg-Ac, pH 7.5).

3.2.7 Viabilitätsexperiment des TALE-PvuII-Konstrukts in *E.coli*

Bei allen Klonierungen und Proteinaufreinigungen der TALE-PvuII-Konstrukte wurden die *E. coli*-Stämme immer mit der PvuII-Methyltransferase vor Selbstrestriktion geschützt. Die spezifische Variante AvrBs3-28-L-wtPvuII-HF wurde in den für die Expression verwendeten XL10Gold-*E. coli*-Stamm ohne Schutz der Methyltransferase transformiert. Zu erwarten war eine hohe Klonanzahl ähnlich wie auf der Kontrollplatte. Abb. 3.2.12 zeigt das Ergebnis des Viabilitätsexperiments. Die überlebende Anzahl der Klone liegt bei ungefähr 2 % im Vergleich zur Transformation mit dem AvrBs3-TAL-Effektor alleine, welches nicht toxisch für die Zellen ist. Sequenzierungen haben große Deletionen des PvuII-Gens bei den 2 % lebensfähigen Klonen ergeben (Daten nicht gezeigt).

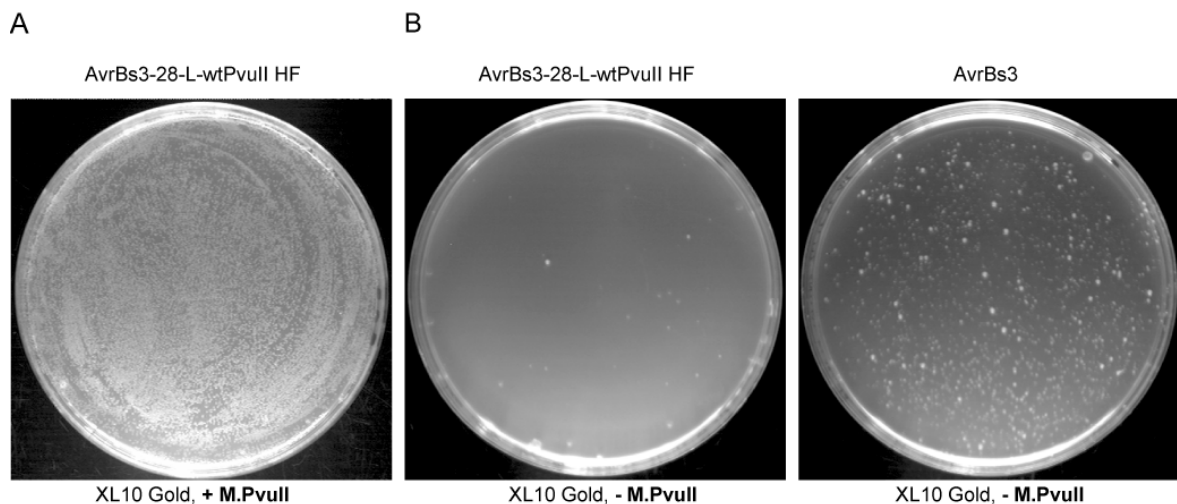


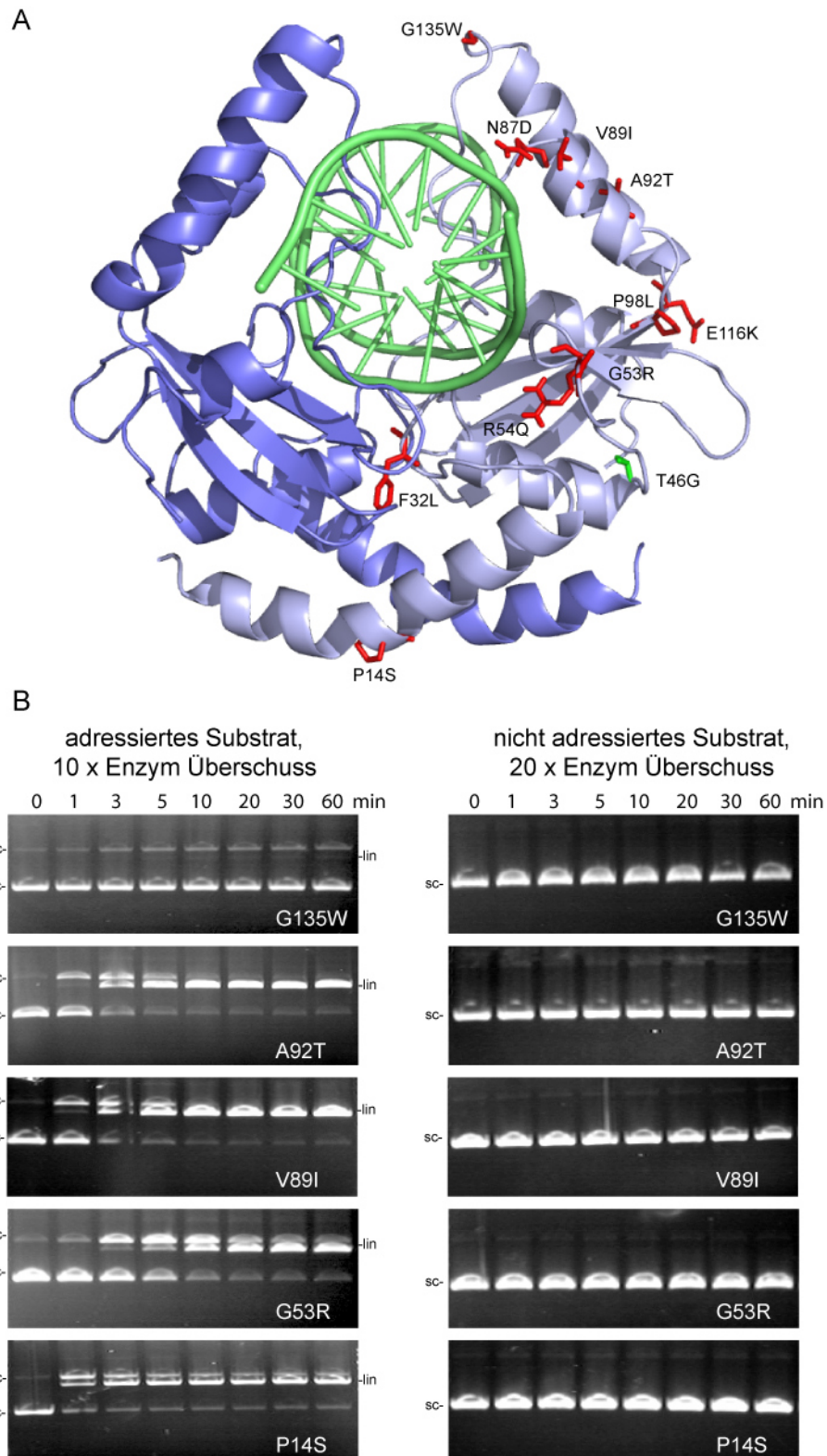
Abb. 3.2.12: Viabilitätstest

A Elektrokompente XL10Gold-*E.coli*-Zellen (Stratagene) mit der PvuII-Methyltransferase wurden mit dem Fusionsenzym AvrBs3-28-L-wtPvuII-HF (100 ng Plasmid-DNA) transformiert. 100 μ l wurden auf Agarplatten mit Amp ausgestrichen. **B** XL10Gold-Zellen ohne PvuII-Methyltransferase wurden mit dem Fusionsenzym (links) und nur mit dem AvrBs3-Protein alleine als Kontrolle (rechts) transformiert.

3.2.8 Neue PvuII-Varianten im Fusionskonstrukt

Die hohe Toxizität des AvrBs3-28-L-wtPvuII-HF-Enzyms sollte in einem zweiten Schritt durch einen Selektionsprozess von neuen PvuII-Varianten gemindert werden (Oelgeschlager *et al.* 1990). Dafür wurde die DNA, die das Fusionsprotein kodiert, in einen Vektor mit einem eukaryotischen CMV-Promotor (pvax-Vektor) eingebracht. Die Selektion findet jedoch über *E. coli*-Zellen statt, ohne den Schutz der PvuII-Methyltransferase. Zehn AvrBs3-PvuII-positive Zellen wurden mit dem Ergebnis, dass jeder Klon eine unterschiedliche einzelne Mutation innerhalb des PvuII-Gens besitzt,

sequenziert. Abb. 3.2.13A zeigt die Mutationen im PvuII-Enzym. Sie sind nicht innerhalb einer spezifischen Domäne lokalisiert, sondern über das ganze Enzym verteilt. Alle zehn PvuII-Varianten wurden anschließend charakterisiert, um ihre Aktivität im Einzelnen zu untersuchen. Dazu wurden die zehn AvrBs3-28-L-wtPvuII-HF-Varianten vom eukaryotischen in den prokaryotischen Vektor pQE30 (Qiagen) umgesetzt und die Enzyme nach der induzierten Expression aufgereinigt. Ihre Aktivität wurde mit Hilfe von Spaltungskinetiken im physiologischen Puffer mit einem 20-fachen Überschuss von Enzym zu DNA überprüft. Fünf von zehn Varianten wiesen keine Aktivität auf. Die Spaltungsaktivität der anderen fünf Varianten ist in Abb. 3.2.13B dargestellt. Unter Verwendung des nicht-adressierten Substrates (Plasmid-DNA mit nur einer PvuII-Sequenz) ist für alle fünf Varianten innerhalb einer Stunde keine Aktivität zu beobachten. Die Spaltungskinetik mit dem adressierten Substrat und einem zehnfachen Überschuss an Enzym, macht die Unterschiede innerhalb der Varianten deutlich. Die G135W-Variante weist die geringste Aktivität auf, A92T und V89I die höchste. Die letzten beiden Varianten zeigen in einem 1:1 Verhältnis ähnlich hohe Spaltungsraten wie das AvrBs3-28-L-PvuII-HF-Enzym (Daten nicht abgebildet).

**Abb. 3.2.13: AvrBs3-PvuII-Varianten**

A die selektierten zehn PvuII-Mutationen (rot) im Fusionsprotein. Jeweils eine Mutation kommt zusätzlich zur T46G-Mutation (grün) in einer AvrBs3-PvuII-Variante vor. **B** Spaltungskinetiken der fünf AvrBs3-PvuII-Varianten mit der höchsten Aktivität. Es wurden 8 nM Plasmid-DNA eingesetzt, auf dem adressierten Substrat wurde ein zehnfacher und auf dem nicht-adressierten Substrat ein 20-facher Enzymüberschuss verwendet.

3.2.9 Viabilitätstests der PvuII-Varianten in *E. coli*

Die neuen TALE-PvuII-Varianten (Abb. 3.2.13) wurden in unterschiedlichen *E. coli*-Stämmen auf ihre Toxizität hin untersucht. Ohne den Schutz der PvuII-Methyltransferase wurde die Anzahl der *E. coli*-Klone gezählt. Die inaktive P98L-Variante diente als Positivkontrolle und das Restriktionsenzym PvuII alleine als Negativkontrolle. Bei zwei Varianten ist eine geringe Toxizität zu beobachten. Die aktivsten Varianten besitzen die höchste Toxizität. Eine Ausnahme stellt die Mutante G53R dar. Diese weist neben ihrer hohen Aktivität eine geringe Toxizität auf (Abb. 3.2.14 und Abb. 3.2.13B). Für die Charakterisierung in eukaryotischen Zellen wurden die Varianten A92T, G53R und G135W ausgewählt und erneut zum Verifizieren in den *E. coli*-Stamm XL10Gold (Stratagene) transformiert.

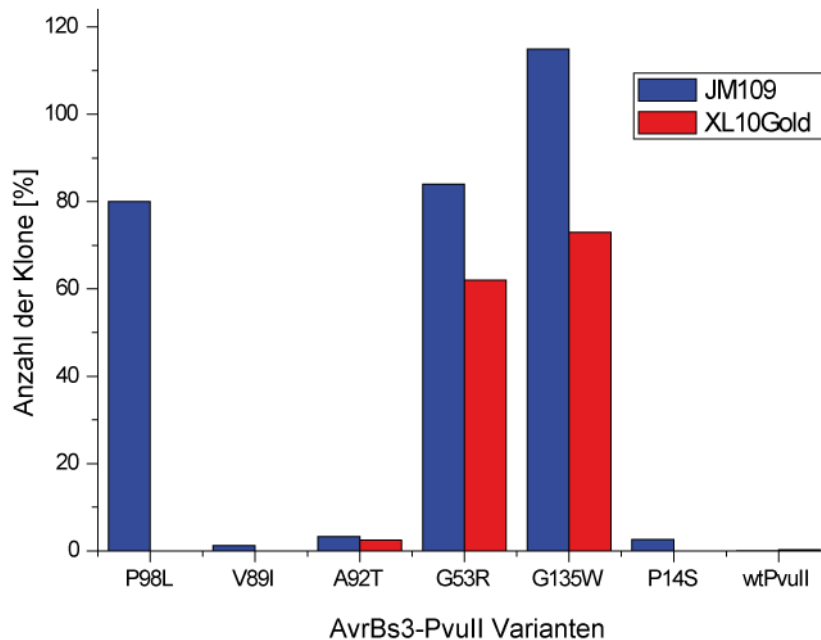


Abb. 3.2.14: Übersicht des Viabilitätstests mit den TALE-PvuII-Varianten

Die elektrokompetenten, PvuII-Methyltransferase (M. PvuII) negativen, JM109-*E. coli*-Zellen wurden jeweils mit 50 ng Plasmid-DNA transformiert und anschließend mit je 50 μ l auf Ampicillin-Agarplatten ausgestrichen. Als Kontrolle dient das nicht toxische AvrBs3-Protein, kodiert auf demselben Vektor-Plasmid. Die Anzahl der Klone wurden auf die Kontrolle normiert.

3.2.10 Toxizitätstest der TALE-PvuII-Varianten in HEK293-Zellen

Die Experimente mit eukaryotischen Zellen (HEK293) wurden in Kooperation mit der Arbeitsgruppe von Prof. Dr. Toni Cathomen in Freiburg durchgeführt. Die Toxizität der TALE-PvuII-Nukleasen wurde überprüft, indem eine Co-Transfektion mit einem zweiten Vektor, welcher für das rot fluoreszierende Protein *mCherry* kodiert, in die Zellen

eingbracht wurde. Die Fluoreszenz innerhalb der Zellen wurde nach zwei und fünf Tagen nach der Transfektion mit Hilfe von FACS (*fluorescence-activated cell sorting*) gemessen. Die Daten wurden auf das Signal der Positivkontrolle I-SceI normiert. Als Negativkontrolle diente eine Zink-Finger-FokI-Nuklease (ZFN) (Alwin *et al.* 2005). Die TALE-PvuII-Mutanten weisen alle eine geringere Toxizität als die ZFN auf. Selbst die Toxizität der aktivsten PvuII-A92T-Variante ist im Vergleich zu der Toxizität in *E. coli*-Zellen (Abb. 3.2.14 und Abb. 3.2.15) in humanen Zelllinien gering.

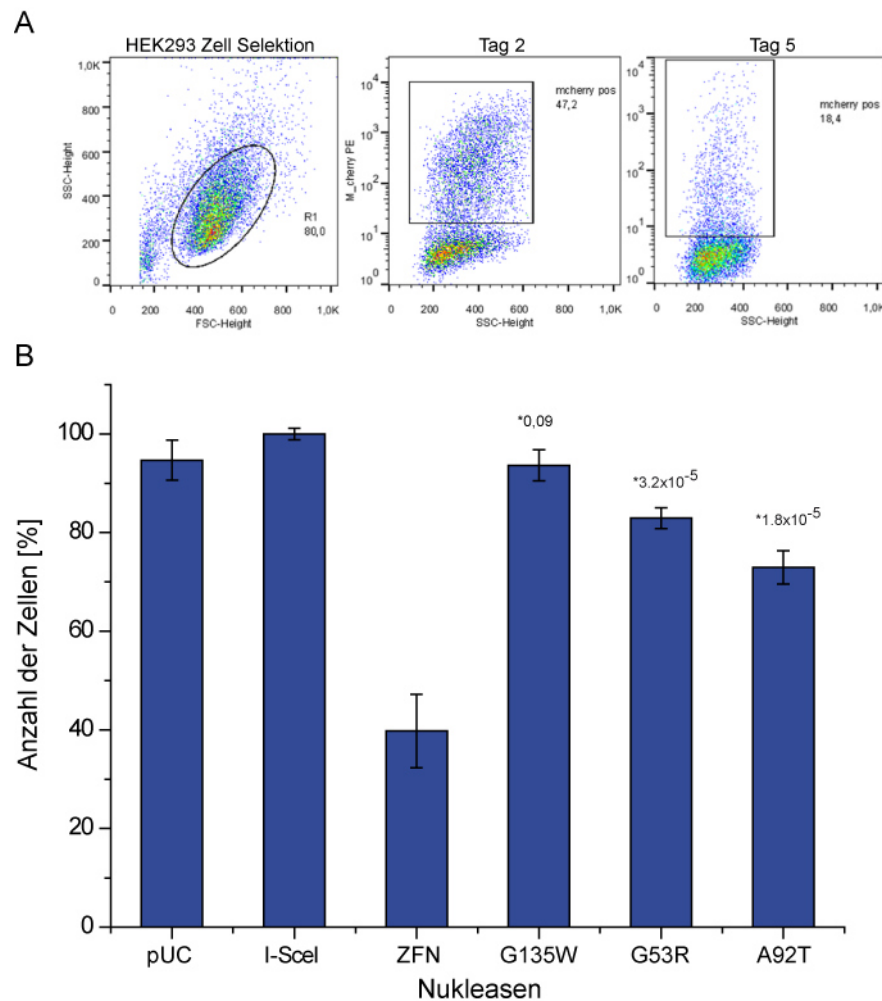


Abb. 3.2.15: Toxizitätstest der TALE-PvuII-Varianten in HEK293-Zellen³

120.000 HEK293-Zellen wurden mit 800 ng TALE-PvuII-DNA und 100 ng *mCherry*-Expressionsvektor co-transfiziert. **A** In der Durchflusszytometrie wird die Population vitaler Zellen (Kreis) durch ihre Streulichtparameter (Vorwärtsstreulicht FSC und Seitenstreulicht SSC) definiert. Anschließend wird die Fluoreszenz des *mCherry*-Proteins nach 2 und 5 Tagen nach der Transfektion via FACS (FACSCalibur, BD Biosciences) bestimmt; hier am Bsp. von G53R. **B** Übersicht der Toxizität. Als Kontrollen dienen der Leervektor pUC, I-SceI als nicht-toxische Positivkontrolle und ZFN-FokI als Negativkontrolle. Die Daten wurden gegen I-SceI normalisiert und sechsfach bestimmt. * gibt *p-value* an. Die durchschnittliche relative Anzahl der Zellen für G135W, G53R und A92T liegen bei 94 %, 83 % und 73 %.

³ Die Experimente in der Zellkultur wurden in Kooperation mit der Arbeitsgruppe von Toni Cathomen in Freiburg durchgeführt.

3.2.11 Aktivitätstest der homodimeren TALE-PvuII-Varianten in HEK293-Zellen

Die Aktivität der TALE-PvuII-Nukleasen in HEK293, eukaryotischen Zellen, wurde mit Hilfe eines Substrates ermittelt, welches zusammen mit dem Nuklease-kodierenden Plasmid co-transfiziert wurde und episomal in der Zelle vorliegt. Wird die PvuII-Sequenz des adressierten Substrates (TALE-6 bp-PvuII-6 bp-TALE) durch die Nuklease gespalten, kann diese durch die zelleigenen Reparaturmechanismen wieder zusammengefügt werden. Dabei können durch *non-homologous-end-joining* (NHEJ) Basenpaare hinzugefügt oder deletiert werden, die zu einem Verlust der PvuII-Erkennungssequenz führen. Zur Überprüfung der Nukleaseaktivität wird das Substrat amplifiziert und anschließend mit PvuII gespalten. Ein gewisser Prozentsatz des Substrats ist durch PvuII nicht mehr spaltbar (aufgrund von NHEJ), und lässt Rückschlüsse auf die Aktivität des Fusionsenzym zu, siehe Abb. 3.2.16.

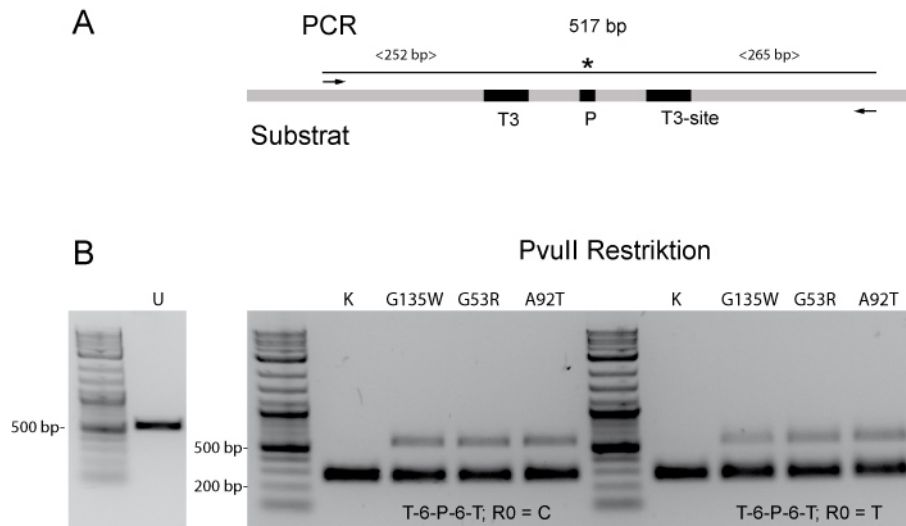


Abb. 3.2.16: *In vivo*-Aktivitätstest der TALE-PvuII-Varianten in HEK293-Zellen

500.000 HEK293-Zellen wurden zusammen mit 1 µg adressierten Substratplasmid (AvrBs3-6bp-PvuII-6bp-AvrBs3) und 4 µg TALE-PvuII-Nuklease transfiziert. Nach drei Tagen wurden die Zellen geerntet und die DNA isoliert. **A** Mit dem aufgereinigten Plasmid wurde eine PCR durchgeführt mit einer resultierenden DNA-Länge von 517 bp. Eine Spaltung mit PvuII (*) führt zu den Größen 252 bzw. 265 bp. **B** Das PCR-Produkt (ungespalten; U) wurde mit PvuII gespalten (100 ng DNA, 20 *units* PvuII-HF, NEB). Die Wildtyp-AvrBs3-Erkennungssequenz auf dem Substrat mit einer T-Base an der nullten *repeat*-Position ist rechts abgebildet. Links wurde die Base T durch C ersetzt.

Alle drei TALE-PvuII-Varianten (G135W, G53R, A92T) sind in den HEK293-Zellen aktiv. Sowohl die Aktivität der drei Varianten untereinander (15-20 %) als auch die Aktivität auf den beiden Substraten ist vergleichbar. Beide Plasmidsubstrate besitzen die adressierte Erkennungssequenz nur mit dem Unterschied, dass neben der Wildtyp-AvrBs3 Erkennungssequenz (R0=T), auch eine Erkennungssequenz gewählt wurde, bei der die Base T an der Position *repeat* 0 durch die Base C substituiert wurde (R0=C).

3.3 TALE-PvuII-Heterodimere

Genomische DNA, die mit hochspezifischen Nukleasen adressiert werden soll, weist um die gewünschte Spaltstelle unterschiedliche Bindungssequenzen auf. Ein homodimeres Fusionsenzym kann aufgrund seiner identischen Bindungsmodule nur an eine Sequenz binden und ist somit nur zu 50 % an die genomische DNA gebunden (Abb. 3.3.1A). Die Bindungsspezifität kann durch eine heterodimere Architektur erhöht werden (Miller *et al.* 2007, Doyon *et al.* 2011). Ein dimeres Spaltmodul mit zwei unterschiedlichen Bindungsmodulen wäre zwar ein heterodimeres Konstrukt, jedoch besitzt dieses die Eigenschaft halbseitig zu binden und zu spalten. (Kapitel 3.2.4). Durch die Generierung von heterodimeren Spaltmodulen kann diese Problematik umgangen werden. Die in dieser Arbeit hergestellten Konstrukte sind an die heterodimeren Fusionsenzyme der ZFN-FokI angelehnt (Miller *et al.* 2007). Der Grundgedanke ist, aus einem homodimeren Spaltmodul ein heterodimeres zu bilden. Dabei werden spezifische Aminosäuren der Nuklease, die sich in den Untereinheiten gegenüberliegen, ausgewählt und durch Aminosäuren mit sauren und basischen Eigenschaften ersetzt. Aufgrund der elektrostatischen Abstoßung der Untereinheiten, die durch die Aminosäuren gleicher Ladungen entstehen, ist die Bildung einer dimeren Nuklease nahezu unmöglich (Abb. 3.3.1).

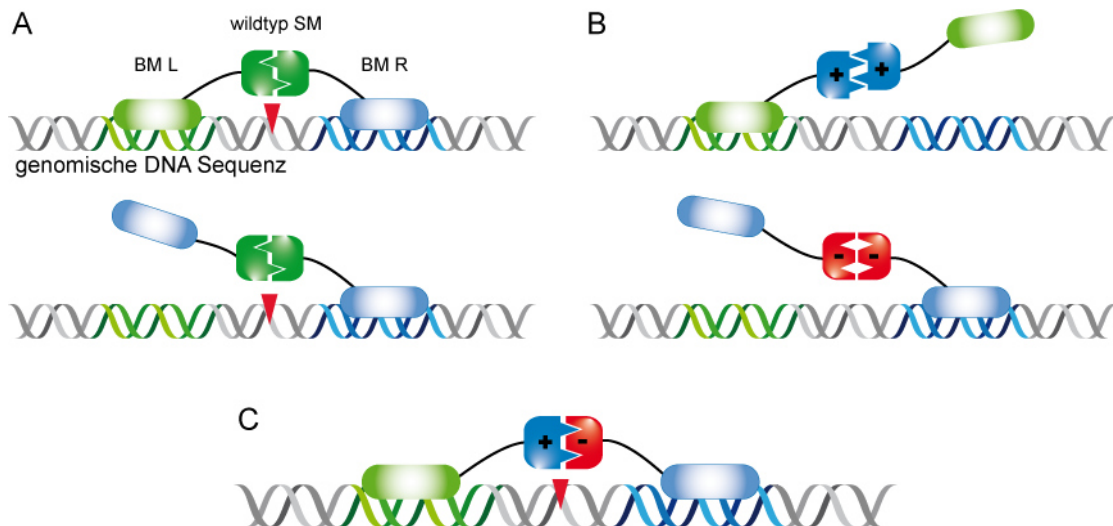


Abb. 3.3.1: Architektur heterodimerer Fusionsenzyme

A Oben: Eine native, homodimere Nuklease, die an zwei unterschiedliche Bindungsmodulen (BM L, grün; BM R, blau) fusioniert ist, spaltet die DNA (roter Keil). Unten: Ein homodimeres Fusionsenzym spaltet bei halbseitiger Bindung die DNA. **B** Zwei identisch modifizierte Nukleaseuntereinheiten können kein katalytisch aktives Dimer bilden. **C** Katalytisch aktives, heterodimeres Fusionsenzym.

3.3.1 Heterodimere mit PvuII und TALE-AvrBs3 & -AvrBs4

Die FokI-Nuklease hat gegenüber PvuII den Vorteil, dass sie in Lösung als monomere Nuklease vorliegt und nur die katalytisch aktiven Untereinheiten auf der DNA dimerisieren (Vanamee, Santagata and Aggarwal 2001). Deshalb sollte zunächst überprüft werden, ob es bei PvuII zu einem Austausch der Untereinheiten kommt. Hierfür wurden zwei verschiedene TALE-Proteine verwendet (AvrBs3 und AvrBs4), die als Bindungsmodule in Zusammenhang mit PvuII dienen (AvrBs3-28-L-wtPvuII-HF und AvrBs4-28-L-wtPvuII-HF). Mit entsprechenden tripartiten Plasmidsubstraten wurden Spaltungskinetiken durchgeführt. In Abb. 3.3.2A wird die Spaltungskinetik des AvrBs3-28-L-wtPvuII-HF mit adressierten (AvrBs3-6bp-P-6bp-AvrBs3) und nicht adressierten (AvrBs4-6bp-P-6bp-AvrBs4) tripartiten Substrat als Kontrolle gezeigt. Trotz der Homologie beider Erkennungssequenzen wird unter physiologischen Bedingungen nur das adressierte Substrat gespalten. Auf dieser Grundlage wurden anschließend tripartite Plasmidsubstrate für heterodimere Fusionsenzyme hergestellt (AvrBs3-6bp-P-6bp-AvrBs4). Das jeweilige homodimere Enzym führt in der Spaltungskinetik hauptsächlich auf dem heterodimeren Substrat einen DNA-Einzelstrangbruch durch. Werden beide Fusionsenzyme miteinander gemischt, so dass sich die Untereinheiten austauschen können, beobachtet man in der Spaltungskinetik auf dem heterodimeren Substrat hauptsächlich einen DNA-Doppelstrangbruch (Linearisierung der Plasmid-DNA, Abb. 3.3.2)

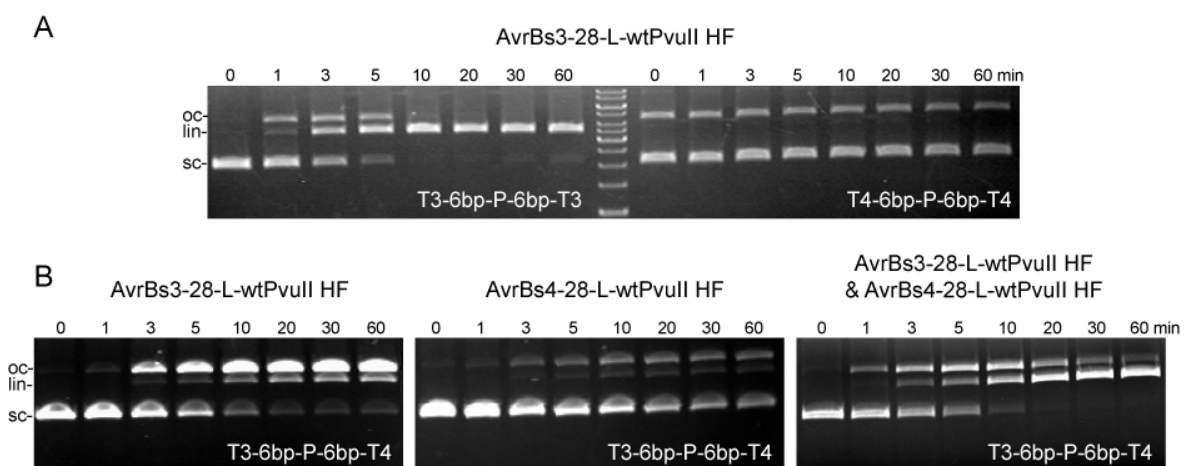


Abb. 3.3.2: Spaltungskinetiken mit wtPvuII-HF und AvrBs3 & AvrBs4

Die Spaltungskinetiken wurden mit äquimolarem Verhältnis von Enzym zu DNA (8 nM) unter physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg²⁺, pH 7.5) in Abhängigkeit von der Zeit durchgeführt. **A** stellt die Kontrolle dar. Spaltungskinetiken des AvrBs3-28-L-wtPvuII-HF-Enzyms wurden mit tripartiten, adressierten Substrat AvrBs3-6bp-PvuII-6bp-AvrBs3 (links) und nicht adressierten Plasmidsubstrat AvrBs4-6bp-PvuII-6bp-AvrBs4 (rechts) aufgenommen. **B** Spaltungskinetiken wurden auf dem Heterosubstrat (AvrBs3-6bp-PvuII-6bp-AvrBs4) einzeln und zusammen mit den Fusionsenzymen AvrBs3-28-L-wtPvuII-HF & AvrBs4-28-L-wtPvuII-HF durchgeführt. (oc = open circle, lin = linearization; sc = supercoiled)

3.3.2 Heterodimere mit unterschiedlichen PvuII-Untereinheiten

Für die Auswahl der spezifischen Aminosäuren der Nuklease, die sich in den Untereinheiten gegenüberliegen und die Dimerisierungsebene bilden, wurden folgenden Positionen ausgewählt. Die eine Untereinheit von PvuII dimerisiert mit der Aminosäure Asparaginsäure an Position 30 (D30) mit der anderen Untereinheit mit Lysin an Position 38 (K38). Die Positionen liegen zentral, jedoch außerhalb des aktiven Zentrums und bilden, aufgrund ihrer unterschiedlichen Ladungen ein stabiles Dimer. Diese im Wildtypenzym vorkommenden, aufeinander abgestimmten Aminosäuren wurden für die Herstellung heterodimerer PvuII-Nukleasen zielgerichtet verändert. Dabei wurden PvuII-Mutanten generiert, die in ihren Untereinheiten entweder nur basische (D30K) oder nur saure Aminosäuren (K38D) enthalten, so dass sie im Einzelnen als homodimere Nukleasen nicht in der Lage sind, eine DNA-Spaltung durchzuführen. Abb. 3.3.3 veranschaulicht die Position der Aminosäuren innerhalb der PvuII-Untereinheiten und die Auswirkung der Mutagenese in Bezug auf die Dimerisierung von PvuII.

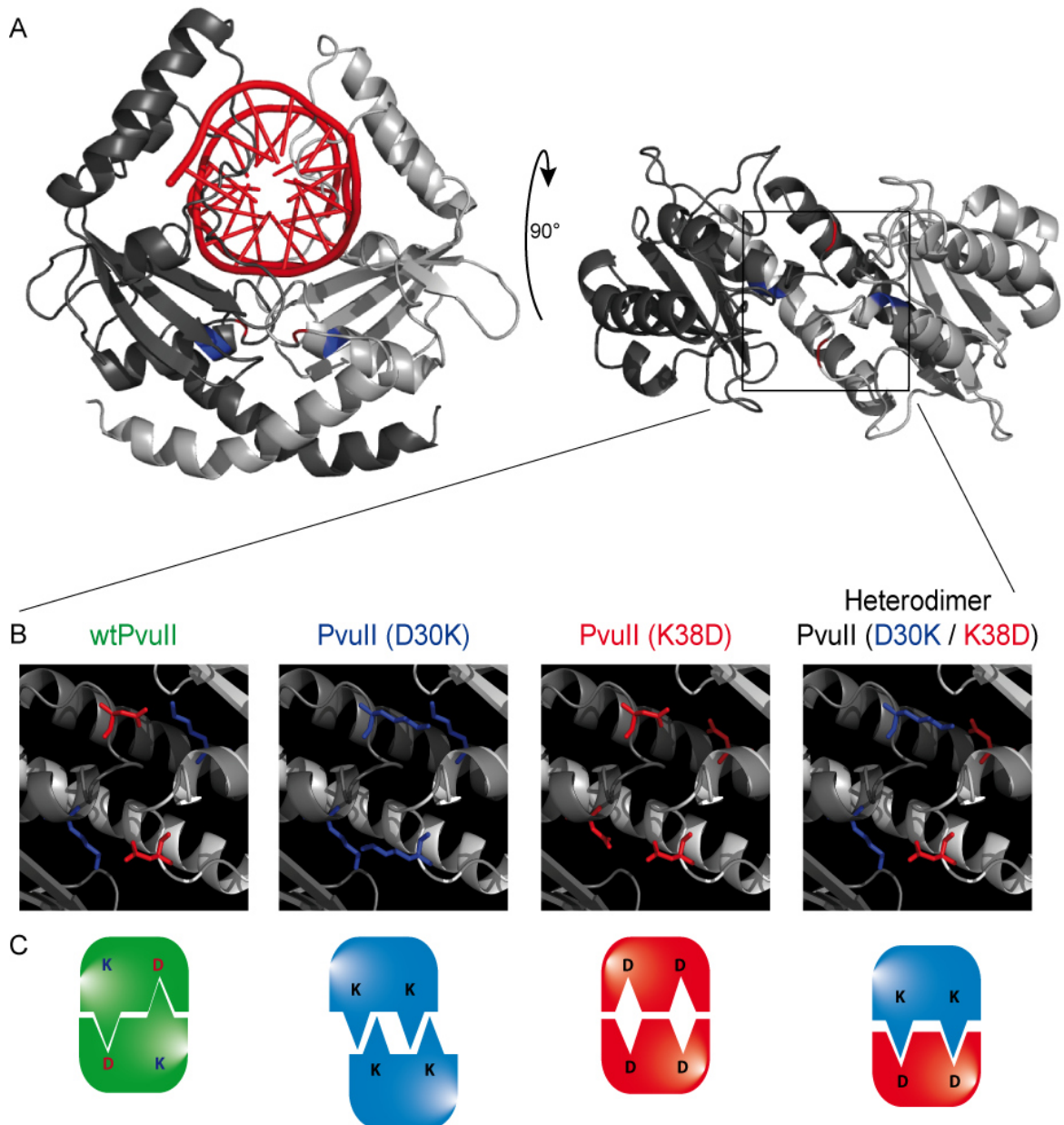


Abb. 3.3.3: PvuII-Heterodimer⁴

A PvuII-Kristallstruktur mit DNA (rot) frontal und von oben ohne DNA dargestellt (pdb code: 1PVI). **B** Vergrößerter Ausschnitt der vier Aminosäurepositionen für die Mutagenese. Die rote Seitenkette stellt die AS Asparaginsäure (D) und die blaue Lysin (K) dar. **C** Auswirkungen der einzelnen Mutationen in Bezug auf die Dimerisierung der PvuII-Untereinheiten.

Die PvuII-Mutanten D30K und K38D wurden im Fusionsprotein mit dem TALE-Protein AvrBs3 verwendet. Die homodimeren Varianten sind bei einem 50-fachen Überschuss an Protein gegenüber der DNA nicht in der Lage, das adressierte Substrat innerhalb von 2 h zu spalten (Abb. 3.3.4A). Unter Verwendung der PvuII-Varianten als Heterodimere ist bereits nach einer Minute eine deutliche Spaltung des Plasmids zu sehen (Abb. 3.3.4B). Bei den Spaltungskinetiken mit dem nicht adressierten Substrat mit nur einer PvuII-

⁴ Die Mutagenese des PvuII-Restriktionsenzym wurde von Nikolas Martin im Rahmen seiner Masterarbeit durchgeführt.

Erkennungssequenz wird mit einem 50-fachen Überschuss an PvuII-Heterodimeren innerhalb von 24 h keine Spaltung beobachtet (Abb. 3.3.4C).

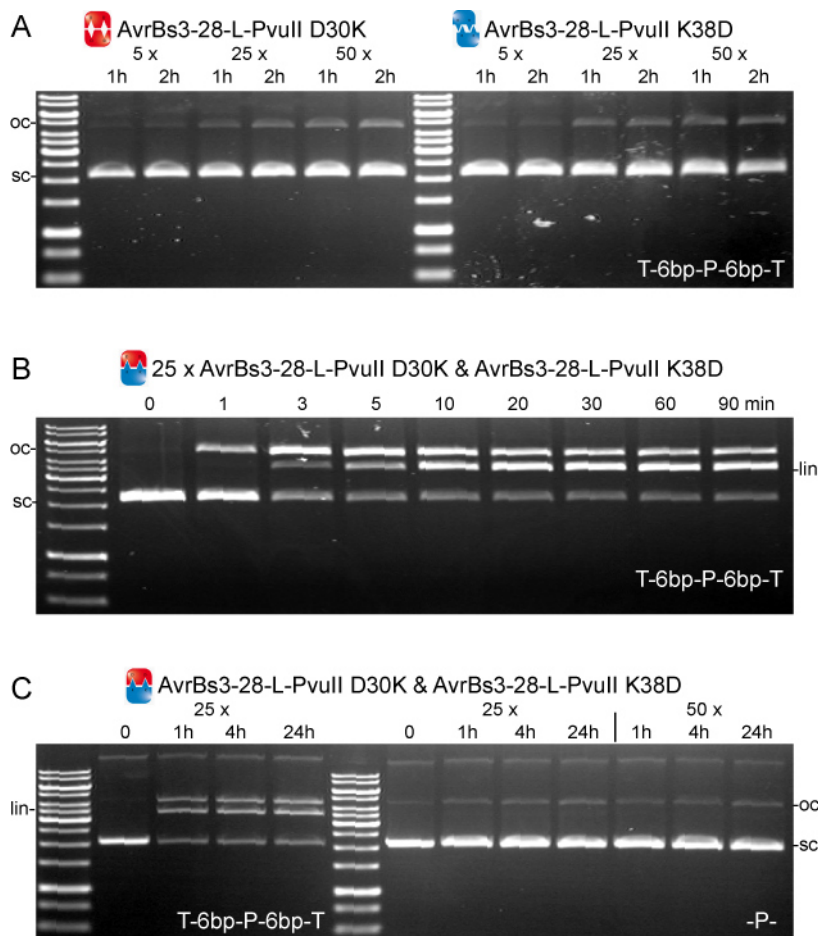


Abb. 3.3.4: Spaltungskinetiken mit heterodimeren TALE-PvuII-Konstrukten

Die Spaltungskinetiken wurden unter physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg²⁺, pH 7.5) und mit 8 nM Plasmid-DNA durchgeführt. **A** zeigt die Aktivität der homodimeren TALE-PvuII-Mutanten D30K und K38D auf dem adressierten Substrat (T-6bp-P-6bp-T) im 5-, 25- und 50-fachen Überschuss. **B** Spaltungskinetik der heterodimeren TALE-PvuII-Varianten im 25-fachen Überschuss (je Variante 200 nM) gegenüber adressiertem Substrat T-6bp-P-6bp-T. **C** das nicht-adressierte Substrat mit nur einer PvuII-Erkennungssequenz wurde bis zu 24 h mit 25- und 50-fachen Überschuss an Heterodimeren inkubiert (rechts). Zur Kontrolle der Aktivität wurde im gleichen Ansatz das adressierte Substrat unter den gleichen Bedingungen verwendet (links).

3.3.3 Aktivitätstest der heterodimeren Varianten in HEK293-Zellen

Die Aktivität der heterodimeren TALE-PvuII-Varianten ist im Vergleich zu der homodimeren Variante (Kapitel 3.2.2) deutlich geringer. *In vitro*-Spaltungskinetiken zeigen, dass mit einem 25-fachen Überschuss an Enzym keine vollständige Spaltung erzielt wird. Nachdem die heterodimeren Varianten in *in vitro*-Spaltungskinetiken aktiv waren, sollte deren *in vivo*-Aktivität überprüft werden. Hierfür wurde die Spaltungsfähigkeit der heterodimeren Konstrukte anhand von *non homologous end-joining* (NHEJ) in HEK293-

Zellen gemessen. Abb. 3.3.5 vergleicht die Aktivität der heterodimeren Varianten mit den homodimeren PvuII-Varianten aus dem Kapitel 3.2.11. Die heterodimere Variante ist sowohl *in vitro* als auch *in vivo* geringer aktiv als die homodimere (*in vivo*: 6 % Aktivität im Vergleich zu 14-20 %).

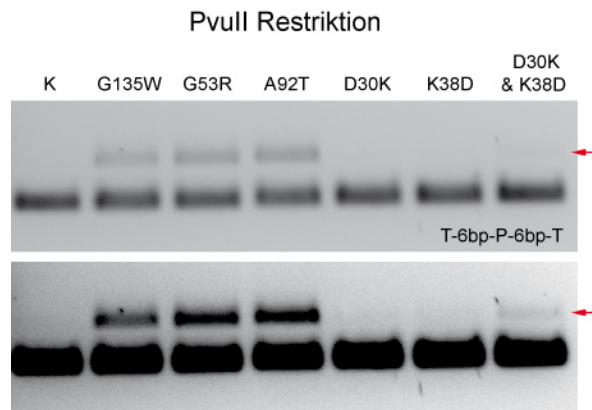


Abb. 3.3.5: *In vivo*-Aktivitätstest der heterodimeren Varianten in HEK293-Zellen

500.000 HEK293-Zellen werden zusammen mit 1 µg adressierten Substratplasmid (AvrBs3-6bp-PvuII-6bp-AvrBs3) und 4 µg TALE-PvuII-Nuklease transfiziert (je 2 µg DNA im heterodimeren Ansatz). Nach drei Tagen werden die Zellen geerntet und die DNA isoliert (QIAamp DNA Mini Kit, Qiagen). Mit dem aufgereinigten Plasmid wird eine PCR durchgeführt mit einer resultierenden DNA-Länge von 517 bp. Eine Spaltung mit PvuII führt zu den Größen von 252 bzw. 265 bp (100 ng DNA, 20 *units* PvuII-HF, NEB). Durch NHEJ wird ein Teil der DNA mit der PvuII-Erkennungssequenz verändert (roter Pfeil). Unteres Bild ist länger belichtet.

3.3.4 Toxizitätstest der heterodimeren Varianten in HEK293-Zellen

Nachdem gezeigt werden konnte, dass die heterodimeren Varianten *in vivo* eine Aktivität aufweisen, sollte des weiteren deren Toxizität überprüft und mit der Toxizität der homodimeren TALE-PvuII-Enzymen in HEK293-Zellen verglichen werden. Dies erfolgte über die Messung der Fluoreszenz des *mCherry*-Proteins, welches zusammen mit der Nukleasen-DNA in den Zellen exprimiert wurde, nach zwei und fünf Tagen. Die für die heterodimeren TALE-PvuII-Varianten D30K und K38D kodierenden Plasmide wurden zusammen in die Zellen transfiziert und zeigen die geringste Toxizität. Die Anzahl der lebenden Zellen liegt im Rahmen der Fehlergrenzen sogar geringfügig über der Anzahl der lebenden Zellen der Kontrolle mit I-SceI (Abb. 3.3.6).

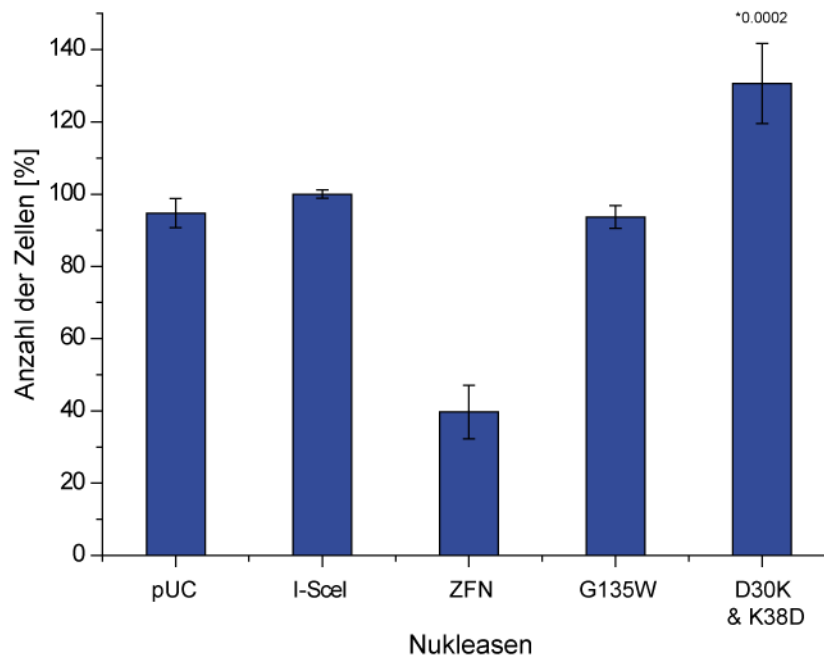


Abb. 3.3.6: Toxizitätstest mit TALE-PvuII-Heterodimeren in HEK293

120.000 HEK293-Zellen werden mit 800 ng TALE-PvuII-DNA und 100 ng *mCherry*-Expressionsvektor co-transfiziert (Bei dem heterodimeren Ansatz werden je Nuklease 400 ng eingesetzt). Die Fluoreszenz des *mCherry*-Proteins wurde nach zwei und fünf Tagen nach der Transfektion mit Hilfe von FACS (FACSCalibur, BD Biosciences) gemessen. Als Kontrollen dienen der Leervektor pUC, I-SceI (Positivkontrolle; nicht toxisch) und ZFN (Negativkontrolle; toxisch). Die Experimente wurden sechsmal wiederholt (n=6) und die Ergebnisse gegen I-SceI normiert. * gibt den *p-value* an. Die Heterodimeren werden zusammen mit der homodimeren TALE-PvuII-G135W-Variante verglichen. Die durchschnittliche relative Anzahl der Zellen für G135W und D30K&K38D liegen zwischen 94 % und 130 %.

3.3.5 Adressierung des humanen Interleukin 2 Rezeptor γ -Gens

Mit den heterodimeren TALE-PvuII-Konstrukten ist es möglich, im Genom eine höhere Spezifität zu erzielen als mit den Homodimeren. Die Adressierung episomaler DNA spiegelt jedoch nur die Aktivität der Nukleasen wider. Interessanter ist die Frage, ob die Konstrukte auch in der Lage sind, in den Zellkern einzudringen, auf dem Genom ihre Zielsequenz zu finden und diese zu spalten. Hierfür wurde das Gen des humanen Interleukin 2 Rezeptors γ (hIL2RG) ausgewählt, da dieses im Fokus von therapeutischen Maßnahmen steht. Mutationen auf dem hIL2RG-Gen führen zu einem schweren kombinierten Immundefekt (*SCID – severe combined immunodeficiency*) mit niedriger T- und B-Zellanzahl (Fischer *et al.* 2010). Das Gen befindet sich auf dem X-Chromosom und besteht aus acht Exons. Insgesamt gibt es in dem Gen vier PvuII-Erkennungssequenzen, wobei sich drei innerhalb von Exons befinden oder in deren Nähe liegen. Für die Adressierung wurde das erste Exon mit einer zentralen PvuII-Erkennungssequenz gewählt. Die entsprechenden, flankierenden TALE-Proteine wurden für diesen Lokus mit

resultierenden 8 bzw. 6 bp Abständen von Robert Morbitzer zur Verfügung gestellt (die Erkennung eines TALE-Proteins beginnt auf der Erkennungssequenz immer mit Thymin), Abb. 3.3.7A. Mit dem homodimeren Fusionsenzym AvrBs3-28-L2-PvuII-HF konnte auf einem tripartiten, analogem Substrat mit 8 und 6 bp Abstand zwischen den Erkennungssequenzen eine Spaltung *in vitro* nachgewiesen werden (Ergebnisse nicht gezeigt). Daher wurde der gleiche Abstand auch für die heterodimere Architektur verwendet. Vor der Durchführung von Zellkulturexperimenten mit den heterodimeren Nukleasen wurde die Zielsequenz als Substratplasmid für eine Spaltungskinetik ebenfalls *in vitro* getestet, um die neuen heterodimeren *ILEx 1up*- und *1 down*-TALE-PvuII-Enzyme auf ihre Aktivität hin zu überprüfen. Abb. 3.3.7B stellt die Kontrolle dar. Die Fusionsenzyme werden in Abb. 3.3.7B als Homodimere und in Abb. 3.3.7C als Heterodimere verwendet. Sowohl als Homodimer als auch als Heterodimer ist unter physiologischen Pufferbedingungen keine Aktivität zu beobachten.

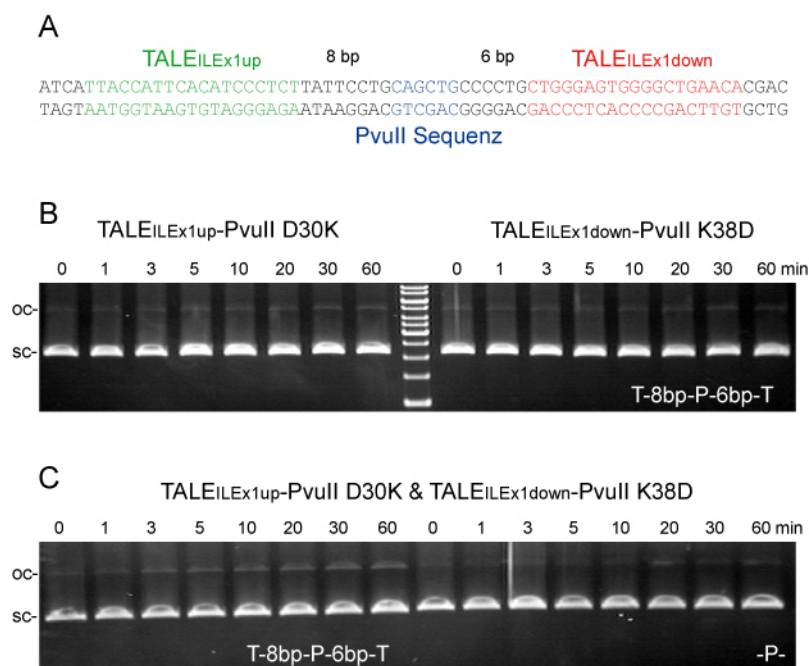


Abb. 3.3.7: Spaltungskinetik auf dem *hIL2RG*-Lokus

A Ausschnitt des *hIL2RG*-Gens im Exon 1 (NCBI Referenz: NC_000023.10; Chromosome X: 70327254 - 70331958). TALE-Proteine wurden generiert (Interleukin-Exon 1, ILEx1), die oberhalb (grün) und unterhalb (rot) der PvuII-Sequenz (blau) im Abstand von 8 und 6 bp binden. **B** zeigt die Kontroll-Spaltungskinetik mit 8 nM Plasmid-DNA aus **A** zusammen mit den homodimeren Varianten in einem 25-fachen Überschuss (200 nM). Die Kinetik wurde unter physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg²⁺, pH 7.5) durchgeführt. **C** veranschaulicht die Spaltungskinetik mit den Heterodimeren mit adressierten *hIL2RG*-Substrat und dem nicht adressierten Substrat, mit nur einer PvuII-Erkennungssequenz, unter denselben Bedingungen wie in **B**.

Die hIL2RG-Sequenz wurde auf einem Plasmidsubstrat neu generiert, so dass die TALE-Bindungssequenzen in einem symmetrischen Abstand von 4 bp von der PvuII-Sequenz entfernt liegen. Unter den gleichen Bedingungen, wie in Abb. 3.3.7, wurden die Spaltungskinetiken mit den heterodimeren und homodimeren Varianten durchgeführt. Eine schwache Aktivität mit Zunahme der linearen Bande ist nur bei den Heterodimeren *ILEx 1 up-* und *1 down-*TALE-PvuII-Enzymen zu beobachten (Abb. 3.3.8).

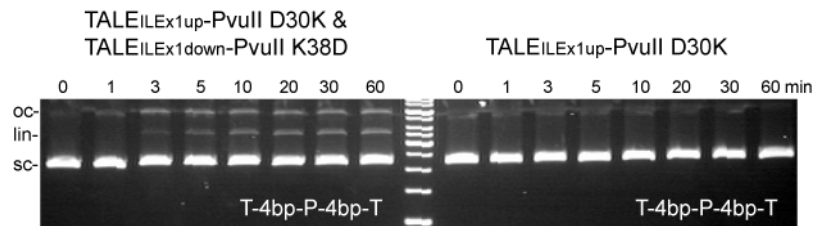


Abb. 3.3.8: Spaltungskinetik mit den hIL2RG-TALE-PvuII-Konstrukten

Der *spacer* zwischen der TALE- und PvuII-Erkennungssequenz auf dem hIL2RG-Substrat wurde von acht und sechs bp auf vier bp verringert im Vergleich zu Abb.3.3.6A. Auf diesem adressierten Substrat wurde eine Spaltungskinetik unter physiologischen Bedingungen (Ionenstärke 150 mM, 1 mM Mg²⁺, pH 7.5) mit Heterodimeren (links) und zur Kontrolle nur mit einem Homodimer TALE-PvuII-D30K (rechts) durchgeführt. 8 nM DNA wurden zusammen mit einem 25-fachen Überschuss an Enzym (200 nM) eingesetzt.

4 Diskussion

4.1 Hochspezifische Nukleasen als Werkzeuge für das *genome engineering*

Hochspezifische Nukleasen bilden innovative und sehr effektive Werkzeuge, um Gene bzw. DNA-Sequenzen in komplexen Genomen spezifisch zu adressieren und somit biologische Prozesse zu studieren (McMahon, Rahdar and Porteus 2012, Wood *et al.* 2011, Dong *et al.* 2011, Flisikowska *et al.* 2011). Neben der Gentherapie, d.h. der Heilung monogenetischer Krankheiten beim Menschen (Silva *et al.* 2011, Rouet, Smih and Jasin 1994), haben sie auch in der Pflanzenzüchtung enorme Vorteile. Hier ist eine Steigerung der Produktivität möglich, da Gene mit positiven Eigenschaften schneller als durch einen Züchtungsprozess kombiniert werden können (Lusser *et al.* 2012, Podevin *et al.* 2012). Die Modifikation von Genomen mit artifiziellen hochspezifischen Nukleasen wurde vom *Nature Methods Journal* zur Methode des Jahres 2011 gekürt (Methods 2012). Eine Vielzahl verschiedener Nukleasen mit unterschiedlichen Architekturen wurden für diesen Zweck bereits generiert (Abb. 1.2). Eine Gemeinsamkeit dieser Nukleasen ist die ausreichend lange Erkennungssequenz, welche nur einmal im Genom vorkommen sollte, damit eine spezifische, adressierte DNA-Spaltung gewährleistet ist. Die Erkennungssequenz und die damit verbundene Spezifität sind abhängig von der Genomgröße. Ein Restriktionsenzym mit einer Erkennungssequenz von 6 Basenpaaren (bp) würde statistisch gesehen alle 4.096 bp (4 Basen; 4^6) binden und spalten. Im humanen Genom mit mehr als 3 Milliarden Basen ($3 \cdot 10^9$ bp) würde ein Restriktionsenzym über 700.000 mal schneiden. Für eine singuläre, spezifische Spaltung bräuchte man im humanen Genom statistisch gesehen Nukleasen mit einer Erkennungssequenz von mindestens 16 bp ($3 \cdot 10^9$ bp / Nuklease 4^{16} bp ≈ 0.7).

ZFN (Zink-Finger-Nukleasen) und TALEN (*transcription-activator-like effector nucleases*) zählen zu den am meisten verwendeten programmierbaren Nukleasen. Beide unterscheiden sich in ihren DNA-Bindungsmodulen, besitzen aber stets das gleiche DNA-Spaltmodul: die unspezifische, katalytisch aktive Nukleasedomäne des Typ II-S-Restriktionsenzym FokI. FokI liegt in Lösung als monomeres Protein vor (Kaczorowski, Skowron and Podhajska 1989, Vanamee *et al.* 2001). Die DNA-bindende Domäne ist über einen flexiblen Linker mit der unspezifischen Nukleasedomäne verbunden. Für die katalytische Aktivität ist die Dimerisierung von FokI-

Nukleasedomänen auf der DNA erforderlich (Bitinaite *et al.* 1998). Die Bildung des Dimers kann unterschiedlich erfolgen (Pernstich and Halford 2012). Zum einen kann eine FokI-Untereinheit spezifisch an die DNA-Sequenz GGATG binden und rekrutiert anschließend eine zweite, ungebundene FokI-Nukleasedomäne aus der Lösung, wodurch eine katalytisch aktive Nuklease entsteht (Abb. 4. 1A). Die Spaltung erfolgt in Anwesenheit von Mg^{2+} . Zum anderen kann das Dimer bei Anwesenheit von zwei oder mehreren Erkennungssequenzen über zwei gebundene FokI-Untereinheiten gebildet werden, wodurch die DNA zwischen den beiden Erkennungssequenzen über eine *loop*-Struktur ausgestülpt und gespalten wird (Abb. 4. 1B) (Catto *et al.* 2006, Rusling *et al.* 2012).

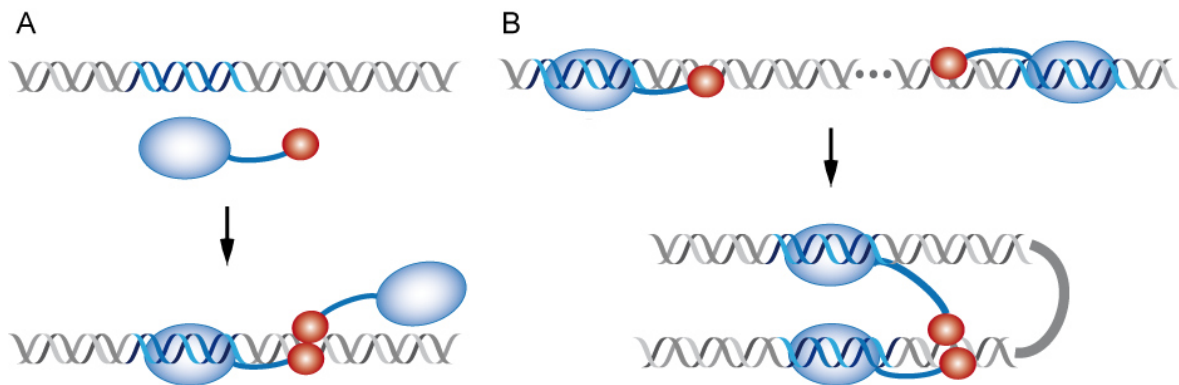


Abb. 4. 1: FokI-Reaktionsschema

A Eine FokI-Untereinheit (UE) bindet mit der nativen Bindedomäne (blau) an die zugehörige DNA-Erkennungssequenz. Eine zweite nicht gebundene UE bildet über die Nukleasedomäne (rot) ein katalytisch aktives Dimer. **B** Zwei gebundene FokI-UE können über eine DNA-*loop*-Bildung Distanzen überwinden und dimerisieren; verändert nach (Halford *et al.* 2011).

Die zwei unterschiedlichen Mechanismen der Dimerisierung verdeutlichen, dass hochspezifische Nukleasen mit FokI als Spaltmodul auch dann eine DNA-Spaltung durchführen können, wenn die Nuklease nur „halbseitig“ an der Erkennungssequenz gebunden hat. Im Falle von ZF-Modulen sind damit erhebliche Spezifitätsverluste verbunden, da die Bindung beispielsweise dreier ZF-Module einer Untereinheit neun, statt 18 Basenpaare abdeckt und keine genaue Adressierung gewährleisten würde. Binden beide Untereinheiten in einem größeren Abstand auf der DNA (unspezifische Bindung), ist FokI dennoch in der Lage, die DNA aufgrund der *loop*-Struktur, zu spalten. Diese Mechanismen werden als gentoxische Aktivität oder *off-target cleavage* bezeichnet (Halford *et al.* 2011). In dieser Arbeit wird für das *gene targeting* eine neue Architektur für hochspezifische Nukleasen eingeführt, mit spezifischen Restriktionsenzymen als Spaltmodul anstelle der unspezifischen FokI-Nukleasedomäne.

Im Gegensatz zu FokI, sind BfiI und PvuII bereits dimere Proteine in Lösung (Athanasiadis *et al.* 1994, Cheng *et al.* 1994, Lagunavicius *et al.* 2003). Diese Eigenschaft reduziert eine potenzielle Quelle der gentoxischen Aktivität. Die Toxizität FokI-basierter Nukleasen konnte zwar durch die Generierung von obligat-heterodimeren FokI-Varianten erheblich gesenkt, aber nicht vollständig gelöst werden (Miller *et al.* 2007, Szczepek *et al.* 2007, Pingoud and Silva 2007)). Die *off-target cleavage*-Aktivität optimierter ZFN wurden 2011 ausführlich analysiert (Gabriel *et al.* 2011; Pattanayak *et al.* 2011).

4.2 Generierung des BfiI-I-SceI-Fusionsenzym

Die außergewöhnliche Aktivitätskontrolle der Katalyse des BfiI-Restriktionsenzym war das Hauptkriterium um BfiI als Spaltmodul in einem Nuklease-Fusionsansatz zu verwenden. Als Bindungsmodul wurde die bereits bewährte inaktivierte I-SceI Homing-Endonuklease verwendet (Fonfara 2011). Der N-Terminus von BfiI bildet die unspezifische Nukleasedomäne und ist ungeeignet für eine Fusion, da sonst der Autoregulationsmechanismus des internen Peptidlinkers (Linker zwischen der Binde- und Nukleasedomäne) nicht gegeben wäre. Des Weiteren bildet sich das Dimer über den N-Terminus aus und ist daher unzugänglich für eine Fusion. Die Fusion erfolgte direkt am C-Terminus an der DNA-Bindedomäne von BfiI, ohne diese dabei zu deletieren oder komplett durch das I-SceI-Bindungsmodul zu ersetzen. Unveröffentlichte Experimente der Arbeitsgruppe von Virginijus Siksnys (Institut für Biotechnologie, Vilnius, Litauen) ergaben einen Verlust der Autoinhibition des Linker bei Deletionen in der Bindedomäne. Um zunächst zu überprüfen, ob BfiI zusammen mit I-SceI in der Lage ist, die DNA zu spalten, wurde neben einer direkten Fusion (2 AS- Linker) weitere, längere Linkervarianten (5 bzw. 10 AS) gewählt, um einen optimalen Abstand zwischen den Proteinen für eine Kooperation zu ermitteln. Ein zu kurzer Linker könnte die Funktion eines Proteins stark einschränken und ein zu langer würde zu DNA-bindungsunabhängigen Reaktionen führen. Alle Linkervarianten sind in der Lage, BfiI-Sequenzen zu erkennen und zu spalten. Für die unterschiedlichen Linkerlängen musste der optimale *spacer* auf der DNA zwischen der I-SceI- und BfiI-Erkennungssequenz ermittelt werden. Durch die Generierung von „halbseitigen“ DNA-Substraten (Bindungssequenzen für I-SceI und BfiI) mit unterschiedlichen *spacern* zwischen den Erkennungssequenzen konnte keine Präferenz gegenüber einer isolierten BfiI-Erkennungssequenz ermittelt werden. Jede BfiI-Sequenz wurde gespalten (Abb. 3.1.1). Aufgrund der räumlichen Trennung der BfiI-Domänen und der Tatsache, dass lediglich die katalytisch aktive Domäne benötigt wird, folgte eine

gerichtete Mutagenese, um die Inaktivierung der Bindungseigenschaften der BfiI-Bindedomäne mit den Mutationen R212A und N279A zu erreichen. Hiermit sollte in einem zweiten Ansatz überprüft werden, ob sich die allosterische Aktivierung der Nukleasedomäne ausschließlich über die Bindung von I-SceI steuern lässt. Ein weiterer Vorteil der Mutagenese ist, dass Abbruchprodukte, die bei der Expression eines Fusionsenzym zu einer freien BfiI-Nuklease führen können, nicht mehr „aktiv“ sind. Es stellte sich heraus, dass die BfiI-Bindedomäne eine große Rolle bei der Allosterie und dem „Flip“ im Reaktionsmechanismus einnimmt (Kapitel 1.4.1), da durch die alleinige Bindung von I-SceI auf dem halbseitigen Substrat keine oder kaum Aktivität beobachtet werden konnte (Abb. 3.1.3).

Der dritte Ansatz, die Nukleasedomäne zu kontrollieren, ist eher „mechanischer“ Natur. Durch zwei *cis*-I-SceI-Erkennungssequenzen auf dem DNA-Substrat, sollen die internen Linker „mechanisch gestreckt“ werden und das aktive Zentrum freigeben, sobald beide I-SceI-Proteine auf der DNA binden. Im ersten Spaltungsexperiment (Abb. 3.1.2) zeigt nur die 10AS-Linkervariante Aktivität auf dem *cis*-Substrat mit 18 bp *spacer*. Die Kontrollen zeigen, dass es sich um *nicking*-Aktivität handelt, wobei der untere Strang mehrmals gespalten wird (Abb. 3.1.3 und Abb. 3.1.8). Hier konnte gezeigt werden, dass es gelungen ist, die katalytische Aktivität der BfiI-Nuklease, unabhängig von BfiI-Erkennungssequenzen, nur mit Hilfe von beiden gebundenen I-SceI-Bindungsmodulen zu steuern. Diese Regulation konnte auch mit den Bindungsmutanten R212A und N279A erzielt werden, ohne dass die katalytische Aktivität beeinträchtigt wird (Abb. 3.1.7). Die Auswahl der Punktmutationen für die Bindungsmutanten resultiert aus nicht veröffentlichten Experimenten der Arbeitsgruppe von Virginijus Siksnys (Institut für Biotechnologie, Vilnius, Litauen). Unter Verwendung von zwei weiteren *cis* I-SceI-Substratplasmiden, mit einem *spacer* von 24 und 30 bp, konnte keine Aktivität mit dem Fusionsenzym BfiI-10AS-I-SceI detektiert werden (Abb. 3.1.4). Hypothetisch gesehen hätte man auf allen drei Substraten keine Aktivität erwartet, da die *spacer*-Länge nicht ausreichend lang für eine „mechanische“ Streckung ist (Abb. 4.2). BfiI hat im geschlossenen Zustand einen Durchmesser von ca. 103 Å, was ungefähr 32 bp entspricht, demnach müsste das Substrat einen *spacer* von mehr als 32 bp aufweisen, um genügend Raum für eine „Streckung“ zu bieten.

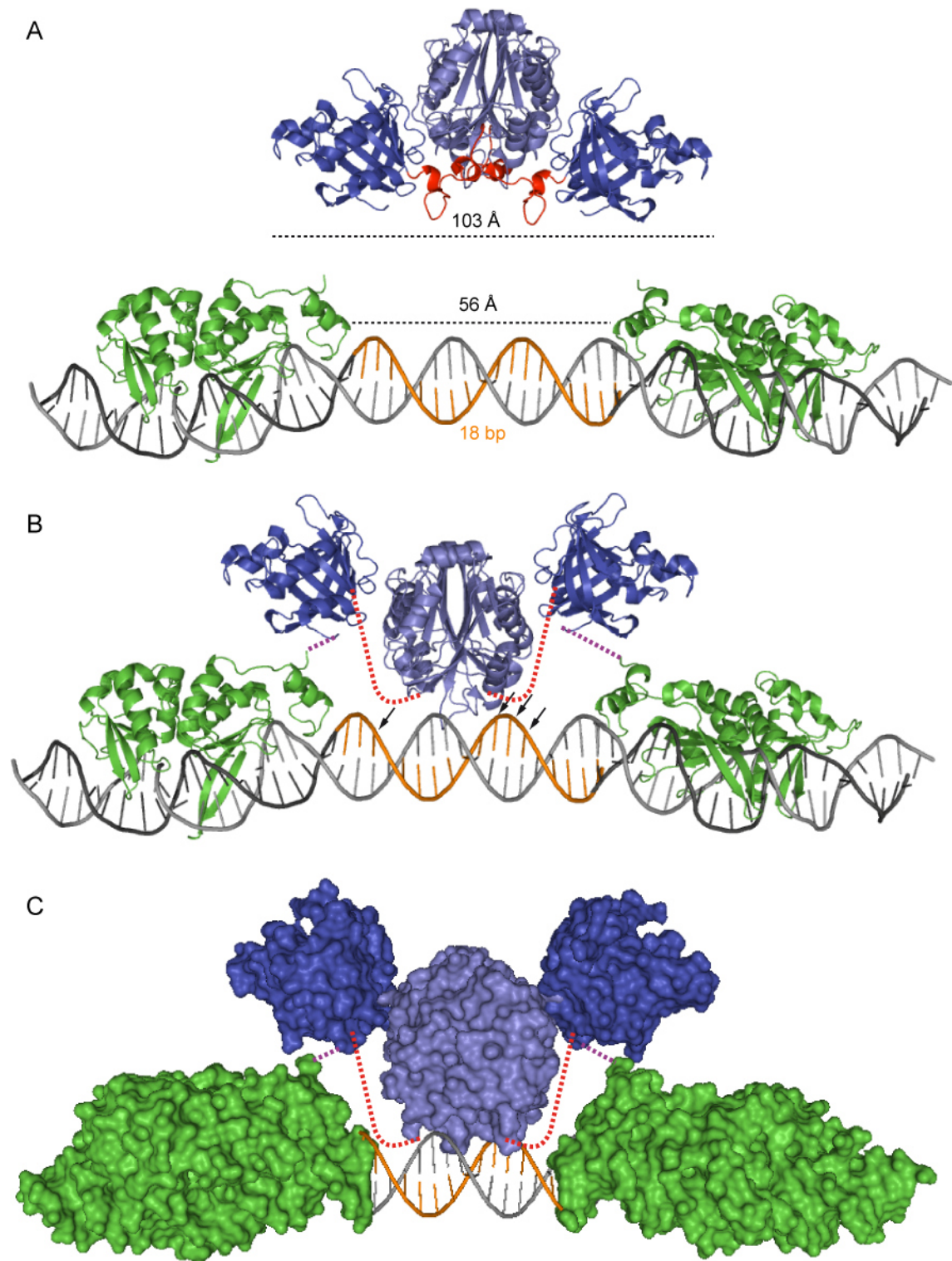


Abb. 4.2: BfiI-I-SceI-cis-Modell

A Größenverhältnisse des BfiI-Apoenzym und der auf der DNA gebundenen I-SceI-Proteine. Der orange markierte DNA-Strang entspricht dem gespaltenen Strang. Der Peptidinterne Linker ist rot dargestellt. **B** Modell des BfiI-I-SceI-Fusionsenzym und der möglichen Interaktion mit der DNA. Die Pfeile markieren die beobachteten Spaltstellen. Die „Streckung“ des internen Linkers wird mit roten Punkten verdeutlicht. Der Linker zwischen den beiden Proteinen ist lila dargestellt. **C** Modell des BfiI-10AS-I-SceI-Enzyms mit der *surface*-Darstellung.

Dieses überraschende Ergebnis warf die Frage auf, wie die Aktivität auf einem Substrat mit 40 bp *spacer* aussehen würde und wie die Spaltung auf dem Substrat mit 18 bp *spacer* erfolgt. Das Modell in der Abb. 4.2 zeigt eine „mechanische“ Streckung nicht in horizontaler, aber dafür in vertikaler Orientierung. Die beiden BfiI-Bindedomänen kommen nicht in Kontakt mit der DNA, allein durch die Flexibilität des internen Linkers kommt es zur Interaktion zwischen der Nukleasedomäne und der DNA. Da der „Flip“-Mechanismus durch die beiden I-SceI-Proteine größtenteils inhibiert ist, kann die Nuklease aufgrund der Orientierung nur einen Strang, diesen jedoch mehrmals, spalten (zugänglicher Einzelstrang ist in orange dargestellt). Die Frage, wie die Aktivität auf dem 40 bp *spacer*-Substrat oder die Aktivitäten der anderen BfiI-Linkervarianten (2 AS und 5 AS) auf dem verschiedenen *cis*-Substraten aussehen würde, steht nicht im Fokus dieser Arbeit.

Es ist gelungen, trotz der neuen Bindungsmodule, die native Regulation der Autoinhibition der BfiI-Nuklease beizubehalten. Unabhängig von BfiI-Erkennungssequenzen wurden gewünschte Spaltungen an definierten Stellen durchgeführt.

4.3 Generierung des AvrBs3-PvuII-Fusionsenzym

In vorangegangenen Arbeiten konnte gezeigt werden, dass Zink-Finger-Konstrukte mit FokI-Nukleasedomänen durch PvuII ersetzt werden können (Schierling *et al.* 2012). In der vorliegenden Arbeit wird gezeigt, dass die Substitution von FokI mit PvuII ebenfalls in TALEN funktioniert und den Vorteil der generellen Spezifität von Restriktionsenzymen mit dem TAL-Effektor kombiniert. Die unspezifische Nukleasedomäne von FokI lässt sich mit der repetitiven DNA-Bindedomäne des TALE-Proteins zu 100 % programmieren. Jedoch treten toxische Nebeneffekte auf, wie im Kapitel 4.1 beschrieben. PvuII braucht immer die eigene Erkennungssequenz CAGCTG, um katalytisch aktiv zu werden. Dies gilt ebenfalls in Kombination mit TALE-Proteinen. Deshalb muss ein entsprechendes TALE-Protein immer in Abhängigkeit einer PvuII-Sequenz generiert werden, was kein Nachteil sein muss, da die 6 bp lange Sequenz im Mittel alle 4.096 bp (4^6) im Genom vorkommt und es ausreichend für eine homologe Rekombination (HR) ist, wenn in der Nähe einer Mutation gespalten wird. Dieses Extraelement der Spezifität konnte schon bei den ZFN mit PvuII (Schierling *et al.* 2012), wie auch in dieser Arbeit gezeigt werden. Die Spezifitätsvorteile gegenüber FokI-Konstrukten überwiegen (Abb. 3.2.15).

Die Herausforderung bei der Generierung von TALE-PvuII-Konstrukten ist, dass das PvuII-Enzym im Fusionsansatz nicht die - statistisch gesehen - alle 4.096 bp

vorkommenden PvuII-Erkennungssequenzen spaltet, sondern nur die eine gewünschte adressierte TALE-PvuII-Erkennungssequenz in einem komplexen Genom. Um dies zu erreichen, gibt es im Prinzip zwei Strategien. Die DNA-Affinität des PvuII-Enzyms kann durch Aminosäuresubstitution gemindert werden, so dass eine verstärkte Abhängigkeit zum Bindungsprotein TALE entsteht und PvuII nur eine Bindung mit der DNA eingehen kann, wenn TALE gebunden hat. Die andere Möglichkeit besteht darin, einen Linker zwischen dem Bindungsmodul TALE und dem Spaltmodul PvuII so zu generieren, dass PvuII nicht in der Lage ist, mit der DNA zu interagieren bis das TALE-Protein an die eigene Zielsequenz gebunden hat und somit PvuII von der inhibierenden Konformation löst.

4.3.1 Spezifität der TALE-PvuII-Konstrukte

PvuII ist nativ ein homodimeres Protein. Als Fusionsprotein mit der TALE-DNA-Bindungsregion bildet es zwei identische TALE-Bindungsproteine mit einer Größe von ca. 203 kDa aus. Die Adressierung genomischer Sequenzen mit einem TALE-PvuII Homodimer erweist sich als nicht besonders vorteilhaft, da die flankierenden Sequenzen einer PvuII-Erkennungssequenz mit einer hohen Wahrscheinlichkeit nie identisch sein werden. Dieses Problem kann mit der Generierung von heterodimeren Konstrukten oder eines monomer agierenden Enzyms behoben werden. Mit der *single-chain(sc)*-PvuII-Variante lassen sich, unter der Voraussetzung, dass das Bindungsmodul alleine die Spezifität erfüllt, monomere hochspezifische Nukleasen generieren.

Verschiedene Versionen von homodimeren und monomeren TALE-PvuII-Nukleasen wurden in dieser Arbeit generiert.

Die TALE-PvuII-Konstrukte wurden auf ihre Spezifität hin auf adressierten tripartiten Substraten, welche eine PvuII-Erkennungssequenz zwischen zwei TALE-AvrBs3-Sequenzen besitzen, und auf nicht-adressierten Substraten mit nur einer PvuII-Sequenz überprüft. Die Variante mit einem kürzeren 28-L Linker erwies sich als spezifischer als die lange Linkervariante 63-L (Abb. 3.2.3). Beim direkten Vergleich des kurzen Linkers im homodimeren und monomeren PvuII-Konstrukt erwies sich die TALE-scPvuII-Variante unspezifischer, mit einer höheren Aktivität auf den nicht-adressierten Substraten (Abb. 3.2.4), woraufhin alle weiteren Experimente mit den homodimeren Konstrukten fortgeführt wurden. Unter physiologischen Bedingungen mit einer Ionenstärke von rund 150 mM konnte keine unspezifische Spaltung mit einem Überschuss an Enzym gegenüber DNA und langen Inkubationszeiten detektiert werden (Abb. 3.2.5). Die Spezifität des Enzyms

wurde mit Hilfe von radioaktiv-markierten adressierten und nicht-adressierten Substraten quantitativ ermittelt, mit dem Resultat einer >34.000-fachen Präferenz gegenüber der adressierten Erkennungssequenz (Abb. 3.2.6). Selbst bei einem 40-fachen Überschuss ist kaum eine unspezifische Aktivität in einem Zeitraum von 24 h zu beobachten (Abb. 3.2.5). Die sterische Hinderung des PvuII-Enzyms durch TALE könnte ein Grund sein, weshalb isolierte PvuII-Erkennungssequenzen vom TALE-PvuII-Fusionsenzym nicht gespalten werden. Die PvuII-T46G-HF-Mutante mit eingeschränkten Bindungseigenschaften (Zhu 2009) in Kombination mit einem kurzen Linker wird höchst wahrscheinlich daran gehindert, alleine einen PvuII-DNA-Komplex einzugehen bzw. einen Übergangszustand bei der Katalyse zu bilden, bis das TALE-Protein an die eigene Zielsequenz gebunden hat. Die konformationsbeeinflussende Eigenschaft des TALE-Proteins auf PvuII kann auch erklären, weshalb die TALE-scPvuII-Architektur unspezifischer ist. Hier kann ein TALE-Protein höchstwahrscheinlich nur eine PvuII-Untereinheit beeinflussen, weshalb die nicht adressierte Spaltung von PvuII-Sequenzen nicht vollständig verhindert werden kann.

Das TALE-PvuII-Fusionsprotein bevorzugt 4 und 6 bp *spacer* auf den adressierten Substraten (Abstand zwischen der TALE- und PvuII-Erkennungssequenz), wobei die Fusion mit dem kürzesten Linker-TALE-L-PvuII einen *spacer* von 2 und 4 bp bevorzugt (Abb. 3.2.8). In der Abb. 3.2.5B ist zum Vergleich die Spaltungskinetik der TALE-28-L-PvuII-P14G-Variante aufgeführt. Dieses Fusionsenzym erwies sich in Kombination mit dem Bindungsprotein I-SceI als eine sehr spezifische Variante (Fonfara 2011). In Kombination mit dem TALE-Protein und dem kurzen 28-L Linker erwies sie sich jedoch als sehr unspezifisch. Das könnte an zwei Gründen liegen: Zum einen wurde die Fusion des Bindungsmoduls I-SceI am C-Terminus von PvuII durchgeführt, die Bindung an TALE-Proteine hingegen am N-Terminus. Die P14G-Mutation könnte deshalb eine andere Auswirkung in dieser Architektur bzw. Orientierung als im TALE-PvuII-Kontext besitzen. Zum anderen wurden die Spaltungskinetiken in einem Puffer mit einer Ionenstärke von mehr als 225 mM (100 mM Kaliumglutamat, 25 mM Tris-Acetat, 100 mM Kaliumchlorid, nicht mit einbezogen sind die Ionen aus den Enzymverdünnungen und den Mg²⁺-Ionen) durchgeführt. Mit hohen Salzkonzentrationen kann die Interaktion von Enzym und DNA vermindert werden, da die Ionen in Konkurrenz mit den Proteinen wechselwirken können. Durch eine hohe Ionenstärke ist es möglich, die Bindung an adressierten, längeren Erkennungssequenzen zu fördern und kürzere, nicht-adressierte Sequenzen zu inhibieren. Die PvuII-Varianten müssen daher an die jeweilige Nuklease-Architektur angepasst werden (siehe Fonfara *et al.* 2012, Schierling *et al.* 2012).

Die meisten Aktivitätsexperimente wurden mit dem TALE-28-L-PvuII-Enzym auf adressierten Substraten mit 6 bp Abstand durchgeführt, da diese Länge bei den ersten Experimenten vielversprechende Resultate erbrachte.

4.3.2 *multi-turnover*-Eigenschaften von TALE-PvuII

Generell wird bei hochspezifischen Fusionsnukleasen eine geringere Spaltungsrate im Vergleich zu den freien Nukleasen beobachtet. Dies könnte mit der Architektur bzw. mit der Struktur erklärt werden, die bewirkt, dass keine ungehinderte Kooperation zwischen den beiden Modulen der Fusionsproteine gewährleistet ist. Des Weiteren besitzen Enzyme mit langen Erkennungssequenzen, wie beispielsweise Homing-Endonukleasen, *single-turnover*-Eigenschaften, so dass sie an ihr Ziel binden und nach der Spaltung am Produkt gebunden bleiben. Daher wird bei Spaltungskinetiken mindestens die gleiche Enzymkonzentration wie die Substratkonzentration verwendet, um eine 100 %-ige Spaltung zu gewährleisten. Native Nukleasen sind bei Spaltungskinetiken mit einem 1000-fachen Überschuss in der Lage, innerhalb von Minuten die Spaltung ihrer Substrate zu katalysieren. Diese hohe Aktivität ist bei Fusionsenzymen nicht zu beobachten. Neben der Erwartung, dass sie an ihrer Zielsequenz gebunden bleiben, kommen noch weitere Faktoren hinzu, die einen Enzymüberschuss bei Spaltungskinetiken erzwingen, um somit innerhalb einer bestimmten Zeit eine vollständige Spaltung zu gewährleisten (siehe Fonfara *et al.* 2012, Schierling *et al.* 2012, Gabsalilow *et al.* 2013). Dazu zählt die Reinheit der Enzyme, um möglichst genaue Konzentrationen zu berechnen und die Tatsache, dass eventuell ein gewisser Prozentsatz des Enzyms inaktiv sein könnte und dadurch einen Überschuss bei den Reaktionen erfordert.

Die *multi-turnover*-Eigenschaft des homodimeren TALE-28-L-PvuII-HF-Fusionsenzym wird in der Abb. 3.2.5 deutlich. Die Spaltungskinetiken im Unter- und Überschuss auf einem adressierten Substratplasmid zeigen deutlich eine vollständige Spaltung innerhalb von wenigen Minuten. Im Unterschuss wurde nur ein Viertel der Enzymkonzentration gegenüber der DNA-Substratkonzentration eingesetzt. Es wurde jedoch eine 100 %-ige Spaltung statt den erwarteten 25 % bei einer *single turnover*-Aktivität beobachtet, was auf eine *multi-turnover*-Eigenschaft zurückzuführen ist. Um dieses Ergebnis zu validieren, wurde eine sensitivere Spaltungskinetik mit radioaktiv markierten Substraten durchgeführt (Abb. 3.2.7). Ein Zehntel der Enzymkonzentration spaltete innerhalb einer Stunde mehr als 60 % der Substrat-DNA und bestätigt die *multi-turnover*-Eigenschaft. Mit Bindungsexperimenten, unter anderem mit *single-molecule*-Studien, wäre es möglich

genau zu analysieren, wie die TAL-Effektoren auf der DNA diffundieren, an die kognate und nicht kognate DNA binden, um eine Erklärung für die *multi-turnover*-Eigenschaft zu geben. Erste Experimente zeigen, dass der N-terminale Bereich essentiell für die Bindung ist und auch verantwortlich für die Interaktion mit nicht-kognater DNA (Gao *et al.* 2012, Meckler *et al.* 2013). PvuII wurde an den C-Terminus vom TALE-Protein fusioniert, wodurch der N-terminale Teil des TALE-Proteins mit der DNA wechselwirken kann, ohne vom PvuII-Enzym negativ beeinflusst zu werden. Erst wenn das TALE-Protein an die eigene Erkennungssequenz bindet, kommt es zur DNA-Spaltung durch PvuII.

In Anbetracht dieser Tatsache wird die Annahme, dass das TALE-Protein konformationsverändernde Eigenschaften auf PvuII haben muss, unterstützt, da bei einem *multi-turnover*-Verhalten eine höhere Wahrscheinlichkeit gegeben ist, mit nicht-adressierten PvuII-Erkennungssequenzen in Kontakt zu treten. Für *in vivo*-Anwendungen ist eine *multi-turnover*-Eigenschaft zwar nicht von Nöten, da in der Zelle ein hoher Überschuss des Enzyms vorliegt. Die konformationsbeeinflussende Eigenschaft des TALE-Proteins auf PvuII ist dennoch sehr von Vorteil.

4.3.3 Spaltungskinetiken auf bipartiten Substraten

Bei FokI-basierten Nukleasen wurde bereits der Nachteil verdeutlicht, dass eine gebundene Untereinheit eine zweite ungebundene rekrutieren und ein aktives Dimer bilden kann. Durch diese Bindung an „halbseitige“, bipartite Erkennungssequenzen, statt an einer tripartiten wird ein Großteil der Spezifität eingebüßt und dadurch genotoxische Aktivität provoziert, wie es speziell bei ZFN der Fall ist. Die TALE-PvuII-Fusionen adressieren und spalten ebenfalls bipartite Sequenzen und bevorzugen dabei überspiralisierte Plasmid-DNA (Abb. 3.2.9). Allerdings ist die Spaltungsrate, verglichen mit tripartiten Sequenzen, geringer. Vermutlich sind die Bindung eines Bindungsmoduls und die nicht-spezifische Interaktion des zweiten Bindungsmoduls ausreichend, PvuII teilweise aus der inhibierenden Konformation zu lösen. Diese Hypothese würde sowohl die geringe Spaltungsrate und die Aktivität des halbseitig gebundenen TALE-PvuII-Homodimers erklären, als auch die höhere unspezifischere Aktivität der TALE-scPvuII-Konstrukte (Abb. 3.2.4). Das TALE-PvuII-Dimer ist im Gegensatz zu FokI-basierten Nukleasen ein Dimer in Lösung. Somit könnte die Bindung eines TALE-Proteins im komplexen Genom mit einer 19 bp langen Erkennungssequenz ausreichen, um spezifisch mit der DNA zu interagieren. Diese Strategie würde mit TALE-FokI-Konstrukten nicht funktionieren, da die FokI-Untereinheiten als Monomere in Lösung vorliegen und auf Grund ihres

katalytischen Mechanismus die DNA in einer *loop*-Konformation spalten können, falls die beiden Untereinheiten entfernt voneinander auf der DNA binden (Kapitel 4.1).

4.3.4 Physiologische Puffer und Stabilität des TALE-PvuII-Enzyms

Die anfänglichen Aktivitätstests der TALE-PvuII-Enzyme wurden unter optimalen Pufferbedingungen für das PvuII-Enzym im NEB4-Acetat-Puffer (New England Biolabs, Puffer Nr. 4; 20 mM Tris-Ac, 50 mM K-Ac, 10 mM Mg-Ac, 1 mM DTT, pH 7.9) durchgeführt. Basierend auf diesem Puffer, wurden die Ionen- und Mg^{2+} -Konzentrationen annähernd den physiologischen Bedingungen angeglichen, um die Spaltungskinetiken und Aktivitätstest unter diesen Bedingungen zu charakterisieren. Nahezu gleiche Resultate konnten in einem zweiten annähernd physiologischen Puffer (KGB-Puffer; McClelland *et al.* 1988) mit unterschiedlicher Zusammensetzung erhalten werden (Abb. 3.2.11). Diese Ergebnisse verdeutlichen, dass insbesondere die Ionenstärke und Mg^{2+} -Konzentration für die Aktivität des TALE-PvuII-Enzyms ausschlaggebend sind.

Das TALE-PvuII-Enzym wies unter physiologischen Bedingungen nach 4 h Inkubation bei 37°C eine ~50 %-ige Aktivität auf, nach 24 h ist es vollständig inaktiv (Abb. 3.2.10). Diese Eigenschaft ist für *gene targeting*-Experimente als positiv zu werten, denn es ist gewünscht, dass die Nukleasen nicht sehr lange aktiv sind, weil sie sonst in Zellen toxische Reaktionen hervorrufen könnten. Eine direkte Übertragung der *in vitro* ermittelten Daten auf Zellen und deren Bedingungen ist nicht möglich, zumal es keinen direkten Einfluss auf die Transkriptions- und Translationsrate gibt, dennoch sind die *in vitro* ermittelten Daten nützlich bei der Charakterisierung der Nukleasen.

4.3.5 Viabilitätsexperimente / Selektion neuer Varianten

Die zuvor *in vitro* ermittelte hohe Spezifität des TALE-28-L-PvuII-HF-Enzyms sollte erneut in *E. coli*-Zellen ohne den PvuII-Methyltransferase-Schutz validiert werden. Jedoch erwies sich dieses Konstrukt als zu toxisch in den *E. coli*-Zellen, mit einer unerwartet niedrigen Überlebensrate von ca. 2 %,

Abb. 3.2.12). Trotz der Tatsache, dass die Induktion der Expression nicht eingeleitet wurde, scheint die Hintergrundexpression durch den T5-Promotor sehr hoch bzw. das Enzym sehr toxisch zu sein. Des Weiteren kann nicht ausgeschlossen werden, dass während der Expression PvuII-Enzyme ohne ihre DNA-bindendes Bindungsmodul entstehen.

Im Falle von eukaryotischen Zellen können der Zellkern und die Chromatinstruktur die Bindung der Nukleasen an die genomische DNA erschweren und eventuell die Toxizität bzw. die Aktivität beeinflussen. Daher muss die Spezifität der artifiziellen Nukleasen in eukaryotischen Zellen ermittelt werden. Für diesen Zweck wurde die TALE-PvuII-Nuklease in ein eukaryotisches Vektorsystem mit einem CMV-Promotor umkloniert. Eine basale niedrige Hintergrundexpression eines toxischen Produkts durch den CMV-Promotor (oder eines entfernteren prokaryotischen Promotors auf dem Plasmid) übt Selektionsdruck auf die *E. coli*-Zellen aus und hilft bei der Isolierung von Nukleasen mit weniger toxischen Eigenschaften (Oelgeschlager *et al.* 1990).

Mit Hilfe dieser Methode wurden zehn TALE-PvuII-Mutanten mit jeweils nur einer einzelnen Mutation innerhalb des PvuII-Gens selektiert, von denen fünf Varianten Aktivität aufweisen. Die Mutationen müssten im PvuII-Enzym mit dem Fusionsenzym verglichen werden, um genauere Aussagen über die Mutationen und deren Auswirkungen im Einzelnen treffen zu können. Daher sind die Eigenschaften der Mutanten noch unklar. Zwei dieser Varianten (A92T, V89I) zeigen ähnlich hohe Spaltungsraten wie das AvrBs3-28-L-PvuII-HF-Enzym. Interessanterweise scheint die P14S-Mutante im Vergleich zur P14G eine höhere Spezifität zu besitzen (Abb. 3.2.5). Dabei unterscheidet sich Serin nur durch ein zusätzliches Kohlenstoffatom gebunden an eine Hydroxygruppe von Glycin, sorgt aber dafür, dass mit einem 20-fachen Enzymüberschuss auf der nicht-adressierten DNA keine Aktivität beobachtet werden kann. Die Viabilitätsexperimente mit den selektierten TALE-PvuII-Mutanten im prokaryotischem Vektor zeigen wie erwartet eine geringe Überlebensrate für die sehr aktiven Varianten A92T und V89I und eine sehr hohe Viabilitätsrate für die weniger aktive Variante G135W. Eine Ausnahme stellt die G53R-Mutante dar, denn sowohl bei den Spaltungskinetiken *in vitro*, als auch bei den *E. coli*-Viabilitätsexperimenten ist für diese Variante eine hohe Aktivität bzw. eine hohe Viabilität zu verzeichnen (Abb. 3.2.14). Drei der eben beschriebenen Mutanten mit unterschiedlichen Spaltungsraten wurden für Experimente in HEK293-Zellen eingesetzt. Die zugehörigen Ergebnisse werden im nächsten Kapitel diskutiert.

4.3.6 TALE-PvuII-*in vivo*-Experimente in HEK293-Zellen

Die TALE-PvuII-Fusionsproteine sind sehr aktiv und *in vitro* sehr spezifisch. Für gentherapeutische Zwecke oder fürs *gene targeting* im Allgemeinen ist sowohl die *in vivo*-Aktivität als auch die Spezifität von großer Bedeutung und sollte in HEK293-Zellen getestet werden. Aktivitätstests zeigten, dass alle drei Fusionskonstrukte, unabhängig von

der *in vitro* ermittelten katalytischen Eigenschaften, nahezu die gleiche Effizienz aufweisen. Erfreulicherweise ist ihre Zytotoxizität im Vergleich zu den ZFN geringer, und sie erreichen fast das Ergebnis der kaum toxischen I-SceI-Kontrolle (Abb. 3.2.15).

Welche Funktionen die Mutationen im Einzelnen haben und wie sie die Spezifität erhöhen kann aufgrund der Datenlage nicht beantwortet werden. Des Weiteren ist unbekannt, ob die Mutationen den gleichen Effekt im PvuII-Enzym alleine erzielen würden. Folgende Hypothesen können hierzu vorsichtig aufgestellt werden: Die G135W-Mutation befindet sich am „Eingang“ des Enzyms, den die DNA passieren muss, um zur PvuII-Bindungsregion zu gelangen. Tryptophan hat im Vergleich zu Glycin eine lange aromatische Seitenkette, somit könnte Tryptophan den „Zugang“ verschließen oder zumindest die Zugänglichkeit für die DNA verringern. Abb. 4.3 stellt jeweils den Wildtyp gegenüber der Mutante je Untereinheit dar. Eine Darstellung der Mutation in beiden Untereinheiten des Enzyms würde den „Eingang“ komplett verschließen und den Zugang zum aktiven Zentrum nur durch die Rotation der Tryptophan-Seitenketten freigeben. Dieser limitierende Zugang würde die geminderte Aktivität bei den Spaltungskinetiken erklären.

Die Mutation A92T befindet sich auf der Oberfläche des Enzyms inmitten einer α -Helix und nimmt vermutlich eher eine passive Rolle ein. Die α -Helix befindet sich unmittelbar neben einer *loop*-Struktur, die mit der DNA interagiert. Dieser *loop* bzw. die Interaktion mit der DNA könnte durch die Mutation beeinträchtigt werden. Dadurch, dass diese Mutation an der Oberfläche ist, besteht ebenfalls die Möglichkeit für eine Interaktion mit dem TALE-Protein an sich.

Für die positiven Ergebnisse der G53R-Mutation gibt es vorläufig nur eine plausible Erklärung. Der Arginin-Rest der G53R-Mutante kann im Gegensatz zum Glycin eine zusätzliche Interaktion mit dem DNA-Phosphat-Rückgrat eingehen. Da bei dieser Mutante eine verminderte Spaltgeschwindigkeit beobachtet wurde, wird die DNA vermutlich „verzerrt“ und liegt eventuell nicht mehr optimal im aktiven Zentrum. Die Reduktion der Spaltungsrate begünstigt die Spezifität dieser Variante.

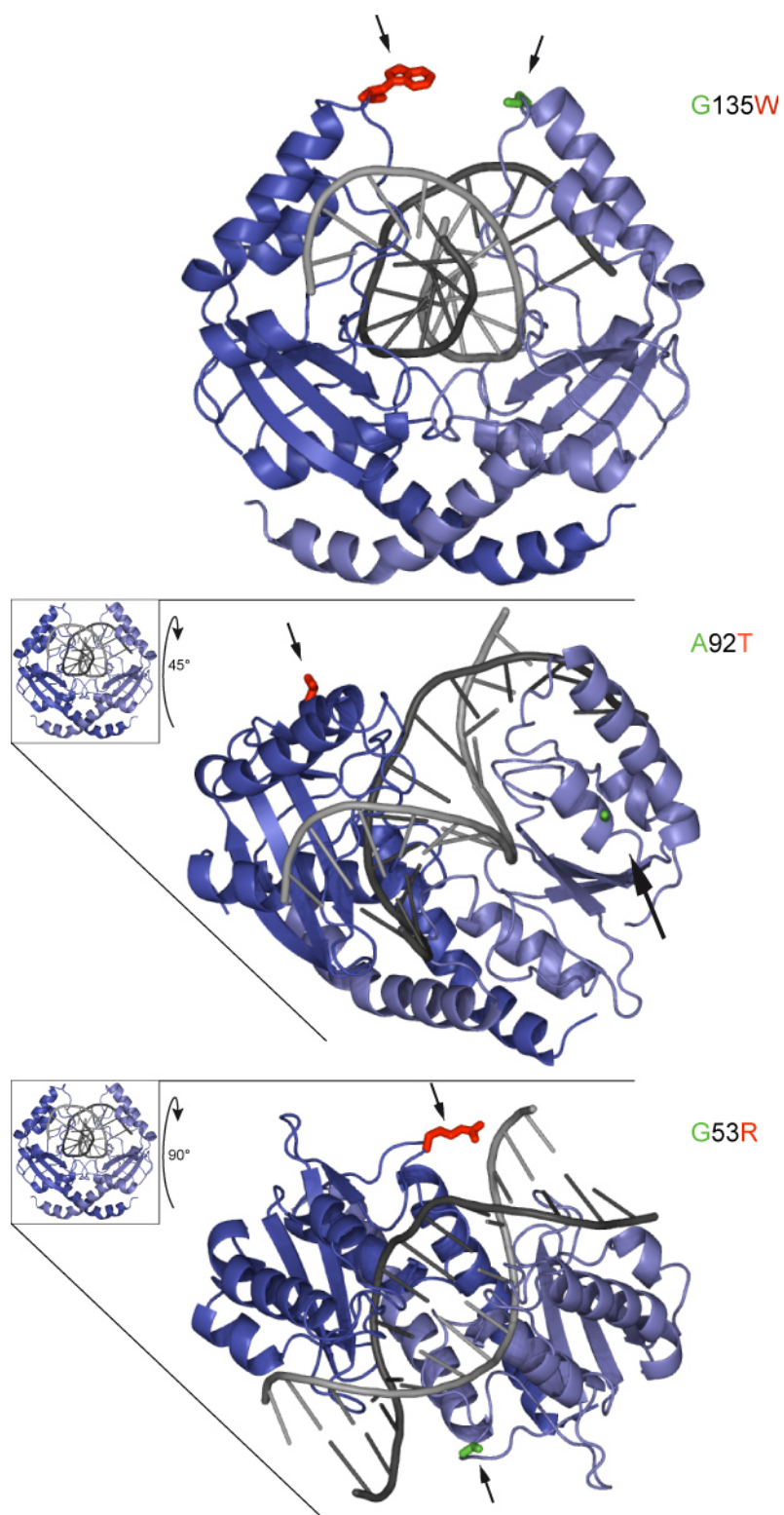


Abb. 4.3: PvuII-Mutanten mit erhöhter Spezifität

Darstellung der Aminosäure(AS)-Position der jeweiligen Mutation im Vergleich zum Wildtyp. In der dunkelblauen Untereinheit von PvuII ist die Mutation rot abgebildet und steht der zweiten hellblauen PvuII-Untereinheit mit der Wildtyp-AS in grün gegenüber (jeweils markiert durch die schwarzen Pfeile).

4.3.6.1 TALE-repeat 0-Limitierung

Die Spaltung der Substrate die am *repeat 0* mit einem Cytosin beginnen, anstatt eines nativ in der Erkennungssequenz vorkommenden Thymins, wurde nicht erwartet (Abb. 3.2.16). Die Base Thymin an der Position *repeat 0* ist bei den *Xanthomonas*-TAL-Effektoren für die eigentliche Aufgabe als Transkriptionsfaktor zwingend erforderlich. Die Webseite <http://boglabx.plp.iastate.edu/TALENT/TALENT> beruht auf den Ergebnissen von Cermak *et al.* 2011, auf der nützliche Hinweise gegeben werden, wie artifizielle TAL-Effektoren aufgebaut sein sollten bzw. welche Erkennungssequenzen sich am besten zum Adressieren eignen. Wiederum wurden in der Arbeit von Reyon *et al.* 2012 über 150 TALEN in einem Hochdurchsatzverfahren generiert, bei denen einige TALEN nicht nach den oben genannten Kriterien erstellt wurden, aber dennoch erfolgreich aktiv waren. Es wird diskutiert, dass diese Limitierungen auf der Grundlage von nativ wirkenden monomeren TAL-Effektoren erörtert wurden und einige Faktoren bei verkürzten TALEN kaum Einfluss haben könnten. Um diese Annahme zu bestätigen, müssten weitere Aktivitätsexperimente auf genomischer DNA durchgeführt werden. Sollte die *repeat 0*-Thymin-Limitierung auf verkürzte bzw. artifizielle TALE-Proteine keinen Einfluss haben, dann wäre die „Programmierung“ der TALE-Proteine noch „barrierefreier“. Eine aktuelle Publikation hat durch direkte Evolution des N-terminalen Bereichs der TAL-Effektoren diese Limitierung aufheben können. Jede Base kann an der Position *repeat 0* in der Erkennungssequenz von den neuen TALE-Proteinen erkannt und gebunden werden (Lamb, Mercer and Barbas 2013). Des Weiteren hat diese Gruppe bestätigen können, dass ein Cytosin der beste Ersatz für Thymin an der Position *repeat 0* ist.

4.4 Generierung der heterodimeren AvrBs3-PvuII-Fusionsenzyme

Die Ergebnisse der homodimeren TALE-PvuII-Enzyme implizieren eine ausreichende Spezifität auf bipartiten Substraten gegenüber einzelnen PvuII-Erkennungssequenzen. Diese Spezifität müsste jedoch erst auf genomischer DNA-Ebene validiert werden, dann hätten die TALE-PvuII durch die homodimere Architektur große Vorteile gegenüber FokI-basierten TALEN und ZFN, die wiederum mit zwei Bindungsproteinen mit unterschiedlichen Erkennungssequenzen auf der DNA (hetero-)dimerisieren müssen (Doyon *et al.* 2011). Solche obligat heterodimeren Varianten haben aber nicht nur den Vorteil einer erweiterten Erkennungssequenz, sondern auch reduzierte toxische Effekte. Die Toxizität der TALE-PvuII-Varianten D30K und K38D im HEK293-Toxizitätstest ist

sogar geringer als die der Kontrollen (Abb. 3.3.5). Eine Erklärung für die geringe Toxizität und Aktivität in den Zellen könnte darauf zurückzuführen sein, dass eine Mutante alleine nicht dazu fähig ist, ein katalytisch aktives Dimer zu bilden (Abb. 3.3.3). Zusätzlich muss bei der Co-Transfektion eine hohe Effizienz gegeben sein, um in den *in vivo*-Experimenten Aktivität beobachten zu können (Abb. 3.3.4). Vergleicht man die *in vitro*-Aktivität der heterodimeren Varianten mit denen der homodimeren TALE-PvuII-Architektur, so wird die geringe Spaltungsrate der heterodimeren Architektur für die *in vivo*-Aktivität, unabhängig von einer erfolgreichen Co-Transfektion, bestätigt. Eine niedrige Aktivität minimiert die Wahrscheinlichkeit toxischer Effekte. Inwiefern sich diese geringe Aktivität im zellulären System steuern lässt und ob sie ausreichend für die Adressierung chromosomaler Strukturen ist, muss mit weiteren *in vivo*-Experimenten untersucht werden. Vorbereitend dazu wurde die heterodimere TALE-PvuII-Konstruktion auf die PvuII-Sequenz im 1. Exon des *hIL2RG*-Gens im humanen Genom entwickelt (CCDS14406). Die neu generierten TALE-Proteine flankieren die PvuII-Sequenz mit je 8 und 6 bp Abstand, wegen der Thymin-Base an der Position des *repeat 0* und der optimalen TALE-Proteinlänge von 17.5 *repeats*. Mit dem homodimeren AvrBs3-PvuII-Fusionsenzym konnte das tripartite Substrat mit den 8 und 6 bp *spacer* gespalten werden, allerdings stellte sich später heraus, dass das Dimer nur halbseitig gebunden hatte. Die *in vitro*-Spaltungskinetik (Abb. 3.3.7) zeigt deshalb keine Aktivität, da die *spacer* nicht passend für diese Nuklease-Architektur sind. Um auszuschließen, dass es an der geringen Aktivität der Mutanten im Allgemeinen liegt, wurde die beste ermittelte 4 bp *spacer*-Länge für die 28-L-Linkervariante für dieses Substrat benutzt (Abb. 3.3.8) und damit eine geringe Spaltung nachgewiesen.

Trotz der geringen Aktivität sind die Ergebnisse für die heterodimeren Varianten vielversprechend, da sie in den HEK293-Zellen aktiv und kaum toxisch sind.

5 Fazit

In der vorliegenden Arbeit konnte demonstriert werden, dass das häufig verwendete, unspezifische Spaltmodul FokI durch Restriktionsenzyme in hochspezifischen Nukleasen ersetzt werden kann. Insbesondere bei TALEN stellt dies ein Novum dar.

Es wurden zwei unterschiedliche Ansätze beschrieben. Zum einen mit Fokus auf der Nuklease BfiI mit ihrem speziellen Reaktionsmechanismus, zum anderen auf die programmierbaren Bindungseigenschaften der TALE-Proteine.

Die außergewöhnliche Autoinhibition des BfiI-Enzyms konnte unter die Kontrolle der Bindungsmodule gebracht werden. Die katalytische Reaktion lässt sich über bipartite *cis*-Erkennungssequenzen regulieren und resultiert in einer spezifischen *nicking*-Aktivität. Die Inaktivierung der nativen BfiI-Bindedomäne war notwendig, da sonst die nativen BfiI-Erkennungssequenzen gespalten würden. Jedoch ist es durch die Inaktivierung der BfiI-Bindedomäne gelungen, unabhängig von BfiI-Erkennungssequenzen DNA spezifisch zu spalten.

Eine Anwendung für das *gene targeting* wäre denkbar, wenn die Generierung von Heterodimeren im Fusionsansatz möglich wäre. Wie schnell und ob sich die Untereinheiten des BfiI-Fusionsenzym austauschen lassen, müsste näher untersucht werden. Der Gedanke, diese Daten mit BfiI-TALE-Fusionen zu wiederholen, ist sehr verlockend, zumal eine 100 % programmierbare hochspezifische Nuklease generiert werden würde, ohne dass man die Nachteile von FokI-Fusionen in Kauf nehmen müsste. Im Rahmen einer Masterarbeit wurden diese Experimente von Tim Schneider durchgeführt. Es stellte sich heraus, dass die Fusion des BfiI-Enzyms am N-Terminus des TALE ungeeignet ist. Das BfiI-Enzym kann aufgrund seiner Struktur nur an den N-Terminus des TALE fusioniert werden (Kapitel 4.2). Der N-terminale Bereich des TALE-Proteins ist aber essentiell für die DNA-Bindung bzw. Wechselwirkung mit der DNA und sollte durch einen Fusionsanteil nicht beeinflusst werden. Nukleasen, die an den C-Terminus von TAL-Effektoren fusioniert werden, sind im Vergleich zum N-Terminus deutlich spezifischer (Beurdeley *et al.* 2013).

Dennoch besitzt der BfiI-Mechanismus Potential, vielversprechende Ergebnisse mit den entsprechenden Bindungsmodulen zu liefern, vorausgesetzt, es lassen sich aktive Heterodimere-Varianten generieren.

Die Erkennungssequenz des kleinen PvuII-Restriktionsenzym hingegen ließ sich hervorragend mit den Bindungseigenschaften des AvrBs3-TAL-Effektors erweitern. Es

konnte keine Aktivität für nicht-adressierte, isolierte PvuII-Sequenzen mit einem Überschuss an Enzym innerhalb von 24 h detektiert werden. Unter denselben Bedingungen wurden die adressierten Sequenzen innerhalb von 10 Minuten gespalten. Im Prinzip lassen sich monomer agierende Nukleasen mit scPvuII generieren und hätten aufgrund der geringeren Größe einen Transportvorteil in viralen Systemen. Mit Hilfe von zielgerichteter Mutagenese bzw. direkter Evolution könnte man bei den TALE-scPvuII-Konstrukten die gleiche Spezifität wie bei den Homodimeren erreichen. Ein großer Vorteil der homodimeren Varianten ist der, dass sie in der Lage sind, halbseitige Substrate zu spalten, ohne an Spezifität einzubüßen. Eine besonders hohe Spezifität ist mit heterodimeren TALE-PvuII-Konstrukten gegeben; diese zeigten die geringste Toxizität in HEK293-Zellen. Alle dimeren TALE-PvuII-Konstrukte zeigten Aktivität in zellularen HEK293-Experimenten mit vergleichbar geringer Toxizität wie die I-SceI-Kontrolle.

Zusammengefasst stellt die TALE-PvuII-Architektur eine vielversprechende Alternative zu FokI-basierten Zink-Finger oder TALE-Nukleasen dar. PvuII ist die dritte Alternative zu FokI, neben MutH (Gabsalilow *et al.* 2013) und I-TevI (Beurdeley *et al.* 2013), die an einem TALE-Protein fusioniert wurde. Mit MutH ist die Generierung von monomeren Nukleasen, die einen Einzelstrangbruch durchführen, gelungen. Mit I-TevI wurde ebenfalls ein monomer agierendes Konstrukt gebildet, mit der Fähigkeit, DNA-Doppelstrangbrüche zu generieren.

Das Potential des CRISPR/Cas-Systems eröffnet interessante Möglichkeiten bei der Adressierung genomischer Sequenzen, doch aufgrund der aktuellen Datenlage bezüglich der toxischen Reaktionen (Fu *et al.* 2013, Cradick *et al.* 2013) werden sie die TALEN vermutlich nicht (völlig) verdrängen. Natürlich müssen hochspezifische Nukleasen, unabhängig vom Spaltmodul, *in vivo* auf Genomebene ausführlich untersucht werden, um nähere Aussagen über Toxizität und Spezifität treffen zu können, wie es bereits bei der genomweiten Analyse für ZFN gemacht wurde (Gabriel *et al.* 2011, Pattanayak *et al.* 2011).

Letztendlich kommt es beim *gene targeting* nicht nur auf eine hohe Spezifität, sondern auch auf die Aktivität und die Art der Spaltung an. TALE-PvuII mit den eingeführten Doppelstrangbüchen wäre das Werkzeug der Wahl für das Abschalten von Genen, wohingegen spezifische Einzelstrangbrüche ideal für Gensubstitutionen/Gentherapie wären. Für welche Architektur man sich auch entscheidet, jedes System muss ihrer Zielsequenz entsprechend optimiert werden.

6 Zusammenfassung

Hochspezifische Nukleasen haben sich im Bereich *gene targeting* und *gene therapy* zu unentbehrlichen Werkzeugen entwickelt. Eine auf Zufall basierte Integration von DNA-Sequenzen ins Genom verschiedener Organismen mit den damit verbundenen Nebenwirkungen kann mit Hilfe dieser Werkzeuge vermieden werden. Die Möglichkeit, ein Gen im Genom an einer definierten Position zu reparieren, ein- oder auszuschalten oder zu ersetzen, besitzt neben der therapeutischen Anwendung auch enormes Potential in der Grundlagen- wie auch in der angewandten Forschung. Die einzelnen Funktionen von Genen und deren Wechselwirkungen in einem Modellorganismus lassen sich damit sehr gut studieren. Damit vereinfacht sich auch die Züchtung von Pflanzen und Tieren.

Die Basis der gängigsten Nuklease-Architekturen bildet das unspezifische Spaltmodul FokI, welches in Zink-Finger-Nukleasen (ZFN) am häufigsten verwendet worden ist. Mit sogenannten TAL-Effektoren ist es gelungen, in einer vergleichsweise unkomplizierten Art und Weise spezifische Nukleasen, die sogenannten TALEN, zu generieren. Beide Architekturen lassen sich programmieren, neigen aber durch den FokI-DNA-Spalt-Mechanismus zu unerwünschten Reaktionen. In dieser Arbeit werden zwei alternative Spaltmodule zu FokI mit vielversprechenden Resultaten präsentiert. Die allosterische Aktivierung des BfiI-Restriktionsenzym und die Charakterisierung der TALE-PvuII-Konstrukte.

Mit BfiI-Varianten konnte die Bindung an die Wildtyp-BfiI-Erkennungssequenz unterbunden und ein Einzelstrangbruch mit Hilfe von in *cis* gebundenen Bindungsmodulen durchgeführt werden. Die Toxizität des BfiI-Fusionsenzym ist extrem vermindert, da eine katalytische Aktivität nur durch die langen Erkennungssequenzen der Bindungsmodule dirigiert wird. Weitere Untersuchungen hinsichtlich heterodimerer Varianten werden genauere Aussagen über ihren Nutzen treffen können.

Das TALE-PvuII-Fusionskonstrukt spaltet nur adressierte TALE-PvuII-Erkennungssequenzen und zeigt keine Aktivität gegenüber isolierten PvuII-Sequenzen, trotz Enzymüberschusses und langen Inkubationszeiten. Die *in vitro*-Spaltungskinetiken zeigten, dass die Präferenz für eine adressierte Sequenz gegenüber einer nicht-adressierten Sequenz bei einem Faktor > 34.000 liegt. Darüber hinaus sind die TALE-PvuII-Konstrukte in HEK293-Zellen aktiv und weisen nur eine minimale Zytotoxizität auf. Zusammen mit den vielversprechenden heterodimeren TALE-PvuII-Varianten bleibt ihre Effektivität auf die Anwendung chromosomaler Sequenzen abzuwarten.

7 Abkürzungsverzeichnis

% (v/v)	Volumenprozent pro Volumen	l	Liter
% (w/v)	Gewichtsprozent pro Volumen	m	milli- (10^{-3})
°C	Grad Celsius	M	Molar
μ	Mikro- (10^{-6})	mA	Milliampere
A	Adenin, Alanin	Me ²⁺	Zweiwertiges Metallion
Å	Ångström (10^{-10} m)	min	Minute
Abb.	Abbildung	mRNA	<i>messenger</i> RNA
Amp	Ampicillin	MW	Molekulargewicht
APS	Ammoniumperoxodisulfat	n	nano- (10^{-9})
ATP	Adenosin-5'-triphosphat	nt	Nukleotid
bp	Basenpaare	OD	optische Dichte
BSA	<i>bovine serum albumin</i>	PAA	Polyacrylamid
bzw.	beziehungsweise	PAGE	Polyacrylamid-Gelelektrophorese
C	Cytosin	PCR	<i>polymerase chain reaction</i>
cpm	<i>counts per minute</i>	PMSF	Phenylmethylsulfonylfluorid
cv	Säulenvolumen	rev	rückwärts
Da	Dalton	RNA	Ribonukleinsäure
DMSO	Dimethylsulfoxid	rpm	Umdrehungen pro Minute
DNA	Desoxyribonukleinsäure	RT	Raumtemperatur
dNTP	Desoxyribonukleosidtriphosphat	sc	<i>single-chain</i>
DTE	1.4-Dithioerythrol	SDS	Natriumdodecylsulfat
DTT	1.4-Dithiothreitol	sek	Sekunde
EDTA	Ethylendiamintetraacetat	T	Thymin
<i>et al.</i>	und andere	Tab.	Tabelle
EtOH	Ethanol	TAE	Tris-Acetat-EDTA-Puffer
fwd	vorwärts	TEMED	N,N,N,N-Tetramethylethylendiamin
G	Guanin	TFO	<i>triple-helix forming oligo</i>
g	Gramm	Tris	Tris-(hydroxymethyl)-aminomethan
h	Stunde	ü.N.	über Nacht
IDA	Iminodiessigsäure	U	Unit (Enzymeinheit)
IPTG	Isopropyl-β-thiogalactopyranosid	V	Volt
k	Kilo- (10^3)	W	Watt
Kan	Kanamycin	wt	Wildtyp
KPi	Kaliumphosphatpuffer	x g	x Erdbeschleunigung

8 Abbildungsverzeichnis

Abb. 1.1: Nukleotid und Protein Statistik	9
Abb. 1.2: Übersicht hochspezifischer Nukleasen	13
Abb. 1.3: Zink-Finger-Bindungsmodule Co-Röntgenkristallstruktur	14
Abb. 1.4: I-CreI & I-SceI – Homing-Endonukleasen aus der LAGLIDADG Familie	15
Abb. 1.5: <i>Triple-helix forming Oligo</i> (TFO)	17
Abb. 1.6: Aufbau und Co-Kristallstruktur der TAL-Effektoren	19
Abb. 1.7: Adressierte Spaltung mit der Cas9 Nuklease aus dem CRISPR System.....	21
Abb. 1.8: Restriktionsenzym BfiI.....	24
Abb. 1.9: BfiI Reaktionsschema.....	24
Abb. 1.10: BfiI Reaktionsmechanismus	25
Abb. 1.11: Kristallstrukturen von PvuII und scPvuII.....	26
Abb. 2.1: TALE-PvuII Klonierungsstrategie	41
Abb. 3.1.1: Aktivitätstest BfiI-I-SceI	48
Abb. 3.1.2 BfiI-I-SceI Spaltungskinetiken auf dem <i>head-to-head</i> Substrat.....	49
Abb. 3.1.3: BfiI-10AS-I-SceI Aktivität auf Kontroll-Substraten	50
Abb. 3.1.4: BfiI-10AS-I-SceI Kinetiken auf größeren I-SceI cis Abstandsplasmiden.....	51
Abb. 3.1.5: Mutagenese der BfiI Bindedomäne	52
Abb. 3.1.6: Aktivitätstest der BfiI Einzelmutanten N279A & R212A.....	52
Abb. 3.1.7: Kinetiken der Einzelmutanten im Fusionsansatz.....	53
Abb. 3.1.8: Spaltung der fluoreszenz-markierten Oligos	54
Abb. 3.2.1: AvrBs3-PvuII Fusionsprotein.....	56
Abb. 3.2.2: Proteinaufreinigung des TALE-PvuII Konstrukts.....	57
Abb. 3.2.3: Vergleich der Linkervarianten.....	58
Abb. 3.2.4: Spezifitätsvergleich des Fusionsproteins mit wt- und scPvuII.....	59
Abb. 3.2.5: Kompetitive Spaltungskinetiken.....	60
Abb. 3.2.6: Spezifität des AvrB3-28-L-wtPvuII HF Fusionsenzym	61
Abb. 3.2.7: Spaltungskinetik unter Verwendung radioaktiv markierter DNA.....	62
Abb. 3.2.8: Spaltungsraten unter Verwendung adressierter Substrate unterschiedlicher Länge	63
Abb. 3.2.9: Spaltungskinetik des bipartiten Substrats.....	64
Abb. 3.2.10: Stabilitätstest des AvrBs3-28-L-wtPvuII HF Enzyms	65
Abb. 3.2.11: Vergleich der Spaltungskinetik in physiologisch angepassten Puffern.....	66
Abb. 3.2.12: Viabilitätstest	67
Abb. 3.2.13: AvrBs3-PvuII Varianten.....	69
Abb. 3.2.14: Übersicht des Viabilitätstests mit den TALE-PvuII Varianten	70
Abb. 3.2.15: Toxizitätstest der TALE-PvuII Varianten in HEK293 Zellen.....	71
Abb. 3.2.16: <i>In vivo</i> Aktivitätstest der TALE-PvuII Varianten in HEK293 Zellen.....	72
Abb. 3.3.1: Architektur heterodimerer Fusionsenzyme	73
Abb. 3.3.2: Spaltungskinetiken mit wtPvuII HF und AvrBs3 & AvrBs4	74
Abb. 3.3.3: PvuII Heterodimer	76
Abb. 3.3.4: Spaltungskinetiken mit heterodimeren TALE-PvuII Konstrukten.....	77
Abb. 3.3.5: <i>In vivo</i> Aktivitätstest der Heterodimeren Varianten in HEK293 Zellen	78
Abb. 3.3.6: Toxizitätstest mit TALE-PvuII Heterodimeren in HEK293.....	79
Abb. 3.3.7: Spaltungskinetik auf dem <i>hIL2RG</i> Locus	80
Abb. 3.3.8: Spaltungskinetik mit den <i>hIL2RG</i> TALE-PvuII Konstrukten	81
Abb. 4. 1: FokI Reaktionsschema.....	83
Abb. 4.2: BfiI-I-SceI- <i>cis</i> -Modell	86
Abb. 4.3: PvuII Mutanten mit erhöhter Spezifität.....	95

9 Tabellenverzeichnis

Tabelle 1: Verwendete Kits	29
Tabelle 2: Verwendete Chemikalien	30
Tabelle 3: Puffer & Medien.....	32
Tabelle 4: Verwendete Antibiotika.....	34
Tabelle 6: Fluorophor gelabelte Oligos	34
Tabelle 5: Primer & Oligonukleotide	35
Tabelle 7: Erkennungssequenzen der Plasmidsubstrate	36
Tabelle 8: Verwendete PCR Substrate und deren zur Herstellung verwendeten Primer Paare	37

10 Danksagung

Ein ganz besonderer Dank gilt meinem Doktorvater Prof. Dr. Alfred Pingoud für die wissenschaftliche Unterstützung, für das immer offene Ohr und das Interesse an meiner Arbeit. Ihre Ratschläge und Tipps, die mich sowohl fachlich als auch persönlich weitergebracht haben, habe ich immer genossen. Ich bin stolz, Sie als meinen Doktorvater haben zu dürfen und freue mich darauf, mit Ihnen Istanbul zu entdecken.

Herrn Prof. Dr. Michael Martin danke ich für die freundliche Übernahme des Zweitgutachtens.

Was wäre ich nur ohne Wolle? „Ah Junge, 10 µl 54°C, zack, zack“. Ganz einfach. Mir würde ne gewaltige Portion Coolness fehlen. In den paar Jahren konnte ich immer wieder beobachten, wie die Studenten mit ihren völlig unterschiedlichen Problemen zu dir eilten und noch bevor sie den Satz zu Ende gesprochen hatten, wusstest du, worum es geht. Stets ruhig und geduldig bist du für jeden da. Selbst als ich die Mikrowelle zum Leuchten.....ah lassen wir das. Deine lockere Art hat mir während deiner Betreuung immer die Möglichkeit gegeben, alles auszuprobieren und Fehler machen zu dürfen. Ganz besonders habe ich die Zeit genossen, das klingt jetzt vielleicht nicht ganz passend, als Peter ausgezogen ist. Wir wissen ja beide, dass du die Einsamkeit nicht aushalten konntest und heilfroh warst, als ich Peters Platz einnahm. Neben unseren wissenschaftlichen Diskussionen habe ich den Small Talk über den Tag verteilt sehr genossen. Ich habe eine Menge von dir gelernt. Danke.

Apropos Peter: Ich könnte mich natürlich für die zahlreichen wissenschaftlichen Tipps und Diskussionen bedanken, denn genau das zeichnet dich aus - für gewisse Problemstellungen alternative Herangehensweisen auszutüfteln. Aber mein besonderer Dank gilt deiner Freude- ausstrahlender Art. Ich werde die Gedanken an dich immer mit positiven Zeiten verknüpfen. Deine morgendliche Rezitation der Radiowitze und die gekonnte Versetzung der Pointe wird stets mein Gesicht zum Lächeln bringen.

Ich bedanke mich bei unseren Kooperationspartnern, insbesondere bei Prof. Dr. Virginijus Siksnyis aus Vilnius und Prof. Dr. Toni Cathomen aus Freiburg für ihre Hilfsbereitschaft.

Ein ganz besonderer Dank gilt den ehemaligen Mitgliedern des Instituts, die mich eingearbeitet, unterstützt, mit mir gefeiert und mir einen angenehmen Start in Gießen ermöglicht haben. Erwähnen möchte ich hier meinen Postdoc „Goldjungen“ Benno und Daniel „Hase“.

Sabrina, wie würde wohl das Labor ohne dich aussehen? Ich fand die Salzkrusten in den Elektrophoresekammern künstlerisch gelungen. Ich möchte dir nicht nur für die Ordnung, den nimmer endenden Puffer- und Markervorrat und für deine Unterstützung danken, sondern auch für unser morgendliches Ritual (für dich eher Mittag), den Gang ins CaRé und die damit verbundenen Gespräche.

An dieser Stelle gilt mein Dank ebenfalls allen Angestellten des Instituts. Heike, auf deinen großen praktischen Erfahrungsschatz zurückzugreifen hat mich bereichert und auch vielen Dank an Petra, - ohne dich würden wir alle 3 Jahre länger promovieren.

Vielen Dank für alle (aktuellen) Mitglieder des Instituts, für die stets gute und hilfsbereite Atmosphäre. Die privaten Unternehmungen, uns mit anderen Instituten auf dem „Feld“ zu messen, unterstreichen unseren Zusammenhalt. Besonders möchte ich mich bei Andi und den Mitgliedern und Ersatzmitgliedern von „DwoP“ bedanken. Ein Glück gibt es das Internet, so dass wir nicht auf die Rettung der virtuellen Welt verzichten müssen. Dennoch vermisse ich jetzt schon die langen Nächte im Institut, das Fastfood und unsere „Nerd“-Gespräche.

Ich danke Ina und Karina für ihre Hilfe und Unterstützung bei organisatorischen und bürokratischen Angelegenheiten.

Meinen Studentinnen und Studenten, Praktikantinnen und Praktikanten danke ich für die schöne, als auch für mich lehrreiche Zeit. Besonders möchte ich mich bei Tim Schneider und Nikolas Martin für ihre Mühen bedanken. Viel Erfolg für eure Zukunft.

Meinen Freunden und Vereinsmitgliedern aus dem Ruhrgebiet möchte ich danken, dass sie stets für mich da waren und ein offenes Ohr hatten. Meine kurzen Wochenendbesuche habt ihr immer mit großer Vorfreude gefüllt. Meinem alten Klassenkameraden / Trauzeugen Thorsten möchte ich hier ganz besonders danken. Das sonntagmorgendliche Training mit meinen „Jungs“, die meine „Musik“ spielen, sorgte immer für gute Laune und Motivation. Besser kann man keine Unterstützung finden. Danke.

Anja, Schätzchen, du verdienst einen gaaaanz besonderen Dank. Die Promotion, die langen Nächte im Labor, hätte ich ohne deine Unterstützung nicht so gut meistern können. Dein niemals endender Vorrat an Schokolade sorgte stets für Motivation und Leistungssteigerung. Vielen Dank für die unzähligen Brettspielabende, die mir besonders bei meiner Ankunft eine schnelle und angenehme Eingewöhnung in Gießen ermöglicht haben. Unter den vielen nationalen und internationalen Bewerbern fürs Stipendium im Graduiertenkolleg hast du mich für die engere Auswahl ausgewählt und indirekt dafür gesorgt ein ganz besonderes Mädels kennenzulernen, çok teşekkürler ☺.

Verena, vielen Dank für deine Unterstützung, Verständnis und Geduld während meiner Promotion. Wer hätte am Tag meines Bewerbungsgespräches, an dem ich dich im Sozialraum kennenlernte, gedacht, dass du vor mir „Dr. Yanik“ wirst ;-). Ich bin stolz dich als meine Frau an meiner Seite zu haben. Ich liebe dich!

Als letztes möchte ich meinen Eltern und meiner Familie danken. Stets habt ihr mich bei meinen Wünschen und Taten unterstützt und an mich geglaubt. Ich bin dankbar für meine warmherzige Familie und die Werte, die ihr mir vermittelt habt.

11 Literaturverzeichnis

- Aggarwal, A. K. (1995) Structure and function of restriction endonucleases. *Curr. Opin. Struct. Biol.*, 5, 11-9.
- Alwin, S., M. B. Gere, E. Guhl, K. Effertz, C. F. Barbas, 3rd, D. J. Segal, M. D. Weitzman & T. Cathomen (2005) Custom zinc-finger nucleases for use in human cells. *Mol Ther*, 12, 610-7.
- Antoine, C., S. Muller, A. Cant, M. Cavazzana-Calvo, P. Veys, J. Vossen, A. Fasth, C. Heilmann, N. Wulffraat, et al. (2003) Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience 1968-99. *Lancet*, 361, 553-60.
- Apweiler, R., A. Bairoch & C. H. Wu (2004) Protein sequence databases. *Curr Opin Chem Biol*, 8, 76-80.
- Arnould, S., P. Chames, C. Perez, E. Lacroix, A. Duclert, J. C. Epinat, F. Stricher, A. S. Petit, A. Patin, et al. (2006) Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J Mol Biol*, 355, 443-58.
- Arnould, S., C. Perez, J. P. Cabaniols, J. Smith, A. Gouble, S. Grizot, J. C. Epinat, A. Duclert, P. Duchateau, et al. (2007) Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J Mol Biol*, 371, 49-65.
- Ashworth, J., J. J. Havranek, C. M. Duarte, D. Sussman, R. J. Monnat, Jr., B. L. Stoddard & D. Baker (2006) Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature*, 441, 656-9.
- Ashworth, J., G. K. Taylor, J. J. Havranek, S. A. Quadri, B. L. Stoddard & D. Baker (2010) Computational reprogramming of homing endonuclease specificity at multiple adjacent base pairs. *Nucleic Acids Res*, 38, 5601-8.
- Athanasiadis, A., M. Vlassi, D. Kotsifaki, P. A. Tucker, K. S. Wilson & M. Kokkinidis (1994) Crystal structure of PvuII endonuclease reveals extensive structural homologies to EcoRV. *Nat Struct Biol*, 1, 469-75.
- Ausubel, F., R. Brent, R. Kingston, D. Moore, J. Seidmann, J. Smith & K. Struhl. 2012. *Current protocols in molecular biology*. John Wiley & Sons.
- Baltimore, D. (1978) Limiting science: a biologist's perspective. *Daedalus*, 107, 37-45.
- Barzel, A., E. Privman, M. Peeri, A. Naor, E. Shachar, D. Burstein, R. Lazary, U. Gophna, T. Pupko, et al. (2011) Native homing endonucleases can target conserved genes in humans and in animal models. *Nucleic Acids Res*, 39, 6646-59.
- Bell, O., V. K. Tiwari, N. H. Thoma & D. Schubeler (2011) Determinants and dynamics of genome accessibility. *Nat Rev Genet*, 12, 554-64.
- Beltran, W. A., A. V. Cideciyan, A. S. Lewin, S. Iwabe, H. Khanna, A. Sumaroka, V. A. Chiodo, D. S. Fajardo, A. J. Roman, et al. (2012) Gene therapy rescues photoreceptor blindness in dogs and paves the way for treating human X-linked retinitis pigmentosa. *Proc Natl Acad Sci U S A*, 109, 2132-7.
- Benson, D. A., M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell & E. W. Sayers (2013) GenBank. *Nucleic Acids Res*, 41, D36-42.
- Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell & E. W. Sayers (2011) GenBank. *Nucleic Acids Res*, 39, D32-7.
- Beumer, K., G. Bhattacharyya, M. Bibikova, J. K. Trautman & D. Carroll (2006) Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics*, 172, 2391-403.
- Beurdeley, M., F. Bietz, J. Li, S. Thomas, T. Stoddard, A. Juillerat, F. Zhang, D. F. Voytas, P. Duchateau, et al. (2013) Compact designer TALENs for efficient genome engineering. *Nat Commun*, 4, 1762.

-
- Bitinaite, J., D. A. Wah, A. K. Aggarwal & I. Schildkraut (1998) FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci U S A*, 95, 10570-5.
- Boch, J. & U. Bonas (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol*, 48, 419-36.
- Boch, J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt & U. Bonas (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 326, 1509-12.
- Bogdanove, A. J. & D. F. Voytas (2011) TAL effectors: customizable proteins for DNA targeting. *Science*, 333, 1843-6.
- Bonas, U., R. Schulte, S. Fenselau, G. V. Minsavage, B. J. Staskawicz & R. E. Stall (1991) Isolation of a Gene-Cluster from Xanthomonas-Campestris Pv Vesicatoria That Determines Pathogenicity and the Hypersensitive Response on Pepper and Tomato. *Molecular Plant-Microbe Interactions*, 4, 81-88.
- Bonas, U., R. E. Stall & B. Staskawicz (1989) Genetic and structural characterization of the avirulence gene avrBs3 from Xanthomonas campestris pv. vesicatoria. *Mol Gen Genet*, 218, 127-36.
- Bonetta, L. (2006) Genome sequencing in the fast lane. *Nature Methods*, 3, 141-147.
- Briggs, A. W., X. Rios, R. Chari, L. Yang, F. Zhang, P. Mali & G. M. Church (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res*, 40, e117.
- Buckley, R. H., S. E. Schiff, R. I. Schiff, L. Markert, L. W. Williams, J. L. Roberts, L. A. Myers & F. E. Ward (1999) Hematopoietic stem-cell transplantation for the treatment of severe combined immunodeficiency. *N Engl J Med*, 340, 508-16.
- Bultmann, S., R. Morbitzer, C. S. Schmidt, K. Thanisch, F. Spada, J. Elsaesser, T. Lahaye & H. Leonhardt (2012) Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res*, 40, 5368-77.
- Buttner, D. & U. Bonas (2010) Regulation and secretion of Xanthomonas virulence factors. *FEMS Microbiol Rev*, 34, 107-33.
- Camenisch, T. D., M. H. Brilliant & D. J. Segal (2008) Critical parameters for genome editing using zinc finger nucleases. *Mini Rev Med Chem*, 8, 669-76.
- Catto, L. E., S. Ganguly, S. E. Milsom, A. J. Welsh & S. E. Halford (2006) Protein assembly and DNA looping by the FokI restriction endonuclease. *Nucleic Acids Res*, 34, 1711-20.
- Cavazzana-Calvo, M., E. Payen, O. Negre, G. Wang, K. Hehir, F. Fusil, J. Down, M. Denaro, T. Brady, et al. (2010) Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature*, 467, 318-22.
- Cermak, T., E. L. Doyle, M. Christian, L. Wang, Y. Zhang, C. Schmidt, J. A. Baller, N. V. Somia, A. J. Bogdanove, et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*, 39, e82.
- Chames, P., J. C. Epinat, S. Guillier, A. Patin, E. Lacroix & F. Paques (2005) In vivo selection of engineered homing endonucleases using double-strand break induced homologous recombination. *Nucleic Acids Res*, 33, e178.
- Cheng, X., K. Balendiran, I. Schildkraut & J. E. Anderson (1994) Structure of PvuII endonuclease with cognate DNA. *Embo J*, 13, 3927-35.
- Chevalier, B. S. & B. L. Stoddard (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res*, 29, 3757-74.
- Chial, H. (2008) Rare Genetic Disorders: Learning About Genetic Disease Through Gene Mapping, SNPs, and Microarray Data. *Nature Education*.

-
- Chiang, W. C., T. M. Geel, M. M. Altintas, S. Sever, M. H. Ruiters & J. Reiser (2010) Establishment of protein delivery systems targeting podocytes. *PLoS One*, 5, e11837.
- Choo, Y. & A. Klug (1994) Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc Natl Acad Sci U S A*, 91, 11163-7.
- Christian, M., T. Cermak, E. L. Doyle, C. Schmidt, F. Zhang, A. Hummel, A. J. Bogdanove & D. F. Voytas (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, 186, 757-61.
- Clark, K. J., D. F. Voytas & S. C. Ekker (2011) A TALE of two nucleases: gene targeting for the masses? *Zebrafish*, 8, 147-9.
- Coligan, J., B. Dunn, D. Speicher & P. Wingfield. 2012. *Current protocols in protein science*. John Wiley & Sons.
- Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, et al. (2013a) Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*.
- (2013b) Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339, 819-23.
- Conlan, L. H. & C. M. Dupureur (2002) Multiple metal ions drive DNA association by PvuII endonuclease. *Biochemistry*, 41, 14848-55.
- Cornu, T. I., S. Thibodeau-Beganny, E. Guhl, S. Alwin, M. Eichinger, J. K. Joung & T. Cathomen (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther*, 16, 352-8.
- Cradick, T. J., E. J. Fine, C. J. Antico & G. Bao (2013) CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*.
- Davis, L. & N. Maizels (2011) DNA nicks promote efficient and safe targeted gene correction. *PLoS One*, 6, e23981.
- de Lange, O., T. Schreiber, N. Schandry, J. Radeck, K. H. Braun, J. Koszinowski, H. Heuer, A. Strauss & T. Lahaye (2013) Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytol*.
- DeFrancesco, L. (2011) Move over ZFNs. *Nat Biotechnol*, 29, 681-4.
- Delacote, F., C. Perez, V. Guyot, C. Mikonio, P. Potrel, J. P. Cabaniols, C. Delenda, F. Paques & P. Duchateau (2011) Identification of genes regulating gene targeting by a high-throughput screening approach. *J Nucleic Acids*, 2011, 947212.
- Deng, H., R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, et al. (1996) Identification of a major co-receptor for primary isolates of HIV-1. *Nature*, 381, 661-6.
- Dong, Z., J. Ge, K. Li, Z. Xu, D. Liang, J. Li, J. Li, W. Jia, Y. Li, et al. (2011) Heritable targeted inactivation of myostatin gene in yellow catfish (*Pelteobagrus fulvidraco*) using engineered zinc finger nucleases. *PLoS One*, 6, e28897.
- Doyle, E. L., B. L. Stoddard, D. F. Voytas & A. J. Bogdanove (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. *Trends Cell Biol*.
- Doyon, Y., T. D. Vo, M. C. Mendel, S. G. Greenberg, J. Wang, D. F. Xia, J. C. Miller, F. D. Urnov, P. D. Gregory, et al. (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods*, 8, 74-9.
- Duca, M., P. Vekhoff, K. Oussedik, L. Halby & P. B. Arimondo (2008) The triple helix: 50 years later, the outcome. *Nucleic Acids Res*, 36, 5123-38.
- Dupureur, C. M. & L. H. Conlan (2000) A catalytically deficient active site variant of PvuII endonuclease binds Mg(II) ions. *Biochemistry*, 39, 10921-7.

-
- Eisenschmidt, K., T. Lanio, A. Simoncsits, A. Jeltsch, V. Pingoud, W. Wende & A. Pingoud (2005) Developing a programmed restriction endonuclease for highly specific DNA cleavage. *Nucleic Acids Res*, 33, 7039-47.
- Fischer, A., S. Hacein-Bey-Abina & M. Cavazzana-Calvo (2010) 20 years of gene therapy for SCID. *Nat Immunol*, 11, 457-60.
- Flisikowska, T., I. S. Thorey, S. Offner, F. Ros, V. Lifke, B. Zeitler, O. Rottmann, A. Vincent, L. Zhang, et al. (2011) Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. *PLoS One*, 6, e21045.
- Foley, J. E., M. L. Maeder, J. Pearlberg, J. K. Joung, R. T. Peterson & J. R. Yeh (2009) Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. *Nat Protoc*, 4, 1855-67.
- Fonfara, I. 2011. Generating meganucleases with new specificities by directed evolution and rational engineering approaches. JLU Giessen urn:nbn:de:hebis:26-opus-83734.
- Fonfara, I., U. Curth, A. Pingoud & W. Wende (2012) Creating highly specific nucleases by fusion of active restriction endonucleases and catalytically inactive homing endonucleases. *Nucleic Acids Res*, 40, 847-60.
- Fu, Y., J. A. Foden, C. Khayter, M. L. Maeder, D. Reyon, J. K. Joung & J. D. Sander (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*.
- Gabriel, R., A. Lombardo, A. Arens, J. C. Miller, P. Genovese, C. Kaepfel, A. Nowrouzi, C. C. Bartholomae, J. Wang, et al. (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol*, 29, 816-23.
- Gabsalilow, L., B. Schierling, P. Friedhoff, A. Pingoud & W. Wende (2013) Site- and strand-specific nicking of DNA by fusion proteins derived from MutH and I-SceI or TALE repeats. *Nucleic Acids Res*, 41, e83.
- Gaj, T., C. A. Gersbach & C. F. Barbas, 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*.
- Gaj, T., J. Guo, Y. Kato, S. J. Sirk & C. F. Barbas, 3rd (2012) Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat Methods*, 9, 805-7.
- Gao, H., X. Wu, J. Chai & Z. Han (2012) Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. *Cell Res*, 22, 1716-20.
- Gaspar, H. B., K. L. Parsley, S. Howe, D. King, K. C. Gilmour, J. Sinclair, G. Brouns, M. Schmidt, C. Von Kalle, et al. (2004) Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet*, 364, 2181-7.
- Geel, T. M., G. Meiss, B. T. van der Gun, B. J. Kroesen, L. F. de Leij, M. Zaremba, A. Silanskas, M. Kokkinidis, A. Pingoud, et al. (2009) Endonucleases induced TRAIL-insensitive apoptosis in ovarian carcinoma cells. *Exp Cell Res*, 315, 2487-95.
- Gellhaus, K., T. I. Cornu, R. Heilbronn & T. Cathomen (2010) Fate of recombinant adeno-associated viral vector genomes during DNA double-strand break-induced gene targeting in human cells. *Hum Gene Ther*, 21, 543-53.
- Gimble, F. S., C. M. Moure & K. L. Posey (2003) Assessing the plasticity of DNA target site recognition of the PI-SceI homing endonuclease using a bacterial two-hybrid selection system. *J Mol Biol*, 334, 993-1008.
- Giovannangeli, C., M. Rougee, T. Garestier, N. T. Thuong & C. Helene (1992) Triple-helix formation by oligonucleotides containing the three bases thymine, cytosine, and guanine. *Proc Natl Acad Sci U S A*, 89, 8631-5.
- Grazulis, S., E. Manakova, M. Roessle, M. Bochtler, G. Tamulaitiene, R. Huber & V. Siksnys (2005) Structure of the metal-independent restriction enzyme BfiI reveals

-
- fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc Natl Acad Sci U S A*, 102, 15797-802.
- Greisman, H. A. & C. O. Pabo (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science*, 275, 657-61.
- Grizot, S., J. Smith, F. Daboussi, J. Prieto, P. Redondo, N. Merino, M. Villate, S. Thomas, L. Lemaire, et al. (2009) Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. *Nucleic Acids Res*, 37, 5405-19.
- Grosse, S., N. Huot, C. Mahiet, S. Arnould, S. Barradeau, D. L. Clerre, I. Chion-Sotinel, C. Jacqmarcq, B. Chapellier, et al. (2011) Meganuclease-mediated Inhibition of HSV1 Infection in Cultured Cells. *Mol Ther*, 19, 694-702.
- Gruen, M., K. Chang, I. Serbanescu & D. R. Liu (2002) An in vivo selection system for homing endonuclease activity. *Nucleic Acids Res*, 30, e29.
- Gupta, A., R. G. Christensen, A. L. Rayla, A. Lakshmanan, G. D. Stormo & S. A. Wolfe (2012) An optimized two-finger archive for ZFN-mediated gene targeting. *Nat Methods*, 9, 588-90.
- Halford, S. E., L. E. Catto, C. Pernstich, D. A. Rusling & K. L. Sanders (2011) The reaction mechanism of FokI excludes the possibility of targeting zinc finger nucleases to unique DNA sites. *Biochem Soc Trans*, 39, 584-8.
- Handel, E. M. & T. Cathomen (2011) Zinc-finger nuclease based genome surgery: it's all about specificity. *Curr Gene Ther*, 11, 28-37.
- Hockemeyer, D., F. Soldner, C. Beard, Q. Gao, M. Mitalipova, R. C. DeKolver, G. E. Katibah, R. Amora, E. A. Boydston, et al. (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol*, 27, 851-7.
- Holkers, M., I. Maggio, J. Liu, J. M. Janssen, F. Miselli, C. Mussolino, A. Recchia, T. Cathomen & M. A. Goncalves (2013) Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res*, 41, e63.
- Holt, N., J. Wang, K. Kim, G. Friedman, X. Wang, V. Taupin, G. M. Crooks, D. B. Kohn, P. D. Gregory, et al. (2010) Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol*, 28, 839-47.
- Horton, J. R. & X. Cheng (2000) PvuII endonuclease contains two calcium ions in active sites. *J Mol Biol*, 300, 1049-56.
- Horvath, P. & R. Barrangou (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science*, 327, 167-70.
- Howe, S. J., M. R. Mansour, K. Schwarzwaelder, C. Bartholomae, M. Hubank, H. Kempfski, M. H. Brugman, K. Pike-Overzet, S. J. Chatters, et al. (2008) Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest*, 118, 3143-50.
- Hurt, J. A., S. A. Thibodeau, A. S. Hirsh, C. O. Pabo & J. K. Joung (2003) Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 12271-12276.
- Huston, M. W., N. P. van Til, T. P. Visser, S. Arshad, M. H. Brugman, C. Cattoglio, A. Nowrouzi, Y. Li, A. Schambach, et al. (2011) Correction of murine SCID-X1 by lentiviral gene therapy using a codon-optimized IL2RG gene and minimal pretransplant conditioning. *Mol Ther*, 19, 1867-77.
- Isalan, M., Y. Choo & A. Klug (1997) Synergy between adjacent zinc fingers in sequence-specific DNA recognition. *Proc Natl Acad Sci U S A*, 94, 5617-21.

-
- Ishino, Y., H. Shinagawa, K. Makino, M. Amemura & A. Nakata (1987) Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol*, 169, 5429-33.
- Jansen, R., J. D. van Embden, W. Gaastra & L. M. Schouls (2002) Identification of a novel family of sequence repeats among prokaryotes. *OMICS*, 6, 23-33.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna & E. Charpentier (2012) A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*.
- Jones, J. B., H. Bouzar, R. E. Stall, E. C. Almira, P. D. Roberts, B. W. Bowen, J. Sudberry, P. M. Strickler & J. Chun (2000) Systematic analysis of xanthomonads (*Xanthomonas* spp.) associated with pepper and tomato lesions. *Int J Syst Evol Microbiol*, 50 Pt 3, 1211-9.
- Jose, T. J., L. H. Conlan & C. M. Dupureur (1999) Quantitative evaluation of metal ion binding to PvuII restriction endonuclease. *J Biol Inorg Chem*, 4, 814-23.
- Jurica, M. S. & B. L. Stoddard (1999) Homing endonucleases: structure, function and evolution. *Cell Mol Life Sci*, 55, 1304-26.
- Kaczorowski, T., P. Skowron & A. J. Podhajska (1989) Purification and characterization of the FokI restriction endonuclease. *Gene*, 80, 209-16.
- Kay, S., S. Hahn, E. Marois, G. Hause & U. Bonas (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*, 318, 648-51.
- Kim, J. S., H. J. Lee & D. Carroll (2010) Genome editing with modularly assembled zinc-finger nucleases. *Nat Methods*, 7, 91; author reply 91-2.
- Kim, Y. G., J. Cha & S. Chandrasegaran (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A*, 93, 1156-60.
- Kirsch, R. D. & E. Joly (1998) An improved PCR-mutagenesis strategy for two-site mutagenesis or sequence swapping between related genes. *Nucleic Acids Research*, 26, 1848-1850.
- Kleinstiver, B. P., J. M. Wolfs, T. Kolaczyk, A. K. Roberts, S. X. Hu & D. R. Edgell (2012) Monomeric site-specific nucleases for genome editing. *Proc Natl Acad Sci U S A*, 109, 8061-6.
- Klug, A. (2010) The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Annual Review of Biochemistry*, Vol 79, 79, 213-231.
- Lagunavicius, A., G. Sasnauskas, S. E. Halford & V. Siksnys (2003) The metal-independent type IIs restriction enzyme BfiI is a dimer that binds two DNA sites but has only one catalytic centre. *J Mol Biol*, 326, 1051-64.
- Lamb, B. M., A. C. Mercer & C. F. Barbas, 3rd (2013) Directed evolution of the TALE N-terminal domain for recognition of all 5' bases. *Nucleic Acids Res*.
- Lander, E. S. (2011) Initial impact of the sequencing of the human genome. *Nature*, 470, 187-97.
- Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, 409, 860-921.
- Li, T., S. Huang, W. Z. Jiang, D. Wright, M. H. Spalding, D. P. Weeks & B. Yang (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res*, 39, 359-72.
- Lieber, M. R. (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem*, 79, 181-211.
- Lippow, S. M., P. M. Aha, M. H. Parker, W. J. Blake, B. M. Baynes & D. Lipovsek (2009) Creation of a type IIS restriction endonuclease with a long recognition sequence. *Nucleic Acids Res*, 37, 3061-73.

-
- Liu, L., Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu & M. Law (2012) Comparison of next-generation sequencing systems. *J Biomed Biotechnol*, 2012, 251364.
- Liu, R., W. A. Paxton, S. Choe, D. Ceradini, S. R. Martin, R. Horuk, M. E. MacDonald, H. Stuhlmann, R. A. Koup, et al. (1996) Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*, 86, 367-77.
- Lombardo, A., P. Genovese, C. M. Beausejour, S. Colleoni, Y. L. Lee, K. A. Kim, D. Ando, F. D. Urnov, C. Galli, et al. (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol*, 25, 1298-306.
- Lu, Y. (2004) Recombinant adeno-associated virus as delivery vector for gene therapy--a review. *Stem Cells Dev*, 13, 133-45.
- Lusser, M., C. Parisi, D. Plan & E. Rodriguez-Cerezo (2012) Deployment of new biotechnologies in plant breeding. *Nat Biotechnol*, 30, 231-9.
- Maeder, M. L., S. Thibodeau-Beganny, A. Osiaik, D. A. Wright, R. M. Anthony, M. Eichinger, T. Jiang, J. E. Foley, R. J. Winfrey, et al. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell*, 31, 294-301.
- Magrane, M. & U. Consortium (2011) UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)*, 2011, bar009.
- Mak, A. N., P. Bradley, A. J. Bogdanove & B. L. Stoddard (2013) TAL effectors: function, structure, engineering and applications. *Curr Opin Struct Biol*, 23, 93-9.
- Mak, A. N., P. Bradley, R. A. Cernadas, A. J. Bogdanove & B. L. Stoddard (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*, 335, 716-9.
- Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. Dicarlo, J. E. Norville & G. M. Church (2013) RNA-Guided Human Genome Engineering via Cas9. *Science*.
- Mano, N., F. Mao & A. Heller (2002) A miniature biofuel cell operating in a physiological buffer. *J Am Chem Soc*, 124, 12962-3.
- McClelland, M., J. Hanish, M. Nelson & Y. Patel (1988) KGB: a single buffer for all restriction endonucleases. *Nucleic Acids Res*, 16, 364.
- McMahon, M. A., M. Rahdar & M. Porteus (2012) Gene editing: not just for translation anymore. *Nature Methods*, 9, 28-31.
- McPherson, J. D., M. Marra, L. Hillier, R. H. Waterston, A. Chinwalla, J. Wallis, M. Sekhon, K. Wylie, E. R. Mardis, et al. (2001) A physical map of the human genome. *Nature*, 409, 934-41.
- Meckler, J. F., M. S. Bhakta, M. S. Kim, R. Ovadia, C. H. Habrian, A. Zykovich, A. Yu, S. H. Lockwood, R. Morbitzer, et al. (2013) Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res*, 41, 4118-28.
- Meramveliotaki, C., D. Kotsifaki, M. Androulaki, A. Hountas, E. Eliopoulos & M. Kokkinidis (2007) Purification, crystallization, X-ray diffraction analysis and phasing of an engineered single-chain PvuII restriction endonuclease. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 63, 836-8.
- Methods, N. (2012) Method of the Year 2011. *Nat Methods*, 9, 1.
- Metzger, M. J., A. McConnell-Smith, B. L. Stoddard & A. D. Miller (2011) Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template. *Nucleic Acids Research*, 39, 926-35.
- Miller, J. C., M. C. Holmes, J. Wang, D. Y. Guschin, Y. L. Lee, I. Rupniewski, C. M. Beausejour, A. J. Waite, N. S. Wang, et al. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*, 25, 778-85.

-
- Miller, J. C., S. Tan, G. Qiao, K. A. Barlow, J. Wang, D. F. Xia, X. Meng, D. E. Paschon, E. Leung, et al. (2011) A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*, 29, 143-8.
- Moehle, E. A., J. M. Rock, Y. L. Lee, Y. Jouvenot, R. C. DeKolver, P. D. Gregory, F. D. Urnov & M. C. Holmes (2007) Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A*, 104, 3055-60.
- Moscou, M. J. & A. J. Bogdanove (2009) A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science*, 326, 1501-1501.
- Munoz, I. G., J. Prieto, S. Subramanian, J. Coloma, P. Redondo, M. Villate, N. Merino, M. Marenchino, M. D'Abramo, et al. (2011) Molecular basis of engineered meganuclease targeting of the endogenous human RAG1 locus. *Nucleic Acids Res*, 39, 729-43.
- Mussolino, C., R. Morbitzer, F. Lutge, N. Dannemann, T. Lahaye & T. Cathomen (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res*, 39, 9283-93.
- Oelgeschlager, T., R. Geiger, T. Ruter, J. Alves, A. Fliess & A. Pingoud (1990) Probing the function of individual amino acid residues in the DNA binding site of the EcoRI restriction endonuclease by analysing the toxicity of genetically engineered mutants. *Gene*, 89, 19-27.
- Pace, C. N., F. Vajdos, L. Fee, G. Grimsley & T. Gray (1995) How to Measure and Predict the Molar Absorption-Coefficient of a Protein. *Protein Science*, 4, 2411-2423.
- Paques, F. & P. Duchateau (2007) Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr Gene Ther*, 7, 49-66.
- Pareek, C. S., R. Smoczynski & A. Tretyn (2011) Sequencing technologies and genome sequencing. *J Appl Genet*, 52, 413-35.
- Pattanayak, V., C. L. Ramirez, J. K. Joung & D. R. Liu (2011) Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods*, 8, 765-70.
- Perez-Pinera, P., D. G. Ousterout & C. A. Gersbach (2012) Advances in targeted genome editing. *Curr Opin Chem Biol*, 16, 268-77.
- Perez, E. E., J. Wang, J. C. Miller, Y. Jouvenot, K. A. Kim, O. Liu, N. Wang, G. Lee, V. V. Bartsevich, et al. (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol*, 26, 808-16.
- Pernstich, C. & S. E. Halford (2012) Illuminating the reaction pathway of the FokI restriction endonuclease by fluorescence resonance energy transfer. *Nucleic Acids Res*, 40, 1203-13.
- Pingoud, A., M. Fuxreiter, V. Pingoud & W. Wende (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci*, 62, 685-707.
- Pingoud, A. & A. Jeltsch (1997) Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur J Biochem*, 246, 1-22.
- Pingoud, A. & G. H. Silva (2007) Precision genome surgery. *Nat Biotechnol*, 25, 743-4.
- Pingoud, A. & W. Wende (2011) Generation of novel nucleases with extended specificity by rational and combinatorial strategies. *Chembiochem*, 12, 1495-500.
- Podevin, N., Y. Devos, H. V. Davies & K. M. Nielsen (2012) Transgenic or not? No simple answer! New biotechnology-based plant breeding techniques and the regulatory landscape. *EMBO Rep*, 13, 1057-61.
- Porteus, M. H., T. Cathomen, M. D. Weitzman & D. Baltimore (2003) Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. *Mol Cell Biol*, 23, 3558-65.

-
- Pruett-Miller, S. M., D. W. Reading, S. N. Porter & M. H. Porteus (2009) Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. *PLoS Genet*, 5, e1000376.
- Rahman, S. H., M. L. Maeder, J. K. Joung & T. Cathomen (2011) Zinc-finger nucleases for somatic gene therapy: the next frontier. *Hum Gene Ther*, 22, 925-33.
- Ramirez, C. L., J. E. Foley, D. A. Wright, F. Muller-Lerch, S. H. Rahman, T. I. Cornu, R. J. Winfrey, J. D. Sander, F. Fu, et al. (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods*, 5, 374-5.
- Reyon, D., S. Q. Tsai, C. Khayter, J. A. Foden, J. D. Sander & J. K. Joung (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*, 30, 460-5.
- Rivat, C., G. Santilli, H. B. Gaspar & A. J. Thrasher (2012) Gene therapy for primary immunodeficiencies. *Hum Gene Ther*, 23, 668-75.
- Roberts, R. J., M. Belfort, T. Bestor, A. S. Bhagwat, T. A. Bickle, J. Bitinaite, R. M. Blumenthal, S. Degtyarev, D. T. Dryden, et al. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res*, 31, 1805-12.
- Rogers, F. A., J. A. Lloyd & P. M. Glazer (2005) Triplex-forming oligonucleotides as potential tools for modulation of gene expression. *Curr Med Chem Anticancer Agents*, 5, 319-26.
- Romer, P., S. Hahn, T. Jordan, T. Strauss, U. Bonas & T. Lahaye (2007) Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science*, 318, 645-8.
- Romer, P., T. Strauss, S. Hahn, H. Scholze, R. Morbitzer, J. Grau, U. Bonas & T. Lahaye (2009) Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper Bs3 alleles. *Plant Physiol*, 150, 1697-712.
- Rosen, L. E., H. A. Morrison, S. Masri, M. J. Brown, B. Springstubb, D. Sussman, B. L. Stoddard & L. M. Seligman (2006) Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Research*, 34, 4791-4800.
- Rouet, P., F. Smih & M. Jasin (1994) Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A*, 91, 6064-8.
- Rubelt, M. S., S. Amasheh, T. Grobosch & C. Stein (2012) Liquid chromatography-tandem mass spectrometry for analysis of intestinal permeability of loperamide in physiological buffer. *PLoS One*, 7, e48502.
- Rusling, D. A., N. Laurens, C. Pernstich, G. J. Wuite & S. E. Halford (2012) DNA looping by FokI: the impact of synapse geometry on loop topology at varied site orientations. *Nucleic Acids Res*, 40, 4977-87.
- Samson, M., F. Libert, B. J. Doranz, J. Rucker, C. Liesnard, C. M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cognaux, et al. (1996) Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*, 382, 722-5.
- Sander, J. D., E. J. Dahlborg, M. J. Goodwin, L. Cade, F. Zhang, D. Cifuentes, S. J. Curtin, J. S. Blackburn, S. Thibodeau-Beganny, et al. (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods*, 8, 67-9.
- Sanger, F., S. Nicklen & A. R. Coulson (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74, 5463-7.
- Sapranaukas, R., G. Sasnauskas, A. Lagunavicius, G. Vilkaitis, A. Lubys & V. Siksnys (2000) Novel subtype of type II restriction enzymes. BfiI endonuclease exhibits

-
- similarities to the EDTA-resistant nuclease Nuc of *Salmonella typhimurium*. *J Biol Chem*, 275, 30878-85.
- Sasnauskas, G., S. E. Halford & V. Siksnys (2003) How the BfiI restriction enzyme uses one active site to cut two DNA strands. *Proc Natl Acad Sci U S A*, 100, 6410-5.
- Sasnauskas, G., L. Zakrys, M. Zaremba, R. Cosstick, J. W. Gaynor, S. E. Halford & V. Siksnys (2010) A novel mechanism for the scission of double-stranded DNA: BfiI cuts both 3'-5' and 5'-3' strands by rotating a single active site. *Nucleic Acids Research*, 38, 2399-410.
- Schierling, B., N. Dannemann, L. Gabsalilow, W. Wende, T. Cathomen & A. Pingoud (2012) A novel zinc-finger nuclease platform with a sequence-specific cleavage module. *Nucleic Acids Res*, 40, 2623-38.
- Schmitt, M. W., S. R. Kennedy, J. J. Salk, E. J. Fox, J. B. Hiatt & L. A. Loeb (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proc Natl Acad Sci U S A*, 109, 14508-13.
- Scholze, H., J. Streubel & J. Boch (2011) TAL effectors from *Xanthomonas*: design of a programmable DNA-binding specificity. *Bulletin of Insectology*, 64, S279-S280.
- Segal, D. J., B. Dreier, R. R. Beerli & C. F. Barbas, 3rd (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci U S A*, 96, 2758-63.
- Seligman, L. M., K. M. Chisholm, B. S. Chevalier, M. S. Chadsey, S. T. Edwards, J. H. Savage & A. L. Veillet (2002) Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res*, 30, 3870-9.
- Silva, G., L. Poirot, R. Galetto, J. Smith, G. Montoya, P. Duchateau & F. Paques (2011) Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther*, 11, 11-27.
- Simoncsits, A., M. L. Tjornhammar, T. Rasko, A. Kiss & S. Pongor (2001) Covalent joining of the subunits of a homodimeric type II restriction endonuclease: single-chain PvuII endonuclease. *J Mol Biol*, 309, 89-97.
- Smih, F., P. Rouet, P. J. Romanienko & M. Jasin (1995) Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res*, 23, 5012-9.
- Smith, J., S. Grizot, S. Arnould, A. Duclert, J. C. Epinat, P. Chames, J. Prieto, P. Redondo, F. J. Blanco, et al. (2006) A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res*, 34, e149.
- Spyridaki, A., C. Matzen, T. Lanio, A. Jeltsch, A. Simoncsits, A. Athanasiadis, E. Scheuring-Vanamee, M. Kokkinidis & A. Pingoud (2003) Structural and biochemical characterization of a new Mg(2+) binding site near Tyr94 in the restriction endonuclease PvuII. *J Mol Biol*, 331, 395-406.
- Stein, S., M. G. Ott, S. Schultze-Strasser, A. Jauch, B. Burwinkel, A. Kinner, M. Schmidt, A. Kramer, J. Schwable, et al. (2010) Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med*, 16, 198-204.
- Stieger, K. & B. Lorenz (2010) Gene therapy for vision loss -- recent developments. *Discov Med*, 10, 425-33.
- Stoddard, B. L. (2005) Homing endonuclease structure and function. *Q Rev Biophys*, 38, 49-95.
- (2011) Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure*, 19, 7-15.
- Streubel, J., C. Blucher, A. Landgraf & J. Boch (2012) TAL effector RVD specificities and efficiencies. *Nature Biotechnology*, 30, 593-595.

-
- Sussman, D., M. Chadsey, S. Fauce, A. Engel, A. Bruett, R. Monnat, Jr., B. L. Stoddard & L. M. Seligman (2004) Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J Mol Biol*, 342, 31-41.
- Szczepek, M., V. Brondani, J. Buchel, L. Serrano, D. J. Segal & T. Cathomen (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol*, 25, 786-93.
- Takeuchi, R., A. R. Lambert, A. N. Mak, K. Jacoby, R. J. Dickson, G. B. Gloor, A. M. Scharenberg, D. R. Edgell & B. L. Stoddard (2011) Tapping natural reservoirs of homing endonucleases for targeted gene modification. *Proc Natl Acad Sci U S A*, 108, 13077-82.
- Tang, W. (2013a) Phase 1 dose escalation study of autologous T-cells genetically modified at the CCR5 gene by zinc-finger nucleases in HIV-infected patients, <http://clinicaltrials.gov/ct/show/NCT01044654>.
- (2013b) Study of autologous T-cells genetically modified at the CCR5 gene by zinc finger nucleases in HIV-infected subjects; <http://clinicaltrials.gov/ct2/show/record/NCT01252641>.
- Tebas P, S. D. (2013) Autologous T-cells genetically modified at the CCR5 gene by zinc finger nucleases SB-728 for HIV (zinc-finger); <http://clinicaltrials.gov/ct2/show/NCT00842634>.
- UniProt, C. (2010) The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res*, 38, D142-8.
- (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res*, 40, D71-5.
- Urnov, F. D., J. C. Miller, Y. L. Lee, C. M. Beausejour, J. M. Rock, S. Augustus, A. C. Jamieson, M. H. Porteus, P. D. Gregory, et al. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 435, 646-51.
- Urnov, F. D., E. J. Rebar, M. C. Holmes, H. S. Zhang & P. D. Gregory (2010) Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*, 11, 636-46.
- Valton, J., A. Dupuy, F. Daboussi, S. Thomas, A. Marechal, R. Macmaster, K. Melliand, A. Juillerat & P. Duchateau (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J Biol Chem*, 287, 38427-32.
- Vanamee, E. S., S. Santagata & A. K. Aggarwal (2001) FokI requires two specific DNA sites for cleavage. *J Mol Biol*, 309, 69-78.
- Vasquez, K. M., K. Marburger, Z. Intody & J. H. Wilson (2001) Manipulating the mammalian genome by homologous recombination. *Proc Natl Acad Sci U S A*, 98, 8403-10.
- Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, et al. (2001) The sequence of the human genome. *Science*, 291, 1304-51.
- Vitkute, J., Z. Maneliene, M. Petrusyte & A. Janulaitis (1998) BfiI, a restriction endonuclease from *Bacillus firmus* S8120, which recognizes the novel non-palindromic sequence 5'-ACTGGG(N)5/4-3'. *Nucleic Acids Res*, 26, 3348-9.
- Wang, H., H. Yang, Chikdu S. Shivalila, Meelad M. Dawlaty, Albert W. Cheng, F. Zhang & R. Jaenisch (2013) One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell*.
- Wang, Z., J. Li, H. Huang, G. Wang, M. Jiang, S. Yin, C. Sun, H. Zhang, F. Zhuang, et al. (2012) An integrated chip for the high-throughput synthesis of transcription activator-like effectors. *Angew Chem Int Ed Engl*, 51, 8505-8.
- Watson, J. D. & F. H. Crick (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171, 737-8.

-
- Wiedenheft, B., S. H. Sternberg & J. A. Doudna (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, 482, 331-8.
- Wood, A. J., T. W. Lo, B. Zeitler, C. S. Pickle, E. J. Ralston, A. H. Lee, R. Amora, J. C. Miller, E. Leung, et al. (2011) Targeted genome editing across species using ZFNs and TALENs. *Science*, 333, 307.
- Wright, D. A., S. Thibodeau-Beganny, J. D. Sander, R. J. Winfrey, A. S. Hirsh, M. Eichinger, F. Fu, M. H. Porteus, D. Dobbs, et al. (2006) Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. *Nat Protoc*, 1, 1637-52.
- Yang, B., A. Sugio & F. F. White (2006) Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proc Natl Acad Sci U S A*, 103, 10503-8.
- Zaremba, M., C. Urbanke, S. E. Halford & V. Siksnys (2004) Generation of the BfiI restriction endonuclease from the fusion of a DNA recognition domain to a non-specific nuclease from the phospholipase D superfamily. *J Mol Biol*, 336, 81-92.
- Zhang, F., L. Cong, S. Lodato, S. Kosuri, G. M. Church & P. Arlotta (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*, 29, 149-53.
- Zhu, C., T. Smith, J. McNulty, A. L. Rayla, A. Lakshmanan, A. F. Siekmann, M. Buffardi, X. Meng, J. Shin, et al. (2011) Evaluation and application of modularly assembled zinc-finger nucleases in zebrafish. *Development*, 138, 4555-64.
- Zhu, Z., Blanchard, A., Xu, S.-Y., Wei, H., Zhang, P., Sun, D., and Chan, S.-H. (2009) High fidelity restriction endonucleases. *U.S. Patent 20090029376*.
- Zou, J., M. L. Maeder, P. Mali, S. M. Pruetz-Miller, S. Thibodeau-Beganny, B. K. Chou, G. Chen, Z. Ye, I. H. Park, et al. (2009) Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell*, 5, 97-110.
- Zou, J., C. L. Sweeney, B. K. Chou, U. Choi, J. Pan, H. Wang, S. N. Dowey, L. Cheng & H. L. Malech (2011) Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. *Blood*, 117, 5561-72.