

**Analyse der Mismatch-Toleranz des Streptococcus pyogenes CRISPR-Cas9
Systems**

Inauguraldissertation
zur Erlangung des Grades eines Doktors der Medizin
des Fachbereichs Medizin
der Justus-Liebig-Universität Gießen

vorgelegt von Hoßfeld, Lars Thomas
aus Heidelberg

Gießen 2024

Aus dem Fachbereich Medizin der Justus-Liebig-Universität Gießen
Klinik und Poliklinik für Augenheilkunde

Gutachter: Prof. Dr. Dr. Knut Stieger

Gutachter: Prof. Dr. Andreas Bräuninger

Tag der Disputation: 29.10.2024

1	Einleitung.....	1
1.1	Das CRISPR-Cas-System.....	1
1.1.1	Aufbau und Funktionsweise von <i>CRISPR-Cas</i>	1
1.1.2	Herausforderung <i>Mismatch</i> -Toleranz und <i>Off-Target</i> -Aktivität.....	2
1.1.2.1	Begriffserklärung und klinische Relevanz.....	2
1.1.2.2	Mechanismen der Interaktion von <i>gRNA-Cas9</i> und <i>Target-DNA</i>	2
1.1.2.3	Wissensstand der <i>Mismatch</i> -Toleranz.....	4
1.1.2.4	Erklärungsmodelle der <i>Mismatch</i> -Toleranz.....	6
1.1.2.5	Antizipation und Detektion von <i>Off-Target</i> -Aktivität.....	7
1.1.2.6	Strategien zur Reduktion von <i>Off-Target</i> -Aktivität.....	8
1.1.3	Medizinische Anwendungsbereiche und Potential von <i>CRISPR-Cas</i>	10
1.2	DSB-Reparaturmechanismen und ihre Relevanz für die Geneditierung.....	11
1.2.1	Wahrung der genomischen Stabilität.....	11
1.2.2	<i>Non Homologous End Joining (NHEJ)</i>	11
1.2.3	<i>Microhomology Mediated End Joining (MMEJ)</i>	12
1.2.4	<i>Homology Directed Repair (HDR)</i>	13
1.3	Methoden zum Nachweis von DNA-Editierungen.....	14
1.3.1	DNA basierte Systeme.....	14
1.3.2	Fluoreszenzbasierte Systeme.....	14
1.4	Zielsetzung.....	15
2	Material und Methoden.....	16
2.1	Material.....	16
2.1.1	Geräte.....	16
2.1.2	Verbrauchsmaterialien.....	17
2.1.3	Chemikalien.....	17
2.1.4	Plasmide.....	18
2.1.4.1	Plasmid <i>RLuc8</i>	18
2.1.4.2	Plasmid <i>pRLuc8-AvrII/BsiWI-GFP2</i>	19
2.1.4.3	Plasmid <i>px459</i>	19
2.1.5	Enzyme.....	20
2.1.5.1	Restriktionsenzyme.....	20
2.1.5.2	Weitere Enzyme.....	21
2.1.6	Oligonukleotide.....	21
2.1.6.1	Wildtypsequenzen.....	22
2.1.6.2	Veränderte Wildtypsequenzen.....	22

2.1.6.3	Sequenzierungsprimer	25
2.1.7	Organismen	26
2.1.7.1	Bakterienstämme.....	26
2.1.7.2	Eukaryotische Zellen:	26
2.1.8	Antibiotika für Bakterienkulturen	26
2.1.9	Kulturmedien.....	27
2.1.9.1	Kulturmedien für Bakterien	27
2.1.9.2	Kulturmedien für eukaryotische Zellen	27
2.1.10	Puffer	28
2.1.11	Kit-Systeme	28
2.1.12	Ausgangsmaterialien	28
2.2	Methoden	29
2.2.1	Arbeiten mit <i>DNA</i>	29
2.2.1.1	Verdau mittels Restriktionsenzymen	29
2.2.1.1.1	Nachverdau des Ligationsansatzes	29
2.2.1.2	Hybridisierung von Oligonukleotiden	30
2.2.1.3	Ligation von <i>DNA</i>	30
2.2.1.4	Konzentrationsbestimmung von <i>DNA</i> via Photometrie.....	31
2.2.1.5	Sequenzierung von Plasmidklonen	31
2.2.2	Arbeiten mit Bakterien	31
2.2.2.1	Herstellung von elektrokompetenten <i>E.coli</i> -Bakterien.....	31
2.2.2.2	Transformation von elektrokompetenten <i>E.coli</i> -Bakterien.....	32
2.2.2.3	Herstellung von mit Ampicillin versetzten LB-Agar-Platten:	32
2.2.2.4	Beimpfung flüssiger LB-Kulturen	32
2.2.2.5	Herstellung von Glycerolstocks	33
2.2.2.6	Isolation von Plasmid <i>DNA</i>	33
2.2.3	Arbeiten mit eukaryotischen Zellen.....	33
2.2.3.1	Kultivierung von <i>HEK 293T</i> -Zellen	33
2.2.3.2	Transfektion von <i>HEK 293T</i> -Zellen	33
2.2.3.3	Fluoreszenzmikroskopie.....	34
2.2.4	Arbeiten mit Proteinen	35
2.2.4.1	<i>BRET-ASSAY</i>	35
2.2.5	Statistik und Berechnungen.....	36
2.2.5.1	Statistische Auswertung	36
2.2.5.2	Berechnung der <i>frameshift rate</i>	36
3	Ergebnisse.....	37

3.1	Klonierungen der <i>Target</i>-Sequenzen	37
3.2	Transfektion in <i>HEK 293T</i>-Zellen.....	38
3.3	<i>BRET</i>-Assays	41
3.3.1	<i>Frameshift rates</i> singulärer <i>Mismatches</i>	41
3.3.2	<i>Frameshift rates</i> multipler <i>Mismatches</i>	46
4	Diskussion	49
4.1	Einfluss singulärer <i>Mismatches</i> auf die Aktivität von <i>gRNA-Cas9</i>	49
4.2	Einfluss multipler <i>Mismatches</i> auf die Aktivität von <i>gRNA-Cas9</i>	54
4.3	Variabilität der <i>Mismatch</i> -Toleranz	56
4.4	Möglichkeiten und Einschränkungen des <i>BRET-Reporters</i>	57
4.5	Schlussfolgerungen und Ausblick.....	58
5	Zusammenfassung	59
6	Summary.....	60
7	Abbildungsverzeichnis.....	61
8	Tabellenverzeichnis	62
9	Abkürzungsverzeichnis.....	63
10	Literaturverzeichnis	67
11	Anhang	77
12	Ehrenwörtliche Erklärung.....	82
13	Danksagung	83

1 Einleitung

1.1 Das CRISPR-Cas-System

1.1.1 Aufbau und Funktionsweise von CRISPR-Cas

Der *Clustered Regularly Interspaced Short Palindromic Repeats (C.R.I.S.P.R.)-Locus* ist eine Region im Genom der meisten Prokaryoten. Dieser besteht aus sich wiederholenden *DNA*-Sequenzen ähnlicher Länge mit dazwischen liegenden variablen *DNA*-Sequenzen. Die sich nicht wiederholenden Sequenzen werden als *Spacer* bezeichnet. In unmittelbarer Nähe davon befinden sich *CRISPR-associated genes (Cas)* (Jansen et al. 2002). *Spacer* sind integrierte *DNA*-Bruchstücke aus dem Genom vergangener viraler Infektionen. Die Integration dieser vermittelt eine gewisse Resistenz gegenüber erneuten Virusinfektionen. Das *CRISPR-Cas*-System fungiert somit als adaptives Immunsystem von Prokaryoten (Barrangou et al. 2007). Vermittelt wird diese Immunität durch die Interaktion des *CRISPR-Cas*-Systems mit dem eindringenden viralen Genom. Dafür wird der *CRISPR-Locus* transkribiert. Es entstehen eine *crRNA* und eine *trans-activating crRNA (tracrRNA)*, welche mit dem/den *Cas*-Enzym(en) eine funktionelle Einheit bilden. Die *crRNA* ist komplementär zum *Spacer* und bindet am Genom des eindringenden Virus. Die *tracrRNA* dient der Stabilisierung der *crRNA* und ermöglicht die Komplexbildung mit dem *Cas*-Enzym. Das *Cas*-Enzym fungiert dann als Endonuklease und induziert einen *DNA*-Bruch im viralen Genom (Jiang and Doudna 2017). In diesem Kontext sei die enorme Fülle der unterschiedlichen *CRISPR-Cas* Klassen und Subtypen betont, die sich in der Organisation des *CRISPR-Locus*, der *RNA*-Prozessierung und letztendlich auch der *Cas*-Enzyme unterscheiden (Koonin and Makarova 2017). Da die *crRNA* auch komplementär zum *CRISPR-Locus* selbst ist, muss eine weitere Voraussetzung zur Bindung des *CRISPR-Cas*-Komplexes erfüllt sein, um zwischen eigener und fremder *DNA* zu differenzieren. Dabei handelt es sich um das *protospacer adjacent motif*, kurz *PAM*. Dies ist ein kurzer *DNA*-Abschnitt in unmittelbarer Nähe der Ziel-*DNA* des Virus (*Target*-Sequenz), der vom *Cas*-Enzym gebunden wird und nicht Teil des *CRISPR-Locus* ist. Das *PAM* variiert je nach Organismus und *CRISPR*-Typ (Hille et al. 2018).

Daraus ergibt sich die gentechnische Anwendung von *CRISPR-Cas* als ein *RNA* geführtes System zur gezielten *DNA*-Modifikation. Dabei können *crRNA* und *tracrRNA* zu einer *single guide RNA (gRNA)* fusioniert werden. Die mit dem *Cas*-Enzym verbundene, individuell konstruierbare *gRNA* führt dieses zur gewünschten *Target*-Sequenz (Jinek et al. 2012).

In dieser Arbeit verwendet wurde *SpCas9*. Dieses ist das *Cas*-System von *Streptococcus pyogenes*. Auch hier wurden *crRNA* und *tracrRNA* zu einer *gRNA* vereint. Die *crRNA* von *SpCas9* umfasst eine Länge von 20 Nukleotiden (nt). Das *PAM* von *SpCas9* ist 5'-NGG-3' (Mali et al. 2013). Folglich bestehen die hier verwendeten *Target*- und *Guide*-Sequenzen aus 20 Nukleotiden (2.1.6). *SpCas9* besitzt zwei Endonukleasedomänen. Diese sind *HNH* und *RuvC* und schneiden jeweils einen Strang der *Target-DNA* und induzieren so einen Doppelstrangbruch (DSB) (Jinek et al. 2012).

1.1.2 Herausforderung *Mismatch*-Toleranz und *Off-Target*-Aktivität

1.1.2.1 Begriffserklärung und klinische Relevanz

Idealerweise bindet und schneidet *gRNA-Cas9* eine *DNA*-Sequenz, die vollständig komplementär zur konstruierten *gRNA*-Sequenz ist. Jedoch können auch Sequenzen gebunden und geschnitten werden, welche von dieser perfekten Komplementarität abweichen und abseits der *Target*-Sequenz liegen. Diese werden als *Off-Target*-Sequenzen bezeichnet. Die hierbei abweichenden Nukleotide werden als *Mismatch* bzw. *Mismatches* bezeichnet (Mitchell et al. 2020). Rückblickend auf die ursprüngliche Funktion von *CRISPR-Cas* als ein adaptives Immunsystem von Prokaryoten gegenüber Viren ist dies durchaus sinnvoll. So konnten auch mutierte und genetisch verwandte Bakteriophagen nicht der Immunreaktion entgehen (Semenova et al. 2011). Problematisch ist dies jedoch in der therapeutischen Anwendung von *CRISPR-Cas*. Da die Mechanismen der *Mismatch*-Toleranz noch nicht vollständig verstanden sind, besteht die Gefahr nicht intendierte Genareale zu modifizieren. *Off-Target*-Aktivität in Proto-Onkogenen oder Tumor-Suppressorgenen könnte so zur malignen Transformation der modifizierten Zelle führen (Blattner et al. 2020). Das onkogene Potential unterstützend, kann es durch multiple induzierte DSB zu großen chromosomalen Umstrukturierungen kommen. Beschrieben sind dabei Translokationen, Chromothripsis sowie Aneuploidie (Wienert et al. 2022).

1.1.2.2 Mechanismen der Interaktion von *gRNA-Cas9* und *Target-DNA*

Die Offenlegung der molekularen Prozesse, die zwischen Bindung der *Target-DNA* und letztendlich Aktivierung der Endonukleasedomänen von *Cas9* liegen, sind essentiell für das Verständnis der *Target*-Spezifität, also der Diskriminierung zwischen *On-Target*- und *Off-Target*-Sequenzen. Dabei durchlaufen die verschiedenen Domänen von *Cas9* mehrere Konformationsänderungen und räumliche Umstrukturierungen. Diese bedingen die Regulation und Vermittlung der Endonukleaseaktivität und sind Ansatzpunkte in der Entwicklung von *Cas*-Varianten mit höherer Spezifität (Zhu et al. 2019).

Strukturelle Darstellungen von *Cas9* zeigen eine bilobuläre, halbmondförmige Architektur, bestehend aus einem Nuklease-Lobus und einem α -helikalen-Lobus. Der Nuklease-Lobus umfasst die beiden Endonukleasedomänen *HNH* und *RuvC* sowie eine C-Terminale-Domäne (CTD) (Jinek et al. 2014). Der α -helikale-Lobus wird als *Recognition*-Lobus bezeichnet und besteht aus drei α -helikalen Domänen. Daraus resultieren die Bezeichnungen *REC 1-3* (Jiang and Doudna 2017).

Die Interaktion zwischen *gRNA-Cas9* und der *Target-DNA* beginnt mit der Erkennung des *PAM*. Die Paarung erfolgt dabei nicht über komplementäre Nukleotide, sondern über die Interaktion von basenspezifischen Wasserstoffbrückenbindungen des *DNA*-Doppelstrangs mit zwei Arginin-Überhängen der CTD (Anders et al. 2014). Die Arginin-Überhänge sind starr verankert und zeigen wenig strukturelle Plastizität. Dadurch gewährleistet *Cas9* eine gewisse *PAM*-Spezifität (Anders et al. 2016). Durch diesen Mechanismus ist *Cas9* nun mit dem *PAM* verankert und kann mit der unmittelbar angrenzenden *Target-DNA* interagieren. Besagte Positionierung bedingt das räumliche Zusammenführen des Phosphat-Desoxyribose-Rückgrates des *Target-DNA*-Stranges (*TS*) mit Amid-Gruppen von randständigen Aminosäuren der CTD. Die so entstehenden Wasserstoffbrückenbindungen und elektrostatischen Wechselwirkungen bedingen ein Abknicken des *TS* und damit die Ablösung vom komplementären gegenläufigen *DNA*-Strang. Letzterer wird als *Non-Target-DNA*-Strang (*NTS*) bezeichnet. *TS* und *NTS* trennen sich initial nur in unmittelbarer Nähe des *PAM*. Die weitere Ablösung wird nun durch die schrittweise Basenpaarung von *gRNA* und *TS* vorangetrieben. Es entsteht also ein Heteroduplex aus *gRNA* und *TS* sowie ein ungepaarter *NTS* (Anders et al. 2014). Der Heteroduplex gelangt in einen zentralen Tunnel zwischen *Recognition*- und Nuklease-Lobus. Parallel dazu verläuft der *NTS* in einem gesonderten Tunnel innerhalb des Nuklease-Lobus. So gelingt die Positionierung des *NTS* innerhalb des katalytischen Zentrums von *RuvC* (Jiang et al. 2016).

Das Zusammenbringen der aktiven *HNH*-Endonukleasedomäne und des *TS* ist komplex und abhängig von mehreren Konformationsänderungen des *Recognition*-Lobus (Palermo et al. 2018). *HNH* kann in drei verschiedenen Konformationen vorliegen, die nacheinander im Aktivierungsprozess durchlaufen werden. Ohne Bindung eines *gRNA-DNA*-Heteroduplexes wechselt *HNH* zwischen einer inaktiven und intermediären Form. Erst bei Bindung des Heteroduplexes gelingt eine Überführung in die aktive Form. *HNH* verweilt bzw. arretiert bei unzureichender Komplementarität von *gRNA* und *TS* in der intermediären Form. Dementsprechend fungiert diese als *conformational checkpoint*, den *HNH* erst passieren muss, um katalytisch aktiv zu werden (Dagdas et al. 2017).

Vermittelt wird die Überführung in die aktive Konformation über die drei Helices des *Recognition*-Lobus. *REC 3* kommt hierbei eine tragende Rolle in der *Target*-Spezifität zu. Zunächst haben *REC 3* und *HNH* keinen räumlichen Kontakt zueinander. Dieser wird erst durch den Heteroduplex vermittelt (Palermo et al. 2018). *REC 3* fungiert nun als allosterischer Effektor von *HNH* und erfasst durch intermolekulare Kontakte die distale Komplementarität des Heteroduplexes. Bei ausreichender Komplementarität kommt es zu einer Konformationsänderung von *REC 3*. Dies führt zu einer Auswärtsrotation von *REC 2*. *REC 2* liegt im inaktiven Zustand zwischen dem Phosphat-Rückgrat des Heteroduplexes und *HNH*. Folglich bedingt *REC 2* deren räumliche Trennung und fungiert als sterischer Okkluder. Die Auswärtsrotation ermöglicht so das Zusammenkommen der *HNH*-Endonukleasedomäne mit dem Heteroduplex (Chen et al., 2017). *REC 1* erfährt in dieser Kaskade auch eine räumliche Neuorientierung. *REC 1* nähert sich *HNH* und dem Heteroduplex an und festigt deren Verbindung durch ionische Interaktionen (Palermo et al. 2018).

Die Endonukleasedomäne von *HNH* schneidet letztendlich den *TS*, die von *RuvC* den *NTS*. Dies erfolgt fast zeitgleich. Gewährleistet wird das durch zwei α -Helices, die *HNH* und *RuvC* verbinden. Sie werden als *Linker 1* und *Linker 2* bezeichnet. Eine Aktivierung von *HNH* bedingt somit auch eine Aktivierung von *RuvC* durch Repositionierung. *HNH* kontrolliert folglich allosterisch *RuvC* (Sternberg et al. 2015).

Verantwortlich für das Schneiden des *DNA*-Stranges sind die aktiven katalytischen Residuen von *HNH* und *RuvC*. Diese werden, u.a. koordiniert von zweiwertigen Metallionen, durch die obigen Konformationsänderungen in räumliche Nähe zum Phosphat-Desoxyribose-Rückgrat der *DNA* gebracht. Es resultiert eine nukleophile Attacke bzw. Substitution des Residuums an der jeweiligen kovalenten Bindungen des Phosphat-Desoxyribose-Rückrates mit Unterbrechung dieser (Das et al. 2023).

1.1.2.3 Wissensstand der *Mismatch*-Toleranz

Der in 1.1.2.2 erwähnte *conformational block* verdeutlicht, dass nicht alle von *gRNA*-*Cas9* gebundenen *DNA*-Abschnitte auch geschnitten werden. Die Bindung und anschließende Dissoziation von *Cas9* ist deutlich promisker als die Bindung mit anschließender Endonukleaseaktivität. Letzteres setzt eine ausgeprägtere Komplementarität zwischen *gRNA* und *TS* voraus (Wu et al., 2014). Eine Komplementarität der ersten 5-10 an das *PAM* angrenzenden Nukleotide reicht aus, um *Cas9* fest an die *Target-DNA* zu binden. *Mismatches* in diesem Bereich beeinflussen sowohl die Bindungsaffinität als auch die Endonukleaseaktivität in großem Maße. Dieser Bereich trägt aufgrund seiner Bedeutung den Namen *Seed*. *Mismatches* außerhalb des

Seeds, also distal des *PAM*, beeinflussen kaum die Bindungsaffinität, aber sehr wohl die Endonukleaseaktivität (Boyle et al. 2021).

Welche *Mismatches* toleriert werden, hängt von mehreren Faktoren ab, wie deren Position innerhalb des 20 Nukleotide langen *TS*, die dadurch miteinander interagierenden nicht kanonischen Basenpaare sowie die Anzahl der *Mismatches* (Hsu et al. 2013).

Mismatches proximal des *PAM* werden tendenziell weniger toleriert als solche, die distal davon liegen. Die Abhängigkeit der Toleranz von der Position scheint jedoch nicht ganz linear abzunehmen. Vielmehr findet sich innerhalb des *Seeds* eine zweiteilige Toleranz. Kontraintuitiv zeigen *Mismatches* der Positionen 1-4 eine ausgeprägtere Toleranz als die der Positionen 5-8 (Tsai et al. 2015; Kim et al. 2020a). Ebenfalls bedacht werden müssen *Mismatches* des *PAM* selbst. Von besonderer Relevanz sind 5'-NAG-3', 5'-NGA-3' und 5'-NCG-3'. *SpCas9* ist in der Lage diese zu binden und sie sollten deshalb in *Off-Target*-Analysen integriert werden (Doench et al. 2016). Zudem beeinflusst auch der vor dem *PAM* gelegene *DNA*-Abschnitt, welcher nicht entwunden wird, die Bindungsaffinität von *Cas9* (Boyle et al., 2021; Zhang et al., 2019).

Nicht nur die Position des *Mismatches*, sondern auch die dort miteinander interagierenden Nukleotide von *gRNA* und *TS* beeinflussen dessen Toleranz. In diesem Kontext finden sich in Massensequenzierungen je nach Nukleotidinteraktion deutliche intrapositionelle Unterschiede (Boyle et al., 2017; Fu et al., 2022; Doench et al., 2016).

Auch die Anzahl an *Mismatches* innerhalb der *Target*-Sequenz beeinflusst deren Toleranz. Mit zunehmender Anzahl sinkt die Toleranz dieser (Kim et al. 2020a). Die Kombination singulärer *Mismatches* zu multiplen *Mismatches* ergibt dabei einen kombinatorischen Effekt, der von den Positionierungen und dem Abstand der *Mismatches* zueinander abhängt (Fu et al. 2022).

Es sei jedoch hervorzuheben, dass es sich bei den eben aufgezählten Regeln größtenteils um in vitro, in vivo und in silico ermittelte Generalisierungen handelt. Die Reproduzierbarkeit dieser ist nicht immer gegeben. Ursächlich sind dabei u.a. neben interindividuellen genetischen Variationen (Wienert et al. 2022) auch guide-intrinsische Eigenschaften. Letzteres meint, dass jeder *gRNA-Cas9*-Komplex eine individuelle *Mismatch*-Toleranz inne hat, die unabhängig vom *Mismatch*-Kontext ist und durch die gesamte *gRNA*-Sequenz determiniert wird (Huston et al. 2019; Fu et al. 2022).

1.1.2.4 Erklärungsmodelle der *Mismatch*-Toleranz

Ein Erfolg versprechendes Erklärungsmodell der *Mismatch*-Toleranz liegt in der biophysikalischen Betrachtung der *R-Loop*-Kinetik. Die vollständige Ausbildung des *R-Loops* bedingt die Repositionierung der Endonukleasedomänen und überführt *Cas9* in einen katalytisch aktiven Zustand (Gong et al. 2018). Als *R-Loop* bezeichnet wird die Hybridstruktur aus *gRNA* und *TS*. Die *R-Loop*-Entwicklung beginnt folglich mit Erkennung des *PAM* und anschließend schrittweiser unidirektionaler Paarung der Nukleotide von *gRNA* und *TS* (Szczelkun et al. 2014). Die Überlegung dahinter ist, dass innerhalb des *R-Loop*-Progresses mehrere metastabile Stadien durchlaufen werden. Der Übergang innerhalb dieser wird durch die energetische Landschaft bestimmt. Es wird jeweils der energetisch günstigste Zustand bzw. die niedrigste energetische Barriere angestrebt. Bei perfekter Komplementarität ist demnach die energetische Barriere hinsichtlich des vollständigen *R-Loop*-Progresses und letztendlich Katalyse geringer als die des *R-Loop*-Kollaps mit Dissoziation von *gRNA-Cas9* und *Target-DNA*. *Mismatches* erhöhen in diesem Kontext die energetische Barriere in unterschiedlichem Maße. Ab Überschreitung einer gewissen Schwelle wird dann das *Mismatch* nicht mehr toleriert und die Dissoziation gegenüber dem weiteren *R-Loop*-Progress energetisch favorisiert (Klein et al. 2018). Der Progress vom geschlossenen zum vollständig geöffneten *R-Loop* wird über einen intermediären Zustand erreicht. Dieser wird nach der Hybridisierung von ca. 9-10 Basenpaaren erreicht und liegt somit ziemlich genau im Bereich des *Seeds*. Von der zeitlichen Dynamik fällt diese intermediäre *R-Loop*-Konformation zusammen mit der intermediären *HNH*-Konformation, so dass diese eng verbunden zu sein scheinen. *Mismatches* distal des *Seeds* ermöglichen die stabile Ausbildung des intermediären *R-Loops*, destabilisieren aber den Übergang in den vollständig offenen *R-Loop*. Dieses längere Verbleiben im intermediären Zustand korrespondiert mit der unproduktiven stabilen Bindung von *Cas9* an der *Target-DNA*. *Mismatches* innerhalb des *Seeds* destabilisieren sowohl den intermediären als auch den offenen *R-Loop*. Das erklärt weshalb *Cas9* hier teilweise schneller dissoziiert (Ivanov et al. 2020). Dieses Modell unterstützend, lassen sich auf der Grundlage der Massensequenzierungsreihen *NucleaSeq* und *CHAMP* die positionsbezogenen energetischen Barrieren und deren Beeinflussung durch *Mismatches* errechnen. Dabei werden drei lokale energetische Minima deutlich, welche mit den drei Konformationen des *R-Loops* (geschlossen, intermediär, vollständig offen) zusammenfallen. Zwischen diesen findet sich eine stark fluktuierende Energielandschaft, was deren Relevanz innerhalb der *R-Loop*-Kinetik verdeutlicht (Eslami-Mossallam et al. 2022). Die Stabilität und Kinetik des *R-Loop* ist stark von der jeweiligen *DNA*-Topologie abhängig. Bereiche mit negativem *Torque* bzw. *Supercoiling* erleichtern die *R-Loop*-Progression. Bereiche

mit positivem *Torque* verhalten sich gegensätzlich (Szczelkun et al. 2014; Ivanov et al. 2020). *Torque* bzw. *Supercoiling* ist damit ein wichtiger Faktor der *Mismatch*-Toleranz. So zeigt *Cas9* in Bereichen des negativen *Supercoiling*s eine ausgeprägtere *Mismatch*-Toleranz bzw. Promiskuität. Dies erklärt das häufigere Auftreten von *Off-Target*-Aktivität in Promoter-Bereichen (Ivanov et al. 2020). In diesen induziert die *RNA*-Polymerase im Rahmen der Transkription Bereiche des starken negativen *Supercoiling*s (Ma et al. 2013).

Ergänzt wird das biophysikalische Erklärungsmodell durch strukturelle Analysen. Die *R-Loop*-Formierung ist eng an räumliche Umstrukturierungen der *Recognition*- und Nukleasedomänen gekoppelt. Von Relevanz ist dabei die Architektur des *gRNA-TS*-Heteroduplexes. Nur die *gRNA* ist fest mit *Cas9* verbunden. Dies ermöglicht die Ausbildung von intermolekularen Nukleotidkontakten außerhalb der klassischen Watson-Crick-Basenpaare (A-T und C-G) durch Distorsionen des *TS* (Pacesa et al. 2022b; Pacesa et al. 2022a) Folglich fehlt *Cas9* ein sterischer Mechanismus, der ausschließlich Watson-Crick-Basenpaarungen erzwingt. *Mismatches* können so stabil in die Heteroduplex-Architektur integriert werden. Ein Gegenbeispiel wäre die *DNA*-Polymerase, welche ihre Spezifität über sterische Abweisung von *Mismatches* gewährleistet. Die *Mismatch* induzierten Distorsionen des *TS* können wiederum durch die strukturelle Plastizität der *Cas9*-Domänen kompensiert werden (Pacesa et al. 2022a). Je nach Position und Typ des *Mismatches* ergeben sich so spezifische Veränderungen der Heteroduplex-Architektur, die entweder kompensiert werden können oder selbst ein sterisches Hindernis der *Cas9*-Domänen darstellen und damit nicht toleriert werden (Mitchell et al. 2020). Kurzgefasst erkennt *Cas9* also *Mismatches* über indirektes Auslesen der Heteroduplexkonformation (Pacesa et al. 2022a).

1.1.2.5 Antizipation und Detektion von *Off-Target*-Aktivität

Die gezielte Identifikation von potenzieller *Off-Target*-Aktivität ist Voraussetzung für eine spätere sichere und nebenwirkungsarme therapeutische Applikation von *CRISPR-Cas*. Hierbei unterscheidet man *biased* von *unbiased* Methoden, welche sich gegenseitig ergänzen und meist nacheinander angewandt werden (Martin et al. 2016).

Biased Methoden sind auf Algorithmen basierende *in silico* Vorhersagen. Diese bioinformatischen Tools sind mit den Erkenntnissen der *Mismatch*-Toleranz programmiert. So können anhand der *gRNA*-Sequenz nahezu homologe Areale identifiziert werden, welche das Potential für *Off-Target*-Aktivität haben (Naeem et al. 2020). Ergänzt werden diese von *machine learning*-Applikationen, die über die ihnen zu Grunde liegenden Datensätze hinaus Vorhersagen treffen können (Listgarten et al.

2018; Xiang et al. 2021). Ebenso besteht die Möglichkeit verschiedene Variablen, die Einfluss auf die *Off-Target*-Aktivität nehmen, mittels *Scores* zu integrieren und miteinander zu verrechnen (Naeem et al. 2020). So können beispielsweise Chromatinzugänglichkeit und *guide*-intrinsische Eigenschaften in die *Off-Target*-Antizipation integriert werden (Huston et al. 2019).

Eben genannte Methoden unterliegen jedoch einem *Bias*, da ihre Vorhersagen auf den zu Grunde liegenden experimentellen Datensätzen beruhen. Der nächste Schritt und damit die Elimination des *Bias* ist das Zusammenbringen von *CRISPR-Cas* und *Target-DNA* mit anschließender Betrachtung des Genoms. Dies geht in vitro mit zellfreier DNA, ex vivo innerhalb lebender Zellen oder in vivo im Rahmen von Tiermodellen (Wienert et al. 2022). Eine dieser zahlreichen Möglichkeiten ist *GUIDE-Seq*. In die *CRISPR-Cas* induzierten DSB werden hierbei doppelsträngige Oligonukleotide (dsOGN) integriert. Diese können anschließend selektiv amplifiziert und sequenziert werden. Somit ist eine genomweite Darstellung der *CRISPR-Cas*-Aktivität möglich, wodurch Aussagen über *Off-Target*-Aktivität möglich sind (Tsai et al. 2015).

Letztendlich sind auch *Off-Target*-Areale interindividuell aufgrund von genetischen Variationen, so dass *Off-Target*-Vorhersagen optimalerweise dem Konzept der personalisierten Medizin folgen und oben genannte Methoden jeweils auf von der Patientin oder dem Patienten extrahierten Zellen angewendet werden (Wienert et al. 2022).

1.1.2.6 Strategien zur Reduktion von *Off-Target*-Aktivität

Da *CRISPR-Cas* das Potential hat genetische Erkrankungen zu heilen, sind die Bestrebungen groß, Strategien zur Eingrenzung von *Off-Target*-Aktivität zu entwickeln, um so eine sichere Anwendung zu garantieren. Das Repertoire ist groß und umfasst u.a. Modifikationen der *gRNA*, *Cas*-Variationen, Optimierungen des Einbringens des *gRNA-Cas*-Komplexes in die Zelle sowie *Anti-CRISPR*-Proteine (Naeem et al. 2020; Slaymaker and Gaudelli 2021).

Relevante Modifikationen der *gRNA* umfassen einerseits die Länge. Ein Kürzen dieser am 5'-Ende von 20 auf 17 Nukleotide erhöht die Spezifität bei trotzdem erhaltener *On-Target*-Aktivität. Im Gegensatz dazu führen Reduktionen auf 15-16 Nukleotide zu einem fast vollständigen Verlust der Effizienz. Dennoch zeigen abhängig von Zelltyp und Sequenz einige 17 Nukleotide lange *gRNA* deutliche Einbußen der Effizienz, so dass hier zwischen Spezifität und Effizienz abgewogen werden muss (Zhang et al. 2016; Zhao et al. 2019). Andererseits vielversprechend sind chemische Veränderungen der *gRNA*, die sich aktuell größtenteils noch auf Modifikationen des Desoxyribose- und Phosphat-

Rückgrates beschränken. Eine Ausweitung dieser dürfte zukünftig weitere Erhöhungen der Spezifität ermöglichen (Rozners, 2022).

Ein Versuch der Eindämmung mutagener Effekte durch DSB liegt in der Einführung von *Nickases* (*nCas9*). Diese sind entweder an *HNH* oder *RuvC* mutiert und somit nur fähig einen *DNA*-Strang zu schneiden. Durch die Applikation von zwei *nCas9* mit gegensätzlicher *gRNA* kann so trotzdem ein DSB induziert werden. Vorteilhaft hierbei ist, dass individuelle Einzelstrangbrüche v.a. durch den *base excision repair way* (*BER*) repariert werden. Dieser ist weniger fehlerhaft als *non homologous end joining* (*NHEJ*), welcher v.a. nach DSB induziert wird (Ran et al. 2013a). *Nickases* zeigen deutliche Reduktionen von *Off-Target*-Aktivität. Dennoch finden sich auch hier vermehrt Punktmutationen, sodass auch diese mutagenes Potential besitzen (Saifaldeen et al. 2020).

Insbesondere ambitioniert sind die Bestrebungen in der Entwicklung modifizierter *Cas*-Varianten mit erhöhter Spezifität, welche spezielle Mutationen im Vergleich zum Wildtyp aufweisen. Veränderungen fußen auf der Hypothese der Erhöhung der Spezifität durch Schwächung der Interaktion von *gRNA-Cas9* mit der *Target-DNA* (Lee et al. 2018). Zur Identifikation nützlicher Mutationen gibt es zwei Ansätze. Erstens gezielt induzierte Veränderungen basierend auf den in 1.1.2.4 und 4.1 dargestellten strukturellen Erkenntnissen. Zweitens das Testen zufälliger Mutationen und deren Spezifität in zellbasierten Experimenten (Kim et al. 2019). Beispiele für den strukturbasierten Ansatz sind *eSpCas9* (Slaymaker et al. 2016) und *Super-Fi-Cas9* (Bravo et al. 2022). Im Gegensatz dazu entstand *Sniper-Cas9* in einem *E.coli* basierten Versuch durch positive und negative Selektion (Lee et al. 2018). Es gibt eine Vielzahl weiterer *Cas*-Varianten wie *SpCas9-HF1*, *evoCas9*, *xCas9* und *HypaCas9*. Dazu sei jedoch erwähnt, dass die Spezifität und Effizienz der *Cas*-Varianten deutlich von der *Target*-Sequenz abhängt. Die verschiedenen *Cas*-Varianten präferieren unterschiedliche Nukleotide an unterschiedlichen Positionen und variieren damit in ihrer *Mismatch*-Toleranz (Kim et al. 2020b).

Ebenfalls Einfluss auf die *Off-Target*-Aktivität nimmt die *Cas9*-Expression und damit auch die Dauer der Exposition in der Ziel-Zelle. Somit spielt die Applikationsmethode eine Rolle. Die Applikation als Ribonukleoprotein (RNP) und *mRNA* zeigen kürzere Halbwertszeiten als *DNA*-basierte Methoden (Wienert et al. 2022), wobei es auch hier schon Bestrebungen von sich selbst inaktivierenden *Cas*-Systemen gibt (Wang et al. 2020).

Letztendlich hat der evolutionäre Selektionsdruck nicht nur zur Optimierung des *CRISPR-Cas*-Systems beigetragen, sondern auch Phagen haben Mechanismen entwickelt diesem zu entgehen. Diese exprimieren eine Vielzahl an *Anti-CRISPR*-Proteinen (*ACR*) (Marino et al. 2020). Ein Beispiel ist *AcrIIA4*. Dieses imitiert strukturell das *PAM* von *SpCas9* und konkurriert dadurch mit der *DNA*-Bindung. Zudem blockiert es sterisch die *RuvC*-Domäne. *Anti-CRISPR*-Proteine bieten damit die Möglichkeit der Inaktivierung nach therapeutischer Applikation mit dem Potential der Verringerung von *Off-Target*-Aktivität (Dong et al. 2017).

1.1.3 Medizinische Anwendungsbereiche und Potential von *CRISPR-Cas*

Mit der Applikation von *CRISPR-Cas* eröffnet sich die Möglichkeit der zielgenauen Manipulation von *DNA*-Sequenzen und damit das Potential personalisierter Therapien mit dem Ziel der Heilung genetischer Erkrankungen (Doudna 2020). Das Spektrum der gentechnischen Anwendung von *CRISPR-Cas* geht über das klassische *knock-out* oder *knock-in* von Genen hinaus. Mit *CRISPR activation/interference* (*CRISPRa*, *CRISPRi*) kann gezielt die Expression von Genen moduliert werden und auch epigenetische Modifikationen sind möglich (Liu et al. 2022). Der Vorteil von *CRISPRi* ist, dass hier kein DSB induziert wird und somit die Genexpression nur reversibel beeinflusst wird. Hierfür wird *dCas9* mit einer auf eine Promoter-Region ausgerichteten *gRNA* ausgestattet. Die Bindung von *dCas9* stellt dann einen sterischen Block der *RNA*-Polymerase dar und verhindert die Transkription (Larson et al. 2013). Umgekehrt besteht mit *CRISPRa* durch Fusion von *dCas9* mit Transkriptions-Aktivatoren die Möglichkeit der Steigerung der Genexpression (Perez-Pinera et al. 2013).

Weitere vielversprechende Anwendungen von *CRISPR-Cas*, die ohne Induktion eines DSB auskommen, sind *base edit* und *prime edit*. *Base edit* bietet die Möglichkeit der gezielten Veränderung einzelner Nukleotide und eignet sich somit zur Ergründung und Therapie von Punktmutationen. Hierfür wird eine Desaminase so an *dCas9* gekoppelt, dass diese bei Entwindung der *Target-DNA* mit der zu modifizierenden Base punktgenau in Kontakt kommt (Porto et al. 2020). *Prime edit* hingegen ermöglicht ein größeres Spektrum an Basen-Modifikationen als *base edit* und bietet die Möglichkeit der gezielten Deletion und Insertion von *DNA*-Abschnitten, ohne auf ein *DNA-Template* angewiesen zu sein. Hierfür wird *nCas9* mit einer reversen Transkriptase fusioniert. Die zugehörige *prime edit gRNA* (*pegRNA*) trägt sowohl die *Target* spezifizierende Sequenz als auch die für die reverse Transkriptase als Vorlage dienende Basensequenz (Anzalone et al. 2019).

CRISPR-Cas ist mittlerweile kein theoretisches Konstrukt mehr. Aktuell laufen mehrere klinische Studien mit ersten *in vivo* Applikationen beim Menschen. Vielversprechende Ergebnisse zeigen sich hier u.a. für Therapien des *Non Small Cell Lung Cancer (NSCLC)*, der Akuten Lymphatischen Leukämie (ALL), der Sichelzellanämie, sowie der Transthyretin-Amyloidose. Jedoch wird hier erst in den nächsten Jahren eine Aussage über Langzeitfolgen möglich sein. So könnten nicht detektierte *Off-Target*-Aktivitäten erst Jahrzehnte später klinisch apparent werden (Zheng et al. 2023). Ein weiteres schnell wachsendes Feld ist die diagnostische Anwendung von *CRISPR-Cas*. Einerseits zum Nachweis von *RNA* und *DNA* im Rahmen infektiöser Erkrankungen. Andererseits konnte mit dem Nachweis von humaner *CXCL9-mRNA* sensitiv eine akute Nierentransplantatabstoßung detektiert werden (Kaminski et al. 2021).

1.2 DSB-Reparaturmechanismen und ihre Relevanz für die Geneditierung

1.2.1 Wahrung der genomischen Stabilität

DSB entstehen, wenn das Phosphat-Desoxyribose-Rückgrat beider *DNA*-Stränge in unmittelbarer Nähe unterbrochen wird. Exogene Ursachen hierfür sind ionisierende Strahlung oder radiomimetische Chemikalien. Endogene DSB entstehen u.a. durch reaktive Sauerstoffspezies, die während des Zellmetabolismus anfallen. In proliferierenden Zellen ist v.a. die fragile Replikationsgabel anfällig für DSB. Diese können zum Verlust von Chromosomenabschnitten oder großen Translokationen führen, welche wiederum Zelltod oder Tumorgenese bedingen können. Dem entgegenwirkend haben Zellen komplexe *DNA*-Reparaturmechanismen entwickelt, um somit die genomische Stabilität zu wahren (Aparicio et al. 2014).

Die beiden Hauptreparatur-Mechanismen in Säugetierzellen sind *non homologous end joining (NHEJ)* und *homology directed repair (HDR)*, ferner auch *microhomology mediated end joining (MMEJ)* und *single-strand annealing (SSA)*. Das Verständnis, die Wahl und Kontrolle der verschiedenen Reparaturmechanismen ist essentiell für präzises und sicheres *genome editing*. Je nach intendierter *DNA*-Veränderung werden unterschiedliche Reparaturprozesse angestrebt, um den durch *Cas9* induzierten DSB zu reparieren (Yang et al. 2020).

1.2.2 Non Homologous End Joining (NHEJ)

NHEJ ist der überwiegende DSB-Reparaturweg in Säugetierzellen. Dieser ist unabhängig von einem homologen *DNA-Template* und aktiv in allen Phasen des Zellzyklus. *NHEJ* lässt sich in drei Subschritte unterteilen, die sequentiell durchlaufen werden. Initial die Erkennung der DSB-Enden mit Rekrutierung der *NHEJ*-Maschinerie, folgend die Prozessierung der Enden und letztendlich die Religation (Ali et al. 2022). Die

offen liegenden DSB-Enden werden vom Ku-Heterodimer (Ku70/80) gebunden. Ku70/80 rekrutiert anschließend die Proteinkinase *DNA-PKcs*, welche die Ku-Komplexe beider Enden bindet und so eine Dislokation dieser verhindert. Sofern die beiden Enden nicht direkt religierbar sind, rekrutiert die *DNA-PK* eine Reihe an Proteinen zur Enden-Prozessierung, u.a. die Artemis-Nuklease, Polynukleotide Kinase 3' Phosphatase (PNKP) und *DNA*-Polymerasen wie Pol μ und Pol λ . Die letztendliche Ligierung erfolgt durch einen Komplex aus der Ligase IV und den Co-Faktoren *XRCC4-XLF* (Xue and Greene 2021).

Die Prozessierung der *DNA*-Enden bedingt Insertionen und Deletionen von Nukleotiden (Indels). Damit besitzt *NHEJ* ein mutagenes Potential mit der Gefahr der Disruption von Genen durch Leserasterverschiebungen. Gewünscht ist dies bei *knock-out*-Fragestellungen. So kann, induziert durch *Cas9*, gezielt ein Gen disruptiert werden. Unerwünscht hingegen ist diese fehlerhafte Reparatur an *Off-Target*-Arealen (Stinson et al. 2020). *NHEJ* ist jedoch nicht per se fehlerhaft. Bedingt wird diese Fehlerhaftigkeit durch die Prozessierung der Enden. Direkt kompatible *DNA*-Enden können ohne Prozessierung fehlerfrei ligiert werden (Bétermier et al. 2014). Die von *Cas9* erzeugten DSB sind jedoch heterogen. Dies ist dadurch bedingt, dass *HNH* und *RuvC* ihren *DNA*-Strang nicht zwangsläufig an der gleichen Position schneiden. *HNH* schneidet den *TS* regelrecht an Position 3 distal des *PAM*. *RuvC* hingegen ist sehr viel flexibler und kann den *NTS* an den Positionen 3-5 schneiden. Somit sind neben symmetrischen auch asymmetrische DSB mit 5'-Überhängen möglich (Shou et al. 2018). Tatsächlich scheinen die durch *Cas9* induzierten Mutationen nicht zufällig zu sein. Die durch *NHEJ* und *MMEJ* reparierten DSB zeigen reproduzierbare sequenzabhängige Insertionen oder Deletionen. Von besonderer Bedeutung scheinen dabei die Nukleotide in der Nähe der Schnittstelle zu sein (Allen et al. 2019). Die durch *NHEJ* und *MMEJ* entstehenden Indels sind demnach nicht so willkürlich wie ursprünglich angenommen. Die Vorhersagbarkeit dieser eröffnet die Möglichkeit gezielt *DNA* zu modifizieren, ohne auf ein *DNA-Template* angewiesen zu sein. Klinisch relevant ist dies u.a. bei der Reparatur einzeln mutierter Nukleotide, die durch eine Verschiebung des Leserahmens zu einem Funktionsverlust eines Genes geführt haben (Shen et al. 2018).

1.2.3 Microhomology Mediated End Joining (MMEJ)

MMEJ ist ein alternativer Reparaturmechanismus und wie *NHEJ* ebenfalls unabhängig von einem *DNA-Template*. Hierbei erfolgt die Verbindung der beiden DSB-Enden über das Zusammenlagern mikrohomologer Nukleotidsequenzen, welche sich in unmittelbarer Nähe des DSB befinden. Damit diese einander zugänglich werden, ist jeweils eine Resektion der Enden notwendig. Vermittelt wird diese über die Interaktion

des *MRN*-Komplexes (*MRE11-RAD50-NBS1*) mit *CtIP* (*C-terminal binding protein interacting protein*). *MRE11* induziert einen 5'-ssDNA-Bruch nahe des DSB. Dieser ermöglicht die 5'-3' Exonukleaseaktivität von *MRE11*. Folge sind auf beiden Seiten freiliegende 3'-Überhänge, welche jeweils die mikrohomologen Sequenzen tragen. Die dabei entstehenden Lücken werden von der Polymerase θ gefüllt. Die Verbindung der Enden erfolgt über die Ligase 1 oder 3 (Xue and Greene 2021). Es kommt zur Deletion der den DSB flankierenden Nukleotide, so dass *MMEJ* ebenfalls mutagenes Potential besitzt (Sfeir and Symington 2015). Besagte Deletionen eignen sich zur Therapie von Mikroduplikationserkrankungen. Ein durch *Cas9* induzierter DSB im Zentrum der Duplikatur mit anschließender durch *MMEJ* vermittelter Deletion könnte so einen gesunden Genotyp herstellen (Iyer et al. 2019). Ferner gibt es auch Protokolle co-applizierte Gensequenzen mittels *MMEJ* in das Genom einzufügen. Dies stellt eine Alternative zum *HDR* assistierten *knock-in* dar (Sakuma et al. 2016). *MMEJ* ist v.a. während der S- und G2-Phase des Zellzyklus aktiv (Black et al. 2019). Allerdings ist auch ein Auftreten während der M- und G1-Phase möglich (Taleei and Nikjoo 2013).

1.2.4 Homology Directed Repair (HDR)

HDR nutzt ein homologes *DNA-Template* zur präzisen Reparatur des DSB. Dafür notwendig ist eine ausgeprägte Resektion der DSB-Enden mit Formierung langer 3'-ssDNA-Überhänge. Die initiale Resektion der Enden überlappt mit der von *MMEJ* und wird durch den *MRN-CtIP*-Komplex vermittelt. Bei *HDR* erfolgt jedoch darüber hinaus eine erweiterte Resektion, vermittelt durch *EXO1* und/oder *BLM/DNA2*. Diese Faktoren sind streng innerhalb des Zellzyklus reguliert und auf die S/G2-Phase beschränkt. *RAD51* vermittelt anschließend die Paarung des 3'-Überhanges mit der homologen *dsDNA*, welche dann als *Template* zur *DNA*-Synthese fungiert. Daran beteiligt sind u.a. die Polymerase δ , *PCNA* und *RFC 1-5* (Xue and Greene 2021).

Damit eignet sich *HDR*, um exogen co-applizierte *DNA*-Sequenzen gezielt in ein Gen einzufügen. *HDR* ist nicht so fehlerbehaftet wie *NHEJ* und *MMEJ* und birgt somit das Potential der präzisen Geneditierung. Jedoch sind auch hier im Kontext der *Mismatch-Toleranz* Integrationen des *Templates* in *Off-Target*-Areale beschrieben. Zudem können homologe und non-homologe Reparaturmechanismen um unterschiedliche Enden des gleichen DSB konkurrieren und so die Misintegration des *Templates* begünstigen (Canaj et al. 2019). Dementsprechend gibt es eine Reihe an Bestrebungen, die Effizienz von *HDR* durch Hemmung von *NHEJ* und *MMEJ* zu erhöhen. Diese umfassen u.a. die Suppression von *NHEJ*-Schlüsselfaktoren wie der Ligase IV, die Überexpression von *HDR*-Faktoren wie *RAD51* oder die Beschränkung der Aktivität von *Cas9* auf die G2 und S-Phase. Hier gibt es jedoch einige Sicherheitsbedenken, da *NHEJ* mit der schnell

vermittelten DSB-Reparatur essentiell für den Erhalt der genomischen Stabilität ist und dessen Hemmung demnach fatale Folgen haben könnte (Yang et al. 2020). Dies unterstützend zeigt sich bei einem Mangel an *NHEJ* Faktoren eine Zunahme größerer Deletionen bei *Cas9* vermitteltem *genome editing* (Kosicki et al. 2022).

1.3 Methoden zum Nachweis von *DNA*-Editierungen

1.3.1 *DNA* basierte Systeme

DNA basierte Systeme eignen sich zum Nachweis spezifischer Nukleotidsequenzen innerhalb eines Genoms. Ist der Locus der Editierung bekannt, kann die gewünschte *DNA*-Sequenz durch eine *Polymerase Chain Reaction (PCR)* amplifiziert und anschließend sequenziert werden (Grohmann et al. 2019). Ein Beispiel hierfür ist *Tracking of Indels by DEcomposition (TIDE)*. Hierbei wird ein Bereich von 500-1500 Basenpaaren um die editierte Stelle amplifiziert und danach Sanger-sequenziert. Mit einer bioinformatischen Software können im Anschluss induzierte Sequenzveränderungen im Vergleich zum Referenzgenom identifiziert werden (Brinkman et al. 2014). Nicht geeignet ist die *PCR*-Amplifizierung, wenn der Locus der Editierung nicht bekannt ist. Hier besteht die Möglichkeit der Darstellung des gesamten Genoms durch *whole genome sequencing (WGS)* (Grohmann et al. 2019). Ist nicht bekannt, an welchen Stellen im Genom *CRISPR-Cas* einen DSB induziert, existieren verschiedene Protokolle, diese ohne *WGS* darzustellen. Dazu werden die DSB markiert, selektiv amplifiziert und anschließend sequenziert (Cameron et al. 2017).

1.3.2 Fluoreszenzbasierte Systeme

Eine andere Möglichkeit die Effizienz von Geneditierungen zu analysieren bieten Systeme, die auf dem Nachweis bzw. der Quantifizierung einer Fluoreszenz beruhen (Wettengel et al. 2023). Häufige Anwendung findet dabei das *Green Fluorescent Protein (GFP)*. Die codierenden Sequenzen für dieses können in ein Genom integriert werden und als Marker einer Genexpression dienen. *GFP* zeigt dabei eine stabile grüne Fluoreszenz nach Anregung mit Licht im blauen Wellenlängenbereich (Chalfie et al. 1994). Die codierenden Abschnitte für *GFP* können beispielsweise mit einem *Template* kombiniert werden, das in ein Genom integriert werden soll. Die erfolgreiche Integration kann dann über die Co-Expression von *GFP* nachgewiesen werden (Lin-Shiao et al. 2022). Nach dem gleichen Prinzip kann man sich die Emissionseigenschaften von Luziferasen zu Nutze machen. Im Beisein eines fluorogenen Substrates emittieren diese Licht in einer charakteristischen Wellenlänge (Dragulescu-Andrasi et al. 2011). Die Quantifizierung dieser Emission kann auch zum Nachweis von Geneditierungen genutzt werden, beispielweise durch die Induktion der Expression durch *DNA*-Reparaturprozesse (Yanik et al. 2018).

Abseits dieser einfachen, gibt es noch duale *Reporter*-Systeme, die auf dem Energietransfer zwischen zwei photoaktiven Molekülen beruhen. Dieser Vorgang wird als *Resonance Energy Transfer (RET)* bezeichnet und meint den Energietransfer von einem Donor- auf ein Akzeptormolekül via Dipol-Dipol-Wechselwirkungen. Die Emissionen des Donors überführen den Akzeptor in einen instabilen elektronenreichen Erregungszustand. Das führt wiederum dazu, dass dieser anschließend Licht in einer charakteristischen Wellenlänge emittiert (Kobayashi et al. 2019). Anwendungsformen in der Wissenschaft sind *fluorescence RET (FRET)*, *bioluminescence RET (BRET)* und *chemiluminescence RET (CRET)*. Der Bereich, in dem ein Energietransfer stattfinden kann, liegt bei 10-100 Å. Damit eignen sich diese Methoden zur Erkundung von intrazellulären Proteininteraktionen und Signaltransduktionswegen (Wu and Jiang 2022).

In den Versuchen dieser Arbeit wurde ein *BRET-Reporter* zur Aktivitätsmessung von *gRNA-Cas9*-Komplexen verwendet. Standard *BRET-Reporter* bestehen aus einer Luziferase, die im Beisein eines fluorogenen Substrates als *resonance energy* Donor fungiert sowie einem fluoreszierenden Protein, das als *resonance energy* Akzeptor fungiert (Dragulescu-Andrasi et al. 2011).

1.4 Zielsetzung

Ziel dieser Arbeit war die Erprobung der *Mismatch*-Toleranz von *gRNA-Cas9*-Komplexen mittels eines *BRET-Reporter*-Systems. Dafür wurden verschiedene Sequenzen im *Reporter* mit einfacher oder mehrfacher Sequenzveränderung hergestellt und zusammen mit *gRNA-Cas9* in *HEK 293T*-Zellen exprimiert. Eine Aktivität des *gRNA-Cas9*-Komplexes führte durch Induktion eines Doppelstrangbruches mit anschließender Verschiebung des Leserahmens zu einer eingeschränkten Expression von *GFP 2* und somit zu einer veränderten *BRET-Ratio* (Wimmer et al. 2021). Somit ließ sich die *Mismatch*-Toleranz mittels *BRET-Assay* quantifizieren. Die erhobenen Daten dienen dem Ziel, den Einfluss positions- und nukleotidspezifischer Veränderungen auf die Aktivität von *gRNA-Cas9* zu erheben. Dies ist v.a. bedeutsam in der Ergründung von *Off-Target*-Arealen und hat große Relevanz in der Planung und Sicherheit von *CRISPR-Cas9* im Kontext des *genome editing* (Wienert et al. 2022).

2 Material und Methoden

2.1 Material

2.1.1 Geräte

Alle für die Versuche verwendeten Geräte sind, mit Nennung des Herstellers, in Tabelle 1 alphabetisch gelistet.

Tabelle 1: Verwendete Geräte

Geräte (inkl. Gerätebezeichnung)	Hersteller, Ort
Autoklav (DX65)	Systec, Linden Deutschland
Bunsenbrenner (Fireboy)	Integra Biosciences, Biebertal
Eismaschine (AF80)	Scotsman, Berlin
Elektroporationsküvetten	Eppendorf, Wesseling-Berzdorf BioRad, München
Elektroporationsgerät (Multiporator)	Eppendorf, Wesseling-Berzdorf
Fluoreszenzmikroskop (Biozero)	Keyence, Neu-Isenburg
Gefrierschrank, -20°C (Premium biofresh)	Liebherr, Biberach an der Riß
Gefrierschrank, -80°C (herafreeze HFU 586 Top)	Thermo Fisher Scientific, Bremen
Inkubator (C150)	Binder, Tuttlingen
Kühlschrank, +4°C (KT 3L 17) (TP 1760)	Bosch, Gerlingen Liebherr, Biberach an der Riß
Mikroskop (IT 400)	VWR International, Darmstadt
Multimode Platerreader (Infinite M1000Pro)	Tecan, Männedorf (CH)
Photometriegerät (BioPhotometer)	Eppendorf, Wesseling-Berzdorf
Pipetten (Transferpette)	Brand, Wertheim
Pipetten aus Glas	Hirschmann Laborgeräte, Eberstadt Isolab Laborgeräte, Eschau
Pipettierhilfe (Pipetboy acu 2)	Integra Biosciences, Biebertal
Schüttler	Sartorius AG, Göttingen
Spülmaschine (G7883)	Miele Professional, Gütersloh
Sterilwerkbank für Zellkulturen (MSC-Advantage)	Thermo Fisher Scientific, Bremen

Stickstofftonne (Thermolyne Locator Plus 6)	Thermolyne, Pfullingen
Thermoblock (TB2)	Biometra, Göttingen
Vortex-Geräte (Vortex-Schüttler) (unitexer)	VWR International, Darmstadt LLG Labware, Meckenheim
Waage (Acculab ALC-110.4)	Sartorius AG, Göttingen
Wasseraufreinigungssystem (Arium 611 DI)	Sartorius AG, Göttingen
Wasserbad (TW12)	Julabo, Seelbach
Zellkulturschaber (Zellschaber 16 cm)	Sarstedt, Nümbrecht
Zentrifuge (1-15PK)	Sigma, München
Zentrifuge (4K15)	Sigma, München
Zentrifuge (Micro200)	Hettich Lab Technology, Tuttlingen

2.1.2 Verbrauchsmaterialien

Verbrauchsmaterialien wurden von den Firmen Brandt (Wertheim), Corning (Kaiserslautern), Greiner Bio-one (Frickenhausen) und Sarstedt (Nümbrecht) bezogen.

2.1.3 Chemikalien

In Tabelle 2 finden sich alphabetisch und unter Angabe des Herstellers die verwendeten Chemikalien.

Tabelle 2: Verwendete Chemikalien

Chemikalien	Hersteller, Ort
Accutase	eBioscience, Frankfurt a. M.
Agar	Life Technologies / Thermo Fisher, Darmstadt
Agarose LE	Genaxxon Bioscience, Ulm
Ampicillin	Roth, Karlsruhe
Coelenterazin 400a	Nanolight Inc., Pinetop, AZ, USA
Dinatriumhydrogenphosphat (Na ₂ HPO ₄)	Roth, Karlsruhe
Dulbeco's Modified Eagles Medium (DMEM)	Anprotec, Bruckberg
Ethanol	Roth, Karlsruhe
Fötale Kälberserum (FKS)	PAN-Biotech, Aidenbach

Glukose	Merck, Darmstadt
Glycerin	Merck, Darmstadt
Hefeextrakt	BD Biosciences, Heidelberg
Isopropanol	Roth, Karlsruhe
Kaliumchlorid (KCl)	Roth, Karlsruhe
Kaliumdihydrogenphosphat (KH ₂ PO ₄)	Merck, Darmstadt
L-Glutamin	Anprotec, Bruckberg
Lysogeny Broth Base	Invitrogen, Karlsruhe
Magnesiumchlorid (MgCl ₂)	Merck, Darmstadt
Magnesiumsulfat (MgSO ₄)	Merck, Darmstadt
N-2-Hydroxyethyl-piperazin-N'-2-ethansulfonsäure (HEPES)	Roth, Karlsruhe
Natriumchlorid (NaCl)	Merck, Darmstadt
Penicillin/Streptomycin	Anprotec, Bruckberg
Polyethylenimin (PEI)	Polyscience AG, Cham (CH)
Select Agar	Invitrogen, Karlsruhe
Stickstoff (flüssig)	Linde, München
Tetracyclin	Sigma-Aldrich, Darmstadt
Trypton	BD Biosciences, Heidelberg

2.1.4 Plasmide

2.1.4.1 Plasmid *RLuc8*

Das Plasmid codierend für die Luziferase *RLuc8* basiert auf dem Vektor pcDNA3.1(+). Bezogen wurde dieser von der Firma Thermo Fisher Scientific (Darmstadt). In diesen wurde in der Arbeitsgruppe Experimentelle Ophthalmologie (AG Prof. Stieger) eine nach Loening et al. etablierte Mutante der Luziferase von *Renilla Reniformis* integriert. Diese zeigt im Vergleich zum Wildtyp eine geringere intrazelluläre Inaktivierung sowie verbesserte Emissionseigenschaften und eignet sich somit gut für Luziferase-basierte Assays (Loening et al. 2006). *RLuc8* unterliegt der Kontrolle eines CMV-Promoters. Zudem codiert das Plasmid für eine Ampicillinresistenz, welche zur Selektierung benutzt wurde.

2.1.4.2 Plasmid *pRLuc8-AvrII/BsiWI-GFP2*

Das Plasmid *pRLuc8-AvrII/BsiWI-GFP2* enthält die codierenden Abschnitte für die Luziferase *RLuc8* (Loening et al. 2006), das *Green Fluorescent Protein 2* (*GFP 2*) sowie eine Ampicillinresistenz. Diese stehen unter der Kontrolle eines CMV-Promoters. Ebenfalls vorhanden sind die Schnittstellen der Restriktionsenzyme *AvrII* und *BsiWI* (2.1.5.1). Zwischen diesen erfolgte die Einbettung der *Target-Oligonukleotide* (2.1.6). Das Plasmid wurde in der Arbeitsgruppe Experimentelle Ophthalmologie (AG Prof. Stieger) hergestellt (Wimmer et al. 2021). Die Ampicillinresistenz wurde zur Selektierung genutzt. Abbildung 1 zeigt die zugehörige Plasmidkarte.

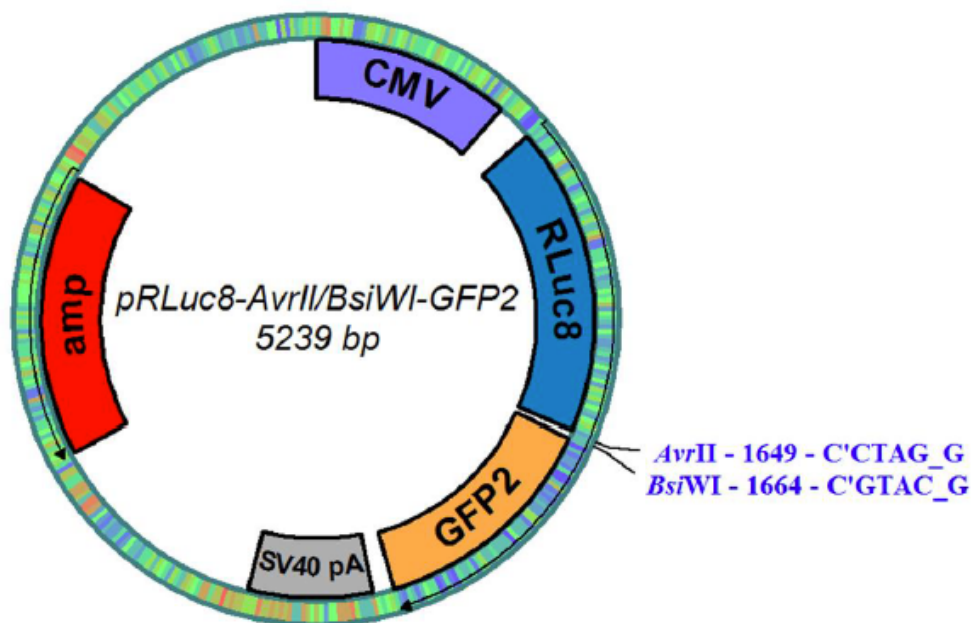


Abbildung 1: Plasmidkarte *pRLuc8-AvrII/BsiWI-GFP2*

Amp: Ampicillinresistenz, CMV: Cytomegalievirus-Promoter, *RLuc8*: Luziferase, *GFP 2*: *Green Fluorescent Protein 2*. Ebenfalls dargestellt sind die Schnittstellen der Restriktionsenzyme *AvrII* und *BsiWI*.

Bild: Dr. Tobias Wimmer, Experimentelle Ophthalmologie Gießen© (AG Prof. Stieger).

2.1.4.3 Plasmid *px459*

Das Plasmid *px459* codiert für den *gRNA-Cas9*-Komplex. Basierendes Plasmid ist *pSpCas9(BB)-2A-Puro (PX459) V2.0* der Arbeitsgruppe von Feng Zhang (Ran et al. 2013b). Bezogen wurde dieses durch die Firma Addgene (Watertown, USA) (Addgene plasmid # 48139 ; <http://n2t.net/addgene:48139> ; RRID:Addgene_48139).

Die codierenden Abschnitte für *Cas9* stehen unter der Kontrolle eines *CBA*-Promoters (*Chicken-Beta-Actin*), die der *gRNA* unter der Kontrolle eines U6-Promoters. Ebenfalls vorhanden sind zwei Schnittstellen des Restriktionsenzym *BbsI*. Zwischen diesen

erfolgte die Einbettung der *gRNA*-Oligonukleotide. Diese waren schon aus Vorarbeiten vorhanden (2.1.12). Abbildung 2 zeigt die zugehörige Plasmidkarte.

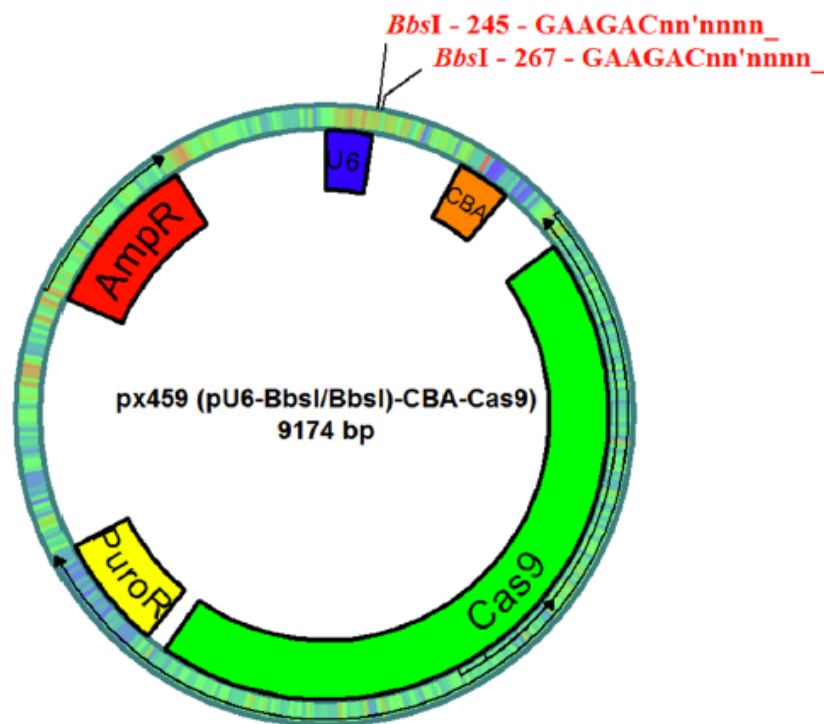


Abbildung 2: Plasmidkarte px459

AmpR: Ampicillinresistenz, U6: U6-Promoter, CBA: *Chicken-Beta-Actin*-Promoter, Cas9: *CRISPR-Associated-Protein 9*, PuroR: Puromycinresistenz. Ebenfalls dargestellt sind die Schnittstellen des Restriktionsenzymys BbsI.

Bild: Dr. Tobias Wimmer, experimentelle Ophthalmologie Gießen© (AG Prof. Stieger).

2.1.5 Enzyme

2.1.5.1 Restriktionsenzyme

In Tabelle 3 gelistet sind die für die Arbeit verwendeten Restriktionsenzyme. Die Schnittstellen sind jeweils durch ↓ und ↑ gekennzeichnet.

Tabelle 3: Verwendete Restriktionsenzyme

Enzym	Restriktionssequenz	Hersteller, Ort
AvrII (5 000 units/ml)	5'----C↓CTAGG----3' 3'----GGATC↑C----5'	New England Biolabs, Frankfurt a. M
BsiWI-HF (20 000 units/ml)	5'----C↓GTACG----3' 3'----GCATG↑C----5'	New England Biolabs, Frankfurt a. M.

2.1.5.2 Weitere Enzyme

Weitere verwendete Enzyme finden sich in Tabelle 4.

Tabelle 4: Weitere Enzyme

Enzyme	Hersteller, Ort
T4 DNA Ligase	New England Biolabs, Frankfurt a. M.

2.1.6 Oligonukleotide

Folgende Oligonukleotide wurden für die Klonierungen verwendet. Diese wurden von der Firma Metabion (Planegg) bezogen. Ein *Target*-Oligonukleotid ist einzelsträngig und besteht aus einer 20 Nukleotide langen *Target*-Sequenz, welche sich an das *PAM* anschließt, sowie den spezifischen Überhängen zur Integration in das Plasmid *pRLuc8-AvrII/BsiWI-GFP2* (2.1.4.2). Dabei gab es zu jedem Wildtyp bzw. *Mismatch* ein *Forward*- und *Reverse*-Oligonukleotid, welche gegenläufig zueinander komplementär waren. Die Hybridisierung dieser generierte einen doppelsträngigen Oligonukleotid-Duplex (2.2.1.2), welcher dann mit dem entsprechenden Plasmid ligiert wurde (2.2.1.3).

Abbildung 3 zeigt exemplarisch für alle *Mismatch*- und Wildtypsequenzen den Heteroduplex von ART t1, welcher aus der Hybridisierung von *Forward*- und *Reverse*-Oligonukleotid entstanden ist (2.2.1.2). Den einzelnen Nukleotiden wurde dabei eine Position (P) zugewiesen. Diese bezieht sich auf das *PAM* (5'-NGG-3') im *Forward* Oligonukleotid. Das erste sich hieran in 3'-5'-Orientierung anschließende Nukleotid hat die Position 1. Folgende Nukleotide wurden fortlaufend bis 20 nummeriert.

P	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	P	A	M	
5'	G	G	A	G	T	C	C	G	C	A	C	C	A	C	G	T	C	G	T	G	T	G	G	3'
3'	C	C	T	C	A	G	G	C	G	T	G	G	T	G	C	A	G	C	A	C	A	C	C	5'

Abbildung 3: Übersichtsschema eines *Target*-Heteroduplexes

P: Position des Nukleotids in Bezug zum *protospacer adjacent motif* (*PAM*). Obere Zeile: *Forward*-Oligonukleotid in 5'-3'-Orientierung. Untere Zeile: *Reverse*-Oligonukleotid in 3'-5'-Orientierung. Nicht dargestellt sind die spezifischen Nukleotid-Überhänge des *Target*-Heteroduplexes zur späteren Integration in das Plasmid *pRLuc8-AvrII/BsiWI-GFP2* (2.2.1.3).

Abseits der *Target*-Oligonukleotide gab es noch *gRNA*-Oligonukleotide. Diese bestanden aus einer 20 Nukleotide langen *gRNA*-Sequenz sowie den spezifischen Überhängen zur Integration in das Plasmid *px459* (2.1.4.3). Diese waren bereits aus Vorarbeiten vorhanden (2.1.12).

2.1.6.1 Wildtypsequenzen

In Tabelle 5 sind die *Target*-Oligonukleotide der Wildtypsequenzen zu finden. Aus diesen wurden durch Veränderungen einzelner Nukleotide die *Mismatch*-Oligonukleotide generiert (2.1.6.2). Das *Forward*-Oligonukleotid (*F*) ist in der oberen, das *Reverse*-Oligonukleotid (*R*) in der unteren Zeile der jeweiligen Sequenz zu finden. Beide sind in der Orientierung 5'-3' aufgeführt. Unter den *Target*-Oligonukleotiden (*t*) der verschiedenen Wildtypen finden sich jeweils die zugehörige *gRNA*-Oligonukleotide (*g*). Das *PAM* der *Target*-Oligonukleotide ist fettgedruckt hervorgehoben.

Tabelle 5: Wildtypsequenzen der *Targets* und zugehörige *gRNA*-Sequenzen

Bezeichnung	F/R	Sequenz (5'-3'-Orientierung)
ART t1	F	CTAGGGGAGTCCGCACCACGTCGTGT GG
	R	GTAC CC ACACGACGTGGTGCGGACTCCC
ART g1	F	CACCGGAGTCCGCACCACGTCGTG
	R	AAACCACGACGTGGTGCGGACTCC
T4 #4461-4462	F	CTAGGGCGAGTCAGGGGGCTCTTGT CGG
	R	GTAC CCG ACAAGAGCCCCCTGACTCGCC
G4 #4463-4464	F	CACCGCGAGTCAGGGGGCTCTTGT
	R	AAACACAAGAGCCCCCTGACTCGC
Rpgr t11	F	CTAGGCAGAGTTTCGCTGAGTGGCTT GG
	R	GTAC CCA AGCCACTCAGCGAAACTCTGC
Rpgr g3	F	CACCCAGAGTTTCGCTGAGTGGCT
	R	AAACAGCCACTCAGCGAAACTCTG
ART t5	F	CTAGGCAGTACCTTTGGGTGGCGGC AGG
	R	GTAC CT GCCGCCACCCAAAGGTA CTGC
ART g5	F	CACCCAGTACCTTTGGGTGGCGGC
	R	AAACCAAACCGACAAGTGCTCTA

2.1.6.2 Veränderte Wildtypsequenzen

In Tabelle 6 abgebildet sind die als *Mismatches* bezeichneten veränderten Wildtypsequenzen. Im Vergleich zur Wildtypsequenz veränderte Nukleotide wurden im *Forward*- (*F*) und *Reverse*- (*R*) Oligonukleotid gekennzeichnet. Das *PAM* ist fettgedruckt hervorgehoben. In Klammern aufgeführt ist die jeweils dem *Mismatch* zugehörige Wildtypsequenz. Die letzte Spalte zeigt zudem, welches Nukleotid an welcher Position (*P*) des späteren Heteroduplexes (2.1.6) verändert wurde, um das *Mismatch* zu generieren. Dies wird durch einen Pfeil symbolisiert.

Tabelle 6: Oligonukleotide der *Mismatch*-Sequenzen

Bezeichnung	F/R	Sequenz (5'-3'-Orientierung)	Veränderung
M 1.1 (ART t1)	F	CTAGGGGAGTCCGCACCACGTCGT <u>C</u> TGG	P 1, G→C
	R	GTAC <u>CC</u> AGACGACGTGGTGCGGACTCCC	P 1, C→G
M 1.2 (ART t1)	F	CTAGGGG <u>C</u> GTCCGCACCACGTCGTGTGG	P 18, A→C
	R	GTAC <u>CC</u> ACACGACGTGGTGCGGAC <u>G</u> CCC	P 18, T→G
M 2 (T4 #4461-4462)	F	CTAGGGCGAGTCAGGGGGCT <u>G</u> TTGT <u>C</u> GG	P 5, C→G
	R	GTAC <u>CC</u> GACAACAGCCCCCTGACTCGCC	P 5, G→C
M 3 (T11)	F	CTAGGCAGAGTTTTCGCTGAG <u>G</u> TGCTTGG	P 5, T→G
	R	GTAC <u>CC</u> AAGCACCTCAGCGAAACTCTGC	P 5, A→C
M 4 (ART t5)	F	CTAGGCAGTACCTTTGGG <u>C</u> GGCGGCAGG	P 7, T→C
	R	GTAC <u>CC</u> TGCCGCC <u>G</u> CCCAAAGGTAAGTGC	P 7, A→G
M 5 (ART t1)	F	CTAGG <u>C</u> GAGTCCGCACCACGTCGTGTGG	P 20, G→C
	R	GTAC <u>CC</u> ACACGACGTGGTGCGGACTC <u>G</u> C	P 20, C→G
M 6 (ART t1)	F	CTAGGGGAG <u>G</u> CCGCACCACGTCGTGTGG	P 16, T→G
	R	GTAC <u>CC</u> ACACGACGTGGTGCGG <u>C</u> CTCCC	P 16, A→C
M 7 (ART t1)	F	CTAGGGGAGTC <u>A</u> GCACCACGTCGTGTGG	P 14, C→A
	R	GTAC <u>CC</u> ACACGACGTGGTGC <u>T</u> GACTCCC	P 14, G→T
M 8 (ART t1)	F	CTAGGGGAGTCCG <u>T</u> ACCACGTCGTGTGG	P 12, C→T
	R	GTAC <u>CC</u> ACACGACGTGGT <u>A</u> CGGACTCCC	P 12, G→A
M 9 (ART t1)	F	CTAGGGGAGTCCGCAG <u>C</u> ACGTCGTGTGG	P 10, C→G
	R	GTAC <u>CC</u> ACACGACGTG <u>C</u> TGCGGACTCCC	P 10, G→C
M 10 (ART t1)	F	CTAGGGGAGTCCGCAC <u>C</u> TGTCGTGTGG	P 8, A→T
	R	GTAC <u>CC</u> ACACGACG <u>A</u> GGTGCGGACTCCC	P 8, T→A
M 11 (ART t1)	F	CTAGGGGAGTCCGCACCAC <u>A</u> TCGTGTGG	P 6, G→A
	R	GTAC <u>CC</u> ACACG <u>A</u> IGTGGTGCGGACTCCC	P 6, C→T
M 12 (ART t1)	F	CTAGGGG <u>G</u> GTCCGCACCACGTCGTGTGG	P 18, A→G
	R	GTAC <u>CC</u> ACACGACGTGGTGCGGAC <u>C</u> CCC	P 18, T→C
M 13 (ART t1)	F	CTAGGGG <u>I</u> GTCCGCACCACGTCGTGTGG	P 18, A→T
	R	GTAC <u>CC</u> ACACGACGTGGTGCGGAC <u>A</u> CCC	P 18, T→A
M 14 (ART t1)	F	CTAGGGG <u>C</u> GT <u>A</u> CGCACCACGTCGTGTGG	P 18, A→C P 15, C→A
	R	GTAC <u>CC</u> ACACGACGTGGTGC <u>G</u> T <u>A</u> CG <u>C</u> CCC	P 18, T→G P 15, G→T
M 15 (ART t5)	F	CTAGG <u>C</u> GTACCTTTGGGTGGCGGCAGG	P 19, A→C
	R	GTAC <u>CC</u> TGCCGCC <u>G</u> CCCAAAGGTAC <u>G</u> C	P 19, T→G

M 16 (ART t1)	F	CTAGG <u>A</u> GAGTCCGCACCACGTCGTGT GG	P 20, G→A
	R	GTAC CC ACACGACGTGGTGC GG ACTC <u>I</u> C	P 20, C→T
M 17 (ART t1)	F	CTAGGGGAG <u>C</u> CCGCACCACGTCGTGT GG	P 16, T→C
	R	GTAC CC ACACGACGTGGTGC GGG CTCCC	P 16, A→G
M 18 (ART t1)	F	CTAGGGGAG <u>A</u> CCGCACCACGTCGTGT GG	P 16, T→A
	R	GTAC CC ACACGACGTGGTGC GG <u>I</u> CTCCC	P 16, A→T
M 19 (ART t1)	F	CTAGGGGAGT <u>C</u> IGCACCACGTCGTGT GG	P 14, C→T
	R	GTAC CC ACACGACGTGGTGC <u>A</u> GA CT CCC	P 14, G→A
M 20 (ART t1)	F	CTAGGGGAGT <u>C</u> GGCACCACGTCGTGT GG	P 14, C→G
	R	GTAC CC ACACGACGTGGTGC <u>C</u> GA CT CCC	P 14, G→C
M 21 (ART t1)	F	CTAGGGGAGTCCG <u>A</u> ACCACGTCGTGT GG	P 12, C→A
	R	GTAC CC ACACGACGTGGT <u>I</u> CGGA CT CCC	P 12, G→T
M 22 (ART t1)	F	CTAGGGGAGTCCG <u>G</u> ACCACGTCGTGT GG	P 12, C→G
	R	GTAC CC ACACGACGTGGT <u>C</u> CGGA CT CCC	P 12, G→C
M 23 (ART t1)	F	CTAGGGGAGTCCG <u>C</u> AACACGTCGTGT GG	P 10, C→A
	R	GTAC CC ACACGACGTG <u>I</u> TGCGGA CT CCC	P 10, G→T
M 24 (ART t1)	F	CTAGGGGAGTCCG <u>C</u> A <u>I</u> CACGTCGTGT GG	P 10 C→T
	R	GTAC CC ACACGACGTG <u>A</u> TGCGGA CT CCC	P 10, G→A
M 25 (ART t1)	F	CTAGGGGAGTCCGCACC <u>C</u> CGTCGTGT GG	P 8, A→G
	R	GTAC CC ACACGACG <u>C</u> GGTGC GG ACTCCC	P 8, T→C
M 26 (ART t1)	F	CTAGGGGAGTCCGCACC <u>C</u> CGTCGTGT GG	P 8, A→C
	R	GTAC CC ACACGACG <u>G</u> GGTGC GG ACTCCC	P 8, T→G
M 27 (ART t1)	F	CTAGGGGAGTCCGCACCAC <u>I</u> TCGTGT GG	P 6, G→T
	R	GTAC CC ACACG <u>A</u> AGTGGTGC GG ACTCCC	P 6, C→A
M 28 (ART t1)	F	CTAGGGGAGTCCGCACCAC <u>C</u> TCGTGT GG	P 6, G→C
	R	GTAC CC ACACGAG <u>G</u> TGGTGC GG ACTCCC	P 6, C→G
M 29 (ART t1)	F	CTAGGGGAGT <u>A</u> CGCACCACGTCGTGT GG	P 15, C→A
	R	GTAC CC ACACGACGTGGTGC <u>G</u> I ACT CCC	P 15, G→T
M 30 (ART t1)	F	CTAGG <u>C</u> GAG <u>G</u> CCGCACCACGTCGTGT GG	P 20, G→C P 16, T→G
	R	GTAC CC ACACGACGTGGTGC GG <u>C</u> CT <u>C</u> G <u>C</u>	P 20, C→G P 16, A→C
M 31 (ART t1)	F	CTAGG <u>C</u> GAGT <u>C</u> AGCACCACGTCGTGT GG	P 20, G→C P 14, C→A
	R	GTAC CC ACACGACGTGGTGC <u>I</u> GA CT <u>C</u> G <u>C</u>	P 20, C→G P 14, G→T
M 32 (ART t1)	F	CTAGGGGAGTCCG <u>I</u> ACCAC <u>A</u> TCGTGT GG	P 12, C→T P 6, G→A

	<i>R</i>	GTAC CCACAC GAT <u>GTGGT</u> ACGGACTCCC	P 12, G→A P 6, C→T
M 33 (ART t1)	<i>F</i>	CTAGGGG <u>GTCCGC</u> ACCACATCGT GTGG	P 6, G→A P 18, A→T
	<i>R</i>	GTAC CCACAC GAT <u>GTGGT</u> GCGGAC <u>ACCC</u>	P 6, C→T P 18, T→A
M 34 (ART t1)	<i>F</i>	CTAGGGG <u>GTCCG</u> TACCACGTCGT GTGG	P 12, C→T P 18, A→T
	<i>R</i>	GTAC CCACAC GACGTGGT <u>ACGGAC</u> ACCC	P 12, G→A P 18, T→A
M 35 (ART t1)	<i>F</i>	CTAGG <u>CGG</u> GTCCGCACCACGTCGT GTGG	P 20, G→C P 18, A→G
	<i>R</i>	GTAC CCACAC GACGTGGTGC GGAC <u>CCGC</u>	P 20, C→G P 18, T→C

2.1.6.3 Sequenzierungsprimer

Der zur Sequenzierung genutzte Primer ist in Tabelle 7 zu finden.

Tabelle 7: Zur Sequenzierung genutzter Primer

Primer	Sequenz (5'-3'-Orientierung)
#4041SeqKas	AGCCCGACGTCGTCCAGATTG

2.1.7 Organismen

2.1.7.1 Bakterienstämme

Zur Klonierung wurden die in Tabelle 8 gelisteten *E.coli*-Stämme verwendet. Diese wurden zuvor elektrokompent gemacht (2.2.2.1).

Tabelle 8: Verwendete Bakterienstämme

Bakterienstamm	Genetische Marker	Hersteller, Ort
XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB lacIqZΔM15 Tn10 (Tetr)]	Takara Clontech, Saint Germain en laye (FR)
Stellar	F-, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ (lacZYA - argF) U169, Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ-	Stratagene, Heidelberg

2.1.7.2 Eukaryotische Zellen:

Die Plasmidexpression erfolgte in *Human Embryonic Kidney cells* (HEK 293T CRL-3216™) der Firma ATCC (American Type Culture Collection).

2.1.8 Antibiotika für Bakterienkulturen

In Tabelle 9 abgebildet sind die zur Selektierung benutzten Antibiotika und zugehörige Konzentrationen.

Tabelle 9: Verwendete Antibiotika

Antibiotikum	Stockkonzentration	Endkonzentration	Hersteller, Ort
Ampicillin	100 mg/ml	100 µg/ml	Roth, Karlsruhe
Tetracyclin	12,5 µg/µl	12,5 µg/ml	Sigma-Aldrich, Darmstadt

2.1.9 Kulturmedien

2.1.9.1 Kulturmedien für Bakterien

In Tabelle 10 sind die für Bakterienkulturen verwendeten Medien aufgelistet. Diese wurden nach dem Ansetzen und vor der Verwendung autoklaviert.

Tabelle 10: Kulturmedien für Bakterien

Kulturmedium	Inhaltsstoffe	Hersteller, Ort
Lysogeny Broth-Medium (LB-Medium)	25 g Lysogeny Broth Base ad 1000 ml ddH ₂ O	Invitrogen, Karlsruhe
LB-Agar	25 g Lysogeny Broth Base 7,5 g Select Agar ad 500 ml ddH ₂ O	Invitrogen, Karlsruhe Invitrogen, Karlsruhe
SOC-Medium:	20 mM Glukose 10 mM MgSO ₄ 10 mM MgCl ₂ 10 mM NaCl 2,5 mM KCl 2 % Trypton 0,5 % Hefeextrakt	Merck, Darmstadt Merck, Darmstadt Merck, Darmstadt Merck, Darmstadt Roth, Karlsruhe BD Biosciences, Heidelberg BD Biosciences, Heidelberg

2.1.9.2 Kulturmedien für eukaryotische Zellen

In Tabelle 11 sind die verwendeten Medien für die Kulturen von eukaryotischen Zellen gelistet.

Tabelle 11: Kulturmedien für eukaryotische Zellen

Kulturmedium	Inhaltsstoffe	Hersteller, Ort
DMEM +++ (Dulbecco's Modified Eagles Medium)	DMEM 1 % Pen/Strep (v/v) 4 mM L-Glutamin 10 % FKS (v/v)	Anprotec, Bruckberg Anprotec, Bruckberg Anprotec, Bruckberg PAN-Biotech, Aidenbach

2.1.10 Puffer

In Tabelle 12 sind verwendete Puffer gelistet.

Tabelle 12: Verwendete Puffer

Puffer	Inhaltsstoffe	Hersteller, Ort
1x PBS	137 mM NaCl 10 mM Na ₂ HPO ₄ 2,7 mM KCl 1,8 mM KH ₂ PO ₄	Merck, Darmstadt Roth, Karlsruhe Roth, Karlsruhe Merck, Darmstadt

In Tabelle 13 sind kommerziell erworbene Puffer zu finden.

Tabelle 13: Kommerziell erhältliche Puffer

Puffer	Hersteller, Ort
10x CutSmart-Puffer	New England Biosciences, Frankfurt am Main
10x Ligase Puffer	New England Biosciences, Frankfurt am Main
S1-Resuspensions-Puffer	Macherey-Nagel, Düren
S2-Lyse-Puffer	Macherey-Nagel, Düren
S3-Neutralizations-Puffer	Macherey-Nagel, Düren

2.1.11 Kit-Systeme

In Tabelle 14 sind die verwendeten Kit-Systeme gelistet. Diese wurden nach Herstellerangaben verwendet.

Tabelle 14: Verwendete Kit-Systeme

Kit-Systeme	Hersteller, Ort
NucleoSpin® Plasmid QuickPure	Macherey-Nagel, Düren
Maxi Plasmid Kit	Qiagen, Hilden

2.1.12 Ausgangsmaterialien

Zu Beginn der Laborarbeiten bereits vorhanden waren die in 2.1.4 aufgeführten Plasmide. Die in 2.1.6.1 dargestellten Wildtyp-Oligonukleotide waren dabei bereits mit dem Plasmid *pRLuc8-AvrII/BsiWI-GFP2* ligiert. Das Schema der Ligation ist 2.2.1.3 zu entnehmen. Ebenfalls vorhanden waren die den Wildtypen zugehörigen *gRNA*-Sequenzen (2.1.6.1). Hierbei jeweils integriert in das Plasmid *px459* (2.1.4.3).

2.2 Methoden

2.2.1 Arbeiten mit DNA

2.2.1.1 Verdau mittels Restriktionsenzymen

Hierfür wurden die in 2.1.5.1 aufgeführten Enzyme nach Herstellerangaben verwendet. Dabei wurden 1000 ng des Plasmides *pRLuc8-AvrII/BsiWI-GFP2* (2.1.4.2) hydrolysiert. Die Inkubation erfolgte im Wasserbad bei 37 °C für mindestens 30 Minuten. Die anschließende Inaktivierung der Enzyme erfolgte für 20 Minuten im Heizblock bei 65 °C. In Tabelle 15 ist die Zusammensetzung des Restriktionsverdau zu finden.

Durch die Restriktionsenzyme entstanden spezifische einzelsträngige Überhänge des linearisierten Plasmides. Diese waren komplementär zu den Überhängen der Oligonukleotid-Duplexe (2.1.6). So konnten letztere jeweils mit dem hydrolysierten Plasmid ligiert werden (2.2.1.3).

Tabelle 15: Zusammensetzung des Restriktionsverdau

Inhalt	Mengen- oder Volumenangabe
Plasmid <i>pRLuc8-AvrII/BsiWI-GFP2</i>	1000 ng
BsiWI-HF (20 000 units/ml)	1 µl
Avr II (5 000 units/ml)	1 µl
Puffer	2 µl
ddH ₂ O	ad 20 µl

2.2.1.1.1 Nachverdau des Ligationsansatzes

Im Anschluss an die durchgeführte Ligation des jeweiligen Heteroduplexes mit dem Plasmid (2.2.1.3) erfolgte ein Nachverdau. Grund hierfür war, dass im Restriktionsverdau nicht alle Plasmide von den Restriktionsenzymen geschnitten wurden. Somit fanden sich im Ligationsansatz noch vereinzelt nicht mit den Oligonukleotidduplexen ligierte Plasmide. Gewählt wurde ein Restriktionsenzym, dessen Bindungssequenz nach erfolgreicher Ligation mit dem Oligonukleotidduplex nicht mehr vorhanden war und somit selektiv unligierte Plasmide geschnitten wurden. Dafür wurde der in Tabelle 16 dargestellte Ansatz zunächst für 1 h im Wasserbad inkubiert und anschließend für 20 Minuten bei 65 °C inaktiviert.

Tabelle 16: Nachverdau des Ligationsansatzes

Inhalt	Volumenangabe
Ligationsansatz (2.2.1.3)	20 µl
Restriktionsenzym BsiWI-HF (20 000 units/ml)	1 µl

2.2.1.2 Hybridisierung von Oligonukleotiden

Zur Generierung von Oligonukleotidduplexen wurden jeweils zwei zueinander komplementäre Oligonukleotide hybridisiert. Die einzelnen Oligonukleotide sind 2.1.6 zu entnehmen. Es wurde nach dem in Tabelle 17 dargestellten Schema vorgegangen. Die Oligonukleotide wurden 10 Minuten bei 98 °C denaturiert und anschließend 30 Minuten bei Raumtemperatur inkubiert, so dass sich die komplementären Oligonukleotide verbinden konnten. Zur späteren Weiterverwendung erfolgte eine Lagerung des fertigen Ansatzes im Kühlschrank bei - 20 °C.

Tabelle 17: Zusammensetzung der Oligonukleotidhybridisierung

Inhalt	Mengenangabe
Oligonukleotid (<i>forward</i>)	8 µl
Oligonukleotid (<i>reverse</i>)	8 µl
ddH ₂ O	4 µl
Gesamtvolumen	20 µl

2.2.1.3 Ligation von DNA

Hierbei wurden 100 ng des nach 2.2.1.1 hydrolysierten Plasmids mit dem jeweiligen Oligonukleotidduplex ligiert. Das molare Verhältnis vom hydrolysierten Plasmid zum Oligonukleotidduplex betrug 8:1. Die für die Ligation verwendete T4-Ligase (2.1.5.2) wurde nach Herstellerangaben verwendet. Der Ligationsansatz ist Tabelle 18 zu entnehmen. Anschließend wurde der Ansatz über Nacht bei 4 °C inkubiert.

Tabelle 18: Ansatz der Ligation

Inhalt	Mengenangabe
Hydrolysiertes Plasmid	100 ng
Hybridisierter Oligonukleotidduplex	5 µl
Ligase Puffer	2 µl
T4-Ligase	1 µl
ddH ₂ O	ad 20 µl

2.2.1.4 Konzentrationsbestimmung von *DNA* via Photometrie

Zur Konzentrationsbestimmung wurde die isolierte und aufgereinigte *DNA* mit ddH₂O verdünnt. Begonnen wurde dabei mit einer Verdünnung im Verhältnis 1:99. Anschließend erfolgte eine photometrische Konzentrationsbestimmung nach dem Lambert-Beer'schen Gesetz. Es wurde darauf geachtet, dass der gemessene Wert im linearen Bereich liegt (Beer 1852). War dies nicht der Fall, wurde eine andere Verdünnung gewählt und die Messung wiederholt.

2.2.1.5 Sequenzierung von Plasmidklonen

Die Sequenzierungen erfolgten durch die externe Firma Microsynth SeqLab. Dieser wurde der in Tabelle 19 dargestellte Ansatz zugesandt. Die Auswertung der Sequenzierungsergebnisse erfolgte mit dem Auswertungsprogramm FinchTV.

Tabelle 19: Ansatz der externen Sequenzierung

Inhalt	Mengen- oder Volumenangabe
Plasmid <i>DNA</i>	1200 ng
Primer (10 pmol/μl)	3 μl
ddH ₂ O	ad 15 μl

2.2.2 Arbeiten mit Bakterien

2.2.2.1 Herstellung von elektrokompenten *E.coli*-Bakterien

Die Herstellung von elektrokompenten *E.coli*-Bakterien erfolgte aus einem bereits vorhandenen Glycerolstock. Aus diesem wurden 5 μl Bakterien-Glycerin-Gemisch mit 5 ml LB Medium und 5 μl Tetracyclin (Endkonzentration: 12,5 μg/ml) angeimpft. Diese Kultur wurde anschließend auf dem Schüttler bei 37 °C und 180 rpm für ca. 16-18 h inkubiert. Besagte Kultur wurde dann in eine größere Kultur überführt. Hierfür wurden 200 ml LB Medium, 5 ml der oben genannten Kultur und 200 μl Tetracyclin (Endkonzentration: 12,5 μg/ml) verwendet. Im Anschluss erfolgte eine weitere Inkubation bei 37 °C und 180 rpm bis eine optische Dichte (OD_{600nm}) von 0,6-0,8 erreicht wurde. Die Bakteriensuspension wurde in 50 ml *Falcons* überführt und für 30 Minuten auf Eis gestellt. Danach erfolgte eine Zentrifugation für 10 Minuten bei 5100 x g und 4 °C. Der Überstand wurde verworfen und das *Pellet* in 50 ml kalten 1 mM HEPES resuspendiert. Nach erneuter Zentrifugation wurde der zuletzt genannte Schritt mit 25 ml 1 mM HEPES wiederholt. Anschließend erfolgte eine weitere Zentrifugation mit Verwurf des Überstandes. Nun wurde das *Pellet* in 10 ml kalten 10 % Glycerin resuspendiert. Nach erneuter Zentrifugation wurde dieser Schritt zweimal mit jeweils 5 und 0,5 ml 10 %

Glycerin wiederholt. Die nun elektrokompetenten Zellen wurden in 50 µl Eppendorfgefäße aliquotiert und bei - 80 °C bis zur Transformation konserviert.

2.2.2.2 Transformation von elektrokompetenten *E.coli*-Bakterien

Zur Transformation wurden elektrokompetente Zellen auf Eis aufgetaut. Zu 50 µl elektrokompetenten *E.coli*-Bakterien (2.2.2.1) wurden 2 µl des nachverdauten Ligationsansatz (2.2.1.1.1) zugefügt. Nach 5-minütiger Inkubation wurde das Gemisch in eine eiskalte Elektroporationsküvette überführt. Dabei wurde insbesondere darauf geachtet, dass sich beim Überführen keine Luftblasen bildeten, um einen Kurzschluss zu vermeiden. Anschließend wurde an die Elektroporationsküvette für 5 ms eine Spannung von 2000 V angelegt. Nach erfolgter Elektroporation wurde der Inhalt der Küvette zügig in 1000 µl SOC-Medium aufgezogen. Nun erfolgte eine Inkubation auf dem Schüttler bei 37 °C und 180 rpm für mindestens eine Stunde, zur Regeneration und Expression der Ampicillinresistenz. Nach der Inkubation erfolgte eine Zentrifugation bei 11000 x g für 2,5 Minuten. Der Überstand wurde bis auf einen Rest von 200-250 µl verworfen. Mit diesem wurden die abzentrifugierten Bakterien vermischt und anschließend 200 µl davon aufgezogen. Diese wurden auf eine auf 37 °C vorgewärmte LB-Agar-Platte ausplattiert, welche mit Ampicillin versetzt war (2.2.2.3), um ein selektives Wachstum der transformierten *E.coli*-Bakterien zu gewährleisten. Zum Schluss erfolgte eine Inkubation für 16-18 h bei 37 °C. Eine längere Inkubation wurde vermieden, um das Auftreten von Satellitenkolonien zu reduzieren.

2.2.2.3 Herstellung von mit Ampicillin versetzten LB-Agar-Platten:

Hierfür wurden in eine leere, autoklavierte 1000 ml Flasche 25 g Lysogeny Broth Base und 7,5 g Select Agar gegeben (2.1.9.1). Anschließend wurde diese bis zur 500 ml Marke mit ddH₂O aufgefüllt und das Gemisch autoklaviert. Nach der Autoklavierung erfolgte eine einstündige Abkühlung im Wasserbad bei konstant 47°C. Erst danach wurde das Gemisch mit 500 µl Ampicillin (2.1.8) versetzt (Stockkonzentration: 100 mg/ml) und durchmischt. Das noch flüssige Gemisch wurde dann in Petrischalen gegossen. In diesen erfolgte die Abkühlung auf Raumtemperatur und damit die Aushärtung des Gemisches.

2.2.2.4 Beimpfung flüssiger LB-Kulturen

Die auf der Agarplatte gewachsenen Bakterienklone wurden zur weiteren Aufzucht in ein Flüssigmedium überführt. Dafür wurden in 50 ml *Falcons* jeweils 5 ml LB Medium und Ampicillin gegeben. Die Endkonzentration des Ampicillins im Flüssigmediums betrug 100 µg/ml. Anschließend wurden mit sterilen Pipettenspitzen Klone der jeweiligen

LB-Agar-Platte gepickt und diese in die 50 ml *Falcons* gegeben. Danach erfolgte eine Inkubation auf dem Schüttler bei 37 °C und 180 rpm für 16-18 h.

2.2.2.5 Herstellung von Glycerolstocks

Zur längeren Aufbewahrung der Bakterien, welche die erfolgreich klonierte Sequenz trugen, wurden diese im Verhältnis 2:1 mit Glycerol versetzt. Dafür wurden 200 µl der nach 2.2.2.4 beimpften Flüssigkulturen mit 100 µl Glycerin zusammengeführt und zu einer Phase vermischt. Dieses Gemisch wurde anschließend bei - 80 °C gelagert.

2.2.2.6 Isolation von Plasmid DNA

Zur Herstellung von hochkonzentrierter und hochreiner *DNA* wurde das *Maxi Plasmid Kit* der Firma Qiagen (2.1.11) nach Herstellerangaben verwendet. *DNA* mit dieser Aufreinigung wurde für die Transfektionen genutzt (2.2.3.2). Ferner erfolgten zu Sequenzierungszecken (2.2.1.5) *DNA*-Aufreinigungen geringerer Reinheit mit und ohne Ladesäulen. Hierfür wurde das *NucleoSpin Plasmid QuickPure Kit* der Firma Macherey und Nagel nach Herstellerangaben verwendet (2.1.11).

2.2.3 Arbeiten mit eukaryotischen Zellen

2.2.3.1 Kultivierung von HEK 293T-Zellen

Die Kultivierung der Zellen erfolgte in einem Brutschrank mit konstanten Bedingungen. Diese waren 37 °C, 5 % CO₂ und mindestens 95 % Luftfeuchtigkeit. Die erworbene Zellreihe wurde bis zur Verwendung in flüssigem Stickstoff gelagert. Zur Kultivierung wurde der gesamte Inhalt von 1 ml mit 10 ml DMEM +++ aufgezogen und in eine mit Kollagen beschichtete Petrischale überführt. Eine Passagierung der Zellen erfolgte alle 3-4 Tage. Dazu wurde das verbrauchte Medium abgenommen und die adhärennten Zellen mit 1 ml Accutase vom Petrischalenboden gelöst. Nach ca. 5-minütiger Inkubation bei 37 °C erfolgte eine Resuspension in 9 ml DMEM +++, um ein Gesamtvolumen von 10 ml zu erreichen. Aus diesem Gesamtvolumen wurden je nach Bedarf neue Zellkulturschalen mit je 1 ml Suspension gefüllt. Diese wurden anschließend noch mit 9 ml DMEM +++ aufgefüllt und in den obigen Brutschrank überführt.

2.2.3.2 Transfektion von HEK 293T-Zellen

Zur Transfektion der *HEK 293T*-Zellen wurde PEI (Polyethylenimin) verwendet. PEI ermöglicht die Komplexierung von *DNA* in positiv geladene Partikel. Damit können diese an die anionische Zellmembran binden und anschließend endozytiert werden (Longo et al. 2013). Zunächst wurden die *HEK 293T*-Zellen in 6-Well-Kulturplatten überführt. Dabei wurden 750 µl Zellsuspension mit 1500 µl DMEM +++ in ein *Well* gegeben. Anschließend erfolgte eine Inkubation im Brutschrank für 12-18 h bei 37 °C, 5 % CO₂ und mindestens 95 % Luftfeuchtigkeit. Vor Hinzugabe des Tabelle 20 zu entnehmenden

Transfektionsansatzes wurde das verbrauchte Medium abgenommen und mit 2000 µl frischen DMEM +++ ersetzt. Nach 10-minütiger Inkubation des Transfektionsansatzes bei Raumtemperatur wurde dieser tröpfchenweise in das entsprechende *Well* pipettiert und die Zellen in den Brutschrank (37 °C, 5 % CO₂ und mindestens 95 % Luftfeuchtigkeit) überführt. Um zytotoxische Effekte von PEI zu minimieren, wurde nach 4-6 h Inkubation das Medium abgenommen und durch neues ersetzt. Beim Pipettieren des Transfektionsansatzes wurde die in Tabelle 20 aufgeführte Reihenfolge gewählt und nach jedem Pipettierschritt sorgfältig gevortext.

Tabelle 20: Ansatz der Transfektion mit Polyethylenimin (PEI)

Inhalt	Volumenangabe
NaCl (150 mM)	100 µl
<i>DNA</i>	Je 2000 ng (2x bei Doppeltransfektion <i>Target + Guide</i>)
PEI (0,1 mg/ml)	200 µl

Für jedes generierte *Mismatch* erfolgte die Transfektion eines *Target*-Plasmids alleine und im zu vergleichenden Ansatz als Doppel-Transfektion mit *Target*-Plasmid und zugehörigem *gRNA-Cas9*-Plasmid. Im Vorfeld erfolgte eine photometrische Konzentrationsbestimmung der zu transfizierenden Plasmide (2.2.1.4). So konnte eine konstante Applikation von 2000 ng gewährleistet werden. Die Doppel-Transfektion von *Target*- und *gRNA-Cas9*-Plasmid erfolgte im Mengenverhältnis 1:1. Zudem wurde in jeder Transfektionsreihe das Plasmid *RLuc8* (2.1.4.1) alleine transfiziert. So konnte im Anschluss die *BRET-Ratio* der allein exprimierten Luziferase *RLuc8* gemessen werden (2.2.4.1). Dies diente der Normierung der errechneten *frameshift rate* (2.2.5.2).

2.2.3.3 Fluoreszenzmikroskopie

Die Kontrolle der Transfektionen erfolgte mittels Fluoreszenzmikroskopie. Zur Anwendung kam das Fluoreszenzmikroskop Biozero der Firma Keyence aus Neu-Isenburg (2.1.1). Grundlage hierbei war, dass die transfizierten *Target*-Plasmide für ein grün fluoreszierendes Protein (*GFP 2*) codierten (2.1.4.2). Die Beurteilung erfolgte 24 h nach durchgeführter Transfektion (2.2.3.2). Dabei wurden zwei Betrachtungsmodi des Gerätes gewählt. Zunächst der Phasenkontrast zur Beurteilung der Konfluenz. Anschließend der Fluoreszenzmodus zur groben Beurteilung der *GFP 2*-Expression.

2.2.4 Arbeiten mit Proteinen

2.2.4.1 BRET-ASSAY

Der *BRET*-Assay wurde zur Quantifizierung der Schneideaktivität der *gRNA*-Cas9-Komplexe verwendet. Das grundlegende Prinzip von diesem ist der Energietransfer zwischen zwei photoaktiven Molekülen mittels *Bioluminescence Resonance Energy Transfer* (1.3.2) (Kobayashi et al. 2019). Im Rahmen des Energietransfers emittieren sowohl Donor als auch Akzeptor Licht in charakteristischen Wellenlängen. Die Intensität des emittierten Lichts lässt sich mit speziell auf die jeweilige Wellenlänge angepassten Emissionsfiltern technisch erfassen. Der so errechnete Quotient aus Akzeptor- und Donoremissionsintensität kann zur Quantifizierung des Energietransfers genutzt werden. Dieser wird als *BRET-Ratio* bezeichnet (Dacres et al. 2012).

Zur Anwendung kam dabei der in der AG Experimentelle Ophthalmologie (AG Prof. Stieger) etablierte *BRET-Reporter pRLuc8-AvrII/BsiWI-GFP2* (2.1.4.2). Auf diesem wurden die codierenden Abschnitte der Luziferase *RLuc8* (Loening et al. 2006) und des *GFP 2* so platziert, dass zwischen diesen eine beliebige *Target*-Sequenz einer Endonuklease eingefügt werden konnte (2.2.1.3). Die Induktion eines DSB innerhalb der *Target*-Sequenz führte über fehlerhafte Reparaturprozesse zu einer Verschiebung des Leserahmens und damit zu einer veränderten Expression von *GFP 2*. Daraus resultierte eine veränderte Akzeptoremissionsintensität und damit eine Veränderung der *BRET-Ratio*. Die *BRET-Ratio* errechnet sich aus I_{GFP2}/I_{RLuc8} . „I“ ist dabei die Lichtintensität in Photonen pro Sekunde (Wimmer et al. 2021).

Zu Beginn des *BRET*-Assays wurde das verbrauchte Medium der nach 2.2.3.2 transfizierten Zellen abgenommen und diese mit 1 ml 1x PBS gewaschen. Dieses wurde wieder abgenommen, 150 µl 1x PBS hinzugegeben und die adhären Zellen mit einem Zellschaber von der 6-Well Platte gelöst. Das Gemisch aus Zellen und PBS wurde dann in Eppendorfgefäße überführt, welche auf Eis gelagert wurden. Zur Lyse der Zellen folgte eine Schockfrostung in flüssigem Stickstoff mit anschließendem ca. 10-minütigen Auftauen bei Raumtemperatur. Die Schockfrostung wurde einmal wiederholt. Weiter wurden alle Ansätze gevortext und auf höchster Stufe für 5 Minuten bei 4 °C zentrifugiert. Der dabei entstandene, die Proteinlösung enthaltene, Überstand wurde zur weiteren Verwendung in neue Eppendorfgefäße überführt, welche ebenfalls direkt auf Eis gelagert wurden. Die folgende *BRET*-Messung erfolgte automatisiert. Hierfür verwendet wurde ein *Multimode Platerereader* der Firma Tecan aus Männedorf (CH) (2.1.1). Dafür wurden in eine weiße 96-Well-Platte (2.1.2) pro Ansatz je 5 µl Proteinlösung in 4 separate *Wells* pipettiert. Innerhalb des Messvorganges wurden automatisiert in jedes *Well* 100 µl Substratgemisch appliziert. Als Substrat der Luziferase fungierte

Coelenterazine 400a (2.1.3) mit einer Stockkonzentration von 1 µg/ml . Dieses wurde im Verhältnis 1:100 mit 1x PBS verdünnt. Da das Substrat lichtinstabil war, wurde dieses bis zur Messung dunkel gelagert und das Behältnis mit lichtdichter Folie umwickelt.

2.2.5 Statistik und Berechnungen

2.2.5.1 Statistische Auswertung

Die statistische Auswertung erfolgte mit den Programmen Microsoft Excel in der Version 16.78 und Prism 9 für MacOS in der Version 9.5.1. der Firma Graphpad Software, LLC. Mit Prism 9 für MacOS wurden die zu vergleichenden *BRET-Ratios* zunächst auf das Vorliegen einer Normalverteilung untersucht. Dies erfolgte mittels eines Normalverteilungstests nach Shapiro-Wilk. Davon ausgehend wurde das Signifikanzniveau ermittelt. Lag eine Normalverteilung vor, wurde ein T-Test durchgeführt. War dies nicht der Fall, wurde ein Mann-Whitney-Test durchgeführt. Eine Signifikanz wurde angenommen, sofern der P-Wert < 0,05 (*p) lag. Stark signifikant waren Werte von $p < 0,01$ (**p). Höchstsignifikant waren Werte von $p < 0,001$ (***p).

Die Errechnung der Mittelwerte und Standardabweichungen erfolgte mittels Microsoft Excel. Jedes klonierte *Target-*, bzw. *Target-* und *Guide-*Plasmid wurde dreimal ($n=3$) transfiziert und die *RLuc8-* und *GFP 2-*Expression im *BRET-Assay* quantifiziert. Die Bestimmung der Emissionen der verschiedenen Ansätze innerhalb eines *BRET-Assays* erfolgte jeweils in Vierfachbestimmung.

2.2.5.2 Berechnung der *frameshift rate*

In Tabelle 21 im Anhang (11) dargestellt sind die Mittelwerte der einzelnen *BRET-Assays* inklusive der kumulierten Standardabweichungen. Die *BRET-Ratio* ergibt sich aus I_{GFP2}/I_{RLuc8} (Wimmer et al. 2021). Zur Berechnung der *frameshift rate (FSR)* muss die Hintergrundaktivität von *RLuc8* alleine bedacht werden. Die gemessene *BRET-Ratio* der allein exprimierten Luziferase *RLuc8* ist gleichzusetzen mit dem kompletten Fehlen von *GFP 2*. Dies entspräche auf die Doppelexpression von *Target-* und *gRNA-Cas9-*Plasmid angewandt einer vollständigen Aktivität in jeder Zelle und damit einer *FSR* von 100 %. Im Gegensatz dazu entspricht die *BRET-Ratio* der alleinigen *Target-*Plasmid-Expression überhaupt keiner Aktivität und damit einer *FSR* von 0 %. Die *FSR* ist somit die Reduktion der *BRET-Ratio* von alleiniger *Target-*Plasmid-Expression zu der Doppelexpression von *Target-* und *gRNA-Cas9-*Plasmid in Relation zu dem oben festgelegten Bereich zwischen 0 und 100 %. Sie entspricht dem prozentualen Anteil der durch die *gRNA-Cas9-*Komplexe induzierten Leserasterverschiebungen in den zugehörigen *Target-*Plasmiden. Somit ist die *frameshift rate* Aktivitätsmaß der *gRNA-Cas9-*Komplexe (Wimmer et al. 2021).

3 Ergebnisse

3.1 Klonierungen der *Target*-Sequenzen

Eine chronologische Übersicht der Klonierungsarbeiten ist in Abbildung 4 dargestellt. Ziel der Klonierungen war die Herstellung von transfektionsfähigen Plasmiden zur späteren Messung der *Frameshift*-Aktivität der *gRNA*-Cas9-Komplexe.

Die Klonierungsarbeiten der *gRNA*-Cas9-Plasmide waren schon aus Vorarbeiten vorhanden (2.1.12). Zur Vollständigkeit sind diese auch in Abbildung 4 aufgeführt. Dafür wurde die jeweilige *gRNA*-Sequenz in das Plasmid *px459* integriert (2.1.4.3). Es wurde nach einem ähnlichen Schema vorgegangen, wobei *BbsI* als Restriktionsenzym verwendet wurde.

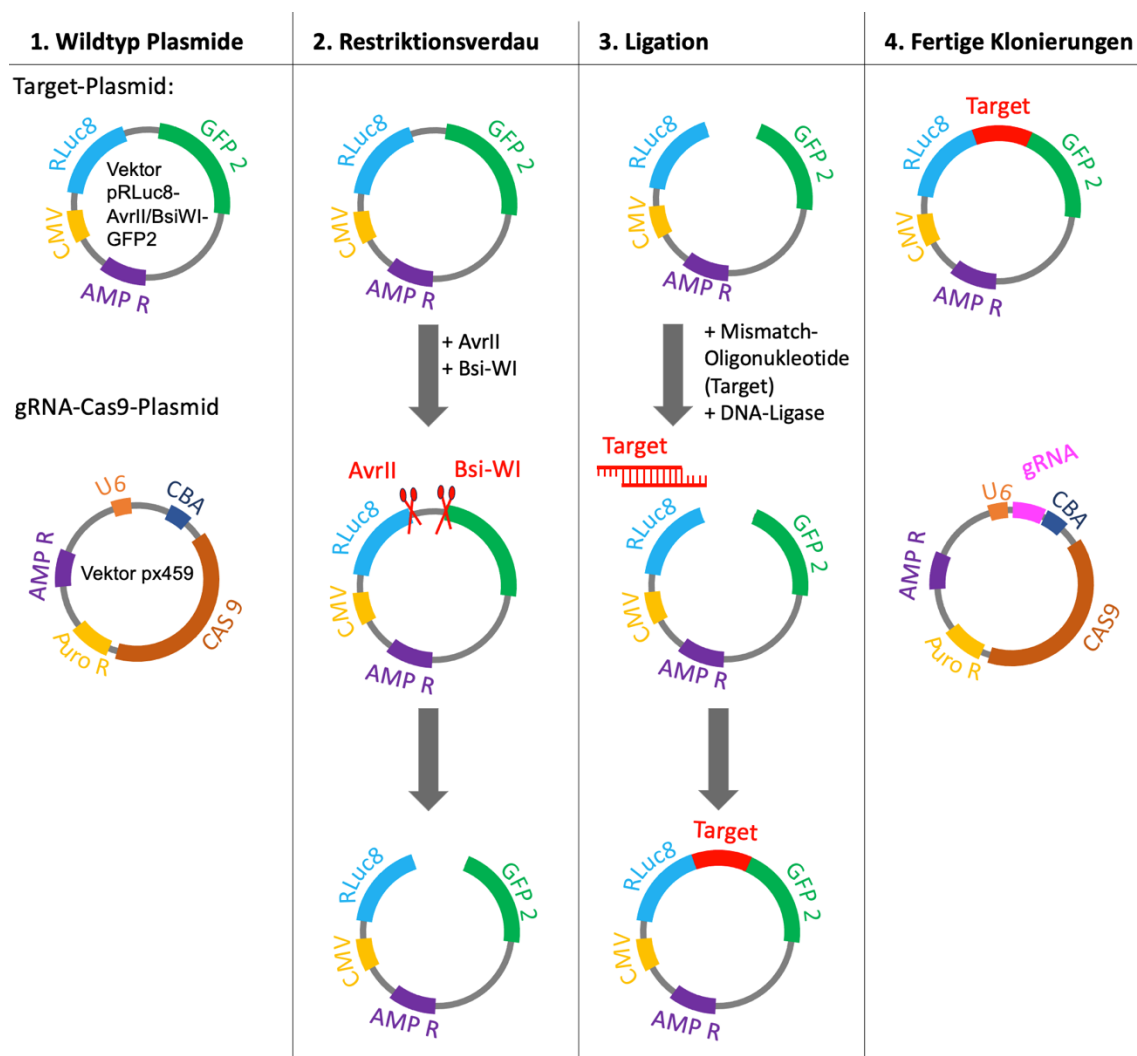


Abbildung 4: Übersichtsschema der Klonierungsarbeiten

AMP R: Ampicillinresistenz, AvrII: Restriktionsenzym, BsiWI: Restriktionsenzym, CMV: Cytomegalie-Virus-Promoter, *RLuc8*: Luziferase, *GFP 2*: *Green Fluorescent Protein 2*, Puro R: Puromycinresistenz, U6: U6-Promoter, *CBA*: *Chicken-Beta-Actin*-Promoter.

Die fertigen Klonierungen wurden zur Verifizierung der erfolgreichen Integration der jeweiligen *Target*-Oligonukleotide an ein externes Labor zur Sequenzierung versendet (2.2.1.5).

In Abbildung 5 ist exemplarisch die Auswertung eines Sequenzierungsergebnisses zu sehen. Nach diesem Schema erfolgte die Überprüfung der erfolgreichen Klonierungen aller veränderten Wildtypsequenzen bzw. *Mismatches*. Bei erfolgreicher Integration folgte eine größere kulturelle Anzucht der plasmidtragenden *E.Coli*-Bakterien mit anschließender Aufreinigung und Isolierung der Plasmide (2.2.2.6).

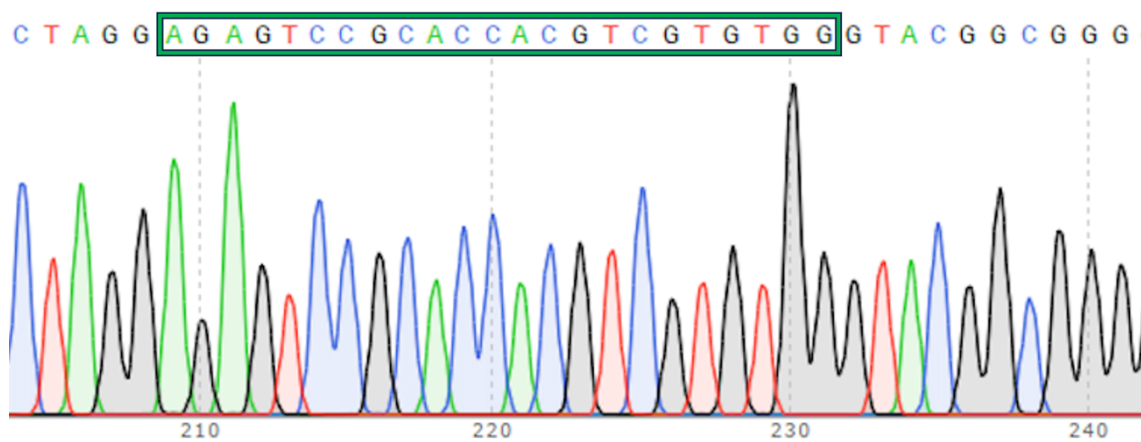


Abbildung 5: Chromatogramm einer Sequenzierung

Dargestellt ist die Sequenzierung des erfolgreich in das Plasmid *pRLuc8-AvrII/BsiWI-GFP2* eingebauten *Target*-Oligonukleotid M 16. Dieses ist grün umrandet. Die Darstellung erfolgt in 5'-3'-Orientierung von links nach rechts.

3.2 Transfektion in *HEK 293T*-Zellen

Ziel der Transfektion war die Genexpression der transfizierten *Target*- und/oder *gRNA*-Cas9-Plasmide. Diese war Voraussetzung für die spätere Quantifizierung der *Mismatch*-Toleranz. Die Transfektion erfolgte in *HEK 293T*-Zellen. Die zuvor isoliert und aufgereinigten Plasmide wurden mittels Polyethylenimin transfiziert (2.2.3.2). Abbildung 6 zeigt exemplarisch die Transfektionsansätze eines *Mismatches*. Einmal das isoliert transfizierte *Target*-Plasmid und im zu vergleichenden Ansatz die Doppeltransfektion von *Target*- und zugehörigen *gRNA*-Cas9-Plasmid.

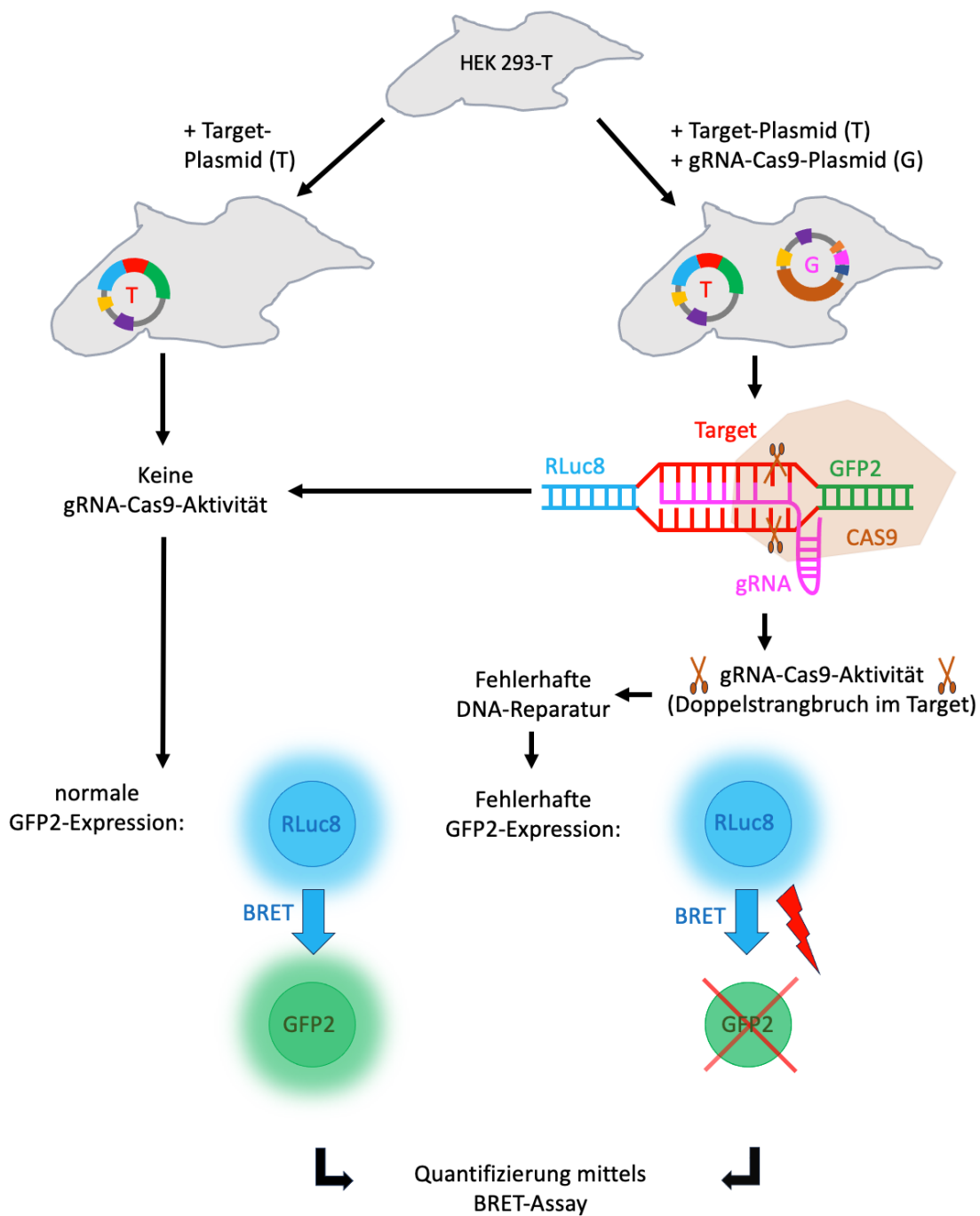


Abbildung 6: Übersicht der Transfektion in HEK 293T-Zellen

Links: isolierte Transfektion des *Target*-Plasmids. Hier zeigt sich keine Aktivität des *gRNA-Cas9*-Komplexes, da das für diesen codierende Plasmid nicht transfiziert wurde. Es resultiert eine regelrechte Expression von *GFP 2*.

Rechts: Doppel-Transfektion von *Target*- und *gRNA-Cas9*-Plasmid. Hier ergaben sich je nach *Target*-Sequenz zwei unterschiedliche Ausgänge: Entweder keine Aktivität des *gRNA-Cas9*-Komplexes, da das *Mismatch* nicht toleriert wurde mit regelrechter Expression von *GFP 2*, oder die Toleranz des *Mismatches* mit Aktivität des *gRNA-Cas9*-Komplexes. Bei einer Aktivität kam es zu einer Induktion eines Doppelstrangbruches mit folgender fehlerhafter *DNA-Reparatur* im *Target*-Plasmid. Dies führte zu einem *Frameshift* mit folglich verminderter *GFP 2*-Expression. Die Reduktion der Expression von *GFP 2* ließ sich anschließend im *BRET*-Assay quantifizieren.

Die Kontrolle der Transfektionen erfolgte mittels Fluoreszenzmikroskopie (2.2.3.3). Grundlage hierbei war, dass die transfizierten *Target*-Plasmide für ein grün fluoreszierendes Protein (*GFP 2*) codierten (2.1.4.2).

In Abbildung 7 dargestellt sind exemplarisch *HEK 293T*-Zellen im Phasenkontrast, sowie Fluoreszenzbilder der Positiv- und Negativkontrolle.

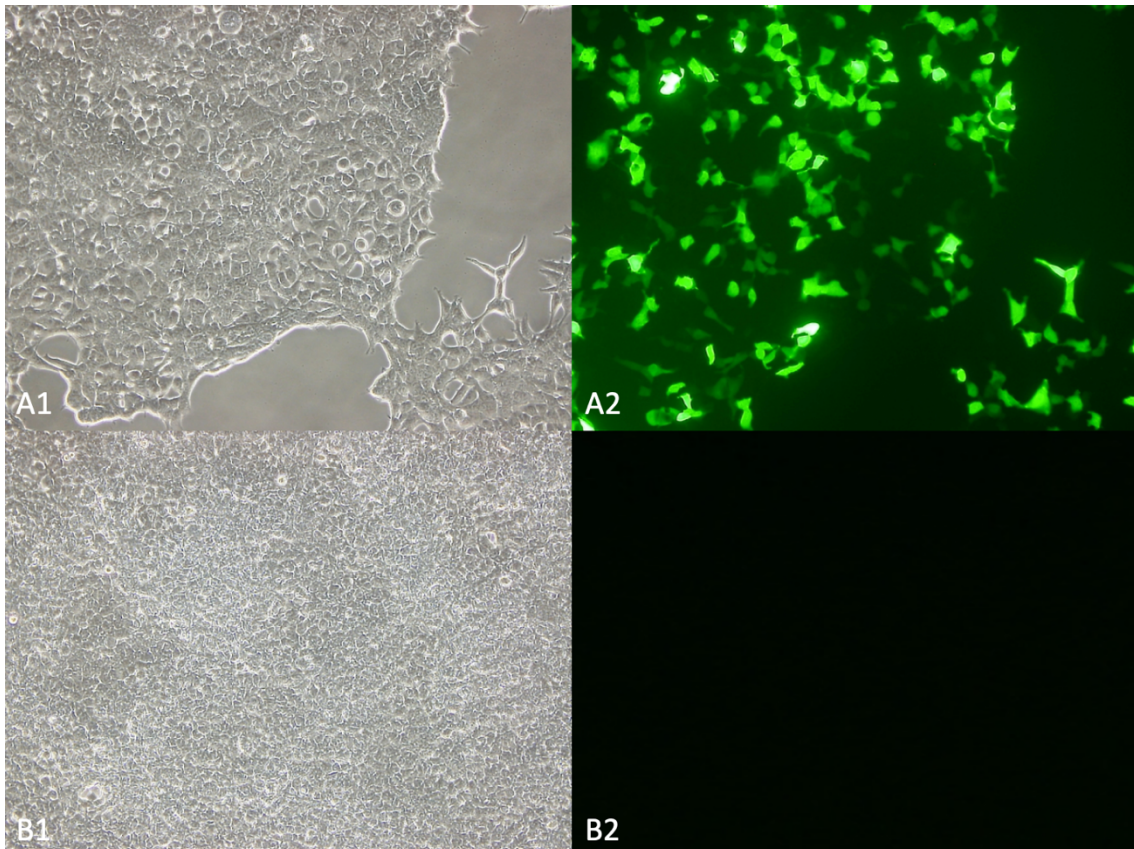


Abbildung 7: Fluoreszenzmikroskopische Kontrolle der Transfektion

A1, B1: Phasenkontrast, A2, B2: Fluoreszenzmodus, A1-A2: Positivkontrolle mit Expression von *GFP 2*.; B1-B2: Negativkontrolle ohne *GFP 2* exprimierendes Plasmid.

So erfolgte jeweils die Überprüfung aller Transfektionsansätze. Bei Darstellung der *GFP 2*-Expression in der Fluoreszenzmikroskopie wurde von einer erfolgreichen Transfektion ausgegangen. Es folgte direkt im Anschluss eine Proteinisolation und anschließend die Verwendung des Isolats im *BRET*-Assay (2.2.4.1).

3.3 BRET-Assays

Ziel der *BRET-Assays* war die Beurteilung der *Mismatch-Toleranz* von *gRNA-Cas9*-Komplexen. Bei einem fluoreszenzmikroskopischen Nachweis der Proteinexpression (2.2.3.3) erfolgte die Isolierung der Proteine und anschließend daraus die Quantifizierung der *RLuc8*- und *GFP 2*-Expression im *BRET-Assay* (2.2.4.1). Dabei erfolgten jeweils Messungen der *BRET-Ratios* der isoliert exprimierten *Target*-Plasmide sowie im zu vergleichenden Ansatz die der Doppelexpression von *Target*- und *gRNA-Cas9*-Plasmid. Pro *BRET-Assay* erfolgte die Messung jedes Ansatzes in Vierfachbestimmung. Pro *Mismatch* erfolgten drei voneinander unabhängige *BRET-Assays*.

Die Mittelwerte der einzelnen *BRET-Assays* inklusive der kumulierten Standardabweichungen und Signifikanzniveaus sind in Tabelle 21 im Anhang (11) zu finden. Die Errechnung der *frameshift rate (FSR)* aus den *BRET-Ratios* ist in 2.2.5.2 dargestellt.

3.3.1 Frameshift rates singulärer Mismatches

Um im Folgenden einzelne Nukleotidinteraktionen zwischen *gRNA* und *Target-DNA* einheitlich und platzsparend zu beschreiben, erfolgt die Darstellung nach dem Schema „rN:dN“. Das Nukleotid der *gRNA* wird mit „rN“ abgekürzt. Das damit interagierende Nukleotid der *Target-DNA* wird mit „dN“ abgekürzt. „N“ meint dabei ein beliebiges Nukleotid. Besagtes Schema orientiert sich an den Darstellungen von Nukleotidinteraktionen in anderen Arbeiten (Kimsey et al. 2018; Pacesa et al. 2022a). Die Position der Interaktion innerhalb der Sequenz soll weiterhin mit „P“ abgekürzt werden (2.1.6). Abbildung 8 zeigt die Interaktion von ART t1 (*Target-DNA*) und ART g1 (*gRNA*). Hiermit soll das Schema visualisiert werden.

P 18 rA:dT P 5 rU:dA

Sequenz		20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	
ART t1	3'	C	C	T	C	A	G	G	C	G	T	G	G	T	G	C	A	G	C	A	C	5'
ART g1	5'	G	G	A	G	U	C	C	G	C	A	C	C	A	C	G	U	C	G	U	G	3'

Abbildung 8: Schema der Beschreibung von Nukleotidinteraktionen

A: Adenin, T: Thymin, C: Cytosin, G: Guanin, U: Uracil. An Position 18 interagiert Adenin der *gRNA* (ART g1) mit Thymin der *Target-DNA* (ART t1). Folglich ergibt sich die Bezeichnung P 18 rA:dT zur Beschreibung dieser Interaktion. Analog dazu P 5 rU:dA.

Abbildung 9 zeigt die *FSRs* von ART t1 (Wildtyp) und den zugehörigen *Mismatches*. Zudem dargestellt sind die Signifikanzniveaus von ART t1 zu seinem jeweiligen *Mismatch*. Die *Mismatches* sind im Gegensatz zu 2.1.6.2 nicht chronologisch nach der Klonierungsgeschichte angeordnet. Stattdessen orientiert sich die Anordnung an deren Positionierung innerhalb der 20 Nukleotide langen *Target*-Sequenz in Relation zum *protospacer adjacent motif (PAM)* (2.1.6).

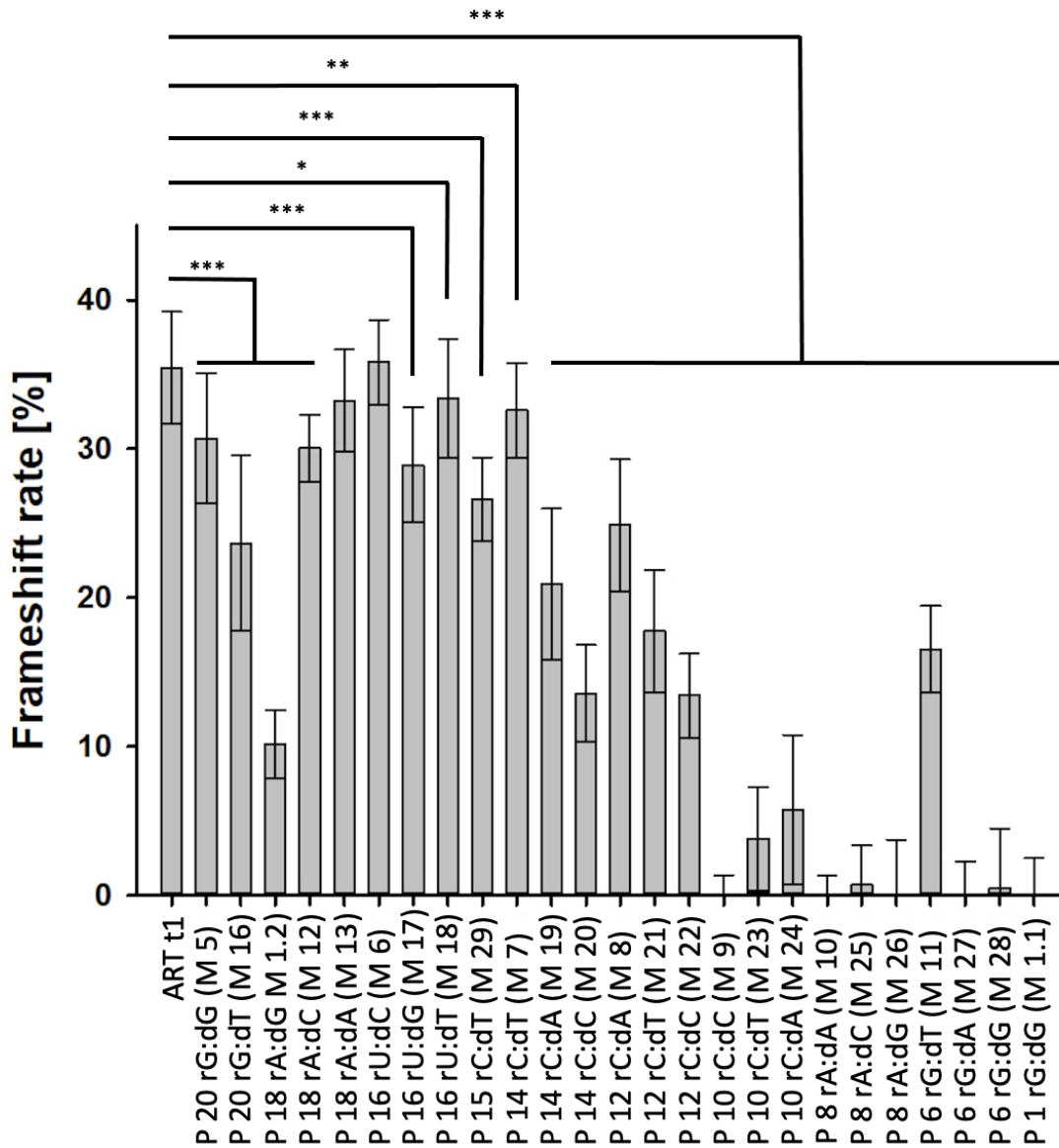


Abbildung 9: Frameshift rates der singulären *Mismatches* von ART t1

Als Referenz abgebildet ist die *FSR* von ART t1. Daneben befinden sich die zugehörigen singulären *Mismatches*. Das Signifikanzniveau zwischen ART t1 und dem jeweiligen *Mismatch* ist sterncodiert dargestellt (* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$). Es ist die einfache Standardabweichung der *FSR* dargestellt. Zudem abgebildet ist die Position und Nukleotidinteraktion des jeweiligen *Mismatches* mit ART g1 (*gRNA*). Das Schema dazu ist 3.3.1 zu entnehmen.

Abbildung 9 ergänzend zeigt Abbildung 10 die Nukleotidsequenz vom *Target*-Wildtyp ART t1 und der korrespondierenden komplementären *gRNA*-Sequenz ART g1. Darunter abgebildet sind die zugehörigen *Mismatch*-Sequenzen mit Hervorhebung der im Vergleich zum Wildtyp veränderten Nukleotide. Auch hier orientiert sich die Anordnung der *Mismatches* an deren Positionierung in Relation zum *PAM*. Für ART t1 und dessen *Mismatches* ist das *Reverse*-Oligonukleotid in 3'-5'-Orientierung von links nach rechts dargestellt (2.1.6). Gegenläufig ist die *gRNA*-Sequenz von ART g1 in 5'-3'-Orientierung von links nach rechts dargestellt. Diese Art der Darstellung wurde gewählt, um die Nukleotidinteraktionen zwischen *gRNA* und *Target-DNA* abzubilden. Das *PAM* ist aus Platzgründen nicht abgebildet. Um die *FSR* zu visualisieren und die Relation zum Wildtyp abzubilden, ist diese farbcodiert dargestellt. Dunkelrot entspricht einer hohen, dem Wildtyp entsprechenden *FSR*. Mit abnehmender *FSR* geht die Farbcodierung über hellrot zu letztendlich weiß. Weiß entspricht einer niedrigen *FSR*.

Sequenz	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	FSR [%]	
ART t1	C	C	T	C	A	G	G	C	G	T	G	G	T	G	C	A	G	C	A	C		35,46 ± 3,76
ART g1	G	G	A	G	U	C	C	G	C	A	C	C	A	C	G	U	C	G	U	G		-
M 5	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30,72 ± 4,36
M 16	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	23,66 ± 5,91
M 1.2	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10,18 ± 2,27
M 12	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30,05 ± 2,27
M 13	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33,24 ± 3,44
M 6	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	35,83 ± 2,84
M 17	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	28,92 ± 3,85
M 18	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33,39 ± 3,98
M 29	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	26,62 ± 2,80
M 7	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	32,60 ± 3,16
M 19	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	20,93 ± 5,06
M 20	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13,59 ± 3,22
M 8	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	24,89 ± 4,42
M 21	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	17,76 ± 4,11
M 22	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	13,44 ± 2,82
M 9	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-1,95 ± 3,30
M 23	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	3,83 ± 3,48
M 24	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	5,75 ± 5,04
M 10	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-1,42 ± 2,76
M 25	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	0,72 ± 2,64
M 26	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-0,86 ± 4,61
M 11	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	16,54 ± 2,90
M 27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-2,74 ± 5,01
M 28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	0,47 ± 4,03
M 1.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-1,37 ± -3,87

Abbildung 10: Target-Sequenzen und farbcodierte FSRs der singulären Mismatches von ART t1

Darstellung der Nukleotidsequenz von ART t1 und zugehörigen *Mismatches* in 3'-5'-Orientierung von links nach rechts. Gegenläufig dazu die Sequenz der korrespondierende *gRNA* (ART g1) in 5'-3'-Orientierung von links nach rechts. Darunter dargestellt sind die zugehörigen *Mismatches* mit Hervorhebung des im Vergleich zu ART t1 veränderten Nukleotids. Die *FSR* wurde farbcodiert. Dunkelrot entspricht hohen Werten. Weiß entspricht niedrigen Werten. Es ist die einfache Standardabweichung der *FSR* dargestellt.

Aus Abbildung 9 und Abbildung 10 geht hervor, dass sich die *FSRs* der jeweiligen *Mismatches* unterscheiden. Es lassen sich interpositionelle, aber auch intrapositionelle Unterschiede ausmachen.

Die *Mismatches* der Positionen 20 bis 14 zeigen *FSRs*, welche größtenteils dem Wildtyp nahekommen. Im Gegensatz dazu zeigen die *Mismatches* der Positionen 10-1 niedrigere Werte der *FSRs*. Die *FSRs* der *Mismatches* von Position 12 liegen zwischen diesen beiden Trends mit mittelhohen Werten.

Abseits dieser interpositionellen Unterschiede finden sich variierende intrapositionelle Unterschiede. Deutlich wird dies zunächst an P 18. Hier zeigen rA:dA und rA:dC *FSRs* von $33,24 \pm 3,44$ % und $30,05 \pm 2,27$ %. Damit liegen diese in einem ähnlichen Bereich wie der Wildtyp mit $35,46 \pm 3,76$ %. Sich davon abhebend zeigt P 18 rA:dG mit $10,18 \pm 2,27$ % eine ca. um den Faktor 3 verminderte *FSR*. Ähnlich findet sich innerhalb von Position 6 mit rG:dT ein Ausreißer. P 6 rG:dA und rG:dG zeigen mit $-2,74 \pm 5,01$ % und $0,47 \pm 4,03$ % niedrige *FSRs*. P 6 rG:dT hingegen zeigt eine *FSR* von $16,54 \pm 2,90$ % und hebt sich damit von den anderen Nukleotidinteraktionen der gleichen Position ab. Ebenfalls intrapositionelle Unterschiede finden sich bei den Positionen 20, 14 und 12. Dem gegenüber gestellt gibt es auch Positionen, in denen sich kaum intrapositionelle Unterschiede ausmachen lassen. Wie innerhalb der Positionen 16 und 8. Hier finden sich für alle Nukleotidkombinationen ähnliche Werte der *FSRs*.

Für P 15 und P 1 wurde jeweils nur eine Nukleotidkombination gemessen. Aussagen über intrapositionelle Unterschiede sind somit nicht möglich. Dennoch lassen sich beide Werte interpositionell in die oben beschriebenen Trends einordnen. P 15 rC:dT zeigt mit $26,62 \pm 2,80$ % eine ähnliche *FSR* wie die umgebenden Positionen 14 und 16. Konträr zeigt P 1 rG:dG mit $1,37 \pm -3,87$ % eine niedrige *FSR* wie die *Mismatches* der Positionen 1-10. Gleichermaßen einordnen lassen sich die vereinzelt singulären *Mismatches* der anderen Wildtypsequenzen. Diese sind in Abbildung 11 dargestellt.

Sequenz	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	FSR [%]
T 4	C	G	C	T	C	A	G	T	C	C	C	C	C	G	A	G	A	A	C	A	40,27 ± 2,72
G 4	G	C	G	A	G	U	C	A	G	G	G	G	G	C	U	C	U	U	G	U	-
M 2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-2,23 ± 4,20
Sequenz	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	FSR [%]
T 11	G	T	C	T	C	A	A	A	G	C	G	A	C	T	C	A	A	C	G	A	52,33 ± 4,49
G 3	C	A	G	A	G	U	U	U	C	G	C	U	G	A	G	U	U	G	C	U	-
M 3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-2,17 ± 4,83
Sequenz	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	FSR [%]
ART t5	G	T	C	A	T	G	G	A	A	A	C	C	C	A	C	C	G	C	C	G	44,47 ± 2,88
ART g5	C	A	G	U	A	C	C	U	U	U	G	G	G	U	G	G	C	G	G	C	-
M 15	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	37,41 ± 3,21
M 4	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	2,07 ± 3,19

Abbildung 11: Target-Sequenzen und farbcodierte *FSRs* der *Mismatches* anderer Wildtypen

Jeweils dargestellt sind die *Target*-Sequenzen (T 4, T 11 und ART t5) in 3'-5'-Orientierung von links nach rechts. Gegenläufig jeweils dazu abgebildet sind die Sequenzen der korrespondierenden *gRNAs* (G 4, G 3, ART g5). Darunter dargestellt sind die zugehörigen *Mismatches* mit Hervorhebung des im Vergleich zum Wildtypen veränderten Nukleotids. Die *FSR* wurde farbcodiert. Dunkelrot entspricht hohen Werten. Weiß entspricht niedrigen Werten. Es ist die einfache Standardabweichung der *FSR* dargestellt.

3.3.2 Frameshift rates multipler Mismatches

Bis jetzt wurden im Ergebnisteil dieser Arbeit nur *Mismatch*-Sequenzen angesprochen, bei denen ein Nukleotid verändert wurde. Abbildung 12 und Abbildung 13 zeigen die Auswirkungen der Kombination von jeweils zwei singular veränderten Nukleotiden innerhalb einer *Target*-Sequenz. Die Doppel-*Mismatches* wurden so konstruiert, dass jeweils zwei bereits gemessene singuläre *Mismatches* miteinander kombiniert wurden, um Aussagen über den kombinatorischen Effekt treffen zu können. Die Anordnung in Abbildung 12 ist so gewählt, dass neben einem Doppel-*Mismatch* die dafür kombinierten singulären *Mismatches* abgebildet sind. Ebenfalls dargestellt sind die Signifikanzniveaus zwischen einem Doppel-*Mismatch* und den zugehörigen singulären *Mismatches*.

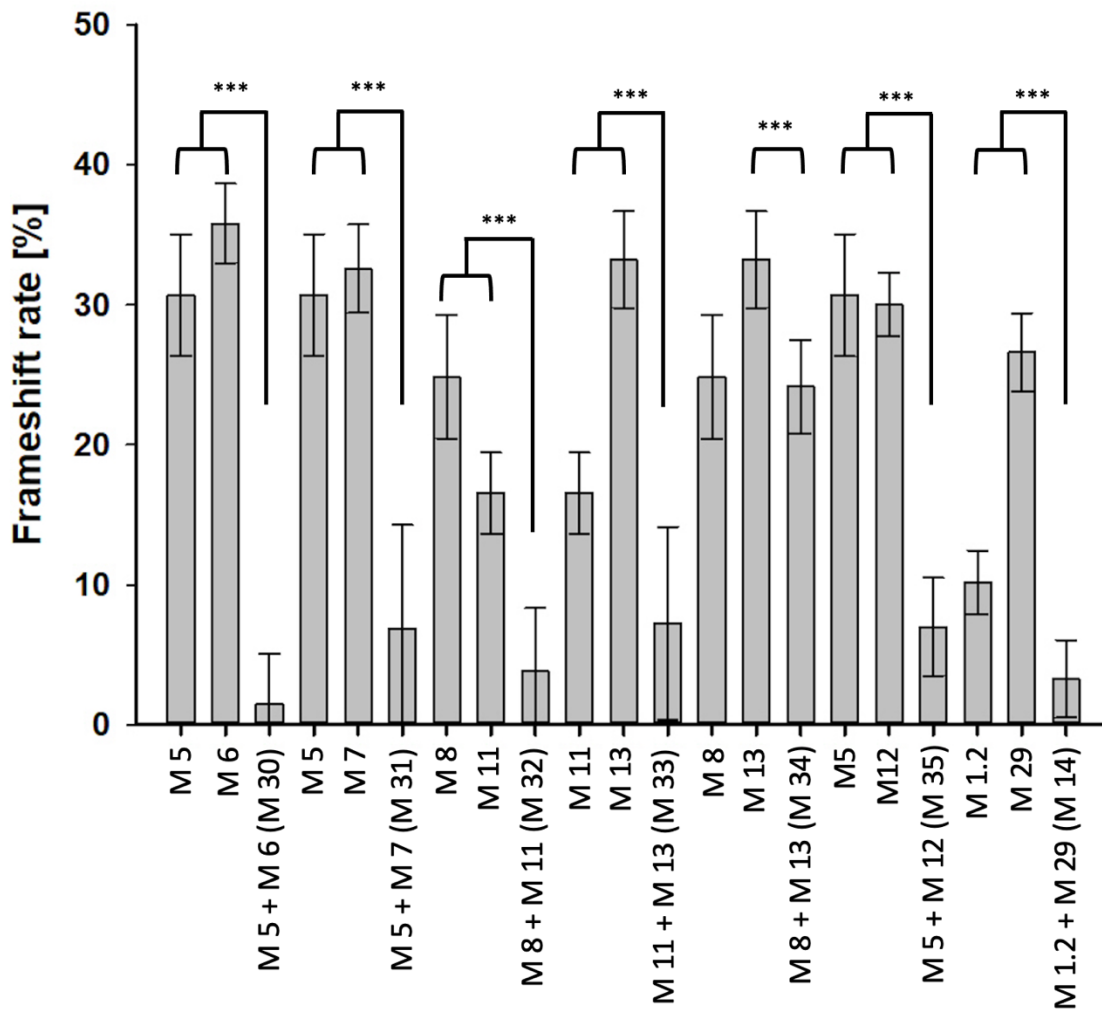


Abbildung 12: Frameshift rates der Doppel-*Mismatches* von ART t1

Die Doppel-*Mismatches* sind gruppiert dargestellt. Links neben einem Doppel-*Mismatch* finden sich jeweils die dafür kombinierten zwei singulären *Mismatches*, inklusive des jeweiligen Signifikanzniveaus. Dieses ist sterncodiert dargestellt (***) ($p < 0,001$). Es ist die einfache Standardabweichung der *FSR* angegeben.

Dies ergänzend zeigt Abbildung 13 die Nukleotidsequenz vom *Target*-Wildtyp ART t1 und der korrespondierenden komplementären *gRNA*-Sequenz ART g1. Darunter abgebildet sind die zugehörigen *Mismatch*-Sequenzen mit Hervorhebung der im Vergleich zum Wildtyp veränderten Nukleotide. Die verschiedenen *Doppel-Mismatches* sind dabei gruppiert, mit den jeweils dafür kombinierten singulären *Mismatches*, angeordnet.

Sequenz	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	FSR [%]
ART t1	C	C	T	C	A	G	G	C	G	T	G	G	T	G	C	A	G	C	A	C	35,46 ± 3,76
ART g1	G	G	A	G	U	C	C	G	C	A	C	C	A	C	G	U	C	G	U	G	-
M 5	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30,72 ± 4,36
M 6	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	35,83 ± 2,84
M 30	G	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,50 ± 3,55
M 5	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30,72 ± 4,36
M 7	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	32,60 ± 3,16
M 31	G	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	6,85 ± 7,42
M 8	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	24,89 ± 4,42
M 11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	16,54 ± 2,90
M 32	-	-	-	-	-	-	-	A	-	-	-	-	-	-	T	-	-	-	-	-	3,80 ± 4,56
M 11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	16,54 ± 2,90
M 13	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33,24 ± 3,44
M 33	-	-	A	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	7,25 ± 6,88
M 8	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	24,89 ± 4,42
M 13	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33,24 ± 3,44
M 34	-	-	A	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	24,16 ± 3,35
M 5	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30,72 ± 4,36
M 12	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30,05 ± 2,27
M 35	G	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6,99 ± 3,50
M 1.2	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10,18 ± 2,27
M 29	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	26,62 ± 2,80
M 14	-	-	G	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	3,27 ± 2,75

Abbildung 13: Target-Sequenzen und farbcodierte FSRs der Doppel-Mismatches von ART t1

Darstellung der Nukleotidsequenz von ART t1 und zugehörigen *Mismatches* in 3'-5'-Orientierung von links nach rechts. Gegenläufig dazu die Sequenz der korrespondierenden *gRNA* (ART g1) in 5'-3'-Orientierung von links nach rechts. Darunter dargestellt sind die zugehörigen *Mismatches* mit Hervorhebung des im Vergleich zu ART t1 veränderten Nukleotids. Die *Doppel-Mismatches* sind gruppiert, mit den jeweils dafür kombinierten zwei singulären *Mismatches*, dargestellt. Die *FSR* wurde farbcodiert. Dunkelrot entspricht hohen Werten. Weiß entspricht niedrigen Werten. Es ist die einfache Standardabweichung der *FSR* angegeben.

Es lässt sich bei allen Doppel-*Mismatches* exklusive M 34 eine Gemeinsamkeit erkennen. Verglichen mit den jeweils singulären *Mismatches* führt die Kombination dieser zu einer ausgeprägten Reduktion der *FSR*. Besonders deutlich wird dies bei M 30, M 31 und M 35. Bei M 30 führt die Kombination aus M 5 mit $30,72 \pm 4,36 \%$ und M 6 mit $35,83 \pm 2,84 \%$ zu einer deutlichen Reduktion der *FSR* auf $1,50 \pm 3,55 \%$. M 31 und M 35 zeigen ähnliche Ausgangswerte und Reduktionen der *FSRs* durch die Kombination dieser. Nicht mit dieser Beobachtung vereinbar ist hingegen M 34. Hier ergibt die Kombination von M 8 mit $24,89 \pm 4,42 \%$ und M 13 mit $33,24 \pm 3,44 \%$ keine ausgeprägte Reduktion der *FSR*. Die *FSR* von M 34 beträgt $24,16 \pm 3,35 \%$ und liegt damit in einem ähnlichen Bereich wie die der dafür kombinierten singulären *Mismatches*.

4 Diskussion

Diese Arbeit behandelt die *Mismatch*-Toleranz von *gRNA-Cas9*-Komplexen. Dafür wurden *Mismatches* generiert, welche sich in einzelnen oder mehreren Nukleotiden vom Wildtyp unterscheiden (2.1.6.2). Zur Quantifizierung der Aktivität der *gRNA-Cas9*-Komplexe an den *Mismatches* wurde ein *BRET-Reporter Assay* verwendet (2.1.4.2). Induziert ein *gRNA-Cas9*-Komplex einen DSB, so kommt es u.a. zu fehlerhaften *DNA*-Reparaturprozessen innerhalb der Zelle mit einer Verschiebung des Leserahmens und damit zu einer veränderten *BRET-Ratio* (Wimmer et al. 2021). Schlussendlich erfolgte aus den gemessenen *BRET-Ratios* die Errechnung einer *frameshift rate (FSR)*, welche als Aktivitätsparameter diente (2.2.5.2).

Hierbei zeigten sich innerhalb der 20 Nukleotide langen *Target*-Sequenzen deutliche inter- und intrapositionelle Unterschiede und daraus schlussfolgernd eine von Position und Nukleotidentität abhängige Toleranz der *Mismatches*.

4.1 Einfluss singulärer *Mismatches* auf die Aktivität von *gRNA-Cas9*

Alle Positionen miteinander vergleichend, wird eine Abnahme der *FSRs* von distal nach proximal deutlich. Visualisiert wird dies durch die Farbcodierung der *FSR* in Abbildung 10 (3.3.1). Anders formuliert werden *Mismatches* nahe des *PAM* weniger toleriert als solche, die weiter distal dieses liegen. Der Einfluss der Position (*P*) des *Mismatches* auf dessen Toleranz steht dabei in Einklang mit bereits etablierten Beobachtungen. Auch andere Arbeiten konnten zeigen, dass distale *Mismatches* tendenziell besser toleriert werden als proximale. Besonders intolerant gegenüber *Mismatches* sind die ersten an das *PAM* angrenzenden Nukleotide. Diese Region wird auch als *Seed* bezeichnet und umfasst bei *SpCas9* ca. die Positionen 1-10 (Hsu et al. 2013; Boyle et al. 2021). Deutlich wird dies in dieser Arbeit an den *Mismatches* der Positionen 10-6 von ART t1. Während *Mismatches* der *P* 12 durchaus noch toleriert werden, zeigt sich ein abrupter Abfall der *FSRs* von *P* 10 und der proximal davon liegenden *Mismatches*. Eine lineare Abnahme der *Mismatch*-Toleranz innerhalb des *Seeds* findet sich jedoch nicht geschlossen in der Literatur wieder. Vielmehr wird eine zweigeteilte Toleranz deutlich. Kontraintuitiv werden hier *Mismatches* der Positionen 1-4 eher toleriert als solche der Positionen 5-8 (Tsai et al. 2015; Kim et al. 2020a). Interessanterweise deuten auch strukturelle Darstellungen von *Cas9* auf eine Zweiteilung des *Seeds* hin. Eine Hybridisierung des Heteroduplexes aus *gRNA* und *TS* über die Position 5 hinaus wird zunächst durch eine Seitenkette von Tyrosin verhindert. Erst eine Auslagerung dieser durch Konformationsänderungen von *Cas9* ermöglicht die weitere Heteroduplexformierung (Pacesa et al. 2022b). Die erhobenen Daten dieser Arbeit lassen keine fundierte Aussage über einen zweigeteilten *Seed* zu, da zu wenig *Mismatches* der Positionen 1-5 konstruiert wurden. Lediglich

M 1.1, M 2 und M 3 liegen in diesem Bereich. Diese zeigen hingegen geringe Werte der *FSRs* und damit keine ausgeprägtere Toleranz in diesem Bereich. Eine Aussage hierüber wäre durchaus von Relevanz, damit in der Antizipation von *Off-Target*-Arealen keine relevanten proximalen *Mismatches* verkannt werden.

Strukturelles Korrelat des *Seeds* ist die Ausbildung einer intermediären *R-Loop*-Konformation, welche ungefähr nach der Hybridisierung eines 9-10 nt langen Heteroduplexes erreicht wird. Diese ermöglicht eine stabile Bindung von *gRNA-Cas9* an der *Target-DNA* und geht dem vollständig geöffneten *R-Loop* mit Repositionierung und Aktivierung der Endonukleasedomänen voraus (Ivanov et al. 2020). Die in dieser Arbeit errechnete *FSR* ermöglicht eine Quantifizierung der Aktivität von *gRNA-Cas9* an den jeweiligen *Mismatch*-Sequenzen. Limitierend ist jedoch, dass diese nur die Endstrecke abbildet. Aussagen über den Einfluss von *Mismatches* auf die strukturelle Dynamik der Interaktion zwischen *gRNA-Cas9* und *Target-DNA* sind folglich nicht möglich. Diese Lücke schließen sequenzielle kristallographische und kryoelektronenmikroskopische Darstellungen (Bravo et al. 2022; Pacesa et al. 2022b; Pacesa et al. 2022a). In Anlehnung hieran impliziert die geringe Toleranz der *Mismatches* von P 6, P 8 und P 10, dass diese den *R-Loop*-Progress in irgendeiner Form behindern (Ivanov et al. 2020), bzw. die intermolekularen Kontakte der interagierenden *Mismatch*-Nukleotide ein sterisches Hindernis darstellen. Letzteres ergibt sich aus der strukturellen Plastizität des *TS*. Dieser ist in der Lage durch Distorsionen intermolekulare Kontakte außerhalb der klassischen Watson-Crick-Bindungen auszubilden (Pacesa et al. 2022a). Besagtes sterisches Hindernis kann sowohl die weitere Hybridisierung des Heteroduplexes als auch die spätere Repositionierung der Endonukleasedomänen beeinträchtigen und somit eine Toleranz des *Mismatches* verhindern (Ricci et al. 2019). Interessant in diesem Kontext ist das *Mismatch* P 6 rG:dT. Dieses zeigt eine *FSR* von $16,54 \pm 2,90$ % und dementsprechend eine mittelhohe Toleranz. Damit steht dieses in Kontrast zu den anderen *Mismatches* der gleichen Position (P 6 rG:dA und P 6 rG:dG), welche kaum toleriert werden. Guanin und Thymin können dabei eine *wobble like* Basenpaarung ausbilden und somit ohne große Distorsionen in den Heteroduplex integriert werden (Pacesa et al. 2022a). Eine klassische Watson-Crick-Basenpaarung wird über sterische Hinderung der Wasserstoff-Atome der Imin-Gruppen verhindert. Bemerkenswert in diesem Kontext ist dennoch die Möglichkeit der transienten Ausbildung einer *watson crick like* Basenpaarung zwischen Guanin und Thymin. Ermöglicht wird diese über Deprotonierung mittels Tautomerisierung oder Ionisierung. Dabei kommt es aus stereochemischer Sicht zu einer nahezu perfekten Imitation der Watson-Crick-Duplexgeometrie (Kimsey et al. 2015). Dieses dynamische Equilibrium aus einer

wobble like und tautomerischen oder anionischen *watson crick like* Konformation (Kimsey et al. 2018) dürfte zur ausgeprägteren Toleranz von P 6 rG:dT beigetragen haben. Aus dem Raster fällt hierbei jedoch M 4 von ART t5. Dieses liegt an einer vergleichbaren Position. Hier interagiert an P 7 rU:dG. Mit einer *FSR* von $2,07 \pm 3,19$ % wird dieses dagegen kaum toleriert. Eben thematisiertes dynamisches Equilibrium besteht jedoch auch zwischen Uridin und Guanin (Kimsey et al. 2015). Dabei sei darauf hingewiesen, dass hier zwar bei beiden *Mismatches* Guanin und Thymin/Uridin interagieren, jedoch an vertauschten Stellen der *gRNA* und des *TS*. Dies ist insofern relevant, da der *TS* und die *gRNA* eine unterschiedliche strukturelle Plastizität besitzen, bzw. die *gRNA* über Verankerung mit *Cas9* in ihrer Beweglichkeit eingeschränkt ist (Pacesa et al. 2022a). Zudem scheinen auch benachbarte Nukleotide Einfluss auf die Ausbildung von Wasserstoffbrückenbindungen zu nehmen, so dass eine isolierte Betrachtung nur einer Position nicht ausreichend ist (Roy et al. 2008; Pacesa et al. 2022a).

Dem hier aufgebauten strukturellen Erklärungsmodell folgend müsste man nun annehmen, dass die Toleranz der *Mismatches* P 6 rG:dA und P 6 rG:dG durch induzierte Distorsionen des *TS* verhindert werde. Dies erscheint jedoch wenig plausibel, da A-G und G-G beide über eine Hoogsteen-Basenpaarung ohne große Distorsionen in den Heteroduplex aus *gRNA* und *TS* integriert werden können (Pacesa et al. 2022a). Dennoch werden diese mit einer *FSR* von $-2,74 \pm 5,01$ % und $0,47 \pm 4,03$ % kaum toleriert. Dies suggeriert, dass die strukturellen Überlegungen der Distorsionen des *TS* durch Non-Watson-Crick-Basenpaarungen alleine nicht ausreichen, um die *Mismatch* Toleranz in ihrer Komplexität zu erfassen. Wohl aber vereinen lassen sich P 6 rG:dA und P 6 rG:dG mit der mechanistischen Rationale der *Mismatch*-Toleranz. Ein *Mismatch* wird demnach toleriert, wenn die durch dieses induzierten strukturellen Veränderungen dennoch eine Repositionierung der Endonukleasedomänen und damit die Überwindung des *conformational checkpoint* ermöglichen (Ricci et al. 2019). Diese Überlegung lässt sich im Umkehrschluss, neben P 6 rG:dA und P 6 rG:dG, auf alle anderen nicht tolerierten *Mismatches* übertragen.

Abseits dieser mechanistischen Einordnung der Ergebnisse, die zusätzlich Arbeiten anderer bedürfen, werden bei der Betrachtung von P 6 intrapositionelle Unterschiede sichtbar. Dies verdeutlicht, dass nicht nur die Position, sondern auch die dort miteinander interagierenden Nukleotide Einfluss auf die Toleranz des *Mismatches* nehmen. Dies unterstreicht die Komplexität der *Mismatch*-Toleranz und findet sich auch so in der Literatur wieder (Huston et al., 2019; Doench et al., 2016).

Neben diesem intrapositionellen Unterschied wurden oben kurz schon interpositionelle Unterschiede erwähnt. Die *Mismatches* dieser Arbeit der Positionen 12-20 werden größtenteils toleriert. Im Gegensatz dazu werden *Mismatches* der Positionen 1-10, mit Ausnahmen, kaum toleriert. Position 12 stellt demnach einen Umschlagspunkt dar. Unter Vernachlässigung des Einflusses der benachbarten Nukleotide (Roy et al. 2008; Pacesa et al. 2022a) interagieren an P 10, P 12 und P 14 jeweils dieselben Nukleotide miteinander, so dass hier am ehesten auch ein Einfluss der Position des *Mismatches* relevant wird. P 12 als Umschlagspunkt ist dabei interessant und in Einklang mit biophysikalischen Überlegungen der *R-Loop*-Kinetik. Die *R-Loop*-Formierung erfolgt über drei energetisch metastabile Stadien. Diese sind der geschlossene, intermediäre und letztendlich offene *R-Loop*. Die intermediäre Konformation entspricht einer Heteroduplexformierung der Positionen 9-12. *Mismatches* proximal der Position 9 sind demnach energetisch ungünstiger und favorisieren eher einen *R-Loop*-Kollaps mit Dissoziation von *Cas9* (Ivanov et al. 2020; Eslami-Mossallam et al. 2022). Dies findet sich in den hier erhobenen Daten an der nicht vorhandenen Toleranz der *Mismatches* von P 6 und P 8 wieder. Weniger vereinbaren hiermit lassen sich jedoch die geringe Toleranz der *Mismatches* von Position 10. Diese liegen ebenfalls in dem beschriebenen lokalenergetischen Minimum (Eslami-Mossallam et al. 2022). Vergleicht man die *gRNA-Cas9*-Aktivitäten an den Positionen 9-12 mit den Daten anderer Veröffentlichungen, so findet sich ein sehr durchwachsenes Bild. Die Daten unterscheiden sich sowohl intrapositionell als auch interpositionell. In Einklang mit dieser Arbeit konnten auch Boyle et al. P 12 als Umschlagspunkt ausmachen (Boyle et al. 2017). Im Gegensatz dazu findet sich in anderen Veröffentlichungen eine weitaus größere Variabilität in diesem Bereich (Doench et al. 2016; Boyle et al. 2021).

Begründen lässt sich diese zwischen verschiedenen *gRNAs* variierende *Mismatch*-Toleranz mit dem Einfluss *guide*-intrinsischer Eigenschaften. Eine isolierte Betrachtung der an einer Position miteinander interagierenden Nukleotide wird somit der *Mismatch*-Toleranz nur eingeschränkt gerecht und auch eine Vergleichbarkeit mit anderen Arbeiten ist dadurch nicht immer ohne Weiteres möglich. Vielmehr muss der Einfluss der gesamten *gRNA*-Sequenz auf einzelne *Mismatches* mitbeachtet werden (Huston et al. 2019). Die Sequenzdeterminanten der *guide*-intrinsischen *Mismatch*-Toleranz sind noch größtenteils unklar. Allerdings konnte ein hoher Anteil an Guanin und ein niedriger Anteil an Thymin innerhalb der *gRNA*-Sequenz mit einer ausgeprägteren *Mismatch*-Toleranz in Verbindung gebracht werden (Fu et al. 2022). Wendet man dies auf die in dieser Arbeit am häufigsten verwendete *gRNA* (ART g1) an, so findet man ein Überwiegen von Guanin gegenüber Thymin im Verhältnis 7:3 (2.1.6.1). Es erscheint demnach plausibel,

dass ART g1 eine höhere *guide*-intrinsische *Mismatch*-Toleranz innehat und sich folglich weitaus promisker als andere *gRNA*-Sequenzen verhalten könnte. Dies ist vom *Mismatch*-Kontext unabhängig und beeinflusst die Toleranz der *Mismatches* über die gesamte Länge des Heteroduplexes (Huston et al. 2019; Fu et al. 2022).

Mit Blick auf die mittleren und distalen *Mismatches* lohnt sich ein erneuter Verweis auf die enge Kopplung von Struktur und Funktion von *Cas9*. Die *R-Loop*-Formierung als dynamischer Prozess geht mit wechselnden intermolekularen Kontakten zwischen *Target-DNA* und *Cas9* einher (Pacesa et al. 2022b). Eine besondere Funktion übernimmt dabei die *REC 3*-Domäne von *Cas9*. Diese kontrolliert allosterisch die *HNH*-Endonukleasedomäne und interagiert v.a. mit dem distalen Heteroduplex (Chen et al., 2017). Durch *Mismatches* induzierte Veränderungen der Heteroduplex-Architektur bedingen neue intermolekulare Kontakte zwischen *REC 3* bzw. *HNH* mit dem *TS*. Diese können *HNH* in seiner Konformationsdynamik einschränken und eine Überführung in den katalytisch aktiven Zustand verhindern (Mitchell et al. 2020; Pacesa et al. 2022a). Bravo et al. konnten zeigen, dass *REC 3* in dieser Dynamik keinen direkten Kontakt mit den P 12-14 des Heteroduplexes ausbildet und in diesem Bereich eine ausgeprägtere *Mismatch*-Toleranz besteht (Bravo et al. 2022). Auch in dieser Arbeit werden *Mismatches* der P 12-14 größtenteils toleriert. Auffallend sind hier jedoch deutliche intrapositionelle Unterschiede. Beispielsweise zeigt P 14 rC:dT mit $32,60 \pm 3,16$ % eine *FSR* nahe des Wildtyps. P 14 rC:dC hingegen wird mit einer *FSR* von $13,59 \pm 3,22$ % weniger toleriert. Ähnlich verhält sich P 12. Dies impliziert, dass ein direkter Kontakt des *Mismatches* mit *REC 3* nicht zwangsläufig notwendig ist, um als solches erkannt und nicht toleriert zu werden. Unterstützt wird dies durch ähnliche intrapositionelle Unterschiede in diesem Bereich in anderen Arbeiten (Boyle et al. 2021).

Im Zusammenspiel zwischen der Architektur des distalen Heteroduplexes und der Konformationsdynamik von *HNH* konnte Position 17 als besonders relevant ausgemacht werden. Vier konsekutive distale *Mismatches* (P 17-20) blockieren *HNH*. Im Gegensatz dazu ist bei nur 3 konsekutiven distalen *Mismatches* (P 18-20) eine Repositionierung von *HNH* in den katalytisch aktiven Zustand noch möglich (Ricci et al. 2019; Mitchell et al. 2020). Strukturelles Korrelat der konsekutiven distalen *Mismatches* ist der Verlust der Basenpaarung mit einem dort geöffneten Heteroduplex. Jedoch induzieren schon zwei konsekutive *Mismatches* der P 16-17 eine ähnliche Öffnung des dahinter liegenden Heteroduplexes (Mitchell et al. 2020). Position 17 wurde in dieser Arbeit nicht abgedeckt, wohl aber die benachbarten Positionen 16 und 18. Die *FSRs* von Position 16 liegen mit rU:dC $35,83 \pm 2,83$ %, rU:dG $28,92 \pm 3,85$ % und rU:dT $33,39 \pm 3,98$ % alle in einem hohen Bereich, der dem Wildtyp nahe kommt. Folglich werden an P 16 alle *Mismatches*

toleriert. Singuläre *Mismatches* der P 16 oder P 17 werden in der Literatur variierend toleriert. Auch die Kombination beider erbringt einen kombinatorischen Effekt im mittleren bis oberen Bereich bei noch zu vernehmender Schneideaktivität (Huston et al. 2019; Boyle et al. 2021; Fu et al. 2022). Eine auf dieser Arbeit fußende Schlussfolgerung des Einflusses von P 16 auf den kombinatorischen Effekt eines Doppel-*Mismatches* P 16 + P 17 ist jedoch aufgrund der fehlenden Daten von P 17 nicht möglich. Dies wäre durchaus von Bedeutung, um weitere Einsicht in diesem strukturell relevanten Bereich zu erlangen.

Obwohl ein konsekutives distales *Mismatch* der P 18-20 die *HNH*-Domäne in ihrer Konformationsdynamik kaum einschränkt und einen katalytisch aktiven Zustand ermöglicht (Mitchell et al. 2020), zeigen sich in dieser Arbeit dennoch unterschiedliche Toleranzen der *Mismatches* von P 18 und P 20. Besonders ins Auge fällt P 18 rA:dG mit einer *FSR* von $10,18 \pm 2,27$ %. Dieses wird deutlich weniger toleriert als P 18 rA:dC und P 18 rA:dA mit einer *FSR* von $30,05 \pm 2,27$ % und $33,24 \pm 3,44$ %. Ähnliche intrapositionelle Schwankungen in diesem Bereich finden sich auch bei anderen Veröffentlichungen (Fu et al., 2022; Doench et al., 2016). Dies unterstreicht, dass Schlussfolgerungen der *Mismatch*-Toleranz fußend auf bereits bekannten Daten nicht immer trivial sind und multiple *Mismatches* sich auch kontraintuitiv verhalten können (Pacesa et al. 2022a). Mit steigender Zahl der *Mismatches* sinkt tendenziell die Toleranz dieser (Kim et al. 2020a). Zudem zeigen konsekutive distale *Mismatches* einen deutlichen kombinatorischen Effekt auf die Toleranz in Bezug zu den dafür kombinierten singulären *Mismatches* (Boyle et al. 2021). Dennoch ermöglicht *Cas9* bei dem konsekutiven *Mismatch* P 18-20 eine Aktivierung der *HNH*-Domäne. Ursächlich hierfür ist ein *Loop* der *RuvC*-Domäne, welcher distale Distorsionen des Heteroduplex stabilisiert und so dennoch die Repositionierung von *HNH* ermöglicht (Bravo et al. 2022). Auf den kombinatorischen Effekt multipler *Mismatches* soll im nächsten Kapitel detaillierter eingegangen werden.

4.2 Einfluss multipler *Mismatches* auf die Aktivität von *gRNA-Cas9*

Die zuvor thematisierten singulären *Mismatches* stellen die einfachste Form der *Mismatches* dar. Sehr viel komplexer und schwieriger in der Antizipation der Toleranz wird es, wenn multiple *Mismatches* bestehen (Boyle et al. 2021). In 3.3.2 sind Sequenzen mit zwei *Mismatches* und ihre *FSRs* dargestellt. Es wurden jeweils zwei bereits gemessene singuläre *Mismatches* zu einem Doppel-*Mismatch* kombiniert, um so Aussagen über einen kombinatorischen Effekt treffen zu können.

Die energetischen Überlegungen der *R-Loop*-Kinetik lassen sich auch auf multiple *Mismatches* übertragen. Jedes *Mismatch* stellt eine energetische Barriere dar und erhöht die freie Energie für die gesamte darauf folgende *R-Loop*-Progression. Mehrere *Mismatches* gehen folglich mit einer höheren freien Energie einher, so dass ein *R-Loop*-Kollaps mit Dissoziation von Cas9 energetisch favorisiert werden kann (Eslami-Mossallam et al. 2022). In Bezug dazu ist der kombinatorische Effekt multipler *Mismatches* jedoch nicht zwangsläufig additiv bzw. linear. Vielmehr hängt dieser von mehreren Faktoren ab. Diese sind die Anzahl (Kim et al. 2020a), sowie die Lokalisation innerhalb der Sequenz und der Abstand der *Mismatches* zueinander (Boyle et al. 2021; Fu et al. 2022).

M 30 und M 35 liegen in dem im letzten Kapitel thematisierten relevanten strukturellen Bereich der Positionen 16-20. Erwähnenswert ist hierbei, dass bei beiden distale singuläre *Mismatches* mit einer hohen Toleranz kombiniert wurden. Bei M 30 führte die Kombination von *FSRs* mit $30,72 \pm 4,36$ % (P 20 rG:dG) und $35,83 \pm 2,84$ % (P 16 rU:dC) zu einer ausgeprägten Reduktion der *FSR* auf $1,50 \pm 3,55$ %. Ähnliches zeigt sich für M 35. Hier erbrachte die Kombination von *FSRs* mit $30,72 \pm 4,36$ % (P 20 rG:dG) und $30,05 \pm 2,27$ % (P 18 rA:dC) eine *FSR* von $6,99 \pm 3,50$ %. Dementsprechend wird ein kombinatorischer Effekt mit geringerer Toleranz der Doppel-*Mismatches* deutlich. Dies deckt sich mit anderen Arbeiten (Huston et al. 2019; Fu et al. 2022). Die distale Lokalisation beider *Mismatches* geht dabei mit einem ausgeprägteren kombinatorischen Effekt einher (Fu et al. 2022). In Anlehnung an den strukturellen Toleranzmechanismus mittels Stabilisierung des distalen Heteroduplexes durch einen *Loop* der *RuvC*-Domäne (Bravo et al. 2022) impliziert dies, dass nicht alle *Mismatch* induzierten Distorsionen durch diesen kompensierbar sind. Vergleichbare kombinatorische Effekte mit ebenfalls noch distal lokalisierten *Mismatches* zeigen sich ebenfalls für M 31 und M 14.

Sind eines oder mehrere der *Mismatches* innerhalb des *Seeds* lokalisiert, so werden diese tendenziell schlechter toleriert (Boyle et al. 2017). Dabei wird v.a. im Bereich des *Seeds* ein *epistasis like* Effekt relevant. Wird ein singuläres *Mismatch* schon schlecht toleriert, so könnte dieses unabhängig von weiteren *Mismatches* eine Toleranz verhindern. Der *epistasis like* Effekt verhält sich i.d.R. invers zum kombinatorischen Effekt (Fu et al. 2022). Da in dieser Arbeit nur größtenteils tolerierte singuläre *Mismatches* kombiniert wurden, ist keine fundierte Aussage über einen *epistasis like* Effekt möglich. Dennoch interessant in diesem Kontext ist M 34. Hier ergibt die Kombination von *FSRs* mit $33,24 \pm 3,44$ % (P 18 rA:dA) und $24,89 \pm 4,42$ % (P 12 rC:dA) eine *FSR* von $24,16 \pm 3,35$ %. Dementsprechend liegt die Toleranz des

Doppel-*Mismatches* ziemlich genau in dem Bereich des dafür kombinierten weniger tolerierten singulären *Mismatches*. Hier ließe sich ein *epistasis like* Effekt spekulieren. Jedoch sei betont, dass aufgrund der geringen Anzahl an Doppel-*Mismatches* in dieser Arbeit dazu keine aussagekräftigen Schlussfolgerungen möglich sind. Nichts destotrotz stellt M 34 innerhalb der Doppel-*Mismatches* eine Ausnahme dar. Die ausgeprägte Toleranz von M 34 verdeutlicht, dass nicht jedes Doppel-*Mismatch* einen starken kombinatorischen Effekt bedingt. Interessanterweise konnten auch Fu et al. für die Kombinationen von P 12 und P 18 (M 34) einen geringeren kombinatorischen Effekt zeigen. Abschließend dazu sei jedoch nochmal auf die *guide*-intrinsische *Mismatch*-Toleranz verwiesen, die selbstverständlich auch für multiple *Mismatches* einer Sequenz gilt (Huston et al. 2019; Fu et al. 2022).

4.3 Variabilität der *Mismatch*-Toleranz

Mismatch-Toleranz bzw. *Off-Target*-Aktivität ist keineswegs trivial und von einer Vielzahl an Faktoren abhängig. Besonders relevant ist das Ausmaß der Expression von *gRNA*-Cas9. Mit Zunahme der Expression kommt es zu einer ausgeprägteren *Mismatch*-Toleranz (Wang et al. 2020). In der späteren therapeutischen Anwendung von *CRISPR*-Cas wird *gRNA*-Cas9 als variable Größe appliziert, um das sich bereits in der Zelle befindende Genom zu modifizieren. Die Kontrolle der *gRNA*-Cas9-Expression ist hier bedeutsam in der Reduktion von *Off-Target*-Aktivität und *Mismatch*-Toleranz (Wang et al. 2015; Yin et al. 2016). In diesem Aspekt liegt ein wichtiger Unterschied zu dieser Arbeit. Hier erfolgte eine Transfektion des *Target*-DNA- und *gRNA*-Cas9-Plasmides im Verhältnis 1:1 (2.2.3.2), so dass eine Expression in ähnlichem Maße anzunehmen ist. Ein Überwiegen der *gRNA*-Cas9-Expression durch ein anderes Mengenverhältnis in der Applikation führt jedoch zu einer deutlich promiskeren *Mismatch*-Toleranz (Huston et al. 2019). Folglich muss betont werden, dass die erhobenen Daten nur für ein *gRNA*-Cas9:*Target*-DNA Verhältnis von 1:1 repräsentativ sind. Neben der hier verwendeten DNA basierten Applikation gibt es noch die Möglichkeit der Applikation als RNP und *mRNA*, welche schneller in der Zelle inaktiviert werden und somit ein anderes Expressionsverhalten zeigen (Wienert et al. 2022). Vorteil des in dieser Arbeit verwendeten *BRET*-Reporters ist, dass dieser einerseits als Expressionskontrolle fungierte und andererseits eventuelle Schwankungen der Expression zwischen den Einzelmessungen durch die Unabhängigkeit der *BRET*-Ratio davon kompensiert wurden (Wimmer et al., 2021; Wimmer, 2018).

Ebenfalls in diesem Kontext relevant ist der Einfluss der DNA-Struktur. Das Ausmaß an *Supercoiling* beeinflusst die Zugänglichkeit von *gRNA*-Cas9 zur DNA und die darauf folgend induzierte DNA-Strang-Separierung (Scott et al. 2018). DNA-Bereiche mit

negativen *Supercoiling* erleichtern den Progress des *R-Loops*. Somit kommt es zu einer ausgeprägteren *Mismatch*-Toleranz bzw. Promiskuität von *gRNA-Cas9* in diesen Bereichen. Relevant wird dies in Promotor- und transkribierten Bereichen (Ivanov et al. 2020). In diesen induziert die *RNA*-Polymerase im Rahmen der Transkription Bereiche des starken negativen *Supercoilings* (Ma et al. 2013). Die Messung der *Mismatch*-Toleranz erfolgte in dieser Arbeit ohne parallele Messung des *Torques* bzw. des *Supercoilings*. Hier ist zu erwähnen, dass die *Target*-Sequenzen zwischen den codierenden Bereichen von *RLuc8* und *GFP 2* lagen (Wimmer et al. 2021). Damit befanden sich diese in einem transkribierten Bereich. Dies muss bei der Übertragung der Ergebnisse dieser Arbeit auf Bereiche anderer *DNA*-Topologie bedacht werden. Besonders relevant wird dies in der späteren klinischen Anwendung von *CRISPR-Cas9*, da hier dann interindividuelle genetische Variationen berücksichtigt werden müssen (Wienert et al. 2022).

4.4 Möglichkeiten und Einschränkungen des *BRET-Reporters*

Wie schon in 1.1.2.5 und 1.3 kurz skizziert, gibt es eine Reihe an Möglichkeiten die Aktivität von *gRNA-Cas9* nachzuweisen, die an unterschiedlichen Stellen in der Planung der späteren therapeutischen Applikation von *CRISPR-Cas* Anwendung finden (Martin et al. 2016). Neben klassischen Sequenzierungsmethoden, die auf dem Nachweis von Indels beruhen, besteht die Möglichkeit der Integration von Oligonukleotiden in durch *gRNA-Cas9* induzierte DSB. Vorteil des Letzteren ist die größere Sensitivität. Nachteil ist jedoch, dass dies von der Integrationseffizienz der verwendeten Zelllinie abhängig ist (Blattner et al. 2020). Dies ist ein Vorteil des in dieser Arbeit verwendeten *BRET-Reporters*, da ein durch *gRNA-Cas9* induzierter DSB keiner anschließenden Amplifikation und Sequenzierung oder Integration von Oligonukleotiden bedarf, sondern durch eine Veränderung der *BRET-Ratio* nachgewiesen wird (Wimmer et al. 2021). Dadurch lassen sich hier jedoch keine Aussagen über *Cas9* induzierte Indels treffen. Dies wäre relevant gewesen, um weitere Aufschlüsse über die Sequenzdeterminanten der Indel-Induktion nach *NHEJ* zu erlangen (Shen et al. 2018; Allen et al. 2019). Der *BRET-Reporter* eignet sich somit gut zur Grundlagenforschung des Einflusses von *Mismatches* (Wimmer et al. 2021). Dennoch kann er die Komplexität des menschlichen Genoms nicht ausreichend abbilden und auch in der genomweiten Antizipation von *Off-Target*-Arealen sind andere etablierte Methoden zu ergänzen (Wimmer et al. 2021; Wienert et al. 2022).

4.5 Schlussfolgerungen und Ausblick

Die Messung der *Mismatch*-Toleranz von *gRNA-Cas9* in dieser Arbeit verdeutlicht die Komplexität dieser. Dabei konnte eine Abhängigkeit der *Mismatch*-Toleranz von der Position und Nukleotididentität gezeigt werden und die Relevanz des *Seeds* reproduziert werden. Zudem führte die Kombination von zwei singulären *Mismatches*, mit einer Ausnahme, zu einem ausgeprägten kombinatorischen Effekt mit geringerer Toleranz des dabei entstandenen Doppel-*Mismatches*. Dennoch ist bei der Übertragung der hier erhobenen Daten auf den Einfluss der *guide*-intrinsischen *Mismatch*-Toleranz (Huston et al. 2019; Fu et al. 2022) und des Expressionsausmaßes von *gRNA-Cas9* in Relation zur *Target-DNA* zu achten (Wang et al. 2020). Der *BRET-Reporter* ist dennoch eine geeignete Methode des Aktivitätsnachweises von *gRNA-Cas9* (Wimmer et al. 2021). Damit reiht er sich in andere Arbeiten, die sich in diesem Kontext ebenfalls Biolumineszenz zu Nutze machen (Chen et al. 2013; Wang et al. 2020). Während die Mechanismen der *Mismatch*-Toleranz noch nicht vollständig verstanden sind, wurden in den letzten Jahren enorme Fortschritte in diesem Bereich gemacht. Die erhobenen Daten finden Anwendung in verschiedenen *Tools* und werden durch *machine learning*-Applikationen ergänzt (Boyle et al. 2021; Eslami-Mossallam et al. 2022; Fu et al. 2022). Dabei findet *CRISPR-Cas* bereits seit Jahren Anwendung in der therapeutischen Applikation am Menschen. Hierbei zeigen sich vielversprechende Ergebnisse im Bereich der Therapie von hereditären Erkrankungen und Malignomen. Problematisch ist jedoch die Frage, inwiefern *Off-Target*-Aktivität klinisch relevant wird und detektiert werden kann (Zheng et al. 2023). Die in 1.1.2.5 und 1.3 beschriebenen Methoden der *Off-Target*-Antizipation und dem Nachweis von Geneditierungen sind nicht ohne Weiteres auf die spätere therapeutische Applikation übertragbar. Insbesondere der Chromatinkontext und zelluläre Reparaturmechanismen bedingen, dass weniger der antizipierten adversen Geneditierungen im Endeffekt phänotypisch bzw. klinisch relevant werden. In den nächsten Jahren ist mit einem wissenschaftlichen Fortschritt in der Entwicklung sensitiver Nachweismethoden von *Off-Target*-Arealen zu rechnen. Die Anwendung dieser auf geneditierte Zellen von Patientinnen und Patienten wird dann weitaus differenziertere Einblicke in die *Mismatch*-Toleranz unter Berücksichtigung des humanzellulären Kontextes ermöglichen (Wienert et al. 2022).

5 Zusammenfassung

Die therapeutische Applikation von *CRISPR-Cas* birgt enormes Potential. Anwendung findet dies u.a. in der Therapie von genetischen Erkrankungen, Malignomen oder auch in der medizinischen Diagnostik. Dementsprechend groß ist das Interesse potentielle Nebenwirkungen zu antizipieren und zu reduzieren. Im Fokus steht hierbei die *Mismatch*-Toleranz von *CRISPR-Cas*. Dieses kann auch *DNA*-Sequenzen binden und modifizieren, die nicht vollständig komplementär zu der konstruierten Erkennungssequenz sind. Damit besteht die Gefahr nicht intendierte Genareale zu modifizieren, insbesondere, da die Mechanismen der *Mismatch*-Toleranz in ihrer Komplexität noch unvollständig verstanden sind.

In dieser Arbeit wurde die *Mismatch*-Toleranz von *gRNA-Cas9* in eukaryotischen *HEK 293T*-Zellen erprobt. Zur Anwendung kam dabei ein *BRET-Reporter*, welcher sich den Mechanismus des *Bioluminescence Resonance Energy Transfer* zu Nutze machte. Eine Aktivität des *gRNA-Cas9*-Komplexes konnte über eine Reduktion des Energietransfers quantifiziert werden und somit Aussagen über die Sequenzdeterminanten der Toleranz eines *Mismatches* getroffen werden. Die einzelnen *Mismatches* unterschieden sich in einem oder mehreren Nukleotiden von der Wildtypsequenz.

Es konnte gezeigt werden, dass die Toleranz eines *Mismatches* einerseits von dessen Position, andererseits von dessen Nukleotidentität abhängig ist. So wurden für die Interaktion verschiedener Nukleotide intrapositionelle Unterschiede deutlich. Auch interpositionell wurden, unabhängig von der Nukleotidentität, Bereiche mit variierender Toleranz deutlich. Des Weiteren wurde der kombinatorische Effekt des Zusammenbringens zweier singulärer *Mismatches* innerhalb einer Sequenz erprobt. Dabei zeichnet sich größtenteils ein deutlicher kombinatorischer Effekt ab, so dass tendenziell mit Zunahme der Anzahl der *Mismatches* die Toleranz dieser abnimmt. Die hier erhobenen Daten tragen zum weiteren Verständnis der *Mismatch*-Toleranz von *CRISPR-Cas9* bei.

6 Summary

Within the application of CRISPR-Cas lies enormous potential. It is currently being used in the treatment of genetic disease or malignant cancers and in medical diagnostics. Therefore there is great interest in anticipating and reducing potential side effects. One major concern is the mismatch tolerance of CRISPR-Cas. It can bind and modify DNA sequences that differ from the constructed recognition sequence. This poses the risk of modifying unintended gene regions, especially since the mechanisms of mismatch tolerance are still incompletely understood.

The goal of this work was to evaluate the mismatch tolerance of *gRNA-Cas9* within eukaryotic HEK 293T cells. Therefore a BRET-Reporter was used, which made use of the mechanism called Bioluminescence Resonance Energy Transfer. Thus the activity of *gRNA-Cas9* could be quantified by measuring the reduction of said energy transfer. Thus statements could be made about the sequence determining the tolerance of a mismatch. The constructed mismatches differed in single or multiple nucleotids from the wildtyp sequence.

It was shown that the tolerance of a specific mismatch depends on both its position and its nucleotide identity. Intrapositional differences were observed for the interaction of different nucleotids. Additionally, there were interpositional differences regardless of the nucleotide identity. Furthermore, the combinatorial effect of combining two singular mismatches was tested. With one exception a noticeable combinatorial effect became apparent. This indicates that the tolerance of a mismatch tends to decrease with an increasing number of mismatches. The obtained data contributes to a better understanding of the sequence determinants of mismatch tolerance in CRISPR-Cas9-systems.

7 Abbildungsverzeichnis

Abbildung 1: Plasmidkarte <i>pRLuc8-AvrII/BsiWI-GFP2</i>	19
Abbildung 2: Plasmidkarte <i>px459</i>	20
Abbildung 3: Übersichtsschema eines <i>Target</i> -Heteroduplexes	21
Abbildung 4: Übersichtsschema der Klonierungsarbeiten	37
Abbildung 5: Chromatogramm einer Sequenzierung	38
Abbildung 6: Übersicht der Transfektion in <i>HEK 293T</i> -Zellen.....	39
Abbildung 7: Fluoreszenzmikroskopische Kontrolle der Transfektion.....	40
Abbildung 8: Schema der Beschreibung von Nukleotidinteraktionen	41
Abbildung 9: <i>Frameshift rates</i> der singulären <i>Mismatches</i> von ART t1	42
Abbildung 10: <i>Target</i> -Sequenzen und farbcodierte <i>FSRs</i> der singulären <i>Mismatches</i> von ART t1	44
Abbildung 11: <i>Target</i> -Sequenzen und farbcodierte <i>FSRs</i> der <i>Mismatches</i> anderer Wildtypen.....	45
Abbildung 12: <i>Frameshift rates</i> der Doppel- <i>Mismatches</i> von ART t1	46
Abbildung 13: <i>Target</i> -Sequenzen und farbcodierte <i>FSRs</i> der Doppel- <i>Mismatches</i> von ART t1	47

8 Tabellenverzeichnis

Tabelle 1: Verwendete Geräte	16
Tabelle 2: Verwendete Chemikalien.....	17
Tabelle 3: Verwendete Restriktionsenzyme	20
Tabelle 4: Weitere Enzyme	21
Tabelle 5: Wildtypsequenzen der <i>Targets</i> und zugehörige <i>gRNA</i> -Sequenzen	22
Tabelle 6: Oligonukleotide der <i>Mismatch</i> -Sequenzen.....	23
Tabelle 7: Zur Sequenzierung genutzter Primer	25
Tabelle 8: Verwendete Bakterienstämme	26
Tabelle 9: Verwendete Antibiotika.....	26
Tabelle 10: Kulturmedien für Bakterien	27
Tabelle 11: Kulturmedien für eukaryotische Zellen	27
Tabelle 12: Verwendete Puffer.....	28
Tabelle 13: Kommerziell erhältliche Puffer.....	28
Tabelle 14: Verwendete Kit-Systeme	28
Tabelle 15: Zusammensetzung des Restriktionsverdau.....	29
Tabelle 16: Nachverdau des Ligationsansatzes.....	30
Tabelle 17: Zusammensetzung der Oligonukleotidhybridisierung	30
Tabelle 18: Ansatz der Ligation.....	30
Tabelle 19: Ansatz der externen Sequenzierung	31
Tabelle 20: Ansatz der Transfektion mit Polyethylenimin (PEI)	34
Tabelle 21: Auflistung der <i>BRET-Ratios</i>	77

9 Abkürzungsverzeichnis

A	Adenin
AG	Aktiengesellschaft
a. M.	Am Main
AMP-R	Ampicillinresistenz
ATCC	American Type Culture Collection
AZ	Arizona
<i>BLM/DNA2</i>	<i>Bloom Helicase / DNA2 Helicase</i>
<i>BRET</i>	<i>Bioluminescence Resonance Energy Transfer</i>
bzw.	Beziehungsweise
C	Cytosin
©	<i>copyright</i>
ca.	Circa
<i>Cas</i>	<i>CRISPR associated</i>
<i>Cas9</i>	<i>CRISPR associated protein 9</i>
<i>CBA</i>	<i>Chicken Beta Actin Promoter</i>
CH	Conföderatio Helvetica
CMV	Cytomegalievirus
CO ₂	Kohlenstoffdioxid
<i>CRISPR</i>	<i>Clustered Regularly Interspaced Short Palindromic Repeats</i>
<i>CRISPRa</i>	<i>Clustered Regularly Interspaced Short Palindromic Repeats activation</i>
<i>CRISPRi</i>	<i>Clustered Regularly Interspaced Short Palindromic Repeats interference</i>
<i>crRNA</i>	<i>CRISPR RNA</i>
CTD	C-Terminale-Domäne von Cas9
<i>CtIP</i>	<i>C-terminal binding protein interacting protein</i>
<i>dCas9</i>	<i>Dead Cas9, katalytisch inaktives Cas9</i>
ddH ₂ O	Doppelt destilliertes Wasser
DMEM	Dulbecco's Modified Eagles Medium
<i>DNA</i>	<i>Deoxyribonucleic acid</i>
<i>DNA-PKc</i>	<i>DNA-dependent protein kinase catalytic subunit</i>
DSB	Doppelstrangbruch
<i>dsDNA</i>	<i>Double stranded DNA</i>
<i>E.Coli</i>	Escherichia Coli Bakterium

<i>Exo1</i>	<i>Exonuclease 1</i>
<i>F</i>	<i>Forward-Oligonukleotid</i>
FKS	Fötale Kälberserum
FR	Frankreich
<i>FSR</i>	<i>Frameshift rate</i>
<i>FSRs</i>	<i>Frameshift rates</i>
G	Guanin
g (Zahl + g)	Gramm
g (Name + g)	Guide
<i>GFP</i>	<i>Green Fluorescent Protein</i>
<i>gRNA</i>	<i>Guide RNA</i>
<i>gRNA-Cas9</i>	Komplex aus <i>Cas9</i> und <i>gRNA</i>
h	Stunde(n)
<i>HDR</i>	<i>Homology Directed Repair</i>
<i>HEK</i>	<i>Human Embryonic Kidney</i>
HEPES	N-2-Hydroxyethyl-piperazin-N'-2-ethansulfonsäure
<i>HNH</i>	<i>HNH-Endonukleasedomäne von Cas9</i>
i.d.R.	In der Regel
Indel	Insertion und/oder Deletion
JLU	Justus-Liebig-Universität Gießen
KCl	Kaliumchlorid
KH ₂ PO ₄	Kaliumdihydrogenphosphat
LB	Luria Broth
LLC	<i>Limited Liability Company</i>
MgCl ₂	Magnesiumchlorid
MgSO ₄	Magnesiumsulfat
mg	Milligramm
ml	Milliliter
mM	Millimolar
<i>MMEJ</i>	<i>Microhomology mediated endjoining</i>
<i>MRN</i>	<i>MRN-Komplex bestehend aus MRE11-RAD50-NBS1</i>
<i>mRNA</i>	<i>Messenger RNA</i>
ms	Millisekunden
N	Beliebiges Nukleotid
NaCl	Natriumchlorid

Na ₂ HPO ₄	Dinatriumhydrogenphosphat
<i>nCas</i>	<i>Cas9</i> mit nur einer funktionalen Endonukleasedomäne, auch <i>Nickase</i> genannt
ng	Nanogramm
nm	Nanometer
nt	Nukleotide
<i>NTS</i>	<i>Non-Target-Strang</i>
OD	Optische Dichte
P	Position eines Nukleotids innerhalb der <i>Target</i> - oder <i>Guide</i> -Sequenz
<i>PAM</i>	<i>Protospacer Adjacent Motif</i>
<i>PCNA</i>	<i>Proliferation cell nuclear antigen</i>
PEI	Polyethylenimin
pmol	Pikomol
Pol μ	<i>DNA</i> -Polymerase mu
Pol λ	<i>DNA</i> -Polymerase delta
Pol θ	<i>DNA</i> -Polymerase theta
Puro-R	Puromycinresistenz
<i>R</i>	<i>Reverse-Oligonukleotid</i>
®	Registered
<i>REC 1-3</i>	<i>Recognition lobe</i> von <i>Cas9</i>
<i>RFC 1-5</i>	<i>Replication factor c 1-5</i>
<i>RLuc8</i>	Nach Loening et al. modifizierte Luziferase von <i>Renilla Reniformis</i>
<i>RNA</i>	<i>Ribonucleic acid</i>
RNP	Ribonukleoprotein
rN:dN	Interaktion zwischen zwei Nukleotiden der gRNA und der Target-DNA
<i>rpm</i>	<i>Rounds per minute</i>
<i>RuvC</i>	<i>RuvC</i> -Endonukleasedomäne von <i>Cas9</i>
s	Sekunde(n)
<i>SpCas9</i>	<i>Streptococcus pyogenes-Cas9</i>
<i>ssDNA</i>	<i>Single stranded DNA</i>
STABW	Standardabweichung
T	Thymin
™	<i>Trade mark</i>
<i>tracrRNA</i>	<i>Trans-activating crRNA</i>

<i>TS</i>	<i>Target-Strang</i>
u.a.	Unter anderem
<i>USA</i>	<i>United States of America</i>
V	Volt
v/v	Volume per volume
x g	Vielfaches der Erdbeschleunigung
<i>XRCC</i>	<i>X-ray cross complementing 4</i>
<i>XLF</i>	<i>XRCC4-like factor</i>
Zahl x	Vielfaches, z.B. fünffach für 5x
z.B.	Zum Beispiel
µg	Mikrogramm
µl	Mikroliter
°C	Grad Celsius
3'	3'-Ende eines Nukleotids
5'	5'-Ende eines Nukleotids
↓ und ↑	Schnittstellen der Restriktionsenzyme
±	Plus-Minus
%	Prozent
→	N → N, Nukleotidwechsel zur Generierung eines Mismatches

10 Literaturverzeichnis

- Ali, A., Xiao, W., Babar, M.E. and Bi, Y. 2022. Double-Stranded Break Repair in Mammalian Cells and Precise Genome Editing. *Genes (Basel)* 13(5). doi: 10.3390/genes13050737.
- Allen, F. et al. 2019. Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nature Biotechnology* 37(1), pp. 64–82. doi: 10.1038/nbt.4317.
- Anders, C., Bargsten, K. and Jinek, M. 2016. Structural Plasticity of PAM Recognition by Engineered Variants of the RNA-Guided Endonuclease Cas9. *Molecular Cell* 61(6), pp. 895–902. doi: 10.1016/j.molcel.2016.02.020.
- Anders, C., Niewoehner, O., Duerst, A. and Jinek, M. 2014. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513(7519), pp. 569–573. doi: 10.1038/nature13579.
- Anzalone, A. V. et al. 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576(7785), pp. 149–157. doi: 10.1038/s41586-019-1711-4.
- Aparicio, T., Baer, R. and Gautier, J. 2014. DNA double-strand break repair pathway choice and cancer. *DNA Repair (Amst.)* 19, pp. 169–175. doi: 10.1016/j.dnarep.2014.03.014.
- Barrangou, R. et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819), pp. 1709–1712. doi: 10.1126/SCIENCE.1138140.
- Beer. 1852. Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten. *Annalen der Physik* 162(5), pp. 78–88. doi: 10.1002/ANDP.18521620505.
- Bétermier, M., Bertrand, P. and Lopez, B.S. 2014. Is Non-Homologous End-Joining Really an Inherently Error-Prone Process? *PLoS Genetics* 10(1). doi: 10.1371/journal.pgen.1004086.
- Black, S.J. et al. 2019. Molecular basis of microhomology-mediated end-joining by purified full-length Pol θ . *Nature Communications* 10(1). doi: 10.1038/s41467-019-12272-9.
- Blattner, G., Cavazza, A., Thrasher, A.J. and Turchiano, G. 2020. Gene Editing and Genotoxicity: Targeting the Off-Targets. *Frontiers in Genome Editing* 2(613252). doi: 10.3389/FGEED.2020.613252.

- Boyle, E.A. et al. 2017. High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. *Proceedings of the National Academy of Sciences of the United States of America* 114(21), pp. 5461–5466. doi: 10.1073/pnas.1700557114.
- Boyle, E.A., Becker, W.R., Bai, H.B., Chen, J.S., Doudna, J.A. and Greenleaf, W.J. 2021. Quantification of Cas9 binding and cleavage across diverse guide sequences maps landscapes of target engagement. *Science Advances* 7(8). doi: 10.1126/sciadv.abe5496.
- Bravo, J.P.K. et al. 2022. Structural basis for mismatch surveillance by CRISPR–Cas9. *Nature* 603(7900), pp. 343–347. doi: 10.1038/S41586-022-04470-1.
- Brinkman, E.K., Chen, T., Amendola, M. and Van Steensel, B. 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research* 42(22). doi: 10.1093/nar/gku936.
- Cameron, P. et al. 2017. Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nature Methods* 14(6), pp. 600–606. doi: 10.1038/nmeth.4284.
- Canaj, H. et al. 2019. Deep profiling reveals substantial heterogeneity of integration outcomes in CRISPR knock-in experiments. *bioRxiv*. doi: 10.1101/841098.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. and Prasher, D.C. 1994. Green Fluorescent Protein as a Marker for Gene Expression. *Science* 263(5148), pp. 802–805. doi: 10.1126/science.8303295.
- Chen, B. et al. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155(7), pp. 1479–1491. doi: 10.1016/j.cell.2013.12.001.
- Chen, J.S. et al. 2017. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550(7676), pp. 407–410. doi: 10.1038/nature24268.
- Dacres, H., Michie, M., Wang, J., Pflieger, K.D.G. and Trowell, S.C. 2012. Effect of enhanced Renilla luciferase and fluorescent protein variants on the Förster distance of Bioluminescence resonance energy transfer (BRET). *Biochemical and Biophysical Research Communications* 425(3), pp. 625–629. doi: 10.1016/j.bbrc.2012.07.133.
- Dagdaz, Y.S., Chen, J.S., Sternberg, S.H., Doudna, J.A. and Yildiz, A. 2017. A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Science Advances* 3(8). doi: 10.1126/sciadv.aao0027.

- Das, A., Rai, J., Roth, M.O., Shu, Y., Medina, M.L., Barakat, M.R. and Li, H. 2023. Coupled catalytic states and the role of metal coordination in Cas9. *Nature Catalysis* 6(10), pp. 969–977. doi: 10.1038/s41929-023-01031-1.
- Doench, J. et al. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology* 34(2), pp. 184–191. doi: 10.1038/nbt.3437.
- Dong, D. et al. 2017. Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature* 546(7658), pp. 436–439. doi: 10.1038/nature22377.
- Doudna, J.A. 2020. The promise and challenge of therapeutic genome editing. *Nature* 578(7794), pp. 229–236. doi: 10.1038/s41586-020-1978-5.
- Dragulescu-Andrasi, A., Chan, C.T., De, A., Massoud, T.F., Gambhir, S.S. and Phelps, M.E. 2011. Bioluminescence resonance energy transfer (BRET) imaging of protein-protein interactions within deep tissues of living subjects. *Proceedings of the National Academy of Sciences of the United States of America* 108(29), pp. 12060–12065. doi: 10.1073/pnas.1100923108.
- Eslami-Mossallam, B. et al. 2022. A kinetic model predicts SpCas9 activity, improves off-target classification, and reveals the physical basis of targeting fidelity. *Nature Communications* 13(1). doi: 10.1038/S41467-022-28994-2.
- Fu, R. et al. 2022. Systematic decomposition of sequence determinants governing CRISPR/Cas9 specificity. *Nature Communications* 13(1). doi: 10.1038/s41467-022-28028-x.
- Gong, S., Yu, H.H., Johnson, K.A. and Taylor, D.W. 2018. DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity. *Cell Reports* 22(2), pp. 359–371. doi: 10.1016/j.celrep.2017.12.041.
- Grohmann, L. et al. 2019. Detection and identification of genome editing in plants: Challenges and opportunities. *Frontiers in Plant Science* 10:236. doi: 10.3389/FPLS.2019.00236.
- Hille, F., Richter, H., Wong, S.P., Bratovič, M., Ressel, S. and Charpentier, E. 2018. The Biology of CRISPR-Cas: Backward and Forward. *Cell* 172(6), pp. 1239–1259. doi: 10.1016/j.cell.2017.11.032.
- Hsu, P.D. et al. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology* 31(9), pp. 827–832. doi: 10.1038/nbt.2647.

- Huston, N.C., Tycko, J., Tillotson, E.L., Wilson, C.J., Myer, V.E., Jayaram, H. and Steinberg, B.E. 2019. Identification of Guide-Intrinsic Determinants of Cas9 Specificity. *The CRISPR Journal* 2(3), pp. 172–185. doi: 10.1089/crispr.2019.0009.
- Ivanov, I., Addison V. Wright, Joshua C. Cofsky, Kevin D. Palacio Aris, Jennifer A. Doudna and Zev Bryant. 2020. Cas9 interrogates DNA in discrete steps modulated by mismatches and supercoiling. *Proceedings of the National Academy of Sciences of the United States of America* 117(11), pp. 5853–5860. doi: 10.1073/pnas.1913445117.
- Iyer, S. et al. 2019. Precise therapeutic gene correction by a simple nuclease-induced double-stranded break. *Nature* 568(7753), pp. 561–565. doi: 10.1038/s41586-019-1076-8.
- Jansen, R., Van Embden, J.D.A., Gaastra, W. and Schouls, L.M. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* 43(6), pp. 1565–1575. doi: 10.1046/j.1365-2958.2002.02839.x.
- Jiang, F. et al. 2016. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* 351(6275), pp. 867–871. doi: 10.1126/science.aad8282.
- Jiang, F. and Doudna, J.A. 2017. CRISPR-Cas9 Structures and Mechanisms. *Annual Review of Biophysics* 46, pp. 505–529. doi: 10.1146/ANNUREV-BIOPHYS-062215-010822.
- Jinek, M. et al. 2014. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* 343(6176). doi: 10.1126/science.1247997.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. 2012. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337(6096), pp. 816–821. doi: 10.1126/science.1225829.
- Kaminski, M.M., Abudayyeh, O.O., Gootenberg, J.S., Zhang, F. and Collins, J.J. 2021. CRISPR-based diagnostics. *Nature Biomedical Engineering* 5(7), pp. 643–656. doi: 10.1038/s41551-021-00760-7.
- Kim, D., Luk, K., Wolfe, S.A. and Kim, J.-S. 2019. Evaluating and Enhancing Target Specificity of Gene-Editing Nucleases and Deaminases. *Annual Review of Biochemistry* 88, pp. 191–220. doi: 10.1146/annurev-biochem-013118-111730.
- Kim, H.K. et al. 2020a. High-throughput analysis of the activities of xCas9, SpCas9-NG and SpCas9 at matched and mismatched target sequences in human cells. *Nature Biomedical Engineering* 4(1), pp. 111–124. doi: 10.1038/s41551-019-0505-1.

- Kim, N. et al. 2020b. Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nature Biotechnology* 38(11), pp. 1328–1336. doi: 10.1038/s41587-020-0537-9.
- Kimsey, I.J. et al. 2018. Dynamic basis for dG· dT misincorporation via tautomerization and ionization. *Nature* 554(7691), pp. 195–201. doi: 10.1038/nature25487.
- Kimsey, I.J., Petzold, K., Sathyamoorthy, B., Stein, Z.W. and Al-Hashimi, H.M. 2015. Visualizing transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature* 519(7543), pp. 315–320. doi: 10.1038/nature14227.
- Klein, M., Eslami-Mossallam, B., Arroyo, D.G. and Depken, M. 2018. Hybridization Kinetics Explains CRISPR-Cas Off-Targeting Rules Engineered systems can increase specificity without losing on-target efficiency. *Cell Reports* 22(6), pp. 1413–1423. doi: 10.1016/j.celrep.2018.01.045.
- Kobayashi, H., Picard, L.-P., Schönege, A.-M. and Bouvier, M. 2019. Bioluminescence resonance energy transfer-based imaging of protein-protein interactions in living cells. *Nature Protocols* 14(4), pp. 1084–1107. doi: 10.1038/s41596-019-0129-7.
- Koonin, E. V. and Makarova, K.S. 2017. Mobile genetic elements and evolution of crispr-cas systems: All the way there and back. *Genome Biology and Evolution* 9(10), pp. 2812–2825. doi: 10.1093/gbe/evx192.
- Kosicki, M., Allen, F., Steward, F., Tomberg, K., Pan, Y. and Bradley, A. 2022. Cas9-induced large deletions and small indels are controlled in a convergent fashion. *Nature Communications* 13(1). doi: 10.1038/s41467-022-30480-8.
- Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S. and Qi, L.S. 2013. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* 8(11), pp. 2180–2196. doi: 10.1038/nprot.2013.132.
- Lee, J.K. et al. 2018. Directed evolution of CRISPR-Cas9 to increase its specificity. *Nature Communications* 9(1). doi: 10.1038/s41467-018-05477-x.
- Lin-Shiao, E. et al. 2022. CRISPR-Cas9-mediated nuclear transport and genomic integration of nanostructured genes in human primary cells. *Nucleic Acids Research* 50(3), pp. 1256–1268. doi: 10.1093/nar/gkac049.
- Listgarten, J. et al. 2018. Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nature Biomedical Engineering* 2(1), pp. 38–47. doi: 10.1038/s41551-017-0178-6.

- Liu, G., Lin, Q., Jin, S. and Gao, C. 2022. The CRISPR-Cas toolbox and gene editing technologies. *Molecular Cell* 82(2), pp. 333–347. doi: 10.1016/j.molcel.2021.12.002.
- Loening, A.M., Fenn, T.D., Wu, A.M. and Gambhir, S.S. 2006. Consensus guided mutagenesis of Renilla luciferase yields enhanced stability and light output. *Protein Engineering, Design and Selection* 19(9), pp. 391–400. doi: 10.1093/protein/gz1023.
- Longo, P.A., Kavran, J.M., Kim, M.-S. and Leahy, D.J. 2013. Transient Mammalian Cell Transfection with Polyethylenimine (PEI). *Methods Enzymol* 529, pp. 227–240. doi: 10.1016/B978-0-12-418687-3.00018-5.
- Ma, J., Bai, L. and Wang, M.D. 2013. Transcription under Torsion. *Science* 340(6140), pp. 1580–1583. doi: 10.1126/science.1235441.
- Mali, P. et al. 2013. RNA-Guided Human Genome Engineering via Cas9. *Science* 339(6121), pp. 823–826. doi: 10.1126/science.1232033.
- Marino, N.D., Pinilla-Redondo, R., Csörgő, B. and Bondy Denomy, J. 2020. Anti-CRISPR protein applications: natural brakes for CRISPR Cas technologies. *Nature Methods* 17(5), pp. 471–479. doi: 10.1038/s41592-020-0771-6.
- Martin, F., Sánchez-Hernández, S., Gutiérrez-Guerrero, A., Pinedo-Gomez, J. and Benabdellah, K. 2016. Biased and unbiased methods for the detection of off-target cleavage by CRISPR/Cas9: An overview. *International Journal of Molecular Sciences* 17(9). doi: 10.3390/ijms17091507.
- Mitchell, B.P. et al. 2020. Spontaneous Embedding of DNA Mismatches Within the RNA:DNA Hybrid of CRISPR-Cas9. *Frontiers in Molecular Biosciences* 7:39. doi: 10.3389/fmolb.2020.00039.
- Naeem, M., Majeed, S., Hoque, M.Z. and Ahmad, I. 2020. Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. *Cells* 9(7). doi: 10.3390/cells9071608.
- Pacesa, M. et al. 2022a. Structural basis for Cas9 off-target activity. *Cell* 185(22), pp. 4067–4081.e21. doi: 10.1016/j.cell.2022.09.026.
- Pacesa, M., Loeff, L., Querques, I., Muckenfuss, L.M., Sawicka, M. and Jinek, M. 2022b. R-loop formation and conformational activation mechanisms of Cas9. *Nature* 609(7925), pp. 191–196. doi: 10.1038/s41586-022-05114-0.

- Palermo, G. et al. 2018. Key role of the REC lobe during CRISPR-Cas9 activation by “sensing”, “regulating”, and “locking” the catalytic HNH domain. *Q Rev Biophys* 51:e91. doi: 10.1017/S0033583518000070.
- Perez-Pinera, P. et al. 2013. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods* 10(10), pp. 973–976. doi: 10.1038/nmeth.2600.
- Porto, E.M., Komor, A.C., Slaymaker, I.M. and Yeo, G.W. 2020. Base editing: advances and therapeutic opportunities. *Nature Reviews Drug Discovery* 19(12), pp. 839–859. doi: 10.1038/s41573-020-0084-6.
- Ran, F.A. et al. 2013a. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154(6), pp. 1380–1389. doi: 10.1016/j.cell.2013.08.021.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F. 2013b. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 8(11), pp. 2281–2308. doi: 10.1038/nprot.2013.143.
- Ricci, C.G., Chen, J.S., Miao, Y., Jinek, M., Doudna, J.A., McCammon, J.A. and Palermo, G. 2019. Deciphering Off-Target Effects in CRISPR-Cas9 through Accelerated Molecular Dynamics. *ACS Central Science* 5(4), pp. 651–662. doi: 10.1021/acscentsci.9b00020.
- Roy, A., Panigrahi, S., Bhattacharyya, M. and Bhattacharyya, D. 2008. Structure, stability, and dynamics of canonical and noncanonical base pairs: Quantum chemical studies. *Journal of Physical Chemistry B* 112(12), pp. 3786–3796. doi: 10.1021/jp076921e.
- Rozners, E. 2022. Chemical Modifications of CRISPR RNAs to Improve Gene Editing Activity and Specificity. *J Am Chem Soc* 144(28), pp. 12584–12594. doi: 10.1021/jacs.2c02633.
- Saifaldeen, M., Al-Ansari, D.E., Ramotar, D. and Aouida, M. 2020. CRISPR FokI Dead Cas9 System: Principles and Applications in Genome Engineering. *Cells* 9(11). doi: 10.3390/cells9112518.
- Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.I.T. and Yamamoto, T. 2016. MMEJ-Assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nature Protocols* 11(1), pp. 118–133. doi: 10.1038/nprot.2015.140.
- Scott, S. et al. 2018. Visualizing structure-mediated interactions in supercoiled DNA molecules. *Nucleic Acids Research* 46(9), pp. 4622–4631. doi: 10.1093/nar/gky266.

- Semenova, E. et al. 2011. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proceedings of the National Academy of Sciences of the United States of America* 108(25), pp. 10098–10103. doi: 10.1073/pnas.1104144108.
- Sfeir, A. and Symington, L.S. 2015. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends in Biochemical Sciences* 40(11), pp. 701–714. doi: 10.1016/j.tibs.2015.08.006.
- Shen, M.W. et al. 2018. Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* 563(7733), pp. 646–651. doi: 10.1038/s41586-018-0686-x.
- Shou, J., Li, J., Liu, Y. and Wu, Q. 2018. Precise and Predictable CRISPR Chromosomal Rearrangements Reveal Principles of Cas9-Mediated Nucleotide Insertion. *Molecular Cell* 71(4), pp. 498-509.e4. doi: 10.1016/j.molcel.2018.06.021.
- Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X. and Zhang, F. 2016. Rationally engineered Cas9 nucleases with improved specificity. *Science* 351(6268), pp. 84–88. doi: 10.1126/SCIENCE.AAD5227.
- Slaymaker, I.M. and Gaudelli, N.M. 2021. Engineering Cas9 for human genome editing. *Current Opinion in Structural Biology* 69, pp. 86–98. doi: 10.1016/j.sbi.2021.03.004.
- Sternberg, S.H., LaFrance, benjamin, Kaplan, matias and Doudna, J.A. 2015. Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* 527(7576), pp. 110–113. doi: 10.1038/nature15544.
- Stinson, B.M., Moreno, A.T., Walter, J.C. and Loparo, J.J. 2020. A Mechanism to Minimize Errors during Non-homologous End Joining. *Molecular Cell* 77(5), pp. 1080-1091.e8. doi: 10.1016/j.molcel.2019.11.018.
- Szczelkun, M.D. et al. 2014. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences of the United States of America* 111(27), pp. 9798–9803. doi: 10.1073/PNAS.1402597111.
- Taleei, R. and Nikjoo, H. 2013. Biochemical DSB-repair model for mammalian cells in G1 and early S phases of the cell cycle. *Mutation Research - Genetic Toxicology and Environmental Mutagenesis* 756(1–2), pp. 206–212. doi: 10.1016/j.mrgentox.2013.06.004.

- Tsai, S.Q. et al. 2015. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature Biotechnology* 33(2), pp. 187–198. doi: 10.1038/nbt.3117.
- Wang, H. et al. 2020. Development of a Self-Restricting CRISPR-Cas9 System to Reduce Off-Target Effects. *Molecular Therapy - Methods and Clinical Development* 18:18, pp. 390–401. doi: 10.1016/j.omtm.2020.06.012.
- Wang, X. et al. 2015. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nature Biotechnology* 33(2), pp. 175–179. doi: 10.1038/nbt.3127.
- Wettengel, J.M. et al. 2023. A Multifunctional and Highly Adaptable Reporter System for CRISPR/Cas Editing. *International Journal of Molecular Sciences* 24(9). doi: 10.3390/ijms24098271.
- Wienert, B., Cromer, M.K., Hendel, A., Cathomen, T., Fine, E.J. and Wagner, D.L. 2022. CRISPR nuclease off-target activity and mitigation strategies. *Frontiers in Genome Editing* 4. doi: 10.3389/fgeed.2022.1050507.
- Wimmer T. 2018. *Entwicklung eines Systems zur Kontrolle des "Vascular Endothelial Growth Factor" Spiegels im Auge*. doi: 10.22029/jlupub-14646.
- Wimmer, T. et al. 2021. A Bioluminescence Resonance Energy Transfer-Based Reporter System: Characterization and Applications. *CRISPR Journal* 4(6), pp. 884–895. doi: 10.1089/crispr.2021.0023.
- Wu, X. et al. 2014. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnology* 32(7), pp. 670–676. doi: 10.1038/nbt.2889.
- Wu, Y. and Jiang, T. 2022. micromachines Review Developments in FRET-and BRET-Based Biosensors. *Micromachines (Basel)* 13(10). doi: 10.3390/mi13101789.
- Xiang, X. et al. 2021. Enhancing CRISPR-Cas9 gRNA efficiency prediction by data integration and deep learning. *Nature Communications* 12(1). doi: 10.1038/s41467-021-23576-0.
- Xue, C. and Greene, E.C. 2021. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. *Trends in Genetics* 37(7), pp. 639–656. doi: 10.1016/j.tig.2021.02.008.

Yang, H. et al. 2020. Methods favoring homology-directed repair choice in response to crispr/cas9 induced-double strand breaks. *International Journal of Molecular Sciences* 21(18). doi: 10.3390/ijms21186461.

Yanik, M. et al. 2018. Development of a Reporter System to Explore MMEJ in the Context of Replacing Large Genomic Fragments. *Molecular Therapy Nucleic Acids* 11, pp. 407–415. doi: 10.1016/j.omtn.2018.03.010.

Yin, H. et al. 2016. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nature Biotechnology* 34(3), pp. 328–333. doi: 10.1038/nbt.3471.

Zhang, J.P. et al. 2016. Different Effects of sgRNA Length on CRISPR-mediated Gene Knockout Efficiency. *Scientific Reports* 6:28566. doi: 10.1038/SREP28566.

Zhang, Q. et al. 2019. The post-PAM interaction of RNA-guided spCas9 with DNA dictates its target binding and dissociation. *Science Advances* 5(11). doi: 10.1126/sciadv.aaw9807.

Zhao, C. et al. 2019. Evaluation of the effects of sequence length and microsatellite instability on single-guide RNA activity and specificity. *International Journal of Biological Sciences* 15(12), pp. 2641–2653. doi: 10.7150/ijbs.37152.

Zheng, R. et al. 2023. Progress and Perspective of CRISPR-Cas9 Technology in Translational Medicine. *Advances Sciences (Weinh.)* 10(25). doi: 10.1002/advs.202300195.

Zhu, X. et al. 2019. Cryo-EM structures reveal coordinated domain motions that govern DNA cleavage by Cas9. *Nature Structural & Molecular Biology* 26(8), pp. 679–685. doi: 10.1038/s41594-019-0258-2.

11 Anhang

In Tabelle 21 finden sich die kumulierten Mittelwerte der *BRET-Ratios*. Zudem zu entnehmen sind die zugehörigen kumulierten Standardabweichungen (STABW), die errechneten frameshift rates (*FSRs*) und Signifikanzniveaus.

Tabelle 21: Auflistung der *BRET-Ratios*

transfizier-tes Plasmid	Mittelwert <i>BRET-Ratio</i>	STABW <i>BRET-Ratio</i>	<i>FSR</i> [%]	STABW <i>FSR</i> [%]	Signifi-kanzni-veau (T / T+G)
ART t1	0,222	0,0069	35,46	3,76	<0,0001
ART t1 + ART g1	0,153	0,0024			
M 1.1	0,252	0,0060	-1,37	-3,87	0,6664
M 1.1+ ART g1	0,2551	0,0061			
M 1.2	0,2177	0,0029	10,18	2,27	0,0205
M 1.2 + ART g1	0,1983	0,0032			
M 5	0,2466	0,0084	30,72	4,36	<0,0001
M 5 + ART g1	0,1793	0,0045			
M 6	0,2345	0,0041	35,83	2,84	<0,0001
M 6 + ART g1	0,1604	0,0042			
M 7	0,2351	0,0043	32,6	3,16	<0,0001
M 7 + ART g1	0,1674	0,0049			
M 8	0,24	0,0062	24,89	4,42	<0,0001
M 8 + ART g1	0,1871	0,0070			

M 9	0,2379	0,006	-1,95	3,3	0,2549
M 9 + ART g1	0,242	0,0034			
M 10	0,2408	0,0047	-1,42	2,76	0,3777
M 10 + ART g1	0,2438	0,0035			
M 11	0,2403	0,0052	16,54	2,9	<0,0001
M 11 + ART g1	0,2051	0,0033			
M 12	0,2401	0,0042	30,05	2,27	<0,0001
M 12 + ART g1	0,1762	0,0024			
M 13	0,2426	0,0047	33,24	3,44	<0,0001
M 13 + ART g1	0,1711	0,0057			
M 14	0,2442	0,0041	3,27	2,75	0,1362
M 14 + ART g1	0,2371	0,0044			
M 16	0,2281	0,0098	23,66	5,91	<0,0001
M 16 + ART g1	0,181	0,0065			
M 17	0,218	0,0060	28,92	3,85	<0,0001
M 17 + ART g1	0,1633	0,0041			
M 18	0,2196	0,0045	33,39	3,98	<0,0001
M 18 + ART g1	0,1559	0,0061			

M 19	0,2181	0,0063	20,93	5,06	<0,0001
M 19 + ART g1	0,1785	0,0072			
M 20	0,2078	0,0023	13,59	3,22	<0,0001
M 20 + ART g1	0,1835	0,0053			
M 21	0,2238	0,0047	17,76	4,11	<0,0001
M 21 + ART g1	0,1892	0,0065			
M 22	0,2219	0,0032	13,44	2,82	<0,0001
M 22 + ART g1	0,196	0,0044			
M 23	0,2189	0,0030	3,8297	3,48	0,0004
M 23 + ART g1	0,2116	0,0059			
M 24	0,2402	0,0058	5,75	5,04	0,0061
M 24 + ART g1	0,2281	0,0089			
M 25	0,2208	0,0023	0,72	2,64	0,2913
M 25 + ART g1	0,2194	0,0045			
M 26	0,2643	0,0074	-0,86	4,61	0,6707
M 26 + ART g1	0,2664	0,0079			
M 27	0,2583	0,0092	-2,74	5,01	0,5899
M 27 + ART g1	0,2646	0,0068			

M 28	0,2643	0,0060	0,47	4,03	0,6707
M 28 + ART g1	0,2632	0,0073			
M 29	0,2583	0,0045	26,62	2,8	<0,0001
M 29 + ART g1	0,1972	0,0045			
M 30	0,2658	0,0061	1,5	3,55	0,4938
M 30 + ART g1	0,2623	0,0058			
M 31	0,2583	0,0048	6,85	7,42	0,0006
M 31 + ART g1	0,2426	0,0163			
M 32	0,2703	0,0044	3,8	4,56	0,0449
M 32 + ART g1	0,2611	0,0101			
M 33	0,2641	0,0063	7,25	6,88	0,0015
M 33 + ART g1	0,2471	0,0149			
M 34	0,2642	0,0056	24,16	3,35	<0,0001
M 34 + ART g1	0,2074	0,0055			
M 35	0,277	0,0057	6,99	3,5	0,0025
M 35 + ART g1	0,2597	0,0066			

T4 #4461-4462	0,2463	0,0039	40,27	2,72	<0,0001
T 4 #4461-4462 + G 4 #4463-4464	0,1581	0,0045			
M 2	0,2424	0,0046	-2,23	4,2	0,5257
M 2 + G 4 #4463-4464	0,2472	0,0078			
T 11	0,0761	0,0015	52,33	4,49	<0,0001
T 11 + G 3	0,0506	0,0016			
M 3	0,2307	0,0061	-2,17	4,83	0,1782
M 3 + G 3	0,2351	0,0077			
ART t5	0,2596	0,005	44,47	2,88	<0,0001
ART t5 + ART g5	0,1563	0,0044			
M 4	0,2593	0,0048	2,07	3,19	0,4911
M 4 + ART g5	0,2545	0,0056			
M 15	0,2375	0,0041	37,41	3,21	<0,0001
M 15 + ART g5	0,1589	0,0054			
Luziferase (<i>RLuc8</i>)	0,0279	0,0036	-	-	-

12 Ehrenwörtliche Erklärung

„Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nichtveröffentlichten 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 sowie ethische, datenschutzrechtliche und tierschutzrechtliche Grundsätze befolgt. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, und dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt und indirekt an der Entstehung der vorliegenden Arbeit beteiligt waren. Mit der Überprüfung meiner Arbeit durch eine Plagiatserkennungssoftware bzw. ein internetbasiertes Softwareprogramm erkläre ich mich einverstanden.“

Ort, Datum

Unterschrift

13 Danksagung

Zunächst gilt mein besonderer Dank Herr Prof. Dr. Dr. Knut Stieger für die Überlassung des Themas dieser Arbeit, für die umfassende Betreuung und fachliche Unterstützung.

Ebenfalls herzlich bedanken möchte ich mich bei Herr Dr. Tobias Wimmer für das Einarbeiten im Labor und die Erläuterung der Methodik dieser Arbeit. Vielen Dank für den fachlichen Austausch, Problemlösungen und die Beantwortung zahlreicher Fragen.

Weiterhin vielen Dank an Frau Prof. em. Birgit Lorenz und Herr Lyubomyr Lytvynchuk MD, PhD für die Bereitstellung der Laborräumlichkeiten.

Herzlicher Dank gilt auch Frau Annabella Janise-Libawski und Frau Dr. Maria Weller für die Beantwortung von Fragen zu Laborabläufen.

Darüber hinaus gilt mein Dank allen weiteren Mitarbeiterinnen und Mitarbeitern der AG Experimentelle Ophthalmologie (AG Prof. Stieger). Vielen Dank für die tolle Atmosphäre und bereichernde Zeit.

Abschließend möchte ich meinen Eltern Mechthild und Frank für ihre Unterstützung danken. Vielen Dank für die finanzielle Ermöglichung des Medizinstudiums und dieser Arbeit. Vielen Dank auch an meine Schwester Pia und meine Freundin Lara für die Unterstützung in allen Lebenslagen.