Development of transposon-mediated gene delivery systems with improved safety for the use in gene and cell therapy

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich Biochemie, Chemie und Pharmazie

der Johann Wolfang Goethe-Universität

in Frankfurt am Main

von

Maximilian Amberger

aus Mainz

Frankfurt am Main 2022

(D 30)

Vom Fachbereich Biochemie, Chemie und Pharmazie (14) der

Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: Prof. Dr. Rolf Marschalek

                   Prof. Dr. Zoltán Ivics

Datum der Disputation: 28.11.2023

E R K L Ä R U N G

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im

Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Darmstadt, den …..

Versicherung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation über

"Development of transposon-mediated gene delivery systems with improved safety for the use in gene and cell therapy"

selbständig angefertigt und mich anderer Hilfsmittel als der in ihr

angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen

aus anderen Schriften mit Angabe der betreffenden Schrift

gekennzeichnet sind.

Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis

beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung

in Anspruch genommen zu haben.

Darmstadt, den ....

1. <u>Acknowledgements</u>

My deepest gratitude goes to the people that accompanied me throughout this project. It goes without saying that finishing such an endeavor would not have been possible without your support.

First and foremost I would like to thank Prof. Dr. Zoltán Ivics for giving me the opportunity to work on such an exciting project and providing the most supportive guidance I could have wished for. Your door was (literally) always open for when help was needed and the prioritization of your students was wholeheartedly felt at all times. I'm leaving with wonderful memories of exciting scientific discussions, but also deep conversations about the latest developments in the film industry and innumerable laughs while enjoying excellent food and a vast variety of hot and cold beverages.

Many thanks also to Prof. Dr. Rolf Marschalek for agreeing to provide university supervision. But beyond that for your genuine interest, taking your time to discuss my data with me engaging in exciting scientific discussions and providing insightful input, eye opening perspectives and valuable suggestions.

To Dr. Csaba Miskey I'm most thankful for providing the essential know-how about NGS, bioinformatics, everything lab-related and always being willing to help me out of often tricky situations with experimental work. But most importantly for turning me into a chili pepper aficionado with his willingness to share his remarkable knowledge and seeds to get me started.

To all my colleagues at PEI, thank you for making my stay so memorable. Matthias Ochmann for being the best office mate and especially for being away on holidays when we moved offices, giving me the choice for the best spot for my desk. Nicolas Sandoval Villegas for the Chilean humor, Spanish practice and making me laugh every time you insist on preferring Belgian over German beer. Wasifa Nurieva for your cheerfulness, the treasure chest of sweets and making sure I don't starve at work. Dr. Lacramioara Botezatu for being such great company in the cell culture lab and easing my fear of FACS, and Dr. Johanna Wildemann for all the fun times in and out of the PEI and giving me a wonderful reason to visit Berlin more often.

Additional gratitude goes to my family, extended family and friends for always being supportive, providing counsel as well as encouragement when needed and although the subject might have seemed confusing, never doubting that in the end everything would turn out right and I would be successful.

2. <u>**Table of contents**</u>

Table of contents

3. <u>**Summary**</u>

Nowadays, highly specialized genetic engineering tools are available, especially for the use in the medical context. This work has primarily focused on transposon-based genome engineering. Transposons are DNA segments that can change their position in the genome by a process known as transposition. Depending on the nature of their mobile phase (RNA or DNA), a distinction is made between class I and class II transposons. This work addressed class II transposons. These transpose as a DNA unit from one position to the next and follow a so-called cut-and-paste mechanism. Class II transposable elements consist of two essential components: Transposon-specific recognition sequences that flank the transposon and the transposase enzyme that recognizes the flanking sequences and catalyzes the mobilization from excision to integration at another position. In nature, autonomous DNA transposons encode their transposase between said recognition sequences and thus, upon successful transposition, take with them all necessary components for additional transposition steps. For modern biotechnology, this class of transposons has proven to be extremely useful because, as described, only two components are required for transposition. How these components occur is irrelevant as long as they come together. For this reason, transposase and recognition sequences can be administered separately from each other into a cell, and the sequence between the two recognition sequences can be replaced by any other. This makes it possible, for example, to stably integrate therapeutic genes into the genome of a host cell and thus to produce gene and cell therapy products. However, potential risk factors have to be taken into account. Since the position in which the transposon is integrated is not predetermined, but the transposase itself determines a suitable integration site, there is a risk that important genome components are interrupted and thus their function is affected with potentially serious consequences such as oncogenesis. Chapter I of this thesis has addressed the *piggyBac* transposon system. This widely used system is capable of efficiently performing the genome modifications just described and has already been used in clinical trials for the production of CAR T-cell products. However, the *piggyBac* transposase has the problematic property of preferentially integrating transposons near transcriptional start sites. This increases the likelihood of causing genotoxic effects, making it potentially dangerous to use as a tool for clinical application. It has been shown in the past that the *piggyBac* transposase shows physical interactions with BET proteins (e.g. BRD4) through Co-IP experiments. Representatives of these proteins are part of the transcriptional activation complex and are abundant at transcriptional start sites. Accordingly, it was previously proposed that this interaction mediates the biased integration preference

Summary

towards these positions. For the first chapter of this thesis, the goal was to inhibit this interaction and determine if this changes the integration preference. As part of a collaboration in the lead-up to this project, a region between V390 and P410 in the *piggyBac* structure was identified based on structural analysis, which was suspected to be responsible for the interaction. In this work, this entire structure was subjected to an alanine scan and the resulting mutants were analyzed for their interaction capacity through a series of Co-IP experiments with BRD4. This resulted in a total of four positions identified where substitution with alanine resulted in a significant reduction in interaction capacity with BRD4: F395A, F397A, L403A, and K407A by 61%, 90%, 70%, and 96%, respectively, compared with the wildtype control. However, in subsequent transposition assays, it was found that all generated mutants had lost their transposition activity. Since the role of said structure in the interaction with BRD4 was confirmed however by the Co-IP experiments, multiple sequence alignments with related members of the *piggyBac* superfamily and structural analysis obtained in collaboration were used to propose alternative substitutions that are either found in members at said positions or do not significantly alter the conformation of the structure but maximally affect the protein surface with charged and/or long side chains. The mutants were generated, subjected to transposition assays, and the active representatives were used in Co-IP experiments. In total, seven substitutions were identified (E380F, V390K, T392Y, M394R, K407C, K407Q, and K407V) which simultaneously exhibited transposition activity and showed reduced interaction capacity with BRD4 (reduction by 65%, 53%, 95%, 92%, 67%, 62%, and 97%, respectively). Each of the aforementioned mutants were used for further transposition experiments and integration libraries were generated from the genomic DNA of the resulting cell populations. It was determined via next-generation-sequencing, if the integration preferences of the respective mutants in comparison to the *piggyBac* wildtype and in relation to their proximity to transcription start sites had changed. Based on 1.6 to 3.2 thousand integration events identified per generated mutant, these integration preferences were assessed. In the immediate range 200 base pairs up- and downstream from known human transcriptional start sites, all mutants used exhibited a reduced integration bias. At a wider observation window 3000 base pairs up- and downstream from transcriptional start sites, further mutants with the substitutions M394R, T392Y and V390K showed a reduction in integration frequency of 17.3%, 1.5% and 5.4%, respectively, compared to the wildtype. Of particular note was the mutant with the M394R substitution, which showed a substantial reduction in all window sizes analyzed with a maximum of 65% less integration preference in the immediate vicinity of transcriptional start sites. Structural analyses showed that the side chain of M394 has no intermolecular interactions

Summary

with neighboring side chains, protruding at the surface of the _piggyBac_ molecule and thus consolidating the hypothesis of achieving the desired effect by surface modification with simultaneous preservation of the 3D structure. Thus, the goal of the first project was achieved. The substitution M394R theoretically generates a significant safety advantage over the wildtype transposase in terms of integration near transcriptional start sites in the tested cell system. Going further, it is suggested to generate mutants with combinations of the analyzed substitutions to determine potential additive properties and to achieve further eventual reductions in integration bias. In addition, it is recommended to compare the generated datasets with respect to integration preferences in additional potentially critical regions with the generated control dataset of the wildtype transposase to identify possible differences and thus further characterize the mutants.

Chapter II of this thesis was dedicated to the overall safety improvement for transposon-based gene modification and addressed the time point after the transgene has already been integrated and serious side effects may occur. In addition to the aforementioned risk of oncogenesis, other potentially lethal side effects such as cytokine release syndrome or graft-versus-host disease may occur within gene and cell therapies which are gaining attention. An elegant method to deal with the eventuality is to incorporate a so-called suicide-switch into the cell product. Thus, if conventional treatment methods do not resolve the side effects, the transgene cell product can be eliminated from the patient by activating the switch. Several representatives of suicide-switches are currently being tested for this purpose or are already in use. However, all systems developed so far have significant drawbacks such as incomplete elimination modified cells, incompatibility upon administration of the substances required for activation, or even severe side effects caused by these substances. With this in mind, the aim of this second chapter was to develop a novel suicide-switch that addresses these drawbacks, can be stably introduced into cells via transposition, and reliably leads to cell death of the modified cells once activated. A system based on CRISPR/Cas9 technology was developed and termed CRISPR Induced Suicide-Switch (CRISISS). CRISPR/Cas9 is a programmable nuclease that can find and generate double strand breaks at homologous DNA sequences to a predetermined and co-administered guide RNA sequence. This revolutionary technology is both highly precise and flexible due to its dependence on the guide RNA sequence. However, in this work CRISPR/Cas9 was used in a highly unconventional way. A sequence that guides the Cas9 nuclease to _Alu_ elements was chosen as the guide RNA. These are short, repetitive sequences, which are distributed over the human genome in more than one million copies and make up about 11% of our genetic material. If Cas9 causes double strand breaks within these elements,

Summary

they are too numerous to be repaired by DNA repair mechanisms, the genome becomes fragmented and the cell dies. However, in order to function as a suicide switch, inducibility is required so that it is activated exclusively when necessary, without background activity and thus causing unintended cell death. For this purpose, a transcriptional as well as post-translational control mechanism was added. Transcription of the Cas9 nuclease was regulated using a tet-on system so that transcription occurs only in the presence of doxycycline. Furthermore, a version of the Cas9 nuclease called arC9 was used, which, by fusion with an estrogen receptor ligand binding domain, adopts a conformation that prevents nuclease function and induces a conformational change that allows double strand break generation only in the presence of 4-Hydroxytamoxifen. Together with an expression cassette for the *Alu*-specific guide RNA, an expression cassette for the reverse tetracycline controlled transactivator, and a neomycin resistance gene expression cassette, all components were arranged between transposase-specific recognition sequences on a plasmid to allow integration of the construct into the host cell genome after introduction into host cells together with the respective transposase. The system was tested in HeLa cells. To confirm successful transposition, cells were selected with G418 for 14 days, generating a resistant cell population suspected to contain the suicide-switch. First, conditional-only expression of the arC9 nuclease was confirmed by addition of 1 µg/ml doxycycline through western blot analyses. Furthermore, the suicide-switch was induced by additional addition of 200 nM 4-Hydroxytamoxifen, protein extracts were extracted and assayed for the presence of phosphorylated KAP1 protein to confirm the induced DNA damage response. Only in the presence of doxycycline, 4-Hydroxytamoxifen, and *Alu*-specific guide RNA (cell populations containing all suicide cassette components except *Alu*-specific guide RNA were generated as a negative control) was phosphorylated KAP1 protein detected and as early as 24 hours. Flow cytometric assays for detection of Annexin-V, further confirmed the onset of apoptosis. Monoclonal cell lines were subsequently generated from both described conditional Cas9 expressing cell populations. These were used to perform extensive growth and survival experiments to determine the effect of suicide-switch induction on cell proliferation and survival. It was found that after induction, the cell line harboring an active switch lose their ability to divide and die. Between 24 and 48 hours after induction, a halt in cell division was detected, after which the dying process occurred. Within 5 days post induction, >99% of all cells were eliminated. In the absence of both inducers, no significant differences in survival were determined compared with the control cell line without *Alu*-specific guide RNAs. Microscopic examinations of the <1% surviving cell fraction revealed a senescence-associated phenotype and showed no signs of resumption of the cell division process. Thus, the second

Summary

chapter of this thesis also achieved its goal in developing a functional CRISPR/Cas9-based suicide-switch that can be inserted into human cells via transposition, is highly dependent in its activity on the necessary induction signals, and exhibits excellent elimination capabilities in the context tested. For the future, it is encouraged to confirm the efficiency and the function of the presented system in clinically relevant cells and to perform tests including a therapeutic gene. If successful, the possibility of further characterization in animal studies might be considered.

In summary, this work presents two independent approaches aiming to improve the safety of the rapidly growing array of technologies for the production and implementation of gene and cell therapeutics. Further development and testing of the presented systems is recommended to fully exploit their promising potential.

4. <u>**Zusammenfassung**</u>

Für die gezielte Genommodifizierung, vor allem im medizinischen Kontext, stehen heutzutage hochgradig spezialisierte Methoden zur Verfügung. Diese Arbeit hat sich in erster Linie mit Transposon-basierter Genommodifikation beschäftigt. Als Transposons werden DNA-Segmente bezeichnet, die durch einen als Transposition bezeichneten Prozess ihre Position im Genom verändern können. In Abhängigkeit der Natur ihrer mobilen Phase (RNA oder DNA) wird zwischen Klasse-I und Klasse-II Transposons unterschieden. Diese Arbeit widmete sich Klasse-II Transposons. Diese transponieren als DNA-Einheit von einer Position zur nächsten und folgen einem sogenannten cut-and-paste Mechanismus. Klasse-II Transposon-Einheitenen bestehen aus zwei essenziellen Komponenten: Transposon-spezifischen Erkennungssequenzen, welche das Transposon flankieren und dem Transposase-Enzym, welches die flankierenden Sequenzen erkennt und die Mobilisierung von Exzision bis Integration an einer anderen Position katalysiert. Autonome DNA-Transposons kodieren ihre Transposase in der Natur zwischen besagten Erkennungssequenzenen und nehmen somit bei erfolgreicher Transposition jegliche nötige Information für weitere Transpositionsschritte mit. Für die moderne Biotechnologie haben sich diese Klasse von Transposons als außerordentlich nützlich erwiesen, da wie beschrieben nur zwei Komponenten für die Transposition nötig sind. Wie diese Komponenten vorkommen, ist nicht von Bedeutung, solange sie zusammenkommen. Aus diesem Grund können Transposase und Erkennungssequenzen separat von einander in eine Zelle verabreicht werden, wobei die Sequenz zwischen den beiden Erkennungssequenzen durch eine beliebige andere ersetzt werden kann. Dadurch ist es möglich, beispielsweise therapeutische Gene stabil in das Genom einer Wirtszelle zu integrieren und somit Gen- und Zelltherapieprodukte herzustellen. Potentielle Risikofaktoren müssen jedoch berücksichtigt werden. Da die Position, in der das Transposon integriert wird, nicht vorbestimmt werden kann, sondern die Transposase selbst einen geeigneten Integrationsort bestimmt, besteht die Gefahr, dass wichtige Genomkomponenten unterbrochen und somit ihre Funktion beeinflusst werden mit potentiell gravierenden Folgen wie beispielsweise Onkogenese. Kapitel I dieser Arbeit hat sich dem *piggyBac* Transposon-System gewidmet. Dieses weitverbreitete System ist in der Lage, effizient die eben beschriebene Genommodifizierungen durchzuführen und wurde bereits im Rahmen von klinischen Studien für die Produktion von CAR-T-Zell-Produkten eingesetzt. Jedoch besitzt die *piggyBac* Transposase die problematische Eigenschaft, Transposons bevorzugt nahe Transkriptionsstartpunkten zu integrieren. Dies erhöht die Wahrscheinlichkeit, genotoxische Effekte zu verursachen, was den Einsatz als Werkzeug für klinische Anwendung erschwert. In der Vergangenheit wurde gezeigt, dass die *piggyBac* Transposase in Co-IP

Zusammenfassung

Experimenten physische Interaktionen mit BET-Proteinen (z.B. BRD4) aufweist. Vertreter dieser Proteine sind Teil des Transkriptions-Aktivierungskomplexes und kommen vermehrt nahe dieser Transkriptionsstartpunkte vor. Demnach wurde vorgeschlagen, dass diese Interaktion Grund für die voreingenommene Integrationspräferenz gegenüber diesen Positionen ist. Im ersten Kapitel dieser Arbeit wurde das Ziel verfolgt, diese Interaktion zu inhibieren und zu ermitteln, ob sich dadurch die Integrationspräferenz ändert. Im Rahmen einer Kooperation im Vorlauf dieses Projekts wurde auf Basis von Strukturanalysen eine Region zwischen V390 und P410 in der *piggyBac* Struktur identifiziert, welche vermutet wurde, für die Interaktion verantwortlich zu sein. In dieser Arbeit wurde diese gesamte Struktur einem Alanin-Scan unterzogen und die daraus entstehenden Mutanten mittels Co-IP Experimenten mit BRD4 auf ihre Interaktionskapazität analysiert. Daraus wurden insgesamt vier Positionen identifiziert, bei denen eine Substitution mit Alanin eine erhebliche Reduktion in der Interaktionskapazität mit BRD4 mit sich führte: F395A, F397A, L403A und K407A um jeweils 61%, 90%, 70% und 96% im Vergleich zu der Wildtyp-Kontrolle. In anschließenden Transpositionsassays wurde jedoch festgestellt, dass alle generierten Mutanten ihre Transpositionsaktivität verloren hatten. Da jedoch die Rolle der besagten Struktur in der Interaktion mit BRD4 durch die Co-IP Experimente bestätigt werden konnte, wurden anhand von multiplen Sequenzalignments mit verwandten Vertretern der *piggyBac* Superfamilie und durch eine in Kooperation erhaltene Strukturanalyse alternative Substitutionen vorgeschlagen, die entweder bei Vertretern an besagten Positionen zu finden sind, oder die Struktur in ihrer Konformation nicht maßgeblich verändern, jedoch die Proteinoberfläche mit Ladungen und/oder langen Seitenketten größtmöglich beeinflussen. Die Mutanten wurden generiert, Transpositionsassays unterzogen und die aktiven Vertreter in Co-IP Experimenten eingesetzt. Insgesamt wurden sieben Substitutionen identifiziert (E380F, V390K, T392Y, M394R, K407C, K407Q und K407V), welche zugleich Transpositionsaktivität aufwiesen und reduzierte Interaktionskapazität mit BRD4 zeigten (Reduktion um jeweils 65%, 53%, 95%, 92%, 67%, 62% und 97%). Mit jeder der genannten Mutanten wurden für weitere Transpositionsversuche eingesetzt und aus der genomischen DNA der resultierenden Zellpopulationen Integrations-Libraries angefertigt, um mittels Next-Generation-Sequencing die Integrationspräferenzen der jeweiligen Mutanten im Vergleich zu dem *piggyBac* Wildtyp und in Bezug auf ihre Nähe zu Transkriptionsstartpunkten zu bestimmen. Auf Basis von 1,6 bis 3,2 Tausend analysierten Integrationspositionen pro generierte Mutante wurden diese Integrationsfrequenzen bestimmt. Im unmittelbaren Bereich 200 Basenpaare auf- und abwärts von bekannten humanen Transkriptionsstartpunkten zeigten alle eingesetzten Mutanten eine deutlich reduzierte Integrationspräferenz. Bei einem größer

gezogene Beobachtungsfenster 3000 Basenpaare auf- und abwärts von Transkriptionsstartpunkten wiesen weiterhin Mutanten mit den Substitutionen M394R, T392Y und V390K eine Reduktion in der Integrationsfrequenz von jeweils 17,3%, 1,5% und 5,4% im Vergleich zum Wildtyp auf. Besonders hervorzuheben ist dabei die Mutante mit der Substitution M394R, welche in allen analysierten Fenstergrößen eine erhebliche Reduktion vorwies, mit einem Maximum von 65% weniger Integrationspräferenz im unmittelbaren Umfeld von Transkriptionsstartpunkten. Strukturanalysen zeigten, dass die Seitenkette von M394 keine intermolekularen Interaktionen mit benachbarten Seitenketten aufweist, an der Oberfläche des *piggyBac* Moleküls hinausragt und somit die Hypothese festigt, durch Oberflächenmodifikation mit gleichzeitiger Erhaltung der 3D-Struktur den erwünschten Effekt zu erzielen. Das Ziel des ersten Projektes wurde somit vollumfänglich erreicht. Die Substitution M394R generiert ein in Bezug auf die Integration nahe Transkriptionsstartpunkten in dem getesteten Zellsystem einen erheblichen Sicherheitsvorteil gegenüber der Wildtyptransposase. Weiterführend wird vorgeschlagen, Mutanten mit Kombinationen der ermittelten Substitutionen zu generieren, um potentiell additive Eigenschaften zu ermitteln und weitere eventuelle Reduktionen in der Integrationspräferenz zu erreichen. Zusätzlich wird empfohlen, die generierten Datensätze in Bezug auf Integrationspräferenzen in weiteren potentiell kritischen Regionen mit dem generierten Kontrolldatensatz der Wildtyptransposase zu vergleichen, um mögliche Unterschiede zu ermitteln und die Mutanten somit weiter zu charakterisieren.

Kapitel II dieser Arbeit hat sich der Verbesserung der allgemeinen Sicherheit transposonbasierter Genmodifikation gewidmet und den Zeitpunkt adressiert, nachdem das Transgen bereits integriert wurde und schwerwiegende Nebenwirkungen eintreten können. Neben der erwähnten Gefahr der Onkogenese können weitere potentiell tödliche Nebenwirkungen wie Zytokin-Freisetzungssyndrom oder Graft-versus-Host-Reaktion bei an Aufmerksamkeit gewinnenden Gen- und Zelltherapien auftreten. Eine elegante Methode für den Fall des Eintretens ist das Einbauen einer sogenannten Suizid-Kassette in das Zellprodukt. Somit kann, falls herkömmliche Behandlungsmethoden die Nebenwirkungen nicht beheben, das Transgene Zellprodukt durch Aktivierung der Suizid-Kassette aus dem Patienten eliminiert werden. Einige Vertreter von Suizid-Kassetten werden aktuell für diesen Zweck erprobt oder sind bereits im Einsatz. Jedoch weisen alle bislang entwickelten Systeme entscheidende Nachteile auf wie beispielsweise unvollständige Eliminierung aller modifizierten Zellen, Unverträglichkeit bei Verabreichung der für die Aktivierung nötigen Stoffe oder sogar schwere Nebenwirkung dieser. Mit diesem Hintergrund war das Ziel dieses zweiten Kapitels eine

Zusammenfassung

neuartige Suizid-Kassette zu entwickeln, die besagte Nachteile adressiert, mittels Transposition stabil in Wirtszellen eingebracht werden kann und zuverlässig zum Zelltod der modifizierten Zellen führt, sobald sie aktiviert wird. Es wurde ein System auf Basis von CRISPR/Cas9 Technologie entwickelt, welche mittels Transposition in das Genom von Wirtszellen eingefügt werden kann. Bei CRISPR/Cas9 handelt es sich um eine programmierbare Nuklease, die mittels einer vorab bestimmten und beigesteuerten guide-RNA Sequenz zu homologen DNA-Sequenzen finden und an besagten Stellen DNA Doppelstrangbrüche generieren kann. Diese revolutionäre Technologie ist auf Grund der Abhängigkeit der guide-RNA Sequenz hoch präzise und flexibel zugleich. Im Rahmen dieser Arbeit wurde CRISPR/Cas9 jedoch auf eine äußerst unkonventionelle Art und Weise verwendet. Als guide-RNA wurde eine Sequenz gewählt, welche die Cas9-Nuklease zu *Alu*-Elementen führt. Dabei handelt es sich um kurze, repetitive Sequenzen, welche verteilt über das humane Genom in über einer Million Kopien vertreten und etwa 11% unseres Erbguts ausmachen. Verursacht Cas9 innerhalb dieser Elemente Doppelstrangbrüche, sind diese zu zahlreich, um von Reparaturmechanismen behoben zu werden, das Genom wird fragmentiert und die Zelle stirbt. Um jedoch als Suizid-Kassette fungieren zu können, bedarf es an Induzierbarkeit, sodass diese nur wenn notwendig aktiviert wird und es andernfalls zu keiner Hintergrundaktivität und somit unbeabsichtigtem Zellsterben kommt. Für diesen Zweck wurde ein transkriptioneller sowie post-translationaler Kontrollmechanismus eingesetzt. Die Transkription der Cas9 Nuklease wurde mittels eines tet-on Systems geregelt, sodass die Transkription nur in Anwesenheit von Doxycyclin geschieht. Des Weiteren wurde eine Version der Cas9 Nuklease, genannt arC9, verwendet, welche durch die Fusion mit einer Östrogenrezeptor ligandenbindende Domäne eine Konformation annimmt, die die Nukleasefunktion verhindert und nur nach Zugabe von 4-Hydroxytamoxifen eine Konformationsänderung herbeiführt, die Doppelstrangbruchgenerierung erlaubt. Zusammen mit einer Expressionskassette für die *Alu*-spezifische guide-RNA, einer Expressionskassette für den reversen Tetracyclin kontrollierten Transaktivator und einer Neomycin-Resistenzgen Expressionskassette wurden alle Bestandteile zwischen transposasespezifischen Erkennungssequenzen auf einem Plasmid arrangiert, um nach Einbringung in Wirtszellen zusammen mit der dazugehörigen Transposase, die Integration des Konstrukts in das Wirtszellgenom zu ermöglichen. Das System wurde in HeLa Zellen getestet. Um die erfolgreiche Transposition zu bestätigen, wurden die Zellen 14 Tage mit G418 selektiert, wodurch eine resistente Zellpopulation generiert wurde, die vermutet wurde, die Suizid-Kassette zu enthalten. Zunächst wurde die ausschließlich konditionelle Expression der arC9 Nuklease durch Zugabe von 1 µg/ml Doxycyclin mittels Western Blot Analysen bestätigt. Des

Zusammenfassung

Weiteren wurde die Suizid-Kassette mittels zusätzlicher Zugabe von 200 nM 4-Hydroxytamoxifen induziert, Proteinextrakte extrahiert und auf die Anwesenheit von phosphoryliertem KAP1 Protein untersucht, um die potentiell eingeleitete DNA-Schadensantwort zu bestätigen. Nur in Anwesenheit von Doxycyclin, 4-Hydroxytamoxifen und *Alu*-spezifischen guide-RNA (als Negativkontrolle wurden Zellpopulationen generiert, welche alle Suizid-Kassettbestandteile außer die guide-RNA enthielten, generiert) wurde phosphoryliertes KAP1 Protein bereits nach 24 Stunden detektiert. Durchflusszytometrische Untersuchungen zur Detektierung von Annexin-V bestätigten weiterhin den Eintritt in die Apoptose. Von beiden generierten konditionell Cas9 exprimierenden Zellpopulationen wurden anschließend monoklonale Zelllinien generiert. Mit diesen wurden umfangreiche Wachstums- und Überlebensversuche durchgeführt, um den Effekt der Suizid-Kassettinduktion auf die Zellteilung und die Überlebensfähigkeit der Zelllinien zu ermitteln. Dabei wurde festgestellt, dass nach Induktion die Zelllinien ihre Teilungsfähigkeit verlieren und sterben. Zwischen 24 und 48 Stunden nach Induktion wurde das Anhalten der Zellteilung festgestellt, wonach der Sterbeprozess eintrat. Innerhalb von 5 Tagen nach Induktion wurden >99% aller Zellen eliminiert. In Abwesenheit von beiden Induktoren wurden keine signifikanten Unterschiede bei den Überlebenszahlen im Vergleich zu der Kontrollzelllinie ohne *Alu*-spezifische guide-RNA ermittelt. Lichtmikroskopische Untersuchungen der <1% überlebenden Zellfraktion offenbarten einen Seneszenz-assoziierten Phänotyp und zeigten keine Anzeichen einer Wiederaufnahme des Zellteilungsprozesses. Somit erreichte auch das zweite Kapitel dieser Arbeit sein Ziel zur Entwicklung einer funktionellen CRISPR/Cas9 basierten Suizid-Kassette, welche mittels Transposition in humane Zellen eingefügt werden kann, hochgradig abhängig in seiner Aktivität von den nötigen Induktionssignalen ist und hervorragende Eliminierungsfähigkeiten im getesteten Kontext aufweist. Zukunftsgerichtet wird der Weg geebnet, die Effizienz und die Funktion des vorgestellten Systems in klinisch relevanten Zellen zu bestätigen und samt eines therapeutischen Gens zu erproben. Falls erfolgreich, wäre die Möglichkeit zu erwägen, in Tierversuchen weitere Charakterisierungen vorzunehmen.

Zusammenfassend stellt diese Arbeit zwei unabhängige Ansätze vor, um die rasant wachsende Palette an Technologien für die Herstellung und Implementierung von Gen- und Zelltherapeutika mit zusätzlichen Angehensweisen zu erweitern, die deren Sicherheit für zukünftige Vorhaben potentiell verbessern könnten. Eine Weiterentwicklung und Weitererprobung der vorgestellten Systeme wird empfohlen, um deren vielversprechendes Potential vollumfänglich auszuschöpfen.

5. <u>General introduction</u>

The knowledge about DNA, the code that every living being harbors and which condenses all essential information for their existence, has come a long way. First discovered in 1869 by Friedrich Miescher and termed "nuclein" [1] it took close to a century to describe the structure by Watson and Crick in 1953 [2]. Since then, breakthroughs have been developed at an ever increasing pace: discovering that the code of bases is read in triplets which code for one amino-acid (aa) [3], deciphering the code (reviewed in [4]), cutting (reviewed in [5]), ligating (reviewed in [6]), sequencing (reviewed in [7]) and multiplying [8] DNA gave birth to the disciplines of molecular and synthetic biology. Today, in a process referred to as genetic engineering, we are able to purposefully and precisely make genetic altercations from introducing entire genes to edit individual bases to our advantage. Applications of genetic engineering range from adding functions to single cell systems enabling industrial-scale production of novel substances, include optimization of crops and livestock in agriculture adding valuable traits and bypassing time consuming and imprecise breeding, to modern medicine with gene and cell therapy, where disorders can either be genetically corrected or treated by reprogrammed cells (e. g., by immune cells genetically modified to recognize and combat cancer).

At the core of gene and cell therapy, there are the tools that allow for this process, of which there are several with distinct advantages and disadvantages, applications and limitations. A general categorization into two groups can be made here: i) tools that produce permanent changes within the genome of their target and thus lead to stable changes that are passed down after cell division and ii) tools that transiently shuttle genetic material into the target cells nucleus, achieving the desired effect only while transiently present. Within both categories, a further, more technical distinction can be made: viral and non-viral systems.

5.1. Viral vectors

Viral systems have been a widely implemented and obvious choice. Due to their nature, viruses possess the innate ability to enter target cells and sequestrate their machinery to ultimate express their genes, multiply their genome and thus ensure their own existence. Within the context of genetic engineering and gene and cell therapy, these characteristics can be used as a natural advantage. By adding transgenes to their genome, viruses can be used as a vehicle to shuttle genetic material with a potential to alter a cellular function into the target cells. Depending on the viruses' tropism and replication properties, this can be done for a variety of needs. Instances where stable integration of a therapeutic gene is required to ensure lifelong expression in target

cells maintained throughout cell division cycles and passed on to the next generation after mitosis, retroviruses haven been used extensively as integrating viral vector platforms in gene replacement therapy and most prominently in adoptive immunotherapy, leading the adoptive immunotherapy landscape (reviewed in [9]). When integration of the therapeutic gene is not required, non-integrating viral vectors such as adenoviruses and adeno-associated viruses (AAVs) have been established as a reliable and efficient vector technology. Episomal persistence paired with high transduction efficiencies and broad tropism profiles ensure efficient long lasting transgene expression in non-dividing cell types (reviewed in [9]). While several advantageous qualities have cemented viral vectors in preclinical and clinical research, culminating in being the tool of choice of therapeutic products enjoying regulatory approval today, distinct and difficult to overcome disadvantages plague these systems push for broader application and acceptance: namely, associated risks and costs. The most prominent risk factor contributing to the necessity of stringent safety assessments are the potential hazardous effects that can be caused by integrating viral vectors. Their preferential integration into transcriptional start sites (TSSs) and into actively transcribed genes [10,11] contributes to a significant safety risk regarding genotoxicity and tumorigenicity due to their increased potential of disrupting expression patterns ensuring cellular homeostasis. Additionally, the need for substantial safety screening and elaborate infrastructure for GMP-conform manufacturing adds a considerable cost-component, limiting accessibility. Constituting the currently most expensive medicines in the world, products have been withdrawn from entire continental markets due to low demand [9] and are expected to skip other markets entirely. Lastly, concerning non-integrating viral vectors, preexisting humoral and cellular immunities directed towards the viral particles can significantly lower therapeutic efficiencies within affected populations and negatively interfere with clinical outcomes [12,13].

5.2.   Non-integrating non-viral vectors

Non-viral systems encompass all alternative tools that do not rely on viral transduction to introduce foreign genetic material into desired cells. While mostly lacking the ability to independently enter cells and tissues efficiently, technological advances have pushed them into a position of challenging their viral counterparts with regards to delivery efficiency while often surpassing them with respect to specificity and safety, adding the important benefit of lower production and implementation costs. Although arguably more varied than viral alternatives, a similar classification into systems that produce permanent genomic changes and systems that do not modify their host's genome can be made. Concerning the former, the simplest yet especially in recent times one of the most discussed emerging technologies is the delivery of

naked nucleic acid such mRNA-based approaches and in this context the less known DNA-based approaches. Certainly, since the emergence of the Sars-CoV2 pandemic and the impact vaccine development has had on the public eye, mRNA technology has established itself as a robust, scalable, modular and economical method of eliciting an immune response against an infectious agent [14,15]. The underlying technology however, holds far greater potential and is expected to play an important role in the coming gene therapy landscape, as immune responses directed towards the mRNA-encoded antigen can be tailored towards an unlimited number of targets, expanding the utilities for vaccine development aimed at other infectious diseases but additionally encompassing tumors. This is achieved by i) chemically modifying mRNA to enhance its biological stability and translational capacity while reducing its immunogenicity [16], ii) packaging into lipid nanoparticles to enable cellular uptake and *in vivo* administration [17] and iii) advanced sequencing technologies and bioinformatically aided diagnostic tools allowing for the rapid identification of patient specific tumor-associated neoantigens and subsequent translation for manufacturing of tailored mRNA-based tumor vaccines [18]. Less prominent but quickly gaining attention, episomal vectors are emerging as an alternative to nonintegrative non-viral vector technology that share the low potential for insertional mutagenesis and low production cost but add the advantage of long-term transgene expression (reviewed in [19]). The latest generation of episomal vectors are circular DNA molecules that share multiple similarities with conventional expression plasmids: simple production, high stability, efficient delivery into cells, large cargo capacity and a modular design, allowing for a virtually unlimited combination of transgenes paired with promoters of varying strength and/or cell type specificity adaptable to the individual needs. Importantly however, they address the typical problem of plasmid loss after mitotic segregation, negatively affecting long-term nuclear persistence and thus, transgene expression. They achieve this by including scaffold/matrix attachment regions, AT-rich chromosomal elements that tether the episomal DNA to the nuclear scaffold or matrix, paired with an origin of replication ensuring low copy number episomal replication, leading to vector and expression persistence through multiple generations [20]. Therapeutic gene delivery using this technology has not arrived in form of clinical trials yet, but important advances are being made to potentially add it to the repertoire for biomedical applications [21–23].

Several needs however, can currently only be met by the category of tools that produce permanent genomic modifications. Within the non-viral realm of technologies two major representatives can be highlighted that bring distinct properties and functions: transposons and designer nucleases.

5.3. Transposons: integrating non-viral vectors

Transposons or transposable elements (TEs), commonly known as "jumping genes", were first discovered and described by Barbara McClintock [24], earning her a Nobel Prize in 1983. Traces of them and active copies can be found across all domains of life making up large portions of their hosts genome and have been described as being an integral part evolution through genome rearrangements and diversification, horizontal gene transfer and domestication (reviewed in [25,26]), their key characteristic being able to mobilize and change their location from one position to another. Focusing on eukaryotic TEs, a general distinction and categorization into two separate classes is made depending on the composition of their mobile phase. Class I TEs or retrotransposons move via an intermediate RNA molecule and exhibit replicative "copy-and-paste" transposition mechanism [27]. To effectively mobilize, their transposition involves transcription, reverse transcription and integration. Class II transposons or DNA-transposons mobile molecule remains DNA and can be either replicative [28] or non-replicative ("cut-and-paste") [29]. Non-replicative autonomous DNA TEs are typically two component elements: a gene coding for a transposase protein and repetitive flanking regions termed terminal inverted repeats (TIRs) containing recognition and binding sequences for the transposase. As such, once expressed, the transposase can recognize its own gene and is able to excise the entire sequence from left-TIR to right-TIR and reintegrate it into a novel location where the process can repeat itself (reviewed in [30]). The obvious conflict however, between being a strong driving force of evolution and having the potential of causing significant detrimental effects in its host genome, led to a majority of TEs to either accumulate mutations until they became inactive over the course of millions of years or to find and adopt key cellular and organismic functions within its host [31]. The potential to become a powerful tool for genetic engineering was not evident however until just over two decades ago, when the reconstruction out of fossil DNA sequences within fish genomes awakened the first DNA transposon showcasing activity in vertebrates, including humans, after a 10 million year slumber [32]: *Sleeping Beauty* (SB) [33]. SB presented the first viable alternative to viral vectors for efficient transgenesis. This is achieved by constructing a two-component system, replacing the transposase gene on a conventional plasmid with a transgene of choice and supplying the transposase *in trans*, on a separate expression plasmid. Once expressed, the transposase recognizes the TIRs and catalyzes transposition of the synthetic transposon harboring the transgene (reviewed in [34]).
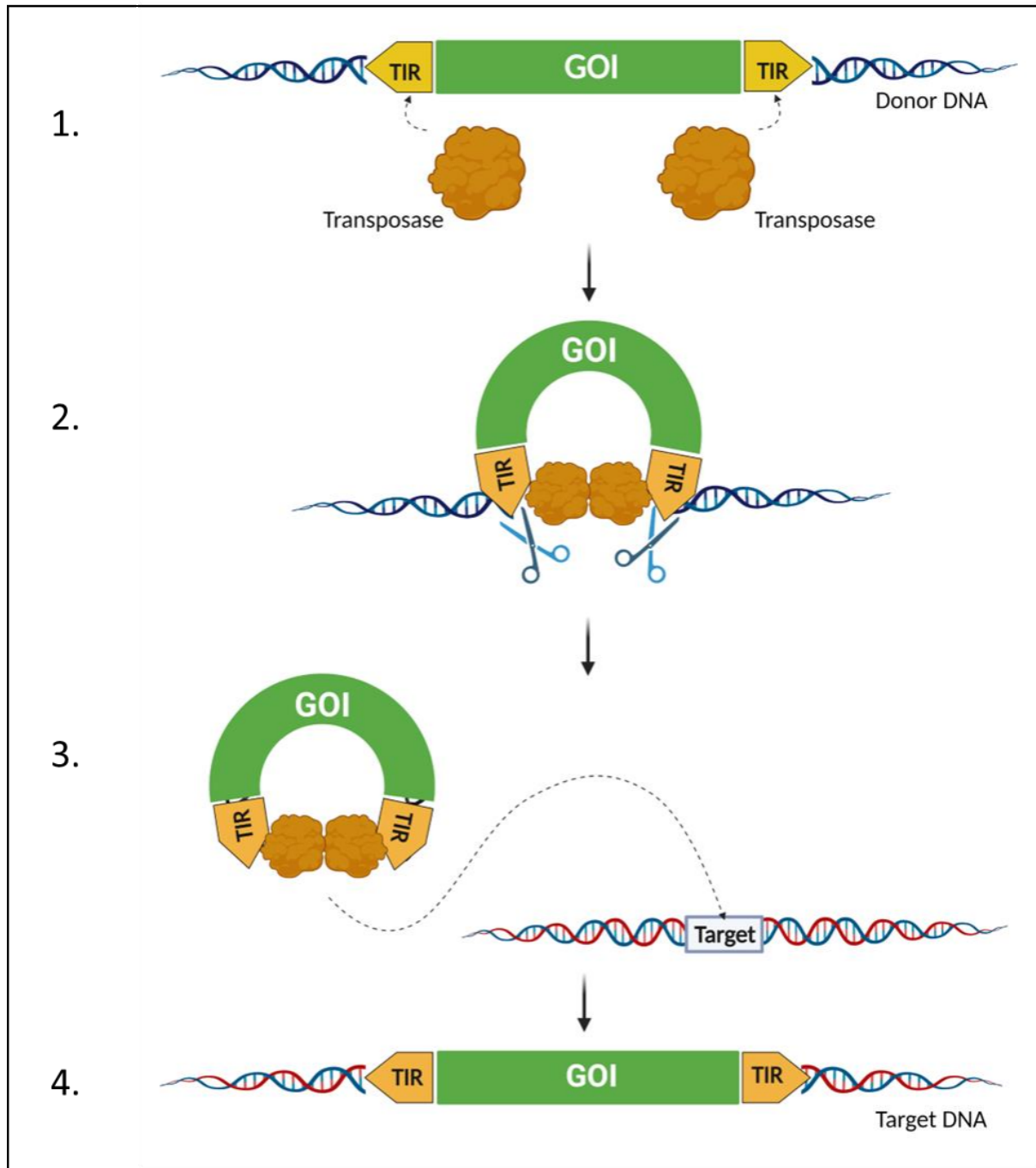
Figure 1*: Schematic overview of the basic principle of cut-and-paste transposition in transposon system based genetic engineering to shuttle a gene of interest (GOI) into a target cells genome. 1: Transposase enzyme recognizes TIRs flanking the GOI sequence, 2: Transposase enzyme introduces DSBs at the TIR and excises the transposon from the donor DNA strand, 3: The transposase carrying the transposon finds a suitable target on the target DNA strand to integrate its cargo, 4: Transposon containing GOI is integrated into a new location. (Figure was created using Biorender.com).*

Following SB, other transposon systems emerged showcasing utility in genetic engineering and are establishing transposon based vector technology in a growing landscape of non-viral gene and cell therapy clinical trials trying to highlight their distinct advantages over viral vectors (reviewed in [35]). Details will be discussed in following sections of this work.

5.4.   Designer nucleases

Finally, constituting an important second class of non-viral/hybrid genome engineering tools counting with a Nobel Prize winning discovery, designer nucleases such as zinc finger nucleases, transcription activator-like effector nucleases (TALENs) and most prominently clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein (Cas)9 have become indispensable for targeted genome modifications (reviewed in [36]). Different to all beforehand mentioned systems, designer nucleases lack the ability to directly introduce transgenes into their hosts. Rather, they allow for precise double strand break (DSB) induction at sequence-specific locations depending on the design of their DNA binding domains. Usually, these breaks get repaired via the non-homologous end joining (NHEJ) pathway introducing indels, which can be exploited to efficiently knock out genes through resulting frame-shit mutations. Through the supplementation with oligonucleotides or AAV vectors containing sequences of interest flanked by homologous sequences to the flanking regions of the intended cut, the homology directed repair (HDR) pathway can be forced however, effectively adding the possibility of introducing transgenic sequences with this tool as well. Similarly to transposon technologies, details and the current state of designer nucleases, especially CRISPR/Cas9, will be discussed later on.

5.5.   General safety concerns with integrative vector technology

It is obvious that the possibility of dangers associated with the disruption of a genome will always be present when foreign genetic material is inserted into a cell's genome, independently of the use of viral or non-viral methods describes previously. It became tragically apparent in the early days of gene and cell therapy, when first generation gammaretroviral vectors were used to introduce an IL-2 receptor γ gene to patient derived hematopoietic stem cells, manufacturing a cell product to administer to patients in order to treat X-linked severe combined immunodeficiency (SCID-X1) [37–39]. While immune reconstitution was achieved, several patients developed insertional mutagenesis driven leukemia [40,41], hampering the advance of gene therapy worldwide at the time. It became evident, that vectors used at the time possessed several characteristics which would have been in heavy conflict with today's safety standards. A central consideration when a foreign sequence is introduced into a genome and by definition causes an interruption is, where this interruption occurs. Growing knowledge and new assays, especially the advent of next-generation sequencing (NGS) and accompanying bioinformatics allow today to better define and identify so called genomic safe harbors: genomic sites where integrative events would not be expected to cause malignant transcriptome alterations. Unfortunately for the first SCID-X1 trials, this knowledge was lagging behind the drive to

provide revolutionary lifesaving therapies. Integrating patterns for the first gammaretroviral vectors were wrongfully expected to be random and thus primarily target noncoding regions. It was merely ensured that vector batches were free of replication-competent viruses, which was assumed to be the main risk factor of using a retrovirus vector. However, further research showed that integration preference towards TSSs, proto-oncogenes, cancer associated insertion sites and growth controlling genes [10,42], paired with enhancer sequences within the long terminal repeat (LTR) sequences affecting expression patterns of neighboring genes [40] represent a significant safety threat to gammaretroviral gene therapy applications. The overall better safety profile compared to other retroviruses has encouraged a preferred use of lentiviral vectors (derived of HIV-1) as a viral gene delivery vehicle and up to date no similar adverse effects as observed with the first gammaretroviral clinical trials have been registered. Yet, alternative splicing, the generation of aberrant transcripts and viral-cellular fusion transcripts [43,44] have been observed in lentivirus-modified cells highlighting the need for further development to reduce existing risks as therapies become more widely available. In a very recent example within a chimeric antigen receptor (CAR) T-cell clinical trial using a transposon engineered cell product, unexpectedly 2 out of 10 patients developed T-cell lymphoma originating from cells of the therapeutic transplantation [45,46]. As a highly surprising event, since no such development had ever been observed in CAR T-cell trials and mature cells have been shown to be highly resistant to malignant transformation [47], the underlying causes are still subject of investigation. The ideal gene therapy tool would efficiently allow for a targeted single insertion at a predefined genomic safe harbor. CRISPR/Cas9 has been hailed as a technology that could meet these criteria. However, a recent focus on chromosomal translocations that have been observed within pilot clinical trials employing CRISPR/Cas9 technology [48], underline the need for continuous safety screening and stringent regulatory evaluations concerning the risk/benefit ratio in a rapidly advancing landscape of novel gene therapies treading into uncharted territory.

5.6.   General project target

With the general risks linked to the advance of genetic engineering discussed beforehand in mind, especially concerning the foreseeable exponential growth of translational medicine in the gene and cell therapy field due to the rise and establishment of breakthrough technologies, this work aims to contribute to improving the overall safety aspect of current state of the art procedures. Two key time points within a therapeutic scenario were considered for this purpose: before the products delivery during manufacturing and after the product is administered to the patient. For the first time point, the *piggyBac* (PB) transposon system was taken into

consideration. PB is, similarly to SB, an attractive transposon system and alternative to viral vectors for permanent transgene insertion and has been translated to clinical trials. It exhibits however, in contrast to SB, a preference for integrating its cargo into transcriptional start sites (TSSs) of its host, which constitutes a significant risk of inducing genotoxic effects. Chapter I of this work discusses and compiles important background information, addresses a hypothesis explaining this behavior and presents and discusses an experimental approach to modify the integration profile of the PB transposase, yielding several PB transposase variants with reduced integration biases. For the second time point, a novel suicide-switch was conceptualized, developed and tested which enables rapid and effective cell killing of genetically modified cells. Within a therapeutic scenario, a suicide-switch component within a cell product would allow to clear a patient from the product if severe side effects were to develop. Chapter II introduces the novel concept of unconventionally using CRISPR/Cas9 to maximize target numbers within the human genome and conditioning its activity to transcriptional and translational induction, causing genome fragmentation and cell death upon activation. Background information, study design and results providing a proof-of-concept regarding the systems performance will be introduced, presented and discussed in detail in the respective chapter.

6. <u>Chapter I: Disrupting PB transposase – BRD4 interaction influencing insertion site preferences and transposition efficiency</u>

6.1. Introduction chapter I

6.1.1. Transposon system-based gene transfer

As briefly mentioned previously, transposon gene transfer systems based on autonomous DNA TEs constitute an attractive alternative to virus-based systems, offering an extensive list of notable advantages. Current transposon systems are remarkably simple to produce and implement for efficient integrative gene transfer protocols. Their simplicity lies in the two-component nature of their mechanism. For transposition, only two main components are required: the transposase enzyme and the transposon. The form in which these components are delivered to the target cell to induce transposition is highly flexible and can be adjusted depending on the applications needs. The most widely adopted delivery method is the classical two-plasmid delivery. The transposase expression cassette is delivered on a conventional expression plasmid and the transposon composed of a gene of interest flanked by transposase specific TIRs is delivered separately on a second plasmid. Manufacturing and propagation of plasmid DNA is economical, scalable and together with the implementation to produce transgenic organisms can be conducted in standard BSL1 laboratories, facilitating the access where higher biosafety levels cannot be met or infrastructure costs prevent from meeting them. Transfer into target cells can be achieved with standard transfection or electroporation methods. The same advantages are true for the remaining forms in which transposon system components can be delivered. This can be in form of miniature plasmids, doggybone (DB) DNA and for the transposase additionally mRNA and recombinant protein (reviewed in [35]). These design changes address the shortcomings of plasmid DNA with respect to delivery efficiency, transposition efficiency, toxicity and genomic safety. Standard plasmid DNA typically contains, besides the expression cassette for the transposase or the transposon harboring the gene of interest, a backbone consisting of sequences needed for bacterial replication including antibiotic selection markers. These sequences often containing unmethylated CpG dinucleotides can be recognized by Toll-like receptor 9 and stimulate the innate immune system [49]. Additionally, while only playing a role for plasmid manufacturing, they contribute to the overall plasmid size, which negatively correlates to the transfection/electroporation rate as

larger molecules cross the cell membrane less efficiently. Miniature plasmids such as plasmids free of antibiotic resistance markers (pFARs) [50], minicircles (MCs) [51] and DBs [52] address this issue while simultaneously, concerning the plasmid carrying the transposon, reduce the distance between the TIRs, improving transposition efficiency [53]. For the transposase, the use of recombinant protein or mRNA circumvents this problem altogether. Additionally, an important safety aspect is added. The significantly lower half time of mRNA and protein compared to DNA limits the transposase expression to a shorter time window, reducing the risk of transposon remobilization that might compromise genomic integrity. Overall, these properties have allowed the SB and PB transposon systems to join the group of genetic engineering tools used to manufacture cell products used in clinical trials (reviewed in [35,54]). Subject of this chapter will be the PB transposon system, which will be introduced in detail in the following section.

6.1.2. The *piggyBac* transposon system

The PB transposon is a naturally active DNA TE that made its first appearance in the laboratories of the 70's when it was noticed that, for unknown reasons, the occlusion bodies of certain nuclear polyhedrosis viruses changed their morphology when they were passaged in cabbage looper moth (*Trichoplusia ni*) cell lines [55]. Host cell DNA insertions disrupting viral genes were identified to be the probable cause for these phenotypical changes and as insertions could be reverted by continuous passage in the same cell line and homologous sequences to the interrupting insertions were found to be represented in multiple copies across the *T. ni* genome [56], it was suspected that this phenomenon was caused by an active mobile DNA element. Suspicions were definitively confirmed several years later, when the full TE sequence was determined [57], the responsible transposase enzyme identified and the transposon named PB [58]. The PB transposase is a member of a large superfamily of PB-like elements [59]. The enzyme is 594 aa in length and coded within a single 2475 open reading frame (ORF) flanked by terminal sequences (5' end 311 bp long 3' end 235 bp in length) harboring asymmetric inverted repeats. These contain 13-bp long TIRs and 19-bp long sub-terminal repeats (STRs), separated by 3 bp in the left and 31 bp in the right terminal sequence [60]. The PB transposase itself can be divided into several domains, starting at the N-terminus with an acidic N-terminal domain followed by a dimerization and DNA-binding domain, a catalytic domain containing the catalytic triad of three aspartates at positions 268, 346 and 447, an insertion domain interrupting the catalytic domain sequence and a cysteine-rich C-terminal domain [61] containing a nuclear localization signal predicted to span positions 551 to 571 [62].
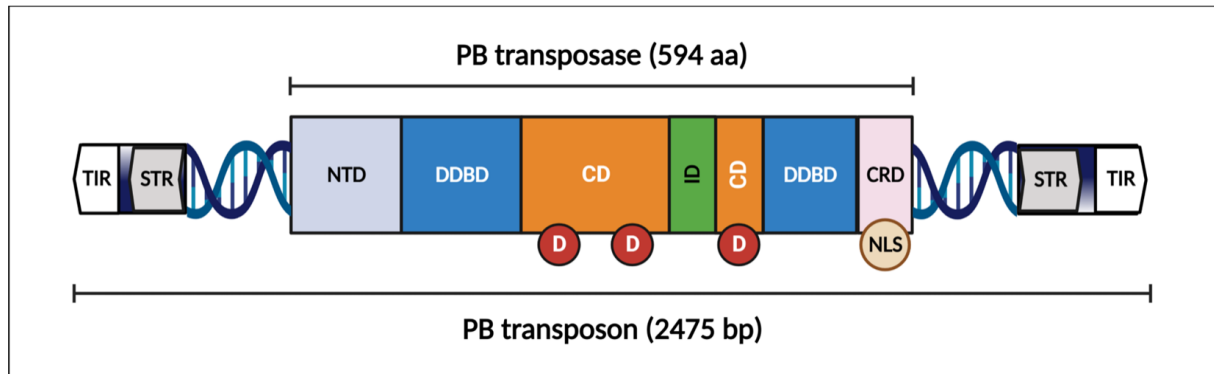
Figure 2: Overview of the structure of the PB-TE spanning 2475 bp and containing an ORF for a 594 aa long PB transposase protein flanked by terminal sequences containing TIRs (white arrow shapes) and STRs (grey arrow shapes). The PB transposase itself is composed of an N-terminal domain (NTD, light blue), a dimerization and DNA-binding domain (DDBD, blue), a catalytic domain (CD, orange) containing the three Ds (red circles), and a cysteine rich domain (CRD) containing the predicted NLS (beige circle). (Figure was created using Biorender.com)

PB transposition occurs almost exclusively into TTAA tetranucleotides [57]. From a biochemical perspective, the transposition itself begins with a PB transposase-mediated hydrolytic nick at the 3'-ends of the PB transposon. The exposed 3'-OHs start a nucleophilic attack on the 5'-end of the TTAA sequence on the complementary strand creating a phosphodiester bond and hairpin on the transposon ends effectively producing a DSB, liberating the transposon from the donor DNA [63]. This process leaves complementary TTAA overhangs at the 5'-ends of the donor DNA which can anneal to each other, repairing the DSB on the donor DNA through ligation instead of DNA synthesis, producing the unique attribute of the PB-system amongst eukaryotic DNA transposons of not generating footprints at the excision site [64,65]. At the ends of the excised transposon, the hairpins protect the transposon ends from degradation following a similar mechanism as bacterial transposons Tn5 and Tn10 [66,67]. Before integration occurs, the hairpins are hydrolytically resolved, re-exposing the 3'-OHs which start a nucleophilic attack on the 5'-Ts of the TTAA target site, creating a nick. Resulting TTAA overhangs at the target site pair to the resolved TTAA flaps on the transposon ends, creating TTAA target site duplications on each side of the now integrated transposon [63]. The TTAA target site specificity, although not strictly needed for transposon excision [68], is suggested to be essential for transposon integration due to the high distortability a TTAA tetranucleotide exhibits. This enables the melting and necessary severe distortion needed for the scissile phosphodiester bonds at the 5'-Ts of the TTAA target site to reach the catalytic pocket of the PB transposase and the exposed 3'-OH of the transposon ends [61]. It is hypothesized, that PB randomly probes target DNA until the requirements of target site flexibility are met, which in the vast majority of cases is a TTAA tetranucleotide [61].

As previously described for transposon-based genetic engineering tools, the PB transposon system can be split. Placing a gene of interest between the PB TIRs while providing the PB transposase separately allows for transgene transfer besides in insects and infecting baculoviruses in yeast [63], plants [69], mammals [70] and *in vitro* [63].

Since its discovery, several efforts have succeeded in further improving PB transposition efficiency in comparison to the wildtype isolate. Besides shortening and optimizing the TIR sequences enhancing overall transposition efficiency [71,72] and codon optimization [73], several hyperactive PB mutants have been engineered [74,75] exhibiting transgene transfer rates that rival viral vectors. In contrast to viral vectors however, PB shares a unique trait of transposon systems of not having a strict upper cargo capacity, in contrast to viral particles with a structural packaging limit. PB has been shown to mobilize giant bacterial artificial chromosomes of over 100 kb in length while maintaining the integrity of its cargo [53,76].

PB has been used for a variety of applications. The earliest and most widely adopted application that takes special advantage of PBs ability to efficiently mediate transposition in a vast range of organisms, is the use as a powerful mutagenic element for forward genetic screens [77]. To do so, the PB system is modified to incorporate reporter and mutagenic cassettes depending on the intended usage that ranges from gene trapping, poly(A) trapping, oncogene trapping to promoter/enhancer trapping *in vitro* as well as *in vivo* (reviewed in [78]). In order to identify the insertions that produced a phenotypical impact throughout the experiment, insertions can be retrieved via PCR-based methods [77,79], mapped to the respective genome after sequencing and then assessed. The PB transposon system has furthermore been a valuable tool to efficiently generate transgenic animals. This can be achieved fairly efficiently by microinjection of transposon DNA and transposase DNA or mRNA (avoiding potential background integration of transposase gene at spontaneous DSB sites, leading to detrimental lifelong transposon remobilization) into fertilized eggs and has been successfully demonstrated for a variety of animals, namely zebrafish, mice, rats, pigs and cattle [70,80–83].

The efficiency with which transposon systems shuttle large complex gene cassettes into a target cell's genome has proven to be particularly useful for reprogramming somatic cells to induced pluripotent stem cells (iPSCs) by introducing the necessary reprogramming factors (reviewed in [84]). The reprogramming efficiency is significantly higher than with alternative non-integrating non-viral methods, is comparable to the efficiency of viral vectors and adds several significant advantages. Besides the cost reduction in contrast to viral methods, the unique property of PB transposition described previously that leaves no footprint after excision, allows

to seamlessly remove reprogramming factors after successful reprogramming, generating footprint-free and factor-free iPSCs [85,86]. The generation of a excision proficient/integration deficient PB mutant [87] would additionally allow for a more efficient reprogramming factor removal since 40-60% of excised transposons reintegrate when using the wildtype PB [88]. Also, efficient transient transgenesis in general is facilitated.

Importantly, PB holds a substantial potential for gene and cell therapy applications (reviewed in [35,89]). Within the scope of preclinical studies, PB has been used in numerous occasions to successfully mobilize therapeutic genes and achieve functional correction of genetic disease models and full functionality of therapeutic cell products. Concerning functional correction, the most prominent models are disorders where a functional gene is introduced into the target cells genome to replace or complement a defective one for sustained transgene expression. Genes include full-length dystrophin providing functional amelioration of muscles in Duchenne muscular dystrophy mouse models [90], the cystic fibrosis transmembrane conductance regulator gene for functional correction of cystic fibrosis in pigs [91], full length factor VIII correcting mouse models of hemophilia A [92], the argininosuccinate synthetase gene and ornithine transcarbamylase gene for correction of urea cycle defects in mice [93] and the ATP-binding cassette 4 gene correcting familial intrahepatic cholestasis type 3 in mouse models [94].

Besides replacement gene therapies, immune cell engineering aiming for adoptive immunotherapy solutions to treat several forms of malignant tumors has seen a substantial rise in recent years. Especially CAR therapy, which targets cancer specific cell surface markers, has taken off as a breakthrough therapy with several T-cell products gaining regulatory approval to treat hematological malignancies and several more in clinical trials including efforts to expand towards solid tumors (reviewed in [95]). Currently, the vast majority are being generated via retroviral gene transfer including all approved products for the world market. However, severe bottlenecks concerning the cost of these therapies are starting to solidify as therapies move away from academic research and small early phase clinical trials to a larger patient group after product approval [96]. In an effort to reduce cost and simplify the manufacturing procedure, transposon systems have successfully been introduced into the adoptive immunotherapy landscape and a number of preclinical studies have proven feasibility of generating therapeutic CAR products using non-viral transposon technology alternatives including PB. This encompasses CD19- and CD116/CD131-specific CAR T-cells aimed at B-cell malignancies [97–101], human granulocyte-macrophage colony-stimulating factor receptor-specific CAR T-cells against acute myelocytic leukemia [102], human epidermal growth factor receptor-specific CAR T-cells targeting non-small cell lung cancer [103], glypican-3 and EGFRvIII-directed

CAR T-cells against hepatocellular carcinoma [104,105] and mesothelin-targeted CAR T-cells against large solid tumors [106].

Providing solid preclinical data suggesting efficacy and cost-effective clinical translation, PB entered the clinical trial stage in recent years with several CAR T-cell clinical trials. The majority adopted the well-characterized CD19 B-cell-specific cell surface marker which has been employed by most viral vector-based methods and is the target of most approved CAR T-cell products such as Kymriah (Novartis), Yescarta (Gilead Sciences), Tecartus (Gilead Sciences) and Breyanzi (Brystol Myers Squibb). These include the CARTELL study in Australia (NHMRC identifier: 1102172) and two more recent studies from Japan and China (UMIN Clinical Trials registry ID: UMIN000030984 and ClinicalTrials.gov ID: NCT04289220 respectively). Additionally, in USA-based trials BCMA is being targeted as a multiple myeloma-specific marker in a phase I/II study (ClinicalTrials.gov Identifier: NCT03288493) and PSMA as a target for prostate and salivary gland cancer (ClinicalTrials.gov Identifier: NCT04249947). CD116-specific CAR T-cells will be used as an alternative to CD19 in Japan to treat leukemia patients (JRCT ID: jRCT2033210029). For solid tumors, a phase I clinical trial employed EGFR-specific CAR T-cells to treat patients with NSCLC (ClinicalTrials.gov Identifier: NCT03182816) in China. While tolerance is generally high, toxicities low and initial reports on response rates favorable [107–109], in an unexpected development 2 out of 10 patients enrolled in the CARTELL study developed CD19-CAR T-cell product derived T-cell lymphoma [45,46], with one patient tragically dying. Genome and transcriptome sequencing data analysis show a higher-than-expected numbers of transposon insertions in one of the lymphoma cases, structural variants, copy number variations and transcriptional readthrough from the CAR gene, affecting global gene expression levels in both lymphoma cases. None of the before mentioned data match with known drivers for lymphomagenesis, including a shared integration into the BACH2 gene, causing downregulation. BACH2 haploinsufficiency is rather associated to immunodeficiency and autoimmunity [110]. However, a preference to integrate into the BACH2 gene has been described before [111] for PB and is a feature shared with the murine leukemia virus (MLV) retrovirus, raising the question if a correlation does exist.

6.1.3. *piggyBac*'s biased integration pattern

While the mechanism that causes these tragic lymphoma cases is currently unknown and research is being conducted to elucidate the underlying cause for other malignant T-cell transformations, it is clear that particular care has to be taken when venturing into unexplored territory translating preclinical data into the clinic. The PB transposon system, while containing the previously discussed advantages of a non-viral gene transfer tool and several unique

advantages, also exhibits certain problematic characteristics that are not found in the alternative SB transposon system. While current consensus is that this was not the cause behind the lymphoma cases, PB integrates its cargo preferentially into 5'-transcriptional regulatory regions of genes [111,112] following a similar integration pattern as gammaretroviral vectors such as MLV [10,111], which was linked to insertional mutagenesis in early clinical trials [40,41]. The mechanism behind the MLV integration preference is known: the integrase (IN) of gammaretroviruses including MLV was shown to physically interact with members of the bromodomain and extraterminal domain (BET) proteins BRD2, BRD3 and BRD4 which acts as a targeting factor for MLV integration near BET-chromatin binding sites [113–115]. Since BET proteins bind to acetylated histone tails [116] which associate with TSSs and are part of the transcriptional activation complex, integrations are tethered towards these locations. Inhibiting BET proteins accordingly leads to a reduction in integration bias towards these sites [113,115]. Interestingly, the PB transposase has been shown to physically interact with BET proteins as well [111]. A similar mechanism tethering the PB transposase has been proposed and structural analysis show distinct structural similarities between a beta hairpin found responsible for the MLV-IN/BET interaction [117] and an initially predicted (unpublished data, Prof. Dr. Thomas Krey, Universität zu Lübeck) and later experimentally confirmed [61] beta hairpin found in the PB transposase between aa V390 and P410 located within the insertion domain, which interrupts the catalytic domain (Fig. 2).
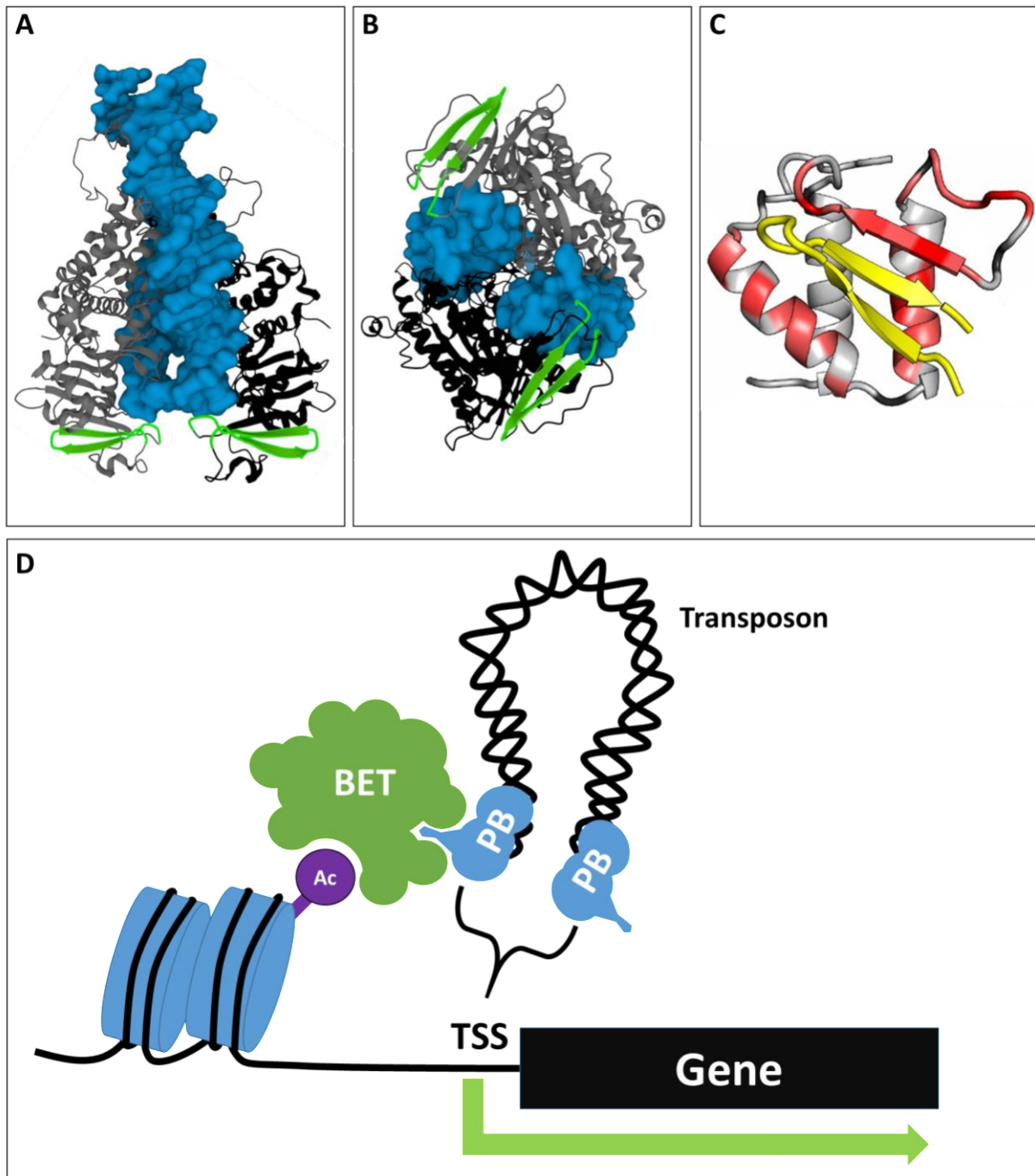
***Figure 3**: Structural basis for PB integration bias towards TSSs. **A:** Side view of PB strand transfer complex (STC) (Protein Data Bank entry 6X67, [61]). Blue: Transposon DNA structure, Grey: PB monomer protein structure, Black: PB monomer protein structure, Green: beta hairpin of both PB monomers spanning aa V390-P410 exhibiting similar structural characteristics to the MLV-IN beta hairpin (yellow structure in C), **B:** Bottom view of PB-STC (Protein Data Bank entry 6X67, [61]), **C:** MLV-IN interaction with extraterminal domain of BET protein (adapted from [117]). Yellow: MLV integrase hairpin structure, red: BRD4 extra terminal domain, **D:** Depiction of hypothesis explaining PB insertional bias towards TSSs. BET proteins (green protein) associated to acetylated (violet) histone (blue cylinders) interact with PB transposase (light blue protein) during target site selection, guiding integration towards TSSs.*

Thus, it is proposed that the tethering mechanism guiding gammaretroviral integration into TSSs and promoter regions can be similarly applied to PB transposition (Fig. 3D). It is hypothesized, that modifying the hairpin to disrupt PB-BET interaction while maintaining

transposition activity intact, might reduce the biased insertion preference of PB transposase and produce a more random and thus safer integration profile. The first project of this work aimed at experimentally characterizing the beta hairpin's role in the PB-BET interaction, generate mutations that disrupt the interaction while retaining transpositional activity and observe the effect of the mutations at potentially redirecting PB integrations away from TSSs.

6.1.4. Significance chapter I

As discussed, it is clear that PB is an extremely efficient and versatile genetic engineering tool for stable transgene delivery. However, due to the insertion bias it shares with the first generation gammaretroviral vectors [111] that proved problematic during the first gene therapy clinical trials, it severely lags behind established viral gene transfer systems and the SB transposon system with regards to adoption in clinical applications today. Modifying the integration preferences away from TSSs would substantially improve the system's safety profile and might alleviate concerns that were further strengthened with the occurrence of lymphoma cases after the first in-human clinical trial employing PB technology and encourage taking advantage of attractive safety features unique to the PB system such as footprint-free excision. Furthermore, it would provide a valuable proof-of-concept concerning rational design-based retargeting of a transposase system, highlighting the potential for additional improvements of existing gene transfer systems concerning their underlying performance and characteristics.

6.2.   Methods chapter I

6.2.1.   Site-directed mutagenesis

All generated PB mutants were generated via site-directed mutagenesis PCR using either the Q5® High-Fidelity DNA Polymerase (NEB) for all insect PB (iPB) mutats, and the Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher) for the mouse codon optimized (mPB) mutants. Unless stated otherwise, the employed oligonucleotide primers included a 5'-phosphate to facilitate downstream ligation reactions and were manually designed with back-to-back orientation (Tab. 14 for details) considering and ensuring suitable annealing temperatures, which were calculated using the Tm Calculator tool from NEB (https://tmcalculator.neb.com). PCR conditions were set according to the manufacturer's instructions. PCR products were separated on 0,7% agarose gel, excised, purified and eluted in 10 µl dH$_2$O. 5 µl of the eluted sample was taken and the linearized double-stranded DNA contained was re-circularized by overnight ligation at 16 °C using the T4 DNA ligase (NEB) following the manufacturers recommendations. When non-phosphorylated oligonucleotide primers were employed in the mutagenesis PCR, this step was preceded by a phosphorylation reaction using a polynucleotide kinase (NEB) according to the manufacturer's recommendation, followed by DNA purification. 2 µl of the ligation reaction was directly transformed into chemically competent Subcloning Efficiency™ *E. coli* DH5α (Thermo Fisher) according to manufacturer's instructions followed by a 1 h incubation at 37 °C and 700 rpm in SOC medium. After the incubation period, 100 µl were plated on LB-Agar plates containing 100 µg/ml ampicillin and incubated overnight in an incubator at 37 °C. Colonies were picked the subsequent day to inoculate 5 ml LB-medium containing 100 µg/ml ampicillin for overnight incubation at 37 °C and 200 rpm. Plasmid DNA was extracted from the liquid cultures using an extraction kit according to manufacturer's recommendations and samples were sent for sequencing (Eurofins) to confirm the presence of the desired mutations and identify colonies that harbor them.

6.2.2.   Determination of PB-BRD4 interaction via co-immunoprecipitation

HEK293T cells, seeded at a quantity of $2,5 \times 10^5$ cells per well of a 6-well plate 24 h prior, were co-transfected with 1 µg of an expression plasmids for C-terminally HA-tagged PB transposase (Tab. 13) and 1 µg of an expression plasmid for N-terminally Myc-tagged BRD4 expression plasmid (Tab. 13) using the transfection reagent Lipofectamine® 3000 (Thermo Fisher) according to the manufacturer's recommendations. For each sample corresponding to either transfections with the Myc-BRD4 expression plasmid in combination with either a PB-

Mutant-HA expression plasmid, the iPB-HA/PB-3X-HA expression plasmids as a positive control, or lacking a PB-expression plasmid as a negative control, transfections were performed in duplicates on two independent wells. 48 h after transfection, cells were washed once with PBS and harvested directly in 350 µl of Lysis buffer #3 (Tab. 5) by scraping the adherent cells with a cell scraper and resuspending the cell suspension with a micropipette. Once all cells of one sample were detached, the content was moved to the second well containing the corresponding duplicate and the process was repeated. The 350 µl lysate containing all cells of one duplicate co-transfection sample was transferred to a microcentrifuge tube and placed in a rotating shaker at 20 rpm for 30 min. at 4 °C. To finalize the cell lysis, samples were sonicated using a Bioruptor® Plus sonication device (Diagenode, 3 cycles, each consisting of 15s on and 15 s off at LOW setting). Cell debris were sedimented by centrifugation at 20 000 × g for 15 min. in a cooled microcentrifuge at 4 °C. 320 µl of the supernatant were moved to a clean microcentrifuge tube and centrifugation at 20 000 × g was repeated for 10 min. at 4 °C to sediment any remaining debris. 300 µl of the supernatant were moved to a clean microcentrifuge tube. 50 µl of each 300 µl sample were transferred to a fresh microcentrifuge tube to serve as the input sample and stored at -20 °C. The remaining 250 µl were placed at 4 °C while the washing of the Protein G Sepharose™ 4 Fast Flow beads (Cytiva) intended for sample pre-clearing was carried out: 6 × n µl (where n equals the number of samples) of sepharose bead suspension were washed with 17 × n µl of Lysis Buffer #3 by resuspension with a micropipette using previously modified pipette tips (tip end cut off manually to increase the tip diameter). Sepharose beads were sedimented by centrifugation at 20 000 × g for 10 s at 4 °C. The supernatant was removed, discarded and the washing process repeated three times. The final sepharose-bead pellet was resuspended in 13 × n µl of Lysis Buffer #3. 12 µl of pre-washed sepharose beads were added to each 250 µl sample and incubated in a rotating shaker at 20 rpm for 30 min. at 4 °C for pre-clearing to remove any trace amounts of contaminating immunoglobulins. After incubation, sepharose beads were sedimented at 20 000 × g for 10 min. at 4 °C. The supernatant was carefully removed and transferred to a fresh microcentrifuge tube, where 12 µl of c-Myc Monoclonal Antibody-Agarose Beads (Takara) were added. The co-immunoprecipitation was carried out overnight in a rotating shaker at 20 rpm at 4 °C. The next day, samples were centrifuged at 20 000 × g for 10 min. at 4 °C to sediment the c-Myc Monoclonal Antibody-Agarose Beads. The supernatant was discarded carefully and the remaining pellet was washed five times similarly to the pre-washing of the sepharose beads except that 300 µl of Lysis buffer #3 were used for re-suspension per sample for each washing step. To finalize the last washing step, only 230 µl of Lysis buffer #3 were removed to leave

the washed c-Myc Monoclonal Antibody-Agarose Beads in 70 µl of Lysis buffer #3. To the 70 µl, constituting one Co-IP sample, 35 µl of 3X SDS sample loading buffer were added. The 50 µl previously frozen input samples were thawed and 25 µl of 3X SDS sample loading buffer were added. Co-IP and Input samples were incubated for 5 min. at 95 °C, loaded on an 10% polyacrylamide gel and subjected to an SDS-PAGE. Separated proteins were transferred to a nitrocellulose membrane for Western Blot (WB) analysis. The nitrocellulose membrane was blocked overnight in blocking buffer (Tab. 5) at 4°C in a shaking incubator at 55 rpm. The next day the blocking buffer was discarded, the membranes cut between 35 and 25 kDa, and individually incubated with the corresponding primary antibody (anti-H3 for the loading control, anti-HA to detect HA-tagged PB transposase protein (Tab. 4)) for 90 min. at room temperature on a shaking incubator at 55 rpm. After incubation with the primary antibodies, these were removed and membranes washed three times with TBS-T Buffer for 10 min. at room temperature on a shaking incubator at 55 rpm. Afterwards, incubation with the corresponding HRP-conjugated secondary antibodies (Tab. 4) was carried out similarly to primary antibody incubation except for a 60 min. incubation time instead. After secondary antibody incubation was finalized, membranes were similarly washed three times in TBS-T Buffer except for a 5 min. incubation per washing step. Finally, protein detection was carried out via chemoluminescent detection (Tab. 11) using a CCD-Imager (Tab. 10). After detection, the top membrane half was stripped of antibodies (https://www.abcam.com/protocols/western-blot-membrane-stripping-for-restaining-protocol) for re-probing with Myc-specific primary antibodies and secondary HRP-conjugated antibodies (Tab. 4) to detect Myc-tagged BRD4. Antibody incubation steps and chemoluminiscent detection was carried out as previously described. If necessary, quantification of the WB signals was carried our using ImageJ's (Tab. 17) gel analysis tools.

6.2.3. Transposition assays

The transposition activity of all PB transposase variants was assessed in HeLa cells by performing colony forming assays by co-transfecting the corresponding PB expression plasmid alongside a plasmid harboring a PB transposon containing a neomycin resistance gene (Tab. 13), followed by selection with G418 (Tab. 8), colony staining and counting. For that purpose, $2{,}5 \times 10^5$ HeLa cells were seeded the day before transfection per sample on a well of a 6-well plate. The next day, 500 ng of PB expression plasmid were co-transfected with 50 ng of pXL-Neo plasmid via lipofection (Lipofectamine® 3000 Transfection Kit according to manufacturer's instructions). 48 h after transfection, cells were trypsinized and resuspended in growth medium vigorously to produce a single cell suspension. 10% of the suspension were

moved to a 10 cm cell culture plate and selection was started by cultivating cells in growth medium containing 1 mg/ml G418. Selection was continued for 14 days with regular medium changes to remove dead cells. After the selection period, the medium was aspirated and remaining cell colonies washed once with PBS. Afterwards, cells were fixed by covering the plates with a 4% PFA solution and a 2h incubation at room temperature. Following fixing, the PFA solution was removed and replaced with a 0,2% methylene blue solution. Incubation was carried out for 2h at room temperature to stain the fixed cell colonies. After the incubation period, the staining solution was removed and plates were rinsed in dH$_2$O to remove excess staining solution. The plates were then dried at room temperature overnight. To determine the colony numbers, each plate was covered with wheat flower to increase the contrast and the bottom of each plate was scanned. Files were imported into ImageJ (Tab. 17) to digitally determine colony numbers and the relative area covered by cells on each plate in the following manner:

- Select plate with an oval and clear non-selected component (Edit → Clear Outside)
- Convert Image type to 8-bit (Image → Type → 8-bit)
- Adjust selection threshold automatically on the plate containing the most colonies of the experiment, apply same threshhold for the remaining plates (Image → Adjust → Threshold → Auto/manually adjust → Apply)
- Re-select plate with an oval shape
- Modify image to increase individual colony separation (Process → Binary → Watershed)
- Analyze image (Analyze → Analyze Particles → Size: 0-Infinity, Circularity: 0.00 – 1.00, Show: Nothing, checkmark on "display results", "summarize" and "exclude on edges" → OK)
- Export results and plot

Where specified, results were shown normalized to the corresponding negative control. Negative values, resulting from a lower value than the corresponding negative control, were set to 0.

6.2.4. Multiple sequence alignment and secondary structure rendering

aa sequences were retrieved from the Protein Database of the National Center for Biotechnology Information (NCBI) and were subjected to multiple sequence aligment using the MAFFT version 7 online service (https://mafft.cbrc.jp/alignment/server/, [118,119]). To add the secondary structure rendering and depict residue conservation, the alignment files were

downloaded and submitted to ESPript 3.0 (https://espript.ibcp.fr/, [120]) where entries in the Protein Data Bank (PDB, http://www.rcsb.org/pdb/) were used as input files.

6.2.5. Insertion library generation

To determine the genomic integration site preferences of the generated PB mutants, an insertion library was generated out of various cell lines harboring PB transposons introduced with each PB mutant variant under investigation and subsequently sequenced on using the Illumina platform.

Initially, HeLa cells that were seeded one day prior at a density of $2,5 \times 10^5$ cells per well on a six well plate, were co-transfected with 500 ng of either an expression plasmid of an mPB mutant (see 6.2.1.) or the mPB expression plasmid for the wildtype PB transposase as a control, alongside with 50 ng of the pXL-Neo plasmid using the TransIT®-LT1 Transfection Reagent (Mirus) according to manufacturer's instructions which if integrated into the cells genome introduces a resistance to G418 and thus allows for antibiotic selection of cells containing PB transposons in their genome. For the generation of the mPB-based insertion library, each co-transfection was performed in triplicates in order to triple the overall number of transposition events. 24 h post transfection, cells were moved from each individual well on the 6-well plate to a 10 cm cell culture plate and subjected to antibiotic selection by adding 1 mg/ml G418 into the cell culture medium. Selection was maintained for 2 weeks with regular medium changes to remove dead cells. After the two-week selection period, 5×10^6 cells from each 10 cm plate were harvested and genomic DNA was isolated following the kit's (Tab. 11) manufacturer's instructions. The DNA concentration of each sample was determined with a spectrophotometer (Tab. 10). DNA from each sample of one triplicate was pooled in equal ratios to one final sample with a combined amount of 2 µg of DNA.

The 2 µg for each sample were diluted in the genomic DNA isolation kit's provided elution buffer to a total volume of 50 µl and subjected to shearing via sonication using an M220 Focused Ultrasonicator (Covaris) according to the manufacturer's instructions at the following settings in order to generate fragments with an average length of 600 bp: Peak Power 50.0, Duty Factor 5.0, Cycles/Burst 200 and Duration 80 s. The resulting DNA fragments were isolated using magnetic beads (Beckman Coulter) according to manufacturer's instructions and dissolved in 85 µl H_2O.

In the following reactions, each followed by isolation using magnetic beads, fragments were end-repaired using the and dA-tailed using the NEBNext® Ultra™ II End Repair/dA-Tailing Module (NEB) following manufacturer's protocol to generate single dA-overhangs, and ligated

to linker-sequences (Tab. 14) using a T4 DNA-Ligase (NEB) according to the manufacturer's recommendations. Isolated ligation products after magnetic bead purification were dissolved in 20 µl H_2O.

The resulting DNA fragments were subjected to a nested PCR. The first round is to amplify and enrich the junction between PB-transposon ends and genomic sequence using PB-TIR and linker-sequence specific primers (#79 and #80, Tab. 14). 5 µl of the isolated DNA solution following the ligation step was used as a template while following the NEBNext® High-Fidelity 2X PCR Master Mix (NEB) manufacturer's instruction to set up the 50 µl reaction sample. The program for the first round of PCR is described in the following table.

Table 1: PCR program used to enrich junctions between PB-transposon ends and genomic sequences

Step	Temperature (°C)	Duration (s)	Cycles
Initial Denaturation	98	30	1
Denaturation	98	10	
Annealing	67	30	10
Extension	72	30	
Denaturation	98	10	
Annealing	62 (1 °C/s ramp rate)	30	15
Extension	72	30	
Final Extension	72	300	1
Store	4	∞	1

Following the first PCR-round, PCR products were purified using magnetic beads and dissolved in 20 µl H_2O. 5 µl of each sample were used as a template for the second round of PCR, which is designed to further enrich transposon-genomic junction sequences and introduce (through a combination of the employed PCR-Primers #81 - #92, Tab. 14) 6-bp barcodes, specific to each sample, to enable parallel sequencing and subsequent read-sample association during analysis and Illumina adapter sequences for downstream processing. The program for the second round of PCR is described in the following table.

Table 2: *PCR program used to enrich junctions between PB-transposon ends and genomic sequences and to introduce sample specific barcode and Illumina adapter sequences for downstream processing.*

Step	Temperature (°C)	Duration (s)	Cycles
Initial Denaturation	98	30	1
Denaturation	98	10	
Annealing	63 (1 °C/s ramp rate)	30	15
Extension	72	30	
Final Extension	72	300	1
Store	4	∞	1

The resulting PCR products were separated on a 1,5% agarose gel and the products ranging between 200 and 500 bp in length were excised and gel-purified. After purification, the DNA content of each sample was determined using a Qubit® 2.0 Fluorometer (Invitrogen) according to the manufacturer's instruction and 30 ng of each sample was pooled into a final sample, constituting the finalized insertion library.

6.2.6. Next-generation sequencing and data analysis

All NGS-services and data subsequent data analyses were conducted by the Genomics Core Facility of the Paul-Ehrlich-Institute (Dr. Csaba Miskey and Tanja Leyendecker).

6.3.   Results chapter I

6.3.1.   Co-immunoprecipitations of alanine scan PB-mutants

Based on previous research, where single aa within a structural hairpin of the murine MLV integrase were identified to mediate the protein-protein interaction between the MLV integrase and BET-proteins and inhibiting said interaction shifted the transgene integration pattern away from a biased one towards TSSs [113,115], a similar hairpin structure was predicted within the PB transposase based on structural modeling (unpublished data, Prof. Dr. Thomas Krey) and later confirmed via cryo-EM [61]. This provided a structural explanation for the known protein-protein interaction between PB transposase and BET proteins ([111]) which drove the hypothesis of being able to shift the integration preference of the PB transposase from TSS to a more random distribution if disturbing the PB-BET interaction in a similar way to the MLV integrase study is achieved.

Consequently, the aa sequence in question ranging from V390 to P410 was subjected to an alanine scan and each resulting alanine replacement mutant used in Co-IP experiments (see 6.2.2.) to observe its interaction behavior with BRD4. The resulting protein bands on the WB membrane were quantified according to their relative intensity in comparison to the signal measured with the corresponding wildtype PB transposase protein, which was subjected to the same Co-IP experiments as a control. Additionally, a PB transposase mutant containing 3 C-terminal mutations (iPB-C3, [121]) was included. This hyperactive mutant had previously shown a significantly less biased integration pattern towards TSS than its counterparts [121] and was generated similarly to the alanine substitution mutants via site directed mutagenesis using the pRRL.PPT.SF.piggybacHApre plasmid as a template (see 6.2.1.).
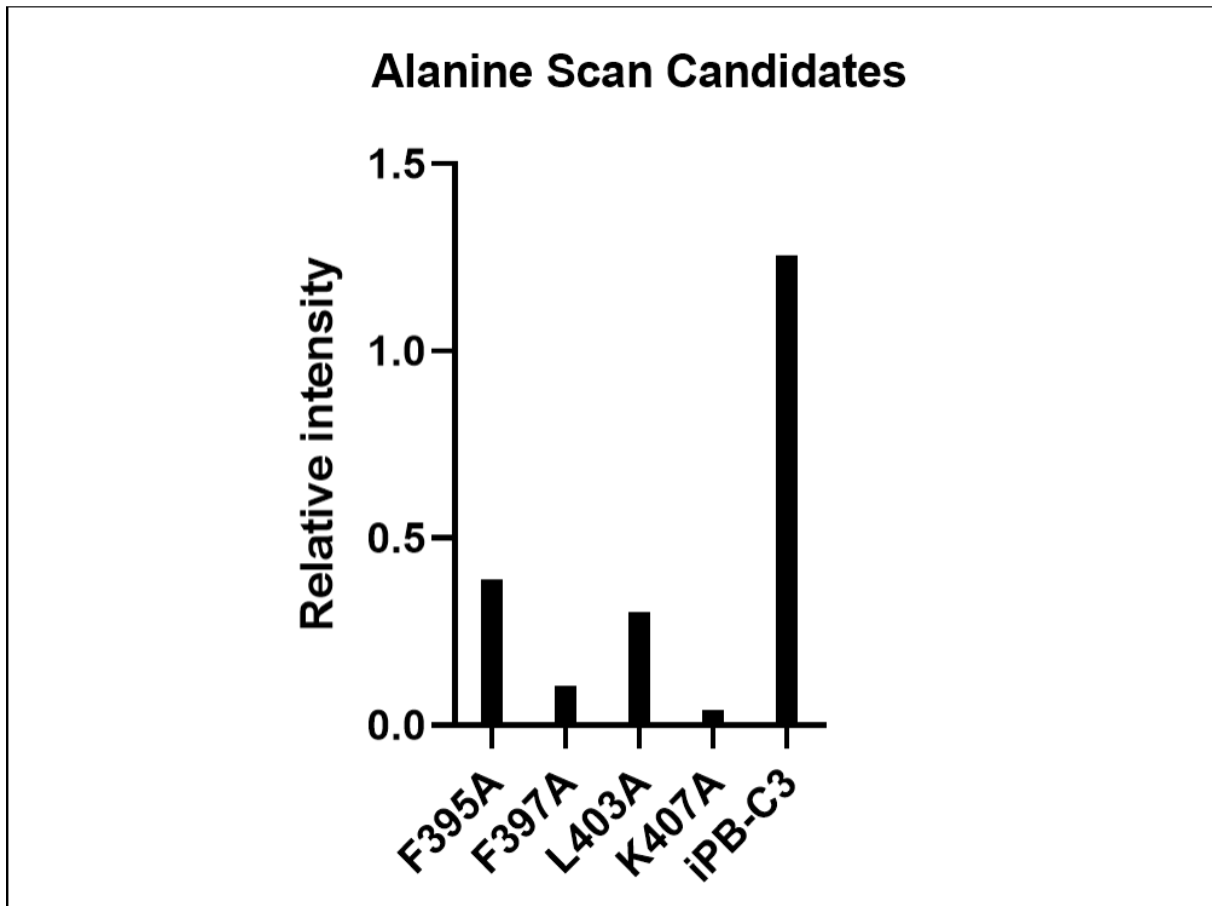
Figure 4*: Excerpt of relative signal intensity of Co-immunoprecipitated HA-tagged PB mutant protein bands normalized with the signal intensity obtained with the internal wildtype HA-tagged iPB control. Shown are the values that were observed to show a visible decrease in signal strength and additionally the value generated with the HA-tagged iPB-C3 transposase. Values below 1.0 indicate a weaker signal intensity relative to the control while values above 1.0 indicate a stronger signal intensity relative to the control.*

Out of the 20 resulting mutants with single alanine substitutions throughout the hairpin structure (for full Co-IP WBs see supplementary material), 4 exhibited a notable reduction in signal strength compared to the control, while the HA-tagged iPB-C3 transposase exhibited a stronger relative signal strength (Fig. 4). Mutant PB-HA-K407A exhibited a 96% reduction in relative signal strength, PB-HA-F395A a 61% reduction, PB-HA-L403A a 70% reduction and PB-HA-F397A a 90% reduction, while iPB-C3-HA increased its signal strength by 26%.

Given these preliminary results that identified critical sequence positions that influence PB-BET interactions, mutants were subjected to transposition assays to assess if they retain their transpositional activity. However, mutant transposases lacked apparent transpositional activity when harboring an alanine substitution at positions determined to influence BRD4 interaction via Co-IP. Thus, alternative aa substitutions were considered and investigated as described in the following sections.

6.3.2. Multiple sequence alignment and structural disruption-based PB-mutagenesis

Given that the PB-mutants generated through the alanine scan with reduced BRD4 interaction behavior appeared to have lost their transpositional activity, alternative substitutions were investigated with the goal of finding replacements both transpositionally active and exhibiting a decrease in protein-protein interaction with BRD4 when subjected to Co-IP experiments. To find suitable substitutions, the aa sequence of the PB transposase (GenBank accession number ABC67521.1) was subjected to a protein BLAST search and resulting hits with decreasing sequence similarity were picked (63% similarity GenBank accession number: GBP63866.1, 60% sequence similarity GenBank accession number: ADV17598.1, 40% sequence similarity GenBank accession number: ADU33112.1, 39% sequence similarity GenBank accession number: TGZ37512.1). Alongside the original aa sequence of the PB transposase and sequences of two additional members of the PB superfamily (*FoxyPB* (unpublished sequence) and PiggyBat [122]), they were subjected to a multiple sequence alignment and subsequent secondary structure rendering based on the recently described PB transposase 3D structure ([61], see 6.2.4. for methods).
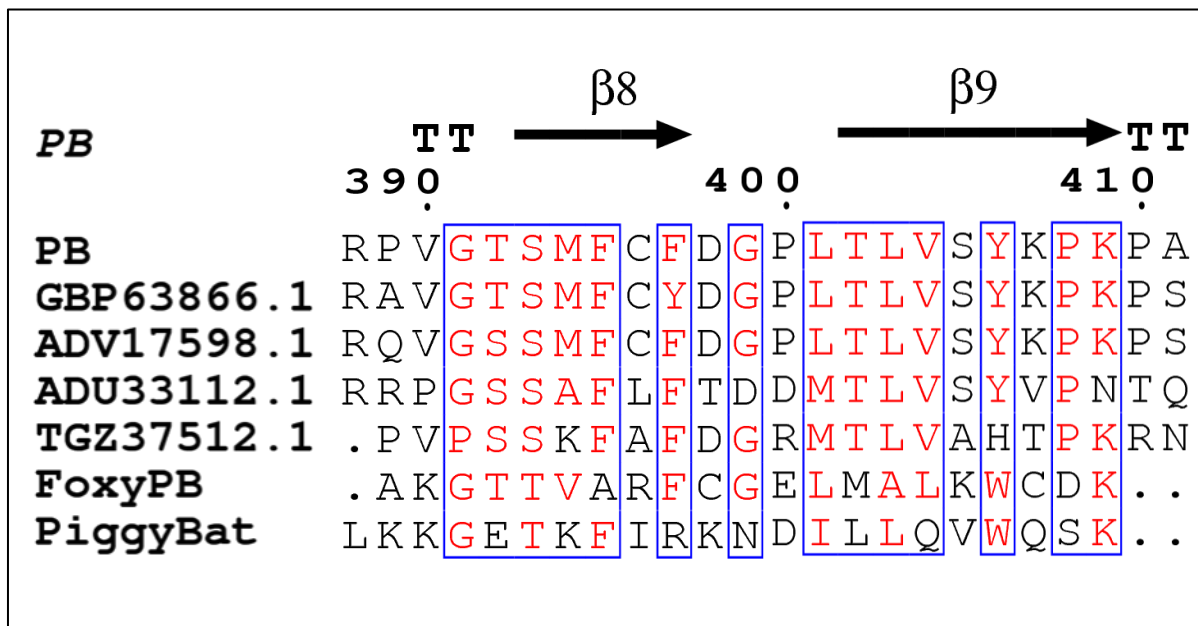
Figure 5: *Excerpt of a multiple sequence alignment between the PB transposase sequence and selected members of the* piggyBac *superfamily, showing the structural hairpin suspected to mediate PB-BET interaction between V390 and P410. Corresponding positions of turns (T) and beta sheets 8 and 9 (β8 and β9). Sequence conservation across all sequences is indicated with blue rectangles while individual aa conservation with red letters.*

As Co-IP experiments with alanine replacement mutants had provided transpositionally inactive PB-transposases yet identified positions that influenced the interaction with BRD4, potential alternative substitutions were identified at given positions based on aa found in selected representatives of the *piggyBac* superfamily. Selected were substitutions V390K,

F397R, F397Y, Y406W, K407C, K407Q, K407Q, K407V, and P410_A411del (Fig. 5) based on their occurrence in selected members, both the PiggyBat and FoxyPB sequence and the finding that PiggyBat exhibits a weaker insertion bias towards TSSs [122]. Additionally, several mutants were generated based on structural predictions done by Prof. Dr. Thomas Krey (University of Lübeck, unpublished data) with the premise of selecting aa substitutions adding steric hindrance to the PB-BET interaction without disrupting the hairpin structure itself. These substitutions were E380F, V381D, S387Q, T392Y, M394E, M394R and C396R.

All mentioned substitutions were introduced into the an HA-tagged iPB-C3 sequence via site-directed mutagenesis (see 6.2.1.), to maximize a potential reduction in insertion bias toward TSS, given that iPB-C3 had shown a significant drop in previous studies and additionally improving transposition efficiency as it is described as a hyperactive mutant in comparison to the wildtype [121].

6.3.3. Transposition assays of new PB-mutants

Generated PB mutants based on the previously described multiple sequence alignment and structural predictions were subjected to a transposition assay to assess their transpositional activity as a requirement for downstream Co-IP experiments.
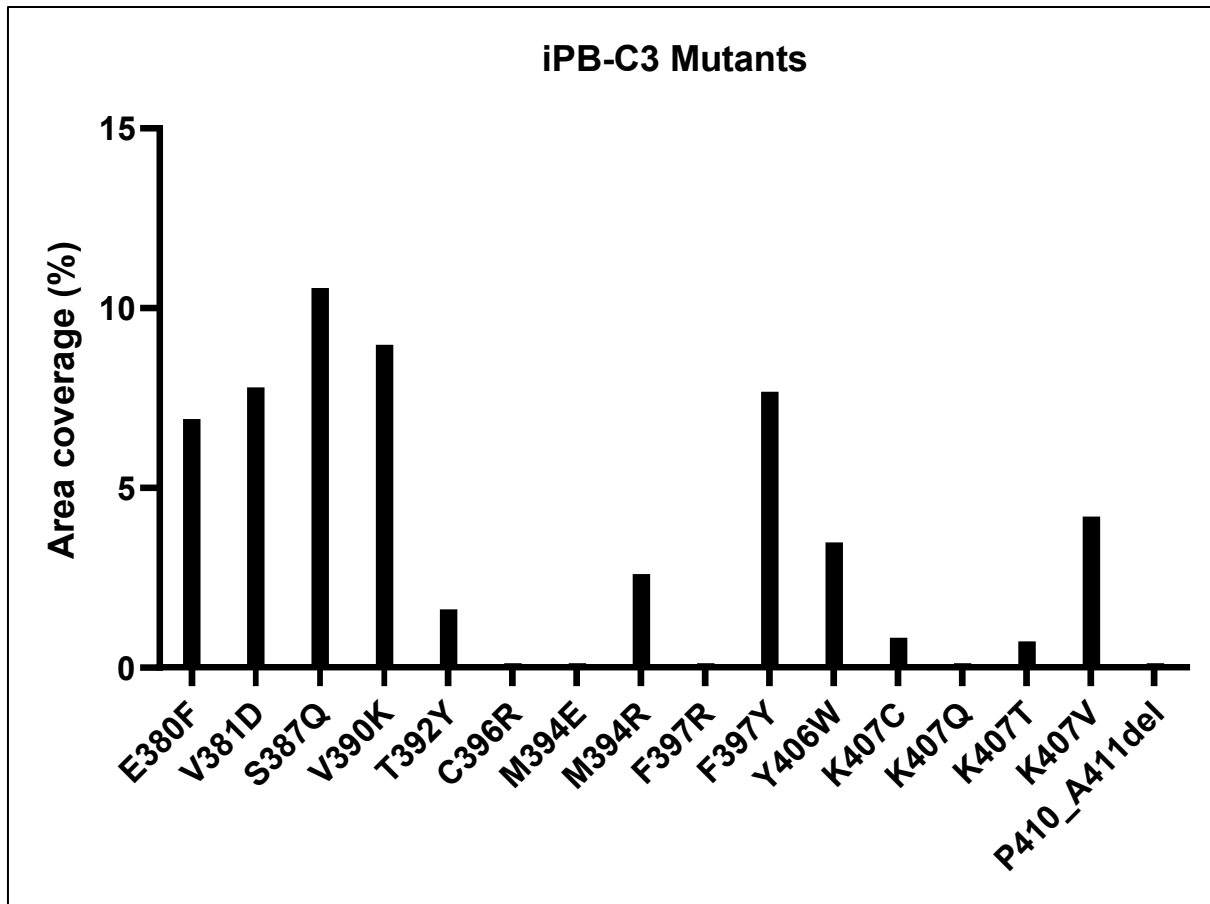
iPB-C3 Mutants

Figure 6: Representation of the transpositional activity of iPB-C3 Mutants generated on the basis of multiple sequence alignments and structural predictions. Values indicate the percentage of area covered by resistant cells on the corresponding cell culture plate after 2 weeks of antibiotic selection, normalized by subtracting the percentage of area coverage by the corresponding negative control.

Over a total of 16 mutants tested, 11 exhibited transpositional activity: E380F, V381D, S387Q, V390K, T392Y, M394R, F397Y, Y406W, K407C, K407T and K407V (Fig. 6). Accordingly, these mutants were selected for downstream Co-IP experiments to assess their protein-protein interaction with BRD4.

6.3.4. Co-immunoprecipitations of new PB-mutants

Mutants identified to have kept or regained their transpositional activity (Fig. 6) were used as input in additional Co-IP experiments (for methodology see 6.2.2.). Resulting protein bands corresponding to the Co-immunoprecipitated iPB-C3-mutants that showed a visibly lower band intensity in comparison to the iPB-C3 control were quantified and normalized with the corresponding band intensity of the corresponding iPB-C3 positive control.

iPB-C3 Mutants

AA substitution	Co-IP Signal Intensity Decrease
E380F	65%
V390K	53%
T392Y	95%
M394R	92%
K407C	67%
K407Q	62%
K407V	97%
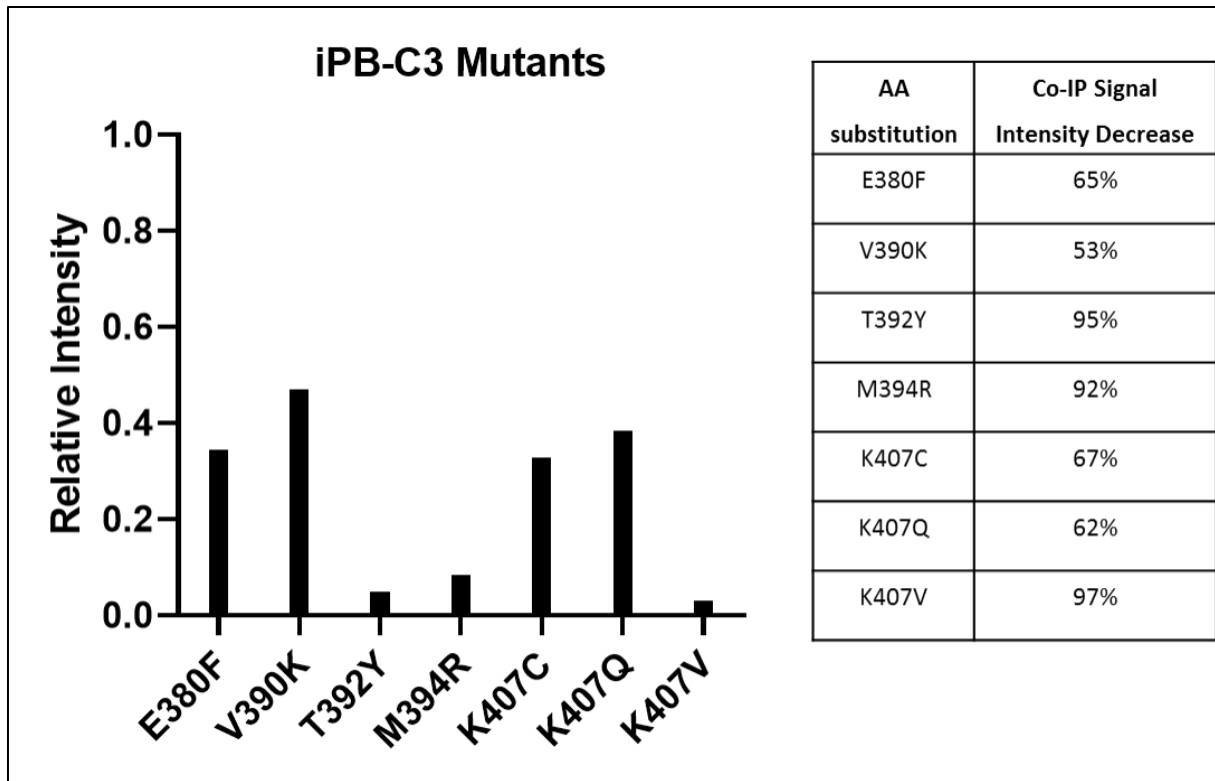
Figure 7*: Excerpt of relative signal intensity of Co-immunoprecipitated HA-tagged iPB-C3 mutant protein bands normalized with the signal intensity obtained with the internal HA-tagged iPB-C3 control. Shown are samples that exhibited a measurable decrease in signal strength. Values below 1.0 indicate a weaker signal intensity relative to the control.*

Quantification resulted in the identification of a total of 8 aa substitutions that weakened the protein-protein interaction between the PB-transposase and BRD4 (Fig. 7). These 8 substitutions were subsequently introduced into a mouse codon-optimized PB sequence termed mPB (Tab. 13) via site-directed mutagenesis (see 6.2.1.) to generate an HA-tag-free version that could possibly influence target site selection while maximizing transposition activity, given that the C-terminally added HA-tag had a previously shown severe negative impact on transposition efficiency (unpublished data) limiting transposition events and thus reducing the number of integrations that can be rescued in downstream NGS analysis. Additionally, given that the C-terminus contains the DNA binding domain of the PB transposase [61], it was excluded that the close proximity of a foreign HA-tag could influence target site selection in any way. Also, since surprisingly the three C-terminal mutations contained in iPB-C3 did not reduce the PB-BRD4 interaction (Fig. 4), they were excluded from the generation of mPB mutants used for downstream integration library generation.

6.3.5. Novel PB-mutants exhibit a reduced integration preference for TSSs

To determine if the mutations determined via the Co-IP experiments produced an effect on influencing the integration preference, integration libraries of cells harboring transposons introduced with each mutant were generated and sequenced.

Table 3: Number of recovered integration events for each mPB-Mutant and the WT-PB transposase (data were determined by Dr. Csaba Miskey).

mPB-Mutant	Number of recovered integrations
mPB-E380F	2318
mPBK-407C	1557
mPBK-407Q	1647
mPBK-407V	3272
mPB-M394R	2429
mPB-T392Y	2637
mPB-V390K	2974

Recovered integration events (Tab. 3) within 3 kb up- and downstream of TSSs were quantified and normalized with respect to the total number of recovered integrations and normalized frequencies for each mutant compared to each other and the WT.
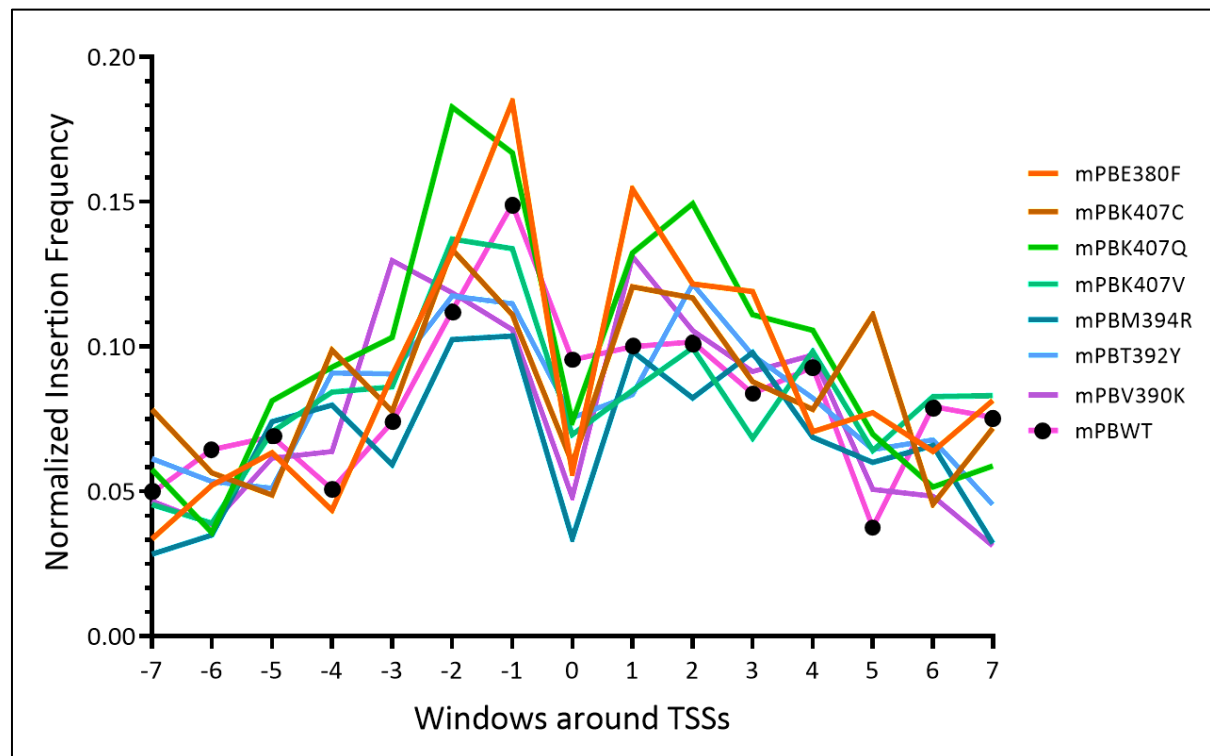
Figure 8: Insertion frequencies of PB-Transposons within a 6 kb window around TSSs obtained for each mPB-Mutant and the WT-mPB transposase (mPBWT, pink line with black dots) as an internal control. Insertion frequencies obtained with each mutant are normalized with the total number of integrations obtained with each mutant. Every datapoint represents the normalized insertion frequency within 400 bp (200 bp up and downstream of each x-axis value) windows around TSSs (value "0" on x-axis). (Sequencing data contained in this figure were produced by Dr. Csaba Miskey).

When comparing the integration profiles within the 6 kb window surrounding TSSs, a noticeable drop within the 400 bp window directly surrounding the TSSs is observed in all integration profiles of the generated PB-Mutants (X axis datapoint 0, Fig. 8), indicating a lower insertion frequency in the direct proximity of that region. Additionally, a different distribution pattern can be observed. While the insertion frequency of the WT-PB showcases a singular peak within the first downstream window from the TSSs, mutants appear to exhibit two peaks, up and downstream of TSSs.
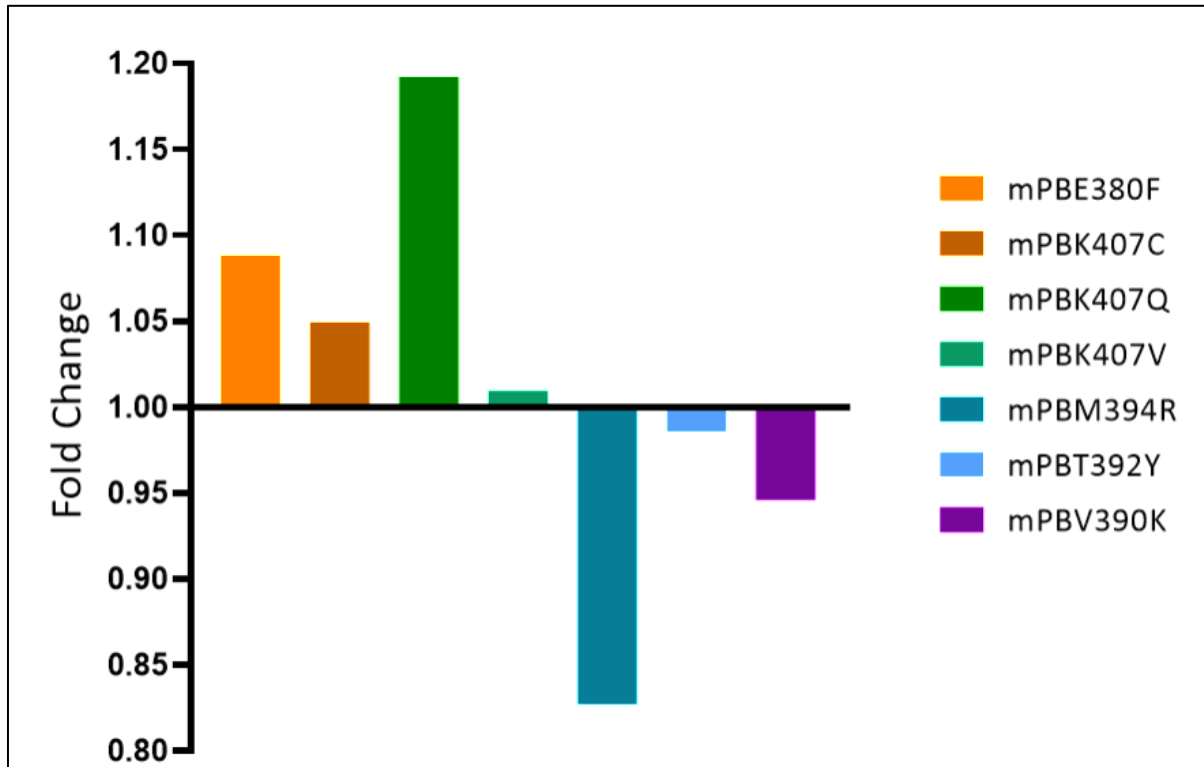
Figure 9: *Fold change of integrations into 6 kb window around TSSs produced with mPB-Mutants in comparison to the WT PB-transposase. Values above 1 represent a bigger relative fraction of integrations detected within the 6 kb window, values lower than 1 a lower relative fraction.*

Quantifying the overall integration frequency into the 6 kb window around TSS obtained for each PB-mutant and normalizing it with the integration frequency of the WT-PB transposase used as a control (Fig. 9), three mutants can be identified that exhibit a lower integration preference into that area in contrast to the WT control: mPB-M394R, mPB-T392Y and mPB-V390K. mPB-V390K exhibited a 5,4% reduction in insertion frequency into this window, mPB-T392Y a 1,5% reduction and mPB-M394R the largest reduction with a 17,3% smaller proportion of integration events. The remaining mutants all exhibit larger proportion of integration events occurring into the 15 kb window surrounding TSSs. These increments are: 8,8% for mPB-E380F, 5% for mPB-K407C, 19,2% for mPB-K407Q and 1% for mPB-K407V.
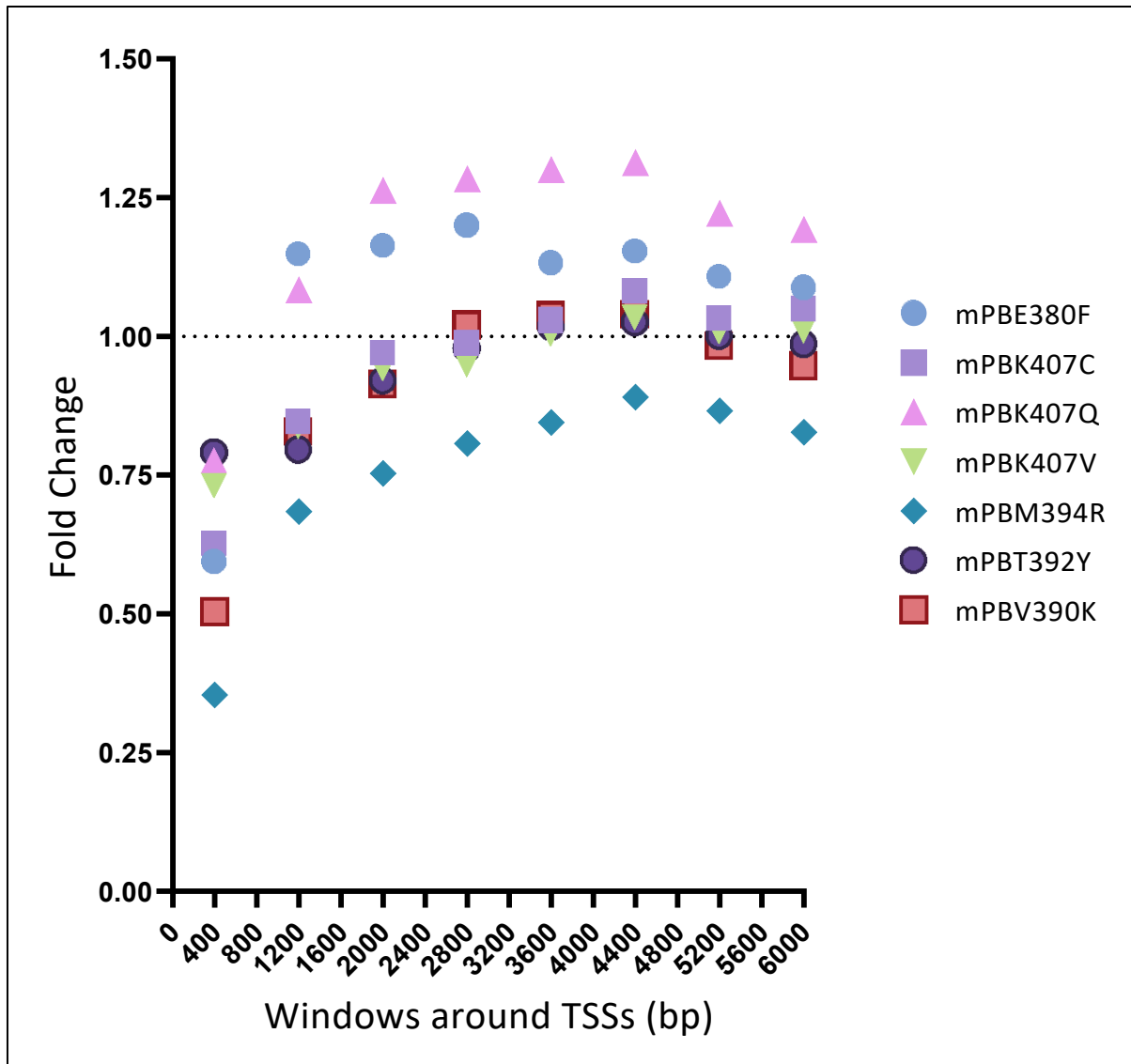
Figure 10*: Fold change of integrations into windows around TSSs of varying sizes (between 400 and 6000 bp) produced with mPB-Mutants in comparison to the WT PB-transposase. Values above 1 represent a bigger relative fraction of integrations detected within each window, values lower than 1 a lower relative fraction.*

Dissecting the 6 kb window surrounding the TSSs into smaller windows and determining the fold change difference in the proportion of integrations within these windows (Fig. 10) highlights the reduction in integration bias occurring within the smallest window of 400 bp around TSSs exhibited by all generated PB-Mutants. Within that window, mPB-E380F showed a reduction of 41%, mPB-K407C of 37%, mPB-K407Q of 22%, mPB-M394R of 65%, mPB-T392Y of 21% and mPB-V390K of 50%. Within every window up to the 2 kb window, a majority of mutants still exhibit a lower integration preference in comparison to the WT-PB transposase. Interestingly, mPB-M394R is the only mutant that exhibits a reduction within every window size tested and shows the strongest reduction amongst mutants in every window.

6.4.   Discussion chapter I

The PB transposon system is a highly efficient, flexible and easy to use non-viral genetic engineering tool. As discussed previously however, it is significantly held back from broader clinical translation by exhibiting a problematic integration pattern for its cargo, preferentially hitting TSSs. It is safe to assume, that the close-to-random integration pattern exhibited by the alternative SB transposon system is a contributing factor favoring its adoption for human applications, reflected in the larger number of clinical trials conducted in the past and being conducted currently (reviewed in [35]). Considering the disadvantage, this work represents a rational approach to modifying PB integration pattern based on disrupting the physical interaction interplay of the PB transposase and the BRD4 protein, which co-localizes at TSSs. Previously unpublished data preceding this work predicted a hairpin within the PB transposase (Prof. Dr. Thomas Krey, Universität zu Lübeck) with structural similarities to a hairpin structure found in the MLV integrase, shown to be responsible for the respective interaction with BET proteins, mediating integration bias for MLV itself [117]. The hairpin structure has since then been confirmed experimentally [61].

6.4.1.   Alanine scan through PB-hairpin structure

The initial approach of conducting an alanine scan through the hairpin structure, although not providing active transposases, generated valuable insights and a proof-of-concept, that the interaction properties between PB and BRD4 could be influenced by modifying the hairpin's composition. This provided the first confirmation, that the structure at hand is a major contributor to the interaction as predicted. On the other hand, it highlighted the importance of the hairpin's integrity for the transposase function. This is expected, since the hairpin structure was shown to be an essential component of the insertion domain, spanning from residues 373 to 433 of the PB transposase and interrupting the catalytic domain [61]. As such, two of the four alanine scan mutants can be predicted to have influenced the hairpin structure and function significantly: F395A and F397A.
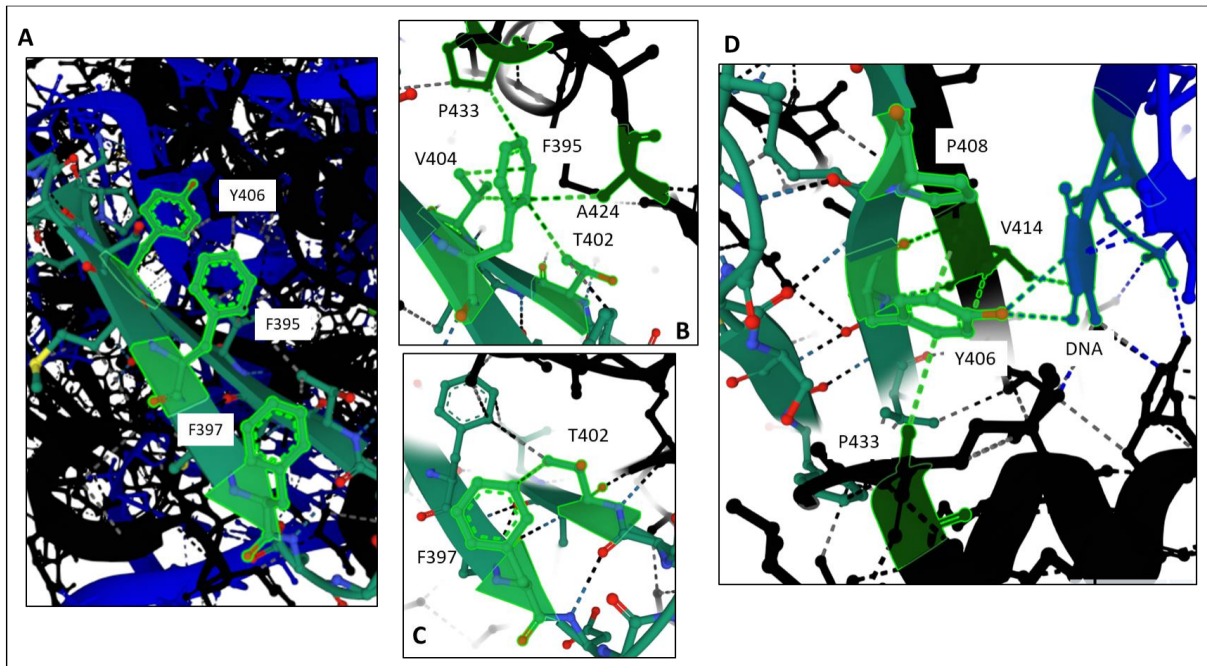
Figure 11: *Excerpt of the cryo-EM structure of the PB strand transfer complex showcasing the hairpin spanning from residues V390 to P410 (green) and the transposon DNA (blue) (Protein Data Bank entry 6X67, [61]). Black: remaining PB transposase protein. Red balls: Oxygen, Blue balls: Nitrogen, Green balls: Carbon **A:** Side view of the hairpin structure (green) highlighting aromatic residues F395, F397 and Y406. **B:** Detailed view of hairpin residue F395 and its hydrophobic contacts with T402, V404, A424 and P433. **C:** Detailed view of hairpin residue F397 and its hydrophobic contact with T402. **D:** Detailed view of hairpin residue Y406 and its hydrogen bond mediated interaction with the transposon DNA (blue) as well as the hydrophobic contact with P408, V414 and P433.*

As depicted (Fig. 11, A) F395 and F397 are positioned alongside Y406 in an orientation that would suggest attractive parallel-displaced pi stacking between the aromatic rings of the respective residues, contributing substantially to the hairpin's β-sheet stability and orientation. Y406 orientation is of particular importance, since it is directly involved in transposon-DNA interaction with two hydrogen bonds, exerts a hydrophobic contact with V414 which is also implicated in transposon-DNA interaction, and exhibits additional hydrophobic contacts with P408, V414 and P433 (Fig. 11, D). Follow-up experiments replacing F397 with Y and R corroborate this hypothesis: F397Y did not substantially influence transposition activity as the aromatic ring in Y397 might have replaced the stacking properties of F397. F397R however generated an inactive transposase variant. Moreover, F395 exhibits a substantial number of hydrophobic contacts with residues V404, A424, P433 (Fig. 11, B) and sharing a contact with F397 at T402 (Fig. 11, C). This observation highly suggests essential roles in structure integrity, directly linked to transposase function by transposon-DNA interaction through Y406. Accordingly, replacement of F395 and F397 with an A might have critically compromised said interplay and lead to a loss of transposition function. It can be assumed, that the hairpin integrity itself is critical in mediating BRD4 interaction but that F395 and F397 are not directly involved in said interaction The remaining two alanine replacement mutants displaying reduced

interaction behaviors with BRD4 in the presented Co-IP experiments L403A and K407A exhibit a less obvious reason for disrupting the hairpin structure and the transposase function.
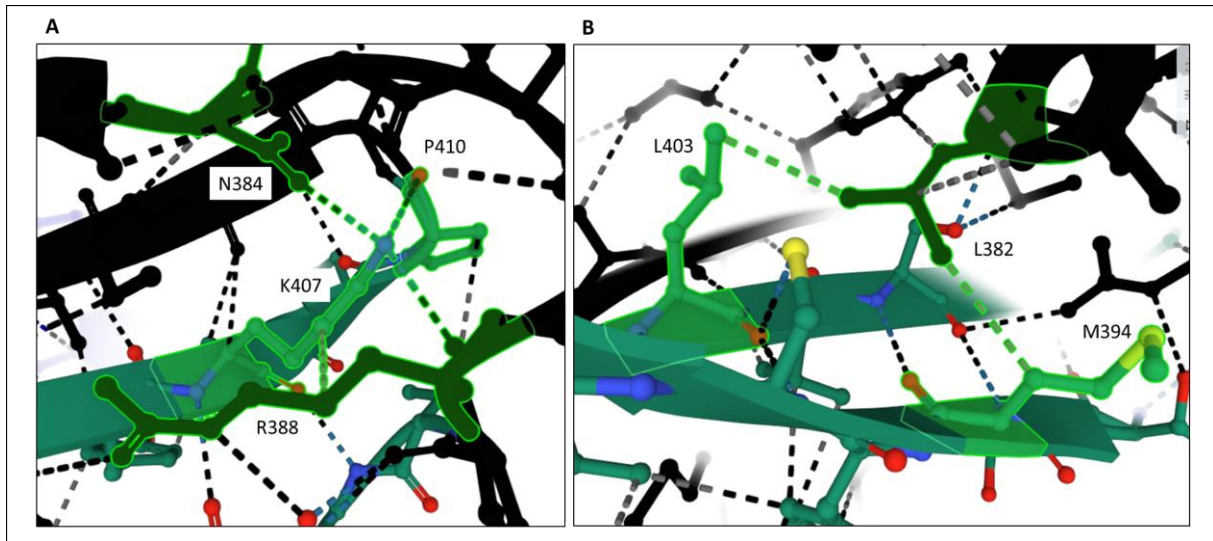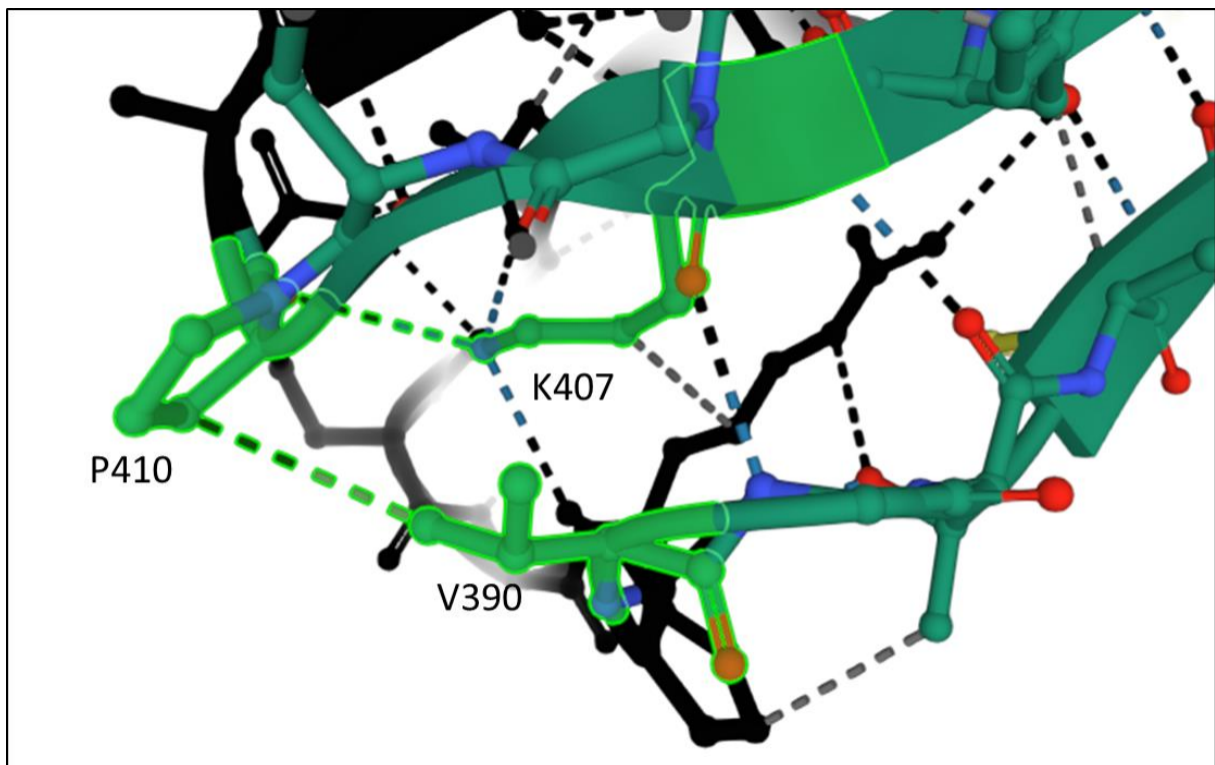

Figure 12: Excerpt of the cryo-EM structure of the PB strand transfer complex showcasing a section of the hairpin spanning from residues V390 to P410 (green) (Protein Data Bank entry 6X67, [61]). Black: remaining PB transposase protein. Red balls: Oxygen, blue balls: Nitrogen, yellow balls: sulfur A: Detailed view of hairpin residue K407 and its hydrogen bonds with P410, N384 and R388 as well as hydrophobic contact with R388. B: Detailed view of hairpin residue L403 and its hydrophobic contact with L382 as well as the hydrophobic contact between L382 and M394.

It can be assumed that K407 has a significant influence over the overall hairpin stability, being involved in a hydrophobic contact with R388 but mainly exhibiting three hydrogen bonds with N384, R388 and P410 respectively (Fig. 12, A). Except for the hydrogen bond with P410 however, none of the interactions are between K407 and a component of the hairpin itself. Thus, K407A might have influenced the hairpin's orientation significantly losing its possible function as an anchor with N384, R388 and P410 and thus affected BRD4 interaction as well as transposition efficiency. It may however not have influenced the hairpin's integrity itself, which may contribute to the transpositional efficiency rescue by employing other substitutions at position K407 as discussed in the following section. L403A's effect is less obvious however and subject to greater speculation. Through L403's hydrophobic contact with L382 it appears to contribute to the hairpin's stability. Interestingly however, it could play an important role in stabilizing L382 which in turn stabilizes M394 through a hydrophobic contact. M394 was shown to play a major role in BRD4 interaction (Fig. 7) and the mutant M394R was the candidate which had the greatest impact in shifting PB's integration pattern away from TSS (Fig. 9 & 10) as discussed in a later section.

6.4.2. Alternative PB-hairpin mutants

As experimental data strongly indicated that the hairpins structural integrity is essential for maintaining transposition function, two strategies were followed to generate alternative aa replacement mutants that would not compromise the overall structure. The first strategy encompassed finding alternative aa substitutions that were present in the highly conserved hairpin structure across members of the PB superfamily (Fig. 5) at positions that had shown a reduced interaction behavior in Co-IP experiments when substituted with alanine. The second strategy suggested introducing large and charged aa residues at positions that would not compromise the hairpins structural integrity but influence the structures surface through which the BRD4 interaction is mediated. Out of all candidates tested, substitutions E380F, V390K, T392Y, M394R, K407C, K407Q, K407V had a substantial impact on reducing BRD4 interaction as seen in follow-up Co-IP experiments while simultaneously maintaining transpositional activity. Interestingly, K407C, K407Q and K407V did not suppress the transposition activity, in contrast to K407A. While losing its hydrogen bonds with N384, R388 and P410 (Fig. 12), the replacing residues might have maintained the hydrophobic contact with R388 providing enough stability to maintain function yet influenced the structure sufficiently enough to influence BRD4 interaction.

Figure 13: Excerpt of the cryo-EM structure of the PB strand transfer complex showcasing a section of the hairpin spanning from residues V390 to P410 (green) (Protein Data Bank entry 6X67, [61]). Black: remaining PB transposase protein. Red balls: Oxygen, blue balls: Nitrogen, yellow balls: sulfur. A detailed view of aa V390s hydrophobic contact with P410 as well as K407 hydrogen bond with P410 are shown.

Out of the remaining four positions, V390 is the only aa residue that is involved in an interaction with a member of the hairpin sequence in form of a hydrophobic contact with P410 (Fig. 13). The observation that both aa residues that interact with the P410 residue are implicated in reducing the BRD4 interaction when mutated (V390 and K407, Fig. 7), together with the observation that a P410_A411del produced an inactive PB transposase (Fig. 6), strongly suggest that both transpositional activity and BRD4 interaction are linked to the hairpins integrity since prolines are known to heavily influence protein folding, disrupting secondary structures by inhibiting alpha-helix or beta-sheet conformation through its restricted phi angular range in peptide bond formation (reviewed in [123]).
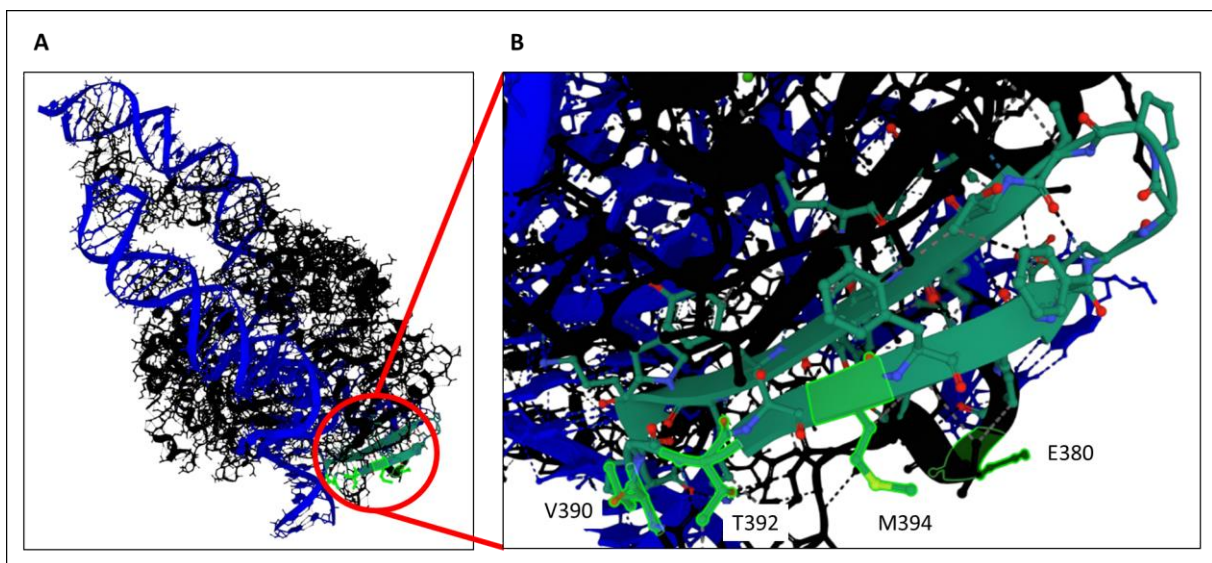
*Figure 14: PB strand transfer complex (A) (Protein Data Bank entry 6X67, [61]) containing hairpin structure spanning from V390 to P410 (green). Blue: Transposon DNA, black: remaining PB transposase protein. **B**: Detailed view of hairpin structure (green) and the positions of residues E380 V390, T392 and M394 (highlighted in green). Red balls: oxygen, blue balls: nitrogen, yellow balls: sulfur, green balls: carbon.*

The remaining three side chains of M380, T392, M394 together with V390 are all situated at the protein's surface (Fig. 14), thus having the innate ability to be involved in protein-protein interactions. Furthermore, side chains of E380, T392 and M394 do not showcase any intermolecular interactions with aa members of the hairpin suggesting minimal to no involvement in its integrity. Thus, they represent optimal candidates to be exchanged for charged and/or bulky side chains, modifying the hairpins surface without compromising its structure.

Accordingly, all mutations at discussed positions that produced a visible drop in BRD4 interaction signal during the Co-IP experiments while maintaining transposition activity were used for integration library generation and subsequent analysis to determine potential shifts in insertion preferences.

6.4.3. Impact of PB-mutants on integration bias

Following the hypothesis that the insertion bias PB exhibits towards TSSs is based on a physical interaction with BET proteins that tethers transposition events towards BET-binding sites, generating transpositionally active PB-mutants with measurable reductions in BRD4 interaction as a BET representative, led in all instances to the desired outcome of reducing the insertion frequency within the strict proximity of TSSs in comparison to the wildtype transposase (Fig. 9). Within a 400 bp window around TSS insertion frequencies were reduced between 21% and 65%. Surprisingly, increasing the size of the analyzed windows around TSSs, the reduction is not maintained for all mutans. E380F and K407Q exhibit an increment of insertional bias if considering window sizes between 800 bp and 600 bp around TSSs. A correlation between reduction of insertional bias and reduction of BRD4 interaction behavior can however be generally made with one exception: V390K. Out of all employed PB-mutants it exhibited the lowest reduction in BRD4 interaction in Co-IP experiments (Fig. 7), yet belonged to the only three mutants that exhibited a reduction in interaction bias throughout the entire 6 kb window analyzed around TSSs together with M394R and T392Y. The former showed together with K407V the highest reduction in interaction behavior in Co-IP experiments, surpassing 90% (Tab. 7). While not exhibiting an integration bias reduction throughout the entire 6 kb window, K407V remained closest to neutral and up to a window size of 2,8 kb around TSSs exhibited a reduced insertion preference. The remaining two mutants showed higher insertion frequencies in all window sizes except the smallest of 400 bp around TSS, correlating with the smallest reduction in BRD4 interaction behavior seen in Co-IP experiments (Fig. 10). Considering the lack of correlation between insertion profile shift away from TSSs and reduction in BRD4 interaction exhibited by V390K it is important to note key factors which might influence this single exception. Generally, Co-IP experiments are highly useful to gain insights on protein-protein interactions. The procedure itself however is cumbersome and dependent on several handling steps throughout a multi-day protocol, making quantifiable results at the end of the procedure highly dependent on a high degree of consistency during the experimental steps. While highly speculative and also applying to every other sample, it needs to be acknowledged that quantitative Co-IP results as presented in this work have to be interpreted with caution. For the project itself however, they served the important purpose of pre-screening possible candidates to use for the generation of insertion libraries and effectively lead to the discovery of several insertion preference modified PB-mutants. A more likely explanation for the correlation exception exhibited by V390K can be that only one BET protein member was used to pre-screen mutants via Co-IP to find suitable candidates: BRD4. It is known that PB

additionally interacts with BRD2 which is also implicated in recruiting transcription factors and regulatory complexes, activating gene expression [124–126] and thus also is involved in the hypothesis of how PB is tethered towards TSSs. The determination of interaction properties with BRD2 were not subject of this work. Nevertheless, the generated mutants based on Co-IP experiments with BRD4 might have in turn influenced the interactions with BRD2 in unknown ways. While the BRD4 interaction after introducing mutation V390K presented the least drop of all mutants tested, it is unknown how the interaction with BRD2 may have been shifted, potentially adding to the reduction of integration bias. Therefore, it would be highly interesting to produce Co-IP data similarly to the one presented in this work but with BRD2, to gain further insight into the mechanisms behind insertion preference selection and possibly introducing another factor influencing mutant design to further improve the reduction of insertional bias towards TSSs.

Interestingly, Co-IP experiments revealed that iPB-C3, which includes three C-terminal mutations (S509G, N538K and N570S) and were implicated previously in reducing the integration bias towards TSSs [121] did not exhibit an effect on reducing BRD4 interaction in Co-IP experiments (Fig. 4). Since the hypothesis driving this project is based on a correlation between BET protein interaction and insertional bias towards TSSs, the three C-terminal mutations were not included in the mutants used in the library generation. However, since a reduction in insertional bias was experimentally proven previously [121], there is strong evidence that additional factors besides BET interaction properties can influence PB's integration preferences.

Changing the integration preference, despite being the primary project goal, was not the only modification of the insertion profile produced by the presented PB-mutants that was observed. The wildtype PB used as an internal control for the integration library preparation exhibited a single peak in insertion frequency near TSSs as expected (Fig. 9). However, every single mutant tested exhibited two: directly up- and downstream of the TSSs position (Fig. 9). This pattern is highly reminiscent of the integration profile seen with retroviral vectors such as MLV, where a link between the MLV integrases preference to insert its cargo into nucleosome-associated DNA [127,128] has been proposed [111]. In contrast to the MLV integrase, PB is known to integrate preferably into nucleosome-free DNA [111]. Since nucleosome-depleted DNA is associated to TSSs of actively transcribed genes [129], this observation poses a highly valuable insight into a possible secondary effect of the mutations generated in this work: not only retargeting insertions away from TSSs but also effectively from TSSs of actively expressed genes. This secondary effect would, if confirmed, present an additional safety benefit adding to

the risk reduction goal presented here. It is expected that the bioinformatic dataset generated throughout this project contains all necessary information to answer not only this question but additionally, the integration preferences into the proximity of other problematic regulatory-associated regions that PB is known to favor such as DNaseI hypersensitivity sites, H3K4me3 marks and RNA-Polymerase II-bound regions [68,112,130–133]. Specifically, integration preference into the BACH2 gene would be of great interest since it is a problematic preference shared with MLV [111] and has been found in all instances of lymphomagenesis of therapeutic CAR T-cell products generated with PB transposase technology [134]. It is projected that moving forward, these insights will become increasingly relevant when deciding which genetic engineering tool to choose and that he ground work presented here for the PB transposon system provides valuable insights for the goal of improving the safety aspects of the system.

6.5. Conclusion and outlook chapter I

This project set out to confirm the hypothesis that the insertional bias towards TSSs that the PB transposon system exhibits is linked to a physical interaction between the PB transposase protein and BET proteins that co-localize at these sites and that disrupting said interaction influences the insertion profile by reducing the tethering effect. For this purpose, the proposed hairpin element identified within the PB transposase spanning aa's V390-P410 was subjected to an alanine scan, followed by Co-IP experiments that experimentally confirmed the structure's central role in the interaction behavior. Additionally, it suggested a critical role of the hairpins structural integrity for transpositional activity, which was confirmed during the course of this work by published cryo-EM structures for the PB transposase revealing the structural basis for PB unique transposition mechanism and experimentally confirmed the hairpin as part of the transposase's insertion domain which, together with the dimerization, DNA binding and catalytic domains, direct cleavable phosphodiester bonds within the TIRs to the active site's DDD motif [61]. Accordingly, rational design modifications were conducted by investigating alternative aa substitutions at interaction-relevant sites identified via alanine scan and found in PB-superfamily members and prediction-based substitutions theoretically maintaining hairpin integrity while introducing large surface changes with respect to size and/or charge. Both approaches were successful and yielded active PB-variants with comparable transpositional activity as seen by the number of recovered integrations during the insertion library generation (Tab. 3), and reduced BRD4 interaction profiles (Fig. 7). Finally, the initial hypothesis was confirmed via integration library generation and insertion site analysis, where a substantially lower numbers of integrations close to TSSs were observed. Furthermore, one particular mutant was generated (mPB-M394R) which exhibited an exquisite reduction in integration bias in every window size up to 6 kb around TSSs.

While this achievement is expected to represent a substantial improvement over the wildtype PB integration profile with respect to lowering the risk of insertional mutagenesis caused by TSS disruption affecting gene expression, this study exposes the potential of further reducing the insertional bias to close the gap as much as possible to an ideally random integration site distribution. For this purpose, several strategies should be considered for future developments. First, as a total of seven transpositionally active PB mutants were generated throughout this project that exhibited a modified insertion profile, all of them a strong reduction within the immediate proximity of TSSs, it should be considered to produce PB versions that include combinations of said mutations. As discussed, it is expected that the disruption of PB-BRD4 interaction was caused by the modification of the protein's surface maintaining hairpin

integrity, it is likely that the disruptive effect is additive and thus can be further improved. To test this assumption, the total potential of insertion bias reduction mediated by PB-BET disruption could be experimentally assessed by analyzing integration libraries that are generated while PB-BET interaction is artificially yet fully disrupted by the addition of BET inhibitors such as JQ1. These results could be compared with this study's results to determine the integration bias difference, indirectly defining the leftover potential. Second, the interaction behavior with other members of the BET family (BRD2 and BRD3) should be assessed, as interaction with BRD2 has previously been described [111], could hypothetically also be implicated in a tethering effect and it remains unclear, if a BRD4 interaction disruption caused by the introduced mutations would have the same effect with regards to related proteins. Accordingly, mutations described here that did not influence BRD4 interaction might have influenced interactions with other BET members and thus should be considered. Third, through the initial description of iPB-C3 an unintended yet independent proof of mutation-caused redirection of insertions has been described that significantly reduced integrations within TSSs, CpG islands and within genes [121]. As described in this study, the three C-terminal mutations included in iPB-C3 did not have an effect on reducing BRD4 interaction (Fig. 4). This strongly suggest that there are additional factors influencing PB's target site selection besides the one addressed here. Thus, effort should be made to identify the mechanism behind iPB-C3's phenotype with respect to target site distribution, and then to combine these mutants to potentially achieve an additive impact.

Overall, this work presents the first instance of experimentally confirming a mechanism behind the PB-BRD4 interaction and describes the generation of functional PB-variants that include modifications reducing said interaction leading to the desired outcome of changing PB's risk associated integration profile. It is predicted that this contribution will lead to an effort to further improve PB's potential to become a safe alternative for a transposon-based genetic engineering tool, which is currently under the spotlight due to the still unresolved association between PB gene transfer and the development of lymphoma cases, increasing the repertoire of cost effective and thus more widely adoptable technologies to provide gene and cell therapy.

7. **Chapter II: development of a novel, highly effective CRISPR/Cas9-based cellular suicide-switch**

7.1. **Introduction chapter II**

7.1.1. **Suicide-switches**

While modern gene and cell therapies bring novel treatment possibilities to previously untreatable conditions, potential dangers are extensively being discussed and subjected to stringent evaluation by the regulatory authorities when considering clinical trial and product approval. While efforts are being led to make vector technology safer, including the first chapter of this thesis, options will always be required to address unforeseeable negative side effects that might arise once the transgenic cell product has been administered to the patient, independently of how refined and optimized technologies might become with regards to their safety profile. An elegant way of including an exit strategy to the product itself has been the introduction of suicide-switches to the therapeutic vectors that are employed. Ideally, these switches will remain dormant within the modified cells until needed. If serious adverse effects were to occur and the removal of the transgenic cells could reduce or eliminate them, the suicide-switch can be activated, killing all cells harboring it and effectively clearing the patient of the living drug. Several options for such suicide-switches exist today, each with its advantages and disadvantages and they will be discussed in the following sections.

7.1.2. **Herpes Simplex Virus thymidine kinase**

The thymidine kinase (TK) is a phosphotransferase that catalyzes the first of three phosphorylations of thymidine to yield thymidine monophosphate, a precursor building block for DNA synthesis. The cellular TK is highly specific in choosing thymidine as its substrate. The TK belonging to the Herpes Simplex Virus (HSV) however, exhibits a broader specificity in its substrate choice, including purine and pyrimidine analogs [135,136]. If the HSV-TK phosphorylates these analogs that can be administered as a prodrug *in vitro* or *in vivo* such as the antiviral drug ganciclovir, the phosphorylated product will be further phosphorylated by cellular thymidylate kinases and nucleoside diphosphate kinases and finally block DNA synthesis once incorporated in nascent DNA by generating dead-end complexes [137], ultimately leading to cell cycle arrest [138] and apoptosis [139]. This trait has led the HSV-TK to be used as an efficient cytostatic agent and suicide gene to alleviate side effects by eliminating cell therapy products from affected patients [140]. Nevertheless, several

shortcomings require improvements for it to be an ideal solution. As a viral protein, the HSV-TK represents a xenoantigen and its intracellular presence required for its function can lead to strong responses in immunocompetent patients and rejection of transplanted cell products [141]. This would be less of a disadvantage in immunocompromised patients. However, as a weakened immune system enhances the risk of viral infections which might call for ganciclovir administration, the therapeutic cell population would unintentionally be affected. Additionally, since the HSV-TK leads to a blockade in DNA synthesis, the killing mechanism is cell-cycle dependent and not suitable for non-dividing cells such as notably chemoresistant, dormant, cancer-initiating stem cells (reviewed in [142]).

7.1.3. Inducible Caspase 9

The current state-of-the-art suicide-switch available is the inducible caspase 9 (iCasp9) system [143]. It was originally developed as a safety switch more suitable for adoptive therapy than the HSV-TK system and consists of a fusion protein composed of the caspase 9 protein that initiates apoptosis and the human FK506 binding protein (FKPB) [144]. The fusion protein remains in its monomeric form until the small molecule inducer AP1903 is added and binds to FKPB, causing cross-linkage between monomers and thus prompting dimerization, upon which caspase 9 initiates the apoptotic pathway. In contrast to the HSV-TK system, as iCasp9 is based on endogenous proteins, no immune response against the switches components is triggered. Additionally, AP1903 has been thoroughly tested and been found to be safe and well tolerated, not causing any drug related adverse effects [145]. This is not the case for ganciclovir, which is associated with hematological and neurological toxicities (reviewed in [146]). Furthermore, iCasp9 acts considerably faster than HSV-TK, generating detectable levels of the apoptotic marker Annexin-V 30 min. upon drug administration [147], since it is not dependent on cell division and directly induces an apoptotic cascade. Accordingly, the iCasp9 system has transitioned into the clinical world and has been successfully employed to handle graft-versus-host disease (GvHD) in recipients that received donor T cells that were genetically modified to specifically include the iCas9 safety-switch for this scenario [148]. Although GvHD was terminated, only 90% of the total modified cells were effectively killed and the remaining cells were still detected in long-term follow-up 3,5 years later. While GvHD did not recur, 90% cell clearance might not be sufficient to treat all potential adverse effects, especially malignant transformations, and complete cell clearance has not been observed *in vivo* (reviewed in [149]).

7.1.4. Epitope-based elimination

A third system that has been frequently used in clinical settings is formally not considered a suicide-switch but similarly allows for depletion of gene modified cell product upon

administration of a compound when adverse effects warrant it: the use of an epitope-based elimination marker. For this purpose, a cell surface marker which is ideally non-immunogenic and not present either in non-modified cells of the same type or anywhere else is added to the transgene cassette, effectively labeling every therapeutic cell. This is not only useful to quantify, monitor, select and sort the product during manufacturing and once administered. It also allows riddance by administration of marker specific clinically approved therapeutic antibodies, leading to antibody- or complement dependent cytotoxicity toward marked cells. Tested surface markers include CD20 [150] a truncated version of the human epidermal growth factor (hu-EGFRt) [151] and a combined target epitope of CD20 and CD34 termed RQR8 [152]. While efficiency and speed of elimination of epitope-based elimination markers has been shown to be equal to the iCasp9 system [147], potential bottlenecks are to be expected when used in heavily immunocompromised patients [153], worsening of adverse immune reactions including cytokine release syndrome (CRS) (reviewed in [154]) and poor biodistribution and tissue penetration [155].

Given the shortcomings of the most widely used suicide mechanisms for gene and cell therapeutic applications described previously, this work aims at providing an alternative that addresses them, in particular the incomplete cell elimination. For this purpose, a novel approach will be pursued to engineer an inducible CRISPR/Cas9 technology-based suicide-switch.

7.1.5. CRISPR/Cas9- and sgRNA-mediated targeting

The CRISPR/Cas9 system, now known worldwide as a Nobel-Prize winning breakthrough genetic engineering tool to generate precise genome modifications, is a repurposed component of bacterial and archaeal adaptive immune defense against invasive nucleic acids including plasmids and viruses. For that purpose, the CRISPR system is able to detect incoming foreign nucleic acids by comparing them with a roster of previously encountered sequences (CRISPR array), add novel foreign sequences to the list and fend off invading agents by fragmenting and neutralizing their DNA [156,157] or RNA [158] if matching sequences are encountered. This process is usually categorized into three distinct stages. In the first stage, adaptation, invading nucleic acid are recognized and short sequence fragments called protospacers are acquired and incorporated into the CRISPR array by Cas1 and Cas2, resulting in spacer sequences within the array [159]. In the second step, expression, the CRISPR arrays are transcribed forming precursor CRISPR (pre-cr) RNA and processed into shorter creRNA via Cas6 nuclease-mediated fragmentation [160]. creRNA contains complementary sequences to the invading agent, spacer sequences and a repeat sequence which allows association to transactivating crRNA (tracrRNA) [161]. Both guide (g) RNAs associate to the Cas9 nuclease (in the most

widely used type II systems [162]) and form a complex, which probes nucleic acids for sequences matching the spacer within the crRNA and a suitable protospacer-adjacent motif (PAM) sequence, consisting of a trinucleotide NGG [163]. In the third and final step, interference, if matching sequences and suitable PAM motif are available, Cas9 cleaves the target DNA [157], fragmenting the intruder.

As mentioned before, this system has been exploited in the recent years to become a versatile tool for genetic engineering. By supplying synthetic gRNAs, Cas9 can be easily "programmed" to cut any sequence of interest, only limited by the availability of an adjacent PAM sequence. The development of a chimera of tracrRNA and crRNA which can be supplied as a single guide (sg) RNA, further improved ease of use [157]. Genetic engineering with the CRISPR/Cas9 system usually revolves around the way cells overcome DSBs: Non-homologous end joining (NHEJ) or homology directed repair (HDR). If a DSB is repaired via NHEJ, insertions or deletions (indels) are often introduced as a byproduct of the repair mechanism. If these indel mutations occur within a target gene, frameshifts will cause disruption of gene expression, making this method highly effective at knocking out genes or screening for unknown functions by observing the resulting phenotypic changes. HDR on the other hand can be used to introduce DNA sequences into a target cell's genome. By supplying synthetic DNA fragments with flanking sequences that match up- and downstream of the generated DSB, HDR can use these fragments to repair the cuts, effectively introducing the provided foreign sequences stably into the genome (reviewed in [164]).
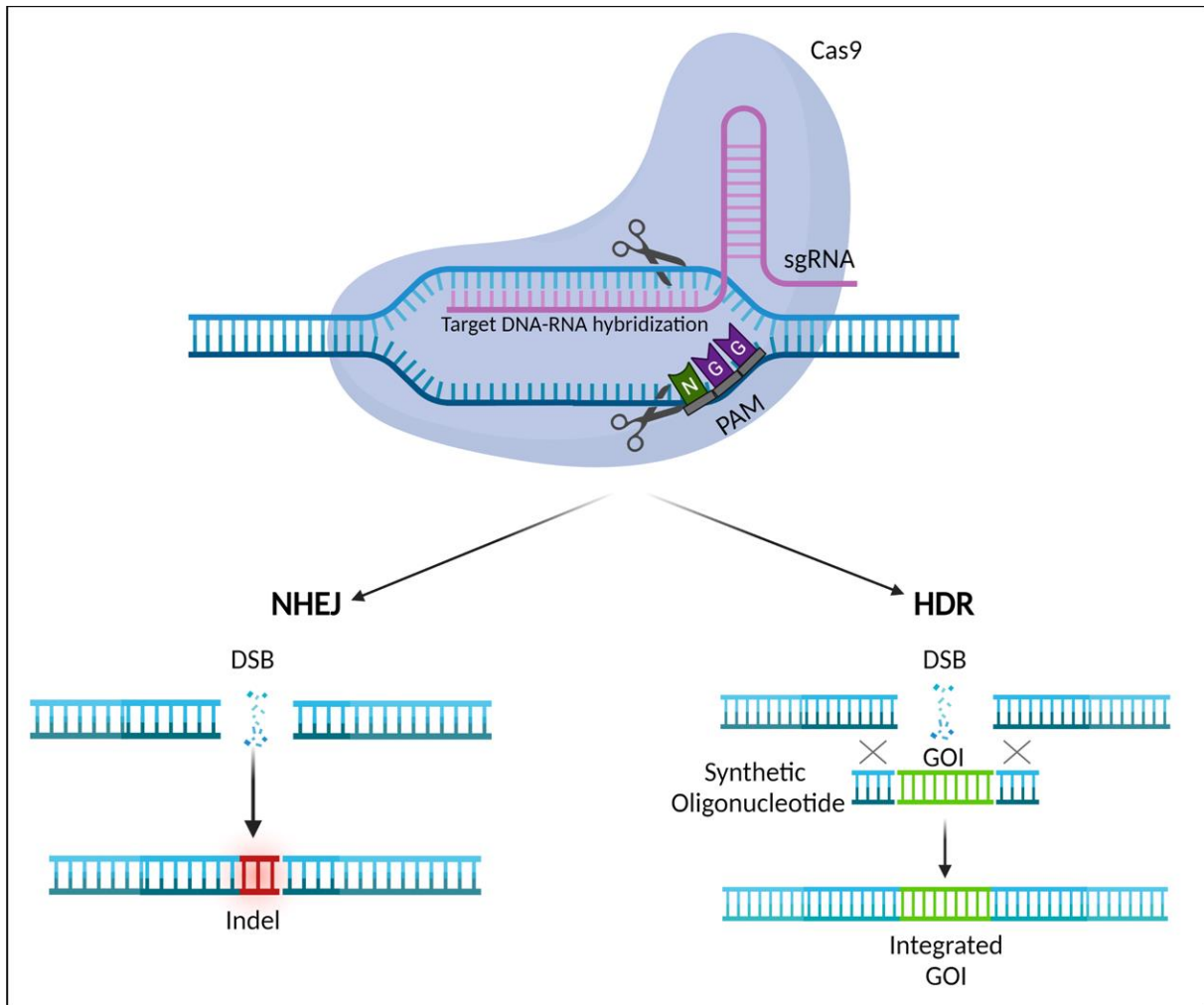
Figure 15*: Schematic overview of CRISPR/Cas9-mediated genome modifications. The CRISPR/Cas9 complex is guided to the target site via the target site complementary sgRNA sequence. After hybridization, Cas9 mediates DSB induction next to the NGG PAM sequence. After DSB induction, the damaged site is either repaired via NHEJ, introducing indels or via HDR when sequences with matching flanking regions to the break are available, opening up the possibility of introducing a GOI into the damaged site. (Figure was created using Biorender.com).*

This method has converted CRISPR/Cas9 into a valuable option to introduce therapeutic sequences for clinical applications such as integrative viral vectors and transposons, with the added benefit of precisely defining the location of the insertion beforehand. Accordingly, several preclinical studies have been conducted (reviewed in [165]), several of them culminating in first-in-human clinical trials using CRISPR/Cas technology (reviewed in [166]). Most preclinical studies address gene correction of monogenic disease models which include cystic fibrosis [167,168], sickle cell anemia [169–171], beta-thalassemia [172–176], Huntington's disease [177–179], Duchenne muscular dystrophy [180,181], hemophilia A/B [182,183], amyotrophic lateral sclerosis [184] and chronic granulomatous disease [185,186], but also multifactorial diseases such as diabetes [187], cardiovascular diseases [188], AIDS (reviewed in [189]) and cancer (reviewed in [190]).

Given that the technology is fairly new, the amount of clinical research that utilizes CRISPR/Cas9 is remarkable. Even more impressive is the fact, that a substantial number of clinical trials is already being conducted making use of its potential. Pilot studies translated the knowledge accumulated over the years of checkpoint inhibitors in cancer therapy, such as the effect of blocking the binding of the programmed death-ligand (PD-L1), expressed by numerous tumors, to the programmed cell death protein 1 (PD-1), with anti PD-1 antibody administration [191,192] to CIRISPR technology. Pilot studies showed the feasibility and safety of *ex vivo* CRISPR modification of patients' T-cells to knock-out PD-1 and re-administering them (ClinicalTrials.gov Identifier NCT02793856, [193]), circumventing the need of PD-1 blockade via therapeutic antibodies. Another early trial successfully piloted combining the PD-1 knockout with an additional endogenous T-cell receptor (TCR) knockout, reducing TCR mispairing and enhancing the expression of a synthetic cancer-specific TCR (ClinicalTrials.gov Identifier: NCT03399448, [48]). Furthermore, several clinical trials involving site-directed CAR T-cell integration into the T-cell receptor α-chain constant locus or replacing endogenous TCRs with synthetic TCRs or CARs are currently being planned/conducted (reviewed in [190]). Most recently, within the context of the ongoing SARS-CoV-2 pandemic, in an effort to provide long-lasting immunity and prevent T-cell exhaustion, PD-1 and ACE2 knockout T-cells will be generated with CRISPR/Cas9 and re-administered to the patient in a soon-to-be-recruiting phase I clinical trial (ClinicalTrials.gov Identifier: NCT04990557).

While incredibly adaptable and representing a tool of incomparable precision within genetic engineering, a growing number of concerns might have to be addressed before CRISPR/Cas9-based gene therapy will be available to the masses. One area of concern is the occurrence of *off-target* effects. This phenomenon can potentially cause substantial and difficult to detect effects on the host cells' genome, as unintended DSB at unforeseeable sites can lead to similar disruptive consequences as on the target site. Homology between the ~20 nucleotide (nt) long sgRNA sequence and genomic sequences other than the target site can typically cause this. Mismatches of up to 5 bp have been tolerated [194]. There is, however, a correlation between the number, distribution and position of mismatches in the sgRNA-DNA recognition and the efficiency of *off-target* cleavage [194,195] allowing for current *in silico* models to make predictions on *off-target* cleavage probabilities and suggest optimized sgRNA design to maximize the chances of avoiding them (reviewed in [196]). Another substantial risk factor receiving growing attention in recent years, is the generation of large *on-target* structural variants (SV) and unwanted chromosomal rearrangements. Since CRISPR/Cas9 formally only mediates DSB generation and repair leading to the desired modification is mediated by

endogenous cellular mechanisms, the existence of multiple DSB within a genome generated either spontaneously, through multiple sgRNAs utilized in the experiment, through unintended *off-target* cleavages or a combination of several mentioned factors, can lead to the joining of distant cleavage sites and chromosomal translocations [197], chromosome loss [198], SV generation [199], chromosomal truncations [200] and chromothripsis [201]. Instances have indeed been reported in clinical trials [48]. So far, no malignant transformations have been reported in patients treated with CRISPR/Cas9-modified cells. However, these phenomena represent a substantial risk that have been heavily associated with oncogenesis and congenital disorders and have generated an urgent call on thoroughly assessing the safety CRISPR-based therapies before transitioning to larger patient groups.

7.1.6. Creation of an inducible CRISPR/Cas9 system that targets *Alu* elements and kills cells

For the intended purposes of this work, generating a CRISPR/Cas9-based suicide-switch, a highly unconventional usage of this tool was tested which is not affected by the before mentioned risks and shortcomings, but might actually benefit from them. Instead of exploiting the unique qualities that allow precise programmable DSB generation at single genomic sites, sgRNAs were employed that maximize the number of potential genomic targets, hypothesizing that reaching a sufficient number of CRISPR/Cas9-induced DSBs present the cell with an irreparable amount of DNA damage and thus inducing cell death. To achieve the goal of efficiently maximizing sites that can be targeted with an individual sgRNA sequence, *Alu* retrotransposon elements were chosen as a target.

Alu elements constitute highly repetitive, primate-specific short interspersed elements (SINE) of the retroelement class of transposable elements, which exceed one million copies in the human genome [202]. They represent an active SINE and make up almost 11% of our entire genome and belong to what is colloquially termed "junk DNA", although their function is starting to be elucidated as contributing to genetic diversity, influencing gene expression, but also driving disease development through insertional mutagenesis (reviewed in [203]). Their presence in such high copy numbers and repetitive sequence nature, however, position them as a prime candidate to achieve maximal potential target sites through a specifically designed sgRNA which is expected to overwhelm the DNA damage repair capacity of the affected cell and in turn cause cell death (Fig. 16).
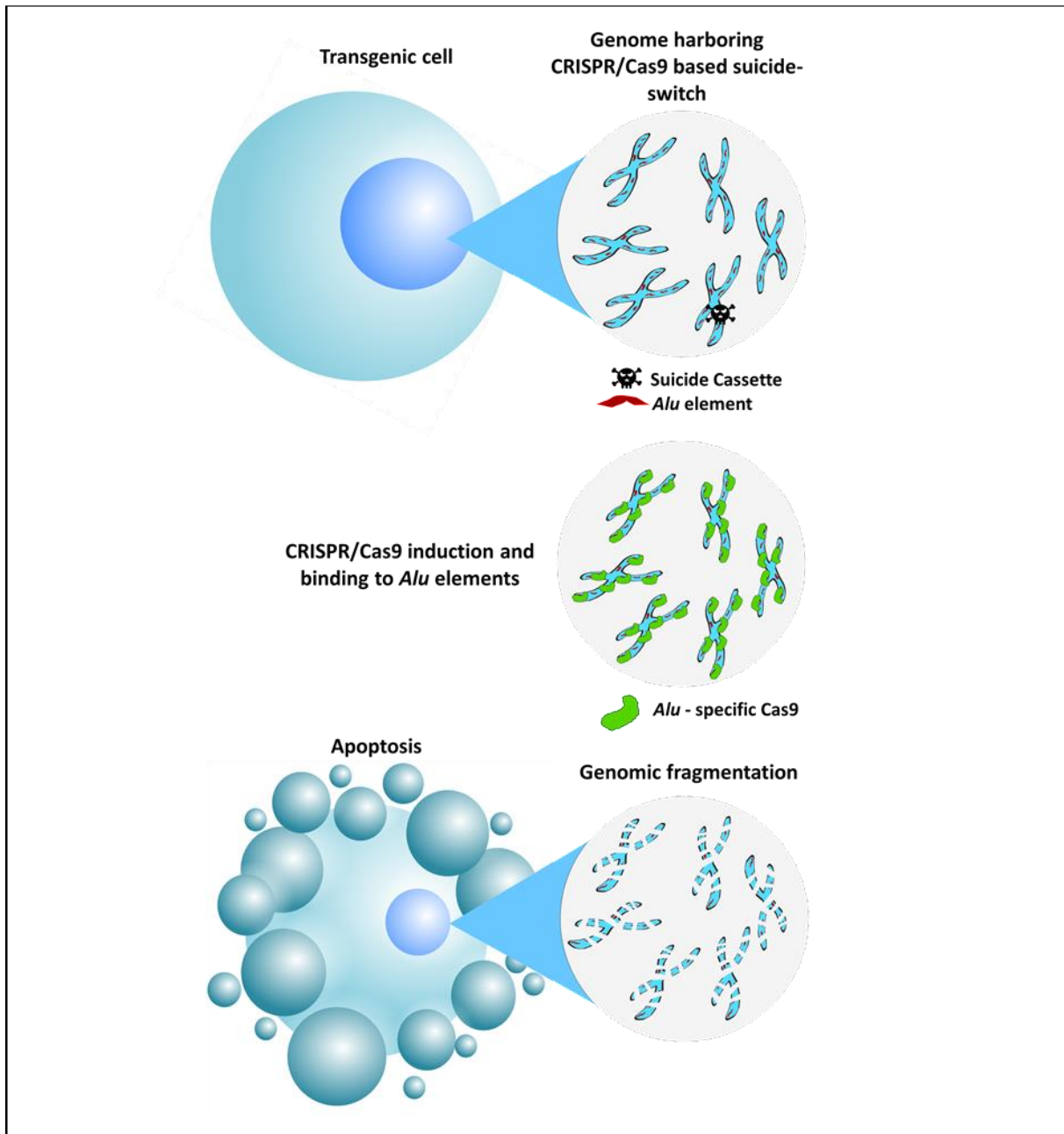
Figure 16: Overview of the envisioned CRISPR/Cas9 based suicide cassette. Modified cells harbor the inducible Alu-specific Cas9 suicide-switch within them. Once induced, the Alu-specific Cas9 is expressed and binds to the repetitive Alu elements present in multiple copies throughout the human genome, generating DSBs and effectively fragmenting the host cells genome, leading to apoptosis.

7.1.7. DNA damage induced apoptosis

Inducing cell death through the means of producing a high number of DSBs is a procedure that has a long history in medicine. Alongside chemotherapy and surgery, radiotherapy is a pillar of current cancer treatment. By focusing ionizing radiation onto cancerous tissue, critical numbers of DSBs are generated (reviewed in [204]). To prevent genomic instability, if the DNA damage exceeds the potential of the cell's DNA repair mechanisms, the p53 pathway leads to an accumulation of pro-apoptotic proteins that trigger programmed cell death (reviewed in [205]). Independently from the work presented here, other groups have succeeded recently in targeting

Alu elements with CRISPR/Cas9 with the purpose of inducing cell killing [206,207]. By supplementing Cas9 with *Alu*-specific sgRNAs, they were able to efficiently kill cancer cells in proof-of-concept studies, presenting the possibility of using this method to target difficult-to-treat cancer types. However, for the intended purposes of this work, their design does not represent an option for a suicide-switch, since the key feature of inducibility is missing. Thus, to fulfil the requirements for a working suicide-switch, a major challenge of this work was to engineer an *Alu*-targeted CRISPR/Cas9 system that can be stably introduced into a target cell, persists for an indefinite time without exhibiting unintended leakage and presents a reliable inducibility leading to complete clearance of cells harboring the switch. To achieve stable integration, the CRISPR/Cas9-based suicide-switch was designed as part of a non-viral SB transposon vector, enabling virus-free genetic engineering and circumventing viral vectors' cargo capacity limits, since a full-length CRISPR/Cas9 system would be close to the maximum cargo size of most viral vectors (reviewed in [208]) and adding a hypothetical therapeutic gene would exceed their capacity. Additionally, when envisioning a safety mechanism for potential use in gene therapy, the intrinsic advantages of non-viral gene transfer technology add an attractive benefit to the concept, which is the close-to-random integration pattern of the SB transposon system in comparison to viral systems or alternative transposon systems such as PB (as discussed in Chapter I of this thesis) [111]. This further decreases the possibility of insertional mutagenesis and thus reduces the likelihood having to use the suicide-switch. Lastly, concerning the inducibility, post-translational and transcriptional regulation were explored in order to achieve optimal induction response while maintaining no measurable leakiness, ensuring a sustained lack of unintended activity which would compromise cell fitness long-term. In an effort to identify regions within the Cas9 protein that allow insertions of foreign aa sequences without compromising sgRNA-dependent Cas9-mediated cleavage, a Cas9 mutant was generated termed arC9, which allows for reversible allosteric control over its DSB induction capability [209]. arC9 is devoid of a nuclear localization signal, but includes the Human Estrogen Receptor α Ligand Binding Domain (ER-LBD; residues 302–552 of ESR1) [210]. Through the ER-LBD, arC9 exhibits a structural conformation which inhibits arC9 cleavage and, since the NLS is missing, localizes in the cytoplasm. However, upon induction with the small molecule 4-Hydroxytamoxifen (4-HT), binding with the ER-LBS occurs, which produces a conformational change habilitating arC9 cleavage and doubles as an NLS, facilitating nuclear translocation. The proof-of-concept study showed remarkable induction capacity coupled with unmeasurable background activity. However, a single genomic site was targeted with an individual sgRNA. It is speculative how the background activity will translate

to a system where thousands of potential target sites will be available. Additionally, taking long-term uses into consideration, having a foreign protein of bacterial origin constantly expressed in a therapeutic cell poses a significant immunogenicity problem, threatening clearance of the product by the immune system once transplanted. For Cas9, this has been described to be the case (reviewed in [211]). Accordingly, a second control mechanism was explored in the form of a transcriptional control by putting arC9 expression under a tet-on system [212], conditioning arC9 transcription to doxycycline (DOX) presence.

7.1.8. Significance chapter II

Untargeted insertion of foreign genetic material inevitably disrupts the natural sequence composition of the genome of the hosts' cells. This can be accompanied with deleterious effects which, in a worst-case scenario, can lead to oncogenic transformation. In addition, in light of the recent advances in cell-based immunotherapy, important complications such as CRS accompanied by neurotoxicity call for extra careful consideration. A suicide-switch would provide an effective way of installing a final exit option into a therapeutic cell product in case the potential adverse effects outweigh the product's benefit. As discussed, current suicide-switch options still come with substantial room for improvement and in some cases, with unfixable roadblocks. This work aims at providing an unconventional alternative addressing these.

7.2. Methods chapter II

7.2.1. Cloning of *AluY1*-specific sgRNAs containing arC9 expression plasmids

For the construction of *AluY1* targeting arC9 expression plasmids, the plasmid pT2HB_arC9_COMBI_scaffold (Tab. 13) was linearized by digestion with BbsI (BbsI-HF, NEB) and subjected to agarose gel electrophoresis. The band corresponding to the fully digested plasmid was extracted and purified using a gel purification kit (Tab. 11) according to manufacturer's instructions. The purified DNA was dephosphorylated (Antarctic Phosphatase, NEB) according to manufacturer's recommendations and ligated to phosphorylated *AluY1*-specific sgRNA oligonucleotides (Tab. 14) using a T4 DNA ligase (NEB) following the manufacturer's instructions over night at 16 °C. 5 µl of the ligation reaction were directly transformed into chemically competent Subcloning Efficiency™ *E. coli* DH5α (Thermo Fisher) and incubated on Agar plates containing 100 µg/ml Ampicillin at 37 °C overnight. The next day, colonies were picked and 4 ml of LB liquid medium were inoculated for overnight growth in a shaking incubator at 37 °C and 200 rpm. The following day, plasmid preparations were conducted using a plasmid miniprep kit (Tab. 11) according to manufacturer's instructions and purified samples were sent for sanger sequencing (Eurofins) alongside sequencing primers (Tab. 14) to confirm the insertion of the *AluY1*-specific sgRNA. The sequence-verified version of this plasmid was termed pT2HB_arC9_COMBI_Alu.

7.2.2. Attempt to generate monoclonal HeLa cells harboring the arC9_COMBI_Alu transgene via *Sleeping Beauty* transposition

HeLa cells that were seeded a day before at a density of $2,5 \times 10^5$ cells per well on a 6-well plate were co-transfected with 1 µg of pT2HB_arC9_COMBI_Alu alongside 300 ng of the pCMV(CAT)T7-SB100X expression plasmid using the Lipofectamine® 3000 Transfection Kit (Thermo Fisher), according to the manufacturer's instructions. In parallel, similar transfections were conducted, yet missing the transposase expression plasmid component as a negative control, to determine the time point where transient expression, visualized by monitoring mCherry expression via fluorescence microscopy, is no longer taking place while stable expression, product of SB-mediated transposition of the transgene cassette into the HeLa genome, persists. One day post transfection, cells were passaged from each 6-well to individual 10 cm plates per transfection sample and mCherry fluorescence monitored daily. Single-clone picking was not attempted and experiment discontinued due to the suspicion of leaky arC9 expression driving cell death.

7.2.3. Cloning of pTOV_T11_Neo_arC9_Alu/scaff

The arC9 expression cassette was introduced together with the hU6 driven CRISPR scaffold ± *Alu*-specific sgRNA sequences into the pTOV_T11_Neo plasmid (Tab. 13) to enable SB-mediated transposition, DOX dependent arC9 expression via a tet-on system and G418 mediated selection of modified cells. For that purpose, the pTOV_T11_Neo plasmid was linearized with NotI-HF and SalI-HF (NEB) according to the manufacturer's instructions followed by agarose gel electrophoresis to confirm complete digestion. The band with the expected size was excised and the DNA contained was purified using a gel purification kit (Tab. 11). The arC9 gene was PCR-amplified using the pT2HB_arC9_COMBI_Alu plasmid as the template and oligonucleotide primers #93 and #94 (Tab. 14), introducing a stop codon at the 5'-end and matching restriction sites NotI and SalI for insertion into the previously linearized plasmid pTOV_T11_Neo, using a high fidelity Q5 Polymerase (NEB) following the manufacturer's instructions. For calculating optimal annealing temperatures, the tool Tm Calculator (https://tmcalculator.neb.com, NEB) was used. The PCR product was purified using a purification kit (Tab. 11). The eluted product was digested with NotI-HF and SalI-HF (NEB) according to the manufacturer's recommendations to generate matching overhangs. After digestion, the sample was heat inactivated for 20 min. at 80 °C, subjected to a 1,2 % agarose gel electrophoresis to confirm the correct size and afterwards excised and purified as described previously. The eluted product was subjected subjected alongside the eluted product from the pTOV_T11_Neo plasmid to a ligation reaction (T4 DNA Ligase, NEB) according to the manufacturer's recommendations for 16 °C overnight. The next day, 5 µl of the ligation reaction were transformed into chemically competent *E. coli*, plated, picked and grown, followed by plasmid preparations as previously described. Purified plasmids were sent for Sanger sequencing together with sequencing primers (Tab. 14) to verify the introduction of the arC9 coding sequence. Once a plasmid was identified, that harbored the correct sequence, this plasmid was linearized with NheI-HF and AgeI-HF (NEB) according to manufacturer's recommendations, separated on an agarose gel, excised and purified as previously described. The hU6 driven CRISPR scaffold ± *Alu*-specific sgRNA sequences were PCR amplified using plasmids pT2HB_arC9_COMBI_Alu or pT2HB_arC9_COMBI_scaff respectively as templates, with primers #95 and #96 (Tab. 14) in order to introduce NheI and AgeI restrictions sites. The cloning process, to introduce the hU6 driven CRISPR scaffold ± *Alu*-specific sgRNA sequences into the previously linearized plasmid was repeated similarly to the previously described procedure, except with restriction enzymes NheI-HF and AgeI-HF (NEB). The final

purified and sequence verified samples constitute plasmids pTOV_T11_Neo_arC9_Alu and pTOV_T11_Neo_arC9_scaff.

7.2.4. Generation of polyclonal HeLa cells harboring transcriptionally and post-translationally inducible suicide-switch CRISISS

To generate polyclonal HeLa cells harboring CRISISS suicide-switch, cells were transfected with 500 ng pTOV_T11_Neo_arC9_Alu and 300 ng SB100X expression plasmid (Tab. 13). For that purpose, the day before transfection, $2,5 \times 10^5$ HeLa cells were seeded on a well of a 6-well plate and transfected the day after via lipofection using the Lipofectamine® 3000 Transfection Kit (Thermo Fisher) following the manufacturer's instructions. The following day, cells were transferred to a 10 cm cell culture plate. Selection was started immediately by culturing cells in growth medium containing 1 mg/ml G418. Medium was regularly changed and cells washed with PBS to remove dead cells. Selection lasted a minimum of 14 days to eliminate cells exhibiting G418 resistance due to transient expression and not stable transgene integration.

7.2.5. arC9 expression confirmation via WB Analysis

$2,5 \times 10^5$ HeLa cells expected to harbor CRISISS (polyclonal or monoclonal, see 7.2.4. & 7.2.8. for details) were seeded on wells of a 6-well plate. DOX was directly added to the growth medium at a concentration of 1 µg/ml to induce arC9 expression. After 48h, growth medium was removed, cells were washed once with PBS and harvested directly in 250 µl of RIPA buffer (Tab. 5) by scraping the adherent cells with a cell scraper and resuspending the cell suspension with a micropipette. The cell suspension was moved to a microcentrifuge tube and placed in a rotating shaker at 20 rpm for 30 min. at 4 °C. Cell debris were sedimented by centrifugation at 20 000 × g for 10 min. in a cooled microcentrifuge at 4 °C. 200 µl of the supernatant was moved to a clean microcentrifuge tube. 20 µl were mixed with 10 µl 3X SDS sample loading buffer (Tab. 5) and heated at 95 °C for 5 min. Afterwards, the samples totaling 30 µl volume were subjected to an SDS-PAGE on an 8% polyacrylamide-gel. Once electrophoresis was complete, proteins were blotted on a nitrocellulose membrane. Membranes were blocked overnight at 4 °C on a shaking incubator in blocking buffer. The following day, the blocking buffer was discarded, the membrane cut below 100 kDa and individually incubated with the corresponding primary antibody (the lower membrane half with an anti-H3 or anti-Actin for the loading control, anti-Cas9 to detect arC9 (Tab. 4) for 90 min. at room temperature on a shaking incubator at 55 rpm. After incubation with the primary antibodies, these were removed and membranes washed three times with TBS-T for 10 min. at room temperature on a shaking incubator at 55 rpm. Afterwards, incubation with the corresponding HRP-conjugated secondary

antibodies (Tab. 4) was carried out similarly to primary antibody incubation except for a 60 min. incubation time instead. After secondary antibody incubation was finalized, membranes were similarly washed three times in TBS-T except for a 5 min. incubation per washing step. Finally, protein detection was carried out via chemoluminescent detection (Tab. 11) according to the manufacturer's instructions using a CCD-Imager (Tab. 10).

7.2.6. CRISISS induction and colony-forming kill-assay

Cells harboring the transposon containing the DOX and 4-HT inducible suicide-switch were seeded on cell culture plates as varying cell densities depending on setup, (disclosed in corresponding result section) and induced with 1 µg/ml DOX and 200 nM 4-HT (added directly into the cell culture medium). Medium and inducers were changed every two days for the entirety of the induction period. To determine the effectivity of the suicide mechanism after the induction period, the remaining cells on the culture plates were fixed, stained and analyzed similarly as described previously (see 6.2.3.).

7.2.7. Flow-cytometry based viability assay

The documentation of the apoptosis and cell death dynamics after CRISISS induction was performed via a flow-cytometry based viability assay. Wildtype HeLa cells were seeded on two 10 cm cell culture plates at densities of 1×10^6 and 2×10^6 per plate, containing the cells intended for the unstained samples, apoptosis background samples and the positive dead cell control samples. Cells for the remaining samples, namely the wildtype HeLa cells for the apoptosis positive control, the HeLa pTOV scaff + DOX induction, the HeLa pTOV Alu + DOX induction, the HeLa pTOV scaff +DOX +4-HT induction and the HeLa pTOV Alu +DOX +4-HT induction were seeded on wells of a 6-well plate at the following densities per well: 5×10^5 for the measurements at day 1 post induction, $2,5 \times 10^5$ for day 2 post induction, $1,25 \times 10^5$ cells for day 3 post induction and 5×10^4 for day 4 post induction. Induction with 1 µg/ml DOX and 200 nM 4-HT was started immediately after seeding on the corresponding samples. The medium was changed on day 2 post induction containing fresh DOX and 4-HT at the required concentrations. Apoptosis induction intended for the positive control was performed by adding camptothecin directly to the medium of the corresponding sample to a final concentration of 10 µM on the day before the corresponding measurement. On each measurement day, for the control samples except the apoptosis positive control, 1 million cells were harvested from one of the two 10 cm plates containing wildtype HeLa cells. The remaining cells were replated to be used for the same measurement 2 days later. On day 2 post induction the same procedure was repeated with the other 10 cm plate containing wildtype HeLa cells. For the positive dead-cell control, cells were harvested and heated to 65 °C for 20 min. For the remaining samples,

the supernatant of their corresponding well was removed each measurement day and dead cells were sedimented by centrifugation at 250 × g for 5 min. Remaining adherent cells were harvested and pooled with the corresponding dead cell pellet. Subsequently, all samples were washed with PBS once. The unstained samples and the apoptosis positive control sample were resuspended in 100 µl FACS buffer (Tab. 5) and passed through a cell strainer. The remaining samples were resuspended in 100 µl Zombie working solution (1:500 dilution in PBS) and incubated for 20 min. at RT in the dark. All samples were washed twice with 4 °C FACS buffer. Subsequently, staining of apoptotic cells was performed by resuspending corresponding cell samples in 1 ml annexin binding buffer (Tab. 5), mixing 100 µl of the resulting cell suspension with 5 µl of Pacific Blue™ labeled Annexin V (Tab. 11) by vortexing and incubating samples for 10 min. at RT in the dark. After the incubation period, 400 µl annexin binding buffer was added and the resulting 505 µl passed through a cell strainer. Flow cytometric measurements were performed on a BD LSR II Flow Cytometry Cell Analyzer (Tab. 10) and the Analysis using FloJo™ v10 (Tab. 17).

7.2.8. Generation of CRISISS harboring monoclonal HeLa cell line

The monoclonal HeLa cell line was generated based on the polyclonal cell line described in 7.2.4. Cells were seeded at a cell density of 5×10^3 per 10 cm cell culture plate. One day after seeding, solitary single cells were identified via microscopy and marked to monitor development daily. Once single cells developed into well-defined colonies, these were picked under the microscope with a 10 µl micropipette and moved to individual wells on a 24-well plate, where expansion was begun. Colonies were regularly trypsinized and cells transferred to a bigger well format (12-well, then 6-well). To ensure monoclonality, the entire process was repeated culminating at a cell population stemming from two consecutive single-colony picking processes. Finally, to confirm the presence and expression of the suicide-cassette in the generated monoclonal cell line, arC9 expression was induced and measured via WB analysis (see 7.2.5.).

7.2.9. Detection of KAP1 phosphorylation after CRISISS induction

DNA damage occurring through the induction of the CRISISS system was determined by detecting DNA-damage response via WB based on the detection of phosphorylated KAP1 (pKAP1). For that purpose, monoclonal HeLa cell line harboring CRISISS including *Alu*-specific sgRNAs (arC9 Alu HeLa clone #1) and without *Alu*-specific sgRNAs (arC9 scaff HeLa clone #1) were induced with 200 nM 4-HT and 1µg/ml DOX. Cells were harvested 1 day post induction (dpi), 2 dpi and 3 dpi alongside uninduced cells of both cell lines as controls and lysed in RIPA buffer including protease and phosphatase inhibitors (Tab. 5) at 1X concentration and

for 30 min. at 4 °C in a rotating shaker. Subsequently, a WB analysis was carried out in a similar way as described before with a few exceptions: i) 5 µg of total protein was loaded onto the acrylamide gel per sample ii) a 10% polyacrylamide gel was used iii) bovine serum albumin was used for the blocking buffer and the antibody dilutions (Tab. 4).

7.2.10. Determination of CRISISS kill-kinetics

To determine the kinetics of cells harboring CRISISS upon activation, 2×10^5 cells of cell lines arC9 Alu HeLa clone #1, arC9 scaff HeLa clone #1 and a polyclonal HeLa control cell line were seeded on wells of a 6-well plates and immediately induced by adding 4-HT and DOX at concentrations of 200 nM and 1 µg/ml respectively to the cell culture medium. For a period of 5 dpi, for each cell line, the total amount of living cells per well was determined after washing three wells per cell line with PBS to remove dead cells and cell debris, harvesting them via trypsinization, resuspending with cell culture medium to generate single cell suspensions and counting with an automated cell counter (Tab. 10). For each cell line triplicates were run in parallel.

7.3. Results chapter II

7.3.1. Generation of monoclonal HeLa cell lines harboring a post-translationally inducible CRISPR/Cas9 based suicide-switch

The first attempt of generating a CRISPR/Cas9-based suicide-switch that kills its host cells by fragmenting their genome through *Alu* element targeting was conducted by transfecting the transposon plasmid pT2_arC9_COMBI_Alu into WT HeLa cells. After transfection, cells were kept in culture for an additional 14 days in order to dilute non-integrated episomal DNA leading to transient expression. After 14 days colonies still exhibiting mCherry expression, the fluorescent marker fused to arC9 via a T2A sequence, were identified via fluorescent microscopy and monitored daily for growth.
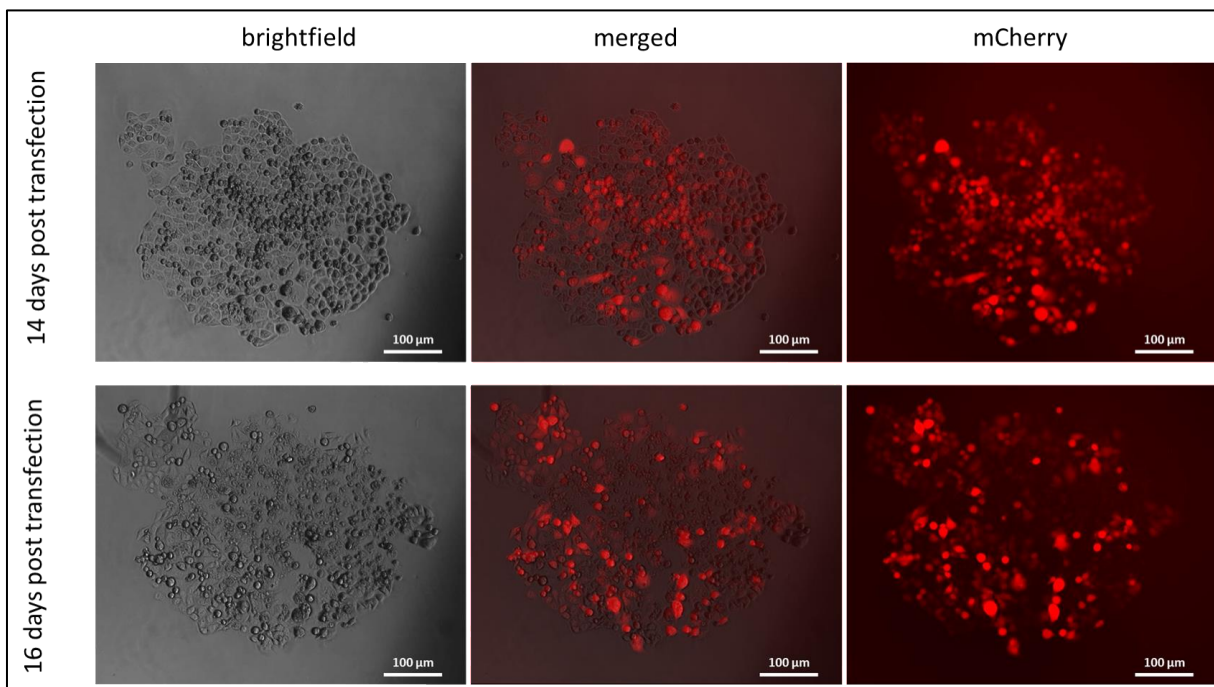
Figure 17: *Exemplary microscopic following of a HeLa cell colony exhibiting mCherry expression at 14 days and 16 days after co-transfection with the transposon plasmid pT2_arC9_COMBI_Alu and the transposase expression plasmid pCMV(CAT)T7-SB100X.*

However, as shown as an example depicted in Fig. 17, fluorescent cell colonies appeared to lose fitness and eventually die. This behavior is best seen in the middle section of the colony in Fig. 17 in the brightfield image and in the disappearance of red fluorescent in that section in the merged/mCherry channels. This behavior made clonal expansion impossible yet provided early indications, that presumable leakiness of the system was sufficient for cell killing. To circumvent this issue, an additional regulatory switch in form of a transcriptional control was added to the system.

7.3.2. Generation of CRISISS, a translational and post-transcriptionally inducible CRISPR/Cas9-based suicide-switch

Since generating a cell line harboring a post-translationally inducible suicide-switch proved unsuccessful presumably due to leakiness of the system, an additional control mechanism was added in form of a tet-on system driving arC9 expression. Additionally, the mCherry selection marker was discarded and exchanged for a neomycin resistance gene, facilitating elimination of cells lacking the desired transgene cassette (see 7.2.3.). In parallel, the process was repeated with an identical transposon except for the lack of sgRNAs to be used as a negative control with respect to the role of the *Alu*-targeting in the suicide-mechanism.
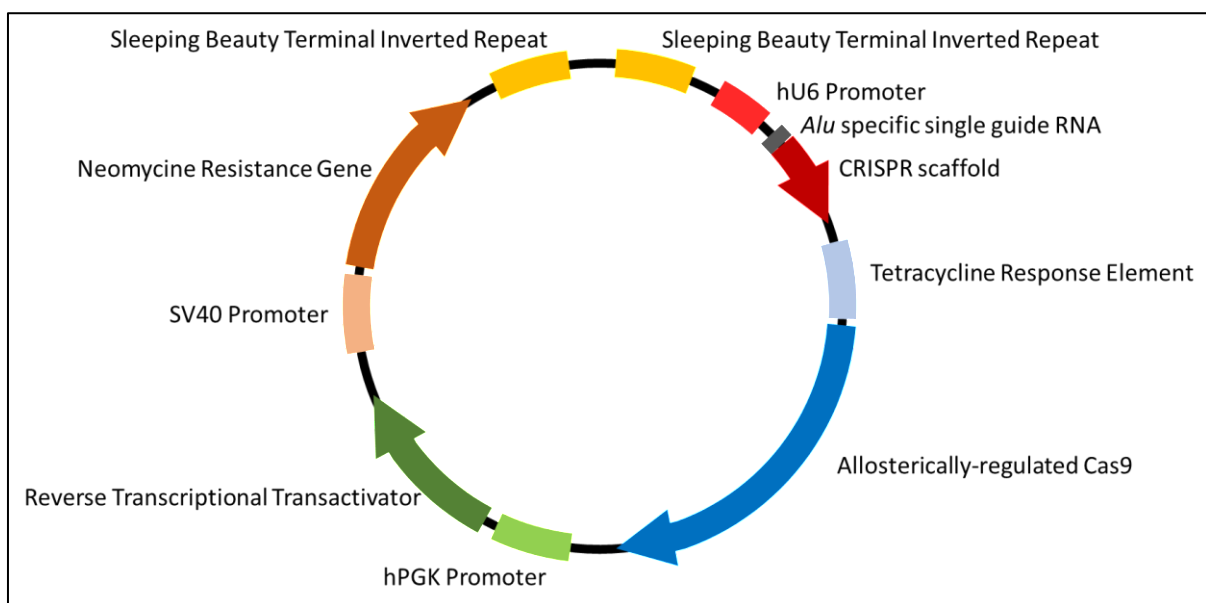
Figure 18: *Schematic representation of pTOV_T11_Neo_arC9_Alu plasmid, depicting the components of the transcriptionally and post-translationally inducible suicide-switch CRISISS. Components include the human U6 promoter (light red) driving Alu-specific sgRNA (dark grey) and CRISPR scaffold (dark red) transcription, a tetracycline response element (light blue) driving transcription of arC9 (blue), a human PGK promoter (light green) driving transcription of the reverse transcriptional transactivator (green) and an SV40 promoter (beige) driving the neomycin resistance gene (orange) transcription. These components are flanked by SB-TIR's (yellow).*

To test the functionality of the constructed plasmids, they were transfected into HeLa cells and induction was initiated by adding 1 µg/ml of DOX. 48h post-transfection, protein extracts were subjected to WB Analysis.
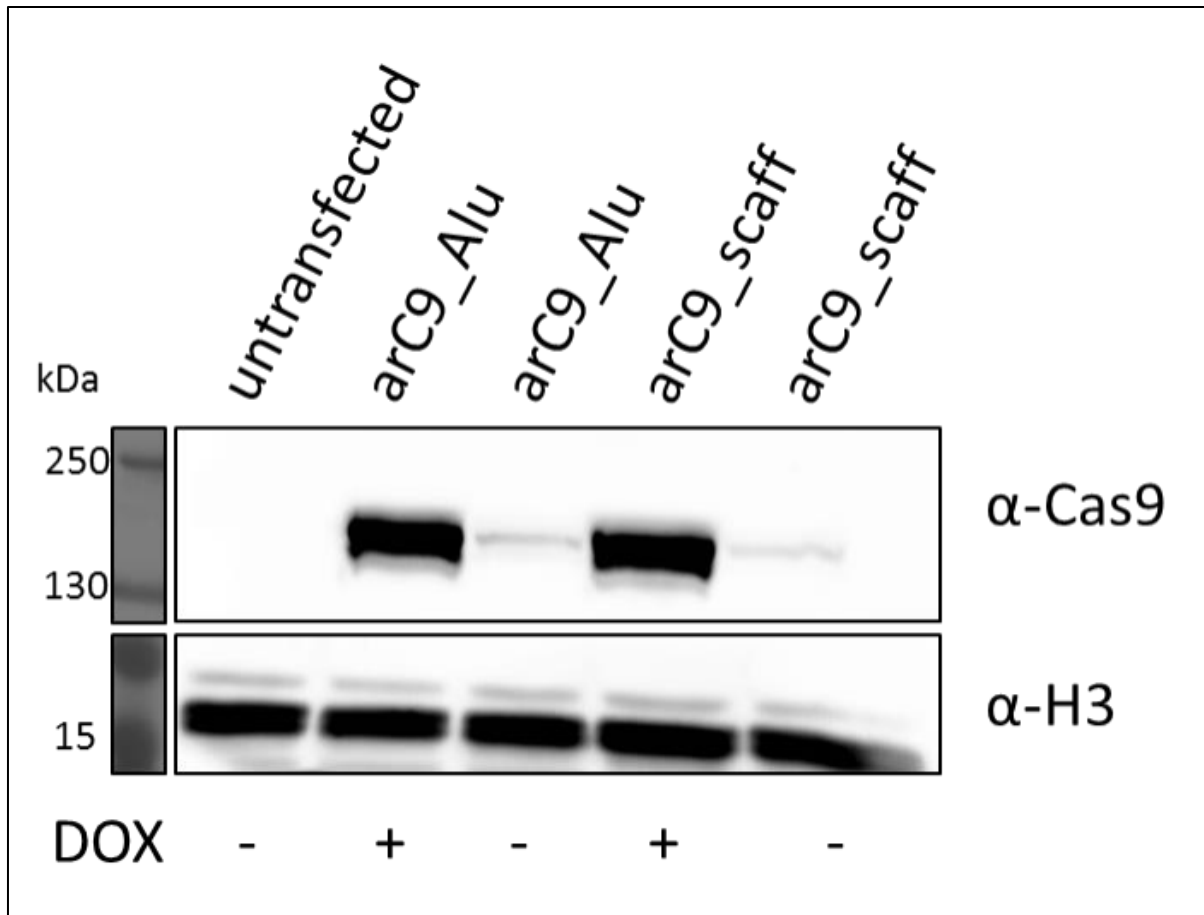
Figure 19*: WB analysis of protein extracts stemming from HeLa cells 48 h post transfection with the transcriptionally and post-translationally inducible suicide-switch with (arC9_Alu) and excluding (arC9_scaff) the Alu-specific sgRNAs to confirm DOX-dependent arC9 expression. Protein extract stemming from untransfected HeLa cells was used as a negative control. As a loading control, H3 was detected. Both proteins were detected on the same membrane with an exposure time of 3 min.*

The WB Analysis (Fig. 19) shows a strong signal between 130 kDa and 250 kDa, indicating a DOX-dependent expression of arC9 in the corresponding samples. Protein extracts of uninduced samples show a faint signal at a similar size, suggesting the presence of arC9 and thus minimal DOX-independent arC9 expression leakage.

7.3.3. Polyclonal HeLa cell cultures harboring CRISISS exhibit solid, DOX-dependent arC9 expression and unmeasurable leaky expression

As a follow up, transposons harboring the components of CRISISS (Fig. 18) that had been confirmed to enable DOX-dependent arC9 expression (Fig. 19) were stably introduced into the genome of HeLa cells via SB-mediated transposition (see 7.2.4. for methods). After selection with G418 for a total period of 14 days two stable G418-resistant polyclonal cell cultures were established, in which the functional integration of the transgene cassettes was determined by inducing arC9 transcription via addition of 1 µg/ml of DOX to the cell growth medium for 48h followed by WB Analysis to detect the presence of arC9.
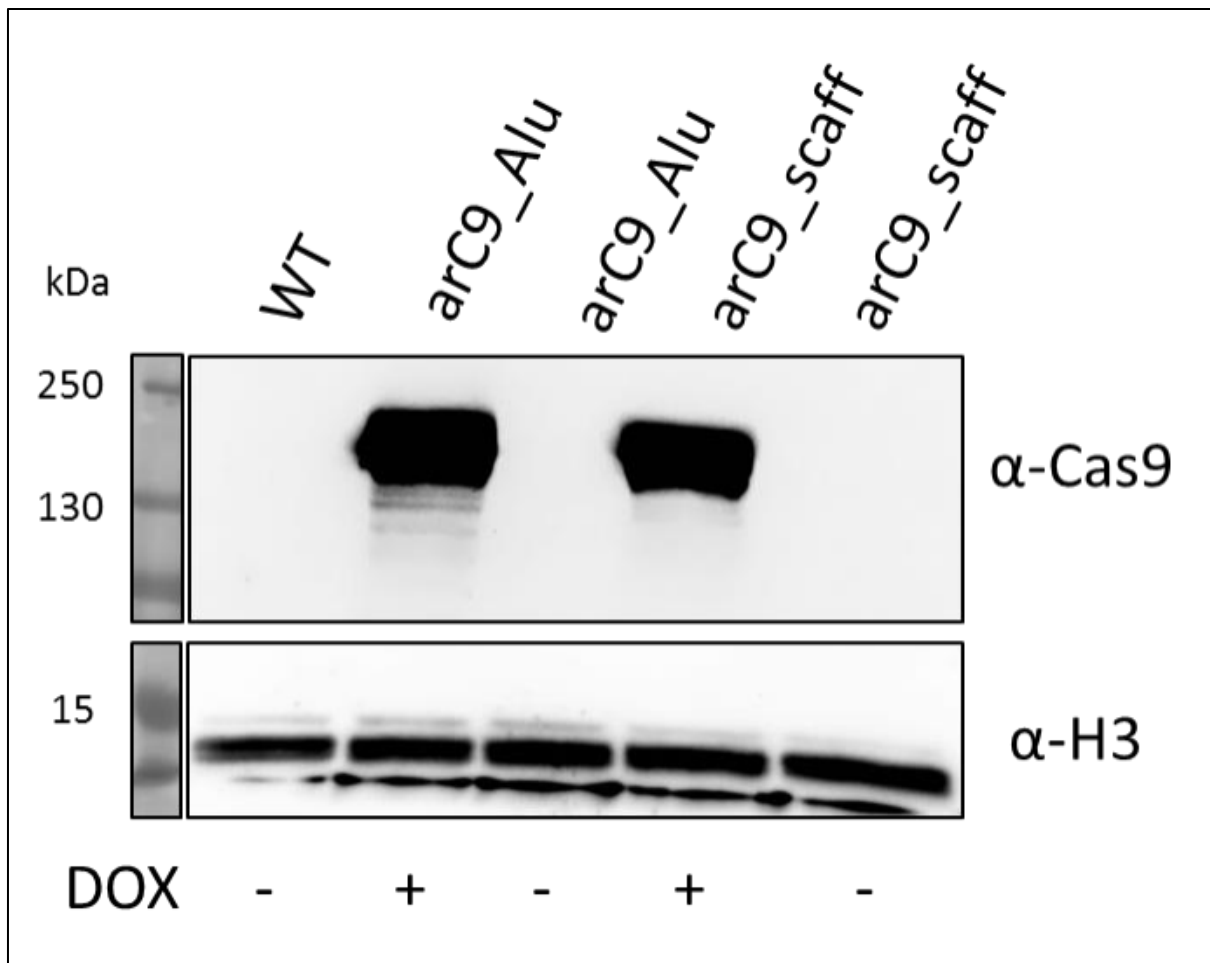
Figure 20: *WB analysis of protein extracts stemming from two polyclonal HeLa cell lines harboring the transcriptionally and post-translationally inducible suicide-switch with (arC9_Alu) and excluding (arC9_scaff) the Alu-specific sgRNAs to confirm DOX-dependent arC9 expression. A protein extract stemming from wildtype HeLa cells was used as a negative control. As a loading control, H3 was detected. For arC9 detection, the membrane was exposed for a total of 5 min., for loading control detection for 9 s.*

The WB Analysis (Fig. 20) shows DOX-dependent expression of arC9, as seen with the presence of two protein bands in the corresponding lanes between 130 and 250 kDa. Additionally, and in contrast to the previously conducted DOX induction after transient introduction of both transposons into HeLa cells (Fig. 19), no detectable levels of leaky arC9 expression were observed in the absence of DOX as seen with the lack of a specific protein band in the corresponding sample lanes, similar to the negative WT control.

7.3.4. CRISISS induction and colony-forming kill-assay show CRISISS-dependent cell killing but generates a CRISISS resistant cell population

The previously described polyclonal HeLa cell populations expressing arC9 in a DOX-dependent manner were used to assess the feasibility of the suicide mechanism and to generate preliminary data to gain an early insight on the effect of targeting *Alu* elements with an inducible designer nuclease. For this purpose, 5×10^5 cells of each polyclonal cell population were

81

induced with 200 nM 4-HT and DOX at a concentration of 1 µg/ml and grown on a 10 cm cell culture plate in triplicates.
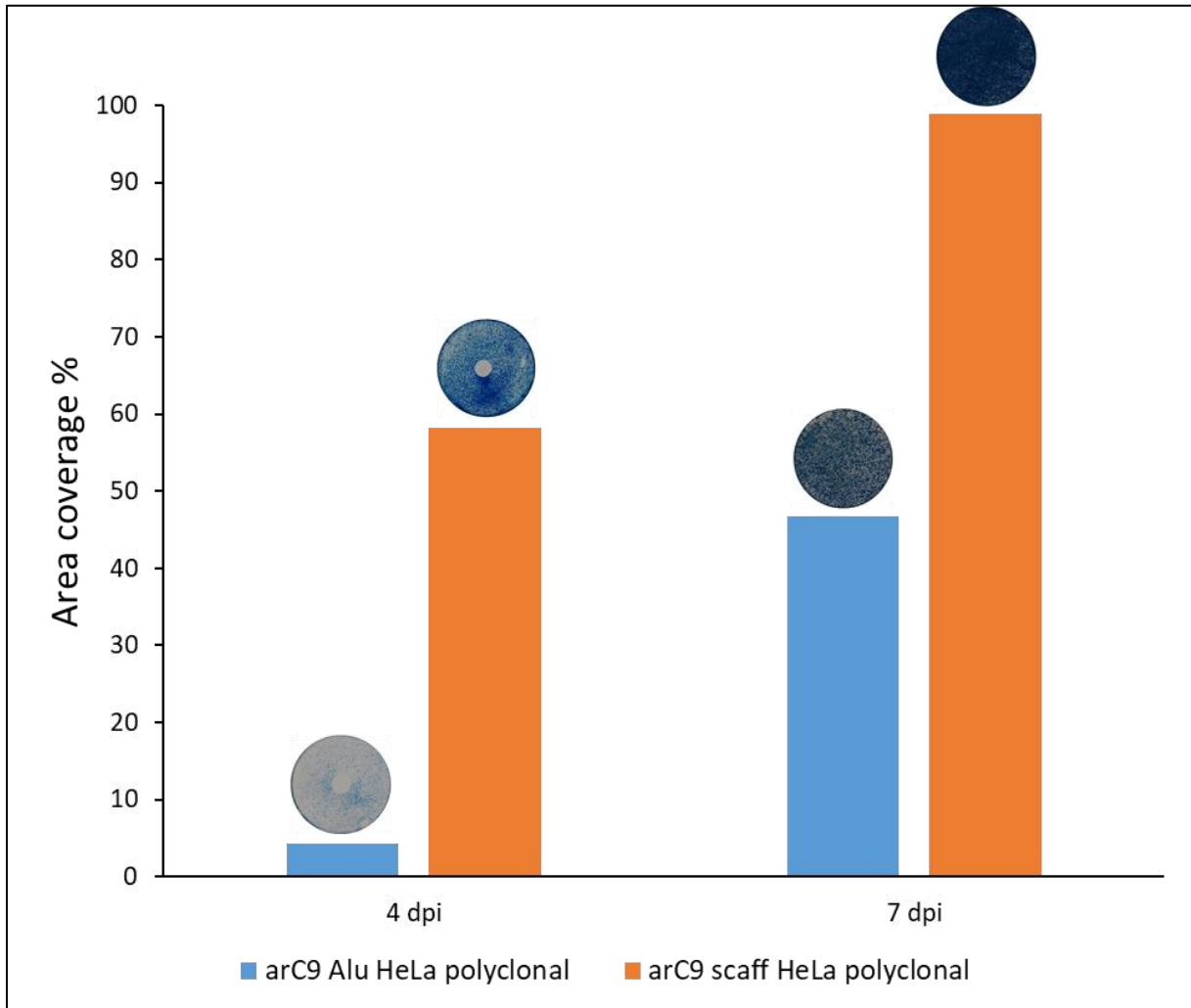
Figure 21*: Area coverage on a 10 cm cell culture plate of polyclonal HeLa cell lines harboring CRISiSS with (blue) and without (orange) sgRNAs, 4 and 7 dpi with 200 nM 4-HT and 1 µg/ml DOX. Corresponding stained plates are depicted above each column. Cell-less circles in both 4 dpi plates were caused due to the direct addition of 4-HT diluted in EtOH directly to the growth medium, breaking surface tension and killing cells in that area. This problem was avoided at 7 dpi by tilting the plate while adding the 4-HT dilution.*

As depicted (Fig. 21), plates corresponding to the first replicate were stained and analyzed 4 days post-induction. While the cells lacking the *Alu*-specific sgRNAs repopulated the plates' surface to a total coverage of 58%, only 4% of coverage was measured on the plate containing the cells that harbored *Alu*-specific sgRNAs. However, at 7 days post-induction, the same cells repopulated the plate to a total surface coverage of 47% under continuing induction with 4-HT and DOX. The cells lacking the *Alu*-specific sgRNAs repopulated the plate to reach 99% confluency.

Hypothesizing that the cell population presumed to harbor a CRISISS cassette including *Alu*-specific sgRNAs would nonetheless have a fraction that did not exhibit DOX-dependent arC9

expression and as a result appeared to regain fitness and repopulate the cell culture plate, the third replicate was kept in culture under induction conditions with regular medium changes adding fresh 4-HT and DOX every 48h for a period of two weeks. This established a third cell population apparently resistant to CRISISS induction.

To confirm the previously described hypothesis, protein extracts of this third cell population were generated and analyzed via WB, probing for arC9 expression. As controls, protein extracts of two polyclonal cell populations (described in 7.3.3.) were included in the analysis after 48h induction with 1 µg/ml DOX, uninduced, and for the cells lacking *Alu*-specific sgRNAs with 1 µg/ml DOX and 200 nM 4-HT.
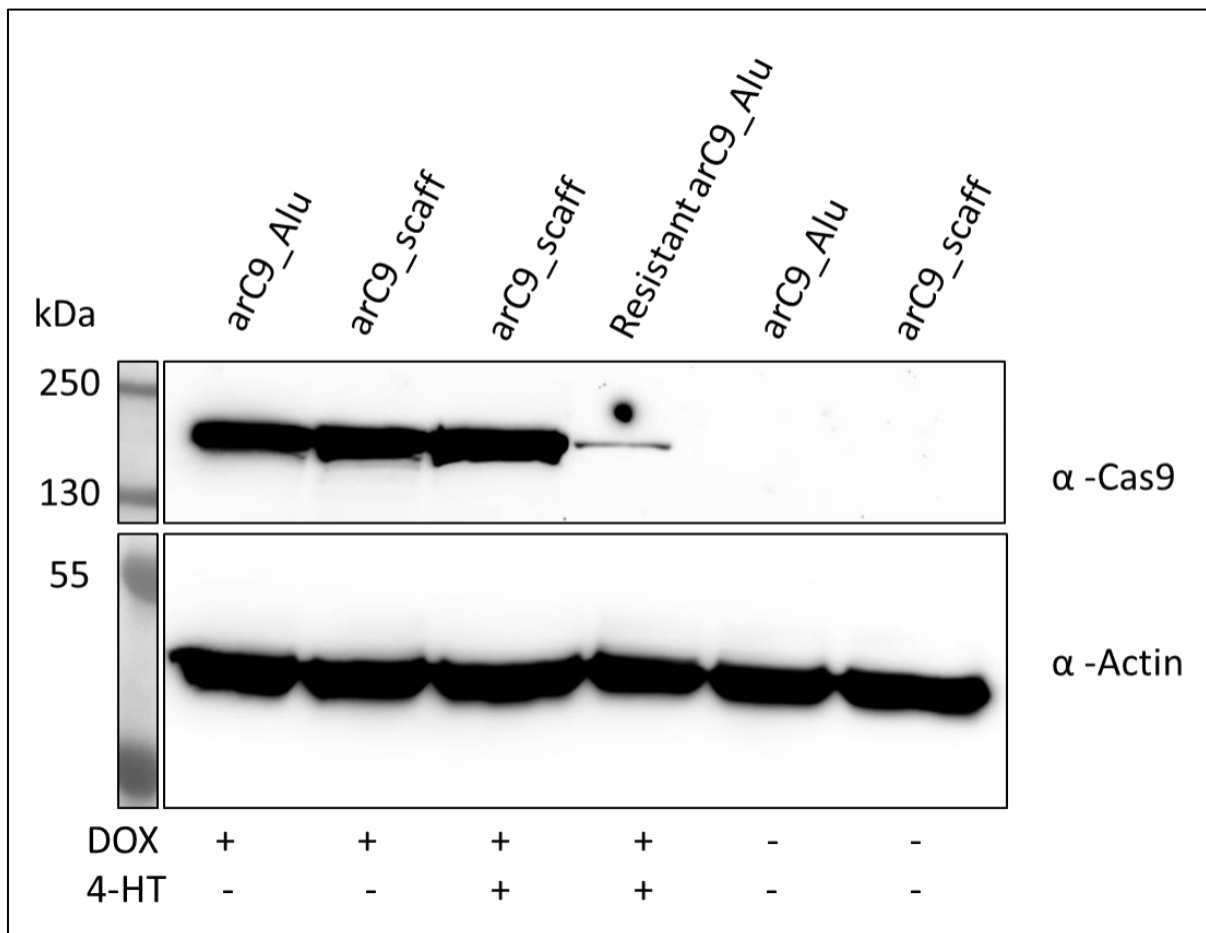
Figure 22*: WB analysis of protein extracts stemming from three HeLa cell populations harboring the transcriptionally and post-translationally inducible suicide-switch with (arC9_Alu), excluding (arC9_scaff) the Alu-specific sgRNAs and generated by prolonged selection of arC9_Alu with 4-HT and DOX to confirm DOX-dependent arC9 expression. As a loading control, actin was detected. For arC9 detection, the membrane was exposed for a total of 10 min., for loading control detection for 1 min. Samples were loaded on an 8% polyacrylamide gel.*

The analysis presented in Fig. 22 generated visible protein band signals at the expected size of arC9 for samples arC9_Alu and arC9_scaff when induced with DOX for 48h and a similar protein band signal for arC9_scaff when induced for 48h with DOX and 4-HT. The cells

generated by selecting the polyclonal arC9_Alu with DOX and 4-HT produced a protein band signal at the expected size only visible when exposing the membrane for a minimum of 10 min and noticeably fainter than the first three samples.

7.3.5. Time-limited flow cytometry-based viability assay shows *Alu*-specific, sgRNA-dependent induction of apoptosis and cell killing

Taking into consideration the observation, that although a fraction of the polyclonal arC9_Alu cells express arC9 (though at notably lower levels), but nonetheless survives double induction with 4-HT and DOX and repopulates the cell culture, a time-limited flow cytometry-based viability assay was conducted to observe the apoptosis induction and cell killing caused by CRISISS. Since the previous colony-forming kill-assay indicated cell killing until day 4 post induction followed by repopulation by the remaining resistant cells, the 4-day timeframe was chosen as the observation period.
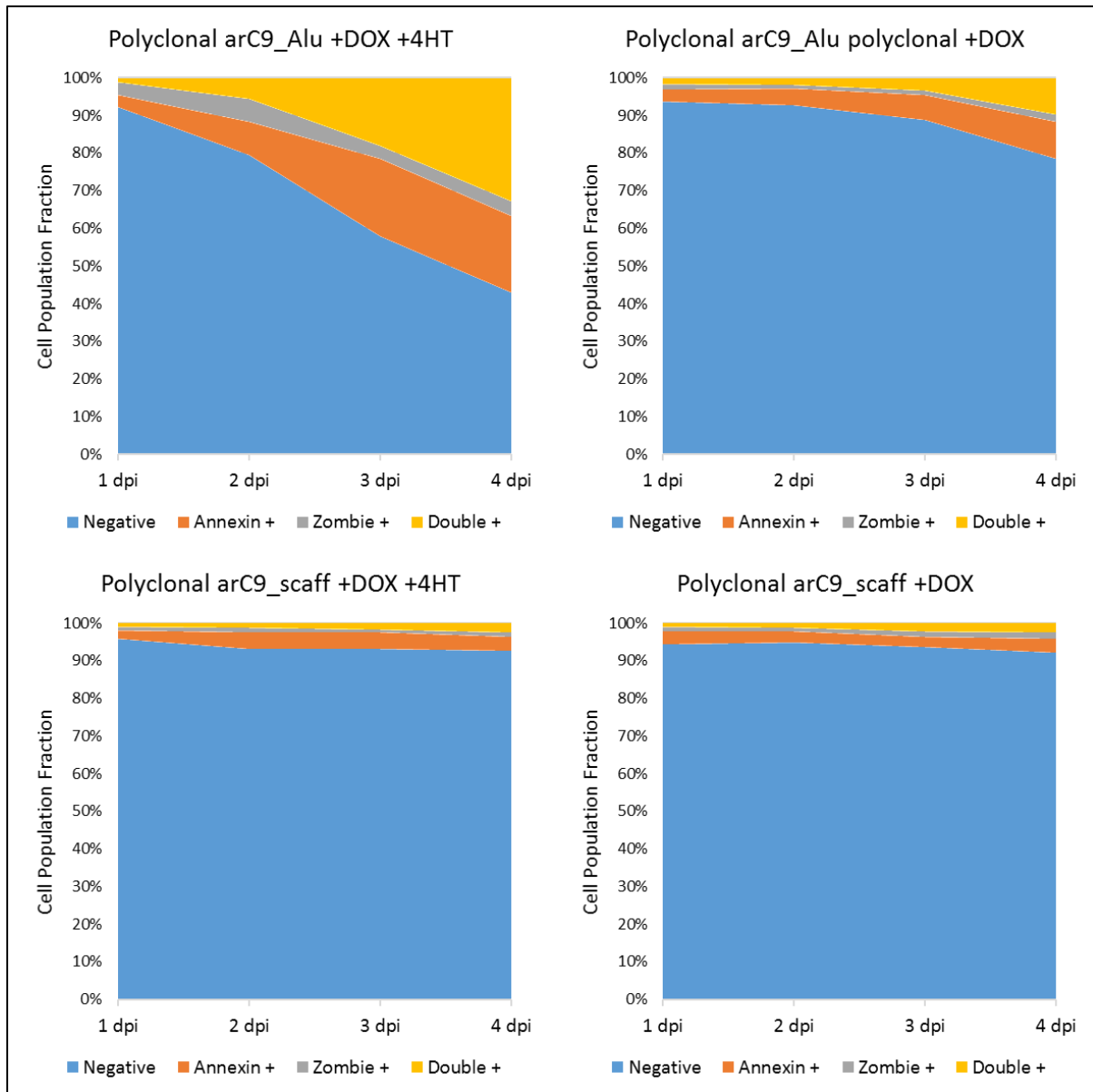
Figure 23*: Time-limited flow-cytometric assay to determine the apoptosis induction and cell killing through CRISISS. Shown are the two polyclonal arC9_Alu and arC9_scaff cell populations after single (+DOX) and double (+DOX +4-HT) and the respective cell fractions, namely unstained (blue), Annexin V positive (orange), Zombie positive (grey) and Annexin V and Zombie positive (yellow) out of the total cell population over a course of 4 dpi. The Annexin V positive fraction represents apoptotic cells, double positives for Annexin V and Zombie represent dead cells.*

When *Alu*-specific sgRNAs were absent (polyclonal arC9_scaff), no differences were discernible between the cell populations composition within samples induced with DOX and samples induced with both DOX and 4-HT across the 4-dpi timeframe. Additionally, cell population composition did not seem to change over time. However, when *Alu*-specific sgRNAs were present (polyclonal arC9_Alu), a growing fraction of apoptotic and dead cells was detected as time progressed. Differences were observed dependent on the induction: single induction with DOX led to an overall smaller and slower increasing fraction of apoptotic and dead cells while double induction with DOX and 4-HT led to a faster increment of the apoptotic

and dead cell population fractions making up 57% of the total cell population fraction at 4 dpi in contrast to 21,4% for the sample induced with DOX only.

7.3.6. Generation of a monoclonal HeLa cell line harboring CRISISS exhibits solid DOX-dependent arC9 expression

Since the established polyclonal cell populations harboring the transcriptionally and post-translationally inducible suicide-switch provided data indicating that DOX- and 4-HT-dependent cell killing was taking place but surviving cells showed a substantial loss of the capacity for DOX-dependent arC9 expression and were able to repopulate and grow under double induction growth conditions, clones were picked out of the two polyclonal cultures, generating monoclonal cell lines (see 7.2.8. for methods).

To confirm the functional integration of the SB-transposon harboring CRISISS, protein extracts of the newly generated monoclonal cell lines were extracted and analyzed via WB, probing for arC9 expression after 48h induction with 1 µg/ml DOX.
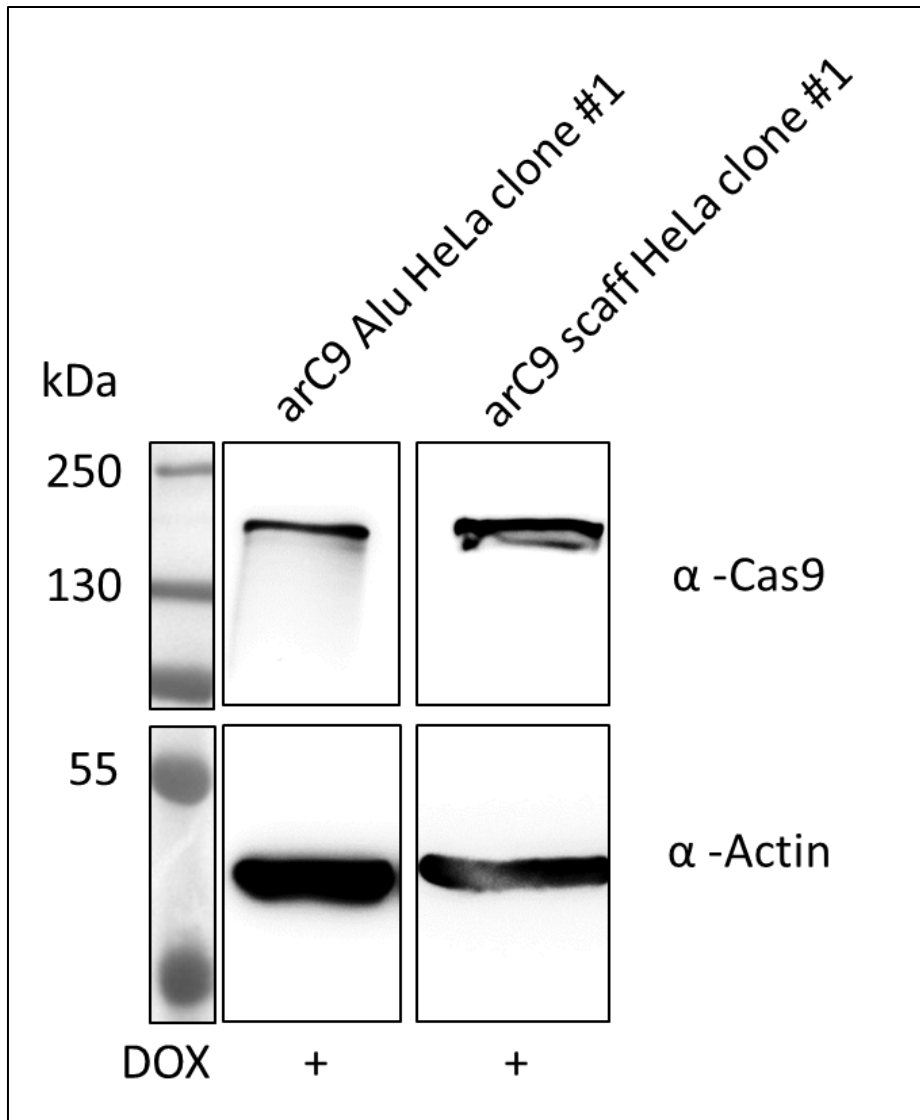
***Figure 24**: WB analysis of protein extracts stemming from two monoclonal HeLa cell lines harboring the transcriptionally and post-translationally inducible suicide-switch including (arC9_Alu HeLa clone #1) and excluding (arC9_scaff HeLa clone #1) the Alu-specific sgRNAs. As a loading control, actin was detected. For arC9 detection, the membrane was exposed for a total of 2 to 6 min., for loading control detection for 30 s to 6 min. Samples were loaded on an 8% polyacrylamide gel.*

As shown in Fig. 24, monoclonal cell lines maintained DOX-dependent arC9 expression, represented by protein band signals in lane 1 and 2, confirming results observed in the origin polyclonal cell populations (Fig. 20).

7.3.7. CRISISS induction produces strong KAP1 phosphorylation while no leakage is observed when uninduced

Based on the hypothesis that CRISISS induces an irreparable amount of DSBs in the host cell when induced causing cell death, WB based detection of pKAP1 probing for DNA damage response was conducted on the two monoclonal cell lines arC9 Alu HeLa clone #1 and arC9 scaff HeLa clone #1 in order to: i) confirm CRISISS-induced DNA damage when induced, ii) measure potential leakage leading to DNA damage when uninduced, iii) gain insight about the

timeframe in which DNA damage occurs and iv) link DNA damage directly to the presence of a sgRNA.
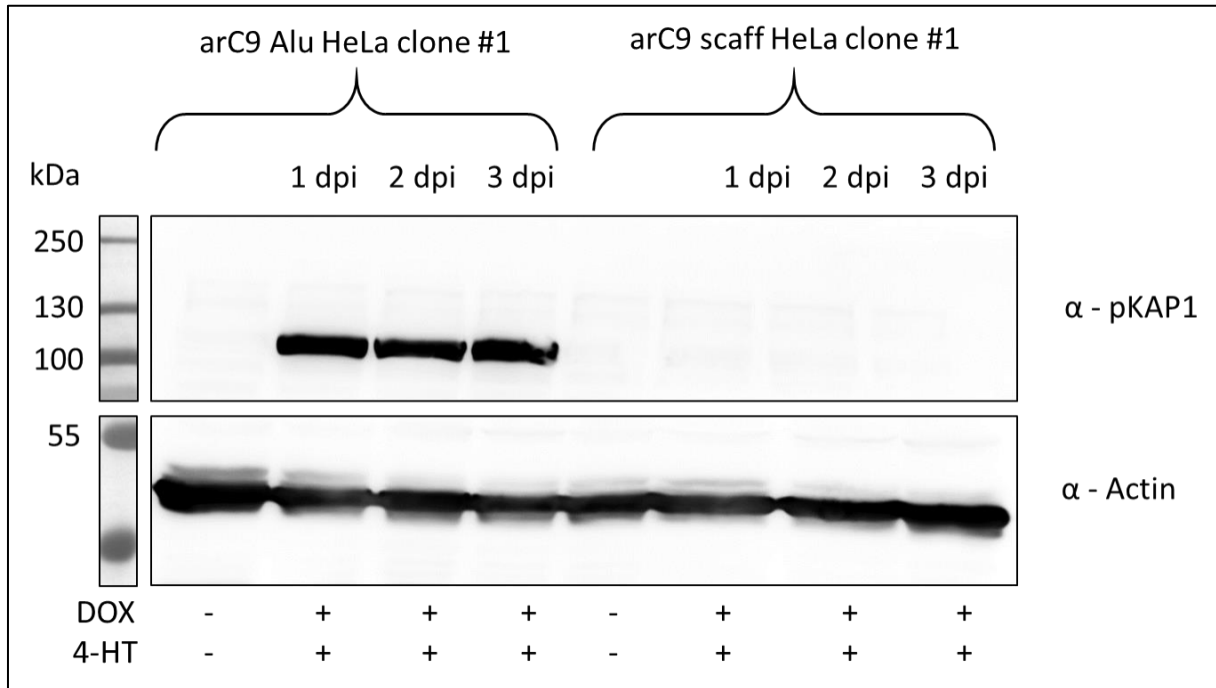
Figure 25: *WB based assay to detect the presence of pKAP1 in cell lines arC9 Alu HeLa clone #1 and arC9 scaff HeLa clone #1 after CRISISS induction over a time period of 3 days. A total of 5 µg of total protein was loaded per sample on an 8% polyacrylamide gel. To detect the presence of pKAP1 the membrane was exposed for 1 min and 26 s. For the loading control (actin) detection, the membrane was exposed for 1 min.*

When uninduced, the cell extracts of the cell line harboring CRISISS (arC9 Alu HeLa clone #1) did not contain any detectable levels of pKAP1. Upon induction however, pKAP1 was detected at all timepoints at equal levels. Cell extracts of the cell line lacking *Alu*-specific sgRNAs (arC9 scaff HeLa clone #1) did not display any pKAP1-specific signals, regardless of induction and timepoint of extraction from day 1 to 3 post-induction.

7.3.8. Colony forming and overall survival assays after CRISISS induction

Using the established monoclonal cell lines that displayed induction and *Alu*-sgRNA dependent DNA damage responses (Fig. 25), overall survival and colony forming assays were performed 5 dpi to assess the efficiency of CRISISS (with and without *Alu*-specific sgRNAs) at eliminating cells and to simultaneously exclude a negative impact on cellular growth and fitness when uninduced.
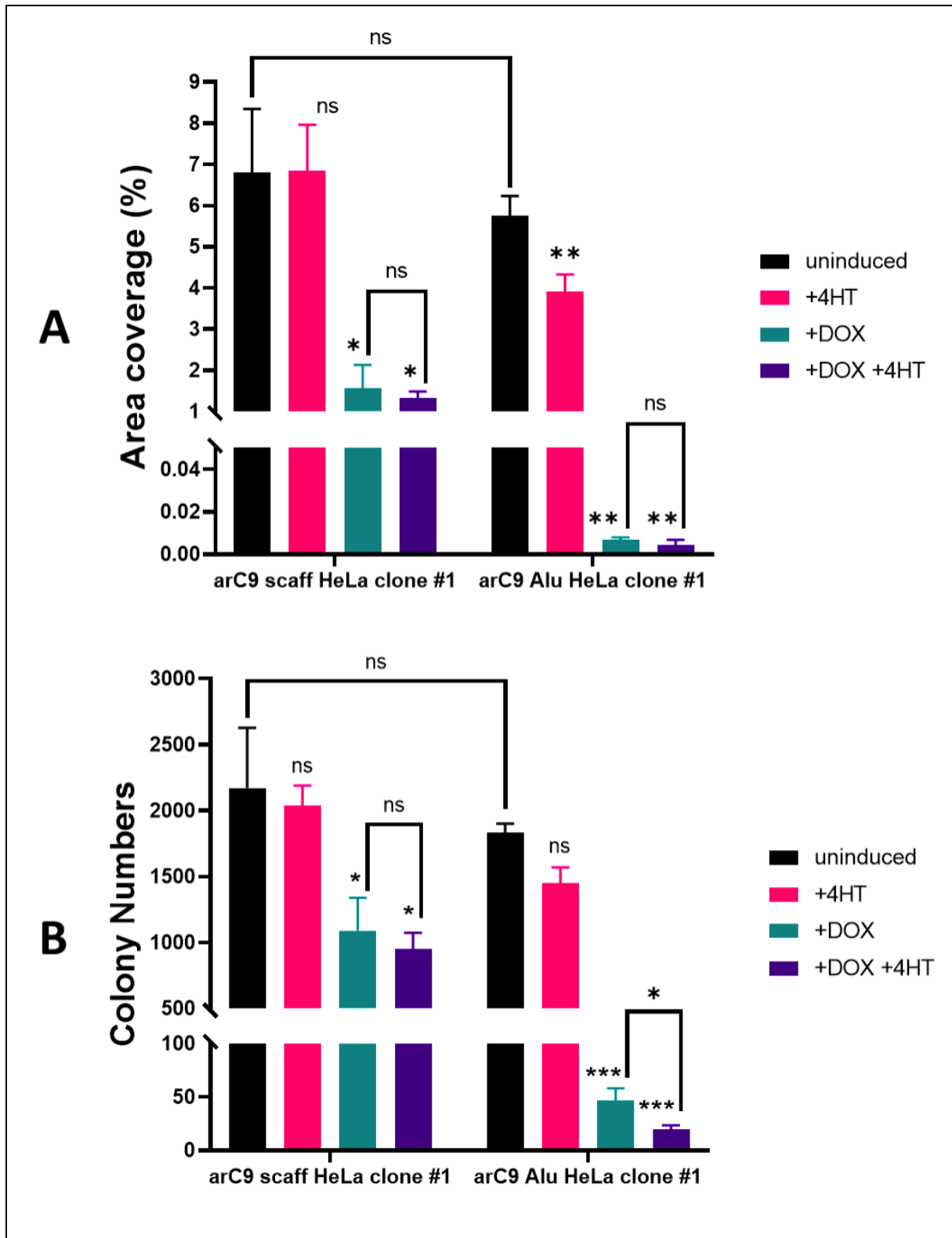
Figure 26: *Assay to determine CRISISS effectivity. **A**: The percentage of area covered by cells corresponding to the monoclonal cell lines harboring CRISISS with (arC9 Alu HeLa clone #1) and without (arC9 scaff HeLa clone #1) Alu-specific sgRNAs on a 10 cm cell culture dish after single induction with 4-HT (pink), DOX (petrol), double induction with 4-HT and DOX (violet) and uninduced (black) are shown. **B**: The total number of visible colonies formed by cells corresponding to the monoclonal cell lines harboring CRISISS with and without Alu-specific sgRNAs on a 10 cm cell culture dish after single induction with 4-HT (pink), DOX (petrol), double induction with 4-HT and DOX (violet) and uninduced (black) are shown. Significances were calculated with respect to the corresponding negative control (uninduced sample, black). Error bars: SD, ns: p>0,05, * = p<0,05, ** =p<0,01, *** = p<0,001 (Welch's t-test).*

There was no significant difference measured between the overall survival, measured by relative area covered by cells (Fig. 26, A), and colony numbers (Fig. 26, B) of cell lines arC9 scaff HeLa clone #1 and arC9 Alu HeLa clone #1 in the absence of induction (Fig. 26, black bars). In addition, arC9 scaff HeLa clone #1 remained unaffected concerning overall survival rate and colony numbers, when exposed to single induction with 4-HT (Fig 26, pink bars). Significant reduction of the numbers of surviving cells and colonies was observed however when DOX was added. This was observed regardless of 4-HT presence, as the DOX induced samples exhibit no significant difference to the DOX + 4-HT induced sample. When *Alu*-specific sgRNAs are present, survival rates were significantly affected in all induction conditions. The weakest effect was observed when cells were induced with 4-HT only, where area covered by cells and colony numbers dropped by an average of 32% and 21% respectively. Inducing with DOX and a full induction with DOX and 4-HT however, lead to an almost complete cell killing. Cells induced with DOX exhibited an average total reduction in area coverage and colony numbers of 99,9% and 97% respectively. When full induction with DOX and 4-HT took place, this effect reached an average maximum reduction of 99,9% and 98,9% with respect to area coverage and colony numbers. Interestingly, statistically significant differences between arC9 Alu HeLa clone #1 samples induced with DOX and samples induced with both DOX + 4-HT could only be observed when comparing the total colony numbers and not the total area coverage.

Hypothesizing that the reduction in cell viability of cell line arC9 scaff HeLa clone #1 when in presence of DOX despite not containing sgRNA sequences and having ruled out DNA damage (Fig. 25) might be caused by the overexpression of a foreign protein, a polyclonal WT HeLa cell line was used in an identical overall survival/colony forming assay.
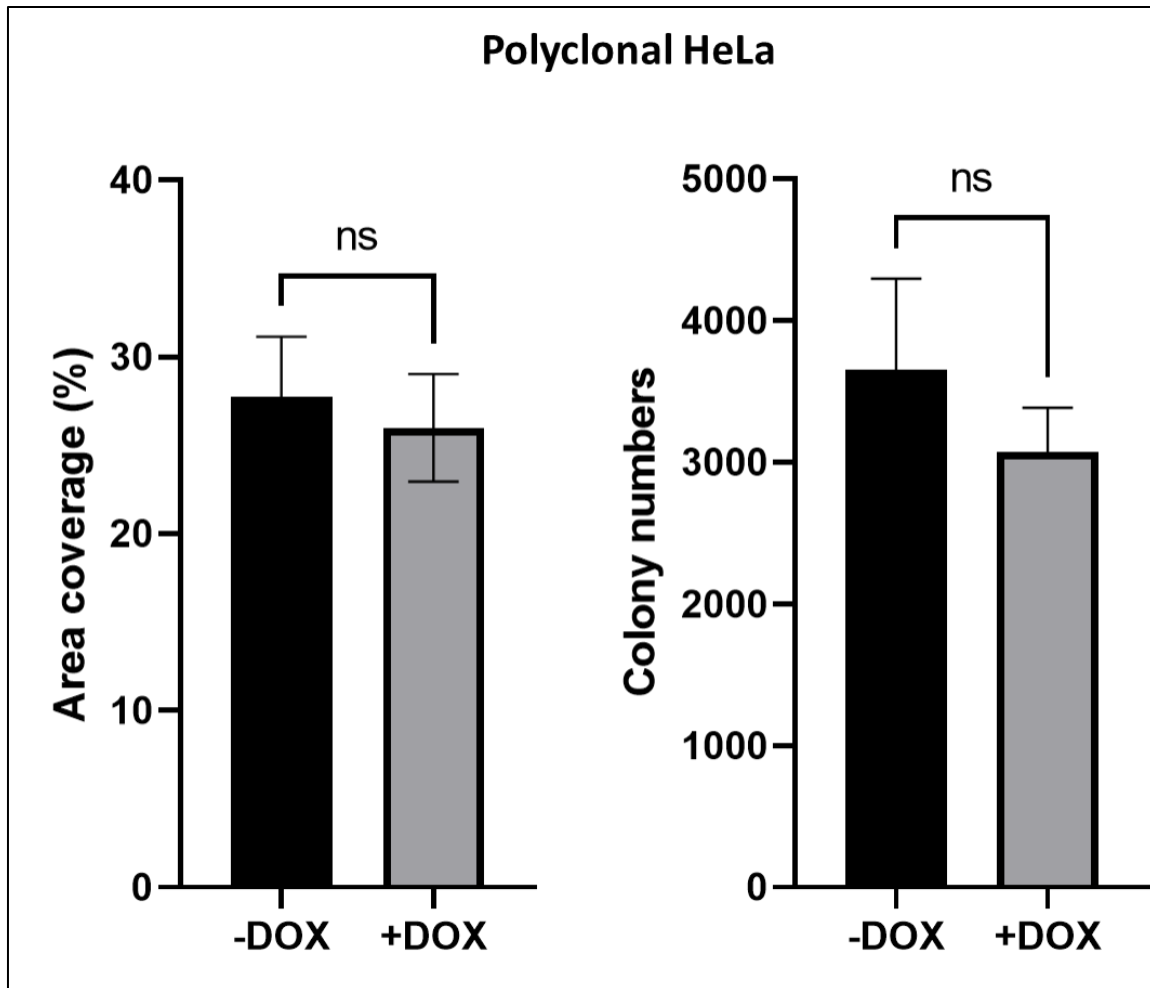
Figure 27: *Assay to determine the impact of DOX on a polyclonal WT HeLa cell.* **Left***: The percentage of area covered by cells on a 10 cm cell culture dish uninduced and after single induction with and without DOX are shown.* **Right***: The total number of visible colonies formed by cells on a 10 cm cell culture dish, uninduced and after single induction with DOX are shown. Significances were calculated with respect to the corresponding negative control (uninduced sample, black). Error bars: SD, ns: p>0,05 (Welch's t-test).*

There were no significant differences observed between uninduced samples and samples induced with DOX with regards to area coverage and colony formation (Fig. 27) and thus no significant impact of DOX on cell viability.

Lastly, since CRISISS efficiency seemed to near 100% but trace amounts of colony numbers and area coverage could still be measured after the induction period, the remaining cells were inspected microscopically after a slightly longer 7-day induction period with DOX and 4-HT, in order to rule out signs of resistance and recovery.
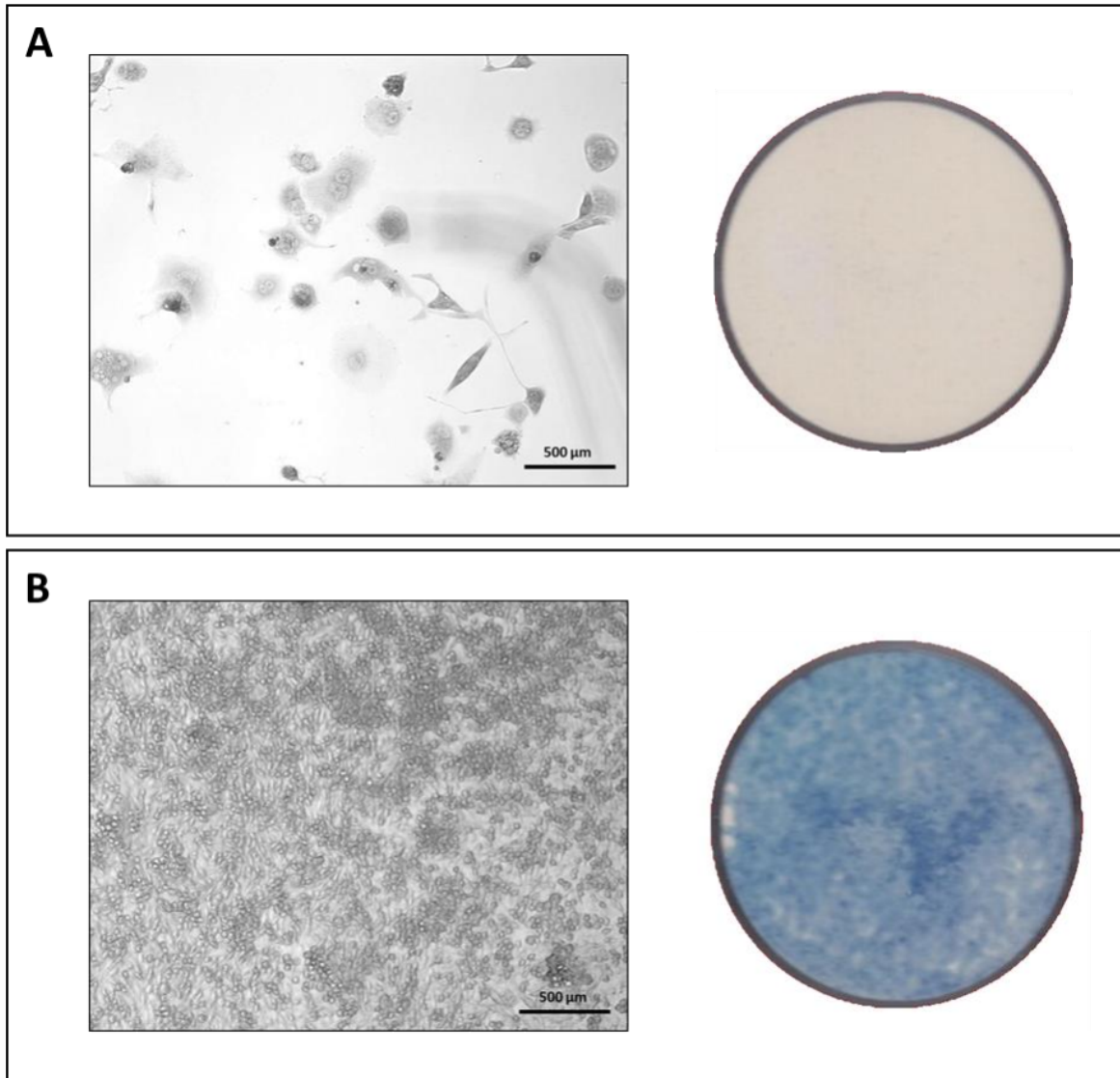
Figure 28*: Exemplary microscopy (left) and methylene blue stained plate scan (right) images showing the resulting cellular morphology (left) and cell growth (right) after 7 days incubation in a 6-well plate well of arC9 Alu HeLa clone #1 with double induction (DOX + 4-HT) (**A**) and uninduced (**B**). 5×10^5 cells were seeded on day 0.*

While the uninduced sample reached complete confluency on the well of a 6-well plate within 7 days (Fig. 28 B, right), the induced sample exhibited no signs of recovery (Fig. 28 A, right) in comparison to previous observations after a 5-day induction time (Fig. 27). Interestingly however, when observing the remaining trace amounts of cells after double induction with DOX and 4-HT under a light microscope, noticeable morphological differences were seen. Instead of colonies and counted as such previously after staining (Fig. 27), the observed remaining cells on these plates were individual senescent cells, multiple times larger in size (Fig. 28 A, left). No other phenotype was observed.

7.3.9. CRISISS induction and kill kinetics

After confirming that CRISISS is capable of killing nearly all cells within a cell population carrying it and seemingly completely stopping cell division in the surviving cells, an assay to determine the kinetics of the cell killing mechanism was conducted. For this purpose, 2×10^5 cells of the monoclonal arC9 Alu HeLa clone#1 were seeded on a 6-well plate and induced with DOX and 4-HT to start CRISISS. Every 24h for a total of 7 days, the number of living cells remaining on the dish was assessed and compared with the results obtained from an identical setup with the control cells harboring CRISISS but excluding the *Alu*-specific sgRNAs, and a third control consisting of a wildtype HeLa cell line. Besides comparing the total number of cells, the dynamics of the fraction each cell line occupied during the observation period out of the total number of surviving cells, considering that they were equally distributed in the beginning, was assessed.
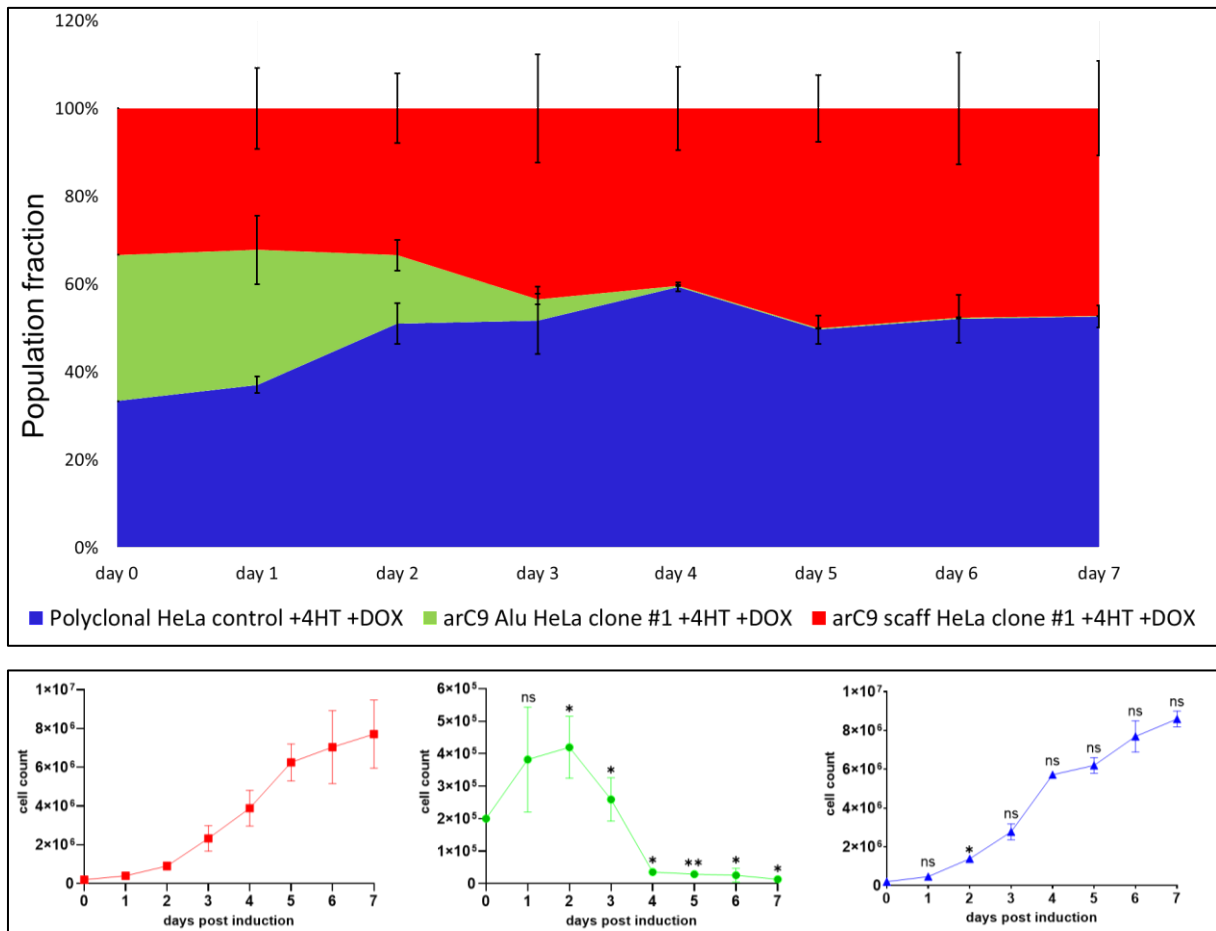
Figure 29: *Determination of population dynamics and cell proliferation of monoclonal HeLa cells after suicide-switch induction over a period of 7 dpi. Top half: Representation of the average fractions of the total population each cell line occupies, after an initial equal distribution corresponding to 1/3 of the total, each. Bottom half: Number of the average surviving cells counted every 24h after induction over a period of 7 dpi. Significances were calculated with respect to the data of the arC9 scaff HeLa clone #1 cell line (red). Error bars: SD, ns: p>0,05, * = p<0,05, ** = p<0,01 (Welch's t-test).*

93

As seen in Fig. 29, control cell lines arC9 scaff HeLa clone #1 and polyclonal HeLa both exhibited exponential growth until plateauing due to the limits imposed by the physical size of the 6-well plate. Overall, except for day 2 post-induction, no significant differences were measured regarding growth rate. At 7 dpi, both control cell lines made up an average of almost precisely 50% of the total remaining cell numbers, each. However, the remaining cell line arC9 Alu HeLa clone #1, while exhibiting significantly similar cell numbers compared to the control cell lines on day 1 post induction yet stark variances between replicates, halted cell division on day 2 post induction and started dying on day 3 post-induction. An average reduction of 82,2% in cell numbers compared to the originally seeded number was reached on day 4 post induction, where this decline is further enhanced to an average of 91,5% when compared to the average maximum number of living cells present at day 3 post induction. Cell death continued until the end of the experiment at day 7 post-induction, where an average death of 96,8% of the average maximum number of living cells measured on day 3 post-induction was observed. Measured by the fraction that each cell line represents, beginning with day 4 post-induction, the control cell lines made up an average of 99,6% of the total living cell population. At day 7 and at the end of the experiment, this fraction grew to an average of 99,9%.

7.4. Discussion chapter II

7.4.1. Post-translationally controlled suicide-switch

The initial results, basing the inducibility concept of a suicide-switch on post-translational induction through the use of the allosteric Cas9 switch arC9 [209] provided critical and valuable information for the project's continuation and final success. While unexpected, since Oakes et al. reported no significant leakage levels in their original study, when employing *Alu*-specific sgRNAs which theoretically match to thousands of targets in the human genome instead of a single target as used in their study, background leakage was determined to be the probable cause that prevented cells suspected to harbor the switch to proliferate for extended periods of time. Attempts to rear a cell line harboring the system based on the fluorescent marker expression mCherry fused to arC9 via a T2A sequence, failed due to premature cell death as exemplary depiction in Fig. 17 shows. Despite based exclusively on this observation and not conclusively proven experimentally, essential preliminary conclusions were drawn that drastically influenced downstream decisions: i) while unintentional, targeting *Alu* elements with a designer nuclease causes cell death, which was an unknown at the time the observation as related publications were not published yet and ii) it is essential that any background activity is strictly prevented as it opens up the possibility of not only killing a potential cell product but also of unintentionally inducing severe DNA damage which in a worst case scenario could lead to malignant transformation reversing the role of a safety switch to a risk factor.

7.4.2. Effect of coupling post-translational and transcriptional control on suicide-switch expression

Accordingly, the decision was made to include an additional induction mechanism in form of a tet-on system allowing transcriptional control of the allosteric Cas9 switch arC9. Conceptually, this provided several advantages over the original design. Combining two inducible systems should hypothetically lead to a drastic reduction in background activity since it would be limited to the background activity of the first stage transcriptional switch. Second, from a practical perspective, switching from a fluorescent mCherry selection marker to an antibiotic-based neomycin resistance marker allows for more rapid selection and less biased polyclonal cell line generation compared to individual clone picking. Additionally, it addresses the immunogenicity issue that would have been expected when arC9 is under the regulation of a constitutive promoter [211]. Initial transient expression tests 48 h after transfection of the new plasmids into HeLa cells provided promising results exhibiting strong DOX dependent arC9 expression measured via WB analysis (Fig. 19). Background expression in the absence of DOX was

observed to be substantially lower and was expected to be reduced further when changing the setting from a transient one, where cells are flooded with expression plasmids, to a stable one, where due to the nature of SB transposition, <12 suicide cassette copies are expected to integrate per cell [213]. The generation of a stable, transgenic, polyclonal cell population confirmed this by providing no signal of background expression in the absence of DOX, as seen in the corresponding WB analysis (Fig. 20), despite overexposure of the WB membrane. Furthermore, WB analysis were conducted after a 2-week culturing period to ensure sufficient dilution of plasmid DNA, stable expression stemming from SB-mediated genomic integration and elimination of G418-sensitive cells without stable transposon insertions. In contrast to the initial experiments attempting to generate cell lines lacking the second stage of transcriptional control, neither growth inhibition nor cell killing could be observed suggesting a lack of background arC9 activity.

7.4.3. Initial CRISISS induction in polyclonal cell cultures

The first pilot experiment to gain an insight on CRISISS' performance upon double induction with DOX and 4-HT produced both expected and unexpected results. As expected, G418-resistant polyclonal cells assumed to carry a full CRISISS cassette appeared to rapidly decline in numbers and exhibited a significantly slower growth rate than the control cell line lacking *Alu*-specific sgRNAs. This behavior seemed to peak on day 4 post-induction, where, in contrast to the control cells lacking *Alu* specific sgRNAs and reaching >50% confluency, the experimental cells displayed only <5% confluency. This accounts for a reduction of ~90% in cell growth of a cancerous cell line, which initially provided a promising insight into the potential utility of CRISISS. Unexpectedly however, the remaining cells quickly repopulated the dish and reached close to 50% confluency on day 7 post-induction. Since epigenetic silencing is rarely observed in transposon-modified cells [214] and G418 selection was continued indicating unaltered expression of the transgenic neomycin resistance gene which is part of the CRISISS cassette, it was hypothesized that a small fraction of cells that randomly incorporated incomplete segments of the CRISISS cassette through random recombination events sufficient for conferring G418 resistance but not to provide functional CRISISS action (either because these cells lack expression of Cas9 or the sgRNA or both), repopulated the dish after clearance of CRISISS harboring cells.
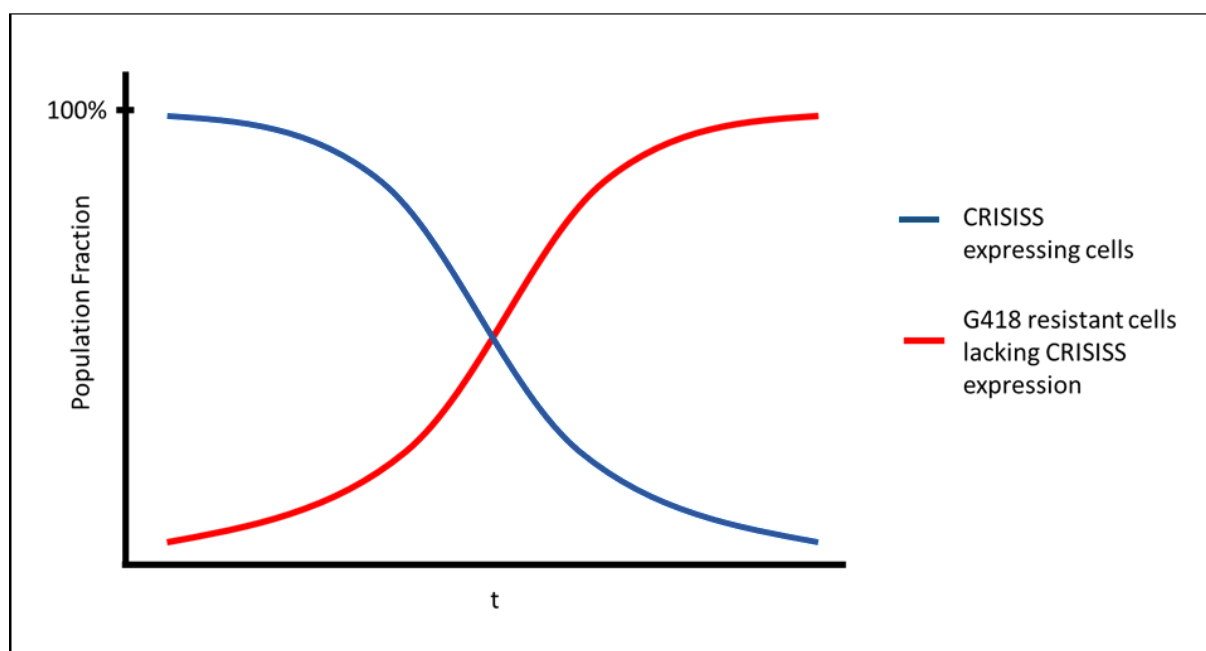
Figure 30*: Hypothetical development of two population fractions over time, including (blue line) and excluding (red) the CRISISS system while under double induction with DOX and 4-HT. (Figure was created using Biorender.com).*

In such a scenario as depicted in Fig. 30, the observed culture dish confluency would decrease as the CRISISS+ cells which constitute a majority of the cell population at the beginning die, until the G418-resistant cells, which constitute a minority of the cell population at the beginning, surpass the number of CRISISS+ cells. At this critical point, which is hypothesized to occur at day 4 post-induction, the cell population would transit to a regular growth behavior and repopulate the cell culture dish, as was observed on day 7 post-induction.

The previously described hypothesis was ultimately validated by continuing induction with DOX and 4-HT until a cell population emerged with apparent resistance to CRISISS induction and exhibiting G418 resistance. This cell population exhibited severely reduced DOX-dependent arC9 expression as determined via WB analysis in comparison to the original non-induced cells that were used as controls (Fig. 22). Interestingly, the arC9 signal obtained from the DOX- and 4-HT-resistant cells was substantially weaker yet not zero as initially expected. However, assuming that random recombination events lead to the incorporation of a functional neomycin resistance gene, the same argument can be made for non-functional or incomplete fragments of the CRISISS cassette. Since functional testing by inducing CRISISS did not cause cell death but WB analysis shows minimal expression, two scenarios would explain this observation: i) an incomplete incorporation of the arC9 gene, leading to non-functional arC9 protein but maintaining the necessary epitope for WB based detection ii) missing or incomplete incorporation of the *Alu*-specific sgRNA.

Yet, the preliminary observations concerning CRISISS-induced cell killing within a timeframe of 4 dpi were taken as an observation window to experimentally and formally confirm CRISISS-mediated apoptosis induction via flow cytometric analysis to detect apoptosis-correlated Annexin V staining and the transition into dead cells with Annexin V and the live-cell non-permeant Zombie Aqua™ dye. As described in the corresponding results section, CRISISS induction produced a growing population of apoptotic and dead cells within the chosen time window of 4 dpi (Fig. 23) in a time-dependent manner, confirming CRISISS killing effect and inducibility. Interestingly, CRISISS+ cells exposed to single induction with DOX also exhibited apoptosis and cell death, however less pronounced and delayed in comparison to the double induced sample. Considering the observations made with the initial suicide-switch that lacked a transcriptional switch in form of a tet-on system however, these initial results were expected, since substantial leakage of the allosteric switch was assumed initially and proposed to be the reason behind the impossibility of rearing a cell line harboring an only post translational inducible suicide-switch based on arC9-mediated *Alu* targeting. This observation confirms the hypothesis and warrants the need for a second regulatory element needed for induction.

7.4.4. CRISISS performance on a monoclonal cell line

In line with the obtained results, the following characterization of CRISISS performance that was presented in this work required the generation of the monoclonal cell lines harboring the CRISISS system with and without *Alu*-specific sgRNAs. Due to the methodology that was employed to rear them, which included two rounds of single clone picking, substantial expansion to establish sufficient material for downstream material and thus went through multiple passages (>10) and lengthy culture time (>1 month), allowed to get early insights on the long-term tightness of the system. As confirmed experimentally in downstream experiments (Fig. 26), differences in fitness affecting growth rates in comparison to the control cell line which excluded the *Alu*-specific sgRNAs were not observed and did not hinder cell line establishment during the rearing period.

Meeting expectations, CRISISS induction led to a strong DNA damage response as indicated with the pKAP1 assay results (Fig. 25). Remarkably, no visible increment in pKAP1 signal strength could be observed over the course of the 3 dpi period samples suggesting DNA damage to reach critical levels between 0 and 24h post induction. As such, this not only confirms CRISISS induced DNA damage but it also strongly suggests to achieve early cell cycle arrest and/or apoptosis induction. KAP1, being a transcriptional repressor of cell cycle control proteins p21 and Gadd45α as well as proapoptotic proteins Bax, Puma and Noxa loses its

repressor function when phosphorylated at serine 824 [215], which was detected in the pKAP1 assay. Additionally, through the employment of the control cell line lacking *Alu*-specific sgRNAs, the induced DNA damage can be confirmed to be strictly sgRNA dependent. Lastly but of high importance, the uninduced CRISISS containing sample showed no detectable pKAP1 levels. This experimentally confirms that background activity is not only unmeasurable by probing for leaky arC9 expression via WB, but also does not produce the slightest indication of unintended DNA damage, further strengthening the tightness of the system.

Definitive experimental confirmation of CRISISS performance was achieved by inducing the suicide-switch and performing colony forming assays with the generated monoclonal cell lines. As expected from previous observations, no leakage could be suggested when uninduced and full induction lead to a near complete cell elimination within the observed 5-day induction period, proving the suitability as an efficient first-of-its-kind CRISPR/Cas9 based suicide-switch system. The follow-up observations that monitored the marginal numbers of surviving cells identified senescent phenotypes without exceptions. For a suicide-switch system this proves highly valuable as while a close-to-complete cell killing was achieved, the remaining cells lose their proliferative properties. In an *in vivo* scenario it would be thus likely highly beneficial for two main reasons: i) considering the ultimate threat of malignant transformation of a gene therapeutic cell product, CRISISS showed that within the context of a cancerous cell line as an *in vitro* model, besides near-complete cell clearance, a potent tumor-suppressive state (reviewed in [216,217]) was achieved in the remaining cells completely stopping proliferative properties and ii) the general coupling to a senescence-associated secretory phenotype [218] would suggest a swift clearance by the immune system *in vivo* (reviewed in [219,220]) by recruiting macrophages, natural killer cells, neutrophils and T-lymphocytes through pro-inflammatory protein secretion [221].

Besides showcasing its effectiveness as a suicide-switch, the colony forming assay also confirmed the strict necessity of combining the two inducible systems at hand by exposing the leaking properties of each. Single induction with DOX driving arC9 expression led to slightly lower yet comparable killing performance as achieved with double induction. This was expected, since the arC9 post-translational switch by itself proved to be insufficient for that precise reason at the start of the project. The argument that the post-translational switch may not be necessary is however unsubstantiated for two important reasons: i) as seen during the flow cytometric apoptosis induction analysis it was clearly shown, that double induction in comparison to single DOX induction aided with CRISISS activation speed in the polyclonal setting, and, importantly ii) single induction with 4-HT also lead to a less pronounced yet

significantly reduced cell viability in comparison to the controls, highly suggesting leakiness of the tet-on system based transcriptional control. A slight indication of this was seen during the transient expression assays without DOX induction (Fig. 19) that fell under the detection limit after integration and stable transgene expression was achieved. Slowly progressing DNA damage induced through a leaky, genome-fragmenting suicide-switch must be avoided at all cost, and experimental data presented in this work suggests that the combination of both inducible systems achieves said goal.

Surprisingly however, the established monoclonal cell line employed as a negative control for the CRISISS switch lacking *Alu*-specific sgRNAs seemed to be impacted by the addition of DOX in the colony forming and overall survival assays exhibiting reduced numbers of colonies and area coverage after the induction period (Fig. 26). Seemingly contradictory, the preceding polyclonal cell population this monoclonal line was established from, did not show any signs different behavior when exposed to DOX in previous assays measuring apoptosis induction and cell death (Fig. 23). Additionally, following up with an independent assay designed to test potential toxicity of DOX on wildtype HeLa cells, no such effect could be observed (Fig. 27). It is therefore highly likely, that the observed apparent reduction on cell fitness is not attributed to an onset of cell death upon DOX induction since this would have been detected on the flow cytometric analyses mentioned previously (Fig. 23), but rather a slowing in cell proliferation which would have passed unnoticed. No conclusive data exists which independently describes the effects of overexpressing Cas9 as observed in this study. However, it could be argued that it would not be surprising if a measurable effect like the one at hand might be caused by the potential additional stress the cellular system is exposed when overexpressing a foreign protein. It is clear however, that this phenomenon is not related to potential substantial off-target activity of Cas9, as this would have been apparent in the negative control when detecting phosphorylated KAP1 in response to DNA damage (Fig. 25).

Favorable insights were lastly gained by measuring CRISISS kinetics upon induction. The expected DNA damage mediated cell cycle arrest was observed between 24 and 48h post induction, representing approximately one cell division after induction with the employed cell model. Massive cell death occurred between 72 and 96h post induction reducing on average the initially seeded cell numbers by approximately 82% and by approximately 92% with respect to the highest cell numbers detected on day 2 post induction. While in comparison to established suicide-switches 96h represents a slightly longer period for cell killing, considering that cell proliferation arrest was documented between 24 and 48h and the onset of DNA damage repair pathways was detected in a period between 0 and 24h post induction, it can nevertheless be

expected that the slightly delayed killing itself will not lead to significant complications with typical adverse effects, especially considering that while killing takes longer, no resistance and survival of proliferative cells was observed.

7.5.　　　Conclusion and Outlook Chapter II

Overall, the presented work showcases a proof of concept for a functional suicide-switch in an *in vitro* HeLa cell model. This represents the first time, CRISPR/Cas9 technology has been employed for this purpose, further expanding the utility of the former prokaryotic adoptive immunity system. The initially set goals of engineering a system that addresses shortcomings of established kill switch systems were met within the scope of this model. First, exceptional levels of cell killing were achieved coupled with no detection of surviving proliferative cells, strictly dependent on both essential components of the system: *Alu*-specific sgRNAs and double induction of translation and post translational activation. Second, no observable levels of unintended background activity when one of the two essential components was missing. And third, designing a system with the potential of being triggered *in vivo*, due to the excellent biodistribution, tissue penetration and tolerability to the small molecule inducers DOX and 4-HT exhibit [222,223]. Both goals could prove valuable and potentially develop into advantages over the established state of the art suicide mechanisms. More complete cell killing than with the iCas9 system, better potential tolerability than with the HSV-TK's ganciclovir dependence and immunogenicity issues and better potential tissue penetration than with epitope-based kill-switches. However, as CRISISS was tested exclusively in a HeLa cell model, the obvious translation into clinically relevant cells will have to be assessed. Nonetheless, current observations and performance strongly suggest that the major challenge concerning the cell type, which was achieved, would be to kill aberrant and resilient by nature cancerous and highly proliferative cells such as HeLa's or as shown recently patient-derived glioblastoma cell lines showcasing high-dose irradiation resistance [206]. It is highly unlikely that primary cells would survive CRISISS-induced genomic fragmentation. Nevertheless, several disadvantages and hurdles must be addressed and overcome moving forward from this pilot study. To show utility in a clinical setting, the neomycin resistance gene, which was used as a selection marker to facilitate cell selection but would constitutes a substantial risk if integrated into a cell product, must be removed and a therapeutic gene added. This is expected to produce potential complications. While removing a neomycin expression cassette would reduce the overall size of the transposon, a typical therapeutic gene such as a full-length CAR-cassette is typically larger, increasing the overall size of the transposon. Adding to the length of CRISPR scaffold, sgRNAs, arC9 expression cassette and reverse tetracycline controlled transactivator (rtTA), the size quickly becomes substantial while simultaneously losing an antibiotic selection marker that efficiently mediated elimination of non-modified cells. A general negative correlation between transposon size and transposition efficiency has been shown for SB [224–227]. Accordingly,

it is to be expected that this would entail a reduction in transposition efficiency with simultaneous difficulty of selecting for modified cells. There are however, potential solutions that could alleviate said bottlenecks. While transposon size and transposition efficiency seem to negatively correlate, it has been suggested following transposition experiments that successfully showed SB-mediated transposition of 100-kb long transposons, that transposition efficiency is mostly influenced by the shortest distance between the TIRs [53], as a shorter distance facilitates synaptic complex formation. On a linear strand of DNA, the distance between the TIRs is equal to the length of the transposon, as the TIRs flank it. On a plasmid however, the shortest distance usually is significantly shorter, since the plasmid backbone itself separating the TIRs can be shorter than the transposon. DNA MCs represent a well-established technology that allows generation of greatly reduced-in-size plasmids completely free of bacterial backbones via site-specific recombination [228]. Thus, MC carrying transposons contain TIRs in significantly closer proximity to each other, which has led to substantial improvements in transposition efficiency [229]. Furthermore, the size reduction itself aids electroporation and transfection efficiency facilitating membrane crossing, reduces the occurrence of innate immune responses targeted at unmethylated CpG motifs associated to bacterial DNA [49] and overall toxicity linked to cytoplasmic DNA [230]. Concerning the strict obligation to remove the antibiotic selection marker from a potential therapeutic transgene cassette, the question remains of how to enable efficient selection. For this purpose it would be highly recommended to maintain common practices to tag either the therapeutic gene or the suicide-switch itself with a non-immunogenic peptide [231] for which the cell product can not only be tracked after administration but which would additionally allow for cell sorting based selection of modified cells.

Lastly, while having substantially reduced the expected immunogenic potential of arC9 by placing it under transcriptional control, immunogenic potential remains for the rtTA, which is required to be expressed continuously for the system to work. Firstly, it is important to note however, that the absence of a detrimental immune response towards the rtTA has been observed in cynomolgus monkeys for up to 5 years [232]. Secondly, recent studies have introduced the possibility of an autoregulatory tet-on strategy which reduces the rtTA expression to background levels in absence of DOX, yet full induction and transgene expression in presence of DOX. This could be achieved by eliminating the constitutive promoter from the rtTA expression cassette, and instead link the rtTA via a self-cleaving peptide sequence to the tet-on driven transgene itself. Minimal background levels of DOX-independent expression proved to be sufficient to start DOX dependent transcription once DOX is added constituting a

positive feedback loop [233]. This would significantly reduce rtTA presence in a potential cell product to minimal levels and would reduce immunogenic issues significantly. On the contrary, local immunogenic responses to the arC9 and rtTA expression following induction would in fact further help the cause of clearing modified cells by attracting the host's immune system in an *in vivo* setting. Additionally, removing the constitutive promoter sequence would further decrease the transposon size increasing transposition efficiency. Overall, this project paved the way with a proof-of-concept study presenting CRISISS performance and presents suggestions to address current disadvantages that would realistically allow for a transition into preclinical scenarios, possible introducing an edge over current suicide-switch alternatives.

8. **Materials**

Table 4: *Antibodies*

Antigen	Clonality	Conjugation	Host	RRID	Supplier
Actin	Monoclonal	Unconjugated	Mouse	AB_2223496	Thermo Fisher
Cas9	Monoclonal	Unconjugated	Mouse	AB_2610639	Thermo Fisher
HA-tag	Monoclonal	Unconjuated	Rabbit	AB_2744968	Thermo Fisher
Histone H3	Monoclonal	HRP	Rabbit	AB_238971	Abcam
Mouse IgG	Polyclonal	HRP	Goat	AB_228313	Thermo Fisher
Myc-tag	Monoclonal	Unconjugated	Rabbit	AB_490778	Cell Signaling Technology
Phosphorylated KAP1	Monoclonal	Unconjugated	Rabbit	AB_2891814	Bethyl laboratories
Rabbit IgG	Polyclonal	HRP	Mouse	AB_228378	Thermo Fisher

Table 5: *Buffers*

Name	Composition/Supplier
10X Taq-Buffer	Thermo Fisher
3X SDS sample loading buffer	188 mM Tris-HCl (pH 6,8) 3% (w/v) SDS 30% (v/v) Glycerol 0,01 % bromphenol blue 15% (v/v) mercaptoethanol 55% (v/v) dH$_2$O
4% PFA	4% paraformaldehyde (w/v) PBS

Materials

5X Phusion™ HF-Buffer	Thermo Fisher
5X Q5® Reaction Buffer	New England Biolabs
Antarctic Phosphatase Buffer (10X)	New England Biolabs
Blocking Buffer	5% (w/v) powedered milk 95% (v/v) TBS-T
CutSmart™ Buffer	New England Biolabs
FACS buffer	1% Fetal Bovine Serum 1% Penicillin/Streptomycin 98% PBS
Lysis buffer #3 pH 7,6 (Recipe from Prof. Dr Thomas Schulz laboratory, Medizinische Hochschule Hannover)	50 mM Tris-HCl 150 mM NaCl 0,5 mM EDTA 1% (v/v) Glycerol 1% (v/v) IGEPAL CA-630 98% (v/v) dH$_2$O
Mild stripping buffer pH 2,2	1,5% (w/v) glycine 0,1% (w/v) SDS 1% (v/v) Tween 20 99% (v/v) dH$_2$O
PBS	In house production
RIPA buffer	150 mM EDTA 50 mM HEPES (pH 7,8) 150 mM NaCl 1% (v/v) IGEPAL CA-630 0,25% (w/v) Na-deoxycholate 1X Protease-inhibitor-cocktail (Roche) 1X PhosSTOP™ (Roche) (only for pKAP1 assay) 99% (v/v) dH$_2$O
SDS Running Buffer	In house production
T4 DNA Ligase Buffer (10X)	New England Biolabs
T4 PNK Reaction Buffer (10X)	New England Biolabs
TBS-T	In house production
Transfer Buffer	In house production

Materials

Table 6: Eukaryotic cell lines

Name	Clonality	Culture Conditions	Source
arC9 Alu HeLa clone #1	Monoclonal	DMEM + 10% FCS + 1% P/S, 37 °C, 5% CO_2, 1 mg/ml G418	See 7.2.8.
arC9 scaff HeLa clone #1	Monoclonal	DMEM + 10% FCS + 1% P/S, 37 °C, 5% CO_2, 1 mg/ml G418	See 7.2.8.
arC9_Alu	Polyclonal	DMEM + 10% FCS + 1% P/S, 37 °C, 5% CO_2, 1 mg/ml G418	See 7.2.4.
arC9_scaff	Polyclonal	DMEM + 10% FCS + 1% P/S, 37 °C, 5% CO_2, 1 mg/ml G418	See 7.2.4.
HEK293T	Polyclonal	DMEM + 10% FCS + 1% P/S, 37 °C, 5% CO_2	Lab stock
HeLa WT	Polyclonal	DMEM + 10% FCS + 1% P/S, 37 °C, 5% CO_2	Lab stock

Table 7: Prokaryotic cell lines

Name	Supplier
Subcloning Efficiency™ _E. coli_ DH5α	Thermo Fisher

Table 8: Chemicals

Name	Supplier
1X PhosSTOP™	Sigma
1X Protease-inhibitor-cocktail	Roche
4-Hydroxytamoxifen	Sigma
Agarose	Sigma
APS	AppliChem
Biozym LE Agarose	Biozym
Bovine Serum Albumin	Thermo Fisher
Bromphenol blue	Sigma
Campthotecin	Merck
Doxycycline Hydrochlorate	Fisher Scientific

EDTA	In-house
G418	InvivoGen
Glycerol	Sigma
Glycine	Roth
HEPES	Thermo Fisher
IGEPAL-CA 630	Sigma
Mercaptoethanol	Sigma
Methylene Blue	Roth
NaCl	Sigma
Na-deoxycholate	Sigma
Powdered milk	Roth
ROTIPHORESE®Gel 30 (37,5:1)	Roth
Sodium dodecyl sulfate	Roth
TEMED	Roth
Tris-HCl	Roth
Tween 20	Thermo Fisher

Table 9: *Enzymes*

Name	Supplier
AgeI-HF	New England Biolabs
Antarctic Phosphatase	New England Biolabs
BbsI-HF	New England Biolabs
NheI-HF	New England Biolabs
NotI-HF	New England Biolabs
Phusion™ High-Fidelity DNA Polymerase	Thermo Fisher
Q5® High-Fidelity DNA Polymerase	New England Biolabs
SalI-HF	New England Biolabs
T4 DNA Ligase	New England Biolabs
T4 Polynucleotide Kinase	New England Biolabs
Taq DNA-Polymerase	Thermo Fisher
Trypsine	In house production

Materials

Table 10: Equipment

Name	Supplier
BD™ LSRII	BD Biosciences
Bioruptor® Plus sonication device	Diagenode
C1000 Touch™ Thermal Cycler	Biorad
ECL Chemocam Imager	Intas
Gel iX20 Imager	Intas
M220 Focused Ultrasonicator	Covaris
NanoDrop™ 2000 Spectrophotometer	Peqlab
Qubit® 2.0 Fluorometer	Invitrogen
TC20 Automated Cell Counter	Biorad
Ti Eclipse Inverted Microscope	Nikon

Table 11: Kits

Name	Supplier
Amersham™ ECL Select™ Western Blotting Detection Reagent	Cytiva
Lipofectamine® 3000 Transfection Kit	Thermo Fisher
NEBNext® High-Fidelity 2X PCR Master Mix	New England Biolabs
NEBNext® Ultra™ II End Repair/dA-Tailing Module	New England Biolabs
Pacific Blue™ Annexin V Apoptosis Detection Kit with PI	Biolegend
QIAprep Spin Miniprep Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
Quick-DNA Miniprep Kit	Zymo
Zombie Aqua™ Fixable Viability Kit	Biolegend
Zymoclean Gel DNA Recovery Kit	Zymo Research
ZymoPURE II Plasmid Midiprep Kit	Zymo Reseach

Materials

Table 12: Growth media

Name	Supplier
Dulbecco's Modified Eagle's Medium (+/- appropriate antibiotic)	Thermo Fisher
LB-Medium (liquid) (+/- appropriate antibiotic)	In house production
LB-Medium (solid) (+/- appropriate antibiotic)	In house production
SOC-Medium	In house production

Table 13: Plasmids

Name	Function	Source
pRRL.PPT.SF.piggybacHApre	Expression of C-terminally HA-tagged PB transposase	Prof. Dr. Thomas Schulz
pRRL.PPT.SF.iPB-C3HApre	Expression of C-terminally HA-tagged iPB-C3 transposase [121]	Generated via mutagenesis PCR based on plasmid pRRL.PPT.SF.piggybacHApre and mutagenesis primers 2-7
pCMV_mPB	Expression vector for mouse codon optimized PB transposase	Lab stock
pCMV(CAT)T7-SB100X	Expression vector for hyperactive *Sleeping Beauty* transposase SB100X	Lab stock [234]
pT2HB_arC9_COMBI_GFP3	*Sleeping Beauty* transposon vector harboring an mCherry tagged arC9 expression cassette and a GFP-specific sgRNA between its terminal inverted repeats	Lab stock
pT2HB_arC9_COMBI_scaffold	*Sleeping Beauty* transposon vector harboring an mCherry tagged arC9 expression cassette but lacking a sgRNA sequence between its terminal inverted repeats	Lab stock
pT2HB_arC9_COMBI_Alu	*Sleeping Beauty* transposon vector harboring an mCherry tagged arC9 expression cassette and an *AluY*-specific sgRNA	Lab stock

	between its terminal inverted repeats	
pTOV_T11_Neo	*Sleeping Beauty* transposon harboring a tetracycline response element followed by a multiple cloning site, a reverse tetracycline transactivator expression cassette and a neomycine resistance gene between its terminal inverted repeats	Lab stock
pTOV_T11_Neo_arC9_Alu	*Sleeping Beauty* transposon harboring an *AluY*-specific sgRNA followed by a tetracycline response element, an arC9 expression cassette, a reverse tetracycline transactivator expression cassette and a neomycine resistance gene between its terminal inverted repeats	Combination of pTOV_T11_Neo and pT2HB_arC9_COMBI_Alu elements
pTOV_T11_Neo_arC9_scaff	*Sleeping Beauty* transposon harboring a tetracycline response element followed by an arC9 expression cassette, a reverse tetracycline transactivator expression cassette and a neomycine resistance gene between its terminal inverted repeats and lacking a sgRNA sequence	Combination of pTOV_T11_Neo and pT2HB_arC9_COMBI_scaff elements
pXL-Neo	Plasmid containing a PB transposon harboring a neomycine resistance gene used for PB transposition assays	Lab stock
pcDNA3.1.Myc-BRD4	Expression of N-terminally Myc-tagged BRD4	Prof. Dr. Thomas Schulz

Materials

Table 14*:* Oligonucleotides used for all work conducted on the PB transposase. UPPER CASE: nucleotides binding to target sequence, lower case: non-binding nucleotides introducing modification to newly synthesized sequence, *italic*: replacement codon introducing mutation for mutagenesis process, underlined: restriction site/cloning overhang [PHO]: phosphate addition to facilitate downstream ligations **bold:** primer specific barcode

#	Name	Sequence (5' → 3')	Description
1	rev_pRRL_PB_colony_pcr	GGTTGATTgtcgacTCAGGC	Serves as a sequencing primer to identify iPB hairpin mutations via sanger sequencing
2	PiggyBac_S508G_fw	CCTGTACATG*ggt*CTGACCTCC	Mutagenesis primer to introduce substitution S508G into the iPB sequence [121]
3	PiggyBac_S508G_rev	TTCCGCATGAATTTCTTCCGGC	Serves as a reverse primer paired with mutagenesis primer number 2
4	PiggyBac_N538K_fw	CATCCTGCCC*aag*GAGGTGCCCG	Mutagenesis primer to introduce substitution N538K into the iPB sequence [121]
5	PiggyBac_N538K_rev	TTGCTGATGTTGTCCCGCAGG	Serves as a reverse primer paired with mutagenesis primer number 4
6	PiggyBac_N570S_fw	GCGGAAGGCC*agc*GCCAGCTGCA	Mutagenesis primer to introduce substitution N570S into the iPB sequence [121]
7	PiggyBac_N570S_rev	CGGATCTTGCTGGGACAGTAGGTGC	Serves as a reverse primer paired with mutagenesis primer number 6
8	PiggyBac_V390A_fw	[PHO]ACCT*gcc*GGCACCAGCATG	Mutagenesis primer to introduce substitution V390A into the iPB sequence
9	PiggyBac_V390A/G391A_rev	[PHO]CTGGACCGGGAGTTCTTCAGC	Serves as a reverse primer paired with mutagenesis primers number 8 and 11

11	PiggyBac_G391A_fw	[PHO]ACCTGTG*gcc*ACCAGCATG	Mutagenesis primer to introduce substitution G391A into the iPB sequence
12	PiggyBac_T392A_fw	[PHO]GC*gcc*AGCATGTTCTGCTTCG	Mutagenesis primer to introduce substitution T392A into the iPB sequence
13	PiggyBac_T392A/S393A_rev	[PHO]CCACAGGTCTGGACCGGG	Serves as a reverse primer paired with mutagenesis primers number 12 and 14
14	PiggyBac_S393A_fw	[PHO]GCACC*gcc*ATGTTCTGCTTCG	Mutagenesis primer to introduce substitution S393A into the iPB sequence
15	PiggyBac_M394A_fw	[PHO]CAGC*gcc*TTCTGCTTCGACG	Mutagenesis primer to introduce substitution M394A into the iPB sequence
16	PiggyBac_M394A/F395A_rev	[PHO]GTGCCCACAGGTCTGGACC	Serves as a reverse primer paired with muagenesis primers number 15 and 17
17	PiggyBac_F395A_fw	[PHO]CAGCATG*gcc*TGCTTCGACG	Mutagenesis primer to introduce substitution F395A into the iPB sequence
18	PiggyBac_C396A_fw	[PHO]TTC*gcc*TTCGACGGCC	Mutagenesis primer to introduce substitution C396A into the iPB sequence
19	PiggyBac_C396A_rev	[PHO]CATGCTGGTGCCCACAGG	Serves as a reverse primer paired with mutagenesis primer number 18
20	PiggyBac_F397A_fw	[PHO]*gcc*GACGGCCCTCTGAC	Mutagenesis primer to introduce substitution F397A into the iPB sequence
21	PiggyBac_F397A/D398A_rev	[PHO]GCAGAACATGCTGGTGCCC	Serves as a reverse primer paired with mutagenesis primers number 20 and 23
23	PiggyBac_D398A_fw	[PHO]TTC*gcc*GGCCCTCTGAC	Mutagenesis primer to introduce substitution D398A into the iPB sequence

24	PiggyBac_G399A_fw	[PHO]*cc*CCTCT GACCCTGGT GTC	Mutagenesis primer to introduce substitution G399A into the iPB sequence paired with primer number 25
25	PiggyBac_G399A_rev	[PHO]*c*GTCGA AGCAGAACA TGCTGG	Mutagenesis primer to introduce substitution G399A into the iPB sequence paired with primer number 24
26	PiggyBac_P400A_fw	[PHO]*gcc*CTGA CCCTGGTGTC CTAC	Mutagenesis primer to introduce substitution P400A into the iPB sequence
27	PiggyBac_P400A/L40 1A_rev	[PHO]GCCGTC GAAGCAGAA CATGC	Serves as a reverse primer paired with mutagenesis primers number 26 and 28
28	PiggyBac_L401A_fw	[PHO]CCT*gcc*A CCCTGGTGTC C	Mutagenesis primer to introduce substitution L401A into the iPB sequence
29	PiggyBac_T402A_fw	[PHO]*gcc*CTGG TGTCCTACAA GC	Mutagenesis primer to introduce substitution T402A into the iPB sequence
30	PB_T402A/L403A/V 404A_rev	[PHO]CAGAG GGCCGTCGA AGC	Serves as a reverse primer paired with mutagenesis primers number 29, 31 and 32
31	PiggyBac_L403A_fw	[PHO]ACC*gcc* GTGTCCTACA AGC	Mutagenesis primer to introduce substitution L403A into the iPB sequence
32	PiggyBac_V404A_fw	[PHO]ACCCTG *gcc*TCCTACAA GC	Mutagenesis primer to introduce substitution V404A into the iPB sequence
33	PiggyBac_S405A_fw	[PHO]*gcc*TACA AGCCCAAGC CC	Mutagenesis primer to introduce substitution S405A into the iPB sequence
34	PiggyBac_S405A/Y4 06A_rev	[PHO]CACCA GGGTCAGAG GGC	Serves as a reverse primer paired with mutagenesis primers number 33 and 35

Materials

35	PiggyBac_Y406A_fw	[PHO]TCC*gcc*AAGCCCAAGC	Mutagenesis primer to introduce substitution Y406A into the iPB sequence
36	PiggyBac_K407A_fw	[PHO]GGTGTCCTAC*gcc*CCCAAGCC	Mutagenesis primer to introduce substitution K407A into the iPB sequence
37	PiggyBac_K407A_rev	[PHO]AGGGTCAGAGGGCCGTCG	Serves as a reverse primer paired with mutagenesis primer number 36
38	PiggyBac_P408A_fw	[PHO]*cc*AAGCCCGCCAAGATGG	Mutagenesis primer to introduce substitution P408A into the iPB sequence paired with primer number 39
39	PiggyBac_P408A_rev	[PHO]*c*CTTGTAGGACACCAGGGTC	Mutagenesis primer to introduce substitution P408A into the iPB sequence paired with primer number 38
40	PiggyBac_K409A_fw	[PHO]CTACAAGCCC*gcc*CCCGCC	Mutagenesis primer to introduce substitution K409A into the iPB sequence
41	PiggyBac_K409A_rev	[PHO]GACACCAGGGTCAGAGGG	Serves as a reverse primer paired with mutagenesis primer number 40
42	PiggyBac_P410A_fw_Alternative	[PHO]CCAAG*gcc*GCCAAGATGG	Mutagenesis primer to introduce substitution P410A into the iPB sequence
43	PiggyBac_P410A_rev_Alternative	[PHO]GCTTGTAGGACACCAGGGTC	Serves as a reverse primer paired with mutagenesis primer number 42
44	PiggyBac_V390K_fw	[PHO]ACCT*aag*GGCACCAGCATG	Mutagenesis primer to introduce substitution V390K into the iPB sequence
45	PiggyBac_F397R_fw	[PHO]*cgg*GACGGCCCTCTGAC	Mutagenesis primer to introduce substitution F397R into the iPB sequence
46	PiggyBac_F397Y_fw	[PHO]*tac*GACGGCCCTCTGAC	Mutagenesis primer to introduce substitution F397Y into the iPB sequence

47	PiggyBac_F397G_fw	[PHO]*ggc*GAC GGCCCTCTGA C	Mutagenesis primer to introduce substitution F397G into the iPB sequence
48	PiggyBac_Y406W_fw	[PHO]TCC*tgg*A AGCCCAAGC	Mutagenesis primer to introduce substitution Y406W into the iPB sequence
49	PiggyBac_K407C_fw	[PHO]GGTGTC CTAC*tgc*CCCA AGCC	Mutagenesis primer to introduce substitution K407C into the iPB sequence
50	PiggyBac_K407Q_fw	[PHO]GGTGTC CTAC*cag*CCC AAGCC	Mutagenesis primer to introduce substitution K407Q into the iPB sequence
51	PiggyBac_K407V_fw	[PHO]GGTGTC CTAC*gtg*CCCA AGCC	Mutagenesis primer to introduce substitution K407V into the iPB sequence
52	PiggyBac_K407T_fw	[PHO]GGTGTC CTAC*acc*CCCA AGCC	Mutagenesis primer to introduce substitution K407T into the iPB sequence
53	PB_PA410/411_del_f w.2	[PHO]AAGAA GATGGTGTA CCTGC	Forward mutagenesis primer to introduce deletion P410_A411del into the iPB sequence
54	PB_PA410/411_del_r ev.2	[PHO]GGGCTT GTAGGACAC CAG	Reverse mutagenesis primer to introduce deletion P410_A411del into the iPB sequence
55	PiggyBac_M394E_fw	[PHO]CAGC*ga a*TTCTGCTTC GACG	Mutagenesis primer to introduce substitution M394E into the iPB sequence
56	PiggyBac_M394R_fw	[PHO]CAGC*cgt*TTCTGCTTCG ACG	Mutagenesis primer to introduce substitution M394R into the iPB and mPB sequence
57	PiggyBac_C396E_fw. 2	[PHO]TTCGAA TTCGACGGCC CTC	Mutagenesis primer to introduce substitution C396E into the iPB sequence

58	PiggyBac_C396R_fw	[PHO]TTC*cgt*TTCGACGGCC	Mutagenesis primer to introduce substitution C396R into the iPB sequence
59	PiggyBac_T392Y_fw	[PHO]GC*tat*AGCATGTTCTGCTTCG	Mutagenesis primer to introduce substitution T392Y into the iPB and mPB sequence
60	PiggyBac_S387Q_fw.2	[PHO]CGG*cag*AGACCTGTGGGC	Mutagenesis primer to introduce substitution S387Q into the iPB sequence
61	PiggyBac_S387Q_rev.2	[PHO]GGAGTTCTTCAGCACCTCGG	Serves as a reverse primer paired with mutagenesis primer number 60
62	PiggyBac_E380F_fw.2	[PHO]CCC*ttt*GTGCTGAAGAACTCCC	Mutagenesis primer to introduce substitution E380F into the iPB sequence
63	PB_E380F/W/V381D_rev.2	[PHO]AATCTCGCGCTTGTTGGACC	Serves as a reverse primer paired with mutagenesis primers number 62 and 65
64	PiggyBac_E380W_fw.2	[PHO]CCC*tgg*GTGCTGAAGAACTCCC	Mutagenesis primer to introduce substitution E380W into the iPB sequence
65	PiggyBac_V381D_fw	[PHO]GAG*gat*CTGAAGAACTCC	Mutagenesis primer to introduce substitution V391D into the iPB sequence
66	SV40_polyA_rev	[PHO]AGTTGTGGTTTGTCCAAACTC	Serves as a sequencing primer to identify mPB hairpin mutations via sanger sequencing
67	mPB_V390K_fw	[PHO]CCC*aag*GGCACCAGCATG	Mutagenesis primer to introduce substitution V390K into the mPB sequence
68	mPB_V390K_rev	[PHO]CCTGGACCTGCTGTTCTTCAGG	Serves as a reverse primer paired with mutagenesis primer number 67

69	mPB_K407C_fw	[PHO]*tgc*CCCA AGCCCGCCA AGATGG	Mutagenesis primer to introduce substitution K407C into the mPB sequence
70	mPB_K407Q_fw	[PHO]*cag*CCC AAGCCCGCC AAGATGG	Mutagenesis primer to introduce substitution K407Q into the mPB sequence
71	mPB_K407V_fw	[PHO]*gtg*CCCA AGCCCGCCA AGATGG	Mutagenesis primer to introduce substitution K407V into the mPB sequence
72	mPB_K407C/Q/V_rev	[PHO]GTAGG ACACCAGGG TCAGGG	Serves as a reverse primer paired with mutagenesis primers number 69, 70 and 71
73	mPB_T392Y_rev.2	[PHO]CCACG GGCCTGGAC C	Serves as a reverse primer paired with mutagenesis primer number 59
74	mPB_M394R_rev	[PHO]GTGCCC ACGGGCCTG	Serves as a reverse primer paired with mutagenesis primer number 56
75	mPB_E380F_fw	[PHO]CC*ttt*GT CCTGAAGAA CAG	Mutagenesis primer to introduce substitution E380F into the mPB sequence
76	mPB_E380F_rev	[PHO]GGATCT CTCTCTTGTT GC	Serves as a reverse primer paired with mutagenesis primer number 75
77	Linker top strand	gtaatacgactcacta tagggctccgcttaag ggactcagacgtgtg ctcttccgatc	Top strand of linker used to ligate with ends of sheared genomic DNA to introduce a known sequence that can be used as a primer target for PCR
78	Linker bottom strand	gatcggaagagcaca cg-SpacerC3	Bottom strand of linker used to ligate with ends of sheared genomic DNA to introduce a known sequence that can be used as a primer target for PCR
79	Linker Primer	GTAATACGA CTCACTATAG GGC	Primer specific to linker

80	PB Uno	CAGACCGAT AAAACACAT GCGTCA	Primer specific to PB-TIR
81	PB5h_bc7_ill	aatgatacggcgacc accgagatctacactc tttccctacacgacgc tcttccgatct**cagat c**AACGTACGT CACAATATG ATTATCTTTC	Illumina-Primer specific to PB-TIR that introduces a 6 nucleotide long barcode sequence to identify reads downstream
82	PB5h_bc8_ill	aatgatacggcgacc accgagatctacactc tttccctacacgacgc tcttccgatct**acttga** AACGTACGT CACAATATG ATTATCTTTC	Illumina-Primer specific to PB-TIR that introduces a 6 nucleotide long barcode sequence to identify reads downstream
83	PB5h_bc9_ill	aatgatacggcgacc accgagatctacactc tttccctacacgacgc tcttccgatct**gatca g**AACGTACGT CACAATATG ATTATCTTTC	Illumina-Primer specific to PB-TIR that introduces a 6 nucleotide long barcode sequence to identify reads downstream
84	PB5h_bc10_ill	aatgatacggcgacc accgagatctacactc tttccctacacgacgc tcttccgatct**tagctt** AACGTACGT CACAATATG ATTATCTTTC	Illumina-Primer specific to PB-TIR that introduces a 6 nucleotide long barcode sequence to identify reads downstream
85	PE_nest_ind7	caagcagaagacgg catacgagat**gatct g**GTGACTGGA	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide

		GTTCAGACGT GTGCTCTTCC GATCT	long barcode sequence to identify reads downstream
86	PE_nest_ind8	caagcagaagacgg catacgagat**tcaag t**GTGACTGGA GTTCAGACGT GTGCTCTTCC GATCT	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide long barcode sequence to identify reads downstream
87	PE_nest_ind9	caagcagaagacgg catacgagat**ctgatc** GTGACTGGA GTTCAGACGT GTGCTCTTCC GATCT	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide long barcode sequence to identify reads downstream
88	PE_nest_ind10	caagcagaagacgg catacgagat**aagct a**GTGACTGGA GTTCAGACGT GTGCTCTTCC GATCT	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide long barcode sequence to identify reads downstream
89	PE_nest_ind11	caagcagaagacgg catacgagat**gtagc c**GTGACTGGA GTTCAGACGT GTGCTCTTCC GATCT	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide long barcode sequence to identify reads downstream
90	PE_nest_ind12	caagcagaagacgg catacgagat**tacaa g**GTGACTGGA GTTCAGACGT	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide long barcode sequence to identify reads downstream

		GTGCTCTTCC GATCT	
91	PE_nest_ind13	caagcagaagacgg catacgagat**ttgact** GTGACTGGA GTTCAGACGT GTGCTCTTCC GATCT	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide long barcode sequence to identify reads downstream
92	PE_nest_ind14	caagcagaagacgg catacgagat**ggaac T**GTGACTGG AGTTCAGAC GTGTGCTCTT CCGATCT	Illumina-Primer specific to linker sequence that introduces a 6 nucleotide long barcode sequence to identify reads downstream
93	arC9_SalI_fw	ttt<u>gtcgac</u> ATGGATTAC AAAGACGAT GACG	Amplification of the arC9 coding sequence from the pT2HB-Vector and addition of a SalI restriction site for cloning into the pTOV-Vector. Additionally used for Sanger sequencing to exclude errors introduced via PCR amplification.
94	arC9_STOP_NotI_rev	ttt<u>gcggccgc</u> *tca*GGATTCGG ATCCGTCG	Amplification of the arC9 coding sequence from the pT2HB-Vector and addition of a Stop-codon followed by a NotI restriction site for cloning into the pTOV-Vector.
95	hU6_NheI_fw	ttt<u>gctagc</u>TTTCC CATGATTCCT TCATATTTGC	Amplification of the hU6 promoter and the CRISPR scaffold ± *Alu*-specific sgRNA from the pT2HB-Vector and addition of a NheI restriction site for cloning into the pTOV-Vector. Additionally used for Sanger sequencing to exclude errors introduced via PCR amplification.

96	gRNAscaff_AgeI_rev	tttaccggtAAAA AAAGCACCG ACTCGG	Amplification of the hU6 promoter and the CRISPR scaffold ± *Alu*-specific sgRNA from the pT2HB-Vector and addition of a AgeI restriction site for cloning into the pTOV-Vector.
97	arC9_Seq1	CTCCAGCAA ACAGTCTTGC C	Used for Sanger sequencing of the arC9 coding sequence to exclude errors introduced via PCR amplification.
98	arC9_Seq2	AAGTGTAAG AACGTAGTG CC	Used for Sanger sequencing of the arC9 coding sequence to exclude errors introduced via PCR amplification.
99	arC9_Seq3	AGAAGATCC TGACCTTCCG	Used for Sanger sequencing of the arC9 coding sequence to exclude errors introduced via PCR amplification.
100	arC9_Seq4	CATCCGGGA CAAGCAGTC CG	Used for Sanger sequencing of the arC9 coding sequence to exclude errors introduced via PCR amplification.
101	arC9_Seq5	TTCGACAATC TGACCAAGG C	Used for Sanger sequencing of the arC9 coding sequence to exclude errors introduced via PCR amplification.
102	arC9_Seq6	TACGGCGGC TTCGACAGCC	Used for Sanger sequencing of the arC9 coding sequence to exclude errors introduced via PCR amplification.
103	sgAluY-1_top	<u>CACC</u>GTCCC AAAGTGCTG GGATTAC	Top strand of Alu-specific sgRNA used throughout the project as CRISISS target (received from Adrian Kovač)
104	sgAluY-1_btm	<u>AAAC</u>GTAAT CCCAGCACTT TGGGAC	Bottom strand of Alu-specific sgRNA used throughout the project as CRISISS target (received from Adrian Kovač)

Table 15: Reagents

Name	Supplier
c-Myc Monoclonal Antibody-Agarose Beads	Takara
dNTP-Mix (10 mM each)	Thermo Fisher

Fetal Bovine Serum	Pan Biotech
GeneRuler 1 kb Plus DNA Ladder	Thermo Fisher
Magnetic Beads	Beckman Coulter
PageRuler™ Plus Prestained Protein Ladder	Thermo Fisher
Protein G Sepharose™ 4 Fast Flow	Cytiva
Q5® High-Fidelity 2X Master Mix	New England Biolabs
TransIT®-LT1 Transfection Reagent	Mirus

Table 16: *Services*

Name	Provider
Next-Generation Sequencing	In-house (Dr. Csaba Miskey)
Sanger Sequencing	Eurofins Genomics

Table 17: *Software*

Name	Provider
ESPript 3.0	[120]
FloJo™ v10	Becton Dickinson
GraphPad Prism 9	GraphPad Software
ImageJ	Wayne Rasband (NIH)
MAFFT version 7	[118,119]
Microsoft Excel	Microsoft
Microsoft Word	Microsoft
Tm Calculator (version 1.13.0)	New England Biolabs

10.   List of tables

11. <u>Abbreviations</u>

%	Percent
° C	Degree Celcius
μg	Microgram
μl	Microliter
μM	Micromolar
4-HT	4-Hydroxytamoxifen
A	Adenine
aa	amino acid
AAV	Adeno-Associated Virus
ACE2	Angiotensin-Converting Enzyme 2
arC9	allosterically-regulated Cas9
ATP	Adenosine Triphosphate
BACH2	Broad complex-tramtrack-bric A brac and Cap'n'collar Homology 2
BCMA	C-cell Maturation Antigen
BET	Bromodomain and Extra-Terminal domain
Bp	Base pair
BRD	Bromodomain containing protein
BSL	Biological Safety Level
C	Cytosine
CAR	Chimeric Antigen Receptor
Cas	CRISPR associated protein
CCD	Charge-Coupled Device
CD	Cluster of Differentiation
cm	Centimeter
Co-IP	Co-Immunoprecipitation
CRD	Cysteine Rich Domain
CRISISS	CRISPR Induced Suicide-Switch
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
CRS	Cytokine Release Syndrome

Abbreviations

Da	Dalton
dA	Deoxyadenosine
DB	Doggybone
DDBD	Dimerization and DNA-Binding Domain
dH_2O	Destilled Water
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DOX	Doxycycline
Dpi	Day(s) post induction
DSB	Double Strand Break
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal Growth Factor Receptor
EM	Electron Microscopy
ER-LBD	Estrogen Receptor α Ligand Binding Domain
ESR	Estrogen Receptor
FACS	Fluorescence Activated Cell Sorting
Fig.	Figure
FKBP	FK506 binding protein
g	Gravity
G	Guanine
GMP	Good Manufacturing Practice
GOI	Gene Of Interest
gRNA	Guide RNA
GvHD	Graft-versus-Host Disease
h	Hours
H3	Histone H3
HA	Human influenza hemagglutinin
HDR	Homology Directed Repair
HRP	Horseradish Peroxidase
HSV	Herpes Simplex Virus
iCasp9	Inducible Caspase 9

Abbreviations

ID	Insertion Domain
IgG	Immunoglobulin
IN	Integrase
Indel	insertions/deletions
iPB	insect PB
iPSC	induced Pluripotent Stem Cell
k	Kilo
kb	Kilobases
LB	Lysogeny Broth
LTR	Long Terminal Repeat
MC	Minicircle
mg	Milligram
min.	Minutes
ml	Milliliter
MLV	Murine Leukemia Virus
mM	Millimolar
mPB	mouse PB
mRNA	messenger RNA
NCBI	National Center for Biotechnology Information
ng	Nanogram
NGS	Next-Generation Sequencing
NHEJ	Non-Homologous End Joining
NLS	Nuclear Localization Signal
nM	Nanomolar
NSCLC	Non-Small Cell Lung Cancer
nt	Nucleotides
NTD	N-Termina Domain
ORF	Open Reading Frame
P	Phosphate
PAGE	Polyacrylamide Gel Electrophoresis
PAM	Protospacer-Adjacent Motif

PB	piggyBac
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PD1	Programmed cell Death protein 1
PDB	Protein Data Bank
PD-L1	Programmed Death-Ligand 1
PFA	Paraformaldehyde
pFAR	Plasmid Free of Antibiotic Resistance
PSMA	Prostate-Specific Membrane Antigen
RNA	Ribonucleic Acid
RIPA	Radioimmunoprecipitation Assay
rpm	Revolutions per minute
rtTA	reverse tetracycline controlled Transactivator
s	Seconds
SARS-CoV2	Severe Acute Respiratory Syndrome Coronavirus 2
SB	Sleeping Beauty
Scaff	Scaffold
SCID-X1	X-linked Severe Combined Immunodeficiency
SDS	Sodium Dodecyl Sulfate
sgRNA	single guide RNA
STC	Strand Transfer Complex
STR	Sub-Terminal Repeat
SV	Structural Variant
T	Thymine
Tab.	Table
TALEN	Transcription Activator-Like Effector Nuclease
TBS-T	Tris-Buffered Saline-Tween
TCR	T Cell Receptor
TE	Transposable Element
tet	Tetracycline
TIR	Terminal Inverted Repeat

TK	Thymidine Kinase
tracrRNA	Transactivating crRNA
TSS	Transcriptional Start-Site
WB	Western Blot
WT	Wildtype

12. <u>**Supplementary material**</u>

Figure 31*: Supplemental WB analysis membranes corresponding to Chapter I Co-IPs #1*
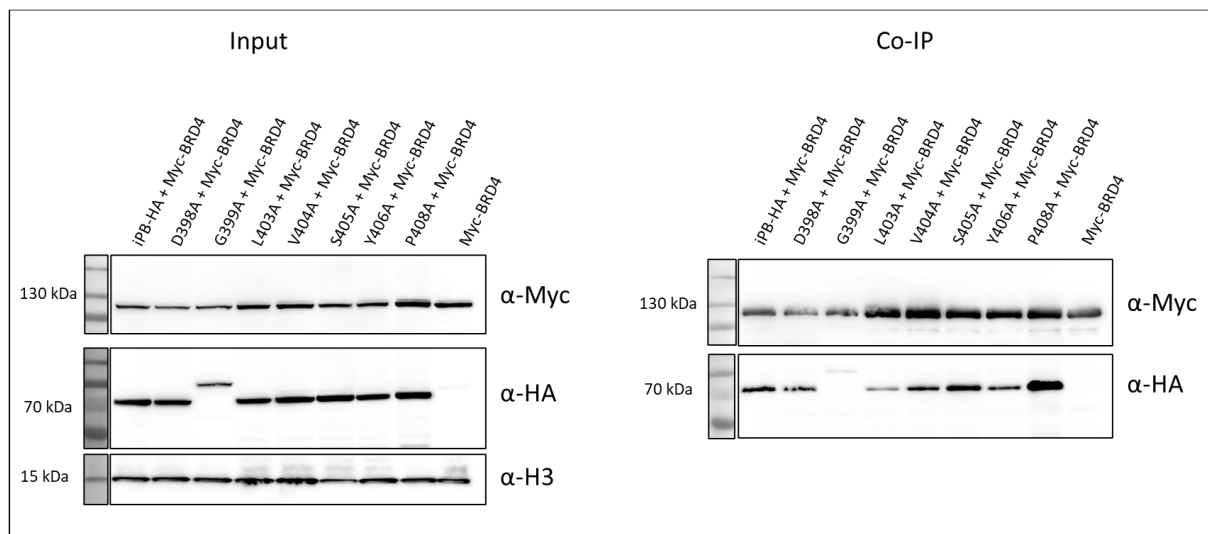
Figure 32*: Supplemental WB analysis membranes corresponding to Chapter I Co-IPs #2*
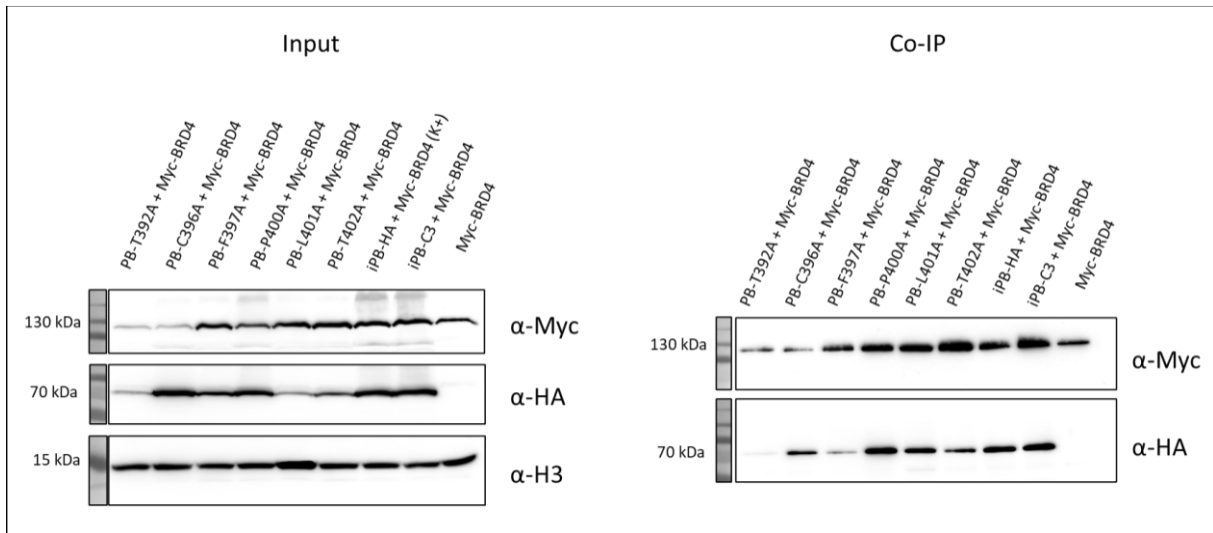
Figure 33: *Supplemental WB analysis membranes corresponding to Chapter I Co-IPs #3*
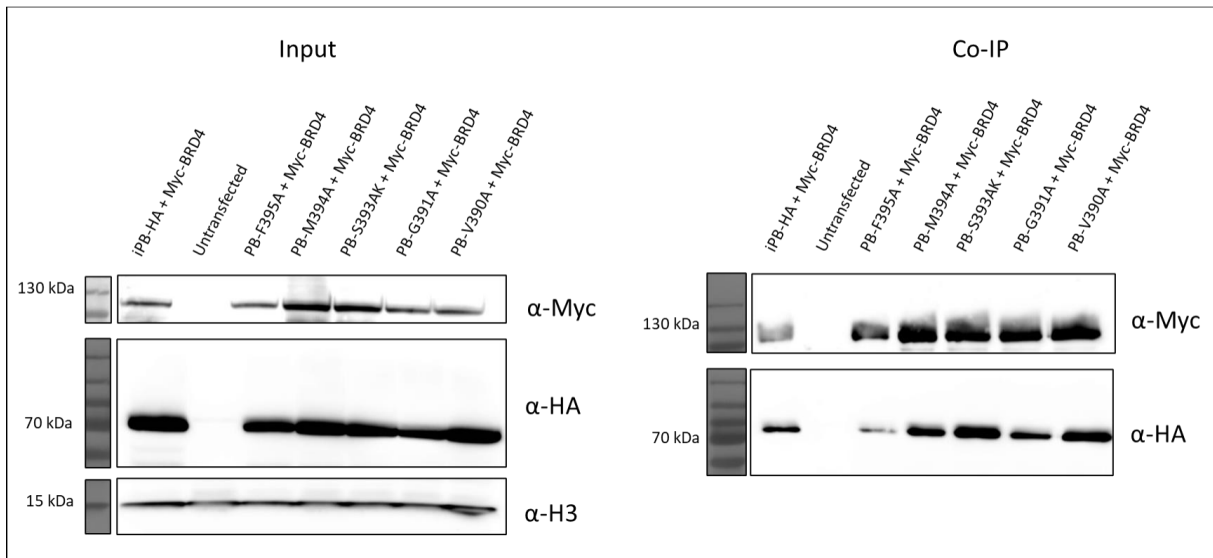
Figure 34*: Supplemental WB analysis membranes corresponding to Chapter I Co-IPs #4*
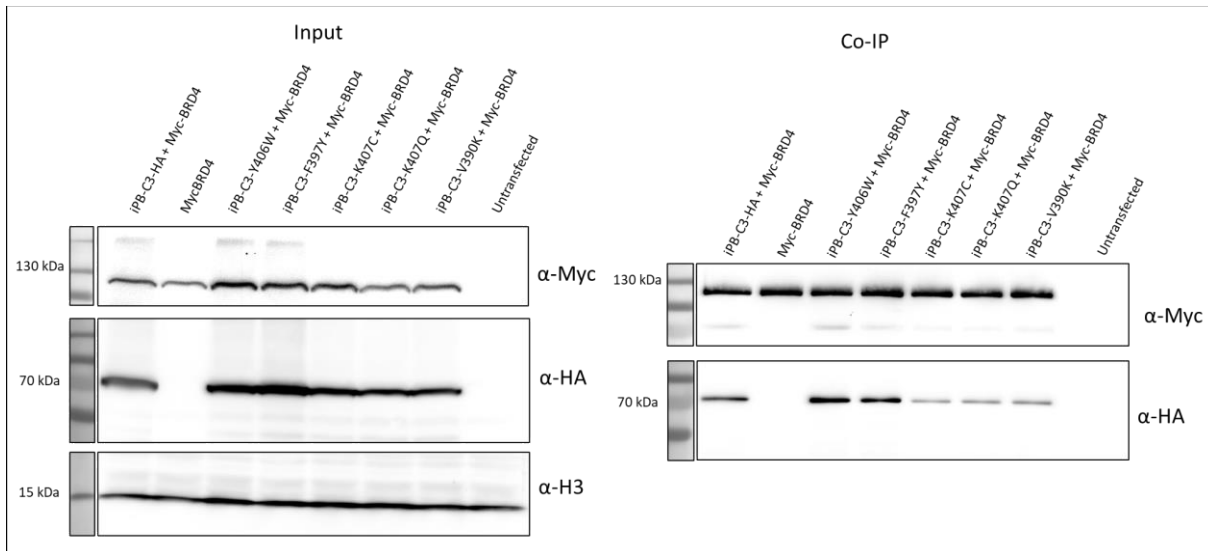
Figure 35*: Supplemental WB analysis membranes corresponding to Chapter I Co-IPs #5*

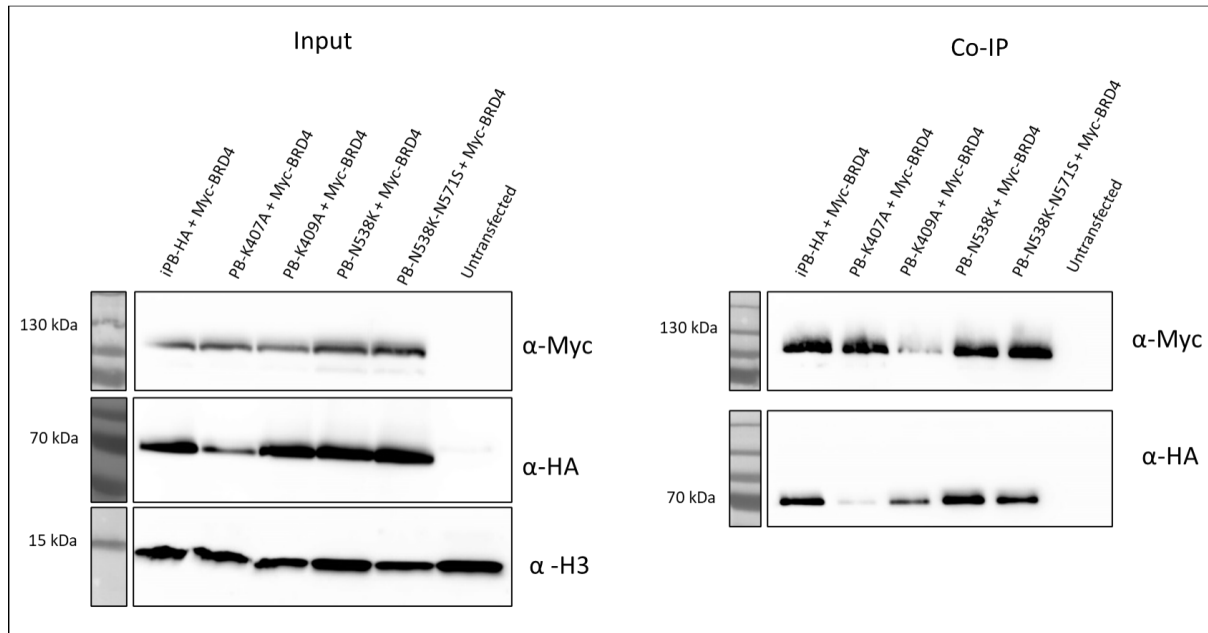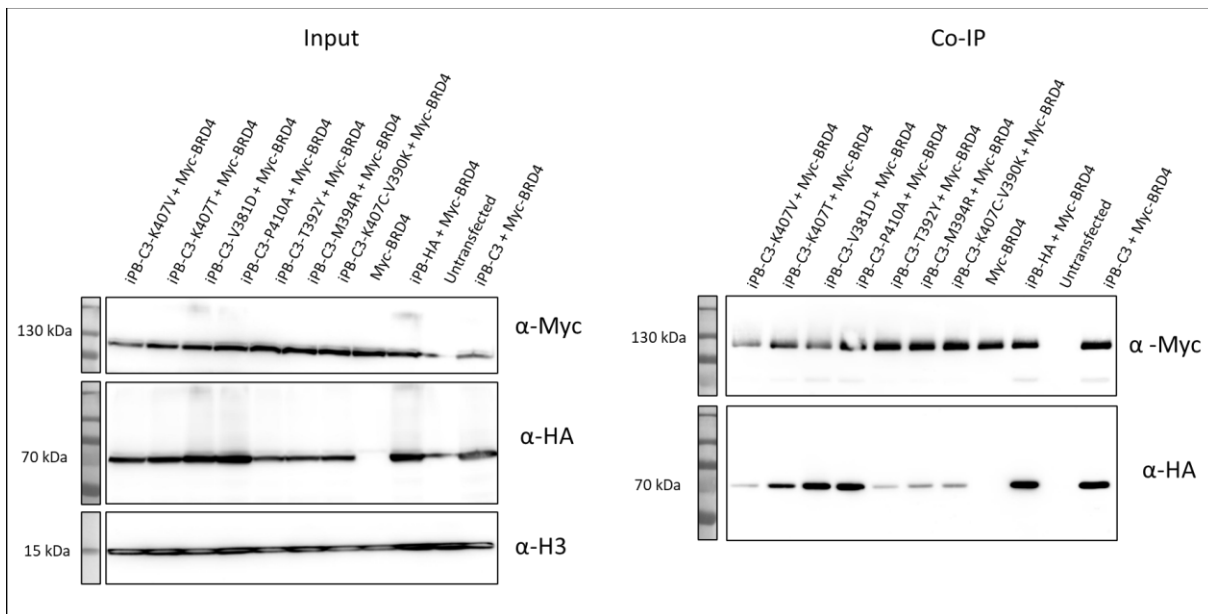***Figure 36****: Supplemental WB analysis membranes corresponding to Chapter I Co-IPs #6*
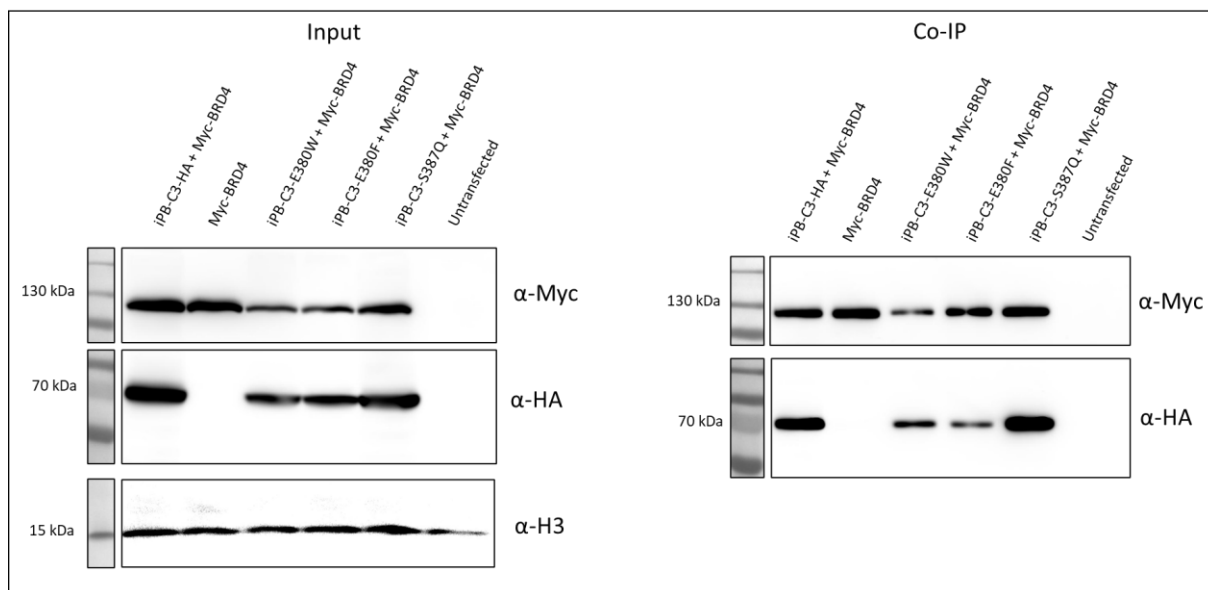
Figure 37: *Supplemental WB analysis membranes corresponding to Chapter I Co-IPs #7*

13.   <u>**References**</u>

1. Miescher, F. Letter to Wilhelm His, Tübingen, February 26th, 1869. *Die histochemischen und physiologischen arbeiten von Friedrich Miescher-aus dem wissenschaftlichen Briefwechsel von F. Miescher* **1869**, 33–38.
2. Watson, J.D.; Crick, F.H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737–738.
3. Crick, F.H.; Barnett, L.; Brenner, S.; Watts-Tobin, R.J. General nature of the genetic code for proteins. *Nature* **1961**, *192*, 1227–1232, doi:10.1038/1921227a0.
4. Nirenberg, M. Historical review: Deciphering the genetic code--a personal account. *Trends Biochem. Sci.* **2004**, *29*, 46–54, doi:10.1016/j.tibs.2003.11.009.
5. Roberts, R.J. How restriction enzymes became the workhorses of molecular biology. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 5905–5908, doi:10.1073/pnas.0500923102.
6. Lehman, I.R. DNA ligase: structure, mechanism, and function. *Science* **1974**, *186*, 790–797, doi:10.1126/science.186.4166.790.
7. Heather, J.M.; Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genomics* **2016**, *107*, 1–8, doi:10.1016/j.ygeno.2015.11.003.
8. Mullis, K.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **1986**, *51 Pt 1*, 263–273, doi:10.1101/sqb.1986.051.01.032.
9. Bulcha, J.T.; Wang, Y.; Ma, H.; Tai, P.W.L.; Gao, G. Viral vector platforms within the gene therapy landscape. *Signal Transduct. Target. Ther.* **2021**, *6*, doi:10.1038/s41392-021-00487-6.
10. Wu, X.; Li, Y.; Crise, B.; Burgess, S.M. Transcription start regions in the human genome are favored targets for MLV integration. *Science* **2003**, *300*, 1749–1751, doi:10.1126/science.1083413.
11. Schröder, A.R.W.; Shinn, P.; Chen, H.; Berry, C.; Ecker, J.R.; Bushman, F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **2002**, *110*, 521–529, doi:10.1016/s0092-8674(02)00864-4.
12. Ahi, Y.S.; Bangari, D.S.; Mittal, S.K. Adenoviral vector immunity: its implications and circumvention strategies. *Curr. Gene Ther.* **2011**, *11*, 307–320, doi:10.2174/156652311796150372.
13. Mingozzi, F.; High, K.A. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* **2013**, *122*, 23–36, doi:10.1182/blood-2013-01-306647.
14. Sahin, U.; Muik, A.; Derhovanessian, E.; Vogler, I.; Kranz, L.M.; Vormehr, M.; Baum, A.; Pascal, K.; Quandt, J.; Maurus, D.; et al. COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses. *Nature* **2020**, *586*, 594–599, doi:10.1038/s41586-020-2814-7.
15. Baden, L.R.; El Sahly, H.M.; Essink, B.; Kotloff, K.; Frey, S.; Novak, R.; Diemert, D.; Spector, S.A.; Rouphael, N.; Creech, C.B.; et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N. Engl. J. Med.* **2021**, *384*, 403–416, doi:10.1056/NEJMoa2035389.
16. Karikó, K.; Muramatsu, H.; Welsh, F.A.; Ludwig, J.; Kato, H.; Akira, S.; Weissman, D. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Mol. Ther.* **2008**, *16*, 1833–1840, doi:10.1038/mt.2008.200.

References

17. Hou, X.; Zaks, T.; Langer, R.; Dong, Y. Lipid nanoparticles for mRNA delivery. *Nat. Rev. Mater.* **2021**, *6*, 1078–1094, doi:10.1038/s41578-021-00358-0.

18. Lancaster, E.M.; Jablons, D.; Kratz, J.R. Applications of Next-Generation Sequencing in Neoantigen Prediction and Cancer Vaccine Development. *Genet. Test. Mol. Biomarkers* **2020**, *24*, 59–66, doi:10.1089/gtmb.2018.0211.

19. Mulia, G.E.; Picanço-Castro, V.; Stavrou, E.F.; Athanassiadou, A.; Figueiredo, M.L. Advances in the Development and the Applications of Nonviral, Episomal Vectors for Gene Therapy. *Hum. Gene Ther.* **2021**, *32*, 1076–1095, doi:10.1089/hum.2020.310.

20. Piechaczek, C.; Fetzer, C.; Baiker, A.; Bode, J.; Lipps, H.J. A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. *Nucleic Acids Res.* **1999**, *27*, 426–428, doi:10.1093/nar/27.2.426.

21. Jenke, A.C.W.; Wilhelm, A.D.; Orth, V.; Lipps, H.J.; Protzer, U.; Wirth, S. Long-term suppression of hepatitis B virus replication by short hairpin RNA expression using the scaffold/matrix attachment region-based replicating vector system pEPI-1. *Antimicrob. Agents Chemother.* **2008**, *52*, 2355–2359, doi:10.1128/AAC.00067-08.

22. Koirala, A.; Makkia, R.S.; Conley, S.M.; Cooper, M.J.; Naash, M.I. S/MAR-containing DNA nanoparticles promote persistent RPE gene expression and improvement in RPE65-associated LCA. *Hum. Mol. Genet.* **2013**, *22*, 1632–1642, doi:10.1093/hmg/ddt013.

23. Rocco, D. de; Pompili, B.; Castellani, S.; Morini, E.; Cavinato, L.; Cimino, G.; Mariggiò, M.A.; Guarnieri, S.; Conese, M.; Del Porto, P.; et al. Assembly and Functional Analysis of an S/MAR Based Episome with the Cystic Fibrosis Transmembrane Conductance Regulator Gene. *Int. J. Mol. Sci.* **2018**, *19*, doi:10.3390/ijms19041220.

24. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U. S. A.* **1950**, *36*, 344–355, doi:10.1073/pnas.36.6.344.

25. Kleckner, N. Transposable elements in prokaryotes. *Annu. Rev. Genet.* **1981**, *15*, 341–404, doi:10.1146/annurev.ge.15.120181.002013.

26. Bourque, G.; Burns, K.H.; Gehring, M.; Gorbunova, V.; Seluanov, A.; Hammell, M.; Imbeault, M.; Izsvák, Z.; Levin, H.L.; Macfarlan, T.S.; et al. Ten things you should know about transposable elements. *Genome Biol.* **2018**, *19*, 199, doi:10.1186/s13059-018-1577-z.

27. Boeke, J.D.; Garfinkel, D.J.; Styles, C.A.; Fink, G.R. Ty elements transpose through an RNA intermediate. *Cell* **1985**, *40*, 491–500, doi:10.1016/0092-8674(85)90197-7.

28. Grabundzija, I.; Messing, S.A.; Thomas, J.; Cosby, R.L.; Bilic, I.; Miskey, C.; Gogol-Döring, A.; Kapitonov, V.; Diem, T.; Dalda, A.; et al. A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat. Commun.* **2016**, *7*, 10716, doi:10.1038/ncomms10716.

29. Greenblatt, I.M. Transposition and replication of modulator in maize. *Genetics* **1966**, *53*, 361–369, doi:10.1093/genetics/53.2.361.

30. Hickman, A.B.; Dyda, F. DNA Transposition at Work. *Chem. Rev.* **2016**, *116*, 12758–12784, doi:10.1021/acs.chemrev.6b00003.

31. Kapitonov, V.V.; Jurka, J. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol.* **2005**, *3*, e181, doi:10.1371/journal.pbio.0030181.

32. Ivics, Z.; Izsvak, Z.; Minter, A.; Hackett, P.B. Identification of functional domains and evolution of Tc1-like transposable elements. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 5008–5013, doi:10.1073/pnas.93.10.5008.

References

33. Ivics, Z.; Hackett, P.B.; Plasterk, R.H.; Izsvák, Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **1997**, *91*, 501–510, doi:10.1016/s0092-8674(00)80436-5.

34. Ochmann, M.T.; Ivics, Z. Jumping Ahead with Sleeping Beauty: Mechanistic Insights into Cut-and-Paste Transposition. *Viruses* **2021**, *13*, doi:10.3390/v13010076.

35. Sandoval-Villegas, N.; Nurieva, W.; Amberger, M.; Ivics, Z. Contemporary Transposon Tools: A Review and Guide through Mechanisms and Applications of Sleeping Beauty, piggyBac and Tol2 for Genome Engineering. *Int. J. Mol. Sci.* **2021**, *22*, doi:10.3390/ijms22105084.

36. Carroll, D. Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* **2014**, *83*, 409–439, doi:10.1146/annurev-biochem-060713-035418.

37. Cavazzana-Calvo, M.; Hacein-Bey, S.; Saint Basile, G. de; Gross, F.; Yvon, E.; Nusbaum, P.; Selz, F.; Hue, C.; Certain, S.; Casanova, J.L.; et al. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **2000**, *288*, 669–672, doi:10.1126/science.288.5466.669.

38. Hacein-Bey-Abina, S.; Le Deist, F.; Carlier, F.; Bouneaud, C.; Hue, C.; Villartay, J.-P. de; Thrasher, A.J.; Wulffraat, N.; Sorensen, R.; Dupuis-Girod, S.; et al. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N. Engl. J. Med.* **2002**, *346*, 1185–1193, doi:10.1056/NEJMoa012616.

39. Gaspar, H.B.; Parsley, K.L.; Howe, S.; King, D.; Gilmour, K.C.; Sinclair, J.; Brouns, G.; Schmidt, M.; Kalle, C. von; Barington, T.; et al. Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* **2004**, *364*, 2181–2187, doi:10.1016/S0140-6736(04)17590-9.

40. Howe, S.J.; Mansour, M.R.; Schwarzwaelder, K.; Bartholomae, C.; Hubank, M.; Kempski, H.; Brugman, M.H.; Pike-Overzet, K.; Chatters, S.J.; Ridder, D. de; et al. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* **2008**, *118*, 3143–3150, doi:10.1172/JCI35798.

41. Hacein-Bey-Abina, S.; Garrigue, A.; Wang, G.P.; Soulier, J.; Lim, A.; Morillon, E.; Clappier, E.; Caccavelli, L.; Delabesse, E.; Beldjord, K.; et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* **2008**, *118*, 3132–3142, doi:10.1172/JCI35700.

42. Cattoglio, C.; Facchini, G.; Sartori, D.; Antonelli, A.; Miccio, A.; Cassani, B.; Schmidt, M.; Kalle, C. von; Howe, S.; Thrasher, A.J.; et al. Hot spots of retroviral integration in human CD34+ hematopoietic cells. *Blood* **2007**, *110*, 1770–1778, doi:10.1182/blood-2007-01-068759.

43. Almarza, D.; Bussadori, G.; Navarro, M.; Mavilio, F.; Larcher, F.; Murillas, R. Risk assessment in skin gene therapy: viral-cellular fusion transcripts generated by proviral transcriptional read-through in keratinocytes transduced with self-inactivating lentiviral vectors. *Gene Ther.* **2011**, *18*, 674–681, doi:10.1038/gt.2011.12.

44. Moiani, A.; Paleari, Y.; Sartori, D.; Mezzadra, R.; Miccio, A.; Cattoglio, C.; Cocchiarella, F.; Lidonnici, M.R.; Ferrari, G.; Mavilio, F. Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. *J. Clin. Invest.* **2012**, *122*, 1653–1666, doi:10.1172/JCI61852.

45. Micklethwaite, K.P.; Gowrishankar, K.; Gloss, B.S.; Li, Z.; Street, J.A.; Moezzi, L.; Mach, M.A.; Sutrave, G.; Clancy, L.E.; Bishop, D.C.; et al. Investigation of product-derived lymphoma following infusion of piggyBac-modified CD19 chimeric antigen receptor T cells. *Blood* **2021**, *138*, 1391–1405, doi:10.1182/blood.2021010858.

References

46. Bishop, D.C.; Clancy, L.E.; Simms, R.; Burgess, J.; Mathew, G.; Moezzi, L.; Street, J.A.; Sutrave, G.; Atkins, E.; McGuire, H.M.; et al. Development of CAR T-cell lymphoma in 2 of 10 patients effectively treated with piggyBac-modified CD19 CAR T cells. *Blood* **2021**, *138*, 1504–1509, doi:10.1182/blood.2021010813.

47. Newrzela, S.; Cornils, K.; Li, Z.; Baum, C.; Brugman, M.H.; Hartmann, M.; Meyer, J.; Hartmann, S.; Hansmann, M.-L.; Fehse, B.; et al. Resistance of mature T cells to oncogene transformation. *Blood* **2008**, *112*, 2278–2286, doi:10.1182/blood-2007-12-128751.

48. Stadtmauer, E.A.; Fraietta, J.A.; Davis, M.M.; Cohen, A.D.; Weber, K.L.; Lancaster, E.; Mangan, P.A.; Kulikovskaya, I.; Gupta, M.; Chen, F.; et al. CRISPR-engineered T cells in patients with refractory cancer. *Science* **2020**, *367*, doi:10.1126/science.aba7365.

49. Hemmi, H.; Takeuchi, O.; Kawai, T.; Kaisho, T.; Sato, S.; Sanjo, H.; Matsumoto, M.; Hoshino, K.; Wagner, H.; Takeda, K.; et al. A Toll-like receptor recognizes bacterial DNA. *Nature* **2000**, *408*, 740–745, doi:10.1038/35047123.

50. Marie, C.; Vandermeulen, G.; Quiviger, M.; Richard, M.; Préat, V.; Scherman, D. pFARs, plasmids free of antibiotic resistance markers, display high-level transgene expression in muscle, skin and tumour cells. *J. Gene Med.* **2010**, *12*, 323–332, doi:10.1002/jgm.1441.

51. Darquet, A.M.; Cameron, B.; Wils, P.; Scherman, D.; Crouzet, J. A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. *Gene Ther.* **1997**, *4*, 1341–1349, doi:10.1038/sj.gt.3300540.

52. Walters, A.A.; Kinnear, E.; Shattock, R.J.; McDonald, J.U.; Caproni, L.J.; Porter, N.; Tregoning, J.S. Comparative analysis of enzymatically produced novel linear DNA constructs with plasmids for use as DNA vaccines. *Gene Ther.* **2014**, *21*, 645–652, doi:10.1038/gt.2014.37.

53. Rostovskaya, M.; Fu, J.; Obst, M.; Baer, I.; Weidlich, S.; Wang, H.; Smith, A.J.H.; Anastassiadis, K.; Stewart, A.F. Transposon-mediated BAC transgenesis in human ES cells. *Nucleic Acids Res.* **2012**, *40*, e150, doi:10.1093/nar/gks643.

54. Amberger, M.; Ivics, Z. Latest Advances for the Sleeping Beauty Transposon System: 23 Years of Insomnia but Prettier than Ever: Refinement and Recent Innovations of the Sleeping Beauty Transposon System Enabling Novel, Nonviral Genetic Engineering Applications. *Bioessays* **2020**, *42*, e2000136, doi:10.1002/bies.202000136.

55. Potter, K.N.; Faulkner, P.; MacKinnon, E.A. Strain selection during serial passage of Trichoplusia in nuclear polyhedrosis virus. *J. Virol.* **1976**, *18*, 1040–1050, doi:10.1128/JVI.18.3.1040-1050.1976.

56. Fraser, M.J.; Smith, G.E.; Summers, M.D. Acquisition of Host Cell DNA Sequences by Baculoviruses: Relationship Between Host DNA Insertions and FP Mutants of Autographa californica and Galleria mellonella Nuclear Polyhedrosis Viruses. *J. Virol.* **1983**, *47*, 287–300, doi:10.1128/JVI.47.2.287-300.1983.

57. Cary, L.C.; Goebel, M.; Corsaro, B.G.; Wang, H.G.; Rosen, E.; Fraser, M.J. Transposon mutagenesis of baculoviruses: analysis of Trichoplusia ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* **1989**, *172*, 156–169, doi:10.1016/0042-6822(89)90117-7.

58. Elick, T.A.; Bauser, C.A.; Fraser, M.J. Excision of the piggyBac transposable element in vitro is a precise event that is enhanced by the expression of its encoded transposase. *Genetica* **1996**, *98*, 33–41, doi:10.1007/BF00120216.

References

59. Bouallègue, M.; Rouault, J.-D.; Hua-Van, A.; Makni, M.; Capy, P. Molecular Evolution of piggyBac Superfamily: From Selfishness to Domestication. *Genome Biol. Evol.* **2017**, *9*, 323–339, doi:10.1093/gbe/evw292.

60. Morellet, N.; Li, X.; Wieninger, S.A.; Taylor, J.L.; Bischerour, J.; Moriau, S.; Lescop, E.; Bardiaux, B.; Mathy, N.; Assrir, N.; et al. Sequence-specific DNA binding activity of the cross-brace zinc finger motif of the piggyBac transposase. *Nucleic Acids Res.* **2018**, *46*, 2660–2677, doi:10.1093/nar/gky044.

61. Chen, Q.; Luo, W.; Veach, R.A.; Hickman, A.B.; Wilson, M.H.; Dyda, F. Structural basis of seamless excision and specific targeting by piggyBac transposase. *Nat. Commun.* **2020**, *11*, 3446, doi:10.1038/s41467-020-17128-1.

62. Keith, J.H.; Fraser, T.S.; Fraser, M.J. Analysis of the piggyBac transposase reveals a functional nuclear targeting signal in the 94 c-terminal residues. *BMC Mol. Biol.* **2008**, *9*, 72, doi:10.1186/1471-2199-9-72.

63. Mitra, R.; Fain-Thornton, J.; Craig, N.L. piggyBac can bypass DNA synthesis during cut and paste transposition. *EMBO J.* **2008**, *27*, 1097–1109, doi:10.1038/emboj.2008.41.

64. Coen, E.S.; Carpenter, R.; Martin, C. Transposable elements generate novel spatial patterns of gene expression in Antirrhinum majus. *Cell* **1986**, *47*, 285–296, doi:10.1016/0092-8674(86)90451-4.

65. Weil, C.F.; Kunze, R. Transposition of maize Ac/Ds transposable elements in the yeast Saccharomyces cerevisiae. *Nat. Genet.* **2000**, *26*, 187–190, doi:10.1038/82827.

66. Kennedy, A.K.; Guhathakurta, A.; Kleckner, N.; Haniford, D.B. Tn10 transposition via a DNA hairpin intermediate. *Cell* **1998**, *95*, 125–134, doi:10.1016/s0092-8674(00)81788-2.

67. Bhasin, A.; Goryshin, I.Y.; Reznikoff, W.S. Hairpin formation in Tn5 transposition. *J. Biol. Chem.* **1999**, *274*, 37021–37029, doi:10.1074/jbc.274.52.37021.

68. Li, M.A.; Pettitt, S.J.; Eckert, S.; Ning, Z.; Rice, S.; Cadiñanos, J.; Yusa, K.; Conte, N.; Bradley, A. The piggyBac transposon displays local and distant reintegration preferences and can cause mutations at noncanonical integration sites. *Mol. Cell. Biol.* **2013**, *33*, 1317–1330, doi:10.1128/MCB.00670-12.

69. Johnson, E.T.; Dowd, P.F. A non-autonomous insect piggyBac transposable element is mobile in tobacco. *Mol. Genet. Genomics* **2014**, *289*, 895–902, doi:10.1007/s00438-014-0860-2.

70. Ding, S.; Wu, X.; Li, G.; Han, M.; Zhuang, Y.; Xu, T. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **2005**, *122*, 473–483, doi:10.1016/j.cell.2005.07.013.

71. Li, X.; Harrell, R.A.; Handler, A.M.; Beam, T.; Hennessy, K.; Fraser, M.J. piggyBac internal sequences are necessary for efficient transformation of target genomes. *Insect Mol. Biol.* **2005**, *14*, 17–30, doi:10.1111/j.1365-2583.2004.00525.x.

72. Lacoste, A.; Berenshteyn, F.; Brivanlou, A.H. An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. *Cell Stem Cell* **2009**, *5*, 332–342, doi:10.1016/j.stem.2009.07.011.

73. Cadiñanos, J.; Bradley, A. Generation of an inducible and optimized piggyBac transposon system. *Nucleic Acids Res.* **2007**, *35*, e87, doi:10.1093/nar/gkm446.

74. Yusa, K.; Zhou, L.; Li, M.A.; Bradley, A.; Craig, N.L. A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1531–1536, doi:10.1073/pnas.1008322108.

References

75. Wen, W.; Song, S.; Han, Y.; Chen, H.; Liu, X.; Qian, Q. An efficient Screening System in Yeast to Select a Hyperactive piggyBac Transposase for Mammalian Applications. *Int. J. Mol. Sci.* **2020**, *21*, doi:10.3390/ijms21093064.

76. Li, M.A.; Turner, D.J.; Ning, Z.; Yusa, K.; Liang, Q.; Eckert, S.; Rad, L.; Fitzgerald, T.W.; Craig, N.L.; Bradley, A. Mobilization of giant piggyBac transposons in the mouse genome. *Nucleic Acids Res.* **2011**, *39*, e148, doi:10.1093/nar/gkr764.

77. Chew, S.K.; Rad, R.; Futreal, P.A.; Bradley, A.; Liu, P. Genetic screens using the piggyBac transposon. *Methods* **2011**, *53*, 366–371, doi:10.1016/j.ymeth.2010.12.022.

78. Kawakami, K.; Largaespada, D.A.; Ivics, Z. Transposons As Tools for Functional Genomics in Vertebrate Models. *Trends Genet.* **2017**, *33*, 784–801, doi:10.1016/j.tig.2017.07.006.

79. Potter, C.J.; Luo, L. Splinkerette PCR for mapping transposable elements in Drosophila. *PLoS One* **2010**, *5*, e10168, doi:10.1371/journal.pone.0010168.

80. Shen, D.; Xue, S.; Chan, S.; Sang, Y.; Wang, S.; Wang, Y.; Chen, C.; Gao, B.; Mueller, F.; Song, C. Enhancer Trapping and Annotation in Zebrafish Mediated with Sleeping Beauty, piggyBac and Tol2 Transposons. *Genes (Basel)* **2018**, *9*, doi:10.3390/genes9120630.

81. Li, T.; Shuai, L.; Mao, J.; Wang, X.; Wang, M.; Zhang, X.; Wang, L.; Li, Y.; Li, W.; Zhou, Q. Efficient Production of Fluorescent Transgenic Rats using the piggyBac Transposon. *Sci. Rep.* **2016**, *6*, 33225, doi:10.1038/srep33225.

82. Li, Z.; Zeng, F.; Meng, F.; Xu, Z.; Zhang, X.; Huang, X.; Tang, F.; Gao, W.; Shi, J.; He, X.; et al. Generation of transgenic pigs by cytoplasmic injection of piggyBac transposase-based pmGENIE-3 plasmids. *Biol. Reprod.* **2014**, *90*, 93, doi:10.1095/biolreprod.113.116905.

83. Yum, S.-Y.; Lee, S.-J.; Kim, H.-M.; Choi, W.-J.; Park, J.-H.; Lee, W.-W.; Kim, H.-S.; Kim, H.-J.; Bae, S.-H.; Lee, J.-H.; et al. Efficient generation of transgenic cattle using the DNA transposon and their analysis by next-generation sequencing. *Sci. Rep.* **2016**, *6*, 27185, doi:10.1038/srep27185.

84. Kumar, D.; Anand, T.; Talluri, T.R.; Kues, W.A. Potential of transposon-mediated cellular reprogramming towards cell-based therapies. *World J. Stem Cells* **2020**, *12*, 527–544, doi:10.4252/wjsc.v12.i7.527.

85. Woltjen, K.; Michael, I.P.; Mohseni, P.; Desai, R.; Mileikovsky, M.; Hämäläinen, R.; Cowling, R.; Wang, W.; Liu, P.; Gertsenstein, M.; et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **2009**, *458*, 766–770, doi:10.1038/nature07863.

86. Yusa, K.; Rad, R.; Takeda, J.; Bradley, A. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nat. Methods* **2009**, *6*, 363–369, doi:10.1038/nmeth.1323.

87. Li, X.; Burnight, E.R.; Cooney, A.L.; Malani, N.; Brady, T.; Sander, J.D.; Staber, J.; Wheelan, S.J.; Joung, J.K.; McCray, P.B.; et al. piggyBac transposase tools for genome engineering. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E2279-87, doi:10.1073/pnas.1305987110.

88. Wang, W.; Lin, C.; Lu, D.; Ning, Z.; Cox, T.; Melvin, D.; Wang, X.; Bradley, A.; Liu, P. Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9290–9295, doi:10.1073/pnas.0801017105.

89. Woodard, L.E.; Wilson, M.H. piggyBac-ing models and new therapeutic strategies. *Trends Biotechnol.* **2015**, *33*, 525–533, doi:10.1016/j.tibtech.2015.06.009.

References

90. Iyer, P.S.; Mavoungou, L.O.; Ronzoni, F.; Zemla, J.; Schmid-Siegert, E.; Antonini, S.; Neff, L.A.; Dorchies, O.M.; Jaconi, M.; Lekka, M.; et al. Autologous Cell Therapy Approach for Duchenne Muscular Dystrophy using PiggyBac Transposons and Mesoangioblasts. *Mol. Ther.* **2018**, *26*, 1093–1108, doi:10.1016/j.ymthe.2018.01.021.

91. Cooney, A.L.; Singh, B.K.; Loza, L.M.; Thornell, I.M.; Hippee, C.E.; Powers, L.S.; Ostedgaard, L.S.; Meyerholz, D.K.; Wohlford-Lenane, C.; Stoltz, D.A.; et al. Widespread airway distribution and short-term phenotypic correction of cystic fibrosis pigs following aerosol delivery of piggyBac/adenovirus. *Nucleic Acids Res.* **2018**, *46*, 9591–9600, doi:10.1093/nar/gky773.

92. Matsui, H.; Fujimoto, N.; Sasakawa, N.; Ohinata, Y.; Shima, M.; Yamanaka, S.; Sugimoto, M.; Hotta, A. Delivery of full-length factor VIII using a piggyBac transposon vector to correct a mouse model of hemophilia A. *PLoS One* **2014**, *9*, e104957, doi:10.1371/journal.pone.0104957.

93. Cunningham, S.C.; Siew, S.M.; Hallwirth, C.V.; Bolitho, C.; Sasaki, N.; Garg, G.; Michael, I.P.; Hetherington, N.A.; Carpenter, K.; Alencastro, G. de; et al. Modeling correction of severe urea cycle defects in the growing murine liver using a hybrid recombinant adeno-associated virus/piggyBac transposase gene delivery system. *Hepatology* **2015**, *62*, 417–428, doi:10.1002/hep.27842.

94. Siew, S.M.; Cunningham, S.C.; Zhu, E.; Tay, S.S.; Venuti, E.; Bolitho, C.; Alexander, I.E. Prevention of Cholestatic Liver Disease and Reduced Tumorigenicity in a Murine Model of PFIC Type 3 Using Hybrid AAV-piggyBac Gene Therapy. *Hepatology* **2019**, *70*, 2047–2061, doi:10.1002/hep.30773.

95. Albinger, N.; Hartmann, J.; Ullrich, E. Current status and perspective of CAR-T and CAR-NK cell therapy trials in Germany. *Gene Ther.* **2021**, *28*, 513–527, doi:10.1038/s41434-021-00246-w.

96. Hernandez, I.; Prasad, V.; Gellad, W.F. Accounting for All Costs in the Total Cost of Chimeric Antigen Receptor T-Cell Immunotherapy-Reply. *JAMA Oncol.* **2018**, *4*, 1785–1786, doi:10.1001/jamaoncol.2018.4657.

97. Tanaka, K.; Kato, I.; Tanaka, M.; Morita, D.; Matsuda, K.; Takahashi, Y.; Nakahata, T.; Umeda, K.; Hiramatsu, H.; Adachi, S.; et al. Direct Delivery of piggyBac CD19 CAR T Cells Has Potent Anti-tumor Activity against ALL Cells in CNS in a Xenograft Mouse Model. *Mol. Ther. Oncolytics* **2020**, *18*, 37–46, doi:10.1016/j.omto.2020.05.013.

98. Štach, M.; Ptáčková, P.; Mucha, M.; Musil, J.; Klener, P.; Otáhal, P. Inducible secretion of IL-21 augments anti-tumor activity of piggyBac-manufactured chimeric antigen receptor T cells. *Cytotherapy* **2020**, *22*, 744–754, doi:10.1016/j.jcyt.2020.08.005.

99. Morita, D.; Nishio, N.; Saito, S.; Tanaka, M.; Kawashima, N.; Okuno, Y.; Suzuki, S.; Matsuda, K.; Maeda, Y.; Wilson, M.H.; et al. Enhanced Expression of Anti-CD19 Chimeric Antigen Receptor in piggyBac Transposon-Engineered T Cells. *Mol. Ther. Methods Clin. Dev.* **2018**, *8*, 131–140, doi:10.1016/j.omtm.2017.12.003.

100. Bishop, D.C.; Xu, N.; Tse, B.; O'Brien, T.A.; Gottlieb, D.J.; Dolnikov, A.; Micklethwaite, K.P. PiggyBac-Engineered T Cells Expressing CD19-Specific CARs that Lack IgG1 Fc Spacers Have Potent Activity against B-ALL Xenografts. *Mol. Ther.* **2018**, *26*, 1883–1895, doi:10.1016/j.ymthe.2018.05.007.

101. Hasegawa, A.; Saito, S.; Narimatsu, S.; Nakano, S.; Nagai, M.; Ohnota, H.; Inada, Y.; Morokawa, H.; Nakashima, I.; Morita, D.; et al. Mutated GM-CSF-based CAR-T cells targeting CD116/CD131 complexes exhibit enhanced anti-tumor effects against acute myeloid leukaemia. *Clin. Transl. Immunology* **2021**, *10*, e1282, doi:10.1002/cti2.1282.

<u>*References*</u>

102. Morokawa, H.; Yagyu, S.; Hasegawa, A.; Tanaka, M.; Saito, S.; Mochizuki, H.; Sakamoto, K.; Shimoi, A.; Nakazawa, Y. Autologous non-human primate model for safety assessment of piggyBac transposon-mediated chimeric antigen receptor T cells on granulocyte-macrophage colony-stimulating factor receptor. *Clin. Transl. Immunology* **2020**, *9*, e1207, doi:10.1002/cti2.1207.

103. Li, H.; Huang, Y.; Jiang, D.-Q.; Cui, L.-Z.; He, Z.; Wang, C.; Zhang, Z.-W.; Zhu, H.-L.; Ding, Y.-M.; Li, L.-F.; et al. Antitumor activity of EGFR-specific CAR T cells against non-small-cell lung cancer cells in vitro and in mice. *Cell Death Dis.* **2018**, *9*, 177, doi:10.1038/s41419-017-0238-6.

104. Wang, P.; Qin, W.; Liu, T.; Jiang, D.; Cui, L.; Liu, X.; Fang, Y.; Tang, X.; Jin, H.; Qian, Q. PiggyBac-engineered T cells expressing a glypican-3-specific chimeric antigen receptor show potent activities against hepatocellular carcinoma. *Immunobiology* **2020**, *225*, 151850, doi:10.1016/j.imbio.2019.09.009.

105. Ma, Y.; Chen, Y.; Yan, L.; Cao, H.-X.; Han, S.-Y.; Cui, J.-J.; Wen, J.G.; Zheng, Y. EGFRvIII-specific CAR-T cells produced by piggyBac transposon exhibit efficient growth suppression against hepatocellular carcinoma. *Int. J. Med. Sci.* **2020**, *17*, 1406–1414, doi:10.7150/ijms.45603.

106. Zhang, Z.; Jiang, D.; Yang, H.; He, Z.; Liu, X.; Qin, W.; Li, L.; Wang, C.; Li, Y.; Li, H.; et al. Modified CAR T cells targeting membrane-proximal epitope of mesothelin enhances the antitumor function against large solid tumor. *Cell Death Dis.* **2019**, *10*, 476, doi:10.1038/s41419-019-1711-1.

107. Zhang, Y.; Zhang, Z.; Ding, Y.; Fang, Y.; Wang, P.; Chu, W.; Jin, Z.; Yang, X.; Wang, J.; Lou, J.; et al. Phase I clinical trial of EGFR-specific CAR-T cells generated by the piggyBac transposon system in advanced relapsed/refractory non-small cell lung cancer patients. *J. Cancer Res. Clin. Oncol.* **2021**, *147*, 3725–3734, doi:10.1007/s00432-021-03613-7.

108. Costello, C.L.; Cohen, A.D.; Patel, K.K.; Ali, S.S.; Berdeja, J.G.; Shah, N.; Ganguly, S.; Kocoglu, M.H.; Abedi, M.; Ostertag, E.M.; et al. Phase 1/2 Study of the Safety and Response of P-BCMA-101 CAR-T Cells in Patients with Relapsed/Refractory (r/r) Multiple Myeloma (MM) (PRIME) with Novel Therapeutic Strategies. *Blood* **2020**, *136*, 29–30, doi:10.1182/blood-2020-142695.

109. Nishio, N.; Hanajiri, R.; Ishikawa, Y.; Murata, M.; Taniguchi, R.; Hamada, M.; Nishikawa, E.; Kawashima, N.; Narita, A.; Muramatsu, H.; et al. A Phase I Study of CD19 Chimeric Antigen Receptor-T Cells Generated By the PiggyBac Transposon Vector for Acute Lymphoblastic Leukemia. *Blood* **2021**, *138*, 3831, doi:10.1182/blood-2021-150469.

110. Afzali, B.; Grönholm, J.; Vandrovcova, J.; O'Brien, C.; Sun, H.-W.; Vanderleyden, I.; Davis, F.P.; Khoder, A.; Zhang, Y.; Hegazy, A.N.; et al. BACH2 immunodeficiency illustrates an association between super-enhancers and haploinsufficiency. *Nat. Immunol.* **2017**, *18*, 813–823, doi:10.1038/ni.3753.

111. Gogol-Döring, A.; Ammar, I.; Gupta, S.; Bunse, M.; Miskey, C.; Chen, W.; Uckert, W.; Schulz, T.F.; Izsvák, Z.; Ivics, Z. Genome-wide Profiling Reveals Remarkable Parallels Between Insertion Site Selection Properties of the MLV Retrovirus and the piggyBac Transposon in Primary Human CD4(+) T Cells. *Mol. Ther.* **2016**, *24*, 592–606, doi:10.1038/mt.2016.11.

112. Huang, X.; Guo, H.; Tammana, S.; Jung, Y.-C.; Mellgren, E.; Bassi, P.; Cao, Q.; Tu, Z.J.; Kim, Y.C.; Ekker, S.C.; et al. Gene transfer efficiency and genome-wide integration

profiling of Sleeping Beauty, Tol2, and piggyBac transposons in human primary T cells. *Mol. Ther.* **2010**, *18*, 1803–1813, doi:10.1038/mt.2010.141.

113. Gupta, S.S.; Maetzig, T.; Maertens, G.N.; Sharif, A.; Rothe, M.; Weidner-Glunde, M.; Galla, M.; Schambach, A.; Cherepanov, P.; Schulz, T.F. Bromo- and extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. *J. Virol.* **2013**, *87*, 12721–12736, doi:10.1128/JVI.01942-13.

114. Rijck, J. de; Kogel, C. de; Demeulemeester, J.; Vets, S.; El Ashkar, S.; Malani, N.; Bushman, F.D.; Landuyt, B.; Husson, S.J.; Busschots, K.; et al. The BET family of proteins targets moloney murine leukemia virus integration near transcription start sites. *Cell Rep.* **2013**, *5*, 886–894, doi:10.1016/j.celrep.2013.09.040.

115. Sharma, A.; Larue, R.C.; Plumb, M.R.; Malani, N.; Male, F.; Slaughter, A.; Kessl, J.J.; Shkriabai, N.; Coward, E.; Aiyer, S.S.; et al. BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 12036–12041, doi:10.1073/pnas.1307157110.

116. Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W.B.; Fedorov, O.; Morse, E.M.; Keates, T.; Hickman, T.T.; Felletar, I.; et al. Selective inhibition of BET bromodomains. *Nature* **2010**, *468*, 1067–1073, doi:10.1038/nature09504.

117. Crowe, B.L.; Larue, R.C.; Yuan, C.; Hess, S.; Kvaratskhelia, M.; Foster, M.P. Structure of the Brd4 ET domain bound to a C-terminal motif from γ-retroviral integrases reveals a conserved mechanism of interaction. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2086–2091, doi:10.1073/pnas.1516813113.

118. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066, doi:10.1093/nar/gkf436.

119. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2019**, *20*, 1160–1166, doi:10.1093/bib/bbx108.

120. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320-4, doi:10.1093/nar/gku316.

121. Burnight, E.R.; Staber, J.M.; Korsakov, P.; Li, X.; Brett, B.T.; Scheetz, T.E.; Craig, N.L.; McCray, P.B. A Hyperactive Transposase Promotes Persistent Gene Transfer of a piggyBac DNA Transposon. *Mol. Ther. Nucleic Acids* **2012**, *1*, e50, doi:10.1038/mtna.2012.12.

122. Mitra, R.; Li, X.; Kapusta, A.; Mayhew, D.; Mitra, R.D.; Feschotte, C.; Craig, N.L. Functional characterization of piggyBat from the bat Myotis lucifugus unveils an active mammalian DNA transposon. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 234–239, doi:10.1073/pnas.1217548110.

123. Morgan, A.A.; Rubenstein, E. Proline: the distribution, frequency, positioning, and common functional roles of proline and polyproline sequences in the human proteome. *PLoS One* **2013**, *8*, e53785, doi:10.1371/journal.pone.0053785.

124. Denis, G.V.; McComb, M.E.; Faller, D.V.; Sinha, A.; Romesser, P.B.; Costello, C.E. Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines. *J. Proteome Res.* **2006**, *5*, 502–511, doi:10.1021/pr050430u.

125. Denis, G.V.; Vaziri, C.; Guo, N.; Faller, D.V. RING3 kinase transactivates promoters of cell cycle regulatory genes through E2F. *Cell Growth Differ.* **2000**, *11*, 417–424.

References

126. Sinha, A.; Faller, D.V.; Denis, G.V. Bromodomain analysis of Brd2-dependent transcriptional activation of cyclin A. *Biochem. J.* **2005**, *387*, 257–269, doi:10.1042/BJ20041793.

127. Roth, S.L.; Malani, N.; Bushman, F.D. Gammaretroviral integration into nucleosomal target DNA in vivo. *J. Virol.* **2011**, *85*, 7393–7401, doi:10.1128/JVI.00635-11.

128. Maskell, D.P.; Renault, L.; Serrao, E.; Lesbats, P.; Matadeen, R.; Hare, S.; Lindemann, D.; Engelman, A.N.; Costa, A.; Cherepanov, P. Structural basis for retroviral integration into nucleosomes. *Nature* **2015**, *523*, 366–369, doi:10.1038/nature14495.

129. Struhl, K.; Segal, E. Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* **2013**, *20*, 267–273, doi:10.1038/nsmb.2506.

130. Wilson, M.H.; Coates, C.J.; George, A.L. PiggyBac transposon-mediated gene transfer in human cells. *Mol. Ther.* **2007**, *15*, 139–145, doi:10.1038/sj.mt.6300028.

131. Liang, Q.; Kong, J.; Stalker, J.; Bradley, A. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. *Genesis* **2009**, *47*, 404–408, doi:10.1002/dvg.20508.

132. Jong, J. de; Akhtar, W.; Badhai, J.; Rust, A.G.; Rad, R.; Hilkens, J.; Berns, A.; van Lohuizen, M.; Wessels, L.F.A.; Ridder, J. de. Chromatin landscapes of retroviral and transposon integration profiles. *PLoS Genet.* **2014**, *10*, e1004250, doi:10.1371/journal.pgen.1004250.

133. Wang, H.; Mayhew, D.; Chen, X.; Johnston, M.; Mitra, R.D. "Calling cards" for DNA-binding proteins in mammalian cells. *Genetics* **2012**, *190*, 941–949, doi:10.1534/genetics.111.137315.

134. Wilson, M.H.; Gottschalk, S. Expect the unexpected: piggyBac and lymphoma. *Blood* **2021**, *138*, 1379–1380, doi:10.1182/blood.2021012349.

135. Chen, M.S.; Walker, J.; Prusoff, W.H. Kinetic studies of herpes simplex virus type 1-encoded thymidine and thymidylate kinase, a multifunctional enzyme. *J. Biol. Chem.* **1979**, *254*, 10747–10753.

136. Balzarini, J.; Bohman, C.; Clercq, E. de. Differential mechanism of cytostatic effect of (E)-5-(2-bromovinyl)-2'-deoxyuridine, 9-(1,3-dihydroxy-2-propoxymethyl)guanine, and other antiherpetic drugs on tumor cells transfected by the thymidine kinase gene of herpes simplex virus type 1 or type 2. *J. Biol. Chem.* **1993**, *268*, 6332–6337.

137. Reardon, J.E. Herpes simplex virus type 1 and human DNA polymerase interactions with 2'-deoxyguanosine 5'-triphosphate analogues. Kinetics of incorporation into DNA and induction of inhibition. *J. Biol. Chem.* **1989**, *264*, 19039–19044.

138. Wei, S.J.; Chao, Y.; Hung, Y.M.; Lin, W.C.; Yang, D.M.; Shih, Y.L.; Ch'ang, L.Y.; Whang-Peng, J.; Yang, W.K. S- and G2-phase cell cycle arrests and apoptosis induced by ganciclovir in murine melanoma cells transduced with herpes simplex virus thymidine kinase. *Exp. Cell Res.* **1998**, *241*, 66–75, doi:10.1006/excr.1998.4005.

139. Beltinger, C.; Fulda, S.; Kammertoens, T.; Meyer, E.; Uckert, W.; Debatin, K.M. Herpes simplex virus thymidine kinase/ganciclovir-induced apoptosis involves ligand-independent death receptor aggregation and activation of caspases. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 8699–8704, doi:10.1073/pnas.96.15.8699.

140. Bonini, C.; Ferrari, G.; Verzeletti, S.; Servida, P.; Zappone, E.; Ruggieri, L.; Ponzoni, M.; Rossini, S.; Mavilio, F.; Traversari, C.; et al. HSV-TK gene transfer into donor lymphocytes for control of allogeneic graft-versus-leukemia. *Science* **1997**, *276*, 1719–1724, doi:10.1126/science.276.5319.1719.

141. Riddell, S.R.; Elliott, M.; Lewinsohn, D.A.; Gilbert, M.J.; Wilson, L.; Manley, S.A.; Lupton, S.D.; Overell, R.W.; Reynolds, T.C.; Corey, L.; et al. T-cell mediated rejection

of gene-modified HIV-specific cytotoxic T lymphocytes in HIV-infected patients. *Nat. Med.* **1996**, *2*, 216–223, doi:10.1038/nm0296-216.

142. Zhao, J. Cancer stem cells and chemoresistance: The smartest survives the raid. *Pharmacol. Ther.* **2016**, *160*, 145–158, doi:10.1016/j.pharmthera.2016.02.008.

143. Straathof, K.C.; Pulè, M.A.; Yotnda, P.; Dotti, G.; Vanin, E.F.; Brenner, M.K.; Heslop, H.E.; Spencer, D.M.; Rooney, C.M. An inducible caspase 9 safety switch for T-cell therapy. *Blood* **2005**, *105*, 4247–4254, doi:10.1182/blood-2004-11-4564.

144. Clackson, T.; Yang, W.; Rozamus, L.W.; Hatada, M.; Amara, J.F.; Rollins, C.T.; Stevenson, L.F.; Magari, S.R.; Wood, S.A.; Courage, N.L.; et al. Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 10437–10442, doi:10.1073/pnas.95.18.10437.

145. Iuliucci, J.D.; Oliver, S.D.; Morley, S.; Ward, C.; Ward, J.; Dalgarno, D.; Clackson, T.; Berger, H.J. Intravenous safety and pharmacokinetics of a novel dimerizer drug, AP1903, in healthy volunteers. *J. Clin. Pharmacol.* **2001**, *41*, 870–879, doi:10.1177/00912700122010771.

146. McGavin, J.K.; Goa, K.L. Ganciclovir: an update of its use in the prevention of cytomegalovirus infection and disease in transplant recipients. *Drugs* **2001**, *61*, 1153–1183, doi:10.2165/00003495-200161080-00016.

147. Marin, V.; Cribioli, E.; Philip, B.; Tettamanti, S.; Pizzitola, I.; Biondi, A.; Biagi, E.; Pule, M. Comparison of different suicide-gene strategies for the safety improvement of genetically manipulated T cells. *Hum. Gene Ther. Methods* **2012**, *23*, 376–386, doi:10.1089/hgtb.2012.050.

148. Di Stasi, A.; Tey, S.-K.; Dotti, G.; Fujita, Y.; Kennedy-Nasser, A.; Martinez, C.; Straathof, K.; Liu, E.; Durett, A.G.; Grilley, B.; et al. Inducible apoptosis as a safety switch for adoptive cell therapy. *N. Engl. J. Med.* **2011**, *365*, 1673–1683, doi:10.1056/NEJMoa1106152.

149. Gargett, T.; Brown, M.P. The inducible caspase-9 suicide gene system as a "safety switch" to limit on-target, off-tumor toxicities of chimeric antigen receptor T cells. *Front. Pharmacol.* **2014**, *5*, 235, doi:10.3389/fphar.2014.00235.

150. Griffioen, M.; van Egmond, E.H.M.; Kester, M.G.D.; Willemze, R.; Falkenburg, J.H.F.; Heemskerk, M.H.M. Retroviral transfer of human CD20 as a suicide gene for adoptive T-cell therapy. *Haematologica* **2009**, *94*, 1316–1320, doi:10.3324/haematol.2008.001677.

151. Wang, X.; Chang, W.-C.; Wong, C.W.; Colcher, D.; Sherman, M.; Ostberg, J.R.; Forman, S.J.; Riddell, S.R.; Jensen, M.C. A transgene-encoded cell surface polypeptide for selection, in vivo tracking, and ablation of engineered cells. *Blood* **2011**, *118*, 1255–1263, doi:10.1182/blood-2011-02-337360.

152. Philip, B.; Kokalaki, E.; Mekkaoui, L.; Thomas, S.; Straathof, K.; Flutter, B.; Marin, V.; Marafioti, T.; Chakraverty, R.; Linch, D.; et al. A highly compact epitope-based marker/suicide gene for easier and safer T-cell therapy. *Blood* **2014**, *124*, 1277–1287, doi:10.1182/blood-2014-01-545020.

153. Ericson, S.G.; Guyre, C.A.; Benoit, N.E.; Meehan, K.R.; Mills, L.E.; Fanger, M.W. Anti-body-dependent cellular cytotoxicity (ADCC) function of peripheral blood polymorphonuclear neutrophils (PMN) after autologous bone marrow transplantation (ABMT). *Bone Marrow Transplant.* **1995**, *16*, 787–791.

154. Hansel, T.T.; Kropshofer, H.; Singer, T.; Mitchell, J.A.; George, A.J.T. The safety and side effects of monoclonal antibodies. *Nat. Rev. Drug Discov.* **2010**, *9*, 325–338, doi:10.1038/nrd3003.

References

155. Tabrizi, M.; Bornstein, G.G.; Suria, H. Biodistribution mechanisms of therapeutic monoclonal antibodies in health and disease. *AAPS J.* **2010**, *12*, 33–43, doi:10.1208/s12248-009-9157-5.

156. Garneau, J.E.; Dupuis, M.-È.; Villion, M.; Romero, D.A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadán, A.H.; Moineau, S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **2010**, *468*, 67–71, doi:10.1038/nature09523.

157. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821, doi:10.1126/science.1225829.

158. Samai, P.; Pyenson, N.; Jiang, W.; Goldberg, G.W.; Hatoum-Aslan, A.; Marraffini, L.A. Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* **2015**, *161*, 1164–1174, doi:10.1016/j.cell.2015.04.027.

159. Nuñez, J.K.; Kranzusch, P.J.; Noeske, J.; Wright, A.V.; Davies, C.W.; Doudna, J.A. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* **2014**, *21*, 528–534, doi:10.1038/nsmb.2820.

160. Wang, R.; Preamplume, G.; Terns, M.P.; Terns, R.M.; Li, H. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. *Structure* **2011**, *19*, 257–264, doi:10.1016/j.str.2010.11.014.

161. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602–607, doi:10.1038/nature09886.

162. Heler, R.; Samai, P.; Modell, J.W.; Weiner, C.; Goldberg, G.W.; Bikard, D.; Marraffini, L.A. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* **2015**, *519*, 199–202, doi:10.1038/nature14245.

163. Mojica, F.J.M.; Díez-Villaseñor, C.; García-Martínez, J.; Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading)* **2009**, *155*, 733–740, doi:10.1099/mic.0.023960-0.

164. Sun, W.; Liu, H.; Yin, W.; Qiao, J.; Zhao, X.; Liu, Y. Strategies for Enhancing the Homology-Directed Repair Efficiency of CRISPR-Cas Systems. *CRISPR J.* **2022**, *5*, 7–18, doi:10.1089/crispr.2021.0039.

165. Kotagama, O.W.; Jayasinghe, C.D.; Abeysinghe, T. Era of Genomic Medicine: A Narrative Review on CRISPR Technology as a Potential Therapeutic Tool for Human Diseases. *Biomed Res. Int.* **2019**, *2019*, 1369682, doi:10.1155/2019/1369682.

166. Lim, J.M.; Kim, H.H. Basic Principles and Clinical Applications of CRISPR-Based Genome Editing. *Yonsei Med. J.* **2022**, *63*, 105–113, doi:10.3349/ymj.2022.63.2.105.

167. Schwank, G.; Koo, B.-K.; Sasselli, V.; Dekkers, J.F.; Heo, I.; Demircan, T.; Sasaki, N.; Boymans, S.; Cuppen, E.; van der Ent, C.K.; et al. Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **2013**, *13*, 653–658, doi:10.1016/j.stem.2013.11.002.

168. Firth, A.L.; Menon, T.; Parker, G.S.; Qualls, S.J.; Lewis, B.M.; Ke, E.; Dargitz, C.T.; Wright, R.; Khanna, A.; Gage, F.H.; et al. Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. *Cell Rep.* **2015**, *12*, 1385–1390, doi:10.1016/j.celrep.2015.07.062.

169. Hoban, M.D.; Lumaquin, D.; Kuo, C.Y.; Romero, Z.; Long, J.; Ho, M.; Young, C.S.; Mojadidi, M.; Fitz-Gibbon, S.; Cooper, A.R.; et al. CRISPR/Cas9-Mediated Correction

of the Sickle Mutation in Human CD34+ cells. *Mol. Ther.* **2016**, *24*, 1561–1569, doi:10.1038/mt.2016.148.

170. Hoban, M.D.; Lumaquin, D.; Kuo, C.Y.; Romero, Z.; Long, J.; Ho, M.; Young, C.S.; Mojadidi, M.; Fitz-Gibbon, S.; Cooper, A.R.; et al. CRISPR/Cas9-Mediated Correction of the Sickle Mutation in Human CD34+ cells. *Mol. Ther.* **2016**, *24*, 1561–1569, doi:10.1038/mt.2016.148.

171. Dever, D.P.; Bak, R.O.; Reinisch, A.; Camarena, J.; Washington, G.; Nicolas, C.E.; Pavel-Dinu, M.; Saxena, N.; Wilkens, A.B.; Mantri, S.; et al. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **2016**, *539*, 384–389, doi:10.1038/nature20134.

172. Xu, P.; Tong, Y.; Liu, X.; Wang, T.; Cheng, L.; Wang, B.; Lv, X.; Huang, Y.; Liu, D. Both TALENs and CRISPR/Cas9 directly target the HBB IVS2-654 (C T) mutation in β-thalassemia-derived iPSCs. *Sci. Rep.* **2015**, *5*, 12065, doi:10.1038/srep12065.

173. Niu, X.; He, W.; Song, B.; Ou, Z.; Di Fan; Chen, Y.; Fan, Y.; Sun, X. Combining Single Strand Oligodeoxynucleotides and CRISPR/Cas9 to Correct Gene Mutations in β-Thalassemia-induced Pluripotent Stem Cells. *J. Biol. Chem.* **2016**, *291*, 16576–16585, doi:10.1074/jbc.M116.719237.

174. Yang, Y.; Zhang, X.; Yi, L.; Hou, Z.; Chen, J.; Kou, X.; Zhao, Y.; Wang, H.; Sun, X.-F.; Jiang, C.; et al. Naïve Induced Pluripotent Stem Cells Generated From β-Thalassemia Fibroblasts Allow Efficient Gene Correction With CRISPR/Cas9. *Stem Cells Transl. Med.* **2016**, *5*, 8–19, doi:10.5966/sctm.2015-0157.

175. Xie, F.; Ye, L.; Chang, J.C.; Beyer, A.I.; Wang, J.; Muench, M.O.; Kan, Y.W. Seamless gene correction of β-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* **2014**, *24*, 1526–1533, doi:10.1101/gr.173427.114.

176. Song, B.; Fan, Y.; He, W.; Zhu, D.; Niu, X.; Wang, D.; Ou, Z.; Luo, M.; Sun, X. Improved hematopoietic differentiation efficiency of gene-corrected beta-thalassemia induced pluripotent stem cells by CRISPR/Cas9 system. *Stem Cells Dev.* **2015**, *24*, 1053–1065, doi:10.1089/scd.2014.0347.

177. Shin, J.W.; Kim, K.-H.; Chao, M.J.; Atwal, R.S.; Gillis, T.; MacDonald, M.E.; Gusella, J.F.; Lee, J.-M. Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9. *Hum. Mol. Genet.* **2016**, *25*, 4566–4576, doi:10.1093/hmg/ddw286.

178. Monteys, A.M.; Ebanks, S.A.; Keiser, M.S.; Davidson, B.L. CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo. *Mol. Ther.* **2017**, *25*, 12–23, doi:10.1016/j.ymthe.2016.11.010.

179. Yang, S.; Chang, R.; Yang, H.; Zhao, T.; Hong, Y.; Kong, H.E.; Sun, X.; Qin, Z.; Jin, P.; Li, S.; et al. CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. *J. Clin. Invest.* **2017**, *127*, 2719–2724, doi:10.1172/JCI92087.

180. Ousterout, D.G.; Kabadi, A.M.; Thakore, P.I.; Majoros, W.H.; Reddy, T.E.; Gersbach, C.A. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat. Commun.* **2015**, *6*, 6244, doi:10.1038/ncomms7244.

181. Li, H.L.; Fujimoto, N.; Sasakawa, N.; Shirai, S.; Ohkame, T.; Sakuma, T.; Tanaka, M.; Amano, N.; Watanabe, A.; Sakurai, H.; et al. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports* **2015**, *4*, 143–154, doi:10.1016/j.stemcr.2014.10.013.

References

182. Park, C.-Y.; Kim, D.H.; Son, J.S.; Sung, J.J.; Lee, J.; Bae, S.; Kim, J.-H.; Kim, D.-W.; Kim, J.-S. Functional Correction of Large Factor VIII Gene Chromosomal Inversions in Hemophilia A Patient-Derived iPSCs Using CRISPR-Cas9. *Cell Stem Cell* **2015**, *17*, 213–220, doi:10.1016/j.stem.2015.07.001.

183. Guan, Y.; Ma, Y.; Li, Q.; Sun, Z.; Ma, L.; Wu, L.; Wang, L.; Zeng, L.; Shao, Y.; Chen, Y.; et al. CRISPR/Cas9-mediated somatic correction of a novel coagulator factor IX gene mutation ameliorates hemophilia in mouse. *EMBO Mol. Med.* **2016**, *8*, 477–488, doi:10.15252/emmm.201506039.

184. Gaj, T.; Ojala, D.S.; Ekman, F.K.; Byrne, L.C.; Limsirichai, P.; Schaffer, D.V. In vivo genome editing improves motor function and extends survival in a mouse model of ALS. *Sci. Adv.* **2017**, *3*, eaar3952, doi:10.1126/sciadv.aar3952.

185. Flynn, R.; Grundmann, A.; Renz, P.; Hänseler, W.; James, W.S.; Cowley, S.A.; Moore, M.D. CRISPR-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human iPS cells. *Exp. Hematol.* **2015**, *43*, 838-848.e3, doi:10.1016/j.exphem.2015.06.002.

186. Ravin, S.S. de; Li, L.; Wu, X.; Choi, U.; Allen, C.; Koontz, S.; Lee, J.; Theobald-Whiting, N.; Chu, J.; Garofalo, M.; et al. CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Sci. Transl. Med.* **2017**, *9*, doi:10.1126/scitranslmed.aah3480.

187. Yue, J.; Gou, X.; Li, Y.; Wicksteed, B.; Wu, X. Engineered Epidermal Progenitor Cells Can Correct Diet-Induced Obesity and Diabetes. *Cell Stem Cell* **2017**, *21*, 256-263.e4, doi:10.1016/j.stem.2017.06.016.

188. Ding, Q.; Strong, A.; Patel, K.M.; Ng, S.-L.; Gosis, B.S.; Regan, S.N.; Cowan, C.A.; Rader, D.J.; Musunuru, K. Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ. Res.* **2014**, *115*, 488–492, doi:10.1161/CIRCRESAHA.115.304351.

189. Zhang, Z.; Hou, W.; Chen, S. Updates on CRISPR-based gene editing in HIV-1/AIDS therapy. *Virol. Sin.* **2022**, *37*, 1–10, doi:10.1016/j.virs.2022.01.017.

190. Katti, A.; Diaz, B.J.; Caragine, C.M.; Sanjana, N.E.; Dow, L.E. CRISPR in cancer biology and therapy. *Nat. Rev. Cancer* **2022**, *22*, 259–279, doi:10.1038/s41568-022-00441-w.

191. Mok, T.S.K.; Wu, Y.-L.; Kudaba, I.; Kowalski, D.M.; Cho, B.C.; Turna, H.Z.; Castro, G.; Srimuninnimit, V.; Laktionov, K.K.; Bondarenko, I.; et al. Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial. *Lancet* **2019**, *393*, 1819–1830, doi:10.1016/S0140-6736(18)32409-7.

192. Reck, M.; Rodríguez-Abreu, D.; Robinson, A.G.; Hui, R.; Csőszi, T.; Fülöp, A.; Gottfried, M.; Peled, N.; Tafreshi, A.; Cuffe, S.; et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* **2016**, *375*, 1823–1833, doi:10.1056/NEJMoa1606774.

193. Lu, Y.; Xue, J.; Deng, T.; Zhou, X.; Yu, K.; Deng, L.; Huang, M.; Yi, X.; Liang, M.; Wang, Y.; et al. Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nat. Med.* **2020**, *26*, 732–740, doi:10.1038/s41591-020-0840-5.

194. Fu, Y.; Foden, J.A.; Khayter, C.; Maeder, M.L.; Reyon, D.; Joung, J.K.; Sander, J.D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **2013**, *31*, 822–826, doi:10.1038/nbt.2623.

References

195. Hsu, P.D.; Scott, D.A.; Weinstein, J.A.; Ran, F.A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E.J.; Wu, X.; Shalem, O.; et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **2013**, *31*, 827–832, doi:10.1038/nbt.2647.

196. Naeem, M.; Majeed, S.; Hoque, M.Z.; Ahmad, I. Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. *Cells* **2020**, *9*, doi:10.3390/cells9071608.

197. Cho, S.W.; Kim, S.; Kim, Y.; Kweon, J.; Kim, H.S.; Bae, S.; Kim, J.-S. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **2014**, *24*, 132–141, doi:10.1101/gr.162339.113.

198. Papathanasiou, S.; Markoulaki, S.; Blaine, L.J.; Leibowitz, M.L.; Zhang, C.-Z.; Jaenisch, R.; Pellman, D. Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nat. Commun.* **2021**, *12*, 5855, doi:10.1038/s41467-021-26097-y.

199. Owens, D.D.G.; Caulder, A.; Frontera, V.; Harman, J.R.; Allan, A.J.; Bucakci, A.; Greder, L.; Codner, G.F.; Hublitz, P.; McHugh, P.J.; et al. Microhomologies are prevalent at Cas9-induced larger deletions. *Nucleic Acids Res.* **2019**, *47*, 7402–7417, doi:10.1093/nar/gkz459.

200. Cullot, G.; Boutin, J.; Toutain, J.; Prat, F.; Pennamen, P.; Rooryck, C.; Teichmann, M.; Rousseau, E.; Lamrissi-Garcia, I.; Guyonnet-Duperat, V.; et al. CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* **2019**, *10*, 1136, doi:10.1038/s41467-019-09006-2.

201. Leibowitz, M.L.; Papathanasiou, S.; Doerfler, P.A.; Blaine, L.J.; Sun, L.; Yao, Y.; Zhang, C.-Z.; Weiss, M.J.; Pellman, D. Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat. Genet.* **2021**, *53*, 895–905, doi:10.1038/s41588-021-00838-7.

202. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921, doi:10.1038/35057062.

203. Deininger, P. Alu elements: know the SINEs. *Genome Biol.* **2011**, *12*, 236, doi:10.1186/gb-2011-12-12-236.

204. Baskar, R.; Lee, K.A.; Yeo, R.; Yeoh, K.-W. Cancer and radiation therapy: current advances and future directions. *Int. J. Med. Sci.* **2012**, *9*, 193–199, doi:10.7150/ijms.3635.

205. Zio, D. de; Cianfanelli, V.; Cecconi, F. New insights into the link between DNA damage and apoptosis. *Antioxid. Redox Signal.* **2013**, *19*, 559–571, doi:10.1089/ars.2012.4938.

206. Głów, D.; Maire, C.L.; Schwarze, L.I.; Lamszus, K.; Fehse, B. CRISPR-to-Kill (C2K)-Employing the Bacterial Immune System to Kill Cancer Cells. *Cancers (Basel)* **2021**, *13*, doi:10.3390/cancers13246306.

207. Davis, A.; Morris, K.V.; Shevchenko, G. Hypoxia-directed tumor targeting of CRISPR-Cas9 and HSV-TK suicide gene therapy using lipid nanoparticles. *Mol. Ther. Methods Clin. Dev.* **2022**, *25*, 158–169, doi:10.1016/j.omtm.2022.03.008.

208. Ghosh, S.; Brown, A.M.; Jenkins, C.; Campbell, K. Viral Vector Systems for Gene Therapy: A Comprehensive Literature Review of Progress and Biosafety Challenges. *Applied Biosafety* **2020**, *25*, 7–18, doi:10.1177/1535676019899502.

209. Oakes, B.L.; Nadler, D.C.; Flamholz, A.; Fellmann, C.; Staahl, B.T.; Doudna, J.A.; Savage, D.F. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat. Biotechnol.* **2016**, *34*, 646–651, doi:10.1038/nbt.3528.

References

210. Shiau, A.K.; Barstad, D.; Loria, P.M.; Cheng, L.; Kushner, P.J.; Agard, D.A.; Greene, G.L. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell* **1998**, *95*, 927–937, doi:10.1016/S0092-8674(00)81717-1.

211. Mehta, A.; Merkel, O.M. Immunogenicity of Cas9 Protein. *J. Pharm. Sci.* **2020**, *109*, 62–67, doi:10.1016/j.xphs.2019.10.003.

212. Gossen, M.; Freundlieb, S.; Bender, G.; Müller, G.; Hillen, W.; Bujard, H. Transcriptional activation by tetracyclines in mammalian cells. *Science* **1995**, *268*, 1766–1769, doi:10.1126/science.7792603.

213. Miskey, C.; Amberger, M.; Reiser, M.; Prommersberger, S.; Beckmann, J.; Machwirth, M.; Einsele, H.; Hudecek, M.; Bonig, H.; Ivics, Z. Genomic Analyses of SLAMF7 CAR-T Cells Manufactured by Sleeping Beauty Transposon Gene Transfer for Immunotherapy of Multiple Myeloma. *bioRxiv* **2019**, 675009, doi:10.1101/675009.

214. Grabundzija, I.; Irgang, M.; Mátés, L.; Belay, E.; Matrai, J.; Gogol-Döring, A.; Kawakami, K.; Chen, W.; Ruiz, P.; Chuah, M.K.L.; et al. Comparative Analysis of Transposable Element Vector Systems in Human Cells. *Mol. Ther.* **2010**, *18*, 1200–1209, doi:10.1038/mt.2010.47.

215. Li, X.; Lee, Y.-K.; Jeng, J.-C.; Yen, Y.; Schultz, D.C.; Shih, H.-M.; Ann, D.K. Role for KAP1 Serine 824 Phosphorylation and Sumoylation/Desumoylation Switch in Regulating KAP1-mediated Transcriptional Repression*. *J. Biol. Chem.* **2007**, *282*, 36177–36189, doi:10.1074/jbc.M706912200.

216. Hinds, P.; Pietruska, J. Senescence and tumor suppression. *F1000Res.* **2017**, *6*, 2121, doi:10.12688/f1000research.11671.1.

217. Di d'Adda Fagagna, F. Living on a break: cellular senescence as a DNA-damage response. *Nat. Rev. Cancer* **2008**, *8*, 512–522, doi:10.1038/nrc2440.

218. Coppé, J.-P.; Patil, C.K.; Rodier, F.; Sun, Y.; Muñoz, D.P.; Goldstein, J.; Nelson, P.S.; Desprez, P.-Y.; Campisi, J. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **2008**, *6*, 2853–2868, doi:10.1371/journal.pbio.0060301.

219. Burton, D.G.A.; Stolzing, A. Cellular senescence: Immunosurveillance and future immunotherapy. *Ageing Res. Rev.* **2018**, *43*, 17–25, doi:10.1016/j.arr.2018.02.001.

220. Kale, A.; Sharma, A.; Stolzing, A.; Desprez, P.-Y.; Campisi, J. Role of immune cells in the removal of deleterious senescent cells. *Immun. Ageing* **2020**, *17*, 16, doi:10.1186/s12979-020-00187-9.

221. Basisty, N.; Kale, A.; Jeon, O.H.; Kuehnemann, C.; Payne, T.; Rao, C.; Holtz, A.; Shah, S.; Sharma, V.; Ferrucci, L.; et al. A proteomic atlas of senescence-associated secretomes for aging biomarker development. *PLoS Biol.* **2020**, *18*, e3000599, doi:10.1371/journal.pbio.3000599.

222. Gjerde, J.; Gandini, S.; Guerrieri-Gonzaga, A.; Haugan Moi, L.L.; Aristarco, V.; Mellgren, G.; Decensi, A.; Lien, E.A. Tissue distribution of 4-hydroxy-N-desmethyltamoxifen and tamoxifen-N-oxide. *Breast Cancer Res. Treat.* **2012**, *134*, 693–700, doi:10.1007/s10549-012-2074-9.

223. Cunha, B.A.; Sibley, C.M.; Ristuccia, A.M. Doxycycline. *Ther. Drug Monit.* **1982**, *4*, 115–135, doi:10.1097/00007691-198206000-00001.

224. Way, J.C.; Kleckner, N. Transposition of plasmid-borne Tn10 elements does not exhibit simple length-dependence. *Genetics* **1985**, *111*, 705–713, doi:10.1093/genetics/111.4.705.

References

225. Fischer, S.E.; van Luenen, H.G.; Plasterk, R.H. Cis requirements for transposition of Tc1-like transposons in C. elegans. *Mol. Gen. Genet.* **1999**, *262*, 268–274, doi:10.1007/pl00008641.

226. Lampe, D.J.; Grant, T.E.; Robertson, H.M. Factors affecting transposition of the Himar1 mariner transposon in vitro. *Genetics* **1998**, *149*, 179–187, doi:10.1093/genetics/149.1.179.

227. Izsvák, Z.; Ivics, Z.; Plasterk, R.H. Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J. Mol. Biol.* **2000**, *302*, 93–102, doi:10.1006/jmbi.2000.4047.

228. Darquet, A.M.; Cameron, B.; Wils, P.; Scherman, D.; Crouzet, J. A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. *Gene Ther.* **1997**, *4*, 1341–1349, doi:10.1038/sj.gt.3300540.

229. Monjezi, R.; Miskey, C.; Gogishvili, T.; Schleef, M.; Schmeer, M.; Einsele, H.; Ivics, Z.; Hudecek, M. Enhanced CAR T-cell engineering using non-viral Sleeping Beauty transposition from minicircle vectors. *Leukemia* **2017**, *31*, 186–194, doi:10.1038/leu.2016.180.

230. Holstein, M.; Mesa-Nuñez, C.; Miskey, C.; Almarza, E.; Poletti, V.; Schmeer, M.; Grueso, E.; Ordóñez Flores, J.C.; Kobelt, D.; Walther, W.; et al. Efficient Non-viral Gene Delivery into Human Hematopoietic Stem Cells by Minicircle Sleeping Beauty Transposon Vectors. *Mol. Ther.* **2018**, *26*, 1137–1153, doi:10.1016/j.ymthe.2018.01.012.

231. Prommersberger, S.; Reiser, M.; Beckmann, J.; Danhof, S.; Amberger, M.; Quade-Lyssy, P.; Einsele, H.; Hudecek, M.; Bonig, H.; Ivics, Z. CARAMBA: a first-in-human clinical trial with SLAMF7 CAR-T cells prepared by virus-free Sleeping Beauty gene transfer to treat multiple myeloma. *Gene Ther.* **2021**, *28*, 560–571, doi:10.1038/s41434-021-00254-w.

232. Guilbaud, M.; Devaux, M.; Couzinié, C.; Le Duff, J.; Toromanoff, A.; Vandamme, C.; Jaulin, N.; Gernoux, G.; Larcher, T.; Moullier, P.; et al. Five Years of Successful Inducible Transgene Expression Following Locoregional Adeno-Associated Virus Delivery in Nonhuman Primates with No Detectable Immunity. *Hum. Gene Ther.* **2019**, *30*, 802–813, doi:10.1089/hum.2018.234.

233. Ali Hosseini Rad, S.M.; Poudel, A.; Tan, G.M.Y.; McLellan, A.D. Optimisation of Tet-On inducible systems for Sleeping Beauty-based chimeric antigen receptor (CAR) applications. *Sci. Rep.* **2020**, *10*, 13125, doi:10.1038/s41598-020-70022-0.

234. Mátés, L.; Chuah, M.K.L.; Belay, E.; Jerchow, B.; Manoj, N.; Acosta-Sanchez, A.; Grzela, D.P.; Schmitt, A.; Becker, K.; Matrai, J.; et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat. Genet.* **2009**, *41*, 753–761, doi:10.1038/ng.343.

14.   <u>**Declaration of collaborative work**</u>

Except where stated otherwise by reference or acknowledgment, the work presented was generated by myself under the supervision of my advisors during my doctoral studies. All contributions from colleagues are explicitly referenced in the thesis. The material listed below was obtained in the context of collaborative research:

Fig. 3A/B: Structural basis for PB integration bias towards TSSs

Collaboration partner: Prof. Dr. Thomas Krey

Institution: Universität zu Lübeck, Lübeck, Germany

Contributions: Identification of hairpin structure within the PB transposase suspected to mediate PB-BET interaction

My contributions: Visualization and figure generation

Fig. 8: Insertion frequencies of PB-Transposons within a 6 kb window around TSSs obtained for each mPB-Mutant

Collaboration partner: Dr. Csaba Miskey

Institution: Paul-Ehrlich-Institute, Langen, Germany

Contributions: Generation of data visualized in figure

My contributions: Data visualization and figure generation

Tab. 3: Number of recovered integration events for each mPB-Mutant and the WT-PB transposase

Collaboration partner: Dr. Csaba Miskey

Institution: Paul-Ehrlich-Institute, Langen, Germany

Contributions: Generation of data listed in table

Tab. 5: Buffers

Collaboration partner: Prof. Dr. Thomas Schulz, Dr. David Nivia

Institution: Medizinische Hochschule Hannover

Contributions: Recipe for Lysis buffer #3

Declaration of collaborative work

Whenever a figure, table or text is identical to a previous publication, it is stated explicitly in the thesis that copyright permission and/or co-author agreement has been obtained. All previously published figures were modified from either [35] or [117] and therefore do not violate copyright guidelines.

15.   <u>**List of publications**</u>

Ivics Z., **Amberger M.**, Zahn T., Hildt E. Immuntherapien zur Behandlung der chronischen Hepatitis-B-Virusinfektion – eine Übersicht unter besonderer Berücksichtigung von CAR-T-Zellen. *Bundesgesundheitsblatt* 2020 https://doi.org/10.1007/s00103-020-03223-7

Amberger M., Ivics Z. Latest Advances for the Sleeping Beauty Transposon System: 23 Years of Insomnia but Prettier than Ever. *Bioessays* 2020 Sep 16:e2000136. doi: 10.1002/bies.202000136.

Prommersberger, S., Reiser, M., Beckmann, J. Danhof, S., **Amberger, M.**, Quade-Lyssy, P., Einsele, H., Hudecek, M., Bonig, H., Ivics, Z. CARAMBA: a first-in-human clinical trial with SLAMF7 CAR-T cells prepared by virus-free Sleeping Beauty gene transfer to treat multiple myeloma. *Gene Therapy* 2021. https://doi.org/10.1038/s41434-021-00254-w

Sandoval-Villegas, N., Nurieva, W., **Amberger, M.**, Ivics, Z. Contemporary Transposon Tools: A Review and Guide through Mechanisms and Applications of Sleeping Beauty, piggyBac and Tol2 for Genome Engineering. *International Journal of Molecular Sciences* 2021, 22, 5084. https://doi.org/10.3390/ijms22105084

Prommersberger, S.; Monjezi, R.; Botezatu, L.; Miskey, C.; **Amberger, M.**; Mestermann, K.; Hudecek, M.; Ivics, Z. Generation of CAR-T Cells with Sleeping Beauty Transposon Gene Transfer. *Methods in Molecular Biology* 2022, 2521, 41–66. https://doi.org/10.1007/978-1-0716-2441-8_3

16.   *Curriculum Vitae*

Personal Information

Name: Maximilian Amberger

Date of birth: 11.12.1993

Place of birth: Mainz, Germany

Higher Education

11.2018 – 08.2022:   **Doctoral candidate**

Paul-Ehrlich-Institute, Langen, Germany

Johann Wolfgang Goethe – Universität, Frankfurt am Main, Germany

04.2017 – 08.2018:   **Master of Science** in Technical Biology

Technische Universität Darmstadt, Darmstadt, Germany

10.2013 – 12.2016:   **Bachelor of Science** in Biology

Technische Universität Darmstadt, Darmstadt, Germany

Higher Education Entrance Qualification

12.2011:   **University Selection Test**

Viña del Mar, Chile

12.2011:   **Secondary School Graduation**

Deutsche Schule Valparaíso, Viña del Mar, Chile