universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Establishment of recombinant Cas9 production and characterization“

verfasst von / submitted by

Carla Lebada, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra pharmaciae (Mag.pharm.)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 605
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Pharmazie
Betreut von / Supervisor:	Ass.-Prof. Dr. Ulrich Lächelt

Acknowledgements

For the opportunity to gather experiences in such an exciting subject area, I would like to express my sincere thanks to Ass.-Prof. Dr. Ulrich Lächelt. It was a pleasure to work with someone whose joy in his work is so contagious and who always has an open ear and useful input for everybody. I am very grateful for his support within the scope of this thesis and have appreciated his patience, kindness and expertise throughout the whole process.

Furthermore, I would like to thank ao. Univ.-Prof. Mag. Dr. Michael Wirth and ao. Univ.-Prof. Mag. Dr. Franz Gabor for providing their laboratories, equipment and "Schmäh". I enjoyed being part of their team and could not have asked for lovelier colleagues. The graduate students Maria and Kathi always found time to give advice and share their knowledge. Together with Fatlinda and Verena, my workmates who made everyday laboratory life fun, they provided a wonderful working environment.

Finally, with special thanks to my "emotional support team" Philippine, Johanna and Julia, I would like to thank my family and friends for always encouraging me and believing in me, for joint library days, shared despair and euphoria, for their enthusiasm and motivation, for creating wonderful memories in and outside the university throughout my years of study. I am very thankful for all the inspiring people I have met along the way and for all the new friendships I have made.

Abstract

The clustered regularly interspaced short palindromic repeat (CRISPR) - CRISPR associated (Cas)9 system is a sequence-specific genome editing tool that impresses with its ease of programmability. For genetic engineering purposes, the system is typically reduced to two components, a single guide RNA (sgRNA) and an endonuclease Cas9, which assemble to a ribonucleoprotein complex and require nuclear entry. The readily modifiable sgRNA defines the target sequence and guides the complex to the corresponding DNA site within the genome, where Cas9 introduces a double strand break.

This thesis has contributed to the groundwork for the testing of ribonucleoprotein delivery vehicles for Cas9, as well as enhanced Cas9 variants, at the Department of Pharmaceutical Sciences, Division of Pharmaceutical Technology and Biopharmaceutics of the University of Vienna. Within this scope, the expression of recombinant Cas9 in *E. coli* via automated purification was successfully implemented, yielding enzymatically active proteins. Furthermore, an in vitro cleavage assay using a linear EGFP gene and an in vitro cell assay using HeLa GFP-Tub cells were established. As carrier materials for intracellular delivery, 10% succinylated polyethylenimines of different average molecular weight (10, 25 and 70 kDa before succinylation) showed no significant GFP knockout and an artificial oligopeptide termed UV_1 was able to mediate distinct GFP knockout in HeLa GFP-Tub cells when associated with Cas9/sgRNA ribonucleoproteins to form nanoparticles. Subsequent dose titration of UV_1 nanoparticles demonstrated GFP knockout ranging from 5 to 27% for ribonucleoprotein concentrations of 4.7 to 75 nM, with the most promising being 18.75 nM.

Summary

Das System CRISPR-Cas9 (Akronym aus dem Englischen: *clustered regularly interspaced short palindromic repeat (CRISPR) - CRISPR associated (Cas)9*), ist ein Instrument für sequenz-spezifische Genmodifizierung, das mit seiner einfachen Programmierbarkeit besticht. Es handelt sich um ein für gentechnische Zwecke auf zwei Komponenten reduziertes Zusammenspiel aus einer single guide RNA (sgRNA) und einer Cas9 Endonuklease. Diese setzen sich zu einem Ribonukleoprotein-Komplex zusammen und müssen als solcher in den Nukleus gelangen, um ihre Wirkung zu entfalten. Die leicht modifizierbare sgRNA definiert die DNA-Zielsequenz und dirigiert den Komplex zum entsprechenden Angriffsort im Genom, wo das Cas9 Protein einen Doppelstrangbruch einführt.

Die vorliegende Arbeit trug dazu bei, die Grundlage für das Testen von Ribonukleoprotein-Transportvehikeln für Cas9, sowie ferner für fortgeschrittenere Cas9-Varianten, am Department für Pharmazeutische Wissenschaften der Universität Wien in der Abteilung für Pharmazeutische Technologie und Biopharmazie zu schaffen. Im Zuge dessen wurde die Expression von rekombinantem Cas9 in *E. coli* mit automatisierter Aufreinigung unter Gewinnung enzymatisch aktiver Proteine erfolgreich implementiert. Weiters wurden ein In-vitro Cleavage Assay, für den ein lineares EGFP Gen herangezogen wurde, und ein In-Vitro Zellassay mit HeLa GFP-Tub Zellen etabliert. Für die intrazelluläre Verabreichung dienten 10% succinylierte Polyethylenimine verschiedenen durchschnittlichen Molekulargewichts (10, 25 und 70 kDa im unmodifizierten Zustand) und ein künstliches Oligopeptid mit der Bezeichnung UV_1 als Trägermaterialien. In Verbund mit Cas9/sgRNA Ribonukleoproteinen wurden diese zu Nanopartikeln assembliert. Als solche zeigten die succinylierten Polymere keinen signifikanten, das Oligopeptid UV_1 hingegen einen eindeutigen GFP Knockout in HeLa GFP-Tub Zellen. Eine infolgedessen durchgeführte Dosistitration von UV_1-Nanopartikeln zeigte einen GFP-Knockout von 5 bis 27% bei Ribonukleoprotein-Konzentrationen von 4,7 bis 75 nM, wobei sich 18,75 nM als die vielversprechendste Konzentration herausstellte.

Table of contents

1. Introduction

1.1. What is CRISPR-Cas?

The clustered regularly interspaced short palindromic repeat (CRISPR) - CRISPR associated (Cas) system, found in bacteria and archaea, provides its prokaryotic host with sequence-specific adaptive immunity against bacteriophage infection and plasmid transfer. [1] Analysis of 324 archaeal and 12,792 bacterial complete genomes (available as of March 2019) showed CRISPR-Cas loci to be present in 85% of the archaea, including 89 of 92 fully sequenced hyperthermophiles, and only 42% of the bacteria analyzed. [2]

The CRISPR loci consist of multiple short repeats of approximately 30-40 bp regularly intermittent by spacer sequences with a similar length, tied with the presence of CRISPR-associated (Cas) genes encoding Cas proteins. The repetitive sequences are identical, in some exceptions with single base pair substitutions, and usually exhibit loose dyad symmetry, whereas non-repetitive spacers are sequences of viral or plasmid origin, of which each CRISPR locus owns a unique set. [1,3,4] Analysis of 4,210 CRISPR loci and 1,949 Cas loci detected in 2015 showed that Cas loci are 75% adjacent (distance of maximum 530 bp) to a CRISPR array and 22% farther away, while CRISPR loci are 33% adjacent to Cas loci and 56% far off. [5] Cas genes are arranged within operon(s) [6] and encode a set of Cas proteins distinct for each CRISPR-Cas subtype, providing the basis of classification into two classes, six types I-VI and 33 subtypes. [2,5] Type I, II, V target DNA, type VI cleaves RNA, type III aims at both DNA and RNA, whereas type IV target sequences remain to be experimentally determined. [7]

The mechanism of the CRISPR-Cas mediated defense system in bacteria and archaea comprises 3 phases: adaptation, expression and interference. In response to viral or plasmid exposure, selected short DNA sequences of the invading foreign DNA elements, termed protospacers, are extracted and then integrated into the CRISPR array in between repeat sequences during the adaptation phase, resulting in functional spacer sequences. [6–9] In the expression phase, the repeat-spacer sequences are transcribed into precursor CRISPR RNA (pre-crRNA), while Cas genes are transcribed and translated into Cas proteins. The pre-crRNA is cleaved by endonucleases into short fragments of mature CRISPR RNAs (crRNAs), each consisting of a complete or partial spacer sequence flanked by truncated repeat-derived sequences at the 3' or 5' end, or both. [6,9–11] The interference phase is characterized by the crRNA guiding a class 1 Cas protein complex or a class 2 single Cas protein to the

complementary target protospacer of the invading plasmids and phages to cleave target DNA or RNA via nuclease activity, thus conferring immunity. Basis for recognition and base-pairing of the crRNA with the protospacer is on the one hand sequence similarity or identity of spacer and protospacer, in type I and II systems on the other hand a protospacer-adjacent motif (PAM) that flanks the protospacer at the 3' end of the non-complementary strand. [6,8,11,12]

1.2. Early days of CRISPR-Cas9

The evolution of the CRISPR-Cas system as an efficient gene editing tool [13] originated in a publication from 1987. Direct repeats of highly homologous sequences with a length of 29 nucleotides interrupted by non-repetitive spacings of 32 nucleotides were reported during nucleotide sequence analysis of an *E. coli* gene. [14] With increasing availability of whole genome sequencing data, the distinctive motif of multiple repetitive short DNA sequences (21-37 bp) of high homology characterized by regularly intervening non-repetitive spacer sequences of similar length were detected in prokaryotic genomes of both bacteria and archaea, generally occurring in clusters. Acknowledged as a family of DNA repeats in 2000, the acronym CRISPR was established for these clustered regularly interspaced short palindromic repeats (CRISPR) in 2002. [3,15] Additionally, CRISPR-associated (Cas) genes that encode proteins were identified invariably adjacent to CRISPR loci, suggesting a functional relationship. [3] In a 2005 publication, CRISPR spacers were found to derive to a great extent from extrachromosomal preexisting sequences of bacteriophages and conjugative plasmids. A relation of these spacer elements with prevented infectivity was observed in multiple cases, thus implying that CRISPR might play a role in conferring specific immunity against foreign DNA. [4] This hypothesis was supported in 2007 by an experiment with the bacterial strain *Streptococcus thermophilus* (subtype II-A [2]). In response to phage infection, the bacteria showed insertion of new spacers derived from phage genomic sequences into the CRISPR loci, resulting in phage-resistance provided by the Cas enzymes. In this context, mutation analysis revealed resistance specificity to be dependent on the similarity of CRISPR spacer sequences with phage sequences. [8] This first experimental evidence was followed by a series of studies within the next four years to comprehend the function and to identify the underlying mechanisms of the different types of CRISPR-Cas systems in adaptive immunity of prokaryotes. [16]

Investigation of the *Escherichia coli K12* (subtype I-E [2]) CRISPR system, published in 2008, showed that phage resistance requires cleavage of pre-crRNA into small crRNAs of approximately 57 nucleotides that each comprise a spacer sequence flanked by truncated CRISPR repeat sequences and direct the type I Cas protein complex (Cascade complex) to viral nucleic acids. [9] Using a *Staphylococcus epidermis* (subtype III-A [2]) CRISPR-Cas system, it was demonstrated in a publication of the same year that the target of interference is DNA homologous to a spacer sequence within the CRISPR locus, irrespective of spacer sequence orientation. [12] However, evidence of crRNA-guided cleavage of RNA targets, found in *Pyrococcus furiosus* (subtype III-B [2]), followed shortly afterwards in 2009. [11]

In 2011, it was experimentally demonstrated for the first time that the (subtype II-A [2]) CRISPR-Cas system of *Streptococcus thermophilus* can be successfully transferred into *E. coli* via recombinant plasmid transformation and still provide resistance against invading plasmids and phages despite the phylogenetically distant host. [17] Further examination of the type II CRISPR-Cas system with Cas9 isolated from *Streptococcus pyogenes* was published in 2012, substantiating that trans-activating CRISPR RNA (tracrRNA) is essential for both, pre-crRNA processing and cleavage of target DNA. By base-pairing, tracrRNA and mature crRNA form a tracrRNA:crRNA duplex that guides the endonuclease Cas9 to the target DNA. Additionally, evidence was provided that this duplex can be mimicked and replaced by a single chimeric RNA, demonstrating the potential as an easily programmable tool for sequence-specific genome editing that only requires altering of a single guide RNA to change the DNA target site. [6]

Findings from 2013 showed that transfection of human and mouse cells with the type II CRISPR locus of *Streptococcus pyogenes* resulted in heterologous expression of RNA-directed Cas9 nucleases that were able to cleave endogenous genomic loci in mammalian cells. For efficient cleavage of the protospacer, the CRISPR locus required to encode only tracrRNA, pre-crRNA and the Cas9 nuclease. By integrating multiple guiding spacer sequences into the pre-crRNA that targeted protospacers in human or mouse loci, simultaneous editing of several target sites within the mammalian genome was enabled, establishing CRISPR-Cas9 as a tool for multiplex human genome engineering. [18,19]

By outperforming the zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) regarding the ease of target-programming, the CRISPR-Cas9 system became the standard method for targeted genome editing in form of a two-component

system consisting of the endonuclease Cas9 and a guide RNA, enabling high-throughput applications by simply changing the guide RNA to determine the target DNA sequence. [13,20]

1.3. Classification of the CRISPR-Cas system

The challenge of classifying CRISPR-Cas systems is posed by their vast diversity in gene and Cas protein composition, organization of genomic loci and molecular mechanisms. Reason for this is the fast evolution due to the strong selection pressure from changing pathogens like phages as well as anti-CRISPR proteins. Based on criteria such as signature CRISPR-Cas genes, sequence similarity or divergence of the aforementioned gene and Cas protein repertoire, Cas operon structure, phylogeny of highly conserved Cas proteins and available experimental data, the CRISPR-Cas systems are divided into two classes of six types I-VI in total. [2,5,21,22] Each distinct type owns a unique set of Cas proteins [23] and is subclassified into further subtypes labeled by letters. An updated evolutionary classification of CRISPR-Cas systems from 2020 proposes 16 subtypes of class 1 and 17 subtypes of class 2. [2]

Cas genes can be grouped into four functional modules that partially overlap: the **adaptation module** responsible for spacer insertion, the pre-crRNA processing **expression module**, the **interference or effector module** that binds crRNA, consequently recognizes, then cleaves target nucleic acids, and the **signal transduction or ancillary module** of CRISPR-linked genes. [2] Primary focus for classification is on effector modules and Cas protein compositions. [2,5] Class 1 systems are characterized by multiple Cas proteins forming an effector complex and are further subdivided into type I, III and IV. Class 2 systems are distinguished by a single-protein effector module, a large protein that consists of multiple domains, and are subclassified into type II, V and VI depending on their domain architectures. [2,21,22]

The adaptation module is relatively uniform and assembled by Cas1, whose integrase activity is responsible for cleaving invading DNA and CRISPR arrays, and Cas2, forming the structural subunit of the resulting heterohexamer. Additionally, Cas4 endonuclease (type I, II, V subtypes), Cas3 (subtype I-F) , Cas9 (type II), the Csn2 protein (subtype II-A) or a reverse transcriptase (type III subtypes) can be involved, sometimes fused with Cas1 or Cas2. [2,5,7,22]

In class 1 systems, multiple Cas proteins form a crRNA-binding effector complex, consisting of Cas3, Cas5-8, Cas10 and Cas11 in various compositions. Cas5, Cas6 and Cas7 are repeat-associated mysterious proteins (RAMPs). Cas8 represents the large subunit in type I and Cas10

the large subunit in type III. Cas11 forms the small subunit and Cas3 is typical for type I systems. After binding to pre-crRNA, the multi-protein effector complex recruits the nuclease Cas6 (type I, III), or seldom Cas5 (subtype I-C), for processing pre-crRNA to mature crRNA. [2,21,22]

In type I systems, the mature crRNA guides the effector complex, denominated as Cascade (CRISPR-associated complex for antiviral defense), to the target protospacer sequence. There, the recruited Cas3 (in some variants fused with Cas2) unwinds the target DNA via helicase activity and degrades the DNA via HD nuclease activity. **In type III systems**, the effector complex is known as Csm or Cmr complex, depending on the subtype. Under the guidance of crRNA, Cas7 subunits with RNase activity cleave nascent mRNA transcripts and/or single stranded DNA (ssDNA) is nicked via HD nuclease activity of Cas10 during transcription. Regarding the highly divergent **type IV systems**, so far Cas5, Cas7, a large subunit analogue and putative small subunits have been identified. Usually possessing neither adaptation modules nor nucleases for interference, most are defective variants, for which the function of inter-plasmid competition is proposed. [2,21,22,24,25]

Class 2 systems are almost exclusively found in bacteria. Types and subtypes differ significantly in terms of processing pre-crRNA into mature crRNA. For cleavage of the target nucleic acid, however, the nuclease domain/s of the respective single-protein effector complex – Cas9, Cas12 or Cas13 – is/are responsible. [2,22]

In type II and several type V systems, bacterial non-Cas RNase III (expendable in some II-C variants) binds and cleaves pre-crRNA that is base-paired with trans-activating CRISPR RNA (tracrRNA) bound to the effector module. The resulting mature tracrRNA:crRNA duplex remains bound and directs the effector module to the target protospacers, where double strand breaks (DSBs) are introduced by the effector module. Cas9, the effector of type II, possesses a HNH and a RuvC-like nuclease domain which each cleave a single strand of target DNA, whereas the type V effector Cas12 has just one RuvC-like domain that is able to cleave both strands, except for Cas12f (previously Cas14) and Cas12g with ssDNA cleavage activities. **In subtype V-A and type VI systems**, on the other hand, the single-protein effector complex itself contains an RNase domain that catalyzes pre-crRNA maturation. In addition to the RNase domain, the type V-A effector Cas12a (previously Cpf1) contains two DNase domains, including a RuvC-like and a putative nuclease domain. Cas12a is guided by mature crRNA only, without tracrRNA, and cleaves target double stranded DNA (dsDNA). In contrast, the type VI

effector Cas13 stands out for exclusively targeting and cleaving RNA with its two distinct RNase domains: one for pre-crRNA maturation, the other one for crRNA-directed unspecific RNA degradation. The latter catalytic domain is composed of two HEPN (higher eukaryotes and prokaryotes nucleotide-binding) domains, a distinct feature of type VI. [2,21,22,24,26–29]

Lastly, the ancillary module is a diffuse collection of proteins connected with the CRISPR-Cas systems, taking on various functions like signal transduction. With the recent discovery of many ancillary genes, experimental data have yet to be collected. [2] A noteworthy example is provided by cyclic oligoA signal transduction in type III systems. With its cyclic oligoA polymerase activity, the target-bound effector complex Cas10 stimulates the ancillary RNases Csm6 or Csx1 to mediate random RNA degradation. [2,22,24]

1.4. Structure and mechanism of action of CRISPR-Cas9

The well-examined type II system, characterized by its signature multidomain protein Cas9, is widely employed in the field of genome editing, [21] with the most extensively used approach being *Streptococcus pyogenes* Cas9 in conjunction with a single guide RNA (sgRNA). [13,23] The type II systems are subdivided into three subtypes II-A, II-B and II-C based on their CRISPR loci. Each locus generally comprises Cas9, Cas1 and Cas2 genes as well as tracrRNA-encoding DNA. Subtype II-A is distinguished by an additional Csn2 gene, subtype II-B and II-C2 by an additional Cas4 gene. [2,5]

Knowledge about the structure of the Cas9 endonuclease stems from crystal structures and cryofixed electron microscopy structures of the ortholog *Streptococcus pyogenes* (SpyCas9, subtype II-A), *Staphylococcus aureus* (SaCas9, subtype II-A), *Actinomyces naeslundii* (AnaCas9, subtype II-C) and *Francisella novicida* (FnCas9, subtype II-B) Cas9 proteins. [13,23,30,31] Cas9 is structured into two lobes, a nuclease (NUC) and an α-helical recognition (REC) lobe, which are linked by an arginine-rich helix and a loop. The NUC lobe comprises two distinct nuclease domains, on the one hand the HNH endonuclease domain that cleaves the DNA strand complementary to the guiding RNA (target strand) and on the other hand the RuvC-like endonuclease domain for cleavage of the non-complementary DNA strand (non-target strand). [6,13,23,32–34] In its apo form, the nuclease domains of the Cas9 protein are inactive. To experience conformational change and enter a catalytically active state that enables target recognition, Cas9 requires guide RNA (gRNA) binding. [23,33] Also part of the NUC lobe are the

evolutionarily divergent wedge (WED) domain and the PAM-interacting (PI) domain, which consists of a topoisomerase-homology (TOPO) domain and a C-terminal domain (CTD) with a positively charged groove in between. [6,13,23,32–34] The catalytic domains and the bridge helix are well-conserved among Cas9 orthologs. Not only differing in size, Cas9 orthologs also show sequence divergence within the REC lobe, the least conserved element, as well as in the WED domain and in the C-terminal domain of the PI domain. [23,33,34] With their different WED and REC structures, each Cas9 ortholog recognizes different structural features of gRNAs. The WED domain and the REC domains enable this orthogonal recognition of dual and single guide RNA scaffolds mainly via interactions with the RNA backbone, to which the bridge helix also contributes. [23,34] The PI domain is responsible for PAM recognition and differs in its PAM-interacting residues, which determine the PAM specificity of the Cas9 orthologs. Furthermore, the WED domain interacts with the phosphate backbone of the PAM. [13,34]

In type I and II systems, protospacers require the presence of a functional protospacer-adjacent motif (PAM) next to the protospacer to be selected for spacer acquisition, which simultaneously confers specificity for foreign DNA and avoids self-targeting, as repeat sequences lack PAMs. [7,16,35] The PAM-interacting domain of Cas9 is responsible for PAM specificity in type II systems and recognizes a distinct PAM sequence. Then, Cas9 directly interacts with the Cas1-Cas2 complex (and the accessory protein Csn2) to ensure spacer acquisition via dsDNA nuclease activity of Cas1. [7,35]

Cas1 and Cas2 proteins form a heterohexameric integrase complex with two binding sites, one for the protospacer and one for the CRISPR loci DNA sequences. In type II CRISPR-Cas systems, only polarized integration occurs, the integration of new spacers at the CRISPR array end toward the leader sequence, which confers a stronger resistance against the most recently invaded DNA molecules. In type II systems, this is mediated by interaction of a Cas1 α-helix with the leader anchoring sequence (LAS) and does not require the aid of any host factors. After binding the recruited protospacer, the Cas1-Cas2 complex catalyzes ssDNA cleavage and ligation within the CRISPR locus: first at the intersection between leader and repeat sequence, then between repeat and spacer sequence. The 3'-OH ends of the protospacer nucleophilically attach to the 5' ends of the repeat sequences, resulting in spacer dsDNA flanked by two ssDNA repeat sequences that are, presumably by a DNA polymerase, completed to a double strand and ligated. [7]

The repeat-spacer array of the CRISPR locus is transcribed into precursor CRISPR RNA (pre-crRNA). Furthermore, the transcription into small trans-activating CRISPR RNA (tracrRNA), occurs. This small non-coding RNA comprises an anti-repeat sequence and is trans-encoded on the strand opposite the CRISPR array and Cas genes. Molecules of tracrRNA hybridize to the repeats of the pre-crRNA and initiate crRNA maturation. The host RNase III recognizes and processes the RNA duplexes by cleavage of pre-crRNA and tracrRNAs with Cas9 present. The resulting mature duplexes of tracrRNA:crRNA play a significant role in target DNA recognition and trigger DNA cleavage by Watson-Crick base-pairing with protospacers. [6,10]

To simplify the application of Cas9 in genome editing, the tracrRNA:crRNA duplex can be substituted by a chimeric single guide RNA (sgRNA). [6] In 5'→3' direction, an sgRNA consists of a crRNA followed by a linker loop that connects to a tracrRNA. The crRNA comprises a spacer located at the 5' end as well as a repeat sequence, which is linked to the anti-repeat sequence of the tracrRNA. The complementary sequences hybridize, forming a repeat:anti-repeat duplex with a hairpin structure as well as further stem loops in the remaining tracrRNA at the 3' end of the sgRNA. [13,23,33] In the following paragraphs, the term guide RNA (gRNA) is used to refer to both the dual guide complex and the sgRNA.

In the apo state, Cas9 is auto-inhibited due to its sterically blocked and outward facing HNH domain active site. Binding of a gRNA induces the formation of a central channel that confers Cas9 the ability of DNA recognition and binding, with the entering DNA substrates being enclosed by the two lobes and the nuclease domains located on opposite sides. [13,33,34] Scanning of the DNA for protospacers happens via three-dimensional collisions. [36] First requirement for target recognition is the presence of a PAM sequence, [36] whose identification involves base-specific interactions with the C-terminal PI domain. [32,34] SpyCas9, for example, recognizes a 5'-NGG-3' motif in the non-target strand directly downstream of the protospacer, [6,32,36] with N standing for any nucleotide. [35] While the nucleobases of the GG dinucleotide in the non-target strand are detected by the PI domain in the major groove via hydrogen-bonding, only the backbone of the complementary dinucleotide in the target strand experiences interactions. [32] The complex of Cas9 and gRNA quickly dissociates from sequences that do not contain a PAM motif, therefore minimizing the time spent at non-PAM off-target sites and accordingly preventing self-targeting. Only when this complex encounters and binds a PAM sequence, stepwise unwinding of the adjacent potential target DNA is initiated at the PAM, simultaneously followed by stepwise RNA:DNA base-pairing. [36] For this purpose, the

phosphate lock loop of Cas9, in SpyCas9 located in the CTD and in SaCas9 between the WED and RuvC-III domains, forms hydrogen bonds with the +1 phosphate in the target DNA strand backbone. The +1 phosphate, which connects the PAM-complementary sequence with the DNA to be read, is consequently rotated. This introduces a kink in the target DNA strand, which promotes DNA unwinding, stepwise interrogation of the nucleotides by Cas9/gRNA and stepwise formation of the first gRNA:DNA base pairs next to the PAM. [32,34] Each base pair further destabilizes the target DNA duplex and thus enables formation of a gRNA:target DNA heteroduplex via gradual Watson-Crick base-pairing. [32,34] As base-pairing proceeds from the PAM sequence, complementarity of the target DNA with the gRNA is particularly critical for the PAM-proximal seed region of ~8-13 bps. Hence, divergence near the PAM causes merely ephemeral binding. In case of sequence homology for more than ~12 bps, strand separation succeeds because of reaching an energy threshold in favor of further unwinding and base-pairing rather than re-winding. PAM-distal zones, in consequence, are less sensitive to mismatches. [13,34,36] Together with the relocated non-target ssDNA, the resulting gRNA:target DNA heteroduplex forms an R-loop structure. [30]

The R-loop formation of SpyCas9 is characterized by a 30° bend in the DNA helix. Both DNA strands are kinked at the positions +1 and -1 in the phosphodiester backbone. While the gRNA:target ssDNA heteroduplex is located in the central channel, the non-target ssDNA parallels upstream position -5 in a NUC lobe side tunnel. The kink at the +1 phosphate in the target DNA strand is introduced by interaction with Cas9 as described before. Contrarily, the non-target strand shows no direct Cas9 contacts with the -1 phosphate, but several with the -2 and -3 phosphates, which are responsible for the distinct kink at position -1. Further interaction with the -4 phosphate results in additional kinking of the non-target phosphodiester backbone at this position. [30] Binding of both target DNA strands leads to a conformational change, with the REC and NUC lobes shifting toward the central channel. Initiated by a place swap with re- or unfolding, respectively, of the hinge regions L1 and L2 linking the HNH and RuvC domains, R-loop stabilization is conferred. Also, the cleavable phosphate of the non-target strand is positioned in the active site of the RuvC domain. Furthermore, the HNH domain pivots by 180°, placing the catalytic center adjacent to the scissile phosphate of the target strand. [30] Consequently, the RuvC endonuclease domain cleaves the non-target strand 3-4 bp upstream the PAM sequence (with ensuing 3'→5' exonucleolytic cleavage up to 8 bp upstream the PAM) and the HNH endonuclease cuts the

target strand 3-4 bp downstream the PAM-complementary sequence, creating blunt-ended double strand breaks. [6,20] This DNA DSB may activate two DNA repair mechanisms: non-homologous end joining (NHEJ) and homology directed repair (HDR). [13,20] The error-prone NHEJ pathway mediates virtually random insertions or deletions (indels) at the DSB site, which potentially lead to functional gene knockout, e.g. by causing frameshift mutations or premature stop codons. [20,37,38] Optionally, homology-independent insertion of a transgene is possible upon co-delivery of a donor vector with the transgene enclosed by the target sequence. [20,23,37] As NHEJ operates the fastest and is independent of the cell cycle status, it is the main repair pathway for DSBs. [37] The HDR pathway, in contrast, is only active throughout the S or G2 phase and requires the presence of a DNA template flanked by the target sequence in a manner analogous to the DSB site. [20,37–39] Via homologous recombination, desired sequences are integrated into the target site, implementing mutations, correction of mutations, gene knock-in or knockout. [13,20,39] Since the HDR pathway is less prone to errors [20,39] and very precise, [24] but occurs less frequently and only in replicating cells [23,38], efforts are made to raise HDR efficiency, e.g. by manipulating the DSB repair pathway rates towards HDR. [13,38]

1.5. Delivery systems

For successful genome editing by the CRISPR-Cas9 system, the Cas9 protein and the gRNA must assemble into a functional Cas9/gRNA ribonucleoprotein (RNP) complex and enter the nucleus of the target eucaryotic cell. [37,40] This can be achieved via delivery of (1) plasmid DNA (pDNA) encoding Cas9 and gRNA, (2) Cas9 mRNA transcripts and gRNA, (3) Cas9/gRNA RNP complexes or combinations thereof, [39–42] occasionally along with a DNA template for HDR. [37,40] While pDNA requires nuclear delivery for expression of Cas9 and shows genome editing events after 24-48 hours, mRNA only needs to be introduced to the cytosol, initiating instant translation and a more promptly assembly of Cas9/gRNA RNPs. Directly delivering RNPs into the nucleus naturally achieves the fastest genome editing, already detectable after 1 hour. [40,42] Out of these three kinds of biomolecular formats, pDNA is used the most frequently and exhibits the best stability, with a half-live of a few hours in biological liquids. [40,42] However, the concomitant long exposure to Cas9 increases the rate of off-target effects. The susceptibility of ssDNA to randomly incorporate into the genome also raises the possibility of

constitutive Cas9 expression, which would cause an additional increase in off-target events. [40–42] In contrast, the chemically and nucleolytically more unstable mRNA degrades more rapidly, possessing a half-life in the range of a few minutes in biological environments. This lowers the risk of off-target effects due to the shorter-termed exposure to Cas9. Accordingly, direct delivery of RNP complexes offers an even lower off-target effect risk, as RNPs persist only briefly due to fast protein degradation. [40,42] Concerning immunogenicity, RNPs activate the adaptive immune system, whereas bacterial plasmid backbone sequences and particularly exogenous mRNA stimulate the innate immune system. By insertion of chemically modified nucleotides into the mRNA sequence, mRNA immunogenicity and stability can be improved. [40] In terms of production and clinical application, the state of knowledge is more advanced for intracellular delivery of nucleic acids than of proteins. The main advantages of pDNA delivery are comparatively low costs, possible scale-up and ease of production. [40] Although the delivery of mRNA or RNPs is associated with higher expenses, it is upvalued by more efficient, more transient and thus safer genome editing, [37,40,43] and allows approximated prediction of intracellular pharmacokinetics. [40] RNP delivery stands out due to good gene editing results with a diminution of off-target effects, insertional mutagenesis and immunogenicity at the same time. [37,40,41] Furthermore, it is a solid method for cells with low transcriptional and translational rates. [37,40] Regarding delivery vehicle formulations, it is essential to know that upon complexation of the negatively charged gRNA and the positively charged Cas9 protein the net charge of the ribonucleoprotein is negative. [40,41] Also, Cas9 RNPs denature swiftly in organic solvents, posing a challenge for lipid nanoparticle formulations. [40]

The CRISPR-Cas9 delivery methods differ not only in the biomolecular format of their cargo, but also their vehicle, resulting in a classification into (1) physical methods, (2) viral vectors and (3) non-viral vectors. [40,41,44]

The **physical approaches** are usually limited to in vitro and ex vivo use and include microinjection, electroporation, hydrodynamic delivery, micro/nanofluidics, filtroporation, nanotube, iTOP (induced transduction by osmocytosis and propanebetaine) and protoplast transformation. [37,41,45] Microinjection is a well-established standard method performed by direct injection into a single cell using a microneedle, whereas electroporation uses electric current to open nanopores in cell membranes. Both of these in vitro/ex vivo methods are characterized by high intracellular delivery efficiency. [37,41,42] Hydrodynamic delivery, on the other hand, shows poor transfection quotas and is only applied in vivo, generating

hydrodynamic pressure in the blood circulation to confer higher permeability into endothelium and parenchyma. [41] Because of severe adverse reactions from tissue trauma to mortality associated with hydrodynamic delivery, clinical application is refrained from. [41,45]

Viral vectors used for joint or separate transduction of Cas9 gene and gRNA are the ssDNA adeno-associated virus (AAV), the dsDNA adenovirus (AV) and the RNA lentivirus (LV), which are applied in vitro, ex vivo and in vivo. [41,44,45] Also, epstein-barr virus, sendai virus and baculovirus vectors are being investigated for CRISPR-Cas9 delivery. [45] AAV is widely used in gene therapy, mediates persistent transgene expression and is little immunogenic. With its 20-26 nm in diameter, the packaging capacity is restricted to 4.1-4.9 kb, requiring either two separate vectors or insertion of the smaller SaCas9 gene instead of the SpyCas9 gene. AV and LV, in contrast, provoke a severe immunoreaction and are 80-100 nm in diameter. With a genome size of 8 (AV) or 7-12 kb (LV), packaging of additional genetic elements is possible. [20,41,45] However, infecting target cells with the CRISPR-Cas9 viral particles, which are usually generated by transforming HEK293T cells, [41] raises safety concerns regarding immunological response, mutagenesis and tumorigenesis. [42] Furthermore, genomic integration is common for LV and a residual risk for AV and AAV. [41,45]

Non-viral CRISPR-Cas9 delivery systems impress with high biocompatibility and loading capacity, little immunogenicity, capability of upscaling as well as specificity. [42,45] They can be divided into polymeric, lipidic, inorganic and inorganic/organic hybrid, protein-, peptide- and DNA-based delivery systems, as well as bio-derived vesicles. [41] Most formulations consist of multiple components and/or various modifications, which is why only certain characteristics are highlighted in this rough overview.

Often-cited polymers used to create **polymeric delivery systems** include polyethylenimine (PEI), poly(lactic-*co*-glycolic acid) (PLGA), chitosan, polyamidoamine (PAMAM) dendrimers, pH-sensitive poly(β-amino ester)s (PBAEs), ethanolamine-aminated poly (glycidyl methacrylate) (PGEA) as well as poly(amino acids) such as polylysine (PLL), poly(aspartic acid) and derivatives. [40,41,46] Unmodified, their basic structures are polycations and, apart from PEIs and polymethacrylates, biodegradable. [37,40,47–51] However, these polymers are commonly modified to achieve lower toxicity, higher biocompatibility, complex stability as well as efficient delivery at the same time. As their positive charge is unfavorable in vivo in terms of fast elimination via reticuloendothelial system, they are often subject to modification with biocompatible elements, e.g. polyethylene glycol (PEG), polyglutamic acid or polysaccharides.

[37,40,42] To enable efficient RNP binding, dendrimers can be modified with phenylboronic acid (PBA), and PBAEs with carboxylate ligands. Further functionalization can involve incorporation of PEI, PEG, targeting ligands, cell-penetrating peptides (CPPs), lipids, nuclear delivery tags and others. [37,40] PEI is utilized in many non-viral formulations for CRISPR-Cas9 delivery, oftentimes as a coating. In addition to enhancing cellular uptake, as PEI allows endosomal or lysosomal escape by buffering the pH via its proton sponge effect, it can also improve RNP loading capacity. [37,40,41,52] By exploiting covalent and hydrogen bonding, ionic and hydrophobic interactions, up to cyclodextrin-adamantane host-guest interactions, differently structured nanoparticles, e.g. core-shell assemblies, or other organizations are formed. [37,40]

Lipidic delivery systems in the form of liposomes and other lipid nanoparticles (LNPs) can be applied in vitro, ex vivo and in vivo. [37,40,41] Typically, the lipid nanoformulations feature cationic lipids. [37,41,42] LNPs allow encapsulation of pDNA, RNA or RNP by complexing nucleic acids and are taken up by the cells via endocytosis. [37,41] For plasmid delivery, they are based on 1,2-dioleoyl-3-trimethylammoniumpropane (DOTAP) and commonly also contain cholesterol, 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) and 1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-PEG (DSPE-PEG). [37,40] Advancements of LNP systems that enable the co-delivery of mRNA and sgRNA include the development of biodegradable (e.g. cholesteryl-based) lipidoids, zwitterionic amino lipids [40] as well as selective organ targeting (SORT) based on the lipid composition. [37,40] The success of COVID-19 mRNA vaccines and the first phase 1 clinical trial of in vivo CRISPR-Cas9 gene editing emphasize the relevance of LNPs for mRNA delivery. [40] The delivery of RNPs can be successfully mediated by the widely used commercial transfection reagents Lipofectamine 2000, 3000 and RNAi^{MAX} which form cationic liposomes. With the aim of improving RNP delivery, Lipofectamine CRISPR^{MAX}, bioreducible lipids, cationic chalcogen-containing lipids, noncationic lipids and stimuli-triggered liposomes were designed. [37,40,42] Generally, optimization strategies for lipidic delivery systems pursue (multi)functionality via conjugation of lipids to ligands for targeting purposes, use of localization tags, pH-sensitive amino lipids and integration of stimulus-sensitive components. [37,40,42]

Inorganic and inorganic/organic hybrid delivery systems for pDNA or RNP comprise, inter alia, spherical gold nanoparticles (AuNPs), non-spherical gold nanorods (AuNRs) and nanowires (AuNWs), differently coated mesoporous silica nanoparticles (MSNs), metal-organic frameworks (MOFs) and light-sensitive upconversion nanoparticles. [37,40,42] The inert

gold nanoplatforms, which differ in shape, offer the possibility of facile modification. A few examples are gold-thiol bond mediated attachment of oligonucleotides to bind their cargo, cationic polymer coatings made from PEI or poly(N-(N-(2-aminoethyl)-2-aminoethyl) aspartamide (PAsp(DET)) to promote endosomal escape, and functionalization with CPPs, arginine or glutathione. [37,40,41,45] Out of the metal-organic frameworks, the subclass of zeolitic imidazole frameworks (ZIFs) is deployed as CRISPR-Cas9 nanocarriers. ZIF-C, ZIF-8 and the ATP-responsive ZIF-90 are porous self-assemblies of Zn^{2+} and imidazolate-linkers. [37,40] Furthermore, iron oxide nanoparticles, carbon dots (CDs) and $CaCO_3$-based delivery systems are used for pDNA delivery. [40] For RNP delivery, semiconducting copper sulfide nanoparticles, PEG and PEI modified graphene oxide (GO), biodegradable black phosphorus (BP) nanosheets and calcium phosphate nanoparticles are applied. [37,40] Almost all of the aforementioned systems are modified to gain multi-targeting capability, greater loading capacity or other functionalities. [37,40]

Peptidic delivery systems enable RNP delivery and can be distinguished by concept: cell-penetrating peptides (CPPs) that are normally covalently bound to Cas9 and facilitate cellular internalization on the one hand, peptides forming nanocomplexes with RNPs by non-covalent bonding on the other hand. [40,41,46] Examples of CPPs are supercharged polypeptides (SCPs) and low-molecular-weight protamine (LMWP) which mediate efficient cellular uptake when fused to Cas9. Certain SCPs also enable nuclear translocation. [37,40] The category of peptide nanocomplexes includes various types of peptides such as amphiphilic R7L10 peptides, derivatives of endosomolytic peptides and CPPs, lipopeptides and lipid-containing oligo(ethylenamino) amides (lipo-OAA). Lipopeptides are lipid-modified peptides with higher membrane permeability and self-assembly tendencies, among which an oleic aldehyde bound cationic peptide PT_{24}, a blood-brain barrier permeable peptide dNP2 conjugated with caprylic acid, as well as a tandem peptide comprising a targeting peptide for cell-selectivity, a CPP and a palmitoyl tail showed efficient RNP delivery. [37,40] The lipo-OAAs are symmetrical T-shaped artificial peptides with two tyrosine trimers and two fatty acid residues, which obtained the best results when conjugated to hydroxy-stearic acid. [37,40,53] **Protein-based delivery systems** for pDNA delivery employ histone cores, bi-specific antibody derivatives for cell-specific binding and human serum albumin nanoparticles. RNP delivery systems based on proteins are, for instance, chitosan-coated nanoparticles with a red fluorescent protein core, DNAzyme-

based DNA nanoassemblies with a streptavidin scaffold, and tetralysine-conjugated H-chain apoferritin shell nanoparticles. [40,54,55]

Bio-derived vesicles applied in CRISPR-Cas9 delivery are cell-derived extracellular vesicles capable of cell-specific targeting such as tumor or epithelial cell-derived microvesicles, nanovesicles, exosomes and exosome-liposome hybrids. [37,40,42] For Cas9 RNP delivery, leukemia virus (MLV)-like particles and lentivirus (LV)-like particles are used as well. The bio-derived vesicles can be further decorated with chimeric antigen receptors or glycoproteins for cell-specific targeting, e.g. the CD4-tropic HIV-1 envelope glycoprotein. [40,46,56]

DNA-based delivery systems include DNA nanogels, sphere-like nanoclews and nanoflowers. [40,41,46] The frameworks of these DNA nanostructures are formed via hybridization between DNA and are loaded with RNP by hybridization of sgRNA segments with complementary DNA sequences. [37,40,57,58] DNA nanogels are prepared by DNA-crosslinking after RNP embedding, e.g. via DNA linkers that connect DNA-grafted polycaprolactone brushes. [37,40,58] In contrast, DNA nanoclews and nanoflowers are synthesized via rolling circle amplification with subsequent RNP complexation. [37,40,57] Besides modification of the DNA nanovehicles, e.g. by coating them with PEI, it is also possible to incorporate DNA sequences for targeted delivery into the DNA scaffold, e.g. DNA aptamers that target tumor cells or microRNA-responsive elements for tumor cell-specific RNP release. [37,40,57]

For more efficient genome editing and less off-target cutting, responsive and targeted delivery mechanisms play an important part and are frequently employed in various CRISPR-Cas9 delivery systems, [37] of which some examples are presented here. **Stimuli-responsive delivery** enables spatiotemporally controlled release. CRISPR-Cas9 intracellular release can be realized on demand of stimuli such as light, ultrasound, reduction, or pH value. Examples of light-responsiveness include photocaged sgRNA with several nucleobases in the protospacer blocked by caging groups until photolysis by UV light, upconversion nanoparticles that convert NIR into UV/Vis light and are attached to RNPs by photocleavable linkers, liposome-incorporated photosensitive verteporfin which produces singlet oxygen that disrupts the lipid bilayer upon stimulation with 690 nm light. [37,40] Ultrasound as a stimulus is applied in nanoliposome systems conjugated to ultrasound-activated microbubbles, which result in microbubble cavitation-mediated sonoporation upon ultrasound agitation. [37,40,59] Examples for reduction sensitivity are materials that contain disulfide bonds which are reduced and cleaved by the intracellular higher concentrated glutathione. [37,40] The pH differences of

tumorous and healthy tissues, extracellular milieu and intracellular compartments can be exploited to promote intracellular release or endosomal escape. Typically, pH-responsive materials are based on the pH-dependant protonation or deprotonation of ionizable weak acidic or basic residues and consequently change their structure or other properties. [37,40]

Targeted CRISPR-Cas9 delivery systems, inter alia, employ galactose or trimeric N-acetylgalactosamine ligands for receptor-mediated endocytic hepatocyte uptake, tripeptide RGD (Arg-Gly-Asp) to target cancer cells overexpressing $\alpha_V\beta_3$ integrins, the analogue iRGD to target integrins and neuropilin-1, RNA or DNA aptamers like the AS1411 aptamer to target nucleolin receptors in tumor cell membranes, folate to target folate receptors, ARTA to target inter-photoreceptor retinoid-binding protein (IRBP), phenyl boronic acid to target cancer cells overexpressing sialic acid, as well as anti-hEGFR, nuclear localization signals (NLS) and antibodies. Furthermore, cancer cell membrane coated ZIFs, e.g. with MCF-7 cells, and the aforementioned SORT lipid nanoparticles are noteworthy. [37,40,42]

Among over 30 CRISPR-Cas clinical trials ongoing during the year 2022, the majority were conducted ex vivo, as realization of safe, specific and efficient delivery is more complex in vivo. [42] Clinical application of CRISPR-Cas treatments is generally limited by safety, specificity, efficiency, and efficacy. On the one hand, delivery methods must be optimized in terms of tissue or cell specificity, maximizing uptake and minimizing immune response and toxicity, especially for in vivo application. [24,38,39,42,43] In this context, non-viral delivery methods are currently in the focus of research as promising in vivo strategies. [45] On the other hand, improving the editing efficiency of the CRISPR-Cas system itself and reduction of off-target effects by enhancing the target specificity is crucial. [24,38,42,43] To reduce off-target events and increase target specificity, the sgRNA design can be improved, e.g. by ensuring a GC content of 40-60% or truncation to 18-20 bp, and enhanced Cas9 variants can be employed. [42] Fusion of the restriction endonuclease FokI with the nuclease-deactivated Cas9 (dCas9), of which both nuclease domains are mutated and thus non-functional, creates fCas9 that is only catalytically activated upon dimerization. [38,43,60] The nickase nCas9, engineered by mutation of one of the two endonuclease domains, introduces ssDNA nicks. This has the advantage that single off-target nicks are rapidly repaired, while simultaneous nicking of opposite strands at the on-target site results in sticky-end DSBs and activates HDR instead of NHEJ if a suitable DNA template is present. [13,20,38,42,60] Both strategies are based on imperative nuclease pairing and use a dual sgRNA approach targeting adjacent sites on opposite DNA strands to provide

introduction of DSBs with higher specificity, with nCas9 being the more promising one. [13,38,43,60]

1.6. Biomedical application

The CRISPR-Cas system is responsible for remarkable advances in biomedical sciences. Rapid progress of research in CRISPR-Cas technologies has led to a wide range of applications, [46,60] holding great therapeutic potential, especially concerning gene therapy. [24,38]

Using lentiviral delivery of sgRNA libraries, CRISPR-Cas9 has improved **high-throughput genomic screening**. Via loss- or gain-in-function mutations primarily mediated by knockout but also knock-in and transcriptional modulation, functional study of genes with lower off-target rates than RNA silencing is possible. [20,38,43,46,60] Moreover, CRISPR-Cas9 enables the fast generation of **cell and animal models of disease** via induced pluripotent stem cell editing or zygote microinjection, respectively. Due to its ability to perform multiplex genome editing, polygenic diseases such as cancer can be implemented. [16,20,24,38] In terms of **cancer management**, CRISPR-based diagnostic tools are emerging and approaches of inhibiting loss-in-function (LIF) mutations of tumor suppressor genes, gain-in-function (GIF) mutations of proto-oncogenes or checkpoint molecules are being investigated, e.g. holding the potential to prevent human papillomavirus (HPV) infections. Furthermore, CRISPR-Cas9 is used in cancer immunotherapy to engineer more potent CAR T cells. [24,38,43,46] The potential of cell therapy with Cas9-edited T cells has been demonstrated in two recent ex vivo clinical trials in malignant cancers. [45] Another important field of CRISPR-Cas application is the **treatment of human genetic diseases**. In animal models, CRISPR-based treatment of duchenne muscular dystrophy (DMD), hereditary tyrosinemia type 1 (HT1), α1-antitrypsin deficiency, sickle cell anemia, as well as mutation-caused cataract, liver damage and hearing loss showed promising results. [24,38,45,60] In studies using CRISPR-Cas9 to treat patient-derived primary cells, genetic mutations of DMD, cystic fibrosis, Leber's congenital amaurosis, retinitis pigmentosa, fanconi anemia, hemophilia, sickle cell disease and β-thalassemia could be rectified. The latter two hemoglobinopathies are center of ongoing CRISPR-Cas9 clinical trials of ex vivo cell therapy and Leber's congenital amaurosis of an ongoing CRISPR-Cas9 clinical trial in vivo. [24,38,45,60] This year, in November 2023, the first CRISPR-Cas9 based gene editing therapy – exagamglogene autotemcel (Casgevy™, exa-cel) of Vertex Pharmaceuticals and CRISPR Therapeutics – was

approved for treatment of sickle cell disease and transfusion-dependent β-thalassemia in the UK by the Medicine and Healthcare products Regulatory Agency. Approval by the Food and Drug Administration in the US was granted for sickle cell disease in December 2023 and is anticipated for the indication of β-thalassemia in March 2024. Using electroporation, hematopoietic stem cells isolated from the patient are treated with ribonucleoprotein complexes targeting the gene BCL11A and reinfused after depletion of the patient's bone marrow cells, resulting in resumption of fetal hemoglobin expression. [61–63] Another area of application is the CRISPR-based **management of infectious diseases**, for instance the use against the human immunodeficiency virus (HIV), the herpes simplex virus type 1 (HSV-1) and the hepatitis B virus (HBV). HIV, against which CRISPR-Cas9 was able to successfully confer resistance in primary immune cells in prior studies, and HSV-1 are both represented in CRISPR-Cas9 in vivo clinical trials. [24,38,43,45,60]

Moreover, the nuclease-deficient but sequence-specific Cas9 (dCas9) is employed in various fields of application. For **gene regulation**, target genes are repressed (CRISPRi) or activated (CRISPRa) depending on whether dCas9 is used alone or fused to transcriptional repressors or activators, which can be exploited in functional high-throughput genomic screens and as a therapeutic instrument. [13,38,44,46,60] For **epigenome editing**, dCas9 is fused to epigenetic effectors like DNA or histone demethylases and methyltransferases as well as histone acetyltransferases, achieving targeted chromatin manipulation. [13,44,46,60] For **base editing**, the non-functional dCas9 or nickase nCas9 is fused to a cytidine deaminase or adenosine deaminase. The resulting cytosine or adenine base editors introduce precise base conversions more efficiently than HDR and without introducing DSBs. [24,44,46,60] By fusing dCas9 to fluorescent proteins, **genomic imaging** of live cells is possible, allowing the visualization of the dynamics of genomic loci. [13,20,43]

Along with the capability of permanent and inheritable editing of human genomes, ethical and legal issues arise. [24,38] Center of discussion are the unknown risks and consequences, legitimization of therapy or enhancement, as well as the **modification of human germline cells, zygotes and embryos**. [24,42,46] Moreover, CRISPR-based **gene drive technology**, which mediates the conversion of heterozygous into homozygous alleles, is able to interfere with the probability of inheritance within a population. [24,60] As a consequence, human genome editing entails the need for appropriate guidelines and policies on ethical matters to avoid severe implications. [46,60]

1.7. Aim of the thesis

This thesis aims to establish the production and characterization of the recombinant genome editing protein Cas9 at the Department of Pharmaceutical Sciences, Division of Pharmaceutical Technology and Biopharmaceutics of the University of Vienna. In this context, another objective is to create an in vitro EGFP knockout system for Cas9 in human cells, with the perspective to enable the study of nanocarrier materials for targeted delivery of genome editing ribonucleoproteins to the nucleus. In regard of other genome editing proteins, their production can take place analogously. The activity assays, however, would need to be altered respectively.

2. Materials and Methods

2.1. Chemicals and other materials

The following table shows the chemicals, reagents and biological materials used and where these materials were obtained from.

Addgene Europe, UK

- pET28a-Cas9-His, which was a gift from Caixia Gao (Addgene plasmid #98158; http://n2t.net/addgene:98158; RRID:Addgene_98158)

Bio-Rad Laboratories, Inc., USA

- Precision Plus Protein™ Standards, unstained

Carl Roth GmbH + Co. KG, Germany

- Acetic acid 100%
- Acrylamide Solution (30%) – Mix 37.5 : 1
- Ammonium persulfate (APS)
- D(+)-Sucrose
- Dimethyl sulfoxide (DMSO)
- Ethylenediaminetetraacetic acid (EDTA) disodium salt 2-hydrate
- EDTA tetrasodium salt hydrate
- Gentamycin sulfate
- Glycine
- HEPES
- Hydrochloric acid
- Sodium dodecyl sulfate (SDS), ultra pure
- Sodium chloride
- Sodium hydroxide, beads
- Sterile, ready-to-use, nuclease-free, autoclaved, DEPC-treated water
- TRIS

Corning Incorporated, USA

- Fetal Bovine Serum (FBS)

Integrated DNA Technologies, Inc., USA

- Alt-R™ CRISPR-Cas9 sgRNA "sgGFP" (MW 32,405.9): 5'-mG*mA*mC*CAGGAUGGGCACCACCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U-3' ('m_*' = phosphorothioated 2'-O-methyl ribonucleotide)

Jena Bioscience GmbH, Germany

- XbaI, JBSpeed Restriction Enzyme, 10 units/µL (supplied with 10x Universal Buffer)
- Taq Core Kit

Microsynth AG, Switzerland

- pEGFP-Luc forward primer: 5'-CGCCAATAGGGACTTTCCATTG-3'
- pEGFP-Luc reverse primer: 5'-GATAAATAACGCGCCCAACACC-3'

New England Biolabs®, USA

- EnGen® Spy Cas9 NLS, 20 µM (supplied with 10x NEBuffer™ r3.1)

Polysciences, Inc., USA

- Polyethylenimine, branched (30% soln. in water), MW 70 kDa
- Polyethylenimine, MW 10 kDa

Promega Corporation, USA

- BenchTop 1kb DNA Ladder
  BenchTop 100 bp DNA Ladder

Roche Diagnostics International AG, Switzerland (subsidiary of Roche Holding AG, Switzerland)

- cOmplete™ protease inhibitor cocktail tablets, EDTA-free
- DNase I, from bovine pancreas
- RNase A, from bovine pancreas

Sigma-Aldrich®, USA (affiliate of Merck KGaA, Germany)

- Agarose
- Amicon® Ultra-4 Centrifugal Filter Units, NMWL 50 kDa, 4 mL sample volume
- Amicon® Ultra-15 Centrifual Filter Units, NMWL 100 kDa, 15 mL sample volume
- Ampicillin sodium salt
- Bromophenol blue
- Boric acid
- Chloramphenicol
- Coomassie® Brilliant blue G 250
- D(+)-Glucose monohydrate
- Deuterium oxide
- Dialysis tubing cellulose membrane, MWCO 14 kDa, 25 mm
- DL-Dithiothreitol (DTT)
- GelRed® Nucleic Acid Stain 10000X Water
- GenElute™ PCR Clean-Up Kit
- Glycerol 99%, puriss.
- HiLoad® 16/600 Superdex® 200 pg, Cytiva™
- His Buffer Kit, Cytiva™
- His GraviTrap™ (CV 1 mL), Cytiva™
- HisTrap™ High Performance (CV 5 mL), Cytiva™
- Imidazole
- Isopropyl β-D-1-thiogalactopyranoside
- Kanamycin sulfate
- L-Glutamine
- LB agar, Miller
- LB Broth (Lennox)
- Lysozyme, from chicken egg white
- Magnesium chloride hexahydrate
- Methanol
- Novagen® BL21(DE3) Singles™ Competent Cells
- Novagen® BL21(DE3)pLysS Competent Cells
- N,N,N',N'-Tetramethylethylenediamine (TEMED)
- Phenylmethylsulfonyl fluoride (PMSF)
- Polyethylenimine, branched, average MW 25 kDa
- Potassium chloride
- Propidium iodide, ≥94.0%
- Potassium phosphate dibasic
- Potassium phosphate monobasic
- Spectra/Por® Float-A-Lyzer® G2 Dialysis Device, MWCO 100 kDa, 10 mL volume
- Succinic anhydride
- TRIS hydrochloride
- Trypsin-EDTA solution

Takara Bio Inc., USA (former Clontech Laboratories, Inc., USA)

- pEGFP-Luc Vector

Thermo Fisher Scientific Inc., USA

- FastDigest HindIII, Restriction Enzyme
- GeneJET Plasmid Miniprep Kit
- Gibco^{TM} DMEM (1X) with low glucose (1 g/l), L-glutamine and pyruvate

2.2. Buffers

All buffers were prepared with demineralized distilled purified water.

Cas9 Storage Buffer *old* pH 7.4	• 300 mM NaCl • 20 mM TRIS-HCl • 0.1 mM EDTA • 1 mM DTT
Cas9 Storage Buffer *new* pH 7.4	• 300 mM NaCl • 20 mM HEPES • 0.1 mM EDTA • 1 mM DTT
Dialysis HEPES Buffer pH 7.4	• 20 mM HEPES • 200 mM KCl • 10 mM $MgCl_2$ hexahydrate • 1 mM DTT
HEPES buffered glucose solution (HBG)	• 20 mM HEPES • 5% glucose
HisTrap Buffer A pH 7.4	• 20 mM TRIS • 500 mM NaCl • None or 20 mM imidazole
HisTrap Buffer B pH 7.4	• 20 mM TRIS • 500 mM NaCl • 500 mM imidazole

Laemmli Buffer 10x pH 8.4-8.9	• 0.25 M TRIS • 1.92 M glycine • 1% SDS
Loading Buffer 6x for gel electrophoresis	• 8.21 M glycerol • 0.06 M EDTA (using 0.5 M EDTA, pH 8.0) • 0.003 M bromophenol blue (free acid)
Loading Buffer for SDS-PAGE	• 0.02 % w/v bromophenol blue • 0.1 M TRIS, pH 6.5 • 20% w/v glycerol • 100 mM DTT • 4% w/v SDS
Lysis Buffer pH 7.5	• 20 mM TRIS • 20% Sucrose • 0.2 M NaCl • 10 mM $MgCl_2$
PBS stock solution 10x pH 7.4	• 80 g/l NaCl • 2 g/l KCl • 7.65 g/l $Na_2HPO_4 \bullet 2\ H_2O$ • 2 g/l KH_2PO_4
PBS-EDTA	• PBS 1x • 0.1 % EDTA
TBE Buffer 10x (storage at room temperature)	• 0.89 M TRIS base • 0.89 M boric acid • 0.02 M EDTA (using 0.5 M EDTA, pH 8.0)

2.3. Restriction digest of plasmids

Resulting in a total volume of 20 µL, the following substances were pipetted into an Eppendorf tube in the order given: sterile nuclease-free water, 10x universal buffer from Jena Bioscience (final conc. 1x), pDNA (final conc. 12.5 ng/µL) and 10 units/µL restriction enzyme solution of XbaI (final conc. 0.25 units/µL). As a comparison, a 20 µL blank consisting of sterile nuclease-

free water and pDNA (final conc. 12.5 ng/µL) was prepared. After gentle mixing, the solutions were incubated at 37 °C for at least one hour.

Meanwhile, 1% agarose was dissolved in 1x TBE Buffer by microwave heating at 600-800 W. During the process of cooling down, the solution was regularly swirled and GelRed® 1000x in nuclease-free water (final conc. 1x) was added at 50 °C. Once homogenized, the 1% agarose gel was poured and left at room temperature for 20-30 min. The solidified agarose gel was put into a gel box filled with 1x TBE Buffer and was loaded with the BenchTop 1kb DNA Ladder and the pDNA samples. To the latter, 6x Loading Buffer was added (final conc. 1x), before gel electrophoresis was run at 75 V. Subsequently, the gel was exposed to UV light and images of the DNA fragments, visualized by emitted fluorescence, were captured by the Quantum gel documentation system (Vilber Lourmat, France).

2.4. Transformation of *Escherichia coli* with plasmid DNA

For efficient transformation, Novagen® Competent Cells Singles™ kits of the *Escherichia coli* strain BL21(DE3) were used. Purified pET28a-Cas9-His DNA (Addgene plasmid #98158, with a kanamycin resistance marker) [64] was added to the *E. coli* cells, which were then treated pursuant to the enclosed transformation protocol for Singles™ kits. The only change made was the replacement of SOB medium with LB medium (20 g LB Broth Lennox and 5 g NaCl solved in 1 L distilled water, then autoclaved). To select the bacteria that successfully took up the plasmid, the cell suspension was plated on 10 cm kanamycin-containing LB agar plates (40 g LB agar Miller solved in 1 L distilled water, then autoclaved, with kanamycin (final conc. 50 µg/mL) added at 50 °C). After incubation at 37 °C overnight, a single colony was picked to inoculate 5.5 mL LB medium that also contained kanamycin (final conc. 50 µg/mL). In this liquid culture, the bacteria were grown overnight at 37 °C on a Thermo Scientific™ MaxQ™ 4000 shaking incubator (Thermo Fisher Scientific Inc., USA) at 150 rpm with loosely sitting caps to provide enough oxygen. The next day, bacterial glycerol stocks of 1 mL were created by gently mixing 50% glycerol solution with the liquid culture in a ratio of 1:1 (v/v) and subsequent freezing in liquid nitrogen at -160 °C. The glycerol stocks of the transformed *E. coli* were stored at -80 °C.

2.5. Isolation of plasmids

To harvest the transformed *E. coli*, the 5 mL overnight bacterial culture was centrifuged for 10 min at 3,000 rpm and 4 °C using the Sorvall RT7 Refrigerated Benchtop Centrifuge (Thermo Fisher Scientific Inc., USA). After discarding the supernatant, the instructions of the GeneJET Plasmid Miniprep Kit were obeyed to recover the pDNA from the cell pellet. For this purpose, a resuspension solution, lysis solution and neutralization solution were added. Centrifugation was performed at 13,000 rpm and room temperature using the MiniSpin® microcentrifuge (Eppendorf SE, Germany), as were the subsequent centrifugation steps. The supernatant and a wash solution were applied on the GeneJET Spin Column. For elution of the purified DNA, autoclaved nuclease-free water was used instead of the elution buffer of the kit. The plasmid concentration in the collected flow-through was measured against a blank at a wavelength of 260 nm with factor 50.0 (absorption of 1.0 equals 50 µg/mL) by the UV/Vis spectrophotometer Lambda XLS (PerkinElmer, Inc., USA).

2.6. Expression of recombinant Cas9-His

The expression protocol was first tested by expressing recombinant EGFP for proof of concept and, after appropriate adjustments, later performed with recombinant Cas9. For these purposes, the following transformed Novagen® Competent Cells were used:

- *E. coli* strain BL21(DE3)pLysS (with a chloramphenicol resistance marker) containing the EGFP-encoding plasmid pET23a(+)-SV40nls-EGFP,
- *E. coli* strain BL21(DE3) containing the Cas9-His (with a C terminal His_6 tag on the backbone) encoding plasmid pET28a-Cas9-His (Addgene plasmid #98158, with a kanamycin resistance marker). [64]

The methods used for cell disruption and protein purification differed depending on the batch. (a) The EGFP and the first Cas9-His batch are obtained by a manual approach, whereas (b) the second and third Cas9-His batch are yielded using automated systems. The headings are thus marked accordingly with (a) or (b).

2.6.1. Protein expression and cell lysis

After inoculation, the transformed bacteria mentioned above were cultivated at 37 °C in (5 mL) LB medium (20 g LB Broth Lennox and 5 g NaCl solved in 1 L distilled water, then autoclaved) containing appropriate antibiotics (100 µg/mL ampicillin and 50 µg/mL chloramphenicol for pET23a(+)-SV40nls-EGFP; 50 µg/mL kanamycin for pET28a-Cas9-His) with constant shaking at 150 rpm in the shaking incubator to grow overnight. The liquid culture was diluted 1:100 in LB medium with appropriate antibiotics, then grown at 37 °C and 150 rpm until an optical density of 0.6-0.75 was reached at 600 nm (OD_{600} of 1.0 equals approximately 10^9 *E. coli* cells per mL), measured by the UV/Vis spectrophotometer. Subsequently, the bacterial culture was brought to room temperature and induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG, final conc. 1 mM), followed by incubation at 22 °C and 150 rpm overnight.

The next day, the cells were harvested by centrifugation for 1 hour at 4,000 xg using the Sorvall LYNX 6000 Superspeed Centrifuge (Thermo Fisher Scientific Inc., USA). The pellet was resuspended in (50 mL) lysis buffer. RNase (final conc. 10 µg/mL), DNase (final conc. 30 µg/mL), lysozyme (final conc. 1 mg/mL) and phenylmethylsulfonyl fluoride (PMSF, final conc. 1 mM) were added. For improvement, cOmplete™ protease inhibitor cocktail (1 tablet/50 mL extraction solution) was used alternatively to PMSF. 25 mL aliquots were frozen in 50 mL falcons in liquid nitrogen, then stored at -20 °C with parafilm wrapped around the cap.

2.6.2. a) Cell disruption by means of ultrasound

The bacterial suspension was thawed at 25 °C and always kept on ice in between steps. Each aliquot of bacterial suspension was sonicated by the Sonopuls™ HD 2070 ultrasonic homogenizer (BANDELIN electronic GmbH & Co. KG, Germany) at 100 % power three times for 20 sec while on ice. The lysate was transferred into ultracentrifugal tubes and centrifuged for 1 hour at 20,000 rpm and 4 °C. The supernatant was filtered through a 0.45 µm syringe filter, then stored at -20 °C or preferably further processed immediately.

2.6.3. b) Cell disruption by means of high-pressure homogenization

After thawing at 25 °C, the bacterial suspension was homogenized three times under high pressure of 7,000 psi by the LM10 Microfluidizer® (Microfluidics International Corporation, USA) while on ice. The lysate was transferred into ultracentrifugal tubes and centrifuged for 45 min at 14,000 rpm and 4 °C. The supernatant was filtered through a 0.45 µm syringe filter,

then stored at -20 °C or preferably further processed immediately. In between these steps, the protein solution was always kept on ice.

2.6.4. a) Protein purification by manual affinity chromatography

The Cytiva His Buffer Kit and the gravity-flow column Cytiva His GraviTrap™ (column volume (CV) of 1 mL) were used, following the instructions of use accordingly. For the expression of EGFP, a set of phosphate buffers with increasing concentrations of imidazole (20, 40, 60, 80, 100, 250, 500 mM) was prepared using the Cytiva His Buffer Kit, whereas a set of TRIS buffers (20 mM TRIS, 500 mM NaCl, pH 7.4) with imidazole concentrations increasing in smaller steps (20, 30, 40, 50, 60, 80, 100, 250, 500 mM) was fabricated for the expression of Cas9-His. The column was washed with 6 CVs 70% EtOH and 10 CVs distilled water, then equilibrated with 10 CVs binding buffer containing 20 mM imidazole. Before the protein solution was applied to the column, imidazole (final conc. 20 mM) was added to the solution. After subsequent rinsing of the column with 15 CVs binding buffer, stepwise elution was performed with 5 CVs of each buffer in order of increasing imidazole concentrations. Finally, the column was washed with 10 CVs distilled water and 4 CVs 70% EtOH, then stored in 6 CVs 70% EtOH.

Throughout these steps, the flow-through fraction, the wash fraction and the elution fractions were collected to conduct multi wavelength measurements at 214, 260 and 280 nm using the UV/Vis spectrophotometer. For blank values, the buffers containing the corresponding imidazole concentration were measured. By applying the Beer-Lambert law, the protein concentrations of the fractions were determined with a molar absorption coefficient of 55,000 M^{-1} cm^{-1} for EGFP and 120,450 M^{-1} cm^{-1} for Cas9-His.

2.6.5. b) Protein purification by automated affinity chromatography

The Cytiva HisTrap™ High Performance column (CV of 5 mL) was connected to the ÄKTA pure™ chromatography system (Cytiva, Danaher Corporation, USA) situated in the Unichromat 1500 (UniEquip GmbH, Germany), which provided a constant temperature of 4°C. The flow rate was set to 5 mL/min, the pressure limit to 0.50 MPa pre- and 0.30 MPa delta-column. The system was purged with distilled water, then column and injection lines were washed with HisTrap Buffer B followed by HisTrap Buffer A. After adding imidazole (final conc. 20 mM) to the protein solution, it was applied onto the column via pump injection. The sample application was finalized with 10 mL HisTrap Buffer A, then the column was washed with 5 CVs of the same buffer. Gradient elution was performed in three to four elution steps of 5 CVs

each by increasing the imidazole concentration. By changing the ratio of HisTrap Buffer A and B, the imidazole concentrations were adjusted to 40, 200 and 500 mM in the first run (batch 2) and to 40, 60, 250 and 500 mM during the second run (batch 3). Throughout this process, the flow-through fraction, the wash fraction as well as the eluates were collected by the fraction collector and their absorption was detected.

2.6.6. a) Protein purification by Dialysis

After pooling purified fractions containing the proteins of interest, dialysis was performed at 4 °C overnight under constant stirring, followed by two additional dialysis steps for 1-5 hours each. EGFP solutions were dialyzed against PBS Buffer using dialysis tubing cellulose membranes (MWCO 14,000 Da) that required pre-soaking, whereas Cas9-His solutions were dialyzed against HEPES Buffer using Spectra/Por® Float-A-Lyzer® G2 Dialysis Devices (MWCO 100 kDa, 10 mL volume).

2.6.7. b) Protein purification by Size Exclusion Chromatography

Pooled Cas9-His fractions were concentrated with Amicon Ultra centrifugal filter units (MWCO 50,000 or 100,000) by centrifugation at 4°C and 4000 xg using the Benchtop Centrifuge 5810 R (Eppendorf SE, Germany) until the maximum volume of the injection loop's capacity (5 mL) was reached. Subsequently, size exclusion chromatography was conducted at 4°C using the ÄKTA pure™ chromatography system, connected to a HiLoad® 16/600 Superdex® 200 pg column and refrigerated by the Unichromat 1500. The column was washed with distilled water followed by Cas9 Storage Buffer. To rinse the injection loop, 25 mL Cas9 Storage Buffer were used. The sample was then injected into the loop and applied to the column. Via isocratic elution, the protein sample was separated according to size after applying 1 CV of Cas9 Storage Buffer. The eluted fractions were collected, while their absorption was simultaneously detected photometrically.

2.6.8. SDS-PAGE

Purity of the collected fractions was monitored by SDS-PAGE. The running gel was composed of 4 mL distilled water, 2.5 mL 1.5 M TRIS-HCl pH 8.8, 3.3 mL 30% acrylamide-bis, 100 µL 20% SDS, as well as, subsequent to brief vortexing, 33 µL TEMED and 17 µL 25% APS. The stacking gel consisted of 2.425 mL distilled water, 940 µL 0.5 M TRIS-HCl pH 6.8, 625 µL 30% acrylamide-bis, 10 µL 20% SDS, as well as, after brief vortexing, 7.5 µL TEMED and 6 µL 25% APS.

The samples were mixed with Loading Buffer in a ratio of 1:1 (v/v) and incubated at 95 °C using the Eppendorf® Thermomixer Compact (Eppendort SE, Germany). Along with Precision Plus Protein™ Standards, they were loaded into the wells of the gel, which was subsequently run at 30 mA per gel in 1x Laemmli Buffer for 40-45 min. To visualize the bands, the gel was stained with Coomassie protein staining solution (0.3 % w/v Coomassie® Brilliant blue, 50% v/v methanol, 10% v/v acetic acid, in distilled water) for 30-60 min, then incubated in destaining solution (50% v/v methanol, 10% v/v acetic acid, in distilled water) and finally stored in 5% v/v acetic acid (in distilled water).

2.6.9. Concentrating

The pure Cas9-His fractions were concentrated with Amicon Ultra centrifugal filter units (MWCO 50,000 or 100,000 Da) by centrifugation at 4°C and 3,000 xg using the Sorvall RT7 Refrigerated Benchtop Centrifuge (Thermo Fisher Scientific Inc., USA). After addition of glycerol (final conc. 20%) and quantification using the UV/Vis spectrophotometer, the protein solution was aliquoted and frozen in liquid nitrogen, then stored at -80 °C.

2.7. In vitro cleavage assay

2.7.1. Linearization of the target plasmid

Restriction digest of the target plasmid pEGFP-Luc (6367 bp), encoding a fusion protein of EGFP and the firefly *Photinus pyralis* luciferase, [65] was realized according to 2.3, however, applying deviating reaction conditions. At first, pDNA (final conc. 0.33 µg/µL) and restriction enzyme XBaI (final conc. 0.66 units/µL) were incubated in 30 µL for 2 hours at 37 °C. In another approach, the plasmid pEGFP-Luc was purified beforehand with the GeneJET Plasmid Miniprep Kit as described in 2.5. The purified pDNA (final conc. 0.02 µg/µL) and restriction enzyme XBaI (final conc. 0.2 units/µL) were then incubated in 250 µL overnight at 37 °C.

2.7.2. PCR of the target plasmid

In order to amplify a specific gene sequence of the plasmid pEGFP-Luc that includes the EGFP gene, referred to as “IVC EGFP gene (1.5 kb)”, the following PCR primers of 22 nucleotides in length with 50% G/C content, an annealing temperature of 53 °C and similar melting temperatures (T_m) were designed:

Forward primer: 5'-CGCCAATAGGGACTTTCCATTG-3' (T_m 58 °C)

Reverse primer: 5'-GATAAATAACGCGCCCAACACC-3' (T_m 59 °C)

The Jena Bioscience Taq Core Kit test setup was carried out according to the enclosed data sheet: a premix for 12 assays of 20 µL each was prepared by combining sterile nuclease-free water, 10x PCR Buffer (final conc. 1x), 10 mM dNTP mix (final conc. 200 µM), forward and reverse primer (final conc. 300 nM each), 5 units/µL Taq DNA polymerase (final conc. 0.025 units/µL) as well as the non-linearized template DNA pEGFP-Luc (final conc. 0.25 ng/µL) in the order given. After aliquoting 20 µL each into PCR tubes, the tubes were placed into the MJ Research PTC-200 Peltier Thermal Cycler (Bio-Rad Laboratories, Inc., USA). A gradient PCR with varying annealing temperatures of 50.0, 50.3, 50.9, 52.0, 53.4, 55.2, 57.2, 58.9, 60.2, 61.1, 61.8 and 62.0 °C was performed to determine the appropriate one. The thermocycling conditions were set at 95 °C for 2 min for initial denaturation followed by 30 cycles consisting of denaturation at 95 °C for 20 sec, annealing at 50-62 °C for 20 sec and elongation at 72 °C for 90 sec, with the exception of the final elongation step lasting for 10 min. To verify the success of the PCR, samples of the PCR products obtained were mixed with 6x Loading Buffer (final conc. 1x) and, together with the BenchTop 100 bp DNA Ladder, loaded on a 1% agarose gel, which was run at 75 V.

The PCR product with an annealing temperature of 60.2 °C was selected to be purified with the GenElute™ PCR Clean-Up Kit. The preparation instructions and procedure steps of the enclosed technical bulletin were followed, using the MiniSpin® microcentrifuge (Eppendorf SE, Germany) with a maximum speed of 14,000 xg as well as sterile nuclease-free water for elution. Subsequently, the purified IVC EGFP gene (1.5 kb) was subjected to 36 PCR cycles as described above, but with an annealing temperature of 60.2 °C for the entire batch. Concentration of the IVC EGFP gene (1.5 kb) samples was measured at 260 nm with factor 50.0 using the UV/Vis spectrophotometer.

2.7.3. In vitro cleavage assay

The activity of Cas9 protein batches was verified or compared via in vitro cleavage (IVC) assays. Here, the commercially purchased EnGen® Spy Cas9 NLS (20 µM), which was supplied with 10x NEBuffer™ r3.1, in combination with sgRNA served as positive control, whereas samples without sgRNA served as negative control. To achieve cleavage of the linearized pEGFP-Luc (6.4 kb) or the IVC EGFP gene (1.5 kb), Cas9 was combined with the sgRNA “sgGFP” that directs

the endonuclease to the target site of the EGFP gene. The sgGFP sequence reads as shown below ('m_*' = phosphorothioated 2'-O-methyl ribonucleotide).

5'-mG*mA*mC*CAGGAUGGGCACCACCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U-3'

Appropriate stock solutions of Cas9, sgGFP and target DNA were prepared and pipetted together with 10x NEBuffer™ r3.1 and nuclease-free water to create reaction volumes of 30 µL according to the following scheme:

Reaction conditions:	- sgRNA	+ sgRNA	final conc.
Sterile nuclease-free water	24 µL	21 µL	
10x NEBuffer™ r3.1	3 µL	3µL	1x
sgRNA (20 ng/µL)	/	3 µL	2 ng/µL
Cas9 nuclease (1µM = 150 ng/µL)	1 µL	1 µL	5 ng/µL
Pre-incubation for 15min at 37 °C			
Target DNA (150 ng/µL)	2 µL	2 µL	10 ng/µL
Total reaction volume:	30 µL	30 µL	

The final reaction solutions were incubated for 2 hours at 37 °C before adding 6x Loading Buffer for gel electrophoresis (final conc. 1x). Together with the BenchTop 1kb DNA Ladder, all samples were loaded on a 1% agarose gel, which was run at 75 V.

The IVC assay was performed after 1, 3 and 5 freeze-thaw cycles, which were normed to cycles of 5 min on ice, thawing, 3 min at room temperature, 10 min on ice again and subsequent re-freezing.

2.8. Synthesis of delivery carrier materials

2.8.1. 10% Succinylation of 10 kDa, 25 kDa and 70 kDa polyethylenimines

Three types of branched polyethylenimines (PEIs) that differed in average molecular weight (10 kDa, 25 kDa, 70 kDa) were 10% succinylated by the following procedure, assuming every single amine can react (molar ratio of 1:10 based on amines). [52]

Once dissolution of 500 mg PEI (≈ 11.64 mmol amines) in 8.5 mL purified distilled water and 1.5 mL 3 M NaCl was achieved by means of vortexing, 37% HCl was added until a pH of 5 was reached. To this PEI solution, a solution of 116.5 mg succinic anhydride (≈ 1.164 mmol) in 220 µL DMSO was added dropwise, while stirring continuously to avoid precipitation. The resulting reaction mix was stirred for 3 hours at room temperature. Next, dialysis was first carried out against 5 L 0.25 M NaCl overnight, then against 5 L of purified distilled water for 12 hours and again overnight, using dialysis tubing cellulose membranes (MWCO 12,000 or 14,000 Da). The 10% succinylated PEI (PEI-Suc 10%) was obtained by lyophilization. Via ^{1}H-NMR, the percentage of succinylation was determined in deuterium oxide by a 400 MHz NMR spectrometer (Ascend™ 400 magnet system, Avance III console, SampleXpress autosampler, BBFO-PLUS probe, Bruker, USA).

2.8.2. Synthesis of oligopeptide UV_1

Based on the results and methods of Kuhn et al. [53] as well as Lin et al. [66], the artificial oligopeptide “UV_1” was synthesized via solid-phase synthesis and provided by Julian Geritzer BSc.

2.9. Preparation of ribonucleoprotein carriers

2.9.1. Preparation of Cas9 PEI-Suc 10% ribonucleoprotein carriers

For formation of Cas9 ribonucleoprotein (RNP) complexes, a 1.5 µM Cas9-His and a 1.5 µM sgGFP solution were prepared in HBG and mixed 1:1 (v/v) to achieve an equimolar RNP ratio (Cas9 to sgRNA, final conc. 0.75 µM). During incubation for 15 min at room temperature, a dilution series of PEI-Suc 10% in HBG was fabricated corresponding to w/w ratios (polymer to sgRNA) of 4:1, 2:1, 1:1 and 0.5:1 in the final reaction mix. Subsequently, 10 µL 0.75 µM RNP complex solution was added to 10 µL PEI solution and immediately mixed thoroughly by pipetting up and down, followed by incubation for 15 min at room temperature to enable PEI-Suc 10% RNP particle formation. Application in cell assays required preparation in a sterile environment.

2.9.2. Preparation of Cas9 oligopeptide ribonucleoprotein carriers

For formation of Cas9 RNP complexes, a 1.5 µM Cas9-His solution and three sgGFP solutions of 1.5, 2.25 and 3 µM were prepared in HBG. By mixing the Cas9-His solution with one of the

sgGFP solutions in equal proportions (v/v), corresponding molar RNP ratios of 1:1, 1:1.5 or 1:2 (Cas9 to sgRNA, final conc. 0.75 µM to 0.75, 1.125 or 1.5 µM) were achieved. During incubation for 15 min at room temperature, a dilution series of the oligopeptide UV_1 in HBG was fabricated corresponding to N/P ratios (oligopeptide nitrogen to sgRNA phosphate) of 3, 6, 12, 16 and 24, respectively. Subsequently, 10 µL RNP complex solution was added to 10 µL UV_1 solution and immediately mixed thoroughly by pipetting up and down, followed by incubation for 30 min at room temperature to enable UV_1 RNP particle formation. Application in cell assays required preparation under sterile conditions.

2.9.3. Measurement of particle size and zeta potential

Particle size and zeta potential were measured promptly after particle production. To each 20 µL Cas9 RNP carrier sample, 80 µL HBG was added. The resulting total volume of 100 µL (final conc. 75 nM RNP at an equimolar RNP ratio) was filled into folded capillary cells DTS 1060 (Malvern Panalytical Ltd, UK, Netherlands). Hydrodynamic particle size of the Cas9 RNP carriers was determined by the Zetasizer Nano ZS (Malvern Panalytical Ltd, UK, Netherlands) in triplicate measurements with 11 runs each at a fixed backscatter angle of 173°. The temperature was set to 25 °C, the viscosity to 0.8872 cP and the refractive index of the solvent to 1.330. Next, the capillary cells were additionally filled with 700 µL each of a 10 mM NaCl solution, and zeta potential was measured under the same conditions as stated above.

2.10. Cell culture

To analyze the uptake of Cas9/sgGFP RNP nanocarriers into the nucleus, HeLa GFP-Tub cells were used, an immortalized human epithelial cell line of cervical carcinoma with stable expression of the fusion protein GFP tubulin. The cells were cultivated in 10 mL Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10% FBS, 0.1% gentamycin and 12 mg/l L-glutamine, in 75 cm^2 cell culture flasks (Greiner Bio-One International GmbH, Austria) at 37 °C in a humidified CO_2 incubator which provided an atmosphere of 5% CO_2/95% air and a relative humidity of 96%. When reaching a confluence of approximately 70-80%, sub-cultivation was performed under sterile conditions with reagents pre-warmed to 37 °C. For this purpose, spent culture medium was removed by suction. The cell layer was washed using 2 mL PBS-EDTA and then incubated with 2 mL Trypsin-EDTA at 37 °C to detach the adherent cells from the culture surface, promoted by tapping the flask. Trypsinization was stopped by

adding 10 mL of the serum-containing DMEM. The resulting cell suspension was centrifuged at 1,000 rpm for 5 min, forming a cell pellet that was resuspended with 2 mL DMEM. In the culture flask, an aliquot of three drops was seeded in 10 mL of fresh growth medium.

The cryopreservation medium comprised 80% DMEM, 10% FBS and 10% DMSO.

2.11. Cellular delivery of genome editing ribonucleoproteins

HeLa GFP-Tub cells, cultured for up to 30 passages, were seeded in a 96-well microplate (Greiner Bio-One International GmbH, Austria) at a density of 5,000 cells per well, with each well containing a volume of 100 µL DMEM, additionally enriched with 10% FBS, 0.1% gentamycin and 12 mg/L L-glutamine. Subsequent to incubation at 37 °C for 24 hours, the medium was replaced with 80 µL fresh medium, then 20 µL particle solution (freshly prepared according to 2.10) or 20 µL HBG for negative control were added. Ensuing mixing with the medium was standardized to pipetting up and down three times. The cells were treated for 48 hours at 37 °C, forming an 80-90% confluent cell monolayer. In each well, the cells were detached by removing the old medium, washing with 20 µL PBS-EDTA, adding 20 µL Trypsin-EDTA, incubating for 2 min at 37 °C, tapping, stopping the trypsinization with 100 µL serum-containing DMEM and suspending the cells in said medium. The resulting cell suspension of 120 µL per well was transferred into a 24-well microplate (Greiner Bio-One International GmbH, Austria) already pre-filled with 880 µL DMEM per well. The plate was incubated for 72 hours. In each well, the old medium was subsequently removed, the cells were washed with 100 µL PBS-EDTA and treated with 100 µL Trypsin-EDTA for 2 min at 37 °C for dissociation. 500 µL PBS-EDTA with 2% FBS were added per well to stop trypsinization and suspend the cells via pipetting up and down. The cell suspensions were transferred into flow cytometry tubes and put on ice.

Genome editing Cas9/sgGFP RNP uptake into the nucleus of HeLa GFP-Tub cells was evaluated using a Gallios Flow Cytometer (Beckman Coulter, Inc., USA) equipped with an integrated blue laser and a Gallios 561 nm laser option (Beckman Coulter, Inc., USA). Before analysis, 1.2 µL of 1.5 mM propidium iodide stock solution was added to the cell suspension of approximately 600 µL (final conc. 3 µM) and incubated for 10 min. The cytometer was set to detect a cell count of 5,000 cells per sample and determine characteristics of each cell with the fluorescence filters 1-5 at a forward scatter of 50, sideward scatter of 150, and 350 V. As EGFP

exhibits fluorescence excitation/emission maxima at 488/507 nm and propidium iodide bound to nucleic acids at 535/617 nm, blue laser light excited EGFP to emit green fluorescence, while propidium iodide that had permeated dead cells was excited by green laser light to emit red fluorescence. For calculation of the EGFP knockout percentage, dead cells were excluded and the percentage decrease of live green fluorescent cells after treatment with Cas9/sgGFP RNP nanocarriers was compared with the negative control.

3. Results and Discussion

3.1. Restriction digest of plasmids, transformation of *Escherichia coli* with plasmid DNA and isolation of plasmids

Before transforming the *E. coli* strain BL21(DE3) with pET28a-Cas9-His (Addgene plasmid #98158), [64] restriction digest of the plasmid was carried out using the restriction enzyme XbaI:

- **XbaI**: recognizes the depicted restriction site and cuts pET28a-Cas9-His after bp 4368 in 5'→3' direction.

5' – T^CTAG A – 3'
3' – A GATC_^T – 5'

The plasmid pET28a-Cas9-His is 9387 bp in size and has one recognition site for XbaI, located at 4368-4373 bp. By linearizing the plasmid in question with restriction enzyme XbaI, its identity was checked based on the size. For visualization, the pDNA was separated on a 1% agarose gel stained with GelRed®, a more sensitive and less toxic alternative to ethidium bromide, that interacts with nucleic acids by intercalation. When bound to DNA and excited with UV light, GelRed® emits fluorescence of orange colour, enabling the detection of DNA.

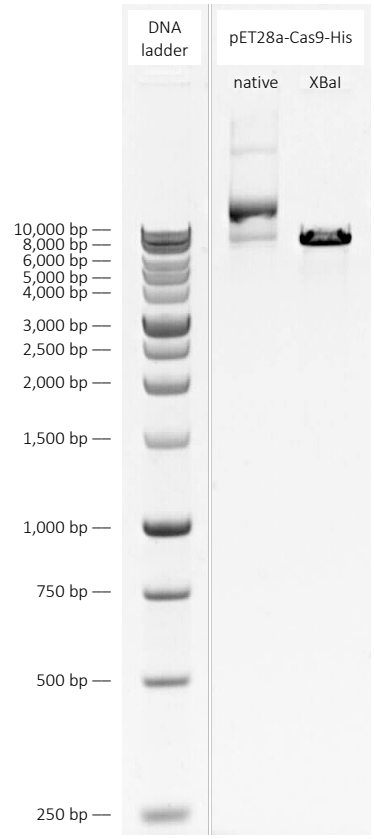

***Figure 1**: Agarose gel post electrophoresis showing plasmid pET28a-Cas9-His in its native form (lane 2) as well as in linearized form after XbaI restriction digest (lane 3)*

Uncut plasmids exist in different conformations, usually appearing as two bands of supercoiled and nicked DNA in the agarose gel. Plasmids with a supercoiled conformation are more compact, therefore experience less resistance and migrate faster in agarose gels. Nicked plasmids on the other hand, obtained by in vivo topoisomerases or harsh purification conditions, are less compact and relatively large in comparison because the superhelical tension is relaxed. Encountering greater resistance, nicked plasmids run slower in agarose gels.

As seen in **Figure 1** lane 2, the undigested plasmid pET28a-Cas9-His showed three bands: The top band occured at low intensity and migrated the least far, indicating that it was nicked

pDNA. The middle band was present with high intensity and migrated further, in all probability being supercoiled pDNA, as this is the native conformation of plasmids in vivo. The bottom band appeared in medium intensity with a size of approximately 9000 bp, which accorded with the linearized pDNA of 9387 bp in length. In native plasmid samples, linearized pDNA is a result of double strand breaks caused by unfavorable conditions (such as spinning or vortexing for too long). Lane 3, in contrast, contained the plasmid cleaved with the single cutter restriction enzyme XBaI and thus exhibited a single band of linearized pDNA, which was consequently the most pronounced.

After this verification of identity, the *E. coli* strain BL21(DE3) was transformed with the plasmid pET28a-Cas9-His. Upon completed transformation, two single colonies were picked for propagation and ensuing plasmid isolation. Colony number 2 yielded a higher concentration of the isolated pDNA and was therefore selected for Cas9-His expression.

3.2. Expression of recombinant proteins

3.2.1. Expression of EGFP

Using the pET23a(+)-SV40nls-EGFP transformed *E. coli* strain BL21(DE3)pLysS, the suitability of the expression protocol was first tested by expressing enhanced GFP. Because of its green fluorescence when exposed to UV light, the workflow could easily be visualized, as shown in **Figure 2**.

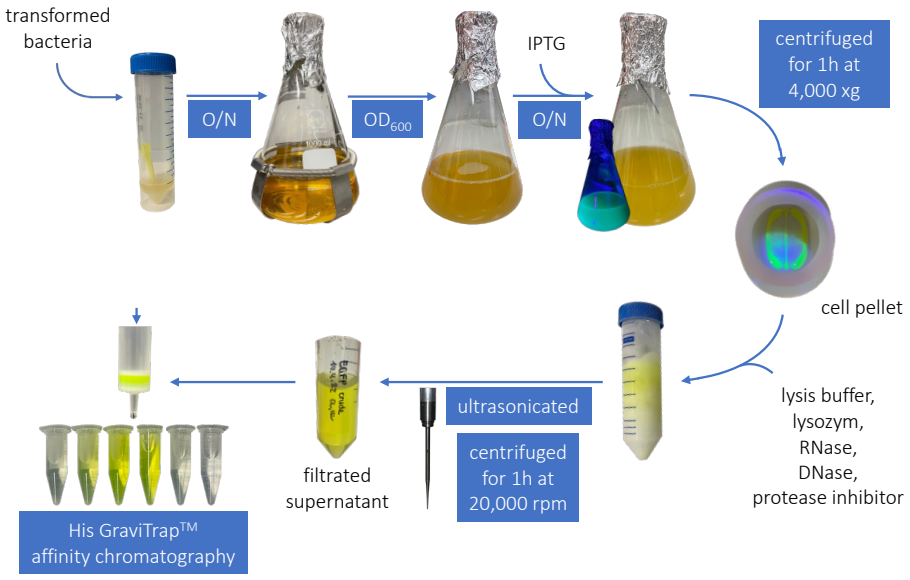

***Figure 2**: The workflow of protein expression is visualized by means of EGFP up to the purification via gravity-flow column. Exposure to UV light enabled the identification of the presence of EGFP proteins that consequently emitted green fluorescence, thus showing that the expression of EGFP proteins and purification by affinity chromatography worked.*

By observing the optical density, bacterial growth in the LB medium was already visible to the naked eye after overnight incubation of the transformed bacteria in 5 mL LB medium as well as while growing in 500 mL LB medium. After addition of IPTG to induce protein expression upon reaching an OD_{600} of 0.6-0.75 and subsequent O/N incubation, the presence of EGFP was unambiguously determined by exposure of the bacterial suspension to UV light, as green fluorescence was consequently emitted. Since EGFP was expressed intracellularly, ensuing centrifugation and removal of the LB medium was performed to obtain only the cells. As expected, EGFP was detected by means of UV light in the remaining cell pellet. Subsequent to cell lysis and disruption, the cell lysate was centrifuged and the cell fragments were discarded. The EGFP-containing supernatant was filtered through a 0.45 µm filter to remove any remaining live E. coli cells or large cell fragments. This filtered supernatant, termed "EGFP crude", exhibited a yellowish green color when exposed to light within the visible spectrum, indicating the presence of EGFP proteins. By applying the crude EGFP solution to the His GraviTrap™ gravity-flow columns, the EGFP proteins were purified by immobilized metal affinity chromatography. The Ni^{2+} ions immobilized by chelation were bound by the histidine-tagged EGFP proteins, which were washed while bound and then eluted by increasing concentration of imidazole, the competitive eluant. The yellowish green color of the eluates obtained at concentrations of 60, 80 and 100 mM imidazole implied the presence of EGFP proteins, which was confirmed by the photometric results, as shown in **Figure 3**.

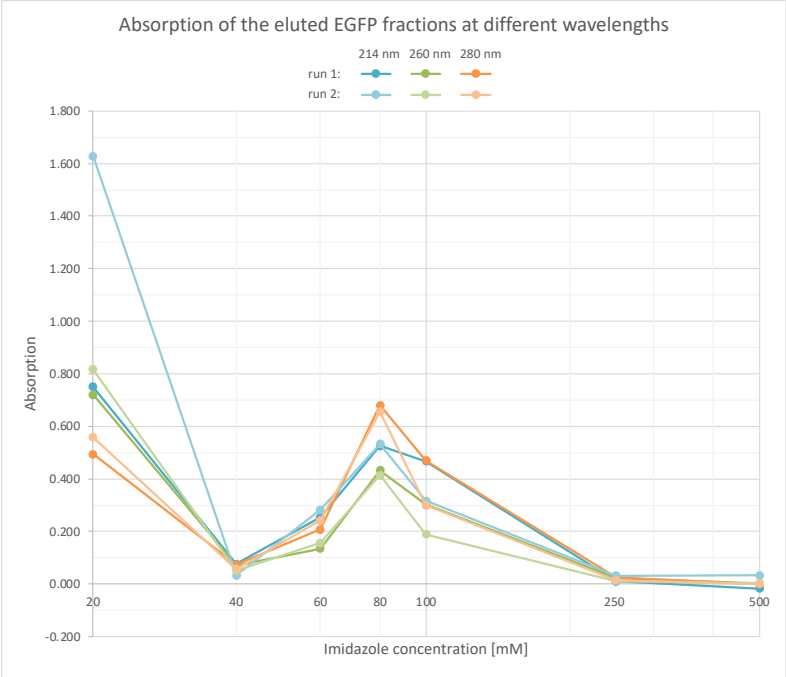

***Figure 3**: Absorption of the wash fraction and six eluted EGFP fractions at wavelengths of 214, 260 and 280 nm plotted against the corresponding imidazole concentrations (20, 40, 60, 80, 100, 250 and 500 mM), with the rich colors representing the first run and the light colors the second run*

Absorption of the wash fraction with 20 mM imidazole as well as ensuing eluate fractions with 40, 60, 80, 100, 250 and 500 mM imidazole was measured by the UV/Vis spectrophotometer at wavelengths of 214, 260 and 280 nm. By subtracting the absorption of the corresponding eluants with respective imidazole concentrations, the net absorption data were obtained and then plotted against the imidazole concentrations contained. UV light with a wavelength of 214 nm is absorbed by peptide bonds and, with higher extinction coefficients, by all aromatic amino acids and proline. [67] Macromolecules such as proteins and nucleic acids are responsible for absorption at 260 nm. The aromatic amino acids tyrosine and tryptophane as well as cystine disulfide bonds absorb UV light at 280 nm. Therefore, the absorption at 280 nm was crucial for determining the presence of EGFP and its concentration by use of the Beer-Lambert law with a molar absorption coefficient of 55,000 M^{-1} cm^{-1} for EGFP. In contrast, all peptides and proteins in solution were detected by measuring the absorption at 214 nm. The ratio of absorption at 260 nm/280 nm helped to estimate the contamination of protein solutions with nucleic acids.

The absorption curves in **Figure 3** were quite similar for both runs. As indicated by the name, it is to be assumed that the wash fractions contain non-histidine-tagged proteins, which were washed off the column with a 20 mM imidazole phosphate buffer after application of the protein solution and immobilization of the histidine-tagged EGFP proteins. At a concentration

of 40 mM imidazole, almost no elution was observed. The phosphate buffers containing 60, 80 and 100 mM imidazole were able to elute EGFP, with the 80 mM imidazole fraction containing the highest amount of EGFP. Hence, the fractions of 250 mM and 500 mM imidazole showed hardly any absorption. Subsequent to pooling the EGFP-containing fractions, SDS-PAGE was performed to determine whether the purification of EGFP via affinity chromatography was achieved.

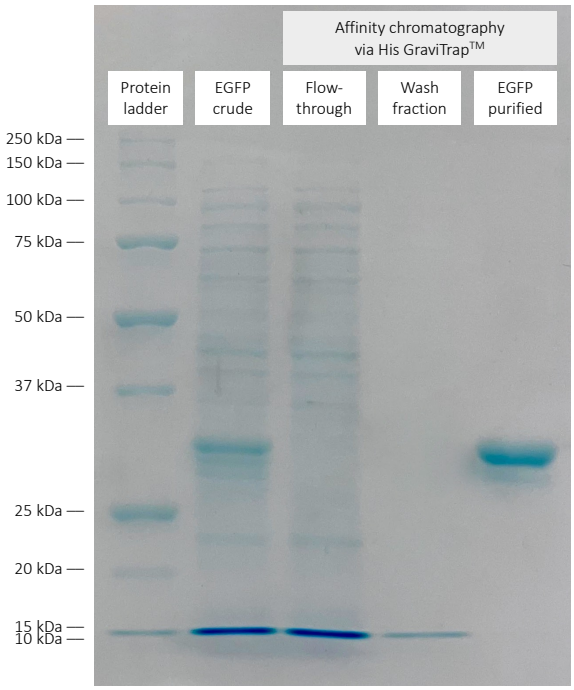

***Figure 4**: SDS-PAGE gel showing the proteins present throughout the purification of EGFP by manually performed affinity chromatography via His GraviTrap™ gravity-flow columns*

The SDS-PAGE gel in **Figure 4** allowed the comparison of proteins present in samples taken before, during and after purification. Lane 2 showed the filtered bacterial lysate before purification, which contained a variety of proteins, among which EGFP with a molecular weight of 26.9 kDa stood out as the strongest band. The flow-through fraction in lane 3 exhibited a similar protein profile but lacked the EGFP band and two bands below of similar size, indicating that EGFP immobilization on the nickel column was accomplished, however, potentially accompanied by impurities. The wash fraction mainly comprised proteins and peptides with a size of 15 kDa and smaller. Lane 5 displayed the purified pooled EGFP fractions, which showed a distinct EGFP band of higher intensity than in the crude solution due to higher concentration as well as a slight band below, indicating a possible impurity.

3.2.2. Expression of Cas9-His

Once the proof of concept for protein expression was demonstrated by means of EGFP, Cas9-His expression was performed analogously with the pET28a-Cas9-His transformed BL21(DE3) *E. coli* colony number 2. Two key differences were the use of other antibiotics, since the plasmids carry different antibiotic resistance genes, and the preparation of other eluants for the manual purification via gravity-flow columns, namely TRIS-buffered imidazole solutions with shorter jumps in concentrations.

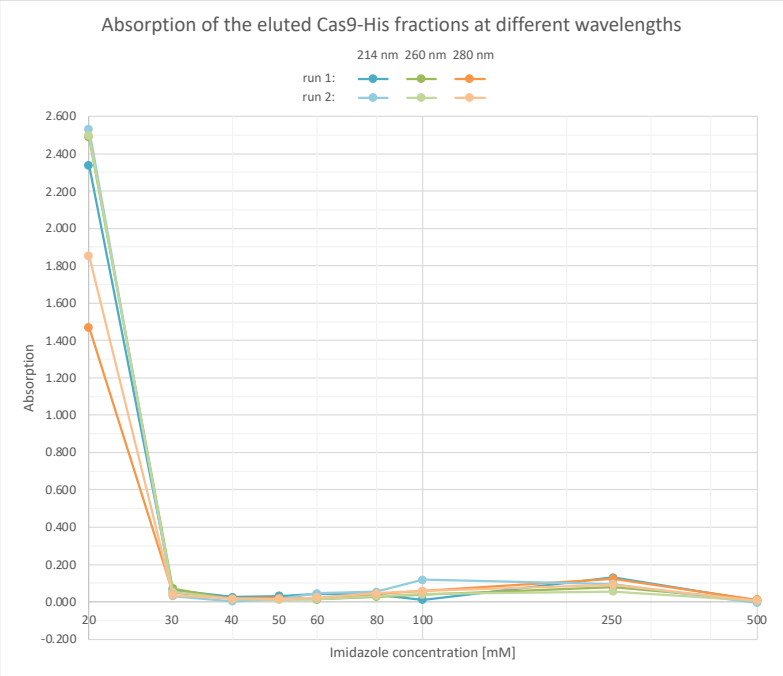

***Figure 5**: Absorption of the wash fraction and eight eluted Cas9-His fractions at wavelengths of 214, 260 and 280 nm plotted against the corresponding imidazole concentrations (20, 30, 40, 50, 60, 80, 100, 250 and 500 mM), with the rich colors representing the first run and the light colors the second run*

Regarding the absorption curves at wavelengths of 214, 260 and 280 nm in **Figure 5**, it must be considered that the molar absorption coefficient for Cas9-His is 120,450 M^{-1} cm^{-1}. Therefore, compared to EGFP, higher absorption is achieved for the same amount of protein or rather less protein is present at the same absorption. Following the elution instructions listed in 2.6.4, the column was washed with 20 mM imidazole solution subsequent to sample application, and then exposed to gradually increasing imidazole concentrations for elution. In comparison to absorption curves previously experienced with EGFP, the wash fraction showed a higher absorption followed by considerably lower absorptions of the eluates, indicating a smaller yield of Cas9-His.

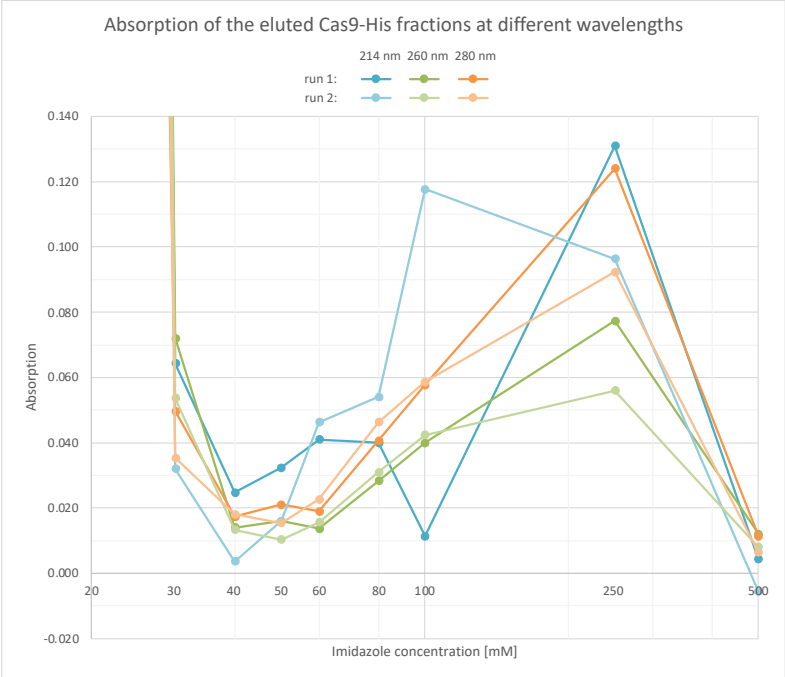

***Figure 6**: Enlarged image detail of the absorption of eight eluted Cas9-His fractions at wavelengths of 214, 260 and 280 nm plotted against the corresponding imidazole concentrations (20, 30, 40, 50, 60, 80, 100, 250 and 500 mM), with the rich colors representing the first run and the light colors the second run*

Providing a more detailed view, **Figure 6** only displays an absorption range of -0.02 to 0.14 on the y axis. For Cas9-His quantification, the absorption of cystine disulfide bonds and aromatic amino acids such as tyrosine and tryptophane at 280 nm was relevant. The slight protein elution at a concentration of 30 mM was probably due to a lag in the wash fraction and thus negligible, as 40 to 60 mM imidazole eluates contained very little protein. Elution of Cas9-His started to occur at an imidazole concentration of 80 mM, increased at 100 mM and peaked at 250 mM. Although the 500 mM imidazole fraction exhibited virtually no absorption, the eluates containing 80, 100, 250 and 500 mM imidazole were dialyzed and pooled for concentration.

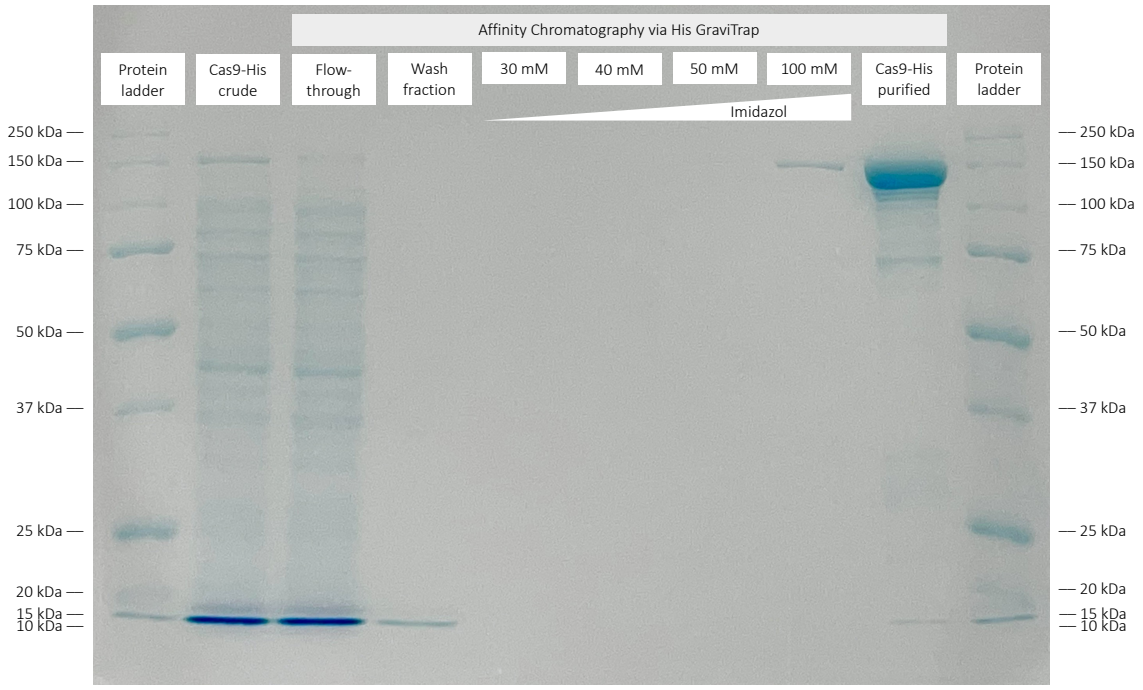

***Figure 7**: SDS-PAGE gel showing the proteins present throughout the purification of the first batch of Cas9-His by manually performed affinity chromatography via His GraviTrap™ gravity-flow columns*

The workflow of the manual purification of Cas9-His was demonstrated by SDS-PAGE of the samples taken throughout the purification process and is shown in **Figure 7**. Since Cas9-His has a molecular weight of 161,703 Da, the bands at the 150 kDa reference level can be assigned to it. The filtered *E. coli* lysate, marked as "crude" Cas9-His solution, contained a range of proteins including Cas9-His, which indicated a successful expression. In the lane of the flow-through, the still slightly visible Cas9-His band represented Cas9 protein that was not immobilized on the His GraviTrap™ columns. In contrast, the wash fraction solely contained proteins and peptides smaller than 20 kDa. None of the eluates at lower imidazole concentrations of 30, 40 and 50 mM comprised any Cas9-His nor other proteins. A previously conducted SDS-PAGE gel loaded with samples of 60 and 80 mM imidazole fractions revealed a barely visible Cas9-His band at 60 mM imidazole and a distinct band at 80 mM imidazole. Out of the higher imidazole concentrations, therefore, only the 100 mM eluate was analyzed here, which exhibited a distinct Cas9-His band, too. In the penultimate lane, the purification product obtained after dialysis, pooling and concentration of the eluates containing Cas9-His showed a large quantity of Cas9-His accompanied by impurities in size ranges of approximately 10-15, 30 and 70 kDa.

To improve the conditions of Cas9-His production, the purification of the second and third batch was carried out according to 2.6.3, 2.6.5 and 2.6.7. Cas9-His was subjected to high-

pressure homogenization and automated chromatography and was only ever exposed to temperatures of 0-4 °C during purification.

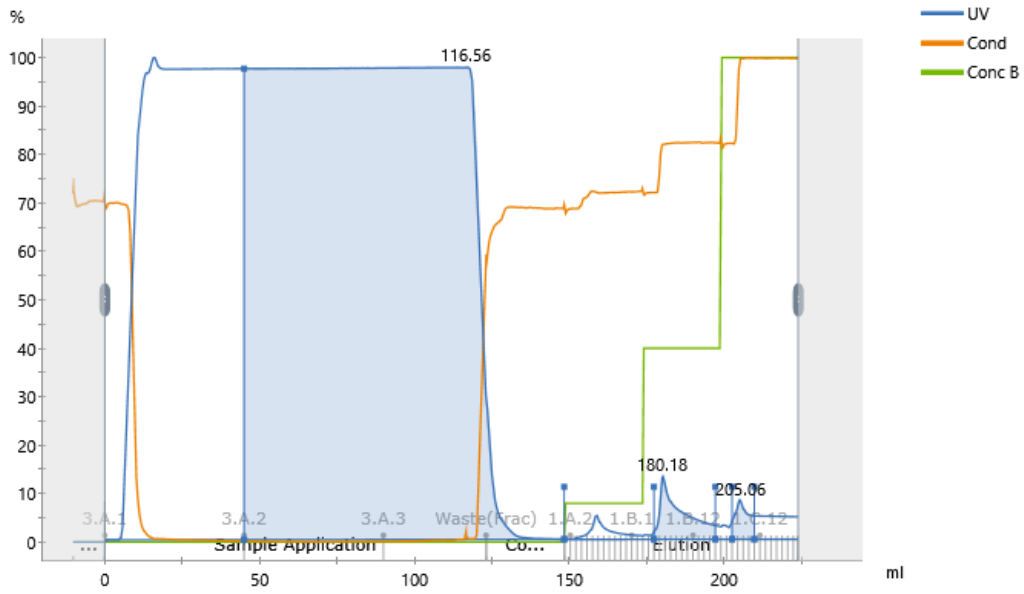

***Figure 8**: Absorption during affinity chromatography of Cas9-His (batch 2) using the ÄKTA pure™ chromatography system connected to a HisTrap™ (CV 5 mL) column, with two peaks at 1.B.5 – 1.C.3 and 1.C.7 – 1.C.10 highlighted in light blue on the right*

Following high-pressure homogenization, the E. coli lysate was subjected to gradient elution affinity chromatography using the ÄKTA pure™ system and a HisTrap™ column. **Figure 8** depicts the absorption of eluted proteins in blue at the corresponding imidazole concentration (as percentage of HisTrap Buffer B) in green. The first major absorption peak represented non-immobilized proteins in flow-through and wash fractions that were collected by rinsing the HisTrap™ column with HisTrap Buffer A (no imidazole). Upon altering the composition of HisTrap Buffer A and B, one peak each was detected at imidazole concentrations of 40, 200 and 500 mM. The eluates of the peaks highlighted in light blue were picked for subsequent size exclusion chromatography using the ÄKTA pure™ system and a HiLoad® 16/600 Superdex® 200 pg column. Cas9 Storage Buffer with a TRIS-HCl buffering system was applied as the mobile phase.

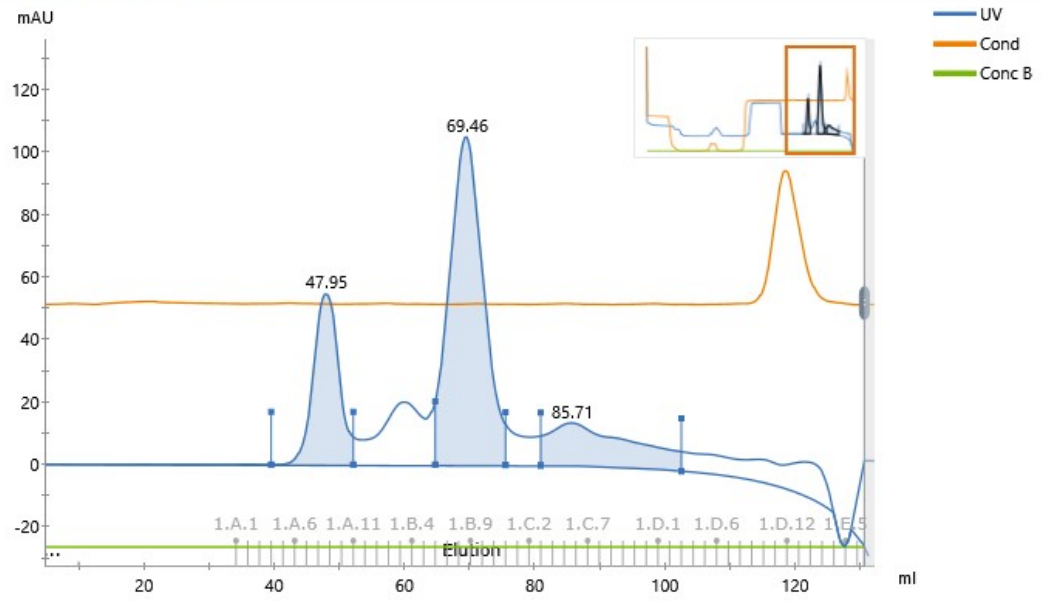

***Figure 9**: Absorption during size exclusion chromatography of Cas9-His (batch 2) using the ÄKTA pure™ chromatography system connected to a HiLoad® 16/600 Superdex® 200 pg column, with two peaks, at 1.A.4 – 1.A.10 and 1.B.6 – 1.B.11, as well as section 1.C.3. – 1.D.2 highlighted in light blue*

The absorption of the fractions eluted during size exclusion chromatography is visualized in **Figure 9** in blue, showing two distinct peaks and two small peaks in the retention range of 30-130 mL. The eluates enclosing the peaks at 47.95 (A4-A10) and 69.46 mL (B6-B11) were each pooled, the latter of which was found to contain Cas9-His, and samples of the other two increases in absorption were taken.

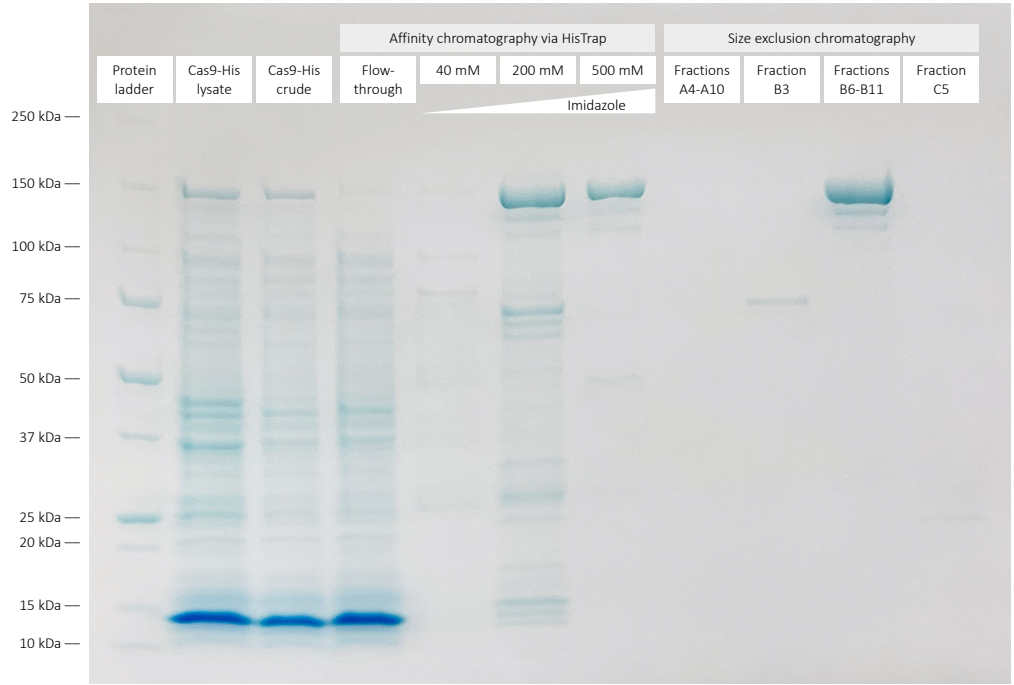

***Figure 10**: SDS-PAGE gel showing the proteins present throughout the purification of the second batch of Cas9-His by automated affinity chromatography via HisTrap™ column (CV 5 mL) and size exclusion chromatography via HiLoad® 16/600 Superdex® 200 pg column performed by the ÄKTA pure™ chromatography system*

To monitor the purification workflow, an SDS-PAGE was performed with samples taken throughout affinity and size exclusion chromatography, shown in **Figure 10**. As expected, the *E. coli* lysate as well as the filtered lysate, referred to as ”crude”, comprised a range of proteins including Cas9-His, to which the bands at the level of the 150 kDa reference were allocated. Representing the non-immobilized proteins, the flow-through showed a similar protein profile, however, containing only a minimal remnant of Cas9-His. This is consistent with the manual purification and implied a successful immobilization of Cas9-His. The affinity chromatography eluate samples (fractions A6, B6 and C8) taken at different imidazole concentrations (40, 200 and 500 mM) showed elution of Cas9-His as of 200 mM imidazole. While the 40 mM eluate in lane 5 comprised hardly any Cas9-His and little impurities, the 200 mM eluate in lane 6 exhibited a prominent Cas9-His band and more pronounced bands of impurities. The sample taken at 500 mM imidazole, in contrast, showed a distinct band of Cas9-His accompanied by fewer impurities in the form of two bands below of similar size and one band around 50 kDa in size. SDS-PAGE analysis of the size exclusion chromatography (SEC) samples revealed that the prominent peak shown in Figure 9, here represented by the fractions B6-B11 in the penultimate lane, contained Cas9-His. Despite size exclusion, the Cas9-His protein was still accompanied by the two protein impurities of slightly smaller size. The other increases in absorption during size exclusion chromatography proved to be contamination. The fractions A4-A10 showed no band, possibly due to an impurity too small in size. Regarding fraction B3, the impurity was around 75 kDa, whereas fraction C5 exhibited a band around 25 kDa. As per 2.6.9, the SEC fraction pool B6-B11 that contained Cas9-His was concentrated and either 0%, 10% or 20% glycerol was added. For storage at -80 °C, 20 µL aliquots were frozen in liquid nitrogen.

Testing the reproducibility, a third batch of Cas9-His was expressed and likewise purified via automated affinity and size exclusion chromatography. Merely the elution steps during the affinity chromatography were altered to imidazole concentrations of 40, 60, 250 and 500 mM. The additional 60 mM step was intended to promote elution of impurities prior to elution of Cas9-His. Raising the concentration of 200 mM to 250 mM imidazole aimed to achieve elution of Cas9-His at once.

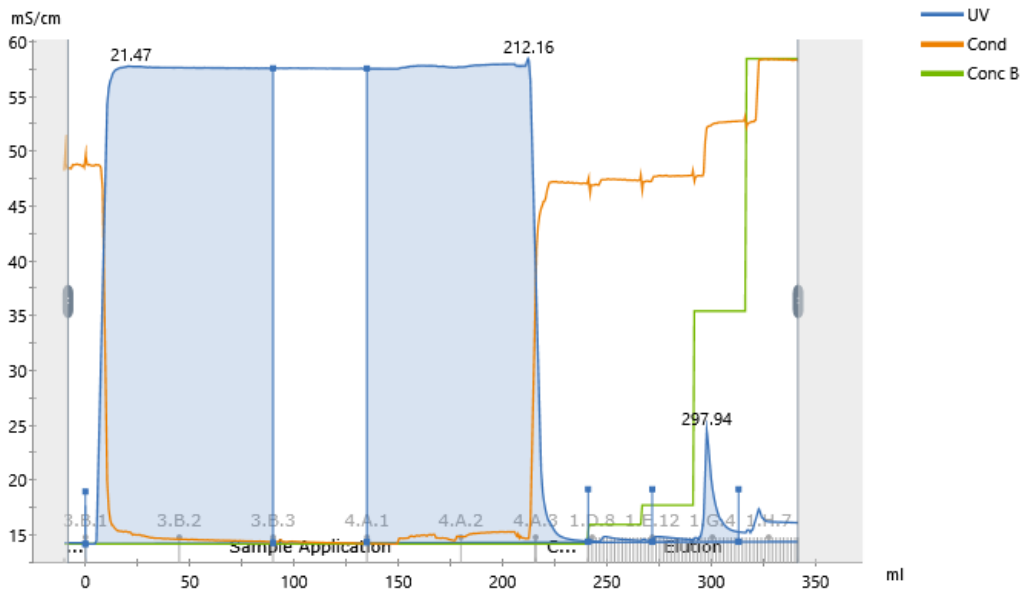

***Figure 11**: Absorption of Cas9-His (batch 3) during affinity chromatography using the ÄKTA pure™ chromatography system connected to a HisTrap™ (CV 5 mL) column, with a distinct peak within section 1.E.12 – 1.G.10 highlighted in light blue on the right*

In contrast to Figure 8 of the second Cas9-His batch, the additional elution step of 60 mM imidazole during the automated affinity chromatography of the third batch lead to minimal increases of the blue absorption curve in **Figure 11** at concentrations of 40 and 60 mM imidazole. The peak at 250 mM imidazole appeared to be slightly more distinct, while the small peak at 500 mM imidazole resembled the one of the second batch.

***Figure 12**: SDS-PAGE gel showing the proteins present throughout the purification of the third batch of Cas9-His by automated affinity chromatography via HisTrap™ column (CV 5 mL) performed by the ÄKTA pure™ chromatography system*

SDS-PAGE analysis of the samples taken before and during affinity chromatography is pictured in **Figure 12**. The protein profiles of the lysates and the flow-through were consistent with previous results. Importantly, the less saturated coloring must be taken into account when comparing with the second batch of Cas9-His, potentially disguising some impurities. Hardly any impurities were therefore visible throughout the elution with imidazole. Remarkably, increasing the 200 mM imidazole eluant to 250 mM imidazole accomplished the elution of Cas9-His proteins all at once, as seen in lane 9. On the downside, the 250 mM imidazole eluate still comprised several impurities. Directly below the prominent Cas9-His band, three very thin bands of similar size appeared. Additionally, four other bands were observed around 70, 75, 35 and 10-15 kDa, of which the two bands around 70-75 kDa were also present in the eluate with an imidazole concentration of 60 mM.

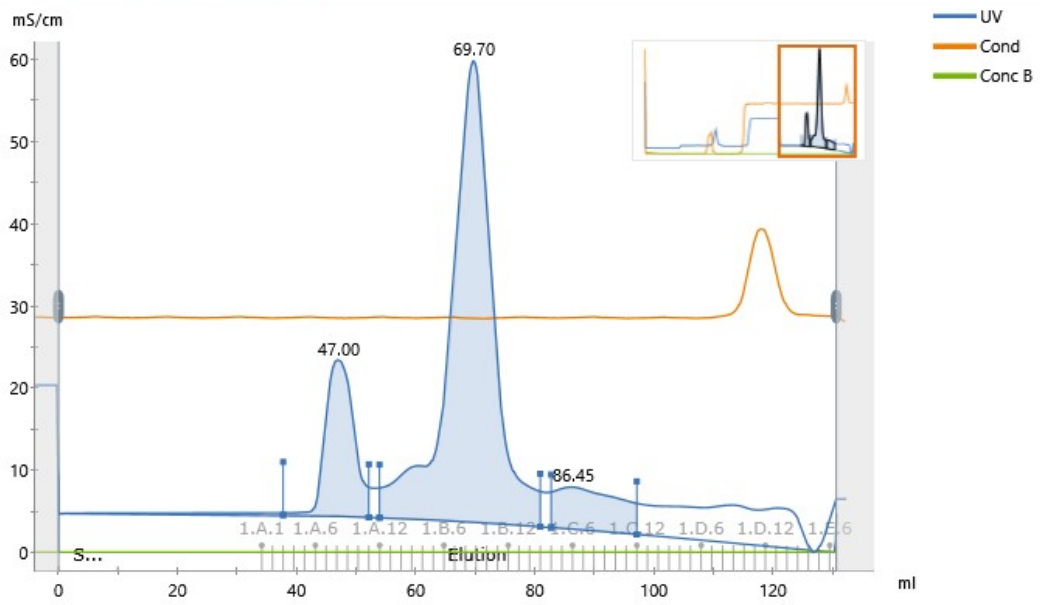

***Figure 13**: Absorption of Cas9-His (batch 3) during size exclusion chromatography using the ÄKTA pure™ chromatography system connected to a HiLoad® 16/600 Superdex® 200 pg column, with two peaks, 1.A.3 – 1.A.10 and 1.A.12 – 1.C.2, and section 1.C.4 – 1.C.11 highlighted in light blue*

Size exclusion chromatography was performed in the same manner as for the second Cas9-His batch, with the only difference being that solely the eluates containing 250 mM imidazole were selected for SEC. As shown in **Figure 13** in blue, the detected absorption peaked at 47.00 and 69.70 mL within the retention range of 30-130 mL. Around 60 and 86 mL, minimal increases were observed. Samples of each increase or peak (A8, B2, B6, C6) were taken and analyzed via SDS-PAGE.

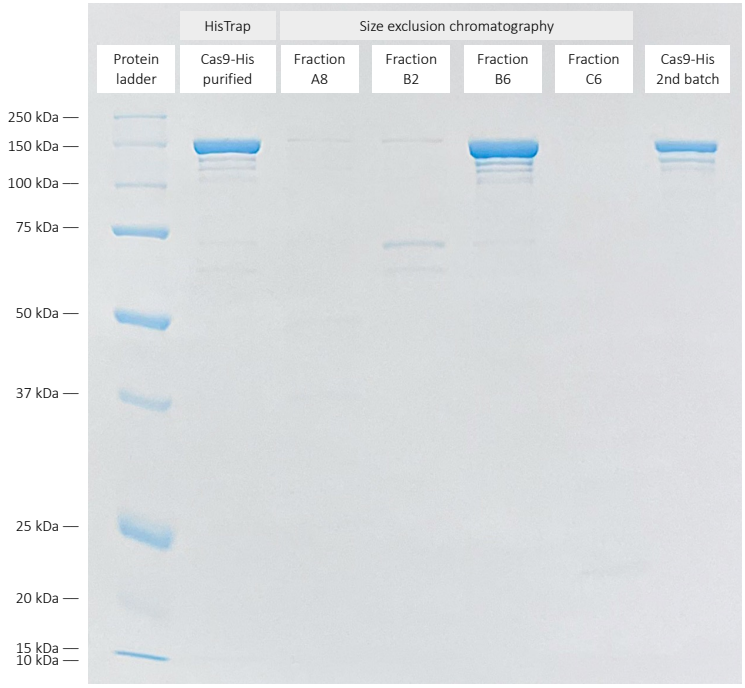

***Figure 14**: SDS-PAGE gel showing the proteins present before and throughout the purification of the third batch of Cas9-His by automated size exclusion chromatography via HiLoad® 16/600 Superdex® 200 pg column performed by the ÄKTA pure™ chromatography system*

The SDS-PAGE gel depicted in **Figure 14** showed the Cas9-His solution before SEC fractionation (lane 2), the SEC fractions (lanes 3-6) as well as the Cas9-His protein solution of the second batch after purification via affinity and size exclusion chromatography (lane 7). As indicated by the bands of Cas9-His at the level of the 150 kDa reference, the fractions A8 and B2 already contained minimal amounts of Cas9-His, which were negligible in relation to fraction B6. Fraction B6 in lane 5 stemmed from the major peak of the SEC absorption graph in Figure 13 and exhibited a visually distinct band of Cas9-His as well as three thin bands of slightly smaller size below. The two bands of 70-75 and 60-70 kDa in size were present in fraction B2 but also to a minor extent in fraction B6 with Cas9-His. In terms of purification, SEC was again able to separate impurities divergent in size from Cas9-His proteins, but unable to separate contaminations of similar size. As the final purified Cas9-His product of the third batch (lane 5) coincides with the second batch (lane 7), purification of Cas9-His proteins appears to be reproducible.

3.3. In vitro cleavage assays

The Cas9 in vitro cleavage assays were performed as described in 2.7, employing target plasmid pEGFP-Luc (6367 bp), which encodes a fusion protein of EGFP and the firefly *Photinus*

pyralis luciferase. [65] To guide Cas9 to the EGFP gene, the sgRNA "sgGFP" was used. Within this test system, the cleavage activity of the recombinantly expressed Cas9-His was examined by first linearizing the target plasmid and then adding the linearized pEGFP-Luc to a pre-incubated mix of Cas9 and sgGFP. Pre-incubation served the assembly to Cas9/sgGFP ribonucleoprotein complexes. To visualize cleavage activity, gel electrophoresis was conducted. As negative control and to monitor the success of linearization, a sample of the linearized target plasmid was used. The IVC samples included commercial Cas9 as positive control. Each Cas9 underwent the IVC assay with, and for negative control purposes additionally without, sgGFP. Since restriction enzyme XBaI cleaves the plasmid after 3,039 bp and Cas9/sgGFP introduces double strand breaks after 646 bp, the linearized target plasmid was expected to be cleaved into two fragments of 2,393 bp and 3,974 bp in size when sgGFP was present.

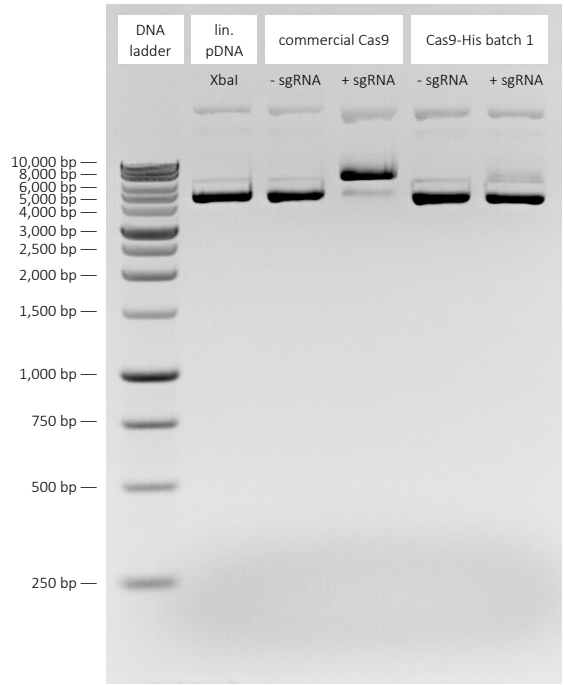

***Figure 15**: 1% Agarose gel displaying an in vitro cleavage assay to determine nuclease activity of Cas9-His, guided by the sgRNA "sgGFP", using target plasmid pEGFP-Luc (6.4 kb) that was subjected to restriction enzyme XbaI*

However, as demonstrated in **Figure 15**, the linearization of the target plasmid did not succeed when incubated for 2 hours with XBaI. Except for lane 1 containing the DNA ladder, each lane showed a band in between 6,000 and 8,000 bp, identified as the linearized plasmid of approximately 6.4 kb in size, as well as one band above and one below. The band significantly larger than 10 kb was present in all sample lanes, remaining despite Cas9 nuclease activity, which suggested a contamination. The band at 5,000 bp represented the supercoiled

plasmids. In lane 2, the linearized plasmid band was very thin, whereas the band of supercoiled pDNA was thicker and more pronounced, indicating that XBaI had not properly digested the target plasmid. For this reason, no fragments were generated upon cleavage by Cas9. As Cas9 is able to cleave plasmids in their linearized as well as supercoiled state, [6] Cas9/sgRNA ribonucleoproteins instead led to an increase of linearized pDNA. In case of the reference Cas9 RNPs (lane 4), a substantial augmentation of linearized pDNA occurred, while RNPs of the first recombinantly expressed Cas9-His batch (lane 6) only showed a slight augmentation. Nonetheless, this pointed to the presence of nuclease activity in the first batch of recombinantly expressed Cas9-His, although minimal in comparison to commercial Cas9.

To eliminate the suspected contaminations, the plasmid pEGFP-Luc was purified using the GeneJET Plasmid Miniprep Kit. The purified target plasmid was then incubated overnight instead of 2 hours with a fivefold higher ratio of restriction enzyme XBaI (see 2.7.1).

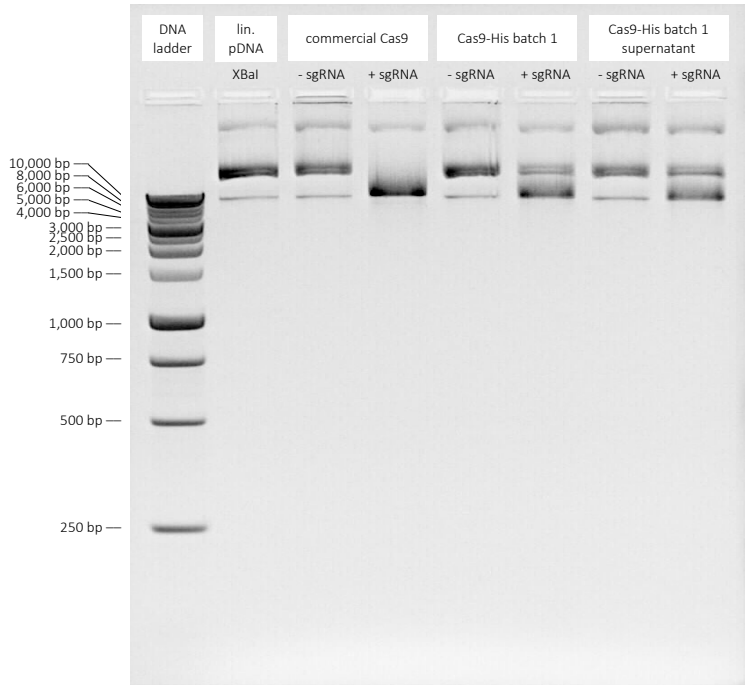

***Figure 16**: 1.5% Agarose gel displaying an in vitro cleavage assay to determine nuclease activity of Cas9-His, guided by the sgRNA "sgGFP", with target plasmid pEGFP-Luc (6.4 kb) that was purified via GeneJET Plasmid Miniprep Kit and then treated with restriction enzyme XbaI*

The subsequent in vitro cleavage assay still revealed similar results, displayed in **Figure 16**. Interestingly, precipitates were visible after thawing the Cas9-His solution. Hence, the supernatant of the protein solution, which was obtained after ten minutes of centrifugation at 14,000 xg, was additionally tested for IVC activity (lanes 7, 8). Besides lane 1 containing the DNA standards, each lane showed three bands, one around 8,000 bp and two significantly larger ones. In all likelihood, the bottom bands were linearized pDNA. The bands in the middle

most likely represented nicked pDNA, as these fragments migrated less far than linearized pDNA. The presence of nicked pDNA might be caused by stress on the plasmids due to the purification. The top band again suggested a contamination, as it was present in all sample lanes including functioning Cas9/sgRNA ribonucleoproteins. Since the linearized pDNA is approximately 6.4 kb large, the plasmid bands appeared to be slightly shifted upwards relative to the DNA ladder. As before, the thin band of linearized plasmids in lane 2 implied that the linearization of the target plasmid had failed, and the Cas9/sgRNA ribonucleoproteins in lane 4, 6, 8 thus showed their IVC activity in form of linearization of the plasmids instead of generating fragments. Commercial Cas9/sgRNA ribonucleoproteins (lane 4) were able to completely convert nicked pDNA and significantly increased the amount of linearized pDNA compared to the negative control of commercial Cas9 without sgRNA. Recombinantly expressed Cas9-His in conjunction with sgRNA (lanes 6, 8) was less efficient, showing an increase of linearized plasmids and imperceptibly changed bands of nicked plasmids. This result again demonstrated the nuclease activity of the first batch of recombinantly expressed Cas9-His, which was lower than the one of commercial Cas9.

The failed linearization of the target plasmid might underlie the inhibition of XBaI by methyl residues that plasmids contain depending on their host. To circumvent this issue, pEGFP-Luc forward and reverse primers, which flank a 1.5 kb gene sequence including the EGFP gene of the target plasmid, were designed and purchased. Using these primers, the linear sequence shown below (with the primers underlined, the EGFP gene marked green, and the target sequence highlighted in red), termed "**IVC EGFP gene (1.5 kb)**", was amplified via gradient PCR at different annealing temperatures.

5'-**CGCCAATAGGGACTTTCCATTG**ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA
GTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT
GGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGG
GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTT
CCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC
TATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGCGCTACCGGTCGCCACCATGGTGAG
CAAGGGCGAGGAGCTGTTCACCG**GGGTGGTGCCCATCCTGGTC**GAGCTGGACGGCGACGTAAACG
GCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGT

TCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTG
CAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAG
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA
AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGC
AACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGC
AGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCG
CCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCT
GAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGCCGGACTCAGATCTC
GAGCTCAAGCTTCGAATTCGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTATCCGCTG
GAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTTCCTGGAACA
ATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTTCGAAATGTCCGTTCG
GTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAATCGTCGTATGCAGTGAAAA
CTCTCTTCAATTCTTTATGCC**GGTGTTGGGCGCGTTATTTATC**-3'

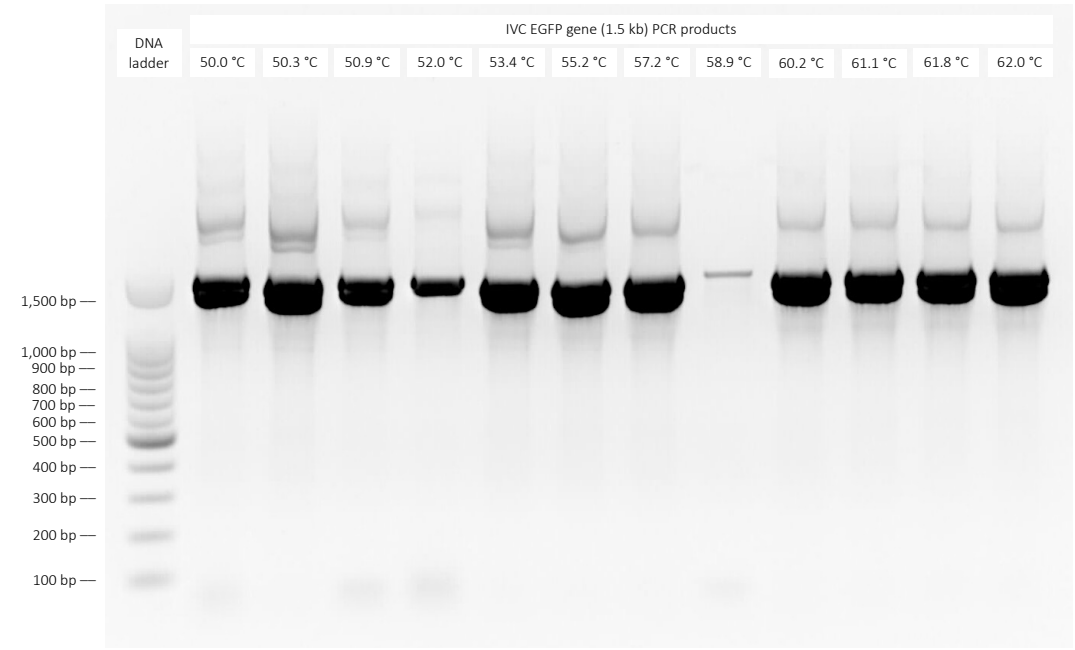

***Figure 17**: 1% Agarose gel showing the PCR products obtained after gradient PCR with varying annealing temperatures, whereby a sequence enclosing the EGFP gene of the plasmid pEGFP-Luc (6.4 kb), termed "IVC EGFP gene (1.5 kb)", was amplified by means of accordingly designed forward and reverse primers*

Figure 17 displays the results of ensuing gel electrophoresis of the linear PCR products obtained. The bands at the level of the 1,500 bp reference represent the IVC EGFP gene, whereas all bands above are impurities. The temperature specifications shown refer to the different annealing temperatures applied in the course of the PCR cycles, nearly all of which yielded prominent IVC EGFP gene bands with high intensities. Exception to this was the

annealing temperature of 58.9 °C, in comparison showing a very thin band of IVC EGFP gene, which suggested that the thermocycler had not been able to control the temperature at this position.

As the best ratio between product and impurities was yielded by PCR cycles with an annealing temperature of 60.2 °C, the corresponding linearized PCR product was used for another PCR run. The resulting IVC EGFP gene (1.5 kb) batch was applied in subsequent IVC assays. Since the DNA was already linear, the step of plasmid linearization could be omitted.

In response to precipitation that occurred in Cas9-His solutions of the first batch, the second batch of Cas9-His was expressed under better conditions, such as automated purification of the His-tagged proteins at 4°C by an ÄKTA pure™ chromatography system. In addition, 0%, 10% or 20% glycerol was added as a cryoprotectant to the purified and concentrated Cas9-His protein solution. To examine whether glycerol improves storage conditions, the nuclease activity of the first batch of Cas9 without and the second batch with 0%, 10% and 20% glycerol were compared with each other after one, three and five normed freeze-thaw cycles. The absence of sgRNA served as a negative control and commercial Cas9 as a positive control. Cleavage of the IVC EGFP gene (1.5 kb) was expected to result in two fragments of 494 and 1013 bp in size.

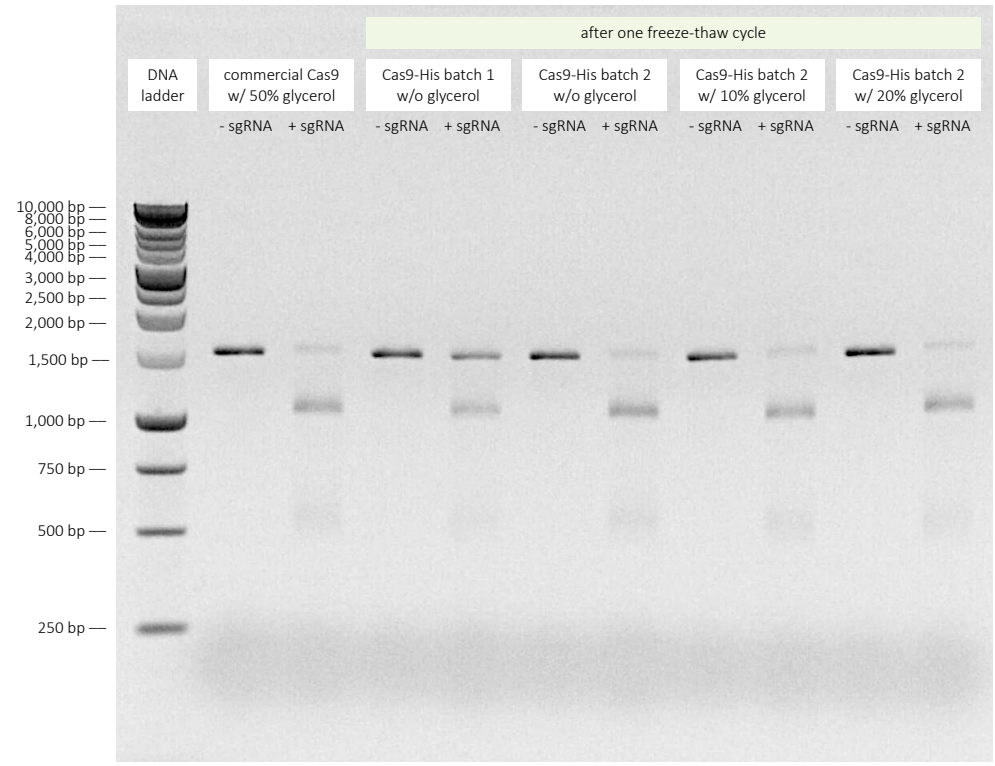

***Figure 18**: 1% Agarose gel of an in vitro cleavage assay using the target IVC EGFP gene (1.5 kb) to determine the nuclease activity of Cas9-His after one freeze-thaw cycle, guided by the sgRNA "sgGFP", and its dependence on the addition of 0%, 10% or 20% glycerol*

Indeed, the agarose gels in Figure 18, Figure 19 and Figure 20 showed the linear IVC EGFP gene (1.5 kb) at the level of the 1,500 bp reference and the fragments generated by cleavage at around 500 and 1,000 bp. As shown in **Figure 18**, the IVC assay after one freeze-thaw cycle revealed nuclease activity present in each Cas9 sample. The Cas9/sgRNA ribonucleoproteins of the second Cas9-His batch (lanes 7, 9, 11) exhibited nuclease activity similar to the commercial Cas9 in conjunction with sgRNA (lane 3), regardless of the glycerol concentration. In contrast, the sample of the first batch of Cas9 (lane 5) again possessed significantly lower nuclease activity, although subjected to only one freeze-thaw cycle. This was indicated on the one hand by slightly less visible bands of the DNA fragments and on the other hand by the distinct target DNA band when complexed with sgGFP.

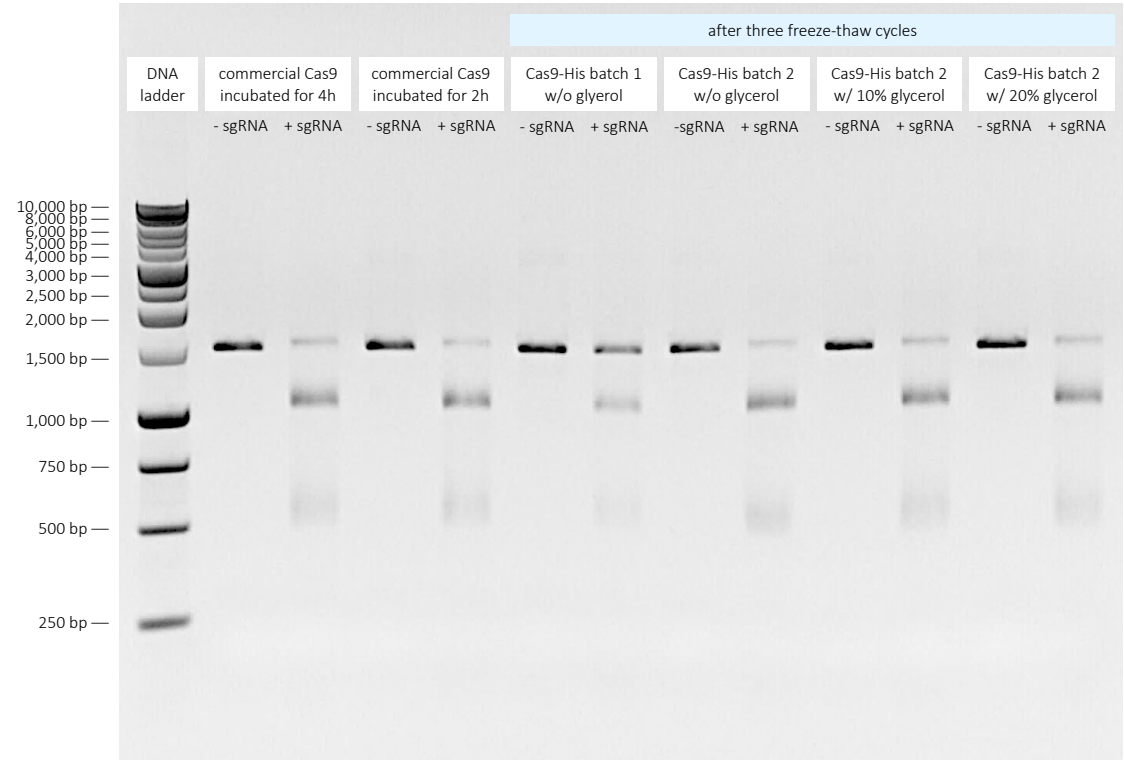

***Figure 19**: 1% Agarose gel of an in vitro cleavage assay using the target IVC EGFP gene (1.5 kb) to determine the nuclease activity of Cas9-His after three freeze-thaw cycles, guided by the sgRNA "sgGFP", and its dependence on the addition of 0%, 10% or 20% glycerol*

For the IVC assay after three freeze-thaw cycles, shown in **Figure 19**, samples of commercial Cas9 with and without sgGFP (lanes 2-5) were incubated for 4 hours in addition to 2 hours with the intention of entirely transforming the target DNA. However, this aim was not achieved and no significant increase in target DNA cleavage was detected. Each ribonucleoprotein sample of the second batch of recombinantly expressed Cas9-His (lanes 9, 11, 13) again showed cleavage activity similar to the positive control, independent of the glycerol concentration. In comparison, the first Cas9-His batch without glycerol yielded a distinctly lower nuclease activity in conjunction with sgRNA (lane 7). As before, this is apparent by the high intensity of the target DNA band and the comparatively lower intensity of the DNA fragment bands.

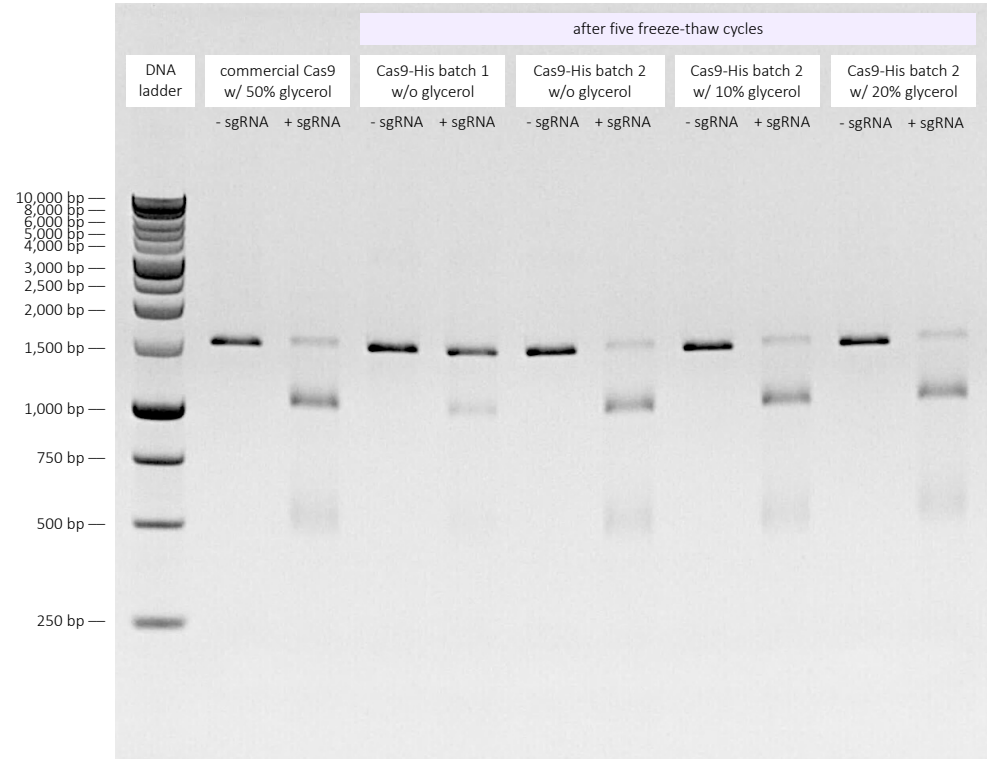

***Figure 20**: 1% Agarose gel of an in vitro cleavage assay using the target IVC EGFP gene (1.5 kb) to determine the nuclease activity of Cas9-His after five freeze-thaw cycles, guided by the sgRNA "sgGFP", and its dependence on the addition of 0%, 10% or 20% glycerol*

The outcome of the IVC assay after five freeze-thaw cycles, displayed in **Figure 20**, was consistent with the previous results. Recombinantly expressed Cas9-His of the second batch in conjunction with sgRNA (lanes 7, 9, 11) once again showed cleavage activity similar to the positive control of commercial Cas9 (lane 3), with visually imperceptible differences between the different glycerol concentrations. The ribonucleoproteins of the first Cas9-His batch (lane 5) also continued to show clearly lower nuclease activity than the positive control and the second batch. It has to be noted that the stronger intensity of the bands in lane 3 appeared to be equally enhanced for each band, which holds the potential to be misleading when making comparisons. This problem was circumvented by the following quantification method.

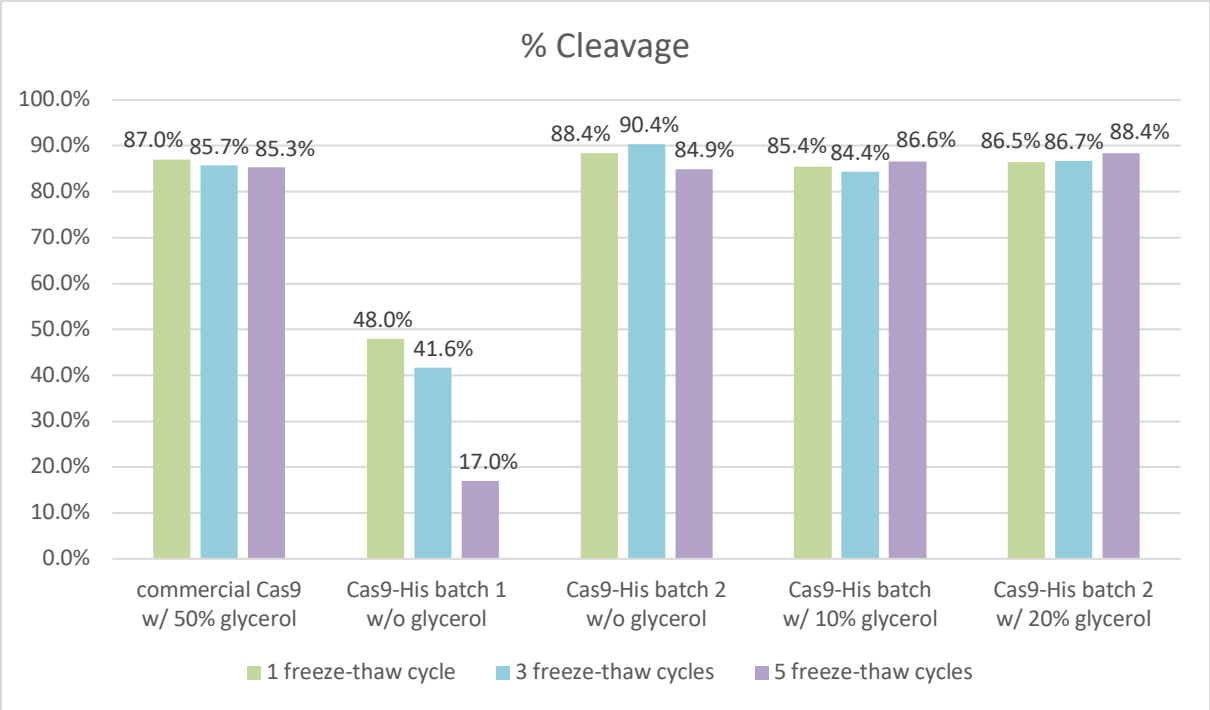

***Figure 21**: Comparison of the Cas9 in vitro cleavage assay results after one, three and five freeze-thaw cycles*

To evaluate the results of the IVC assays, the image processing program ImageJ was used to quantify the three bands of each ribonucleoprotein sample. By putting the sum of the two DNA fragment bands in relation with the total sum of all three fragments, the percentage of cleavage for each Cas9/sgRNA RNP sample was calculated after one, three and five freeze-thaw cycles and summarized in **Figure 21**.

In this graphical representation, it is particularly striking that the first batch of recombinantly expressed Cas9-His possessed distinctly lower cleavage activity that decreased with further freeze-thaw cycles. While the cleavage percentages of the first Cas9-His batch after one and three freeze-thaw cycles were roughly half of the cleavage percentages achieved by the positive control or the second batch, this was reduced to one fifth after five freeze-thaw cycles. Reasons for this might be both the manual protein purification at room temperature as well as storage issues without the cryoprotectant glycerol.

Cleavage percentages of the commercial Cas9 and the second Cas9-His batch all ranged within 84-91%. Thus, automated protein purification at 4°C was worthwhile. With regard to storage conditions, the results showed an effect of glycerol on cleavage activity after five freeze-thaw cycles. With rising glycerol concentration, the Cas9-His RNP cleavage percentages increased from 84.9% (no glycerol) to 86.6% (10% glycerol) to 88.4% (20% glycerol). As cleavage

percentage was the highest for ribonucleoproteins composed of Cas9-His stored in 20% glycerol, the addition of 20% glycerol was selected for future Cas9-His batches.

3.4. PEI-Suc 10%

Via ^{1}H-NMR spectroscopy, samples of the obtained 10% succinylated 10 kDa, 25 kDa and 70 kDa PEIs were analyzed in deuterium oxide. Since D_2O is not 100% pure, a solvent residual peak with a chemical shift of 4.79 ppm was visible in the spectra. [68] The chemical shift of the peak cluster of the different PEIs ranged from approximately 2.60 to 3.70 ppm. For the succinic acid peak, a chemical shift peaking at 2.43-2.44 ppm was found. To verify whether the 10% succinylation was achieved, the yielded percentage of succinylation was determined for each PEI by putting the integrated area of the PEI and succinic acid in proportion. The succinylation of the 10 kDa PEI was the most effective with 9.86% (**Figure 22**), closely followed by the 9.49% succinylated 25 kDa PEI (**Figure 23**). However, the 70 kDa PEI was only 6.93% succinylated (**Figure 24**). A possible explanation is the evaporation of water from the 30% PEI solution used, causing a higher PEI concentration, whereupon a correspondingly larger amount of PEI was used for the succinylation reaction.

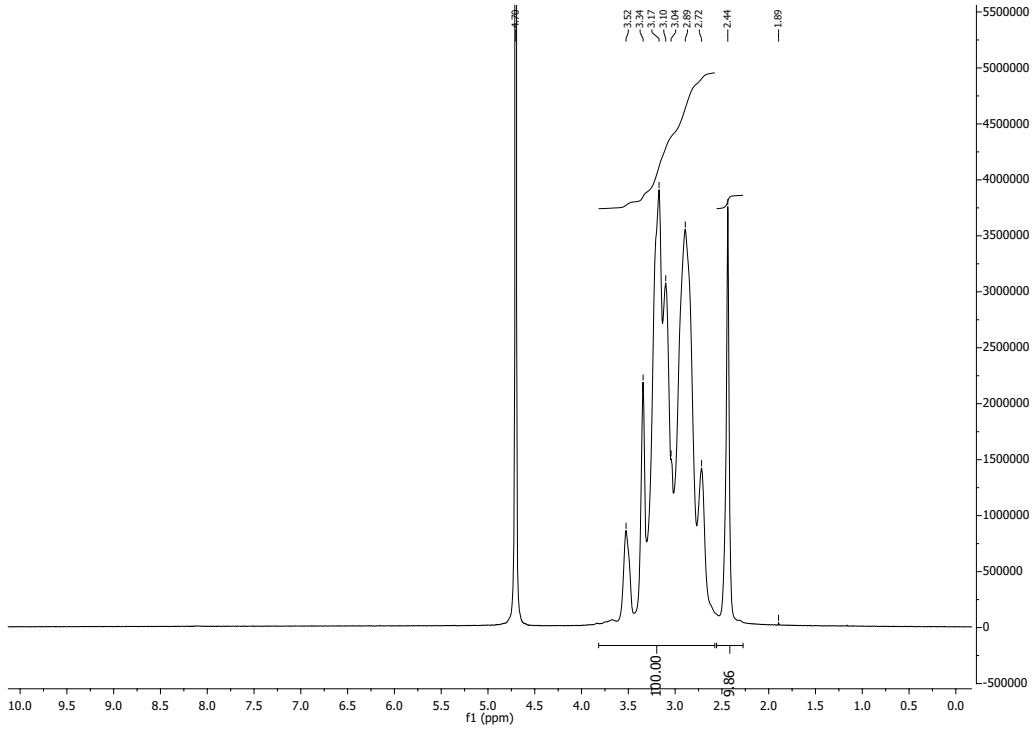

***Figure 22**: ^{1}H-NMR spectrum of 9.86% succinylated 10 kDa PEI*

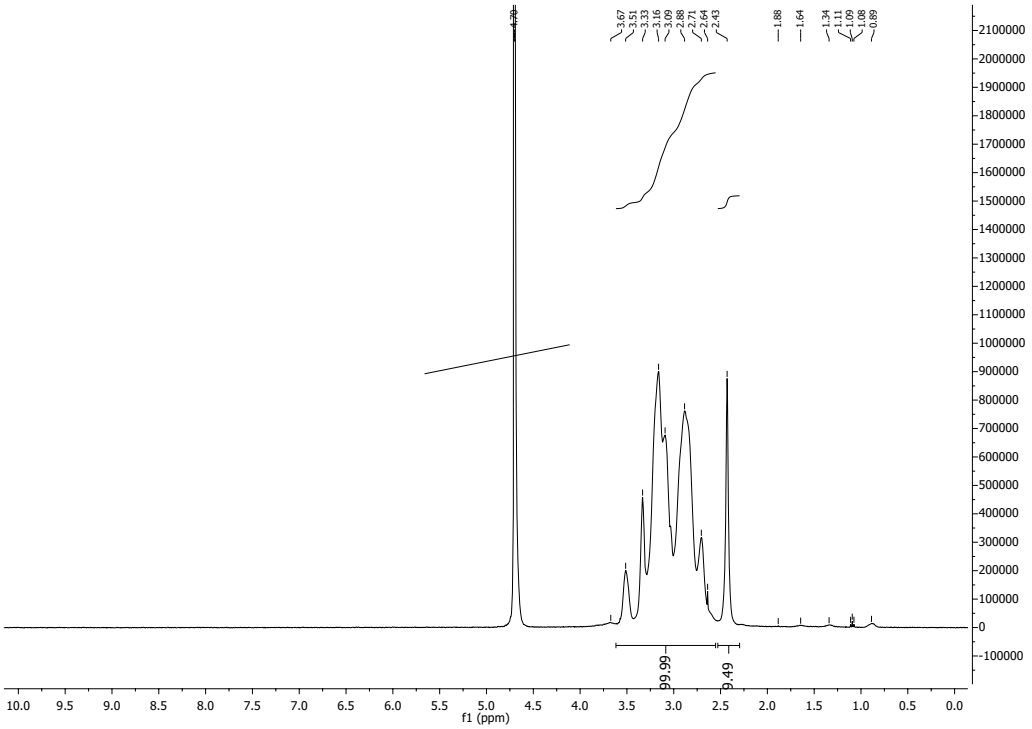

***Figure 23**: [1]H-NMR spectrum of 9.49% succinylated 25 kDa PEI*

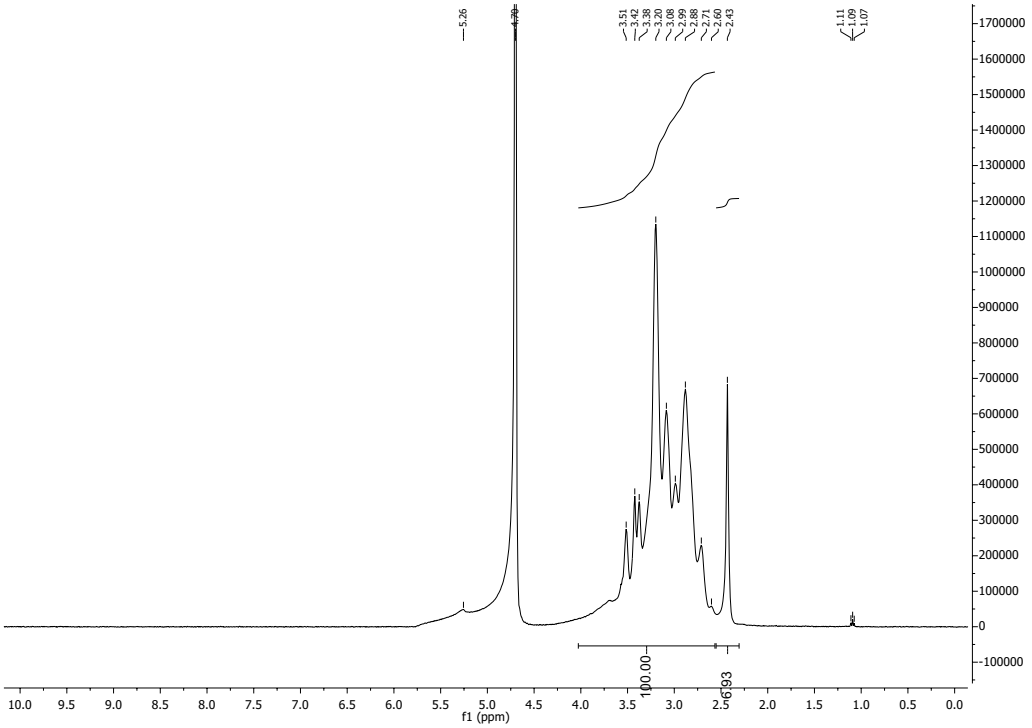

***Figure 24**: [1]H-NMR spectrum of 6.93% succinylated 70 kDa PEI*

3.5. Ribonucleoprotein particles

Subsequent to preparation of Cas9 ribonucleoprotein carriers according to 2.9 with either PEI-Suc 10% or oligopeptide UV_1, the particles were immediately characterized by measuring the diameter and zeta potential using the Zetasizer Nano ZS. Due to a conversion error, the actual values of the Cas9 concentrations and thus RNP ratios deviated by +2% from the target values which are stated in this thesis.

3.5.1. Characterization of Cas9 PEI-Suc 10% ribonucleoprotein carriers

Cas9 PEI-Suc 10% RNP carriers were prepared in HBG using the 10% succinylated 10 kDa, 25 kDa and 70 kDa PEIs. Roughly summarized, pre-incubated equimolar Cas9/sgGFP RNP solutions were added to appropriate dilutions of PEI-Suc 10% at w/w ratios (PEI-Suc 10% to sgRNA) of 4:1, 2:1, 1:1 and 0.5:1.

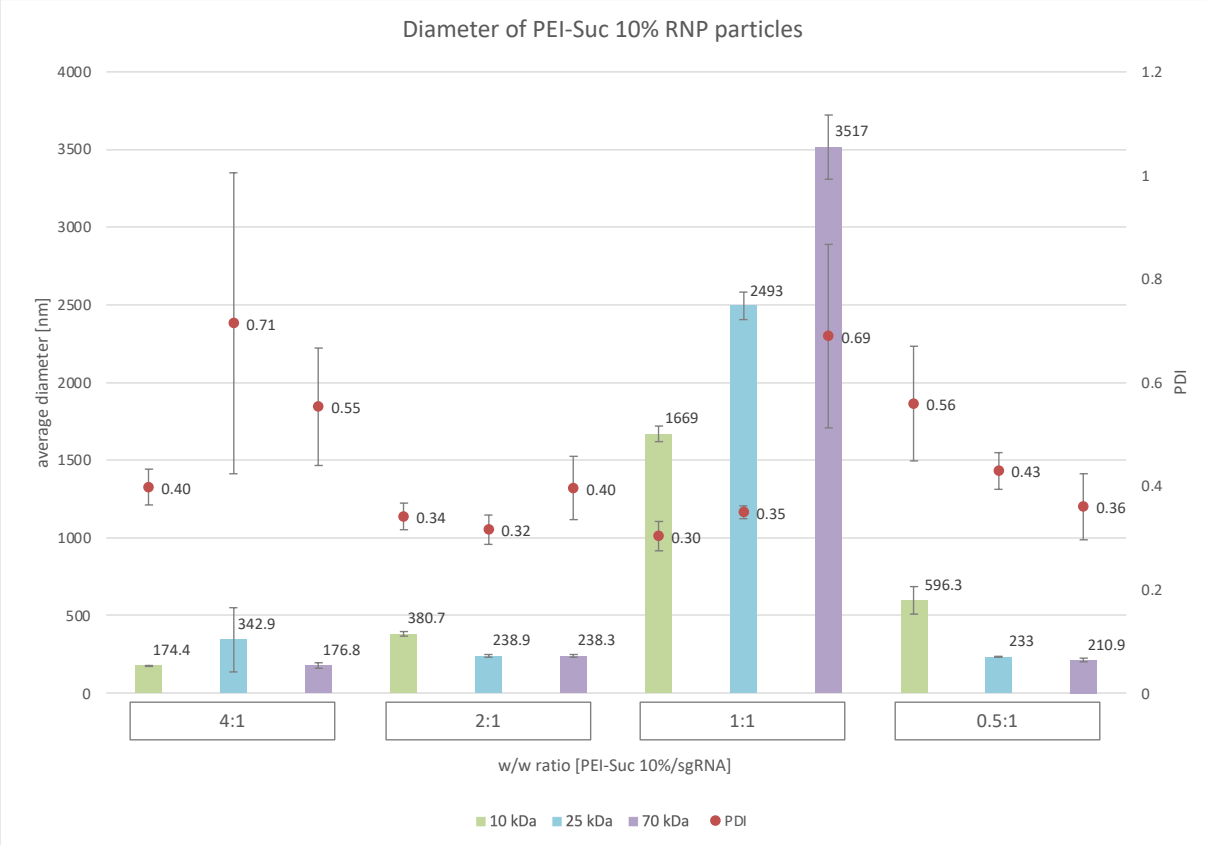

***Figure 25**: Particle size of PEI-Suc 10% Cas9/sgGFP RNP carriers*

For comparison of particle size, the average diameter and polydispersity index (PDI) of each PEI-Suc 10% Cas9/sgGFP RNP particle type were summarized in a chart, seen in **Figure 25**. While w/w ratios of 4:1, 2:1 and 0.5:1 yielded nanoparticles, the w/w ratio of 1:1 resulted in particles much larger than 1 µm in diameter. The latter could underlie several reasons.

Example of one explanation would be the formation of differently charged particles at a w/w ratio of 1:1, which could result in neutralization, leading to a lack of electrostatic repulsive forces and thus increasing the potential of agglomeration. However, the formation process of the particles was not analyzed in this thesis and remains uncertain. At a w/w ratio of 0.5:1, an excess of negatively charged RNPs was potentially present in the solution, since an ethidium bromide exclusion assay of Zintchenko et al. had demonstrated more effective binding of RNA by 9% succinylated PEI at slightly higher w/w ratios. [52] As RNP complexes without PEI exhibited an average diameter of 343.4 nm (SD 162.8 nm) with a PDI of 0.877 (SD 0.127), this additional subpopulation could bias the average size and average PDI of the particles. Zintchenko et al. had also shown promising in vitro gene silencing assay results of 9% succinylated PEI with small interfering RNA (siRNA) at w/w ratios greater than 1:1. [52] Therefore, the particles obtained at a w/w ratio of 4:1 and 2:1 were of particular relevance. 10 kDa and 70 kDa PEI-Suc 10% particles were on average the smallest at a w/w ratio of 4:1 with 174.4 nm and 176.8 nm, respectively, whereas 25 kDa PEI-Suc 10% particles exhibited the smallest average diameter of 238.9 nm at a w/w ratio of 2:1. In regard to particle size distribution, each PDI obtained surpassed the threshold of 0.2, indicating polydispersity. Nonetheless, the nanoparticles with a 4:1 and 2:1 w/w ratio would both be adequate for testing Cas9 nuclease activity during in vitro cell assays.

***Figure 26**: Zeta potential of PEI-Suc 10% Cas9/sgGFP RNP particles*

Figure 26 depicts a diagram showing the average zeta potential of each PEI-Suc 10% Cas9 RNP particle type. A tendency of the zeta potential to decrease with declining w/w ratio, hence with decreasing amount of positively charged polymer while the amount of negatively charged RNPs remained constant, was observed. However, the 25 kDa PEI-Suc 10% particles were an exception, as the zeta potential of -3.49 mV at a w/w ratio of 0.5:1 was higher than at a w/w ratio of 1:1 with -7.89 mV. Interestingly, the PEI-Suc 10% RNP particles with a w/w ratio of 0.5:1 as well as 1:1 showed negative zeta-potentials. While an excess of negatively charged RNP complexes is associated with the particles at a w/w ratio of 0.5:1, [52] it has to be tested whether the negative zeta potential of the particles at a w/w ratio of 1:1 is related to the presence of an RNP subpopulation. Most nanoparticles with a w/w ratio of 2:1 and 4:1 exhibited a positive zeta potential by effectively binding negatively charged RNP complexes.

3.5.2. Characterization of Cas9 oligopeptide ribonucleoprotein carriers

Cas9 oligopeptide UV_1 RNP carriers were prepared in HBG. Roughly outlined, pre-incubated equimolar Cas9/sgGFP RNP solutions were added to appropriate dilutions of oligopeptide UV_1 at different N/P ratios (oligopeptide to sgRNA) of 3, 6, 12, 16, 24 equivalent to w/w

ratios of 4.125:1, 8.25:1, 16.5:1, 22:1, 33:1. Additionally, two samples were prepared using RNP solutions with molar ratios (Cas9/sgRNA) of 1:1.5 and 1:2. By mixing 1:1 (v/v) with the same UV_1 solution used to achieve an N/P ratio of 24 with an equimolar RNP solution, N/P ratios of 16 and 12, respectively, were attained.

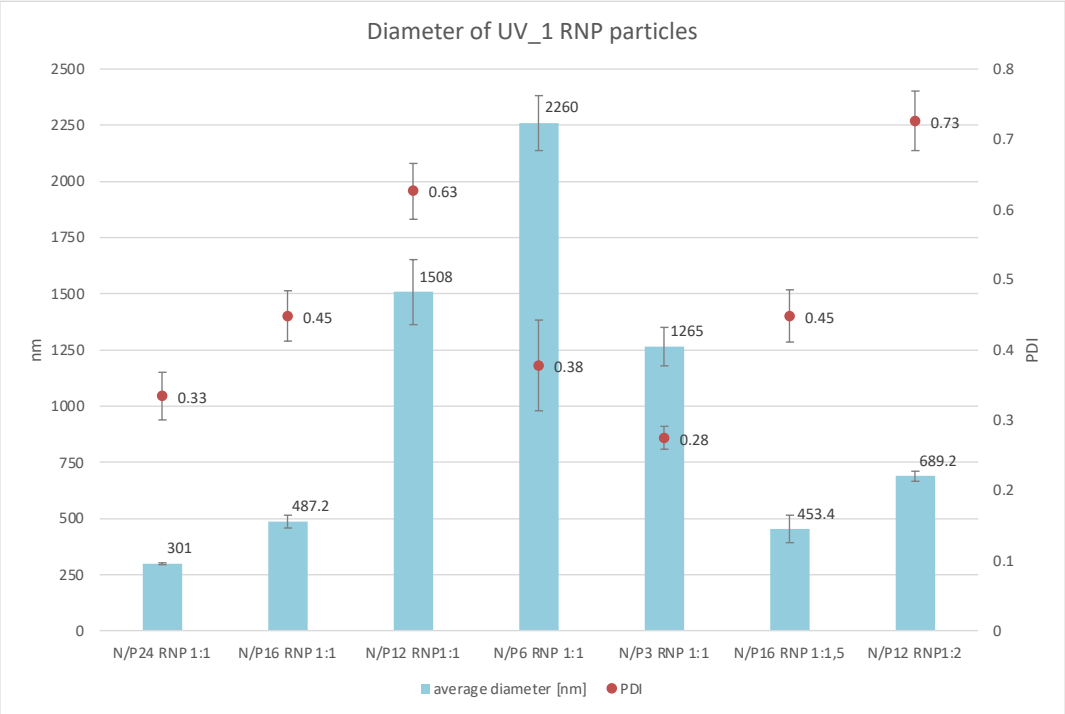

Figure 27: Particle size of UV_1 Cas9/sgGFP RNP particles

Figure 27 shows the average diameters and PDIs of the differently composed UV_1 Cas9/sgGFP RNP particles. N/P ratios of 12, 6 and 3 with equimolar RNP ratio resulted in microparticles. Nanoparticles, on the other hand, were yielded by N/P ratios of 24 and 16 with equimolar RNP ratio, N/P ratio of 16 with RNP ratio of 1:1.5, and N/P ratio of 12 with RNP ratio of 1:2. The average diameter tended to rise with decreasing N/P ratio, as well as with increasing RNP ratio. Furthermore, the particle size correlates with the average zeta potential, which is illuminated in the next paragraph. The smallest particles, with a size of 301 nm on average, were obtained at an N/P ratio of 24 with an equimolar RNP ratio and were therefore employed in cell assays. Particles with an N/P ratio of 16 were the next smallest, exhibiting similar average diameters despite different RNP ratios.

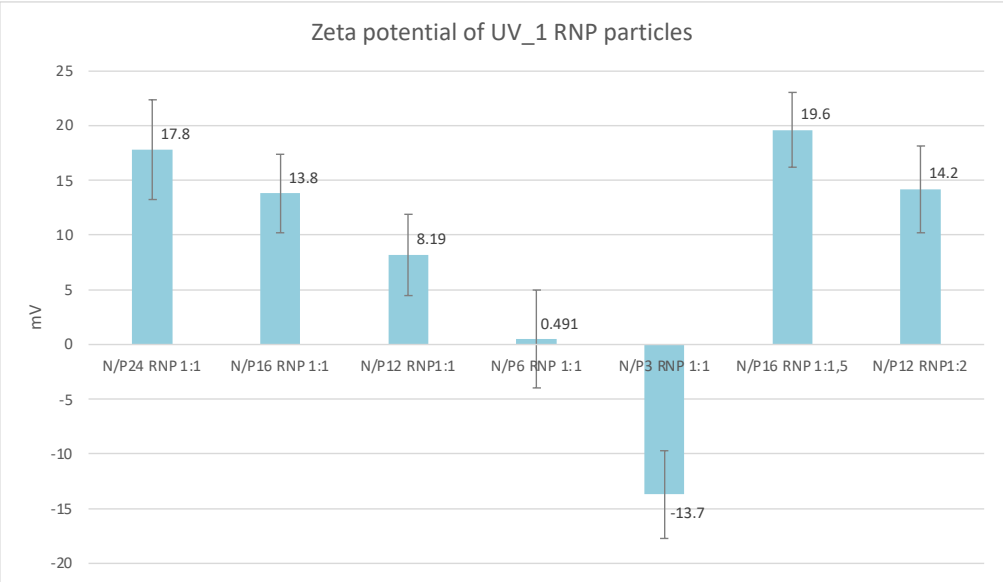

***Figure 28**: Zeta potential of UV_1 Cas9/sgGFP RNP particles*

The average zeta potential of the Cas9 UV_1 RNP particles with varying N/P and RNP ratios, summarized in **Figure 28**, ranged from 19.6 to -13.7 mV. As the positively charged oligopeptide possibly complexed the negatively charged RNPs, the zeta potential rose with increasing N/P ratio. Accordingly, the average zeta potential at an equimolar RNP ratio was 17.8 mV at an N/P ratio of 24, fell with decreasing N/P ratio, and ultimately reached -13.7 mV at an N/P ratio of 3. Despite the negative zeta potential of particles with an N/P ratio of 3 and an equimolar RNP ratio, the presence of a subpopulation of negatively charged RNP complexes cannot be assumed without performing appropriate tests. Compared to equimolar RNP ratios, lower RNP ratios exhibited much higher zeta potentials at the same N/P ratio. Regarding the correlation of zeta potential with particle size, high zeta potentials within the positive range seemed to be associated with small diameters due to repulsive electrostatic forces. The particles with a nearly neutral zeta potential, characterized by an N/P ratio of 6 and an equimolar RNP ratio, were also the largest microparticles with an average diameter of 2260 nm, which could be related to agglomeration of uncharged particles. Interestingly, the particles with RNP ratios of 1:1.5 and 1:2 showed larger average diameters than particles with similar zeta potential but equimolar RNP ratios.

3.6. Cellular delivery

Using a trial-and-error approach, various parameters were probed to establish an in vitro cell assay for Cas9/sgGFP RNP carriers that is based on the gene knockout of GFP in HeLa GFP-Tub

cells and subsequent analysis via flow cytometry. For the latter, dead cells, appropriate cell number and dispersant volume size had to be taken into account. The biggest challenge, however, was to provide sufficient time for the knockout of GFP genes to be implemented at the protein level. On this account, cell seeding density as well as incubation time were varied, and a cell transfer step to larger wells was added. As only equimolar Cas9/sgRNA RNP ratios were used for experiments on cellular delivery, the RNP ratios will not be explicitly mentioned again for ease of reading.

Whilst varying the parameters listed above, HeLa GFP-Tub cells were invariably seeded in 96-well microplates and incubated for 24 hours before treatment. In initial experiments, cell monolayers of 80-90% confluence, product of a seeding density of 10,000 cells/well, were subsequently incubated with Cas9/sgGFP nanoparticles for 24, 48 and 72 hours. Neither RNP nanocarriers of 10 kDa or 25 kDa PEI-Suc 10% nor oligopeptide UV_1 RNP nanocarriers showed detectable GFP knockout. In a different approach, 2,500 and 5,000 cells/well were seeded, resulting in 20% and 40-50% confluence, respectively. Ensuing treatment with UV_1 RNP nanoparticles for 24 hours at RNP concentrations of 37.5 and 75 nM was followed by medium change and incubation for another 48 hours. Only the treatment of the 20% confluent cell monolayers with UV_1 RNP particles at an RNP concentration of 75 nM resulted in a detectable GFP knockout.

By introducing a transfer of cells from 96-well to 24-well microplates, incubation of additional 72 hours was possible. After seeding 5,000 cells/well in a 96-well microplate followed by 24 hours of incubation and 48 or 72 hours of treatment, the monolayers that had formed were detached. After transferring the cell suspensions into a 24-well microplate, the HeLa GFP-Tub cells were able to grow for 72 hours, forming monolayers. Both treatment durations showed explicit GFP knockout. Because of the toxicity of the oligomers, exposure to the Cas9 RNP nanocarriers for 48 hours was preferred to 72 hours, and even achieved better results.

For analysis with the Gallios Flow Cytometer, the HeLa GFP-Tub cells were detached and suspended in colorless PBS-EDTA with 2% FCS. The cell count per sample was set to 3,000 cells for 96-well and 5,000 cells for 24-well microplates. With total cell suspension volumes of 520 µL for 96-well and 600 µL for 24-well microplates, this was usually reached. The final cell assay protocol, as described in 2.11, included using the nucleic acid stain propidium iodide (PI) for detection of dead cells. Propidium iodide does not permeate living cells and emits severalfold increased red fluorescence when intercalated between base pairs of nucleic acids. Exclusion

of this subpopulation of PI^+ cells was intended to counteract any bias of the results due to toxicity of oligomers, as treatment with toxic oligomer formulations may cause cell death-induced unspecific GFP knockout. On the basis of viable cells, percentage of GFP knockout was assessed by calculating the difference between RNP-treated GFP^+ cells and HBG-treated GFP^+ cells in relation to GFP^+ cells after HBG treatment.

$$GFP\ knockout\ [\%] = \frac{GFP^+(HBG) - GFP^+\ (treated)}{GFP^+\ (HBG)} * 100$$

For cellular delivery of equimolar Cas9/sgRNA RNPs, PEI-Suc 10% RNP particles with a w/w ratio of 4:1 were used, since branched 25 kDa PEI-Suc 9% had been found to achieve increased knockout at a w/w ratio of 4:1 (instead of 0.8) in previous gene silencing studies. [52] Moreover, UV_1 RNP particles with an N/P ratio of 24 were selected due to their small diameter.

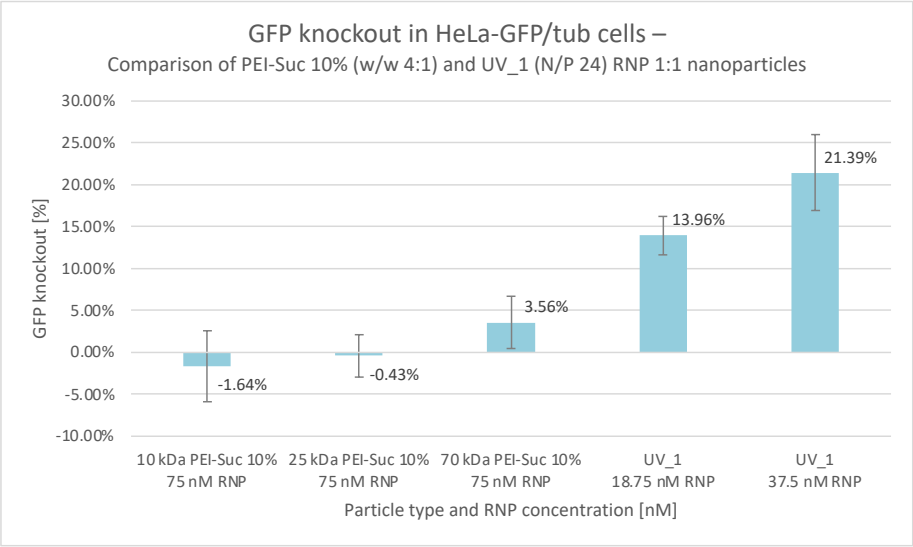

***Figure 29**: Comparison of the GFP knockout percentage obtained by PEI-Suc 10% (w/w 4:1) and UV_1 (N/P 24) RNP particles with a Cas9/sgGFP molar ratio of 1:1, determined by flow cytometry*

In **Figure 29**, GFP knockout data of PEI-Suc 10% RNP carriers at a 75 nM RNP concentration and UV_1 RNP carriers at two different RNP concentrations are summarized. Treatment with RNP nanoparticles consisting of 10 kDa PEI-Suc 10% led to a slightly increased rate of GFP^+ cells instead of a decrease. RNP nanocarrier formulations with 25 kDa PEI-Suc 10% also did not show any GFP knockout in HeLa GFP-Tub cells, despite successful transfection of siRNA with 25 kDa PEI-Suc 10% in previous studies. [52] This result is in accordance with previous GFP knockout experiments of Cas9 RNP vehicles in Neuro2a eGFP/luc cells [53]. 70 kDa PEI-Suc 10% RNP nanoparticles mediated GFP knockout of 3.56%. However, samples containing this polymer showed limited growth, as evidenced by a 10% confluence observed in 24-well plates

and an average cell count of 2,340 instead of anticipated 5,000 cells per well during flow cytometry analysis. This suggests an underlying oligomer toxicity, which might affect the calculation of GFP knockout efficiency. In contrast, distinct GFP knockout was demonstrated by Cas9 RNP formulations comprising the oligopeptide UV_1. At a concentration of 18.75 nM RNP, GFP knockout of 13.96% was detected, which increased to 21.39% at 37.5 nM RNP.

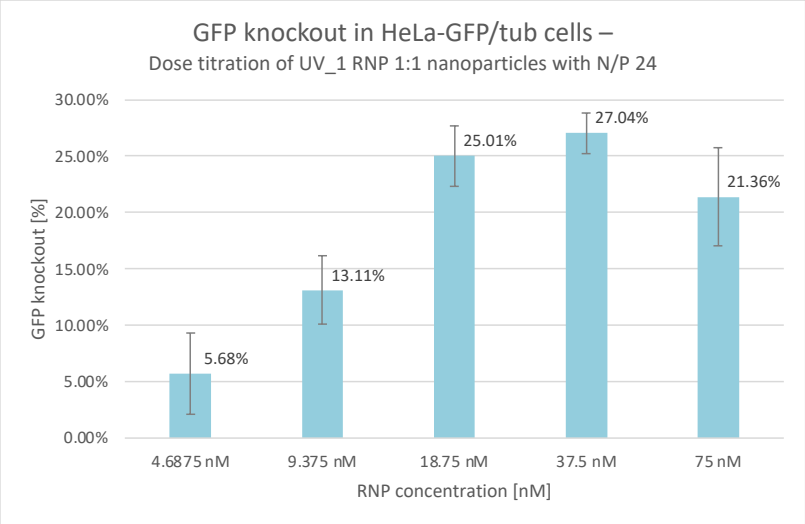

***Figure 30**: Percentage of GFP knockout in HeLa GFP-Tub cells obtained by dose titration of UV_1 RNP nanoparticles with an N/P ratio of 24 and a molar ratio of Cas9/sgGFP of 1:1, determined by flow cytometry*

As the oligopeptide UV_1 was able to achieve GFP knockout at low RNP concentrations, a dose titration was performed. The results of the UV_1-mediated Cas9 RNP cellular delivery at different RNP concentrations and a constant N/P ratio of 24 are displayed in **Figure 30** as GFP knockout percentages (±SD). With increasing RNP concentrations, the percentage of GFP knockout also increased. The initial doubling of the knockout at twice the concentration, from 5.68% at 4.69 nM to 13.11% at 9.38 nM to 25.01% at 18.75 nM RNP, stagnated and reached a peak GFP knockout of 27.04% at 37.5 nM RNP. The UV_1 RNP particles with an RNP concentration of 75 nM induced a 21.36% GFP knockout, yielding lower knockout efficiency than 18.75 and 37.5 nm RNP. In regard of GFP knockout efficiency and oligomer toxicity, the 18.75 nM RNP concentration of UV_1 Cas9/sgGFP RNP nanoparticles was preferred to 37.5 nM and generally the most optimal.

When using HeLa GFP-Tub cells, it must be taken into account that the expression of GFP-fused tubulin might be associated with a slower proliferation rate. Consequently, the proportion of the GFP^+ to the GFP^- population decreases with increasing number of cell passages. Cas9/sgGFP-mediated GFP knockout is thus additionally promoted by natural selection in favor of the GFP^- population. [53]

4. Conclusion

Effectively, a workflow for the analysis of nanocarriers for intracellular Cas9/sgRNA RNP delivery at the Department of Pharmaceutical Sciences, Division of Pharmaceutical Technology and Biopharmaceutics of the University of Vienna was established. Besides the production of recombinant Cas9, an in vitro cleavage assay to determine Cas9 endonuclease activity based on the cleavage of an IVC EGFP gene (1.5 kb) and an in vitro cell assay to assess GFP gene knockout in HeLa GFP-Tub cells were set up, by which additional insights concerning Cas9 stability and RNP delivery carriers were gained.

Stability issues of the recombinantly expressed Cas9 proteins were revealed by in vitro cleavage assays after one, three, and five freeze-thaw cycles. The first batch of Cas9 showed a sharp decline in endonuclease activity with increasing freeze-thaw cycles, ultimately mediating merely 17.0% cleavage. By substitution of manual with automated cell disruption and purification, allowing 4 °C to be maintained during purification, and final addition of 20% glycerol, Cas9 activity was elevated to a steady cleavage of 86.5-88.4%. Without added glycerol, a minimal decrease in Cas9 activity was observed with increasing freeze-thaw cycles, demonstrating its effect as a cryoprotectant.

For the delivery of Cas9/sgGFP RNP complexes, two types of carrier materials, either PEIs with average molecular weight of 10, 25 or 70 kDa that were 10% succinylated (PEI-Suc 10%) or the oligopeptide UV_1, were used at varying w/w or N/P ratios (carrier material to sgRNA) to form nano- or microparticles with RNP complexes of different molar RNP ratios (Cas9 to sgRNA). Based on their characterization, Cas9/sgGFP RNP nanoparticles with an equimolar RNP ratio were prepared for GFP knockout cell assays using 10, 25 and 70 kDa PEI-Suc 10% at a w/w ratio of 4:1 and the oligopeptide UV_1 at an N/P ratio of 24. While the PEI-Suc 10% RNP particles were not able to mediate distinct GFP knockout at an RNP concentration of 75 nM, the UV_1 RNP particles showed promising knockout at 18.75 and 37.5 nM RNP concentrations. In a subsequent dose titration of this UV_1 RNP particle formulation, the knockout effect initially increased with increasing ribonucleoprotein concentrations, culminated in stagnation at 37.5 nM and then decreased again at 75 nM, indicating that 18.75 nM might be the most optimal of the tested RNP concentrations.

5. Abbreviations

AAV ..... adeno-associated virus

APS ..... ammonium persulfate

AV ..... adenovirus

Cas ..... CRISPR-associated

Cascade ..... CRISPR-associated complex for antiviral defense

conc. ..... concentration

CPP ..... cell-penetrating peptide

CRISPR ..... clustered regularly interspaced short palindromic repeats

crRNA … CRISPR RNA

CTD ..... C-terminal domain

CV ..... column volume

dCas9 ..... nuclease-deactivated/nuclease-deficient Cas9

DMD ..... duchenne muscular dystrophy

DMEM ..... Dulbecco's Modified Eagle Medium

DMSO ..... dimethyl sulfoxide

DSB ..... double strand break

dsDNA ..... double stranded DNA

DTT ..... dithiothreitol

EDTA ..... ethylenediaminetetraacetic acid

FBS ..... fetal bovine serum

gRNA ..... guide RNA

HBG ..... HEPES buffered glucose

HDR ..... homology directed repair

HIV ..... human immunodeficiency virus

HSV-1 ..... herpes simplex virus type 1

IVC ..... in vitro cleavage

LB ..... luria broth

lipo-OAA ..... lipid-containing oligo(ethylenamino) amides

LNP ..... lipid nanoparticle

LV ..... lentivirus

MW ..... molecular weight

MWCO ..... molecular weight cutoff

nCas9 ..... Cas9 nickase

NHEJ ..... non-homologous end joining

NLS ..... nuclear localization signal

NMWL ..... nominal molecular weight limit

NUC lobe ..... nuclease lobe

PAM ..... protospacer-adjacent motif

PBAE ..... poly(β-amino ester)

PDI ..... polydispersity index

pDNA ..... plasmid DNA

PEG ..... polyethylene glycol

PEI ..... polyethylenimine

PEI-Suc 10% ..... 10% succinylated PEI

PI ..... propidium iodide

PI domain ..... PAM-interacting domain

PMSF ..... phenylmethylsulfonyl fluoride

pre-crRNA … precursor CRISPR RNA

REC lobe ..... recognition lobe

RNP ..... ribonucleoprotein

SaCas9 ..... *Staphylococcus aureus* Cas9

SCP ..... supercharged polypeptide

SD ..... standard deviation

SDS ..... Sodium dodecyl sulfate

SEC ..... size exclusion chromatography

sgRNA ..... single guide RNA

siRNA ..... small interfering RNA

SORT ..... selective organ targeting

SpyCas9 ..... *Streptococcus pyogenes* Cas9

ssDNA ..... single stranded DNA

TEMED ..... Tetramethylethylenediamine

T_m ..... melting temperature

TOPO domain ..... topoisomerase-homology domain

tracrRNA ..... trans-activating CRISPR RNA

TRIS ..... tris(hydroxymethyl)aminomethane

UV_1 ..... an artificial oligopeptide

WED domain ..... wedge domain

ZIF ..... zeolitic imidazole framework

6. Table of figures

7. References

1. Marraffini LA. CRISPR-Cas immunity in prokaryotes. *Nature* 2015; 526: 55–61.

2. Makarova KS, Wolf YI, Iranzo J, et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol* 2020; 18: 67–83.

3. Jansen R, Embden JDA van, Gaastra W, et al. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 2002; 43: 1565–1575.

4. Mojica FJM, Díez-Villaseñor C, García-Martínez J, et al. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* 2005; 60: 174–182.

5. Makarova KS, Wolf YI, Alkhnbashi OS, et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* 2015; 13: 722–736.

6. Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012; 337: 816–821.

7. McGinn J, Marraffini LA. Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat Rev Microbiol* 2019; 17: 7–12.

8. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007; 315: 1709–1712.

9. Brouns SJJ, Jore MM, Lundgren M, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 2008; 321: 960–964.

10. Deltcheva E, Chylinski K, Sharma CM, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011; 471: 602–607.

11. Hale CR, Zhao P, Olson S, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 2009; 139: 945–956.

12. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 2008; 322: 1843–1845.

13. Wang H, La Russa M, Qi LS. CRISPR/Cas9 in Genome Editing and Beyond. *Annu Rev Biochem* 2016; 85: 227–264.

14. Ishino Y, Shinagawa H, Makino K, et al. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol* 1987; 169: 5429–5433.

15. Mojica FJ, Díez-Villaseñor C, Soria E, et al. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol* 2000; 36: 244–246.

16. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014; 157: 1262–1278.

17. Sapranauskas R, Gasiunas G, Fremaux C, et al. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Res* 2011; 39: 9275–9282.

18. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013; 339: 819–823.

19. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science* 2013; 339: 823–826.

20. Gupta D, Bhattacharjee O, Mandal D, et al. CRISPR-Cas9 system: A new-fangled dawn in gene editing. *Life Sci* 2019; 232: 116636.

21. Chaudhuri A, Halder K, Datta A. Classification of CRISPR/Cas system and its application in tomato breeding. *TAG Theor Appl Genet Theor Angew Genet* 2022; 135: 367–387.

22. Koonin EV, Makarova KS. Origins and evolution of CRISPR-Cas systems. *Philos Trans R Soc Lond B Biol Sci* 2019; 374: 20180087.

23. Jiang F, Doudna JA. CRISPR–Cas9 Structures and Mechanisms. *Annu Rev Biophys* 2017; 46: 505–529.

24. Nidhi S, Anand U, Oleksak P, et al. Novel CRISPR-Cas Systems: An Updated Review of the Current Achievements, Applications, and Future Research Perspectives. *Int J Mol Sci* 2021; 22: 3327.

25. Özcan A, Pausch P, Linden A, et al. Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. *Nat Microbiol* 2019; 4: 89–96.

26. Fonfara I, Richter H, Bratovič M, et al. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 2016; 532: 517–521.

27. East-Seletsky A, O'Connell MR, Knight SC, et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* 2016; 538: 270–273.

28. Liu L, Li X, Wang J, et al. Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. *Cell* 2017; 168: 121-134.e12.

29. Faure G, Shmakov SA, Makarova KS, et al. Comparative genomics and evolution of trans-activating RNAs in Class 2 CRISPR-Cas systems. *RNA Biol* 2019; 16: 435–448.

30. Jiang F, Taylor DW, Chen JS, et al. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* 2016; 351: 867–871.

31. Zhu X, Clarke R, Puppala AK, et al. Cryo-EM structures reveal coordinated domain motions that govern DNA cleavage by Cas9. *Nat Struct Mol Biol* 2019; 26: 679–685.

32. Anders C, Niewoehner O, Duerst A, et al. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 2014; 513: 569–573.

33. Jinek M, Jiang F, Taylor DW, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 2014; 343: 1247997.

34. Nishimasu H, Cong L, Yan WX, et al. Crystal Structure of Staphylococcus aureus Cas9. *Cell* 2015; 162: 1113–1126.

35. Heler R, Samai P, Modell JW, et al. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 2015; 519: 199–202.

36. Sternberg SH, Redding S, Jinek M, et al. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 2014; 507: 62–67.

37. Zhang S, Shen J, Li D, et al. Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics* 2021; 11: 614–648.

38. Jacinto FV, Link W, Ferreira BI. CRISPR/Cas9-mediated genome editing: From basic research to translational medicine. *J Cell Mol Med* 2020; 24: 3766–3778.

39. Mashel TV, Tarakanchikova YV, Muslimov AR, et al. Overcoming the delivery problem for therapeutic genome editing: Current status and perspective of non-viral methods. *Biomaterials* 2020; 258: 120282.

40. Lin Y, Wagner E, Lächelt U. Non-viral delivery of the CRISPR/Cas system: DNA versus RNA versus RNP. *Biomater Sci* 2022; 10: 1166–1192.

41. Lino CA, Harper JC, Carney JP, et al. Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv* 2018; 25: 1234–1257.

42. Sinclair F, Begum AA, Dai CC, et al. Recent advances in the delivery and applications of nonviral CRISPR/Cas9 gene editing. *Drug Deliv Transl Res* 2023; 13: 1500–1519.

43. Shojaei Baghini S, Gardanova ZR, Abadi SAH, et al. CRISPR/Cas9 application in cancer therapy: a pioneering genome editing tool. *Cell Mol Biol Lett* 2022; 27: 35.

44. Horodecka K, Düchler M. CRISPR/Cas9: Principle, Applications, and Delivery through Extracellular Vesicles. *Int J Mol Sci* 2021; 22: 6072.

45. Khoshandam M, Soltaninejad H, Mousazadeh M, et al. Clinical applications of the CRISPR/Cas9 genome-editing system: Delivery options and challenges in precision medicine. *Genes Dis* 2024; 11: 268–282.

46. Zhang C, Quan R, Wang J. Development and application of CRISPR/Cas9 technologies in genomic editing. *Hum Mol Genet* 2018; 27: R79–R88.

47. Chung JJ, Fujita Y, Li S, et al. Biodegradable inorganic-organic hybrids of methacrylate star polymers for bone regeneration. *Acta Biomater* 2017; 54: 411–418.

48. Lee YS, Kim SW. Bioreducible polymers for therapeutic gene delivery. *J Control Release Off J Control Release Soc* 2014; 190: 424–439.

49. Choi YJ, Kang SJ, Kim YJ, et al. Comparative studies on the genotoxicity and cytotoxicity of polymeric gene carriers polyethylenimine (PEI) and polyamidoamine (PAMAM) dendrimer in Jurkat T-cells. *Drug Chem Toxicol* 2010; 33: 357–366.

50. Liu C, Wan T, Wang H, et al. A boronic acid-rich dendrimer with robust and unprecedented efficiency for cytosolic protein delivery and CRISPR-Cas9 gene editing. *Sci Adv* 2019; 5: eaaw8922.

51. Taharabaru T, Yokoyama R, Higashi T, et al. Genome Editing in a Wide Area of the Brain Using Dendrimer-Based Ternary Polyplexes of Cas9 Ribonucleoprotein. *ACS Appl Mater Interfaces* 2020; 12: 21386–21397.

52. Zintchenko A, Philipp A, Dehshahri A, et al. Simple modifications of branched PEI lead to highly efficient siRNA carriers with low toxicity. *Bioconjug Chem* 2008; 19: 1448–1455.

53. Kuhn J, Lin Y, Krhac Levacic A, et al. Delivery of Cas9/sgRNA Ribonucleoprotein Complexes via Hydroxystearyl Oligoamino Amides. *Bioconjug Chem* 2020; 31: 729–742.

54. Zhu X, Lv M-M, Liu J-W, et al. DNAzyme activated protein-scaffolded CRISPR-Cas9 nanoassembly for genome editing. *Chem Commun Camb Engl* 2019; 55: 6511–6514.

55. Pan X, Pei X, Huang H, et al. One-in-one individual package and delivery of CRISPR/Cas9 ribonucleoprotein using apoferritin. *J Control Release Off J Control Release Soc* 2021; 337: 686–697.

56. Hamilton JR, Tsuchida CA, Nguyen DN, et al. Targeted delivery of CRISPR-Cas9 and transgenes enables complex immune cell engineering. *Cell Rep* 2021; 35: 109207.

57. Shi J, Yang X, Li Y, et al. MicroRNA-responsive release of Cas9/sgRNA from DNA nanoflower for cytosolic protein delivery and enhanced genome editing. *Biomaterials* 2020; 256: 120221.

58. Ding F, Huang X, Gao X, et al. A non-cationic nucleic acid nanogel for the delivery of the CRISPR/Cas9 gene editing tool. *Nanoscale* 2019; 11: 17211–17215.

59. Ryu J-Y, Won E-J, Lee HAR, et al. Ultrasound-activated particles as CRISPR/Cas9 delivery system for androgenic alopecia therapy. *Biomaterials* 2020; 232: 119736.

60. Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* 2017; 168: 20–36.

61. Ledford H. Is CRISPR safe? Genome editing gets its first FDA scrutiny. *Nature* 2023; 623: 234–235.

62. Sheridan C. The world's first CRISPR therapy is approved: who will receive it? *Nat Biotechnol*. Epub ahead of print 21 November 2023. DOI: 10.1038/d41587-023-00016-6.

63. Reardon S. FDA Approves First CRISPR Gene Editing Treatment for Sickle Cell Disease. *Scientific American*, https://www.scientificamerican.com/article/fda-approves-first-crispr-gene-editing-treatment-for-sickle-cell-disease/ (accessed 14 December 2023).

64. Liang Z, Chen K, Li T, et al. Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat Commun* 2017; 8: 14261.

65. de Wet JR, Wood KV, DeLuca M, et al. Firefly luciferase gene: structure and expression in mammalian cells. *Mol Cell Biol* 1987; 7: 725–737.

66. Lin Y, Wilk U, Pöhmerer J, et al. Folate Receptor-Mediated Delivery of Cas9 RNP for Enhanced Immune Checkpoint Disruption in Cancer Cells. *Small Weinh Bergstr Ger* 2023; 19: e2205318.

67. Kuipers BJH, Gruppen H. Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. *J Agric Food Chem* 2007; 55: 5445–5451.

68. Gottlieb HE, Kotlyar V, Nudelman A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J Org Chem* 1997; 62: 7512–7515.