



**VILNIUS UNIVERSITY
FACULTY OF CHEMISTRY AND GEOSCIENCES
INSTITUTE OF CHEMISTRY**

Urtė Glibauskaitė

Biochemistry
Bachelor thesis

Cloning and Cell-Free Expression of CRISPR-Cas Nucleases

Supervisor
Dr. Rokas Grigaitis

Vilnius 2023

CONTENTS

ABBREVIATIONS.....	3
INTRODUCTION.....	4
1. LITERATURE REVIEW.....	5
1.1 CRISPR-Cas systems and their classification	5
1.2 CRISPR-Cas mechanisms of action	7
1.2.1 Cas9	11
1.2.2 Cas12a	14
1.3 Overview of CRISPR-Cas genome editing	15
1.3.1 Limitations of CRISPR-Cas genome editing.....	17
1.3.2 Clinical trials	18
1.4 Cell-free protein synthesis	19
1.4.1 CFPS in <i>E. Coli</i> cell-free extracts.....	20
1.4.2 Characterization of CRISPR-Cas nucleases through CFPS	21
2. MATERIALS AND METHODS	23
2.1 Materials	23
2.1.1 Bacterial strains	23
2.1.2 Plasmids.....	23
2.1.3 Oligonucleotides.....	24
2.2 Methods	26
2.2.1 Golden Gate assembly	26
2.2.2 Blunt-end cloning	27
2.2.3 Quick-change mutagenesis.....	27
2.2.4 Preparation of linear gRNA templates	28
2.2.5 Cell-free protein synthesis	29
2.2.6 Protein purification.....	30
2.2.7 Western blot.....	30
3. RESULTS AND DISCUSSION	32
3.1 CRISPR-Cas nucleases selected for the study	32
3.2 Vector design	33
3.3 Cloning of nucleases and gRNA templates	34
3.4 Cell-free protein synthesis	38
3.5 Purification of AsCas12a	41
3.6 Discussion.....	43
CONCLUSIONS	46
SUMMARY	47
ACKNOWLEDGMENTS	48
REFERENCES	49

ABBREVIATIONS

APS – ammonium persulfate
bp – base pair
Cas – CRISPR-associated
CFE – cell-free extract
CFPS – cell-free protein synthesis
CRISPR – clustered regularly interspaced short palindromic repeats
crRNA – CRISPR ribonucleic acid
dNTP – deoxynucleotide triphosphate
DSB – double-stranded break
DTT – dithiothreitol
EDTA – ethylenediaminetetraacetic acid
gRNA – guide ribonucleic acid
HDR – homology-directed repair
HRP – horseradish peroxidase
IPTG – isopropyl β -D-1-thiogalactopyranoside
kb – kilobase
kDa – kilodalton
LB – Luria-Bertani (medium)
mRNA – messenger ribonucleic acid
N – A or G or C or T nucleotide
NHEJ – non-homologous end joining
nt – nucleotide
PAGE – polyacrylamide gel electrophoresis
PBS – phosphate-buffered saline (buffer)
PCR – polymerase chain reaction
PEG – polyethylene glycol
pre-crRNA – precursor CRISPR ribonucleic acid
PVDF – polyvinylidene fluoride
QCM – quick-change mutagenesis
rRNA – ribosomal ribonucleic acid
SDS – sodium dodecyl sulfate
sfGFP – superfolder green fluorescent protein
sgRNA – single guide ribonucleic acid
StrepII – protein purification tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys)
TAE – tris-acetate-EDTA
TBE – tris-borate-EDTA
TEMED – tetramethylethylenediamine
tracrRNA – trans-activating CRISPR ribonucleic acid
V – G or C or A nucleotide
Y – C or T nucleotide

INTRODUCTION

CRISPR (clustered regularly interspaced short palindromic repeat) locus and Cas (CRISPR associated) proteins make up an adaptive immune system against bacteriophages and other mobile genetic elements in bacteria and archaea (Barrangou et al., 2007; Mojica et al., 2005). Upon bacteriophage infection, a short fragment of viral DNA is integrated into the CRISPR (clustered regularly interspaced short palindromic repeat) locus of bacteria or archaea. If the infection recurs, this locus is transcribed and processed into a crRNA (CRISPR RNA) molecule. crRNA then forms a single- or multi-subunit ribonucleoprotein effector complex that recognizes and cleaves complementary phage DNA or RNA.

crRNA-dependent nature of cleavage allows for programmable nucleic acid targeting (Jinek et al., 2012). For this reason, CRISPR-Cas nucleases have been adopted for genome editing (Jinek et al., 2012). However, CRISPR-Cas genome editing has several inherent limitations, the most important of which are related to off-target binding and cleavage (Zischewski et al., 2017). Luckily, numerous phylogenetically diverse CRISPR-Cas nucleases have been identified by bioinformatic means (Makarova et al., 2020). One way to solve the mentioned problems is to find new, more precise CRISPR-Cas nucleases. However, their characterization process requires long and tedious experimental procedures, such as cell cultivation and protein purification (Karvelis et al., 2017; Leenay & Beisel, 2017). To speed up the process, the group of Dr. Stephen Knox Jones Jr. at VU LSC-EMBL Partnership Institute for Genome Editing Technologies is aiming to develop a high-throughput strategy to characterize CRISPR-Cas nucleases. This strategy involves cell-free protein synthesis (CFPS) and microplate-based protein purification. To develop and ensure the efficiency of the characterization process, we selected ten CRISPR-Cas nucleases identified and partially characterized in previous studies: nine Cas12a homologs from ADurb.Bin193 (Adurb193Cas12a), *Acidaminococcus* species (AsCas12a), *Lachnospiraceae* bacterium (LbCas12a), ADurb.Bin336 (Adurb336Cas12a), *Francisella novicada* (Fn3Cas12a), *Francisella novicada* U112 (FnCas12a), *Prevotella ihumii* (PiCas12a), *Prevotella disiens* (PdCas12a) and *Helcococcus kunzii* (HkCas12a) (Bernd Zetsche, 2015; Jacobsen et al., 2020; Tang et al., 2017) along with *Streptococcus pyogenes* Cas9 (SpCas9) (Jinek et al., 2012).

This work demonstrates the utilization of the mentioned high-throughput methods to express and purify the previously introduced nucleases. Consequently, I formulated the objective and tasks listed below:

Objective

1. Create a high-throughput method that allows cloning, expression, and purification of CRISPR-Cas nucleases.

Tasks

1. Clone the genes of the selected CRISPR-Cas nucleases to a customized expression vector.
2. Clone linear crRNA expression templates.
3. Conduct cell-free expression of the CRISPR-Cas nucleases.
4. Affinity-purify the CRISPR-Cas nucleases from the cell-free extracts by utilizing a microplate-based strategy.

1. LITERATURE REVIEW

1.1 CRISPR-Cas systems and their classification

CRISPR (clustered regularly interspaced short palindromic repeats) accompanied by Cas (CRISPR associated) proteins make up an adaptive immune system, first discovered in bacteria and archaea (Barrangou et al., 2007; Ishino et al., 1987; Mojica et al., 2005). Generally, CRISPR-Cas is a programmable RNA-guided system that allows specific recognition and cleavage of foreign nucleic acids and other mobile genetic elements (Jackson et al., 2017; Pinilla-Redondo et al., 2020).

The CRISPR-Cas defense contains three main stages: adaptation, crRNA biogenesis (expression), and interference. During the adaptation stage, the CRISPR-Cas system copies and integrates a small fragment of intruding DNA into the CRISPR locus. (Bhaya et al., 2011). For the expression stage, the integrated fragments, also called spacers, are used as templates to produce pre-crRNAs (CRISPR RNAs) which are processed into crRNAs and form effector complexes with Cas proteins. During the interference stage, the crRNAs facilitate Cas endonucleases to bind and cleave matching sequences.

The three stages (adaptation, expression, and interference) of the CRISPR-Cas mechanism of action and comparison of the sequences and phylogenetic analysis are used to classify the CRISPR-Cas systems. The latest classification, reported in 2020 (Makarova et al., 2020), consists of 2 classes, 6 types, 33 subtypes, and several subtype variants of CRISPR-Cas systems (Table 1.1.).

Table 1.1. CRISPR-Cas classification by types and subtypes. Type IV CRISPR-Cas system possesses an uncertain function in adaptive immunity that appears to lack targeted cleavage activity (Makarova et al., 2020).

Class	Type	Subtypes	Cas endonuclease
1	I	I-A, I-B, I-C, I-D, I-E, I-F, I-G	Cas3
	III	III-A, III-B, III-C, III-D, III-E, III-F	Cas10
	IV	IV-A, IV-B, IV-C	?
2	II	II-A, II-B, II-C	Cas9
	V	V-A, V-B, V-C, V-D, V-E, V-F, V-G, V-H, V-I, V-K, V-U	Cas12
	VI	VI-A, VI-B, VI-C, VI-D	Cas13

Based on the amount of proteins, comprising the effector complex, CRISPR-Cas systems are divided into Class 1 and Class 2 (Makarova et al., 2015) (Fig. 1.1 A). Effector complexes of Class 1 CRISPR-Cas systems are composed of multiple proteins. Comparatively, a single crRNA-binding protein is a characteristic of Class 2 systems. Based on distinct Cas proteins and molecular mechanisms, each CRISPR-Cas class consists of three types (Fig. 1.1. B). According to the differences in CRISPR locus organization and adaptation, expression, and interference schemes, each type is further divided into subtypes (Makarova et al., 2020).

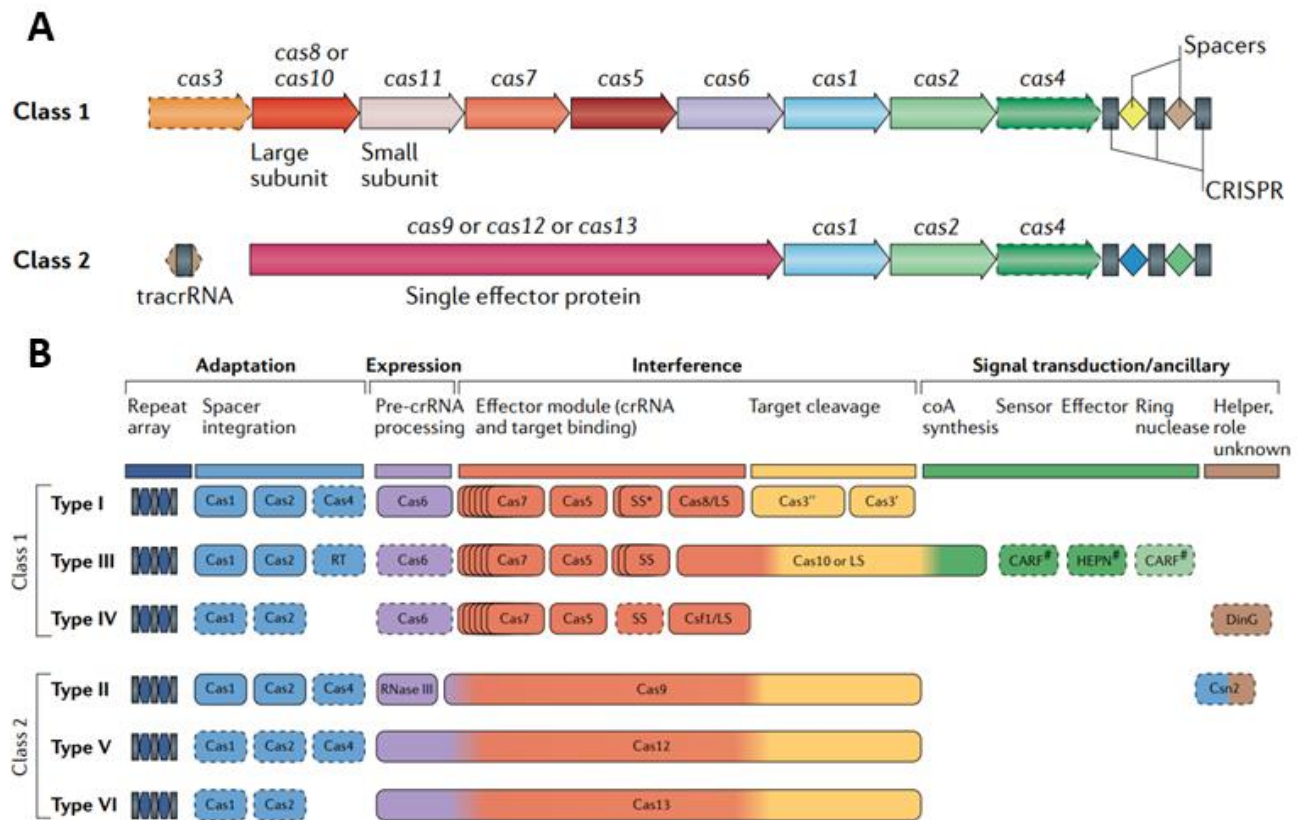


Figure 1.1. Classification of the CRISPR-Cas systems. **A)** Organization of class 1 and class 2 CRISPR-Cas loci. **B)** Classification of the CRISPR-Cas systems based on their genetic, functional, and structural relationships. Dashed outlines indicate missing or dispensable components in some subtypes. An asterisk (*) represents a potential fusion of the protein to a large subunit in some subtypes. A hash (#) indicates that other unknown proteins could be involved in the signalling pathway. SS – small subunit, LS – large subunit, RS – reverse transcriptase. Adapted from (Makarova et al., 2020), tracrRNA – trans-activating RNA.

As mentioned before, to classify CRISPR-Cas systems, CRISPR-Cas immunity stages are used to divide the *cas* genes into functional modules (Makarova et al., 2013, 2020) (Fig. 1.1 B). In most CRISPR-Cas types, the adaptation module includes a Cas1 integrase, an important enzyme for spacer acquisition, and a Cas2 subunit, which is a part of the adaptation complex (Mosterd et al., 2021; Silas et al., 2016). Several CRISPR-Cas subtypes also include a Cas4 nuclease, the Csn2 protein is included in subtype II-A, and reverse transcriptase is found in some type III CRISPR-Cas systems (Staals et al., 2013). Next, pre-crRNA processing is done by proteins, included in the expression module. pre-crRNA processing in class 1 systems is mostly done by Cas6 (Taylor et al., 2019). Bacterial RNase III is responsible for processing in type II systems (Charpentier et al., 2015; Deltcheva et al., 2011). However, the processing of pre-crRNAs in all type VI and most type V systems is done by the effector Cas proteins, containing intrinsic RNase activity (East-Seletsky et al., 2016; Fonfara et al., 2016). Finally, target recognition and cleavage are mediated by the interference (effector) module. The effector module of class 1 CRISPR-Cas systems consists of several proteins which can be combined differently, depending on types and subtypes (T. Y. Liu & Doudna, 2020). These proteins include Cas3, Cas5-Cas8 as well as Cas10 and Cas11. On the contrary, the effector module of class 2 systems contains single proteins, namely Cas9, Cas12, and Cas13 (Makarova et al., 2020). Notably, although class 2 systems contain single proteins in their effector complexes, Cas12f forms a dimeric structure

(Karvelis et al., 2020). However, Cas12f is attributed to class 2 systems due to retained activation mechanism, characteristic among other similar nucleases (e.g., Cas12a).

In summary, all CRISPR-Cas systems share common functional and architectural principles and have signature elements in the *cas* loci. These elements are being used as a reference to classify the new varieties of CRISPR-Cas systems and further understand their origins and properties.

1.2 CRISPR-Cas mechanisms of action

To comprehend CRISPR-Cas mechanisms of action, it is important to understand the structure of the CRISPR locus, composed of *cas* (CRISPR-associated) genes and a CRISPR array (Fig. 1.2). The *cas* genes, found upstream to the CRISPR array, encode Cas proteins. The *cas* genes are followed by the CRISPR array, containing DNA regions, called spacers, and repetitive short DNA sequences (repeats) in between (Bolotin et al., 2005). The spacers, flanked by the repeat sequences, encode short fragments copied from foreign mobile genetic elements. Although absent in some CRISPR-Cas systems, the leader sequence separates the CRISPR array from the *cas* genes and is important in the determination of transcription direction (Rath et al., 2015).

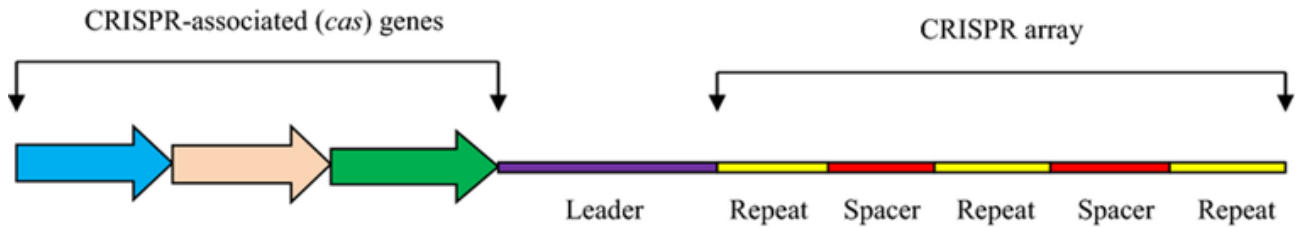


Figure. 1.2. Structure of the CRISPR locus. CRISPR-associated (*cas*) genes are located upstream of the CRISPR array. The *cas* genes and the CRISPR array are linked through the leader sequence. Adapted from (Barman et al., 2020).

Even though brief CRISPR-Cas mechanisms were described in section 1.1, here I will explain the roles of CRISPR locus elements by providing more details on the aforementioned adaptation, expression, and interference stages of CRISPR-Cas.

Adaptation. During the adaptation phase, CRISPR-Cas systems alter the CRISPR loci through a process called spacer acquisition. Spacer acquisition often involves Cas1, a highly conserved Cas protein found in all six CRISPR-Cas types (Koonin et al., 2017). Cas1 interacts with Cas2 to form a heterohexameric complex, called Cas1-Cas2 integrase (Nuñez et al., 2014, 2015). This complex contains two DNA binding regions, one that binds the intruding DNA (also called protospacer) and the other that binds the CRISPR array. Once Cas1-Cas2 integrase captures the protospacer, it catalyzes two cleavage-ligation reactions at the CRISPR array (Fig. 1.3) (Wright et al., 2017; Xiao, Ng, et al., 2017). The first cleavage-ligation reaction occurs at the leader end of the repeat. Subsequently, the second reaction occurs at the spacer end of the repeat.

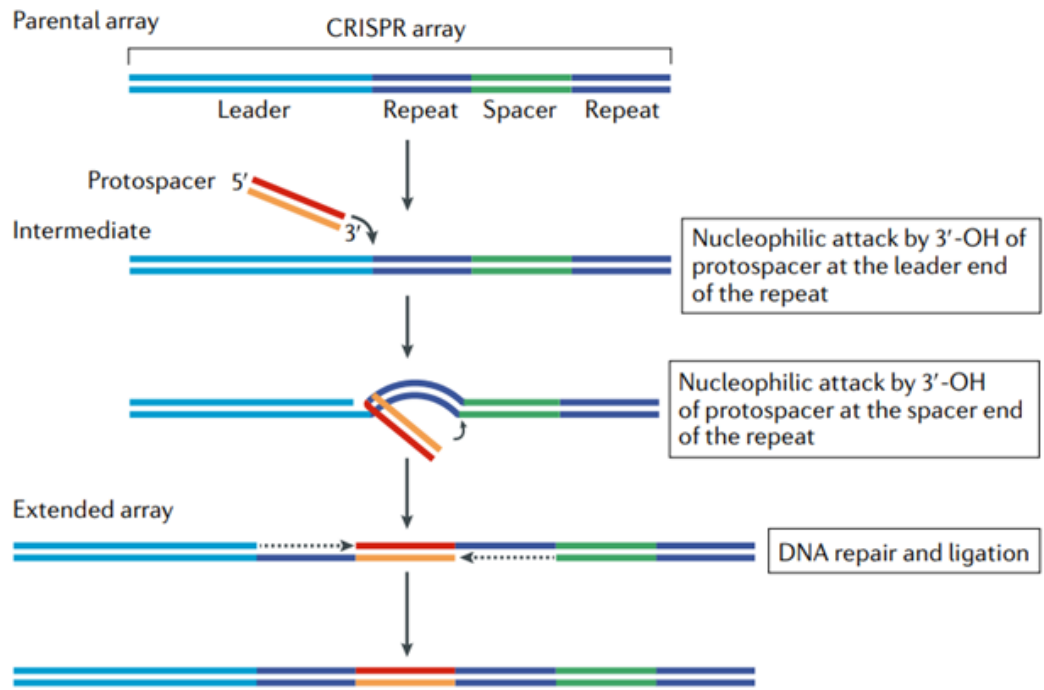


Figure 1.3. Spacer acquisition into the CRISPR locus. Following two subsequent cleavage-ligation reactions, a new spacer gets integrated at the leader end of the CRISPR array. Adapted from (McGinn & Marraffini, 2018).

Following these events, a spacer gets integrated into the leader end of the CRISPR array (Wright et al., 2017; Xiao, Ng, et al., 2017). This kind of spacer acquisition gives information about past infections since the new spacers are predominantly integrated at the proximal end of the CRISPR array (McGinn & Marraffini, 2016).

Notably, several type I, II and V systems use Cas4 nuclease for spacer acquisition, which is thought to be involved in protospacer processing and its directional integration at the CRISPR locus (Makarova et al., 2020; Shiimori et al., 2018). Also, the type III-B CRISPR-Cas system from *Marinomonas mediterranea* contains Cas1 linked to a reverse transcriptase allowing for spacer integration through reverse transcription of RNA-based invaders (Silas et al., 2016). In this case, a foreign RNA molecule is directly ligated into the CRISPR locus and is reverse-transcribed to produce a cDNA (copy DNA) which then serves as a new spacer sequence (Silas et al., 2016). Nevertheless, although the function of the Cas1-Cas2 complex is thought to be well understood, the need to further characterize the adaptation mechanisms in other CRISPR-Cas systems persists.

Expression. The expression stage involves transcription of the CRISPR array to produce a pre-crRNA (precursor CRISPR RNA) which matures to a crRNA and forms an effector complex with respective Cas proteins (Fig 1.4. A) (Marraffini & Sontheimer, 2010).

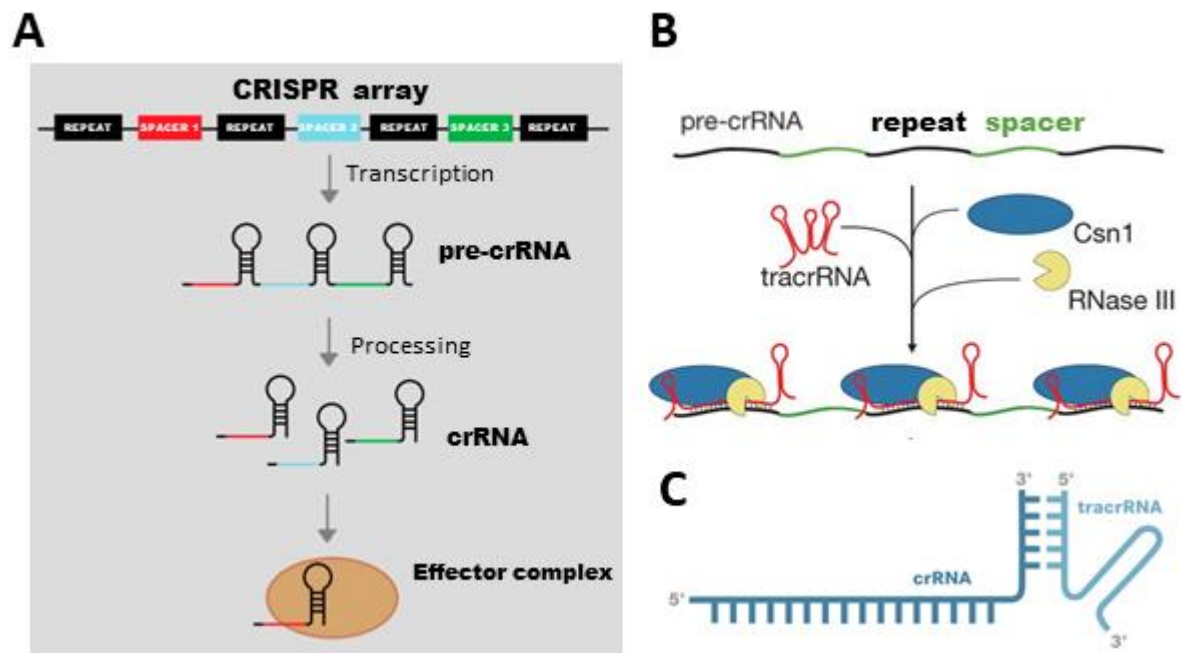


Figure 1.4. Overview of CRISPR-Cas expression stage. **A)** Principle of crRNA biogenesis. The CRISPR array is transcribed into a precursor CRISPR RNA (pre-crRNA) which gets processed into a mature crRNA. Then crRNA binds to single or multiple Cas proteins to form an effector complex. Adapted from (Loureiro & da Silva, 2019). **B)** In the case of type II systems, the tracrRNA base pairs with the repeat region of the pre-crRNA and allows binding of RNase III (associated with Csn1) for further crRNA processing. Adapted from (Deltcheva et al., 2011). **C)** crRNA-tracrRNA duplex, commonly guiding type II CRISPR-Cas nucleases. Adapted from “LubioScience”, 2023.

Depending on the type of CRISPR-Cas system, this long precursor transcript undergoes processing by a Cas6 endoribonuclease, RNase III, or interference-involved proteins (e.g., Cas12, Cas13), containing intrinsic RNase activity (Table 1.2) (Fonfara et al., 2016; Gesner et al., 2011; Niewoehner et al., 2014). Also, the crRNA processing in type II and type III CRISPR-Cas systems involves host endoribonucleases, which trim intermediate crRNAs (Marraffini & Sontheimer, 2010). The processing in type I and III CRISPR-Cas systems mostly depends on Cas6 (Niewoehner et al., 2014). Cas6 hydrolyzes a single phosphodiester bond in the pre-crRNA, generating a crRNA containing the spacer sequence and a palindromic repeat-derived sequence that forms a hairpin structure in some systems. The lack of Cas6 or other characteristic processing proteins in type IV systems suggests that further experiments need to be done to understand the maturation of crRNAs in this type (Makarova et al., 2020). Most type II systems process pre-crRNAs quite differently compared to other types. The CRISPR locus of type II systems contains a sequence of tracrRNA (trans-activating crRNA) (Chylinski et al., 2013). Upon transcription of the tracrRNA sequence, the repeat region of the crRNA transcript base pairs with a complementary sequence of the tracrRNA (Fig. 1.4. B). This duplex region is recognized by RNase III (associated with a Cas protein Csn1) and eventually gets cleaved (Deltcheva et al., 2011). In the case of Cas9, the crRNA processing requires Cas9 to bind and position the RNAs for cleavage by RNase III. Following processing by RNase III, the crRNA-tracrRNA duplex (Fig. 1.4. C) remains bound to Cas9 (Deltcheva et al., 2011). On the contrary to type II systems, type V and VI systems do not require tracrRNA molecules and rely on the endoribonuclease activity of Cas12 and Cas13 proteins for pre-crRNA processing (Fonfara et al., 2016; Shmakov et al., 2015). However, the pre-crRNA of subtype VI-A does not require processing

as the pre-crRNA molecules can be used as guides in effector complexes related to this subtype (East-Seletsky et al., 2017). Although the mechanisms of pre-crRNA processing are different among CRISPR-Cas types, all produced crRNA molecules (or crRNA-tracrRNA duplexes of the type II systems) serve their purpose as gRNAs (guide RNAs) which guide Cas nucleases to their targets.

Table 1.2. The pre-crRNA processing among all six CRISPR-Cas types. The crRNA biogenesis of type IV CRISPR-Cas systems remains uncertain.

CRISPR-Cas type	pre-crRNA processing enzymes	Requires tracrRNA
I	Cas6	-
II	RNase III	+
III	Cas6	-
IV	?	-
V	Cas12	-
VI	Cas13	-

Interference. Based on the class of CRISPR-Cas systems, the effector complexes contain single or multiple protein units (Fig.1.5). During the interference stage, the effector complexes of CRISPR-Cas are guided by their gRNAs towards foreign nucleic acids. Base pairing between gRNA and the target sequence is followed by target cleavage. To avoid self-targeting, CRISPR-Cas systems must recognize PAM (protospacer adjacent motif) sequences right next to the target sequences (Shah et al., 2013) before base pairing between the gRNA and the target sequence. In the case of type VI systems, the target RNA must contain a PFS (protospacer flanking sequence) (Gleditsch et al., 2019).

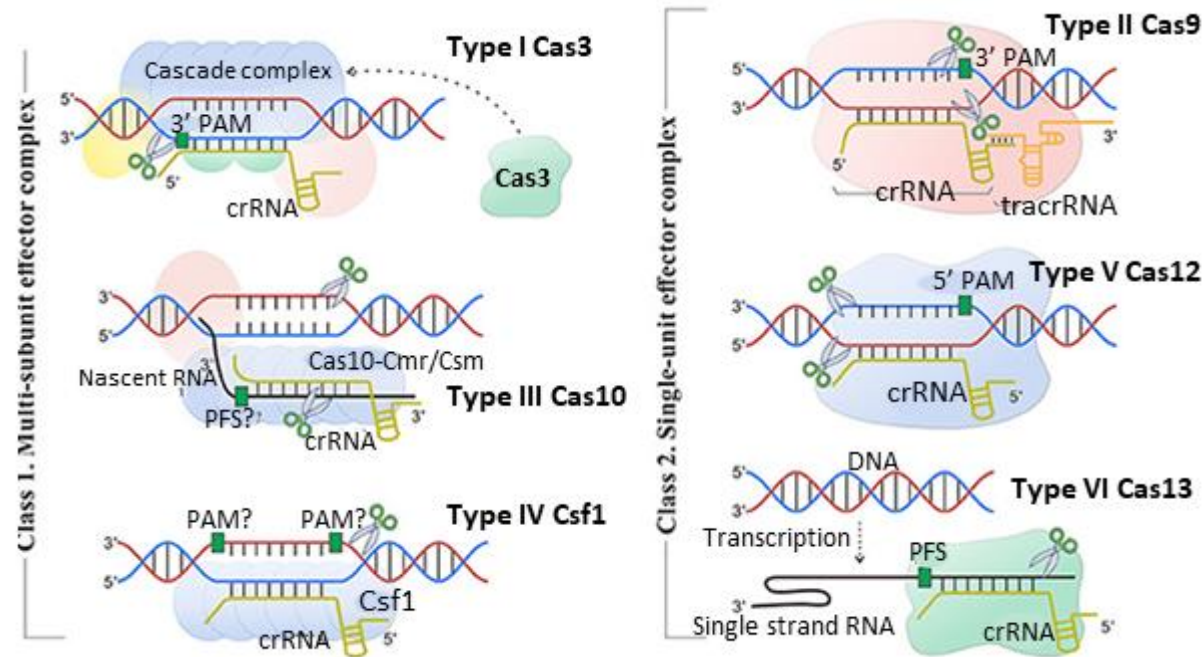


Figure. 1.5. The interference stage of different CRISPR-Cas classes and types. The effector complex of Class 1 CRISPR-Cas systems is comprised of multiple protein units whereas Class 2 effector complexes contain a single protein. Each type contains a signature Cas protein. The interference mechanism of type IV systems remains not fully known. PAM - protospacer adjacent motif, PFS - protospacer flanking sequence. Adapted from (Zhang et al., 2020).

The interference complex of type I CRISPR-Cas systems is composed of multiple proteins, containing Cas proteins bound to a crRNA (Brouns et al., 2008). This ribonucleoprotein (RNP)

complex is also known as a Cascade (CRISPR-associated complex for antiviral defense) system. The Cascade complex recognizes a PAM sequence and then binds to the target. (Sinkunas et al., 2013; Xiao, Luo, et al., 2017). Next, Cas3, a signature protein of type I systems, is recruited to unwind the targeted DNA and degrade one strand of it in the 3' to 5' direction (Brouns et al., 2008). Complete degradation of the invading DNA might be done by other Cascade-independent cellular nucleases (Redding et al., 2015).

Similar to type I systems, the Cascade complex is a characteristic of type III CRISPR-Cas systems (Rouillon et al., 2013). However, subtypes III-A and III-B contain a DNA and RNA binding protein, Cas10 (Rouillon et al., 2013). The Cascade complex of type III CRISPR-Cas systems binds a transcript of invading DNA enabling cleavage of the DNA by Cas10 and Cas7-mediated cleavage of its transcript (Kazlauskienė et al., 2016). Also, studies suggest that Cas10 can produce cyclic oligoadenylates from ATP that activate a non-specific RNase Csm6 (Kazlauskienė et al., 2017). Although Csm6 is not a part of the Cascade complex, it serves an ancillary function by degrading the transcripts in an unspecific manner (Kazlauskienė et al., 2017; Niewoehner et al., 2017).

The multi-subunit interference complex of type IV CRISPR-Cas systems contains its signature protein Csf1 along with Cas5 and Cas7 (Makarova et al., 2020). However, the interference mechanism of type IV nucleases remains not fully known.

The presence of Cas13a protein is linked with type VI CRISPR-Cas systems (Abudayyeh et al., 2016; East-Seletsky et al., 2016). Compared to other class 2 systems, Cas13 contains a unique ability to cleave ssRNA (single-stranded RNA) homologous to the crRNA. Upon binding to the target, the RNase activity of Cas13 is triggered by two HEPN (higher eukaryotes and prokaryotes nucleotide)-binding domains (Abudayyeh et al., 2017; Cox et al., 2017). Also, similar to the Csm6 enzyme from type III systems, Cas13 can cleave ssRNAs in a non-specific manner (Abudayyeh et al., 2016; East-Seletsky et al., 2016).

The interference step of type II CRISPR-Cas systems involves Cas9 bound to a tracrRNA-crRNA hybrid formed during the expression stage (Jinek et al., 2012). Upon binding this duplex to Cas9, the nuclease screens for a PAM site located in the non-targeted strand. Then crRNA of the RNA duplex base pairs with a complementary targeted DNA strand leading Cas9 RuvC and HNH domains to introduce a blunt double-strand break (Gasiunas et al., 2012; Jinek et al., 2012).

Finally, Cas12 is a signature interference-related protein of type V CRISPR-Cas systems (Shmakov et al., 2015). Although Cas12 shares some structural similarities with Cas9, the interference mechanism of Cas12 depends on a single RuvC domain a single crRNA molecule without the requirement of the tracrRNA. Upon PAM recognition, the crRNA of Cas12 binds to a complementary DNA target followed by its cleavage and production of staggered ends (Fonfara et al., 2016).

Since this thesis is focused on Cas9 and Cas12a CRISPR-Cas nucleases, I will further review their properties by providing more detail on structure and mechanisms of action.

1.2.1 Cas9

The signature protein of the type II CRISPR-Cas systems, Cas9, is a multidomain DNA-targeting endonuclease. Targeting of naturally occurring CRISPR-Cas9 nucleases is mediated by a duplex of crRNA and tracrRNA (Jinek et al., 2012). tracrRNA, encoded upstream of the type II CRISPR-Cas locus, binds to a pre-crRNA transcript and activates maturation of the crRNA (Fig. 1.6. A) (Deltcheva et al., 2011). Upon RNase III and tracrRNA-dependent maturation of the crRNA, a crRNA-tracrRNA hybrid remains bound to the Cas9 (Deltcheva et al., 2011). Then Cas9 scans for a

PAM sequence right next to the target and subsequently base pairs with a DNA sequence, complementary to the crRNA (Fig. 1.6. B). Base pairing is followed by an R-loop formation which triggers conformational changes in Cas9 leading to target cleavage (Fig. 1.6. B). The coordinated DNA cleavage is done by two distinct Cas9 domains, RuvC and HNH, introducing a blunt double-stranded break (DSB).

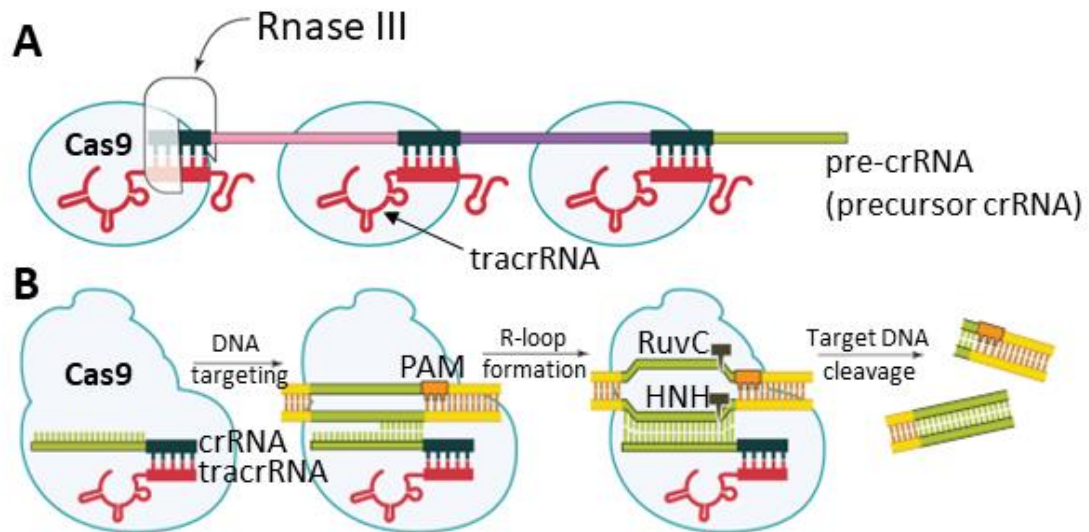


Figure 1.6. Principles of crRNA maturation and RNA-guided target cleavage of type II CRISPR-Cas systems. **A)** Maturation of Cas9 pre-crRNA. RNase III-dependent processing is activated upon tracrRNA binding to the pre-crRNA. **B)** crRNA-tracrRNA guided cleavage of target DNA. Upon PAM identification, crRNA binds to the target sequence. Followed by the R-loop formation, endonuclease domains RuvC and NHN cleave targeted DNA. Adapted from (Doudna & Charpentier, 2014).

To mimic the natural crRNA-tracrRNA hybrid (Fig 1.7 A), the crRNA and tracrRNA can be fused to form a chimeric single-guide RNA (sgRNA) (Fig. 1.7 B) (Jinek et al., 2012). sgRNA retains critical features that allow its binding to Cas9 and targeted DNA. Changing the 20 nt targeting sequence at the 5' end of the sgRNA allows to program Cas9 to target any DNA sequence (Jinek et al., 2012). Nevertheless, both natural crRNA-tracrRNA duplexes and sgRNA chimeras can be utilized for directed genome editing by Cas9.

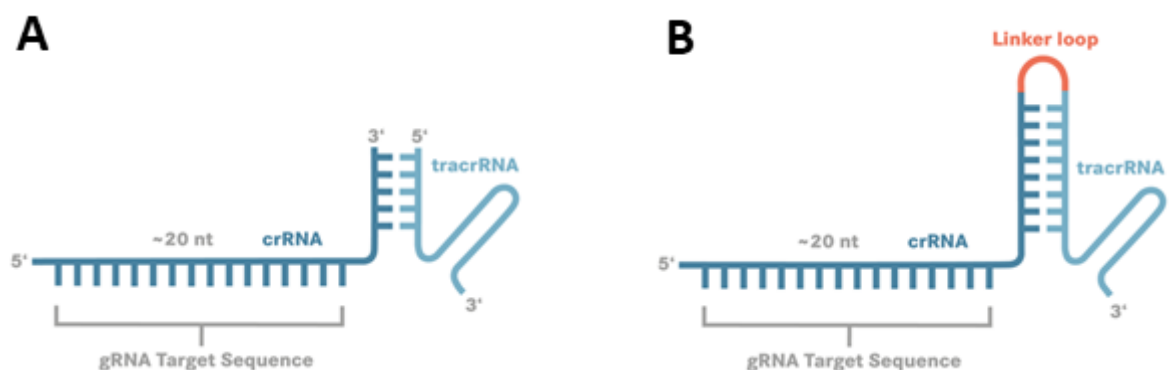


Figure 1.7. Alternative gRNA constructs, accommodated by Cas9. **A)** Naturally occurring crRNA-tracrRNA duplex, containing target sequence on 5' end. **B)** Single-guide RNA (sgRNA). Contains a linker loop to fuse crRNA and tracrRNA. Also contains a target sequence on the 5' end. Adapted from “LubioScience”, 2023.

Although several Cas9 orthologs from distinct subtypes were identified, one of the most characterized Cas9 nucleases comes from *Streptococcus pyogenes* (SpCas9) (Fonfara et al., 2014; Jinek et al., 2012). SpCas9 is comprised of NUC (nuclease) and REC (recognition) domains connected through an arginine-rich bridge helix (BH) (Fig. 1.8) (Jiang et al., 2016; Jinek et al., 2014; Nishimasu et al., 2014). REC I – REC III domains, found in the REC lobe, allow binding of SpCas9 to sgRNA and targeted DNA. The binding of sgRNA to SpCas9 initiates conformational changes to accommodate targeted DNA (Jiang et al., 2015). PAM-interacting domain (PI), found in the NUC lobe, scans for a G-rich (typically 5'-NGG) PAM motif located downstream of the targeted DNA. Following PAM identification, sgRNA base pairing with the targeted strand (TS) causes displacement of the non-targeted strand (NTS) and leads to an R-loop formation (Jiang et al., 2016). R-loop formation triggers conformational changes in the REC subdomains that allosterically activate the HNH domain (Palermo et al., 2018). The activated HNH domain cleaves the NTS of the targeted DNA. Consequently, activation of the HNH domain induces conformational changes in the HNH-RuvC junctions thus activating the RuvC domain to cleave TS afterward (Nierzwicki et al., 2021). As a result, activated HNH and RuvC domains generate a staggered DSB in targeted DNA (Fig. 1.6 B) (Sternberg et al., 2015). Upon cleavage, SpCas9 remains bound to the target DNA and is later displaced for recycling by other cellular factors (Sternberg et al., 2014).

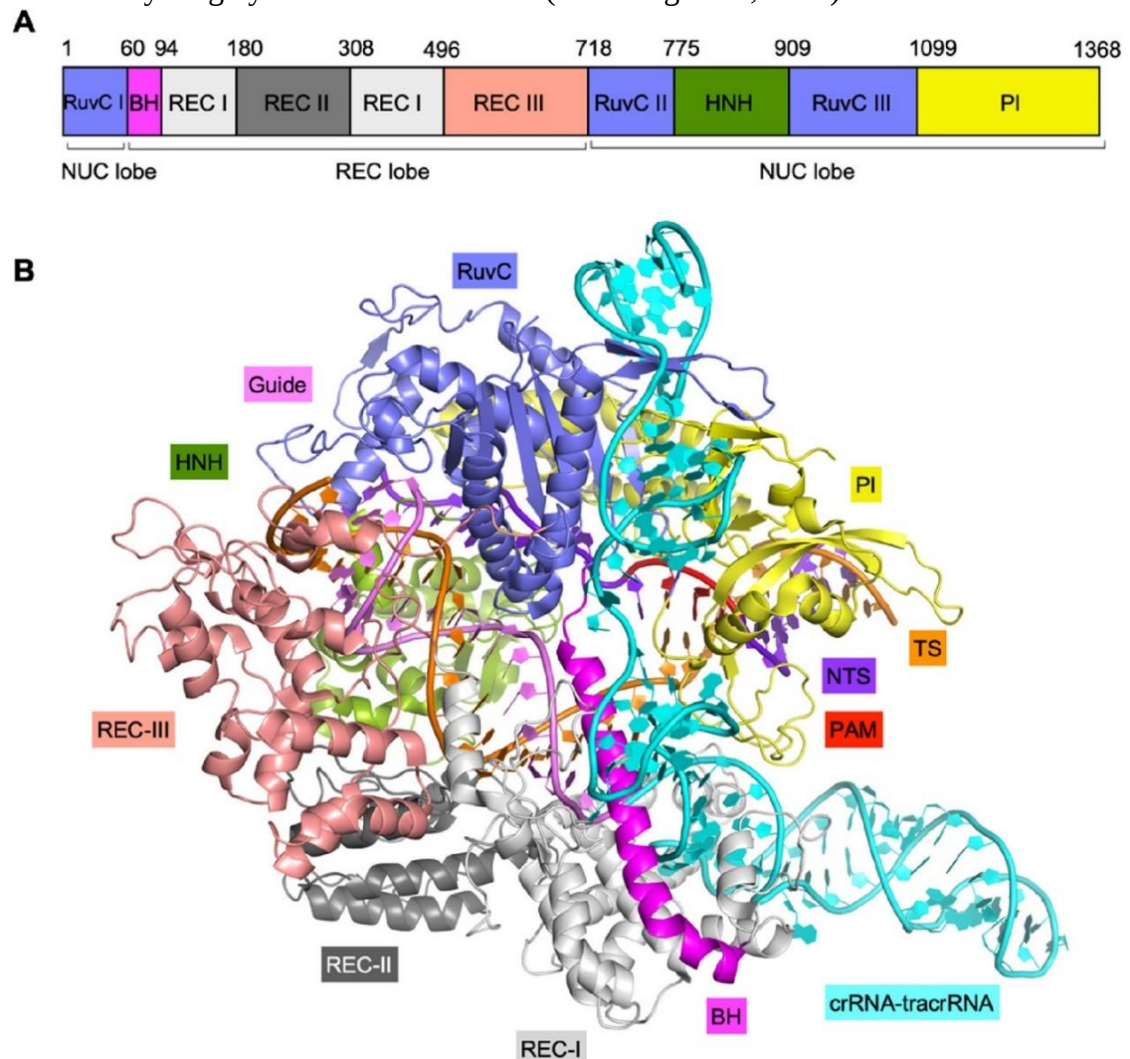


Figure 1.8. Structural organization of SpCas9. **A)** SpCas9 domain organization. **B)** Three-dimensional structure of SpCas9 bound with a sgRNA and target DNA. Abbreviations: BH – bridge helix, TS – targeted strand, NTS – non-targeted strand (contains PAM sequence), PI – PAM interacting domain. Adapted from (Babu et al., 2021; Jiang et al., 2016). PDB ID [5F9R](#).

Extensive structural studies of Cas9 have provided valuable information about Cas9-mediated molecular mechanisms, including PAM recognition as well as target DNA binding and cleavage. However, the repertoire of Cas9-mediated genome engineering is expected to expand once more Cas9 orthologs are characterized.

1.2.2 Cas12a

Cas12a (formerly known as Cpf1) is a member of type V CRISPR-Cas nucleases (Bernd Zetsche, 2015). A few years after the popularity rise of Cas9, Cas12a was discovered as a new promising nuclease for gene editing (Bernd Zetsche, 2015). Cas12a is guided by a single crRNA molecule (Fig. 1.9 A) and utilizes a single RuvC domain to cleave targeted DNA. Upon crRNA hybridization to the target, the endonuclease activity of RuvC produces staggered double-stranded breaks (Fig. 1.9 B) (Yamano et al., 2016).

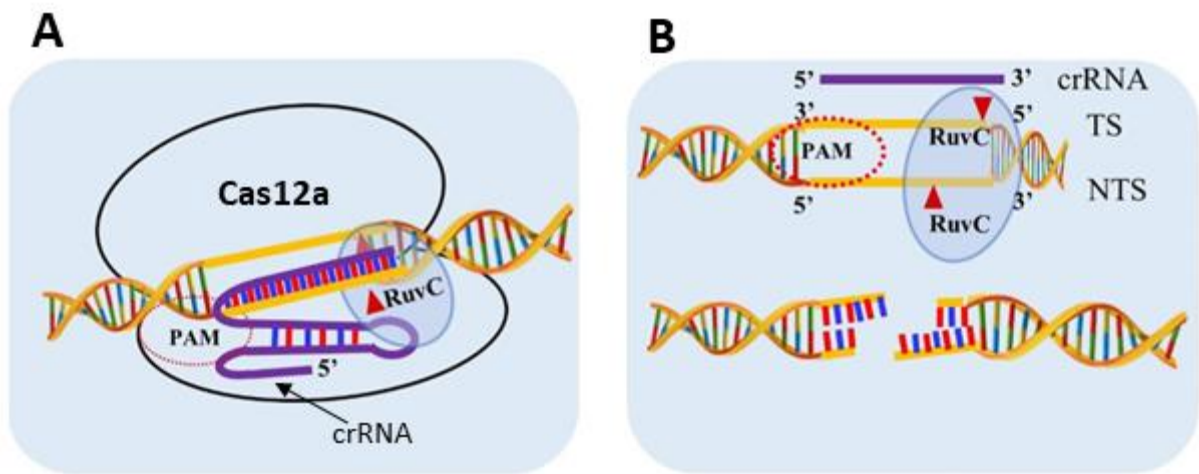


Figure 1.9. Schematics of Cas12a-mediated DNA cleavage. **A)** Cas12a target binding. Cas12a contains a single RuvC domain and is guided by a single crRNA molecule. **B)** Endonuclease activity of Cas12a. NTS (non-targeted strand) cleavage is followed by cleavage of TS (targeted strand). Target cleavage is done by a single RuvC domain, producing a double-stranded staggered break. Adapted from (Bandyopadhyay et al., 2020).

Cryo-EM and crystallographic data (Stella et al., 2017, 2018; Yamano et al., 2016) revealed that a bilobed structure of Cas12a is created by the REC (recognition) and NUC (nuclease) lobes (Fig. 1.10 A). The REC lobe consists of REC1 and REC2 domains, and the NUC lobe is made up of bridge helix (BH) together with PAM-interacting (PI), wedge (WED), and RuvC domains (Fig. 1.10 B). On the contrary to Cas9, the guiding of Cas12a solely depends on a single crRNA molecule, processed by the enzyme itself (Fonfara et al., 2016). The RNase site for crRNA processing is located in the WED-III subdomain and the DNase site is found in the interface of RuvC and NUC domains. Intramolecular base pairing of the 5' repeat region in the crRNA forms a pseudoknot structure which is stabilized by interactions with the RuvC, REC2, and WED domains (Stella et al., 2017; Swarts et al., 2017). This binary interference complex recognizes T-rich PAM sequences (typically 5'-TTTV) and initiates hybridization of the crRNA and target DNA through the WED II-III, REC1, and PAM-interacting domains (Bernd Zetsche, 2015; Dong et al., 2016; Stella et al., 2017). Upon crRNA hybridization to the targeted strand (TS), containing the PAM sequence, the non-target strand (NTS) is positioned towards the DNase site by the PI domain and gets hydrolyzed first. Produced single-stranded DNA flap is displaced by the WED domain and then TS enters the catalytic site generating a staggered cut to the targeted DNA (Fig. 1.9 B).

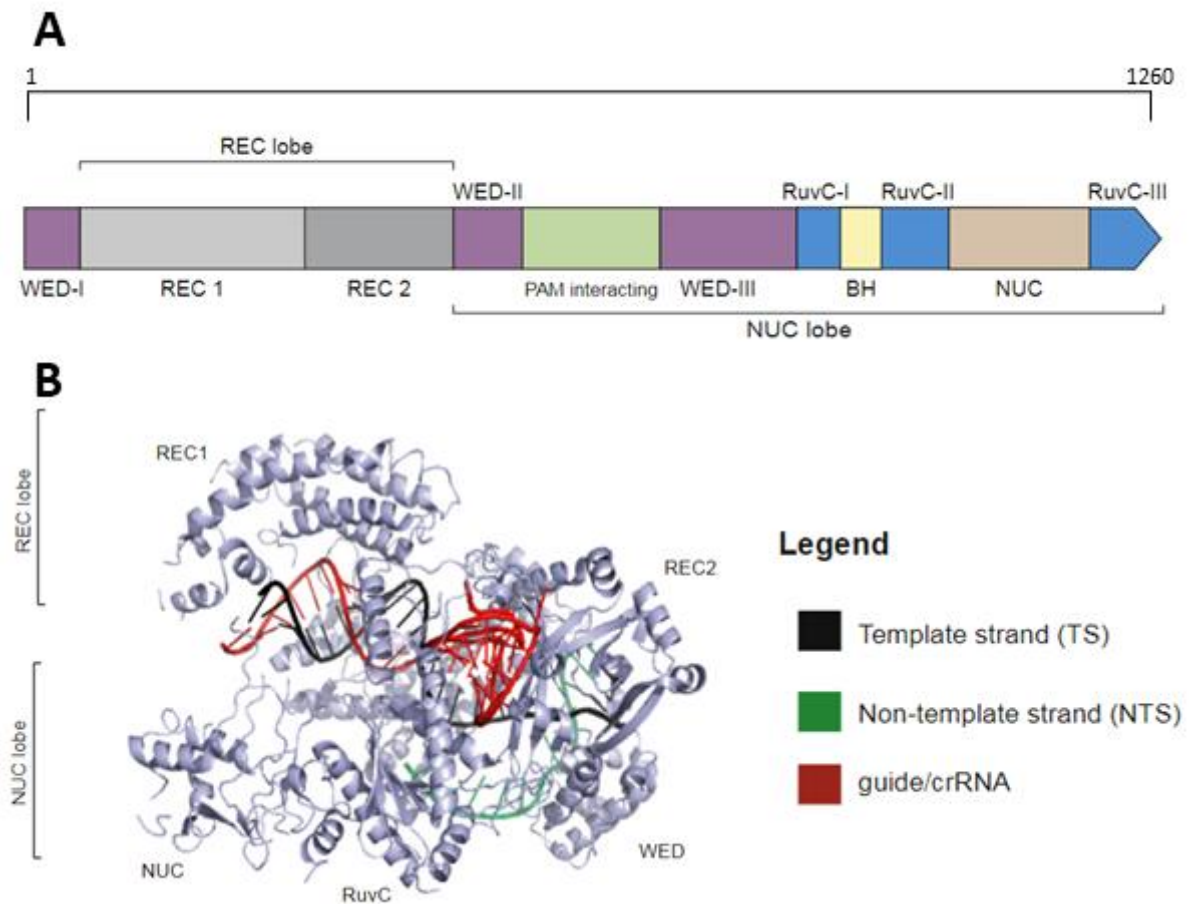


Figure 1.10. Structural organization of Cas12a from *Francisella novicada* U112 (FnCas12a). **A)** FnCas12a domain organization. **B)** Three-dimensional structure of FnCas12a bound with the crRNA and target DNA. Abbreviations: BH – bridge helix, PI – PAM interacting domain, NUC – nuclease (lobe), REC – recognition (lobe). Adapted from (Paul & Montoya, 2020; Stella et al., 2017). PDB ID [5MGA](#).

The discussed features of Cas12a (e.g., T-rich PAM requirement, staggered DSBs in targeted DNA) broaden the versatility of the CRISPR-Cas toolkit that can be applied for precise genome editing.

1.3 Overview of CRISPR-Cas genome editing

CRISPR-Cas-based technology has already been applied for mammalian cell genome editing (Cong et al., 2013; Mali et al., 2013) as well as introducing site-specific transcriptional and epigenetic modifications (Nelson & Gersbach, 2016; Nishida et al., 2016). Also, CRISPR-Cas genome editing brought advances in diagnostics as well as agriculture (Endo et al., 2016; Li et al., 2018). To better understand how these advances were achieved, this section will review the general principles of CRISPR-Cas mediated genome editing along with limitations and therapeutic applications.

As mentioned in section 1.2.1., Cas9 endonucleases can accommodate a synthetic sgRNA (single-guide RNA) that can target any DNA of interest (as long as it is flanked by a respective PAM sequence) (Jinek et al., 2012). Upon PAM recognition, the sgRNA base pairs with targeted DNA and then Cas9 introduces a double-stranded break (DSB) (Jinek et al., 2012). The actual genome editing happens during the cellular repair of the DSB, which occurs in two main pathways – homology-directed repair (HDR) and non-homologous end joining (NHEJ) (Fig. 1.11) (Ran et al., 2013). HDR

utilizes a donor template that is integrated into the DSB site and allows precise corrections of the targeted DNA. However, the HDR pathway is relatively inefficient since it only occurs in dividing cells and requires delivery of the donor template along with Cas9 (Ran et al., 2013). In contrast, during NHEJ the cut DNA is ligated without the requirement of a donor template. Although the occurrence of NHEJ is higher than HDR, this pathway is error-prone and often generates insertions or deletions (indels) at the DSB site (Iliakis et al., 2004).

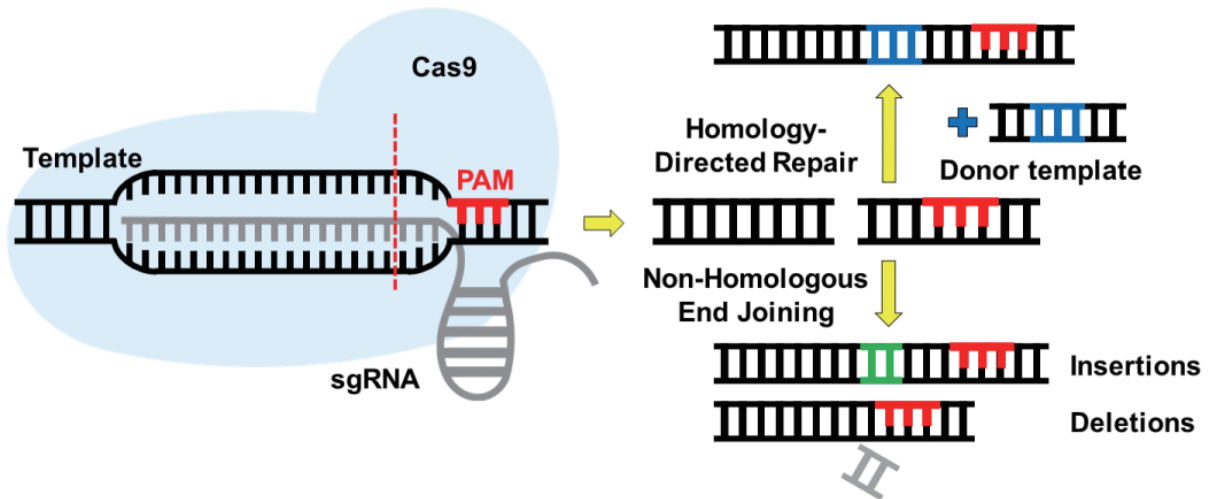


Figure 1.11. Schematics of CRISPR-Cas9 mediated genome editing. Cas9 recognizes a PAM sequence adjacent to the template DNA and introduces a double-stranded break (DSB) after the sgRNA base pairs with the target sequence. Repair of DSB takes place in either homology-directed repair (HDR) or non-homologous end joining (NHEJ) pathway. HDR requires a donor template and allows for precise gene editing, whereas NHEJ occurs through ligation of the DSB which results in insertions or deletions at the site. The red dashed line indicates the cleavage of both strands of the template. Adapted from (Lim & Kim, 2022).

Recent approaches to expand the toolkit of CRISPR-Cas-mediated genome editing resulted in the development of new CRISPR-Cas-based genome editors. Base editors are one of the newly developed genome editing tools composed of Cas9 nickase fused with a cytidine or adenosine deaminase (Fig. 1.12 A) (Gaudelli et al., 2017; Komor et al., 2016). The Cas9 nickases are engineered Cas9 mutants that can contain deactivated either RuvC or HNH domains to cut only a single strand of targeted DNA. Notably, Cas9 nickases are more precise than regular Cas9 and help to avoid NHEJ since they cleave only one strand of the targeted DNA (Shen et al., 2014). The sgRNA of Cas9 nickase works the same way as usual by guiding it to the target sequence. Upon sgRNA hybridization to the targeted DNA, the deaminase fused to the Cas9 nickase enables single nucleotide conversions. To specify, adenine base editors utilize an adenosine deaminase for A (adenosine) conversion to I (inosine) and cytosine base editors utilize a cytidine deaminase to convert C (cytosine) to U (uridine) (Gaudelli et al., 2017; Komor et al., 2016). Following the DNA replication or repair, the A-T base pair is converted to G-C, and the C-G is converted to T-A. This kind of ability to introduce single nucleotide changes in DNA targets holds a large potential to apply base editors for the elimination of disease-causing point mutations (Rees & Liu, 2018).

Another gene editor, comprised of Cas9 nickase (or catalytically impaired Cas9) fused with a reverse transcriptase, is called a prime editor (Fig. 1.12 B) (Anzalone et al., 2019). The reverse transcriptase of the prime editor is paired with a prime editing guide RNA (pegRNA) which guides the Cas9 nickase to its target and simultaneously serves as an RNA template for the reverse

transcriptase (RT). The RNA template of RT can be designed to introduce all types of edits, including substitutions, insertions, and deletions. The more detailed mechanism of prime editors is described by (Anzalone et al., 2019). Whilst further developments remain to be done, prime editors are considered to hold the greatest potential for precise genome editing among other available tools. (Scholefield & Harrison, 2021).

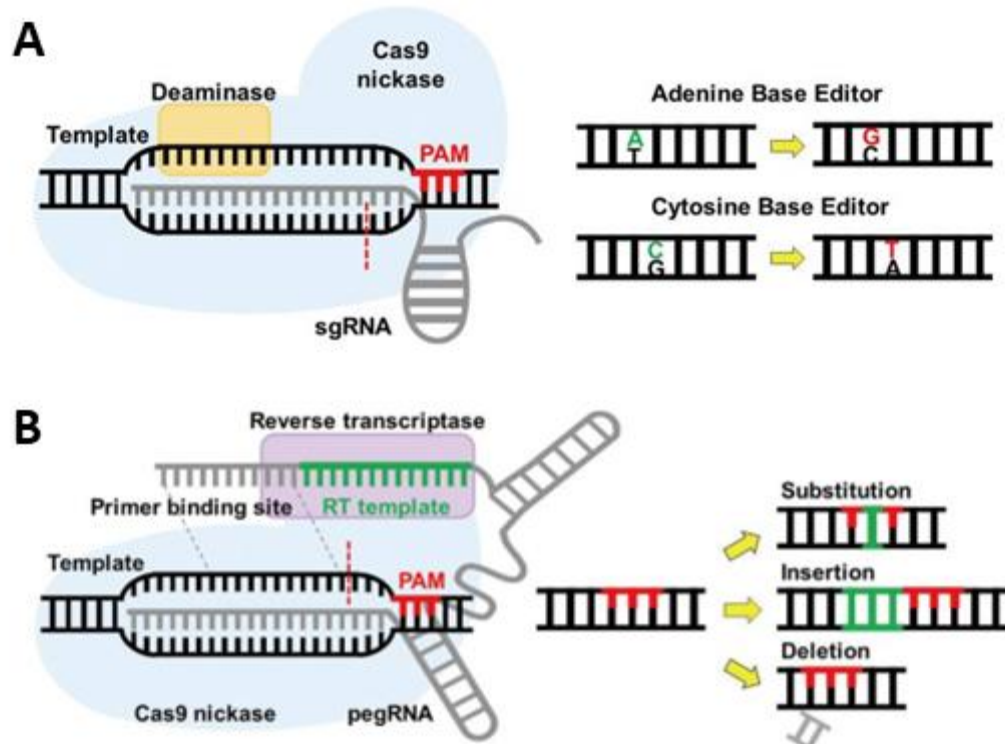


Figure 1.12. Schematics of CRISPR-based genome editors. **A)** Principle of base editors. A Cas9 nickase fused with a respective deaminase enables single nucleotide changes in targeted DNA. **B)** Principle of prime editors. A Cas9 nickase is guided by prime editing guide RNA (pegRNA) containing an adjustable template sequence for reverse transcriptase (RT). The result of prime editing depends on the sequence of the RT template. Adapted from (Lim & Kim, 2022).

Although this section described CRISPR-Cas9-based gene editing tools, it is worth mentioning that similar tools are developed by including other CRISPR-Cas nucleases such as Cas12 and Cas13. While prime editors remain mainly focused on Cas9, Cas12a, and Cas13 were already utilized for base editing (Cox et al., 2017; Gaillochet et al., 2023). Nevertheless, all described CRISPR-Cas-based tools hold the potential for precise genome editing.

1.3.1 Limitations of CRISPR-Cas genome editing

Despite recent advances in CRISPR-Cas technologies, several limitations still exist. One of the major concerns is the off-target activity of CRISPR-Cas nucleases which limits their adaptation for *in vivo* therapeutic applications of complex eukaryotic organisms (H.-C. Yang & Chen, 2018; Zischewski et al., 2017). To reduce off-target effects, CRISPR-Cas nucleases are engineered with increased on-target specificity. Such nucleases include HypaCas9, HF-Cas9, and Sniper Cas9 (Chen et al., 2017; Hu et al., 2018; J. K. Lee et al., 2018). Also, lower off-target effects were demonstrated by the utilization of Cas nickases, containing catalytically inactivated endonuclease domains (Anzalone et al., 2019). Alternatively, off-target effects can be diminished by lowering the duration of the nuclease activity or by directly modulating the activity of Cas proteins upon target locus alterations (Davis et al., 2015; Ghosh et al., 2019).

Another issue regarding genome editing is related to the absence of PAM sequences in the gene loci of interest. This restriction is being decreased through the characterization of phylogenetically diverse CRISPR-Cas nucleases which require different PAM sequences (Gao et al., 2017; Kleinstiver et al., 2015). Also, gene editing through HDR (homology-directed repair) is insufficient due to the rare occurrence of this double-stranded break (DSB) repair pathway in cells (Miyaoaka et al., 2016). HDR can be enhanced by the addition of chemicals, such as KU0060648, SCR7, and NU7441, or by delivery of the donor template together with the nuclease (K. Lee et al., 2017; Maruyama et al., 2016; Robert et al., 2015). However, HDR-independent genome editing can be done by utilization of prime or base editors, which do not require donor templates and were demonstrated to precisely fix mutations in targeted sequences (Anzalone et al., 2019; Rees & Liu, 2018).

Possible immunogenic toxicity raises concerns for CRISPR-Cas technology application for *in vivo* gene therapy. Studies have shown that Cas9 and its gRNAs can trigger immunogenic responses in human cells (Charlesworth et al., 2019; S. Kim et al., 2018). However, to reduce the risk of immunogenic response, CRISPR can be applied to modify cells *ex vivo*. Nevertheless, possible immunogenic toxicity should be taken into account to ensure the safety of CRISPR-Cas-mediated gene therapy.

1.3.2 Clinical trials

CRISPR-based genome editing is currently ongoing clinical trials to treat diseases such as refractory cancer, diabetes, muscular dystrophy, and many more (Lim & Kim, 2022; Y. Yang et al., 2021). These trials include both *in vivo* and *ex vivo* delivery of gene editing agents. For *ex vivo* genome editing, cells are extracted from the patient or donor, modified in the laboratory, and then reinfused into the patient. For *in vivo* gene therapy, the genome editor is directly injected into the patient.

One of the CRISPR-based *ex vivo* clinical trials demonstrated promising results to treat sickle cell disease and beta-thalassemia (Frangoul et al., 2021). Both diseases are related to mutations in the hemoglobin beta-subunit (HBB) gene which cause the absence, reduction, or structural changes of the hemoglobin-beta subunit. This study used CRISPR-Cas9-based technology to restore the production of fetal hemoglobin which would compensate for defective HBB gene. Although these clinical trials are still ongoing (clinicaltrials.gov, [NCT03655678](#), [NCT03745287](#)), high levels of fetal hemoglobin were maintained and the patients avoided disease-related blood transfusions for longer periods. Other CRISPR-based *ex vivo* genome editing trials are done in an attempt to treat refractory cancers (Jing et al., 2018; Lu et al., 2020; Stadtmauer et al., 2020). CRISPR-Cas9 was applied to remove immune checkpoint modulator genes that would result in the enhancement of natural anti-tumor responses. However, these clinical trials were only partially successful. An attempt to treat refractory cancers such as liposarcoma and multiple myeloma was terminated due to low CRISPR-Cas editing frequencies and further progression of tumors (clinicaltrials.gov, [NCT03399448](#)). Nevertheless, another similar clinical trial demonstrated the viability and expansion of edited anti-tumor immune cells without significant off-target effects (clinicaltrials.gov, [NCT02793856](#)). Unfortunately, further disease progression was identified in all 12 examined patients and 11 of them have passed away.

One of the first CRISPR-based *in vivo* clinical trials was done to treat transthyretin amyloidosis (also called ATTR amyloidosis) (Gillmore et al., 2021). ATTR amyloidosis is caused by the accumulation of misfolded transthyretin (TTR) protein in tissues and usually causes diastolic dysfunction or heart failure. CRISPR-Cas9-based pharmaceutical product was injected into patients

to target the TTR gene. 4 weeks after the injection, a 52% and 87% decrease of TTR was reported in low-dose and high-dose groups respectively. However, the clinical trial is still ongoing to determine the safety and long-term durability of this treatment. Even though *ex vivo* clinical trials are more prevalent than *in vivo*, CRISPR-based *in vivo* genome editing is being tested in animals. For example, major successes were demonstrated in attempts to treat retinitis pigmentosa in mice (Goossens et al., 2019; Moreno et al., 2018; Wu et al., 2016). Retinitis pigmentosa refers to a group of inherited diseases that cause progressive ocular degeneration and are related to mutations in several genes. Utilization of CRISPR-Cas tools in mice resulted in successful targeting and correction or depletion of these genes, leading to a decrease in the disease symptoms.

These are just a few examples of how CRISPR-Cas technology is tested to treat various genetic diseases. Although the safety and effectiveness of genome editing by CRISPR-Cas remain to be assessed, ongoing clinical trials and animal testing are contributing to its potential applications in clinical practice.

1.4 Cell-free protein synthesis

Cell-free protein synthesis (CFPS), also known as *in vitro* protein synthesis, is a versatile protein synthesis tool, mostly used in bioengineering and synthetic biology (Silverman et al., 2020; Zubay, 1973). CFPS platforms utilize elements of transcriptional, translational, and metabolic machinery of living cells to produce proteins of interest. These elements include translation initiation and elongation factors, metabolic enzymes, ribosomes, chaperones, and other similar components that supplement the efficiency of CFPS. CFPS systems can accommodate exogenously added DNA and result in the expression of proteins of interest. Depending on the expressed proteins, their yields can range from hundreds of micrograms per milliliter to milligrams per milliliter. (Kim et al., 2011; Zawada et al., 2011). Added amino acids, energy substrates, salts, cofactors, and other biochemical elements act as catalysts for protein production through CFPS. Once the energy substrates are depleted or the accumulated byproducts reach an inhibitory concentration, the CFPS eventually gets terminated. However, the open and flexible nature of the cell-free systems allows to prolong the CFPS by elimination of the inhibitory byproducts and by supplementing the reactions with additional energy components (Gregorio et al., 2019).

Most of the developed CFPS strategies are based on cell-free extracts (CFEs), derived from cells such as *Escherichia coli*, rabbit reticulocytes, or wheat germ (Anastasina et al., 2014; Fogeron et al., 2021; Smolskaya et al., 2020). Compared to cell-based protein synthesis, cell-free platforms have several advantages. First, the utilization of CFEs for recombinant protein production requires fewer experimental steps compared to classical *in vivo* methods (Fig. 1.13). Second, the open nature of cell-free extracts allows easy control and manipulation of the reaction environment. These manipulations include the addition of linear or plasmid DNA templates as well as the replacement or removal of some tRNAs and the inclusion of non-canonical amino acids (Asahara et al., 2021; Ranji Charna et al., 2022; Yokogawa et al., 2010). Third, CFPS systems are optimized toward the production of a desired protein product (including toxic proteins) without the need to support cell viability and growth.

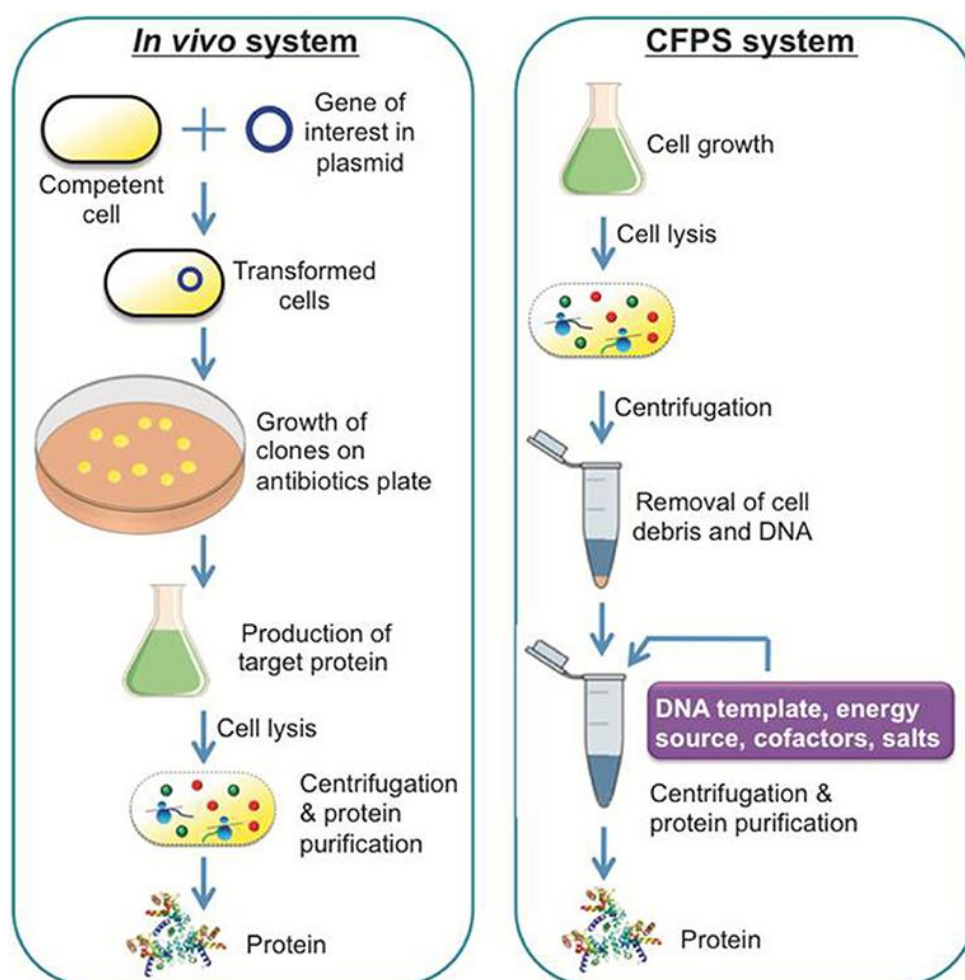


Figure 1.13. Comparison of cell-based and cell-free protein synthesis (CFPS) systems. *In vivo* systems enable the production of recombinant proteins like the CFPS system, although it takes more experimental steps to achieve similar results. Adapted from (Khambhati et al., 2019).

In comparison to CFPS in cell-free extracts, the CFPS platform can be reconstituted from purified elements that compose the so-called PURE (protein synthesis using recombinant elements) system (Kuruma & Ueda, 2015; Shimizu et al., 2001). The PURE system is comprised of 36 purified transcriptional, translational, and metabolic proteins along with ribosomes and other small molecule components such as salts, buffers, and nucleotide triphosphates (NTPs) (Kuruma & Ueda, 2015). Compared to CFE-based systems, the PURE system is highly controlled and lacks undesired proteases or ribonucleases that are usually present in CFEs. The PURE systems were already applied in synthetic biology (Findlay et al., 2016; Kuruma & Ueda, 2015) as well as molecular diagnostics (Pardee, Green, et al., 2016), recombinant DNA replication (van Nies et al., 2018) and therapeutics (Pardee, Slomovic, et al., 2016). However, the costly production of PURE limits its usage, thus cell-free extracts remain to be the more popular alternative for CFPS.

1.4.1 CFPS in *E. Coli* cell-free extracts

One of the most popular prokaryotic CFPS systems comes from *Escherichia coli*. Reasons for *E. Coli* extract utilization include cost-effective large-scale cultivation of *E. coli*, followed by simple and cheap extract preparation. Although many cell-free extract preparation protocols are available, *E. Coli* lysate preparation is generally comprised of four main steps: cell growth, lysis, incubation, and dialysis. To ensure the quality of endogenously produced elements essential for CFPS, cultivated

E. Coli cells are collected in their exponential growth phase. Afterward, cells are lysed by a method of choice, e.g., bead beating or sonication (Kwon & Jewett, 2015; Shin & Noireaux, 2010). Cell lysis is followed by several processing steps, including an incubation step to remove remaining endogenous mRNAs and DNAs, and dialysis to eliminate small molecules.

Following *E. Coli* CFE preparation, the CFPS is initiated by the addition of energy substrates, nucleotides, amino acids, cofactors, and DNA templates. DNA templates can be either plasmid or linear. However, if linear templates are used, they should be protected from degradation by the addition of exonuclease inhibitors, such as GamS (Sitaraman et al., 2004). Also, *E. Coli* cell-free expression is often done by harnessing components of the T7 bacteriophage transcription system (Köhler et al., 1996). This system contains a T7 RNA polymerase (T7 RNAP) that is known to be more processive than endogenous *E. Coli* RNA polymerase (Tabor, 2001) thus providing efficient transcription during CFPS. Depending on the strain used for the CFE production, the T7 RNAP synthesis can be induced before cell lysis (for *E. Coli* BL21 DE3 strains) or added to the CFPS reactions exogenously (T.-W. Kim et al., 2006). Accordingly, added DNA templates for protein synthesis must contain sequences of T7 promoters and T7 terminators to enable transcription by T7 RNAP. This kind of *E. Coli* CFPS system can be used to synthesize various proteins, including enzymes, antibodies, and membrane proteins (Zemella et al., 2015). However, the *E. Coli* CFPS systems are not suitable for the production of proteins with some post-translational modifications, such as nitrosylation, ubiquitylation, methylation, or acetylation (Zemella et al., 2015). In these cases, CFEs derived from eukaryotic cells (e.g., wheat germ, rabbit reticulocytes) should be used (Anastasina et al., 2014; Fogeron et al., 2021).

1.4.2 Characterization of CRISPR-Cas nucleases through CFPS

Although phylogenetical studies revealed a wide diversity of CRISPR-Cas nucleases (Makarova et al., 2020), their adaptation to CRISPR technologies is relatively slow. Such limitation is related to the long and tedious process to characterize the basic properties of the Cas nucleases. Initial characterization experiments usually took days or weeks due to the requirement of cell culturing followed by protein purification (Karvelis et al., 2017; Leenay & Beisel, 2017). To speed up the characterization process, in 2018 Marshall et al. demonstrated the utilization of *E. Coli* CFEs to benchmark the properties of several CRISPR-Cas nucleases. During these experiments, CRISPR-Cas nucleases were expressed in CFEs by the addition of gRNA and nuclease-encoding DNA templates. To measure the activity of these nucleoprotein complexes, a deGFP (slightly modified version of enhanced green fluorescent protein (eGFP)) encoding plasmid was used as a reporter construct (Fig. 1.14 A). Also, a CFPS-based PAM determination method was developed to measure the activity of the Cas nucleases in the presence of target DNAs flanked by a library of different PAM sequences (1.14 B). The produced products are then submitted for new generation sequencing (NGS) which is followed by computational analysis to determine PAM sequences, favoured by the Cas nucleases.

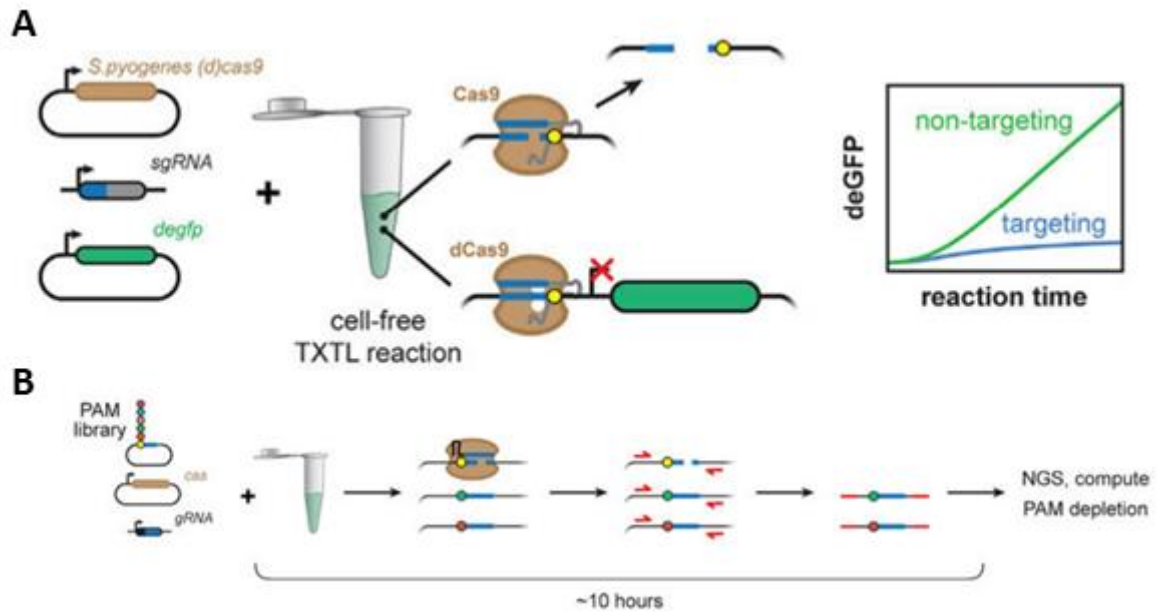


Figure 1.14. Characterization of CRISPR-Cas nucleases through cell-free protein synthesis (CFPS). **A)** Activity measurements of SpCas9 and catalytically dead SpCas9 (dCas9). Reporter plasmid of deGFP, containing a target sequence flanked by a PAM, is added along templates of Cas9, dCas9, and sgRNA. If the targeting occurs, the fluorescence signal decreases. TXTL (stands for transcription-translation) is an alternative name for cell-free protein synthesis. **B)** PAM assay through CFPS. A library of potential PAM sequences flanking the target of a respective Cas nuclease is introduced. Upon CFPS reaction, cleavage products are PCR-amplified and submitted for new-generation sequencing (NGS). Depleted (Cas-favored) PAM sequences are assessed through computational methods. Adapted from (Marshall et al., 2018).

The experiments done in this study also demonstrated that CFPS can be applied to measure the inhibitory activity of anti-CRISPR proteins (Bondy-Denomy et al., 2015) or analyze the gene repression effect of different CRISPR-Cas nucleases. The PAM screen experiments conducted by Marshall et al. allowed to determine the PAM profiles of 5 different Cas12a homologs. Following similar CFPS-based methods, the PAM screen strategy was already utilized to identify PAM profiles across some other Cas12a nucleases. (Jacobsen et al., 2020). This shows that CFPS platforms hold great potential for rapid and scalable characterization of CRISPR-Cas nucleases.

2. MATERIALS AND METHODS

2.1 Materials

2.1.1 Bacterial strains

- *E. Coli* BL21(DE3) Star: $F^-ompT hsdS_B (r_B^-, m_B^-) galdcmrne131 (DE3)$
- *E. Coli* Dh5 α : $F^- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 \phi80dlacZ\Delta M15 \Delta(lacZYA-argF)U169, hsdR17(rK-mK^+), \lambda^-$
- *E. Coli* Mach1: $W \Delta recA1398 endA1 fhuA \Phi80\Delta(lac)M15 \Delta(lac)X74 hsdR(rK-mK^+)$

2.1.2 Plasmids

Table 2.1 provides a brief description of each plasmid used in this work.

Table 2.1. Summary of all plasmids used in this work.

Plasmid name	Description
PRG10	pET type expression vector, encoding sfGFP that contains a C-terminal StrepII-tag.
JPI1.19	A customized pET type vector for Golden Gate cloning, containing two targets of BsaI at MCS followed by a StrepII-tag fused on the C-terminus.
JPI1.04	pET type protein expression vector that includes a gene for Adurb193Cas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.33	pET type protein expression vector that includes a gene for HkCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.08	pET type protein expression vector that includes a gene for PiCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.27	pET type protein expression vector that includes a gene for Adurb336Cas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.10	pET type protein expression vector that includes a gene for SpCas9 (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.11	pET type protein expression vector that includes a gene for AsCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.12	pET type protein expression vector that includes a gene for LbCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.15	pET type protein expression vector that includes a gene for FnCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.16	pET type protein expression vector that includes a gene for Fn3Cas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.17	pET type protein expression vector that includes a gene for PdCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.31	pET type protein expression vector that includes a gene for Adurb336Cas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.32	pET type protein expression vector that includes a gene for PiCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.45	pET type protein expression vector that includes a gene for Adurb193Cas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.
JPI1.46	pET type protein expression vector that includes a gene for HkCas12a (inserted via Golden Gate assembly) and a StrepII-tag fused on C-terminus.

Continued Table 2.1.

pCR®-Blunt	A linear blunt end cloning vector, included in Zero Blunt™ PCR cloning kit ("Thermo Fisher Scientific")
JPI1.21	pCR®-Blunt vector containing cloned template of sgRNA D for SpCas9.
JPI1.23	pCR®-Blunt vector containing cloned template of sgRNA E for SpCas9.
JPI1.24	pCR®-Blunt vector containing cloned template of crRNA D for FnCas12a & Fn3Cas12a.
JPI1.25	pCR®-Blunt vector containing cloned template of crRNA E for FnCas12a & Fn3Cas12a.
JPI1.34	pCR®-Blunt vector containing cloned template of crRNA D for Adurb193Cas12a & HkCas12a.
JPI1.35	pCR®-Blunt vector containing cloned template of crRNA E for Adurb193Cas12a & HkCas12a.
JPI1.36	pCR®-Blunt vector containing cloned template of crRNA D for Adurb336Cas12a.
JPI1.37	pCR®-Blunt vector containing cloned template of crRNA E for Adurb336Cas12a.
JPI1.38	pCR®-Blunt vector containing cloned template of crRNA D for AsCas12a.
JPI1.39	pCR®-Blunt vector containing cloned template of crRNA E for LbCas12a.
JPI1.40	pCR®-Blunt vector containing cloned template of crRNA D for PdCas12a.
JPI1.41	pCR®-Blunt vector containing cloned template of crRNA E for PdCas12a.
JPI1.42	pCR®-Blunt vector containing cloned template of crRNA D for PiCas12a.
JPI1.43	pCR®-Blunt vector containing cloned template of crRNA E for PiCas12a.
JPI1.48	pCR®-Blunt vector containing cloned template of crRNA E for AsCas12a.
JPI1.49	pCR®-Blunt vector containing cloned template of crRNA D for LbCas12a.

Specific utilization of each plasmid is elaborated in further sections of the respective methods.

2.1.3 Oligonucleotides

Oligonucleotides, listed in Table 2.2, were synthesized at "Metabion International". Abbreviations "F" and "R" stand for forward and reverse positions against the PCR templates.

Table 2.2. Sequences of oligonucleotides used in this work.

Primers to amplify CRISPR nuclease genes for Golden Gate cloning		
F	J-Pr5.57	5' -GAGTACGGTCTCAAATGAGCAGCGTGTTCAGCGAC-3'
R	J-Pr4.28	5' -GAGTACGGTCTCACGCTGCTCTTGTACTCACGGTTCTGC-3'
F	J-Pr5.47	5' -GAGTACGGTCTCAAATGACCAACAAGTATAGCCTGAGC-3'
R	J-Pr4.36	5' -GAGTACGGTCTCACGCTGCTTTCTCTTTCTTGCTCTGTTC-3'
F	J-Pr5.45	5' -GAGTACGGTCTCAAATGGCCCCAAAGAAGAAACG-3'
R	J-Pr5.46	5' -GAGTACGGTCTCACGCTTGAAATAATGAAATTAATCCAGTCCTCATTCTTG-3'
F	J-Pr5.44	5' -GAGTACGGTCTCAAATGAAAGTGATGGAAAACATCAGGAGTTCACC-3'
R	J-Pr5.01	5' -GAGTACGGTCTCACGCTTTTCAGGTACGGCTTTCTCTGC-3'
Primers to amplify gRNA templates for blunt-end cloning		
F	J-Pr4.74	5' -AAAAGCACCGACTCGG-3'
R	J-Pr4.75	5' -TAATACGACTCACTATAGGTGATAAG-3'
R	J-Pr4.76	5' -TAATACGACTCACTATAGGACGC-3'
F	J-Pr4.71	5' -GAAATTAATACGACTCACTATAG-3'
R	J-Pr4.72	5' -TCCACATGGCATTCC-3'
R	J-Pr4.73	5' -TCCAGCGTCTCATCTTT-3'
Single-stranded templates of gRNAs		
J-Pr1.44	5' - AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGC TATTTCTAGCTCTAAAACCATGGCATTCCACTTATCACCTATAGTGAGTCGTATT-3'	
J-Pr1.45	5' - AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGC TATTTCTAGCTCTAAAACGCGTCTCATCTTTATGCGTCCATAGTGAGTCGTATTA-3'	
J-Pr4.69	5' - GAAATTAATACGACTCACTATAGAATTTCTACTGTTGTAGATGTGATAAGTGGAATGCCATG TGGA-3'	
J-Pr4.70	5' - GAAATTAATACGACTCACTATAGAATTTCTACTGTTGTAGATGACGCATAAAGATGAGACGC TGGA-3'	

Continued Table 2.2.

Quick-change mutagenesis (QCM) primers		
F	J-Pr5.22	5' -GAATTTCTACTATTGTAGATGTGATAAGTGG-3'
F	J-Pr5.23	5' -GAATTTCTACTATTGTAGATGACGCATAAAGATG-3'
R	J-Pr5.24	5' -CATCTACAATAGTAGAAATTCTATAGTGAGTCG-3'
F	J-Pr5.25	5' -GAATTTCTACTGTGGTAGATGTGATAAGTGG-3'
F	J-Pr5.26	5' -GAATTTCTACTGTGGTAGATGACGCATAAAG-3'
R	J-Pr5.27	5' -CATCTACCACAGTAGAAATTCTATAGTGAGTCG-3'
F	J-Pr5.28	5' -GAATTTCTACTCTTGTAGATGTGATAAGTGG-3'
F	J-Pr5.29	5' -GAATTTCTACTCTTGTAGATGACGCATAAAG-3'
R	J-Pr5.30	5' -CATCTACAAGAGTAGAAATTCTATAGTGAGTCG-3'
F	J-Pr5.31	5' -GAATTTCTACTAAGTGTAGATGTGATAAGTGG-3'
F	J-Pr5.32	5' -GAATTTCTACTAAGTGTAGATGACGCATAAAG-3'
R	J-Pr5.33	5' -CATCTACACTTAGTAGAAATTCTATAGTGAGTCG-3'
F	J-Pr5.34	5' -GAATTTCTACTTCGGTAGATGTGATAAGTGG-3'
F	J-Pr5.35	5' -GAATTTCTACTTCGGTAGATGACGCATAAAG-3'
R	J-Pr5.36	5' -CATCTACCGAAGTAGAAATTCTATAGTGAGTCG-3'
F	J-Pr5.37	5' -GAATTTCTACTTGTGTAGATGTGATAAGTGG-3'
F	J-Pr5.38	5' -GAATTTCTACTTGTGTAGATGACGCATAAAG-3'
R	J-Pr5.39	5' -CATCTACACAAGTAGAAATTCTATAGTGAGTCG-3'
Sequencing primers		
F	ORG1	5' -GTAAAACGACGGCCAG-3'
F	J-Pr4.29	5' -GGCGGATATTCGTTGGGACG-3'
F	J-Pr4.30	5' -CGGTCTGATGCTGAGCAAGATG-3'
F	J-Pr4.31	5' -GGACCGTGATATGAGCTTTTATAGCG-3'
F	J-Pr4.32	5' -GAAGGACCTGATCCCGAAAATGC-3'
F	J-Pr4.33	5' -CATCAACGCGAAAGTTCTGAACG-3'
F	J-Pr4.34	5' -CATTAGCGCGATGGACTTTATTCG-3'
F	J-Pr4.37	5' -GAAGCAGAAGGGTAAAGAGGTGG-3'
F	J-Pr4.38	5' -GAACAACCTGATCAACAAGCTGAG-3'
F	J-Pr4.39	5' -GAGCTGCTGTACGACGATAACG-3'
F	J-Pr4.40	5' -CGAGTGGGACATCTACAAGTTCAAG-3'
F	J-Pr4.41	5' -GCAGCTTTAACATCTTCAAGGATGAG-3'
F	J-Pr4.42	5' -CCAACGGCGAGCGTATCATTC-3'
F	J-Pr4.53	5' -GATCTCAACCTTGACTGGCAG-3'
F	J-Pr4.54	5' -CGGCGAGAAGGTTCAAGGC-3'
F	J-Pr4.55	5' -GCAATCAAGGAGTACCTGGAAAAG-3'
F	J-Pr4.56	5' -GTTTCTCTACAAGTACAAGGGCCC-3'
F	J-Pr4.57	5' -CAAGACCAATAAGAATGTTAATCAAATCGTGC-3'
F	J-Pr4.58	5' -CTGTTTCGGCTTAAGAGCATCAAC-3'
F	J-Pr4.61	5' -CTGAGCAGCTTCAGCATTAGCAAC-3'
F	J-Pr4.62	5' -GTATAACCAGAAACACAAGGACCG-3'
F	J-Pr4.63	5' -GAAGCTGCTGGGCACCAAATAC-3'
F	J-Pr4.64	5' -GCATCAGCAAGTATCCGAACATTAG-3'
F	J-Pr4.65	5' -CATGGAAAACAAATTCCTGTTTCACC-3'
F	J-Pr4.66	5' -CCAGCAAGATCGATCCGGTTAC-3'
Colony PCR primers		
F	ORG1	5' -GTAAAACGACGGCCAG-3'
R	ORG2	5' -CAGGAAACAGCTATGAC-3'
F	ORG3	5' -GGATCTCGACGCTCTCCCT-3'
R	ORG6	5' -ACCCCTCAAGACCCGTTTAG-3'

More detailed purposes of each oligonucleotide are later described in sections of respective methods.

2.2 Methods

2.2.1 Golden Gate assembly

Selected Cas12a homologs were cloned into a customized cell-free expression vector (Table 2.3) through Golden Gate assembly as described previously by (Engler et al., 2008). Briefly, CRISPR nucleases were PCR-amplified from plasmids containing their genes (Table 2.3; section 2.1.2) with corresponding sets of primers (Table 2.3; section 2.1.3) and Phusion Plus (“Thermo Fisher Scientific”) DNA polymerase. To remove the PCR templates, the reactions were treated with DpnI (“Thermo Fisher Scientific”) for 1h at 37°C and 5 min at 80°C. Then the inserts were purified with GeneJET™ PCR Purification Kit (“Thermo Fisher Scientific”) and their length was analyzed by running 1% (m/w) agarose gel electrophoresis in 1X TAE buffer (40mM Tris-acetate, 1mM EDTA, pH 8.0 at 25°C).

Table 2.3. Summary of Golden Gate assembly for nuclease gene cloning.

Nuclease	Insert amplification			Cloning vector	Sequencing primers
	PCR template	PCR primers	PCR settings		
Adurb193Cas12a	JPl1.04	J-Pr5.57, J-Pr4.28	98°C 30s	JPl1.19	J-Pr4.29 – J-Pr4.34
Adurb336Cas12a	JPl1.27	J-Pr5.47, J-Pr4.36	98°C 10s 60°C 30s	JPl1.19	J-Pr4.37 – J-Pr4.42
HkCas12a	JPl1.33	J-Pr5.45, J-Pr5.46	72°C 2min 30s 4°C ∞	JPl1.19	J-Pr4.53 – J-Pr4.58
PiCas12a	JPl1.08	J-Pr5.44, J-Pr5.01		JPl1.19	J-Pr4.61 – J-Pr4.66

20fmol of the cloning vector was combined with equimolar amounts of purified inserts in a 5µL Golden Gate reaction. The reactions contained 1X T4 DNA ligase buffer (“New England BioLabs”; 50mM Tris-HCl, 10mM MgCl₂, 1mM ATP, 10mM DTT, pH 7.5 at 25°C), T4 DNA ligase (“New England BioLabs”) and BsaI restriction enzyme (“New England BioLabs”) and were carried out in a thermocycler under the following settings: 37°C (5 min) and 16°C (5 min) for 30 cycles, followed by 5 min incubation at 80°C afterward.

Assembled plasmids were transformed into chemically competent *E. Coli DH5α* (section 2.1.1) cells prepared by following the “CaCl₂ method” (Sambrook et al., 1989). Transformants were grown overnight at 37°C on LB-agar media (10g/l tryptone, 5g/l yeast extract, 10g/l NaCl, 10g/l bacteriological agar, pH 7.0 at 25°C), containing 50µg/ml of kanamycin. The transformants were screened and selected through colony PCR with primers ORG3 and ORG6 (section 2.1.3). Products of colony PCR were analyzed in 1% (m/w) agarose gel*, containing 1X TAE buffer (40mM Tris-acetate, 1mM EDTA, pH 8.0 at 25°C). Afterward, colonies of interest were inoculated to 8-12ml of LB broth (10g/l tryptone, 5g/l yeast extract, 10g/l NaCl, pH 7.0 at 25°C), containing 50µg/ml of kanamycin, and grown overnight at 37°C with 200 rpm⁻¹ shaking. The plasmids were then purified with the GeneJET™ Plasmid Miniprep Kit (“Thermo Fisher Scientific”). Sequences of the Golden Gate assembly constructs (J.Pl1.31, J.Pl1.32, JPl1.45, JPl1.46; section 2.1.2) were verified by Sanger or whole plasmid Nanopore sequencing (Branton et al., 2008; Sanger et al., 1977). Respective sequencing primers for Golden Gate assembly products are listed in Table 2.3 and in section 2.1.3.

**Note – all mentioned agarose gels contain 0.5X of SYBR™ Safe DNA Gel Stain (“Thermo Fisher Scientific”).*

2.2.2 Blunt-end cloning

Single-stranded DNA oligonucleotides that serve as gRNA templates for SpCas9 and Fn/Fn3Cas12a were PCR-amplified as double-stranded DNAs with Phusion DNA polymerase (homemade). Primer pairs, templates, and optimized PCR settings used for the amplification are listed in Table 2.4 and in section 2.1.3.

Table 2.4. Summary of blunt-end cloning of SpCas9, FnCas12a & Fn3Cas12a gRNA templates.

Nuclease	gRNA template	Insert amplification			Blunt-end cloning vector	Sequencing primer
		PCR template	PCR primers	PCR settings		
SpCas9	D	J-Pr1.44	J-Pr4.74, J-Pr4.75	98°C 30s	pCR®-Blunt	ORG1
	E	J-Pr1.45	J-Pr4.74, J-Pr4.76	98°C 10s 55°C 30s		
FnCas12a, Fn3Cas12a	D	J-Pr4.69	J-Pr4.71, J-Pr4.72	72°C 5s		
	E	J-Pr4.70	J-Pr4.71, J-Pr4.73	4°C ∞		

Amplified inserts were purified with GeneJET™ PCR Purification Kit (“Thermo Fisher Scientific”) and analyzed in a 1% (m/w) agarose gel* containing 1X TAE buffer (40mM Tris base, 20 mM acetic acid, 1mM EDTA disodium salt dihydrate, pH 8.0 at 25°C). The inserts were then cloned into a pCR®-Blunt vector (section 2.1.2) using Zero Blunt™ PCR Cloning Kit (“Thermo Fisher Scientific”). Cloned constructs were transformed via electroporation (using “Bio-Rad Laboratories” MicroPulser™ electroporator with the pulse voltage set to 1.8kV) into electrocompetent *E. Coli Mach1* (section 2.1.1) cells, prepared as described previously (Fuller et al., 2001) and grown overnight at 37°C on LB-agar-IPTG media (10g/l tryptone, 5g/l yeast extract, 10g/l NaCl, 10g/l bacteriological agar, 1mM IPTG, pH 7.0 at 25°C), containing 50µg/ml of kanamycin. Transformants were screened through colony PCR using primers ORG1 and ORG2 (section 2.1.3) and selected based on the length of the PCR product. Afterwards, colonies of interest were grown at 37°C overnight in 4ml of LB broth (10g/l tryptone, 5g/l yeast extract, 10g/l NaCl, pH 7.0 at 25°C), containing 50µg/ml of kanamycin. Then plasmid purification was done using GeneJET™ Plasmid Miniprep Kit (“Thermo Fisher Scientific”). Cloned plasmids (JPl1.24, JPl1.23, JPl1.24, JPl1.25; section 2.1.2) were sequence-verified by Nanopore or Sanger sequencing (Branton et al., 2008; Sanger et al., 1977).

**Note – all mentioned agarose gels contain 0.5X of SYBR™ Safe DNA Gel Stain (“Thermo Fisher Scientific”).*

2.2.3 Quick-change mutagenesis

To adjust for the differences between gRNA sequence requirements of different Cas12a homologs, plasmids containing the sequences of Fn/Fn3Cas12a gRNA molecules (JPl1.24 & JPl1.25; section 2.1.2) were used as templates for quick-change mutagenesis (QCM) (H. Liu & Naismith, 2008) (Table 2.5; sections 2.1.2, 2.1.3). The PCR for the mutagenesis reactions were carried out with

Phusion™ Plus DNA Polymerase (“Thermo Fisher Scientific”). To remove the PCR templates, QCM samples were later digested with DpnI (“Thermo Fisher Scientific”) for 1h at 37°C and incubated at 80°C for 5 min.

Table 2.5. Summary of quick-change mutagenesis for Cas12a gRNA templates.

Nuclease	gRNA template	QCM template	QCM primers	QCM settings	Sequencing primer
AdurbCas12a, HkCas12a	D	JPl1.24	J-Pr5.22, J-Pr5.24	98°C 30s 98°C 10s 60°C 30s 72°C 2min 72°C 3min 4°C ∞ } x35	ORG1
	E	JPl1.25	J-Pr5.23, J-Pr5.24		
Adurb336Cas12a	D	JPl1.24	J-Pr5.25, J-Pr5.27		
	E	JPl1.25	J-Pr5.26, J-Pr5.27		
AsCas12a	D	JPl1.24	J-Pr5.28, J-Pr5.30		
	E	JPl1.25	J-Pr5.29, J-Pr5.30		
LbCas12a	D	JPl1.24	J-Pr5.31, J-Pr5.33		
	E	JPl1.25	J-Pr5.32, J-Pr5.33		
PdCas12a	D	JPl1.24	J-Pr5.34, J-Pr5.36		
	E	JPl1.25	J-Pr5.35, J-Pr5.36		
PiCas12a	D	JPl1.24	J-Pr5.37, J-Pr5.39		
	E	JPl1.25	J-Pr5.38, J-Pr5.39		

After DpnI digestion, QCM products were transformed via electroporation (using “Bio-Rad Laboratories” MicroPulser™ electroporator with the pulse voltage set to 1.8kV) into electrocompetent *E. Coli* Mach1 (section 2.1.1) cells, prepared as described previously (Fuller et al., 2001). Transformants were grown overnight at 37°C on LB-agar-IPTG media (10g/l tryptone, 5g/l yeast extract, 10g/l NaCl, 10g/l bacteriological agar, 1mM IPTG, pH 7.0 at 25°C), containing 50µg/ml of kanamycin. Colonies were screened through colony PCR, using primers ORG1 and ORG2 (section 2.1.3). After analyzing the colony PCR products on 2% (m/w) agarose gel* containing 1X TAE buffer (40mM Tris base, 20 mM acetic acid, 1mM EDTA disodium salt dihydrate, pH 8.0 at 25°C), selected colonies were grown in 4ml of LB broth (10g/l tryptone, 5g/l yeast extract, 10g/l NaCl, pH 7.0 at 25°C) overnight at 37°C with 200 rpm⁻¹ shaking. Plasmids of interest (J.Pl1.34 – J.Pl1.43, J.Pl1.48, JPl1.49; section 2.1.2) were purified from grown cultures using GeneJET™ Plasmid Miniprep Kit (“Thermo Fisher Scientific”) and sequence-verified by Nanopore or Sanger sequencing (Branton et al., 2008; Sanger et al., 1977).

**Note – all mentioned agarose gels contain 0.5X of SYBR™ Safe DNA Gel Stain (“Thermo Fisher Scientific”).*

2.2.4 Preparation of linear gRNA templates

Linear templates of gRNAs of FnCas12a, Fn3Cas12 SpCas9 were PCR-amplified from blunt-end cloned plasmids JPl1.24, JPl1.25, JPl1.21, JPl1.23 (sections 2.1.2, 2.2.2). The remaining linear gRNA templates of Cas12a homologs were amplified from the plasmids that underwent quick-change mutagenesis prior (section 2.2.3). The PCR was done using homemade Phusion DNA polymerase with plasmid templates (section 2.1.2) and custom primers (section 2.1.3) listed in Table 2.6. PCR amplification of gRNA templates for SpCas9 and Cas12a homologs differed by final concentrations of the primers and number of PCR cycles as described in Table 2.6.

Table 2.6. Summary of amplification of linear gRNA templates for SpCas9 and Cas12a homologs.

Nuclease	gRNA template	PCR template	PCR primers	Final concentrations of the primers	PCR settings
SpCas9	D	JPl1.21	J-Pr4.74, J-Pr4.75	0.25µM	98°C 30s 98°C 10s 55°C 30s 72°C 5s 4°C ∞ } x20
	E	JPl1.23	J-Pr4.74, J-Pr4.76		
FnCas12a, Fn3Cas12a	D	JPl1.24	J-Pr4.71, J-Pr4.72	0.5µM	98°C 30s 98°C 10s 55°C 30s 72°C 5s 4°C ∞ } x35
	E	JPl1.25	J-Pr4.71, J-Pr4.73		
AdurbCas12a, HkCas12a	D	JPl1.34	J-Pr4.71, J-Pr4.72		
	E	JPl1.35	J-Pr4.71, J-Pr4.73		
Adurb336Cas 12a	D	JPl1.36	J-Pr4.71, J-Pr4.72		
	E	JPl1.37	J-Pr4.71, J-Pr4.73		
AsCas12a	D	JPl1.38	J-Pr4.71, J-Pr4.72		
	E	JPl1.48	J-Pr4.71, J-Pr4.73		
LbCas12a	D	JPl1.49	J-Pr4.71, J-Pr4.72		
	E	JPl1.39	J-Pr4.71, J-Pr4.73		
PdCas12a	D	JPl1.40	J-Pr4.71, J-Pr4.72		
	E	JPl1.41	J-Pr4.71, J-Pr4.73		
PiCas12a	D	JPl1.42	J-Pr4.71, J-Pr4.72		
	E	JPl1.43	J-Pr4.71, J-Pr4.73		

Linear PCR-amplified gRNA templates were analyzed on 4% (m/w) agarose gel* containing 1X TBE buffer (100mM tris base, 100mM boric acid, 2mM EDTA disodium salt dihydrate, pH 8.0 at 25°C). Cas12a gRNA templates were purified from the same agarose gel using GeneJET™ Gel Extraction Kit (“Thermo Fisher Scientific”) and gRNA templates for Cas9 were purified using GeneJET™ PCR Purification Kit (“Thermo Fisher Scientific”).

**Note – all mentioned agarose gels contain 0.5X of SYBR™ Safe DNA Gel Stain (“Thermo Fisher Scientific”).*

2.2.5 Cell-free protein synthesis

Cell-free protein synthesis of the selected nucleases were carried out in homemade and commercially available cell-free extracts. 125fmol of nuclease encoding plasmids (JPl1.10 – JPl1.12, JPl1.15 – JPl1.17, JPl1.31, JPl1.32, JPl1.45, JPl1.46; section 2.1.2) containing StrepII-tags and 2.5 pmol of their corresponding linear gRNA templates were mixed with 1X energy mix (50mM HEPES-K, 1.5mM ATP, 1.5mM GTP, 0.9mM CTP, 0.9mM UTP, 33mM phosphoenolpyruvate, 0.2mg/ml tRNA from *E. Coli* MRE600 (“Roche”), 0.33mM NAD, 0.27mM coenzyme A, 0.068µM folinic acid, 1.5mM spermidine), 1.25mM L-Leu, 1.5mM L-amino acid mix (Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val), 8mM magnesium glutamate, 150mM potassium glutamate, 1.4% (V/V) PEG8000, 2mM DTT, 1mM IPTG, 10 mg/ml *E. Coli* (DE3) Star (section 2.1.1) cell-free extract (homemade), 27.75µM GamS (homemade). The commercial cell-free extract from NEBExpress® Cell-free *E. coli* Protein Synthesis System (“New England BioLabs”) was mixed with the provided components (Protein Synthesis Buffer, RNase Inhibitor Murine, T7

RNA Polymerase) in addition of 1mM IPTG and 27.75 μ M GamS (homemade). Protein and linear gRNA expression templates were used in identical order and ratio (125fmol:2.5pmol) as in the homemade cell-free extracts. For both extracts, an sfGFP-StrepII encoding plasmid PRG10 (section 2.1.2) was included as a positive control for protein synthesis. Negative controls had no DNA templates included. All reactions were assembled on ice, initiated by transferring them into prewarmed Eppendorf ThermoMixer™ C (“Thermo Fisher Scientific”), and performed for 20h at 30°C, 800 rpm⁻¹. Afterward, part of the reaction samples was collected for protein purification and detection assays, later described in sections 2.2.6 and 2.2.7.

For comparative RNA analysis, part of the CFPS samples were treated with DNase I (“New England BioLabs”) and RNase A (“Thermo Fisher Scientific”) in 1X DNase I reaction buffer (“New England BioLabs”) for 1 hour at 37°C. Nucleic acid samples were denatured in 1X formamide loading solution (48% (V/V) formamide, 10mM EDTA, 1g/l Orange-G) at 70°C for 5 min and resolved in 10% urea-PAGE (AA/BAA ratio 29:1) gel containing 1X TBE buffer (100mM tris base, 100mM boric acid, 2mM EDTA disodium salt dihydrate, pH 8.0 at 25°C). Urea-PAGE gels were post-stained in SYBR™ Gold Nucleic Acid Gel Stain (“Thermo Fisher Scientific”) according to the protocol of the manufacturer and imaged with Amersham™ Typhoon™ (“GE Healthcare Life Sciences”) laser scanner.

2.2.6 Protein purification

Cell-free protein synthesis was followed by protein purification using “IBA Lifesciences” Strep-Tactin™ and Strep-Tactin™XT coated microplates. CFPS samples were spun at 13000 rpm⁻¹ (table-top centrifuge) to separate insoluble precipitates from the liquid fraction. The remaining supernatant was loaded into Strep-Tactin-coated microplates and incubated on ice for 30 min to immobilize proteins of interest. Protein-loaded plates were washed with 100mM Tris-HCl (pH 8.0 at 25°C), 150mM NaCl, and 1mM EDTA solution. Elution was done with buffers, designated for each protein purification system: 100mM Tris-HCl (pH 8.0 at 25°C), 150mM NaCl, 1mM EDTA, 50mM D-(+)-Biotin for Strep-Tactin™XT microplate and 100mM Tris-HCl (pH 8.0 at 25°C), 150mM NaCl, 1mM EDTA, 2.5mM D-Desthiobiotin for Strep-Tactin™. Collected immobilization flowthrough and elution samples were mixed with 1X SDS loading dye (50 mM Tris-HCl pH 6.8, 1% (V/V) β -mercaptoethanol, 20g/l SDS, 10% (V/V) glycerol, 2g/l bromophenol blue) and incubated at 95° for 5 min. Denatured protein samples were resolved in 8% SDS-PAGE (AA/BAA ratio 37.1:1) gel ran in 1X Tris-glycine SDS-PAGE buffer (25mM tris base, 19.2mM glycine, 3.5mM SDS, pH 8.0 at 25°C) and stained with SYPRO™ Ruby Protein Gel Stain (“Thermo Fisher Scientific”). Imaging of fluorescent-stained samples was done using Amersham™ Typhoon™ (“GE Healthcare Life Sciences”) laser scanner and GelDoc Go imaging system (“Bio-Rad Laboratories”). The same protein gel was then post-stained with Coomassie solution (1 g/l Coomassie R-250, 40% (V/V) ethanol, 10% (V/V) acetic acid) and destained by heating in a solution, containing 10% (V/V) ethanol and 7% (V/V) acetic acid.

2.2.7 Western blot

Western blot was done following similar methods as previously described by (Towbin et al., 1979). SDS-denatured protein samples from CFPS and purification assays were resolved in 8% SDS-PAGE (AA/BAA ratio 37.1:1) gel ran in 1X SDS-PAGE buffer (25mM tris base, 19.2mM glycine, 3.5mM SDS, pH 8.0 at 25°C). Four pieces of Whatman filter paper and one piece of PVDF membrane were cut according to the size of the gel. The filter paper pieces were submerged into 1X transfer buffer (160mM tris base, 620mM glycine, 155mM tricine, 2.5mM EDTA, pH 8.0 at 25°C). The same

was done to the PVDF membrane that was activated in 100% methanol prior. A protein “transfer sandwich” was created by stacking the components as follows: two sheets of wet filter paper, wet PVDF membrane, protein gel, two sheets of wet filter paper. The transfer was done for 15 minutes on Thermo Scientific™ Pierce™ Power Blotter (“Thermo Fisher Scientific”) with the current set to 1.3A.

The PVDF membrane was blocked in blocking solution (40g/l skim milk powder, 137mM NaCl, 2.7mM KCl, 10mM Na₂HPO₄, 9mM KH₂PO₄, 0.2% (V/V) Tween 20) and incubated at room temperature with 1:4000 dilution of StrepII-tag Antibody HRP Conjugate (“Sigma-Aldrich”) for 1 hour. The membrane was extensively washed in 1X wash solution (137mM NaCl, 2.7mM KCl, 10mM Na₂HPO₄, 9mM KH₂PO₄, 0.2% (V/V) Tween 20)) and then exposed to the Thermo Scientific™ SuperSignal™ West Femto Maximum Sensitivity Substrate (“Thermo Fisher Scientific”) to start the HRP chemiluminescence reaction. Signal detection was done using the “Bio-Rad Laboratories” ChemiDoc MP Imaging System. Afterward, the PVDF membrane was stained in Ponceau S solution (1 g/l Ponceau S, 5% (V/V) glacial acetic acid) following drying and imaging.

3. RESULTS AND DISCUSSION

3.1 CRISPR-Cas nucleases selected for the study

CRISPR-Cas-based genome editing is still facing challenges, mostly related to off-target effects which can cause unpredicted target cleavage. Such CRISPR-Cas nuclease activity can be followed by unanticipated genomic alterations and rearrangements. The mentioned limitation can be overcome by utilizing different CRISPR-Cas nucleases that inherently have different targeting properties. Although phylogenetical studies revealed a robust variety of CRISPR-Cas nucleases (Makarova et al., 2020), their characterization is usually long and tedious due to lengthy experimental steps, such as cell cultivation and protein purification. To overcome this, we are aiming to develop a high-throughput method to benchmark programmable nucleases (Fig 3.1).

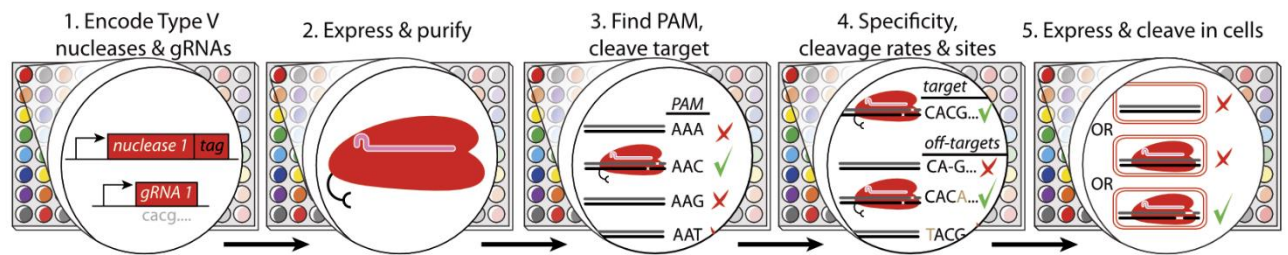


Figure 3.1. High-throughput strategy to benchmark CRISPR-Cas nucleases. Adapted from (Jones, unpublished data).

To ensure the efficiency of the characterization strategy, we selected ten CRISPR-Cas nucleases that were demonstrated to be active *in vitro* (some *in vivo* too) and have their PAM sequence requirements determined (Table 3.1). The set of the selected nucleases includes nine Cas12a homologs from ADurb.Bin193 (Adurb193Cas12a), *Acidaminococcus* species (AsCas12a), *Lachnospiraceae* bacterium (LbCas12a), ADurb.Bin336 (Adurb336Cas12a), *Francisella novicada* (Fn3Cas12a), *Francisella novicada* U112 (FnCas12a), *Prevotella ihumii* (PiCas12a), *Prevotella disiens* (PdCas12a) and *Helcococcus kunzii* (HkCas12a) (Bernd Zetsche, 2015; Jacobsen et al., 2020; Tang et al., 2017) along with Cas9 from *Streptococcus pyogenes* (SpCas9) (Jinek et al., 2012). The prevalence of Cas12a nucleases is related to their independence of tracrRNAs since they require only a single crRNA molecule to form an effector complex. Even though the tracrRNA sequence of SpCas9 is known, less characterized Cas9 homologs would require additional studies for their tracrRNA identification. This would be a lengthy process as the tracrRNA determination through bioinformatic methods faces challenges such as variability of tracrRNA sizes, sequences, and their different locations at the CRISPR arrays (Dooley et al., 2021).

Table 3.1. Nucleases, selected for this work, along with their identified activities and PAM requirements.

Nuclease	Demonstrated activity		PAM sequence (5' → 3')
	<i>in vitro</i>	<i>in vivo</i>	
SpCas9	+	+	NGG
Adurb193Cas12a	+	-	TTTV
AsCas12a	+	+	TTTV
LbCas12a	+	+	TTTV
Adurb336Cas12a	+	-	TTTV

Continued Table 3.1.

Fn3Cas12a	+	-	YTV
FnCas12a	+	+	TTN
PiCas12a	+	+	KKYV
PdCas12a	+	+	TTTV
HkCas12a	+	+	YYV

The planned characterization strategy is comprised of five stages: cloning of the nuclease genes and preparation of their gRNA templates; cell-free expression of the nucleases and their microplate-based purification; determination of the PAM requirements; assessment of the cleavage specificity and dynamics *in vitro*; evaluation of the nuclease activity *in vivo*. Nevertheless, this work was focused on the first two stages, which involve cloning, cell-free expression, and purification of the selected nucleases. The results of these experiments are presented in the further sections.

3.2 Vector design

To proceed with the nuclease cloning, expression, and purification experiments, I designed two custom vectors to clone the genes of the selected nucleases (Fig. 3.2). These vectors only differ by the position of the StrepII affinity tag, which can be fused on either C- or N-terminus of the cloned nucleases. Both vectors contain the following sequences: T7 promoter, T7 terminator, ribosome binding site (RBS), Lac operator, kanamycin resistance gene, a multiple cloning site (MCS) (flanked by two BsaI endonuclease sites), XTEN linker (SGSETPGTSESATPES) (Schellenberger et al., 2009) and the sequence of the StrepII affinity tag. Also, both vectors encode the Lac operator inhibitor protein (LacI), which binds to the Lac operator thus inhibiting the transcription of the cloned nuclease. Synthesis of the cloned protein is induced by the addition of IPTG. The inclusion of BsaI restriction endonuclease sites at the MCS allows for Golden Gate cloning (Engler et al., 2008).

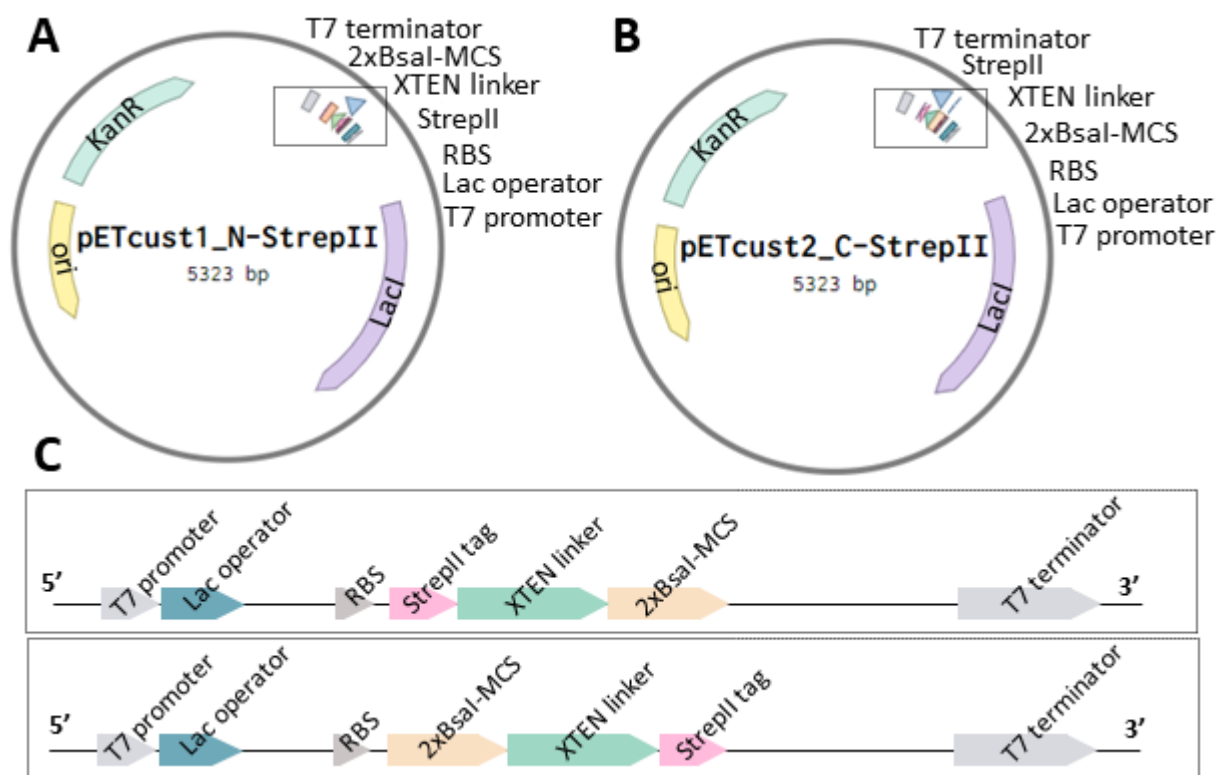


Figure 3.2. Customized cloning vectors, containing sequences required for nuclease cloning, expression, and purification. **A)** A customized vector, containing the StrepII affinity tag on the N-terminus. **B)** A customized vector, containing the StrepII affinity tag on the C-terminus. **C)** Close-ups of the rectangular boxes, containing sequence elements, listed on the right side of each corresponding vector. MCS – multiple cloning site (contains two BsaI sites), RBS – ribosome binding site, ori – origin of replication, KanR – kanamycin resistance gene, LacI – Lac operator inhibitor gene, XTEN linker – peptide (SGSETPGTSESATPES) sequence, bridging between the MCS and the StrepII-tag.

For further experiments, I proceeded with the vector, containing the StrepII-tag on the C-terminus (Fig. 3.2. B). This vector ensures later purification of the full-length nucleases only, as the StrepII-tag is synthesized at the end of the protein translation. If the StrepII-tag was fused at the N-terminus, there would be a chance to purify StrepII-tagged proteins that were not translated entirely. The premature termination of the translation can happen if the mRNA transcripts are degraded on the 3' end, unable to serve as templates to produce full-length proteins.

3.3 Cloning of nucleases and gRNA templates

During this work, I successfully amplified and cloned the genes of Adurb193Cas12a, HkCas12a, PiCas12a, and Adurb336Cas12a (Fig. 3.3) into a customized cloning vector, containing C-terminal StrepII sequence (Fig. 3.2. B; section 2.2.1). The remaining 6 nuclease genes of SpCas9, AsCas12a, LbCas12a, FnCas12a, Fn3Cas12a, and PdCas12a were cloned previously (Grigaitis, unpublished data). Sequences of all 10 constructs were confirmed via sequencing and used for cell-free protein synthesis (CFPS).

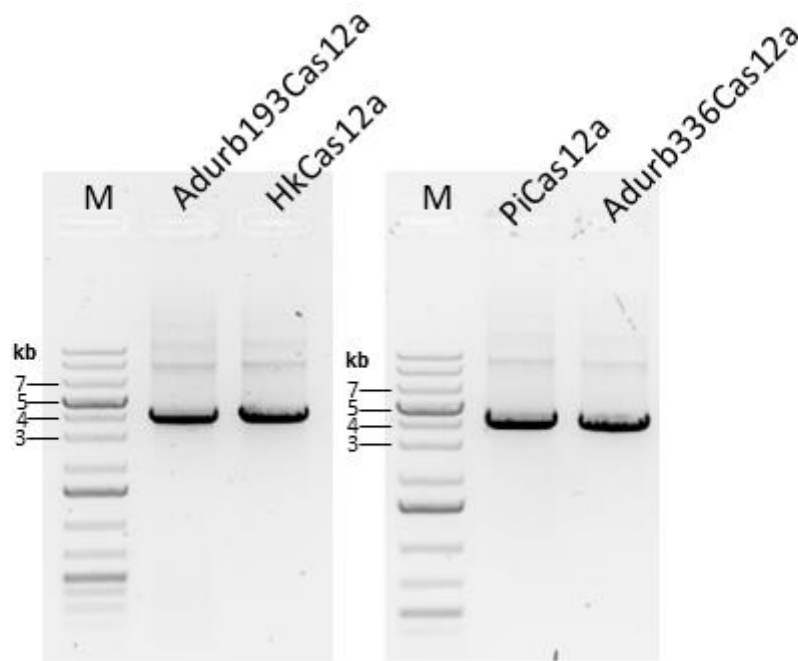


Figure 3.3. PCR-amplified cloning inserts (~4.6kb) of Adurb193Cas12a, HkCas12a, PiCas12a and Adurb336Cas12a. M – DNA size marker.

To prepare templates of gRNAs for their production during nuclease expression through CFPS, I initially amplified linear gRNA templates of SpCas9, FnCas12a, and Fn3Cas12a (Fig. 3.4 B; section 2.2.2). I then used these templates for blunt-end cloning into a commercially available vector (Fig. 3.4 A; section 2.2.2). Each gRNA template has two alternatives (D and E) which differ by the targeting (spacer) sequences (Fig. 3.4 C). More specifically, spacers D and E are 20nt length sequences of gRNAs, essential to target DNA libraries previously described by (Jones et al., 2019). These libraries were used to determine the cleavage and specificity dynamics of several AsCas12a and SpCas9 variants, described in that study. We are planning to use the same DNA libraries for the high-throughput characterization strategy to assess the targeting specificity and dynamics of the newly identified nucleases.

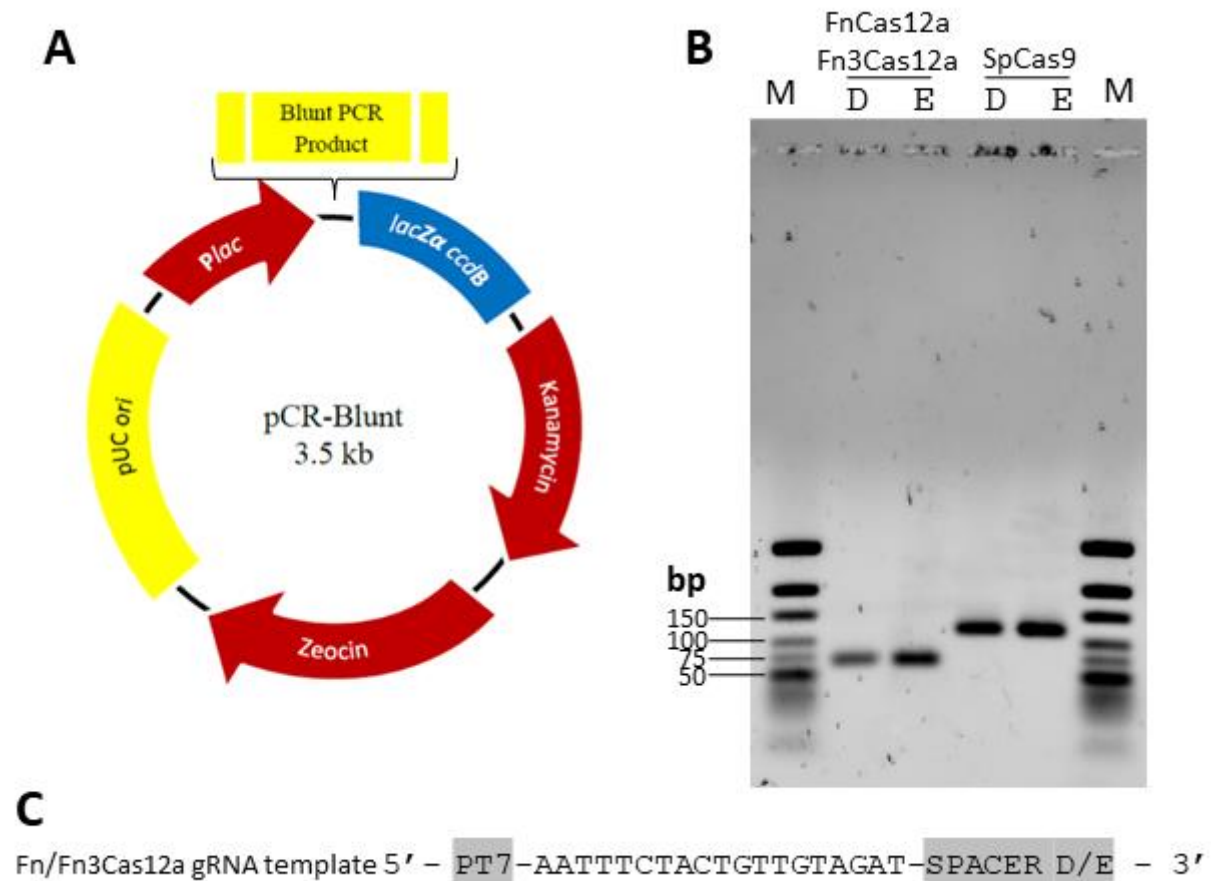


Figure 3.4. Overview of gRNA template cloning. **A)** A plasmid map of pCR™-Blunt vector, designed to clone blunt-end PCR products. Contains Lac promoter (Plac), kanamycin and zeocin resistance, origin of replication (pUC ori), and lethal *E. coli* *ccdB* gene fused to the LacZα (*lacZα-ccdB*) (Bernard et al., 1994). LacZα-*ccdB* expression gets disrupted upon blunt fragment ligation, allowing the growth of only positive recombinants following transformation. Adapted from (Tabatabaei Yazdi et al., 2015) and “Thermo Fisher Scientific”. **B)** PCR-amplified linear gRNA templates D and E of FnCas12a, Fn3Cas12a, and SpCas9. FnCas12a and Fn3Cas12a share the same gRNA sequences. M – DNA size marker. **C)** Linear gRNA template of FnCas12a/Fn3Cas12a, containing alternative spacer variants D and E. PT7 – T7 promoter.

gRNA sequences of selected Cas12a homologs contain a few nucleotide differences. To address these differences, I used plasmids that encode gRNA-D and gRNA-E of Fn/Fn3Cas12a as parental templates for quick-change mutagenesis (QCM) (Fig. 3.5; section 2.2.3).

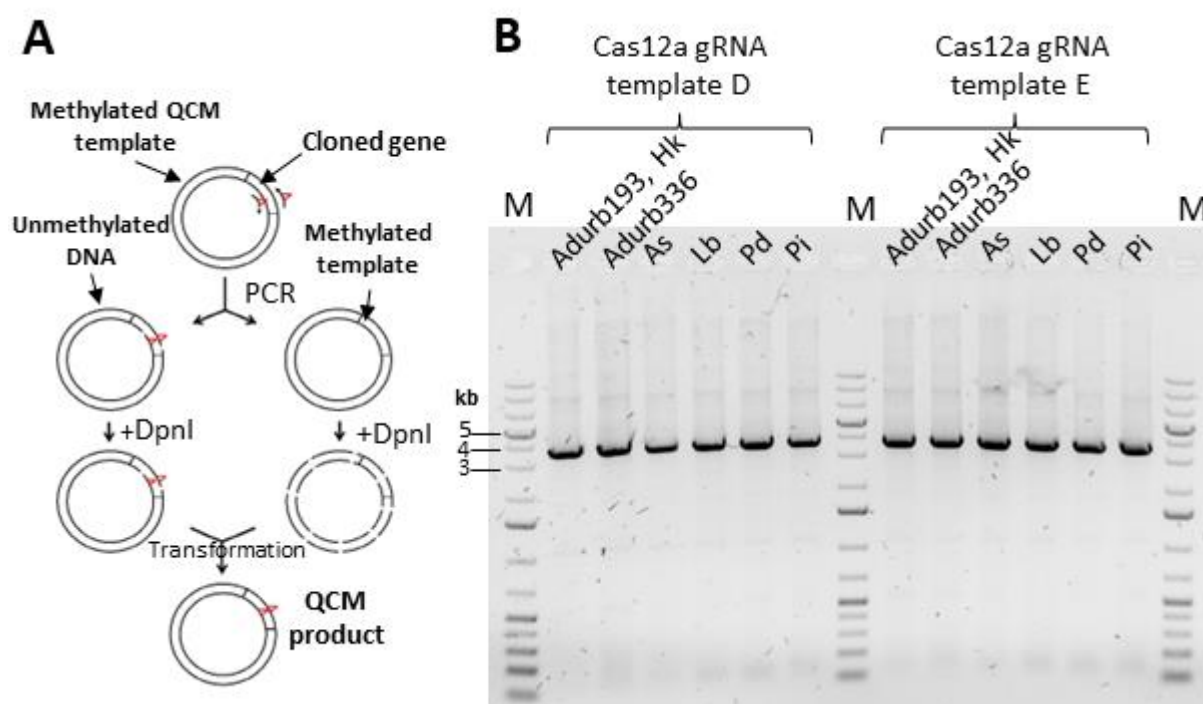


Figure 3.5. Overview of quick-change mutagenesis (QCM). **A)** QCM principle. Overlapping primers, containing the desired mutation, are used on a plasmid template. Following PCR, samples are treated with DpnI to digest the methylated template. QCM product is then transformed into competent cells for nick repair and plasmid amplification. Adapted from (Mandrich, 2015). **B)** QCM for gRNA templates of the selected Cas12a homologs. Plasmids, encoding gRNA-D and gRNA-E for Fn/Fn3Cas12a, were used as initial QCM templates. The products represented on the gel were later transformed into competent *E. Coli* cells. M – DNA size marker.

gRNA template cloning and QCM experiments resulted in a total of 16 sequence-verified plasmids. However, the produced gRNA encoding plasmids are not compatible with gRNA expression in cell-free extracts due to a lack of T7 terminator sequences following the gRNA encoding sequences. Because of the mentioned reason, I used the assembled plasmids to amplify the gRNA templates as linear DNA (Fig. 3.6; section 2.2.4).

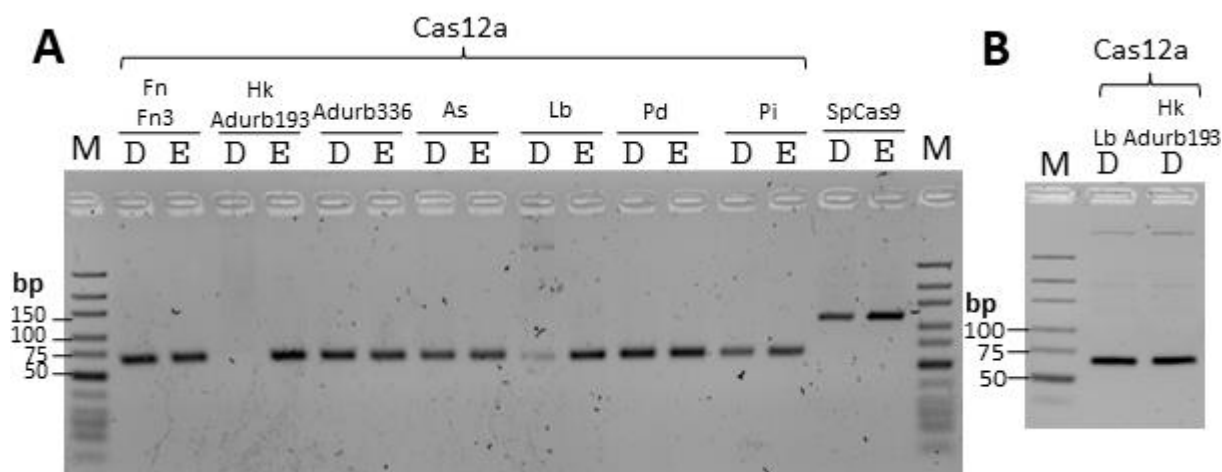


Figure 3.6. Linear gRNA-D and gRNA-E templates of all selected nucleases. **A)** Linear gRNA templates D and E of Cas12a homologs and SpCas9. gRNA-D templates of Adurb193Cas12a and HkCas12a are absent. The amplification of LbCas12a gRNA-D template produced additional unanticipated band. **B)** Linear gRNA-D templates of LbCas12a, HkCas12a, and Adurb193Cas12a. M – DNA size marker.

The amplified linear templates of gRNAs contain T7 promoter sequences (Fig. 3.7) which are later utilized in CFPS for run-off transcription (Loewenstein et al., 2007) to produce gRNAs of the corresponding nucleases.

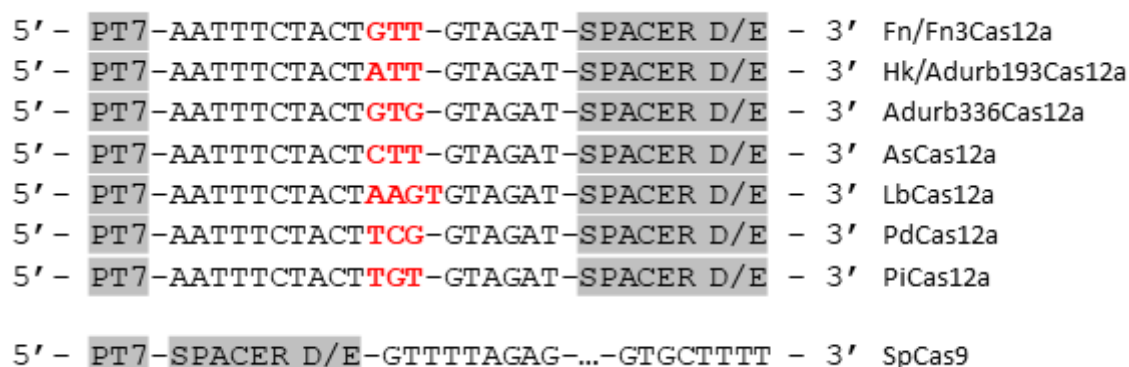


Figure 3.7. Linear gRNA templates of the selected CRISPR-Cas nucleases. The red nucleotides highlight sequence differences between the gRNAs of Cas12a homologs. These differences were addressed through quick-change mutagenesis. FnCas12a and Fn3Cas12a as well as HkCas12a and Adurb193Cas12a share the same gRNA sequences. An ellipsis in the SpCas9 gRNA template sequence indicates omitted nucleotides. PT7 – T7 promoter.

After finishing the cloning experiments of the nucleases and their gRNAs I then proceeded with the CFPS experiments.

3.4 Cell-free protein synthesis

For the initial cell-free protein synthesis (CFPS) of AsCas12a and SpCas9, I tested both homemade and commercial cell-free extracts (CFEs) (section 2.2.5). To express these nucleases, I added AsCas12a-StrepII and SpCas9-StrepII encoding plasmids along their corresponding linear gRNA-D templates into the CFPS reaction mix. Negative control reactions contained no DNA templates whereas the positive control contained a plasmid encoding sfGFP-StrepII, which was

previously used for the homemade CFE optimization experiments (Grigaitis, unpublished data). We selected AsCas12a and SpCas9 since they are known to be active *in vitro* and their expression (and purification) was extensively described in other studies (Jinek et al., 2012; Tang et al., 2017).

Following CFPS, StrepII-tagged proteins of interest were detected via Western blot (Fig 3.8; section 2.2.7). As expected, the negative controls in both extracts contained no signal of proteins of interest. A signal of sfGFP-StrepII was visible in both homemade and commercial CFEs, however, the signal of SpCas9 and AsCas12a was absent in the samples of the homemade extract. Notably, transcription of the sfGFP gene produces shorter mRNA molecules, as the protein is over 5 times smaller (~29 kDa) than AsCas12a (~154 kDa) and SpCas9 (~161 kDa). Long mRNA transcripts are known to be less stable since they are more likely to be degraded by endogenous RNases (Laalami et al., 2014). In this case, the transcripts of the Cas nucleases could have been degraded by RNases present in the homemade CFEs leading to the absence of SpCas9 and AsCas12a.

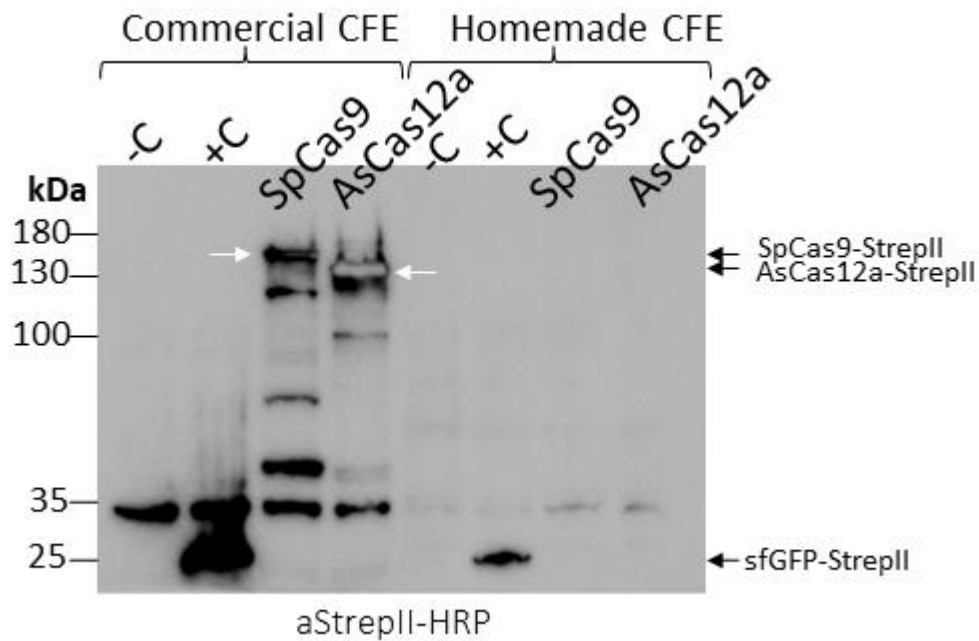


Figure 3.8. Western blot assay of SpCas9, AsCas12a, and sfGFP following cell-free protein synthesis in commercial and homemade cell-free extracts (CFEs). A signal of SpCas9 and AsCas12a is absent in the homemade CFE. White arrows indicate accurate signal positions of the SpCas9 and AsCas12a. -C – negative control, +C – positive control (sfGFP-StrepII), aStrepII-HRP – anti-StrepII-tag antibody conjugated with an HRP (horseradish peroxidase), which produces a chemiluminescence signal for this assay. The protein size markers are placed on the left side of the image.

gRNA observations were done following SpCas9 and AsCas12a expression in the homemade and commercial CFEs (section 2.2.5). I treated the same CFPS samples with either DNase or RNase to compare the backgrounds of nucleic acids after the CFPS. The expected 66nt and 118nt length gRNAs of AsCas12a and SpCas9 are not visible on the gel, possibly due to an excess of ribosomal RNAs and other nucleic acids originating from the CFEs (Fig. 3.9).

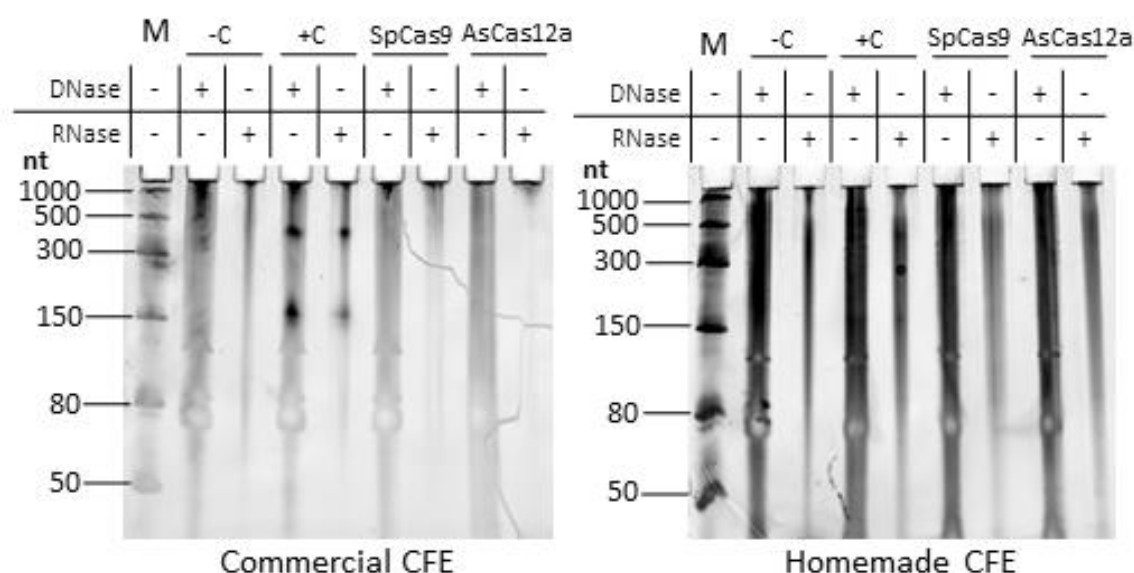


Figure 3.9. Comparative nucleic acid analysis following cell-free protein synthesis (CFPS) of SpCas9 and AsCas12a. Identical CFPS samples, presented in Fig. 3.8., were treated with DNase and RNase, accordingly. DNase-treated samples produce smears on the gel, indicating the abundance of RNAs that might overshadow gRNAs of SpCas9 (expected length 118nt) and AsCas12a (expected length 66nt). -C and +C stand for negative and positive control samples from the initial CFPS experiment, that were treated with DNase and RNase afterward. CFE – cell-free extract, M – RNA size marker.

To remove the overshadowing effect, the presence of gRNAs could be analyzed after purification of the Cas nucleases which would eliminate unwanted nucleic acids. Alternatively, other more sensitive and specific methods, such as Northern blotting (Pougach & Severinov, 2012), can be utilized to detect the gRNAs.

Since the expression of SpCas9 and AsCas12a was successful in the commercial extract only, I used it to express the remaining nucleases (section 2.2.5). This experiment resulted in successful expression of nine out of ten C-terminally tagged nucleases with PiCas12a being absent (Fig. 3.10). Perhaps it was not detected due to StrepII-tag interference with protein folding which might have caused aggregation and subsequent precipitation of the protein. In this case, the StrepII-tag can be fused on the N-terminus. Also, the absence of PiCas12a can be related to inappropriate CFPS conditions. Since the commercial extract comes with a pre-made reaction buffer, optimization of the reaction components is limited. However, to improve protein expression and stability, CFPS reaction temperature might be adjusted, according to the troubleshooting manual of the commercial CFE.

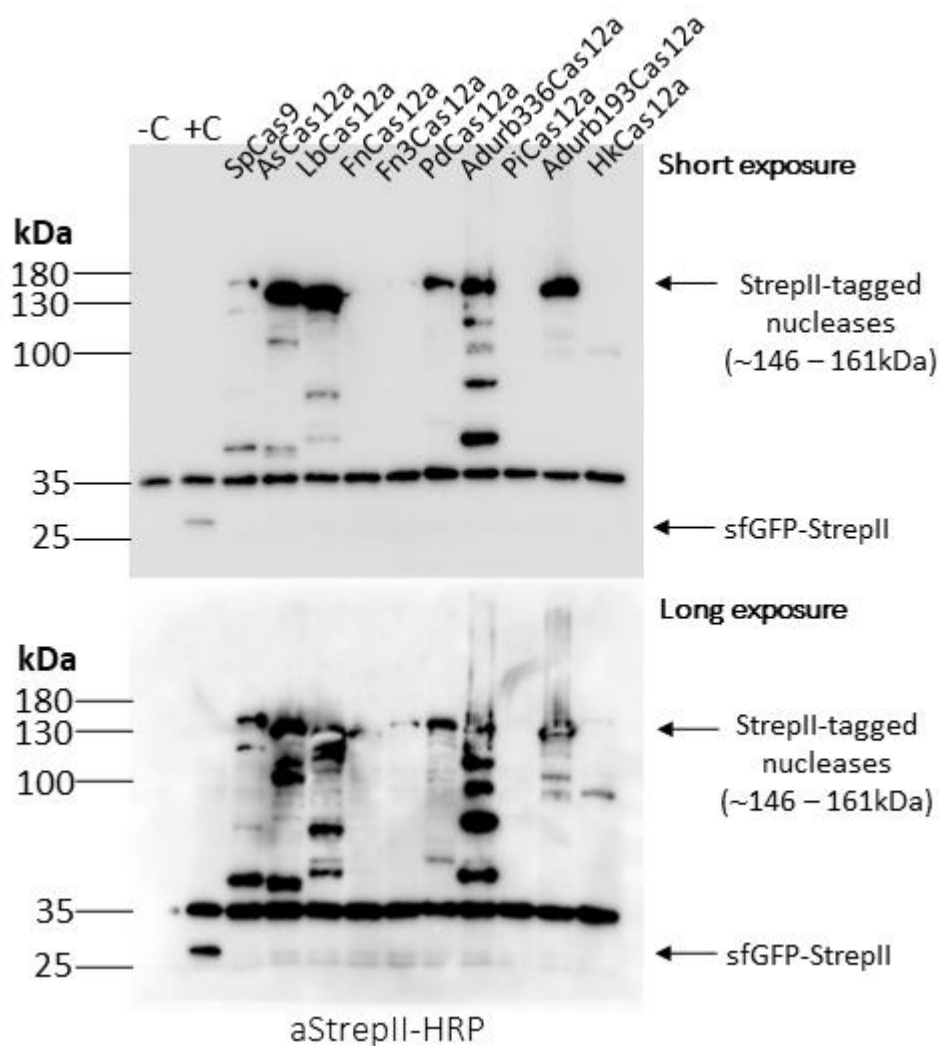


Figure 3.10. Western blot assay of all ten selected nucleases following their expression in the commercial cell-free extract (CFE). Longer exposure presents signals of FnCas12a, Fn3Cas12a, and HkCas12a. However, the signal for PiCas12a is absent. -C – negative control of the cell-free protein synthesis (CFPS), +C – positive control of the CFPS, containing sfGFP-StrepII encoding plasmid, aStrepII-HRP – anti-StrepII-tag antibody conjugated with an HRP (horseradish peroxidase), which produces a chemiluminescence signal for this assay. The protein size markers are placed on the left side of the image.

3.5 Purification of AsCas12a

Following the nuclease expression and detection experiments, I tested a microplate-based strategy for high-throughput nuclease purification (section 2.2.6). For the purification experiment, I used 96-well Strep-Tactin™ and Strep-Tactin™XT coated microplates. These microplates are coated with slightly different versions of modified streptavidin and offer different binding capacities of the StrepII-tagged proteins. Since AsCas12a showed the best expression through CFPS, I simultaneously tested these two purification systems to purify this nuclease.

To detect the purified AsCas12a-StrepII and evaluate the efficiency of the purification, I collected immobilization flowthrough and elution samples throughout each step and used them for Western blot assay (Fig. 3.11; section 2.2.7). Although the purified protein was detected in the elution sample, the Western blot revealed that the immobilization flowthrough contains a large amount of

unbound protein. Further troubleshooting experiments need to be done to determine if the protein does remain unbound or if it saturates the microplate to its binding capacity.

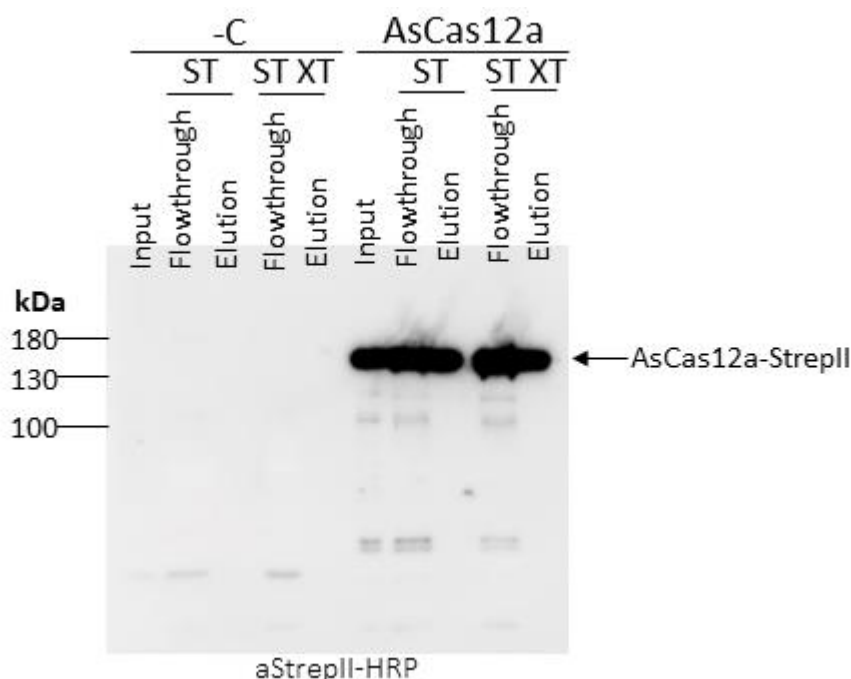


Fig. 3.11. Western blot assay for comparative analysis of protein purification samples from Strep-Tactin™ (ST) and Strep-Tactin™XT (ST XT) coated microplates. Although the blot provided a signal of AsCas12a in the elution samples, a strong signal in the flowthrough sample indicates that a large amount of loaded protein remains unbound. The input contains crude cell-free protein synthesis samples. -C – negative control of cell-free protein synthesis. StrepII-HRP – StrepII-tag antibody conjugated with an HRP (horseradish peroxidase), which produces a chemiluminescence signal for this assay. The protein size markers are placed on the left side of the image.

Following the detection experiment, I used the same samples to evaluate the purity of the eluted protein. To do that, I resolved the samples on a gel and stained them with a fluorescent (SYPRO™ Ruby) and then Coomassie dye (section 2.2.6). Unfortunately, I was unable to evaluate the purity of the elution samples as the characteristic band of AsCas12a was not visible on either fluorescent-stained or Coomassie-stained gel (Fig. 3.12). The absence of the eluted AsCas12 suggests that the amount of the purified protein is probably low as it was detected through Western blot only. However, there is a chance that the lack of eluted AsCas12a is caused by improper sample loading ratios which should be taken into consideration. Also, like in the Western blot, imaging of the fluorescent and Coomassie-stained samples revealed that the immobilization flowthrough contains a lot of unbound protein.

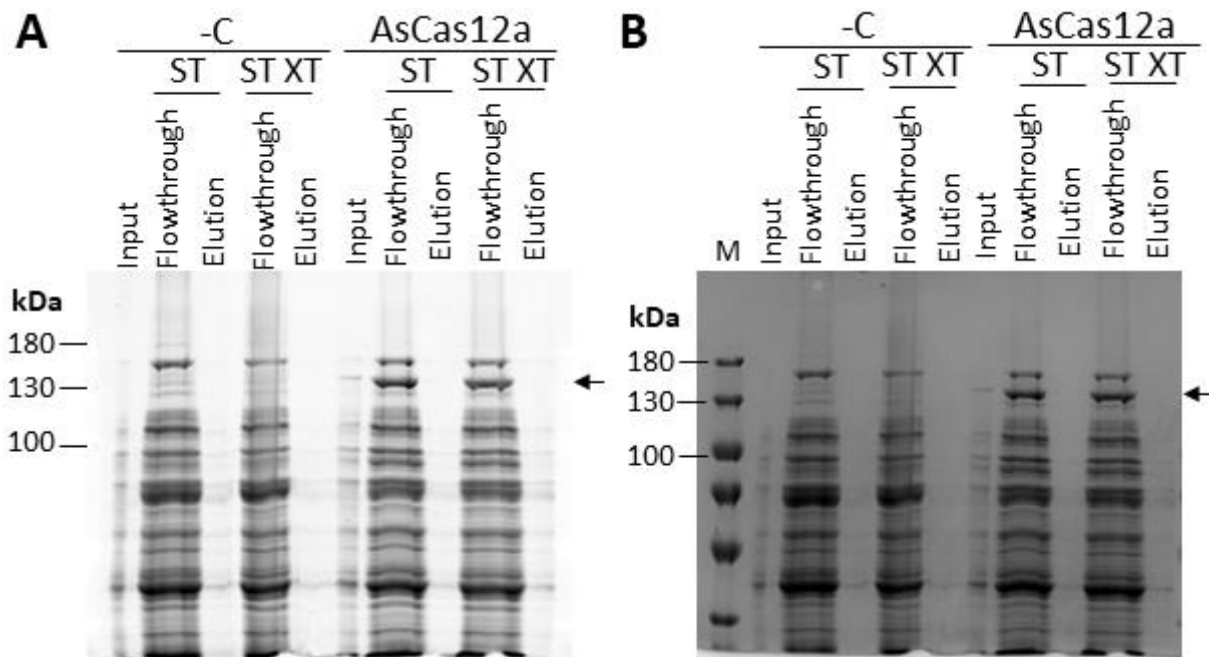


Figure 3.12. Stained purification samples of AsCas12a, collected from Strep-Tactin™ (ST) and Strep-Tactin™XT (ST XT) coated microplates. **A)** Fluorescent (SYPRO™ Ruby) staining. **B)** Coomassie blue staining. Input stands for crude CFPS samples loaded into the microplates. The input samples were compared to immobilization flowthrough and elution fractions. The flowthrough contains unbound AsCas12a and the expected band of AsCas12a is absent in the elution samples. The black arrows indicate the position of a characteristic band of AsCas12a. -C – negative control of CFPS, M – protein size marker.

Conclusively, the Western blot as well as fluorescent and Coomassie staining of AsCas12a purification samples revealed an abundance of AsCas12a remaining in the immobilization flowthrough. This might be caused by improper purification conditions or the limited binding capacity of the microplates. Also, the eluted protein could be detected through Western blot only as the fluorescent and Coomassie-stained samples provided no signal for AsCas12a. These results suggest that the concentration of the purified protein is probably low, and that the microplate-based purification requires further troubleshooting. Notably, the inconclusive results regarding the elution samples could be related to experimental errors which would require repetition of the experiment. Regardless, the conditions of the purification procedure can be optimized by changing the temperature and by adjusting the durations of each purification step.

3.6 Discussion

The cloning, cell-free protein synthesis (CFPS), and microplate-based purification experiments were done as a contribution to develop a high-throughput strategy to characterize and benchmark CRISPR-Cas nucleases. Cell-free protein synthesis (CFPS) platforms were already utilized to determine PAM requirements and assess target cleavage dynamics of several CRISPR-Cas nucleases (Leenay et al., 2016; Marshall et al., 2018). Based on similar methods utilized in these studies, I conducted initial CFPS experiments in homemade and commercial cell-free extracts (CFEs) to express AsCas12a and SpCas9. Western blot analysis with the anti-StrepII-HRP antibody demonstrated successful expression of AsCas12a and SpCas9 in the commercial extracts only. Although the homemade extracts failed to produce the nucleases, an sfGFP, used as a positive control,

was detected in both tested extracts. These results suggest that the synthesis of relatively smaller proteins (e.g., sfGFP) in the homemade CFEs is possible but the reaction conditions need to be optimized to produce proteins of higher molecular weights. For comparison, the molecular weights of the selected CRISPR-Cas nucleases range from around 146 kDa to 161 kDa whereas the sfGFP weighs only around 29 kDa. Accordingly, the mRNA transcripts that encode the nucleases are longer thus less stable and more prone to be degraded by RNases present in the homemade CFEs (Soltani & Bundy, 2022). The assumption regarding the stability of mRNAs is made based on a comparison of the composition of the CFPS reactions. The homemade extracts contained no RNase inhibitors whereas the mouse-derived Murine RNase inhibitor was added to the commercial extract before the CFPS. Conclusively, the homemade CFEs could be optimized by the addition of RNase inhibitors, which were previously demonstrated to improve protein production through CFPS (Scheele & Blackburn, 1979; Soltani & Bundy, 2022).

Nucleic acid analysis of the AsCas12a and SpCas9 CFPS experiment revealed that the samples might contain a large amount of various RNA species (e.g., endogenous rRNA) which overshadow the gRNA molecules produced during the CFPS. To remove the overshadowing effect, the presence of gRNAs can be analyzed after purification of the Cas nucleases which would eliminate the background nucleic acids. The purified nucleases may be denatured in high temperatures to release intact gRNA molecules which could be analyzed by similar methods used in this work. Alternatively, more sensitive and specific RNA detection methods, such as Northern blotting, can be utilized to detect gRNAs produced throughout the CFPS (Jacobsen et al., 2020; Pougach & Severinov, 2012).

Since the commercial CFEs produced AsCas12a and SpCas9, I utilized them to express all ten selected nucleases. Except for PiCas12a, the expression of the remaining nine of them was detected through Western blotting. The absence of PiCas12a can be related to StrepII interference with protein folding which could cause its aggregation and subsequent precipitation. Since the CFPS reaction in the commercial extract is carried out by the addition of a pre-made reaction buffer, adjustments of the reaction components to maintain the protein stability are limited. However, the optimization of PiCas12a expression and detection can be done by adjusting the temperature of the reaction or by using an alternative expression vector that would fuse the StrepII on the N-terminus of the protein. Nevertheless, the successful expression of the majority of the selected nucleases demonstrates the potential to utilize CFEs for the rapid production of other CRISPR-Cas nucleases as well as other proteins of interest. The utilization of CFEs to produce proteins of interest helps to avoid long experimental steps, including transformations and cell culturing, which are characteristic of classical recombinant protein synthesis methods (Silverman et al., 2020).

The attempt to determine the conditions of a microplate-based purification strategy provided inconsistent results. For this experiment, I attempted to purify AsCas12a as it showed the best expression in the previous CPFS experiments. Although I was able to detect the purified AsCas12a through Western blotting, fluorescent and Coomassie staining of the elution samples did not provide the expected signal. Also, the fluorescent and Coomassie staining as well as Western blot revealed an abundance of unbound protein remaining in the flowthrough samples. To optimize the purification conditions, the experiment requires repetition and further troubleshooting. The purification procedure can be improved by changing the temperature and durations of each purification step (e.g., longer duration of protein immobilization), or even by fusing the StrepII-tag on the other terminus of the protein. Although I cannot confidently state that the purification of AsCas12a was successful, once the conditions of the microplate-based purification are optimized, the nucleases produced during CFPS could be purified in a high-throughput manner without the need to utilize costly and low-throughput strategies such as affinity chromatography (Rodriguez et al., 2020).

In summary, this work provided insight into the utilization of high-throughput methods, such as cell-free protein synthesis (CFPS) and microplate-based protein purification. The CFPS experiments done throughout this work provided promising results for rapid production of the CRISPR-Cas nucleases without the need to rely on lengthy classical protein production methods. However, to move to the other stages of the previously described nuclease characterization strategy, cell-free expression, and protein purification remain to be improved. Nevertheless, once all the methods used in this work are optimized, they should contribute to the high-throughput characterization of the CRISPR-Cas nucleases, including identification of their PAM requirements, specificity, and activities both *in vitro* and *in vivo*.

CONCLUSIONS

The results of this work are summarized in the points listed below:

1. The cloning experiments produced four nuclease-encoding plasmids and sixteen gRNA-encoding templates.
2. Cell-free protein synthesis experiments provided successful expression of nine out of ten selected CRISPR-Cas nucleases.
3. Results of the microplate-based purification of AsCas12a showed that not all protein binds to the surface of the microplate. Detection of AsCas12a through Western blot, but not Coomassie blue and fluorescent staining, suggests that only a low concentration of protein was purified.

SUMMARY

VILNIUS UNIVERSITY
FACULTY OF CHEMISTRY AND GEOSCIENCES

URTĖ GLIBAUSKAITĖ

Cloning and Cell-Free Expression of CRISPR-Cas Nucleases

CRISPR (clustered regularly interspaced short palindromic repeats) locus and Cas (CRISPR-associated) proteins are known to comprise a defense system against bacteriophages and foreign mobile genetic elements in bacteria and archaea. During the phage infection CRISPR-Cas machinery processes and incorporates a fragment of viral DNA into the CRISPR locus. If the infection recurs, this fragment is transcribed and processed into a gRNA (guide RNA) molecule that forms an effector complex with a Cas protein (or multiple Cas proteins) and directs it to the target site which then gets cleaved. gRNAs can be engineered to target any sequence of interest making the CRISPR-Cas nucleases potential candidates for precise genome editing.

However, CRISPR-Cas nucleases have limitations related to their off-target effects, which can lead to unanticipated genomic alterations or rearrangements. To fulfill the promise of precise genome editing, such undesirable effects need to be overcome. This can be done by benchmarking newly identified phylogenetically diverse CRISPR-Cas nucleases as they inherently contain different targeting properties. To speed up the characterization process, we are aiming to develop a high-throughput strategy to benchmark CRISPR-Cas nucleases. This strategy includes the utilization of cell-free extracts (CFEs) for nuclease expression as well as microplate-based protein purification.

To ensure the efficiency of the characterization strategy, we selected a set of ten CRISPR-Cas nucleases with identified activities *in vitro*. This set is comprised of ten CRISPR-Cas nucleases (SpCas9, Adurb193Cas12a, AsCas12a, Fn3Cas12a, FnCas12a, LbCas12a, Adurb336Cas12a, PiCas12a, PdCas12a, and HkCas12a) which were cloned into a customized vector, containing StrepII affinity tag. Subsequently, I performed cloning, quick-change mutagenesis, and PCR amplification experiments to produce linear gRNA templates of the selected nucleases. I then used the gRNA templates along the corresponding nuclease-encoding plasmids for cell-free protein synthesis (CFPS). For initial CFPS experiments, I tested both commercial and homemade CFEs to express AsCas12a and SpCas9. The nuclease expression was successful in the commercial extracts only, so I utilized them to express all ten selected nucleases. Following CFPS, the expression of nine out of ten selected nucleases was confirmed, with PiCas12a being absent. Nucleic acid analysis revealed an abundance of CFE-derived RNA molecules which most likely overshadow gRNA molecules, produced during CFPS. To remove the overshadowing effect, gRNAs could be detected after purification of the nucleases or by utilization of alternative RNA detection methods. Finally, to optimize the conditions for a high-throughput affinity purification strategy, I tested Strep-Tactin-coated microplates to purify AsCas12a-StrepII. The purification results suggest that this experiment requires troubleshooting as the amount of the purified protein is probably low and most of it remains unbound to the microplates.

The experiments conducted throughout this work demonstrated the utilization of high-throughput methods, such as cell-free protein synthesis and microplate-based protein purification. These methods can contribute to the rapid benchmarking of phylogenetically diverse CRISPR-Cas nucleases which hold potential as precise genome editors.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor Dr. Rokas Grigaitis for his patience and invaluable help throughout my time at the laboratory. I feel grateful for his competent advice regarding laboratory skills and techniques, which most definitely contributed to this thesis.

I would also like to thank Dr. Stephen Knox Jones Jr. and other teammates of the “Jones! Lab” for their advice, warm welcome, and constant support.

REFERENCES

- Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A., Lander, E. S., Voytas, D. F., Ting, A. Y., & Zhang, F. (2017). RNA targeting with CRISPR–Cas13. *Nature*, 550(7675), Article 7675. <https://doi.org/10.1038/nature24049>
- Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., Shmakov, S., Makarova, K. S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E. S., Koonin, E. V., & Zhang, F. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science (New York, N.Y.)*, 353(6299), aaf5573. <https://doi.org/10.1126/science.aaf5573>
- Anastasina, M., Terenin, I., Butcher, S. J., & Kainov, D. E. (2014). A technique to increase protein yield in a rabbit reticulocyte lysate translation system. *BioTechniques*, 56(1), 36–39. <https://doi.org/10.2144/000114125>
- Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 576(7785), Article 7785. <https://doi.org/10.1038/s41586-019-1711-4>
- Asahara, H., Magnelli, P., Shi, X., Tuckey, C., Zhou, Y., & Samuelson, J. C. (2021). Guidelines for nucleic acid template design for optimal cell-free protein synthesis using an Escherichia coli reconstituted system or a lysate-based system. *Methods in Enzymology*, 659, 351–369. <https://doi.org/10.1016/bs.mie.2021.07.005>
- Babu, K., Kathiresan, V., Kumari, P., Newsom, S., Parameshwaran, H. P., Chen, X., Liu, J., Qin, P. Z., & Rajan, R. (2021). Coordinated Actions of Cas9 HNH and RuvC Nuclease Domains Are Regulated by the Bridge Helix and the Target DNA Sequence. *Biochemistry*, 60(49), 3783–3800. <https://doi.org/10.1021/acs.biochem.1c00354>
- Bandyopadhyay, A., Kancharla, N., Javalkote, V. S., Dasgupta, S., & Brutnell, T. P. (2020). CRISPR-Cas12a (Cpf1): A Versatile Tool in the Plant Genome Editing Tool Box for Agricultural Advancement. *Frontiers in Plant Science*, 11. <https://www.frontiersin.org/articles/10.3389/fpls.2020.584151>
- Barman, A., Deb, B., & Chakraborty, S. (2020). A glance at genome editing with CRISPR-Cas9 technology. *Current Genetics*, 1–18. <https://doi.org/10.1007/s00294-019-01040-3>
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, 315(5819), 1709–1712. <https://doi.org/10.1126/science.1138140>
- Bernard, P., Gabant, P., Bahassi, E. M., & Couturier, M. (1994). Positive-selection vectors using the F plasmid ccdB killer gene. *Gene*, 148(1), 71–74. [https://doi.org/10.1016/0378-1119\(94\)90235-6](https://doi.org/10.1016/0378-1119(94)90235-6)
- Bernd Zetsche, feng zhang. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 163(3), Article 3.
- Bhaya, D., Davison, M., & Barrangou, R. (2011). CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. *Annual Review of Genetics*, 45(1), 273–297. <https://doi.org/10.1146/annurev-genet-110410-132430>
- Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading, England)*, 151(Pt 8), Article Pt 8. <https://doi.org/10.1099/mic.0.28048-0>

Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M. F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K. L., & Davidson, A. R. (2015). Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature*, 526(7571), Article 7571. <https://doi.org/10.1038/nature15254>

Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S. B., Krstic, P. S., Lindsay, S., Ling, X. S., Mastrangelo, C. H., Meller, A., Oliver, J. S., Pershin, Y. V., Ramsey, J. M., ... Schloss, J. A. (2008). The potential and challenges of nanopore sequencing. *Nature Biotechnology*, 26(10), 1146–1153. <https://doi.org/10.1038/nbt.1495>

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., & van der Oost, J. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*, 321(5891), 960–964. <https://doi.org/10.1126/science.1159689>

Charlesworth, C. T., Deshpande, P. S., Dever, D. P., Camarena, J., Lemgart, V. T., Cromer, M. K., Vakulskas, C. A., Collingwood, M. A., Zhang, L., Bode, N. M., Behlke, M. A., Dejene, B., Cieniewicz, B., Romano, R., Lesch, B. J., Gomez-Ospina, N., Mantri, S., Pavel-Dinu, M., Weinberg, K. I., & Porteus, M. H. (2019). Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nature Medicine*, 25(2), Article 2. <https://doi.org/10.1038/s41591-018-0326-x>

Charpentier, E., Richter, H., Oost, J. van der, & White, M. F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiology Reviews*, 39(3), Article 3. <https://doi.org/10.1093/femsre/fuv023>

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., Sternberg, S. H., Joung, J. K., Yildiz, A., & Doudna, J. A. (2017). Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature*, 550(7676), Article 7676. <https://doi.org/10.1038/nature24268>

Chylinski, K., Le Rhun, A., & Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biology*, 10(5), 726–737. <https://doi.org/10.4161/rna.24321>

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121), Article 6121. <https://doi.org/10.1126/science.1231143>

Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J., Joung, J., & Zhang, F. (2017). RNA editing with CRISPR-Cas13. *Science*, eaaq0180. <https://doi.org/10.1126/science.aaq0180>

Davis, K. M., Pattanayak, V., Thompson, D. B., Zuris, J. A., & Liu, D. R. (2015). Small molecule–triggered Cas9 protein with improved genome-editing specificity. *Nature Chemical Biology*, 11(5), Article 5. <https://doi.org/10.1038/nchembio.1793>

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471(7340), 602–607. <https://doi.org/10.1038/nature09886>

Dong, D., Ren, K., Qiu, X., Zheng, J., Guo, M., Guan, X., Liu, H., Li, N., Zhang, B., Yang, D., Ma, C., Wang, S., Wu, D., Ma, Y., Fan, S., Wang, J., Gao, N., & Huang, Z. (2016). The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature*, 532(7600), Article 7600. <https://doi.org/10.1038/nature17944>

Dooley, S. K., Baken, E. K., Moss, W. N., Howe, A., & Young, J. K. (2021). Identification and Evolution of Cas9 tracrRNAs. *The CRISPR Journal*, 4(3), 438–447. <https://doi.org/10.1089/crispr.2020.0093>

- Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, 346(6213), 1258096. <https://doi.org/10.1126/science.1258096>
- East-Seletsky, A., O'Connell, M. R., Burstein, D., Knott, G. J., & Doudna, J. A. (2017). RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes. *Molecular Cell*, 66(3), 373-383.e3. <https://doi.org/10.1016/j.molcel.2017.04.008>
- East-Seletsky, A., O'Connell, M. R., Knight, S. C., Burstein, D., Cate, J. H. D., Tjian, R., & Doudna, J. A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature*, 538(7624), Article 7624. <https://doi.org/10.1038/nature19802>
- Endo, A., Masafumi, M., Kaya, H., & Toki, S. (2016). Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Scientific Reports*, 6. <https://doi.org/10.1038/srep38169>
- Engler, C., Kandzia, R., & Marillonnet, S. (2008). A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLOS ONE*, 3(11), e3647. <https://doi.org/10.1371/journal.pone.0003647>
- Findlay, H. E., Harris, N. J., & Booth, P. J. (2016). In vitro synthesis of a Major Facilitator Transporter for specific active transport across Droplet Interface Bilayers. *Scientific Reports*, 6, 39349. <https://doi.org/10.1038/srep39349>
- Fogeron, M.-L., Lecoq, L., Cole, L., Harbers, M., & Böckmann, A. (2021). Easy Synthesis of Complex Biomolecular Assemblies: Wheat Germ Cell-Free Protein Expression in Structural Biology. *Frontiers in Molecular Biosciences*, 8, 639587. <https://doi.org/10.3389/fmolb.2021.639587>
- Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K. S., Lécrivain, A.-L., Bzdrenga, J., Koonin, E. V., & Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Research*, 42(4), 2577–2590. <https://doi.org/10.1093/nar/gkt1074>
- Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., & Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, 532(7600), Article 7600. <https://doi.org/10.1038/nature17945>
- Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R., Ho, T. W., Kattamis, A., Kernysky, A., Lekstrom-Himes, J., Li, A. M., Locatelli, F., Mapara, M. Y., de Montalembert, M., Rondelli, D., ... Corbacioglu, S. (2021). CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia. *The New England Journal of Medicine*, 384(3), 252–260. <https://doi.org/10.1056/NEJMoa2031054>
- Fuller, S., Takahashi, M., & Hurrell, J. (2001). Short Protocols in Molecular Biology. *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]*, Chapter 11, Unit11.11. <https://doi.org/10.1002/0471142727.mb1111s37>
- Gaillochet, C., Peña Fernández, A., Goossens, V., D'Halluin, K., Drozdzecki, A., Shafie, M., Van Duyse, J., Van Isterdael, G., Gonzalez, C., Vermeersch, M., De Saeger, J., Develtere, W., Audenaert, D., De Vleeschauwer, D., Meulewaeter, F., & Jacobs, T. B. (2023). Systematic optimization of Cas12a base editors in wheat and maize using the ITER platform. *Genome Biology*, 24(1), 6. <https://doi.org/10.1186/s13059-022-02836-2>
- Gao, L., Cox, D. B. T., Yan, W. X., Manteiga, J. C., Schneider, M. W., Yamano, T., Nishimasu, H., Nureki, O., Crosetto, N., & Zhang, F. (2017). Engineered Cpf1 variants with altered PAM specificities. *Nature Biotechnology*, 35(8), Article 8. <https://doi.org/10.1038/nbt.3900>
- Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.

Proceedings of the National Academy of Sciences, 109(39), Article 39. <https://doi.org/10.1073/pnas.1208507109>

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., & Liu, D. R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, 551(7681), Article 7681. <https://doi.org/10.1038/nature24644>

Gesner, E. M., Schellenberg, M. J., Garside, E. L., George, M. M., & Macmillan, A. M. (2011). Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nature Structural & Molecular Biology*, 18(6), 688–692. <https://doi.org/10.1038/nsmb.2042>

Ghosh, D., Venkataramani, P., Nandi, S., & Bhattacharjee, S. (2019). CRISPR–Cas9 a boon or bane: The bumpy road ahead to cancer therapeutics. *Cancer Cell International*, 19(1), 12. <https://doi.org/10.1186/s12935-019-0726-0>

Gillmore, J. D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M. L., Seitzer, J., O’Connell, D., Walsh, K. R., Wood, K., Phillips, J., Xu, Y., Amaral, A., Boyd, A. P., Cehelsky, J. E., McKee, M. D., Schiermeier, A., Harari, O., Murphy, A., ... Lebowitz, D. (2021). CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. *The New England Journal of Medicine*, 385(6), 493–502. <https://doi.org/10.1056/NEJMoa2107454>

Gleditsch, D., Pausch, P., Müller-Esparza, H., Özcan, A., Guo, X., Bange, G., & Randau, L. (2019). PAM identification by CRISPR-Cas effector complexes: Diversified mechanisms and structures. *RNA Biology*, 16(4), 504–517. <https://doi.org/10.1080/15476286.2018.1504546>

Goossens, R., van den Boogaard, M. L., Lemmers, R. J. L. F., Balog, J., van der Vliet, P. J., Willemsen, I. M., Schouten, J., Maggio, I., van der Stoep, N., Hoebe, R. C., Tapscott, S. J., Geijsen, N., Gonçalves, M. A. F. V., Sacconi, S., Tawil, R., & van der Maarel, S. M. (2019). Intronic SMCHD1 variants in FSHD: Testing the potential for CRISPR-Cas9 genome editing. *Journal of Medical Genetics*, 56(12), 828–837. <https://doi.org/10.1136/jmedgenet-2019-106402>

Gregorio, N. E., Levine, M. Z., & Oza, J. P. (2019). A User’s Guide to Cell-Free Protein Synthesis. *Methods and Protocols*, 2(1), Article 1. <https://doi.org/10.3390/mps2010024>

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., Zeina, C. M., Gao, X., Rees, H. A., Lin, Z., & Liu, D. R. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, 556(7699), Article 7699. <https://doi.org/10.1038/nature26155>

Iliakis, G., Wang, H., Perrault, A. R., Boecker, W., Rosidi, B., Windhofer, F., Wu, W., Guan, J., Terzoudi, G., & Pantelias, G. (2004). Mechanisms of DNA double strand break repair and chromosome aberration formation. *Cytogenetic and Genome Research*, 104(1–4), 14–20. <https://doi.org/10.1159/000077461>

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, 169(12), Article 12.

Jackson, S. A., McKenzie, R. E., Fagerlund, R. D., Kieper, S. N., Fineran, P. C., & Brouns, S. J. J. (2017). CRISPR-Cas: Adapting to change. *Science*, 356(6333), Article 6333. <https://doi.org/10.1126/science.aal5056>

Jacobsen, T., Ttofali, F., Liao, C., Manchalu, S., Gray, B. N., & Beisel, C. L. (2020). Characterization of Cas12a nucleases reveals diverse PAM profiles between closely-related orthologs. *Nucleic Acids Research*, 48(10), Article 10. <https://doi.org/10.1093/nar/gkaa272>

Jiang, F., Taylor, D. W., Chen, J. S., Kornfeld, J. E., Zhou, K., Thompson, A. J., Nogales, E., & Doudna, J. A. (2016). Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science*, 351(6275), Article 6275. <https://doi.org/10.1126/science.aad8282>

- Jiang, F., Zhou, K., Ma, L., Gressel, S., & Doudna, J. A. (2015). STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science (New York, N.Y.)*, 348(6242), 1477–1481. <https://doi.org/10.1126/science.aab1452>
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(6096), 816–821. <https://doi.org/10.1126/science.1225829>
- Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E., & Doudna, J. A. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science (New York, N.Y.)*, 343(6176), 1247997. <https://doi.org/10.1126/science.1247997>
- Jing, Z., Zhang, N., Ding, L., Wang, X., Hua, Y., Jiang, M., & Wu, S. X. (2018). Safety and activity of programmed cell death-1 gene knockout engineered t cells in patients with previously treated advanced esophageal squamous cell carcinoma: An open-label, single-arm phase I study. *Journal of Clinical Oncology*, 36(15_suppl), 3054–3054. https://doi.org/10.1200/JCO.2018.36.15_suppl.3054
- Jones, S. K., Hawkins, J. A., Johnson, N. V., Jung, C., Hu, K., Rybarski, J. R., Chen, J. S., Doudna, J. A., Press, W. H., & Finkelstein, I. J. (2019). Massively parallel kinetic profiling of natural and engineered CRISPR nucleases. *BioRxiv*, 696393. <https://doi.org/10.1101/696393>
- Karvelis, T., Bigelyte, G., Young, J. K., Hou, Z., Zedaveinyte, R., Budre, K., Paulraj, S., Djukanovic, V., Gasior, S., Silanskas, A., Venclovas, Č., & Siksnys, V. (2020). PAM recognition by miniature CRISPR–Cas12f nucleases triggers programmable double-stranded DNA target cleavage. *Nucleic Acids Research*, 48(9), 5016–5023. <https://doi.org/10.1093/nar/gkaa208>
- Karvelis, T., Gasiunas, G., & Siksnys, V. (2017). Methods for decoding Cas9 protospacer adjacent motif (PAM) sequences: A brief overview. *Methods (San Diego, Calif.)*, 121–122, 3–8. <https://doi.org/10.1016/j.ymeth.2017.03.006>
- Kazlauskienė, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., & Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science*, eaao0100. <https://doi.org/10.1126/science.aao0100>
- Kazlauskienė, M., Tamulaitis, G., Kostiuk, G., Venclovas, Č., & Siksnys, V. (2016). Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. *Molecular Cell*, 62(2), 295–306. <https://doi.org/10.1016/j.molcel.2016.03.024>
- Kim, H.-C., Kim, T.-W., & Kim, D.-M. (2011). Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source. *Process Biochemistry*, 46(6), 1366–1369. <https://doi.org/10.1016/j.procbio.2011.03.008>
- Kim, S., Koo, T., Jee, H.-G., Cho, H.-Y., Lee, G., Lim, D.-G., Shin, H. S., & Kim, J.-S. (2018). CRISPR RNAs trigger innate immune responses in human cells. *Genome Research*, 28(3), 367–373. <https://doi.org/10.1101/gr.231936.117>
- Kim, T.-W., Keum, J.-W., Oh, I.-S., Choi, C.-Y., Park, C.-G., & Kim, D.-M. (2006). Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system. *Journal of Biotechnology*, 126(4), 554–561. <https://doi.org/10.1016/j.jbiotec.2006.05.014>
- Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., Gonzales, A. P. W., Li, Z., Peterson, R. T., Yeh, J.-R. J., Aryee, M. J., & Joung, J. K. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, 523(7561), Article 7561. <https://doi.org/10.1038/nature14592>

- Köhler, C., Mayer, C., Gröbner, P., & Piendl, W. (1996). Use of T7 RNA Polymerase in an Optimized Escherichia coli Coupled in vitro Transcription-Translation System. *European Journal of Biochemistry*, 236(1), 234–239. <https://doi.org/10.1111/j.1432-1033.1996.00234.x>
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., & Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, 533, 420–424. <https://doi.org/10.1038/nature17946>
- Koonin, E. V., Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology*, 37, 67–78. <https://doi.org/10.1016/j.mib.2017.05.008>
- Kuruma, Y., & Ueda, T. (2015). The PURE system for the cell-free synthesis of membrane proteins. *Nature Protocols*, 10(9), Article 9. <https://doi.org/10.1038/nprot.2015.082>
- Kwon, Y.-C., & Jewett, M. C. (2015). High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific Reports*, 5(1), Article 1. <https://doi.org/10.1038/srep08663>
- Laalami, S., Zig, L., & Putzer, H. (2014). Initiation of mRNA decay in bacteria. *Cellular and Molecular Life Sciences*, 71(10), 1799–1828. <https://doi.org/10.1007/s00018-013-1472-4>
- Lee, J. K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y.-H., Lee, K., Jung, I., Kim, D., Kim, S., & Kim, J.-S. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. *Nature Communications*, 9(1), Article 1. <https://doi.org/10.1038/s41467-018-05477-x>
- Lee, K., Mackley, V. A., Rao, A., Chong, A. T., Dewitt, M. A., Corn, J. E., & Murthy, N. (2017). Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. *ELife*, 6, e25312. <https://doi.org/10.7554/eLife.25312>
- Leenay, R. T., & Beisel, C. L. (2017). Deciphering, Communicating, and Engineering the CRISPR PAM. *Journal of Molecular Biology*, 429(2), 177–191. <https://doi.org/10.1016/j.jmb.2016.11.024>
- Leenay, R. T., Maksimchuk, K. R., Slotkowski, R. A., Agrawal, R. N., Gomaa, A. A., Briner, A. E., Barrangou, R., & Beisel, C. L. (2016). Identifying and Visualizing Functional PAM Diversity across CRISPR-Cas Systems. *Molecular Cell*, 62(1), Article 1. <https://doi.org/10.1016/j.molcel.2016.02.031>
- Li, S.-Y., Cheng, Q.-X., Wang, J.-M., Li, X.-Y., Zhang, Z.-L., Gao, S., Cao, R.-B., Zhao, G.-P., & Wang, J. (2018). CRISPR-Cas12a-assisted nucleic acid detection. *Cell Discovery*, 4(1), 1–4. <https://doi.org/10.1038/s41421-018-0028-z>
- Lim, J. M., & Kim, H. H. (2022). Basic Principles and Clinical Applications of CRISPR-Based Genome Editing. *Yonsei Medical Journal*, 63(2), 105–113. <https://doi.org/10.3349/ymj.2022.63.2.105>
- Liu, H., & Naismith, J. H. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology*, 8(1), 91. <https://doi.org/10.1186/1472-6750-8-91>
- Liu, T. Y., & Doudna, J. A. (2020). Chemistry of Class 1 CRISPR-Cas effectors: Binding, editing, and regulation. *The Journal of Biological Chemistry*, 295(42), 14473–14487. <https://doi.org/10.1074/jbc.REV120.007034>
- Loewenstein, P. M., Song, C.-Z., & Green, M. (2007). The Use of In Vitro Transcription to Probe Regulatory Functions of Viral Protein Domains. In W. S. M. Wold & A. E. Tollefson (Eds.), *Adenovirus Methods and Protocols* (Vol. 131, pp. 15–31). Humana Press. https://doi.org/10.1007/978-1-59745-277-9_2

Loureiro, A., & da Silva, G. J. (2019). CRISPR-Cas: Converting A Bacterial Defence Mechanism into A State-of-the-Art Genetic Manipulation Tool. *Antibiotics*, 8(1), Article 1. <https://doi.org/10.3390/antibiotics8010018>

Lu, Y., Xue, J., Deng, T., Zhou, X., Yu, K., Deng, L., Huang, M., Yi, X., Liang, M., Wang, Y., Shen, H., Tong, R., Wang, W., Li, L., Song, J., Li, J., Su, X., Ding, Z., Gong, Y., ... Mok, T. (2020). Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nature Medicine*, 26(5), 732–740. <https://doi.org/10.1038/s41591-020-0840-5>

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., ... Koonin, E. V. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews. Microbiology*, 13(11), Article 11. <https://doi.org/10.1038/nrmicro3569>

Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., Charpentier, E., Cheng, D., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Scott, D., Shah, S. A., Siksnys, V., Terns, M. P., Venclovas, Č., White, M. F., Yakunin, A. F., ... Koonin, E. V. (2020). Evolutionary classification of CRISPR–Cas systems: A burst of class 2 and derived variants. *Nature Reviews Microbiology*, 18(2), 67–83. <https://doi.org/10.1038/s41579-019-0299-x>

Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2013). The basic building blocks and evolution of CRISPR-CAS systems. *Biochemical Society Transactions*, 41(6), 1392–1400. <https://doi.org/10.1042/BST20130038>

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., & Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*, 31(9), Article 9. <https://doi.org/10.1038/nbt.2675>

Mandrich, L. (2015). The “Evolution” of Mutagenesis. *Cloning & Transgenesis*, 04(03). <https://doi.org/10.4172/2168-9849.1000e121>

Marraffini, L. A., & Sontheimer, E. J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews. Genetics*, 11(3), Article 3. <https://doi.org/10.1038/nrg2749>

Marshall, R., Maxwell, C. S., Collins, S. P., Jacobsen, T., Luo, M. L., Begemann, M. B., Gray, B. N., January, E., Singer, A., He, Y., Beisel, C. L., & Noireaux, V. (2018). Rapid and Scalable Characterization of CRISPR Technologies Using an E. coli Cell-Free Transcription-Translation System. *Molecular Cell*, 69(1), Article 1. <https://doi.org/10.1016/j.molcel.2017.12.007>

Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R., & Ploegh, H. L. (2016). Corrigendum: Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature Biotechnology*, 34(2), 210. <https://doi.org/10.1038/nbt0216-210c>

McGinn, J., & Marraffini, L. A. (2016). CRISPR-Cas Systems Optimize Their Immune Response by Specifying the Site of Spacer Integration. *Molecular Cell*, 64(3), 616–623. <https://doi.org/10.1016/j.molcel.2016.08.038>

Miyaoka, Y., Berman, J. R., Cooper, S. B., Mayerl, S. J., Chan, A. H., Zhang, B., Karlin-Neumann, G. A., & Conklin, B. R. (2016). Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Scientific Reports*, 6(1), Article 1. <https://doi.org/10.1038/srep23549>

- Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, 60(2), Article 2. <https://doi.org/10.1007/s00239-004-0046-3>
- Moreno, A. M., Fu, X., Zhu, J., Katrekar, D., Shih, Y.-R. V., Marlett, J., Cabotaje, J., Tat, J., Naughton, J., Lisowski, L., Varghese, S., Zhang, K., & Mali, P. (2018). In Situ Gene Therapy via AAV-CRISPR-Cas9-Mediated Targeted Gene Regulation. *Molecular Therapy*, 26(7), 1818–1827. <https://doi.org/10.1016/j.ymthe.2018.04.017>
- Mosterd, C., Rousseau, G. M., & Moineau, S. (2021). A short overview of the CRISPR-Cas adaptation stage. *Canadian Journal of Microbiology*, 67(1), 1–12. <https://doi.org/10.1139/cjm-2020-0212>
- Nelson, C. E., & Gersbach, C. A. (2016). Engineering Delivery Vehicles for Genome Editing. *Annual Review of Chemical and Biomolecular Engineering*, 7, 637–662. <https://doi.org/10.1146/annurev-chembioeng-080615-034711>
- Nierzwicki, Ł., Arantes, P. R., Saha, A., & Palermo, G. (2021). Establishing the allosteric mechanism in CRISPR-Cas9. *WIREs Computational Molecular Science*, 11(3), e1503. <https://doi.org/10.1002/wcms.1503>
- Niewoehner, O., Garcia-Doval, C., Rostøl, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L. A., & Jinek, M. (2017). Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature*, 548(7669), 543–548. <https://doi.org/10.1038/nature23467>
- Niewoehner, O., Jinek, M., & Doudna, J. A. (2014). Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. *Nucleic Acids Research*, 42(2), 1341–1353. <https://doi.org/10.1093/nar/gkt922>
- Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K. Y., Shimatani, Z., & Kondo, A. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science (New York, N.Y.)*, 353(6305), Article 6305. <https://doi.org/10.1126/science.aaf8729>
- Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., Ishitani, R., Zhang, F., & Nureki, O. (2014). Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*, 156(5), Article 5. <https://doi.org/10.1016/j.cell.2014.02.001>
- Nuñez, J. K., Harrington, L. B., Kranzusch, P. J., Engelman, A. N., & Doudna, J. A. (2015). Foreign DNA capture during CRISPR–Cas adaptive immunity. *Nature*, advance online publication. <https://doi.org/10.1038/nature15760>
- Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., & Doudna, J. A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nature Structural & Molecular Biology*, 21(6), 528–534. <https://doi.org/10.1038/nsmb.2820>
- Palermo, G., Chen, J. S., Ricci, C. G., Rivalta, I., Jinek, M., Batista, V. S., Doudna, J. A., & McCammon, J. A. (2018). Key role of the REC lobe during CRISPR-Cas9 activation by “sensing”, “regulating”, and “locking” the catalytic HNH domain. *Quarterly Reviews of Biophysics*, 51, e91. <https://doi.org/10.1017/S0033583518000070>
- Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., Ferrante, T., Ma, D., Donghia, N., Fan, M., Daringer, N. M., Bosch, I., Dudley, D. M., O'Connor, D. H., Gehrke, L., & Collins, J. J. (2016). Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell*, 165(5), 1255–1266. <https://doi.org/10.1016/j.cell.2016.04.059>
- Pardee, K., Slomovic, S., Nguyen, P. Q., Lee, J. W., Donghia, N., Burrill, D., Ferrante, T., McSorley, F. R., Furuta, Y., Vernet, A., Lewandowski, M., Boddy, C. N., Joshi, N. S., & Collins, J.

- J. (2016). Portable, On-Demand Biomolecular Manufacturing. *Cell*, 167(1), 248-259.e12. <https://doi.org/10.1016/j.cell.2016.09.013>
- Paul, B., & Montoya, G. (2020). CRISPR-Cas12a: Functional overview and applications. *Biomedical Journal*, 43(1), 8–17. <https://doi.org/10.1016/j.bj.2019.10.005>
- Pinilla-Redondo, R., Mayo-Muñoz, D., Russel, J., Garrett, R. A., Randau, L., Sørensen, S. J., & Shah, S. A. (2020). Type IV CRISPR–Cas systems are highly diverse and involved in competition between plasmids. *Nucleic Acids Research*, 48(4), 2000–2012. <https://doi.org/10.1093/nar/gkz1197>
- Pougach, K., & Severinov, K. (2012). Use of Semi-quantitative Northern Blot Analysis to Determine Relative Quantities of Bacterial CRISPR Transcripts. In K. C. Keiler (Ed.), *Bacterial Regulatory RNA* (Vol. 905, pp. 73–86). Humana Press. https://doi.org/10.1007/978-1-61779-949-5_6
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), Article 11. <https://doi.org/10.1038/nprot.2013.143>
- Ranji Charna, A., Des Soye, B. J., Ntai, I., Kelleher, N. L., & Jewett, M. C. (2022). An efficient cell-free protein synthesis platform for producing proteins with pyrrolysine-based noncanonical amino acids. *Biotechnology Journal*, 17(9), 2200096. <https://doi.org/10.1002/biot.202200096>
- Rath, D., Amlinger, L., Rath, A., & Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie*, 117, 119–128. <https://doi.org/10.1016/j.biochi.2015.03.025>
- Redding, S., Sternberg, S. H., Marshall, M., Gibb, B., Bhat, P., Guegler, C. K., Wiedenheft, B., Doudna, J. A., & Greene, E. C. (2015). Surveillance and Processing of Foreign DNA by the *Escherichia coli* CRISPR-Cas System. *Cell*, 163(4), Article 4. <https://doi.org/10.1016/j.cell.2015.10.003>
- Rees, H. A., & Liu, D. R. (2018). Base editing: Precision chemistry on the genome and transcriptome of living cells. *Nature Reviews. Genetics*, 19(12), Article 12. <https://doi.org/10.1038/s41576-018-0059-1>
- Robert, F., Barbeau, M., Éthier, S., Dostie, J., & Pelletier, J. (2015). Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Medicine*, 7(1), 93. <https://doi.org/10.1186/s13073-015-0215-6>
- Rodriguez, E. L., Poddar, S., Iftekhar, S., Suh, K., Woolfork, A. G., Ovbude, S., Pekarek, A., Walters, M., Lott, S., & Hage, D. S. (2020). Affinity Chromatography: A Review of Trends and Developments over the Past 50 Years. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences*, 1157, 122332. <https://doi.org/10.1016/j.jchromb.2020.122332>
- Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilstein-Edmands, V., Cannone, G., Graham, S., Robinson, C. V., Spagnolo, L., & White, M. F. (2013). Structure of the CRISPR Interference Complex CSM Reveals Key Similarities with Cascade. *Molecular Cell*, 52(1), 124–134. <https://doi.org/10.1016/j.molcel.2013.08.020>
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning: A laboratory manual. *Molecular Cloning: A Laboratory Manual*, Ed. 2. <https://www.cabdirect.org/cabdirect/abstract/19901616061>
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), 5463–5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Scheele, G., & Blackburn, P. (1979). Role of mammalian RNase inhibitor in cell-free protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 76(10), 4898–4902. <https://doi.org/10.1073/pnas.76.10.4898>

- Schellenberger, V., Wang, C., Geething, N. C., Spink, B. J., Campbell, A., To, W., Scholle, M. D., Yin, Y., Yao, Y., Bogin, O., Cleland, J. L., Silverman, J., & Stemmer, W. P. C. (2009). A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. *Nature Biotechnology*, 27(12), Article 12. <https://doi.org/10.1038/nbt.1588>
- Scholefield, J., & Harrison, P. T. (2021). Prime editing – an update on the field. *Gene Therapy*, 28(7), Article 7. <https://doi.org/10.1038/s41434-021-00263-9>
- Shah, S. A., Erdmann, S., Mojica, F. J. M., & Garrett, R. A. (2013). Protospacer recognition motifs: Mixed identities and functional diversity. *RNA Biology*, 10(5), 891–899. <https://doi.org/10.4161/rna.23764>
- Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., & Skarnes, W. C. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature Methods*, 11(4), Article 4. <https://doi.org/10.1038/nmeth.2857>
- Shiimori, M., Garrett, S. C., Graveley, B. R., & Terns, M. P. (2018). Cas4 nucleases define the PAM, length, and orientation of DNA fragments integrated at CRISPR loci. *Molecular Cell*, 70(5), 814–824.e6. <https://doi.org/10.1016/j.molcel.2018.05.002>
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., & Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nature Biotechnology*, 19(8), 751–755. <https://doi.org/10.1038/90802>
- Shin, J., & Noireaux, V. (2010). Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. *Journal of Biological Engineering*, 4, 8. <https://doi.org/10.1186/1754-1611-4-8>
- Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., & Koonin, E. V. (2015). Discovery and functional characterization of diverse Class 2 CRISPR-Cas systems. *Molecular Cell*, 60(3), 385–397. <https://doi.org/10.1016/j.molcel.2015.10.008>
- Silas, S., Mohr, G., Sidote, D. J., Markham, L. M., Sanchez-Amat, A., Bhaya, D., Lambowitz, A. M., & Fire, A. Z. (2016). Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science (New York, N.Y.)*, 351(6276), aad4234. <https://doi.org/10.1126/science.aad4234>
- Silverman, A. D., Karim, A. S., & Jewett, M. C. (2020). Cell-free gene expression: An expanded repertoire of applications. *Nature Reviews Genetics*, 21(3), Article 3. <https://doi.org/10.1038/s41576-019-0186-3>
- Sinkunas, T., Gasiunas, G., Waghmare, S. P., Dickman, M. J., Barrangou, R., Horvath, P., & Siksnys, V. (2013). In vitro reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *The EMBO Journal*, 32(3), Article 3. <https://doi.org/10.1038/emboj.2012.352>
- Sitaraman, K., Esposito, D., Klarmann, G., Le Grice, S. F., Hartley, J. L., & Chatterjee, D. K. (2004). A novel cell-free protein synthesis system. *Journal of Biotechnology*, 110(3), 257–263. <https://doi.org/10.1016/j.jbiotec.2004.02.014>
- Smolskaya, S., Logashina, Y. A., & Andreev, Y. A. (2020). Escherichia coli Extract-Based Cell-Free Expression System as an Alternative for Difficult-to-Obtain Protein Biosynthesis. *International Journal of Molecular Sciences*, 21(3), 928. <https://doi.org/10.3390/ijms21030928>
- Soltani, M., & Bundy, B. C. (2022). Streamlining cell-free protein synthesis biosensors for use in human fluids: In situ RNase inhibitor production during extract preparation. *Biochemical Engineering Journal*, 177, 108158. <https://doi.org/10.1016/j.bej.2021.108158>
- Staals, R. H. J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J. J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P. J.,

- Doudna, J. A., Heck, A. J. R., Yonekura, K., van der Oost, J., & Shinkai, A. (2013). Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus*. *Molecular Cell*, 52(1), Article 1. <https://doi.org/10.1016/j.molcel.2013.09.013>
- Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., Mangan, P. A., Kulikovskaya, I., Gupta, M., Chen, F., Tian, L., Gonzalez, V. E., Xu, J., Jung, I., Melenhorst, J. J., Plesa, G., Shea, J., Matlawski, T., Cervini, A., ... June, C. H. (2020). CRISPR-engineered T cells in patients with refractory cancer. *Science*, 367(6481), Article 6481. <https://doi.org/10.1126/science.aba7365>
- Stella, S., Alcón, P., & Montoya, G. (2017). Structure of the Cpf1 endonuclease R-loop complex after target DNA cleavage. *Nature*, 546, 559–563. <https://doi.org/10.1038/nature22398>
- Stella, S., Mesa, P., Thomsen, J., Paul, B., Alcón, P., Jensen, S. B., Saligram, B., Moses, M. E., Hatzakis, N. S., & Montoya, G. (2018). Conformational Activation Promotes CRISPR-Cas12a Catalysis and Resetting of the Endonuclease Activity. *Cell*, 175(7), Article 7. <https://doi.org/10.1016/j.cell.2018.10.045>
- Sternberg, S. H., LaFrance, B., Kaplan, M., & Doudna, J. A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*, 527(7576), Article 7576. <https://doi.org/10.1038/nature15544>
- Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, 507(7490), Article 7490. <https://doi.org/10.1038/nature13011>
- Swarts, D. C., van der Oost, J., & Jinek, M. (2017). Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. *Molecular Cell*, 66(2), Article 2. <https://doi.org/10.1016/j.molcel.2017.03.016>
- Tabatabaei Yazdi, S. M. H., Yuan, Y., Ma, J., Zhao, H., & Milenkovic, O. (2015). A Rewritable, Random-Access DNA-Based Storage System. *Scientific Reports*, 5. <https://doi.org/10.1038/srep14138>
- Tabor, S. (2001). Expression using the T7 RNA polymerase/promoter system. *Current Protocols in Molecular Biology*, Chapter 16, Unit16.2. <https://doi.org/10.1002/0471142727.mb1602s11>
- Tang, X., Lowder, L. G., Zhang, T., Malzahn, A. A., Zheng, X., Voytas, D. F., Zhong, Z., Chen, Y., Ren, Q., Li, Q., Kirkland, E. R., Zhang, Y., & Qi, Y. (2017). A CRISPR–Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nature Plants*, 3(3), Article 3. <https://doi.org/10.1038/nplants.2017.18>
- Taylor, H. N., Warner, E. E., Armbrust, M. J., Crowley, V. M., Olsen, K. J., & Jackson, R. N. (2019). Structural basis of Type IV CRISPR RNA biogenesis by a Cas6 endonuclease. *RNA Biology*, 16(10), 1438–1447. <https://doi.org/10.1080/15476286.2019.1634965>
- Towbin, H., Staehelin, T., & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America*, 76(9), 4350–4354. <https://doi.org/10.1073/pnas.76.9.4350>
- van Nies, P., Westerlaken, I., Blanken, D., Salas, M., Mencía, M., & Danelon, C. (2018). Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nature Communications*, 9(1), Article 1. <https://doi.org/10.1038/s41467-018-03926-1>
- Wright, A. V., Liu, J.-J., Knott, G. J., Doxzen, K. W., Nogales, E., & Doudna, J. A. (2017). Structures of the CRISPR genome integration complex. *Science*, eaao0679. <https://doi.org/10.1126/science.aao0679>

- Wu, W.-H., Tsai, Y.-T., Justus, S., Lee, T.-T., Zhang, L., Lin, C.-S., Bassuk, A. G., Mahajan, V. B., & Tsang, S. H. (2016). CRISPR Repair Reveals Causative Mutation in a Preclinical Model of Retinitis Pigmentosa. *Molecular Therapy*, 24(8), 1388–1394. <https://doi.org/10.1038/mt.2016.107>
- Xiao, Y., Luo, M., Hayes, R. P., Kim, J., Ng, S., Ding, F., Liao, M., & Ke, A. (2017). Structure Basis for Directional R-loop Formation and Substrate Handover Mechanisms in Type I CRISPR-Cas System. *Cell*, 170(1), Article 1. <https://doi.org/10.1016/j.cell.2017.06.012>
- Xiao, Y., Ng, S., Nam, K. H., & Ke, A. (2017). How Type II CRISPR-Cas establish immunity through Cas1-Cas2 mediated spacer integration. *Nature*, 550(7674), 137–141. <https://doi.org/10.1038/nature24020>
- Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., Fedorova, I., Nakane, T., Makarova, K. S., Koonin, E. V., Ishitani, R., Zhang, F., & Nureki, O. (2016). Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell*, 165(4), Article 4. <https://doi.org/10.1016/j.cell.2016.04.003>
- Yang, H.-C., & Chen, P.-J. (2018). The potential and challenges of CRISPR-Cas in eradication of hepatitis B virus covalently closed circular DNA. *Virus Research*, 244, 304–310. <https://doi.org/10.1016/j.virusres.2017.06.010>
- Yang, Y., Xu, J., Ge, S., & Lai, L. (2021). CRISPR/Cas: Advances, Limitations, and Applications for Precision Cancer Research. *Frontiers in Medicine*, 8. <https://www.frontiersin.org/articles/10.3389/fmed.2021.649896>
- Yokogawa, T., Kitamura, Y., Nakamura, D., Ohno, S., & Nishikawa, K. (2010). Optimization of the hybridization-based method for purification of thermostable tRNAs in the presence of tetraalkylammonium salts. *Nucleic Acids Research*, 38(6), e89. <https://doi.org/10.1093/nar/gkp1182>
- Zawada, J. F., Yin, G., Steiner, A. R., Yang, J., Naresh, A., Roy, S. M., Gold, D. S., Heinsohn, H. G., & Murray, C. J. (2011). Microscale to manufacturing scale-up of cell-free cytokine production—A new approach for shortening protein production development timelines. *Biotechnology and Bioengineering*, 108(7), 1570–1578. <https://doi.org/10.1002/bit.23103>
- Zemella, A., Thoring, L., Hoffmeister, C., & Kubick, S. (2015). Cell-Free Protein Synthesis: Pros and Cons of Prokaryotic and Eukaryotic Systems. *ChemBioChem*, 16(17), 2420–2431. <https://doi.org/10.1002/cbic.201500340>
- Zhang, Y., Zhao, G., Ahmed, F. Y. H., Yi, T., Hu, S., Cai, T., & Liao, Q. (2020). In silico Method in CRISPR/Cas System: An Expedite and Powerful Booster. *Frontiers in Oncology*, 10. <https://www.frontiersin.org/articles/10.3389/fonc.2020.584404>
- Zischewski, J., Fischer, R., & Bortesi, L. (2017). Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnology Advances*, 35(1), 95–104. <https://doi.org/10.1016/j.biotechadv.2016.12.003>
- Zubay, G. (1973). In Vitro Synthesis of Protein in Microbial Systems. *Annual Review of Genetics*, 7(1), 267–287. <https://doi.org/10.1146/annurev.ge.07.120173.001411>