university of groningen / faculty of science and engineering

High-Throughput Proteomics: Screening Methods for CRISPR/Cas9 DNA binding affinity

July 2024

Bachelor's Thesis for Life Science and Technology

Mihai Frăcea[1](S4653572)

1. Bsc Life Sciences and Technology, University of Groningen, 9712 CP Groningen, The Netherlands, m.s.fracea@student.rug.nl

Supervised by Maximilian Fürst

Summary

The RNA-guided CRISPR/Cas9 nuclease is a revolutionary gene-editing tool which dramatically increased the extent to which genetic editing is possible, bringing unprecedented levels of precision and cost efficiency. CRISPR/Cas9 allows for gene editing and expression control and has applications in plant and animal biotechnology, climate and in advancing innovative medical therapies. To address the limitations related to Cas9 specificity and to improve the precision in gene editing, many techniques rely on developing large numbers of Cas9 or guide RNA (gRNA) mutant variants *in vitro* and testing their capabilities. However, probing a large number of variants can often be challenging due to limited throughput, low precision, cost or other underlying factors. Here, we identify the main high-throughput (HT) *in vitro* DNA-Cas9 binding assays and thoroughly evaluate and compare their capabilities, advantages and limitations. Generally, HITS-FLIP and versions derived from it were found to be versatile and more accessible methods that can perform high-throughput assays without overlooking precision, but we suggest each method to be taken under account depending on the experimental setup, as there is no fit-for-all solution.

Table of Contents

Summary **2**
Table of Contents **3**
Introduction **4**
Cas9 is an RNA-guided Endonuclease Able to Cut Double-Stranded DNA 4
CRISPR Cas9 Recognizes and Cleaves Target DNA 5
Overcoming Limitations of CRISPR/Cas9 6
Main **7**
Cas9/gRNA-DNA Binding Affinity Assays 7
EMSA 7
ChIP-seq 9
HITS-FLIP 9
DNA Beacon (FRET) 11
Additional Binding Assays 12
Optical Tweezers 12
Magnetic Tweezers 13
Atomic Force Microscopy 13
Surface Plasmon Resonance & Biolayer Interferometry 13
Other Binding Assays 14
Discussion & Comparison **14**
EMSA 17
ChIP-seq 17
HITS-FLIP 17
Possible Caveat in Calculating Kd in Protein-DNA Binding Assays 18
Improvement of gRNA Design 18
Conclusion **18**
References **20**
Afterword 29

Introduction

Recent advancements in genetic engineering have paved the way for a new era in studying and editing genomes and biological systems as we know them today. Among the multitude of new tools in the "toolbox" of genetic engineering, CRISPR (clustered regularly interspaced short palindromic repeats) and Cas9 (CRISPR-associated protein 9) have dramatically increased the extent to which genetic editing is possible, bringing this area to unprecedented levels of precision and cost efficiency.

The fields in which CRISPR can be applied vary greatly: from agriculture and plant biology, all the way to investigating gene function or gene-editing approaches in conditions such as cancer, cardiovascular and neurodegenerative diseases (Serajian et al, 2021; Zhu et al, 2020; Luthra et al, 2021; Li et al 2021; Chen et al, 2019; Platt et al, 2014). These reasons, alongside its great potential for implementation in other industries, are perhaps why it is estimated that the size of the global CRISPR market is projected to exceed 30 billion U.S. dollars by 2033 (Mikulic, 2024).

Cas9 is an RNA-guided Endonuclease Able to Cut Double-Stranded DNA

Many different types of CRISPR/Cas systems have varying functionalities in manipulating genetic information. However, this paper focuses on CRISPR/**Cas*9***, as it is the most widely used CRISPR system.

CRISPR-Cas9, also called "genetic scissors", is an RNA-guided endonuclease that can correct genetic variants by cutting double-stranded DNA. This feature allows for gene editing and gene expression control (Asmamaw et al, 2021; Anders et al, 2014). The system contains the Cas9 protein which requires a single guide RNA (sgRNA) or guide RNA (gRNA), which consists of two parts: the CRISPR RNA (crRNA) and trans-activating CRISPR RNA (tracrRNA). The crRNA is usually 18-20 base pairs long and identifies and pairs with the target DNA sequence, whereas tracrRNA is an extended sequence with loops that acts as a binding scaffold for the Cas9 nuclease. (Nishimasu et al, 2014; Asmamaw et al, 2021; Anders et al, 2014).

The Cas9 protein consists of several domains with two main functionalities: recognition of target DNA (REC lobe) and cleavage (NUC lobe) (*Figure 1*). Firstly, the REC1 and REC2 (REC lobe) domains are responsible for attaching the gRNA which will recognize and bind the target DNA sequence. Secondly, the HNH and RuvC (NUC lobe) domains are used to cut each DNA strand, whilst the *PAM interacting* (PI) domain (NUC lobe) is responsible for PAM recognition, which is discussed later in this paper (Shao et al, 2016; Asmamaw et al, 2021; Nishimasu et al, 2014).

CRISPR Cas9 Recognizes and Cleaves Target DNA

The CRISPR/Cas9 genome editing mechanism can be divided into three fundamental steps: recognition, cleavage, and repair (Shao et al, 2016). As confirmed by Yang et al (2018) via FRET, the endonuclease spontaneously undergoes three major conformational changes upon transitioning between each of the aforementioned steps.

CRISPR/Cas9 target recognition and cleavage mechanism

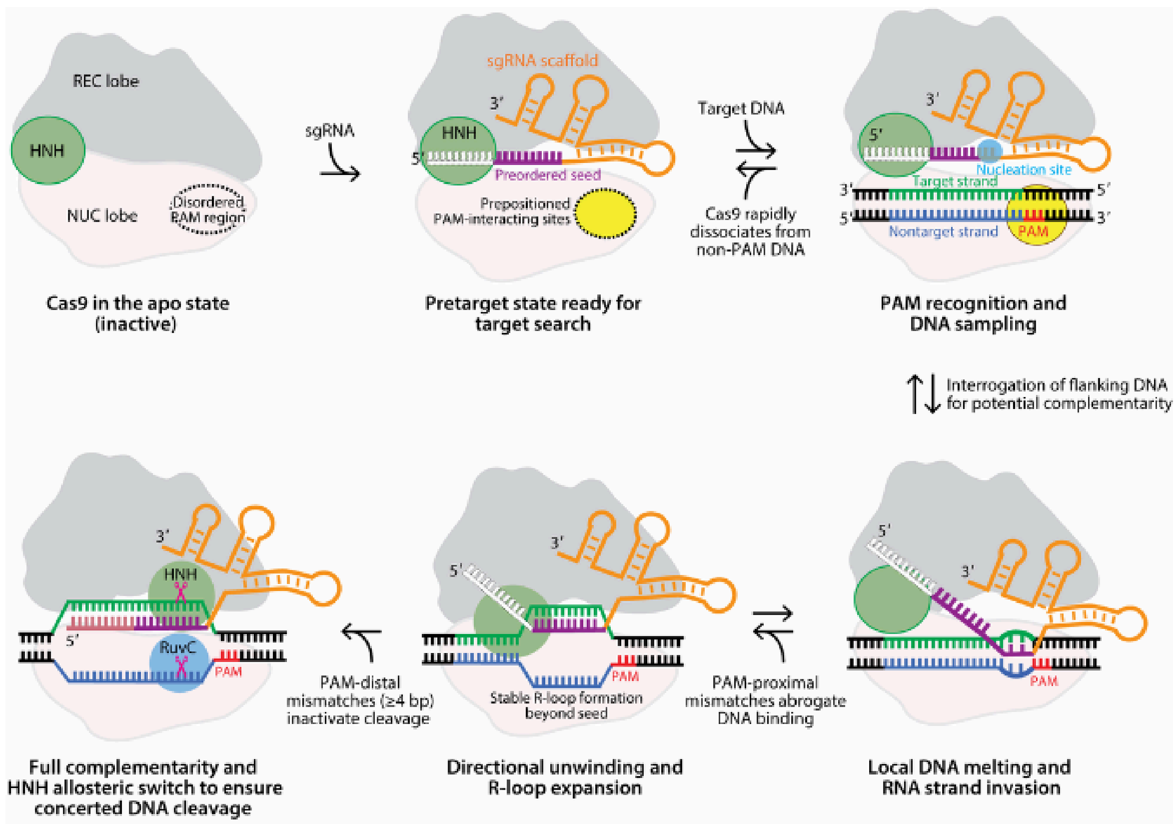

Figure 1 ***CRISPR/Cas9 target recognition and cleavage mechanism*** *Image from Jiang & Doudna (2017)*
Step-wise activation of Cas9 by binding gRNA (or sgRNA) scaffold is followed by PAM sequence recognition by PI domain (yellow). Successful recognition of Cas9 variant-dependent PAM sequence (red) initiates interrogation of upstream DNA sequence (green) for gRNA (purple) complementarity. The emergence of a stable DNA-gRNA hybrid triggers a double-strand break 3 nucleotides upstream of PAM. DNA cleavage is performed by HNH (green; round) and RuvC (blue; round) domains.

As a first step, the PI domain recognizes the PAM sequence within the DNA via a weak interaction with the major groove. PAM recognition is not required for Cas9-DNA binding, yet it is a pivotal step for the later activation of the endonuclease activity (Zhang et al, 2017; Anders et al, 2014) (*Figure 1*). The PAM sequence is a short, conserved DNA sequence (ranging from 2 to 5 base pairs) downstream of the cut site, with its length and PAM sequence nucleotides varying depending on the bacterial species (Ceasar et al, 2016). The most commonly used Cas9 variant

is the one obtained from *Streptococcus pyogenes* (spCas9) and can recognize the PAM sequence as 5'-NGG-3' (where N can be any nucleotide).

Secondly, if the Cas9 identifies a target site with the appropriate PAM, it triggers local DNA unwinding followed by the formation of a gRNA-DNA heteroduplex (R-loop) (Asmamaw et al, 2021; Nishimasu et al, 2014, Anders et al, 2014). Moreover, Cas9 investigates gRNA-target DNA complementary via base pairing to the DNA target sequence (*Figure 1*) upstream of PAM in the gene of interest (Shao et al, 2016). In the absence of the gRNA, the Cas9 protein remains in its inactive form and will not bind DNA neither via PI nor by DNA-gRNA binding (Ceasar et al, 2016; Asmamaw et al, 2021).

Lastly, a stable gRNA-DNA heteroduplex (R-loop) initiates Cas9 nuclease activity (O'Geen et al, 2014; Zhang et al, 2017; Yang et al, 2018) and cleavage (*Figure 1*). Interactions of the stable R-loop with the Cas9 cause the endonuclease to alter its conformation from the open state to an intermediate state. The HNH and RuvC domains will migrate closer to the DNA and induce double-stranded breaks (DSBs) at a site canonically located 3 base pairs upstream of the protospacer adjacent *motif.* A catalytically dead (dCas9) version with mutations in HNH and RuvC presumes the loss of endonuclease function and the capacity to cut the two DNA strands (Ceasar et al, 2016; Gasiunas et al, 2012; Cai et al, 2023; Yang et al, 2018).

Overcoming Limitations of CRISPR/Cas9

The powerful Cas9 editing tool has been extensively used in plant and animal biotechnology, as well as in advancing novel medical therapies. More precisely, the applications of this system have extended to areas such as gene expression regulation, chromatin and epigenetic editing, imaging and genome mapping (Adli, 2018; Yu et al, 2019; Li et al, 2023; Zhu et al, 2020; Luthra et al, 2021; Li et al 2021; Chen et al, 2019; Platt et al, 2014).

Nevertheless, the system is not without limitations. For instance, WT Cas9 is unable to edit areas of the genome that are PAM-free or that contain PAM sequences located farther than 5-20 nucleotides away from the target (Zhao et al, 2017). Moreover, another drawback is the erroneous off-target activity assignable to gRNA-target DNA mismatches. While in some cases the protein activity can severely decrease (by up to 96% in 2-base pairs mismatches) (Anderson et al, 2015), it can also cause off-target binding or cleavage via non-canonical base pairing, base skipping, or distortions in either the DNA backbone or the protein REC lobe (Fu et al, 2013; Pacesa et al, 2022).

These limitations are major setbacks for Cas9 utilization, such as *in vivo* therapies (Atkins et al, 2021; Höijer et al, 2022, Lopes & Prasad, 2024) where off-target activity can lead to toxicity and can even cause transmissible structural variants in offspring (Höijer et al, 2022).

Thus, it is imperative to investigate the Cas9/gRNA-DNA binding affinity *in vitro* to gain more mechanistic insight. This would mean studying how gRNA affects binding and function, investigating the change in conformations and developing Cas9 mutants with improved activity or tailored for each application. Other reasons for preferring *in vitro* probing are the cleaner

high-resolution measurement due to less confounding factors (i.e.transcription factors or DNA binding molecules), no cell fitness effects, control over Cas9 variant, spacer and Cas9 inhibitors, and unveiling rare binding sites by increasing enzyme concentration (Orenstein & Shamir, 2017; Tsai et al, 2017). Many strategies identified in the literature rely on screening custom/mutant Cas9 and selecting the most preferred variants.

However, Cas9/gRNA-DNA binding screening techniques often improve screening precision or throughput at the expense of the other because to obtain a larger number of sample data, the methodologies often exhibit lower complexity of the aspect being measured or decreased number of data points for a single molecule. This leaves a gap in contemporary research regarding critical evaluation of *in vitro* Cas9/gRNA-DNA binding screening that can offer both *precise* and *high-throughput* data.

This thesis aims to answer the following research questions: What are the main *in vitro* protein-DNA binding techniques that can be utilized in High-Throughput Cas9/gRNA-DNA binding assays? What are the factors contributing to their underlying differences? Where possible, to what extent is the precision of Cas9 screening exchanged for higher data throughput?

Main

Several *in vitro* techniques have been developed to study Cas9/gRNA-DNA binding, with some of them offering more insight into DNA-protein binding force, equilibrium association constant (K_a), equilibrium dissociation constant (K_d), or other relevant parameters. All of the techniques evaluated in this thesis are merely Cas9/gRNA-DNA techniques measuring different binding parameters and can all be utilized to study the effect of gRNA, target DNA or Cas9 variants on binding.

Moreover, whilst some techniques study the Cas9 activity including DNA cleavage, this paper will focus on DNA binding. For a detailed view on the advantages and disadvantages of each technique, consult the *Discussion* section (*Table 1*).

Cas9/gRNA-DNA Binding Affinity Assays

EMSA

EMSA-based assays deploy a more classical approach, by incubating catalytically deactivated Cas9-gRNA (dCas9) to a randomly generated library of target-DNA sequences, followed by observing the migration of the bound versus unbound target DNA fragments in an electrophoresis gel (*Figure 2)*. The ribonucleoprotein complexes (RNP) comprising of Cas9 bound to a gRNA are then isolated from the gel and the DNA, gRNA and Cas9 variant can be separated and analyzed (Anders et al, 2014). To test the effect of gRNA on binding affinity, EMSA can be performed using the same Cas9 variant and target DNA with a library of different gRNA molecules.

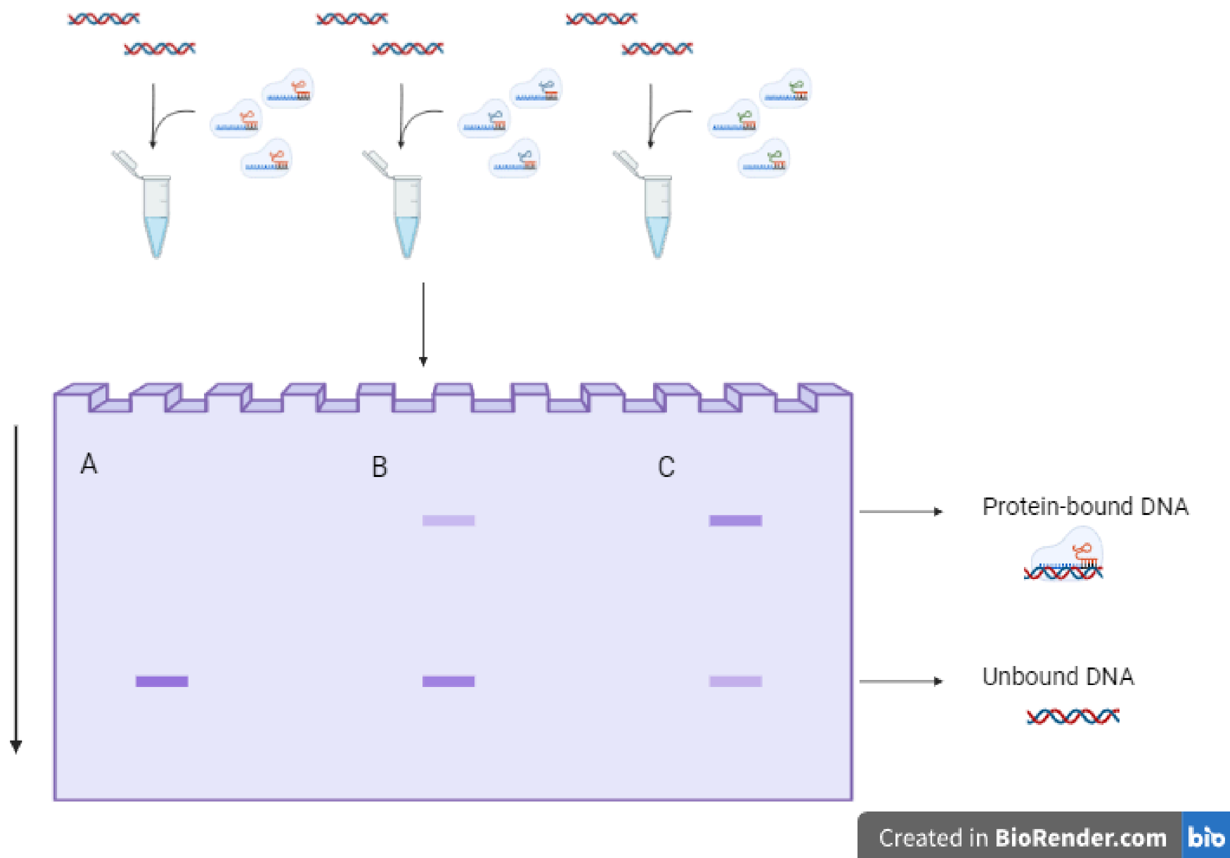

Figure 2 ***EMSA overview*** *Illustration created using Biorender.com*
Target DNA sequences (either identical or randomized) are incubated with Cas9 with various gRNAs. The mixture is purified and a mobility shift assay is performed via gel electrophoresis. (A) DNA does not bind the Cas9-gRNA combination (RNP) tested for; (B) DNA binds with low affinity the RNP combination it is tested for; (C) DNA binds with high affinity (or lower Kd) the RNP combination it is tested for.

Improved versions can assess strand-by-strand Cas9/gRNA-DNA K_d (Richardson et al, 2016) or can be coupled with high-throughput (HT) sequencing to i) train *in silico* biophysical models that predict the position-specific energy contribution of each nucleotide to binding (e.g. SELEX-seq and SelexGLM; Zhang et al, 2017) or ii) characterize consensus motif sequence (MEM and Perl scripts in *Bind-n-Seq*; Zykovich et al, 2009). Although widely used for transcription factors so far, the HT-Selex variant of this technique functions by multiplexed sequencing of all nucleotide sequences that bind to the protein of interest (i.e. dCas9). During the enrichment steps (removal of low-affinity sequences and amplification of the remaining fragments), the multiplexed sequencing is repeated several times, offering a relatively wide view of all sequences that can be bound by one protein.

The outcomes of EMSA-derived techniques (see *Discussion*) can be used to identify which DNA-RNP combinations bind with higher affinities by assuming that the most abundant RNP combinations have the highest binding stability and binding force.

ChIP-seq
Chromatin Immunoprecipitation followed by high-throughput sequencing is one of the most used methods to assess dCas9/gRNA-DNA binding preference on either fractions or whole genomes (Duan et al, 2014; O'Geen et al, 2014). This method is derived from the study of transcription factors and relies on immobilizing and mapping the dCas9 after it is bound to its DNA target, usually followed by tagging the dCas9 with an antibody. Needless to say, the technique offers a quantitative assessment on which RNP-DNA combination yields the highest amount of bound complex. Furthermore, some techniques deploy Cas9 target sequence identification via NGS (Lopes & Prasad, 2024; Savic et al, 2015; Duan et al, 2014; O'Geen et al, 2014).

Although techniques including cleavage analysis on top of the Cas9/gRNA-DNA binding assay (Extru-seq Kwon et al, 2023; Digenome-seq Kim et al, 2015; DIG-seq Kim et al, 2018; nDigenome-seq Kim et al, 2020; SITE-seq Cameron et al, 2017; CIRCLE-seq Tsai et al, 2017; CHANGE-seq Lazzarotto et al, 2020; Lopes & Prasad, 2024) are not the subject of this paper, Extru-seq, CHANGE-seq, and DIG-seq benefit from higher clinical validation rates in using chromatin. Out of the aforementioned, Extru-seq mimics cell-based strategies with outmost fidelity by incubating the Cas9 RNP with live cells, followed by rapid RNP-DNA isolation via extrusion of the mix through a filtering membrane, ensuring measurements in cell-like chromatin and epigenetic landscape (Kwon et al, 2023).

HITS-FLIP
The technique described by Boyle et al (2017) allows for very high-throughput quantitative measurement of dCas9 association and dissociation by fluorescence tracking. The system functions by immobilizing a large pool of mutated versions of a well-characterized 20 bp DNA sequence and visualizing in real-time the binding and unbinding of fluorescently-labeled dCas9-gRNA complexes to the fluorescently-labeled DNA targets. K_a and K_d are calculated from the observed on-rate and off-rate of the fluorescently labeled dCas9 RNP to the DNA sequences (*Figure 3A*). During association (incubation of RNP and DNA pool for 12h) and dissociation (washing the flow cell with dCas9-free buffer), images were collected to track the on and off rates (*Figure 3B*). The sequences are flanked by adapters, allowing for mass parallel sequencing, sequence spacial mapping and biochemical profiling of the DNA pool using a GAIIx instrument (*Figure 3C*).

Another method designed by Nutiu et al (2011) functions on a similar principle but uses a step-wise concentration gradient. After creating randomized k-mers, sequencing, and spatial mapping of sequences within the pool, the fluorescently tagged Cas9 or dCas9 is introduced in well with DNA at different concentrations (*Figure 3D*). This allows measuring K_a and K_d and differentiation of high versus low-affinity binding motifs (following the hypothesis that low-concentration Cas9 RNP will bind high-affinity motifs first). The assay is relatively easy to use, as it does not consist of many intermediate steps.

Another similar, albeit fluorescence-free, version of this method deploys filtering of the dCas9-bound DNA using a nitrocellulose filter (Boyle et al, 2020). After RNP-DNA incubation and mix filtration, the unbound DNA will pass through the filter and will be collected for

sequencing and quantification (*Figure 3E*). The DNA sequences are collected at different time points and barcoded with different timepoint primers, allowing for quantitative results of the highest binding preferences among all RNP combinations to all DNA sequences using the sequencer. This allows for measuring fraction bound (f_{final}) and the observed rate constant (k_{obs}). Boyle et al (2020) also include the flanking DNA sequences in data analysis, studying not only the impact of gRNA and target DNA on binding but also the relevance of the genomic context. The sequences surrounding the 5' and 3' ends of the target DNA and PAM sequence, respectively, are divided into 3-mers and their nucleic acid composition is correlated with the observed impact they have on the Cas9 association rate for the same gRNA (*Figure 3F*).

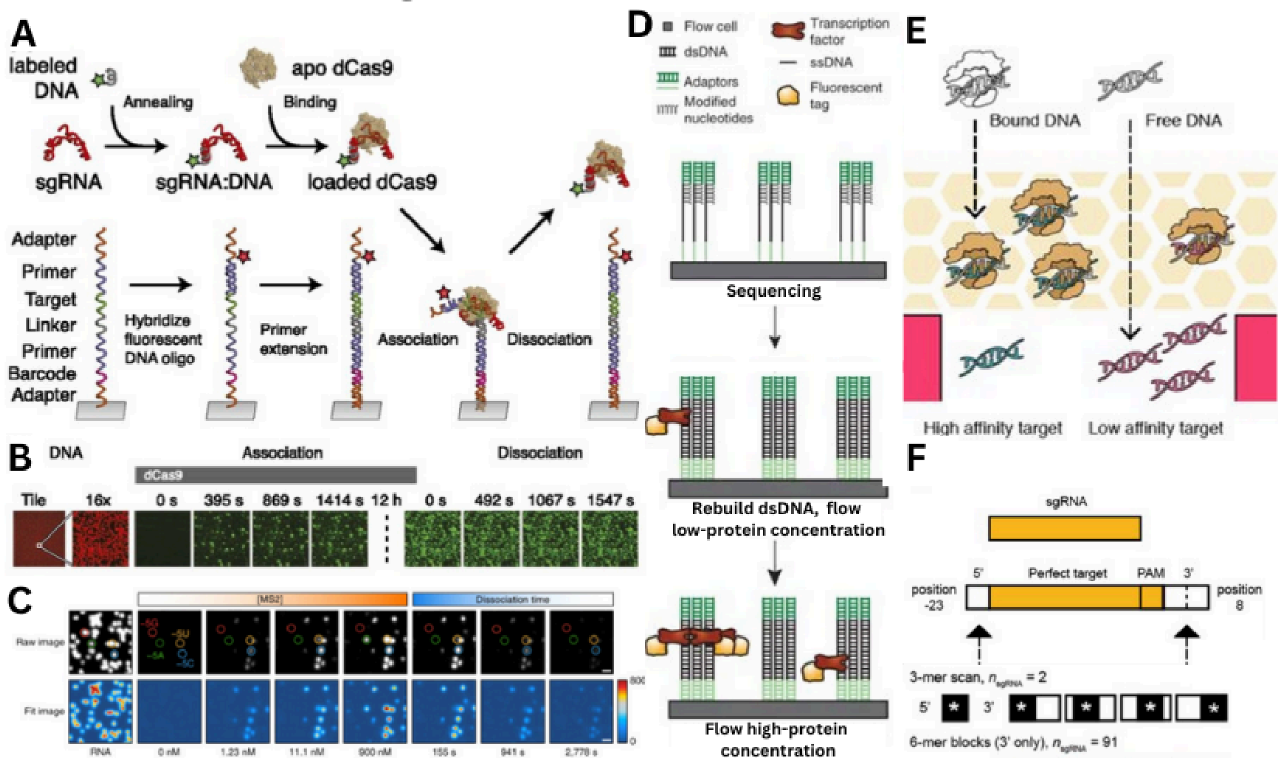

Figure 3 ***Biochemical profiling strategies*** *Images from Boyle et al (2017) (A&B), Buenrostro et al (2014) (C), Nutiu et al (2011) (D), (Boyle et al, 2020) (E&F)*

(A) *Experimental design of the HT biochemical profiling procedure. The Cy3 label (orange dye; green in the figure) is added to the gRNA before being loaded into the dCas9 for RNP formation. The Cas9 will associate and then dissociate from the Alexa Fluor 647-labeled DNA (red).*

(B) *Tracking association and dissociation of dCas9 RNP (green) from the DNA pool (red) in the HT flow cell. Images were taken at time intervals during the 12h incubation of the DNA with the dCas9 RNP ("Association"), followed by washing of the flow cell with dCas9-free buffer ("Dissociation"). Images taken at 532-nm excitation; 16x magnification.*

(C) *Example of GAIIx instrument principle of usage and capability to track single nucleotide polymorphisms (SNP). The example in the figure shows the binding profiling of fluorescently tagged MS2 coat protein SNP variants (red, yellow, blue, and green circles). MS2 is bound at increasing concentrations to the RNA, followed by a washing step to determine dissociation constants (Buenrostro et al, 2014).*

(D) Nutiu et al (2011) HITS-FLIP overview. After random k-mers sequencing, (optional washing step) the dsDNA are rebuilt and the fluorescently-tagged protein is added in different concentrations to the flow cell, allowing measurement of high vs low-affinity binding motifs.

(E) The key principle in HITS-FLIP by Boyle et al (2020): unbound DNA undergoes vacuum-filtration through the nitrocellulose filter whilst RNP-bound DNA is trapped. The free DNA is collected at various time intervals and a fraction of it is sequenced

(F) HITS-FLIP Boyle et al (2020): The region flanking the target DNA is divided into 3-mers and their nucleic acid composition is correlated with the measured Cas9-DNA association rates for the same gRNA.

DNA Beacon (FRET)

DNA Beacon strategies usually entail the activation of a fluorescent signal upon DNA-Cas9 binding. As seen in *Figure 4*, the steps of this method comprise of i) gRNA-bound Cas9 (or dCas9) binds to the target DNA and initiates R-loop formation; ii) denaturing separates the dsDNA strands that are labeled with a Fluorophore and a Quencher, respectively. The Beacons fluorescence intensity baseline is low because of the proximity between the fluorescence tag and the quencher. iii) when Cas9 binding is complete the recorded fluorescence intensity reaches maximal values (due to separation from quencher) (*Figure 4*). Target DNA sequences are short, consisting of the protospacer, PAM region, and a short flanking region downstream of the PAM region (≈15 bp) (Mekler et al, 2020; Mekler et al, 2019; Mekler et al, 2017).

Moreover, the DNA Beacons follow one of the models illustrated in *Figure 4*: Beacon 1 contains a disrupted non-complementary strand which will dissociate upon Cas9 binding to the complementary strand, whilst the 2nd Beacon model consists of continuous oligonucleotide sequence on the non-complementary strand which will only bend upon Cas9 association (Mekler et al, 2019; Mekler et al, 2017). Despite the discrepancies, the binding kinetics do not differ significantly between these two models (Mekler et al, 2016).

DNA Beacon strategy is used to investigate the binding affinity and K_d of distinct Cas9/gRNA-DNA during the binding of different DNA samples, namely fully double-stranded, partially single-stranded, or “bubbled model substrates” (Mekler et al, 2016).

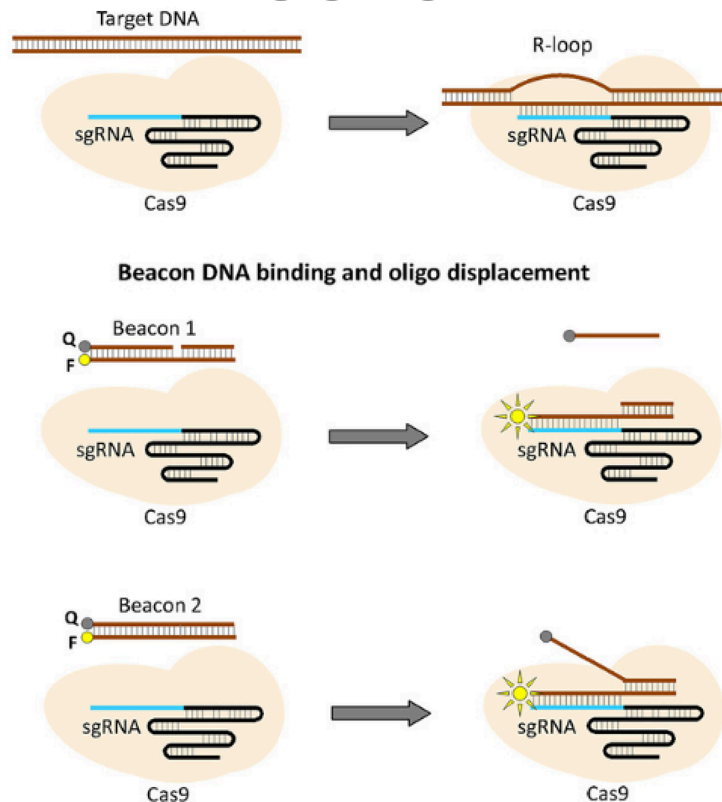

Figure 4 ***Cas9-DNA binding signaling via DNA Beacon assay*** *Image from Mekler et al (2019)*
Strands of DNA Beacon (target DNA) are labeled with a fluorophore (F) and a quencher (Q), respectively. Upon Cas9 association and R-loop formation, F and Q will separate and the registered fluorescence will increase, compared to the baseline level which is low due to Q and F proximity. The non-complementary strand can be continuous or discontinuous which will cause the bending or complete dissociation of the non-complementary strand upon Cas9 binding.

Additional Binding Assays

Other promising techniques were evaluated for this study, although not included due to various limitations. However, out of the binding assays that can be deemed promiscuous, this thesis mentions below some biomechanical assays that quantify the binding force between the DNA and gRNA-Cas. These can be used to measure the force required to unbind different gRNA from the target DNA, offering a biophysical approach to assessing the RNP-DNA binding affinity.

Optical Tweezers

A variant of a single-molecule high-resolution assay of gRNA-dCas9 binding to DNA is the Optical Tweezer technique. Briefly, this allows manipulation of biomolecules with high precision, whilst measuring the force needed for separation of interacting molecules and their length variation as a function of the force applied. In this context, a high-power laser beam can be used to trap a bead connected to the target DNA sequence, while the binding dCas-gRNA complex is immobilized to a surface or by another beam. The laser is then uni-directionally moved to stretch the complex until the gRNA-dCas9-DNA breaks. This technique can be coupled with

single-molecule FRET (smFRET) for higher result validation or measurement of local stretching within the complex (Choudhary et al, 2019) or with imaging techniques for localizing bound proteins with sub-base-pair precision (Choudhary et al, 2019; Killian et al, 2018). Multiple laser traps can be used for a multiplexed assay (Noom et al, 2007). Despite being able to measure displacements up to Ångstrom spatial resolution and microsecond temporal resolution, and the force required to unbind dCas9/gRNA from target DNA at the piconewton level (Capitanio & Pavone; 2013), the assay does not offer a high-throughput variant that can measure more than just several gRNA-Cas9-DNA complexes at once.

Magnetic Tweezers

Magnetic Tweezers deploy a strategy similar to that of Optical Tweezers, the only difference being that the method of applying strain force to the gRNA-Cas9-DNA complex is via the DNA-bound bead which is metallic and attracted to a low-intensity magnet instead of being trapped in a laser beam. Despite being relatively low throughput compared to the biochemical assays in *Table 1*, variants still manage to obtain several hundred reads per experiment run (Feiz et al, 2024; Liang et al, 2021; De Vlaminck et al, 2011; Berghuis et al, 2016). Similar to Optical Tweezers, this category of single-molecule assays can measure high-resolution biomechanical aspects of DNA-protein binding, such as force required to unbind dCas9/gRNA from target DNA at the piconewton level, dissociation time, protein folding, protein displacement up to nanometer spatial resolution upon straining and submillisecond temporal resolution in displacement tracking.

Atomic Force Microscopy

Protein-DNA Atomic Force Microscopy (AFM) studies interactions by topographical imaging of target DNA with a low throughput. AFM uses a cantilever to scan with high-resolution the surface of biomolecules and can detect the dCas9-gRNA complexes bound to DNA. Using AFM to analyze a single protein-DNA binding reaction can reveal the specificity and affinity of protein binding, protein-induced DNA bending, stoichiometry of protein binding, changes in DNA structure and can form 3D images of the protein-DNA complex (Balderston et al, 2021; Mohan Bangalore & Tessmer, 2018; Josephs et al, 2015).

Surface Plasmon Resonance & Biolayer Interferometry

Surface Plasmon Resonance (SPR) is a highly-sensitive method to monitor real-time molecule binding. SPR measures the change in the refractive index at a Au/glass protein-coated surface that occurs upon substrate binding. In the context of this thesis, the dCas9 with different gRNAs can be immobilized on the surface and the target DNA could be flowed over the surface, as performed similarly in other protein-substrate experiments. SPR allows for measurement of association and dissociation rate, modeling of interaction kinetics and equilibrium binding analysis (Stahelin, 2013; Blay et al, 2020). In a CRISPR-Cas9 detection tool used by Zheng et al (2022), gRNA efficiency was measured with high precision using target DNA concentration as low as 1.3 fM and offered more sensitive binding affinity quantification than qPCR.

Immobilization can occur on microarrays in order to ensure higher throughput, but the technique is still not available in HT options comparable to the techniques in *Table 1*. Moreover, SPR also

entails other major disadvantages. For instance, SPR cannot discriminate between specific and non-specific binding, orientation of binding protein or DNA substrate, is very sensitive to any condition affecting the refractive index (including small movement, dCas9/gRNA complex unfolding), is highly sensitive to temperature and requires precise gear and methodology for calibration (Blay et al, 2020; Stahelin, 2013).

Another method similar to SPR is Biolayer Interferometry (BLI). This technique measures the phase change alteration between the layer of immobilized protein and target-DNA-bound immobilized protein. This technique is easier to use but less sensitive than SPR, also offers a medium-throughput assay of binding affinity, but is more suitable for small molecules (Overackeret al, 2021; Blay et al, 2020). However, medium-throughout CRISPR experiments suggest that BLI can be also be used in dCas9 assays to investigate PAM identification and crRNA efficiency (Müller-Esparza et al, 2020).

Other Binding Assays

Other techniques were not included in this evaluation due to either lack of extensive peer-review - DNA–Protein Binding Force Chip (BiFo-Chip; Severin & Gaub, 2012; Severin et al, 2011), parallelization of HT binding force screening (HT, but measures the rupture of the weakest binding DNA-RNP and compares it to other DNA complexes with a well-defined rupture values of other DNA duplexes; it does not measure binding preference and offers low-precision binding force quantification; Limmer et al, 2014) or rely on DNA distortion (Henneman et al, 2018).

Discussion & Comparison

Albeit their ability to study the gRNA-Cas9-DNA binding affinity, the assays presented in this paper are characterized by a great degree of diversity in terms of design, measurement, strengths and limitations (see *Table 1* for comparative evaluation and the section below for additional features; all comparative evaluations for each technique are based on the information, results and limitations presented in the corresponding articles).

Evaluation of Cas9-DNA Binding Screening Assays				
HT Cas9 binding screening method (*in vitro)*	**Description**	**Measurement**	**Advantages**	**Disadvantages**
EMSA	Gel migration comparison Cas9-DNA bound vs unbound DNA	Cas9/gRNA-DNA binding sites	EMSA: cost-effective, accessible Can be used to study small molecule interference on gRNA-DNA bond Preserves DNA-RNP complex structure	Usually not very high throughput No data on Cas9 preference to bind high-affinity sequences compared to the pool of all viable sequences for binding

		(applicable to the following EMSA)	Allows analysis of multiple gRNA simultaneously (applicable to the following EMSA)	No epigenetic validation (applicable to the following EMSA)
	Variants include a sequencing step of bound nucleic acids	Can measure strand-by-strandCas9/gRNA-DNA Kd (Richardson et al, 2016)		
		SELEX-seq; Bind-n-Seq - sequencing of bound nucleic acids	Measures position-specific energy contribution of each nucleotide in gRNA to DNA binding	
		HT-Selex counts sequences that can bind 1 protein being studied (i.e. Cas9)	HT Selex: **high-throughput data** Allows analysis of multiple gRNA simultaneously	Mass sequencing can be expensive HT-Selex design does not allow for all possible sequences to go through the enrichment step because only a fraction is selected for enrichment whilst the rest is sequenced => **lower data precision for some** target sequences
ChIP	Immobilize the Cas9 after binding target DNA, followed by tagging the Cas9 with an antibody	Binding protein locus	Higher data validation: chromatin usage instead of purified DNA Allows analysis of multiple gRNA simultaneously	Limited by the availability of suitable ChIP-seq grade antibodies: solved by CETCh-seq Cannot distinguish between direct and indirect DNA binding Labor-intensive **Low data validation** in **high-throughput** variants
	Numerous variants include a sequencing step (NGS) (CETCh-seq; Lopes & Prasad, 2024; Savic et al, 2015; Duan et al, 2014; O'Geen et al, 2014)	Sequence bound to gRNA (quantitative and qualitative)	Can be HT	Usually **low precision** Usually **throughput is lower** than in other techniques
	Variants include cleavage assays (Extru-seq, Digenome-seq, DIG-seq, nDigenome-seq, SITE-seq, CIRCLE-seq, CHANGE-seq, Extru-seq, DIG-seq,, SITE-seq, CIRCLE-seq, CHANGE-seq) Extru-seq: binding+digestion assays; similar to ChiP	Sequence bound to gRNA-DNA & cleavage activity	Extru-seq:Highest clinical validation. Cells are incubated with active Cas9 and chromatin is extruded and incubated with Cas9 RNP	Extru-seq:includes DNA cleaving step by Cas9
HITS-FLIP	HITS-FLIP Boyle et al (2017) - visual tracking of association and dissociation of tagged dCas9 to tagged DNA gRNA pool	gRNA-DNA Ka & Kd; off-target vs on-target on-rate binding seq via NGS	**High-Throughput** (≈150000) and **precise** Visual tracking of binding and unbinding of as little as 1 single-nucleotide polymorphism gRNA mutant to dCas9	Fluorescent tags have a risk of photobleaching due to length of experiments
	HITS-FLIP Nutiu et al (2011) - sequencing of random k-mer pool followed by fluorescent-based	Cas9/gRNA-DNA Ka & Kd; tracking of binding to	**Very High-Throughput** (≈100 mil measurement//run)	**Moderate precision** due to:

	binding measurement by increasing DNA-binding protein concentration	random k-mers	Could measure multiple gRNAs simultaneously (to the same dsDNA) Allows for experiment re-run without re-sequencing or resynthesizing dsDNA pool Lower cost per protein for multiplexed sequential assay (up to 8 proteins) No intermediate steps (reduces bias)	Fluorescence is subject to photobleaching, presumably by 35% median between measurements (can be corrected for, but is inefficient) Fluorescence intensity influenced by tile position on sequencer flow cell (can be corrected)
	HITS-FLIP Boyle et al (2020) - HT off and on-target assay by quantifying unbound DNA that is filtered through a nitrocellulose filter	f_{final} & K_{obs} (estimates) off and on-target assay (binding preference)	**High-Throughput** (≈10^3 off-targets/RNP/experiment) Not affected by photobleaching Allows analysis of multiple gRNA simultaneously Inclusion of genomic context outside of target sequence in the off-target assay Ease-of-use compared to fluorescence HITS-FLIP	Has a high detection threshold => might miss out some results **Moderate precision**- only a fraction of unbound DNA is sequenced Does not differentiate between productive and unproductive bound states Caveat: Cannot differentiate the RNPs that dissociate rapidly from high-affinity sites
DNA Beacon	Quenching-based signaling of Cas9 binding to short target DNA	Kd of only effector Cas9 at R-loop formation stage Can measure Cas9-gRNA assembly	Allows parallel assays of different gRNAs to the same Cas9 or of different Cas9 to the same gRNA Can use fully double-stranded, partially single-stranded, or bubbled model DNA substrates Sensitive: Baseline and maximal fluorescence values differ up to x50; can detect very low affinity binding targets Does not alter Cas9 structure with addittional tags, allowing **more precise** measurement of engineered Cas9 activity By design, it measures affinity of only functional Cas9-gRNA that reach R-loop formation step. Suitable for investigating PAM-specificities even in non-complementary spacer (Mekler et al, 2020)	**Low Throughput** (compared to other methods in this paper) Time-dependent: binding events detected only if they occur within a few hours (not specified) False positive if DNA concentration is too high and DNA-RNP affinity is too strong No proven clinical validation, despite claims (Mekler et al, 2020; Mekler et al, 2019)

Table 1 ***Comparative Description of Cas-DNA Binding Screening Assays (Detailed overview)***
The table comprises a general description of the Cas9-DNA binding affinity assays identified in the literature, the type of data recorded, advantages and disadvantages (columns). Supplementary features and contextualization in the Discussion section.

EMSA

A limitation of all the EMSA-based assays is the lack of qualitative data about the Cas9-DNA binding energy or a comparison between the preference of Cas9 to bind certain sequences over other viable ones. Furthermore, EMSA-based assays utilize purified DNA and not genomic DNA with other co-factors, offering both advantages (emphasis on the binding mechanism can be used better to train biophysical models) and disadvantages (lower binding mechanism validation due to lack of, for example, epigenetic elements involved, indirect DNA binding or not taking into account co-factors interfering with binding).

A limitation of HT-Selex stems from its design, mainly since during each step, only part of the bound fragments are sequenced, whilst only the remaining fractions will be used for the enrichment cycles. More specifically, the alteration of the data accuracy (by removing bound sequences from the pool during enrichment cycles) is exchanged for higher data throughput (Jolma et al, 2010; Orenstein & Shamir, 2016).

A strong point of this technique is the low number of sample preparation steps between the binding and the mobility assay that would alter the DNA-protein complex. This allows for a high degree of 3D structure conservation. For this reason, EMSA has been widely used in investigating the interference of small molecules in Cas9 binding. (Wang et al, 2022; Chen et al, 2022; Lu et al, 2021).

ChIP-seq

Among the major advantages is the increased validation due to utilizing chromatin instead of purified gDNA, which leads to more realistic testing conditions and significant breakthroughs in the effect of the chromatin state (open/closed) on the Cas9 binding (Wu et al, 2014).

To avoid a major limitation of this technique, the restricted availability of ChIP-seq grade antibodies (e.g. antibodies that tag custom variants of Cas9), CETCh-seq makes use of adding a Flag epitope to the Cas9. In this manner, only one generic anti-Flag antibody is required, making it possible to screen the position of virtually any protein, including mutated Cas9 on the DNA, without the need for custom antibodies or Cas9 overexpression (Lopes & Prasad, 2024; Savic et al, 2015). An automated version, iDeal ChIP-Seq Kit for Transcription Factors, has been shown to significantly reduce hands-on time and experiment duration to test dCas9 off-target binding. However, the read distributions by Campenhout et al (2018) show lower throughput and result fidelity, compared to the non-automated methodology.

HITS-FLIP

A major advantage of the methods described by Boyle et al (2017) is the high resolution, which allows visual tracking of the association/dissociation of as little as 1 dCas9-DNA complex of interest (*Figure 2C*). This allows for comparative DNA binding assays among many mutated versions of the gRNA, including SNPs (*Figure 2C*). Another advantage not discussed by Boyle et al (2017) but considered in this paper is the barcode sequence which allows for tracking of the target DNA, with potential in custom experiments in which one would need to use identical

target DNA sequences. This feature can be used either for testing the same target DNA in different epigenetic states or for testing dCas9 variants with the same gRNA.

Possible Caveat in Calculating K_d in Protein-DNA Binding Assays

Lastly, a caveat identified in most techniques presented in this paper is the calculation of Kd (Jarmoskaite et al, 2020). The equilibrium state infers that the system should not vary in time. Determining equilibrium dissociation constant (Kd) implies that there should be no difference in the amount of bound complex over time at the same concentration, opening the discussion on the precision of the measurements which can vary greatly, as seen in Gillmore et al (2021). Vulnerabilities can appear due to the controls used to measure Kd (Jarmoskaite et al, 2020), which are not followed over longer periods. In this regard, the most affected techniques are the ones based on fluorescence, EMSA, or ChIP (which do not allow long-term tracking due to photobleaching, reagent toxicity, or inefficiency). Other techniques that exploit biophysical approaches for studying the binding of molecules, rather than biochemical techniques, such as isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR) may have an upper hand.

Improvement of gRNA Design

The design of gRNA is a critical factor that can impact genome editing when using CRISPR systems. Studies suggest that small changes in gRNAs targeting the same loci can also make an impact on their efficiency, underlining the importance of *in silico* predictions (Lopes & Prasad, 2024; Shalem et al, 2014; Wang et al, 2014; Doench et al, 2016). Besides the fact that increased GC content increases gRNA efficacy, according to Lopes & Prasad (2024), modifying its length can also have an impact on Cas9 specificity. Adding two guanine nucleotides to the 5' end, truncating the sgRNA (Cho et al, 2013; Fu et al, 2014), chemical modifications to the gRNA (Ryan et al, 2018) or incorporation of modified nucleic acids can also increase gRNA/Cas9 specificity *in vitro* (Cromwell et al, 2018).

Conclusion

Selecting a high-throughput DNA binding affinity assay is a crucial step when designing *in vitro* (d)Cas9/gRNA-DNA experiments due to the need to test the efficacy of gRNA, target DNA or (d)Cas9 variant. Here, we critically evaluate DNA binding assays that can be used for Cas9/gRNA-DNA, with focus on their underlying differences and in what manner is HT exchanged for lower precision or vice versa. Although HITS-FLIP is a technique that offers both high precision and relatively high throughput when compared to the other techniques presented in this thesis, an ideal methodology or assessment paradigm is still unclear. However, as each method comprises of advantages and disadvantages, we recommend deploying complementary methods to increase the validation of the affinity computation and to cover a larger set of measurements for each Cas9/gRNA-DNA combination.

Lastly, monitoring the relevant literature in HT proteomics is indispensable in the current and future gene editing landscape and future research into real-time, HT, high-resolution binding assays is neccessary to unlock the full potential of gene editing.

References

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. In Nature Communications (Vol. 9, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1038/s41467-018-04252-2

Anders, C., Niewoehner, O., Duerst, A. et al. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature 513, 569–573 (2014). https://doi.org/10.1038/nature13579

Anderson, E. M., Haupt, A., Schiel, J. A., Chou, E., Machado, H. B., Strezoska, Ž., Lenger, S., McClelland, S., Birmingham, A., Vermeulen, A., & Smith, A. van B. (2015). Systematic analysis of CRISPR–Cas9 mismatch tolerance reveals low levels of off-target activity. In Journal of Biotechnology (Vol. 211, pp. 56–65). Elsevier BV. https://doi.org/10.1016/j.jbiotec.2015.06.427

Asmamaw, M., & Zawdie, B. (2021). Mechanism and Applications of CRISPR/Cas-9-Mediated Genome Editing. Biologics: Targets and Therapy, 15, 353–361. https://doi.org/10.2147/BTT.S326422

Atkins, A., Chung, C.-H., Allen, A. G., Dampier, W., Gurrola, T. E., Sariyer, I. K., Nonnemacher, M. R., & Wigdahl, B. (2021). Off-Target Analysis in Gene Editing and Applications for Clinical Translation of CRISPR/Cas9 in HIV-1 Therapy. In Frontiers in Genome Editing (Vol. 3). Frontiers Media SA. https://doi.org/10.3389/fgeed.2021.673022

Balderston, S., Taulbee, J.J., Celaya, E. et al. Discrimination of single-point mutations in unamplified genomic DNA via Cas9 immobilized on a graphene field-effect transistor. Nat Biomed Eng 5, 713–725 (2021). https://doi.org/10.1038/s41551-021-00706-z

Berghuis, B. A., Köber, M., van Laar, T., & Dekker, N. H. (2016). High-throughput, high-force probing of DNA-protein interactions with magnetic tweezers. In Methods (Vol. 105, pp. 90–98). Elsevier BV. https://doi.org/10.1016/j.ymeth.2016.03.025

Boyle, E., Becker, W., Bai, H., Chen, J., Doudna, J., & Greenleaf, W. (2020). Quantification of Cas9 binding and cleavage across diverse guide sequences maps landscapes of target engagement. Science Advances, 7. https://doi.org/10.1101/2020.09.09.290668

Boyle, E. A., Andreasson, J. O. L., Chircus, L. M., Sternberg, S. H., Wu, M. J., Guegler, C. K., Doudna, J. A., & Greenleaf, W. J. (2017). High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. In Proceedings of the National Academy of Sciences (Vol. 114, Issue 21, pp. 5461–5466). Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.1700557114

Blay, V., Tolani, B., Ho, S. P., & Arkin, M. R. (2020). High-Throughput Screening: today's biochemical and cell-based approaches. In Drug Discovery Today (Vol. 25, Issue 10, pp. 1807–1821). Elsevier BV. https://doi.org/10.1016/j.drudis.2020.07.024

Buenrostro, J. D., Araya, C. L., Chircus, L. M., Layton, C. J., Chang, H. Y., Snyder, M. P., & Greenleaf, W. J. (2014). Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. In Nature Biotechnology (Vol. 32, Issue 6, pp. 562–568). Springer Science and Business Media LLC. https://doi.org/10.1038/nbt.2880

Cai R, Lv R, Shi X, Yang G, Jin J. CRISPR/dCas9 Tools: Epigenetic Mechanism and Application in Gene Transcriptional Regulation. International Journal of Molecular Sciences. 2023; 24(19):14865. https://doi.org/10.3390/ijms241914865

Cameron, P., Fuller, C. K., Donohoue, P. D., Jones, B. N., Thompson, M. S., Carter, M. M., Gradia, S., Vidal, B., Garner, E., Slorach, E. M., Lau, E., Banh, L. M., Lied, A. M., Edwards, L. S., Settle, A. H., Capurso, D., Llaca, V., Deschamps, S., Cigan, M., … May, A. P. (2017). Mapping the genomic landscape of CRISPR–Cas9 cleavage. In Nature Methods (Vol. 14, Issue 6, pp. 600–606). Springer Science and Business Media LLC. https://doi.org/10.1038/nmeth.4284

Campenhout, C. V., Anne-Clémence Veillard, Zelisko-Schmidt, A., Gueydan, C., Kruys, V., & Sabatel, C. (2018). Optimize the selection of guide RNA by ChIP to keep CRISPR on-target. Unpublished. https://doi.org/10.13140/RG.2.2.27239.96161

Capitanio, M., & Pavone, F. S. (2013). Interrogating Biology with Force: Single Molecule High-Resolution Measurements with Optical Tweezers. In Biophysical Journal (Vol. 105, Issue 6, pp. 1293–1303). Elsevier BV. https://doi.org/10.1016/j.bpj.2013.08.007

Ceasar SA, Rajan V, Prykhozhij SV, Berman JN, Ignacimuthu S. Insert, remove or replace: a highly advanced genome editing system using CRISPR/Cas9. Biochim Biophys Acta Mol Cell. 2016;1863(9):2333–2344. doi:10.1016/j.bbamcr.2016.06.009

Cha, J., & Lee, J. Y. (2021). A novel high-throughput single-molecule technique: DNA curtain. In Journal of the Korean Physical Society (Vol. 78, Issue 5, pp. 442–448). Springer Science and Business Media LLC. https://doi.org/10.1007/s40042-020-00031-9

Chen, K., Wang, Y., Zhang, R., Zhang, H., & Gao, C. (2019). CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. In Annual Review of Plant Biology (Vol. 70, Issue 1, pp. 667–697). Annual Reviews. https://doi.org/10.1146/annurev-arplant-050718-100049

Chen, S., Chen, D., Liu, B., & Haisma, H. J. (2022). Modulating CRISPR/Cas9 genome-editing activity by small molecules. In Drug Discovery Today (Vol. 27, Issue 4, pp. 951–966). Elsevier BV. https://doi.org/10.1016/j.drudis.2021.11.018

Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S., & Kim, J.-S. (2013). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. In Genome Research (Vol. 24, Issue 1, pp. 132–141). Cold Spring Harbor Laboratory. https://doi.org/10.1101/gr.162339.113

Choudhary, D., Mossa, A., Jadhav, M., & Cecconi, C. (2019). Bio-Molecular Applications of Recent Developments in Optical Tweezers. In Biomolecules (Vol. 9, Issue 1, p. 23). MDPI AG. https://doi.org/10.3390/biom9010023

Cromwell, C. R., Sung, K., Park, J., Krysler, A. R., Jovel, J., Kim, S. K., & Hubbard, B. P. (2018). Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity. In Nature Communications (Vol. 9, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1038/s41467-018-03927-0

De Vlaminck, I., Henighan, T., van Loenhout, M. T. J., Pfeiffer, I., Huijts, J., Kerssemakers, J. W. J., Katan, A. J., van Langen-Suurling, A., van der Drift, E., Wyman, C., & Dekker, C. (2011). Highly Parallel Magnetic Tweezers by Targeted DNA Tethering. In Nano Letters (Vol. 11, Issue 12, pp. 5489–5493). American Chemical Society (ACS). https://doi.org/10.1021/nl203299e

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., & Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. In Nature Biotechnology (Vol. 34, Issue 2, pp. 184–191). Springer Science and Business Media LLC. https://doi.org/10.1038/nbt.3437

Duan, J., Lu, G., Xie, Z. et al. Genome-wide identification of CRISPR/Cas9 off-targets in human genome. Cell Res 24, 1009–1012 (2014). https://doi.org/10.1038/cr.2014.87

Feiz, M. S., Wubulikasimu, Y., Cnossen, J., Papini, F. S., Bera, S. C., Quack, S., & Dulin, D. (2024). A high-throughput correlative magnetic tweezers-TIRF assay to investigate protein-nucleic acids interactions. In Biophysical Journal (Vol. 123, Issue 3, p. 286a). Elsevier BV. https://doi.org/10.1016/j.bpj.2023.11.1782

Fu, Y., Foden, J., Khayter, C. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31, 822–826 (2013). https://doi.org/10.1038/nbt.2623
Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., & Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. In Nature Biotechnology (Vol. 32, Issue 3, pp. 279–284). Springer Science and Business Media LLC. https://doi.org/10.1038/nbt.2808

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. In Proceedings of

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

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., Lebwohl, D. (2021). CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. In New England Journal of Medicine (Vol. 385, Issue 6, pp. 493–502). Massachusetts Medical Society. https://doi.org/10.1056/nejmoa2107454

Henneman, B., Heinsman, J., Battjes, J., Dame, R.T. (2018). Quantitation of DNA-Binding Affinity Using Tethered Particle Motion. In: Dame, R. (eds) Bacterial Chromatin. Methods in Molecular Biology, vol 1837. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-8675-0_14

Höijer, I., Emmanouilidou, A., Östlund, R. et al. CRISPR-Cas9 induces large structural variants at on-target and off-target sites in vivo that segregate across generations. Nat Commun 13, 627 (2022). https://doi.org/10.1038/s41467-022-28244-5

Jarmoskaite, I., AlSadhan, I., Vaidyanathan, P. P., & Herschlag, D. (2020). How to measure and evaluate binding affinities. In eLife (Vol. 9). eLife Sciences Publications, Ltd. https://doi.org/10.7554/elife.57264

Jiang, F., & Doudna, J. A. (2017). CRISPR–Cas9 Structures and Mechanisms. In Annual Review of Biophysics (Vol. 46, Issue 1, pp. 505–529). Annual Reviews. https://doi.org/10.1146/annurev-biophys-062215-010822

Jolma, A., Kivioja, T., Toivonen, J., Cheng, L., Wei, G., Enge, M., Taipale, M., Vaquerizas, J. M., Yan, J., Sillanpää, M. J., Bonke, M., Palin, K., Talukder, S., Hughes, T. R., Luscombe, N. M., Ukkonen, E., & Taipale, J. (2010). Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. In Genome Research (Vol. 20, Issue 6, pp. 861–873). Cold Spring Harbor Laboratory. https://doi.org/10.1101/gr.100552.109

Josephs, E. A., Kocak, D. D., Fitzgibbon, C. J., McMenemy, J., Gersbach, C. A., & Marszalek, P. E. (2015). Structure and specificity of the RNA-guided endonuclease Cas9 during DNA interrogation, target binding and cleavage. In Nucleic Acids Research (Vol. 43, Issue 18, pp. 8924–8941). Oxford University Press (OUP). https://doi.org/10.1093/nar/gkv892

Killian, J. L., Inman, J. T., & Wang, M. D. (2018). High-Performance Image-Based Measurements of Biological Forces and Interactions in a Dual Optical Trap. In ACS Nano (Vol. 12, Issue 12, pp. 11963–11974). American Chemical Society (ACS). https://doi.org/10.1021/acsnano.8b03679

Kim, D., & Kim, J.-S. (2018). DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA. In Genome Research (Vol. 28, Issue 12, pp. 1894–1900). Cold Spring Harbor Laboratory. https://doi.org/10.1101/gr.236620.118

Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H. R., Hwang, J., Kim, J.-I., & Kim, J.-S. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. In Nature Methods (Vol. 12, Issue 3, pp. 237–243). Springer Science and Business Media LLC. https://doi.org/10.1038/nmeth.3284

Kim, D. Y., Moon, S. B., Ko, J.-H., Kim, Y.-S., & Kim, D. (2020). Unbiased investigation of specificities of prime editing systems in human cells. In Nucleic Acids Research (Vol. 48, Issue 18, pp. 10576–10589). Oxford University Press (OUP). https://doi.org/10.1093/nar/gkaa764

Kwon, J., Kim, M., Hwang, W., Jo, A., Hwang, G.-H., Jung, M., Kim, U. G., Cui, G., Kim, H., Eom, J.-H., Hur, J. K., Lee, J., Kim, Y., Kim, J., Bae, S., & Lee, J. K. (2023). Extru-seq: a method for predicting genome-wide Cas9 off-target sites with advantages of both cell-based and in vitro approaches. In Genome Biology (Vol. 24, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1186/s13059-022-02842-4

Lazzarotto, C.R., Malinin, N.L., Li, Y. et al. CHANGE-seq reveals genetic and epigenetic effects on CRISPR–Cas9 genome-wide activity. Nat Biotechnol 38, 1317–1327 (2020). https://doi-org.proxy-ub.rug.nl/10.1038/s41587-020-0555-7

Liang, L., Wang, Z., Qu, L., Huang, W., Guo, S., Guan, X., Zhang, W., Sun, F., Yuan, H., Zou, H., Liu, H., & Yu, Z. (2021). Single-molecule multiplexed profiling of protein–DNA complexes using magnetic tweezers. In Journal of Biological Chemistry (Vol. 296, p. 100327). Elsevier BV. https://doi.org/10.1016/j.jbc.2021.100327

Li, J., Xu, R., Qin, R., Liu, X., Kong, F., & Wei, P. (2021). Genome editing mediated by SpCas9 variants with broad non-canonical PAM compatibility in plants. In Molecular Plant (Vol. 14, Issue 2, pp. 352–360). Elsevier BV. https://doi.org/10.1016/j.molp.2020.12.017

Li, T., Yang, Y., Qi, H., Cui, W., Zhang, L., Fu, X., He, X., Liu, M., Li, P., & Yu, T. (2023). CRISPR/Cas9 therapeutics: progress and prospects. In Signal Transduction and Targeted Therapy (Vol. 8, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1038/s41392-023-01309-7

Limmer, K., Pippig, D. A., Aschenbrenner, D., & Gaub, H. E. (2014). A Force-Based, Parallel Assay for the Quantification of Protein-DNA Interactions. In F. Leng (Ed.), PLoS ONE (Vol. 9, Issue 2, p. e89626). Public Library of Science (PLoS). https://doi.org/10.1371/journal.pone.0089626

Lopes, R., & Prasad, M. K. (2024). Beyond the promise: evaluating and mitigating off-target effects in CRISPR gene editing for safer therapeutics. In Frontiers in Bioengineering and Biotechnology (Vol. 11). Frontiers Media SA. https://doi.org/10.3389/fbioe.2023.1339189

Lu, W.-T., Trost, C. N., Müller-Esparza, H., Randau, L., & Davidson, A. R. (2021). Anti-CRISPR AcrIF9 functions by inducing the CRISPR–Cas complex to bind DNA non-specifically. In Nucleic Acids Research (Vol. 49, Issue 6, pp. 3381–3393). Oxford University Press (OUP). https://doi.org/10.1093/nar/gkab092

Luthra, R., Kaur, S., & Bhandari, K. (2021). Applications of CRISPR as a potential therapeutic. In Life Sciences (Vol. 284, p. 119908). Elsevier BV. https://doi.org/10.1016/j.lfs.2021.119908

Mekler, V., Kuznedelov, K., & Severinov, K. (2020). Quantification of the affinities of CRISPR–Cas9 nucleases for cognate protospacer adjacent motif (PAM) sequences. In Journal of Biological Chemistry (Vol. 295, Issue 19, pp. 6509–6517). Elsevier BV. https://doi.org/10.1074/jbc.ra119.012239

Mekler, V., Kuznedelov, K., Minakhin, L., Murugan, K., Sashital, D. G., & Severinov, K. (2019). CRISPR–Cas molecular beacons as tool for studies of assembly of CRISPR–Cas effector complexes and their interactions with DNA. In Methods in Enzymology (pp. 337–363). Elsevier. https://doi.org/10.1016/bs.mie.2018.10.026

Mekler, V., Minakhin, L., & Severinov, K. (2017). Mechanism of duplex DNA destabilization by RNA-guided Cas9 nuclease during target interrogation. In Proceedings of the National Academy of Sciences (Vol. 114, Issue 21, pp. 5443–5448). Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.1619926114

Mekler, V., Minakhin, L., Semenova, E., Kuznedelov, K., & Severinov, K. (2016). Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3′-terminal segment of guide RNA. In Nucleic Acids Research (Vol. 44, Issue 6, pp. 2837–2845). Oxford University Press (OUP). https://doi.org/10.1093/nar/gkw138

Mikulic, M. (2024). CRISPR-Cas Gene Market Size Global 2033 Forecast. Statista. https://www.statista.com/statistics/950816/crispr-genome-editing-market-size-globally/#statisticContainer

Mohan Bangalore, D., & Tessmer, I. (2018). Unique insight into protein-DNA interactions from single molecule atomic force microscopy. In AIMS Biophysics (Vol. 5, Issue 3, pp. 194–216). American Institute of Mathematical Sciences (AIMS). https://doi.org/10.3934/biophy.2018.3.194

Müller-Esparza, H., Osorio-Valeriano, M., Steube, N., Thanbichler, M., & Randau, L. (2020). Bio-Layer Interferometry Analysis of the Target Binding Activity of CRISPR-Cas Effector Complexes. In Frontiers in Molecular Biosciences (Vol. 7). Frontiers Media SA. https://doi.org/10.3389/fmolb.2020.00098

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. In Cell (Vol. 156, Issue 5, pp. 935–949). Elsevier BV. https://doi.org/10.1016/j.cell.2014.02.001

Noom, M., van den Broek, B., van Mameren, J. et al. Visualizing single DNA-bound proteins using DNA as a scanning probe. Nat Methods 4, 1031–1036 (2007). https://doi.org/10.1038/nmeth1126

Nutiu, R., Friedman, R., Luo, S. et al. Direct measurement of DNA affinity landscapes on a high-throughput sequencing instrument. Nat Biotechnol 29, 659–664 (2011). https://doi-org.proxy-ub.rug.nl/10.1038/nbt.1882

O'Geen, H., Henry, I., Bhakta, M., Meckler, J., & Segal, D. (2014). A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. Nucleic Acids Research, 43, 3389 - 3404. https://doi.org/10.1093/nar/gkv137

Overacker, R.D., Plitzko, B. & Loesgen, S. Biolayer interferometry provides a robust method for detecting DNA binding small molecules in microbial extracts. Anal Bioanal Chem 413, 1159–1171 (2021). https://doi.org/10.1007/s00216-020-03079-5

Pacesa, M., Lin, C.-H., Cléry, A., Saha, A., Arantes, P. R., Bargsten, K., Irby, M. J., Allain, F. H.-T., Palermo, G., Cameron, P., Donohoue, P. D., & Jinek, M. (2022). Structural basis for Cas9 off-target activity. In Cell (Vol. 185, Issue 22, pp. 4067-4081.e21). Elsevier BV. https://doi.org/10.1016/j.cell.2022.09.026

Platt, R. J., Chen, S., Zhou, Y., Yim, M. J., Swiech, L., Kempton, H. R., Dahlman, J. E., Parnas, O., Eisenhaure, T. M., Jovanovic, M., Graham, D. B., Jhunjhunwala, S., Heidenreich, M., Xavier, R. J., Langer, R., Anderson, D. G., Hacohen, N., Regev, A., Feng, G., … Zhang, F. (2014). CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. In Cell (Vol. 159, Issue 2, pp. 440–455). Elsevier BV. https://doi.org/10.1016/j.cell.2014.09.014

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. In Cell (Vol. 163, Issue 4, pp. 854–865). Elsevier BV. https://doi.org/10.1016/j.cell.2015.10.003

Richardson, C., Ray, G., DeWitt, M. et al. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol 34, 339–344 (2016). https://doi.org/10.1038/nbt.3481

Ryan, D. E., Taussig, D., Steinfeld, I., Phadnis, S. M., Lunstad, B. D., Singh, M., Vuong, X., Okochi, K. D., McCaffrey, R., Olesiak, M., Roy, S., Yung, C. W., Curry, B., Sampson, J. R., Bruhn, L., & Dellinger, D. J. (2017). Improving CRISPR–Cas specificity with chemical

modifications in single-guide RNAs. In Nucleic Acids Research (Vol. 46, Issue 2, pp. 792–803). Oxford University Press (OUP). https://doi.org/10.1093/nar/gkx1199

Savic D, Partridge EC, Newberry KM, Smith SB, Meadows SK, Roberts BS, Mackiewicz M, Mendenhall EM, Myers RM. CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Res. 2015 Oct;25(10):1581-9. doi: 10.1101/gr.193540.115. Epub 2015 Sep 9. PMID: 26355004; PMCID: PMC4579343

Serajian S, Ahmadpour E, Oliveira SMR, Pereira MdL, Heidarzadeh S. CRISPR-Cas Technology: Emerging Applications in Clinical Microbiology and Infectious Diseases. Pharmaceuticals. 2021; 14(11):1171. https://doi.org/10.3390/ph14111171

Severin, P. M. D., & Gaub, H. E. (2012). DNA–Protein Binding Force Chip. In Small (Vol. 8, Issue 21, pp. 3269–3273). Wiley. https://doi.org/10.1002/smll.201201088

Severin, P. M. D., Ho, D., & Gaub, H. E. (2011). A high throughput molecular force assay for protein–DNA interactions. In Lab on a Chip (Vol. 11, Issue 5, p. 856). Royal Society of Chemistry (RSC). https://doi.org/10.1039/c0lc00302f

Shao M, Xu TR, Chen CS. The big bang of genome editing technology: development and application of the CRISPR/Cas9 system in disease animal models. Dongwuxue Yanjiu. 2016 Jul 18;37(4):191-204. doi: 10.13918/j.issn.2095-8137.2016.4.191. PMID: 27469250; PMCID: PMC4980067

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelson, T., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84–87. https://doi.org/10.1126/science.1247005

Stahelin, R. V. (2013). Surface plasmon resonance: a useful technique for cell biologists to characterize biomolecular interactions. In W. Bement (Ed.), Molecular Biology of the Cell (Vol. 24, Issue 7, pp. 883–886). American Society for Cell Biology (ASCB). https://doi.org/10.1091/mbc.e12-10-0713

Tome, J., Ozer, A., Pagano, J. et al. Comprehensive analysis of RNA-protein interactions by high-throughput sequencing–RNA affinity profiling. Nat Methods 11, 683–688 (2014). https://doi.org/10.1038/nmeth.2970

Tsai, S., Nguyen, N., Malagon-Lopez, J. et al. CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. Nat Methods 14, 607–614 (2017). https://doi-org.proxy-ub.rug.nl/10.1038/nmeth.4278

Wang, T., Wei, J. J., Sabatini, D. M., and Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80–84. https://doi.org/10.1126/science.1246981

Wu, X., Scott, D., Kriz, A. et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol 32, 670–676 (2014). https://doi.org/10.1038/nbt.2889

Xiaoshen Wang, Xuzichao Li, Yongjian Ma, Jiaqi He, Xiang Liu, Guimei Yu, Hang Yin, Heng Zhang, Inhibition mechanisms of CRISPR-Cas9 by AcrIIA17 and AcrIIA18, Nucleic Acids Research, Volume 50, Issue 1, 11 January 2022, Pages 512–521, https://doi.org/10.1093/nar/gkab1197

Yang, M., Peng, S., Sun, R., Lin, J., Wang, N., & Chen, C. (2018). The Conformational Dynamics of Cas9 Governing DNA Cleavage Are Revealed by Single-Molecule FRET. In Cell Reports (Vol. 22, Issue 2, pp. 372–382). Elsevier BV. https://doi.org/10.1016/j.celrep.2017.12.048

Yaron Orenstein, Ron Shamir, Modeling protein–DNA binding via high-throughput in vitro technologies, Briefings in Functional Genomics, Volume 16, Issue 3, May 2017, Pages 171–180, https://doi.org/10.1093/bfgp/elw030

Yu, J. S. L., & Yusa, K. (2019). Genome-wide CRISPR-Cas9 screening in mammalian cells. In Methods (Vols. 164–165, pp. 29–35). Elsevier BV. https://doi.org/10.1016/j.ymeth.2019.04.015

Zhang, L., Rube, H. T., Bussemaker, H. J., & Pufall, M. A. (2017). The effect of sequence mismatches on binding affinity and endonuclease activity are decoupled throughout the Cas9 binding site. Cold Spring Harbor Laboratory. https://doi.org/10.1101/176255

Zhao, D., Feng, X., Zhu, X. et al. CRISPR/Cas9-assisted gRNA-free one-step genome editing with no sequence limitations and improved targeting efficiency. Sci Rep 7, 16624 (2017). https://doi.org/10.1038/s41598-017-16998-8

Zheng, F., Chen, Z., Li, J., Wu, R., Zhang, B., Nie, G., Xie, Z., & Zhang, H. (2022). A Highly Sensitive CRISPR‐Empowered Surface Plasmon Resonance Sensor for Diagnosis of Inherited Diseases with Femtomolar‐Level Real‐Time Quantification. In Advanced Science (Vol. 9, Issue 14). Wiley. https://doi.org/10.1002/advs.202105231

Zhu, H., Li, C. & Gao, C. Applications of CRISPR–Cas in agriculture and plant biotechnology. Nat Rev Mol Cell Biol 21, 661–677 (2020). https://doi.org/10.1038/s41580-020-00288-9

Zykovich Artem, Korf Ian, Segal David J., Bind-n-Seq: high-throughput analysis of in vitro protein–DNA interactions using massively parallel sequencing, Nucleic Acids Research, Volume 37, Issue 22, 1 December 2009, Page e151, https://doi.org/10.1093/nar/gkp802

Afterword

In the course of this academic bachelor's thesis, AI tools were utilized with caution to enhance the literature research (ChatGPT, Perplexity.ai, consensus.app, elicit.com) and to offer contextual synonyms (ChatGPT).