Exploring Novel AcrIIA5 Variants as Anti-CRISPR therapy : Design and Artificial Intelligence based AlphaFold Prediction for Targeting a Spectrum of Cas9 Proteins

[1]Sahal Sabilil Muttaqin, [1]Qi Li, [2]Serafina Soehianto

[1]School of Biological Sciences, The University of Edinburgh
[2]Institute of Genetics and Cancer, The University of Edinburgh

Supervisor : Dr. Christopher W. Wood

Abstract

The CRISPR-Cas9 gene-editing system has revolutionized genome engineering. However, its extensive use necessitates measures to prevent unintended modifications to the genome. AcrIIA5 offers a natural approach to regulate CRISPR-Cas9, and leveraging AI-based AlphaFold Prediction can optimize its inhibitory potential. This innovative strategy aims to broaden AcrIIA5's effectiveness across Type II CRISPR classes, meeting the critical need for precise genome editing control and advancing our understanding of anti-CRISPR proteins. Our study utilized a multi-step bioinformatics approach to explore AcrIIA5's structural and functional aspects. In this study, it was discovered that the Alphafold Model N-Terminal intrinsically disordered region (IDR), versatile Cas9 inhibitor region, exhibits several mutations and have a low pLDDT around 30 to 60 score. The Alphafold model identifies AcrIIA5-5 as having an exceptionally low E value of of 8.35E-98, indicating its high compatibility as an experimental model with its given sequence. In docking, AcrIIA5 (6LKF) and AcrIIA5-10 scored strongly at -286.31 and -283.47 for spyCas9, while AcrIIA5 (6LKF) and AcrIIA5-3 achieved notable scores of -330.27 and -267.67 for st1Cas9. Similarly, AcrIIA5 (6LKF) and AcrIIA5-2 displayed promising scores around -302.74 and -290.35 for Nme1Cas9, confirming their strong binding affinities with their respective Cas9 receptor proteins. AcrIIA5 inhibits the RuvC domain in St1Cas9 and Nme1Cas9, while interacting with the REC and PAM-Interacting domain in spyCas9. Its adaptability relies on the associated receptor protein. Leveraging AlphaFold for screening AcrIIA5 variants offers a viable method to develop specific inhibitors for various anti-CRISPR therapeutic applications.

Introduction

The CRISPR-Cas9 gene-editing system, derived from bacteria and archaea, utilizes clustered, regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) genes genes (H. Wang et al., 2016). The system involves adaptation, expression, and interference phases, with Cas proteins recognizing a protospacer-adjacent motif (PAM) and facilitating targeted genome editing(Marraffini & Sontheimer, 2010; Westra et al., 2012). CRISPR-Cas9's simplicity and versatility have revolutionized gene editing, transcription regulation, genome imaging, and epigenetic modification (Ahmad & Amiji, 2018; Arabi et al., 2022). In recent development, the UK's Medicines and Health Care Products Regulatory Agency approved the world's first CRISPR/Cas9-based gene therapy in 2023 for sickle cell disease and beta-thalassemia (Wilkinson, 2023). The therapy, Casgevy, demonstrated significant efficacy in reducing pain crises and transfusion needs (Wilkinson, 2023).

However, the widespread adoption of CRISPR/Cas9 has led to an urgent demand for countermeasures to prevent unintended genome modifications(Piergentili et al., 2021). Among the potential solutions explored, AcrIIA5 stands out as a promising candidate. Its unique role in binding to the Cas9-sgRNA-DNA complex orchestrates a mechanism akin to a molecular brake, effectively inhibiting nuclease activity within the RuvC domain and preventing unforeseen alterations to the genome(Liang et al., 2020). Prevention of DNA binding and subsequent sgRNA cleavage by AcrIIA5 is proposed by *in vivo* study (Garcia et al., 2019). Nevertheless, *in vitro* experiments did not observe sgRNA degradation by AcrIIA5, suggesting the involvement of other nucleases for sgRNA degradation(An et al., 2020). Further investigations are needed to fully elucidate the precise mechanism of AcrIIA5-mediated Cas9 inhibition(An et al., 2020; Garcia et al., 2019). What sets AcrIIA5 apart is its distinctive capacity not only to bind but also to actively interfere with nuclease activity, introducing an additional layer of control that enhances the precision of genome editing processes (An et al., 2020; Liang et al., 2020). Moreover, AcrIIA5's mechanism reveals intriguing parallels with other anti-CRISPR strategies, suggesting a broader, nature-endorsed approach to regulating the formidable CRISPR-Cas9 system(An et al., 2020). Notably, AcrIIA5 emerges as a powerful tool in addressing off-target effects in CRISPR/Cas9-based cytosine and adenine base editors (CBEs and ABEs), efficiently suppressing the activity of base editing systems at different

ratios and demonstrating its potential to "shut off" nuclease activity while controlling base editing precision(An et al., 2020).

In this research, we utilise an artificial intelligence-based AlphaFold (Jumper et al., 2021) for refining AcrIIA5 variants using its structure prediction. This innovative approach aims to enhance our understanding of AcrIIA5's unique binding mechanisms, enabling the design of broader Type-II anti-CRISPRs. Anticipated outcomes include an augmented ability to regulate various CRISPR-Cas9, specifically SpyCas9, St1Cas9, and Nme1Cas9 (Müller et al., 2016; Nishimasu et al., 2015; Sun et al., 2019), expanding the utility of AcrIIA5. This research holds significance for refining gene-editing tools, ensuring heightened control and safety. The hypothesis asserts that leveraging AI predictions can visualize and optimize AcrIIA5's inhibitory potential, leading to designing new robust variants, effectively broadening its reach across diverse Type II CRISPR classes and addresses the critical need for precise control in genome editing technologies especially in therapeutic applications.

Methods

1. Retrieval of protein structure

The retrieval of protein structure involved acquiring the Anti-CRISPR AcrIIA5 structure (PDB : 6LKF), Cas9-sgRNA-AcrIIA4 anti-CRISPR complex (PDB : 5VZL), St1Cas9-sgRNA-tDNA20-AcrIIA6 monomeric assembly (PDB : 6RJ9), and Nme1Cas9-sgRNA-ACRIIC5 complex (PDB : 8HJ4) from the RCSB PDB website. The corresponding AcrIIA5 FASTA data was then extracted to perform protein analysis through blast analysis.

2. BLASTp Analysis

The AcrIIA5 sequence obtained from the RCSB PDB was subjected to BLASTp analysis to identify AcrIIA5 homologs. Based on the BLASTp analysis, a total of 67 sequences were selected as AcrIIA5 homologs. These homolog sequences were then aligned using the NCBI multiple sequence alignment tool. Additionally, pairwise alignment was performed on three Cas9 structures to determine the percentage of identity, identifying Cas9 sequence variations.

3. Selection of Top 10 AcrIIA5 Sequences and Create Phylogenetic Tree

The identification and selection of the top 10 AcrIIA5 sequences were carried out. Subsequently, a phylogenetic tree representing diverse branches was constructed based on these sequences using the MEGA software(Tamura et al., 2021). Further analysis will be conducted to interpret the data.

4. AlphaFold Protein Structure Prediction

The top 10 AcrIIA5 sequences were analysed into AlphaFold for protein structure prediction. top-ranked prediction was chosen and the structure in PDB format was downloaded. The noteworthy outputs generated by AlphaFold, including Percentage of Identity and E-value. We also measure predicted local distance difference test (pLDDT) score of notable residues in AcrIIA5's folded region to estimate of the extent to which the prediction aligns with an experimental structure (Mariani et al., 2013). Furthermore, we ensure the alignment of the first 20 residues of the N-Terminal Intrinsically Disorder Region (IDR) across the ligands in order to determine the number of variations in residues(An et al., 2020).

5. Protein Docking Using HDock

The HDOCK server utilizes a hybrid docking strategy to predict the binding complexes between two molecules, such as a receptor protein and a ligand protein. In this study, we employed the HDock online server to execute protein docking between the predicted AcrIIA5 structure and the chosen protein target(Yan et al., 2020). The docking calculations carried out by HDOCK encompass the determination of the docking score and confidence score. The Docking Score furnishes an assessment of the anticipated binding affinity between the molecules, thereby providing insights into the strength of the interaction. The confidence score denotes the dependability of the anticipated complex and signifies the level of confidence linked to the docking outcomes.These metrics play a crucial role in assessing the quality and reliability of the results obtained from molecular docking(Yan et al., 2017).

6. Visualize AcrIIA5 Alignment and AcrIIA5-Protein target Interaction

In our investigation, we employ a visual representation of the alignment of the AcrIIA5 structure through a comparative analysis of the experimental structure of AcrIIA5 and the models generated by Alphafold using Pymol. Secondly, the visualization of the top six ligand choices was carried out through the utilization of the PyMOL software(Lill & Danielson,

2011), enabling a comprehensive scrutiny of their interactions with the Cas9 proteins in three distinct species. The evaluation of the precision and dependability of the computational predictions was achieved by means of a comparative analysis between the original AcrIIA5 Protein Data Bank (PDB) structure and the AcrIIA5 structure predicted by AlphaFold. This examination aimed to offer valuable insights into the manners in which these ligands bind and the potential functional consequences thereof. Additionally, by comparing these interactions with those of the receptor original ligand, we sought to discern any similarities or differences in the way each ligand engages with the Cas9 protein.

Results and Discussion

1. AcrIIA5 Protein Alignment Analysis

BLAST searches identify local sequence similarities, with alignment extent determined by query coverage. Higher query coverage leads to a lower E-value, a statistical measure indicating the likelihood of a significant match. Lower E-values signify more meaningful alignments, with increased query coverage reducing the chance of a random alignment. (Newell et al., 2013). Percent identity in bioinformatics quantifies similarity between sequences, representing the percentage of identical residues. Higher values indicate greater similarity, implying a closer evolutionary or functional relationship, while lower values suggest more distant connections or functional disparities.(Newell et al., 2013).

No	Scientific name	Query Cover	Per Identity	E-value	Accession
1	hypothetical protein PP213_gp25 [Streptococcus phage CHPC877]	100%	100%	4,00E-95	Select seq ref|YP_010645967.1|
2	TPA: anti-CRISPR protein AcrIIA5 [Escherichia coli]	98%	100%	3,00E-93	HDG8206169.1
3	anti-CRISPR protein AcrIIA5 [Streptococcus phage Sfi21]	100%	95.71%	7,00E-91	Select seq ref|NP_049988.1|
4	anti-CRISPR protein AcrIIA5 [Streptococcus thermophilus]	100%	93.57%	3,00E-89	Select seq gb|MCE2227820.1|
5	anti-CRISPR protein AcrIIA5 [Streptococcus phage TP-778L]	100%	90.71%	2,00E-86	Select seq ref|YP_008772110.1|
6	anti-CRISPR protein AcrIIA5 [Streptococcus oralis]	100%	89.29%	5,00E-86	Select seq ref|WP_268699057.1|
7	anti-CRISPR protein AcrIIA5 [Streptococcus thermophilus]	100%	90.71%	5,00E-85	Select seq gb|MCE2320809.1|
8	anti-CRISPR protein AcrIIA5 [Streptococcus sp. WB01_FAA12]	100%	84.29%	3,00E-82	Select seq ref|WP_169447340.1|
9	AcrIIA5 [Streptococcus phage D1126]	100%	87.14%	1,00E-81	Select seq gb|AVO22762.1|
10	anti-CRISPR protein AcrIIA5 [Streptococcus mitis]	100%	82.86%	1,00E-80	WP_125448047.1

Table 1. Top 10 Anti-CRISPR Homologous AcrIIA5

We conducted a comprehensive analysis by subjecting a series of experimental AcrIIA5 sequences to the BLASTp algorithm, which allowed us to identify and compare 10 sequences that bear striking similarities. According to table 1, the hypothetical protein PP213_gp25 from Streptococcus phage CHPC877 exhibits the highest query cover and percent identity of 100% based on BLASTp analysis. It also has the lowest E-value of 4E-95. This suggests that the sequences have a perfect match with experimental structure of AcrIIA5 (PDB: 6LKF), underscoring a high degree of conservation. As a result, we will proceed to utilize the structure of 6LKF as the representative for this sequence. The AcrIIA5 protein derived from *Escherichia coli*, an anti-CRISPR protein, exhibits a strong resemblance to the query sequence, with a coverage of 98%, percent identity of 100%, and E-value of 3E-93. This observation emphasizes a remarkable level of similarity in the protein's sequence. Similarly, the anti-CRISPR proteins obtained from various Streptococcus phages and strains also showcase different degrees of similarity. The coverage percentages of these proteins range from 90% to 100%, while the identity percentages span from 82.86% to 95.71%. These results imply that although these proteins share significant similarities in their sequences, slight variations in identity could potentially indicate functional diversity or adaptation to distinct host environments(Romero & Arnold, 2009).

2. Comparison of Cas9 Pairwise Alignment

Pairwise Sequence Alignment, a fundamental bioinformatics method, is essential for precisely identifying and characterizing similarity regions in biological sequences. This tool provides valuable insights into functional, structural, and evolutionary relationships by comparing proteins or nucleic acids. It plays a crucial role in discerning conserved regions, indicating shared ancestry or analogous biological functions (Andrade et al., 2014).

In type II CRISPR-Cas systems, the Cas9 endonuclease protein forms a complex with either the tracrRNA:crRNA duplex or a designed single-guide RNA (sgRNA), resulting in the creation of the Cas9 ribonuclease protein (Cas9 RNP). This assembled complex is then directed to the target DNA, leading to the specific cleavage of double-stranded DNA. *Streptococcus pyogenes* Cas9 (SpyCas9), extensively studied within the type II-A CRISPR-Cas system, has been thoroughly characterized and widely applied in diverse biotechnological applications. In contrast, *Neisseria meningitidis* Cas9 (Nme1Cas9), the Cas9 homolog in the type II-C CRISPR-Cas system, is recognized for its minimal off-target effects compared to SpyCas9. Additionally, Nme1Cas9 is approximately 250 amino acids shorter than SpyCas9, simplifying its introduction into living cells. The Cas9 proteins sourced from *Streptococcus thermophilus* (St1Cas9) exhibit a distinct ability to recognize longer protospacer adjacent motifs (PAMs). Intriguingly, the utilization of St1Cas9, as opposed to the conventional SpyCas9, results in significantly diminished off-target mutagenesis under conditions of comparable on-target cleavage efficiency(Müller et al., 2016; Zhang & Marchisio, 2021).

In our research, we have effectively employed this methodological approach to investigate and analyze three distinct Cas9 proteins originating from the three different bacteria. The analysis of pairwise sequence alignment of Cas9 proteins has yielded valuable insights into the similarities of their sequences (Figure 1). The dataset presents three comparisons between distinct Cas9 proteins, each of which is identified by a unique Protein Cas9 ID. In the initial comparison, St1Cas9 protein (PDB : 6RJ9) was aligned with spyCas9 protein (PDB : 5VZL), revealing a sequence identity of 25%. The second comparison involved Nme1Cas9 protein (PDB : 8HJ4) against spyCas9 protein (PDB : 5VZL), resulting in a slightly higher identity of 27.38%. Notably, the third comparison between St1Cas9 protein (PDB : 6RJ9) and Nme1Cas9 protein (PDB : 8HJ4) demonstrated a comparable identity of 27.07%. These findings indicate a moderate level of sequence conservation among the examined Cas9 proteins. The observed identities suggest that these proteins share common sequence motifs, although with substantial variations. These variants may include amino acid substitution of critical domains as the primary factor and reliance on the presence of a strict PAM sequence(Allemailem et al., 2023).

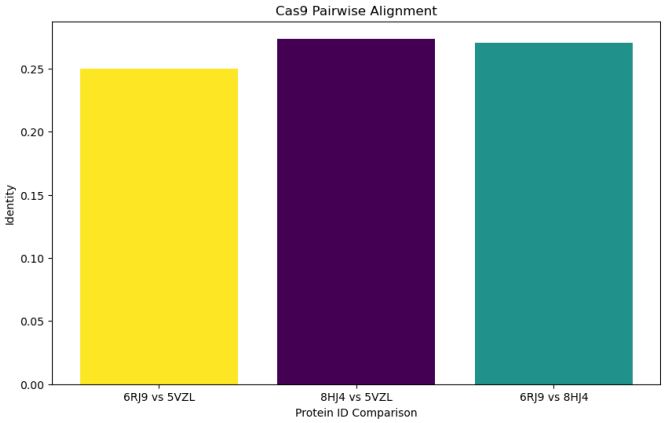


Figure 1. Cas9 Protein pairwise alignment

3. AcrIIA5 Phylogenetic Tree

We generated a phylogenetic tree using BLASTp reveals the intricate diversity of AcrIIA5 proteins across top 10 data in different species. The phylogenetic associations among AcrIIA5 orthologs become more evident in figure 2. Notably, the hypothetical protein PP213_gp25 from Streptococcus phage CHPC877 and the AcrIIA5 protein derived from *Escherichia coli* display the closest phylogenetic proximity, suggesting a potential shared evolutionary lineage(Hall,

2013). Additionally, the examination of anti-CRISPR proteins in Figure 2 uncovers noteworthy patterns. The phylogenetic clustering of anti-CRISPR protein AcrIIA5 from Streptococcus phage TP-778L with that from *Streptococcus thermophilus* implies a common evolutionary origin or conserved genetic elements. Similarly, the grouping of anti-CRISPR protein AcrIIA5 from *Streptococcus sp*. WB01_FAA12 and anti-CRISPR protein AcrIIA5 from *Streptococcus mitis*, exhibiting identical phylogenetic relationships, suggests shared evolutionary trajectories or functional convergence within this specific subset of AcrIIA5 proteins(Hall, 2013).

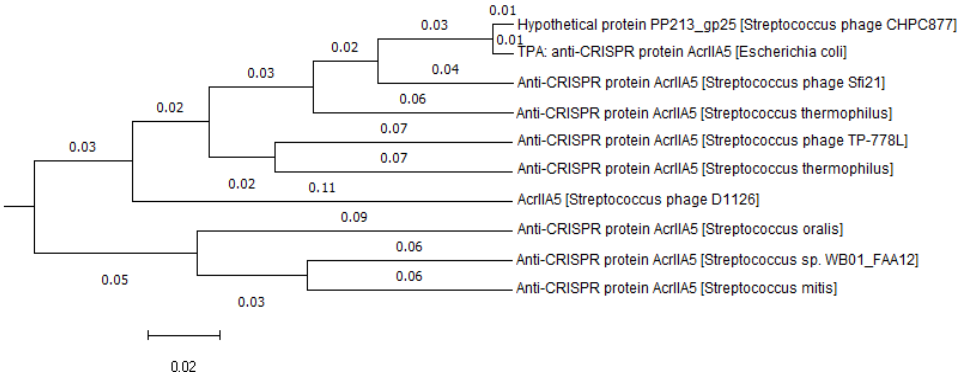
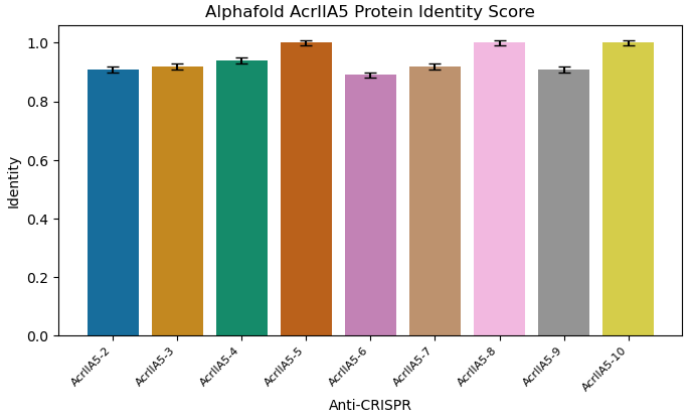


Figure 2. Phylogenetic Tree of Top 10 AcrIIA5

4. AlphaFold Data analysis



a

Anti-Crispr	E value
AcrIIA5-2	3,13E-88
AcrIIA5-3	3,54E-90
AcrIIA5-4	1,05E-91
AcrIIA5-5	8,35E-98
AcrIIA5-6	4,13E-89
AcrIIA5-7	5,02E-90
AcrIIA5-8	5,88E-98
AcrIIA5-9	4,83E-88
AcrIIA5-10	1,44E-98

b

Figure 3. Alphafold AcrIIA5 protein a) Identity score b) E-value score

The genetic engineering field has been significantly transformed by the introduction of the CRISPR-Cas system. Furthermore, the existence of anti-CRISPR proteins, such as AcrIIA5, introduces a level of intricacy to the governing regulatory mechanisms of this system(Zhang & Marchisio, 2021). For the purpose of this research, we utilize AlphaFold, an artificial intelligence tool for predicting AcrIIA5 structures based on sequences data from NCBI (Jumper et al., 2021), to investigate the projected structures of different variants of AcrIIA5. Our analysis reveals several intriguing findings. Notably, AcrIIA5-5, AcrIIA5-8, and AcrIIA5-10 exhibit perfect matches with 100% identity, indicating a high level of structural conservation. On the other hand, AcrIIA5-6, while not a perfect match, demonstrates a substantial 89% identity, suggesting a noteworthy similarity in its predicted structure.

Further analysis of a range of E values for different AcrIIA5 orthologs was generated by Alphafold explains worth noting that AcrIIA5-2, AcrIIA5-3, AcrIIA5-4, AcrIIA5-6, AcrIIA5-7, AcrIIA5-9, and AcrIIA5-10 all exhibit low E values, ranging from 3.13E-88 to 1.44E-98. This range implies robust and statistically significant similarities among these AcrIIA5 Anti-

CRISPR proteins. AcrIIA5-5, on the other hand, stands out with an even lower E value of 8.35E-98, emphasizing an exceptionally strong match, suggesting a heightened similarity between the query and the database hit(Newell et al., 2013).

5. Alphafold Modelling

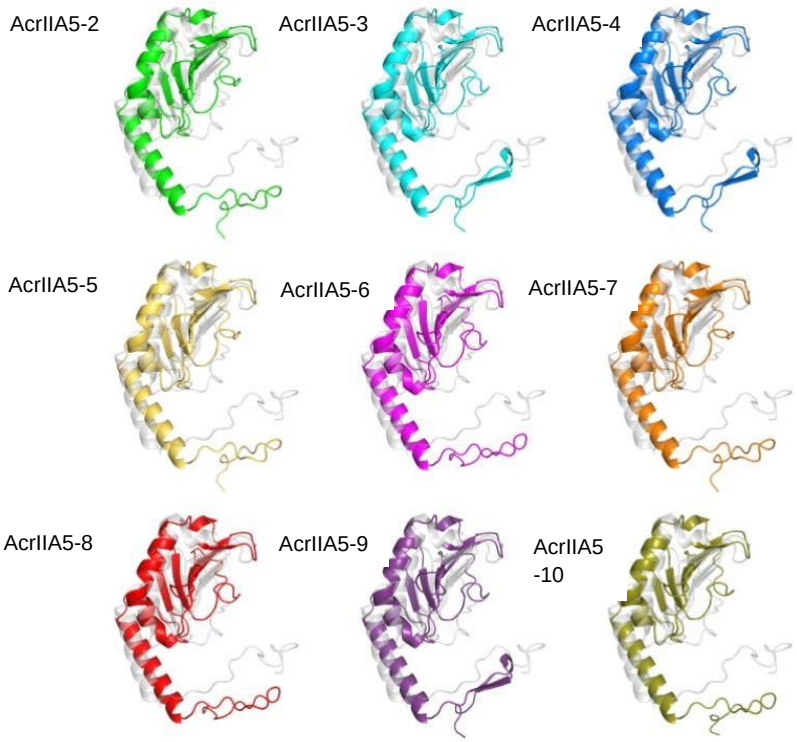


Figure 5. Structure alignment of AcrIIA5 and and its alphafold models

The Pymol alignment results reveal a notable similarity between the AcrIIA5 variants modeled by AlphaFold and the AcrIIA5 structure obtained from the Protein Data Bank (PDB) (Figure 5). The high degree of similarity observed through the alignment suggests that the AlphaFold models have successfully predicted AcrIIA5 structures with a level of accuracy comparable to experimental data.

The data in Figure 5 reveals diverse conformations of the N-terminal intrinsically disordered region (IDR) in AcrIIA5. This IDR's structural plasticity is associated with AcrIIA5's broad inhibitory range against different Cas9 homologs, including type II-A and II-C Cas9 homologs, and modest inhibition against type II-B Cas9. The IDR's inherent flexibility facilitates the formation of adaptable interaction surfaces, allowing AcrIIA5 to bind and inhibit Cas9 indiscriminately across various targets. AcrIIA5's broad-spectrum activity benefits phage survival in hosts with multiple Cas9 proteins. The suppression of various Cas9 targets hints at potential ambiguous complexes, with the IDR of AcrIIA5 remaining disordered even when bound to Cas9. This region plays a crucial role as a primary interface for inhibiting the nuclease activity of Cas9-sgRNA complex with complete inhibition of Cas9 requires an IDR of full length(An et al., 2020).

The N-terminal IDR with its 22 residues, especially the first 20, are crucial for inhibitory function. Removing the first 20 residues eliminates inhibition, and further truncations significantly reduce or abolish Acr activity. The optimal IDR length for maximum inhibition is at least 20 amino acid residues(An et al., 2020). It was observed that the predicted model produced by Alphafold exhibits a low pLDDT, ranging from 30 to 60 score, for the initial 1-18 residues of IDR. The pLDDT metric represents the model's estimation of its performance on the local Distance Difference Test (lDDT-Cα), which serves as an indicator of local accuracy(Tunyasuvunakool et al., 2021).

Further examination of the N-terminal intrinsically disordered region (IDR) (Supplementary Table 1) reveals that AcrIIA5-6, AcrIIA5-8, and AcrIIA5-10 exhibit a Y3F and S10A mutation. Subsequently, AcrIIA5-2, AcrIIA5-5, and

AcrIIA5-7 possess an R18I mutation. Additionally, all of the predicted proteins exhibit an N20D mutation. This perhaps stems from their inhibition of distinct Cas9 variants, resulting in varying degrees of amino acid flexibility as previously mentioned(An et al., 2020).

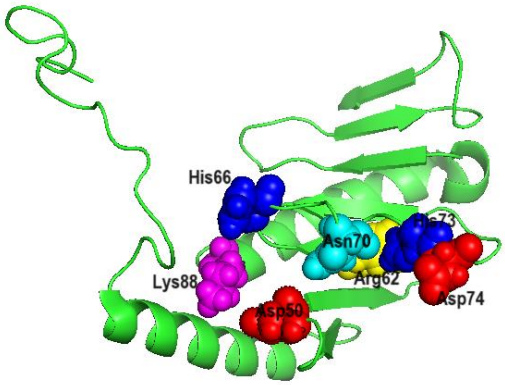


Figure 6. The specific residues that impact the in vivo functionality of AcrIIA5

	Amino acids PLDDT score						
Anti-CRISPR	Asp50	Arg62	His66	Asn70	His73	Asp74	Lys88
AcrIIA5-2	81,44	90,88	78,62	55,97	49,53	45,56	84,56
AcrIIA5-3	82,12	90,12	77,19	58,69	50,56	50,88	84,5
AcrIIA5-4	82,12	90,12	77,19	58,69	50,56	50,88	84,5
AcrIIA5-5	81,44	90,88	78,62	55,97	49,53	45,56	84,56
AcrIIA5-6	81,38	90,25	78,44	58	49,75	47,44	84,44
AcrIIA5-7	81,44	90,88	78,62	55,97	49,53	45,56	84,56
AcrIIA5-8	81,38	90,25	78,44	58	49,75	47,44	84,44
AcrIIA5-9	82,12	90,12	77,19	58,69	50,56	50,88	84,5
AcrIIA5-10	82,25	90,75	80,56	55,53	51,56	47,75	85,06

Table 2. The pLDDT scores of prominent residues on folded region

We assess the likelihood of the AcrIIA5 model acquiring a mutation that affects Cas9 inhibition based from previous research findings(Garcia et al., 2019). Based on the conducted alignment analysis, it has been observed that there is an absence of any mutation occurring within the residues that exert an impact on the functionality of AcrIIA5 variants (figure 6). These particular residues are located within the folded region, specifically in the β1– β2 and β3– β4 loop region, which are recognized for their conservation. More precisely, His66, Asn70, His73, and Asp74 are located in the area between the β3 and β4 loops. Asp50, Arg62, and Lys88, on the other hand, are placed in the β1–β2 loop, β3 strand, and β4–β2 loop, respectively(An et al., 2020). It suggests that the predicted model of AcrIIA5 will exhibits functional inhibition of Cas9.

Furthermore, we examine the pLDDT scores assigned to the spesific amino acids within the predicted AcrIIA5 proteins (Tunyasuvunakool et al., 2021), as shown in the table 2, reflect different levels of confidence in the accuracy of the structural predictions. Generally, these proteins exhibit a high degree of confidence, as evidenced by scores ranging from 81.38 to 82.25 for Asp50 and 90.12 to 90.88 for Arg62. The scores for His66, which range from 77.19 to 80.56, also tend to fall within the high-confidence range, except for AcrIIA5-10, which obtains an exceptionally high score. On the other hand, the confidence levels for Asn70, His73, and Asp74 are relatively lower, with scores ranging from 45.56 to 58.69, all falling into the low or very low categories. Nonetheless, Lys88 consistently demonstrates a remarkably high level of confidence, with scores ranging from 84.44 to 85.06. AcrIIA5-10 stands out due to its remarkably high pLDDT score for His66, distinguishing it from other AcrIIA5 proteins. These ratings reflect varying degrees of reliability in the structure predictions for the individual amino acids within the AcrIIA5 proteins(Tunyasuvunakool et al., 2021).

6. Docking Analysis

In the current study, we carried out an in-depth investigation by employing HDock webtool for docking analysis(Yan et al., 2020). To ensure comprehensive results, a total of 10 AcrIIA5 ligands were utilized for each Cas9 variant, leading to a remarkable accumulation of 100 protein models for each ligand. Consequently, an impressive total of 3000 proteins was obtained. From this extensive dataset, we carefully selected the two most promising ligands that exhibited a robust binding affinity for each Cas9 protein, culminating in a total of six exceptional candidates. Subsequently, to further enhance our understanding of the binding phenomenon, we employed PyMol, a powerful software tool, for visualizing the intricate binding interactions(Lill & Danielson, 2011). By employing this cutting-edge technology, we were able to gain valuable insights into the molecular interactions and spatial arrangements of the ligands and the respective Cas9 proteins, ultimately shedding light on the underlying mechanisms governing their binding behavior.

In the spyCas9 docking, our selection process concentrated on two specific ligands, AcrIIA5 (PDB ID: 6LKF) and AcrIIA5-10. These alternatives manifested noteworthy docking scores of -286.31 and -283.47, respectively. Furthermore, the confidence scores affiliated with these selections were documented as 0.9352 and 0.9386, respectively (Supplementary Table 2). Subsequently, we elected to prioritize AcrIIA5 (6LKF) and AcrIIA5-3 based on their binding affinities with st1Cas9. These choices exhibited docking scores of -330.27 and -267.67, respectively, with corresponding confidence scores of 0.9735 and 0.9180 (Supplementary Table 3). Notably, the exclusion of AcrIIA6 as a ligand in this specific context was warranted due to its propensity to bind with AcrIIA6 rather than the Cas9 protein. Ultimately, our selected Nme1Cas9 inhibitors were AcrIIA5 (6LKF) and AcrIIA5-2, demonstrating highly favorable docking scores of approximately -302.74 and -290.35, respectively. Their associated confidence scores were determined as 0.9550 and 0.9431 (Supplementary Table 4). Based on the data provided, it can be observed that the Alphafold modelling approach exhibits a slightly lower performance in both docking and confidence score parameters when compared to the original experimental structure.

7. Visualization of AcrIIA5-Cas9 Interaction

Streptococcus pyogenesis Cas9

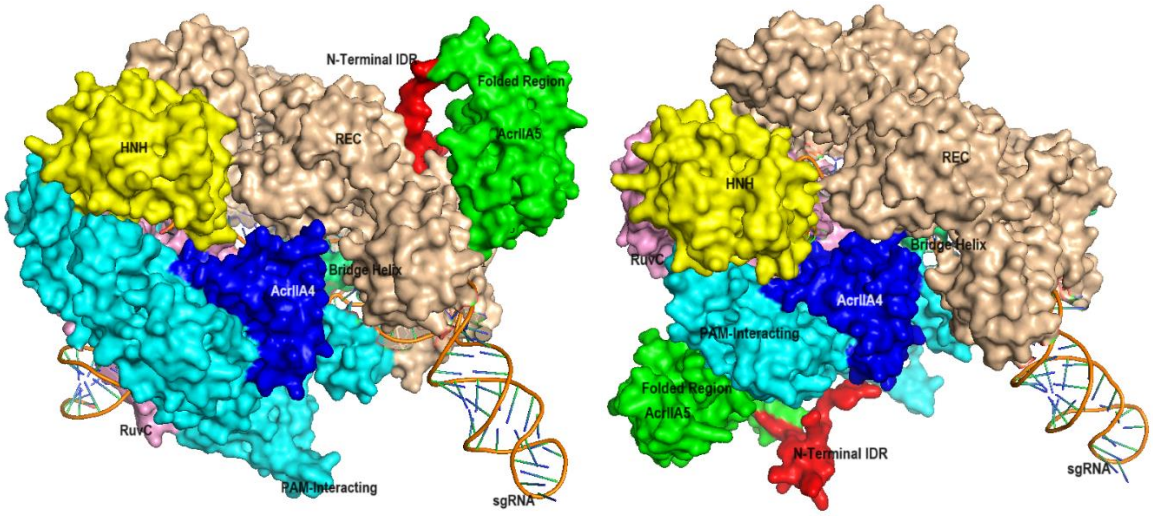


Figure 7. SpyCas9-sgRNA-AcrIIA4 anti-CRISPR complex with With a) AcrIIA5 (6LKF) and b) AcrIIA5-10.

Cas9 is an intricate and elaborate protein with a multitude of domains, which can be primarily classified into two prominent lobes, namely the alpha-helical recognition (REC) lobe and the nuclease (NUC) lobe. The REC lobe plays a crucial role in identifying and attaching itself to the crRNA, demonstrating its inherent ability to recognize and bind to this vital molecule. Conversely, the NUC lobe encompasses various domains that contribute to its functional complexity. Notably, the PAM-interacting domain (PID) within the NUC lobe exhibits a remarkable propensity to interact with the protospacer adjacent motif (PAM) present on the target DNA, thereby establishing a critical link between Cas9 and its designated DNA target. Furthermore, this remarkable lobe houses the HNH and RuvC endonuclease domains, both of which play instrumental roles in the pivotal process of cleaving the target DNA, thus enabling Cas9 to fulfill its intricate genetic editing functions. Overall, the multifaceted architecture of Cas9 highlights its remarkable adaptability and versatility in recognizing, binding, and ultimately cleaving target DNA sequences, thereby underscoring its indispensable role in contemporary genetic engineering endeavors(Hwang et al., 2023). The REC1-3 domains'

collective conformational dynamics are essential for aiding the recognition and response to nucleic acid binding events. The complex interaction of structural changes allows for the accurate control of the HNH conformational transition, which is necessary for the effective cutting of target DNA. In addition, the REC1–3 domains have the function of actively stabilizing and securing the HNH at the cleavage site, essentially immobilizing it. The combination of many factors leads to the creation of a fully operational and catalytically active CRISPR-Cas9 complex(Palermo et al., 2018).

The REC lobe is composed of three regions which are the bridge helix (limegreen), along with the REC1 and REC2 areas (wheat color). The NUC lobe consists of the RuvC (pink), HNH (yellow), and PAM-interacting (cyan) domains(Nishimasu et al., 2015). The AcrIIA4 protein (blue) binds specifically to the crIIA4 region at the protospacer adjacent motif (PAM)–interacting cleft, which is situated between the recognition (REC) lobe and the nuclease lobe. The PAM binding pocket is entirely occupied by the AcrIIA4 molecule, thus preventing DNA recognition through interactions established between the beta-3 strand of AcrIIA4 and the Cas9 PAM-binding residues (Figure 7a-b)(Shin et al., 2017).

The analysis based on AcrIIA5 (6LKF) modelling (Figure 7a – green and red color) has revealed that AcrIIA5 also establishes a connection with the REC lobe. It is worth noting that the REC lobe displays significant variability in sequence across different Cas9 homologs. The REC lobe is known to play a vital role in the dimerization induced by AcrIIA5. Considering that the region on the surface of Cas9, engaged by AcrIIA5, is not essential for its functionality, it is highly probable that the inhibition of DNA binding activity can be attributed to the dimerization facilitated by AcrIIA5. It is important to acknowledge that there might be vital functional surfaces hidden within the Cas9 dimer interface. Additionally, essential conformational changes could potentially be impeded as well. These factors could contribute to the observed inhibition of DNA binding activity(Davidson et al., 2020). In this case, AcrIIA5 attaches to the REC lobe, inhibiting the crucial conformational alterations in Cas9 required for its endonuclease activity. However, in the Alphafold model of AcrIIA5-10 (Figure 7b – green and red color), this protein functions by inhibiting PAM-interacting, hence preventing Cas9 from cleaving DNA while still allowing DNA binding. This model is being a same result as the experimental discovery(Hynes et al., 2017).

Furthermore, The PAM duplex attaches to a positively charged cleft on the C-terminal PAM-interacting domain(Anders et al., 2014). The binding of AcrIIA5-10 (Figure 7b) to the carboxyl-terminal domain (CTD), the nuclease lobe suggests that this specific domain is likely involved in the recognition of distinct guide RNA and PAM motifs (Jiang & Doudna, 2015). Additionally, the C-terminal domain (CTD) plays a crucial role in facilitating the binding of Cas and its subsequent degradation(Johnson, 2020). It is suggested that blocking the activity of the CTD could potentially disrupt the recognition process of RNA and PAM motifs. Consequently, inhibiting the functionality of the CTD may have implications for the overall efficiency and accuracy of the recognition mechanism of CRISPR-Cas9.

Neisseria meningitidis

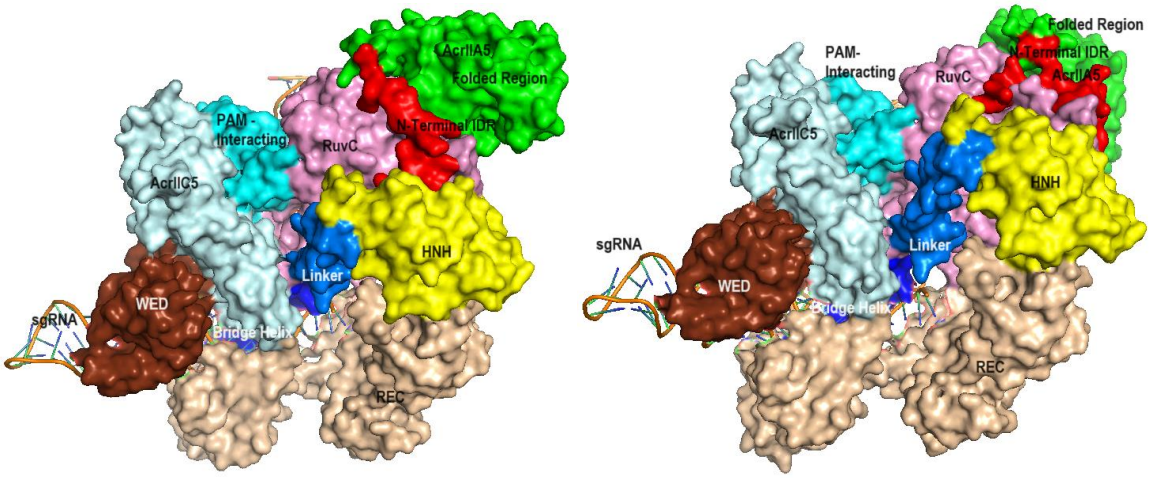


Figure 8. Nme1Cas9-AcrIIC2-sgRNA complex with a)AcrIIA5 (6LKF) b) AcrIIA5-2

Nme1Cas9 is composed of two lobes: the NUC lobe and the REC lobe (Figure 8a-b). The NUC lobe has distinct regions such as RuvC-I (Pink), Bridge Helix (lightblue), RuvC-II (pink), RuvC-III (pink), Linker 1 and Linker 2 (Blue), HNH (Yellow), WED (Chocolate), and PAM-Interacting (cyan). On the other hand, the REC lobe consists of REC1 and REC2 (wheat color)(Sun et al., 2019). AcrIIC5 (bluewhite) is inserted into the narrow space between the WED and PI domains of Cas9. One end of the anti-CRISPR molecule interacts with the phosphate lock loop and a connecting region between the RuvC and BH domains (Figure 8a-b).. AcrIIC5 interacts with type II-C Cas9 at the interface between the WED and PI domains. It imitates the structure and electrical charge of a DNA double helix, effectively occupying the DNA binding pocket and hiding the PAM recognition site. Therefore, AcrIIC5 renders Nme1Cas9 inactive by obstructing its ability to attach to DNA(Sun et al., 2023). Unexpectedly, the inhibitory effects of AcrIIA5 (6LKF) (as depicted in Figure 8a, with green and red) are observed on both the RuvC and HNH domains of Cas9. Meanwhile, the existing literature depicts wherein the AcrIIA5 molecule is known to impede the functioning of the RuvC nuclease domain of Cas9 independently from its effects on the HNH nuclease domain(Song et al., 2019).

Moreover, AcrIIA5-2 (Figure 8b) bind RuvC independently without HNH on different orientation to AcrIIA5 (6LKF). The Cas9 RuvC domain specifically cleaves the strand of the target DNA that is not complimentary(Nishimasu et al., 2014). Both AcrIIA5 inhibits the RuvC domains by preventing DNA recognition, binding, unwinding, and cleavage. Additionally, it suppresses the function of the RuvC domain of Cas9, regardless of the HNH domain(Choudhary et al., 2023).

Streptococcus thermophilus

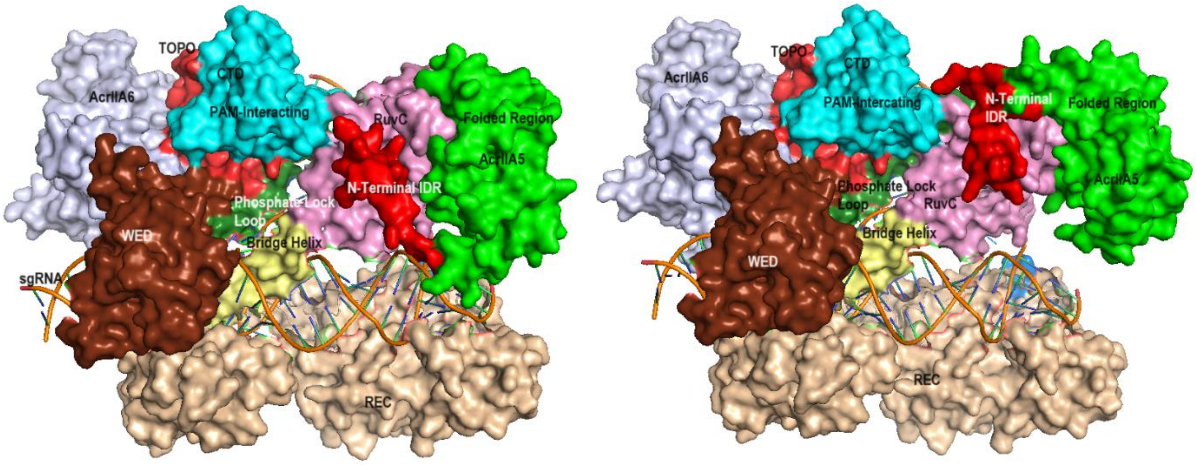


Figure 9. St1Cas9-AcrIIA6-sgRNA complex with a) AcrIIA5 (6LKF) b) AcrIIA5-3

The St1Cas9 protein consists of two lobes which are a recognition (REC) lobe and a nuclease (NUC) lobe (Figure 9a-b). The two lobes are joined by a bridge helix abundant in arginine (yellow) and a linker region that is incompletely resolved in the cryo-EM density maps. The HNH domain is not completely discernible this structure. The RuvC-like domain (pink) as a model reconstruction are not fully visible in 3D structure. The phosphate lock loop (forest color) serves to connect the nuclease domains with the a/b wedge domain (WED – chocolate color). The C-terminal region of St1Cas9 comprises the topoisomerase-homology (TOPO) domain (red) and C-terminal domains (CTD) (cyan). AcrIIA6 has the capacity to obstruct St1Cas9 through allosteric inhibition and also has the ability to promote the formation of St1Cas9 dimers notably in PAM-Interacting domain. Its mechanism of action involves modifying the dynamics of St1Cas9 in relation to PAM binding. Moreover, AcrIIA6 diminishes the affinity of St1Cas9 for DNA binding, resulting in the inhibition of DNA binding within cells (Fuchsbauer et al., 2019).

It is worth noting that the PI domain can be further divided into two distinct parts, namely the Topoisomerase-homology (TOPO) domain and the C-terminal domain, similar to the structure found in SpCas9 (Nishimasu et al., 2015). The PAM helix, which plays a crucial role in the CRISPR-Cas9 system, is positioned within a groove that is positively charged. This groove is formed between the C-terminal domain (CTD) and the Topo-homology domain, which are collectively known as the PID(Tsui & Li, 2015). The protein known as Cas9 possesses a protein called the Topo domain, which exhibits homology with topoisomerases, a class of enzymes involved in regulating the topology of DNA. The vital role of topoisomerases lies in their ability to manage the supercoiling of DNA during essential processes such as replication

and transcription(Y. Wang et al., 2019). The TOPO domain present in Cas9 potentially shares functional similarities with topoisomerases, although its specific role within the context of Cas9 remains to be fully elucidated.

Based on this visualization of the docking process (Figure 9a-b), it can be observed that AcrIIA5 (6LKF) engages in interactions and forms bindings with both RuvC and DNA in a clockwise direction. On the other hand, AcrIIA5-3 only exhibits bindings with RuvC, but in a counterclockwise direction. It is highly advantageous for AcrIIA5 to attain a clockwise position as this enables it to effectively bind to the essential structure of st1Cas9. In a broader perspective, AcrIIA5 exhibits the potential to impede the activity of RuvC in both Nme1Cas9 and St1Cas9, whereas in the case of SpyCas9, it presents two possibilities, specifically the obstruction of REC and PAM-Interacting. This intricate interplay between AcrIIA5 and Cas9 highlights the diverse mechanisms by which AcrIIA5 exerts its inhibitory effects, thereby shedding light on the complexity of the CRISPR-Cas system.

Future Prospective

AcrIIA5 has the potential to be used as a form of phage therapy due to its capability to specifically target different types of antimicrobial resistant (AMR) bacteria, serving as a substitute for antibiotics(Choudhary et al., 2023). Nevertheless, in order to put this therapy into practice, it is necessary to provide evidence of its feasibility and effectiveness. Moreover, the anti-CRISPR exhibits potential in reducing off-target effects by disrupting the CRISPR system. Recent research indicates that AcrIIA5 has the ability to efficiently suppress multiple base editing mechanisms and minimize unintended effects in human cells. Base editing, a technique that combines dCas9 with a cytidine deaminase, has demonstrated high efficacy in rectifying certain genetic variants associated with diseases, while avoiding the introduction of detrimental double-strand breaks in DNA. Nevertheless, it has been discovered that it produces unintended single-nucleotide variations in rice and mouse embryos. By employing AcrIIA5 to regulate base editing events, the enzymatic activity can be deactivated, hence diminishing off-target consequences (Liang et al., 2020). This mechanism has the potential to function as a deterrent for unintended modifications, especially in CRISPR-Cas gene therapies that provide treatment for diverse genetic disorders (Davidson et al., 2020). The method of administering these medicines needs to be determined, either by utilizing lipid nanoparticles (LNP) or Adeno-Associated Virus (AAV) as delivery mechanisms (Vyas & Harish, 2022). Gene drive systems possess the capacity to effectively combat human diseases, such as insect-borne diseases like malaria(Jia & Patel, 2021). Nevertheless, considering the capacity to genetically modify entire species using gene drive technology, it is imperative to implement precautions to ensure the effective implementation of gene drives. Although AcrIIA2 and AcrIIA4 have demonstrated strong inhibition of gene drive in *Saccharomyces cerevisiae* by deactivating Cas9 (Basgall et al., 2018), further investigation is required to assess the effectiveness of AcrIIA5 in suppressing gene drive when paired with Cas9 RNP to form Cas9-AcrIIA5 systems. Lastly, the target range of AcrIIA5 can be expanded and its effectiveness can be optimized by employing a directed evolution method in its design. Prior research demonstrates that the validation of the Acr functional selects method enables the targeted development of developed SpyCas9 inhibitors through the utilization of a highly varied phage mutagenesis pool(Jingrui (Priscilla) Wang, 2020). In the end, it is necessary to validate all AlphaFold models by doing real experimental research in order to obtain anti-CRISPR based therapy in various application above.

Conclusion

One can use an artificial intelligence method to examine several AcrIIA5 orthologs and determine the most potent anti-CRISPR. In this study, we have discovered that the N-Terminal intrinsically disordered region (IDR) of the Alphafold model displays several mutations. Nevertheless, this region has the capacity to impede diverse receptor proteins, making it a versatile region with inherent advantages. The AlphaFold model identifies AcrIIA5-5 as having an exceptionally low E value, indicating its great compatibility as an experimental model with its given sequence. The chosen ligands for this study have demonstrated outstanding docking scores and confidence scores, confirming their strong binding affinities with their respective receptor protein (Cas9). These ligands hinder the activity of the RuvC domain in St1Cas9 and Nme1Cas9, but they interact with the REC and PAM-Interacting domain in spyCas9. It can be inferred that the adaptability of AcrIIA5 is contingent upon the receptor protein it associates with. In conclusion, the utilization of Alphafold to optimize the screening of AcrIIA5 variant structures offers a viable method to acquire specific inhibitors for diverse anti-CRISPR therapeutics and applications.

References

Ahmad, G., & Amiji, M. (2018). Use of CRISPR/Cas9 gene-editing tools for developing models in drug discovery. In *Drug Discovery Today* (Vol. 23, Issue 3). https://doi.org/10.1016/j.drudis.2018.01.014

Allemailem, K. S., Almatroodi, S. A., Almatroudi, A., Alrumaihi, F., Al Abdulmonem, W., Al-Megrin, W. A. I., Aljamaan, A. N., Rahmani, A. H., & Khan, A. A. (2023). Recent Advances in Genome-Editing Technology with CRISPR/Cas9 Variants and Stimuli-Responsive Targeting Approaches within Tumor Cells: A Future Perspective of Cancer Management. In *International Journal of Molecular Sciences* (Vol. 24, Issue 8). https://doi.org/10.3390/ijms24087052

An, S. Y., Ka, D., Kim, I., Kim, E. H., Kim, N. K., Bae, E., & Suh, J. Y. (2020). Intrinsic disorder is essential for Cas9 inhibition of anti-CRISPR AcrIIA5. *Nucleic Acids Research*, *48*(13). https://doi.org/10.1093/nar/gkaa512

Anders, C., Niewoehner, O., Duerst, A., & Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*, *513*(7519). https://doi.org/10.1038/nature13579

Andrade, H., Area, I., Nieto, J. J., & Torres, Á. (2014). The number of reduced alignments between two DNA sequences. *BMC Bioinformatics*, *15*(1). https://doi.org/10.1186/1471-2105-15-94

Arabi, F., Mansouri, V., & Ahmadbeigi, N. (2022). Gene therapy clinical trials, where do we go? An overview. In *Biomedicine and Pharmacotherapy* (Vol. 153). https://doi.org/10.1016/j.biopha.2022.113324

Basgall, E. M., Goetting, S. C., Goeckel, M. E., Giersch, R. M., Roggenkamp, E., Schrock, M. N., Halloran, M., & Finnigan, G. C. (2018). Gene drive inhibition by the anti-CRISPR proteins AcrIIA2 and AcrIIA4 in saccharomyces cerevisiae. *Microbiology (United Kingdom)*, *164*(4). https://doi.org/10.1099/mic.0.000635

Choudhary, N., Tandi, D., Verma, R. K., Yadav, V. K., Dhingra, N., Ghosh, T., Choudhary, M., Gaur, R. K., Abdellatif, M. H., Gacem, A., Eltayeb, L. B., Alqahtani, M. S., Yadav, K. K., & Jeon, B. H. (2023). A comprehensive appraisal of mechanism of anti-CRISPR proteins: an advanced genome editor to amend the CRISPR gene editing. In *Frontiers in Plant Science* (Vol. 14). Frontiers Media SA. https://doi.org/10.3389/fpls.2023.1164461

Davidson, A. R., Davidson, A. R., Lu, W. T., Stanley, S. Y., Wang, J., Mejdani, M., Trost, C. N., Hicks, B. T., Lee, J., & Sontheimer, E. J. (2020). Anti-CRISPRs: Protein Inhibitors of CRISPR-Cas Systems. In *Annual Review of Biochemistry* (Vol. 89). https://doi.org/10.1146/annurev-biochem-011420-111224

Fuchsbauer, O., Swuec, P., Zimberger, C., Amigues, B., Levesque, S., Agudelo, D., Duringer, A., Chaves-Sanjuan, A., Spinelli, S., Rousseau, G. M., Velimirovic, M., Bolognesi, M., Roussel, A., Cambillau, C., Moineau, S., Doyon, Y., & Goulet, A. (2019). Cas9 Allosteric Inhibition by the Anti-CRISPR Protein AcrIIA6. *Molecular Cell*, *76*(6), 922-937.e7. https://doi.org/10.1016/j.molcel.2019.09.012

Garcia, B., Lee, J., Edraki, A., Hidalgo-Reyes, Y., Erwood, S., Mir, A., Trost, C. N., Seroussi, U., Stanley, S. Y., Cohn, R. D., Claycomb, J. M., Sontheimer, E. J., Maxwell, K. L., & Davidson, A. R. (2019). Anti-CRISPR AcrIIA5 Potently Inhibits All Cas9 Homologs Used for Genome Editing. *Cell Reports*, *29*(7). https://doi.org/10.1016/j.celrep.2019.10.017

Hall, B. G. (2013). Building phylogenetic trees from molecular data with MEGA. *Molecular Biology and Evolution*, *30*(5). https://doi.org/10.1093/molbev/mst012

Hwang, S., Shah, M., Garcia, B., Hashem, N., Davidson, A. R., Moraes, T. F., & Maxwell, K. L. (2023). Anti-CRISPR Protein AcrIIC5 Inhibits CRISPR-Cas9 by Occupying the Target DNA Binding Pocket. *Journal of Molecular Biology*, *435*(7). https://doi.org/10.1016/j.jmb.2023.167991

Hynes, A. P., Rousseau, G. M., Lemay, M. L., Horvath, P., Romero, D. A., Fremaux, C., & Moineau, S. (2017). An anti-CRISPR from a virulent streptococcal phage inhibits Streptococcus pyogenes Cas9. *Nature Microbiology*, *2*(10), 1374–1380. https://doi.org/10.1038/s41564-017-0004-7

Jia, N., & Patel, D. J. (2021). Structure-based functional mechanisms and biotechnology applications of anti-CRISPR proteins. In *Nature Reviews Molecular Cell Biology* (Vol. 22, Issue 8, pp. 563–579). Nature Research. https://doi.org/10.1038/s41580-021-00371-9

Jiang, F., & Doudna, J. A. (2015). The structural biology of CRISPR-Cas systems. In *Current Opinion in Structural Biology* (Vol. 30). https://doi.org/10.1016/j.sbi.2015.02.002

Jingrui (Priscilla) Wang. (2020). *A Method for the Directed Evolution and Protein Engineering of Synthetic CRISPR-Cas9 Inhibitors*. University of Toronto.

Johnson, W. E. (2020). Not Your Typical Anti-CRISPR. In *Cell Host and Microbe* (Vol. 28, Issue 1). https://doi.org/10.1016/j.chom.2020.06.016

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873). https://doi.org/10.1038/s41586-021-03819-2

Liang, M., Sui, T., Liu, Z., Chen, M., Liu, H., Shan, H., Lai, L., & Li, Z. (2020). AcrIIA5 Suppresses Base Editors and Reduces Their Off-Target Effects. *Cells*, *9*(8). https://doi.org/10.3390/cells9081786

Lill, M. A., & Danielson, M. L. (2011). Computer-aided drug design platform using PyMOL. *Journal of Computer-Aided Molecular Design*, *25*(1). https://doi.org/10.1007/s10822-010-9395-8

Mariani, V., Biasini, M., Barbato, A., & Schwede, T. (2013). lDDT: A local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, *29*(21). https://doi.org/10.1093/bioinformatics/btt473

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

Müller, M., Lee, C. M., Gasiunas, G., Davis, T. H., Cradick, T. J., Siksnys, V., Bao, G., Cathomen, T., & Mussolino, C. (2016). Streptococcus thermophilus CRISPR-Cas9 systems enable specific editing of the human genome. *Molecular Therapy*, *24*(3). https://doi.org/10.1038/mt.2015.218

Newell, P. D., Fricker, A. D., Roco, C. A., Chandrangsu, P., & Merkel, S. M. (2013). A Small-Group Activity Introducing the Use and Interpretation of BLAST. *Journal of Microbiology & Biology Education*, *14*(2). https://doi.org/10.1128/jmbe.v14i2.637

Nishimasu, H., Cong, L., Yan, W. X., Ran, F. A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., & Nureki, O. (2015). Crystal Structure of Staphylococcus aureus Cas9. *Cell*, *162*(5). https://doi.org/10.1016/j.cell.2015.08.007

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). https://doi.org/10.1016/j.cell.2014.02.001

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*. https://doi.org/10.1017/S0033583518000070

Piergentili, R., Del Rio, A., Signore, F., Umani Ronchi, F., Marinelli, E., & Zaami, S. (2021). Crispr-cas and its wide-ranging applications: From human genome editing to environmental implications, technical limitations, hazards and bioethical issues. In *Cells* (Vol. 10, Issue 5). https://doi.org/10.3390/cells10050969

Romero, P. A., & Arnold, F. H. (2009). Exploring protein fitness landscapes by directed evolution. In *Nature Reviews Molecular Cell Biology* (Vol. 10, Issue 12). https://doi.org/10.1038/nrm2805

Shin, J., Jiang, F., Liu, J. J., Bray, N. L., Rauch, B. J., Baik, S. H., Nogales, E., Bondy-Denomy, J., Corn, J. E., & Doudna, J. A. (2017). Disabling Cas9 by an anti-CRISPR DNA mimic. *Science Advances*, *3*(7). https://doi.org/10.1126/sciadv.1701620

Song, G., Zhang, F., Zhang, X., Gao, X., Zhu, X., Fan, D., & Tian, Y. (2019). AcrIIA5 Inhibits a Broad Range of Cas9 Orthologs by Preventing DNA Target Cleavage. *Cell Reports*, *29*(9). https://doi.org/10.1016/j.celrep.2019.10.078

Sun, W., Yang, J., Cheng, Z., Amrani, N., Liu, C., Wang, K., Ibraheim, R., Edraki, A., Huang, X., Wang, M., Wang, J., Liu, L., Sheng, G., Yang, Y., Lou, J., Sontheimer, E. J., & Wang, Y. (2019). Structures of Neisseria meningitidis Cas9 Complexes in Catalytically Poised and Anti-CRISPR-Inhibited States. *Molecular Cell*, *76*(6). https://doi.org/10.1016/j.molcel.2019.09.025

Sun, W., Zhao, X., Wang, J., Yang, X., Cheng, Z., Liu, S., Wang, J., Sheng, G., & Wang, Y. (2023). Anti-CRISPR AcrIIC5 is a dsDNA mimic that inhibits type II-C Cas9 effectors by blocking PAM recognition. *Nucleic Acids Research*, *51*(4). https://doi.org/10.1093/nar/gkad052

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7). https://doi.org/10.1093/molbev/msab120

Tsui, T. K. M., & Li, H. (2015). Structure Principles of CRISPR-Cas Surveillance and Effector Complexes. *Annual Review of Biophysics*, *44*. https://doi.org/10.1146/annurev-biophys-060414-033939

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R.,

Kohl, S. A. A., … Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, *596*(7873). https://doi.org/10.1038/s41586-021-03828-1

Vyas, P., & Harish. (2022). Anti-CRISPR proteins as a therapeutic agent against drug-resistant bacteria. In *Microbiological Research* (Vol. 257). Elsevier GmbH. https://doi.org/10.1016/j.micres.2022.126963

Wang, H., La Russa, M., & Qi, L. S. (2016). CRISPR/Cas9 in Genome Editing and beyond. *Annual Review of Biochemistry*, *85*. https://doi.org/10.1146/annurev-biochem-060815-014607

Wang, Y., Rakela, S., Chambers, J. W., Hua, Z. C., Muller, M. T., Nitiss, J. L., Tse-Dinh, Y. C., & Leng, F. (2019). Kinetic Study of DNA Topoisomerases by Supercoiling-Dependent Fluorescence Quenching. *ACS Omega*, *4*(19). https://doi.org/10.1021/acsomega.9b02676

Westra, E. R., Swarts, D. C., Staals, R. H. J., Jore, M. M., Brouns, S. J. J., & Van Der Oost, J. (2012). The CRISPRs, they are A-Changin': How prokaryotes generate adaptive immunity. In *Annual Review of Genetics* (Vol. 46). https://doi.org/10.1146/annurev-genet-110711-155447

Wilkinson, E. (2023). UK regulator approves "groundbreaking" gene treatment for sickle cell and β thalassaemia. *BMJ (Clinical Research Ed.)*, *383*, p2706. https://doi.org/10.1136/bmj.p2706

Yan, Y., Tao, H., He, J., & Huang, S. Y. (2020). The HDOCK server for integrated protein–protein docking. *Nature Protocols*, *15*(5). https://doi.org/10.1038/s41596-020-0312-x

Yan, Y., Zhang, D., Zhou, P., Li, B., & Huang, S. Y. (2017). HDOCK: A web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Research*, *45*(W1). https://doi.org/10.1093/nar/gkx407

Zhang, Y., & Marchisio, M. A. (2021). Type II anti-CRISPR proteins as a new tool for synthetic biology. In *RNA Biology* (Vol. 18, Issue 8). https://doi.org/10.1080/15476286.2020.1827803