In Silico Discovery of Potential Novel CRISPR-Cas12a Inhibitors

Bethany Miller, PharmD Candidate and Maria Kontoyianni, PhD

Southern Illinois University Edwardsville School of Pharmacy

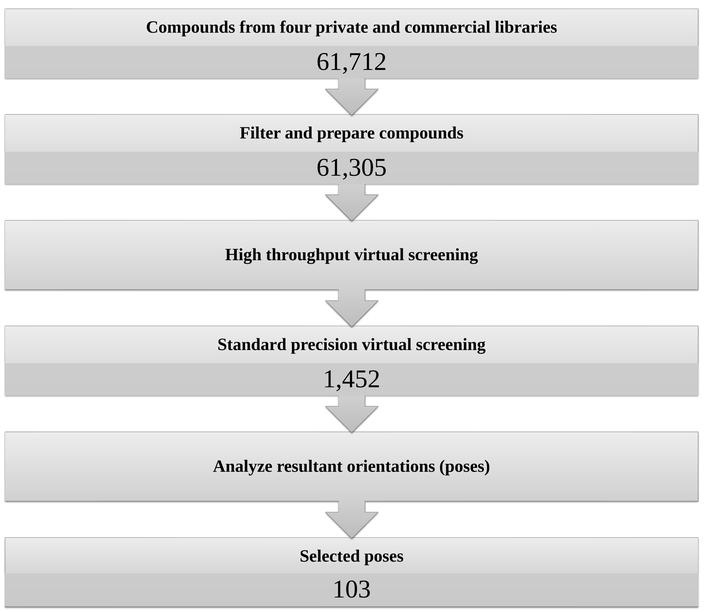
Background

CRISPR technology is becoming increasingly popular for use in gene editing, but the precise control of CRISPR-Cas systems is critical for it to be both successful and safe. Two classes of CRISPR-Cas exist: Class 1 and class 2, which are further divided into *types* characterized by distinct organization of the effectors and unique signature proteins. Class 2 includes the prevalent type II, with Cas 9 being its effector, whereas Cas 12(a-c) and Cas 13 are the effectors for the rare types V and VI. Cas 9 and Cas 12 proteins are homologous to nucleases, and share a domain that belongs to the RuvC-like endonuclease family. Further, the structures of several Cas9 proteins, Cas12a (Cpf1) and Cas12b (C2c1) have been resolved complexed with guide RNA, target DNA, and trascRNA. Even though CRISPR-Cas9 was the first CRISPR tool to be utilized for gene editing, CRISPR-Cas12a is believed to be more precise.[1] In addition, small molecule inhibitors of Cas9 have been identified,[2] whereas nothing has been reported for Cas12a to date.

External and internal controls preventing CRISPR systems from causing unintended off-site effects exist; however, internal controls are limited by the challenges of genetic engineering. In contrast, external controls can be relatively straightforward through the discovery of small molecule inhibitors.[3] Natural antiCRISPR (Acr) proteins inhibit Cas effectors, thus functioning as a defense mechanism against the CRISPR-Cas system.[4] The purpose of this study was to identify inhibitors of CRISPR-Cas12a that would result in similar actions as the Acr proteins.

Methods

Using virtual screening, we targeted the same binding site as that of the Acr proteins on Acidaminococcus sp. Cpf1 (AsCpf1), a CRISPR-Cas12a species. A receptor grid was generated with an inner box (defining where the ligand center can move) of 10 Å, and an outer box (where all of the ligand atoms must be contained) of 20 Å. The grid was defined by residues Asn178, Gly783, and Lys607. The antiCRISPR protein AcrVa1 acts on these conserved amino acids in the WED and REC domains of CRISPR Cas12a, with Asn178 lying in the REC domain and Gly783 in the WED domain. Lys 607, found in the PI domain, was also used in the receptor grid generation, as it has been found to be critical for PAM recognition[4] (see Figure 1). Private and commercially available compound libraries were first filtered using Rapid Elimination of Swill (REOS)[5] filters, and then prepared with LigPrep to produce accurate, low energy 3-dimensional conformations. Subsequently, a two-step Glide docking (high throughput, followed by standard precision) workflow was employed in order to identify hits based on energetics and visual inspections.



Results

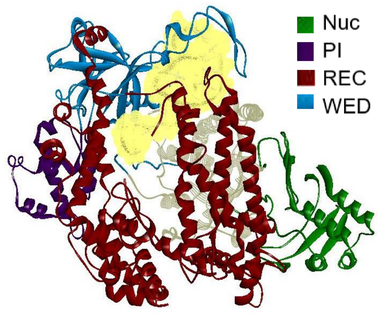
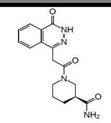
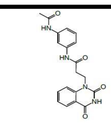
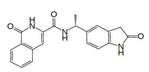
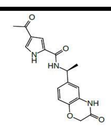
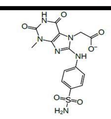
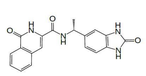
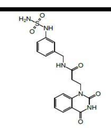


Figure 1. Depiction of the AsCpf1 structure employed in this study. The PAM-interacting (PI) domain plays an important role in PAM recognition, and the WED domain is reported to also interact with the PAM sequence. The REC lobe recognizes the crRNA-DNA heteroduplex, and the Nuc domain is responsible for DNA cleavage. Shown in yellow is the binding site for anti-CRISPR protein AcrVa1.

Table 1. Representative compounds identified as potential hits by the virtual screening experiments of one of the libraries (Enamine DNA)

Compound	Docking Score[a]	Rank #[b]
	-7.70983	3
	-7.07486	11
	-6.95448	20
	-6.85315	29
	-6.84771	30
	-6.83884	31
	-6.78191	38

[a]The docking score evaluates the free energy of binding and ensures that unfavorable but energetically accessible protonation and tautomeric states are penalized accordingly. [b]Relative to the lowest energy pose.

A.

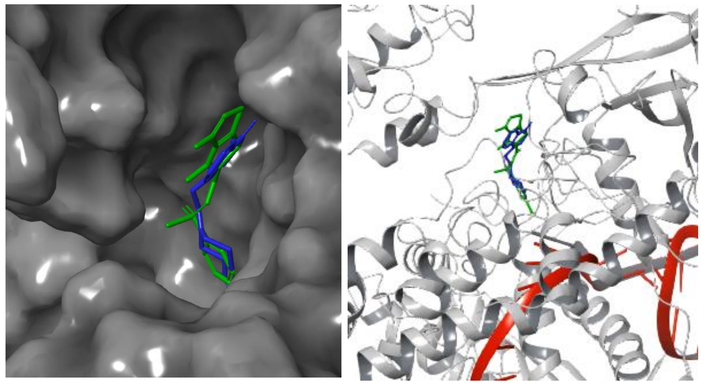
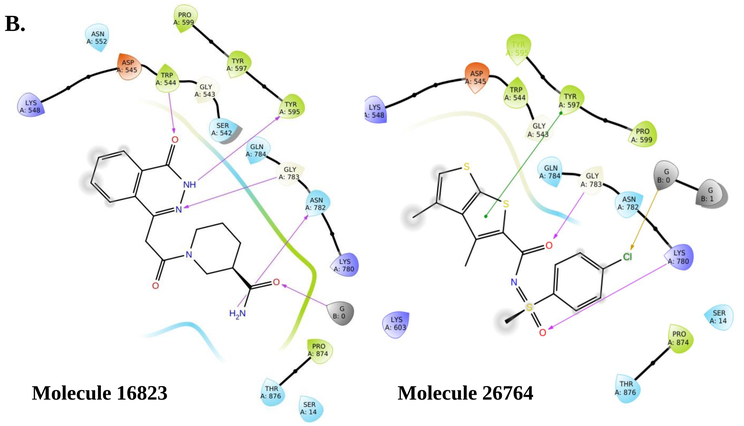


Figure 2. A. 3-Dimensional complexes of two representative hits within the AcCpf1 binding pocket. B. Corresponding 2D diagrams with respective interactions of the same hits.

B.



Molecule 16823 **Molecule 26764**

Discussion

Hits were selected using as criteria the presence of interactions with the residues discussed in Methods, along with energetics (docking score). Target-ligand complexes of two identified hits are depicted in Figures 2A and B. Specifically, compound 16823 is observed to hydrogen bond with Gly783, Trp544, Tyr595, and with the CRISPR guide RNA molecule. The backbone of Gly783 acts as a hydrogen bond donor to the nitrogen atom on the phthalazinone ring, while the amide forms a hydrogen bond with the RNA. In compound 26764, the chloro-substituent forms a halogen bond with RNA, the thioenothiophene group forms weak hydrophobic interactions with tryptophan, and the protein's glycine backbone is predicted to hydrogen bond with the amide group of the molecule. Out of 61,712 starting compounds, 103 potential hits were identified. However, it is desirable for a CRISPR inhibitor to be able to penetrate the Central Nervous System. Consequently, 19 out of these 103 hits have properties within ranges (see below) deemed necessary for blood brain barrier penetration.

Properties	Optimal Ranges for CNS
Molecular weight	<450
cLog P	Between 1.5 and 2.7
HBD	≤ 3
HBA	≤ 7
PSA A^2	< 60 - 70
# rot bond	< 8
pK$_a$	7.5-10.5

1. Swarts, D.C.; Jinek M., Cas9 versus Cas12a/Cpf1: Structure-function comparisons and implications for genome editing. *WIREs RNA*, **2018**, 9:e1481.
2. Maji, B.; Gangopadhyay, S. A.; Lee, M.; Clemons, P. A.; Wagner, B. K.; Choudhary, A., A High-throughput platform to identify small-molecule inhibitors of CRISPR-Cas9. *Cell* **2019,** *177, 1067-1079*
3. Tian, B.; Minero, G.; Fock, J.; Dufva, M.; Hansen, M.; CRISPR-Cas12a based internal negative control for nonspecific products of exponential rolling circle amplification, *Nucleic Acids* **2020,** *48*, (5), e30.
4. Zhang, H.; Li, Z.; Daczkowski, C. M.; Gabel, C.; Mesecar, A. D.; Chang, L., Structural Basis for the Inhibition of CRISPR-Cas12a by Anti-CRISPR Proteins. *Cell Host Microbe* **2019**, *25*, (6), 815-826 e4.
5. Walters, W. P.; Stahl, M. T.; Murcko, M. A., Virtual screening—an overview. *Drug Discovery Today* **1998**, *3*, 18.