

Study on relaxing the PAM recognition requirements of Cas9 nuclease

放松Cas9核酸酶PAM识别要求的研究

G09: 李彦娇、任晓朴、刘拓宇、赵月磊

汇报人: 任晓朴

生物技术研究所

2022.01.23

Study on relaxing the PAM recognition requirements of Cas9 nuclease

放松Cas9核酸酶PAM识别要求的研究

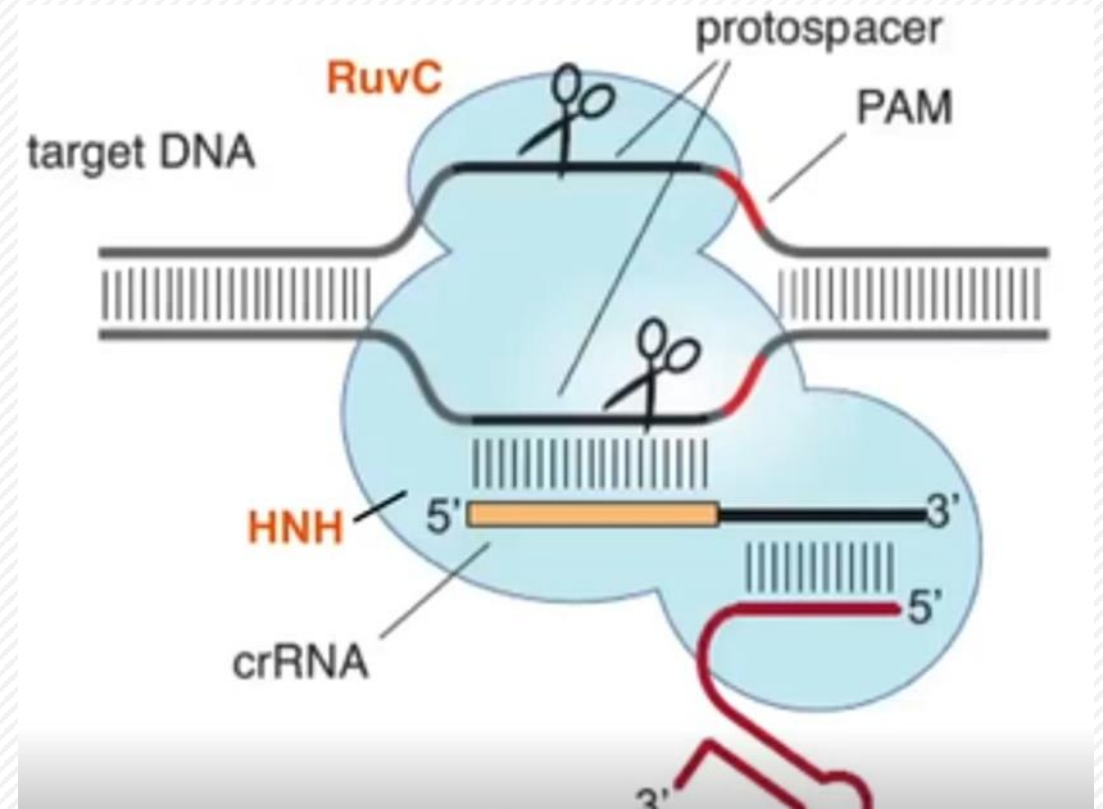
➤ 背景

➤ 天然同源基因挖掘

➤ 蛋白质工程

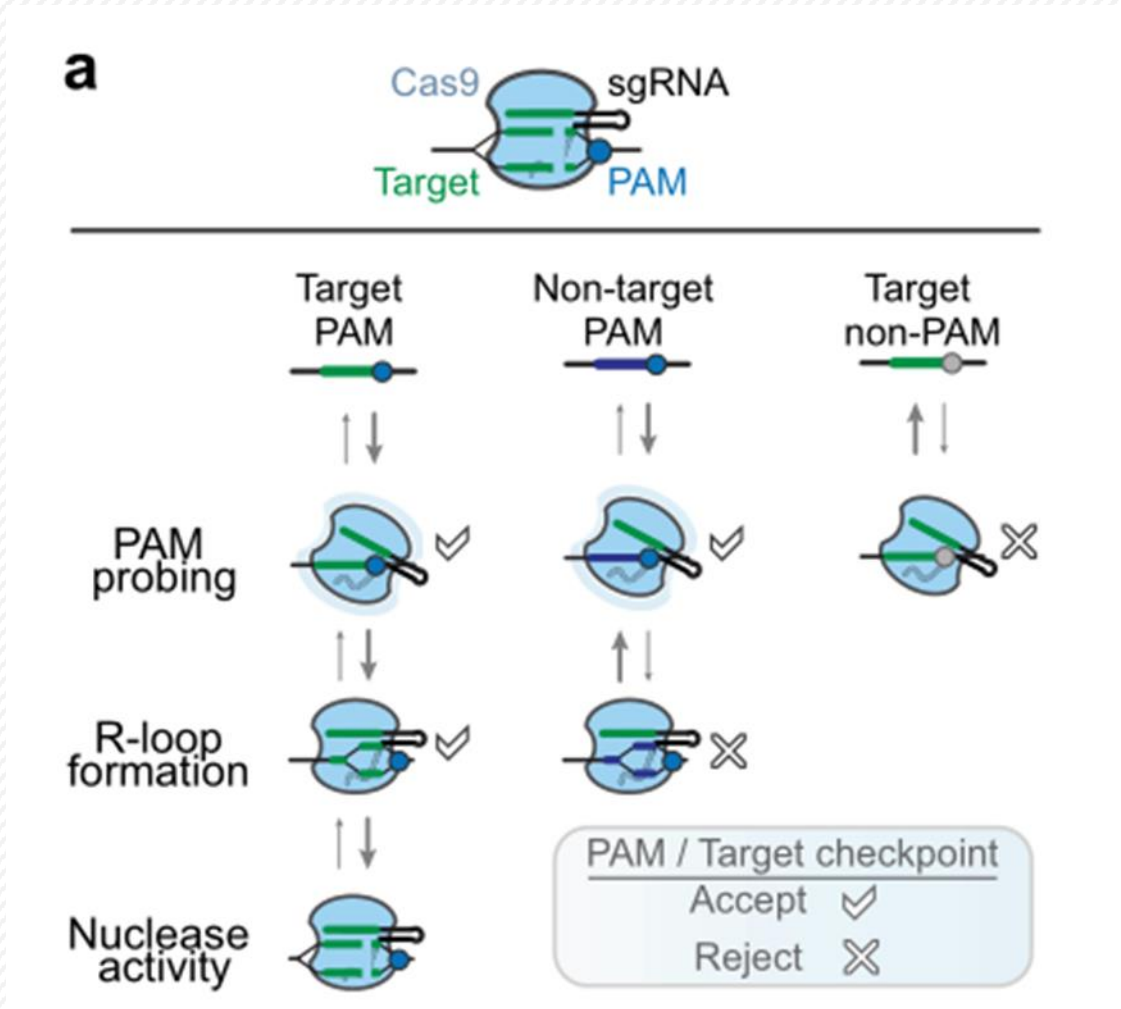
一、背景

➤ CRISPR-Cas9核酸酶能够在各种各样的生物体和细胞类型中进行有效的基因组编辑，Cas9的识别靶位点是由单导RNA (sgRNA)控制的，sgRNA可以与目标原间隔序列互补，在此之前也需要识别一个短的特定原间隔邻近基序 (**PAM**)。化脓性链球菌Cas9 (SpyCas9)是迄今为止应用最广泛、功能最强的Cas9，主要识别NGG PAM。



一、背景

➤ 1. Cas成功靶向的要素



成功的靶向需要两个要素：

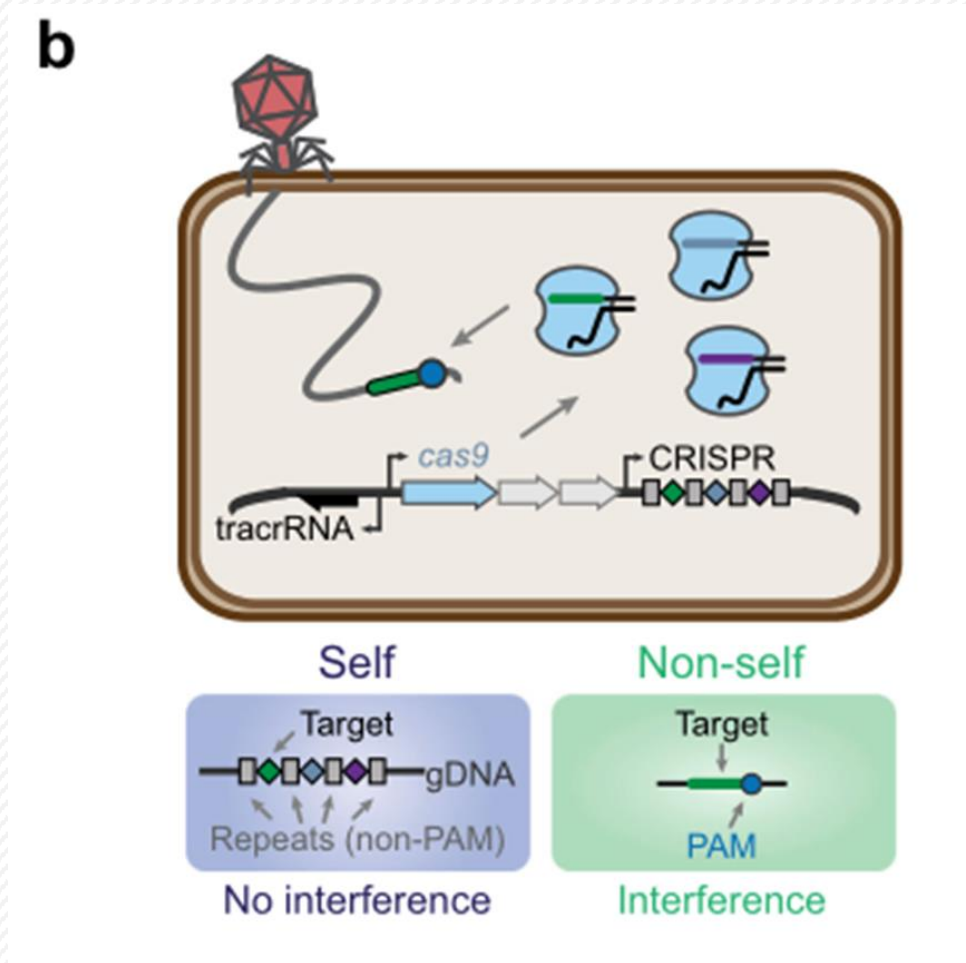
1. 靶点序列旁边的短序列，通常称为原间隔邻近基序（PAM）；

2. sgRNA和靶序列之间的广泛互补性。

与sgRNA完全互补但缺少PAM的序列将被核酸酶忽略。因此，PAM充当CRISPR-Cas靶向的把关。

一、背景

➤ 2.PAM的关键性



PAM在CRISPR-Cas系统的自然功能中起着至关重要的作用。PAM使这些原核免疫系统能够区分外来遗传物质（**非自身**）中的DNA靶点和CRISPR阵列（**自身**）中编码的相同DNA序列（产生sgRNA）。如果没有PAM要求，CRISPR-Cas系统将以其CRISPR阵列为目标，**导致潜在的灾难性自身免疫反应**。这一要求限制了用CRISPR靶向任何序列的能力，并努力来放松PAM要求，甚至几乎任何序列都可以被识别为PAM。

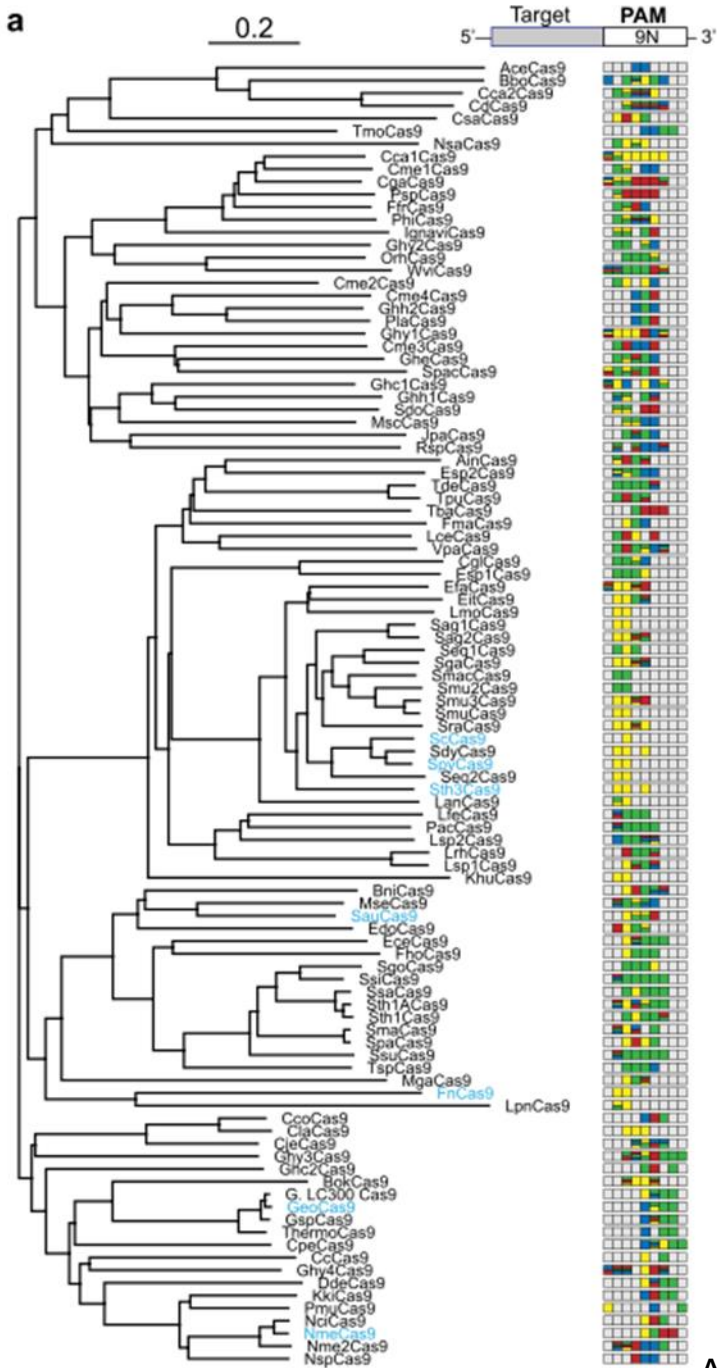
二、天然同源基因挖掘

- ➔ 现在已经了解到不同类型的cas蛋白及其功能，但是在所有CRISPR效应蛋白中，特征较好的蛋白是Cas9。用以调节基因表达、控制转录、染色体成像、药物递送，甚至治疗人类眼科和血液病，也是未来的前沿。尽管这项技术在药物开发方面有很大的前景，疗效和长期安全性仍然是临床试验中的主要问题，为了提高靶向特异性，鉴定CRISPR/Cas9同源物。

The screenshot shows the UniProtKB search interface. The search bar contains the query 'name:cas9 AND reviewed:yes'. The results page is titled 'UniProtKB 2021_04 results' and shows 14 results. A red box highlights the 'Filter by' section, which includes 'Reviewed (14) Swiss-Prot'. Below the filter, a table of results is visible, with the first three rows highlighted. The table columns are: Entry, Entry name, Protein names, Gene names, Organism, and Length.

Entry	Entry name	Protein names	Gene names	Organism	Length
G3ECR1	CAS9_STRTR	CRISPR-associated endonuclease Cas9	cas9 csn1	Streptococcus thermophilus	1,409
A0Q5Y3	CAS9_FRATN	CRISPR-associated endonuclease Cas9	cas9 FTN_0757	Francisella tularensis subsp. novicida (strain U112)	1,629
J7RUAS	CAS9_STAAU	CRISPR-associated endonuclease Cas9	cas9	Staphylococcus aureus	1,053

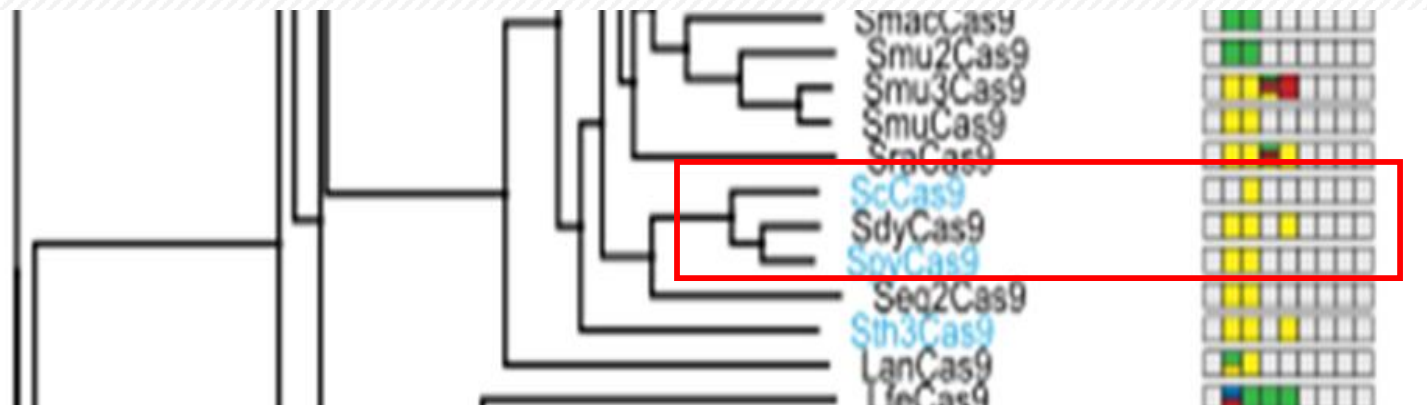
Gasiunas, G. et al. A catalogue of biochemically diverse CRISPR-Cas9 orthologs. Nat. Commun. 11, 5512 (2020).



二、天然同源基因挖掘

- ➔ SpyCas9 NGG PAM (N=any base)
- ➔ Sth1Cas9 NNATAAW PAM (W=A, T)
- ➔ SauCas9 NNGRRT PAM (R=A, G)
- ➔ NmeCas9 NNNNGATT PAM.

➔ 到目前为止，已经在测序的基因组和元基因组中鉴定出900多个不同的Cas9同源物，更多的同源物可能等待着进一步的测序工作来发现，它们具有不同的PAM谱、蛋白质大小和最适的活性温度。



A=绿色, C=蓝色, G=黄色, T=红色

Collias D, Beisel CL. CRISPR technologies and the search for the PAM-free nuclease. Nat Commun. 2021 Jan 22;12(1):555.

三、蛋白质工程

Family & Domainsⁱ

Domains and Repeats

Feature key	Position(s)	Description	Actions	Graphical view	Length
Domain ⁱ	770 – 921	HNH Cas9-type PROSITE-ProRule annotation	Add BLAST		152

Region

Feature key	Position(s)	Description	Actions	Graphical view	Length
Region ⁱ	1 – 62	RuvC-I 1 Publication	Add BLAST		62
Region ⁱ	56 – 718	Recognition lobe 1 Publication	Add BLAST		663
Region ⁱ	56 – 73	ARM 1 Publication	Add BLAST		18
Region ⁱ	718 – 765	RuvC-II 1 Publication	Add BLAST		48
Region ⁱ	925 – 1102	RuvC-III 1 Publication	Add BLAST		178
Region ⁱ	1099 – 1368	PAM-interacting domain (PI) 1 Publication	Add BLAST		270

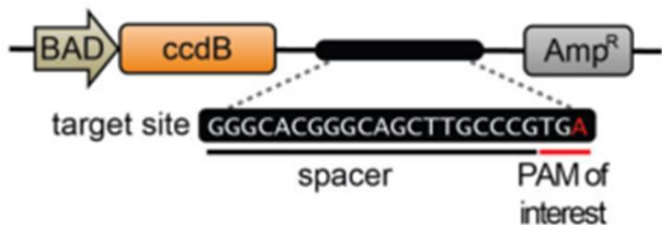
序列号: Q99ZW2

三、蛋白质工程

Library of Cas9/sgrNA plasmids with randomly mutated PI domains



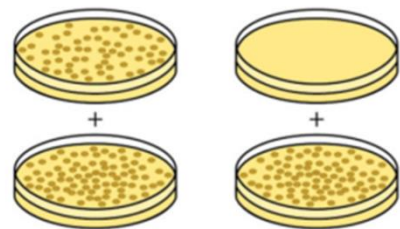
Selection plasmid with PAM of interest



1) transform Cas9/sgrNA plasmid into cells harboring positive selection plasmid:

Cas9/sgrNA cuts selection plasmid Cas9/sgrNA does not cut selection plasmid

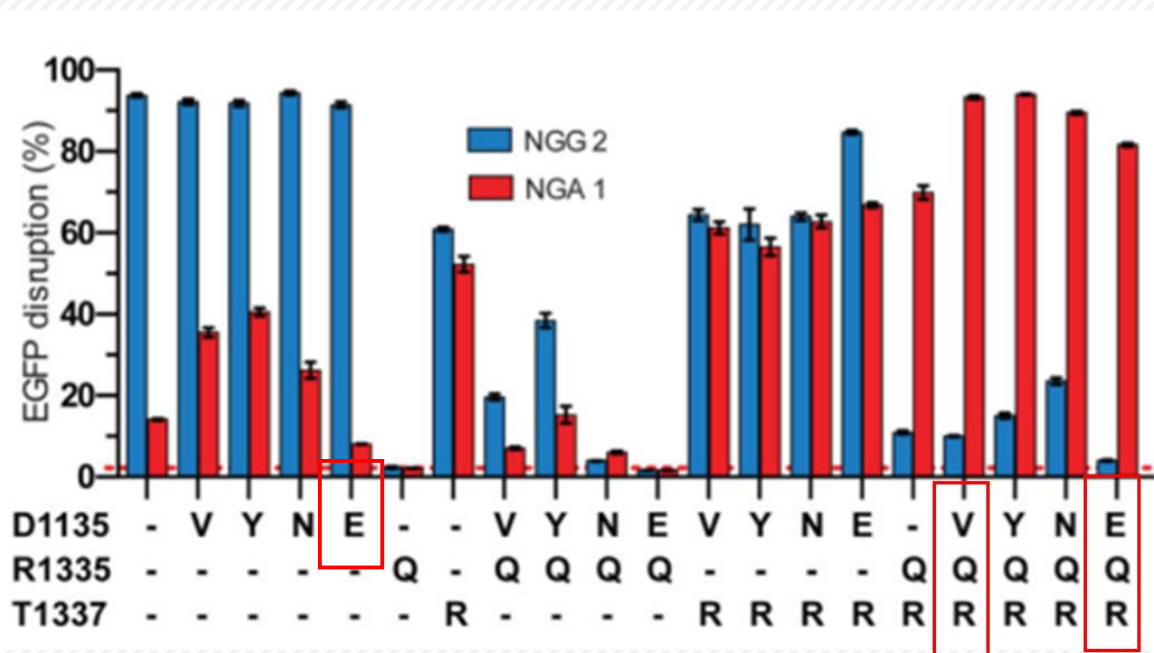
2) plate on media:
 selective (Cm + arabinose)
 non-selective (Cm only)



3) determine relative activity by calculating survival percentage:

$$\text{survival \%} = \frac{\text{selective counts}}{\text{non-selective counts}}$$

细菌阳性选择系统



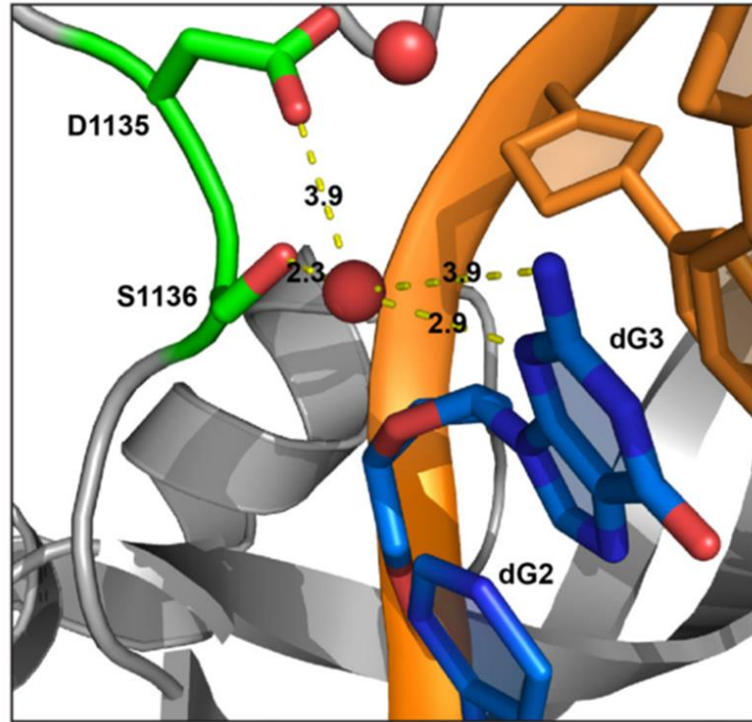
➔ 两个文库中存活的克隆替换频率最高的是D1135V/Y/N/E、R1335Q和T1337R。

使用基于人类细胞的EGFP破坏试验，对这些突变体进行进一步分析，VQR、EQR组合突变更好的区分NGG和NGA PAMs。

➔ Cas9切割编码毒性基因的选择质粒，细胞可以生存。

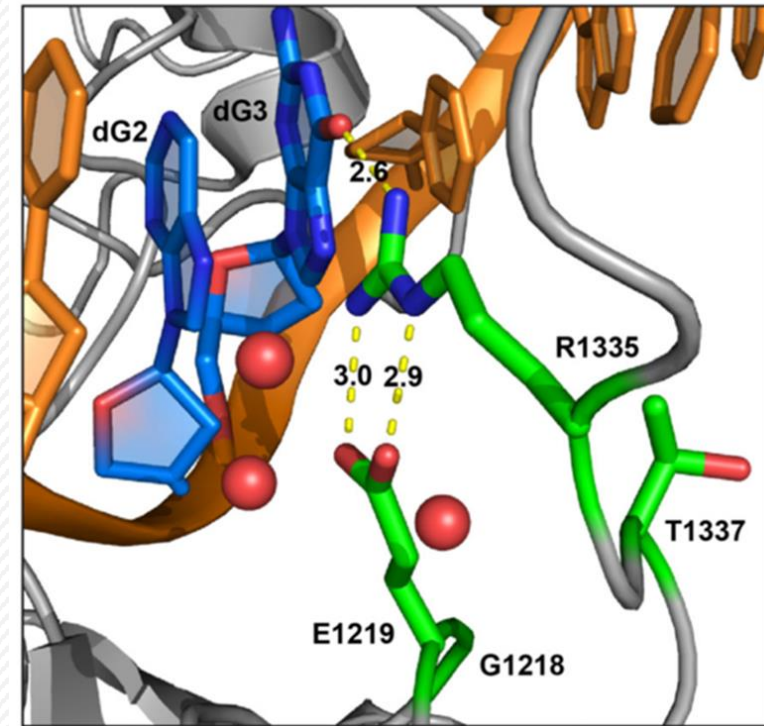
三、蛋白质工程

➤ VRER: D1135V、 G1218R、 R1335E、 T1337R



S1136、**D1135**

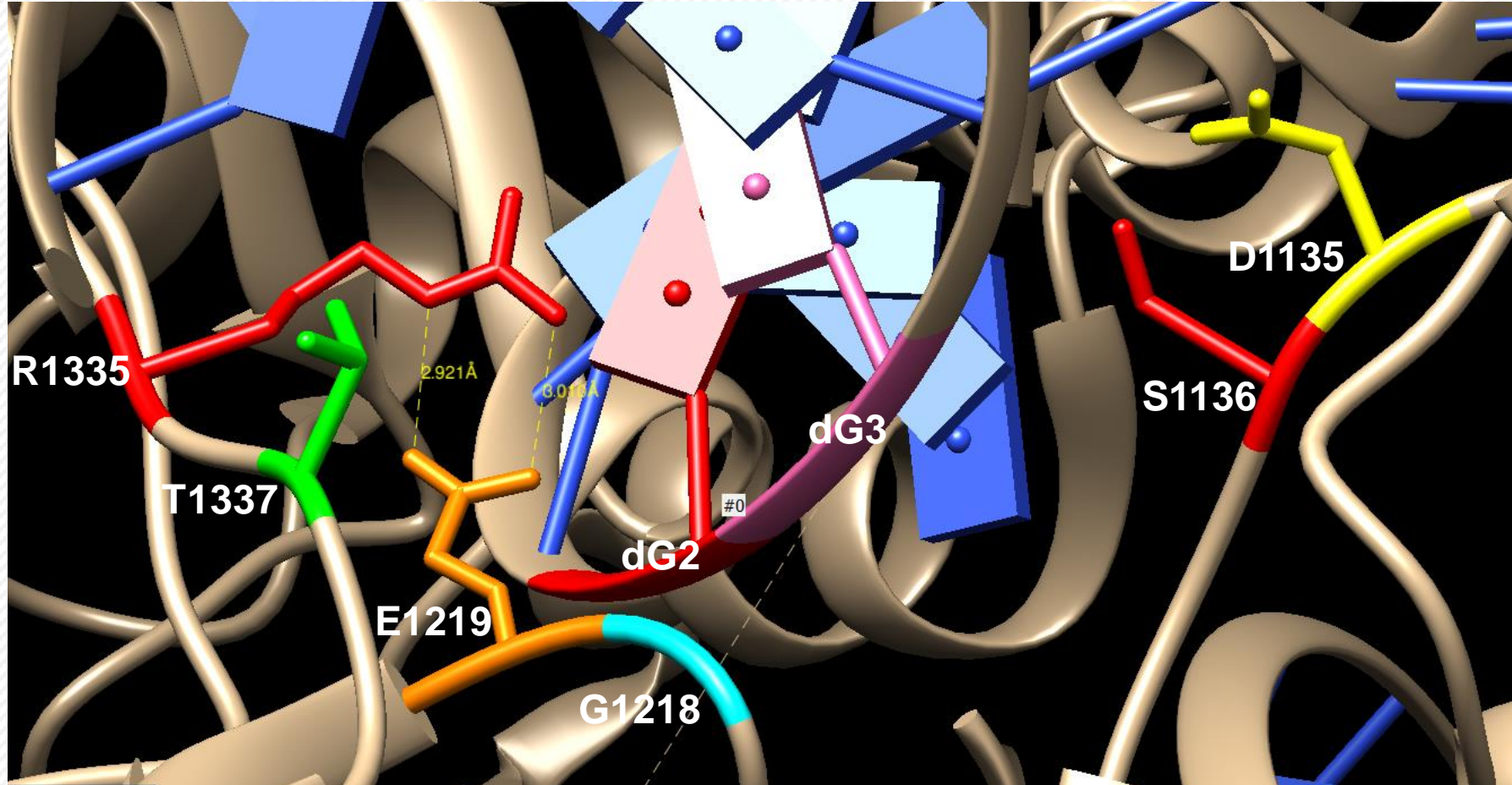
在PAM的第3个碱基上形
成一个水介导的小沟接触



E1219、**R1335**、**G1218**、**T1337**

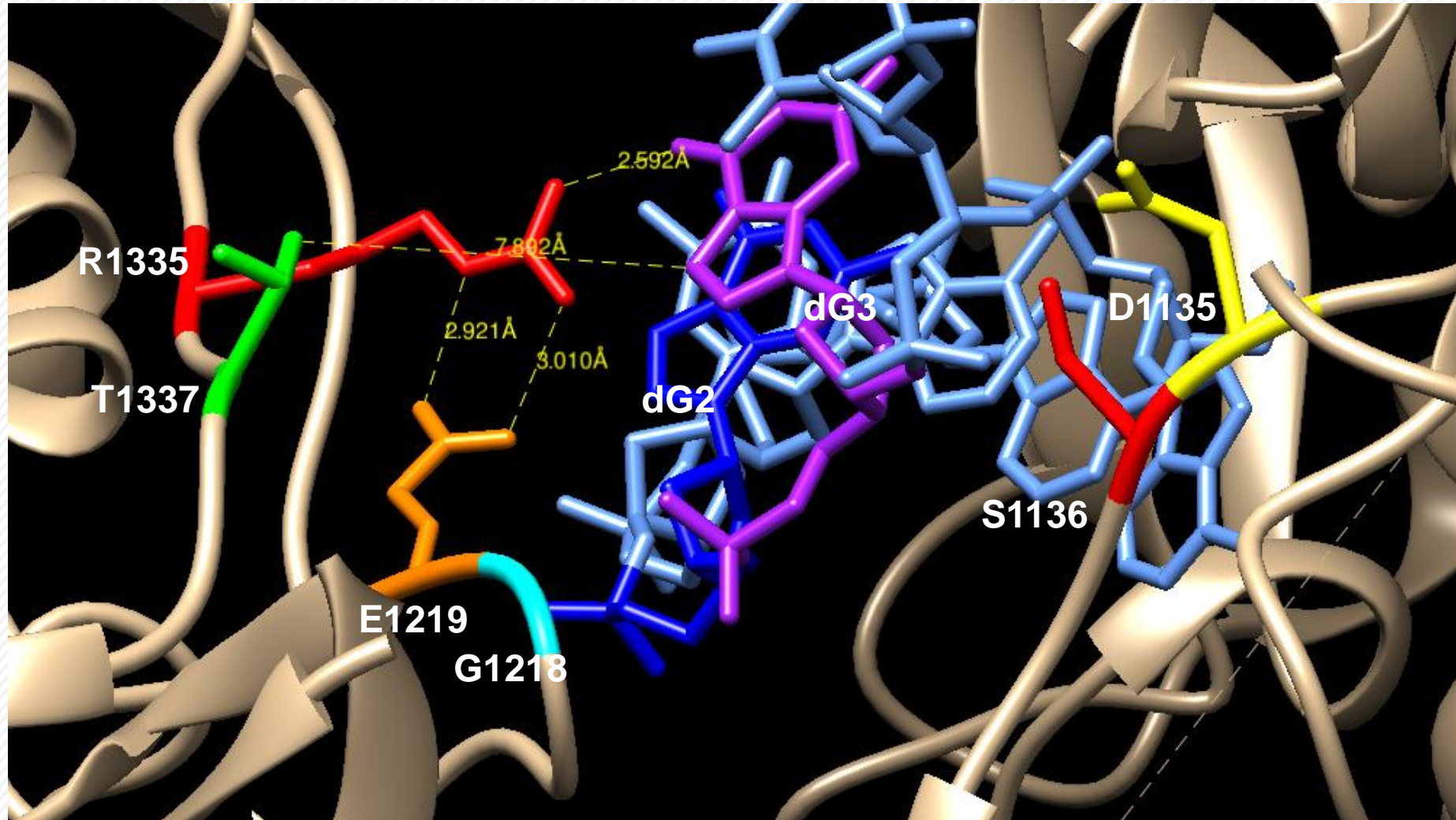
直接与PAM的第3个碱基产生特定的大沟接触

三、蛋白质工程



软件: Chimera
结构来源: PDB (4UN3)

三、蛋白质工程



软件: Chimera
结构来源: PDB (4UN3)

欢迎交流指正