Recitation 8 **Mar 30, April 1**

CRISPR-cas9 to study the function of PCSK9 in controlling LDL-C levels

Acronyms related to CRISPR-cas9 system:

Cas (CRISPR Associated Protein): in the type II CRISPR system, our focus, is the Cas9 nuclease.

CRISPR: Clustered Regularly Interspaced Short Palendromic Repeat, a region in bacterial genomes used in defense against pathogens.

gRNA targeting sequence: The 20 nucleotides that precede the PAM sequence in the genomic DNA.

sgRNA: Single guide RNA is a synthetic RNA composed of a targeting sequence and scaffold sequence derived from the endogenous crRNA and tracrRNA. It is used to target Cas9 to a specific genomic locus in the genome engineering experiments.

HDR (Homology Directed Repair): DNA repair mechanism that uses a template to repair double strand (DS) DNA breaks.

NHEJ (Non-homologous end joining): DNA repair system that does not use a template and that often introduces InDels.

InDel (insertion/deletion) types of mutation that can result in disruption of a gene by shifting the ORF (open reading frame) and or creating premature stop codons during DS break repairs by the NHEJ process.

PAM: Protospacer Adjacent Motif that is required for Cas9 to bind target DNA and must immediately follow the target sequence at the 3'-end.

Protospacer: the 20 nt sequence necessary for Cas9 to bind the target DNA

Target sequence: genomic target of the gRNA sequence. The 20 nucleotides that are incorporated into the gRNA plus the PAM sequence.

I Controlling genome editing

A number of creative approaches for genome editing have been designed and used, each with its own challenges. In all cases an endonuclease is targeted to a given DNA locus and the targeting results in "efficient" double stranded DNA cleavage. (See power point for the description of the designed Zinc Fingers that bind to DNA and the TALENS, that use the Tal effector to bind to DNA. Both constructs append their DNA binders to the FokI nuclease.) The discovery of the CRISPR system and the engineering of this system in the last 3 years has revolutionized genome editing. Changes are continually being made to this system resulting in the explosion of papers on this topic. We will focus on a brief overview of the CRISPR-cas9 system that is used in the paper you are assigned to read this week.

CRISPR-cas9 System (description is adapted from Keasling et al in Metabolic Engineering 34, 44-59 2016).

Historically, CRISPRs were discovered in the *E. coli* genome as palindromic repeats interspaced with short DNA sequences (1987). CRISPR and the associated Cas proteins serve as a three step bacterial defense mechanism against viruses. A cartoon of the process in *S. pyogenes* is shown in Figure 1. In the first step, the Cas proteins coded for by the *cas* genes integrate short DNA spacers (colored circles) that come from viral DNA or plasmid DNA into the CRISPR array in the host genome composed of identical palindromic repeats (purple rectangles).

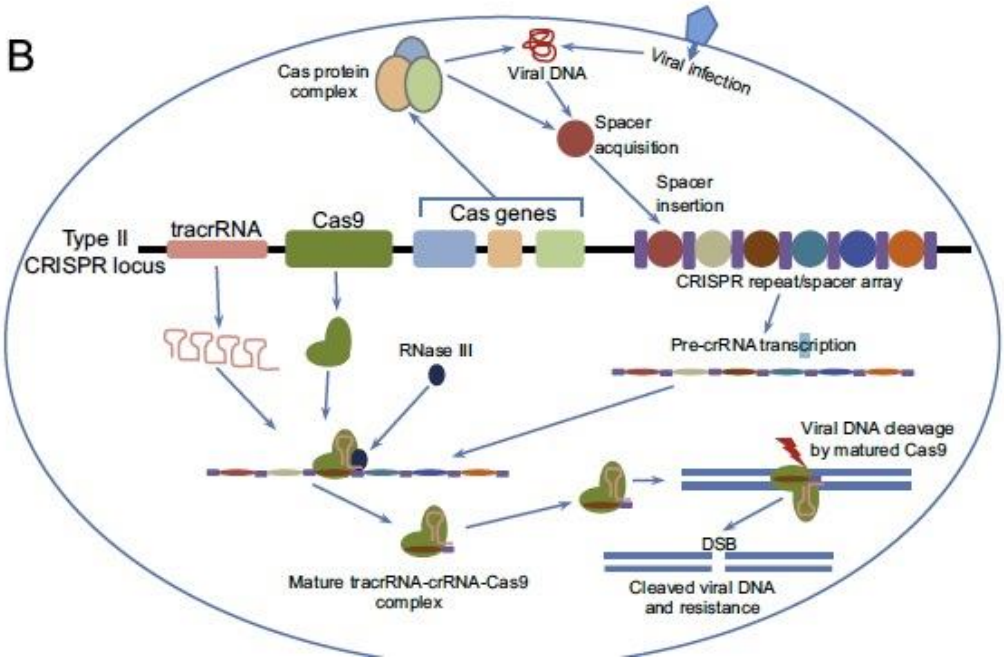
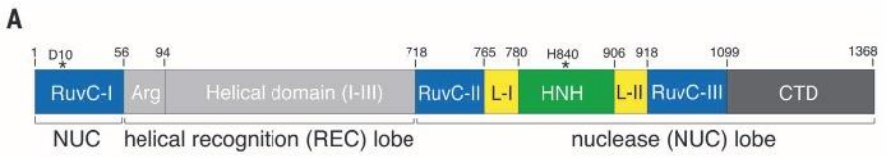


Figure 1 Taken from Metabolic Engineering 34, 44-59 2016. CRISPR-cas bacterial immunity is established by integrating short spacers, which are homologous to invading phage DNA, into the CRISPR array in the bacterial genome. Next, the CRISPR array is transcribed to form a long RNA transcript, CRISPR RNA (crRNA), that then hybridize with tracrRNA that undergoes maturation via RNase III. For type II immunity shown above, the endonuclease Cas9 is guided by the crRNA-tracrRNA hybrid to a foreign-invading DNA and cleaves it by introducing a double strand break. This process leads to resistance to the invading species.

In the second step, the CRISPR array is transcribed to form CRISPR RNA (pre-crRNA transcription). This crRNA undergoes maturation and duplex formation with a complementary trans acting RNA, the tracrRNA. The Cas9 protein binds and RNase III trims the RNA to form the mature tracrRNA-crRNA-Cas9 complex. In the third step, this Cas9 is guided by the tracrRNA/crRNA to a foreign DNA locus, also called the protospacer, and cleaves it by introducing a DS break. The RNA-guided endonuclease activity of the Cas proteins then destroy the genome of the invader, resulting in host immunity.

The best studied CRISPR system to date is type-II from *S pyogenes* and involves the endonuclease Cas9. A carton of the protein is shown below and of note is the α helical lob and the nuclease domain lobe that



S. pyogenes Cas9

contains two nuclease domains deisgnated Ruv and HNH. Structures via XRD and cryoEM have recently been solved (Science 2014 Mar 14 and Science 2016).

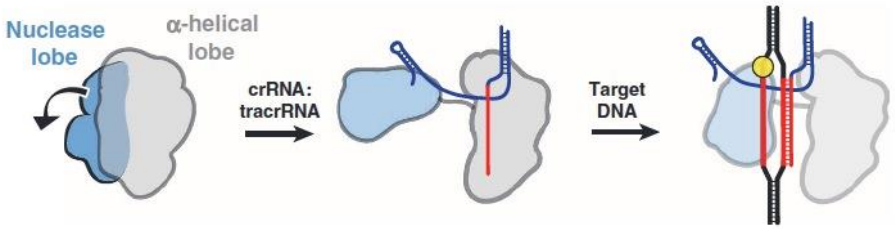


Figure 2 (above) Model for RNA induced conversion of Cas into a structurally activated DNA surveillance complex. Upon binding the crRNA/tracrRNA guide, the two structural lobes of Cas9 reorient such that the two nucleic acid binding clefts face each other. This movement generates a central DNA binding channel which allows access to DS DNA. Target DNA binding in this channel and PAM-dependent R-loop formation result in further structural changes in which the nuclease domain undergoes further movement to fully enclosed the DNA target.

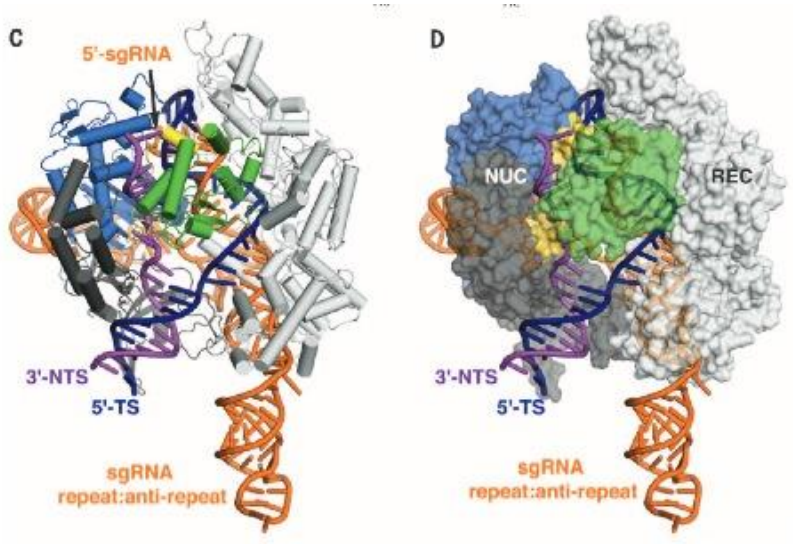


Figure 3 from Science 2016 351 867. Cas9 with sgRNA shown in Figure 4. Color coding is based on the Cas9 catoon in A above. Gray is the helical domain; blue, the RUV nuclease; green, the HNH nuclease; orange, sgRNA; purple, displaced non-targeted DNA strand (20 nt) with the 3' adjacent PAM motif; blue, target DNA strand (20 nts).

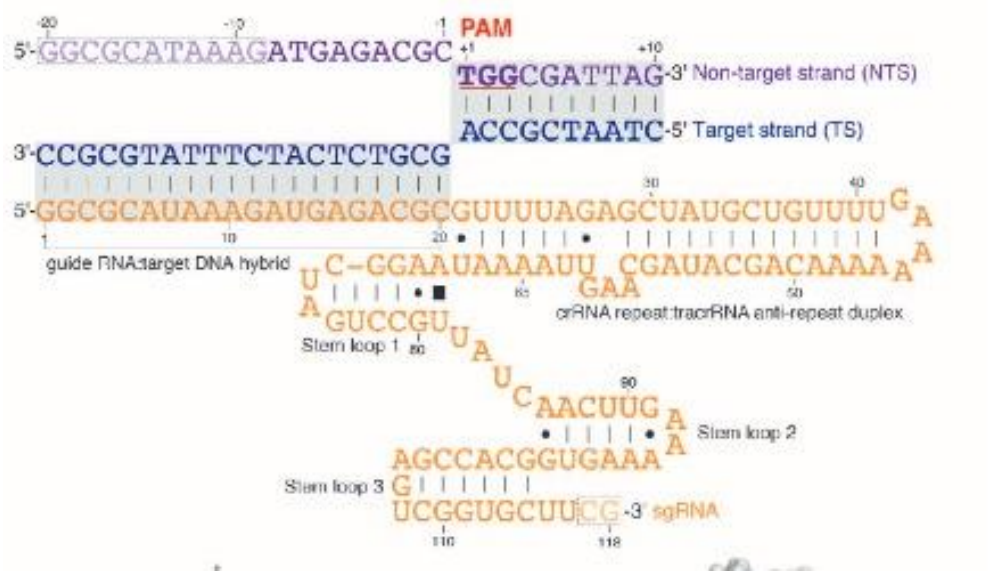


Figure 4 from Science 2016 351 867. Schematic diagram of the sgRNA-target DNA complex. The target DNA strand and the displaced non-target strand are colored dark blue and purple, respectively. The PAM sequence is underlined in red (TGG) and the sgRNA sequence is highlighted in orange. Note by the definitions of the first page that in this case the gRNA and tracrRNA have been engineered into an sgRNA.

The most important parameters for target recognition and successful cleavage are the base pairing between the target and crRNA and the protospacer adjacent motif (PAM) that is a short sequence next to the crRNA homologous sequence on a target DNA. The Cas9 nuclease requires a PAM sequence of 3' nucleotides (NGG). Recently the tracrRNA and the crRNA have been fused into a guide RNA (gRNA) for Cas9. This system has become the two component genome engineering tool of choice: sgRNA/Cas9.

Engineering continues on this system with a focus on the following major issues: 1. The delivery methods to cells (this is discussed in the PCSK9 paper); 2. Minimization of the off target effects (also discussed in the PCSK9 paper you are assigned); 3. How to control the balance between the NHEJ an HDR mechanism of DS break repair depending on the purpose of the engineering (In the PCSK9 paper, the NHEJ method predominates); 4. To better understand Cas9 activity and target selection with sgRNA (the new structures recently published will likely help).

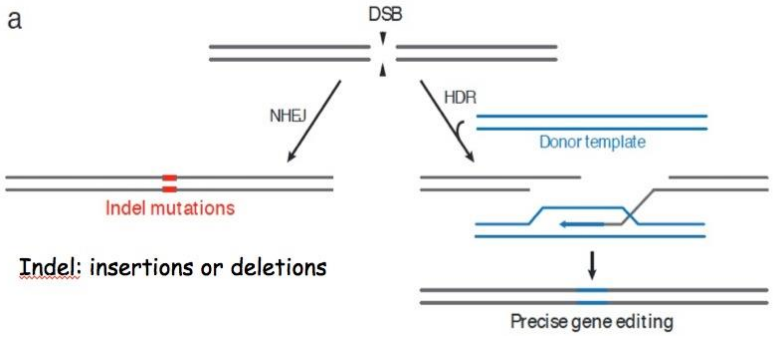


Figure 5 (from Ann Rev Micro 69, 209-228 2015). The two mechanisms of repair of DS DNA breaks: NHEJ and HDR.

II. PCSK9 (proprotein convertase subtilisn/kexin type9)

PCSK9 is a protein secreted into the plasma that binds to the hepatic LDLR and prevents LDLR recycling by its degradation in the lysosome. If the LDLR concentration is reduced, the concentration of LDL-C in the plasma is increased and one has greater risk for coronary disease.

The PCSK9 system was discovered, as with the LDLR, by a gain of function mutation in a patient that presented with lower LDLR levels and as a consequence elevated LDL-C in the plasma. A loss of function patient, also presented. In this case a single substitution (R46L) resulted in 15% reduced LDL-C and 45% reduction in coronary disease. The reduced coronary disease in these patient, even with other risk factors like hypertension, diabetes and smoking, placed PCSK9 as a potential therapeutic target in dealing with coronary disease. The main thesis is that lifelong LDL reduction levels are much more effective in lowering coronary disease than statins which are given after coronary disease has been diagnosed, that is, much latter in life.

PCSK9 is a 75 Kda zymogen that must be autocatalytically processed to be secreted. PCSK9 is a serine protease, homologous to subtilisn and yeast kexin. There are 8 human isozymes of this protein. The domains of the protein are given below.

N- signal peptide----prosegment-----catalytic domain----C, H rich, C terminal domain

The paper for recitation this week Circulation 2015 132, 1648-66, was chosen for two reasons. It demonstrates the importance of PCSK9 as a therapeutic target and it allows introduction to the CRISPR-cas9 technology that has taken the scientific community by storm.

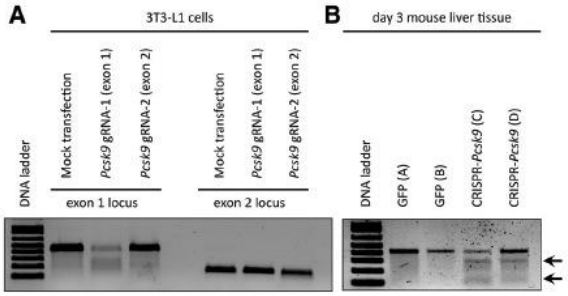


Figure 6 On-target and off target effects in mouse cells and livers receiving clustered regularly interspaced short palindromic repeats (CRSPR)-Cas9. A. Surveyor assays performed with genomic DNA from 3T3-L1 cells transfected with Cas9 and a guide RNA targeting *pcsk9* exon 1 (gRNA-1) or a guide RNA targeting *pcsk9* exon 2 (gRNA-2). B. Surveyor assays performed with genomic DNA liver samples taken from mice 3 days after receiving a control adenovirus expressing green fluorescent protein (GFP) A and B or an adenovirus expressing Cas9 and gRNA. Arrows show the cleavage products resulting from the Surveyor assays; The intensity of the cleavage product bands relative to the uncleaved product band corresponds to the mutagenesis rate.

The investigators initially screened candidate CRISPR guide RNA targeting sequences in exon 1 and exon 2 of mouse *pcks9* gene in 3T3-L1 cells. In the Surveyor assays shown in Figure 6, they showed that after 3 days, the guide RNA targeting exon 1 displayed 50% mutagenesis at the on target PCSK9. They analyzed the liver DNA of the mouse after 3 days and found a wide variety of Indels with one or two bp insertions or deletions. They searched for off-target sites of cleavage (shown in the paper and in SI) and found none, within the limits of detection.

Determination of the effects of CRISPR editing. To look at the effects of editing, they repeated the experiments with a modified protocol and looked at levels of PCSK9, triglycerides, total cholesterol (Figure 7) and complete lipoprotein profiling of pooled plasma samples (Figure

8). Finally they also quantitated levels of LDLR with Western blotting (data in paper, but not shown). The levels of LDLR for mice treated with CRISPR-Pcks9 was higher compared to the control mice.

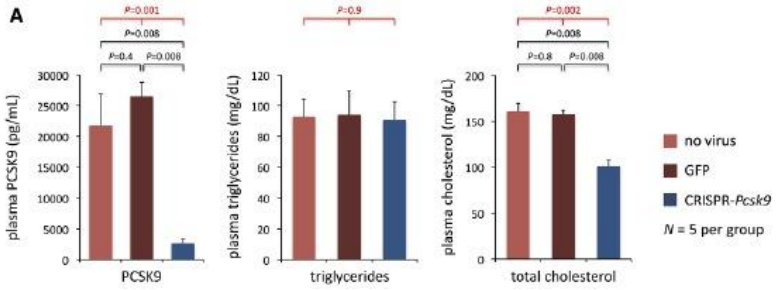


Figure 7 Levels of plasma PCKS9, triglycerides, and plasma cholesterol of the controls (no virus and GFP) and the CRISPR-PCKS9.

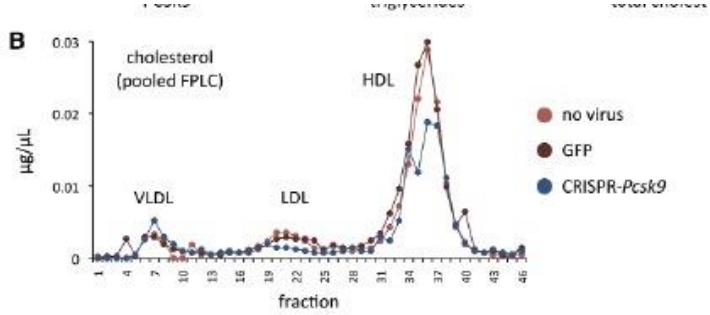
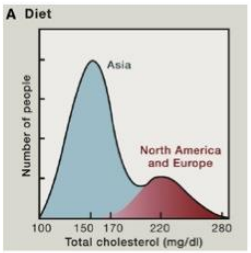
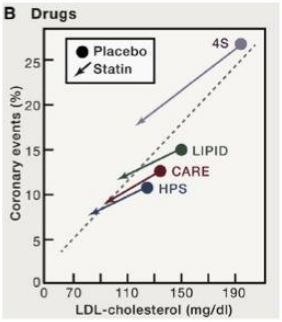


Figure 8 Levels of lipoproteins in the mice as described in the three sets of experiments in Figure 5.

Conclusions: With coronary disease being the number one killer world wide, a safe and effective PCSK9 genome editing therapy might have a significant impact on human health. This single treatment could potentially mimic the loss of function mutations observed with patients. Thus therapeutics targeting PCSK9 could potentially compliment statins. As noted in the Brown and Goldstein review, the coronary events are much lower when the levels of LDL-C are maintained throughout life (Japanese fisherman), rather than subsequent to statin treatment (Figure 9). Recently a paper has been published in Circulation Feb 15, 2016 by Bergeron, An, Ding and Krasss in press on line, which summarizes the different approaches to therpeutics, including two monoclonal antibodies to PCKS9 in clinical trials.



Japanese fisherman low coronary disease and low Ch



4 different studies with 4000 to 10000 people

Figure 9 Japanese fisherman have lower cholesterol levels and lower coronary disease incidents. The studies suggest having low LDL-C ones entire life is much more effective than lowering LDL-C with statins after diagnosis.

5.08J Biological Chemistry II

Spring 2016