# Cas9 functional domains

Functional domains tend to exhibit evolutionary conservation of structure. This allows prediction of function in unknown proteins when conserved domains are identified. HNH and RuvC are examples of nuclease domains that are found in many other proteins.

- **Recognition lobe (REC)** – Rec I, II, III helical domains, important for binding the guide RNA
- **Nuclease lobe (NUC)**
  - HNH: nuclease with 1 metal ion catalytic mechanism, cleaves target strand

    His 840 is conserved; H840A creates "nickase" by mutating HNH activity
  - RuvC: nuclease with 2 metal ion catalytic mechanism, cleaves non-target DNA strand

    Asp 10 is conserved; D10A creates "nickase" by mutating RuvC activity

    *H840A plus D10A creates "dead Cas9", Cas9 with no nuclease activity
  - C-terminal domain: PAM recognition and DNA unwinding
- **Arginine-rich bridge helix** – important for properly aligning the spacer sequence for interrogation

**Cas9 Function:**

PAM recognition region is disordered and hidden until Cas9 binds guide RNA. Binding of guide RNA causes the greatest conformational change and is the critical step in creating the "surveillance complex" (Cas9 domain model: PDB 4zt0)

The guide RNA is "held" primarily by the Rec helical domains interacting with the stem loop 1 of tracrRNA and the repeat-antirepeat duplex between tracrRNA and crRNA. This is the 'handle' that Cas9 uses to hold the guide RNA in place.

Upon guide RNA binding, 2 important conformational changes take place. The first 10 nucleotides (seed sequence) of the "spacer" sequence is opened up and "pre-ordered" by binding along the arginine bridge helix within the central tunnel between the REC and NUC lobes. This allows quick interrogation of DNA to determine complementarity. The PAM binding site in the C-terminal domain (arginines 1333 and 1335) becomes ordered and accessible. The PAM binding arginines are critical for binding DNA and exposing the flanking DNA to the seed sequence for interrogation.

The PAM binding arginines are located on a hairpin nestled in a positively charged groove of the C-terminal domain. Opposite these residues in this groove is the phosphate lock loop (K1107-S1109) which is critical for unwinding the DNA once bound by PAM sequence to the arginines.

Binding of the arginines to the PAM sequence ('NGG') it in the nontarget strand causes the phosphate lock loop to bind to the backbone of the target strand and through an *ATP-independent (non-helicase) conformational change*, produces a kink in the target strand, melts the DNA, turns the target strand toward the seed sequence within the central tunnel (between lobes) and throws the nontarget strand through a positively charged side tunnel within the NUC lobe, thereby separating it from the target strand. If the target strand does not match the seed sequence perfectly, the DNA is released. If there is a perfect match (nucleotides 11-20 don't have to be a perfect match, only 1-10), the HNH domain cleaves the target strand and the RuvC domain cleaves the nontarget strand and the DNA is released.