INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student	:	**Vineet Anand**
Roll Number	:	**166106104**
Programme of Study	:	**Ph.D.**

Thesis Title: **Understanding the defence mechanism of CRISPR-Cas subtype I-C of *Leptospira interrogans***

Name of Thesis Supervisor(s)	:	**Prof. Manish Kumar**
Thesis Submitted to the Department/ Center	:	**BSBE**
Date of completion of Thesis Viva-Voce Exam	:	**27-09-2024**
Key words for description of Thesis Work	:	CRISPR-Cas, *Leptospira,* DNase, RNase

SHORT ABSTRACT

Leptospira interrogans is a bacterium that causes Leptospirosis. However, its pathogenicity is poorly understood due to the lack of effective conventional genetic manipulation tools. *L. interrogans* serovar (sv.) Copenhageni encodes two CRISPR-Cas systems - I-B and I-C in its genome. The presence of the CRISPR-Cas system in the genome of pathogenic *Leptospira* has been hypothesized as the reason for low success in genetic manipulation and, hence, deciphering its virulent gene function. The CRISPR-Cas I-B locus in *L. interrogans* has been extensively studied. However, CRISPR-Cas I-C lacks an array element, making it an interesting subject to investigate its Cas protein's role in CRISPR biology. The CRISPR-Cas I-C locus encodes all sets of proteins required for adaptation (LinCas1C, LinCas2C, and LinCas4C), maturation (LinCas5C), and interference (LinCas7C, LinCas8C, LinCas3C, and LinCas3C′). Moreover, the genes encoding these proteins are transcriptionally active. The nucleotide sequence and encoded amino acids of the *cas* genes show low sequence similarity with subtype I-B and are phylogenetically distantly related. The adaptation Cas protein (LinCas1C, LinCas2C, and LinCas4C) exhibits metal-ion-dependent DNase and metal-ion-independent RNase activity under *in vitro* conditions. However, neither rLinCas1C nor LinCas2C demonstrates any activity towards small DNA oligos, except for rLinCas4C. The crystal structure analysis of rLinCas2C shows that it adopts a dimeric conformation, and each subunit exhibits the characteristic ferredoxin fold. The crystal structure of rLinCas2C indicates its existence in a catalytically inactive conformational state. The substitution mutation of active site residues (Tyr7, Asp8, Arg33, and Phe39) of rLinCas2C with Ala and deletion mutation of RNA recognition loop (Asp60, Lys62, Thr63, and Asp64) resulted in compromised DNase activity. The catalytic activity observed in adaptation Cas proteins (LinCas1C and LinCas4C) of sub-type I-C suggests

their potential involvement in recognizing/processing foreign nucleic acid and integrating it as a spacer into the CRISPR array. While the catalytic function of LinCas2C might be necessary for biological processes distinct from the CRISPR-Cas-associated function. The rLinCas5C, a maturation Cas protein, exhibits DNase activity and is contingent on the large size of the DNA substrates (≥3kb), as it displayed no activity on short ssDNA and dsDNA oligos (≤36 nt). This study also revealed that within subtype I-B, which possesses a complete set of Cas proteins and array elements, the mature crRNA-rLinCas6 complex is protected from the endoribonuclease activity of rLinCas5C. *In silico* analysis discovered that among Cas5C orthologs, LinCas5C exclusively contains β-sheet insertion in the C-terminal, and Phe141 is one of the catalytic triad equivalents to a more conserved histidine residue. Deleting additional β-sheet insertions and substituting Phe141 with His141 in LinCas5C resulted in the gain of nuclease activity compared to the wild-type. The rLinCas5C exhibits non-canonical cleavage activity on the non-cognate repeat RNAs and pre-crRNAs. This raises a question about CRISPR-Cas I-C interference capability. To address this, a plasmid interference assay was conducted, which revealed that the Cas proteins of CRISPR-Cas I-C of *Leptospira* involved in the interference could functionally intervene in the foreign plasmids carrying predicted protospacer adjacent motifs in the spacers. Overall, this study provides a comprehensive understanding of the *Leptospira* CRISPR-Cas I-C adaptation and maturation Cas proteins' biochemical activity and sheds light on the interference capability of Cas proteins in a heterologous host (*E. coli*). These findings lay the groundwork for leveraging the endogenous CRISPR-Cas systems of *L. interrogans* for genome editing in the near future.