MECHANISTIC INSIGHTS INTO RNA-GUIDED GENOME EDITING NUCLEASES

DHVANI SANDIP VORA

DEPARTMENT OF BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY DELHI

MARCH 2023

Certificate

This is to certify that the thesis titled **'Mechanistic Insights into RNA-guided Genome Editing Nucleases'** being submitted by **Ms. Dhvani Sandip Vora** to the Indian Institute of Technology Delhi for the award of the degree of **'Doctor of Philosophy'**, is a record of the bonafide research work carried out by her, which has been prepared under my supervision in conformity with the rules and regulations of the Indian Institute of Technology Delhi. The research reports and the results presented in this thesis have not been submitted for any degree or diploma in any other University or Institute.

Dr. D. Sundar
Institute Chair Professor
Department of Biochemical Engineering and Biotechnology
Indian Institute of Technology Delhi

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. D. Sundar, for his patient guidance and support by giving his time, ideas and feedback. I could only complete the journey of my PhD because of the motivation provided by him. I have improved by trying to imbibe his scientific and personal principles.

I am grateful to Prof. Ashok Kumar Patel, Prof. Ravikrishnan Elangovan and Prof. Preeti Srivastava, the members of my Students' Research Committee, for providing support and expertise in shaping this thesis. Completing this endeavor would not have been possible without Dr. Jaspreet Kaur Dhanjal, who has greatly influenced my professional and personal growth during my time as a PhD student.

I express my gratitude to UGC for the financial support during my PhD and to IIT Delhi for having a student-friendly and supportive administration. I also appreciate the IITD Kailash hostel staff and mess staff for their hospitality. Special thanks to my colleagues and friends from the lab- Dr. Vidhi Malik, Dr. Neetu Singh, Dr. Atul Jaiswal, Dr. Moolchand, Shashank Yadav, Yugesh Verma, Sakshi Bhandari, Navaneethan Radhakrishnan, Yogesh Kalakoti, Vipul Kumar, Seyad Shefrin and Pragya Kesarwani; I have learned a lot from each of them on both scientific and personal fronts. I have been able to broaden my perspective and find friendships in many other visiting scholars, undergraduate and graduate project students in my research group.

I also acknowledge the DAAD for the binational PhD fellowship and express gratitude to Jun.-Prof. Michael Boettcher for hosting me during my stay at the Martin Luther University, Germany. The training to perform CRISPR/Cas9 related experiments has deepened my understanding of the subject and broadened my perspectives.

I would like to acknowledge my parents and sister for their love and encouragement in this challenging pursuit; I could strive for betterment with the support of their faith and affection. I would also like to thank my friends, fellow PhD students and hostel mates who made my time at IIT Delhi an enjoyable experience.

Dhvani Sandip Vora

Abstract

Efficient manipulation of genes in organisms has propelled research in understanding mechanisms that govern life. New techniques that allow editing in single-celled organisms as well as complex life forms such as animals and plants are gaining traction. At the heart of such advances are custom-designed nucleases like zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the CRISPR/Cas9 system.

A naturally occurring adaptive defence mechanism found in prokaryotes, the CRISPR/Cas9 system has been repurposed as an RNA-guided DNA targeting machinery. This transformative technology has shown great promise for biology, genetics and medicine. However, it's widespread application for safe and effective genome editing and transcription modulation has been curbed by its inadequate specificity. In this thesis, the key factors involved in determining the efficacy of CRISPR/Cas9 have been studied that can be incorporated in guide RNA design tools to improve their predictions. Further, the study of the mechanism of on-target activity by Gaussian accelerated molecular dynamics solution has been carried out to determine sub-molecular interactions essential for driving Cas9 activity.

सार

जीवों में जीनों के कुशल क्रम परिवर्तन ने जीवन को नियंत्रित करने वाले तंत्रों को समझने में अनुसंधान को प्रेरित किया है। नई तकनीकें जो एकल-कोशिका वाले जीनोम के साथ-साथ जानवरों और पौधों जैसे जटिल जीवन रूपों में संपादन की अनुमति देती हैं, आकर्षण प्राप्त कर रही हैं। इस तरह की प्रगति के केंद्र में अनुकूलित किए गए न्यूक्लियीज़ जैसे ज़िंक फिंगर न्यूक्लियीज़ (TALEN), ट्रांसक्रिप्शन एक्टीवेटर-लाइक इफ़ेक्ट न्यूक्लीज़ (ZFN) और क्रिस्पर-कस 9 सिस्टम हैं।

प्रोकैरियोट्स में पाया जाने वाला एक स्वाभाविक रूप से होने वाला अनुकूली रक्षा तंत्र, क्रिस्पर-कस 9 प्रणाली को एक आरएनए-निर्देशित डीएनए लक्ष्यीकरण मशीनरी के रूप में पुनर्निर्मित किया गया है। इस परिवर्तनकारी तकनीक ने जीव विज्ञान, आनुवंशिकी और चिकित्सा के लिए बहुत आशाजनक कार्य किये है। हालाँकि, यह सुरक्षित और प्रभावी जीनोम संपादन और प्रतिलेखन समायोजन के लिए व्यापक अनुप्रयोग है, इसकी अपर्याप्त विशिष्टता प्रतिरोध बन गयी है।

प्रस्तुत शोध प्रबन्ध में क्रिस्पर-कस 9 की प्रभावशीलता का निर्धारण करने में शामिल प्रमुख कारकों का अध्ययन किया गया है। जिन्हें उनकी भविष्यवाणियों को बेहतर बनाने के लिए गाइड डिज़ाइन टूल में शामिल किया जा सकता है। इसके अलावा, कस 9 गतिविधि को चलाने के लिए आवश्यक उप-आण्विक, परस्पर प्रभाव निर्धारित करने के लिए गॉसियन त्वरित आणविक गतिशीलता समाधान द्वारा ऑन-टारगेट गतिविधि के तंत्र का अध्ययन किया गया है।

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMs	Activation Maps
AUC	Area Under the Curve
Cas9	CRISPR-associated nuclease 9
CRISPR	Clustered Regularly Interspersed Palindromic Repeats
crRNA	CRISPR RNA
DSB	Double-stranded break
GaMD	Gaussian accelerated molecular dynamics
HDR	Homology-directed repair
MLP	Multilayer Perceptron
NHEJ	Non-homologous end joining
nt	nucleotide
PAM	Protospacer Adjacent Motif
pdf	Probability Density Function
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Networks
ROC	Receiver Operating Characteristic
sgRNA	single guide RNA
SHAP	SHapley Additive exPlanations
TALENs	Transcription Activator-like Nucleases
tDNA	target DNA
tracrRNA	trans-CRISPR RNA
ZFNs	Zinc Finger Nucleases