Mapping of RNA G-quadruplexes in mirtrons and understanding their biological role

Uzma Salim

Kusuma School of Biological Sciences

INDIAN INSTITUTE OF TECHNOLOGY, DELHI

November 2024

Mapping of RNA G-quadruplexes in mirtrons and understanding their biological role

by

Uzma Salim

Kusuma School of Biological Sciences

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY, DELHI

November 2024

CERTIFICATE

This is to certify that the thesis titled "Mapping of RNA G-quadruplexes in mirtrons and understanding their biological role", being submitted by Ms. Uzma Salim in the Indian Institute of Technology, Delhi for the award of the degree of "Doctor of Philosophy" is a record of the bonafide research carried out by him, which has been prepared under my supervision and guidance in conformity with rules and regulation of the Indian Institute of Technology Delhi, India. The results prescribed in it have not been submitted in part or full to any other University or Institute for the award of any Degree/Diploma.

Dr. Vivekanandan Perumal

Professor,

Kusuma School of Biological Sciences

Indian Institute of Technology Delhi

New Delhi-110016, India

Acknowledgments

I want to express my deepest gratitude to my supervisor, Prof. Vivekanandan Perumal, for his unwavering support, guidance, and mentorship throughout my doctoral journey. His invaluable feedback has shaped my research direction and enhanced this thesis's quality. His unwavering dedication and expertise have profoundly influenced my academic and personal growth, which I sincerely appreciate.

My heartfelt thanks extend to the members of my SRC committee: Prof. Manidipa Banerjee, Prof. Ashok Kumar Patel, and Prof. Ritu Kulshreshtha. I deeply admire their expertise and the insightful comments and suggestions they provided, which have significantly contributed to the refinement of this work.

I am incredibly grateful to Prof. Manoj B. Menon and Dr. Sonam Dhamija for their tremendous support and invaluable assistance in completing my Ph.D. thesis. Their guidance, encouragement, and willingness to share their expertise have been pivotal in overcoming numerous challenges throughout my research. This work would not have been possible without their unwavering help and dedication.

I would also like to thank all the faculty members of KSBS, including Prof. Tapan Chaudhary, Prof. C.S. Dey, Prof. James Gomes, Prof. Aditya Mittal, Prof. Biswajit Kundu, Prof. Archana Chug, Prof. Amitabh Mukhopadhyay, Prof. Tapan Naik, Prof. Shilpi Minocha, and Prof. Anita Roy, for their availability for scientific discussions and suggestions. I am particularly grateful for the use of the facilities available in their labs, which have been instrumental in completing my work.

I want to thank my lab members for creating a supportive and collaborative environment that greatly contributed to the success of my research. First, I would like to mention Dr. Ashish Kumar for being a fantastic senior and friend. From work to trip discussions, from reagents to food discussions, you have been my support system throughout my Ph.D. journey. I would also like to acknowledge the three incredible girls from my lab: Sneha Chaudhary, Priya Nagpal, and Rishika Kaushal. Thank you for never letting me fall, being there in my difficult times, and ensuring we find a solution together. I have had the best of times with you, even in the most challenging phases

Acknowledgments

of my Ph.D. I express my appreciation to the seniors in my lab: Dr. Bahni Biswas, Dr. Shivani Kumar, Dr. Nandhini Srinivasan, Dr. Kiruthika Shankar, Dr. Shivaksh Ahluwalia, Dr. Dipannita Ghosh, Dr. Shruti Chowdhari, Dr. Arijit Pal, Dr. Jayant, Dr. Mishi, and Dr. Priyanka Verma. Their constant support, valuable insights, constructive feedback, and willingness to share their expertise have been instrumental in completing my thesis. I am deeply thankful to the juniors in my lab: Belal Hafiz, Sanket Mukerjee, Auroni Deep, Sandhaya Mahinto, Jyoti Rawat, Iswarya, Prachi, and Temesgen, for their enthusiasm and support throughout my Ph.D. journey. Their fresh perspectives, assistance with experiments, and positive energy have significantly contributed to the progress of my research. Their eagerness to help and collaborate has provided invaluable support, making the research process more enriching and enjoyable. The camaraderie and encouragement of my lab members have made my research experience both enriching and enjoyable. Thank you all for making this challenging journey easier, filled with gossip, fun, outings, tea sessions, and tons of memories.

I want to express my deepest gratitude to my friends Pragya, Dr. Firdos Ehtram, Dr. Prachi Verma, Dr. Priyanka Nair, Dr. Aquib Ehtram, Akanksha, Dr. Kritika, Dr. Ramesh Kumar, and Anand for standing by my side during the most challenging phase of my life. Your unwavering support, encouragement, and companionship have been my constant source of strength and happiness throughout my PhD journey. The cherished memories we created in the lab and during our tea sessions are invaluable treasures to me. I am profoundly grateful to each of you; your presence has made an immeasurable difference in my life. Thank you from the bottom of my heart.

I am profoundly grateful to my wonderful friends from the institute, Pragati, Maryam, Malvika Ramani, Dibyakanti, Tapabrata, Vikas, Saurabh, Shumayila, Shubham, Medha, Devanshi, Yamini, Saurabh, Ankit, Gagandeep, Ashok, Amit, Sai, and Rupam, for always being there for me. My PhD journey was genuinely incredible because of the considerable part you all played in it.

I also sincerely thank Saurav Shrivastava and Siddharth for taking time out from their work to assist me with flow cytometry and fluorescence microscopy. Your help was crucial, and I am deeply thankful for your generosity and expertise.

Acknowledgments

Moreover, I am thankful to all Kusuma School of Biological Sciences students for being so helpful and for sharing resources and instruments. Your willingness to make the labs available and offer assistance whenever needed has been invaluable. I am truly blessed to have had such a supportive and collaborative community around me.

I extend my appreciation to the Kusuma School of Biological Sciences for providing the necessary resources and a conducive research environment. I thank KSBS staff members and Ankit for their administrative and technical support. I want to thank UGC for the fellowship.

I want to extend my heartfelt thanks to my long-time friends Sunil and Sachin for the countless scientific and 'not so scientific' discussions over tea, food, and trips. Your friendship has been a beacon of joy and comfort. Thank you for bringing a smile to my face whenever I was feeling down, helping me in every possible way during times of great need, and ensuring I stayed grounded with your views and antics, even if it meant getting on my nerves sometimes. I appreciate your patience and understanding, especially when I was equally irritating. Your unwavering support, humor, and camaraderie have been pillars of strength for me. I am deeply grateful for every moment we've shared, every laugh, every debate, and every adventure. Thank you from the bottom of my heart for being such incredible friends.

I want to extend my heartfelt thanks to my friends Charu, Priyanka, and Nikita. Our continuous meetups, birthday and wedding celebrations gave me some of the best memories of my life during my PhD tenure. Your companionship brought immense joy and balance to my life.

I am also profoundly grateful to my friends Indrani, Ratul, Ansh, Manish, and Jai. Our meetings and every event celebration brought joy and happiness during the challenging phases of my PhD. Your presence made those challenging times bearable and filled my life with positivity.

I want to express my profound appreciation to Sandeep, Ashish, Tajinder, Tabish, and Panshul. Your help and support during the most challenging phase of my life were invaluable, and I will always be grateful for your kindness and assistance.

I am thankful to Shahid Jamal's family, including Rana Mami, Aasia, Umar, Saad, and Rabia, for being an extended family.

Acknowledgments

I am deeply grateful to my family for their unwavering love, encouragement, and understanding. Your support has been a constant source of patience, strength, and motivation, enabling me to navigate the complexities of doctoral research. My father has been immensely supportive, always ensuring my well-being. From packing my lunch to dropping and picking me up from the metro station, he made sure everything was available for me. Dad, your selflessness and care mean the world to me. My mother, my first teacher, and my most significant source of strength, I am here because of you. Your guidance and unconditional love have been my foundation. You will always be with me. My elder sister, Shazia, has been like a second mother to me. Your wisdom, care, and support have been indispensable. My younger sister, Nida, is my little bundle of joy. Your laughter and energy have always lifted my spirits. I am incredibly thankful to my sisters for their unwavering support and love throughout my life. Their presence has made all the difference. My sisters have been my pillars of strength, always there with encouragement and understanding. I am truly blessed to have them in my life. Words cannot adequately express how thankful I am to each of you. Your love and support have been the bedrock of my journey, and I am forever indebted to you all. Thank you from the bottom of my heart.

This thesis would not have been possible without the collective support of these individuals and institutions. Thank you for being an integral part of my academic and personal journey.

Abstract

Mirtrons represent the predominant class of non-canonical miRNAs derived from introns through a Drosha-independent, splicing-dependent pathway. Unregulated splicing of introns containing hairpins can adversely impact Dicer/Ago-mediated canonical miRNA biogenesis. Among invertebrates, the 3' tailing of mirtron hairpins by terminal uridyl transferases (TUTases) suppresses mirtron biogenesis. However, the mechanisms regulating mirtron biogenesis in vertebrates remain poorly understood. G-quadruplexes are nucleic acid secondary structures formed by G-rich sequences in both DNA and RNA, and they have emerged as significant splicing regulators. Our study found that human mirtrons are enriched with rG4s specifically in the 5' arm. This enrichment suggests a potential role for rG4s in mirtron biogenesis. Careful analysis revealed that while the 5' arms of plant and invertebrate mirtrons are enriched with uracils (Us), the 5' arms of vertebrate mirtrons are enriched with guanines (Gs). This shift in nucleotide enrichment could indicate an evolutionary adaptation in mirtron biogenesis. Further analysis revealed that most mammalian mirtrons contain an RNA G-quadruplex (rG4), a feature not observed in plants or invertebrate mirtrons. Interestingly, almost all rG4s in mammalian mirtrons were located in the 5' arm. Predicted rG4s in human mirtrons form a G-quadruplex structure *in vitro*, and rG4 formation in the 5' arm facilitates Drosha-independent splicing for the excision of mirtrons. Disruption of rG4s in the 5' arm inhibits splicing and maturation in DROSHA knockout cells, whereas mutations outside the rG4-motif do not impact mirtron biogenesis. Thus, our findings indicate that rG4s in the 5' arm are critical regulatory elements in the evolutionary landscape of mammalian mirtrons. Extending this knowledge to human-infecting herpesviruses revealed that more than 50% of Herpes Simplex Virus (HSV) pre-miRNAs contain an rG4. Further analysis identified hsv1-mir-H27 and its analog hsv2-mir-H20 as putative mirtrons. Expression of hsv1-mir-H27 in DROSHA knockout cells demonstrated that the pre-miRNA could be processed in a Drosha-independent manner, confirming that this virus-encoded miRNA is a mirtron. Predicted splice sites for hsv1-mir-H27 lead to successful splicing of the mirtron. The rG4 in hsv1-mir-H27 pre-miRNA forms a G-quadruplex structure *in vitro*, and like human mirtrons, rG4 formation in the 5' arm facilitates Drosha-independent splicing. The sequential disruption of rG4s in the 5' arm inhibits splicing in DROSHA knockout cells, while mutations outside the rG4-motif do not affect mirtron biogenesis. This indicates that rG4 formation in the 5' arm of pre-miRNA significantly contributes to its

splicing. Hsv1-mir-H27 is, thus, the first human-infecting viral mirtron to be identified. This work revealed *NFE2L1* as a bona fide target of hsv1-miR-H27. Originating from the ICP0 locus, hsv1-miR-H27 expression is associated with increased cellular apoptosis, indicating its potential role in modulating cell death pathways. It enhances pro-apoptotic caspase 3/7 activity, suggesting involvement in HSV1-induced apoptosis through the hsv1-miR-H27-*NFE2L1* axis. Thus, hsv1-mir-H27 can be the potential unidentified non-coding factor responsible for ICP0-induced cellular apoptosis. However, the apoptosis induced by hsv1-miR-H27 is minimal compared to infection models, implying that additional viral proteins or stress factors may amplify this response. These could include oxidative stress, proteotoxic stress from infection, or other virus-derived elements affecting the *NFE2L1-PSMB6* pathway to enhance cell death. Further investigation is needed to fully understand the role of these amplifiers in the apoptosis response.

This research is a significant step in understanding mirtron biogenesis in mammals and viruses. It reveals additional roles for rG4s in small RNA biology and suggests a potentially crucial role for hsv1-miR-H27 in the viral manipulation of host cell processes. These implications underscore the importance of our findings in RNA biology and pave the way for numerous exciting avenues of research into the role of mirtrons, their processing and biogenesis, and their downstream effects in regulating cellular processes.

सार

मिट्रॉन गैर-विहित माइक्रोआरएनए के प्रमुख वर्ग का प्रतिनिधित्व करते हैं जो ड्रोसा-स्वतंत्र, स्प्लिसिंग-निर्भर मार्ग के माध्यम से इंट्रॉन से प्राप्त होते हैं। हेयरपिन युक्त इंट्रॉन के अनियमित स्प्लिसिंग से डिसर/एगो-मध्यस्थ विहित माइक्रोआरएनए बायोजेनेसिस पर प्रतिकूल प्रभाव पड़ सकता है। अकशेरुकी जीवों में, टर्मिनल यूरिडाइल ट्रांसफ़ेरेस (TUTases) द्वारा मिट्रॉन हेयरपिन की 3' टेलिंग मिट्रॉन बायोजेनेसिस को दबा देती है। हालाँकि, कशेरुकियों में मिट्रॉन बायोजेनेसिस को विनियमित करने वाले तंत्रों को अभी भी ठीक से समझा नहीं गया है। जी-क्वाड्रप्लेक्स डीएनए और आरएनए दोनों में जी-समृद्ध अनुक्रमों द्वारा निर्मित न्यूक्लिक एसिड द्वितीयक संरचनाएँ हैं, और वे महत्वपूर्ण स्प्लिसिंग विनियामक के रूप में उभरे हैं। हमारे अध्ययन में पाया गया कि मानव मिट्रॉन विशेष रूप से 5' भुजा में आरएनए जी-क्वाड्रुप्लेक्स से समृद्ध होते हैं। यह संवर्धन मिट्रॉन बायोजेनेसिस में आरएनए जी-क्वाड्रुप्लेक्स की संभावित भूमिका का सुझाव देता है। सावधानीपूर्वक विश्लेषण से पता चला कि पौधे और अकशेरुकी मिट्रॉन की 5' भुजाएँ यूरेसिल (यू) से समृद्ध हैं, जबकि कशेरुकी मिट्रॉन की 5' भुजाएँ ग्वानिन (जी) से समृद्ध हैं। न्यूक्लियोटाइड संवर्धन में यह बदलाव मिट्रॉन बायोजेनेसिस में एक विकासवादी अनुकूलन का संकेत दे सकता है। आगे के विश्लेषण से पता चला कि अधिकांश स्तनधारी मिट्रॉन में एक आरएनए जी-क्वाड्रप्लेक्स होता है, जो पौधों या अकशेरुकी मिट्रॉन में नहीं देखा गया है। दिलचस्प बात यह है कि स्तनधारी मिट्रॉन में लगभग सभी आरएनए जी-क्वाड्रुप्लेक्स 5' भुजा में स्थित थे। मानव मिट्रॉन में पूर्वानुमानित आरजी4 इन विट्रो में एक जी-क्वाड्रप्लेक्स संरचना बनाते हैं, और 5' भुजा में जी-क्वाड्रुप्लेक्स गठन मिट्रॉन के छांटने के लिए ड्रोशा-स्वतंत्र स्प्लिसिंग की सुविधा प्रदान करता है। 5' भुजा में आरएनए जी-क्वाड्रुप्लेक्स का विघटन द्रोषा नॉकआउट कोशिकाओं में स्प्लिसिंग और परिपक्वता को बाधित करता है, जबकि आरएनए जी-क्वाड्रुप्लेक्स-मोटिफ के बाहर उत्परिवर्तन मिरट्रॉन बायोजेनेसिस को प्रभावित नहीं करते हैं। इस प्रकार, हमारे निष्कर्ष संकेत देते हैं कि 5' भुजा में आरएनए जी-क्वाड्रुप्लेक्स स्तनधारी मिरट्रॉन के विकासवादी परिदृश्य में महत्वपूर्ण विनियामक तत्व हैं। इस ज्ञान को मानव-संक्रमित हर्पीज वायरस तक विस्तारित करने से पता चला कि हर्पीज सिम्प्लेक्स वायरस (HSV) के 50% से अधिक प्री-माइक्रोआरएनए में आरएनए जी-क्वाड्रुप्लेक्स होता है। आगे के विश्लेषण ने hsv1-mir-H27 और इसके एनालॉग hsv2-mir-H20 को संभावित मिरट्रॉन के रूप में पहचाना। द्रोषा नॉकआउट कोशिकाओं में hsv1-mir-H27 की अभिव्यक्ति ने प्रदर्शित किया कि प्री-माइक्रोआरएनए को ड्रोशा-स्वतंत्र तरीके से संसाधित किया जा सकता है, जिससे पुष्टि होती है कि यह वायरस-एनकोडेड माइक्रोआरएनए एक मिरट्रॉन है। hsv1-mir-H27 के लिए पूर्वानुमानित

सार

स्प्लिस साइट्स मिरट्रॉन के सफल स्प्लिसिंग की ओर ले जाती हैं। hsv1-mir-H27 प्री-माइक्रोआरएनए में आरएनए जी-क्वाड्रुप्लेक्स इन विट्रो में एक जी-क्वाड्रप्लेक्स संरचना बनाता है, और मानव मिरट्रॉन की तरह, 5' आर्म में आरएनए जी-क्वाड्रुप्लेक्स गठन ड्रोशा-स्वतंत्र स्प्लिसिंग की सुविधा देता है। 5' आर्म में आरएनए- जी-क्वाड्रुप्लेक्स का क्रमिक विघटन द्रोषा नॉकआउट कोशिकाओं में स्प्लिसिंग को रोकता है, जबकि आरएनए-जी-क्वाड्रुप्लेक्स-मोटिफ के बाहर उत्परिवर्तन मिरट्रॉन बायोजेनेसिस को प्रभावित नहीं करते हैं। यह दर्शाता है कि प्री-माइक्रोआरएनए की 5' आर्म में आरएनए जी-क्वाड्रुप्लेक्स गठन इसके स्प्लिसिंग में महत्वपूर्ण रूप से योगदान देता है। इस प्रकार, Hsv1-mir-H27 पहला मानव-संक्रमित वायरल मिरट्रॉन है जिसकी पहचान की गई है। इस कार्य ने NFE2L1 को hsv1-miR-H27 के एक वास्तविक लक्ष्य के रूप में प्रकट किया। ICP0 लोकस से उत्पन्न, hsv1-miR-H27 अभिव्यक्ति बढ़ी हुई कोशिकीय एपोप्टोसिस से जुड़ी है,कोशिका मृत्यु पथों को संशोधित करने में इसकी संभावित भूमिका को दर्शाता है। यह प्रो-एपोप्टोटिक कैस्पेज़ 3/7 गतिविधि को बढ़ाता है, जो hsv1-miR-H27-NFE2L1 अक्ष के माध्यम से HSV1-प्रेरित अपोप्टोसिस में भागीदारी का सुझाव देता है। इस प्रकार, hsv1-mir-H27 ICP0-प्रेरित सेलुलर अपोप्टोसिस के लिए जिम्मेदार संभावित अज्ञात गैर-कोडिंग कारक हो सकता है। हालाँकि, संक्रमण मॉडल की तुलना में hsv1-miR-H27 द्वारा प्रेरित अपोप्टोसिस न्यूनतम है, जिसका अर्थ है कि अतिरिक्त वायरल प्रोटीन या तनाव कारक इस प्रतिक्रिया को बढ़ा सकते हैं। इनमें ऑक्सीडेटिव तनाव, संक्रमण से प्रोटिओटॉक्सिक तनाव या कोशिका मृत्यु को बढ़ाने के लिए NFE2L1-PSMB6 मार्ग को प्रभावित करने वाले अन्य वायरस-व्युत्पन्न तत्व शामिल हो सकते हैं। अपोप्टोसिस प्रतिक्रिया में इन एम्पलीफायरों की भूमिका को पूरी तरह से समझने के लिए आगे की जांच की आवश्यकता है।

यह शोध स्तनधारियों और वायरस में मिर्ट्रोन जैवजनन को समझने में एक महत्वपूर्ण कदम है। यह छोटे आरएनए जीव विज्ञान में आरएनए- जी-क्वाड्रुप्लेक्स की अतिरिक्त भूमिकाओं को प्रकट करता है और मेजबान कोशिका प्रक्रियाओं के वायरल हेरफेर में hsv1-miR-H27 के लिए संभावित रूप से महत्वपूर्ण भूमिका का सुझाव देता है। ये निहितार्थ आरएनए जीव विज्ञान में हमारे निष्कर्षों के महत्व को रेखांकित करते हैं और मिर्ट्रोन की भूमिका, उनके प्रसंस्करण और जैवजनन, और सेलुलर प्रक्रियाओं को विनियमित करने में उनके डाउनस्ट्रीम प्रभावों में अनुसंधान के कई रोमांचक मार्गों का मार्ग प्रशस्त करते हैं।

Table of Contents

List of Figures

List of Tables

Annexures

List of Abbreviations

AFP	Alpha fetoprotein
AKT	Protein Kinase B
AREs	AU rich elements
ASSP	Alternative Splice Site Prediction
ATF6	Activating Transcription Factor 6
ATRX	Alpha thalassemia/mental retardation syndrome X linked
B-DNA	B form DNA
BLV	Bovine Leukemia Virus
BMVC	Benzimidazole Methyl Viologen Conjugate
BRACO-19	N,N′- (9- (4- (Dimethylamino)phenylamino)acridine-3,6-diyl)bis (3-(pyrrolidin-1-yl)propanamide) trihydrochloride
bZIP	Basic Leucine Zipper
CaMKII	Calcium/Calmodulin dependent Protein Kinase II
CASP8 mediate	Caspase 8 mediated
CD	Circular Dichroism
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A
cDNA	Complementary DNA
CMV	Cytomegalovirus
CNC-bZIP	Cap ‘n’ Collar basic Leucine Zipper

Abbreviations

CpG	Cytosine phosphate Guanine
cRNA	Complementary RNA
DBR1	Debranching Enzyme Homolog 1
DDX41	DEAD-Box Helicase 41
DEGs	Differentially Expressed Genes
DNA	Deoxyribonucleic Acid
DSC	Differential Scanning Calorimetry
EBV	Epstein Barr Virus
EcoRI	Escherichia coli Restriction Enzyme I
EDTA	Ethylenediaminetetraacetic Acid
FACS	Fluorescence Activated Cell Sorting
FAK	Focal Adhesion Kinase
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
GFP	Green Fluorescent Protein
GPI	Glycosylphosphatidylinositol
HSV-1	Herpes Simplex Virus Type 1
ICP0	Infected Cell Protein 0
ICP34	Infected Cell Protein 34.5

Abbreviations

ICP4	Infected Cell Protein 4
ICP6	Infected Cell Protein 6
IFN	Interferon
IL-6	Interleukin 6
LATs	Latency Associated Transcripts
MDV-1	Marek's Disease Virus Type 1
mRNA	Messenger RNA
MTA2	Metastasis Associated 1 Family Member 2
MUC1	Mucin 1
NaCl	Sodium Chloride
NFE2L1	Nuclear Factor Erythroid 2 Like 1
NMR	Nuclear Magnetic Resonance
OSM	Oncostatin M
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDS	Pyridostatin
PI3K	Phosphoinositide 3 kinase
PI3K-Akt	Phosphoinositide 3 kinase Protein Kinase B pathway
PIGT	Phosphatidylinositol Glycan Anchor Biosynthesis Class T

Abbreviations

PTEN	Phosphatase and Tensin Homolog
qPCR	Quantitative Polymerase Chain Reaction
RISC	RNA Induced Silencing Complex
RLU	Relative Light Units
RNA	Ribonucleic Acid
RNAi	RNA Interference
SNP	Single Nucleotide Polymorphism
SPR	Surface Plasmon Resonance
tRNA	Transfer RNA
TUTase	Terminal Uridylyltransferase
UTR	Untranslated Region
UV	Ultraviolet
VEGFA	Vascular Endothelial Growth Factor A
XhoI	Xanthomonas holcicola Restriction Enzyme I
YWHAE	Tyrosine 3 Monooxygenase/Tryptophan 5 Monooxygenase Activation Protein Epsilon
Z-DNA	Z form DNA