ISOLATION AND MOLECULAR CHARACTERIZATION OF PROMOTERS FROM *GORDONIA*

JANANEE JAISHANKAR

DEPARTMENT OF BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY DELHI

JUNE 2021

ISOLATION AND MOLECULAR CHARACTERIZATION OF PROMOTERS FROM *GORDONIA*

by

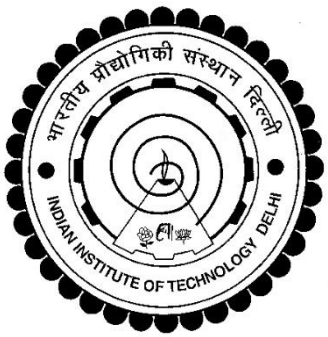
JANANEE JAISHANKAR

DEPARTMENT OF BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY

Submitted

In fulfilment of the requirements of the degree of Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

JUNE 2021

CERTIFICATE

This is to certify that the thesis titled "**Isolation and molecular characterization of promoters from *Gordonia***" being submitted by **Ms. Jananee Jaishankar** to the Indian Institute of Technology Delhi for the award of the degree of **Doctor of Philosophy** is a record of bonafide research work carried out by her under my supervision and guidance in conformity with the rules and regulations of Indian Institute of Technology Delhi.

The results presented in this thesis have not been submitted in part or full to any other University or Institute for the award of any other degree or diploma.

Date

Dr. Preeti Srivastava
Associate Professor
Department of Biochemical Engineering and Biotechnology
Indian Institute of Technology Delhi
New Delhi- 110016

ACKNOWLEDGEMENTS

There have been innumerable people who have helped me directly and indirectly towards my Ph.D. I would like to acknowledge everyone for their significant contribution in shaping my research career and helping me navigate tough times.

I am highly grateful to my supervisor, Prof. Preeti Srivastava for helping me grow both professionally and personally. I greatly admire her passion for science and her hard-working nature which have been one of the biggest motivations throughout my Ph.D. Words cannot express how her encouragement has given me the much-needed push when things were not going smoothly, while she also patiently answered my doubts. She has also provided me with several exciting opportunities to present my work. I will always remember the interesting discussions we had during our meetings and the sheer joy on her face whenever a paper got accepted or I received an award. I hope to carry forward the enthusiasm she has for research and teaching in my future endeavours.

I would like to express gratitude to my research committee members: Prof. Saroj Mishra, Prof. Biswajit Kundu and Prof. Shilpi Sharma for evaluating my work from time-to-time and giving valuable suggestions for improving my work. I would also like to thank Prof. D. Sundar, Head of Department, and other faculty members of DBEB for teaching interesting courses and offering support during the last five years.

I am extremely thankful to everyone of RNA-I lab for their help and assistance during the last five years. My seniors, Dr. Pooja Murarka and Dr. Divya Singhi have been immensely supportive both in managing the lab environment and for their suggestions related to work. I am also grateful to my good friend Rohit, for his continuous motivation and for cheering me up when things did not work out. Also, thanking him for all the times he has collected my samples during late night hours. I would definitely miss all the lunch and coffee time conversations with Rohit, Divya and Pooja ma'am.

I would further like to show my appreciation to the current members of RNA-I lab: Arif for being a great friend and good lab partner, Aditi for being such a lively person and for her help with Western Blotting, Deeksha for all the gossips and phone conversations about random things, Krishnapriya for her continuous encouragement, Shabnam for her understanding nature, Jyoti for her motivation, and Priyanka and Sourabh for their support.

I am also thankful to my mentees Jayaram, Prarthana and Abhishek; spending time discussing their work has given me new ideas for my own work. It was fun working with them. I would also like to acknowledge my other mentees who supported me in the lab. Thanks are also due to Ayushi, Bintu, Mansi, and other BTPs and MTPs for maintaining a lively environment in the lab.

Extending my gratitude to Dr. Pooja Singh, Ankur sir and Sabita for their help during the initial days in the lab.

Special thanks to Amit bhaiya for maintaining a clean lab environment and for his assistance in processing of bills.

I am also grateful to the members of DBEB and CBME labs for allowing me to use specific instruments. I would also like to thank Upma for her help in qPCR, Shashi for help with the bioreactor, and Srishti for her suggestions related to Western Blotting.

I cherish the time I have spent with amazing batchmates including Nudrat, Manju, Srishti, Rabab, Shashi, Rishabh, Arun, Navodit and Rohit, meeting in the hallway and supporting each other throughout has definitely been helpful.

I would also like to thank my husband Ashwin for being so supportive during the final stages of my thesis. His constant encouragement and advice helped me complete my doctorate successfully. I truly appreciate my in-laws for helping me stay calm and am thankful to them for their continuous encouragement and motivation.

I am highly grateful to the people without whom nothing in my life would have ever been possible: Mom, Dad and my extremely supportive sister Nandhini. I cannot express in words how much they mean to me, the way in which they always motivate me, listen to my never-ending rants, help me stay positive and persevere. My accomplishments are only because of their never-ending love, care and blessings. I am fortunate for their continuous belief in me which has helped me in the pursuit of excellence. Many many thanks to Nandhini, for taking most of the family responsibility on her shoulders and giving me a dose of positivity every time I require. I dedicate this thesis to them.

Special thanks to Paati (my grandmother), a person who has always believed I will do something great in life. I hope you are watching me happily from heaven and showering your blessings for my thesis submission.

Finally, I would like to express my humble gratitude to God, for always being my strength and guiding light.

Jananee Jaishankar

ABSTRACT

Novel regulatable promoters serve as effective tools for the construction of gene expression systems. In this aspect, promoters with varying strength and regulatability are required for use in different process-specific applications. Understanding the promoter functionality and characteristics is therefore essential for their exploitation. In this study, a previously isolated stationary phase promoter from *Gordonia* sp. IITR100 was engineered by repeated rounds of random mutagenesis to obtain a strong synthetic promoter, that offers stationary phase inducibility and higher strength as compared to the wild-type promoter in *E. coli*.

To isolate additional promoters from *Gordonia* sp. IITR100, the whole-genome sequence was determined, facilitating the identification of core promoter sequences. Further, the consensus sequence of the core promoter regions was deduced, which revealed the presence of conserved extended -10 region. Such a feature has also been reported in other members of Actinobacteria including *Corynebacterium* spp., *Rhodococcus* spp. and *Mycobacterium* spp. The whole-genome sequence also served as the basis for carrying out microarray profiling to identify the differentially expressed genes in the bacterium. This led to the isolation of two promoters: P*glx* and P*dsbA*, driving the expression of genes encoding glyoxalase and disulfide bond formation protein, respectively.

Microarray analysis also served as the basis for determining the expression pattern of 794 divergent gene pairs in *Gordonia* sp. IITR100. Several expression patterns were observed in these gene pairs and one promising gene pair controlled by divergent promoters was shortlisted for further experiments. Detailed molecular characterization of these two promising divergent promoters: P*maiA* and P*hyd* (driving the expression of genes encoding maleate cis-trans isomerase and hydantoinase, respectively) was carried out. A dual-reporter promoter probe vector has been constructed in the study, which was utilized to demonstrate the co-expression ability of these divergent promoters. Key aspects of their regulation were also

studied, which revealed their biological significance. The evidence of the transcription of genes driven by the truncated derivatives of promoter fragments under different conditions serves as useful tools for fine-tuning their expression levels.

Based on the results of promoters obtained in the study, two strong promoters: P*hyd* and synthetic promoter were used to develop useful gene expression systems for therapeutic proteins in Actinobacteria. A shuttle expression vector was constructed to demonstrate the expression of single-chain analog of insulin (SCI-57). Further, a novel method for the expression of two subunits of Human Chorionic Gonadotropin (hCG) was developed using the divergent promoters working in opposite orientations, which pave the way for using Actinobacterial hosts for therapeutic protein production.

Ultimately, this study has led to the isolation of diverse promoters with different efficacy. The characterization of divergent promoters is noteworthy and provides several insights into their applicability. More studies on the identification of such systems will offer alternative approaches for co-expression of genes.

सार

नवीन नियमन-योग्य प्रमोटर जीन अभिव्यक्ति प्रणालियों के निर्माण के लिए प्रभावी उपकरण के रूप में कार्य करते हैं। इस संदर्भ में विभिन्न प्रक्रिया-विशिष्ट अनुप्रयोगों में उपयोग के लिए विभिन्न शक्ति एवं नियामकता वाले प्रमोटरों की आवश्यकता होती है। प्रमोटर की कार्यक्षमता और विशेषताओं को समझना उनके उपयोग के लिए आवश्यक है। इस अध्ययन में गॉर्डोनिया प्रजाति IITR100 से पूर्व-पृथक्कृत स्थिर अवस्था प्रमोटर को एक मजबूत कृत्रिम प्रमोटर प्राप्त करने के लिए पुनरावृत्त यादृच्छिक उत्परिवर्तन द्वारा इंजीनियर किया गया जो *ई. कोलाई* में प्रमोटर के प्राकृतिक रूप की तुलना में स्थिर अवस्था प्रेरकता और उच्च शक्ति प्रदान करता है।

गॉर्डोनिया प्रजाति IITR100 से अतिरिक्त नवीन प्रमोटरों को अलग करने के लिए, संपूर्ण-जीनोम अनुक्रम निर्धारित किया गया, जिससे मुख्य प्रमोटर अनुक्रमों की पहचान की जा सकी। इसके आधार पर, मुख्य प्रमोटर अनुक्रमों का सर्वसम्मत अनुक्रम निकाला गया, जिससे संरक्षित विस्तारित -10 क्षेत्र की उपस्थिति का पता चला, जो एक्टिनोबैक्टीरिया के अन्य सदस्यों जैसे कि *कोरिनेबैक्टीरियम* प्रजातियों, *रोडोकोकस* प्रजातियों और *माइकोबैक्टीरिया* प्रजातियों में भी बताया गया है। माइक्रोएरे प्रोफाइलिंग द्वारा जीन की विभेदक अभिव्यक्ति प्रोफाइल का पता लगाया गया, जिसके कारण दो प्रमोटरों: P*glx* और P*dsbA*, की पहचान की गयी जो क्रमशः ग्लाइऑक्सालेज़ और डाइसल्फ़ाइड बंध निर्माता प्रोटीन को अभिव्यक्त करते हैं।

माइक्रोएरे प्रोफाइलिंग ने *गॉर्डोनिया* प्रजाति IITR100 में अपसरित जीन युग्म की अभिव्यक्ति प्रोफ़ाइल को निर्धारित करने के आधार के रूप में भी कार्य किया। अभिव्यक्ति के कई प्रतिरूप देखे गए और आगे के अध्ययन के लिए एक जीन युग्म का चयन किया गया। दो आशाजनक अपसरित प्रमोटरों: PmaiA और Phyd, जो क्रमशः मैलेट सिस-ट्रांस आइसोमेरेज़ और हाइडेंटोइनेज़ को अभिव्यक्त करने वाले जीन के अनुलेखन के लिए ज़िम्मेदार हैं, का विस्तृत निरूपण किया गया। अध्ययन में एक दोहरे

रिपोर्टर प्रमोटरधारी वाहक का निर्माण किया गया है, जिसका उपयोग अपसरित प्रमोटरों की सह-अभिव्यक्ति क्षमता को प्रदर्शित करने के लिए किया गया। उनके विनियमन के प्रमुख पहलुओं का भी अध्ययन किया गया, जिससे उनके जैविक महत्व का पता चला। विभिन्न परिस्थितियों में खण्डित किए गए प्रमोटर अंशों सहित अपसरित प्रमोटरों द्वारा संचालित जीन के अनुलेखन के प्रमाण उनके अभिव्यक्ति स्तरों को ठीक करने के लिए उपयोगी उपकरण के रूप में कार्य करते हैं।

अध्ययन में प्राप्त प्रमोटरों के परिणामों के आधार पर, दो मजबूत प्रमोटर: P*hyd* और कृत्रिम प्रमोटर का उपयोग एक्टिनोबैक्टीरिया में उपयोगी जीन अभिव्यक्ति प्रणाली विकसित करने के लिए किया गया। इन प्रमोटरों का उपयोग करते हुए एक अभिव्यक्ति वाहक का निर्माण इंसुलिन (SCI-57) के एकल-श्रृंखला अनुरूप की अभिव्यक्ति को प्रदर्शित करने के लिए किया गया था। विपरीत दिशाओं में काम कर रहे अपसरित प्रमोटरों का उपयोग कर मानव कोरियोनिक गोनाडोट्रोपिन (hCG) के दो उप-इकाइयों की अभिव्यक्ति के लिए एक नई विधि विकसित की गई, जो चिकित्सीय प्रोटीन उत्पादन के लिए एक्टिनोबैक्टीरियल मेजबानों का उपयोग करने का मार्ग प्रशस्त करती है।

अंततः, इस अध्ययन ने भिन्न-भिन्न प्रभावकारिता वाले नवीन प्रमोटरों को पृथक् किया है। अपसरित प्रमोटरों का अभिलक्षण उल्लेखनीय है और यह उनकी प्रयोज्यता में अंतर्दृष्टि प्रदान करता है। ऐसी प्रणालियों की पहचान पर आधारित अधिक अध्ययन जीन की सह-अभिव्यक्ति के लिए वैकल्पिक दृष्टिकोण प्रदान करेंगे।

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND SYMBOLS

Δ	Delta/deletion
µg	Microgram
µl	Microlitre
Amp	Ampicillin
APS	Ammonium per sulphate
ArgR	Arginine Repressor
ATCC	American Type Culture Collection
BLAST	Basic local alignment search tool
bp	Base pair
°C	degree Celsius
cAMP	Cyclic adenosine monophosphate
cDNA	Complimentary DNA
CRP	Catabolite repressor protein
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dNTP	Deoxy ribose nucleotide triphosphate
EB	Elution buffer
ECL	Enhanced Chemiluminescence
EDTA	Ethylene diamine tetra acetic acid
Fig.	Figure
GFP_{uv}	Green fluorescent protein
hCG	Human Chorionic Gonadotropin
HCl	Hydrochloric acid
HRP	Horseradish peroxidase
IPTG	Isopropyl β-D thio galactopyranoside
Kan	Kanamycin
kDa	kilo Dalton
LA	Luria agar
LB	Luria Broth
MCC	Microbial culture collection
min	Minute
ml	Millilitre
mM	Milimolar
MTCC	Microbial type culture collection

NCBI	National Center for Biotechnology Information
OD	Optical density
ONPG	O-nitrophenyl β-D galactopyranoside
Ori	Origami
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
P*dsbA*	Promoter transcribing gene encoding disulfide bond formation protein
P*glx*	Promoter transcribing gene encoding glyoxalase
P*hyd*	Promoter transcribing gene encoding hydantoinase
P*maiA*	Promoter transcribing gene encoding maleate cis-trans isomerase
PMSF	Phenylmethane sulfonyl fluoride
P*trc*	Trc promoter
PVDF	Polyvinylidene fluoride
qPCR	Quantitative PCR
qRT-PCR	Quantitative Reverse transcriptase PCR
RACE	Rapid amplification of c-DNA ends
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolution per minute
SCI-57	Single-chain insulin analog-57
SDS	Sodium dodecyl sulphate
SFI	Specific fluorescence intensity
TAE	Tris-glacial acetic acid-EDTA
TBST	Tris buffered saline containing Tween-20
Tet	Tetracycline
Tris	Tris (hydroxymethyl) amino methane
UTR	Untranslated region
UV	Ultraviolet
X-gal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
α	Alpha
β	Beta