DEVELOPMENT OF NOVEL EXPRESSION SYSTEMS IN *ESCHERICHIA COLI* AND CORYNEBACTERIA

by

MD. JAVED EQUBAL

Department of Biochemical Engineering and Biotechnology

Thesis submitted
in fulfillment of the requirements of the degree of
Doctor of Philosophy
to the

Indian Institute of Technology Delhi

JULY, 2013

DEDICATED TO

MY DEAR PARENTS

CERTIFICATE

*This is to certify that the thesis titled "**Development of Novel Expression Systems in Escherichia coli and Corynebacteria**" being submitted by **Md. Javed Equbal** to the Indian Institute of Technology, Delhi, for the award of degree of **Doctor of Philosophy**, has been prepared under my supervision and guidance in conformity with the rules and regulations of Indian Institute of Technology, Delhi. The research report and the results presented in this thesis have not been submitted to any other University or Institute for the award of any other degree or diploma.*

(Gopal Prasad Agarwal)
Professor
Deptt. of Biochemical Engineering
& Biotechnology
Indian Institute of Technology Delhi
New Delhi (India) – 110016.

ACKNOWLEDGEMENTS

The work described in this thesis has been done at Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi.

A journey is easier when you travel together. Interdependence is certainly more valuable than independence. This thesis is the result of six years of work whereby I have been accompanied and supported by many people. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

I am extremely grateful to the Department Doctoral Program Selection Committee for providing me this privilege to be part of this institute where I carried out my research work that has enabled me to reach to a point from where I could easily embrace the long and arduous road of higher education.

From the formative stages of this thesis to the final draft, I owe an immense debt of gratitude to my research supervisor ***Prof. Gopal P. Agarwal*** *for his keen interest and constant encouragement. His sound advice, incessant discussion, attention to details, untiring help and careful guidance are invaluable. He kept my spirits up and without his lifting me up when this thesis seemed interminable, I doubt it would ever have been completed. I am grateful to him for not only guiding me in scientific research but also for sharing his experience and knowledge with me and other aspects of scientific research. Despite his busy schedules, he always gave patient hearing to my work and his instant*

ideas used to take me out of big problems. I have learnt a lot through his intelligence and experience. His critical analysis and wide knowledge helped formulating this thesis in a good way. He has been a constant source of motivation, help and support. It is indeed a profound privilege to work under his distinguished guidance.

My sincere acknowledgement to demised Prof. J. K. Deb who played a key role in this PhD project. I am also very thankful to Dr. Preeti Srivastava for her valuable suggestions on different aspects of the thesis.

I am sincerely thankful to the members of my Scientific Research Committee Prof. Saroj Mishra, Prof. K. J. Mukherjee (JNU), and Prof. Tapan K. Chaudhuri (KSBS, IITD) for timely review of the research progress. Their keen interest, critical analysis and useful ideas provided good basis for research

My sincere thanks to Prof. Ruth Luthi-Carter for providing me an opportunity to work in her lab at EPFL, Switzerland as visiting fellow.

My sincere thanks go to a number of people who helped me in various ways. I am thankful to Mr. Rajiv Dahiya for providing me various materials and equipments to carry out work, to Amit for maintaining the lab. Mrs. Sunita Dang, Mr. Sitaram and Ms. Kirti Arora are thanked for handling important office matters. I am also thankful to Mrs. Neera Verma for providing me various books and protocol manuals required time to time in research.

I sincere thanks to my seniors Shaikh Ziauddin Ahammad, Vinod Pathak, Rupali Walia, Richa Baranwal, Raju Sankarayan, Mohd. Asif

Shah, and Bhawna Madan, and also to my batchmates and juniors Prabha Arya, Sucharita Sen, Gayathri Ramachandrane, Shweta Kamathan, Jyoti Batra, Swati Ojha, Muthu Mareeswaran, Satyendra Singh, Arun Kumar Shukla, Sunil Kumar, Bhuvanesh Aravind, Pooja Singh, Sikha Gupta, Anuj Yadav, Shilpi Jain, Ritesh Agarwal, Sharaddha Maiti, Neha Nagpal, Shivani Sharma, Rahul Agrawal, Neeti Kalyani, Tushar Das, Rashi Gupta, Younus Rather, Rishabh Gangwar, Surbhi Goyal, Mohd. Shazia Faridi, Krishna Saharan, for their supports in various ways.

I would also like to acknowledge the immense support and love provided by my friends Naushad Alam Siddiqui, Afzal Hingora, Maroof Ali, Ghulam Rather, Aqeel Ahmad, Jamilur Rahman, Intekhab Alam, Puneet Juneja, Anil Sangwan, Salahuddin Ahamad, Zuber Ali Ansari and Irfan Khursheed Shah, during this research work.

My vocabulary is not wide enough to reflect all my sense of regards and infinite gratitude to my parents Mr. Abdul Rahim and Mrs. Amna Khatoon for their love, affection, blessing, pivotal role in shaping my academic career, guidance for the right path, tremendous moral boosting and sacrifice which brought me to reach their dreams come true. I am and will ever be indebted to my family members Shahin Sultana, Yasmin Sultana, Nasrin Sultana, Md. Zainul Abedin, Mariyam Sultana, Roqaiya Sultana, Abdullah, Hamza Rahman, Abdul Hannan, Atiqur Rahman for their affection, support, care and blessings that enabled me to reach my present position.

Financial support in the form of fellowships by Department of Biotechnology and Indian Institute of Technology Delhi are duly acknowledged.

Regret and regards to those who have been close enough to be mentioned but not included by name in this acknowledgement due to hurry in writing. I also expect their grant of forgiveness and acknowledge their help and support.

*Above all, I am thankful to Almighty **ALLAH** who bestowed on me all the powers to do anything in life.*

Md. Javed Equbal

ABSTRACT

Bacteria are most commonly used expression systems for recombinant protein production. The gram-negative *Escherichia coli* has been the dominant organism used as a host for protein production both at laboratory and industrial levels. Many promoter systems have been developed for *E. coli* but a system based on T7 RNA polymerase dependent promoter is the most preferred one due to its ability to produce recombinant proteins at high levels. Few T7 based systems are available for *E. coli* but the existing ones are dependent on external inducers, viz. IPTG, L-arabinose, or heat. These external inducers are expensive and may impose toxic and deleterious effects on the host organisms.

Although, the usefulness of *E. coli* as an expression host cannot be underestimated, but this host is incapable of secretory expression of recombinant proteins. Also, the bacterial endotoxin makes it unsuitable for the production of therapeutic proteins else the endotoxin must be removed from the product through an expensive process. These limitations greatly impose a barrier on cost-effectiveness of therapeutic protein production in *E. coli*. In this regard, the industrially important GRAS species of corynebacteria viz. *Corynebacterium acetoacidophilum*, appear to be alternative hosts for recombinant protein production. But, despite many efforts, a strong promoter-based system in corynebacteria was not reported so far.

Thus, here a novel T7 RNA polymerase dependent two plasmid based system has been developed for *E. coli* and corynebacteria which gets auto-induced at a particular growth

phase. The system is auto-induced and therefore eliminates the requirement of addition of external inducer as well as monitoring of the culture.

In the two plasmid based *E. coli* system, one of the plasmids namely pJE-*fts*RNAP carried T7 RNA polymerase gene under the control of growth-phase dependent promoter, P-*ftsZ*, a Kan^R selection marker and p15A origin of replication. The second plasmid used was T7 promoter based pRSET-A which carried Amp^R selection marker and pUC origin of replication and therefore was compatible with pJE-*fts*RNAP. The foreign gene, *sf*GFP was cloned in this plasmid to get the recombinant plasmid pRSET-*sf*GFP. The expression of recombinant protein was successfully demonstrated in this system which was found to be auto-induced in a growth-phase dependent manner and thus was independent of external inducers. The expression levels of late log or stationary phase was found to be several fold higher than that of lag and/or early log phase. Thus the newly developed T7 system is a useful addition in the existing T7 expression systems of *E. coli*.

A T7 RNA polymerase and T7 promoter dependent two plasmid based system was developed in corynebacteria using two compatible corynebacterial replicons pBL1 and pCG1. The system was found to be auto-induced in a growth-phase dependent and external inducer independent manner. The shuttle plasmid, pJEC-*fts*RNAP, carried the T7 RNA polymerase gene under auto-induced *ftsZ* promoter. Then, a series of T7 promoter based shuttle plasmids were constructed. The *sf*GFP gene was cloned into T7 promoter based shuttle plasmid pBJ4 which was found to be compatible with pJEC-*fts*RNAP. The T7 promoter derived expression of *sf*GFP was successfully demonstrated

in *Corynebacterium* sp. and this is the very first report of a T7 expression system for *Corynebacterium* sp.

In continuation with that work the tightly regulated P-*ara*BAD based system was also developed for corynebacteria. Thus, the well-established tightly regulated *ara*BAD promoter of *E. coli* was exploited to develop an *ara*BAD promoter based system in corynebacteria. In this study, the *ara*BAD promoter based shuttle plasmid was constructed for corynebacteria and recombinant protein expression was subsequently demonstrated. Initially, the system was not found to be responsive with L-arabinose induction due to the lack to corresponding sugar transporter. This problem was overcome with the cloning of low-affinity *ara*E transporter of *E. coli* into *Corynebacterium* sp.

The expression systems established in the present study would be not only useful for high level production of recombinant proteins (T7 based) but also for studying physiological expression levels for metabolic engineering studies (P-*ara*BAD based).

CONTENTS