



UNIVERSITI PUTRA MALAYSIA

**DEVELOPMENT OF SECRETORY EXPRESSION SYSTEM USING
SYNTHETIC SIGNAL PEPTIDES *spA* AND *spD* FOR PROTEIN
PRODUCTION IN *Escherichia coli***

VITHYA A/P VELAITHAN

FBSB 2011 49

**DEVELOPMENT OF SECRETORY EXPRESSION SYSTEM USING
SYNTHETIC SIGNAL PEPTIDES spA AND spD FOR PROTEIN
PRODUCTION IN *Escherichia coli***

By

VITHYA A/P VELAITHAN

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,
in Fulfillment of the Requirement for the Degree of Master of Science**

September 2011

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Master of Science

**DEVELOPMENT OF SECRETORY EXPRESSION SYSTEM USING
SYNTHETIC SIGNAL PEPTIDES spA AND spD FOR PROTEIN
PRODUCTION IN *Escherichia coli***

By

VITHYA A/P VELAITHAN

September 2011

Chairman: Raha Abdul Rahim, PhD

Faculty: Biotechnology and Biomolecular Sciences

Recombinant protein expression is very important in biotechnology. Successful protein expression depends on the expression host, vector and target protein. *Escherichia coli* being a popular expression host, is still plagued by various problems in expression like formation of inclusion bodies, incorrect folding and low soluble protein yield. These problems can be circumvented by using different promoters, different host strains, co-expression of chaperones, reduction of culture temperature and secretion of proteins into periplasm and culture medium. In this study, periplasmic protein secretion was investigated by using novel synthetic signal peptides, spA and spD. The two signal peptides were designed based on signal peptides of *Bacillus* spp. for secretion of heterologous proteins. They were amplified and ligated to genes coding for the green fluorescent protein (GFP) and cyclodextrin glucanotransferase (CGTase) to construct secretion cassettes spA*GFP, spD*GFP and spA*CGT and

spD*CGT. These secretion cassettes were first cloned into pCR[®]-Blunt II-TOPO cloning vector. The cassettes were then sub-cloned into pET-32b(+) expression vector to construct pAGFP, pDGFP, pACGT and transformed into competent *E. coli* BL21(DE3) and BL21(DE3)pLysS cells. The cloning of secretion cassette spD*CGT into pET-32b(+) was not successful. GFP without the signal peptide was cloned into similar pET-32b(+) as a control and the construct was named pGFP. SDS-PAGE and western blotting results for recombinant GFP clones showed successful expression and secretion into the periplasm. Fluorescence analysis for GFP clones showed that the secreted GFP was not fluorescent while cytoplasmically expressed GFP was fluorescent. Induction temperature also affected the secretion of recombinant GFP as better secretion was attained at 37°C. SDS-PAGE analysis for recombinant clone BL-pACGT showed that CGTase was detected in both cytoplasm and periplasm but in Western blotting only cytoplasmic expression was detected. However, the positive control, (CGTase with native signal peptide) was detected in both cytoplasm and periplasm. Cell growth analysis for recombinant clones GFP and CGTase did not show any adverse effect to the secretion host. These results show that the synthetic signal peptides, spA and spD, could direct recombinant proteins to the periplasmic space.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk Ijazah Master Sains

**PEMBANGUNAN SISTEM EXPRESI SEKRETORI MENGGUNAKAN
PEPTIDA ISYARAT SINTETIK spA dan spD UNTUK PENGHASILAN
PROTEIN DI DALAM *Escherichia coli***

By

VITHYA A/P VELAITHAN

September 2011

Pengerusi: Raha Abdul Rahim, PhD

Fakulti: Bioteknologi dan Sains Biomolekul

Penghasilan protein rekombinan adalah penting untuk sektor bioteknologi. Penghasilan protein yang baik bergantung kepada factor-faktor seperti perumah, vektor serta protein yang ingin diekspreskan. Walaupun *E. coli* merupakan strain yang popular, ia menghadapi masalah seperti pembentukan badan inklusi, penghasilan struktur protein kuartenari yang tidak tepat, dan juga penghasilan protein pada tahap yang rendah. Masalah-masalah tersebut boleh diatasi dengan menggunakan kaedah seperti, penggunaan promoter berbeza, strain perumah berbeza, ko-ekspresi penghantar, pengurangan suhu kultur serta rembesan protein ke periplasma dan media.

Dalam kajian ini, rembesan protein dikaji dengan peptida isyarat sintetik spA dan spD. Peptida isyarat sintetik tersebut dihasilkan berdasarkan struktur asid amino peptida isyarat dari *Bacillus* sp. Peptida isyarat spA dan spD digandakan dan kemudiannya diligasikan dengan gen penanda GFP (green fluorescent protein) serta gen penanda

CGTase (cyclodextrin glucanotransferase) untuk menghasilkan kaset rembesan spA*GFP, spD*GFP, spA*CGT dan spD*CGT. Kaset-kaset itu diklonkan ke dalam vektor pengklonan, pCR[®]-Blunt II-TOPO. Seterusnya, kaset rembesan tersebut diklonkan ke dalam vektor ekspresi pET-32b(+) untuk menghasilkan konstruk pGFP, pAGFP, pDGFP serta pACGT. Konstruk tersebut dimasukkan ke dalam perumah *E. coli* BL21(DE3) dan BL21(DE3)pLysS. Kaset rembesan spD*CGT tidak dapat dimasukkan ke dalam perumah *E. coli* BL21(DE3) serta BL21(DE3)pLysS. Gen GFP tanpa peptide isyarat diklon ke dalam pET-32b(+) sebagai kawalan dan konstruk tersebut dinamakan pGFP. Keputusan kajian SDS-PAGE dan western blotting untuk klon rekombinan GFP menunjukkan bahawa pengekspresan serta rembesan GFP ke periplasma berjaya dilakukan. Analisis pendafluor juga menunjukkan bahawa rembesan GFP ke periplasma tidak menghasilkan pendafluor tetapi GFP di dalam sitoplasma menghasilkan pendafluor. Suhu induksi turut mempengaruhi rembesan GFP dan suhu 37°C menghasilkan rembesan yang baik. Kajian SDS-PAGE bagi klon rekombinan BL-pACGT menunjukkan bahawa CGTase berjaya dikesan di dalam sitoplasma serta periplasma tetapi western blotting cuma dapat mengesan CGTase di dalam sitoplasma. Kawalan positif (CGTase dengan peptide isyaratnya sendiri) dapat dikesan di dalam sitoplasma serta periplasma. Manakala, kajian pertumbuhan sel bagi klon rekombinan GFP dan CGTase tidak menunjukkan sebarang kesan tindakbalas negative terhadap perumah *E. coli*. Keputusan kajian menunjukkan bahawa peptida isyarat spA serta spD berjaya merembeskan protin penanda ke ruang periplasma.

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor Prof. Dr. Raha Abd. Rahim for giving me the opportunity to work under her supervision and to enhance my knowledge in this challenging field of study. I would not have succeeded if not for her guidance and support during difficulties. It has been a pleasure to work under her. I am also thankful for the guidance and support from my co-supervisors, Prof. Datin Paduka Dr. Khatijah Yusoff and Dr. Sieo Chin Chin.

Life as a master's student doing full research was only made bearable with the fantastic group of friends at the Microbial Biotechnology Lab. Failures and successes were shared and hard times were made easy. I am truly happy to have this group of fun, supportive friends and I am obliged to mention them here. So thank you, Shamsiah, Noreen, Ali, Adelene, Azmi, Tannaz, Abang Alai, Kak Yanti, Farzaneh, Shawal, Menaga, Kak Erni, Omid, Yee, and Sahar.

Last but not least, I am truly indebted to my family, my father Mr. Velaithan, mom, Mrs. Semila Devi, sister, Vidyarthini, uncle, Mr. Sivan and my aunt, Ms. Santha, who helped me get through these tough times. Failures are mine to learn and grow from but my successes belong to you.

APPROVAL

I certify that an Examination committee has met on date of viva voce to conduct the final examination of Vithya A/P Velaithan on her degree of Master of Science thesis entitled “Development Of Secretary Expression System Using Synthetic Signal Peptides spA and spD For Protein Production In *Escherichia coli*” in accordance with Universiti Putra Malaysia (Higher Degree) Act 1980 and Universiti Pertanian (Higher Degree) Regulations 1981. The committee recommends that the student be awarded the degree of Master of Science.

Members of the Examination Committee were as follows:

Name of Chairperson, PhD

Title

Name of Faculty

Universiti Putra Malaysia
(Chairman)

Name of Chairperson, PhD

Title

Name of Faculty

Universiti Putra Malaysia
(Internal Examiner)

Name of Chairperson, PhD

Title

Name of Faculty

Universiti Putra Malaysia
(Internal Examiner)

Name of Chairperson, PhD

Title

Name of Department/Faculty

Name of Organisation
(External Examiner)

....., PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia
Date:

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Raha Abdul Rahim, PhD

Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

Khatijah Yusoff, PhD

Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

Sieo Chin Chin, PhD

Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

BUJANG BIN KIM HUAT, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or other institutions.



VITHYA A/P VELAITHAN

Date: 20 September 2011

TABLE OF CONTENTS

	PAGE
ABSTRACT	ii
ABSTRAK	iv
ACKNOWLEDGEMENTS	vi
APPROVAL	vii
DECLARATION	ix
LIST OF TABLES	xiv
LIST OF FIGURES	xv
LIST OF ABBREVIATIONS	xix
CHAPTER	
1 INTRODUCTION	1
2 LITERATURE REVIEW	6
2.1 Recombinant protein expression	6
2.2 <i>Escherichia coli</i> as recombinant protein production host	7
2.3 Cytoplasmic and secretory protein expression	7
2.4 Protein secretion pathways in Gram-negative bacteria	8
2.5 Type II secretion system (Sec pathway)	9
2.5.1 Stage I (Targeting of preproteins)	10
2.5.1.1 SecB dependent targeting of preproteins	11
2.5.1.2 SRP dependent targeting of preproteins	13
2.5.2 Stage II (Translocation of preproteins)	13
2.5.3 Stage III (Release of preproteins)	15
2.6 Signal peptides	15
2.6.1 Signal peptides domain structure	16
2.6.2 Signal peptide conformation and membrane Interaction	17
2.6.3 Species-specific variations in signal peptides	18
2.6.3.1 Variations in N-region	19
2.6.3.2 Variations in hydrophobic region (H-region)	19
2.6.3.3 Variations in carboxyl region(C-region)	20
2.6.3.4 Significance of signal peptides variations	20
2.6.4 Signal peptide prediction method	20
2.7 Heterologous protein expression in <i>Escherichia coli</i>	22
2.7.1 Utilizing heterologous signal peptides for protein secretion	22

2.7.2	Enhancing periplasmic protein secretion	24
2.7.3	General strategies for optimizing of protein expression	25
2.7.3.1	Optimization of genetic elements in expression vector	26
2.7.3.2	Selection of suitable <i>Escherichia coli</i> expression host	29
2.7.3.3	Optimization of culture condition	30
2.8	Reporter proteins	31
2.8.1	Green fluorescent protein (GFP)	32
2.8.2	Cyclodextrin glucanotransferase (CGTase)	33
3	MATERIALS AND METHOD	35
3.1	Bacterial strains, plasmids and primers	35
3.2	Construction of synthetic signal peptides	36
3.3	In silico analysis of synthetic signal peptides, spA and spD	36
3.4	Preparation of <i>E. coli</i> glycerol stock	36
3.5	Plasmid DNA extractions from <i>E. coli</i>	37
3.6	Measurement of DNA concentration	38
3.7	Preparation of <i>E. coli</i> competent cells	38
3.8	Agarose gel electrophoresis of DNA	39
3.9	Restriction enzyme digestion and ligation	39
3.10	Transformation of <i>E. coli</i> TOP 10 and <i>E. coli</i> BL21(DE3) and BL21(DE3)pLysS	40
3.11	Purification of PCR products and plasmids	40
3.12	Polymerase chain reaction (PCR)	40
3.12.1	Polymerase chain reaction (PCR) amplification of GFP genes	41
3.12.2	Polymerase chain reaction (PCR) amplification of CGTase gene	41
3.12.3	Polymerase chain reaction (PCR) amplification of synthetic signal peptides	43
3.13	Construction of sp*reporter cassettes	43
3.14	Cloning of sp*reporter cassette in pCR [®] -Blunt II TOPO cloning vector	43
3.15	Sub-cloning of sp*reporter cassettes into expression vector, pET-32b(+)	44
3.16	Protein induction and extraction	45
3.16.1	Extracellular fraction	45
3.16.2	Periplasmic fraction	45
3.16.3	Cytoplasmic fraction	46
3.17	SDS-PAGE	46
3.18	Western Blot	48

3.19	Fluorescence analysis of recombinant GFP clones	49
3.20	Quantification of GFP	49
3.21	Quantification of CGTase	50
3.22	Cell growth analysis	50
4	RESULTS AND DISCUSSION	51
4.1	Construction of synthetic signal peptides	51
4.1.1	Signal peptide sequence alignment in ClustalW multiple alignment programme	51
4.1.2	Signal peptide prediction using SignalP 3.0 software	53
4.2	In silico analysis of new signal peptides spA and spD	57
4.2.1	Net charge, hydrophobicity and amino acid composition analysis	57
4.2.2	Analysis of synthetic signal peptides secondary structure conformation	62
4.3	Construction of signal peptide-reporter cassettes	63
4.3.1	PCR amplification of synthetic signal peptides spA and spD	63
4.3.2	PCR amplification of green fluorescent protein (GFP) gene	64
4.3.3	PCR amplification of cyclodextrin glucanotransferase (CGTase) gene	66
4.4	Construction of secretion cassettes	67
4.5	Cloning of sp*GFP cassettes into <i>E. coli</i> by pCR [®] -Blunt II-TOPO cloning vector	67
4.5.1	Verification of recombinant GFP clones by colony PCR	67
4.5.2	Plasmid extraction verification	70
4.5.3	Restriction enzyme digestion	71
4.6	Cloning of sp*CGT cassettes into <i>E. coli</i> by using pCR [®] -Blunt II-TOPO cloning vector	73
4.6.1	Plasmid extraction verification for recombinant CGT clones	73
4.6.2	Restriction enzyme digestion	74
4.7	Sub-cloning of GFP secretion cassettes into <i>E. coli</i> expression vector, pET- 32b(+)	75
4.7.1	Transformation into <i>E. coli</i> BL21(DE3)pLysS	75
4.7.2	Transformation into <i>E. coli</i> BL21(DE3)	78
4.8	Sub-cloning of CGT secretion cassette into <i>E. coli</i> expression vector, pET- 32b(+)	79
4.8.1	Transformation into <i>E. coli</i> BL21(DE3)	79
4.8.2	Transformation into <i>E. coli</i> BL21(DE3)pLysS	81

4.9	Protein expression analysis for recombinant clones	83
4.9.1	SDS-PAGE and western blot analysis of recombinant GFP protein from <i>E. coli</i> BL21(DE3)pLysS	83
4.9.2	SDS-PAGE and western blot analysis of recombinant GFP protein from <i>E. coli</i> BL21(DE3)	86
4.9.2.1	Effect of induction temperature and inducer concentration	88
4.9.2.2	Western blotting analysis for effect of post-induction temperature on GFP expression	92
4.9.2.3	Fluorescence study of recombinant GFP clones	94
4.9.2.4	Quantification of GFP expression	99
4.10	CGTase clones	101
4.10.1	SDS-PAGE and Western blot analysis of recombinant CGTase expression	101
4.10.2	Quantification of CGTase	106
4.11	Cell growth studies of recombinant clones	108
4.11.1	GFP clones	108
4.11.2	CGTase clones	110
5	CONCLUSION AND RECOMMENDATIONS	112
	REFERENCES	115
	APPENDIX	127
	BIODATA OF STUDENT	145