CHARACTERISTICS OF ELASTASE STRAIN K-REPEAT-IN-TOXIN FUSION PROTEIN OVEREXPRESSED FROM NEWLY CONSTRUCTED GENETIC TOOLS

NURUL HAZWANI BINTI SHAMSUDIN

UNIVERSITI PENDIDIKAN SULTAN IDRIS

2021

CHARACTERISTICS OF ELASTASE STRAIN K-REPEAT-IN-TOXIN FUSION PROTEIN OVEREXPRESSED FROM NEWLY CONSTRUCTED GENETIC TOOLS

NURUL HAZWANI BINTI SHAMSUDIN

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (BIOLOGY)

FACULTY SCIENCE AND MATHEMATICS
UNIVERSITI PENDIDIKAN SULTAN IDRIS

2021

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

UNIVERSITI PENDIDIKAN SULTAN IDRIS
SULTAN IDRIS EDUCATION UNIVERSITY

Please tick (√)
Project Paper
Masters by Research
Master by Mixed Mode
PhD

INSTITUTE OF GRADUATE STUDIES

DECLARATION OF ORIGINAL WORK

This declaration is made on the 9th day of Dec 20 21

i. Student's Declaration:

I, Nurul Hazwani Binti Shamsudin, P20131001201, Faculty of Science and Mathematics (PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled Characterization of Elastase Strain K-Repeat-in-Toxin Fusion Protein Overexpressed from Newly Constructed Genetic Tools is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

Signature of the student

ii. Supervisor's Declaration:

I Associate Prof. Dr. Muhammad Aqil Aryan Wong (SUPERVISOR'S NAME) hereby certifies that the work entitled Characterization of Elastase Strain K-Repeat-in-Toxin Fusion Protein Overexpressed from Newly Constructed Genetic Tools (TITLE) was prepared by the above named student, and was submitted to the Institute of Graduate Studies as a * partial/full fulfillment for the conferment of Doctor of Philosophy (Biology) (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.

Date

Signature of the Supervisor

INSTITUT PENGAJIAN SISWAZAH /
INSTITUTE OF GRADUATE STUDIES

BORANG PENGESAHAN PENYERAHAN TESIS/DISERTASI/LAPORAN KERTAS PROJEK
DECLARATION OF THESIS/DISSERTATION/PROJECT PAPER FORM

Tajuk / *Title*: Characterization of Elastase Strain K-Repeat-in-Toxin Fusion Protein Overexpressed from Newly Constructed Genetic Tools

No. Matrik /*Matric's No.*: P20131001201

Saya / *I* : Nurul Hazwani Binti Shamsudin

(Nama pelajar / *Student's Name*)

mengaku membenarkan Tesis/Disertasi/Laporan Kertas Projek (Kedoktoran/Sarjana)* ini disimpan di Universiti Pendidikan Sultan Idris (Perpustakaan Tuanku Bainun) dengan syarat-syarat kegunaan seperti berikut:-
acknowledged that Universiti Pendidikan Sultan Idris (Tuanku Bainun Library) reserves the right as follows:-

1. Tesis/Disertasi/Laporan Kertas Projek ini adalah hak milik UPSI.
   The thesis is the property of Universiti Pendidikan Sultan Idris
2. Perpustakaan Tuanku Bainun dibenarkan membuat salinan untuk tujuan rujukan dan penyelidikan.
   Tuanku Bainun Library has the right to make copies for the purpose of reference and research.
3. Perpustakaan dibenarkan membuat salinan Tesis/Disertasi ini sebagai bahan pertukaran antara Institusi Pengajian Tinggi.
   The Library has the right to make copies of the thesis for academic exchange.
4. Sila tandakan (√) bagi pilihan kategori di bawah */ Please tick (√) for category below:-*

☐	**SULIT/*CONFIDENTIAL***	Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub dalam Akta Rahsia Rasmi 1972. / *Contains confidential information under the Official Secret Act 1972*
☐	**TERHAD/*RESTRICTED***	Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan ini dijalankan. / *Contains restircted information as specified by the organization where research was done.*
☑	**TIDAK TERHAD / *OPEN ACCESS***	

(Tandatangan Pelajar/ Signature)

(Tandatangan Penyelia / *Signature of Supervisor*)
& (Nama & Cop Rasmi / *Name & Official Stamp*)

Tarikh: _____

Catatan: Jika Tesis/Disertasi ini **SULIT** @ **TERHAD**, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan ini perlu dikelaskan sebagai **SULIT** dan **TERHAD**.

Notes: If the thesis is CONFIDENTAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction.

ACKNOWLEDGEMENT

All praises to Allah the Almighty for giving me the strength, health, passion and patience to complete this research project.

I would like to express my heartfelt gratitude to my supervisor, Dr. Wong Chee Fah, who constantly conveyed a positive spirit of adventure in regards to research and give me so much advise to improve myself in all aspects. I would not been possible to complete this project without his persistent help, invaluable guidance and overwhelming kindness.

I would also like to express a deepest appreciation to my co-supervisor, Prof. Dr. Raja Noor Zaliha Raja Abd. Rahman and Dr. Mohd Shukuri Mohamad Ali for their great assistance, insightful discussion and suggestions.

A million thanks to all EmTech members, Dr. Ely, Suzana, Adura Syamimi, Farhani, Nur Shakila, Zarir, Aisyah, Fatin Farziana, Nadzmi, Wahidah, Atie, Asmah and Aswhini who have been extraordinarily helpful, tolerant and supportive. This beautiful friendship is a priceless treasure and will be cherished forever.

A deepest appreciation to my beloved family and husband for their unduly love and prayers. Thank you for being unconditionally supportive and understanding throughout this challenging study period. Words can't describe how grateful and thankful I am to you in my life.

And for those who involved directly or indirectly in this project, I sincerely thank you.

ABSTRACT

This study was intended to construct a new *Escherichia coli-Pseudomonas* shuttle vector for overexpression of elastase strain K in both *E. coli* and *Pseudomonas* as well as for rapid purification using new RTX-tag. A 6.5 kb novel shuttle vector, designated as pSIT/RTX, was constructed from pCon2(3) as to improvise the expression of pCon2(3). pSIT/RTX was employed with tightly regulated promoter $P_{T7(A1/O4/O3)}$ for controlling gene expression, stabilizing fragment (SF) for replication and maintenance of plasmid in *E. coli* and *P. aeruginosa*, *attB* gene for genome integration, elastase strain K as passenger enzyme and RTX-tag which is located at C-terminal for rapid purification. *E. coli* TOP10/pSIT/RTX was chosen to proceed with purification as the highest amount of proteolytic activity was detected at 12 h after incooperation with 0.6 mM IPTG (Isopropyl β- d-1-thiogalactopyranoside). Elastase strain K-RTX fusion protein was purified using Ca^{2+} as novel ligand in immobilized-metal affinity chromatography (IMAC) with 28 % recovery and 3.8 fold. The estimated molecular weight as observed on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was 73 kDA, with optimal temperature and pH were 40^{o}C and pH6, respectively. The proteolytic activity was significantly enhanced by increasing the concentration of Na^{+} and Cu^{2+} ions and more stable in phenylmethylsulfonyl fluoride (PMSF), Tween20 and Triton-X-100. On the downside, Ni^{2+}, Zn^{2+}, n-dodecane, n-tetradocane, dithiothreitol (DTT) and SDS showed strong inhibition on the proteolytic activity. Elastase strain K exhibited preference towards 25% (v/v) of Dimethyl sulfoxide (DMSO), methanol and pyridine as their uniqueness as an organic solvent tolerant enzyme. Congo-red as the most specified substrate for elastase recorded the lowest release of its product. As a concluding remark, experimental work conducted in this study had indeed highlighted several achievements, novelties and findings including construction of vectors which had led to the overexpression of elastase strain K by constructed vectors, the using of RTX-tag for purification via IMAC and most importantly, the remarkable stability of elastase strain K in hydrophilic organic solvents.

PENCIRIAN ELASTASE STRAIN K MENGULANGI DALAM TOKSIN PROTIN LAKURAN PENGEKSPRESAN DARI PEMBINAAN ALAT GENEETIK BARU

ABSTRAK

Tumpuan utama kajian bertujuan untuk membina vector ulang alik *Escherichia coli-Pseudomonas* yang baru bagi tujuan pengekspresan elastase strain K di dalam *E. coli* dan *Pseudomonas* dan bagi penulenan pantas menggunakan RTX-tag baru. 6.5 kb Vektor ulang-alik baru, ditetapkan sebagai pSIT/RTX dibina dari pCon2(3) bagi penambaikkan penekspresan pCon2(3). pSIT/RTX mengandungi promoter kawalan ketat, $P_{T7(A1/O4/O3)}$ untuk meningkatkan pengekspresan gen, serpihan pnstabilan (SF) untuk mereplikasi and mengekalkan plasmid di dalam *E. coli* dan *Pseudomonas aeruginosa*, gen *attB* bagi integrasi genom, elastase strain K sebagai enzim penumpang dan RTX-tag an terletak di terminal C untuk penulenan pantas. *E. Coli* TOP10/pSIT/RTX telah dipilih untuk meneruskan kajian penulenan setelah aktiviti proteolitik yang tertinggi di kesan selepas 12 jam dengan kerjasama 0.6 mM IPTG (Isopropyl β- d-1-thiogalactopyranoside). Elastase strain K-RTX ditulenkan menggunakan Ca^{2+} sebagai ligan terbaru dalam komatografi keafinan logam tidak bergerak (IMAC) dengan 28 % pemulihan dan 3.8 lipatan. Berat molekul yang dianggarkan pada SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) ialah 73 kDA. Optimum suhu dan pH adalah masing-masing 40ºC dan pH6. Aktiviti proteolitik adalah ketara dengan peningkatan kepekatan ion Na^{+} dan Cu^{2+} dan lebih stabil pada PMSF (phenylmethylsulfonyl fluoride), Tween20 dan Triton-X-100. Disebaliknya, Ni^{2+}, Zn^{2+}, n-dodekana, n-tetradocane, DTT (dithiothreitol) dan SDS menunjukan kesan penceratan yang kuat pada aktiviti proteolitik. Elastase strain K menunjukan keutamaan terhadap 25 % (v/v) daripada DMSO (Dimethyl sulfoxide), methanol, piridina kerana keunikan sebagai enzim pelarut organic. Congo-red sebagaimana yang dinyatakan substrat bagi elastase merekodkan pelancaran produk yang terendah. Kesimpulannya, eksperimen yan dijalankan dalam kajian ini telah mencapai beberapa pencapaian dan penemuan baru termasuklah pembinaan vector yang mampu mengekspresikan elastase strain K, pengunaan ta RTX bagi proses penulenan melalui IMAC dan yang penstabilan elastase strain K di dalam pelarut organik hidrofilik.

CONTENTS

CHAPTER 3 METHODOLOGY

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVATIONS

%,	percent
°C	degree Celcius
A	Ampere
A_{280}	absorbance at 280 nm
A_{440}	absorbance at 440 nm
A_{4513}	absorbance at 513 nm
A_{495}	absorbance at 495 nm
A_{600}	absorbance at 600 nm
bp	base pair
dH_2O	distilled water
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
g	gram
g/L	gram per litre
h	hour
IPTG	isopropyl-β-D-galactoside
kb	kilobase pair
kDa	kiloDalton
L	litre
LB	Luria Bertani
M	molar
mg	milligram
mg/mL	milligram per milliliter

min	minute
mL	milliliter
mM	millimolar
MW	molecular weight
nm	nanometer
OD_{600}	Optical density at 600 nm
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PMSF	phenylmethylsulfonyl fluoride
rpm	rotation per minute
RTX	Repat-in-toxin
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
T_m	thermal denaturation
U	unit
U/mL	unit per milliliter
V	volt
v/v	volume per volume
w/v	weight per volume
µg	microgram
µL	microlitre
µm	micrometer

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

The primary goal of gene vector construction is to tune vector expression to the maximum level of protein production while minimizing the input cost and time. Nonetheless, in many circumstances, lesser protein production than maximum production is desirable. At a certain point, proteins could become harmful and toxin to the growth of the cell which indirectly make protein ineffectively folded and unable to express at high levels. However, a rapid development in molecular biology field has now made vector is possible to overexpress at desired points.

Presently, Escherichia coli is the most common expression system, used in recombinant protein production due to the well-characterized genome and a variety of mature tools obtainable for genetic manipulation (Liu et al., 2013). Unfortunately,

genes from Pseudomonas aeruginosa and other Pseudomonas host cannot be expressed at an adequate level in E. coli (Raftari et al., 2013). The ability for cloning vector to be able to replicate in both E. coli as well as Pseudomonas is due to the two different origins of replication (Maj et al., 2013) which are rep (pMB1) that actives in E. coli and stabilizing fragment (SF) for Pseudomonas (Wong et al., 2017; Riley et al., 2013). Therefore, in this research, stabilizing fragment (SF) from the pUCP19 multiple cloning sites (MCS) was used as it conserved all of the needed features including: (i) blue-white screening for recombinant, (ii) the employment of commercial available primer for sequencing and PCR amplification of inserts and (iii) regulated expression from Plac.

The presence of mini-stabilization fragment (mSF) in pSS124 was useful in controlling the cloning, sequencing and expression of gene (Riley et al., 2013). The presence of 1.8kb *Pst*I stabilizing fragment (SF) of pRO1614 becomes an essential genetic tool in gene manipulation. Wood and Ohman (2015) reported that pBR322 which containing stabilizing fragment (SF) was not only able to replicate in *E. coli* and *P. aeruginos*a, but also in *P. putida*, *P. fluorescens* and *Klebsiella pneumoniae*.

Besides SF, an insertion of strong promoter can optimize the yield of gene expression (Pitzer et al., 2016). Considering that, well-studied promoter that suitable for a high level of protein expression in *E. coli* was selected for this research. $T7_{(A1/O4/O3)}$ consists of *A1* which is known as a stronger promoter compared with other commonly used promoter including *E. coli* operons P_{lac}, P_{tac} and P_{bla} (Haris, 2014). The strength modified promoter $T7_{(A1/O4/O3)}$, used in this project was designed based on segments originally created by Lanzer and Bujard (1988). Lanzer and Bujard (1988) constructed a suite of promoter by inserting *lacI* operator sequence into two sites; first site was a

sequence that consisted of 29 bp native *E. coli* lac operon in position homologous to P_{lac} *(O3)* and the other one was *O4* which was a sequence carried 17 bp core region of the wild type lac operator as spacer between -10 and 133 hexamer. In consequence of the specific interaction of operator/promoter region and DNA polymerase, the spacing between -10 and -35 is generally tightly conserved (Fulcrand et al., 2016). Comprehensively, $T7_{(A1/O4/O3)}$ promoter is reported to be repressed tightly because it contains two *lac* operator sites for binding of Lac repressor expressed by $lacI^q$ (Lanzer and Bujard, 1988).

According to Zhang and colleges (2016), gene fusion technology offers an analyzing expression of many genes encoded by bacterial genome in order to maintain the plasmid in the particular host. In order to enhance the integration of vector in the *Pseudomonas*, a certain gene is needed. For example, specialized mini-CTX vectors containing the insertion of gene cassette of 30 bp *attB* sequence have been used to construct *P. aeruginosa* host strains allowing regulated expression from *T7* and *lac* promoters and also for studying gene expression using *lac* and *lux*-based reporter genes (Gilbertsen and Williams, 2014).

Nowadays, scientist is eager to over-express their protein in order to facilitate the next steps in research like purification, crystallization and protein structure. Therefore, in this research, as to enhance rapid purification, repeat-in-toxin (RTX-tagged) was added to the construct. RTX-tagged belongs to the Family 1.3 lipase. Family 1.3 lipase is composed of 617 amino acid residues (Cheng et al., 2014) that distinguish the residue from other lipases not only in amino acid sequence but also in their secretion mechanism (Morgan et al., 2017, Chenal et al., 2015). Family 1.3 lipase

is classified under type 1 secretion system (T1SS) (Cheng et al., 2014). This make the family to have the ability to be secreted by a three components ATP-binding cassette (ABC) transporter system (one step pathway) (Morgan et al., 2017, Chenal et al., 2015). Protein that secreted under ABC transporter has several repeats of nine-residue GGXGXDXUX sequence motif (X: amino acid; U: hydrophobic residue) at C-terminal secretion signal (Sotomayor-Pérez et al., 2015). The first six residues of this motif form a loop and each Ca^{2+} ion binds between pairs of the loop. These repeated form a β-roll structure motif (Sotomayor-Pérez et al., 2015)

Resulted from the characteristics, a construction of genetic and molecular tools consists of: (1) 1.8 kb *Pst*I stabilizing fragment (SF) (ATCC 87110) for replication and maintenance of plasmid in both *E. coli* and *P. aeruginosa*, (2) a tightly regulated $T7_{(A1/O4/O3)}$ promoter/repressor system for control of gene expression (Wong et al., 2017), (3) the 30 bp *attB* sequence for integration of gene cassettes into *Pseudomonas* genome (Yu et al., 2014), and (4) RTX -tag for purification purpose is done.

1.2 Problem Statements

Recently, various techniques have been developed for over-expressing and purifying the desired recombinant protein in a bulk quantity due to the enormous potential for biotechnology application (Al-Hejin et al., 2018). Yet, the progress of technique development is still continuously being made due to some problems often found throughout the process such as poor growth of the host, inclusion body (IB) formation, protein inactivity and even the failure to obtain any protein at all.

In 2017, Wong and colleagues reported that constructed vector known as pCon2(3), which harboured two expression cassettes was unable to exhibit greater protease activity in *E. coli*. However, in *Pseudomonas aeruginosa*, PAO1, pCon2(3) was able to overexpress greater elastinolytic activity. This is consequence from fact that pCon2(3)/*E.coli*TOP10 is capable to utilize *Lac* promoter for protein expression only.

Therefore, the improvised shuttle expression vector is design in this study in order to achieve a higher expression of protease in both *E. coli* and *Pseudomonas* strains. The new construct pSIT/RTX will greatly assist by modified *T7* promoter/repressor, $T7_{(A1/O4/O3)}$. The fact that tightly regulated *T7* promoter system consist of two lac promoter sites, *O3* and *O4* make it efficiently in controlling the expression level (Haris, 2014) and important to avoid leaky expression (Osasumwen, 2017). This may result in higher rate of promoter clearance that lower the occupancy of the promoter by RNA polymerase, thereby increase the repression factor (Osasumwen, 2017).

Another focus of this study is to visualize the functionality and specificity of new and novel RTX tag in purify recombinant protein. Thus, a widely employed method utilizes immobilized metal-affinity chromatography (IMAC) is employ for single-step purification to attain relatively more than 80 % pure protein (Hong et al., 2017). IMAC interaction mechanism is based on interaction between affinity-tag consisting of polyhistidine residues (Sun et al., 2014; Cimen et al., 2016)) and transition metal ions like Co^{2+}, Ni^{2+}, Cu^{2+}, Ca^{2+} and Zn^{2+}. Therefore, modified purification step should be develop to suit with target protein and RTX-tag.

1.3 Objective of the Research

The specific objectives of the project are:

1) To construct a new *Escherichia coli-Pseudomonas* shuttle expression vector
2) To overexpress recombinant protein from constructed vector *in E. coli* and *Pseudomonas*
3) To purify and characterize recombinant protein using a new and novel RTX-tag