UNIVERSITI PUTRA MALAYSIA

EXPRESSION OF CHALCONE SYNTHASE AND CHALCONE ISOMERASE FROM Polygonum minus Huds. IN Escherichia coli, GENE CHARACTERIZATION, AND STRUCTURE ELUCIDATION OF CHALCONE ISOMERASE

FATIN LYANA BINTI AZMAN SHAH

FBSB 2021 8

EXPRESSION OF CHALCONE SYNTHASE AND CHALCONE ISOMERASE FROM *Polygonum minus* Huds. IN *Escherichia coli*, GENE CHARACTERIZATION, AND STRUCTURE ELUCIDATION OF CHALCONE ISOMERASE

By

FATIN LYANA BINTI AZMAN SHAH

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

December 2020

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

EXPRESSION OF CHALCONE SYNTHASE AND CHALCONE ISOMERASE FROM *Polygonum minus* Huds. IN *Escherichia coli*, GENE CHARACTERIZATION, AND STRUCTURE ELUCIDATION OF CHALCONE ISOMERASE

By

FATIN LYANA BINTI AZMAN SHAH

December 2020

Chairman : Associate Professor Suriana binti Sabri, PhD
Faculty : Biotechnology and Biomolecular Sciences

Flavonoids are commonly found in plants and possessed vast benefits particularly towards the improvement of human health. These compounds that are attracting scientific attention, have been traditionally extracted from plant sources and synthesized chemically. Nevertheless, methods such as heat reflux and Soxhlet extractions are commercially infeasible, laborious, waste plant sources, and are environmentally unsafe due to the use of toxic solvents.

Polygonum minus Huds. is a household-plant with flavonoids as one of its major compounds, hence engineering a plant's phenylpropanoid pathway into microbial hosts for flavonoid production is an option, where these metabolites are produced without facing the issues mentioned. This study aims to identify and isolate chalcone isomerase (CHI) from *P. minus*, to express chalcone synthase (CHS) and CHI in *Escherichia coli*, and to predict the three-dimensional (3D) structure of CHI protein from *P. minus*. The Open Reading Frame (ORF) of CHI gene was isolated through PCR amplification from *P. minus* cDNA. Various bioinformatics analyses were carried out to analyze the identified *Pm*CHI protein sequence. Then, CHS and CHI genes from *P. minus* were codon-optimized, synthesized, and individually cloned into an expression vector.

The recombinant vectors were later transformed into *E. coli* BL21(DE3). Temperature differences applied during the growth of recombinant *E. coli* BL21(DE3) expressing these enzymes were investigated, followed by western blot analysis and partial purification of these enzymes. As the 3D structure of *Pm*CHS was predicted previously, Yet Another Scientific Artificial Reality Application (YASARA) software was used to construct the *Pm*CHI model by

using the CHI protein model of *Deschampsia antartica* Desv. (PDB ID: 5YX4) as template. The predicted model was analyzed, and previously theorized active site residues were highlighted. Then, the predicted model was validated by using validation programs and servers available online.

ORF of *PmCHI* was identified to be ~700-bp long, encoding for 236 amino acid residues, with a predicted molecular weight of 25.4 kDa. Conserved residues involved in the active site cleft of CHI enzyme group are present in *Pm*CHI protein sequence and was identified to belong to the class of Type I of CHI. *Pm*CHI also comprises of more than 50% hydrophobic residues, lacks any signal peptide and transmembrane helices. The expression of CHS and CHI were observed at 16 °C and succeeding partial purification showed the potentials of these proteins to be fully purified for enzyme characterization studies.

The Ramachandran plot showed that 93.9% of residues in the predicted *Pm*CHI model were present in the most favored regions. Likewise, Verify3D and ERRAT showed that the 3D model of *Pm*CHI gave values that are within the acceptable range of a good structure. As a conclusion, this study would provide more information on *Pm*CHI and *Pm*CHS enzymes, that can be incorporated into the flavonoid-producing pathway and engineered into a recombinant host, where significant yields of flavonoid compounds are expected.

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

PENGEKSPRESIAN KALKON SINTASE DAN KALKON ISOMERASE DARIPADA *Polygonum minus* Huds. DI DALAM *Escherichia coli*, PENCIRIAN GEN DAN PENENTUAN STRUKTUR KALKON ISOMERASE

Oleh

FATIN LYANA BINTI AZMAN SHAH

Disember 2020

Pengerusi : Profesor Madya Suriana binti Sabri, PhD
Fakulti : Bioteknologi dan Sains Biomolekul

Flavonoid yang biasanya dijumpai dalam tumbuhan, mempunyai banyak kebaikan terutamanya terhadap penambahbaikan kesihatan manusia. Sebatian yang menarik perhatian saintifik ini, secara tradisionalnya diekstrak daripada tumbuhan dan disintesis secara kimia. Kaedah seperti pengekstrakan menggunakan refluks haba dan Soxhlet, berkos tinggi, membazirkan masa dan sumber tumbuhan dan menggunakan pelarut toksik berbahaya terhadap alam sekitar.

Polygonum minus Huds. ialah tumbuhan yang mempunyai flavonoid sebagai antara sebatian utama, jadi kejuruteraan laluan fenilpropanoid daripada tumbuhan ke dalam perumah mikrob untuk penghasilan flavonoid adalah satu pilihan, tanpa masalah yang disebut. Misi kajian ini adalah untuk mengecam dan mengasing kalkon isomerase (CHI) daripada *P. minus*, untuk mengekspres kalkon sintase (CHS) dan CHI di dalam *Escherichia coli* dan untuk meramal struktur tiga-dimensi (3D) protein CHI daripada *P. minus*. Rangka baca terbuka (ORF) gen *CHI* telah diasingkan melalui amplifikasi PCR daripada cDNA *P. minus*. Analisis bioinformatik telah dilakukan untuk menganalisis jujukan protein *Pm*CHI yang dikenalpasti. Gen *CHS* and *CHI* daripada *P. minus* telah dikodon-optimum, disintesis dan diklon secara berasingan ke dalam vektor pengekspresan.

Vektor-vektor rekombinan ini seterusnya ditransformasi ke dalam *E. coli* BL21(DE3). Perbezaan suhu ketika pertumbuhan *E. coli* BL21(DE3) rekombinan yang mengekspres enzim-enzim ini juga telah dikaji, dianalisis menggunakan blot western dan penulenan separa enzim-enzim juga dijalankan. Dengan kejayaan ramalan struktur 3D *Pm*CHS sebelum ini, Perisian Yet Another

Scientific Artificial Reality Application (YASARA) telah difokuskan untuk membina model *Pm*CHI menggunakan model CHI *Deschampsia antartica* Desv. (PDB ID: 5YX4) sebagai templat. Model ini dianalisis dan penumpuan terhadap residu di tapak aktif yang diteorikan telah dilakukan. Model yang diramalkan telah dinilai menggunakan program dan pelayan yang terdapat dalam talian.

ORF *Pm*CHI berukuran ~700-bp, mengenkodan 236 residu asid amino dan diramalkan mempunyai jirim molekul 25.4 kDa. Jujukan protein *Pm*CHI mempunyai residu terlindung yang hadir dalam klef tapak aktif kumpulan enzim ini dan tergolong di dalam kelas Jenis I bagi enzim CHI. *Pm*CHI mempunyai lebih 50% residu hidrofobik, tiada isyarat peptida dan heliks transmembran. Pengekspresan enzim CHS dan CHI dilihat pada 16 °C dan penulenan separa menunjukkan potensi mereka untuk ditulenkan sepenuhnya bagi kajian tentang sifat protein.

Plot Ramachandran menunjukkan 93.9% residu dalam model *Pm*CHI terletak dalam kawasan paling terbaik. Program Verify3D dan ERRAT juga menunjukkan model 3D *Pm*CHI memberi nilai yang diterima untuk struktur yang baik. Kesimpulannya, kajian ini menyediakan lebih informasi tentang enzim *Pm*CHI dan *Pm*CHS, yang boleh dimasukkan ke dalam laluan yang menghasilkan flavonoid dan perumah rekombinan, di mana penghasilan flavonoid pada kadar yang tinggi dijangkakan.

ACKNOWLEDGEMENTS

Firstly, all praise be to Allah for giving me a good health and the opportunity in completing my postgraduate study. I would like to express my sincerest gratitude towards my supervisor, Associate Professor Dr. Suriana Sabri, for her patience, wisdom, and guidance during this research journey. The ideas, suggestions and positive encouragements that she had given were among the reasons for the completion of this study. Only God can repay all the of kindness that she has shown to me.

I would also like to express my deepest appreciation to my supervisory committee, Associate Professor Dr. Syarul Nataqain Baharum, Associate Professor Dr. Goh Hoe Han, and Associate Professor Dr. Siti Nurbaya Oslan for their continuous support and generous contributions along the journey. I would like to convey my biggest gratitude and thankfulness to Associate Professor Dr. Adam Leow Thean Chor and Dr. Ahmad Bazli Ramzi for their immense support and invaluable guidance in solving issues faced whilst completing this study.

I am immensely grateful towards Professor Dr. Normah Md Noor, Dr. Maizom Hassan, and Puan Rafidah Ahmad from the Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia that have contributed greatly in this research. I am forever thankful to the principal investigators of the Enzyme and Microbial Technology (EMTEch) group in UPM and INBIOSIS, UKM for their infinite wisdom, insightful ideas, and suggestions that are essential to this study.

Special thanks towards my friends and colleagues from the Protein Engineering Lab in Biotech 3, Sayangku Nur Ariati, Nurul Syazwani, Noordiyanah, Sharifah Hidayah, Nurul Hazwani, Gayathrie, Jonathan Maiangwa, Siti Hajar, Fatin Liyana, and Nur Najihah for their assistance and generous support. My sincerest gratitude to Ku Nurul Aqmar Ku Bahaudin from the Metabolomics lab in INBIOSIS, UKM for her cooperation and knowledge that have assisted me greatly.

I would like to express my sincerest appreciation to both of my parents, Azman Shah Md Hashim and Hamidah Mat Amin for giving their utmost love and encouragement. I am immensely blessed to have a group of supportive friends, Adlina, Saleha, Aishah, Farhana and Liyana, whom have given countless support and encouragement throughout the years. Finally, I am grateful towards the Jabatan Pertahanan Awam (JPA) Malaysia that had awarded a scholarship that was of a great help and necessity during my postgraduate study.

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

Suriana binti Sabri, PhD
Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

Siti Nurbaya binti Oslan, PhD
Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

Syarul Nataqain binti Baharum, PhD
Associate Professor
Institute of Systems Biology
Universiti Kebangsaan Malaysia
(Member)

Goh Hoe Han, PhD
Associate Professor
Institute of Systems Biology
Universiti Kebangsaan Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 06 May 2021

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: _____ Date: _____

Name and Matric No: Fatin Lyana binti Azman Shah, GS44228

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature:	
Name of Chairman of Supervisory Committee:	Associate Professor Dr. Suriana binti Sabri
Signature:	
Name of Member of Supervisory Committee:	Associate Professor Dr. Siti Nurbaya binti Oslan
Signature:	
Name of Member of Supervisory Committee:	Associate Professor Dr. Syarul Nataqain binti Baharum
Signature:	
Name of Member of Supervisory Committee:	Associate Professor Dr. Goh Hoe Han

TABLE OF CONTENTS

		Page
ABSTRACT		i
ABSTRAK		iii
ACKNOWLEDGEMENTS		v
APPROVAL		vi
DECLARATION		viii
LIST OF TABLES		xiv
LIST OF FIGURES		xv
LIST OF APPENDICES		xvii
LIST OF ABBREVIATIONS		xviii
CHAPTER		
1	**INTRODUCTION**	1
	1.1 Introduction	1
	1.2 Research objectives	2
2	**LITERATURE REVIEW**	3
	2.1 Flavonoids	3
	2.2 The health benefits of flavonoids	3
	2.3 *P. minus* as a source of flavonoids	6
	2.4 Naringenin production through flavonoid biosynthesis pathway in *P. minus*	7
	2.5 Chalcone synthase	9
	2.6 Chalcone isomerase	11
	2.7 Recombinant expression of chalcone synthase and chalcone isomerase in a prokaryotic system	13
	2.8 Purification of recombinantly-expressed chalcone synthase and chalcone isomerase enzymes through affinity chromatography	16
	2.9 Three-dimensional (3D) protein structure of chalcone synthase from *P. minus*	16
3	**PARTIAL ISOLATION AND BIOINFORMATICS ANALYSIS OF CHALCONE ISOMERASE FROM *Polygonum minus* Huds.**	18
	3.1 Introduction	18
	3.2 Materials and methods	19
	3.2.1 Partial isolation and identification of chalcone isomerase (*CHI*) from *P. minus*	19
	3.2.1.1 Re-amplification of *CHI* from *P. minus*	19
	3.2.1.2 Agarose gel electrophoresis for DNA fragment analysis	19

3.2.1.3 Purification of the amplified DNA product 20
3.2.1.4 Cloning of *CHI* into *pEASY*-T5-Zero plasmid 20
3.2.1.5 Transformation of the recombinant plasmid into *E. coli* 20
3.2.1.6 Direct PCR analysis of the *E. coli Trans*1-T1 transformants 21
3.2.1.7 Extraction of recombinant plasmid from transformed *E. coli* 21
3.2.1.8 Nucleotide sequencing analysis of extracted recombinant plasmids 22
3.2.2 Bioinformatics analysis of chalcone isomerase from *P. minus* (*Pm*CHI) 22
3.3 Results and discucssion 23
3.3.1 Partial isolation and identification of *PmCHI* 23
3.3.1.1 Re-amplification of *CHI* from *P. minus* 23
3.3.1.2 Cloning of putative *CHI* into *pEASY*-T5-Zero plasmid 24
3.3.1.3 Nucleotide and protein sequence analysis of *PmCHI* 26
3.3.2 Bioinformatics analysis of *CHI* from *P. minus* 28
3.3.2.1 Protein sequence analysis of *Pm*CHI 28
3.3.2.2 Multiple sequence alignment of *Pm*CHI 29
3.3.2.3 Phylogenetic analysis of *Pm*CHI 30
3.3.2.4 Amino acid composition 32
3.3.2.5 Signal peptide detection in *Pm*CHI 32
3.3.2.6 Subcellular localization of *Pm*CHI 33
3.3.2.7 Transmembrane helices prediction of protein 35
3.3.2.8 Instability index 36
3.3.2.9 Hydrophobicity profile 37
3.4 Conclusion 39

4 EXPRESSION OF CHALCONE SYNTHASE (CHS) AND CHALCONE ISOMERASE (CHI) FROM *Polygonum minus* Huds. IN *Escherichia coli* 40
4.1 Introduction 40
4.2 Materials and methods 41
4.2.1 Codon optimization and gene synthesis of *CHS* and *CHI* from *P. minus* 41
4.2.2 Cloning of *CHS* and *CHI* in *E. coli* TOP10 41
4.2.2.1 Preparation of *E. coli* TOP10 competent cells 41

4.2.2.2 Transformation of pET-28b/co*CHS* and pET-28b/co*CHI* into *E. coli* TOP10 42
4.2.2.3 Screening of positive transformants by plasmid extraction and nucleotide sequencing 42
4.2.3 Cloning of *CHS* and *CHI* in *E. coli* BL21 (DE3) 42
4.2.3.1 Transformation of recombinant plasmids into *E. coli* BL21(DE3) 42
4.2.3.2 Verification of recombinant plasmids through restriction enzyme digestions 43
4.2.4 Heterologous protein expression in *E. coli* 43
4.2.4.1 Expression study of chalcone synthase and chalcone isomerase in *E. coli* BL21(DE3) at different temperatures 43
4.2.4.2 Determination of protein concentration 44
4.2.4.3 Recombinant protein detection by SDS-PAGE analysis 44
4.2.5 Western blot analysis 45
4.2.6 Partial purification of CHS and CHI 46
4.2.6.1 Preparation of crude lysates 46
4.2.6.2 Partial purification of the recombinant proteins 46
4.3 Results and discussion 46
4.3.1 Transformation of recombinant plasmids pET-28b/co*CHS* and pET-28b/co*CHI* into *E. coli* TOP10 and *E. coli* BL21(DE3) 46
4.3.1.1 Verification of transformed recombinant plasmids in *E. coli* TOP10 47
4.3.1.2 Confirmation of cloned recombinant plasmids in *E. coli* BL21(DE3) through restriction enzyme digestion 47
4.3.2 Expression of chalcone synthase and chalcone isomerase in *E. coli* BL21(DE3) 48
4.3.3 Western blot analysis of chalcone synthase and chalcone isomerase 51
4.3.4 Partial purification of chalcone synthase (CHS) and chalcone isomerase (CHI) 52
4.4 Conclusion 55

5 COMPUTATIONAL ANALYSIS OF THE THREE-DIMENSIONAL (3D) STRUCTURE OF CHALCONE ISOMERASE FROM *Polygonum minus* Huds. (*Pm*CHI) 56
5.1 Introduction 56
5.2 Materials and methods 57
5.2.1 *Pm*CHI protein sequence 57
5.2.2 Three-dimensional (3D) structure prediction of CHI protein from *P. minus* through homology modeling 57
5.2.3 Verification of predicted *Pm*CHI protein structure 57
5.3 Results and discussion 58
5.3.1 Secondary structure of *Pm*CHI protein 58
5.3.2 Analysis of the predicted structure of *Pm*CHI through homology modeling 59
5.3.3 Verification of predicted *Pm*CHI protein structure 61
5.4 Conclusion 67

6 CONCLUSION AND RECOMMENDATIONS FOR FUTURE RESEARCH 68
6.1 Conclusion 68
6.2 Recommendations for future research 68

REFERENCES 70
APPENDICES 86
BIODATA OF STUDENT 94
LIST OF PUBLICATIONS 95

LIST OF TABLES

Table		Page
2.1	Health benefits of flavonoids	5
2.2	Expression of chalcone synthase and chalcone isomerase in *E. coli* and suggested conditions used	15
3.1	Transmembrane prediction of CHI protein from *P. minus* by using the TMHMM server	35
4.1	The composition of SDS-PAGE gel	44

LIST OF FIGURES

Figure		Page
2.1	Basic flavonoid structure	3
2.2	Flavonoid production pathway	8
2.3	Catalytic mechanism of chalcone synthase (CHS)	10
2.4	Partial pathways for flavonoid and isoflavonoid biosynthesis	12
2.5	Three-dimensional (3D) structure of *Pm*CHS	17
3.1	PCR product of *CHI* from *P. minus*	24
3.2	Colony PCR products harboring *pEASY*-T5-Zero/*CHI* insert	25
3.3	Non-linearized plasmid extraction of *pEASY*-T5-Zero/*CHI*	25
3.4	Multiple sequence alignment of amplified *CHI* fragments from colonies 1, 2, and 3 harboring *pEASY*-T5-Zero/*CHI* plasmids	27
3.5	Sequence alignment between CHI protein sequences amplified from *P. minus* (in this study) and provide from INBIOSIS, UKM	28
3.6	Full-length nucleotide and amino acid sequence of ORF from *P. minus*	29
3.7	Multiple sequence alignment of *Pm*CHI with CHIs from various organisms	30
3.8	Evolutionary relationships of various CHIs based on deduced amino acid sequences	31
3.9	Prediction of output SignalP Neural Network of CHI protein from *P. minus*	33
3.10	The posterior probability for *Pm*CHI constructed from the TMHMM server	36
3.11	Hydrophobicity of chalcone isomerase (CHI) from *P. minus*, according to Kyte & Doolittle	38
4.1	Verification of recombinant host *E. coli* BL21(DE3) harboring pET-28b/co*CHS* and pET-28b/co*CHI* by restriction enzyme digestion	48

4.2	SDS-PAGE analysis of supernatants and pellets of *E. coli* BL21 (DE3) expressing CHS protein at 25 °C, 20 °C and 16 °C and 0.55 mM IPTG	49
4.3	SDS-PAGE analysis of strain *E. coli* BL21(DE3)/pET-28b-co*CHI* at 25 °C, 20 °C and 16 °C and 0.55 mM IPTG	50
4.4	Western blot and SDS-PAGE analyses of recombinants *E. coli* BL21(DE3)/pET-28b-co*CHS* and *E. coli* BL21(DE3)/pET-28b-co*CHI*	52
4.5	Chromatogram for the partial purification of chalcone synthase	53
4.6	Chromatogram for the partial purification of chalcone isomerase	53
4.7	SDS-PAGE analysis of the partial protein purification of chalcone synthase	54
4.8	SDS-PAGE analysis of the partial protein purification of chalcone isomerase	54
5.1	Secondary structure of *Pm*CHI	58
5.2	Alignment of amino acid residues of *Pm*CHI and *Da*CHI (PDB ID: 5YX4)	59
5.3	The 3D structure of *Pm*CHI predicted through YASARA	60
5.4	Location of conserved residues involved in CHI active sites observed in *Pm*CHI	61
5.5	Ramachandran plot of predicted *Pm*CHI protein model	62
5.6	Verify3D plot of predicted *Pm*CHI structure	64
5.7	ERRAT plot of predicted *Pm*CHI structure	66

LIST OF APPENDICES

Appendix		Page
1	List of chemicals, kits, antibiotics, reagents, enzymes, and antibodies used	86
2	Schematic representation of *pEasy*-T5-Zero harboring *CHI* gene from *P. minus*	87
3	Schematic representation of pET-28b(+) plasmid harboring codon-optimized *CHS* gene from *P. minus*	88
4	Schematic representation of pET-28b(+) plasmid harboring codon-optimized *CHI* gene from *P. minus*	89
5	Standard curve for Bovine serum albumin (BSA) using absorbance 595 nm	90
6	Sequence alignment of co*CHS* and cloned co*CHS* in pET-28b transformed into *E. coli* TOP10	91
7	Sequence alignment of co*CHI* and cloned co*CHI* in pET-28b transformed into *E. coli* TOP10	93

LIST OF ABBREVIATIONS

α	Alpha
Å	Angstrom
A_{600nm}	Optical density at wavelength 600 nanometer
APS	Ammonium persulfate
bp	Base pair
β	Beta
BSA	Bovine serum albumin
$CaCl_2$	Calcium chloride
cDNA	Complementary deoxyribonucleic acid
CHS	Chalcone synthase
CHI	Chalcone isomerase
°C	Degree celsius
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EC	Enzyme Commission
EDTA	Ethylenediaminetetraacetic acid
g	Gram
h	Hour
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilobase
kDa	Kilodalton
LB	Luria-Bertani
µg	Microgram

µL	Microliter
µm	Micrometer
min	Minute
mA	Milliamps
mM	Millimolar
M	Molar
NaCl	Sodium chloride
PCR	Polymerase chain reaction
P. minus	*Polygonum minus*
PmCHI	CHI gene from *Polygonum minus*
*Pm*CHI	CHI protein from *Polygonum minus*
PmCHS	CHS gene from *Polygonum minus*
*Pm*CHS	CHS protein from *Polygonum minus*
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel
TAE	Tris-acetate-EDTA
TBS	Tris-buffered saline
TBSTT	Tris-buffered saline containing Tween-20
V	Volt
v/v	Volume per volume
w/v	Weight per volume

CHAPTER 1

INTRODUCTION

1.1 Introduction

Flavonoids are classified as secondary metabolites produced through evolved phenylpropanoid metabolism, resulting in various inter-related flavonoid structures (Trantas et al., 2015). Flavonoids are mostly seen in plants as their coloring pigments (Khoo et al., 2017). These compounds proved to be beneficial as they have cholesterol-reducing (Kim et al., 2011), antioxidant, antidiabetic, antimicrobial, and anticancer properties (Hedlund et al., 2003; Cushnie & Lamb, 2005; Hossain et al., 2016; Chu et al., 2018). Economic prospects of flavonoids are enormous as they have potentials in pharmaceutical, nutraceutical, and agricultural sectors and were forecasted to reach up to USD 1.05 billion in revenue by 2021 (Hichri et al., 2011; Khoo et al., 2017; "Global flavonoid market," n.d.).

Current methods of flavonoid production depend on plant extractions and chemical synthesis. These processes are often time-consuming, commercially infeasible, unsafe towards the environment due to the use of harmful chemicals and are wasteful in terms of natural resources (Trantas et al., 2015). Hence, a solution to these issues is to provide an alternate source capable of producing flavonoids by engineering phenylpropanoid pathway into microbial hosts, such as *E. coli* and *Saccharomyces cerevisiae*.

Polygonum minus Huds. (synonym: *Persicaria minor*) or 'kesum' in Malay is a plant commonly found in the Southeast-Asia region. The high antioxidant capacity of *P. minus* is due to the high polyphenol, vitamin C, and β-carotene contents in this plant (Christapher et al., 2015). Also, aside from terpenes, flavonoids are the major compounds in *P. minus* (Khairudin et al., 2013). There are several enzymes involved in the flavonoid biosynthetic pathway of *P. minus*, starting from phenylalanine- or tyrosine-ammonia lyase (PAL or TAL), cinnamate-4-hydroxylase (C4H), 4-coumarate:CoA-ligase (4CL), followed by chalcone synthase (CHS) and isomerization into the final product by chalcone isomerase (CHI) which produces naringenin, a flavanone and universal precursor for many other essential flavonoid compounds.

Absence of the detailed information on the CHI protein sequence from *P. minus* hinders the characterization of this protein and for the complete assembly of the flavonoid biosynthesis pathway in microbial hosts for heterologous flavonoid production. On the other hand, the CHS protein sequence from *P. minus* has previously been identified (Roslan et al., 2013). However, there is no current

study conducted to investigate the heterologous expression of these enzymes from *P. minus* in a prokaryotic system (*E. coli*), and their subsequent purifications; which would later be very beneficial in future research focusing on the characterization of these proteins.

Determination of the three-dimensional structure of a protein is an important area in the proteomics study as it can deliver information on the functions of this protein (Waterhouse et al., 2018). The computational approach is a current method used for structure determination, as it will provide comprehensive structural information based on a novel protein sequence, where the structure has not been determined experimentally (Centeno et al., 2005). Previously, Roslan et al. (2013) had successfully predicted the 3D structure of CHS by computational analysis. Hence, the absence of a detailed 3D structure of CHI protein from *P. minus* could be solved by using structure prediction programs (such as homology modeling) and gives further insight on the working mechanism of this enzyme.

1.2 Research objectives

This study is divided into three different chapters, each focusing on a single objective. The specific objectives for each chapter in this study are:

1. To identify and isolate chalcone isomerase (CHI) from *P. minus*
2. To express chalcone synthase (CHS) and chalcone isomerase (CHI) in *E. coli*
3. To predict the three-dimensional (3D) structure of CHI protein from *P. minus* through homology modeling

REFERENCES

Aftabuddin, M., & Kundu, S. (2007). Hydrophobic, hydrophilic, and charged amino acid networks within protein. *Biophysical Journal*, *93*(1), 225–231.

Al-Ishaq, R. K., Abotaleb, M., Kubatka, P., Kajo, K., & Büsselberg, D. (2019). Flavonoids and their anti-diabetic effects: Cellular mechanisms and effects to improve blood sugar levels. *Biomolecules*, *9*(9), 1-35.

Alcaráz, L. E., Blanco, S. E., Puig, O. N., Tomás, F., & Ferretti, F. H. (2000). Antibacterial activity of flavonoids against methicillin-resistant *Staphylococcus aureus* strains. *Journal of Theoretical Biology*, *205*(2), 231–240.

Armenteros, J. J. A., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., & Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance*, *2*(5), 1–14.

Austin, M. B., & Noel, J. P. (2003). The chalcone synthase superfamily of type III polyketide synthases. *Natural Product Reports*, *20*(1), 79–110.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology*, *10*(5), 411–421.

Barh, D., Barve, N., Gupta, K., Chandra, S., Jain, N., Tiwari, S., Leon-Socairos, N., Canizalez-Roman, A., dos Santos, A. R., Hassan, S. S., Almeida, S., Ramos, R. T. J., de Abreu, V. A. C., Carneiro, A. R., de Castro Soares, S., de Paula Castro, T. L., Miyoshi, A., Silva, A., Kumar, … Azevedo, V. (2013). Exoproteome and secretome derived broad spectrum novel drug and vaccine candidates in *Vibrio cholerae* targeted by Piper betel derived compounds. *PloS One*, *8*(1), 1-10.

Bednar, R. A., & Hadcock, J. R. (1998). Purification and characterization of chalcone isomerase from soybeans. *Journal of Biological Chemistry*, *263*(20), 9582–9588.

Berg, J. M., Tymoczko, J. L., & Stryer, L. (2002). *Biochemistry* (5th ed.). New York, NY: W. H. Freeman and Company.

Bertin, C., Yang, X., & Weston, L. A. (2003). The role of root exudates and allelochemicals in the rhizosphere. *Plant and Soil*, *256*(1), 67–83.

Boyle, A. L. (2018). 3 - Applications of *de novo* designed peptides. In S. Koutsopoulos (Ed.), *Peptide Applications in Biomedicine, Biotechnology and Bioengineering* (pp. 51–86). Sawston, CB: Woodhead Publishing.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Bioc*hemistry, *72*(1-2), 248–254.

Broadway, N. (2012). Protein expression. *Materials & Methods*, *2*(123). Retrieved on 25th August, 2019 from http://www.labome.com/method/RecombinantProtein-Expression-Vector-Host-Systems.html.

Bucolo, C., Leggio, G. M., Drago, F., & Salomone, S. (2012). Eriodictyol prevents early retinal and plasma abnormalities in streptozotocin-induced diabetic rats. *Biochemical Pharmacology*, *84*(1), 88–92.

Catarino, M. D., Alves-silva, J. M., Pereira, O. R., & Cardoso, S. M. (2015). Antioxidant capacities of flavones and benefits in oxidative-stress related diseases. *Current Topics in Medicinal Chemistry*, *15*(2), 105-119.

Centeno, N. B., Planas-Iglesias, J., & Oliva, B. (2005). Comparative modelling of protein structure and its impact on microbial cell factories. *Microbial Cell Factories*, *4*(20), 1–11.

Chang, A. Y., Chau, V. WY., Landas, J. A., & Pang, Y. (2017). Preparation of calcium competent *Escherichia coli* and heat-shock transformation. *JEMI Methods*, *1*, 22–25.

Chen, M., Zhu, W. J., You, X., Liu, Y. D., Kaleri, G. M., & Yang, Q. (2015). Isolation and characterization of a chalcone isomerase gene promoter from potato cultivars. *Genetics and Molecular Research*, *14*(4), 18872–18885.

Chen, X. Q., Hu, T., Han, Y., Huang, W., Yuan, H. B., Zhang, Y. T., Du, Y., & Jiang, Y. W. (2016). Preventive effects of catechins on cardiovascular disease. *Molecules*, *21*(12), 1–7.

Cheng, A. X., Zhang, X., Han, X. J., Zhang, Y. Y., Gao, S., Liu, C. J., & Lou, H. X. (2018). Identification of chalcone isomerase in the basal land plants reveals an ancient evolution of enzymatic cyclization activity for synthesis of flavonoids. *New Phytologist*, *217*(2), 909–924.

Cheng, H., Li, L., Cheng, S., Cao, F., Wang, Y., & Yuan, H. (2011). Molecular cloning and function assay of a chalcone isomerase gene (*Gb*CHI) from *Ginkgo biloba*. *Plant Cell Reports*, *30*(1), 49–62.

Cherrak, S. A., Mokhtari-Soulimane, N., Berroukeche, F., Bensenane, B., Cherbonnel, A., Merzouk, H., & Elhabiri, M. (2016). *In vitro* antioxidant versus metal ion chelating properties of flavonoids: A structure-activity investigation. *PLoS ONE*, *11*(10), 1–21.

Christapher, P. V., Parasuraman, S., Christina, J. M. A., Asmawi, M. Z., & Vikneswaran, M. (2015). Review on *Polygonum minus*. Huds, a commonly used food additive in Southeast Asia. *Pharmacognosy Research*, *7*(1), 1-6.

Chu, L. L., Dhakal, D., Shin, H. J., Jung, H. J., Yamaguchi, T., & Sohng, J. K. (2018). Metabolic engineering of *Escherichia coli* for enhanced production of Naringenin 7-Sulfate and its biological activities. *Frontiers in Microbiology*, *9*, 1–13.

Clark, D. P., & Pazdernik, N. J. (2013). Chapter e13 - Protein Synthesis. In D. P. Clark & N. J. Pazdernik (Eds.), *Molecular Biology (Second edition)* (pp. e250–e255). Cambridge, MA: Academic Press.

Claudot, A. C., & Drouet, A. (1992). Preparation and assay of chalcone synthase from walnut tree tissue. *Phytochemistry*, *31*(10), 3377–3380.

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, *2*(9), 1511–1519.

Conrad, A. C., & Mathabatha, M. F. (2017). Characterization and expression analyses of chalcone synthase (*CHS*) and anthocyanidin synthase (*ANS*) genes in *Clivia miniata*. *Transcriptomics*, *4*(2), 1-15.

Cushnie, T. P. T., & Lamb, A. J. (2005). Antimicrobial activity of flavonoids. *International Journal of Antimicrobial Agents*, *26*(5), 343–356.

Dao, T. T. H., Linthorst, H. J. M., & Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. *Phytochemistry Reviews*, *10*(3), 397–412.

Das, P., & Rawal, S. K. (2016). Cloning, expression and purification of chalcone synthase from *Solanum tuberosum*. *IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)*, *2*(3), 7–11.

Deng, H., Jia, Y., & Zhang, Y. (2018). Protein structure prediction. *International Journal of Modern Physics. B*, *32*(18), 1-14.

Dixon, R. A., & Harrison, M. J. (1990). Activation, structure, and organization of genes involved in microbial defense in plants. *Advances in Genetics*, *28*(C), 165–234.

Dixon, R. A., Richard Blyden, E., Robbins, M. P., Van Tunen, A. J., & Mol, J. N. (1988). Comparative biochemistry of chalcone isomerases. *Phytochemistry*, *27*(9), 2801–2808.

Dobson, C. M. (1999). Protein misfolding, evolution and disease. *Trends in Biochemical Sciences*, *24*(9), 329–332.

Druka, A., Kudrna, D., Rostoks, N., Brueggeman, R., Von Wettstein, D., & Kleinhofs, A. (2003). Chalcone isomerase gene from rice (*Oryza sativa*) and barley (*Hordeum vulgare*): Physical, genetic and mutation mapping. *Gene*, *302*(1–2), 171–178.

Eichenberger, M., Lehka, B. J., Folly, C., Fischer, D., Martens, S., Simön, E., & Naesby, M. (2017). Metabolic engineering of *Saccharomyces cerevisiae* for *de novo* production of dihydrochalcones with known antioxidant, antidiabetic, and sweet tasting properties. *Metabolic Engineering*, *39*, 80–89.

Eisenberg, D., Luthy, R., & Bowie, J. U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in Enzymology*, *277*, 396–404.

Emanuelsson, O. (2002). Predicting protein subcellular localisation from amino acid sequence information. *Briefings in Bioinformatics*, *3*(4), 361–376.

Fernández-Fernández, Á. D., & Corpas, F. J. (2016). *In silico* analysis of *Arabidopsis thaliana* peroxisomal 6-phosphogluconate dehydrogenase. *Scientifica*, *2016*, 1-9.

Ferrer, J. L., Jez, J. M., Bowman, M. E., Dixon, R. A., & Noel, J. P. (1999). Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nature Structural and Molecular Biology, 6*, 775–784.

Fiser, A. (2010). Template-based protein structure modeling. *Methods in Molecular Biology*, *673*, 73–94.

Galvano, F., La Fauci, L., Vitaglione, P., Fogliano, V., Vanella, L., & Felgines, C. (2007). Bioavailability, antioxidant and biological properties of the natural free-radical scavengers cyanidin and related glycosides. *Annali Dell'Istituto Superiore Di Sanita*, *43*(4), 382–393.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In J. M. Walker (Ed.). *The Proteomics Protocols Handbook* (pp. 571-607). New York, NY: Humana Press.

Gay, G., Wagner, D. T., Keatinge-Clay, A. T., & Gay, D. C. (2014). Rapid modification of the pET-28 expression vector for ligation independent cloning using homologous recombination in *Saccharomyces cerevisiae*. *Plasmid*, *76*, 66–71.

Gensheimer, M., & Mushegian, A. (2004). Chalcone isomerase family and fold: no longer unique to plants. *Protein Science*, *13*(2), 540–544.

Global-Flavonoids-Market-will-reach-USD-104763-million-in-2021-Zion-Market-Research-457514 @ www.econotimes.com. (n.d.). Retrieved on 21st July, 2017 from http://www.econotimes.com/Global-Flavonoids-Market-will-reach-USD-104763-million-in-2021-Zion-Market-Research-457514

Goldwasser, J., Cohen, P. Y., Yang, E., Balaguer, P., Yarmush, M. L., & Nahmias, Y. (2010). Transcriptional regulation of human and rat hepatic lipid metabolism by the grapefruit flavonoid naringenin: Role of PPARα, PPARγ and LXRα. *PLoS ONE*, *5*(8), 1–9.

Graf, B. A., Milbury, P. E., & Blumberg, J. B. (2005). Flavonols, flavones, flavanones, and human health: Epidemiological evidence. *Journal of Medicinal Food*, *8*(3), 281–290.

Gregoret, L. M., & Sauer, R. T. (1998). Tolerance of a protein helix to multiple alanine and valine substitutions. *Folding and Design*, *3*(2), 119–126.

Guo, D., Gao, Y., Liu, F., He, B., Jia, X., Meng, F., Zhang, H., & Guo, M. (2019). Integrating molecular characterization and metabolites profile revealed CtCHI1's significant role in *Carthamus tinctorius* L. *BMC Plant Biology*, *19*(1), 1–13.

Guo, J., Zhou, W., Lu, Z., Li, H., Li, H., & Gao, F. (2015). Isolation and functional analysis of chalcone isomerase gene from purple-fleshed sweet potato. *Plant Molecular Biology Reporter*, *33*(5), 1451–1463.

Guruprasad, K., Reddy, B. V. B., & Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Engineering, Design and Selection*, *4*(2), 155–161.

Hage, D. S., Anguizola, J. A., Bi, C., Li, R., Matsuda, R., Papastavros, E., Pfaunmiller, E., Vargas, J., & Zheng, X. (2012). Pharmaceutical and biomedical applications of affinity chromatography: Recent trends and developments. *Journal of Pharmaceutical and Biomedical Analysis*, *69*, 93–105.

Harmon, A. W., & Patel, Y. M. (2003). Naringenin inhibits phosphoinositide 3-kinase activity and glucose uptake in 3T3-L1 adipocytes. *Biochemical and Biophysical Research Communications*, *305*(2), 229–234.

Hedlund, T. E., Johannes, W. U., & Miller, G. J. (2003). Soy isoflavonoid equol modulates the growth of benign and malignant prostatic epithelial cells *in vitro*. *Prostate*, *54*(1), 68–78.

Hemleben, V., Dressel, A., Epping, B., Lukačin, R., Martens, S., & Austin, M. B. (2004). Characterization and structural features of a chalcone synthase mutation in a white-flowering line of *Matthiola incana* R. Br. (Brassicaceae). *Plant Molecular Biology*, *55*(3), 455–465.

Heneman, K., & Zidenberg-Cherr, S. (2008). Some facts about flavonol. Retrieved on 7th March, 2017 from https ://nutri tion.ucdav is.edu/sites /g/files /dgvns k426/files /conte nt/ infos heets /fact-pro-flavo nol.pdf.

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., & Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *Journal of Experimental Botany*, *62*(8), 2465–2483.

Hopp, T. P., & Woods, K. R. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(6 I), 3824–3828.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). WoLF PSORT: Protein localization predictor. *Nucleic Acids Research*, *35*(SUPPL.2), 585–587.

Hossain, M. K., Dayem, A. A., Han, J., Yin, Y., Kim, K., Saha, S. K., Yang, G. M., Choi, H. Y., & Cho, S. G. (2016). Molecular mechanisms of the anti-obesity and anti-diabetic properties of flavonoids. *International Journal of Molecular Sciences*, *17*(4), 1-32.

Hrazdina, G., & Jensen, R. A. (1992). Spatial organization of enzymes in plant metabolic pathways. *Annual Review of Plant Physiology and Plant Molecular Biology*, *43*(1), 241–267.

Hwang, E. I., Kaneko, M., Ohnishi, Y., & Horinouchi, S. (2003). Production of plant-specific flavanones by *Escherichia coli* containing an artificial gene cluster. *Applied and Environmental Microbiology*, *69*(5), 2699–2706.

Idicula-Thomas, S., & Balaji, P. V. (2005). Understanding the relationship between the primary structure of proteins and its propensity to be soluble on overexpression in *Escherichia coli*. *Protein Science*, *14*(3), 582–592.

Ignatova, Z. (2005). Monitoring protein stability *in vivo*. *Microbial Cell Factories*, *4*(23), 1–6.

Imelda, F., Faridah, D. N., & Kusumaningrum, H. D. (2014). Bacterial inhibition and cell leakage by extract of *Polygonum minus* Huds. leaves. *International Food Research Journal*, *21*(2), 553–560.

Jaspard, E., Macherel, D., & Hunault, G. (2012). Computational and statistical analyses of amino acid usage and physico-chemical properties of the twelve late embryogenesis abundant protein classes. *PLoS ONE*, *7*(5), 1-20.

Jez, J. M., Bowman, M. E., Dixon, R. A., & Noel, J. P. (2000). Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nature Structural Biology*, *7*(9), 786–791.

Jez, J. M., Bowman, M. E., & Noel, J. P. (2002). Role of hydrogen bonds in the reaction mechanism of chalcone isomerase. *Biochemistry*, *41*(16), 5168–5176.

Jez, J. M., & Noel, J. P. (2000). Mechanism of chalcone synthase. pK_a of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. *Journal of Biological Chemistry*, *275*(50), 39640–39646.

Jez, J. M., & Noel, J. P. (2002). Reaction mechanism of chalcone isomerase: pH dependence, diffusion control, and product binding differences. *Journal of Biological Chemistry*, *277*(2), 1361–1369.

Jiang, C., Schommer, C. K., Kim, S. Y., & Suh, D.-Y. (2006). Cloning and characterization of chalcone synthase from the moss, *Physcomitrella patens*. *Phytochemistry*, *67*(23), 2531–2540.

Jung, U. J., Kim, H. J., Lee, J. S., Lee, M. K., Kim, H. O., Park, E. J., Kim, H. K., Jeong, T. S., & Choi, M. S. (2003). Naringin supplementation lowers plasma lipids and enhances erythrocyte antioxidant enzyme activities in hypercholesterolemic subjects. *Clinical Nutrition*, *22*(6), 561–568.

Kalman, M., & Ben-Tal, N. (2010). Quality assessment of protein model-structures using evolutionary conservation. *Bioinformatics*, *26*(10), 1299–1307.

Kaneko, M., Hwang, E. I., Ohnishi, Y., & Horinouchi, S. (2003). Heterologous production of flavanones in *Escherichia coli*: potential for combinatorial biosynthesis of flavonoids in bacteria. *Journal of Industrial Microbiology and Biotechnology*, *30*(8), 456–461.

Khairudin, K., Sukiran, N. A., Goh, H-H., Baharum, S. N., & Noor, N. M. (2013). Direct discrimination of different plant populations and study on temperature effects by Fourier transform infrared spectroscopy. *Metabolomics*, *10*, 203-211.

Khan, S., Ullah, M. W., Siddique, R., Nabi, G., Manan, S., Yousaf, M., & Hou, H. (2016). Role of recombinant DNA technology to improve life. *International Journal of Genomics*, *2016*, 1-14.

Khoo, H. E., Azlan, A., Tang, S. T., & Lim, S. M. (2017). Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food and Nutrition Research*, *61*(1), 1-21.

Khor, B. Y., Tye, G. J., Lim, T. S., Noordin, R., & Choong, Y. S. (2014). The structure and dynamics of BmR1 protein from *Brugia malayi*: *In silico* approaches. *International Journal of Molecular Sciences*, *15*(6), 11082–11099.

Kim, A., Chiu, A., Barone, M. K., Avino, D., Wang, F., Coleman, C. I., & Phung, O. J. (2011). Green tea catechins decrease total and low-density lipoprotein cholesterol: a systematic review and meta-analysis. *Journal of the American Dietetic Association*, *111*(11), 1720–1729.

Koduri, P. K. H., Gordon, G. S., Barker, E. I., Colpitts, C. C., Ashton, N. W., & Suh, D. Y. (2010). Genome-wide analysis of the chalcone synthase superfamily genes of *Physcomitrella patens*. *Plant Molecular Biology*, *72*(3), 247–263.

Krieger, E., Darden, T., Nabuurs, S. B., Finkelstein, A., & Vriend, G. (2004). Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins*, *57*(4), 678–683.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, *305*(3), 567–580.

Krystek, S. R., Metzler, W. J., & Novotny, J. (1995). Hydrophobicity profiles for protein sequence analysis. *Current Protocols in Protein Science*, *00*(1), 1–13.

Kumar, S., & Pandey, A. K. (2013). Chemistry and biological activities of flavonoids: An overview. *The Scientific World Journal*, *2013*, 1-16.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874.

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680–685.

Land, H., & Humble, M. S. (2018). YASARA: A tool to obtain structural guidance in biocatalytic investigations. In U. T. Bornscheuer & M. Höhne (Eds.), *Protein Engineering: Methods and Protocols* (pp. 43–67). New York, NY: Humana Press.

Laskowski, R. A. (2001). PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Research*. *29*(1), 221-222.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291.

Lee, S. Y. (1996). High cell-density culture of *Escherichia coli*. *Trends in Biotechnology*, *14*, 98–105.

Lei, W., Tang, S. H., Luo, K. M., & Sun, M. (2010). Molecular cloning and expression profiling of a chalcone synthase gene from hairy root cultures of S*cutellaria viscidula* Bunge. *Genetics and Molecular Biology*, *33*(2), 285-291.

Li, F., Jin, Z., Qu, W., Zhao, D., & Ma, F. (2006). Cloning of a cDNA encoding the *Saussurea medusa* chalcone isomerase and its expression in transgenic tobacco. *Plant Physiology and Biochemistry*, *44*(7–9), 455–461.

Lila, M. A. (2004). Anthocyanins and human health: An *in vitro* investigative approach. *Journal of Biomedicine and Biotechnology*, *2004*(5), 306–313.

Lim, H. J., Nguyen, T. T. H., Kim, N. M., Park, J. S., Jang, T. S., & Kim, D. (2016). Inhibitory effect of flavonoids against NS2B-NS3 protease of ZIKA virus and their structure activity relationship. *Biotechnology Letters*, *39*(3), 415-421.

Lin, Z., Feng, M., Dos Santos, C. N., Yu, M., Xiang, B., Zhou, B., & Bengio, Y. (2017). A structured self-attentive sentence embedding. *ArXiv*, 1–15

Liu, X., Ahmad, N., Yang, L., Fu, T., Kong, J., Yao, N., Dong, Y., Wang, N., Li, X., Wang, F., Liu, X., Liu, W., & Li, H. (2019). Molecular cloning and functional characterization of chalcone isomerase from *Carthamus tinctorius*. *AMB Express*, *9*(132), 1-12.

Loake, G. J., Choudhary, A. D., Harrison, M. J., Mavandad, M., Lamb, C. J., & Dixon, R. A. (1991). Phenylpropanoid pathway intermediates regulate transient expression of a chalcone synthase gene promoter. *Plant Cell*, *3*(8), 829–840.

Loke, K.-K., Rahnamaie, T. R., Yeoh, C.-C., Goh, H.-H., Hussein, Z.-A. M., Zainal, Z., Ismail, I., & Noor, N. M. (2017). Transcriptome analysis of *Polygonum minus* reveals candidate genes involved in important secondary metabolic pathways of phenylpropanoids and flavonoids. *PeerJ*, *5*, 1-20.

López-Llano, J., Campos, L. A., & Sancho, J. (2006). α-helix stabilization by alanine relative to glycine: Roles of polar and apolar solvent exposures and of backbone entropy. *Proteins: Structure, Function, and Bioinformatics*, *64*(3), 769–778.

Ma, W., Wu, Y., Wu, M., Ren, Z., & Zhong, Y. (2013). Cloning, characterization and expression of chalcone synthase from medicinal plant *Rhus chinensis*. *Journal of Plant Biochemistry and Biotechnology*, *24*(1), 18–24.

Mahmood, T., & Yang, P.-C. (2012). Western blot: technique, theory, and trouble shooting. *North American Journal of Medical Sciences*, *4*(9), 429–434.

Makita, H., Tanaka, T., Fujitsuka, H., Tatematsu, N., Satoh, K., Hara, A., & Mori, H. (1996). Chemoprevention of 4-nitroquinoline 1-oxide-induced rat oral carcinogenesis by the dietary flavonoids chalcone, 2-hydroxychalcone, and quercetin. *Cancer Research*, *56*, 4904–4909.

Marais, J. P. J., Deavours, B., Dixon, R. A., & Ferreira, D. (2006). The stereochemistry of flavonoids. In E. Grotewold (Ed.), *The Science of Flavonoids* (pp. 1–46). New York, NY: Springer.

Mehdy, M. C., & Lamb, C. J. (1988). Chalcone isomerase cDNA cloning and mRNA induction by fungal elicitor, wounding and infection. *The EMBO Journal*, *6*(6), 1527–1533.

Messaoudi, A., Belguith, H., & Ben Hamida, J. (2013). Homology modeling and virtual screening approaches to identify potent inhibitors of VEB-1 β-lactamase. *Theoretical Biology and Medical Modelling*, *10*(1), 1–10.

Mielcke, T. R., Muradás, T. C., Filippi-Chiela, E. C., Amaral, M. E. A., Kist, L. W., Bogo, M. R., Mascarello, A., Neuenfeldt, P. D., Nunes, R. J., & Campos, M. M. (2017). Mechanisms underlying the antiproliferative effects of a series of quinoxaline-derived chalcones. *Scientific Reports*, *7*(1), 1–16

Miyahisa, I., Kaneko, M., Funa, N., Kawasaki, H., Kojima, H., Ohnishi, Y., & Horinouchi, S. (2005). Efficient production of (2*S*)-flavanones by *Escherichia coli* containing an artificial biosynthetic gene cluster. *Applied Microbiology and Biotechnology*, *68*(4), 498–504.

Mol, J. N. M., Robbins, M. P., Dixon, R. A., & Veltkamp, E. (1985). Spontaneous and enzymic rearrangement of naringenin chalcone to flavanone. *Phytochemistry*, *24*(10), 2267–2269.

Moraes, J. P. A., Pappa, G. L., Pires, D. E. V., & Izidoro, S. C. (2017). GASS-WEB: A web server for identifying enzyme active sites based on genetic algorithms. *Nucleic Acids Research*, *45*(W1), W315–W319.

Morita, Y., Takagi, K., Fukuchi-Mizutani, M., Ishiguro, K., Tanaka, Y., Nitasaka, E., Nakayama, M., Saito, N., Kagami, T., Hoshino, A., & Iida, S. (2014). A chalcone isomerase-like protein enhances flavonoid production and flower pigmentation. *Plant Journal*, *78*(2), 294–304.

Moustafa, E, & Wong, E. (1967). Purification and properties of chalcone-flavanone isomerase from soya bean seed. *Phytochemistry*, *6*, 625–632.

Nakai, K., & Imai, K. (2019). Prediction of protein localization. In S. Ranganathan, M. Gribskov, K. Nakai, & C. B. Schönbach (Eds.), *Encyclopedia of Bioinformatics and Computational Biology* (pp. 53–59). Cambridge, MA: Academic Press.

Negi, V.S., Borthakur, D. (2016). Heterologous expression and characterization of mimosinase from *Leucaena leucocephala*. In A. G. Fett-Neto (Ed.), *Biotechnology of Plant Secondary Metabolism* (pp. 59-77). New York, NY: Humana Press.

Ngaki, M. N., Louie, G. V., Philippe, R. N., Manning, G., Pojer, F., Bowman, M. E., Li, L., Larsen, E., Wurtele, E. S., & Noel, J. P. (2012). Evolution of the chalcone-isomerase fold from fatty-acid binding to stereospecific catalysis. *Nature*, *485*(7399), 530–533.

Nielsen, H. (2017). Predicting secretory proteins with SignalP. In D. Kihara (Ed.), *Methods in Molecular Biology* (pp. 59–73). New York, NY: Humana Press.

Othman, A., Mukhtar, N. J., Ismail, N. S., & Chang, S. K. (2014). Phenolics, flavonoids content and antioxidant activities of 4 Malaysian herbal plants. *International Food Research Journal*, *21*(2), 759–766.

Panche, A. N., Diwan, A. D., & Chandra, S. R. (2016). Flavonoids: An overview. *Journal of Nutritional Science*, *5*(47), 1-15.

Pandey, R. P., Parajuli, P., Koffas, M. A. G., & Sohng, J. K. (2016). Microbial production of natural and non-natural flavonoids: Pathway engineering, directed evolution and systems/synthetic biology. *Biotechnology Advances*, *34*(5), 634–662.

Park, S. H., Lee, C. W., Cho, S. M., Lee, H., Park, H., Lee, J., & Lee, J. H. (2018). Crystal structure and enzymatic properties of chalcone isomerase from the Antarctic vascular plant *Deschampsia antarctica* Desv. *PLoS ONE*, *13*(2), 1–17.

Parker, J. (2001). Leader peptide. In S. Brenner & J. H. Miller (Eds.), *Encyclopedia of Genetics* (pp. 1077–1078). Cambridge, MA: Academic Press.

Pelley, J. W. (2007). 3 - Protein Structure and Function. In J. W. Pelley (Ed.), *Elsevier's Integrated Biochemistry* (pp. 19–28). Maryland Heights, MO: Mosby.

Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, *8*(10), 785-786.

Peterson, J. J., Beecher, G. R., Bhagwat, S. A., Dwyer, J. T., Gebhardt, S. E., Haytowitz, D. B., & Holden, J. M. (2006). Flavanones in grapefruit, lemons, and limes: A compilation and review of the data from the analytical literature. *Journal of Food Composition and Analysis*, *19*, S74–S80.

Pitakdantham, W., Sutabutra, T., Chiemsombat, P., & Pitaksutheepong, C. (2010). Isolation and characterization of chalcone synthase gene isolated from *Dendrobium sonia* Ersakul. *Pakistan Journal of Biological Science*, *13*(20), 1000–1005.

Pope, B., & Kent, H. M. (1996). High efficiency 5 min transformation of *Escherichia coli*. *Nucleic Acids Research*, *24*(3), 536–537.

Ralston, L., Subramanian, S., Matsuno, M., & Yu, O. (2005). Partial reconstruction of flavonoid and isoflavonoid biosynthesis in yeast using soybean type I and type II chalcone isomerases. *Plant Physiology*, *137*(4), 1375–1388.

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, *7*(1), 95–99.

Ren, C., Tang, X., Chen, C., Chen, J., Pei, J., Wu, Y., & Wu, Q. (2019). Cloning and expression analysis of a new chalcone isomerase gene during flowering in safflower. *Turkish Journal of Botany*, *43*(2), 143–150.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges, *5*(April), 1–17.

Roslan, N. D., Tan, C.-S., Ismail, I., & Zainal, Z. (2013). cDNA cloning and expression analysis of the chalcone synthase gene (*CHS*) from *Polygonum minus*. *Australian Journal of Crop Science*, *7*(6), 777–783.

Roslan, N. D., Yusop, J. M., Baharum, S. N., Othman, R., Mohamed-Hussein, Z.-A., Ismail, I., Noor, N. M., & Zainal, Z. (2012). Flavonoid biosynthesis genes putatively identified in the aromatic plant *Polygonum minus via* Expressed Sequences Tag (EST) analysis. *International Journal of Molecular Sciences*, *13*(12), 2692–2706.

Samarghandian, S., Azimi-Nezhad, M., & Farkhondeh, T. (2017). Catechin treatment ameliorates diabetes and its complications in streptozotocin-induced diabetic rats. *Dose-Response*, *15*(1), 1–7

Sambrook, J., Fritsch, E. R., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (2nd ed.). New York, NY: Cold Spring Harbor Laboratory Press.

Sanmugavelan, R., Teoh, T. C., Roslan, N., & Mohamed, Z. (2018). *In vitro* and *in silico* studies of chalcone synthase variant 2 in *Boesenbergia rotunda* and its substrate specificity. *Turkish Journal of Biology*, *42*(3), 213–223.

Schröder, J., & Schäfer, E. (1980). Radioiodinated antibodies, a tool in studies on the presence and role of inactive enzyme forms: Regulation of chalcone synthase in parsley cell suspension cultures. *Archives of Biochemistry and Biophysics*, *203*(2), 800–808.

Schwede, T. (2013). Protein modeling: What happened to the "protein structure gap"? *Structure*, *21*(9), 1531–1540.

Schwede, T., Sali, A., Honig, B., Levitt, M., Berman, H. M., Jones, D., Brenner, S. E., Burley, S. K., Das, R., Dokholyan, N. V., Dunbrack, R. L. J., Fidelis, K., Fiser, A., Godzik, A., Huang, Y. J., Humblet, C., Jacobson, M. P., Joachimiak, A., Krystek, S. R., J., … Wilson, I. A. (2009). Outcome of a workshop on applications of protein models in biomedical research. *Structure*, *17*(2), 151-159.

Scott, M. S., Calafell, S. J., Thomas, D. Y., & Hallett, M. T. (2005). Refining protein subcellular localization. *PLoS Computational Biology*, *1*(6), 0518–0528.

Shah, F. L. A., Ramzi, A. B., Baharum, S. N., Noor, N. M., Goh, H.-H., Leow, T. C., Oslan, S. N., & Sabri, S. (2019). Recent advancement of engineering microbial hosts for the biotechnological production of flavonoids. *Molecular Biology Reports*, *46*, 6647-6659.

Shahat, A. A., & Marzouk, M. S. (2013). 13 - Tannins and related compounds from medicinal plants of Africa. In V. Kuete (Ed.), *Pharmacology and Chemistry* (pp. 479–555). Amsterdam, AMS: Elsevier.

Sheik, S. S., Sundararajan, P., Hussain, A. S. Z., & Sekar, K. (2002). Ramachandran plot on the web. *Bioinformatics*, *18*(11), 1548–1549.

Shen, C. H. (2019). *Diagnostic Molecular Biology*. Amsterdam, AMS: Elsevier.

Shiloach, J., & Fass, R. (2005). Growing *E. coli* to high cell density – a historical perspective on method development. *Biotechnology Advances*, *23*(5), 345–357.

Sivoňová, M. K., Kaplán, P., Tatarková, Z., Lichardusová, L., Dušenka, R., & Jurečeková, J. (2019). Androgen receptor and soy isoflavones in prostate cancer (Review). *Molecular and Clinical Oncology*, *10*(2), 191–204

Sperschneider, J., Williams, A. H., Hane, J. K., Singh, K. B., & Taylor, J. M. (2015). Evaluation of secretion prediction highlights differing approaches needed for oomycete and fungal effectors. *Frontiers in Plant Science*, *6*, 1–14.

Stepanić, V., Matijašić, M., Horvat, T., Verbanac, D., Chlupáćová, M. K., Saso, L., & Žarković, N. (2019). Antioxidant activities of alkyl substituted pyrazine derivatives of chalcones - *In vitro* and *in silico* study. *Antioxidants*, *8*(4), 1-11.

Sun, W., Meng, X., Liang, L., Jiang, W., Huang, Y., He, J., Hu, H., Almqvist, J., Gao, X., & Wang, L. (2015). Molecular and biochemical analysis of chalcone synthase from *Freesia hybrida* in flavonoid biosynthetic pathway. *PLoS ONE*, *10*(3), 1–18.

Sun, W., Shen, H., Xu, H., Tang, X., Tang, M., Ju, Z., & Yi, Y. (2019). Chalcone isomerase a key enzyme for anthocyanin biosynthesis in *Ophiorrhiza japonica*, *Frontiers in Plant Science*, *10*, 1–12.

Tapas, A. R., Sakarkar, D. M., & Kakde, R. B. (2008). Flavonoids as nutraceuticals: A review. *Tropical Journal of Pharmaceutical Research*, *7*(3), 1089–1099.

Tian, J., Shen, H., Zhang, J., Song, T., & Yao, Y. (2011). Characteristics of chalcone isomerase promoter in crabapple leaves (*M*.cv. 'royalty') and transient expression assay modified in onion epidermal cell. *African Journal of Biotechnology*, *10*(50), 10232–10240.

Toomula, N., Kumar, D. S., & Kumar, V. V. L. P. (2011). Computational methods for protein structure prediction and its application in drug design. *Journal of Proteomics and Bioinformatics*, *4*(12), 289–293.

Trantas, E. A., Koffas, M. A. G., Xu, P., & Ververidis, F. (2015). When plants produce not enough or at all: metabolic engineering of flavonoids in microbial hosts. *Frontiers in Plant Science*, *6*, 1–16.

Trantas, E., Panopoulos, N., & Ververidis, F. (2009). Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. *Metabolic Engineering*, *11*(6), 355–366.

Umar, K. M., Abdulkarim, S. M., Radu, S., Abdul Hamid, A., & Saari, N. (2012). Engineering the production of major catechins by *Escherichia coli* carrying metabolite genes of *Camellia sinensis*. *The Scientific World Journal*, *2012*, 1–7.

van Dijk, E., Hoogeveen, A., & Abeln, S. (2015). The hydrophobic temperature dependence of amino acids directly calculated from protein structures. *PLoS Computational Biology*, *11*(5), 1–17.

Vera, A., González-Montalbán, N., Arís, A. and Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering*, *96*(6), 1101-1106.

Vikram, P., Chiruvella, K. K., Ripain, I. H. A., & Arifullah, M. (2014). A recent review on phytochemical constituents and medicinal properties of kesum *(Polygonum minus* Huds.). *Asian Pacific Journal of Tropical Biomedicine*, *4*(6), 430–435.

Wallin, E., & von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, *7*(4), 1029–1038.

Wan Hassan, W. E. (2006). *Healing herbs of Malaysia*. Kuala Lumpur, WP: Federal Land Development Authority (FELDA).

Wang, L., Liu, X., Meng, X., Wu, G., & Xu, F. (2018). Cloning and expression analysis of a chalcone isomerase (*CnCHI*) gene from *Chamaemelum nobile*. *Biotechnology*, *17*(1), 19–25.

Wang, W., Wang, H. L., Wan, S. B., Zhang, J. H., Zhang, P., Zhan, J. C., & Huang, W. D. (2012). Chalcone isomerase in grape vine: gene expression and localization in the developing fruit. *Biologia Plantarum*, *56*(3), 545–550.

Wang, Y., Chen, S., & Yu, O. (2011). Metabolic engineering of flavonoids in plants and microorganisms, *91*, 949–956.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL : homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*, 296–303.

Williamson, G., & Manach, C. (2005). Bioavailability and bioefficacy of polyphenols in humans. II. Review of 93 intervention studies. *The American Journal of Clinical Nutrition*, *81*, 243S-255S.

Winkel-Shirley, B. (2001). Flavonoid Biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology*, *126*(2), 485–493.

Yaacob, K. B. (1990). Essential oil of *Polygonum minus* Huds. *Journal of Essential Oil Research*, *2*(4), 167–172.

Yahyaa, M., Ali, S., Davidovich-Rikanati, R., Ibdah, M., Shachtier, A., Eyal, Y., Lewinsohn, E., & Ibdah, M. (2017). Characterization of three chalcone synthase-like genes from apple (*Malus x domestica* Borkh.). *Phytochemistry*, *140*, 125–133.

Yamamoto, S., Sobue, T., Kobayashi, M., Sasaki, S., & Tsugane, S. (2003). Soy, isoflavones, and breast cancer risk in Japan. *Journal of the National Cancer Institute*, *95*(12), 906–913.

Yan, B. X., & Sun, Y. Q. (1997). Glycine residues provide flexibility for enzyme active sites. *The Journal of Biological Chemistry*, *272*(6), 3190–3194.

Yin, Y. C., Zhang, X. D., Gao, Z. Q., Hu, T., & Liu, Y. (2019). The research progress of chalcone isomerase (CHI) in plants. *Molecular Biotechnology*, *61*(1), 32–52.

Yu, C.-S., Cheng, C.-W., Su, W.-C., Chang, K.-C., Huang, S.-W., Hwang, J.-K., & Lu, C.-H. (2014). CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. *PloS One*, *9*(6), 1-9.

Ziaei, S., & Halaby, R. (2017). Dietary isoflavones and breast cancer risk. *Medicines*, *4*(4), 1-11.