

A COMPUTATIONAL DOCKING ANALYSIS OF
CHALCONE SYNTHASE RECEPTOR AND
SUBSTRATES IN *Boesenbergia rotunda*

RAGAVENTHAN A/L SANMUGAVELAN

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

**A COMPUTATIONAL DOCKING ANALYSIS OF
CHALCONE SYNTHASE RECEPTOR AND
SUBSTRATES IN *Boesenbergia rotunda***

RAGAVENTHAN A/L SANMUGAVELAN

**DISSERTATION SUBMITTED IN FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER
OF SCIENCE**

**INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2019

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **RAGAVENTHAN A/L SANMUGAVELAN**

(I.C/Passport No:

Matric No: **SGR160014**

Name of Degree: **MASTER OF SCIENCE**

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

**A COMPUTATIONAL DOCKING ANALYSIS OF CHALCONE SYNTHASE
RECEPTOR AND SUBSTRATES IN *Boesenbergia rotunda***

Field of Study: **BIOINFORMATICS**

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature

Date:

Subscribed and solemnly declared before,

Witness’s Signature

Date:

Name:

Designation:

A COMPUTATIONAL DOCKING ANALYSIS OF CHALCONE SYNTHASE RECEPTOR AND SUBSTRATES IN *Boesenbergia rotunda*

ABSTRACT

Boesenbergia rotunda is locally known as Chinese ginger. Its rhizomes resemble fingers, making it known as fingerroot. Previous studies show that its rhizome extracts exhibited anticancer, antiviral, antibacterial and antioxidant properties. Numerous secondary metabolites can be found in the extract which are derived from the flavonoid biosynthetic pathway that involves a number of enzymes. Chalcone synthase (CHS) of *B. rotunda* (*BrCHS*) belongs to the type III polyketide synthase. It is a key enzyme involves in the initial stage of the flavonoid biosynthetic pathway. It has broad substrate specificity for diverse starter molecules and generating corresponding products. In this study, computational methods were employed to investigate the enzymatic activity of five *BrCHS* receptor variants with cinnamoyl-CoA and *p*-coumaroyl-CoA along with malonyl-CoA, feruloyl-CoA and caffeoyl-CoA as the ligand substrates. Homology models of the five variants of *BrCHS* receptor were built using YASARA software. ProtParam and SOPMA tools in ExPasy webserver were used to predict molecular mass and analyse the secondary structures of *BrCHS* receptor variants, respectively. HADDOCK 2.2 web server was used for molecular docking the *BrCHS* receptor variants with the ligand and docked energies were used for the docked conformation analysis. Then, 10 ns molecular dynamics simulations were performed using GROMACS 5.1.4 to verify the docking results based on the root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), the radius of gyration and binding free energies. The predicted molecular masses of the five *BrCHS* receptor variants are in the range of 42 – 44 kDa and secondary structure analysis revealed that the variants mainly comprise of α -

helices and random coils. The docking results showed that cinnamoyl-CoA has a higher binding affinity towards *BrCHS* receptor variants 1, 2 and 3 than *p*-coumaroyl-CoA. On the other hand, *p*-coumaroyl-CoA has a higher binding affinity towards *BrCHS* variants 4 and 5. Trajectory analysis based on RMSD, RMSF and radius of gyration revealed that the protein-ligand complexes were stable throughout the 10 ns simulations. In addition, binding free energy profiles showed a good agreement with the docking results and some experimental data. This study further enhances understandings on the substrate specificity and catalytic activity of the *BrCHS* in chalcone production.

Keywords: Chalcone synthase, *Boesenbergia rotunda*, molecular docking, molecular dynamics, cinnamoyl-CoA.

University of Malaya

ANALISIS DOK BERKOMPUTER BAGI RESEPTOR CHALCONE SYNTHASE DAN SUBSTRAT DALAM *BOESENBERGIA ROTUNDA*

ABSTRAK

Boesenbergia rotunda dikenali sebagai halia Cina dalam kalangan masyarakat tempatan. Tumbuhan tersebut juga dikenali sebagai Temu Kunci disebabkan oleh rizomnya yang merupai jari. Kajian-kajian lalu menunjukkan bahawa ekstrak rizomnya menunjukkan ciri-ciri antikanser, antivirus, antibakteria dan antioksidan. Pelbagai metabolit sekunder didapati dalam ekstrak tersebut yang berasal daripada laluan biosintesis flavonoid yang melibatkan pelbagai jenis enzim. Chalcone synthase (CHS) daripada *B. rotunda* atau *BrCHS* merupakan enzim jenis III poliketide sintesis. Enzim tersebut terlibat pada peringkat awal laluan biosintetik flavonoid. Enzim tersebut mempunyai pengkhususan substrat yang luas bagi molekul-molekul pemula dan menghasilkan produk-produk yang sepdannya. Dalam kajian ini, kaedah pengkomputeran telah digunakan untuk menyiasat aktiviti enzimatik bagi lima varian reseptor *BrCHS* dengan cinnamoyl-CoA dan *p*-Coumaroyl-CoA bersama-sama dengan malonyl-CoA, feruloyl-CoA dan caffeoyl-CoA sebagai substrat ligan. Model-model homologi bagi lima varian reseptor *BrCHS* telah dibina melalui perisian 'YASARA,' 'ProtParam' dan 'SOPMA' di dalam web 'Expasy', masing-masing telah digunakan untuk meramal jisim molekul dan menganalisis struktur sekunder varian-varian *BrCHS*. 'HADDOCK 2.2' telah digunakan untuk dok varian-varian *BrCHS* dengan ligan-ligan dan skor dok telah digunakan untuk analisis dok. Kemudian, simulasi molekul dinamik telah dilakukan melalui 'GROMACS v5.1.4' selama 10 ns bagi mengesahkan keputusan analisis dok berdasarkan punca min sisihan kuasa dua (RMSD), punca min fluktuasi kuasa dua (RMSF), jejari legaran dan tenaga ikatan bebad. Jisim molekul yang diramalkan bagi varian-varian reseptor *BrCHS* adalah

di antara 42 – 44 kDa dan analisis struktur sekunder didapati bahawa varian-varian tersebut mengandungi kebanyakannya heliks alfa dan gegelung rawak. Keputusan dok menunjukkan bahawa cinnamoyl-CoA mempunyai daya perikatan yang lebih tinggi terhadap jenis *BrCHS* 1, 2 dan 3 daripada *p*-coumaroyl-CoA. Manakala, *p*-Coumaroyl-CoA mempunyai daya perikatan yang lebih tinggi terhadap jenis *BrCHS* 4 dan 5. Hasil analisis trajektori berpandukan RMSD, RMSF dan jejari legaran didapati bahawa kompleks-kompleks protein-ligan stabil sepanjang simulasi bagi 10 ns. Tambahan pula, keputusan bagi tenaga ikatan bebas menunjukkan persetujuan yang baik dengan keputusan dok dan beberapa data eksperimen. Justeru itu, kajian ini akan meningkatkan pemahaman tentang pengkhususan substrat dan aktiviti enzimatik *BrCHS* dalam pengeluaran chalcone.

Kata kunci: Chalcone synthase, *Boesenbergia rotunda*, molekul dok, molekul dinamik, cinnamoyl-CoA.

ACKNOWLEDGEMENTS

First and foremost, praises and thanks to the almighty God for His showers of blessings and granting me the capability to complete my research work successfully. I would like to express my sincere thanks to my supervisor, Dr Teoh Teow Chong for his invaluable guidance, constant supervision and support in completing this endeavor. His sincerity, vision, motivation and dynamism have deeply inspired me. It was a great privilege and honor to work under his guidance.

Furthermore, I am thankful for the University of Malaya's research grant RP032C-15AFR for supporting this research work. In addition, I would like to acknowledge my thanks to Dr Teh Ser Huy for providing the amino acid sequences of the five variants of *BrCHS* receptor. My sincere gratitude to the lab officers for substantial assistance throughout this project.

I would like to express my deepest thanks to my beloved family. Without their encouragement and unconditional love, this project would not be completed successfully. Lastly, I am grateful for all the people who have supported me to complete the research work directly or indirectly.

TABLE OF CONTENTS

ABSTRACT	iii
ABSTRAK	v
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	viii
LIST OF FIGURES	xi
LIST OF TABLES	xiii
LIST OF SYMBOLS AND ABBREVIATIONS	xv
LIST OF APPENDICES	xvii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	5
2.1 <i>Boesenbergia rotunda</i>	5
2.1.1 Traditional Uses	6
2.1.2 Pharmaceutical Properties	6
2.1.3 Chalcone Synthase	8
2.2 Multiple Sequence Alignment (MSA).....	10
2.3 Molecular Modelling	11
2.3.1 Homology Modelling	11
2.4 Molecular Docking	12
2.4.1 HADDOCK 2.2 Web Server	13
2.5 Molecular Dynamics.....	14
2.5.1 GROMACS	16
2.6 Binding Free Energy.....	16
2.6.1 GMXPBSA 2.1	17

CHAPTER 3: METHODOLOGY	19
3.1 Hardware.....	19
3.2 Software and Web Server	19
3.3 General Workflow	20
3.4 Molecular Modelling	21
3.4.1 Ligand Preparation	21
3.4.2 Homology Modelling	23
3.5 Multiple Sequence Alignment (MSA).....	24
3.6 Molecular Docking	25
3.6.1 Docked Conformation Analysis	25
3.7 Molecular Dynamics (MD) Simulation.....	26
3.7.1 Trajectory Analysis	27
3.7.1.1 Calculation of Binding Free Energy.....	28
CHAPTER 4: RESULTS	30
4.1 Molecular Modelling	30
4.1.1 Homology Modelling	30
4.2 Multiple Sequence Alignment (MSA).....	37
4.3 Molecular Docking Analysis	43
4.4 Trajectory Analysis of Molecular Dynamics Simulations	63
4.4.1 Root-Mean-Square Deviation (RMSD).....	63
4.4.2 Root-Mean-Square Fluctuation (RMSF).....	66
4.4.3 Radius of Gyration	69
4.4.4 Binding Free Energy	72

CHAPTER 5: DISCUSSION	93
5.1 Molecular Modelling	93
5.1.1 Homology Modelling	93
5.2 Multiple Sequence Alignment (MSA).....	95
5.3 Docked Conformation Analysis	97
5.4 Molecular Dynamics Simulation	104
5.4.1 Trajectory Analysis	104
5.4.1.1 Binding Free Energy	106
CHAPTER 6: CONCLUSION.....	111
REFERENCES.....	112
LIST OF PUBLICATIONS AND PAPERS PRESENTED	133
APPENDIX.....	135

LIST OF FIGURES

Figure 2.1	: <i>Boesenbergia rotunda</i> (fingerroot) with its rhizomes and flower	6
Figure 2.2	: Structures of pinostrobin (left) and boesenbergin B (right) isolated from <i>B. rotunda</i>	7
Figure 2.3	: Flavonoid biosynthetic pathway	9
Figure 3.1	: Brief workflow of the research project	20
Figure 3.2	: Structures of ligands	22
Figure 3.3	: Workflow of calculation steps in GMXPBSA 2.1 tool	28
Figure 4.1	: Homology model of the <i>BrCHSv1</i> receptor	31
Figure 4.2	: Superimposition of homology models of <i>BrCHS</i> receptor variants with its respective template models	32
Figure 4.3	: Ramachandran plots of <i>BrCHS</i> receptor variants after homology modelling	35
Figure 4.4	: Ramachandran plots of <i>BrCHS</i> receptor variants for post-minimization	36
Figure 4.5	: Multiple sequence alignment (MSA) of five variants of <i>BrCHS</i> with <i>Medicago sativa</i> , <i>Oryza sativa</i> , <i>Zea mays</i> , <i>Curcuma longa</i> , <i>Curcuma alismatifolia</i> and <i>Musa acuminata</i>	38
Figure 4.6	: Cartoon rendering of <i>BrCHSv2</i> receptor	42
Figure 4.7	: Docked conformations of <i>BrCHSv1</i> with the ligands	45
Figure 4.8	: Docked conformations of <i>BrCHSv2</i> with the ligands	49
Figure 4.9	: Docked conformations of <i>BrCHSv3</i> with the ligands	53
Figure 4.10	: Docked conformations of <i>BrCHSv4</i> with the ligands	57
Figure 4.11	: Docked conformations of <i>BrCHSv5</i> with the ligands	61
Figure 4.12	: RMSD of <i>BrCHS</i> receptor variants with the substrate ligands after 10 ns simulation	64
Figure 4.13	: RMSF of <i>BrCHS</i> receptor variants with the substrate ligands after 10 ns simulation	67
Figure 4.14	: Radius of gyration of <i>BrCHS</i> receptor variants with the substrate ligands after 10 ns simulation	70

Figure 4.15	:	Interactions of <i>BrCHSv1</i> with the ligands after 10 ns simulation	75
Figure 4.16	:	Interactions of <i>BrCHSv2</i> with the ligands after 10 ns simulation	79
Figure 4.17	:	Interactions of <i>BrCHSv3</i> with the ligands after 10 ns simulation	83
Figure 4.18	:	Interactions of <i>BrCHSv4</i> with the ligands after 10 ns simulation	87
Figure 4.19	:	Interactions of <i>BrCHSv5</i> with the ligands after 10 ns simulation	91

University of Malaya

LIST OF TABLES

Table 2.1	: Taxonomical classification of <i>B. rotunda</i>	5
Table 3.1	: List of software and web servers	19
Table 3.2	: Templates used for homology modelling	24
Table 3.3	: List of clusters and corresponding structures formed for the docked complex	26
Table 4.1	: List of homology models, their RMSDs and templates used	30
Table 4.2	: Molecular weight and the isoelectric point of <i>BrCHS</i> receptor variants	33
Table 4.3	: Composition of secondary structures of <i>BrCHS</i> receptor variants	34
Table 4.4	: Ramachandran plot summary from RAMPAGE analysis	37
Table 4.5	: Docked energy of <i>BrCHS</i> variant 1 receptor with the ligands	43
Table 4.6	: List of residues of <i>BrCHSv1</i> formed interactions with the ligands	44
Table 4.7	: Docked energy of <i>BrCHS</i> variant 2 receptor with the ligands	47
Table 4.8	: List of residues of <i>BrCHSv2</i> formed interactions with the ligands	48
Table 4.9	: Docked energy of <i>BrCHS</i> variant 3 receptor with the ligands	51
Table 4.10	: List of residues of <i>BrCHSv3</i> formed interactions with the ligands	52
Table 4.11	: Docked energy of <i>BrCHS</i> variant 4 receptor with the ligands	55
Table 4.12	: List of residues of <i>BrCHSv4</i> formed interactions with the ligands	56
Table 4.13	: Docked energy of <i>BrCHS</i> variant 5 receptor with the ligands	59
Table 4.14	: List of residues of <i>BrCHSv5</i> formed interactions with the ligands	60
Table 4.15	: Binding free energy of <i>BrCHS</i> variant 1 receptor with the ligands	73
Table 4.16	: List of interactions formed between <i>BrCHSv1</i> and the ligands	74
Table 4.17	: Binding free energy of <i>BrCHS</i> variant 2 receptor with the ligands	77
Table 4.18	: List of interactions formed between <i>BrCHSv2</i> and the ligands	78
Table 4.19	: Binding free energy of <i>BrCHS</i> variant 3 receptor with the ligands	81

Table 4.20	: List of interactions formed between <i>BrCHSv3</i> and the ligands	82
Table 4.21	: Binding free energy of <i>BrCHS</i> variant 4 receptor with the ligands	85
Table 4.22	: List of interactions formed between <i>BrCHSv4</i> and the ligands	86
Table 4.23	: Binding free energy of <i>BrCHS</i> variant 5 receptor with the ligands	89
Table 4.24	: List of interactions formed between <i>BrCHSv5</i> and the ligands	90

University of Malaya

LIST OF SYMBOLS AND ABBREVIATIONS

2D	:	Two-dimensional
3D	:	Three-dimensional
<i>BrCHS</i>	:	Chalcone synthase of <i>Boesenbergia rotunda</i>
CHARMM	:	Chemistry at HARvard Macromolecular Mechanics
CHI	:	Chalcone isomerase
CHS	:	Chalcone synthase
CID	:	Compound Identifier
CoA	:	Coenzyme-A
<i>CfCHS</i>	:	Chalcone synthase of <i>Coleus forskohlii</i>
GROMACS	:	GRoningen MACHine for Chemical Simulation
HADDOCK	:	High Ambiguity Driven protein-protein DOCKing
HPC	:	High Performance Computing
MD	:	Molecular dynamics
MM-PBSA	:	Molecular mechanics Poisson-Boltzmann surface Area
MSA	:	Multiple sequence alignment
NMR	:	Nuclear Magnetic Resonance
NPT	:	Constant Number of particles, Pressure and Temperature
NVT	:	Constant Number of particles, Volume and Temperature
PBC	:	Periodic Boundary Conditions
PDB	:	Protein Data Bank
pI	:	Isoelectric point
PKS III	:	Polyketide synthase type III
PME	:	Particle mesh Ewald
RMSD	:	Root-mean-square deviation

RMSF : Root-mean-square fluctuation
SOPMA : Self-Optimized Prediction Method with Alignment
TIP3P : Transferable intermolecular potential with 3 points
VMD : Visual Molecular Dynamics
YASARA : Yet Another Scientific Artificial Reality Application

University of Malaya

LIST OF APPENDICES

Appendix A	: Flavonoid biosynthetic pathway	135
Appendix B	: In-house amino acid sequences of <i>BrCHS</i> receptor variants	136
Appendix C	: Amino acid sequences used in multiple sequence alignment	137
Appendix D	: Amino acid sequences of templates for homology modelling ...	138

University of Malaya

CHAPTER 1: INTRODUCTION

Flavonoids are natural products that are widely found in flowers, vegetables and fruits. They belong to secondary metabolites that consist of variable polyphenolic structures (Panche et al., 2016). Flavonoids play a crucial role in plant biological functions such as pollination, regulation of the nodulation, protection against infection and ultraviolet (UV) radiation (Mandal et al., 2010; Mahajan et al., 2011; Verdan et al., 2011). Numerous studies have been carried out on investigating its pharmaceutical properties, for instance, antioxidant, antimicrobial, anti-inflammatory, anticancer activities (Chang et al., 1993; Cao et al., 1997; Lee et al., 2009; D'Mello et al., 2011).

Chalcones belongs to the subclass of flavonoids which contain basic flavonoid skeleton structure with the absence of 'ring C' (Panche et al., 2016). They play a vital role in plants in terms of pigmentation of flowers and defense mechanism (Batovska & Todorova, 2010). Naringenin chalcone, pinocembrin chalcone, licochalcone, cardamomin, isobavachalcone and brousochalcone A are the few examples of naturally occurring chalcones in plants (Cheng et al., 2001; Fu et al., 2004; Hatzieremia et al., 2006; Nishimura et al., 2007; Orlikova et al., 2011). Chalcones exhibit potential antifungal, anticancer, cardioprotective, antimicrobial and antioxidant activities (Gafner et al., 1996; Liu & Go, 2007; Boumendjel et al., 2008; Zhong et al., 2015).

Chalcone synthase (CHS, EC 2.3.1.74) belongs to the superfamily of type III polyketide synthase (PKS) involve in producing different chalcones and numerous secondary metabolites such as auronones, stilbenes and flavonoids (Abe et al., 2006). The enzyme is the precursor of the flavonoid biosynthetic pathway that produces diverse plant metabolites (Yu et al., 2012). Stilbene synthase (STS) (Austin et al., 2004), pentaketide chromone synthase (Morita et al., 2007) and 2-pyrone synthase (Jez et al., 2000) are the few examples of members of the superfamily.

Numerous CHS enzymes belong to the superfamily were reported to date. For instance, CHS enzymes were found in *Sorghum bicolor* (Lo et al., 2002), *Physcomitrella patens* (Jiang et al., 2006), *Pinus sylvestris* (Fliegmann et al., 1992), *Glycine max* (Wingender et al., 1989) and *Silybum marianum* (Sanjari et al., 2015). Moreover, *CHS* genes are also expressed in *Zea mays* (Han et al., 2016), *Gerbera hybrid* (Deng et al., 2014), *Freesia hybrida* (Sun et al., 2015) and *Oncidium* orchid (Liu et al., 2012). Other than that, CHS genes are also found in *Petunia hybrid* (van der Meer et al., 1990), *Solanum lycopersicum* (tomato) (España et al., 2014), *Arabidopsis thaliana* (Feinbaum & Ausubel, 1988) and *Citrus sinensis* (L.) Osbeck (Moriguchi et al., 1999).

Generally, CHS uses three molecules of malonyl-CoA as extender substrates and one molecule of *p*-coumaroyl-CoA as the most preferential starter substrate (Ferrer et al., 1999). The enzyme has shown to exhibit starter substrate and product diversities. CHS prefers a wide range of aliphatic and aromatic thioesters, for instance, feruloyl-CoA, phenylacetyl-CoA, benzoyl-CoA, caffeoyl-CoA and many more to produce diverse novel polyketides (Christensen et al., 1998; Morita et al., 2000; Katsuyama et al., 2010). Moreover, the number of polyketide chain elongations and cyclization process can result in product diversity of the enzyme. Small changes in the residues at the active site may contribute to the functional diversity of the CHS enzyme (Jez et al., 2001b).

Boesenbergia rotunda is locally known as fingerroot or Chinese keys. It is traditionally used as a remedy for stomachache, muscle pain and cure for parasitic infections (Tushar et al., 2010; Eng-Chong et al., 2012). It contains secondary metabolite compounds that exhibit potential pharmaceutical properties such as antioxidant, anticancer and antimicrobial activities. Panduratin A, boesenbergin A, pinocembrin, cardamonin are the phytochemicals extracted from the plant (Supinya et al., 2003; Tan et al., 2015).

CHS from fingerroot belongs to the superfamily of the enzyme. It was reported that the five *BrCHS* receptor variants were predominantly detected and highly expressed in a tissue-specific manner from *B. rotunda* (Nurnadiah, 2017). In an *in vitro* study, it was reported that variant 2 of CHS from the plant was only active towards cinnamoyl-CoA but not *p*-coumaroyl-CoA (Sanmugavelan et al., 2018). Besides that, pinocembrin was one of the major compounds detected using high-performance liquid chromatography (HPLC) (Tan et al., 2015). However, very little is known about the origin or source of this functional diversity, and the factors that influence them. It is assumed that different *BrCHS* variants have different substrate preference.

In silico technology has given new insight to explore the substrate specificity of protein receptors (Roche & McGuffin, 2016). Homology modelling and site-directed mutagenesis studies revealed the importance of residues in active site cavity on the functional diversity of CHS from *Physcomitrella patens* (Rahman et al., 2012). *In silico* approach has been utilized to analyze substrate specificity of chalcone synthase gene from *Coleus forskohlii* using homology modelling and molecular docking (Awasthi et al., 2016).

Besides that, molecular dynamics simulations have been used to study the catalytic mechanism and efficiency of CHS from basal land plants (Liou et al., 2018). Thus, this research is carried out to investigate the substrate specificity and mechanism of *BrCHS* receptor variants using an *in silico* approach. The information obtained from this study would provide an insight into the production of novel polyketides from *BrCHS*s via *in vitro* biotechnology.

Hence, the objectives of the study are as follows:

- 1) To build models of the five *BrCHS* receptor variants by homology modelling,
- 2) To perform the molecular docking simulation of substrate ligands to *BrCHS* receptor variants,
- 3) To perform molecular dynamics simulations of the docked complexes and verify the docked conformation analysis using binding free energy,
- 4) To infer the mechanism of *BrCHS* receptor variants to its products such as naringenin and pinocembrin based on docked conformations and energies.

University of Malaya

CHAPTER 2: LITERATURE REVIEW

2.1 *Boesenbergia rotunda*

Fingerroot (Chinese keys) is a local name for *Boesenbergia rotunda* which is commonly used as food and traditional medicine. It is also known as ‘Krachai’ in Thailand, ‘Ao Chun Jiang’ in China and ‘Temu Kunci’ in Malaysia (Veldkamp, 2013). *B. rotunda* belongs to the family of Zingiberaceae. Table 2.1 shows the taxonomical classification of the plant based on the (Integrated Taxonomic Information System, 2018).

Table 2.1: Taxonomical classification of *B. rotunda*.

Rank	Taxa
Kingdom	Plantae
Division	Tracheophyta
Class	Magnoliopsida
Order	Zingiberales
Family	Zingiberaceae
Genus	Boesenbergia
Species	<i>Boesenbergia rotunda</i>

(Integrated Taxonomic Information System, 2018).

It has a strong aromatic rhizome which its colour depends on the variety of the plant. The yellow variety has bright-yellow rhizomes while other varieties have red and black rhizomes. The rhizomes resemble fingers that grow from the central part of the plant as shown in Figure 2.1. It is a perennial plant that grows up to 40 cm. The plant has red leaf sheath and broad, light green leaves. The leaves have width and length up to 11 cm and 50 cm respectively. It's tubular, pink flowers usually grow at the base of foliage (Eng-Chong et al., 2012; Ongwisespaiboon & Jiraungkoorskul, 2017). *B. rotunda* is natively distributed from China to West Malesia. It usually grows naturally in damp, shaded areas either in lowland or hill slopes (Ching et al., 2007).



Figure 2.1: *Boesenbergia rotunda* (fingerroot) with its rhizomes and flower (Adapted from National Parks Boards of Singapore, 2013; Lim, 2016).

2.1.1 Traditional Uses

B. rotunda is an edible plant that is used as a condiment in food among Asian people because of its strong aromatic flavour. This plant helps in treating stomachache, gout, muscle pain and rheumatism. It is used as an important ingredient in preparing “Jamu” among Indonesians that serves as a tonic for women after labour (Stevensen, 1999). In addition, it is used to cure inflammatory diseases, for instance, dermatitis, dental carries and fungal and parasitic infections (Tushar et al., 2010; Eng-Chong et al., 2012). Besides that, the herbal plant is used as an aphrodisiac in Thailand (Ongwisepaiboon & Jiraungkoorskul, 2017).

2.1.2 Pharmaceutical Properties

Various phytochemical compounds were found in different parts of the plant. Flavonoids such as pinocembrin, panduratin, pinostrobin, boesenbergin A, boesenbergin B, cardamonin and alpinetin are generally found in the plant (Ching et al., 2007; Isa et al., 2013; Tan et al., 2015). Figure 2.2 shows the structures of pinostrobin and

boesenbergin B isolated from fingerroot (Eng-Chong et al., 2012). Other than that, polyphenols including naringin, caffeic acid, *p*-coumaric acid, kaempferol and quercetin are also found in the fingerroot (Jing et al., 2010). Camphor, borneol, neryl acetate, geraniol, rosephenone and terpinyl valerate are the examples of essential oils found in the plant (Zaeoung et al., 2005; Sukari et al., 2008; Baharudin et al., 2015).

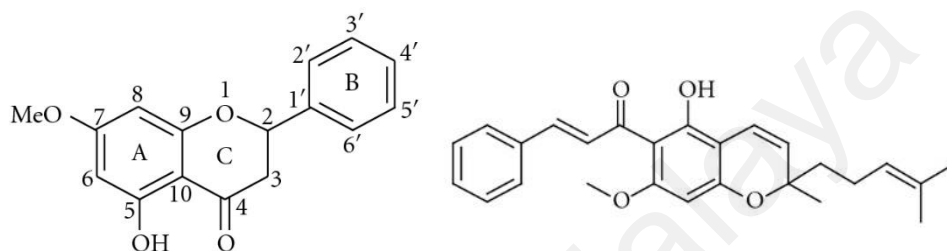


Figure 2.2: Structures of pinostrobin (left) and boesenbergin B (right) isolated from *B. rotunda* (Adpated from Eng-Chong et al., 2012).

Previous studies show that the rhizomes of *B. rotunda* exhibited potential inhibitory activities such as antibacterial, antiparasitic, antioxidant, anticancer, antiviral, antifungal and anti-inflammatory activities. The ethanolic extract of *B. rotunda* showed potential antibacterial activity by inhibiting *Bacillus subtilis*, *Staphylococcus aureus* and *Staphylococcus epidermidis* with values of minimum inhibitory concentration (MIC) from 0.04 to 25 mg/mL (Jitvaropas et al., 2012). Salama et al. (2012) reported that the plant extract inhibited the progression of liver cirrhosis induced by thioacetamide in a rat model.

It was reported that the medicinal plant is effective as an anticancer due to the presence of quercetins. It reduced the proliferative activity in cancer cell lines including colon, breast, ovarian, cervical cancer cell lines (Jing et al., 2010). In addition, boesenbergin A of the fingerroot contributed to significant anti-inflammatory, antioxidant and cytotoxic

activities (Isa et al., 2012). Pinocembrin isolated from the plant serves as a glucosidase inhibitor and anti-glycation agent (Potipiranun et al., 2018).

2.1.3 Chalcone Synthase

Chalcone synthase (CHS, EC 2.3.1.74) belongs to type III polyketide synthase enzyme (PKS) superfamily including pyrone synthase, acridone synthase, stilbene synthase, benzalacetone synthase, bibenzyl synthase, benzophenone synthase, curcuminoid synthase and olivetol synthase (Abe & Morita, 2010). It is a key enzyme which involves in the initial stage of flavonoid biosynthesis.

Several models have been used to establish the pathway. For instance, petunia (*Petunia hybrid*) and maize (*Zea mays*) (Koes et al., 1987; Han et al., 2016). CHS produces chalcone via the condensation of one CoA-linked molecule and three molecules of malonyl-CoA. The general reaction mechanism of CHS which was proposed by (Jez et al., 2001b) comprises three main steps which are loading, decarboxylation and elongation. Interaction of CoA-linked starter molecule with Cys164 residues of CHS will initiate the loading process. It is followed by the decarboxylation of malonyl-CoA by His303 and Asn336 in the catalytic triad (Jez & Noel, 2000; Abe et al., 2003). Two additional rounds of decarboxylation and elongation process continue with two molecules of malonyl-CoA and result in the tetraketide intermediate molecule.

The intermediate tetraketide undergoes several cyclization reactions resulting in the production of chalcone (Ferrer et al., 1999). Naringenin chalcone, as the product will be isomerized into a flavanone by chalcone isomerase (CHI) (Dao et al., 2011). Figure 2.3 shows the general pathway of flavonoid biosynthesis (Winkel-Shirley, 2001; Falcone Ferreyra et al., 2012; Kanehisa et al., 2017).

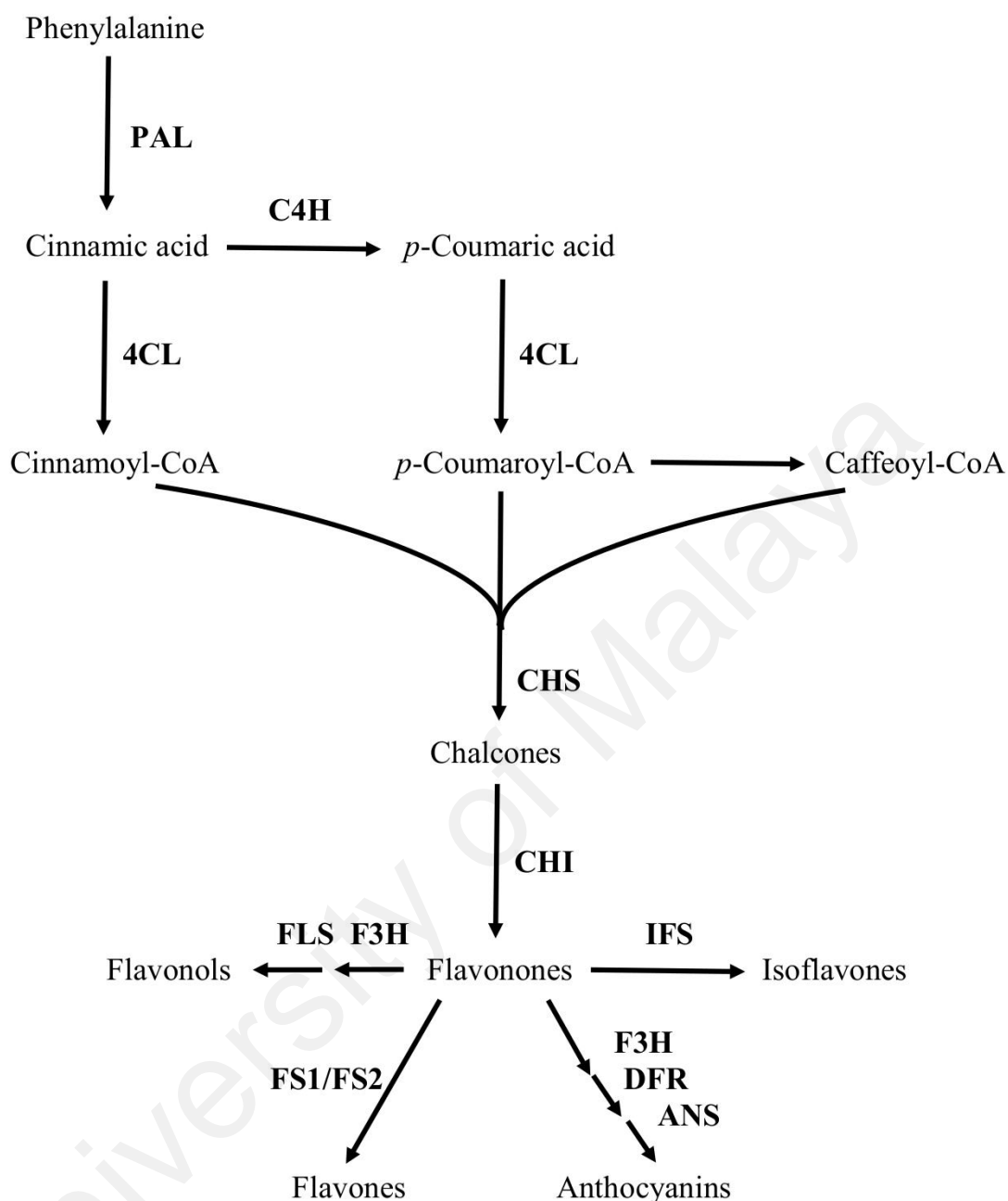


Figure 2.3: Flavonoid biosynthetic pathway (Winkel-Shirley, 2001; Falcone Ferreyra et al., 2012; Kanehisa et al., 2017). PAL, phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylase; 4CL, 4-coumarate-CoA-ligase; CHS, chalcone synthase; CHI, chalcone isomerase; FS1/FS2: flavone synthase 1 and 2; F3H, flavanone 3-hydroxylase; FLS, flavonol synthase; IFS, isoflavone synthase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase.

Ferrer and his colleagues (Ferrer et al., 1999) reported that alfalfa CHS2 comprises two domains, upper and lower domains. Upper domain serves as catalytic machinery, while lower domain serves as a space for the formation of chalcone. CHS has broad

substrate specificity for starter molecules and generates corresponding products. In many plants, *p*-coumaroyl-CoA is one of the most preferred starter substrates for CHS. With hexanoyl-CoA and benzoyl-CoA as starter molecules, alfalfa CHS2 formed tetraketide lactone and phlorobenzophenone respectively as the major products (Jez et al., 2001a).

Besides that, parsley CHS produces phlorobutyrophenone and tetraketide lactone when accepts feruloyl-CoA and butyryl-CoA, respectively as the starter molecule (Schüz et al., 1983). Meanwhile, CHS in *Fragaria vesca* generates triketide lactones when reacting with substrates such as isovaleryl-CoA and isobutyryl-CoA (Song et al., 2015). CHS2 from *Medicago sativa* yields phlorobenzyl ketone and methylpyrone as products when accepts phenylacetyl-CoA as starter molecule (Morita et al., 2000).

2.2 Multiple Sequence Alignment (MSA)

Multiple sequence alignment (MSA) is a bioinformatics tool that compares amino acid or nucleotide sequences to identify regions of similarity. It is a fundamental step in phylogenetic constructions and analysis of protein structure and functions (Edgar & Batzoglou, 2006). There are several methods available such as ClustalW (Thompson et al., 1994), PROBCONS (Christen et al., 2005), T-COFFEE (Notredame et al., 2000), Clustal Omega (Sievers et al., 2011) and MUSCLE (Edgar, 2004). As the latest MSA algorithm in the Clustal family, Clustal Omega algorithm produces a pairwise alignment of amino acid sequences using k-tuple method (Daugelaite et al., 2013). Sievers et al. (2011) reported that Clustal Omega, though has similar accuracy but it performs better than other packages.

2.3 Molecular Modelling

2.3.1 Homology Modelling

A methodology that generates a 3D model of a protein structure from its amino acid sequences that shares similarities is called homology modelling. It is a multi-step process of starting with template identification, followed by multiple sequence alignments, the building of a target model based on the 3D structure of the template, model refinement and lastly, model validation (Vyas et al., 2012). Identity between target and template sequences determines the quality of a model. If the sequence similarities are over 50%, the models built are accurate enough for further analysis. MODELLER (Sali & Blundell, 1993), YASARA (Yet Another Scientific Artificial Reality Application) (Krieger & Vriend, 2014) and SWISS-MODEL (Waterhouse et al., 2018) are the several programs and server used in homology modelling.

In the past years, homology modelling method was extensively used in drug discovery processes such as the study of the catalytic activity of enzymes, protein functions and biological role of mutations in protein mechanisms (Cavasotto & Phatak, 2009). Simulations of protein-protein docking, ligand search for a known binding site and antigenic epitopes prediction are the common applications of homology modelling (Vyas et al., 2012). Besides that, homology models are used for modelling substrate specificity. For instance, (Lukk et al., 2012) discovered the divergent substrate specificities of a group of dipeptide epimerases via homology modelling and molecular docking.

It is common that models generated contain errors and need to be refined and validated before subjected to further analysis. Errors in a model can be evaluated based on the calculation of root-mean-square deviation (RMSD) of backbone atoms and Z-score (Vyas et al., 2012). Tools such as PROCHECK (Laskowski et al., 1993), WHATIF (Vriend,

1990) and VERIFY3D (Eisenberg et al., 1997) are used to check protein stereochemistry and sequence fitness to the model.

2.4 Molecular Docking

In pharmaceutical research, molecular docking has been an increasingly important approach to elucidate the protein-ligand interactions. This method enables researchers to study biochemical processes based on the interaction of small molecules to proteins at the binding site (Meng et al., 2011).

Protein-ligand docking is one of the popular techniques among different docking types. This docking approach involves the prediction of conformation and orientation of ligands and its binding affinity for the binding site of the target protein. Information on the binding site can be obtained by comparing target protein to a family of proteins that are common in terms of functionality. In some cases, the information on binding sites is not yet known. There are several programs such as POCKET (Levitt & Banaszak, 1992) and GRID (Kastenholz et al., 2000) which can be used in predicting the binding sites and this approach is called blind docking (Meng et al., 2011).

There are various docking programs available. For examples, Genetic Optimization for Ligand Docking (GOLD) (Jones et al., 1997), DOCK (Kuntz et al., 1982; Allen et al., 2015), AutoDock (Morris et al., 2009), HADDOCK (Dominguez et al., 2003) and AutoDock Vina (Trott & Olson, 2010). Docking process includes mainly search algorithms and scoring functions. Search algorithms such as Genetic algorithms and Monte Carlo methods give the possible binding poses of the ligand within the target protein. On the other hand, scoring functions such as empirical, force-field-based and knowledge-based scoring functions predict and rank the docked conformations based on their binding free energies (Huang & Zou, 2006).

Molecular docking approach provides information that is useful in designing and developing more potent drugs and selective analogues. In addition, the approach is used in bioremediation field to discover possible enzymes that degrade pollutants (Liu et al., 2018). Docking technique has led to the discovery of laccase from *Ceriporiopsis subvermispora* which is a potential enzyme in the biotransformation of herbicide diuron (Vieira et al., 2015). Current advances in molecular docking make it is possible for flexible docking and modelling of the quaternary structure of complexes such as protein-protein complexes (de Ruyck et al., 2016). In medicinal chemistry, molecular docking approach sheds the light of structural information and underlying mechanisms of G protein-coupled receptors (GPCRs) (Bartuzi et al., 2017).

2.4.1 HADDOCK 2.2 Web Server

High Ambiguity Driven protein-protein DOCKing or HADDOCK 2.2 is a user-friendly web server that uses the data-driven approach in generating docking poses (Dominguez et al., 2003; van Zundert et al., 2016). It deals with different types of classes such as protein-ligand, protein-protein and protein-nucleic acids. It integrates information from mutagenesis, nuclear magnetic resonance (NMR) experiments and mass spectrometry. Haddock 2.2 web server allows small changes in the conformation of the receptor upon binding of a ligand (Spiliotopoulos & Caflisch, 2014). The clusters after docking are ranked according to their HADDOCK score along with other standard energy terms including van der Waals, electrostatics, desolvation and restraints violation energies (van Zundert et al., 2016). The best structures are chosen with the lowest HADDOCK score (Dominguez et al., 2003).

It provides multiple interfaces with different level of control over protocols of docking. Firstly, the Easy interface comprises the basic level of control over docking procedures on single and mixed molecules types. Guru interface allows users to control over up to

5000 parameters. Meanwhile, the Expert interface provides options for specifying the protonation state of each histidine residue and flexibility regions of a protein. In addition, this program also provides other interfaces such as prediction, refinement, multi-body and file upload interfaces (van Zundert et al., 2016).

2.5 Molecular Dynamics

With advances in technology, high-performance computers and methodologies of refined protein design make molecular dynamics tools play a crucial role in drug discovery. Molecular dynamics (MD) is a method using Newtonian physics to study the interaction and motion of atoms in molecules (De Vivo et al., 2016). The forces within interactions and energy profile of the system are estimated using a force field. MD trajectories produced give information on the position and velocities of the atoms over time (De Vivo et al., 2016).

Currently, there is much software available for molecular dynamics simulations. For instance, GRONingen Machine for Chemical Simulation (GROMACS) (Abraham et al., 2015), Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) (Plimpton, 1995), Assisted Model Building with Energy Refinement (AMBER) (Weiner & Kollman, 1981; Case et al., 2005), Nanoscale Molecular Dynamics (NAMD) (Phillips et al., 2005) and others.

Force fields, such as GROMOS (Christen et al., 2005), CHARMM (Brooks et al., 1983), Optimized Potentials for Liquid Simulations (OPLS) (Jorgensen & Tirado-Rives, 1988) and AMBER (Weiner & Kollman, 1981) are commonly used in MD simulations. The forces that cause motions on the atoms of the system can be categorised based on types of interactions. Intramolecular interactions include stretching, bending and torsion of the bond between atoms. Simple springs are used to describe the bond stretching and

atomic angles. Meanwhile, the sinusoidal function describes the dihedral angles or torsional. Non-bonded interactions include the electrostatic and van der Waals interactions between atoms. Both interactions are also known as Coulomb and Lennard-Jones potentials, respectively (Durrant & McCammon, 2011; De Vivo et al., 2016).

Periodic boundary conditions (PBC) are used in MD to describe bulk properties of complexes within a finite size system. Particle-mesh Ewald (PME) method, on the other hand, increases the efficiency of the MD simulations (de Souza & Ornstein, 1997). Thermostats are the system in which its temperature is kept constant throughout the simulation as if it is immersed in a thermostatic bath. The pressure of the simulated system is also controlled via the scaling of the system volume appropriately (De Vivo et al., 2016).

Various properties can be obtained via the MD trajectories such as energy profiles, free energy and others. Many tools are invented to carry out the trajectory analysis of MD simulations. For examples, AmberTools (Case et al., 2005), MDTraj (McGibbon et al., 2015), MDAAnalysis (Michaud-Agrawal et al., 2011; Gowers et al., 2016) and ProDy (Bakan et al., 2011).

Currently, the routine simulations can be carried out to the microsecond or even millisecond scale and usage of GPU in simulations (Hospital et al., 2015). Molecular dynamics aids in refining structure predictions via longer simulations particularly for *ab initio* protein structure. Besides that, MD simulations assist in understanding the energetics and allosteric transition details of a protein (Hospital et al., 2015).

Molecular docking coupled with the simulations add an advantage in improvising the virtual screening or docking. MD simulations are employed in food carbohydrate research such as interactions of carbohydrates with protein and inclusion complexation in

nutraceuticals and cosmetic fields (T. Feng, 2015). In biomedical research, MD simulations have been extensively utilized to assess the toxicity of nanomaterials (Selvaraj et al., 2018).

2.5.1 GROMACS

GROningen Machine for Chemical Simulation or GROMACS is freely available software that is commonly used to perform molecular dynamics. It is widely used in investigating bonded interactions among molecules such as nucleic acids, lipids and proteins. In addition, it is further utilised in research on non-biological systems such as polymers (Abraham et al., 2015). Simulations can be run in parallel using the message-passing interface (MPI) protocol.

This program supports a range of force field for instances GROMOS96 (van Gunsteren et al., 1996), GROMOS53a6 (Oostenbrink et al., 2004), Amber94 (Cornell et al., 1995), Amber99 (Wang et al., 2000), CHARMM27 (Vanommeslaeghe et al., 2010) and others. It allows the user to run simulations in modern cloud computing environments. The gromacs pre-processor (grompp) is employed prior to energy minimization or molecular dynamics simulations. The pre-processor not only read a molecular topology file, but it also validates and expands its molecular information into an atomic description. GROMACS can be used for calculations of QM/MM and the free energy of a molecule (Pirhadi et al., 2016).

2.6 Binding Free Energy

Binding free energy is a computational method that uses the principle of statistical thermodynamics to estimate the free energies of the protein-ligand complex (Thomas & Andreas, 2010). Calculations of binding free energy help in determining the binding affinity of a small molecule to its receptor or stability of the protein-ligand complex (Du

et al., 2016). Free energy methods take into account both the entropic and energetic contributions (de Ruiter & Oostenbrink, 2011).

There are several methods available for the calculation of binding free energies. For instance, alchemical calculation, endpoint and path sampling methods (Du et al., 2016). Free energy perturbation (FEP) approach belongs to the alchemical calculation method which estimates the absolute or relative binding affinities (Zwanzig, 1954; Chodera et al., 2011). Endpoint method is an efficient approach to estimating the binding free energy of a protein-ligand complex (Du et al., 2016). Molecular mechanics generalized born surface area (MM-GBSA) (Hou et al., 2011), molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) (Kollman et al., 2000) and linear interaction energy (LIE) (Aqvist et al., 1994) are the approaches using endpoint method. LIE method calculates the binding free energy that has a linear dependence on the changes in ligand-surrounding energies (Aqvist et al., 1994).

On the other hand, MM-PBSA and MM-GBSA methods estimate the binding free energy by calculating the changes in entropic contributions, solvation free and molecular mechanic energies (Kollman et al., 2000; de Ruiter & Oostenbrink, 2011). The endpoint method is a relatively fast way and has higher accuracy than the scoring and docking methods (Singh & Warshel, 2010; Hou et al., 2011; Genheden & Ryde, 2015).

2.6.1 GMXPBSA 2.1

GMXPBSA 2.1 is a program using Bash/Perl scripts to calculate the binding free energies of molecular dynamics trajectories of protein-protein or protein-ligand complexes. It uses MM-PBSA methods for the binding free energies of the complexes. It is also a freely available program under the General Public License (GPL). This tool utilises GROMACS and Adaptive Poisson-Boltzmann Solver (APBS) (Baker et al., 2001)

which is a Poisson-Boltzmann equation solver to calculate binding free energies using the frames extracted from trajectory files. In addition, it is useful in comparing binding free energies of different complexes trajectories and ranking the relative affinity of different ligands with the same receptor (Paissoni et al., 2014, 2015).

University of Malaya

CHAPTER 3: METHODOLOGY

3.1 Hardware

The computer with the dual interface of Ubuntu v16.04 and Windows 10, respectively was used in this study. The device has memory of 8 GB RAM with 2.50 GHz Intel(R) Core(TM) CPU i5-7200U processors and internal storage of 1 TB.

3.2 Software and Web Server

Table 3.1 shows the list of software and web servers used in the research.

Table 3.1: List of software and web servers.

Software	Web server
GROMACS v5.1.4	Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/)
YASARA Structure v18.4.24	SwissParam (http://www.swissparam.ch/)
GMXPBSA 2.1	PubChem Compound (https://pubchem.ncbi.nlm.nih.gov/search/search.cgi)
Discovery Studio Client v4.5.0.15701	RAMPAGE Server (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php)
PyMOL v1.3	HADDOCK 2.2 (http://milou.science.uu.nl/services/HADDOCK2.2/haddockserver-easy.html)
LigPlot ⁺ v1.4.5	SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)
Visual Molecular Dynamics (VMD) v1.9.3 Chimera v1.12	ExpASy (Compute pI/Mw Tool) (https://web.expasy.org/compute_pi/)

3.3 General Workflow

A brief workflow of the computational biochemical analysis used in the research is shown in Figure 3.1.

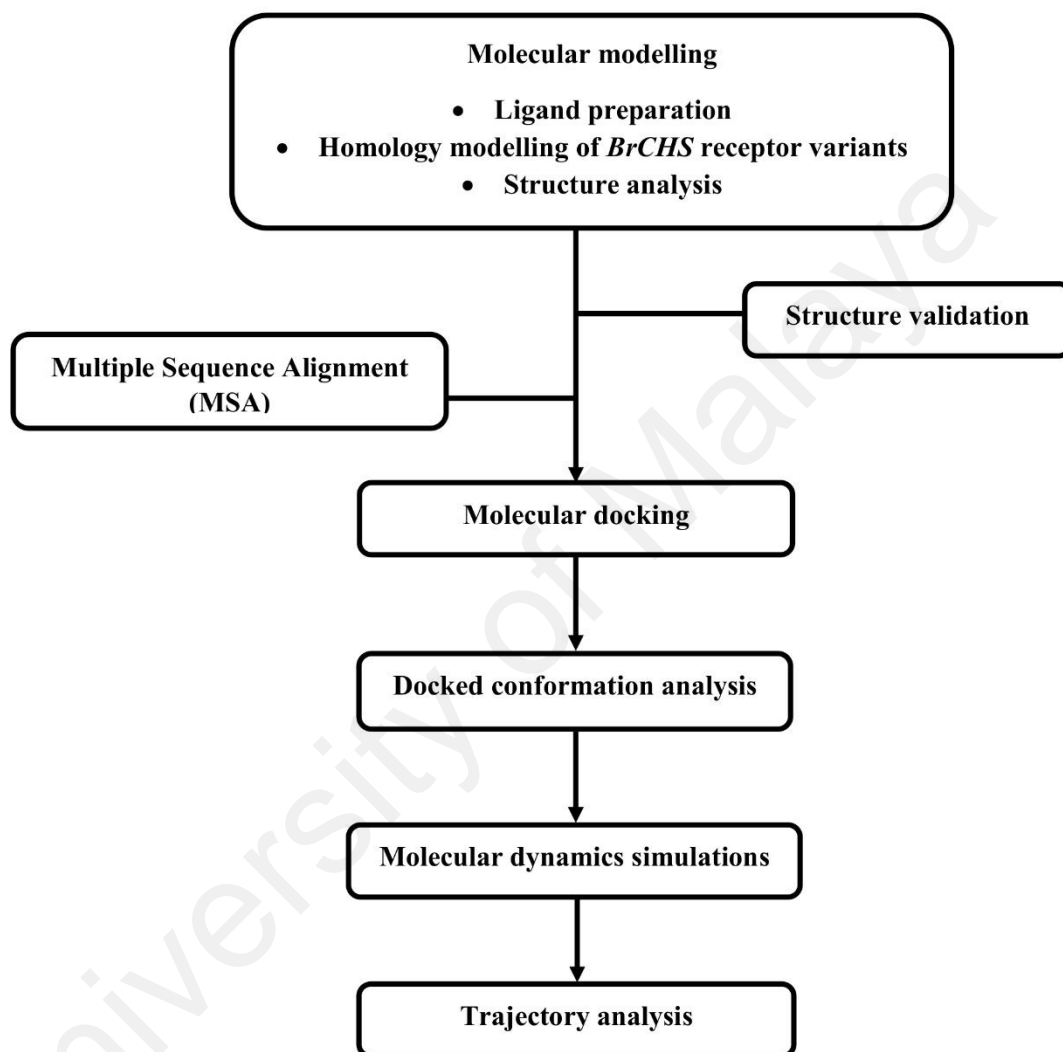


Figure 3.1: Brief workflow of the research project.

3.4 Molecular Modelling

3.4.1 Ligand Preparation

Five ligands [caffeoyl-CoA (CID: 5280336), cinnamoyl-CoA (CID: 6444037), *p*-coumaroyl-CoA (CID: 5462161), feruloyl-CoA (CID: 11966129) and malonyl-CoA (CID: 644066)] were chosen as substrate ligands based on the reference KEGG flavonoid biosynthetic pathway (https://www.genome.jp/kegg-bin/show_pathway?ko00941) (Kanehisa et al., 2017) (Appendix A). Meanwhile, acetyl-CoA (CID: 444493) and coenzyme-A (CoA) (CID: 46936280) were selected as reference ligands. All the ligands were retrieved from PubChem Compound (<https://pubchem.ncbi.nlm.nih.gov/search/search.cgi>) (Kim et al., 2015) and were chosen based on the molecular weight and chemical formula. All the ligands were optimized using the CHARMM force field in Discovery Studio Client v4.5.0.15071 (Accelrys Inc., Dassault Systemes, BIOVIA Corp) (Vanommeslaeghe et al., 2010). Figure 3.2 shows the structures of the substrate and reference ligands.

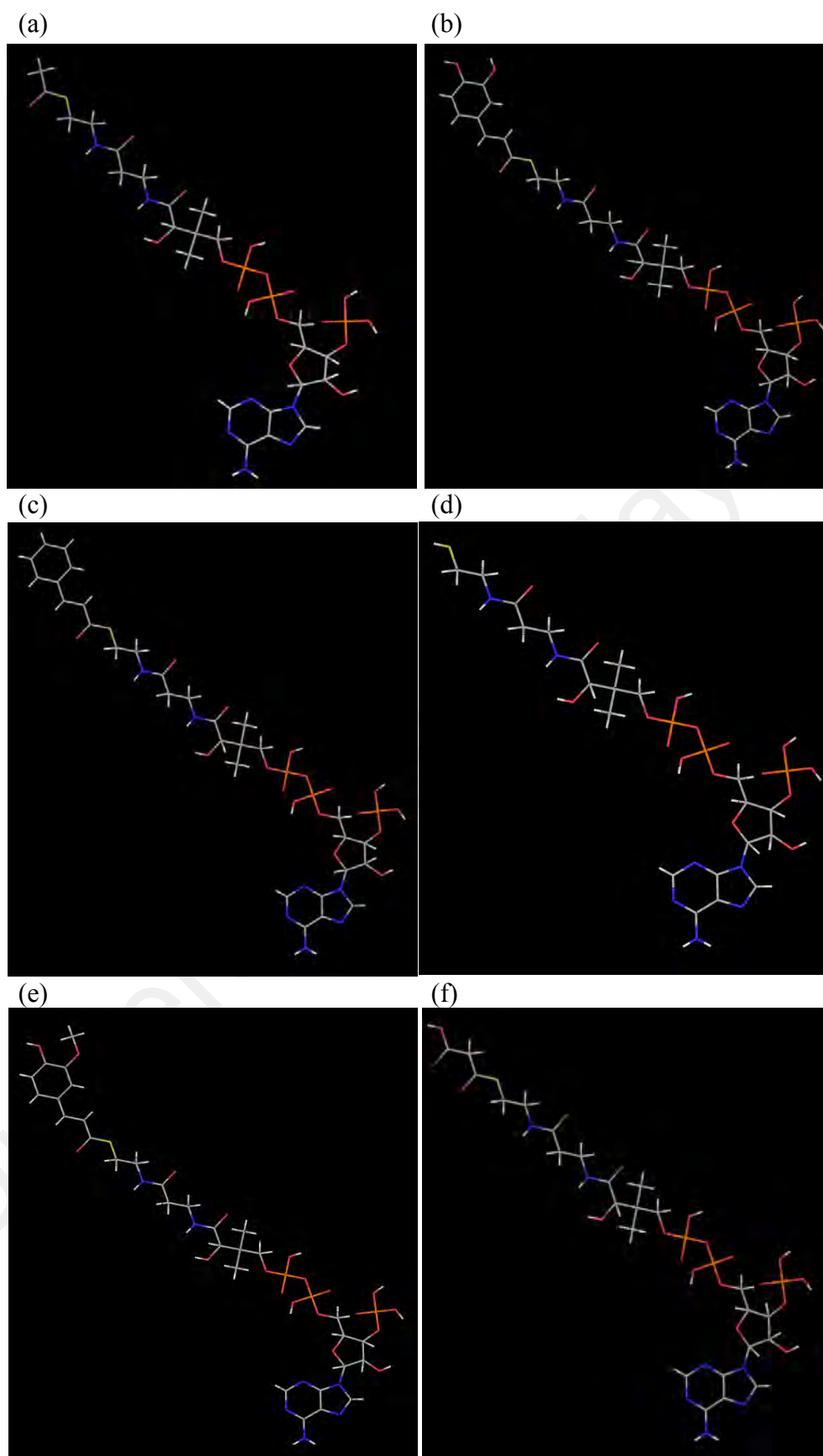


Figure 3.2: Structures of ligands. (a) Acetyl-CoA; (b) caffeoyl-CoA; (c) cinnamoyl-CoA; (d) CoA; (e) feruloyl-CoA; (f) malonyl-CoA; (g) *p*-coumaroyl-CoA.

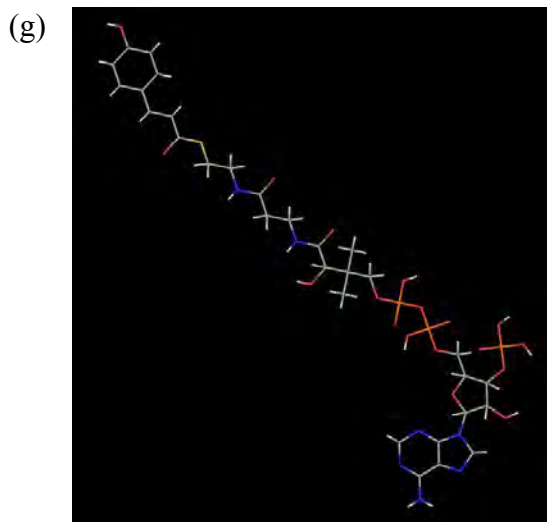


Figure 3.2, continued.

3.4.2 Homology Modelling

Protein models of five *BrCHS* receptor variants were built using homology modelling module in YASARA Structure software v18.4.24 (Krieger et al., 2002; Land & Humble, 2017) with in-house amino acids sequences (Appendix B). The software used a CASP (Critical Assessment of Structure Prediction) approved protocol that automatically handles all the modelling steps from an amino acid sequence until the production of a refined high-resolution model (Krieger et al., 2009). A position-specific scoring matrix (PSSM) from related sequences was built after the amino acids sequences PSI-BLASTed (Altschul et al., 1997) against Uniprot. The PDB for potential modelling templates were then searched using the profile. The templates were ranked based on the structural quality and alignment score according to WHATCHECK from the PDB Finder 2 database.

Several models were generated for the target sequences based on each of the top scoring five templates (4YJY, 1I88, 4WUM, 1I86, 1JWX) as shown in Table 3.2. The amino acids sequences for the template models are as shown in Appendix D. The sequence identity between templates and the five variants of *BrCHS* receptor ranged from 77% to 84%. The final homology models for the five *BrCHS* variants receptor were

obtained after the software automatically combined only the best parts from the generated models.

Table 3.2: Templates used for homology modelling.

Template	PDB ID	Total score	Blast E-value	Cover (%)	Resolution (Å)
1	4YJY-A	951.85	1e-153	100	1.86
2	1I88-B	899.00	1e-154	99	1.45
3	4WUM-C	893.14	6e-157	99	1.77
4	1I86-A	884.33	5e-155	99	1.50
5	1JWX-A	879.92	2e-152	99	1.63

The homology models were validated using Ramachandran plots generated via RAMPAGE server (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) (Lovell et al., 2003) for its stereochemical quality. PyMOL v1.3 (Schrödinger, LLC) was used to visualize the protein receptors. The homology models were then minimized for 100 ps in GROMACS v5.1.4 (Abraham et al., 2015) to energy convergence of 0.01 kJ/mol using steepest-descent and conjugate gradient, respectively. The quality of the minimized models was validated using the Ramachandran plot via RAMPAGE server (Lovell et al., 2003). The predicted molecular masses and theoretical isoelectric points of the *BrCHS* receptor variants were then evaluated using the pI/Mw tool in ExPASy (https://web.expasy.org/compute_pi/) (Gasteiger et al., 2005). The secondary structure of the protein receptors was predicted using SOPMA server (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (Geourjon & Deleage, 1995).

3.5 Multiple Sequence Alignment (MSA)

In-house amino acid sequences of the five *BrCHS* receptor variants were aligned with amino acid sequences from *Medicago sativa* (*MsCHS2*; L02902.1), *Oryza sativa* (*OsCHS*; AB000801.2), *Zea mays* (*ZmCHS*; NM_001155550.1), *Curcuma longa* (*CIPKS9*; JN017186.1), *Curcuma alismatifolia* (*CaCHS*; GU140082.1) and *Musa*

acumiata (*MaPKSIII3*; GU724609.1) (Appendix C). Multiple Sequence Alignment (MSA) was performed using Clustal Omega web server (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (Sievers et al., 2011) with default parameters including 20 × 20 Gonnet matrix as substitution matrix to perform multiple sequence alignment. Then, it was confirmed by the putative active site as reported by Jez and Noel (2000).

3.6 Molecular Docking

Chain A of each receptor variants was selected for docking with the selected substrate ligands. Molecular docking was performed using the easy interface module of Haddock 2.2 web server (<http://haddcok.science.uu.nl/services/HADDOCK/haddockserver-easy.html>) (van Zundert et al., 2016). The active site residues of the receptors were defined in the interface. The pdb files of both receptors and ligands were uploaded and the job was submitted.

3.6.1 Docked Conformation Analysis

After a successful docking run, clusters were sorted by HADDOCK score and numbered according to cluster size. The clusters with the lowest HADDOCK score were selected. Each cluster gives four conformations of protein-ligand. The first top structures of the chosen clusters were downloaded and used for further analysis. Table 3.3 shows the list of clusters and corresponding structures formed for the docked complex. The docked energies \pm standard deviation (SD) were tabulated. The docked conformations were analyzed in terms of van der Waals, electrostatic interactions, and a number of hydrogen bonds formed. LigPlot⁺ v1.4 (Laskowski & Swindells, 2011) software was used to view the 2D structures of the docked complexes. Besides that, Discovery Studio Client v4.5.0.15071 (Accelrys Inc., Dassault Systemes, BIOVIA Corp) was used to analyze

other interactions such as π -interactions, salt bridge and repulsive interactions in between protein and ligands.

Table 3.3: List of clusters and corresponding structures formed for the docked complex.

Ligand	Variant				
	<i>BrCHSv1</i>	<i>BrCHSv2</i>	<i>BrCHSv3</i>	<i>BrCHSv4</i>	<i>BrCHSv5</i>
Malonyl-CoA	8 clusters (166 structures)	4 clusters (156 structures)	12 clusters (168 structures)	7 clusters (164 structures)	8 clusters (168 structures)
Cinnamoyl-CoA	7 clusters (178 structures)	7 clusters (174 structures)	7 clusters (175 structures)	5 clusters (192 structures)	7 clusters (186 structures)
<i>p</i>-Coumaroyl-CoA	8 clusters (154 structures)	7 clusters (167 structures)	9 clusters (165 structures)	8 clusters (178 structures)	11 clusters (178 structures)
Caffeoyl-CoA	10 clusters (143 structures)	7 clusters (178 structures)	10 clusters (172 structures)	4 clusters (187 structures)	5 clusters (178 structures)
Feruloyl-CoA	10 clusters (162 structures)	10 clusters (165 structures)	10 clusters (155 structures)	7 clusters (171 structures)	10 clusters (167 structures)
Acetyl-CoA	4 clusters (175 structures)	6 clusters (182 structures)	1 clusters (194 structures)	2 clusters (190 structures)	1 clusters (194 structures)
CoA	3 clusters (191 structures)	1 clusters (197 structures)	3 clusters (191 structures)	2 clusters (199 structures)	3 clusters (197 structures)

3.7 Molecular Dynamics (MD) Simulation

GROMACS v5.1.4 (Abraham et al., 2015) was used for system preparation and simulations. The initial structures for the simulations were obtained from the docked results. All the simulations were performed using CPU at the High Performance Computing (HPC) cluster by the Data Intensive Computer Centre (DICC) of University of Malaya.

The coordinates of docked complexes were separated into two individual PDB files. One for ligands and another for the protein. Chimera v1.12 (Pettersen et al., 2004) was used to convert Protein Data Bank (PDB) files of ligands into .mol2 format. The GROMACS topologies for ligands were generated via SwissParam server

(<http://www.swissparam.ch/>) (Zoete et al., 2011) using the CHARMM force field (Vanommeslaeghe et al., 2010). The GROMACS topologies for the protein molecules were generated using the pdb2gmx module with CHARMM22/CMAP (MacKerell et al., 1998; MacKerell et al., 2004). The pdb2gmx module added the missing hydrogen atoms to the protein structure by default. The TIP3P model (Neria et al., 1996) was used to fill the explicit water molecules into the dodecahedron at a distance of 1.0 nm. A proper number of sodium ions were added to the system to neutralize its charge.

Prior to simulations, the solvated systems were minimized in order to remove any steric hindrance due to the presence of added hydrogen atoms of the protein-ligand complex. Firstly, the systems were minimized using position restraint method and then followed by the steepest descent and conjugate gradient methods. Equilibration of the systems started with an NVT (Constant Number of particles, Volume and Temperature) run for 100 ps which is directly equilibrated to 300 K (Berendsen et al., 1984). It was followed by NPT (Constant Number of particles, Pressure and Temperature) run for 100 ps at 1.0 atm and 300 K to equilibrate pressure of the systems using Parrinello-Rahman barostat (Parrinello & Rahman, 1981). Particle-Mesh Ewald (PME) method (Essmann et al., 1995) was used for the calculation of the long-range electrostatic interactions with a cut-off of 1.0 nm. The molecular dynamics simulations of the systems were carried out for 10 ns at the temperature of 300 K and pressure of 1.0 atm with a time step of 2 fs for integration.

3.7.1 Trajectory Analysis

The trajectory files were inspected using VMD v1.9.3 (Humphrey et al., 1996). Post-processing tool, trjconv was used to correct the periodicity of the protein molecules. Several GROMACS modules were used to analyze the properties of the trajectory files after the simulations (Abraham et al., 2015). Gmx rms module was used for root-mean-

square deviation (RMSD) calculations and checking the protein stability. Root-mean-square fluctuation (RMSF) of the proteins were obtained using the `gmx rmsf` module which calculates the fluctuation of C- α atoms coordinates from the average position. `Gmx gyrate` module was used to measure the compactness of the proteins. The interactions in the protein-ligand complexes were analyzed using LigPlot⁺ v1.4 (Laskowski & Swindells, 2011) and Discovery Studio Client v4.5.0.15071 (Accelrys Inc., Dassault Systemes, BIOVIA Corp).

3.7.1.1 Calculation of Binding Free Energy

The binding free energies of the complexes were calculated using GMXPBSA 2.1 module (Paissoni et al., 2015). Figure 3.3 shows the workflow of the calculations steps in GMXPBSA 2.1.

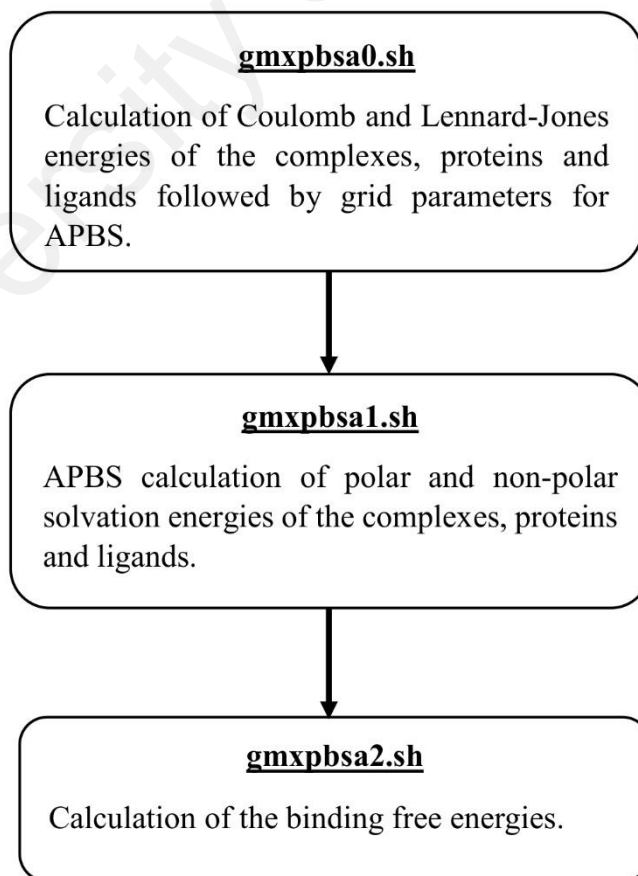


Figure 3.3: Workflow of calculation steps in GMXPBSA 2.1 tool.

The binding free energies of protein-ligand complexes were calculated as in Equation 3.1.

$$\Delta G_{binding} = G_{binding} - (G_{protein} + G_{ligand}) \quad (3.1)$$

The free energy term was calculated based on the average over the considered structures. Equation 3.2 shows the calculation for the free energy term.

$$\langle G \rangle = \langle E_{MM} \rangle + \langle G_{MM} \rangle - T \langle S_{MM} \rangle \quad (3.2)$$

The energetic term E_{MM} was calculated based on Equation 3.3 below. E_{int} denotes the energies of the bond, angle and torsional angle. E_{coul} and E_{LJ} denote the intermolecular electrostatic and Lennard-Jones energies, respectively.

$$E_{MM} = E_{int} + E_{coul} + E_{LJ} \quad (3.3)$$

The solvation term was calculated using Equation 3.4. The polar contribution, G_{polar} was calculated using the non-linearized or linearized Poisson-Boltzmann equation (Baker et al., 2001). Whereas, nonpolar contribution, $G_{nonpolar}$ was calculated based on the Equation 3.5, where $\gamma = 0.0227 \text{ kJ mol}^{-1}$ and $\beta = 0 \text{ kJ mol}^{-1}$ (Brown & Muchmore, 2009).

$$G_{solv} = G_{polar} + G_{nonpolar} \quad (3.4)$$

$$G_{nonpolar} = \gamma SASA + \beta \quad (3.5)$$

The binding free energy profiles of protein-ligand complexes were extracted and tabulated.

CHAPTER 4: RESULTS

4.1 Molecular Modelling

4.1.1 Homology Modelling

Table 4.1 shows the RMSDs of the homology models of the *BrCHS* variant receptors and the templates used for generating the final homology models. All the receptor variants used 4YJY as the template except *BrCHSv3*. Figure 4.1 depicts the 3D structures of *BrCHSv1* receptor after homology modelling in YASARA. All the receptor variants are homodimers. Structure of chain A of the *BrCHSv1* receptor consists of α -helix (red), β -sheets (yellow) and loop (green). Figure 4.2 shows the superimposition of the *BrCHS* receptor variants with its respective template structures.

Table 4.1: List of homology models, their RMSDs and templates used.

Homology model	RMSD (Å)	Template (s) [PDB ID]
<i>BrCHSv1</i>	0.478	4YJY-A, 4WUM-C (Hybrid model)
<i>BrCHSv2</i>	0.167	4YJY-A
<i>BrCHSv3</i>	0.466	1JWX-A
<i>BrCHSv4</i>	0.155	4YJY-A
<i>BrCHSv5</i>	0.551	4YJY-A, 1I88-B, 1JWX-A (Hybrid model)

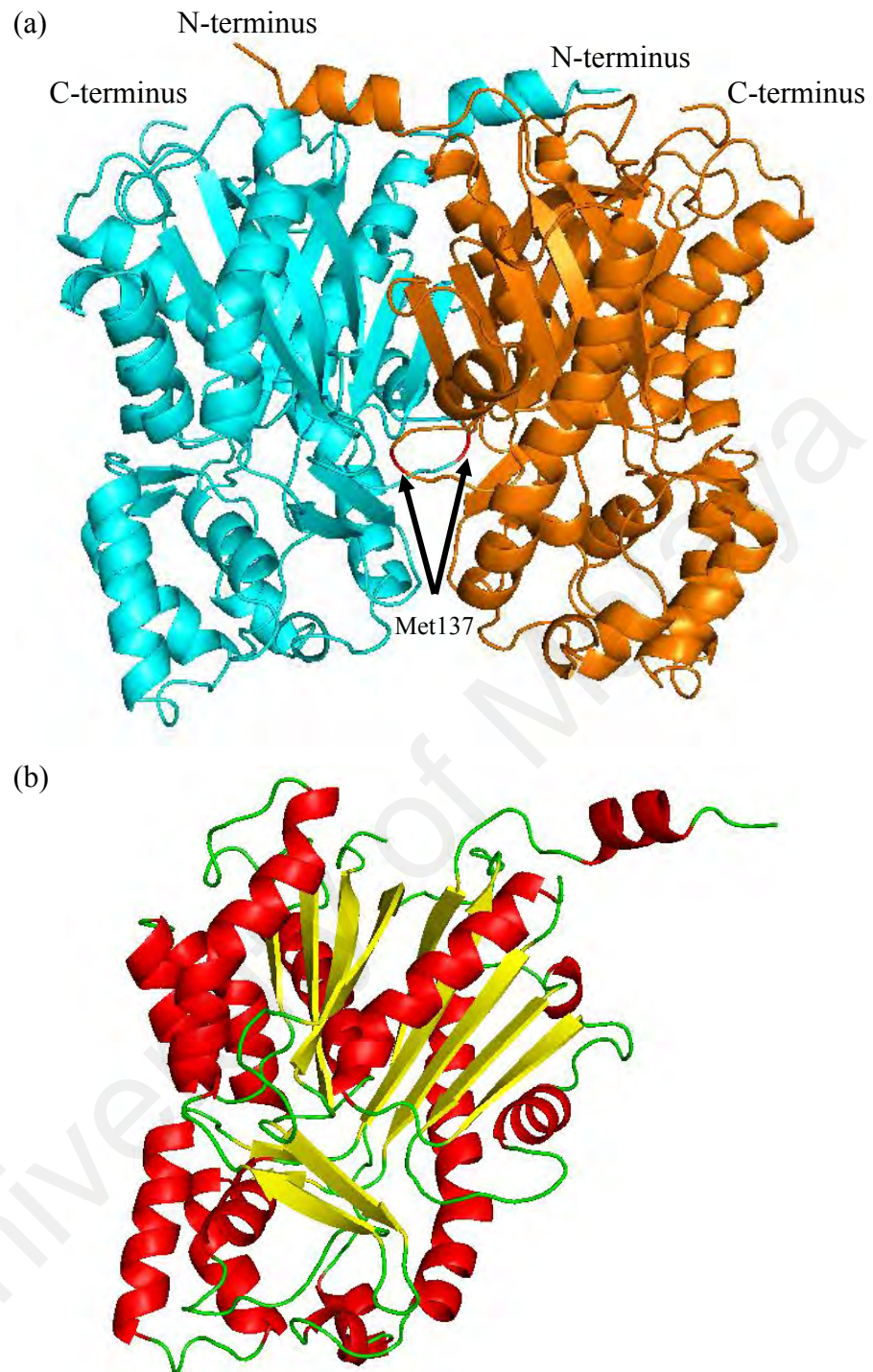


Figure 4.1: Homology model of the *BrCHSv1* receptor. (a) Homodimeric structure with two monomers A (cyan) and B (orange); (b) chain A of the receptor.

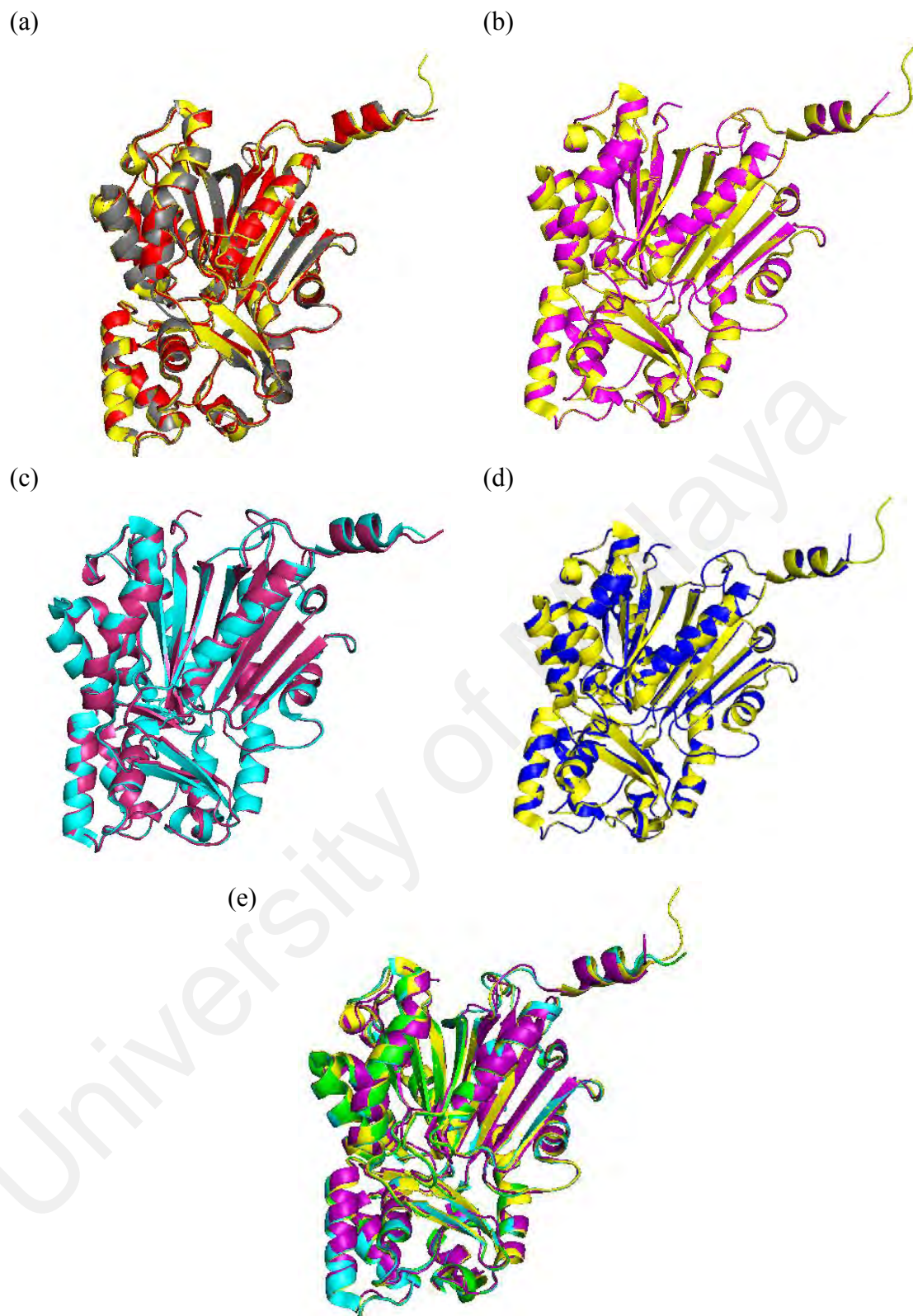


Figure 4.2: Superimposition of homology models of *BrCHS* receptor variants with its respective template models. 4YJY (yellow); 1JWX (cyan); 4WUM (grey); 1I88 (green.); (a) *BrCHSv1* (red); (b) *BrCHSv2* (magenta); (c) *BrCHSv3* (pink); (d) *BrCHSv4* (blue); (e) *BrCHSv5* (purple).

The molecular weights and theoretical isoelectric points (pI) of the *BrCHS* receptor variants are shown in Table 4.2. *BrCHSv3* recorded the highest molecular weight which is 43.27 kDa. Whereas, the molecular weight of *BrCHSv5* is the lowest which is 42.97 kDa. The isoelectric points of the receptors are similar except for *BrCHSv3* which is 6.53.

Table 4.2: Molecular weight and the isoelectric point of *BrCHS* receptor variants.

Variant	Molecular weight (kDa)	Theoretical isoelectric point (pI)
<i>BrCHSv1</i>	43.00	5.92
<i>BrCHSv2</i>	43.01	5.92
<i>BrCHSv3</i>	43.27	6.53
<i>BrCHSv4</i>	42.99	5.92
<i>BrCHSv5</i>	42.97	5.92

The structure analysis via SOPMA revealed the secondary structures of the homology models. Table 4.3 depicts the percentages of the secondary structures of the receptors that mainly consist of α -helix, extended strand, β -turn and random coil. *BrCHSv4* receptor comprises the highest α -helix structures of 45.27%. On the other hand, *BrCHSv2* consist of the highest β -turn and random coil structures which are 7.67% and 34.27%, respectively. The highest percentage of the extended strand was recorded by *BrCHSv5*.

Table 4.3: Composition of secondary structures of *BrCHS* receptor variants.

Variant	Percentage (%)			
	α -helix	Extended strand	β -turn	Random coil
<i>BrCHSv1</i>	43.73	15.60	6.65	34.02
<i>BrCHSv2</i>	42.20	15.86	7.67	34.27
<i>BrCHSv3</i>	41.94	16.62	7.42	34.02
<i>BrCHSv4</i>	45.27	15.35	7.16	32.23
<i>BrCHSv5</i>	44.50	16.88	7.16	31.46

Ramachandran plots of the homology models before and after minimization are shown in Figure 4.3 and 4.4, respectively. All the pre- and post-minimized homology models contain more than 90% of residues in the favored regions. Based on Table 4.4, both pre- and post-minimized homology models contain more than 95% of the residues number in the favored region. Among the pre-minimized homology models, only *BrCHSv5* contains 0.3% residues in the outlier region. In contrast, the post-minimization receptor variants contain residues number with the range of 0.3% to 0.5% in the outlier region except for *BrCHSv3* receptor which contains none. Thus, post-minimized models were chosen for the molecular docking.

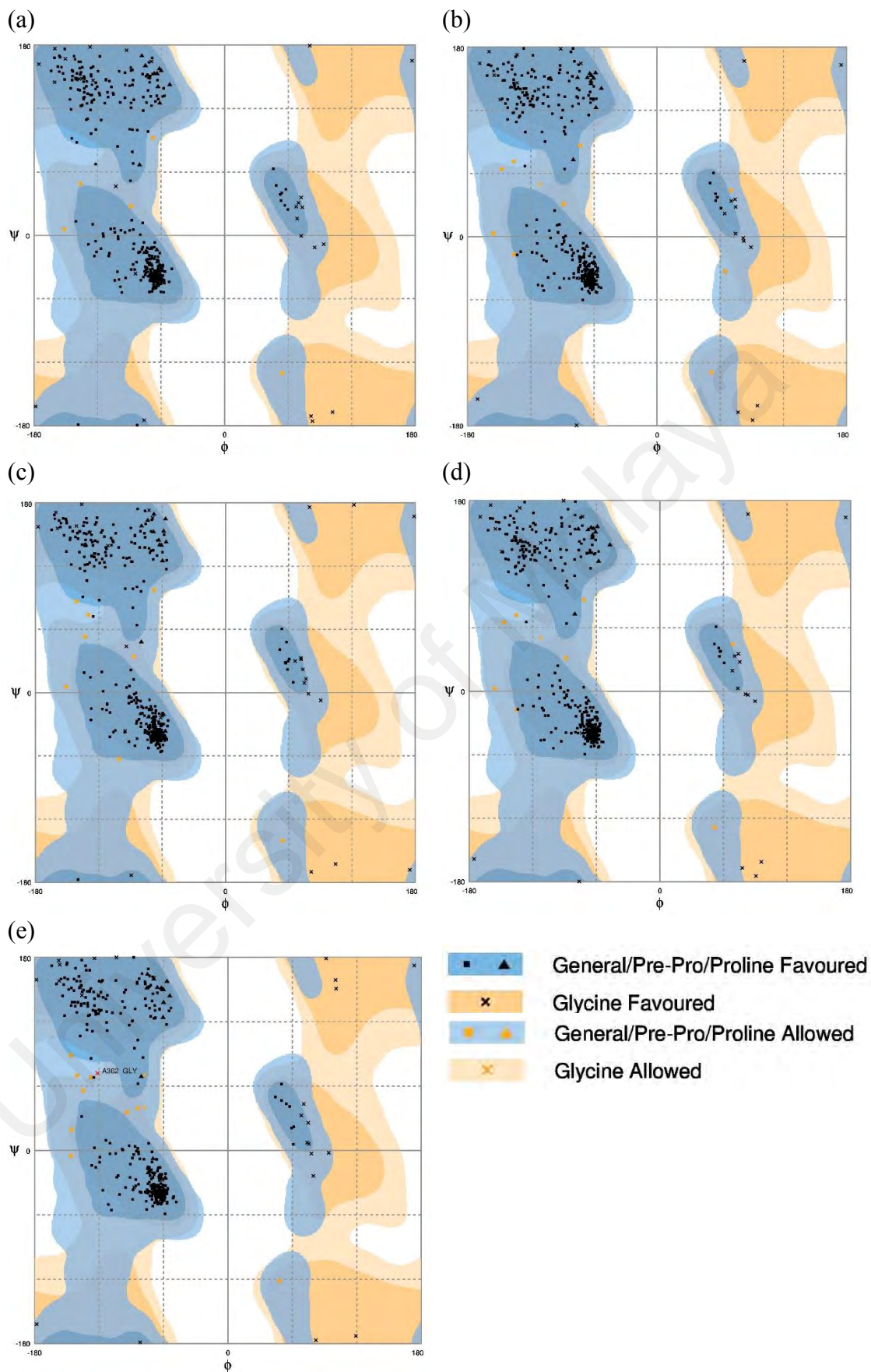


Figure 4.3: Ramachandran plots of *BrCHS* receptor variants after homology modelling. (a) Variant 1; (b) variant 2; (c) variant 3; (d) variant 4; (e) variant 5.

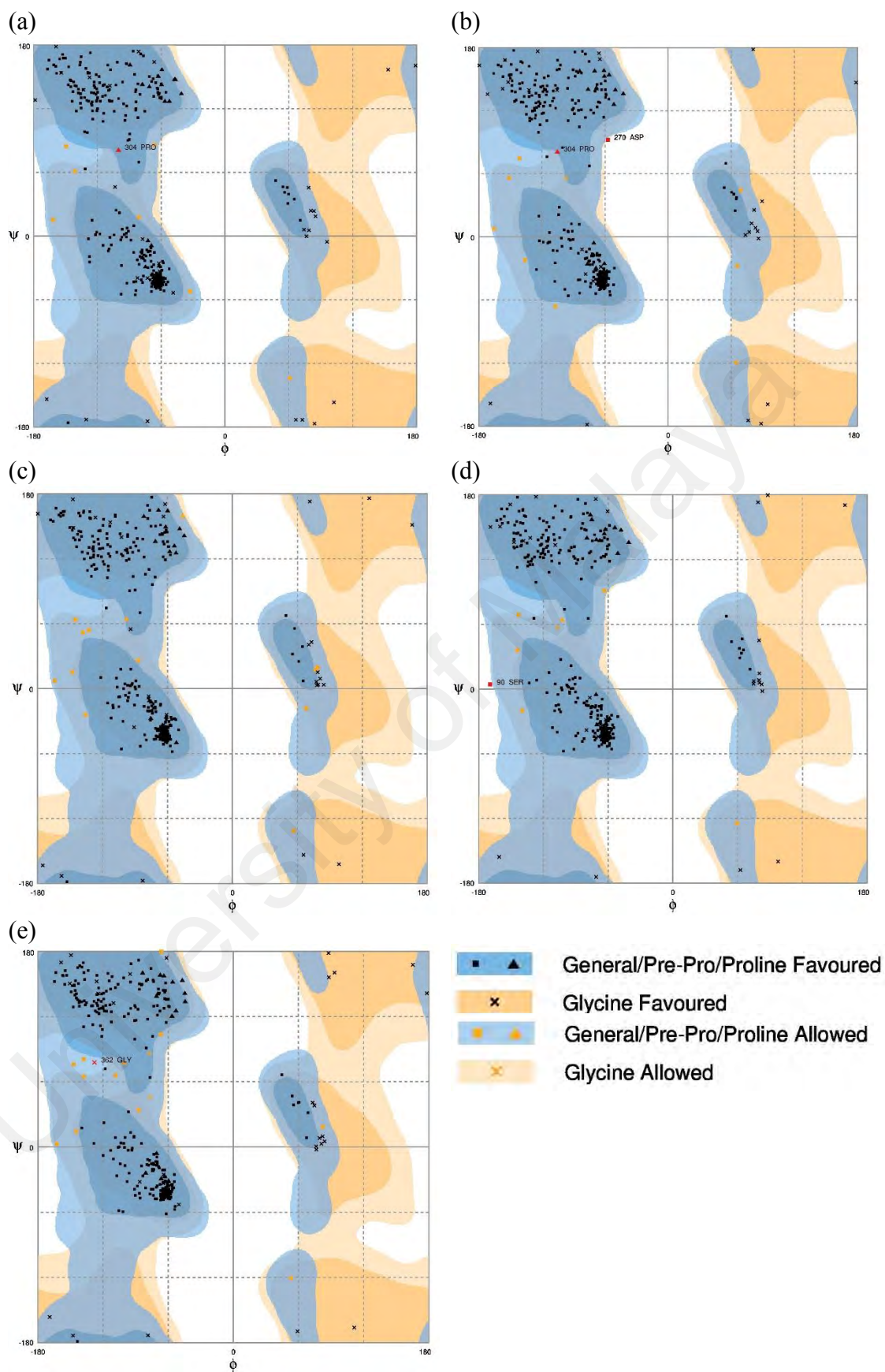


Figure 4.4: Ramachandran plots of *BrCHS* receptor variants for post-minimization. (a) Variant 1; (b) variant 2; (c) variant 3; (d) variant 4; (e) variant 5.

Table 4.4: Ramachandran plot summary from RAMPAGE analysis.

Variant	Percentage of residues number (%)					
	Before minimization			After minimization		
	Favored region	Allowed region	Outlier region	Favored region	Allowed region	Outlier region
<i>BrCHSv1</i>	98.7	1.3	0	97.9	1.8	0.3
<i>BrCHSv2</i>	97.4	2.6	0	97.2	2.3	0.5
<i>BrCHSv3</i>	97.9	2.1	0	96.6	3.4	0
<i>BrCHSv4</i>	97.7	2.3	0	97.9	1.8	0.3
<i>BrCHSv5</i>	96.9	2.8	0.3	96.1	3.6	0.3

4.2 Multiple Sequence Alignment (MSA)

Figure 4.5 shows the multiple sequence alignment result of the *BrCHS* receptor variants with the CHS receptors from six different plant species. All the *BrCHS* receptor variants shared percentage identity in the range of 77% to 89% with the selected CHS amino acid sequences. Three amino acid residues namely Cys164, His303 and Asn336 are highly conserved in the active site of the receptor of CHS enzyme family. MSA revealed that the conserved catalytic triad is maintained in the *BrCHS* receptor variants as well. Besides that, GFGPG loop of the CHS receptors also presents as a highly conserved region. Gatekeeper residues, Phe215 and Phe265 also conserved in all of the receptors. Figure 4.6 shows the *BrCHSv2* receptor in cartoon rendering with residues involved in the catalytic triad, gatekeepers, the CoA binding tunnel, coumaroyl-binding pocket, cyclization pocket and active site geometry.

CLUSTAL O(1.2.4) multiple sequence alignment

```

sp|P30074|CHS2_MEDSA      ----MVSVSEIRKAQRAEGPATILAI GTANPANCVEQSTYPDFYFKITNSEHKTELKEKF 56
sp|Q2R3A1|CHS1_ORYSJ     -MAAAVTVEEVRRQRAEGPATVLAIGTATPANCVYQADYPDYFTRITKSEHMVELKEKF 59
tr|B6T9S4|B6T9S4_MAIZE  MAGATVTVDEVKGRATGPATVLAIGTATPANCVYQADYPDYFTRITKSDHLTDLKEKF 60
tr|G9F7X4|G9F7X4_CURLO  ---MAKLVTEIRKSQRAEGPAAVLAIGTATPPNVVYQADYPDYFTRITRSEHLVELKEKF 57
tr|K9JFE2|K9JFE2_9LILI  ---MTKLVTEIRRSQRAEGPAAVLAIGTANPPNVVYQADYPDYFTRITRSEHLTELKEKF 57
tr|D5KZK0|D5KZK0_MUSAC  ----MAKLAEIRQSQRAEGSATVLAIGTATPVNVLYQADYPDYFTRITKSEHLTELKEKF 56
BrCHS_var3              ----MAKVQEIRLRQRAEGPAAILAI GKATPTNVVYQADYADYFRTKSEHLTELKEKF 56
BrCHS_var5              ----MAKVQEIRLRQRAEGPAAILAI GKATPTDVVYQADYADYFTRITKSEHLTELKEKF 56
BrCHS_var4              ----MAKVQEIRLRQRSEGPAAILAI GKATPTNVVYQADYADYFTRITKSEHLTELKEKF 56
BrCHS_var1              ----MAKVQEIRQRQRAEGPAAILAI GKATPTNVVYQADYADYFTRITKSEHLTELKEKF 56
BrCHS_var2              ----MAKVQEIRQRQRAEGPAAILAI GKATPTNVVYQADYADYFTRITKSEHLTELKEKF 56
                          : *:*  **:* * *::****.*.* : : *:* * **:::*.*.* .:*****

sp|P30074|CHS2_MEDSA      QRMCDKSMIKRRYMYLTHEEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAIKE 116
sp|Q2R3A1|CHS1_ORYSJ     KRMCDKSQIRKRYMHLTHEEILQENPNMCAYMAPSLDARQDIVVVEVPKLGKAAAQKAIKE 119
tr|B6T9S4|B6T9S4_MAIZE  KRMCDKSMIRKRYMHLTHEEFLSENPSMAYMAPSLDARQDVVVEVPKLGKAAAQKAIKE 120
tr|G9F7X4|G9F7X4_CURLO  KRMCDKSMIRKRHMYLTHEEILRENPKMAYMEASLDARQDIVVVEVPRLGKEAAVKAIKE 117
tr|K9JFE2|K9JFE2_9LILI  KRMCDKSMIRKRHMYLTHEEILRENPKMAYMEASLDARQDIVVVEVPRLGKEAAVKAIKE 117
tr|D5KZK0|D5KZK0_MUSAC  KRMCDKSMIHKRYMHINEEILKENPNVCAYMAPSLDARQDIVVVEVPKLGKEAAVKAIKE 116
BrCHS_var3              KRMCDKSMIRKRYMHLTHEEILKENPNMSAYMEPSLDERQDIVVVEVPKLGKEAAKAIKE 116
BrCHS_var5              KRMCDKSMIRKRYMHLTHEEILKENPNMCAYMEPSLDERQDIVVVEVPKLGKEAAKAIKE 116
BrCHS_var4              KRMCDKSMIRKRYMHLTHEEILKENPNMCAYMEPSLDVRQDIVVVEVPKLGKEAAKAIKE 116
BrCHS_var1              KRMCDKSMIRKRYMHVTHEEILKENPNMCAYMEPSLDERQDIVVVEVPKLGKEAAKAIKE 116
BrCHS_var2              KRMCDKSMIRKRYMHVTHEEILKENPNMCAYMEPSLDERQDIVVVEVPKLGKEAAKAIKE 116
                          :***** *::*::::.*:* * **.:. **  *** **:::*.***:* ** *****

```

Figure 4.5: Multiple sequence alignment (MSA) of five variants of *BrCHS* with *Medicago sativa*, *Oryza sativa*, *Zea mays*, *Curcuma longa*, *Curcuma alismatifolia* and *Musa acuminata*. Highly conserved residues were highlighted as follows: catalytic triad (yellow); gatekeepers (black box); GFGPG loop (red box); coumaroyl-binding pocket (green); cyclization pocket (blue); active site geometry (magenta).


```

sp|P30074|CHS2_MEDSA      FEMVWTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPLGISDYN 296
sp|Q2R3A1|CHS1_ORYSJ     FQMVSASQTILPDSEGAIDGHLREVGLTFHLLKDVPLISKNIERALGDAFTPLGISDWN 299
tr|B6T9S4|B6T9S4_MAIZE  FQLVSAAQTILPDSEGAIDGHLREVGLTFHLLKDVPLISKNIERALEDAFEPLGISDWN 300
tr|G9F7X4|G9F7X4_CURLO  FELVSASQTILPDSEGAIDGHLREAGLTFHLLKDVPLISKNIEKSLVEAFKPLGISDWN 297
tr|K9JFE2|K9JFE2_9LILI  FELVSASQTILPDSEGAIDGHLREVGLTFHLLKDVPLISKNIEKSLTEAFKPLGISDWN 297
tr|D5KZK0|D5KZK0_MUSAC  FQLVSASQTILPDSEGAIDGHLREVGLTFHLLKDVPLISKNIERSLAEAFKPLGISDWN 296
BrCHS_var3              FELVSASQTILPDSEGAIVAHLEVLGLTFHLLKNVPVLISKNIEKILVEAFAPLGIDDWN 296
BrCHS_var5              FELVSASQTILPDSEGAIGGHLREVGLTFHLLKGVPLVLISKNIEKILVEAFAPLGIDDWN 296
BrCHS_var4              FELVSASQTILPDSEGAIDGHLREVGLTFHLLKDLPLVLISKNIEKSLVEAFAPLGIDDWN 296
BrCHS_var1              FELVSASQTILPDSEGAIDGHLREVGLTFHLLKDVPLISKNIEKSLVEAFAPLGIDDWN 296
BrCHS_var2              FELVSASQTILPDSEGAIDGHLREVGLTFHLLKDVPLISKNIEKSLVEAFAPLGIDDWN 296
*::*  :::**  *****  .****.*****. :*  :::***  :  *  :**  ***: .*:

```



```

sp|P30074|CHS2_MEDSA      SIFWIAHPGGPAILDQVEQKLALKPEKMNATREVLSEYGNMSSACVLFILDEMRRKSTQN 356
sp|Q2R3A1|CHS1_ORYSJ     SIFWVAHPGGPAILDQVEAKVGLDKERMTRATHVLSYGNMSSACVLFILDEMRRKSAED 359
tr|B6T9S4|B6T9S4_MAIZE  SIFWVAHPGGPAILDQVEARVGLDKARMTRATHVLSYGNMSSACVLFILDEMRRKSAED 360
tr|G9F7X4|G9F7X4_CURLO  SLFWIAHPGGPAILDQVEAKLALDKDKMKATRNVLSYGNMSSACVLFILDEMRRRSAAE 357
tr|K9JFE2|K9JFE2_9LILI  SLFWIAHPGGPAILDQVEAKLALNKDKMKATREVLSEYGNMSSPCVLFILDEMRRRSAAE 357
tr|D5KZK0|D5KZK0_MUSAC  SIFWIAHPGGPAILDQVEAKLGLLEKEKMKATREVLKEYGNMSSACVLFILDEMRRKSAED 356
BrCHS_var3              SIFWIAHPGGPAILDQVEAKLVLEKEKMAATRQVLSEYGNMSSASVMFILDEMRRKSAQE 356
BrCHS_var5              SIFWIAHPGGPAILDQVEAKLLEKEKLAATRQVLSEYGNMSSACVIFILDEMRRKSAQE 356
BrCHS_var4              SIFWIAHPGGPAILDQVEAKLLEKEKMAATRQVLSEYGNMSSACVIFILDEMRRKSAQE 356
BrCHS_var1              SLFWIAHPGGPAILDQVEAKLLEKEKMAATRQVLSEYGNMSSACVIFILDEMRRKSAQE 356
BrCHS_var2              SIFWIAHPGGPAILDQVEAKLLEKEKMAATRQVLSEYGNMSSACVIFILDEMRRKSAQE 356
*::*:*****  *****  ::  *  .  ::  ***.***.*****  .*:*****:::***:

```

Figure 4.5, continued.

sp	P30074	CHS2_MEDSA	GLKTTGEGLEWGVLF	GF	GP	LT	IETVVLRSVAI	-----	389
sp	Q2R3A1	CHS1_ORYSJ	GHATTGEGMDWGVLF	GF	GP	LT	VETVVLHSVPITAGAAA	--	398
tr	B6T9S4	B6T9S4_MAIZE	GQATTGEGLDWGVLF	GF	GP	LT	VETVVLHSVPITGAPTAA		401
tr	G9F7X4	G9F7X4_CURLO	GKATTGEGLEWGVLF	GF	GP	LT	VETVVLHSVPI SAAATH	--	396
tr	K9JFE2	K9JFE2_9LILI	GKAPPGEGVEWGVFF	GF	GP	LT	VETVVLHRVPI SAGATP	--	396
tr	D5KZK0	D5KZK0_MUSAC	GKATTGEGLEWGVLF	GF	GP	LT	VETVVLHSIPIAVH	----	392
		BrCHS_var3	GKATTGEGFNWGVLF	GF	GP	LT	VETVVLHSKPINH	-----	391
		BrCHS_var5	GKATTGEGLNWGVLF	GF	GP	LT	VETVVLHSKPINH	-----	391
		BrCHS_var4	GKATTGEGLNWGVLF	GF	GP	LT	VETVVLHSKPINH	-----	391
		BrCHS_var1	GKTTTGEGLNWGVLF	GF	GP	LT	VETVVLHSKPINH	-----	391
		BrCHS_var2	GKATTGEGLNWGVLF	GF	GP	LT	VETVVLHSKPINH	-----	391

* *** . : *** : ***** : * : ***** : *

Figure 4.5, continued.

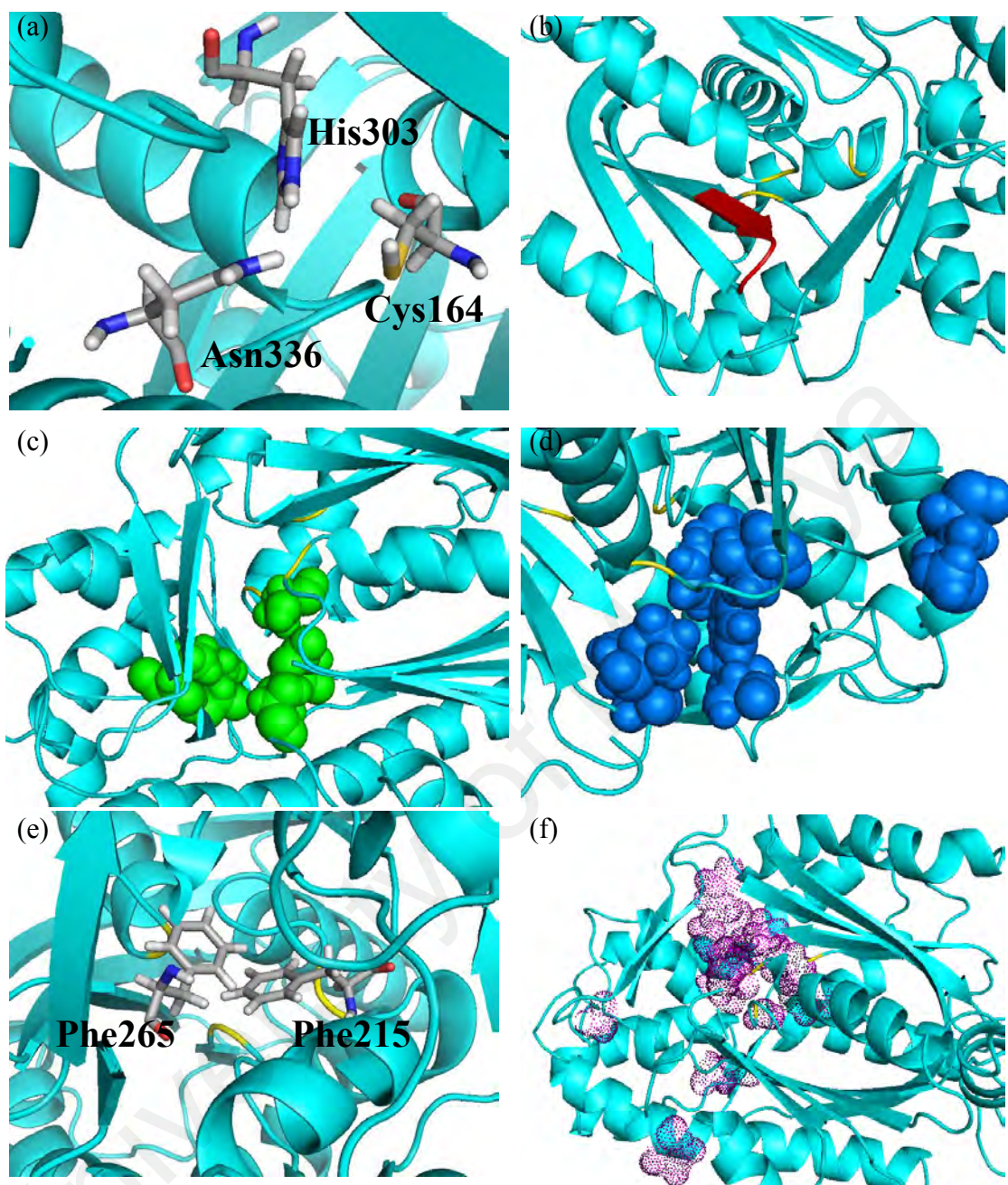


Figure 4.6: Cartoon rendering of *BrCHSv2* receptor. (a) Catalytic triad residues; (b) GFGPG loop (red); (c) coumaroyl-CoA binding pocket (green spheres); (d) cyclisation pocket (blue spheres); (e) gatekeepers; (f) geometry of the active site (magenta dots).

4.3 Molecular Docking Analysis

Table 4.5 shows the docked energies of the *BrCHSv1* receptor with the ligands. Cinnamoyl-CoA recorded more negative docked energy compared to *p*-coumaroyl-CoA. Hence, *BrCHSv1* has higher binding affinity for cinnamoyl-CoA than *p*-coumaroyl-CoA. On the other hand, malonyl-CoA has the highest binding affinity towards the receptor as showed the most negative docked energy.

Table 4.5: Docked energy of *BrCHS* variant 1 receptor with the ligands.

Ligand	Docked energy (kcal/mol)	Electrostatic energy (kcal/mol)	van der Waals energy (kcal/mol)
Malonyl- CoA	-60.9 ± 4.1	-187.3 ± 42.1	-19.8 ± 1.8
Cinnamoyl- CoA	-51.2 ± 1.6	-102.5 ± 21.1	-22.6 ± 2.2
<i>p</i> -Coumaroyl- CoA	-46.9 ± 1.1	-45.6 ± 17.2	-26.3 ± 2.1
Caffeoyl-CoA	-49.1 ± 4.9	-37.6 ± 6.1	-25.5 ± 3.7
Feruloyl-CoA	-59.2 ± 3.7	-151.0 ± 6.5	-31.1 ± 3.9
Acetyl-CoA (Reference)	-68.2 ± 5.0	-188.5 ± 17.2	-18.5 ± 4.6
CoA (Reference)	-79.8 ± 5.3	-264.0 ± 18.8	-20.8 ± 2.5

Table 4.6 shows the residues of *BrCHSv1* formed interactions with the ligands. Cys164 of the receptor formed π -alkyl and π -sulfur interactions with cinnamoyl-CoA and *p*-coumaroyl-CoA, respectively. Lys62 of the receptor formed a salt bridge with cinnamoyl-CoA and malonyl-CoA, respectively. Figure 4.7 depicts the interaction of the ligands with the *BrCHSv1* receptor. Malonyl-CoA formed hydrogen bonds with the catalytic triad of *BrCHSv1* as shown in Figure 4.7 (b). Cinnamoyl-CoA formed only one hydrogen bond, whereas *p*-coumaroyl-CoA formed three hydrogen bonds with the *BrCHSv1* receptor.

Table 4.6: List of residues of *BrCHSv1* formed interactions with the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Lys62 (2.63, 2.83) Cys164 (3.20) His303 (2.68) Asn336 (2.60)	Val210, Leu214, Ile254, Gly305, Gly306, Pro307, Ala308, Ile309, Phe373, Gly374	Arg58 (π - Cation)	Lys62	-
Cinnamoyl-CoA	Ala308 (2.68)	Leu214, Phe215, Gly305, Gly306, Pro307, Ile309, Asn336, Gly374, Pro375	Lys62 (Alkyl) Cys164, Ile254, Phe265 (π - Alkyl)	Lys62	-
<i>p</i>-Coumaroyl-CoA	Tyr160 (3.23) Gln161 (3.34) Gly305 (2.98)	Thr132, Gly163, Leu214, Ile254, Asp255, Gly256, Phe265, His303, Asn336, Gly374, Pro375	Cys164 (π - Sulfur) Phe215 (π -Pi T- shaped)	-	-
Caffeoyl-CoA	Gln161 (2.78) Gly305 (2.86)	Gly163, Leu214, Ile254, Asp255, Gly256, Leu267, Asn336, Pro375	Cys164, Phe215 (π -Alkyl) Phe265 (π - sulfur)	-	-
Feruloyl-CoA	Arg156 (2.61, 2.64) Gly256 (2.96) Leu258 (2.85) Arg259 (2.68)	Thr132, Val135, Met137, Pro138, Tyr142, Met158, Tyr160, Gln161, Gly163, His257, Phe265	Leu263, Pro375 (π -Alkyl)	Arg259	-
Acetyl-CoA (Reference)	Arg58 (2.79)	Lys62, Leu214, Phe215, Leu267, Pro272, Gly305, Gly306, Ala308, Ile309, Asn336	-	Arg58 Lys62	-
CoA (Reference)	Arg58 (2.64, 2.91) Lys62 (2.61) Cys164 (3.22) His303 (3.08)	Leu214, Phe215, Ile254, Leu267, Val271, Pro272, Gly305, Ala308, Ile309, Gly374	Val210 (Alkyl)	Arg58 Lys62 His303	Arg58 Asn336

- : Absence

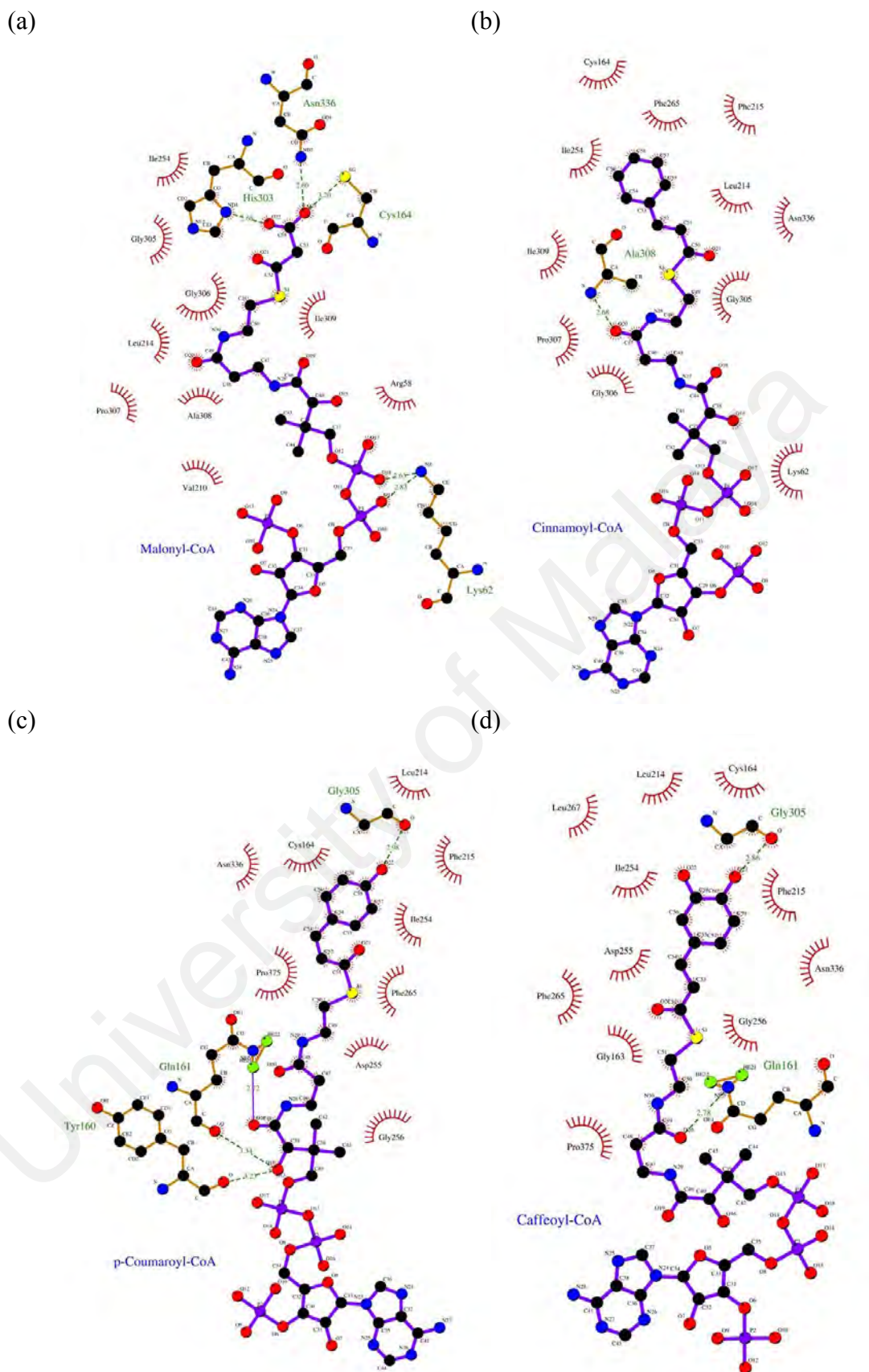
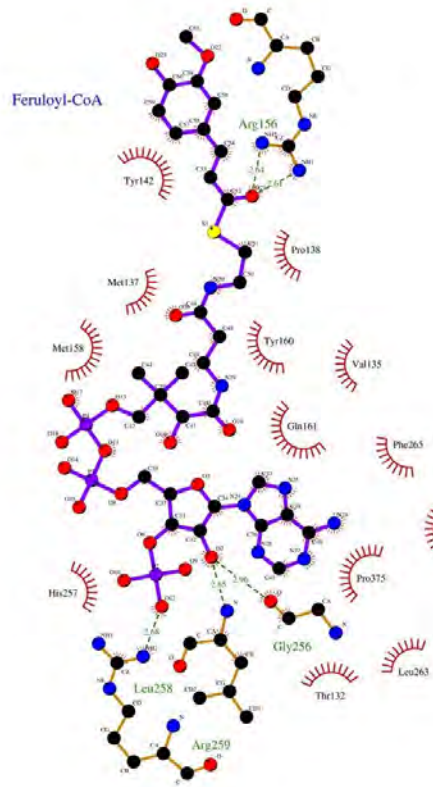


Figure 4.7: Docked conformations of *BrCHSvI* with the ligands. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

(e)



(f)



(g)

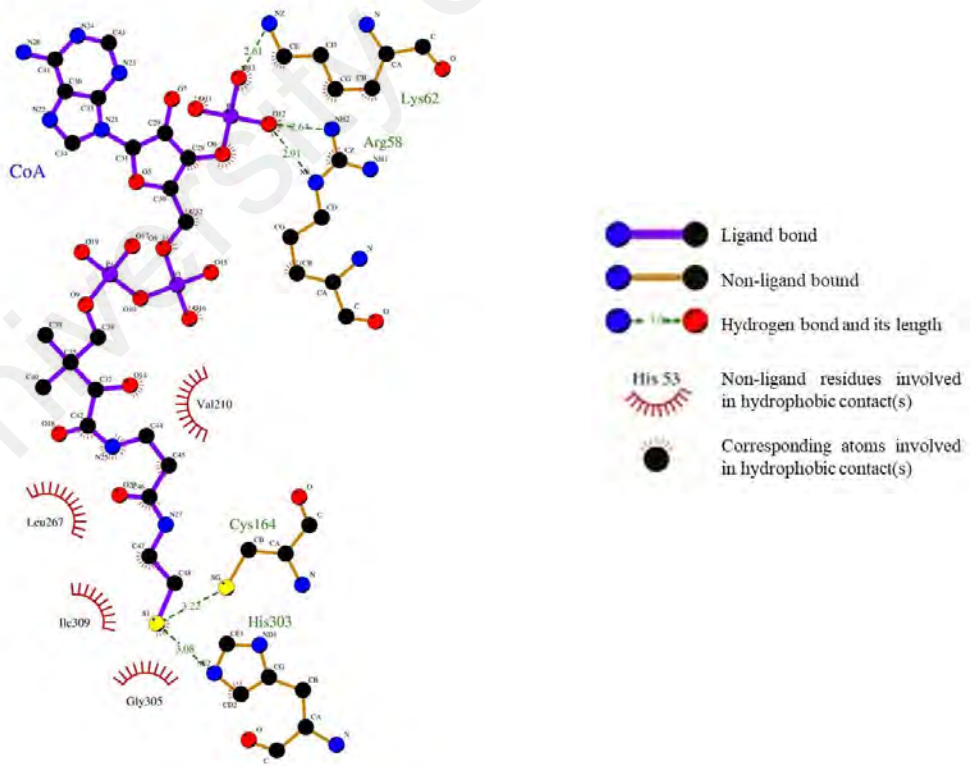


Figure 4.7, continued.

Docked energies of the *BrCHSv2* receptor with the ligands are shown in Table 4.7. Malonyl-CoA recorded the most negative docked energy among the substrate ligands and followed by cinnamoyl-CoA. *p*-Coumaroyl-CoA recorded the least negative docked energy among the substrate ligands. Thus, *p*-coumaroyl-CoA has the least binding affinity towards the *BrCHSv2* receptor compared to other ligands.

Table 4.7: Docked energy of *BrCHS* variant 2 receptor with the ligands.

Ligand	Docked energy (kcal/mol)	Electrostatic energy (kcal/mol)	van der Waals energy (kcal/mol)
Malonyl- CoA	-62.7 ± 3.4	-223.9 ± 9.0	-18.0 ± 5.2
Cinnamoyl- CoA	-54.0 ± 3.0	-64.3 ± 14.7	-24.7 ± 5.8
<i>p</i>-Coumaroyl- CoA	-49.3 ± 4.3	-33.5 ± 14.0	-27.3 ± 1.8
Caffeoyl-CoA	-51.2 ± 2.0	-82.1 ± 20.6	-23.8 ± 3.9
Feruloyl-CoA	-48.0 ± 1.4	-85.0 ± 28.7	-34.6 ± 2.4
Acetyl-CoA (Reference)	-63.5 ± 3.6	-212.2 ± 2.0	-18.1 ± 6.8
CoA (Reference)	-83.7 ± 5.8	-289.1 ± 4.1	-21.0 ± 1.3

Table 4.8 shows the type of interactions formed between *BrCHSv2* and the ligands. Gln161 of the receptor involved in the repulsive interaction with *p*-coumaroyl-CoA. Residues Lys62 and Arg259 formed a salt bridge with malonyl-CoA and cinnamoyl-CoA, respectively. Hydrogen bonds and non-hydrophobic interactions of the docked *BrCHSv2* complexes are depicted in Figure 4.8. Malonyl-CoA and *p*-coumaroyl-CoA, respectively formed hydrogen bonds with the Cys164 residue of the receptor. In addition, malonyl-CoA formed hydrogen bonds with other catalytic triad residues which are His303 and Asn336. Cinnamoyl-CoA formed more hydrogen bonds with *BrCHSv2* receptor than *p*-coumaroyl-CoA.

Table 4.8: List of residues of *BrCHSv2* formed interactions with the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Lys62 (2.44, 2.56) Cys164 (3.16) His303 (3.06) Asn336 (2.80)	Val210, Leu214, Phe215, Ile254, Gly305, Gly306, Pro307, Ala308, Ile309, Gly374	Lys62 (Alkyl)	Lys62	-
Cinnamoyl-CoA	Gln161 (2.63) Leu258 (2.62) Arg259 (2.73)	Thr132, Gly163, Phe215, Gly256, His257, Phe265, Gly305, Asn336, Pro375	Cys164, Ile254, His303 (π - Alkyl)	Arg259	-
<i>p</i>-Coumaroyl-CoA	Cys164 (3.13)	Met137, Val210, Leu214, Ile254, Asp255, Gly256, Phe265, Gly305, Asn336, Pro375	Leu258 (Alkyl) Phe215 (π - π Stacked) Cys164 (π - sulfur)	-	Gln161
Caffeoyl-CoA	Arg58 (2.76, 3.06) Cys164 (3.33)	Val210, Leu214, Phe215, Ile254, Val271, His303, Gly305, Gly306, Pro307, Ile309, Asn336	Val212 (Alkyl)	Arg58 Lys62	-
Feruloyl-CoA	Gln161 (2.62, 2.72) Leu258 (2.64) Thr378 (2.82)	Cys164, Phe165, Phe215, Ile246, Ile254, Asp255, Gly256, His257, Leu263, Ser338	Phe265 (π - π Stacked) Pro375 (π -Alkyl, π -Sigma)	Arg259	-
Acetyl-CoA (Reference)	Lys55 (2.62) Lys62 (3.26) Cys164 (3.13) His303 (2.76)	Met59, Leu206, Val210, Leu214, Phe215, Ile254, Val271, Gly306, Pro307, Ile309, Asn336, Gly374	Lys62, Ala308 (Alkyl) Lys55 (π -Alkyl, π -cation)	Lys55	Ala308
CoA (Reference)	Arg58 (2.55, 2.73, 3.35) Lys62 (2.59, 2.62) Cys164 (3.13) Asn336 (3.28)	Phe215, Ile254, Pro272, His303, Gly305, Gly306, Ala308, Ile309	Val210 (Alkyl) Lys55 (π -Alkyl)	Arg58 Lys62	-

-: Absence

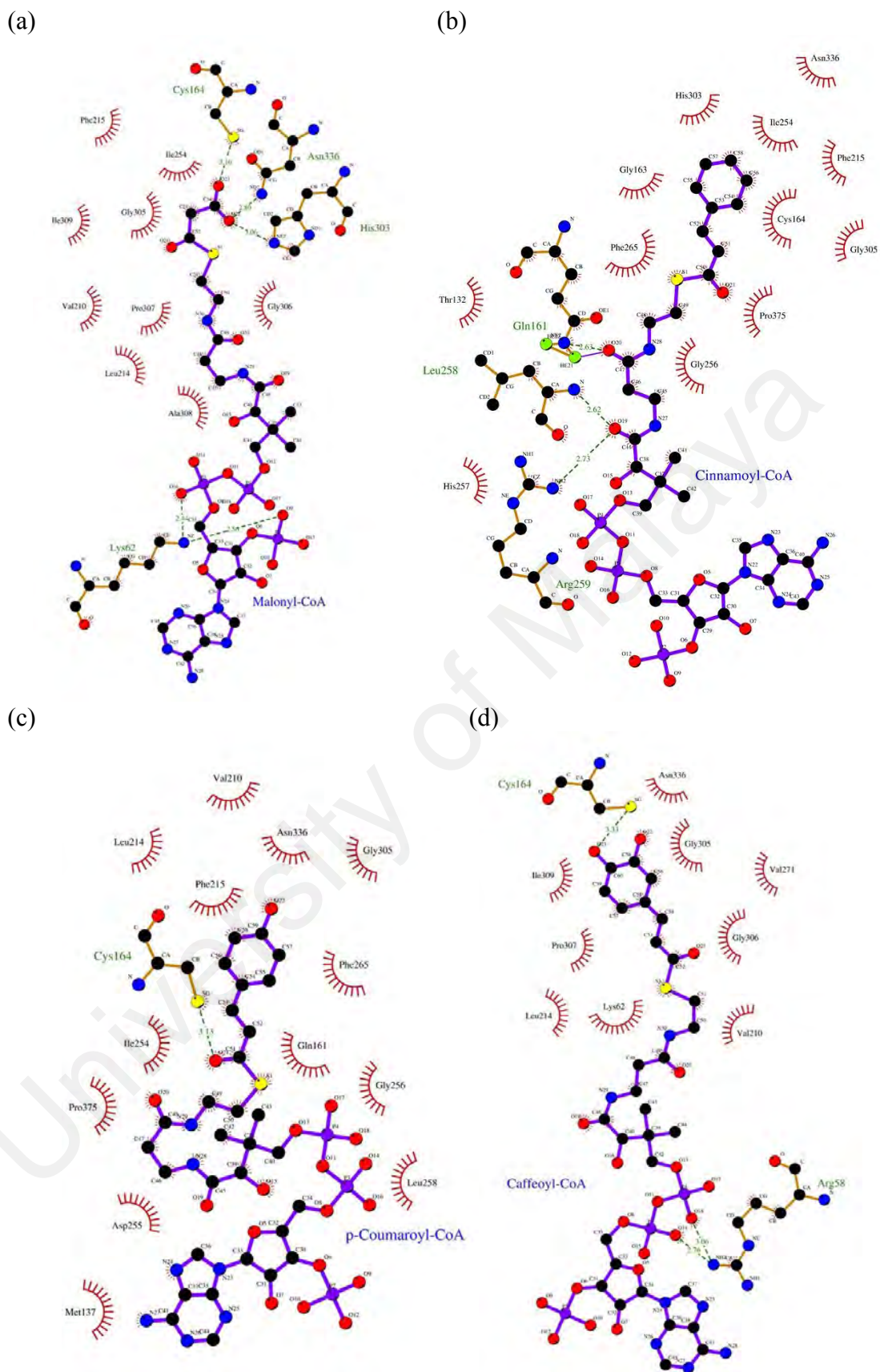
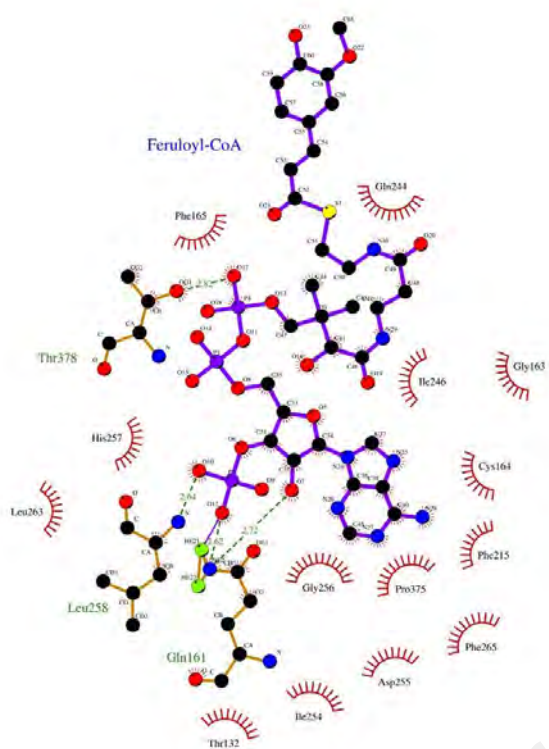
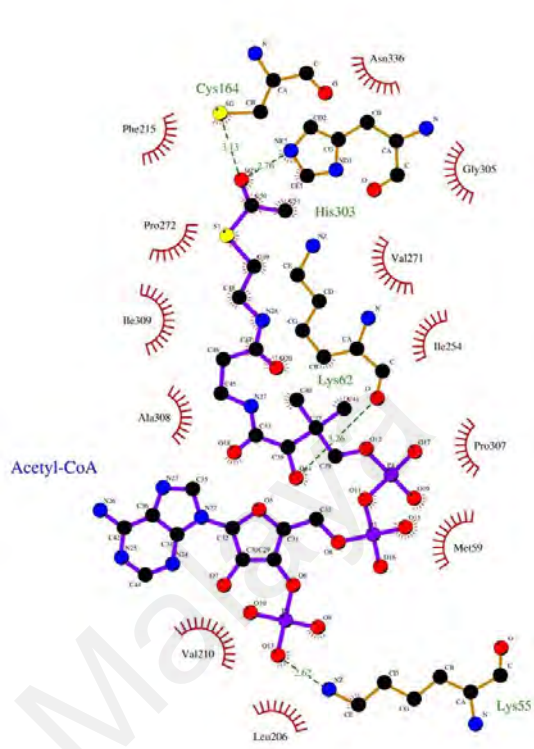


Figure 4.8: Docked conformations of *BrCHSv2* with the ligands. (a) Malonyl-CoA; (b) Cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

(e)



(f)



(g)

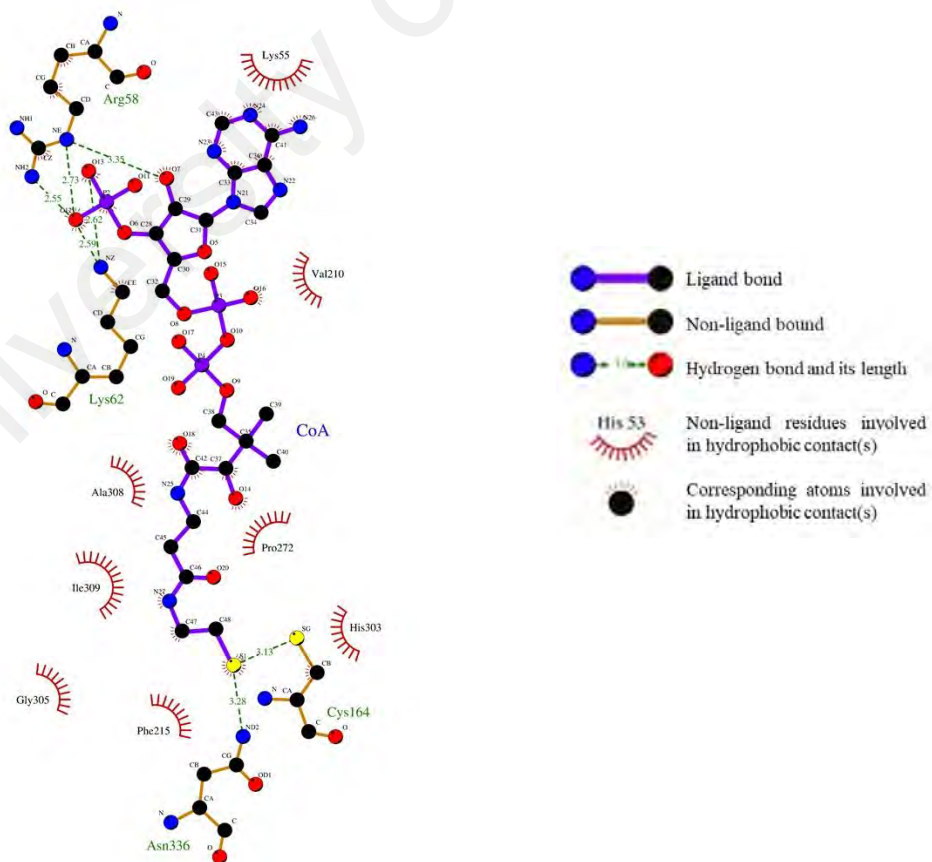


Figure 4.8, continued.

Table 4.9 depicts the docked energies of the *BrCHSv3* receptor with the ligands. Malonyl-CoA has the most negative docked energy among the substrate ligands. Cinnamoyl-CoA recorded as the second most negative docked energy. *p*-Coumaroyl-CoA recorded the least negative docked energy. Hence, *p*-Coumaroyl-CoA has lower binding affinity for *BrCHSv2* receptor than cinnamoyl-CoA.

Table 4.9: Docked energy of *BrCHS* variant 3 receptor with the ligands.

Ligand	Docked energy (kcal/mol)	Electrostatic energy (kcal/mol)	van der Waals energy (kcal/mol)
Malonyl- CoA	-84.1 ± 6.7	-280.4 ± 59.1	-17.9 ± 2.0
Cinnamoyl- CoA	-61.2 ± 2.6	-154.0 ± 27.0	-16.7 ± 2.2
<i>p</i>-Coumaroyl- CoA	-52.8 ± 6.2	-63.6 ± 42.7	-22.0 ± 4.4
Caffeoyl-CoA	-56.6 ± 4.7	-92.7 ± 8.5	-23.8 ± 3.8
Feruloyl-CoA	-57.1 ± 5.7	-149.2 ± 44.3	-28.8 ± 5.0
Acetyl-CoA (Reference)	-83.3 ± 3.3	-301.9 ± 17.9	-13.9 ± 2.9
CoA (Reference)	-92.4 ± 2.8	-305.7 ± 17.6	-18.8 ± 5.4

Residues of *BrCHSv3* involved in the interactions with the ligands are as shown in Table 4.10. Residue Ile254 of the receptor formed π -alkyl interaction with cinnamoyl-CoA and *p*-coumaroyl-CoA, respectively. Salt bridge formed in between Arg58 of the receptor and the ligand (cinnamoyl-CoA, malonyl-CoA, acetyl-CoA, CoA and *p*-coumaroyl-CoA). Figure 4.9 shows the interactions formed between ligands and *BrCHSv3* receptor. Five hydrogen bonds were found in the malonyl-CoA-*BrCHSv3* docked complex. The Arg58 residue of the *BrCHSv3* formed three hydrogen bonds with cinnamoyl-CoA. Whereas, *p*-coumaroyl-CoA formed two hydrogen bonds with the receptor.

Table 4.10: List of residues of *BrCHSv3* formed interactions with the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Lys62 (2.47, 2.81) Cys164 (3.02) His303 (2.78) Asn336 (2.81)	Leu214, Phe215, Ile254, Val271, Pro272, Gly305, Gly306, Pro307, Ala308, Ile309, Gly374	-	Arg58 Lys62	Asn270
Cinnamoyl-CoA	Arg58 (2.64, 2.70, 2.74)	Cys164, Val210, Leu214, Phe215, Gly305, Gly306, Ile309, Phe373, Gly374	Ile254, Leu271, His303 (π - Alkyl)	Arg58	-
<i>p</i>-Coumaroyl-CoA	Lys55 (2.87) Cys164 (3.13)	Met59, Leu206, Asp207, Val210, Leu214, Phe215, Leu267, Pro272, His303, Gly305, Gly306, Ile309, Asn336, Gly374	Cys164, Ile254 (π -Alkyl)	Lys55 Arg58	-
Caffeoyl-CoA	Lys62 (2.77) Arg194 (2.91) Asn270 (3.10)	Thr132, Gly163, Leu214, Phe215, Ile254, Pro272, Gly305, Gly306, Pro307, Asn336, Ser338, Gly374	Ala308 (Alkyl) Cys164, Phe265, Pro375 (π - Alkyl)	Lys62	-
Feruloyl-CoA	Leu258 (3.13) Arg259 (2.54)	Thr132, Val135, Met137, Gly163, Cys164, Arg194, Phe265	Ala256, Gln161 (π -Sigma) Ala256, Leu263, Pro375 (π - Alkyl)	Arg259	-
Acetyl-CoA (Reference)	Arg58 (2.59, 2.65) Lys269 (2.86) Val271 (2.77)	Cys164, Phe215, Ile254, Phe265, Asn270, Pro272, Gly305, Gly306, Ile309, Asn336, Ser338	Leu267 (alkyl), His303 (π - sulfur)	Arg58 Lys62	-
CoA (Reference)	Arg58 (2.69, 2.78) Lys62 (2.62) Cys164 (3.23) Ala308 (3.09)	Met59, Leu214, Phe215, Ile254, Leu267, Val271, Pro272, Gly305, Gly306, Pro307, Ile309, Asn336	His303 (π - Sulfur) Val210 (Alkyl)	Arg58 Lys62	-

-: Absence

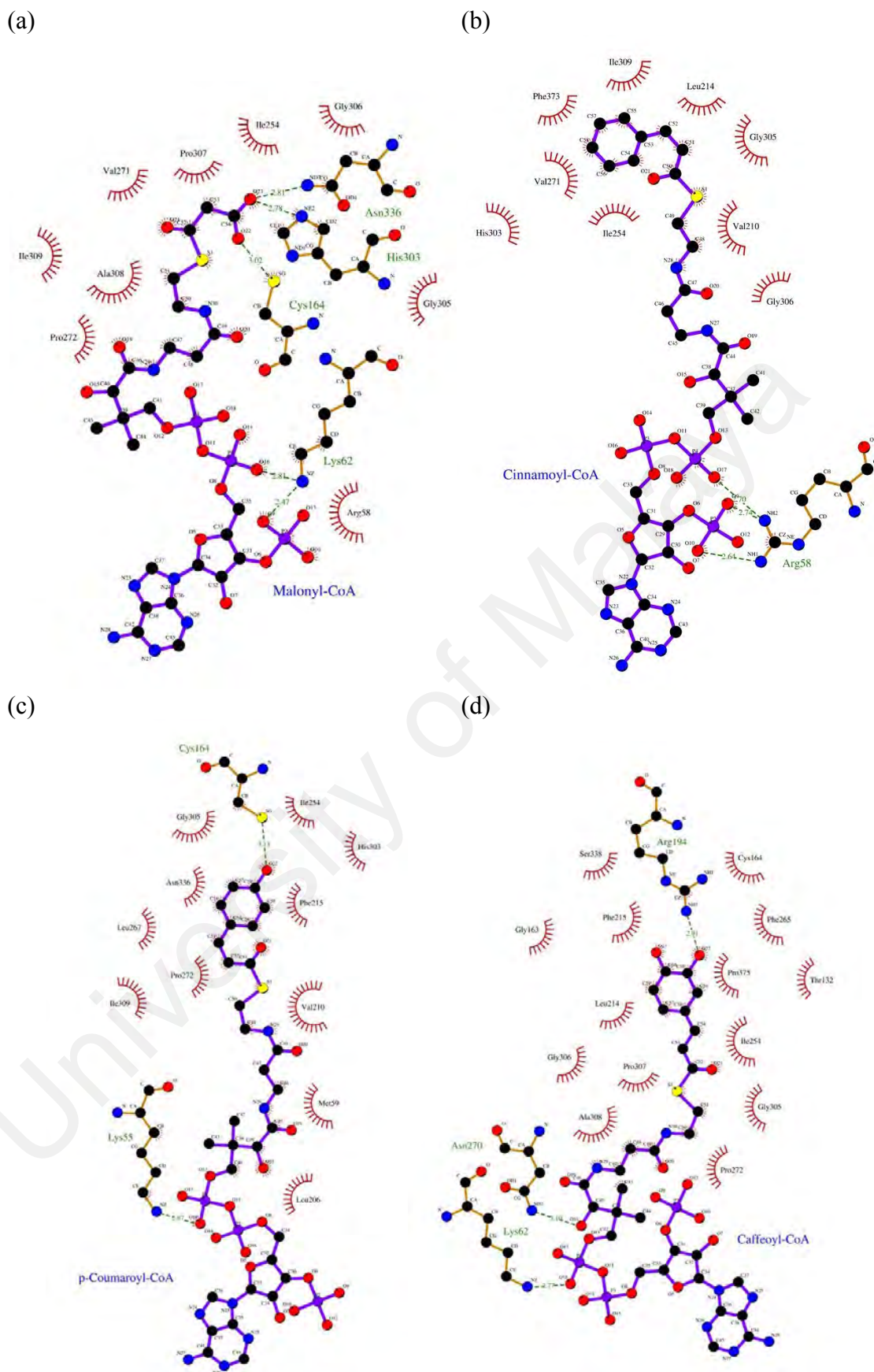
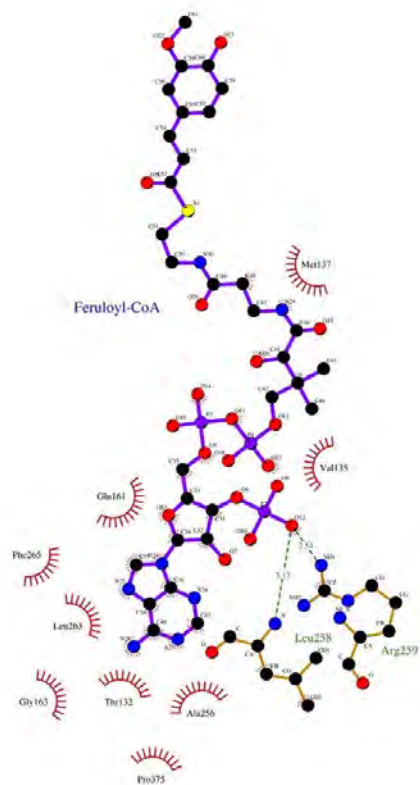
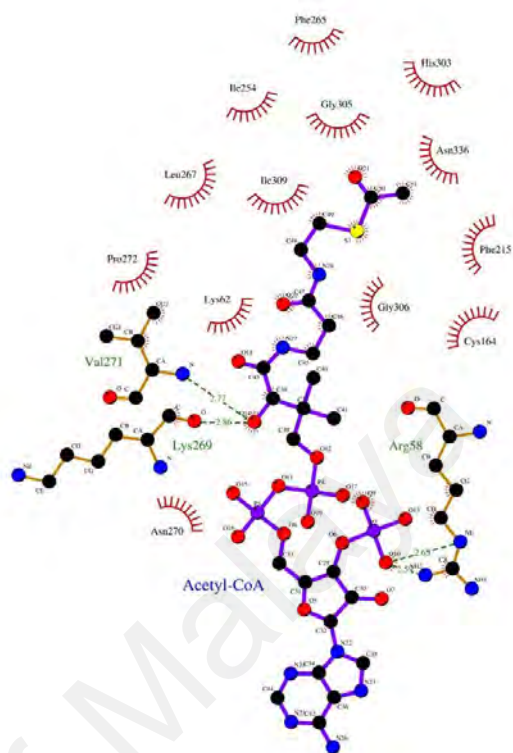


Figure 4.9: Docked conformations of *BrCHSv3* with the ligands. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

(e)



(f)



(g)

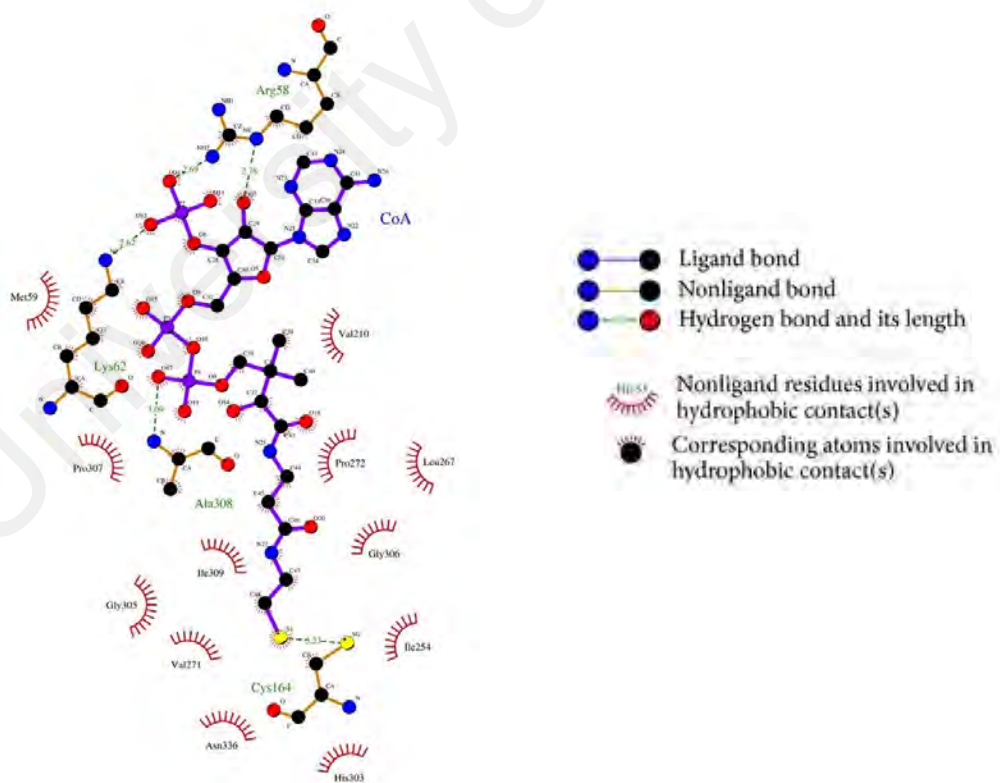


Figure 4.9, continued.

Docked energy profile of the docked *BrCHSv4* complexes is shown in Table 4.11. Malonyl-CoA recorded the most negative docked energy which is -91.5 ± 4.1 kcal/mol and has the highest binding affinity for *BrCHSv4* receptor. *p*-Coumaroyl-CoA showed more negative docked energy compared to cinnamoyl-CoA. Therefore, *p*-coumaroyl-CoA has higher binding affinity for *BrCHSv4* receptor than cinnamoyl-CoA.

Table 4.11: Docked energy of *BrCHS* variant 4 receptor with the ligands.

Ligand	Docked energy (kcal/mol)	Electrostatic energy (kcal/mol)	van der Waals energy (kcal/mol)
Malonyl- CoA	-91.5 ± 4.1	-333.4 ± 40.0	-19.1 ± 4.7
Cinnamoyl- CoA	-61.8 ± 2.6	-104.0 ± 42.5	-24.7 ± 1.6
<i>p</i>-Coumaroyl- CoA	-65.0 ± 7.9	-89.2 ± 30.2	-23.8 ± 2.3
Caffeoyl-CoA	-58.8 ± 2.9	-76.0 ± 38.4	-28.6 ± 1.2
Feruloyl-CoA	-50.5 ± 6.7	-146.3 ± 54.4	-27.4 ± 4.5
Acetyl-CoA (Reference)	-74.2 ± 1.9	-194.4 ± 53.1	-25.3 ± 8.5
CoA (Reference)	-76.6 ± 3.3	-265.2 ± 27.2	-21.1 ± 2.4

Table 4.12 shows the interactions formed between *BrCHSv4* and the ligands. The Cys164 residue of the receptor formed π -alkyl interactions with cinnamoyl-CoA and *p*-coumaroyl-CoA, respectively. π - π stacked interaction was seen between *p*-Coumaroyl-CoA and the receptor. Hydrogen bonding interactions formed by substrate ligands with the *BrCHSv4* receptor are shown in Figure 4.10. *p*-Coumaroyl-CoA formed only one hydrogen bond with the Cys164 residue of the *BrCHSv4* receptor. On the other hand, a hydrogen bond interaction was seen between the Lys62 residue of the receptor and cinnamoyl-CoA. Malonyl-CoA formed two hydrogen bonds with residues in the active site of the receptor which are His303 and Asn336.

Table 4.12: List of residues of *BrCHSv4* formed interactions with the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	His303 (2.68) Asn336 (2.67)	Cys164, Leu214, Phe215, Ile254, Phe265, Pro272, Gly305, Gly306, Pro307, Ala308, Ile309, Gly374	-	Arg58 Lys62	-
Cinnamoyl-CoA	Lys62 (2.75)	Val210, Leu214, Ile254, Leu271, Gly305, Gly306, Pro307, Asn336	Cys164, Phe215, Phe265, Pro375 (π -Alkyl) Met59 (Alkyl) Gly305 (Sulfur- X)	Arg58 Lys62	-
<i>p</i>-Coumaroyl-CoA	Cys164 (3.12)	Gly163, Phe165, Val210, Leu214, Phe215, Ile254, Leu271, Pro272, Gly305, Gly306, Ile309, Gly374	Phe265 (π - π Stacked) Cys164, Pro375 (π -Alkyl) Lys62 (Alkyl)	Arg58 Lys62	-
Caffeoyl-CoA	Arg58 (2.75) Lys62 (2.90) His303 (3.24) Asn336 (2.85)	Cys164, Val210, Leu214, Phe215, Leu267, Gly305, Gly306, Pro307, Gly374	Phe265, Ile254, Leu271 (π - Alkyl)	Arg58 Lys62	-
Feruloyl-CoA	Arg156 (2.54, 3.17) Met158 (3.22) Gly256 (2.74) Arg259 (2.48, 2.74)	Thr132, Pro152, Tyr160, Gly163, Cys164, His257, Leu258	Gln161 (π - Sigma) Leu263, Pro375 (π -Alkyl)	Arg259	-
Acetyl-CoA (Reference)	Lys55 (2.63) Arg58 (2.76, 2.90) His303 (2.73)	Met59, Lys62, Cys164, Leu206, Leu214, Phe215, Leu271, Gly305, Gly306, Pro307, Ile309, Asn336	Arg58 (π -Alkyl) Ala308 (Alkyl) Phe265 (π - Sulfur)	Lys55	-
CoA (Reference)	Arg58 (2.73) Lys62 (2.66) Asn336 (2.95)	Lys55, Met59, Cys164, Gly211, Leu214, Phe215, Phe265, Leu267, Gly305	His303 (π - Sulfur) Val210 (Alkyl)	Arg58 Lys62	-

-: Absence

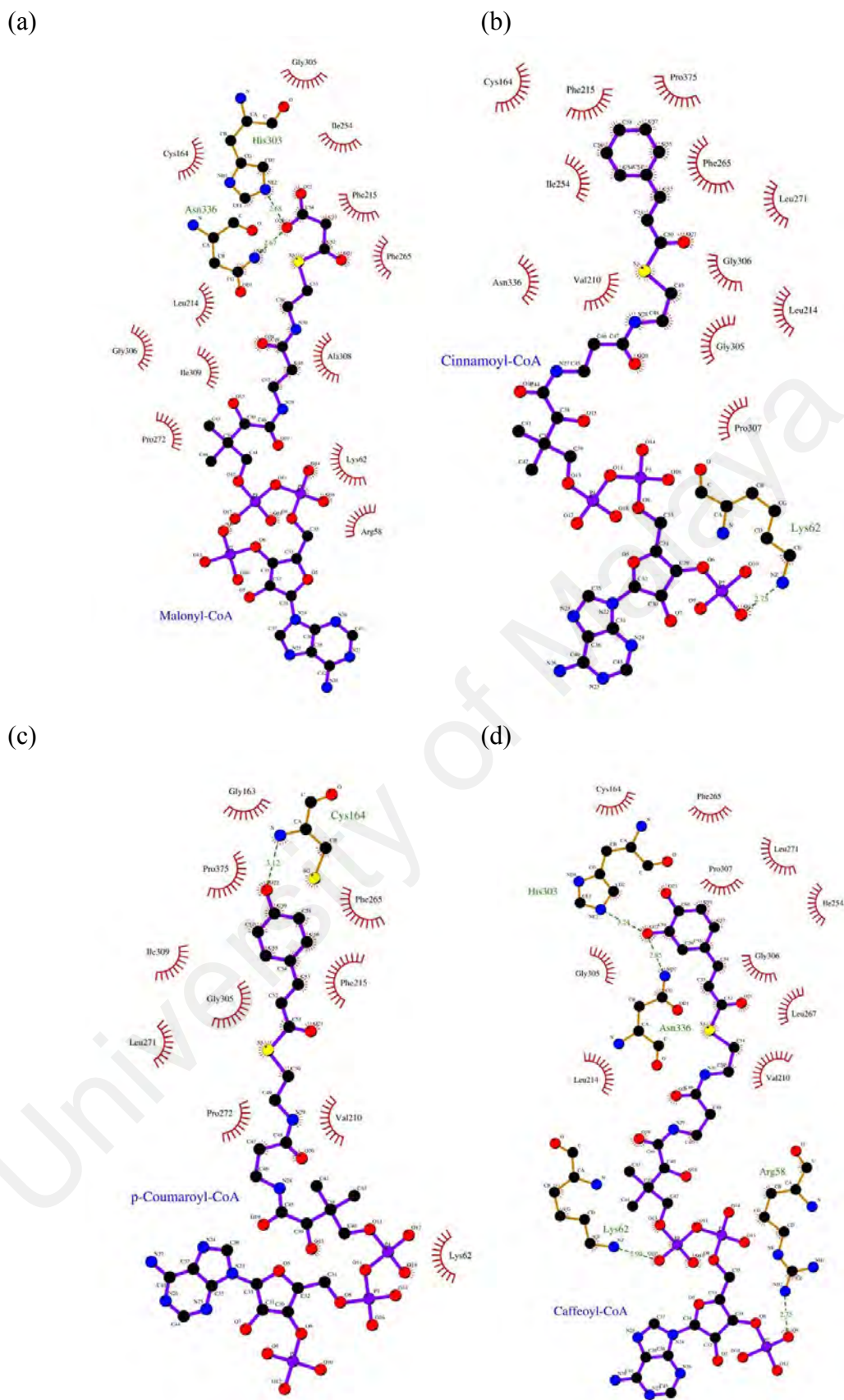
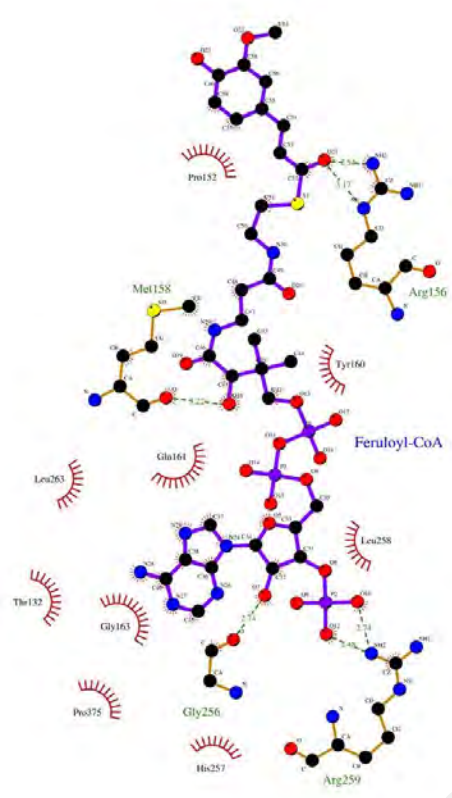
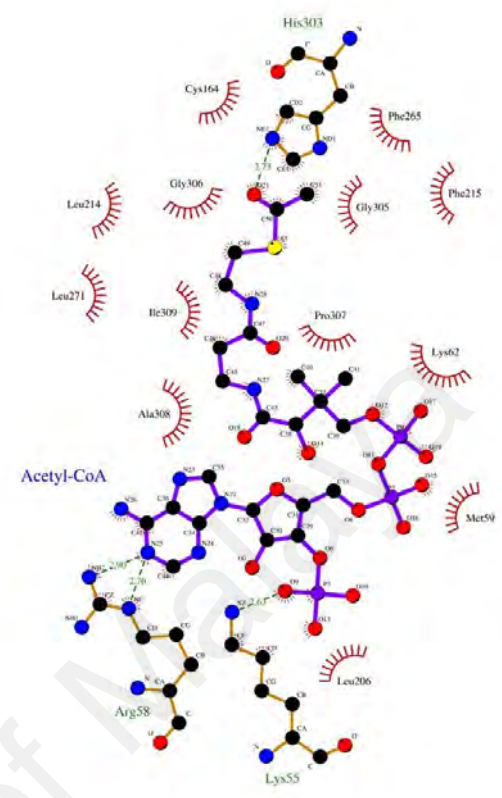


Figure 4.10: Docked conformations of *BrCHSv4* with the ligands. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

(e)



(f)



(g)

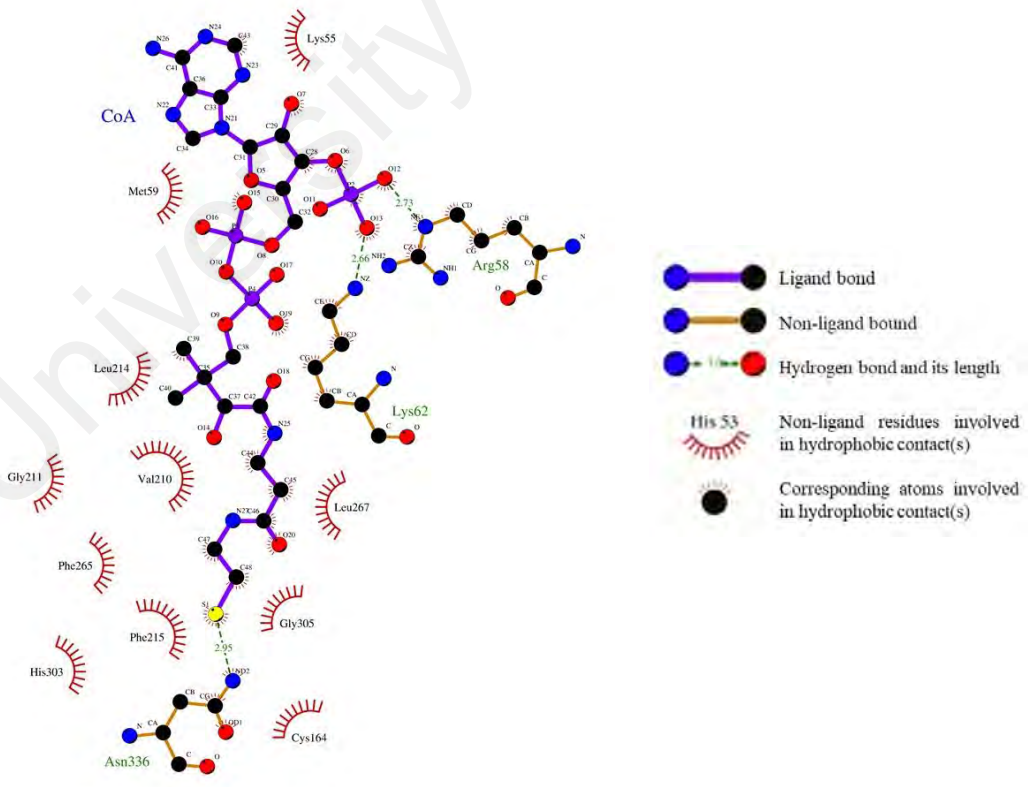


Figure 4.10, continued.

Table 4.13 shows the docked energies of the *BrCHSv5* receptor with the ligands. The most negative docked energy was recorded by malonyl-CoA. This shows that malonyl-CoA has the highest binding affinity towards *BrCHSv4* receptor. *p*-Coumaroyl-CoA recorded more negative docked energy compared to cinnamoyl-CoA. Thus, cinnamoyl-CoA has lower binding affinity for *BrCHSv5* receptor than *p*-coumaroyl-CoA.

Table 4.13: Docked energy of *BrCHS* variant 5 receptor with the ligands.

Ligand	Docked energy (kcal/mol)	Electrostatic energy (kcal/mol)	van der Waals energy (kcal/mol)
Malonyl- CoA	-71.9 ± 2.8	-202.7 ± 13.6	-21.6 ± 4.2
Cinnamoyl- CoA	-53.6 ± 1.7	-39.5 ± 17.1	-19.4 ± 2.3
<i>p</i>-Coumaroyl- CoA	-63.7 ± 9.2	-73.2 ± 24.9	-27.0 ± 3.1
Caffeoyl-CoA	-61.3 ± 2.9	-60.5 ± 15.8	-28.5 ± 1.8
Feruloyl-CoA	-66.1 ± 4.6	-198.0 ± 18.0	-30.0 ± 3.1
Acetyl-CoA (Reference)	-77.6 ± 2.5	-249.9 ± 11.1	-12.9 ± 1.5
CoA (Reference)	-89.7 ± 6.2	-301.6 ± 10.9	-16.2 ± 3.5

Residues of *BrCHSv5* involved in the interactions with the ligands are shown in Table 4.14. *p*-Coumaroyl-CoA formed salt bridge interaction with the residues Arg58 and Lys62 of the receptor, respectively. π -sigma interaction was observed between the Ile254 residue of the receptor and *p*-coumaroyl-CoA. The Cys164 residue of the receptor formed π -sulfur interaction with malonyl-CoA. Figure 4.11 depicts the interactions formed in the ligand-bound *BrCHSv5* complexes. Malonyl-CoA formed a hydrogen bond with Asn336 residue at the catalytic triad of *BrCHSv1*. Both cinnamoyl-CoA and *p*-coumaroyl-CoA formed one hydrogen bond with the receptor, respectively.

Table 4.14: List of residues of *BrCHSv5* formed interactions with the ligands.

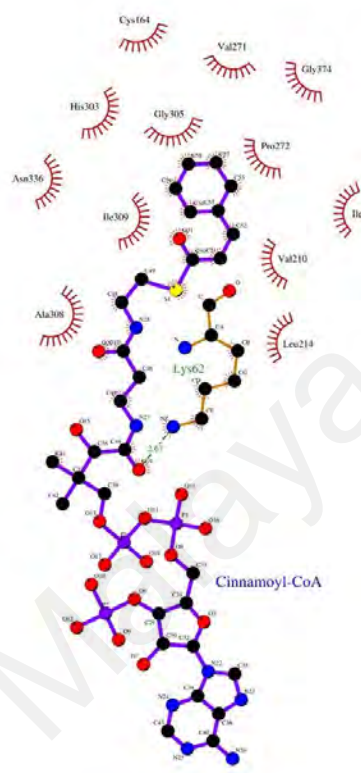
Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Arg58 (2.66, 2.70) Gly305 (3.29) Asn336 (3.04)	Cys164, Leu214, Phe215, Ile254, Val271, Pro272, Gly306, Gly374, Pro375	His303 (π - Sulfur) Lys62 (Alkyl) Ala308 (Alkyl)	Arg58	-
Cinnamoyl-CoA	Lys62 (2.67)	Val210, Leu214, Phe215, Val271, Pro272, Gly305, Gly306, Ala308, Ile309, Asn336, Gly374	His303, Cys164 (π -Alkyl) Ile254 (π -Alkyl)	-	-
<i>p</i>-Coumaroyl-CoA	Arg58 (2.79)	Cys164, Val210, Leu214, Phe215, Phe265, Leu267, His303, Gly305, Gly306, Ala308, Ile309, Gly374, Pro375	Ile254 (π -sigma) Arg58, Lys62 (Alkyl)	Arg58 Lys62	-
Caffeoyl-CoA	Lys62 (2.83)	Met59, Cys164, Val210, Leu214, Phe215, Ile254, Leu267, His303, Gly305, Gly306, Ala308, Ile309, Asn336, Gly374	-	Arg58 Lys62	-
Feruloyl-CoA	Leu258 (2.63) Arg259 (2.58, 2.86)	Gly161, Gly163, Cys164, Phe215, Ile246, Glu251, Ile254, Gly255, His257, Leu263, Gly376	Leu263 (π - Sigma) Pro375 (π - Alkyl)	Arg259	Gly256
Acetyl-CoA (Reference)	Arg58 (2.67) Lys62 (2.63, 3.29) His303 (3.13) Asn336 (2.74)	Cys164, Val210, Leu214, Phe215, Leu267, Pro272, Gly305, Gly306, Ala308, Ile309	-	Arg58 Lys62	-
CoA (Reference)	Arg58 (2.66, 2.93) Lys62 (2.48) Gly305 (3.13) Asn336 (2.89)	Met59, Leu214, Phe215, Leu267, Val271, Gly306, Ala308, Ile309	Val210 (Alkyl) His303 (π - sulfur)	Arg58 Lys62	-

-: Absence

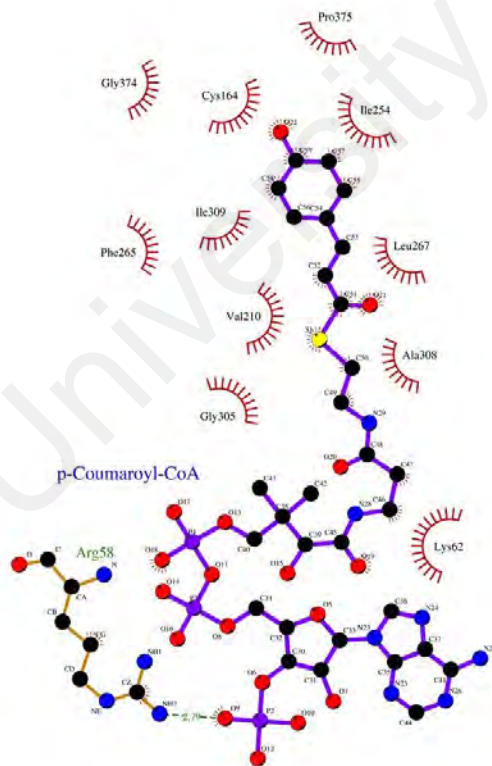
(a)



(b)



(c)



(d)

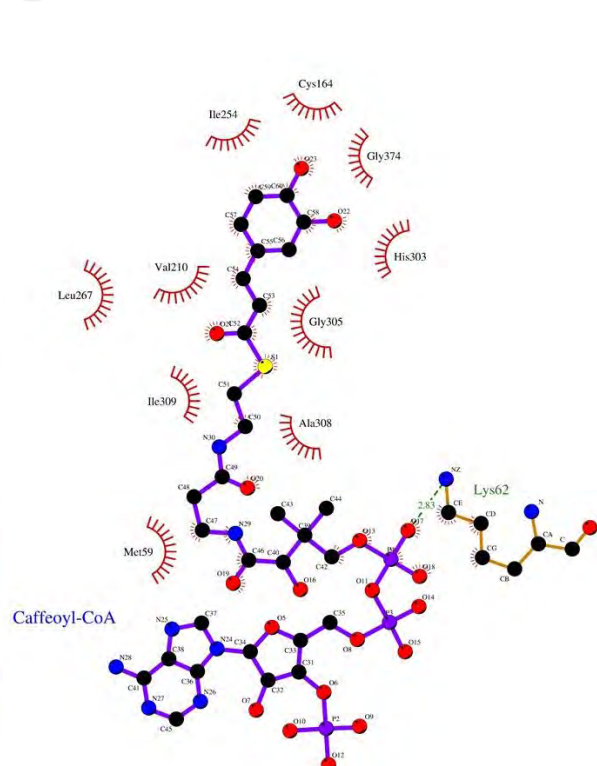


Figure 4.11: Docked conformations of *BrCHSv5* with the ligands. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

4.4 Trajectory Analysis of Molecular Dynamics Simulations

4.4.1 Root-Mean-Square Deviation (RMSD)

The stability of protein-ligand complexes measured by RMSD and is shown in Figure 4.12 for all the *BrCHS* receptor variants. *BrCHSv1*-malonyl-CoA complex showed fluctuations from 2 to 6 ns and became stable at the end of the simulation. All the ligand-protein complexes fluctuated from 0.05 to 0.25 nm. Complexes of *BrCHSv2* with ligands showed fluctuations from 0.1 to 0.2 nm throughout 10 ns simulations. However, only complex with malonyl-CoA showed high fluctuation for the last 5 ns of the simulations. Fluctuations in the range of 0.1 to 0.25 nm were observed for the complexes of *BrCHSv3*. Meanwhile, feruloyl-CoA showed high fluctuations compared to other ligands which are about 0.25 nm. The complex of *BrCHSv4*-Cinnamoyl-CoA showed slightly higher fluctuations than other ligands. On the other hand, ligand-bound *BrCHSv5* complexes also showed the same trend of fluctuations as in *BrCHSv2* which is in the range of 0.1 – 0.25 nm.

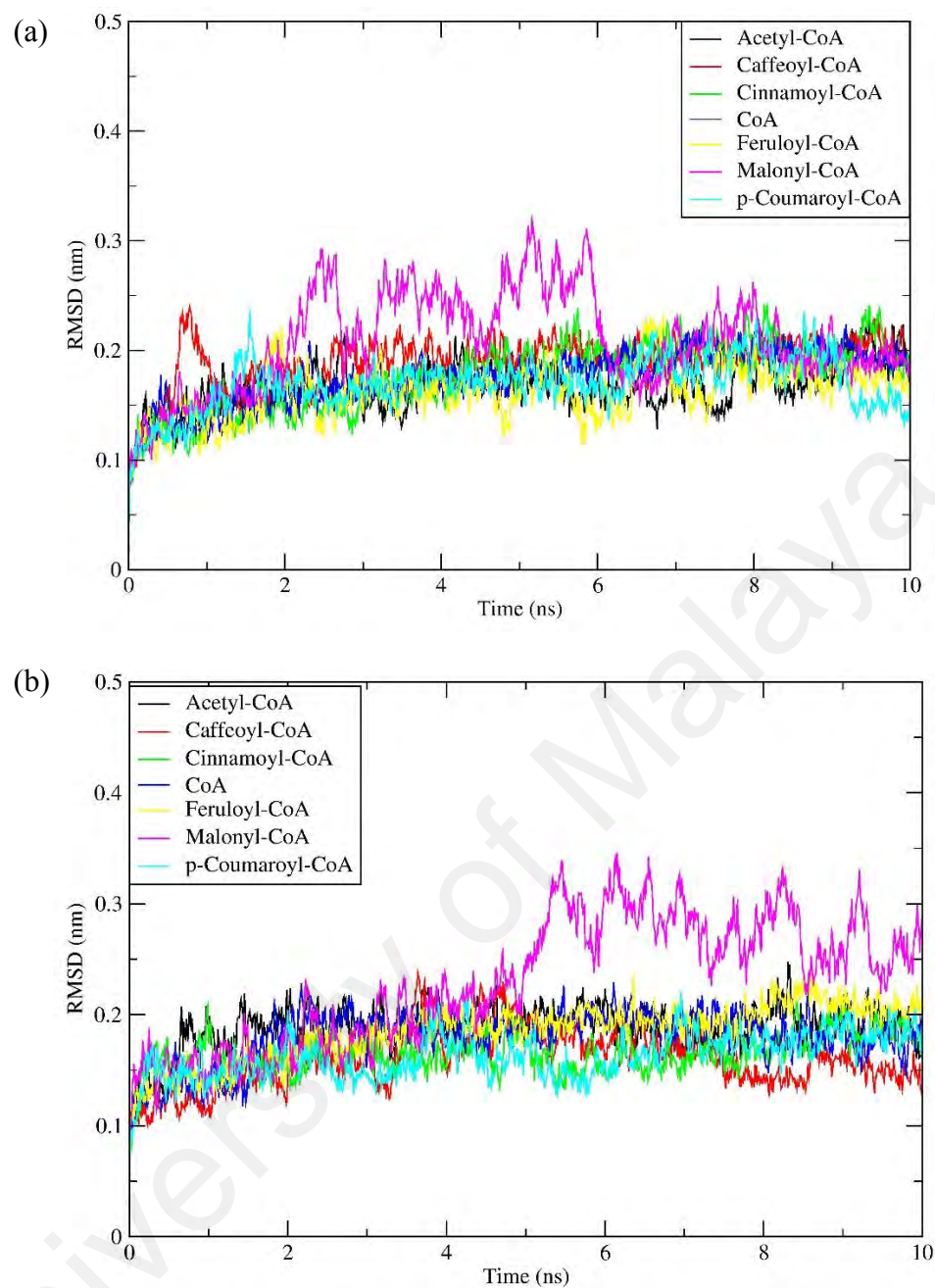


Figure 4.12: RMSD of *BrCHS* receptor variants with the substrate ligands after 10 ns simulation. (a) Variant 1; (b) variant 2; (c) variant 3; (d) variant 4; (e) variant 5.

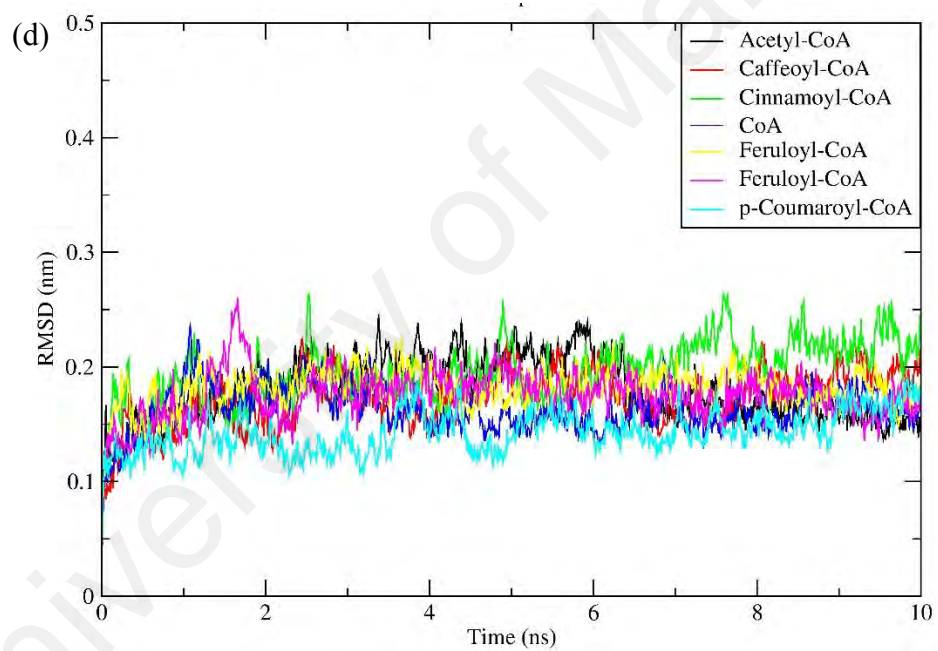
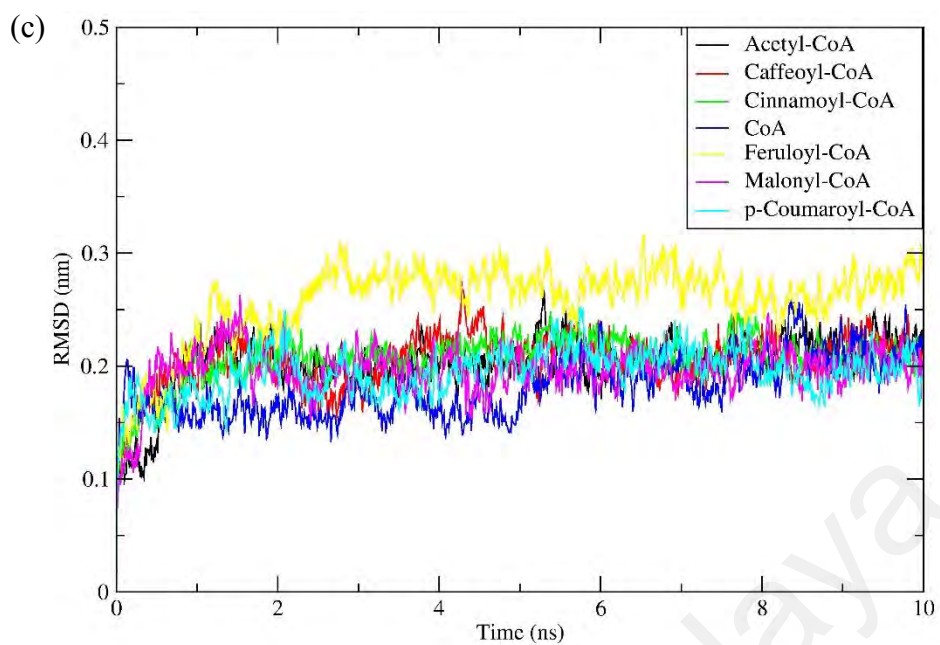


Figure 4.12, continued.

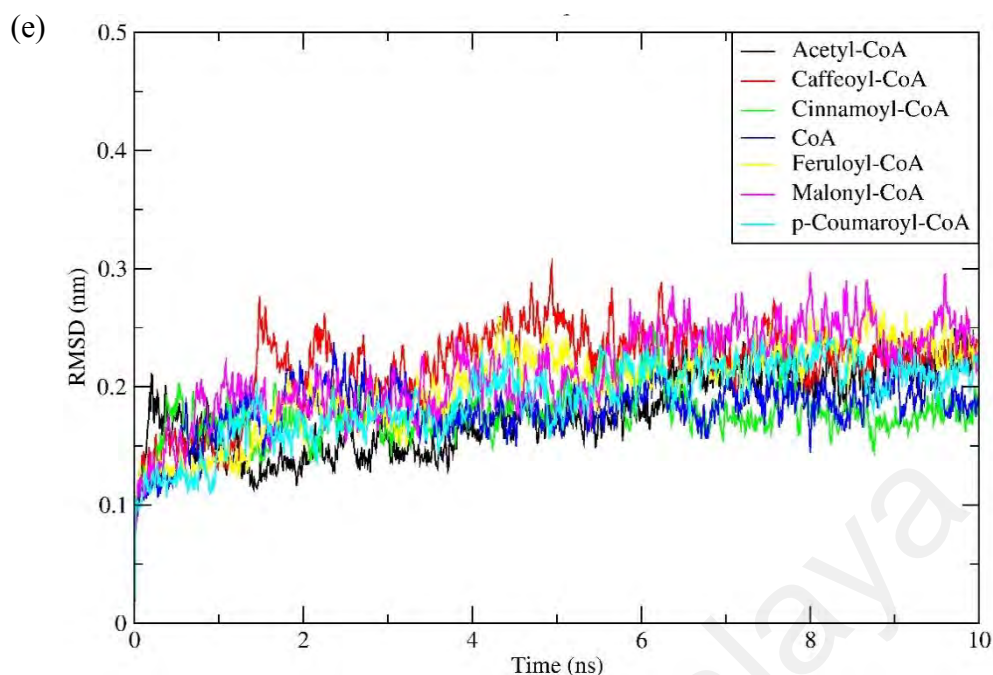


Figure 4.12, continued.

4.4.2 Root-Mean-Square Fluctuation (RMSF)

RMSF of all residues of *BrCHS* complexes were calculated throughout the 10 ns simulation and depicted in Figure 4.13. All the residues of *BrCHSv1* receptors showed a similar trend of fluctuations except for CoA bound receptor. The complex of *p*-coumaroyl-CoA bound with *BrCHSv1*, *BrCHSv2*, *BrCHSv3* and *BrCHSv5* showed higher fluctuations compared to cinnamoyl-CoA. Whereas, CoA and *p*-coumaroyl-CoA bound *BrCHSv2* complexes showed slightly higher fluctuations than other ligands. A similar trend of fluctuations was observed in all residues of the ligand bound to *BrCHSv3* and *BrCHSv4* complexes. Only *p*-coumaroyl-CoA bound to *BrCHSv5* receptor showed slightly higher fluctuations compared to other ligand-bound complexes.

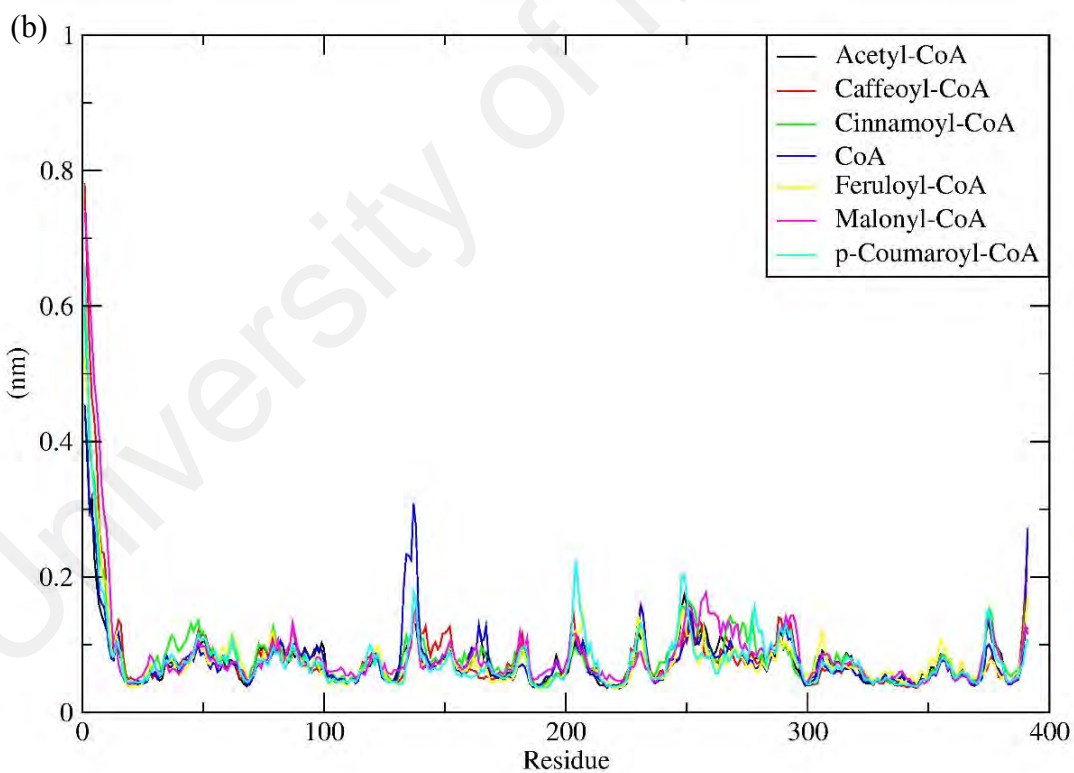
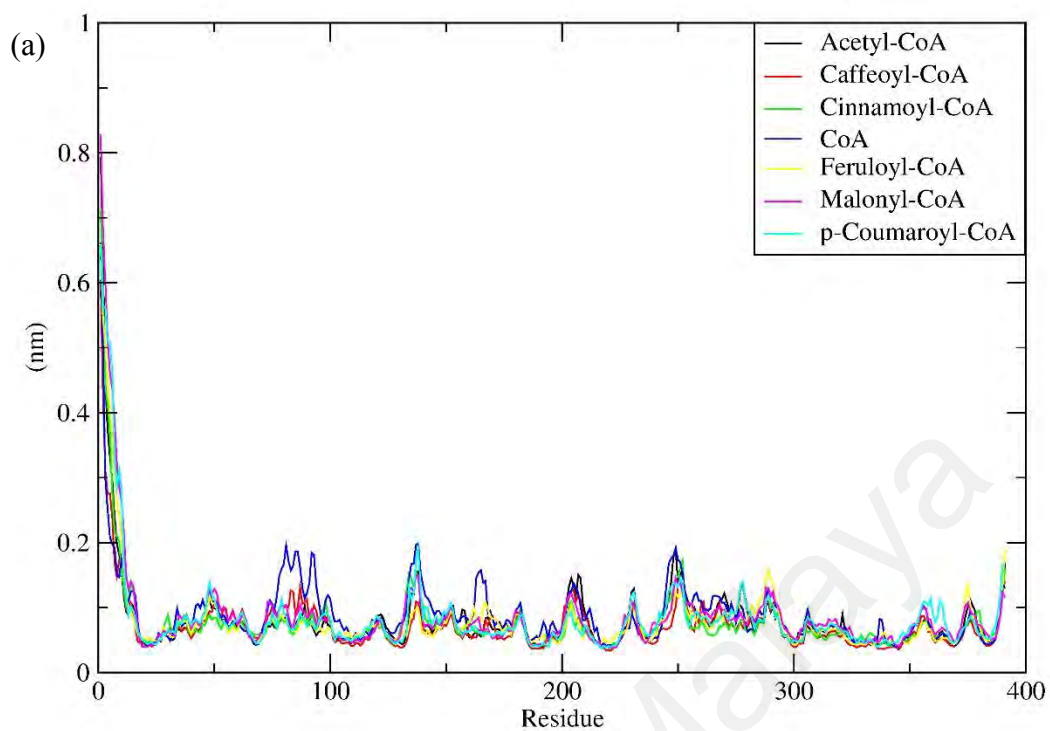


Figure 4.13: RMSF of *BrCHS* receptor variants with the substrate ligands after 10 ns simulation. (a) Variant 1; (b) variant 2; (c) variant 3; (d) variant 4; (e) variant 5.

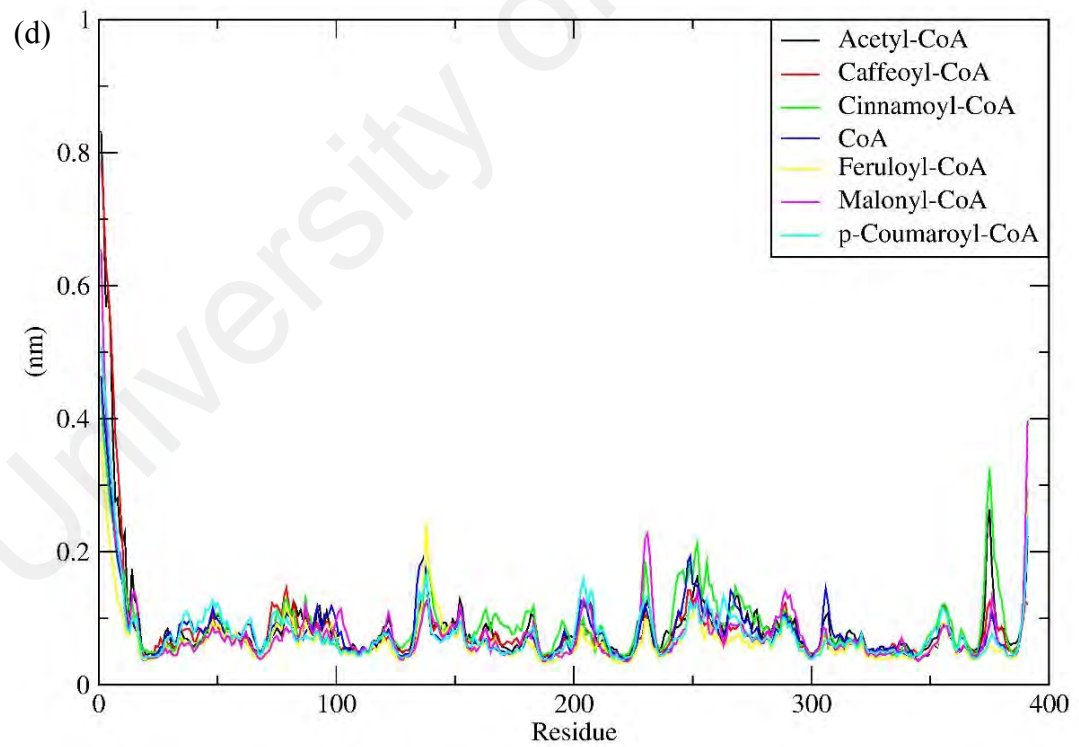
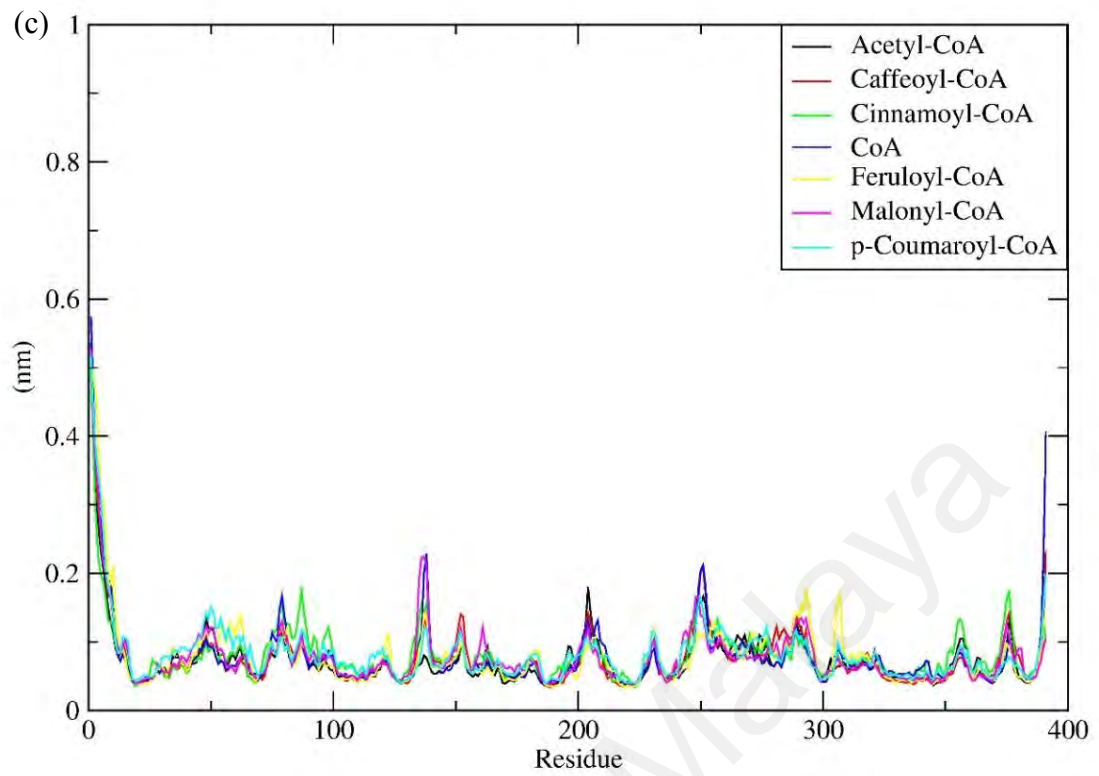


Figure 4.13, continued.

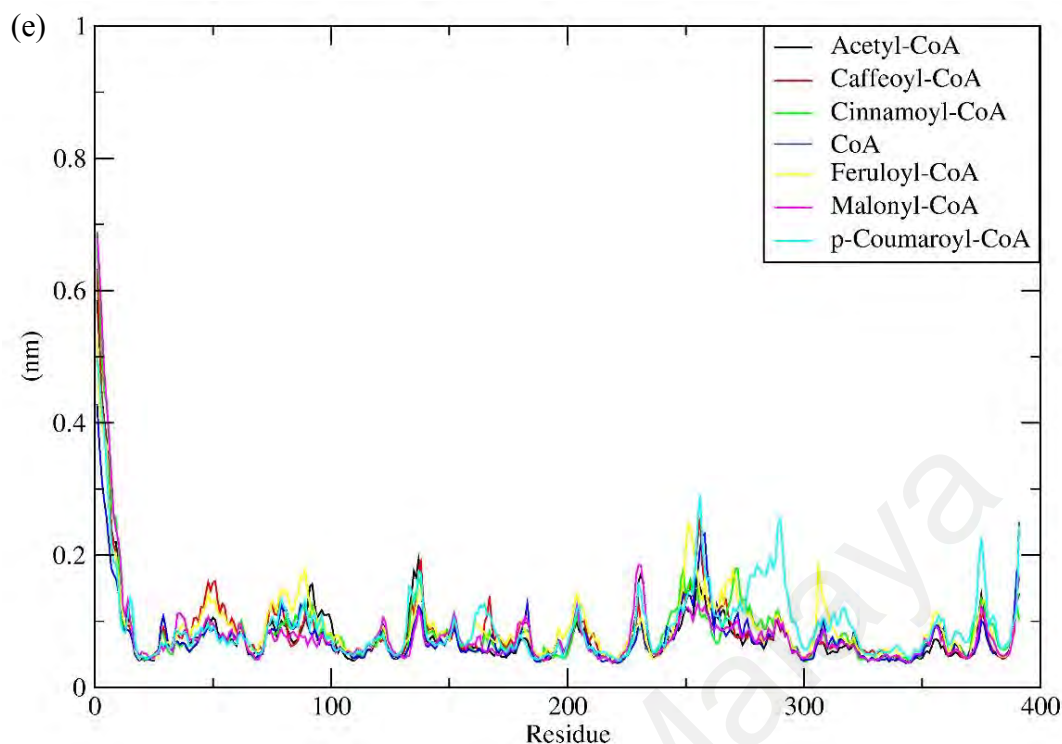


Figure 4.13, continued.

4.4.3 Radius of Gyration

The radius of gyration of all five *BrCHS* variant complexes shown as in Figure 4.14. Ligands bound *BrCHSv1* and *BrCHSv4* complexes showed a minor downward trend of fluctuations in the radius of gyration. *p*-Coumaroyl-CoA bound *BrCHSv4* complex showed a slight decrease in the fluctuation of the radius of gyration towards the last 1 ns of simulation. The radius of gyration of *BrCHSv2* and *BrCHSv3* complexes showed a similar trend of fluctuations ranging from 2.075 to 2.125 nm. Cinnamoyl-CoA and *p*-coumaroyl-CoA bound *BrCHSv2* complexes showed a similar way of fluctuations towards the last 5 ns of simulations. On the other hand, the radius of gyration of the *BrCHSv5* complexes a stable trend of fluctuations was seen except for cinnamoyl-CoA. The radius of gyration of cinnamoyl-CoA-*BrCHSv5* complex showed a sharp decrease in fluctuations in the last 4 ns of MD simulation.

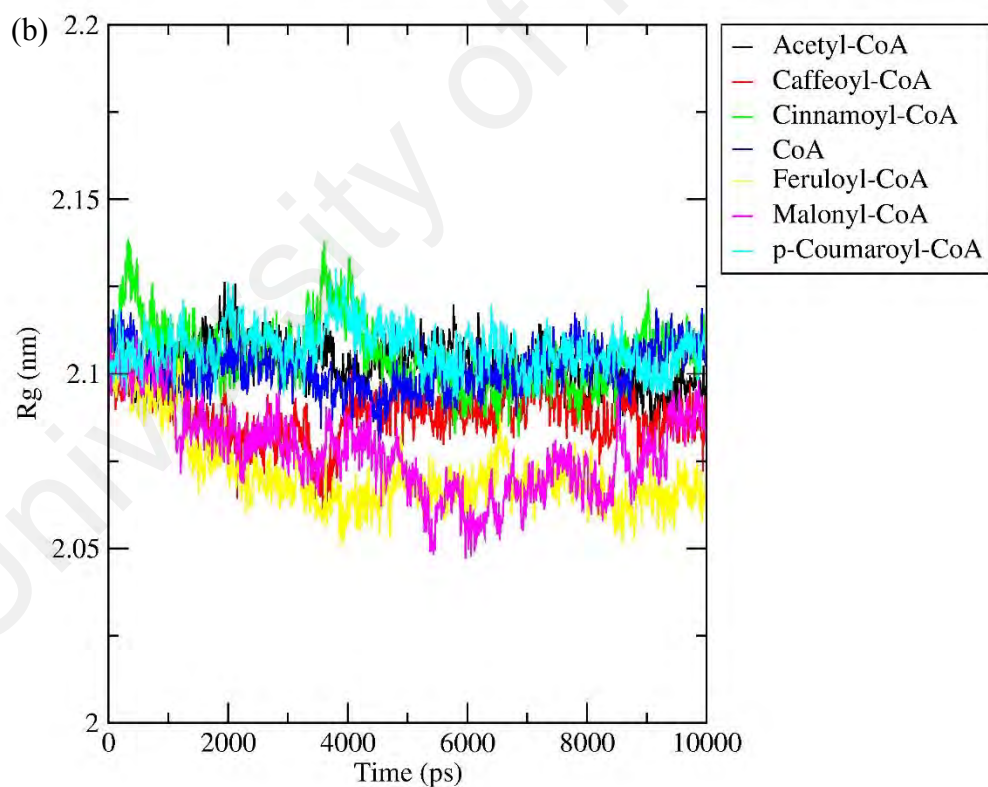
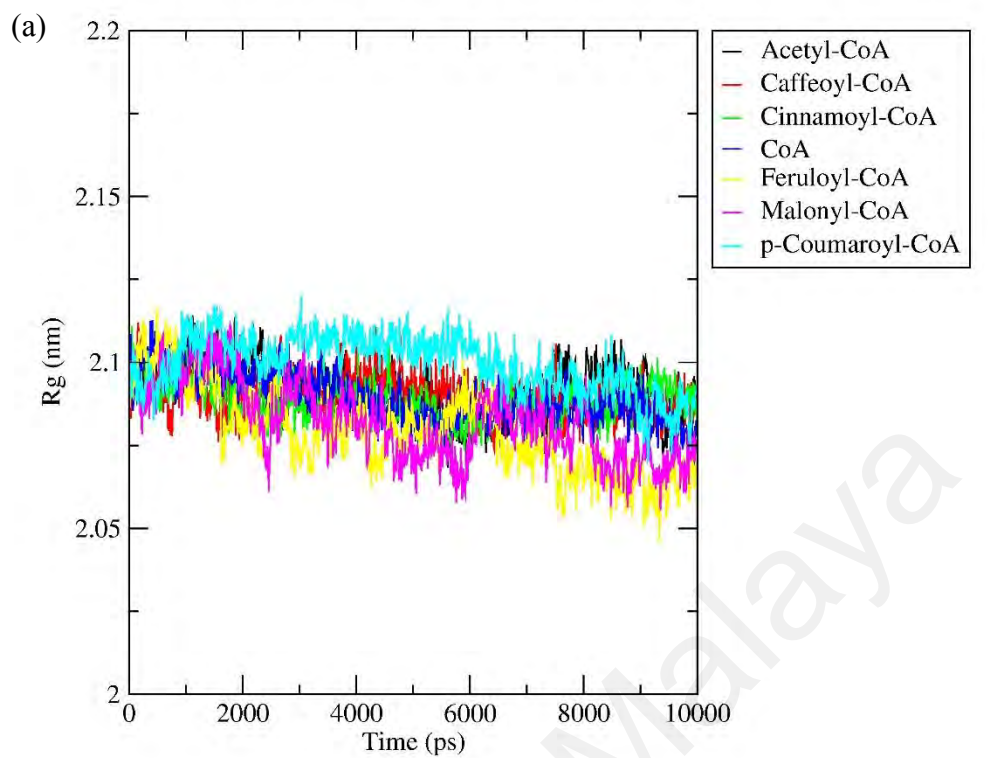


Figure 4.14: Radius of gyration of *BrCHS* receptor variants with substrate ligands after 10 ns simulation. (a) Variant 1; (b) variant 2; (c) variant 3; (d) variant 4; (e) variant 5.

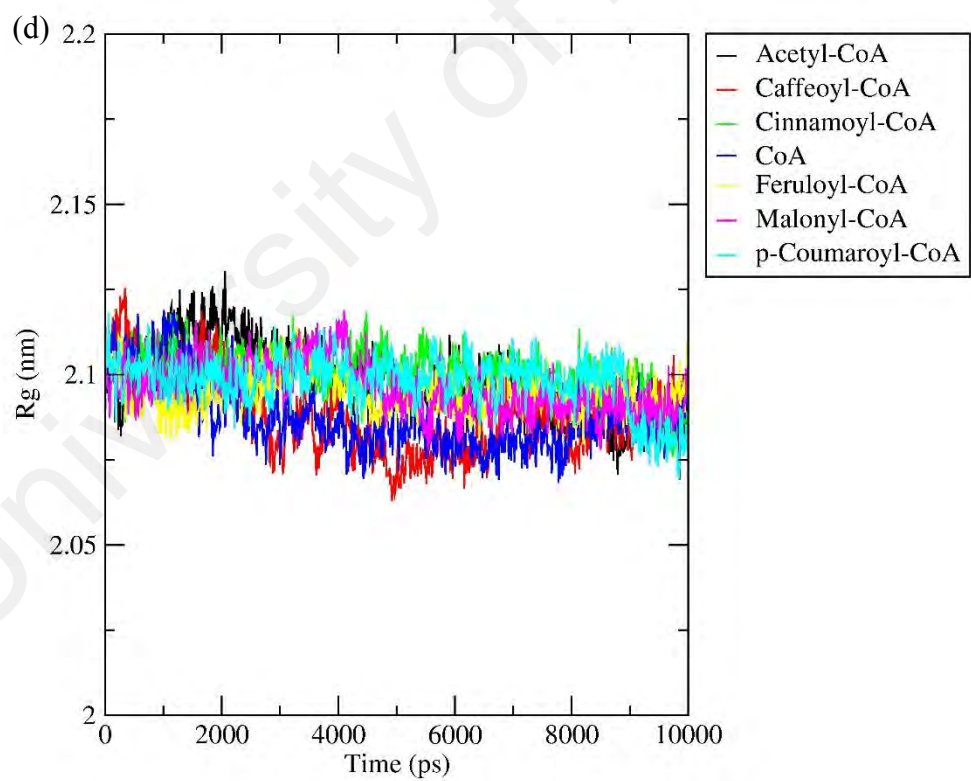
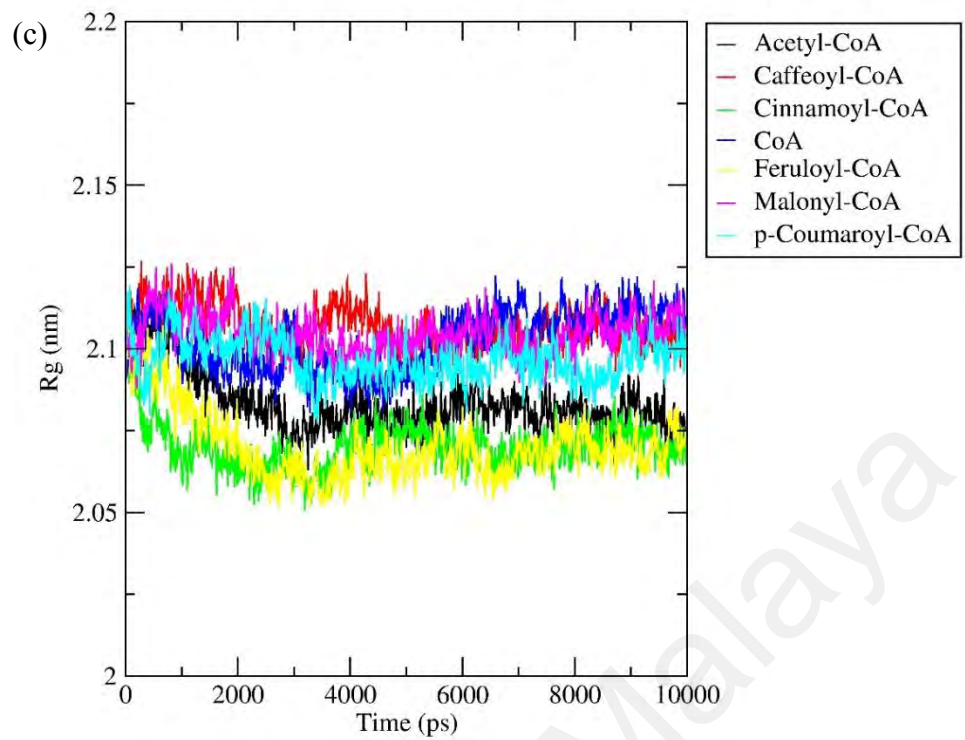


Figure 4.14, continued.

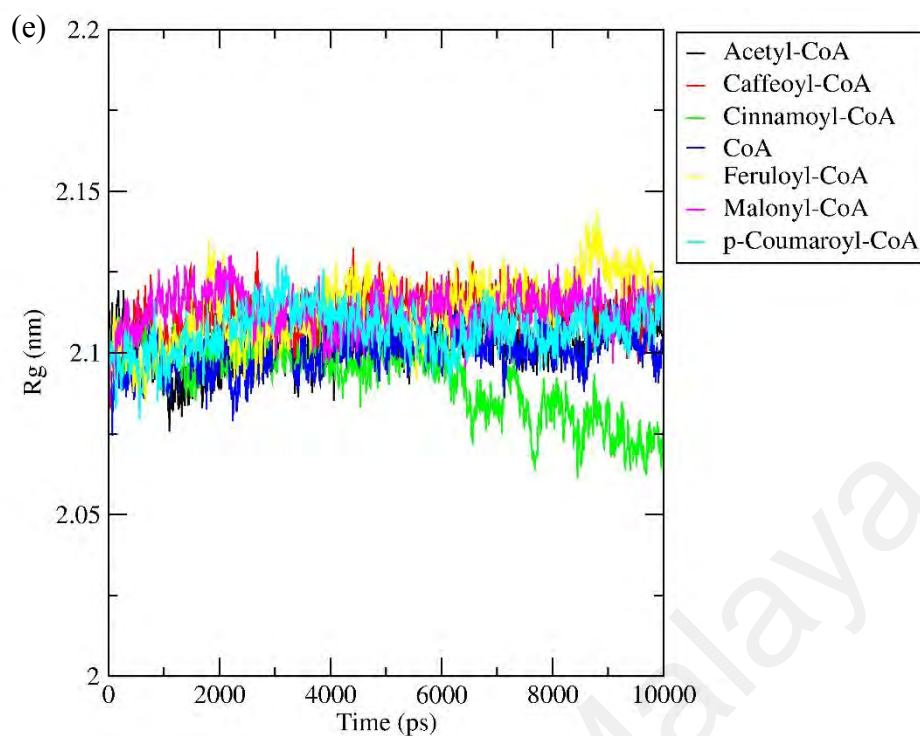


Figure 4.14, continued.

4.4.4 Binding Free Energy

Binding free energies of the protein-ligand complexes were calculated throughout the 10 ns simulation. Binding free energy profile of ligand bound to *BrCHSv1* complexes is shown in Table 4.15. The most negative binding free energy was recorded by cinnamoyl-CoA-*BrCHSv4* complex which is -182.958 ± 2.836 kJ/mol. It is followed by *p*-coumaroyl-CoA with the binding free energy of -134.110 ± 4.120 . Hence, *BrCHSv1* has a higher binding affinity towards cinnamoyl-CoA than *p*-coumaroyl-CoA.

Table 4.15: Binding free energy of *BrCHS* variant 1 receptor with the ligands.

Ligand	Overall Coulombic (kJ/mol)	Overall Lennard- Jones (kJ/mol)	Polar solvation (kJ/mol)	Non-polar solvation (kJ/mol)	Binding free energy (kJ/mol)
Malonyl- CoA	-54.106 ± 33.496	-154.823 ± 1.988	201.987 ± 28.620	-25.346 ± 0.134	-32.288 ± 5.631
Cinnamoyl- CoA	-300.557 ± 11.139	-228.677 ± 2.028	376.174 ± 9.542	-29.898 ± 0.072	-182.958 ± 2.836
<i>p</i>-Coumaroyl- CoA	-422.736 ± 18.966	-169.344 ± 2.113	482.451 ± 16.675	-24.481 ± 0.106	-134.110 ± 4.120
Caffeoyl-CoA	971.081 ± 12.790	-170.369 ± 2.298	-546.781 ± 10.770	-22.945 ± 0.194	230.985 ± 4.642
Feruloyl-CoA	937.699 ± 11.972	-219.002 ± 2.466	-503.530 ± 9.785	-27.848 ± 0.183	187.319 ± 4.168
Acetyl-CoA (Reference)	-38.954 ± 13.558	-104.089 ± 2.073	126.596 ± 10.226	-16.716 ± 0.142	-33.163 ± 3.593
CoA (Reference)	-345.549 ± 14.405	-131.489 ± 1.726	396.769 ± 10.673	-22.063 ± 0.079	-102.332 ± 4.256

Table 4.16 shows the residues of *BrCHSv1* involved in the interactions with the ligands. Residues Lys55 and Arg58 of the receptor formed salt bridge with cinnamoyl-CoA. Cys164 of the *BrCHSv1* formed π -alkyl interaction with cinnamoyl-CoA, whereas π -sulfur interaction with *p*-coumaroyl-CoA. Ile254, on the other hand, formed π -alkyl interaction with both cinnamoyl-CoA and *p*-coumaroyl-CoA, respectively. Arg58 of the receptor formed unfavorable bonds with malonyl-CoA. Interactions formed in the *BrCHSv1* complexes are shown in Figure 4.15. Cinnamoyl-CoA formed more hydrogen bonds with the receptor compared to *p*-coumaroyl-CoA.

Table 4.16: List of interactions formed between *BrCHSv1* and the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Arg58 (2.63, 3.22) Lys62 (2.56, 2.68) His303 (3.25) Ala308 (3.21)	Cys164, Val210, Leu214, Gly305, Gly306, Pro307, Ile309	Phe373 (π - sulfur) Lys55 (π -Cation, π -alkyl) Leu296 (Alkyl)	Arg58 Lys62	Arg58
Cinnamoyl-CoA	Thr50 (3.17) Lys55 (2.61, 2.64) Arg58 (2.79, 2.84) Gly306 (2.88)	Glu51, Glu54, Val210, Leu214, Phe215, Phe265, Gly305, Ile309, Gly374, Pro375	Lys62, Arg58, Met59 (Alkyl) Thr50, Ile254, Cys164 (π -alkyl)	Lys55 Arg58	Arg58
<i>p</i>-Coumaroyl-CoA	Arg58 (2.69, 2.72) Lys62 (2.73)	Val210, Leu214, Phe215, Phe265, Gly305, Gly306, Gly374, Pro375	Cys164 (π - sulfur) Lys55, Arg58, Met59 (Alkyl) Ile254 (π -Alkyl)	Arg58 Lys62	-
Caffeoyl-CoA	Arg259 (2.58, 2.89)	Thr132, Asp136, , Gln161, His257, Leu258, Leu263, Phe265	Gly256 (Amide- π stacked) Met158 (alkyl) Met137, Pro375 (π -alkyl)	Arg259	Arg259
Feruloyl-CoA	Gly139 (3.03) Gln161 (3.25) Leu258 (3.02, 3.06) Arg259 (2.59, 2.81)	Met137, Pro138, Asn155, Arg156, Met158, Gly256, Pro375	Leu258 (π - Alkyl) Arg156 (π -donor HB)	Arg259	Arg259
Acetyl-CoA	Lys57 (2.79) Arg58 (2.61)	Met59, Ser63, Val210, Leu214	-	Arg58 Lys62	-
(Reference)	Lys62 (2.66, 2.83)				
CoA	Arg58 (2.63, 2.69)	Glu51, Glu54, Val210,	Met59, Lys62 (Alkyl)	Arg58 Lys62	Arg58
(Reference)	Lys62 (2.61, 2.62)	Leu214, Ile254, Phe265, Leu267	Lys55 (π -Alkyl)		

-: Absence

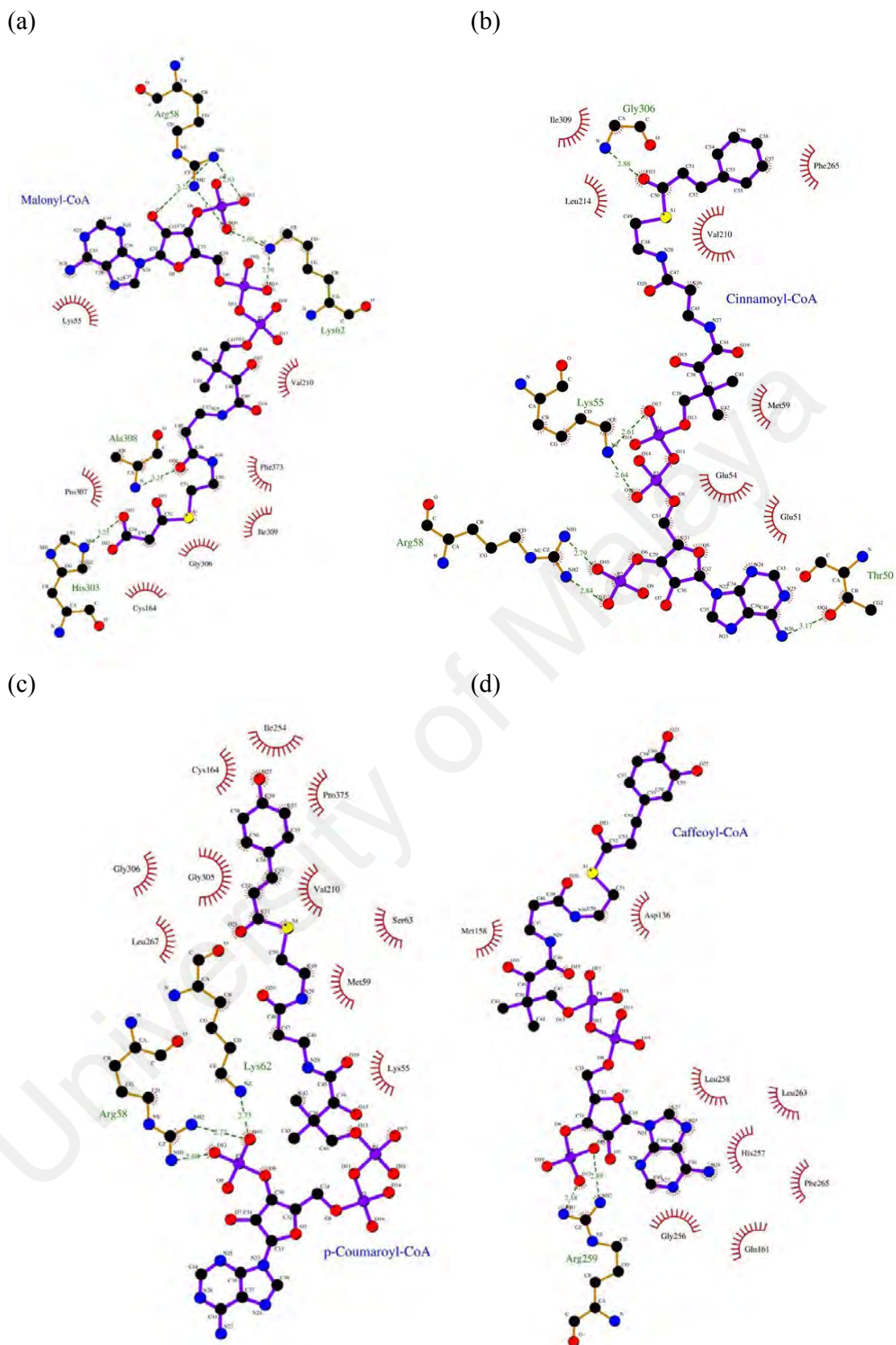


Figure 4.15: Interactions of *BrCHSvI* with the ligands after 10 ns simulation. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

Table 4.17 shows the binding free energies of ligands with the *BrCHSv2* receptor. Malonyl-CoA recorded less negative binding free energy than *p*-coumaroyl-CoA. Cinnamoyl-CoA has the most negative binding free energy which is -181.575 ± 4.214 kJ/mol. Therefore, cinnamoyl-CoA has higher binding affinity for *BrCHSv2* than *p*-coumaroyl-CoA.

Table 4.17: Binding free energy of *BrCHS* variant 2 receptor with the ligands.

Ligand	Overall Coulombic (kJ/mol)	Overall Lennard-Jones (kJ/mol)	Polar solvation (kJ/mol)	Non-polar solvation (kJ/mol)	Binding free energy (kJ/mol)
Malonyl- CoA	-556.254 ± 21.039	-147.609 ± 2.358	679.662 ± 14.341	-26.778 ± 0.162	-50.979 ± 8.008
Cinnamoyl-CoA	-704.122 ± 16.247	-135.843 ± 2.147	682.663 ± 12.591	-24.272 ± 0.108	-181.575 ± 4.214
<i>p</i>-Coumaroyl-CoA	150.178 ± 19.429	-207.103 ± 2.150	11.703 ± 14.650	-26.749 ± 0.141	-71.971 ± 6.456
Caffeoyl-CoA	-478.579 ± 12.723	-129.525 ± 3.049	477.562 ± 9.291	-21.952 ± 0.279	-152.495 ± 3.550
Feruloyl-CoA	1121.062 ± 16.406	-163.049 ± 1.994	-844.146 ± 11.352	-20.962 ± 0.240	92.905 ± 6.787
Acetyl-CoA (Reference)	-532.167 ± 21.094	-198.975 ± 1.904	599.112 ± 18.114	-26.991 ± 0.141	-159.020 ± 4.554
CoA (Reference)	-589.711 ± 13.829	-148.138 ± 2.237	585.248 ± 10.464	-23.264 ± 0.101	-175.865 ± 4.100

Table 4.18 shows the interactions formed between *BrCHSv2* and the ligands. Residues Lys55, Arg58 and Lys62 of the receptor formed salt bridge with cinnamoyl-CoA and between *p*-coumaroyl-CoA. Cys164 of the receptor formed π -alkyl interaction with cinnamoyl-CoA and *p*-coumaroyl-CoA, respectively. A repulsive interaction was formed by the Glu54 residue of *BrCHSv2* with *p*-coumaroyl-CoA. Figure 4.16 depicts the interactions formed between the ligands and *BrCHSv2* receptor after the simulations. Hydrogen bonding was observed between Lys55 and Lys62 residues of the receptor with the cinnamoyl-CoA, malonyl-CoA and *p*-coumaroyl-CoA.

Table 4.18: List of interactions formed between *BrCHSv2* and the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Lys55 (2.74) Lys62 (2.56, 2.84) Asp207 (2.76) Asp270 (2.97)	Cys164, Leu206, Val210, Leu214, Phe215, Gly305, Gly306, Asn336	Pro272, Ala308 (π -alkyl)	Lys55 Lys62	Asp311
Cinnamoyl-CoA	Lys55 (2.72, 3.10) Arg58 (2.60, 2.68) Lys62 (2.67, 2.80)	Asp207, Val210, Gly211, Leu214, Phe215, Ile254, Phe265, Leu267, His303, Gly305, Asn336, Gly374, Pro375	Cys164 (π -alkyl)	Lys55 Arg58 Lys62	-
<i>p</i>-Coumaroyl-CoA	Lys55 (2.63) Arg58 (2.57, 3.10) Lys62 (2.64)	Val210, Leu214, Phe215, Phe265, Leu267, Gly305, Leu377	Lys55, Arg58, Met59 (alkyl) Cys164, Ile254, Val271 (π -alkyl)	Lys55 Arg58 Lys62	Glu54
Caffeoyl-CoA	Arg58 (2.60, 2.64), Lys62 (2.64, 2.71)	Cys164, Val210, Leu214, Phe215, Ile254, Phe265, Gly305, Gly306, Pro375	Lys55 (π - Cation) Leu267(π -alkyl) Ala308 (Alkyl)	Arg58 Lys62	-
Feruloyl-CoA	Gln244 (3.03, 3.13) Gly256 (2.83) His257 (3.07)	Ser243, Ile254, Phe265, Thr378	Ile246 (Alkyl) His257 (π - π stacked) Pro375 (π -alkyl)	-	-
Acetyl-CoA (Reference)	Arg58 (2.67, 2.69) Lys62 (2.58, 2.70)	Cys164, Val210, Ile254, Leu267, Val271, Pro272, His303, Gly305, Gly306, Phe373	Glu54 (Amide- π Stacked) Met59, Lys62, Pro307 (Alkyl) Lys55 (π -Alkyl)	Arg58 Lys62	-
CoA (Reference)	Arg58 (2.74, 3.03), Lys62 (2.63, 2.68) Asp207 (2.73, 2.92)	Glu54, Leu214, Ile254, Leu267, Val271, Gly305, Gly306, Ile309	Met59, Leu206, Val210 (Alkyl) Lys55 (π -alkyl)	Arg58 Lys62	-

-: Absence

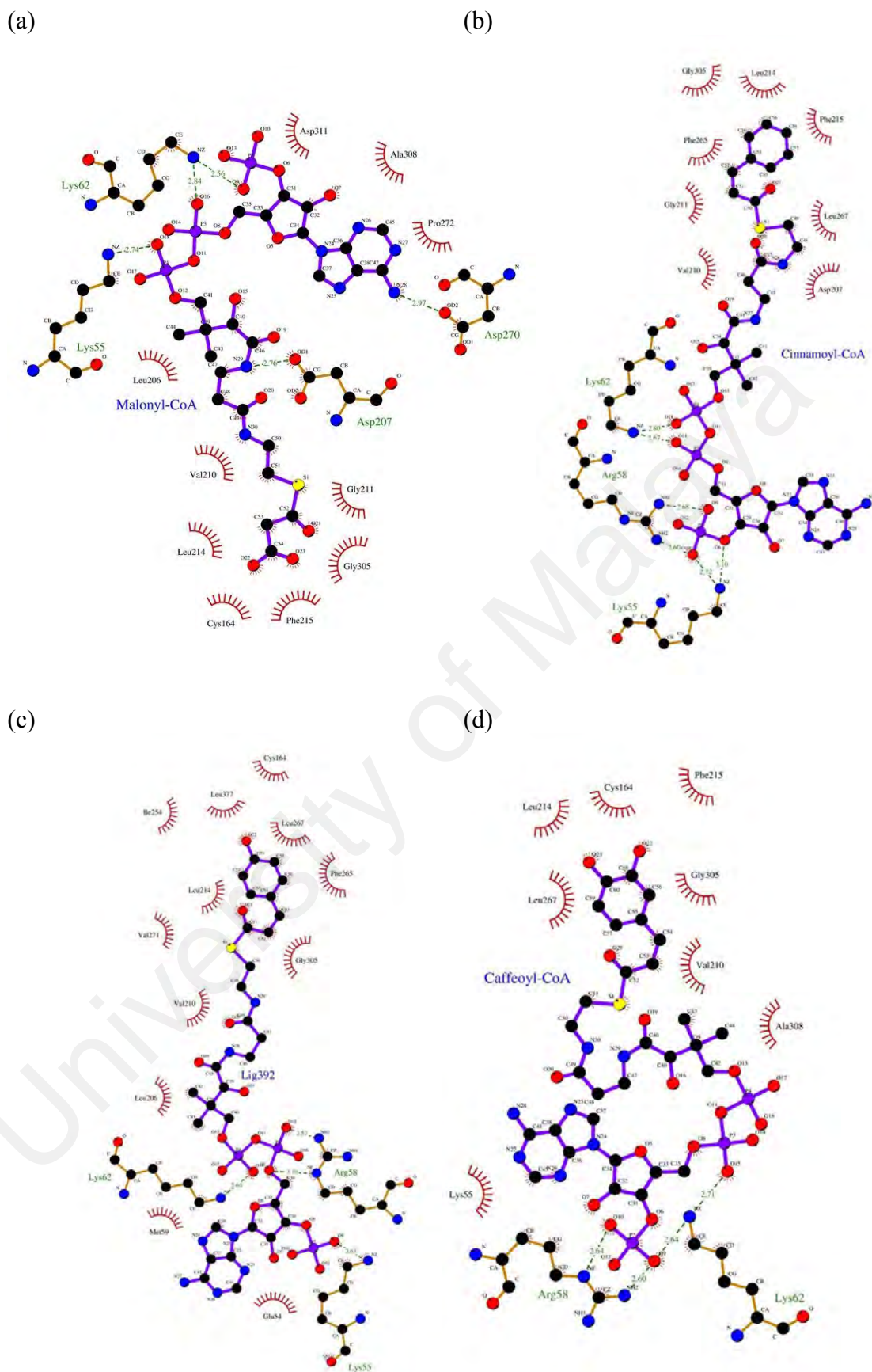


Figure 4.16: Interactions of *BrCHSv2* with the ligands after 10 ns simulation. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

Binding free energies of ligand bound to *BrCHSv3* complexes are shown as in Table 4.19. Malonyl-CoA recorded more negative binding free energy compared to *p*-coumaroyl-CoA. Cinnamoyl-CoA recorded the most negative binding free energy among the substrate ligands which is -255.234 ± 4.826 kJ/mol. Hence, *BrCHSv3* has a higher binding affinity towards cinnamoyl-CoA than *p*-coumaroyl-CoA.

Table 4.19: Binding free energy of *BrCHS* variant 3 receptor with the ligands.

Ligand	Overall Coulombic (kJ/mol)	Overall Lennard-Jones (kJ/mol)	Polar solvation (kJ/mol)	Non-polar solvation (kJ/mol)	Binding free energy (kJ/mol)
Malonyl- CoA	-1864.798 ± 18.275	-101.637 \pm 2.513	1782.147 \pm 12.304	-20.758 \pm 0.156	-205.046 \pm 5.985
Cinnamoyl-CoA	-1590.867 ± 17.595	-198.490 \pm 2.488	1563.171 \pm 13.887	-29.048 \pm 0.162	-255.234 \pm 4.826
<i>p</i>-Coumaroyl-CoA	-1416.737 ± 23.597	-165.740 \pm 1.878	1463.582 \pm 18.326	-25.700 \pm 0.097	-144.596 \pm 5.842
Caffeoyl-CoA	-2046.944 ± 14.081	-148.681 \pm 1.886	1970.232 \pm 11.318	-26.893 \pm 0.075	-252.286 \pm 4.569
Feruloyl-CoA	-440.750 \pm 13.165	-168.575 \pm 2.508	663.352 \pm 9.956	-23.730 \pm 0.172	30.297 \pm 4.700
Acetyl-CoA (Reference)	-1795.403 ± 11.819	-170.547 \pm 2.980	1715.303 \pm 9.231	-24.995 \pm 0.157	-275.642 \pm 3.850
CoA (Reference)	-1864.798 ± 18.275	-101.637 \pm 2.513	1782.147 \pm 12.304	-20.758 \pm 0.156	-205.046 \pm 5.985

Residues involved in the interaction between *BrCHSv3* and the ligands are as shown in Table 4.20. Cinnamoyl-CoA formed more van der Waals interactions with the receptor. The phe215 residue of the receptor formed π -stacked interactions with cinnamoyl-CoA and *p*-coumaroyl-CoA, respectively. Figure 4.17 shows the interaction of ligands with the *BrCHSv3* receptor after 10 ns simulation. Malonyl-CoA formed hydrogen bonds with Asn270 and His303 residues of the *BrCHSv1*. Cinnamoyl-CoA formed more hydrogen bonds with the receptor compared to *p*-coumaroyl-CoA.

Table 4.20: List of interactions formed between *BrCHSv3* and the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Arg58 (2.58, 2.62) Lys62 (2.57, 2.74) Asn270 (2.99) His303 (2.76)	Cys164, Ile254, Leu267, Pro272, Gly305, Ala308, Asn336	-	Arg58 Lys62	-
Cinnamoyl-CoA	Gln51 (2.76) Arg58 (2.48, 2.65, 2.72, 2.74, 3.09) Lys62 (2.62)	Glu54, Cys164, Val210, Leu214, Ile254, Phe265, Ile267, His303, Gly305, Gly306, Asn336, Pro375	Phe215 (π - π T- shaped) Lys55, Arg58, Met59 (Alkyl)	Arg58 Lys62	-
<i>p</i>-Coumaroyl-CoA	Arg58 (2.60, 2.66) Lys62 (2.60) Asp207 (2.96)	Arg194, Thr197, Leu206, Val210, Gly211, Gln212, Phe215, Ile254, Leu267, Pro375	Phe215 (π - π stacked) Lys55, Met59, Lys62 (Alkyl)	Arg58 Lys62	-
Caffeoyl-CoA	Lys55 (2.53) Arg58 (2.64, 2.92, 2.93) Lys62 (2.70, 2.78) Asn270 (3.15) His303 (3.25)	Gly163, Cys164, Val210, Leu214, Phe215, Ile254, Phe265, Val271, Gly305, Gly306, Asn336, Ser338	Ala308 (Alkyl)	Lys55 Arg58 Lys62	-
Feruloyl-CoA	Gly139 (2.88) Arg156 (2.74) Arg259 (2.66, 2.71)	Met137, Pro138, Gln161, Gly163, Phe165, Ile254, Gly376	Tyr142 (π - sulfur) Tyr160 (Alkyl)	Arg259	-
Acetyl-CoA (Reference)	Arg58 (2.56, 2.78) Lys62 (2.62, 2.72) Val271 (3.06)	Cys164, Ile254, Phe265, Leu267, Asn270, His303, Gly305, Ile309, Phe373, Gly374	Lys55 (π -cation) Met59, Val210 (Alkyl) Lys55 (π -alkyl)	Arg58 Lys62	Arg58
CoA (Reference)	Lys55 (3.27) Arg58 (2.64, 2.98) Lys62 (2.58, 2.63)	Leu214, Ile254, Phe265, Leu267, Val271, Gly305, Gly306, Ala308, Ile309, Asn336	Phe215 (π - sulfur)	Arg58 Lys62	-

-: Absence

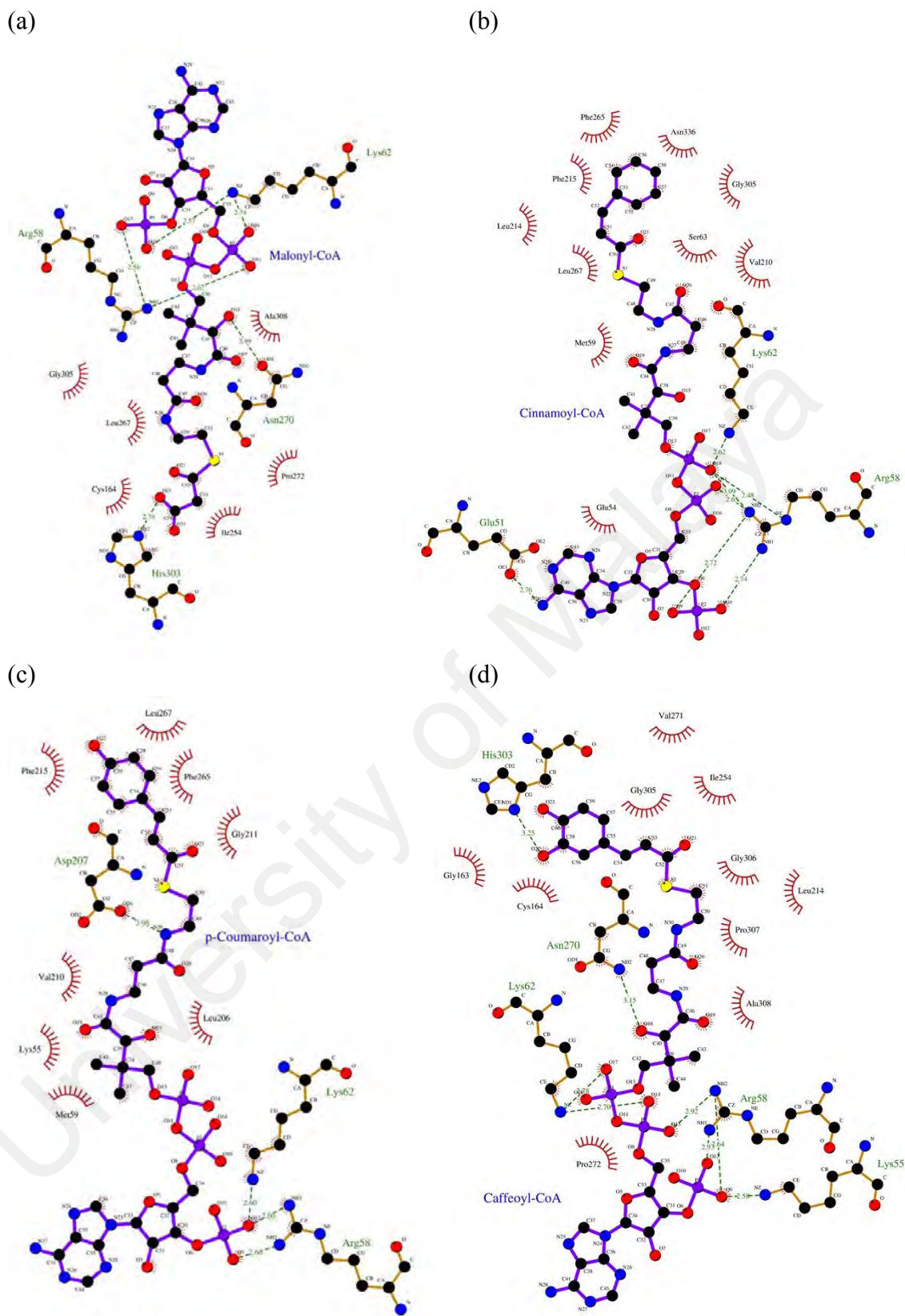
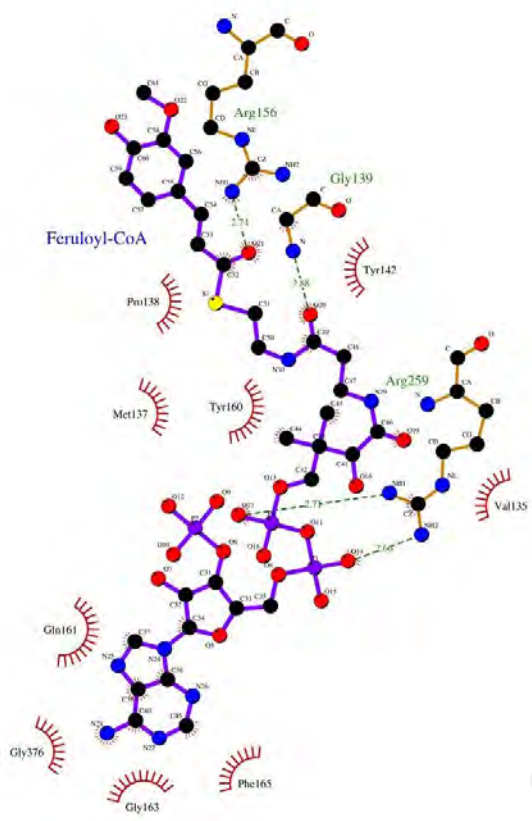
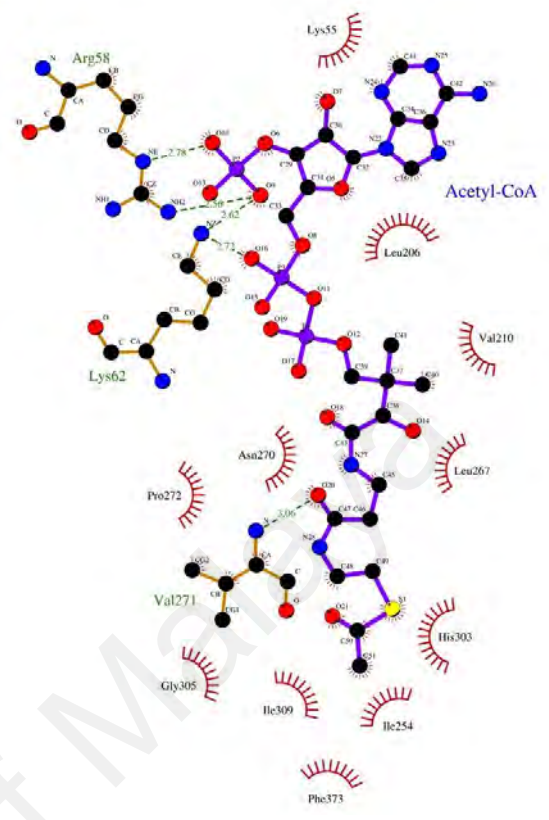


Figure 4.17: Interactions of *BrCHSv3* with the ligands after 10 ns simulation. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

(e)



(f)



(g)

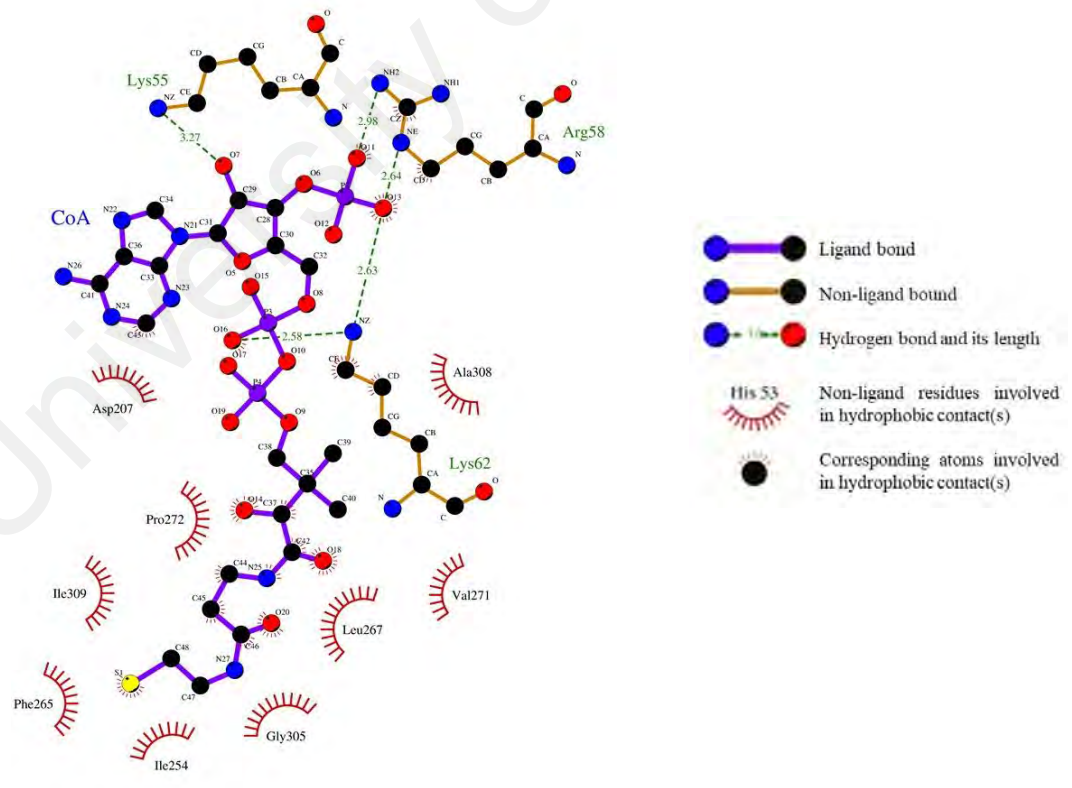


Figure 4.17, continued.

Binding free energies of ligand bound *BrCHSv4* complexes are shown in Table 4.21. *p*-Coumaroyl-CoA recorded the most negative binding free energy among the substrate ligands. It is followed by cinnamoyl-CoA and malonyl-CoA with binding free energies of -94.321 ± 4.319 kJ/mol and -88.861 ± 4.181 kJ/mol, respectively. Thus, *p*-coumaroyl-CoA has higher binding affinity for *BrCHSv4* than cinnamoyl-CoA.

Table 4.21: Binding free energy of *BrCHS* variant 4 receptor with the ligands.

Ligand	Overall Coulombic (kJ/mol)	Overall Lennard-Jones (kJ/mol)	Polar solvation (kJ/mol)	Non-polar solvation (kJ/mol)	Binding free energy (kJ/mol)
Malonyl- CoA	-485.831 ± 13.362	-133.279 ± 2.414	554.283 ± 10.338	-24.034 ± 0.100	-88.861 ± 4.181
Cinnamoyl-CoA	-325.135 ± 18.911	-159.894 ± 1.861	414.344 ± 15.676	-23.637 ± 0.078	-94.321 ± 4.319
<i>p</i>-Coumaroyl-CoA	-1011.086 ± 21.150	-153.084 ± 1.984	959.235 ± 18.150	-23.631 ± 0.092	-228.567 ± 5.018
Caffeoyl-CoA	-180.722 ± 14.371	-142.166 ± 1.871	282.113 ± 12.410	-22.372 ± 0.088	-63.147 ± 3.004
Feruloyl-CoA	28.796 ± 11.398	-163.658 ± 1.623	81.561 ± 9.274	-22.852 ± 0.068	-76.153 ± 3.227
Acetyl-CoA (Reference)	-631.249 ± 14.282	-133.325 ± 2.124	685.591 ± 11.179	-21.969 ± 0.059	-100.951 ± 3.953
CoA (Reference)	-467.968 ± 14.810	-176.822 ± 2.107	550.994 ± 12.330	-26.172 ± 0.100	-119.967 ± 4.225

Table 4.22 shows the residues of *BrCHSv4* formed interaction with the ligands. *p*-Coumaroyl-CoA formed more van der Waals interactions with the receptor. π - π stacked interaction was found between *p*-coumaroyl-CoA and Phe265 residue of the receptor. The hydrogen bonds and non-hydrophobic interactions of the *BrCHSv4* complexes are depicted in Figure 4.18. Hydrogen bonding was observed between Arg58 and Lys62 residues of the receptor and ligand. *p*-Coumaroyl-CoA formed a hydrogen bond with the Cys164 residue of the *BrCHSv4* receptor.

Table 4.22: List of interactions formed between *BrCHSv4* and the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Glu51 (2.91) Arg58 (2.70, 2.72) Lys62 (2.54, 2.72)	Leu214, Phe215, Phe265, Leu267, Gly305	Glu54 (Amide- π stacked) Lys55 (π -alkyl)	Lys55 Arg58 Lys62	-
Cinnamoyl-CoA	Lys55 (2.98) Arg58 (2.62, 2.76) Lys62 (2.58)	Val210, Phe215, Phe265, Gly305, Gly306, Pro307	Leu214 (π -alkyl) Lys55, Met59 (alkyl) Cys164 (π - sulfur)	Lys55 Arg58 Lys62	-
<i>p</i>-Coumaroyl-CoA	Lys55 (2.75) Arg58 (2.67, 2.95, 3.17) Lys62 (2.64, 2.92) Cys164 (3.05)	Gly163, Leu214, , Leu271, His303, Gly305, Gly306, Ile309, Phe373, Gly374	Phe265 (π - π stacked) Pro375 (π -alkyl)	Lys55 Arg58 Lys62	-
Caffeoyl-CoA	Arg58 (2.61, 2.67) Lys62 (2.66, 2.70)	Asp207, Val210, Gly211, Leu214, Phe215, Leu267	Phe265 (π - π stacked) Lys55, Met59 (Alkyl)	Arg58 Lys62	-
Feruloyl-CoA	Arg58 (2.61) Lys62 (2.67, 2.73, 3.00)	Asp207, Val210, Gly211, Phe215, Ile254, Phe265, Gly305, Gly306, Pro307	Met64 (Alkyl) Leu214, Leu267 (π -alkyl)	Arg58 Lys62	-
Acetyl-CoA (Reference)	Lys55 (2.60) Arg58 (2.57, 2.68) Lys62 (2.66, 2.83)	Asp207, Val210, Leu214, Phe215, Leu267, Gly305, Gly306, Ala308	Glu54 (π -anion) Arg58 (π -cation)	Lys55 Arg58 Lys62	-
CoA (Reference)	Arg58 (2.57, 2.64), Lys62 (2.54, 2.86, 3.23)	Gly211, Leu214, Phe215, Gly305	Phe265 (π - sigma) Lys55 (π -alkyl, π -cation) Ala308 (Alkyl)	Arg58 Lys62	Arg58

-: Absence

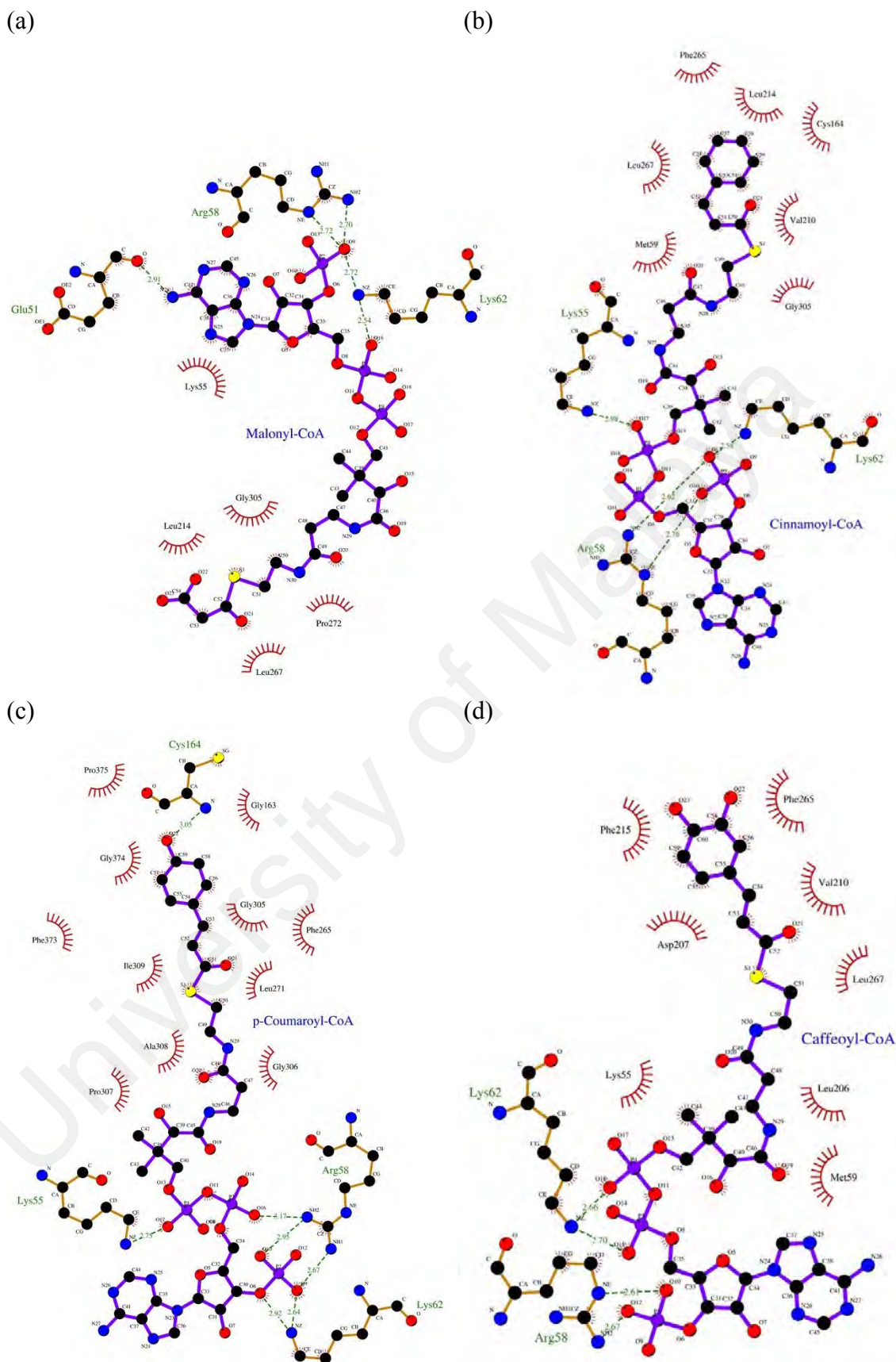
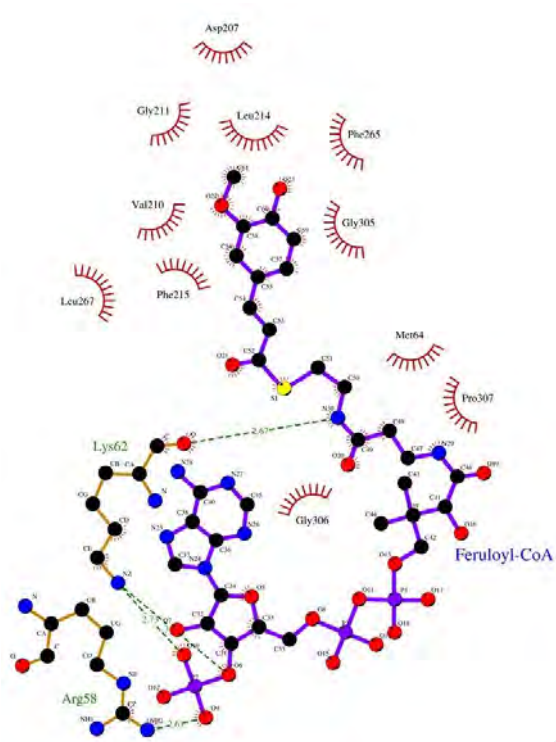
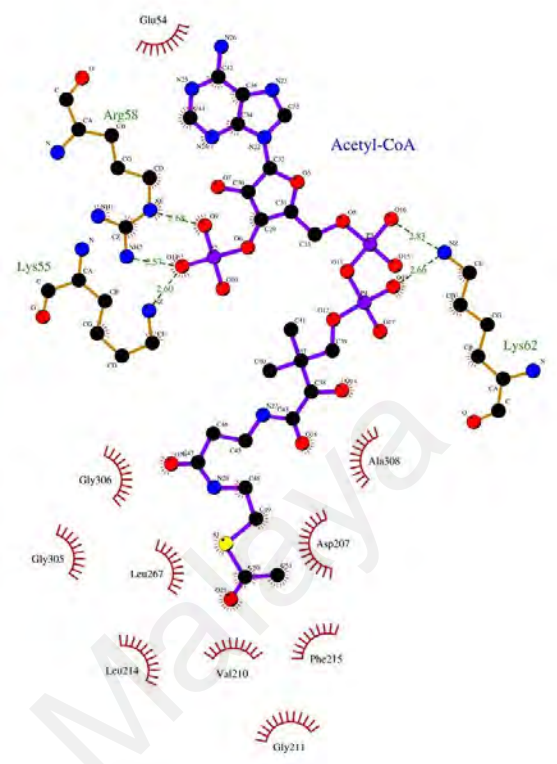


Figure 4.18: Interactions of *BrCHSv4* with the ligands after 10 ns simulation. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

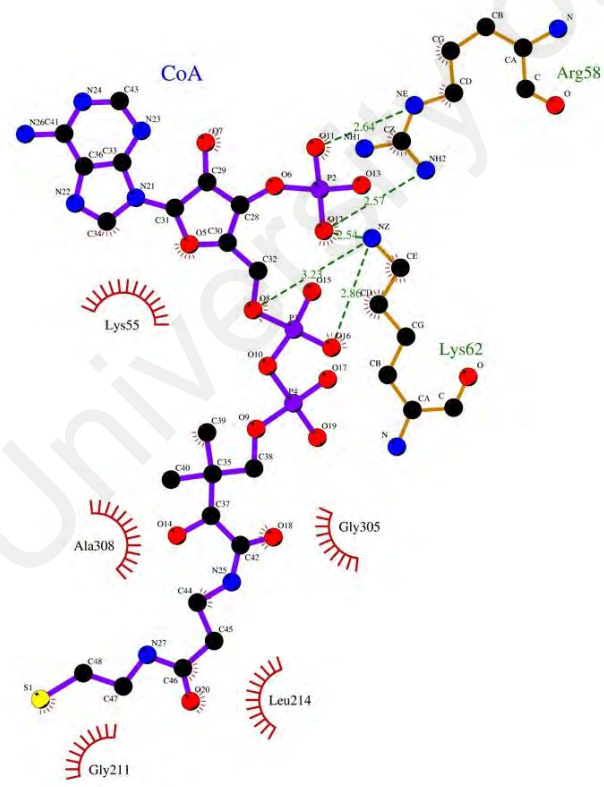
(e)



(f)



(g)








-  Ligand bond
-  Non-ligand bond
-  Hydrogen bond and its length
-  His 53 Non-ligand residues involved in hydrophobic contact(s)
-  Corresponding atoms involved in hydrophobic contact(s)

Figure 4.18, continued.

Table 4.23 shows the binding free energy profile of *BrCHSv5* with substrate ligands. Cinnamoyl-CoA recorded the most negative binding free energy among substrate ligands which is -290.644 ± 7.159 kJ/mol. *p*-Coumaroyl-CoA recorded more negative binding free energy than malonyl-CoA. Hence, *BrCHSv5* has a higher binding affinity towards cinnamoyl-CoA than *p*-coumaroyl-CoA.

Table 4.23: Binding free energy of *BrCHS* variant 5 receptor with the ligands.

Ligand	Overall Coulombic (kJ/mol)	Overall Lennard-Jones (kJ/mol)	Polar solvation (kJ/mol)	Non-polar solvation (kJ/mol)	Binding free energy (kJ/mol)
Malonyl- CoA	-463.423 ± 23.804	-96.035 ± 2.129	448.403 ± 20.817	-19.779 ± 0.166	-130.834 ± 4.844
Cinnamoyl-CoA	-1210.530 ± 32.212	-105.369 ± 2.533	1045.975 ± 26.411	-20.719 ± 0.163	-290.644 ± 7.159
<i>p</i>-Coumaroyl-CoA	-207.996 ± 13.829	-149.704 ± 1.799	233.969 ± 11.001	-24.803 ± 0.067	-148.534 ± 3.259
Caffeoyl-CoA	-256.898 ± 16.723	-151.988 ± 2.296	300.515 ± 13.017	-22.168 ± 0.123	-130.539 ± 4.426
Feruloyl-CoA	490.854 ± 11.320	-131.815 ± 2.233	-289.898 ± 9.582	-16.720 ± 0.263	52.421 ± 3.421
Acetyl-CoA (Reference)	-869.965 ± 24.525	-106.107 ± 2.467	791.678 ± 22.479	-20.930 ± 0.145	-205.324 ± 4.881
CoA (Reference)	-1640.212 ± 17.981	-89.714 ± 2.459	1372.206 ± 13.672	-20.738 ± 0.055	-378.458 ± 4.673

Table 4.24 shows the residues of *BrCHSv5* interacted with the ligands. Cinnamoyl-CoA formed more salt bridge interactions with the receptor compared to *p*-coumaroyl-CoA. Malonyl-CoA formed a repulsive interaction with the Arg58 residue of the receptor. Hydrogen bonding between ligands and the *BrCHSv5* receptor is shown in Figure 4.19. *p*-Coumaroyl-CoA formed fewer hydrogen bonds with the receptor compared to cinnamoyl-CoA.

Table 4.24: List of interactions formed between *BrCHSv5* and the ligands.

Ligand	Type of interaction				
	Hydrogen bond (length, Å)	van der Waals	π -interaction	Salt bridge	Repulsive interaction
Malonyl-CoA	Arg58 (2.61, 2.77, 3.18) Lys62 (2.55)	Leu214, Leu267, Val271, Gly305, Ala308	Lys62 (Alkyl)	Lys55 Arg58 Lys62	Arg58
Cinnamoyl-CoA	Lys55 (2.57, 2.61) Arg58 (2.58, 2.90, 3.27) Lys62 (2.69) Asp207 (2.83)	Cys164, Phe165, Leu206, Val210, Lys269, Gly305	Leu267, Val271, Pro272 (π -alkyl)	Lys55 Arg58 Lys62	-
<i>p</i>-Coumaroyl-CoA	Lys55 (2.67) Arg58 (2.77, 2.85) Lys62 (2.83)	Cys164, Phe165, Val210, Leu214, Phe215, Ile254, Phe265, Pro272, His303, Gly305, Gly306	Arg58, Met59 (Alkyl) Val271 (π -alkyl)	Lys55 Arg58	-
Caffeoyl-CoA	Lys55 (2.65) Arg58 (2.77, 3.23) Lys62 (2.60, 2.60)	Leu206, Asp207, Gly211, Leu214, Phe215, Phe265,	Met59, Val210 (π -alkyl)	Lys55 Arg58 Lys62	-
Feruloyl-CoA	Asp136 (3.15) Gln161 (2.90) Arg259 (2.65, 2.80)	Val135, Met137, Ile254, Gly256	Tyr160 (π - π T-shaped) Phe265, Pro375 (Alkyl) Leu258, Leu263 (π -alkyl)	Arg259	-
Acetyl-CoA (Reference)	Arg58 (2.73, 3.15) Lys62 (2.55, 3.18) Asp207 (2.79, 3.02)	Ile254, Phe265, Ile267, Gly376, Leu377	Met59, Lys62, Val210 (Alkyl)	Lys55 Arg58 Lys62	-
CoA (Reference)	Lys55 (2.56, 2.66) Arg58 (2.67, 2.80) Lys62 (2.68, 2.70)	Cys164, Leu214, Phe265, Gly305, Gly306, Pro375	Phe215 (π -sulfur) Val210, Ala308 (Alkyl)	Lys55 Arg58 Lys62	-

-: Absence

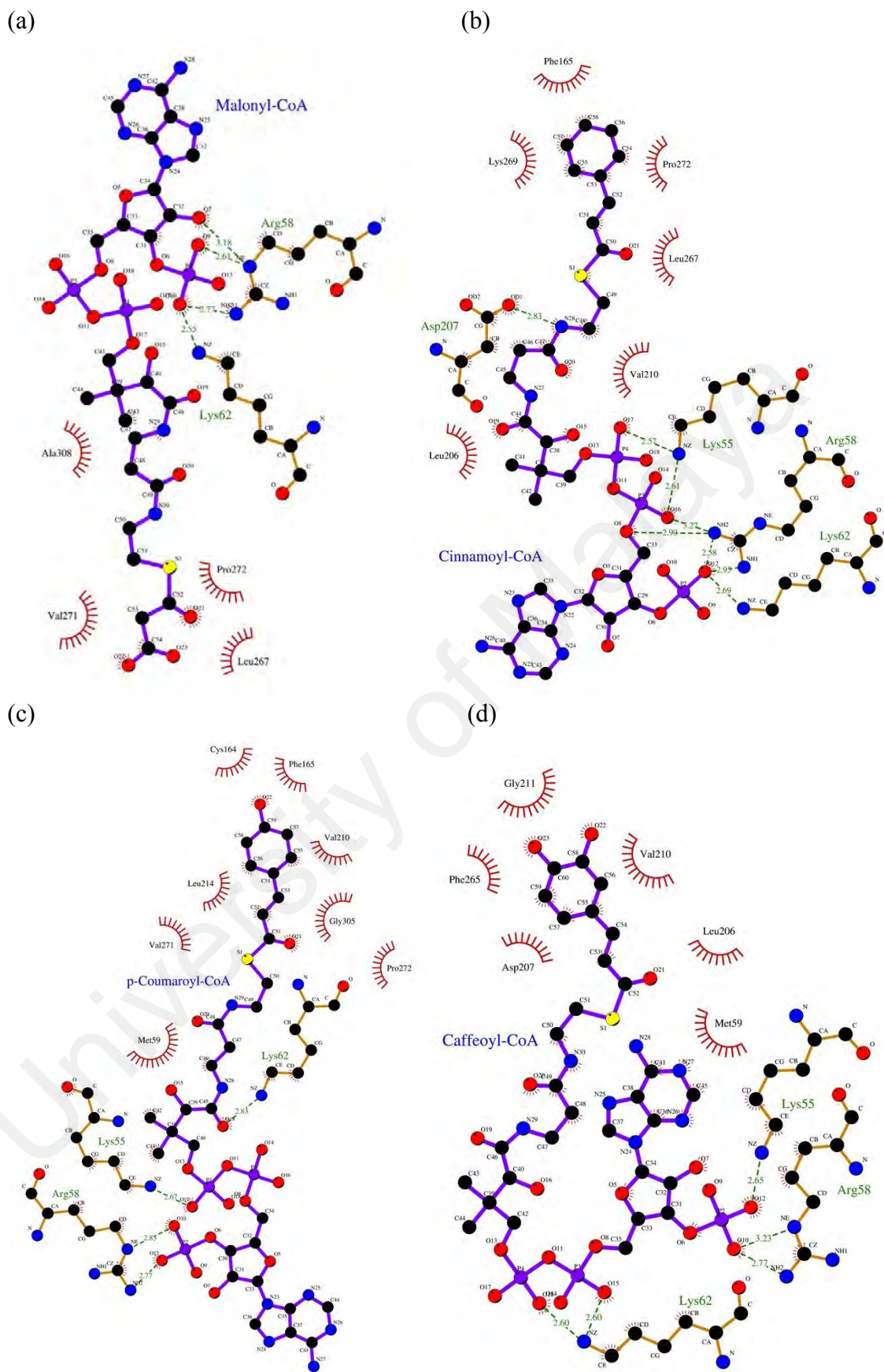
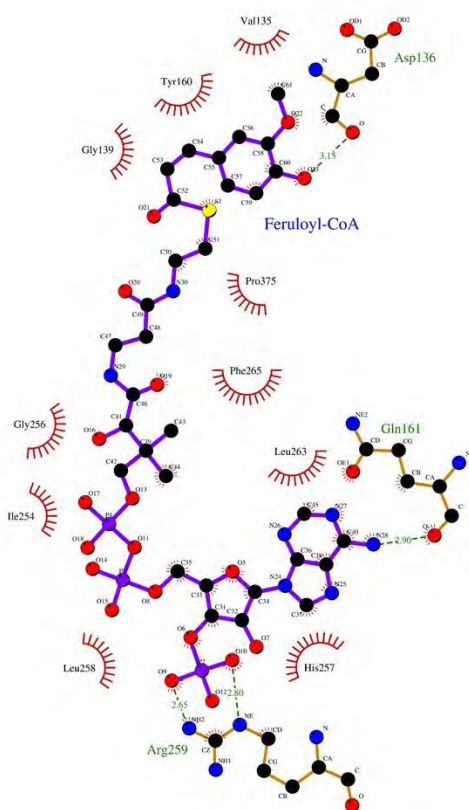
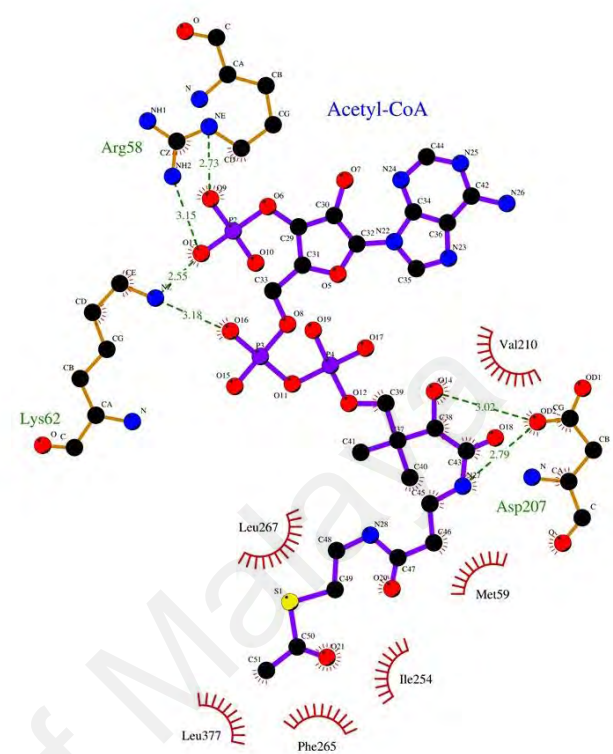


Figure 4.19: Interactions of *BrCHSv5* with the ligands after 10 ns simulation. (a) Malonyl-CoA; (b) cinnamoyl-CoA; (c) *p*-coumaroyl-CoA; (d) caffeoyl-CoA; (e) feruloyl-CoA; (f) acetyl-CoA; (g) CoA.

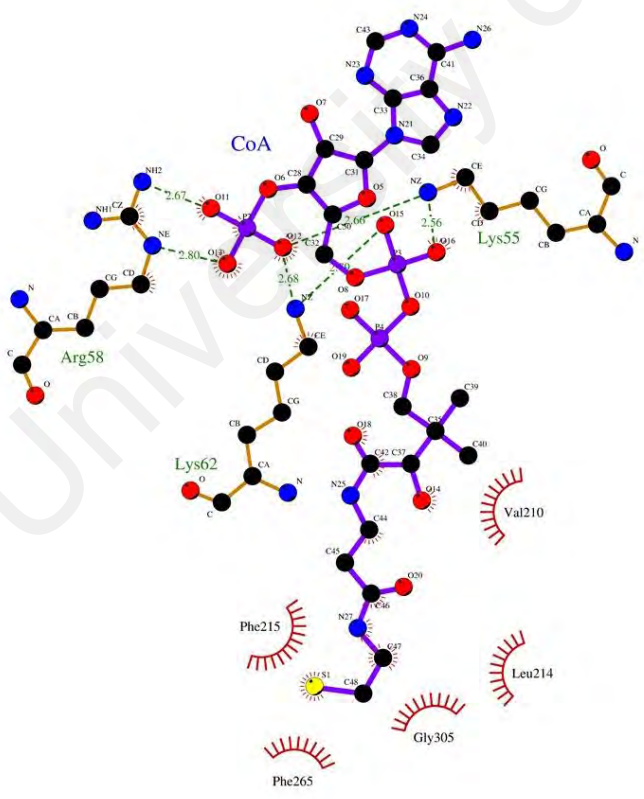
(e)



(f)



(g)



- Ligand bond
- Non-ligand bond
- Hydrogen bond and its length
- Non-ligand residues involved in hydrophobic contact(s)
- Corresponding atoms involved in hydrophobic contact(s)

Figure 4.19, continued.

CHAPTER 5: DISCUSSION

5.1 Molecular Modelling

5.1.1 Homology Modelling

Chalcone synthase (CHS) is known as the simplest enzyme in the family of type III polyketide synthase (PKS) in terms of its structure and mechanism (Austin & Noel, 2002). It is the key enzyme in the initial stage flavonoid biosynthetic pathway. The three-dimensional structures of the five variants of *BrCHS* show that it shared similar homodimeric structures. Stilbene synthase (STS), an enzyme of PKS type III family also shares similar homodimeric structure with CHS (Tropf et al., 1995). The N-terminal helix of both monomers A and B entwines each other. The *cis*-peptide bond between Met137 and Pro138 caused Met137 to become as a knob on the surface and it forms the cyclization pocket by protruding into a hole at the surface of the adjacent monomer (Ferrer et al., 1999).

The predicted molecular weights of *BrCHS* receptor variants are within the range of 42 to 44 kDa. The result shows a good agreement that fell within the protein size range of 40 to 45 kDa for CHS from different plant species (Abe & Morita, 2010). For instance, the predicted molecular masses of CHS from *Curcuma longa* and *Lamiophlomis rotata* 42.75 kDa and 42.80 kDa, respectively (Wannapinpong et al., 2013; Q. Feng et al., 2015). Secondary structure analysis revealed that the theoretical isoelectric points for all the *BrCHS* variants were similar except for *BrCHSv3*. *BrCHSv3* recorded slightly higher isoelectric point which is 6.53 due to the presence of a higher amount of positively-charged amino acid residues of arginine and lysine compared to other variants. The computed pI which is less than 7 indicates these protein receptors were acidic in nature.

The secondary structure of homology models is mainly made of α -helices, random coils and β -strands. The former two structures comprise a higher percentage of extended strand and β -turns among the *BrCHS* receptor variants. The results show similar structural compositions with the other CHSs from *Scutellaria viscidula* (*SvCHS*) and *Grewia asiatica* (*GaCHS2*). The percentage for α -helices and random coils in the secondary structure of *SvCHS* are 43.6% and 37.4%, respectively (Lei et al., 2010). Meanwhile, the secondary structure of *GaCHS2* consisted of 42.16% of α -helices and 30.08% of random coils, respectively (Wani et al., 2017). The high percentage of α -helices may be due to the presence of amino acid residues such as alanine, phenylalanine, leucine and isoleucine (Mbah et al., 2012).

Ramachandran plot analysis revealed that the homology models of five *BrCHS* receptor variants contain more than 90% of residues in the most favorable region. (Morris et al., 1992) stated that a protein model with 90% over residues present in the most favourable region is considered as a good quality model. The results show that all the homology models were found to be of good quality. Only *BrCHSv5* contained 0.3% residue numbers in the outlier region which is Gly362.

Energy minimization was carried out in order to reduce steric collisions and strains without any significant alteration in the overall structure of the receptor models (Messaoudi et al., 2013). The refined models were assessed using the Ramachandran plot. The results revealed that all the models consist of more than 95% of residues in the favored region. Only *BrCHSv3* does not contain any residues in the outlier region. *BrCHSv5* contain the same residue in the outlier region before and after energy minimization which is Gly362. Residues Pro304 and Ser90 were found in the outlier region in the *BrCHS* receptor variants 1 and 4, respectively. Meanwhile, *BrCHSv2* contains residues Asp270 and Pro304 in the outlier region. The residues Ser90, Asp270,

Gly362 which is in the outlier regions were located far from the binding site domains in the receptors and will not affect the binding of the ligands towards the receptors. Whereas, Pro304 is involved in the shaping of active site geometry of the receptors (Ferrer et al., 1999). It may affect the interaction of ligand at the binding site of the protein. All the post-minimized models of *BrCHS* receptor variants have good stereochemical quality and reliable. Therefore, these models were chosen for molecular docking.

5.2 Multiple Sequence Alignment (MSA)

Information on the active site of a protein receptor can be obtained by comparing the target protein with proteins in the same family that share similar functionality (Meng et al., 2011). Multiple sequence alignment showed that three main amino acid residues (Cys164, His303 and Asn336) are conserved in *BrCHS* receptor variants as well as in the CHSs of other plants (Abe & Morita, 2010). The abovementioned amino acid residues form a catalytic triad in the binding site domain of *BrCHS* receptor variants. Residues Asp311 and Glu314 of the receptor interact with Arg328 that help to position residues Cys164, His303 and Asn336 in a correct topology at the active site (Fukuma et al., 2007). Alfalfa CHS2 also contain similar conserved active site residues as in *BrCHS*s. Besides that, the catalytic triad residues also highly conserved in other CHS-like enzymes, for instance, stilbene synthase and bibenzyl synthase (Jez & Noel, 2000). This catalytic triad plays a vital role in the catalytic mechanism of CHS enzyme.

In general, the active site of CHS comprises three interconnected cavities that intersect with the catalytic triad residues. The three cavities are the CoA-binding tunnel, coumaroyl-binding pocket and cyclization pocket (Ferrer et al., 1999). The CoA-binding tunnel is connected to the protein surface that allows the contact of a substrate to the buried active site (Morita et al., 2010). On the other hand, coumaroyl-binding pocket comprises five residues which are Ser133, Glu192, Thr194, Thr197 and Ser338. The

residues showed a high degree of conservation in the multiple sequence alignment. The coumaroyl-binding pocket is found in the CoA-binding tunnel at its lower left space (Ferrer et al., 1999). Seven residues (Thr132, Met137, Phe215, Ile254, Gly256, Phe265 and Gly375) form the cyclization pocket that assists in the elongation of polyketide during the catalytic mechanism (Ferrer et al., 1999).

There are few amino acid residues showing strong conservation in all CHS-like enzymes. The residues (Pro138, Gly163, Gly167, Leu214, Asp217, Gly262, Pro304, Gly305, Gly306, Gly335, Gly374, Pro375 and Gly376) involve in the formation of the active site geometry in CHS (Ferrer et al., 1999). These residues were found in the *BrCHS* receptor variants as well that shapes the geometry of the active site. The residues, Gly163 and Phe165 are highly conserved and surrounding one of the catalytic triad residues, Cys164 (Mallika et al., 2011).

The GFGPG loop is a strictly conserved region that presents in all the *BrCHS* receptor variants. The GFGPG loop consists of residues Gly372, Phe373, Gly374, Pro375 and Gly376. It is also a signature loop not only strictly conserved in CHS but also in other enzymes in the CHS superfamily. Other than that, the loop was also detectable in all non-CHS plants (Mallika et al., 2011). The role of the loop might be different among the different enzymes in the family though it is conserved (Suh et al., 2000). The signature loop is found at the sidewall of the active site of the enzyme. Ferrer et al. (1999) reported that GFGPG loop is a part of the sidewall of alfalfa CHS's active site and plays a crucial role during cyclization reactions. The ionic or hydrogen bond between Glu380 and Arg172 maintains the configuration of the loop relative to one of the catalytic triad residues, Cys164 (Fukuma et al., 2007).

Two phenylalanine residues, Phe215 and Phe265 also were found as conserved residues in the multiple sequence alignment along with CHS from other plant species.

Both residues are called “gatekeepers” because they were situated at the entrance of the active site (Jez et al., 2002; Abe et al., 2003). During polyketide elongation, the gatekeepers serve as mobile steric gates that separate CoA-binding tunnel from the cyclization pocket by inhibiting the lower portion of the opening between the two sites (Ferrer et al., 1999; Morita et al., 2010). Fukuma et al. (2007) reported that two highly conserved residues Arg68 and Glu33 help in positioning Phe215 in the correct configuration at the active site.

There are several residues that are different among the *BrCHS* receptor variants. The most prominent changes were seen in the *BrCHSv3* and *BrCHSv5* receptors. *BrCHSv3* contains two residues Tyr194 and Gly256 which are replaced with Arg194 and Ala256, respectively. The residue 194 of the receptor is part of CoA binding pocket and was replaced with more bulkier and positively charged residue, arginine. Whereas, changes of Gly256 to Ala256 which is slightly bulkier than glycine residue might affect the cyclization reaction of the variant receptor. On the other hand, the Thr132 residue of *BrCHSv5* is replaced with Ser132 which is less bulky with the methyl group. The Ser132 residue is a part of the cyclization pocket and might affect the elongation of polyketide intermediates.

5.3 Docked Conformation Analysis

Molecular docking is the most widely used method for modelling and calculation of the interactions between ligand and protein (Sousa et al., 2013). The docking efficiency can be increased when the location of the active site is known prior to docking processes (Meng et al., 2011). The docking is a quick method for predicting the most favorable structure of the protein-ligand complex and evaluating its binding affinity (Du et al., 2016; Chandel et al., 2018).

CHS belongs to type III polyketide synthase that produces diverse products by catalyzing different starter molecules. It catalyzes both CoA-linked aliphatic and aromatic starter molecules. For instance, CHS from parsley accepts *p*-coumaroyl-CoA as well as hexanoyl-CoA, butyryl-CoA and benzoyl-CoA as the substrate molecules (Schüz et al., 1983). In addition, CHS from *Scutellaria baicalensis* produced various polyketides after using isobutyryl-CoA, isovaleryl-CoA and phenylacetyl-CoA as starter molecules (Morita et al., 2000). It also used methylmalonyl-CoA as extender substrate along with the *p*-coumaroyl-CoA as a substrate to yield unnatural aromatic polyketide (Abe et al., 2002). The substrate specificity of CHSs is influenced by relatively small changes in the protein structures and alters its condensation reaction (Eckermann et al., 1998). Therefore, this study is conducted mainly to investigate the substrate preference of CHS of *Boesenbergia rotunda* whether *p*-coumaroyl-CoA or cinnamoyl-CoA.

Cinnamoyl-CoA, *p*-coumaroyl-CoA, feruloyl-CoA and caffeoyl-CoA are the CoA esters that involved in the phenylpropanoid pathway to produce phenylpropanoids (Dixon et al., 2002). Besides that, the ligands also are preferred substrates for CHS in some plants (Christensen et al., 1998). Therefore, these ligands were chosen as substrates for the *BrCHS* variant receptors. In addition, malonyl-CoA was selected as a substrate because it is an extender molecule that plays a vital role in the mechanism of CHS.

Docking results revealed that the five *BrCHS* receptor variants preferred different substrate ligands. The binding affinity of the substrate was determined by the docked energies. The more negative the docked energy of docked complexes, the higher its binding affinity. Docked conformation analysis revealed the interactions formed between the docked complexes and the residues contributed to the respective interactions. The protein-ligand interactions explain the binding affinity of the substrate molecules towards the receptors. Generally, the interactions between protein and ligand are noncovalent

contacts such as hydrogen bonds and van der Waals contacts (Zhou et al., 2012; Salentin et al., 2014).

p-Coumaroyl-CoA is the main substrate ligand for most of CHS from different plant species (Lanz et al., 1991; Ferrer et al., 1999; Jez & Noel, 2000). The enzyme utilizes one molecule of *p*-coumaroyl-CoA and three molecules of malonyl-CoA and undergoes a series of intramolecular Claisen condensation producing naringenin chalcone (Ferrer et al., 1999; Jez et al., 2000). The naringenin chalcone will be converted into naringenin by chalcone isomerase (CHI) in plants (Dao et al., 2011). Meanwhile, in an *in vitro* environment, naringenin chalcone will be converted into naringenin non-enzymatically (Sun et al., 2015).

p-Coumaroyl-CoA is a more preferred substrate for *BrCHSv4* and *BrCHSv5* compared to cinnamoyl-CoA. This is because of the docked energies of *p*-coumaroyl-CoA are more negative than cinnamoyl-CoA. The binding affinity is contributed by several interactions. *BrCHSv4* formed a hydrogen bond with *p*-coumaroyl-CoA via Cys164 residue. This is similar to another docking study by Awasthi et al. (2016) that a hydrogen bond (2.47 Å) is formed between Cys164 of *CfCHS* and *p*-coumaroyl-CoA. The residues Arg58 and Lys62 are mainly found at the outside of the CoA binding tunnel and located in the lower portion of the tunnel (Jez et al., 2001b). Arg58 and Lys62 residues of *BrCHSv4* and *BrCHSv5* receptors formed salt bridge interactions with *p*-coumaroyl-CoA, respectively and contributed more negative electrostatic energy. Ferrer et al. (1999) also reported that Phe215 involved in van der Waals interaction and Phe265 separates the cyclization pocket from coumaroyl binding pocket. *p*-Coumaroyl-CoA formed van der Waals interaction with Phe215 and Phe265 residues of *BrCHSv5*, respectively. Besides that, Phe265 of *BrCHSv4* formed π - π stacked interaction with *p*-coumaroyl-CoA. It is a special case of hydrophobic contact that is commonly formed by the aromatic ring of a ligand with the

aromatic ring of amino acids such as phenylalanine, tyrosine, tryptophan and histidine (de Freitas & Schapira, 2017). Besides that, *p*-Coumaroyl-CoA formed π -alkyl interaction with residues Ile254 and Pro375 of *BrCHSv4*, respectively. Residues Ile254 and Pro375 of the receptors play a vital role in its cyclization process upon binding of ligands. Repulsive interaction may cause a significant decrease in binding affinity of ligand towards its protein receptor. Gln161 of *BrCHSv2* formed an unfavorable interaction with *p*-coumaroyl-CoA and recorded the least negative electrostatic energy than cinnamoyl-CoA (Sanmugavelan et al., 2018). Thus, it reduced the binding affinity of *p*-coumaroyl-CoA towards *BrCHSv2*.

On the other hand, cinnamoyl-CoA also preferred for CHS as a starter molecule. Cinnamoyl-CoA is different from *p*-coumaroyl-CoA which has an additional hydroxyl group at the position 3. CHS from *Physcomirella patens* accepts cinnamoyl-CoA as a starter molecule and converts it into its corresponding chalcones (Jiang et al., 2006). It produces pinocembrin through an intermediate molecule called pinocembrin chalcone by the action of CHI or non-enzymatically in the *in vitro* environment (Guo et al., 2016). Enzymatic assay and HPLC analysis of an *in vitro* studies revealed that *BrCHSv2* yielded pinocembrin chalcone when malonyl-CoA and cinnamoyl-CoA were added together (Sanmugavelan et al., 2018). Besides that, cinnamoyl-CoA recorded the most negative docked energy of *CfCHS* from *Coleus forskohlii* and showed that it has a higher binding affinity towards the receptor (Awasthi et al., 2016).

BrCHS receptor variants 1 to 3 prefer cinnamoyl-CoA as substrate ligand compared to *p*-coumaroyl-CoA. This is due to docked energies of the ligand is more negative than *p*-coumaroyl-CoA that corresponds to its higher binding affinity of cinnamoyl-CoA towards the receptors. Arginine is more likely to form hydrogen bonds than lysine due to its side chains consisting of guanidinium group with three nitrogen atoms (de Freitas & Schapira,

2017). Residue Arg259 of *BrCHSv2* formed hydrogen bonds with cinnamoyl-CoA along with residues Gln161 and Leu258 and responsible for more negative electrostatic energy than *p*-coumaroyl-CoA and increased its binding affinity. Other than that, more hydrogen bonds formed between cinnamoyl-CoA and *BrCHSv3* resulted in a more negative electrostatic energy and further strengthen the binding of the protein-ligand complex. Besides that, Ala308 is a highly conserved residue that forms hydrogen bonds with CoA-linked thioesters via amide nitrogen of its backbone (Ferrer et al., 1999). Residue Ala308 of *BrCHSv1* interact with cinnamoyl-CoA via hydrogen bond that results it recorded more negative docked energy. Arg58, Lys62 and Arg259 are the residues involved in the electrostatic interaction with cinnamoyl-CoA and contributed to its higher electrostatic energies.

Ferrer et al. (1999) reported that van der Waals contact more prevalent interaction between CoA-linked starter molecules and CHS. Phe373, Gly374, Pro375 and Gly376 are the residues forming the GFGPG loop that plays a vital role in the cyclization reaction. Phe373 and Gly374 residues of *BrCHSv3* formed van der Waals contact with cinnamoyl-CoA. In addition, the same ligand formed van der Waals interaction with Pro375 of *BrCHSv1*. Residues of the cyclization pocket which are Thr132, Met137, Phe215, Ile254, Gly256 and Phe265 also formed van der Waals interactions with the cinnamoyl-CoA. Other than that, residues of coumaroyl binding pocket (Arg194 and Ser338) and coumaroyl binding tunnel (Gly305, Gly306, Gly374 and Pro375) involved in the interaction of van der Waals with the ligands. Besides that, residues Pro138, Gly163 and Leu214 that shape the active site geometry of the receptor also formed van der Waals interaction with the receptors. The interaction formed by the residues strengthen the binding of receptor and cinnamoyl-CoA. Cinnamoyl-CoA-*BrCHSv2* complex has π -alkyl interactions with Cys164 and His303 residues which also responsible for its binding affinity.

Malonyl-CoA is an extender substrate involved in the decarboxylation reaction mechanism of *BrCHS* receptor variants. The substrate ligand has recorded the most negative docked energies with all the *BrCHS* receptor variants. The results show that it has a higher binding affinity towards the protein receptor. Malonyl-CoA contact with the catalytic triad residues of *BrCHS* receptor variants via hydrogen bonding. This interaction contributed to the highest docked energy compared to other substrate ligands. Malonyl-CoA formed more hydrogen bonds ranging between 2.5 and 3.2 Å. This shows that bonds with moderate strength are mostly electrostatic and contributed to higher electrostatic energies (Jeffrey, 1997). Hydrogen bonding between malonyl-CoA and Asn336 of all the receptors was observed and support its role in decarboxylation. The reaction occurs independently in the absence of transferring malonyl moiety to Cys164 (Kreuzaler et al., 1978). The side chain amide of the residue formed a hydrogen bond with the carbonyl oxygen of the thioester (Ferrer et al., 1999). The Gly305 residue of *BrCHSv5* formed a hydrogen bond with malonyl-CoA. Malonyl-CoA interacted with Lys58 of *BrCHSv1* via cation- π interaction that results in more negative electrostatic energy.

In a previous study by (Hrazdina et al., 1976) revealed that parsley CHS accepts feruloyl-CoA and produced products such as methylpyrone and tetraketide lactone. CHS2 from *Hordeum vulgare* preferred feruloyl-CoA and caffeoyl-CoA than cinnamoyl-CoA and *p*-coumaroyl-CoA. Both caffeoyl-CoA and feruloyl-CoA converted into respective chalcone intermediate products at a higher rate (Christensen et al., 1998). The small molecules are different from cinnamoyl-CoA in term of substitutions in its ring structure. Caffeoyl-CoA contains two hydroxyl groups at the 3 and 4 positions. Meanwhile, feruloyl-CoA has similar substitutions as in caffeoyl-CoA, but its 3 position is O-methylated (Austin & Noel, 2002). The substitutions made the ligands become bulkier than cinnamoyl-CoA and *p*-coumaroyl-CoA. It might affect the binding affinity of the caffeoyl-CoA and feruloyl-CoA with the receptors.

Feruloyl-CoA is a more favourable ligand towards *BrCHSv1* and *BrCHSv5* than cinnamoyl-CoA and *p*-Coumaroyl-CoA, respectively because of its more negative docked energies. A high number of hydrogen bonds formed between feruloyl-CoA and the receptors contributed to a more negative electrostatic energy and resulted in higher docked energies than cinnamoyl-CoA. The residue Gly256 of *BrCHSv5* formed an unfavorable bond with feruloyl-CoA and affects its binding affinity. Hence, it recorded less negative docked energy than malonyl-CoA. Arg259 is the only one residue of the receptors formed a salt bridge with feruloyl-CoA and contributed to a more negative electrostatic energy. Caffeoyl-CoA recorded more negative docked energy with the receptors of *BrCHSv1*, *BrCHSv2* and *BrCHSv3* than *p*-coumaroyl-CoA. Salt bridge formed by residues Arg58 and Lys62 of *BrCHSv2* with caffeoyl-CoA contributed to a more negative electrostatic energy. Whereas, caffeoyl-CoA formed more hydrogen bonds with *BrCHSv3* than *p*-coumaroyl-CoA. These interactions have increased the binding affinity of the receptors for caffeoyl-CoA than *p*-coumaroyl-CoA.

Acetyl-CoA and CoA were used as reference ligands for molecular docking. Both acetyl-CoA and CoA are aliphatic small molecules. Acetyl-CoA and CoA were chosen as reference ligands as both shared basic chemical structure with the substrate ligands. CHS2 of *Gerbera hybrida* used acetyl-CoA as starter molecule and catalyze further condensation with only two molecules of malonyl-CoA. The reaction has yielded methylpyrone (Eckermann et al., 1998). Cation- π interaction is an electrostatic interaction because of the presence of an electron cloud of π systems that is negatively charged (Mahadevi & Sastry, 2013). Based on the docking results, Lys55 residue of *BrCSHv2* formed cation- π interactions with acetyl-CoA and contributed to a more negative electrostatic energy. The catalytic triad residues of the receptors formed the hydrogen bonds with the reference ligands which is in the range of 2.5 to 3.2 Å. The hydrogen bonds contributed a more negative docked energy. Besides that, hydrogen bonds also

were formed between Arg58 and Lys62 residues of *BrCHSv5* and reference ligands. The Gly305 residue of *BrCHSv5* formed a hydrogen bond with CoA. The hydrogen bonds and salt bridge interactions strengthened the binding between receptors and the reference ligands.

5.4 Molecular Dynamics Simulation

Molecular dynamics (MD) simulations were performed using the docked complexes to elucidate the dynamic behavior of the ligands against the five *BrCHS* receptor variants. The simulations take account of the effect of solvent and protein flexibility which is lacking in the molecular docking procedure (Xie et al., 2015). MD simulations of 10 ns were performed with the docked complexes. A long duration of MD simulations is not always necessary to achieve better predictions (Hou et al., 2011). Therefore, a short period of MD simulations was chosen. Dodecahedron shape was chosen because it is close to being a spherical macromolecule that requires fewer solvent molecules to fill the periodic box given and saves 29% of the central processing unit (CPU) time during simulation (Abraham et al., 2015).

5.4.1 Trajectory Analysis

The protein stability relative to its conformation is measured by means of RMSD. It is determined by the deviations produced by the protein throughout the simulation (Aier et al., 2016). Smaller RMSD value indicates that the protein-ligand structure is less deviated from its initial structure and has greater stability during the entire MD simulation (Rohini & Srikumar, 2013; Zhao et al., 2015). Most of the *BrCHS* variant complexes converged towards the end of simulation with smaller fluctuations. This indicates that the complexes were stable during the course of the simulation. *BrCHSv3*-feruloyl-CoA complex showed higher RMSD values compared to other ligands with the *BrCHSv3* receptor. This shows

that the complex is more stable. *BrCHSv2*-Malonyl-CoA complex showed high fluctuations towards the last 5 ns of the simulations and indicates that the complex is not stable. *BrCHSv4*-cinnamoyl-CoA complex also not stable as showed higher fluctuations towards the end of the simulation.

RMSF is a measure for assessing the residues involved in the structural fluctuations of the complex during the simulation (Liao et al., 2014). Residues with high fluctuations in RMSF show high flexibility that reflects the unstable interactions (Firdayani et al., 2018). In the *BrCHSv1*-*p*-coumaroyl-CoA complex, residue Pro138 showed a high level of fluctuation. High level of fluctuation also observed in the residue Pro375 of both *BrCHSv4*-cinnamoyl-CoA and *BrCHSv5*-*p*-coumaroyl-CoA complex, respectively. The residues Pro138 and Pro375 are the key residues that form the active site geometry of the receptor. This will eventually affect the binding of the ligand to the active site of the receptor. On the other hand, well-structured regions with limited movements are reflected by the low RMSF value. The catalytic triad residues showed a low RMSF value which is lower than 0.2 nm. It indicates that the residues are distorted minimally upon binding of the ligands during the simulations.

The radius of gyration is a characteristic that assesses the compactness changes of ligand bound protein complexes (Liao et al., 2014). The radius of gyration profiles for ligand-bound complexes of *BrCHSv2* and *BrCHSv3* showed no significant changes in the fluctuations. Besides that, *BrCHSv5* complexes showed a similar trend of fluctuation except for cinnamoyl-CoA-*BrCHSv5*. Thus, the compactness of the complexes was not changed and the complexes were stable throughout the MD simulations. *BrCHSv1* and *BrCHSv4* complexes showed a decrease in the radius of gyration towards the end of MD simulations. The same trend was observed in the cinnamoyl-CoA-*BrCHSv5* complex. The compactness of ligand-protein complexes is higher when the radius of gyration is lower

and strengthen the interactions between ligand and receptor (Sivashanmugam et al., 2013; Liao et al., 2014).

5.4.1.1 Binding Free Energy

Calculation of binding free energy based on molecular dynamics simulation is the potential approach that provides an accurate estimation of protein-ligand binding affinities. MM-PBSA is an effective method to calculate the binding free energy of a receptor-ligand complex (Wang et al., 2013). The approach enhances understanding of binding affinity contributed by the electrostatic and van der Waals interactions along with changes in entropy and solvation (Wang et al., 2016). In addition, MM-PBSA also a useful method for rationalizing the observed differences between docked and post-simulation structures (Genheden & Ryde, 2015). The single-trajectory method was used for the calculation of MM-PBSA using GMXPBSA 2.1 tool (Paissoni et al., 2014). This method can decrease the noise significantly and cancel errors as the internal energy can be cancelled between the receptor, ligand and protein-ligand complex. Besides that, the computational cost of this method is less expensive compared to the multiple trajectories method (Hou et al., 2011).

Electrostatics and van der Waals interactions are the major contributors to the interaction energies of the protein-ligand complexes. Coulomb and Lennard-Jones terms were used to calculate the electrostatic and van der Waals interactions, respectively (Meng et al., 2011). The van der Waals energy describes the protein steric interaction and shape complementarity between the ligand and the protein. On the other hand, Coulomb energy accounts for the electrostatic interactions within the protein in the bound and unbound states (Gräter et al., 2005). The binding free energy with negative value reflects the spontaneous protein-ligand binding process (Du et al., 2016). The more negative the binding free energy, the greater the binding affinity of ligands to its protein receptor. The

binding affinity of the ligands with receptors can be explained further in terms of its interactions. The protein-ligand interactions contribute to its binding free energy significantly (Saini, 2017).

Cinnamoyl-CoA recorded the most negative binding free energy with *BrCHSv1*, *BrCHSv2* and *BrCHSv3* compared to other substrate ligands. It showed that it is a more preferred substrate for the above-mentioned receptors. It shows a good agreement with the docking results that the receptor variants have a higher binding affinity towards cinnamoyl-CoA. The Coulomb energy is found to be the main energy term favoring binding of cinnamoyl-CoA to the receptors. Hydrogen bonds with distance ranges of 2.2 – 2.5 Å and 2.5 – 3.2 Å are categorised as “strong, mostly covalent” and “moderate, mostly electrostatic,” respectively (Jeffrey, 1997). More hydrogen bonds ranging from 2.60 Å to 3.17 Å were formed within the protein-ligand complex. For instance, Gly306 of *BrCHSv1* and Gln51 of *BrCHSv3* formed hydrogen bonds with cinnamoyl-CoA, respectively. On the other hand, cinnamoyl-CoA formed a hydrogen bond with Gln51 of *BrCHSv3*. Residue Arg58 of *BrCHSv3* formed a strong hydrogen bond (2.48 Å) with cinnamoyl-CoA. The hydrogen bonds along with the salt bridge interactions also have contributed to a more negative Coulomb energy. The van der Waals interactions also play a role in the stability of protein-ligand complex and enhanced by the presence of stronger interactions for instance hydrogen bonds (Vladilo & Hassanali, 2018). A high number of van der Waals interactions formed between cinnamoyl-CoA and the receptors. It has increased the binding affinity of the receptors towards cinnamoyl-CoA. For instance, residues Leu214, Pro305, Gly374 and Pro375 of the receptors (*BrCHSv1*, *BrCHSv2*, *BrCHSv3*) that shape the active site geometry have formed van der Waals interaction with the ligand and further strengthen its binding. Awasthi et al. (2016) reported that Phe266 of *CfCHS* interacted with cinnamoyl-CoA via π - π stacked interaction. The same

interaction was observed between the residue Phe215 of *BrCHSv3* receptor and cinnamoyl-CoA and increased the binding affinity.

On the other hand, *p*-coumaroyl-CoA recorded the most negative binding free energy with *BrCHSv4*. This supports the docking result that *p*-coumaroyl-CoA is the preferred substrate for the *BrCHSv4*. *p*-Coumaroyl-CoA formed hydrogen bonds with the receptor in the range of 2.64 Å to 3.05 Å which strengthen the binding. High numbers of hydrogen bonds and salt bridge formed between the receptor and ligand have contributed to more negative Coulomb energy and resulted in a higher binding affinity of the receptor for *p*-coumaroyl-CoA than cinnamoyl-CoA. Moreover, the ligand has retained the hydrogen bond with the Cys164 residue of *BrCHSv4* as in the docked structure. Unfavorable bond was formed between Glu54 of *BrCHSv2* and *p*-coumaroyl-CoA. The repulsion interaction decreased the Coulomb energy and binding free energies of the complex became less negative. Hence, the ligands became a less preferred substrate for the *BrCHSv2* receptor.

In contrast to the docking result, cinnamoyl-CoA has a higher binding affinity towards *BrCHSv5* than *p*-coumaroyl-CoA. The ligand initially recorded less negative docked energy than *p*-coumaroyl-CoA-*BrCHSv5* complex. The radius of gyration of cinnamoyl-CoA-*BrCHSv5* complex decreased rapidly than in *p*-coumaroyl-CoA towards the end of simulations. This indicates that the compactness of the complex has been increased and strengthened the interaction between cinnamoyl-CoA and *BrCHSv5*. Thus, it has contributed to its higher binding affinity towards the receptor. On the other hand, residue pro375 showed high fluctuations in the RMSF profile of the *p*-coumaroyl-CoA-*BrCHSv5* complex. It resulted in the distortion of the active site geometry of the receptor and reduced the binding affinity of *p*-coumaroyl-CoA towards receptor. In addition, the increased binding affinity of *BrCHSv5* for cinnamoyl-CoA contributed by the hydrogen bonds and salt bridge interaction. Aspartic acid acts as a hydrogen bond donor because it

is capable of forming hydrogen bonds thrice than glutamine (de Freitas & Schapira, 2017). Cinnamoyl-CoA formed hydrogen bonds with the residue Asp207 of *BrCHSv5* receptor. The number of salt bridge interactions increased at the end of the simulation and mainly formed by Lys55, Arg58 and Lys62 residues of the receptor with cinnamoyl-CoA.

Malonyl-CoA recorded less negative binding free energies with all five variants of *BrCHS* receptor than cinnamoyl-CoA. It is also a preferred substrate ligand for the receptors and validates the docking results. Arg58 of *BrCHSv1*, *BrCHSv3*, *BrCHSv4* and *BrCHSv5* receptors formed a repulsive bond with malonyl-CoA as well. It has decreased the Coulomb energy and resulted in less negative binding free energies than cinnamoyl-CoA. Hence, the receptors have less binding affinity for malonyl-CoA than cinnamoyl-CoA. Residues Arg58 and Lys62 of the receptors formed strong hydrogen bonds that contributed to its negative binding free energy.

Caffeoyl-CoA showed the same trend of preference as in the docking result. It recorded more negative binding free energies with *BrCHSv2* and *BrCHSv3* than in *p*-coumaroyl-CoA. This indicates that it has a higher binding affinity towards the receptor variants compared to *p*-coumaroyl-CoA. Only caffeoyl-CoA-*BrCHSv1* complex recorded positive binding free energy due to the formation of repulsive interaction by Arg259. The repulsive interaction resulted in a more positive Coulomb energy and reduced its binding affinity. This shows that it is not a favourable ligand for the *BrCHSv1* receptor.

Results showed that feruloyl-CoA is the least favorable ligand among the substrate ligand. The ligand recorded positive binding free energies with all *BrCHS* receptor variants except for *BrCHSv4*. Though docking results showed that feruloyl-CoA has more preference for *BrCHSv1* and *BrCHSv5* than cinnamoyl-CoA, but the binding free energy results contradict to the docking results. Therefore, feruloyl-CoA is not a preferred substrate ligand for *BrCHSv1*, *BrCHSv2*, *BrCHSv3* and *BrCHSv5* as it recorded a more

positive binding free energies. Arg259 of the *BrCHSv1* receptor formed repulsive interaction with feruloyl-CoA. It has resulted in more positive Coulomb energy decrease of electrostatic energy and thus, reduced its binding affinity towards the receptor.

Both reference ligands (acetyl-CoA and CoA) recorded more negative binding free energies and indicating that the ligands have a higher binding affinity towards the *BrCHS* receptor variants. The binding free energy profiles of the reference ligands are in good agreement with the docking results of both ligands which have a higher binding affinity towards the *BrCHS* receptor variants. Its binding free energy is mainly contributed by more negative Coulomb energy due to high numbers of hydrogen bonds and salt bridge interactions.

Binding free energy calculation using the MM-PBSA method is useful indeed for post-processing of the docked complexes and rationalize the observed differences in the protein-ligand interactions. It can be summarized that cinnamoyl-CoA is the preferred substrate for *BrCHSv1*, *BrCHSv2*, *BrCHSv3* and *BrCHSv5* receptors. Meanwhile, *p*-coumaroyl-CoA is the only preferred substrate for the *BrCHSv4* receptor.

CHAPTER 6: CONCLUSION

Chalcone synthase of *Boesenbergia rotunda* shared high similarity with CHS from other plants of the same superfamily. The homology models of the five variants of *BrCHS* receptor showed good stereochemical quality as recorded more than 90% of residues in the favored region of the Ramachandran plots. The five receptor variants have molecular masses in the range of 42 – 44 kDa. Multiple sequence alignment revealed the presence of highly conserved catalytic triad (Cys164, His303 and Asn336), gatekeeper residues (Phe215 and Phe265) and GFGPG loop. Molecular docking results showed that cinnamoyl-CoA has a higher binding affinity towards *BrCHSv1*, *BrCHSv2* and *BrCHSv3* receptors as shown by docking results. Whereas, *BrCHSv4* and *BrCHSv5* receptors have a higher binding affinity for *p*-coumaroyl-CoA. Trajectory analysis based on the RMSD, RMSF and radius of gyration showed that the protein-ligand complexes were stable throughout the 10 ns simulations. Based on the binding free energy profiles, cinnamoyl-CoA has a higher binding affinity towards *BrCHSv1*, *BrCHSv2* and *BrCHSv3* receptors. On the other hand, *p*-coumaroyl-CoA has a higher binding affinity towards *BrCHSv4* as recorded more negative binding free energy. Contradicting to the docking result, cinnamoyl-CoA has a higher binding affinity towards *BrCHSv5*. In conclusion, molecular docking and molecular dynamics studies revealed that cinnamoyl-CoA is the preferred substrate for *BrCHSv1*, *BrCHSv2*, *BrCHSv3* and *BrCHSv5*. Whereas, *p*-coumaroyl-CoA is the only preferred substrate for *BrCHSv4*. The findings of this research are useful for the production of pharmaceutically important polyketides. The information would pave way for further *in vitro* and *in silico* studies of chalcone synthase and discloses the underlying pathway mechanism of flavonoid biosynthesis.

REFERENCES

- Abe, I., Tanaka, H., & Noguchi, H. (2002). Enzymatic formation of an unnatural hexacyclic C35 polyprenoid by bacterial squalene cyclase. *Journal of the American Chemical Society*, 124(49), 14514-14515.
- Abe, I., Sano, Y., Takahashi, Y., & Noguchi, H. (2003). Site-directed mutagenesis of benzalacetone synthase. The role of the Phe215 in plant type III polyketide synthases. *Journal of Biological Chemistry*, 278(27), 25218-25226.
- Abe, I., Watanabe, T., Morita, H., Kohno, T., & Noguchi, H. (2006). Engineered biosynthesis of plant polyketides: Manipulation of chalcone synthase. *Organic Letters*, 8(3), 499-502.
- Abe, I., & Morita, H. (2010). Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. *Natural Product Reports*, 27(6), 809-838.
- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2, 19-25.
- Aier, I., Varadwaj, P. K., & Raj, U. (2016). Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. *Scientific Reports*, 6(1).
- Allen, W. J., Balias, T. E., Mukherjee, S., Brozell, S. R., Moustakas, D. T., Lang, P. T., . . . Rizzo, R. C. (2015). DOCK 6: Impact of new features and current docking performance. *Journal of Computational Chemistry*, 36(15), 1132-1156.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389-3402.
- Aqvist, J., Medina, C., & Samuelsson, J.-E. (1994). A new method for predicting binding affinity in computer-aided drug design. *Protein Engineering, Design and Selection*, 7(3), 385-391.
- Austin, M. B., & Noel, J. P. (2002). The chalcone synthase superfamily of type III polyketide synthases. *Natural Product Reports*, 20(1), 79-110.

- Austin, M. B., Bowman, M. E., Ferrer, J.-L., Schroder, J., & Noel, J. P. (2004). An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. *Chemistry & Biology*, 11(9), 1179-1194.
- Awasthi, P., Jamwal, V. L., Kapoor, N., & Rasool, S. (2016). Homology modeling and docking study of chalcone synthase gene (*CfCHS*) from *Coleus forskohlii*. *Cogent Biology*, 2(1).
- Baharudin, M. A., Hamid, S. A., & Susanti, D. (2015). Chemical composition and antibacterial activity of essential oils from three aromatic plants of the zingiberaceae family in Malaysia. *Journal of Physical Science*, 26(1), 71-81.
- Bakan, A., Meireles, L. M., & Bahar, I. (2011). ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics*, 27(11), 1575-1577.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*, 98(18), 10037-10041.
- Bartuzi, D., Kaczor, A. A., & Matosiuk, D. (2017). Signaling within Allosteric Machines: Signal Transmission Pathways Inside G Protein-Coupled Receptors. *Molecules*, 22(7), 1188.
- Batovska, D., & Todorova, I. (2010). Trends in utilization of the pharmacological potential of chalcones. *Current Clinical Pharmacology*, 5(1), 1-29.
- Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*, 81(8), 3684-3690.
- Boumendjel, A., Bocard, J., Carrupt, P.-A., Nicolle, E., Blanc, M., Geze, A., . . . Dumontet, C. (2008). Antimitotic and antiproliferative activities of chalcones: Forward structure-activity relationship. *Journal of Medicinal Chemistry*, 51(7), 2307-2310.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4(2), 187-217.

- Brown, S. P., & Muchmore, S. W. (2009). Large-scale application of high-throughput molecular mechanics with Poisson– Boltzmann surface area for routine physics-based scoring of protein– ligand complexes. *Journal of Medicinal Chemistry*, 52(10), 3159-3165.
- Cao, G., Sofic, E., & Prior, R. L. (1997). Antioxidant and prooxidant behavior of flavonoids: Structure-activity relationships. *Free Radical Biology and Medicine*, 22(5), 749-760.
- Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., . . . Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16), 1668-1688.
- Cavasotto, C. N., & Phatak, S. S. (2009). Homology modeling in drug discovery: Current trends and applications. *Drug Discovery Today*, 14(13-14), 676-683.
- Chandel, T. I., Zaman, M., Khan, M. V., Ali, M., Rabbani, G., Ishtikhar, M., & Khan, R. H. (2018). A mechanistic insight into protein-ligand interaction, folding, misfolding, aggregation and inhibition of protein aggregates: An overview. *International Journal of Biological Macromolecules*, 106, 1115-1129.
- Chang, W. S., Lee, Y. J., Lu, F. J., & Chiang, H. C. (1993). Inhibitory effects of flavonoids on xanthine oxidase. *Anticancer Research*, 13(6A), 2165-2170.
- Cheng, Z., Lin, C., Hwang, T., & Teng, C. (2001). Brousochalcone A, a potent antioxidant and effective suppressor of inducible nitric oxide synthase in lipopolysaccharide-activated macrophages. *Biochemical Pharmacology*, 61(8), 939-946.
- Ching, A. Y. L., Tang, S. W., Sukari, M. A., Ee, G., Lian, C., Rahmani, M., & Khalid, K. (2007). Characterization of flavonoid derivatives from *Boesenbergia rotunda* (L.). *Malaysian Journal of Analytical Sciences*, 11(1), 154-159.
- Chodera, J. D., Mobley, D. L., Shirts, M. R., Dixon, R. W., Branson, K., & Pande, V. S. (2011). Alchemical free energy methods for drug discovery: progress and challenges. *Current Opinion in Structural Biology*, 21(2), 150-160.
- Christen, M., Hunenberger, P. H., Bakowies, D., Baron, R., Burgi, R., Geerke, D. P., . . . van Gunsteren, W. F. (2005). The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*, 26(16), 1719-1751.

- Christensen, A. B., Gregersen, P. L., Schroder, J., & Collinge, D. B. (1998). A chalcone synthase with an unusual substrate preference is expressed in barley leaves in response to UV light and pathogen attack. *Plant Molecular Biology*, 37(5), 849-857.
- Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., . . . Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19), 5179-5197.
- D'Mello, P., Gadhwal, M., Joshi, U., Shetgiri, P., K. M. Kundnani College of Pharmacy, P., & Mumbai. (2011). Modeling of COX-2 inhibitory activity of flavonoids. *International Journal of Pharmacy and Pharmaceutical Sciences*, 3(4), 33-40.
- 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.
- Daugelaite, J., O' Driscoll, A., & Sleator, R. D. (2013). An overview of multiple sequence alignments and cloud computing in bioinformatics. *ISRN Biomathematics*, 2013, 1-14.
- de Freitas, R. F., & Schapira, M. (2017). A systematic analysis of atomic protein-ligand interactions in the PDB. *Medicinal Chemistry Communication*, 8(10), 1970-1981.
- de Ruiter, A., & Oostenbrink, C. (2011). Free energy calculations of protein-ligand interactions. *Current Opinon in Chemical Biology*, 15(4), 547-552.
- de Ruyck, J., Brysbaert, G., Blossey, R., & Lensink, M. F. (2016). Molecular docking as a popular tool in drug design, an in silico travel. *Advances and Applications in Bioinformatics and Chemistry*, 9, 1-11.
- de Souza, O. N., & Ornstein, R. L. (1997). Effect of periodic box size on aqueous molecular dynamics simulation of a DNA dodecamer with particle-mesh Ewald method. *Biophysical Journal*, 72(6), 2395-2397.
- De Vivo, M., Masetti, M., Bottegoni, G., & Cavalli, A. (2016). Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of Medicinal Chemistry*, 59(9), 4035-4061.

- Deng, X., Bashandy, H., Ainasoja, M., Kontturi, J., Pietiäinen, M., Laitinen, R. A. E., . . . Teeri, T. H. (2014). Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of *Gerbera hybrida*. *New Phytologist*, *201*(4), 1469-1483.
- Dixon, R. A., Achnine, L., Kota, P., Liu, C.-J., Reddy, M. S. S., & Wang, L. (2002). The phenylpropanoid pathway and plant defence—a genomics perspective. *Molecular Plant Pathology*, *3*(5), 371-390.
- Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: A protein–protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731-1737.
- Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., . . . Liu, S.-Q. (2016). Insights into protein–ligand interactions: Mechanisms, models, and methods. *International Journal of Molecular Sciences*, *17*(2), 144.
- Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, *9*(1), 71.
- Eckermann, S., Schröder, G., Schmidt, J., Strack, D., Edrada-Ebel, R., Helariutta, Y., . . . Schröder, J. (1998). New pathway to polyketides in plants. *Nature*, *396*(6709), 387-390.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792-1797.
- Edgar, R. C., & Batzoglou, S. (2006). Multiple sequence alignment. *Current Opinion in Structural Biology*, *16*(3), 368-373.
- Eisenberg, D., Lüthy, R., & Bowie, J. U. (1997). [20] VERIFY3D: Assessment of protein models with three-dimensional profiles. *Macromolecular Crystallography Part B*, 396-404.
- Eng-Chong, T., Yean-Kee, L., Chin-Fei, C., Choon-Han, H., Sher-Ming, W., Li-Ping, C. T., . . . Yusof, R. (2012). *Boesenbergia rotunda*: From ethnomedicine to drug discovery. *Evidence-Based Complementary and Alternative Medicine*, *2012*, 1-25.

- España, L., Heredia-Guerrero, J. A., Segado, P., Benítez, J. J., Heredia, A., & Domínguez, E. (2014). Biomechanical properties of the tomato (*Solanum lycopersicum*) fruit cuticle during development are modulated by changes in the relative amounts of its components. *New Phytologist*, 202(3), 790-802.
- Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *Journal of Chemical Physics*, 103(19), 8577-8593.
- Falcone Ferreyra, M. L., Rius, S. P., & Casati, P. (2012). Flavonoids: Biosynthesis, biological functions, and biotechnological applications. *Frontiers in Plant Science*, 3, 222.
- Feinbaum, R. L., & Ausubel, F. M. (1988). Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. *Molecular and Cellular Biology*, 8(5), 1985-1992.
- Feng, Q., Gui-Gong, G., Yang, Z., Hui-Chun, X., Lan, J., Jun, S., & Zhi, C. (2015). Molecular cloning and expression profiling of a chalcone synthase gene from *Lamiophlomis rotata*. *Journal of Genetics*, 94(2), 193-205.
- Feng, T., Li, M., Zhou, J., Zhuang, H., Chen, F., Ye, R., ... Fang, Z. (2015). Application of molecular dynamics simulation in food carbohydrate research—a review. *Innovative Food Science & Emerging Technologies*, 31, 1–13.
- 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 Biology*, 6, 775.
- Firdayani, F., Arsianti, A., Churiyah, C. C., & Yanuar, A. (2018). Molecular docking and dynamic simulation studies of benzoylated emodin into hbv core protein. *Journal of Young Pharmacists*, 10, S20-S24.
- Fliegmann, J., Schroder, G., Schanz, S., Britsch, L., & Schroder, J. (1992). Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. *Plant Molecular Biology*, 18(3), 489-503.
- Fu, Y., Hsieh, T. C., Guo, J., Kunicki, J., Lee, M. Y., Darzynkiewicz, Z., & Wu, J. M. (2004). Licochalcone-A, a novel flavonoid isolated from licorice root (*Glycyrrhiza glabra*), causes G2 and late-G1 arrests in androgen-independent PC-3 prostate cancer cells. *Biochemical and biophysical research communications*, 322(1), 263-270.

- Fukuma, K., Neuls, E. D., Ryberg, J. M., Suh, D.-Y., & Sankawa, U. (2007). Mutational analysis of conserved outer sphere arginine residues of chalcone synthase. *The Journal of Biochemistry*, 142(6), 731-739.
- Gafner, S., Wolfender, J.-L., Mavi, S., & Hostettmann, K. (1996). Antifungal and antibacterial chalcones from *Myrica serrata*. *Planta Medica*, 62(01), 67-69.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook*, 571-607.
- Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, 10(5), 449-461.
- Geourjon, C., & Deleage, G. (1995). SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics*, 11(6), 681-684.
- Gowers, R. J., Linke, M., Barnoud, J., Reddy, T. J. E., Melo, M. N., Seyler, S. L., . . . O, B. (2016). *MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations*. Paper presented at the Proceedings of the 15th Python in Science Conference, Austin, Texas.
- Gräter, F., Schwarzl, S. M., Dejaegere, A., Fischer, S., & Smith, J. C. (2005). Protein/ligand binding free energies calculated with quantum mechanics/molecular mechanics. *The Journal of Physical Chemistry B*, 109(20), 10474-10483.
- Guo, L., Chen, X., Li, L.-N., Tang, W., Pan, Y.-T., & Kong, J.-Q. (2016). Transcriptome-enabled discovery and functional characterization of enzymes related to (2S)-pinocembrin biosynthesis from *Ornithogalum caudatum* and their application for metabolic engineering. *Microbial Cell Factories*, 15(1).
- Han, Y., Ding, T., Su, B., & Jiang, H. (2016). Genome-wide identification, characterization and expression analysis of the chalcone synthase family in maize. *International Journal of Molecular Sciences*, 17(2), 161.
- Hatzieremia, S., Gray, A. I., Ferro, V. A., Paul, A., & Plevin, R. (2006). The effects of cardamonin on lipopolysaccharide-induced inflammatory protein production and MAP kinase and NFkappaB signalling pathways in monocytes/macrophages. *British Journal of Pharmacology*, 149(2), 188-198.

- Hospital, A., Goñi, J. R., Orozco, M., & Gelpi, J. L. (2015). Molecular dynamics simulations: advances and applications. *Advances and Applications in Bioinformatics and Chemistry*, 8, 37-47.
- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*, 51(1), 69-82.
- Hrazdina, G., Kreuzaler, F., Hahlbrock, K., & Grisebach, H. (1976). Substrate specificity of flavanone synthase from cell suspension cultures of parsley and structure of release products in vitro. *Archives of Biochemistry and Biophysics*, 175(1), 392-399.
- Huang, S.-Y., & Zou, X. (2006). An iterative knowledge-based scoring function to predict protein–ligand interactions: I. Derivation of interaction potentials. *Journal of Computational Chemistry*, 27(15), 1866-1875.
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33-38.
- Integrated Taxonomic Information System, ITIS. (2018). *Boesenbergia rotunda* (L.) Mansf. Retrieved on June 2, 2018 from https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=506504#null
- Isa, N. M., Abdelwahab, S. I., Mohan, S., Abdul, A. B., Sukari, M. A., Taha, M. M. E., . . . Mustafa, M. R. (2012). In vitro anti-inflammatory, cytotoxic and antioxidant activities of boesenbergin A, a chalcone isolated from *Boesenbergia rotunda* (L.) (fingerroot). *Brazilian Journal of Medical and Biological Research*, 45(6), 524-530.
- Isa, N. M., Abdul, A. B., Abdelwahab, S. I., Abdullah, R., Sukari, M. A., Kamalidehghan, B., . . . Mohan, S. (2013). Boesenbergin A, a chalcone from *Boesenbergia rotunda* induces apoptosis via mitochondrial dysregulation and cytochrome c release in A549 cells in vitro: Involvement of HSP70 and Bcl2/Bax signalling pathways. *Journal of Functional Foods*, 5(1), 87-97.
- Jeffrey, G. A. (1997). *An introduction to hydrogen bonding* (Vol. 32). New York: Oxford university press.
- Jez, J. M., & Noel, J. P. (2000). Mechanism of Chalcone Synthase: pKa 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., Austin, M. B., Ferrer, J.-L., Bowman, M. E., Schröder, J., & Noel, J. P. (2000). Structural control of polyketide formation in plant-specific polyketide synthases. *Chemistry & Biology*, 7(12), 919-930.
- Jez, J. M., Bowman, M. E., & Noel, J. P. (2001a). Structure-guided programming of polyketide chain-length determination in chalcone synthase. *Biochemistry*, 40(49), 14829-14838.
- Jez, J. M., Ferrer, J. L., Bowman, M. E., Austin, M. B., Schröder, J., Dixon, R. A., & Noel, J. P. (2001b). Structure and mechanism of chalcone synthase-like polyketide synthases. *Journal of Industrial Microbiology and Biotechnology*, 27(6), 393-398.
- Jez, J. M., Bowman, M. E., & Noel, J. P. (2002). Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. *Proceedings of the National Academy of Sciences*, 99(8), 5319-5324.
- 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.
- Jing, L. J., Mohamed, M., Rahmat, A., & Bakar, M. F. A. (2010). Phytochemicals, antioxidant properties and anticancer investigations of the different parts of several gingers species (*Boesenbergia rotunda*, *Boesenbergia pulchella* var *attenuata* and *Boesenbergia armeniaca*). *Journal of Medicinal Plant Research*, 4(1), 27-32.
- Jitvaropas, R., Saenthaweesuk, S., Somparn, N., Thuppia, A., Sireeratawong, S., & Phoolcharoen, W. (2012). Antioxidant, antimicrobial and wound healing activities of *Boesenbergia rotunda*. *Nat Prod Commun*, 7(7), 909-912.
- Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. Edited by F. E. Cohen. *Journal of Molecular Biology*, 267(3), 727-748.
- Jorgensen, W. L., & Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6), 1657-1666.
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, 45(D1), D353-D361.

- Kastenholz, M. A., Pastor, M., Cruciani, G., Haaksma, E. E. J., & Fox, T. (2000). GRID/CPCA: A new computational tool to design selective ligands. *Journal of Medicinal Chemistry*, 43(16), 3033-3044.
- Katsuyama, Y., Hirose, Y., Funa, N., Ohnishi, Y., & Horinouchi, S. (2010). Precursor-Directed Biosynthesis of Curcumin Analogs in *Escherichia coli*. *Bioscience, Biotechnology, and Biochemistry*, 74(3), 641-645.
- Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., . . . Bryant, S. H. (2015). PubChem Substance and Compound databases. *Nucleic Acids Research*, 44(D1), D1202-D1213.
- Koes, R. E., Spelt, C. E., Mol, J. N. M., & Gerats, A. G. M. (1987). The chalcone synthase multigene family of *Petunia hybrida* (V30): Sequence homology, chromosomal localization and evolutionary aspects. *Plant Molecular Biology*, 10(2), 159-169.
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., . . . Cheatham, T. E. (2000). Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research*, 33(12), 889-897.
- Kreuzaler, F., Light, R. J., & Hahlbrock, K. (1978). Flavanone synthase catalyzes CO₂ exchange and decarboxylation of malonyl-CoA. *FEBS Letters*, 94(1), 175-178.
- Krieger, E., Koraimann, G., & Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA-A self-parameterizing force field. *Proteins*, 47(3), 393-402.
- Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., . . . Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins: Structure, Function, and Bioinformatics*, 77(S9), 114-122.
- Krieger, E., & Vriend, G. (2014). YASARA View - Molecular graphics for all devices - from smartphones to workstations. *Bioinformatics*, 30(20), 2981-2982.
- Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., & Ferrin, T. E. (1982). A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*, 161(2), 269-288.
- Land, H., & Humble, M. S. (2017). YASARA: A tool to obtain structural guidance in biocatalytic investigations. *Protein Engineering*, 43-67.

- Lanz, T., Tropf, S., Marner, F. J., Schröder, J., & Schröder, G. (1991). The role of cysteines in polyketide synthases. Site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. *Journal of Biological Chemistry*, 266(15), 9971-9976.
- Laskowski, R., MacArthur, M. W., Moss, D. S., & Thornton, J. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283-291.
- Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *Journal of Chemical Information and Modeling*, 51(10), 2778-2786.
- Lee, J.-Y., Jeong, K.-W., Shin, S., Lee, J.-U., & Kim, Y. (2009). Antimicrobial natural products as β -ketoacyl-acyl carrier protein synthase III inhibitors. *Bioorganic & Medicinal Chemistry*, 17(15), 5408-5413.
- 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 *Scutellaria viscidula* Bunge. *Genetics and Molecular Biology*, 33(2), 285-291.
- Levitt, D., & Banaszak, L. (1992). POCKET: A computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *Journal of Molecular Graphics*, 10(4), 229-234.
- Liao, K. H., Chen, K.-B., Lee, W.-Y., Sun, M.-F., Lee, C.-C., & Chen, C. Y.-C. (2014). Ligand-based and structure-based investigation for alzheimer's disease from traditional chinese medicine. *Evidence-Based Complementary and Alternative Medicine*, 2014, 1-16.
- Lim, T. K. (2016). *Boesenbergia rotunda Edible medicinal and non-medicinal plants: Volume 12, modified stems, roots, bulbs* (pp. 214-232). Switzerland. Springer International Publishing.
- Liou, G., Chiang, Y.-C., Wang, Y., & Weng, J.-K. (2018). Mechanistic basis for the evolution of chalcone synthase catalytic cysteine reactivity in land plants. *Journal of Biological Chemistry*, 293(48), 18601-18612.
- Liu, X.-J., Chuang, Y.-N., Chiou, C.-Y., Chin, D.-C., Shen, F.-Q., & Yeh, K.-W. (2012). Methylation effect on chalcone synthase gene expression determines anthocyanin pigmentation in floral tissues of two *Oncidium* orchid cultivars. *Planta*, 236(2), 401-409.

- Liu, X., & Go, M.-L. (2007). Antiproliferative activity of chalcones with basic functionalities. *Bioorganic & Medicinal Chemistry*, 15(22), 7021-7034.
- Liu, Z., Liu, Y., Zeng, G., Shao, B., Chen, M., Li, Z., . . . Zhong, H. (2018). Application of molecular docking for the degradation of organic pollutants in the environmental remediation: A review. *Chemosphere*, 203, 139-150.
- Lo, C., Coolbaugh, R. C., & Nicholson, R. L. (2002). Molecular characterization and in silico expression analysis of a chalcone synthase gene family in *Sorghum bicolor*. *Physiological and Molecular Plant Pathology*, 61(3), 179-188.
- Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., . . . Richardson, D. C. (2003). Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins: Structure, Function, and Bioinformatics*, 50(3), 437-450.
- Lukk, T., Sakai, A., Kalyanaraman, C., Brown, S. D., Imker, H. J., Song, L., . . . Jacobson, M. P. (2012). Homology models guide discovery of diverse enzyme specificities among dipeptide epimerases in the enolase superfamily. *Proceedings of the National Academy of Sciences*, 109(11), 4122-4127.
- MacKerell, A. D., Bashford, D., Bellott, M. L. D. R., Dunbrack Jr, R. L., Evanseck, J. D., Field, M. J., . . . & Joseph-McCarthy, D. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry B*, 102(18), 3586-3616.
- MacKerell, A. D. (2004). Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, 25(13), 1584-1604.
- Mahadevi, A. S., & Sastry, G. N. (2013). Cation- π Interaction: Its role and relevance in chemistry, biology, and material science. *Chemical Reviews*, 113(3), 2100-2138.
- Mahajan, M., Ahuja, P. S., & Yadav, S. K. (2011). Post-transcriptional silencing of flavonol synthase mRNA in tobacco leads to fruits with arrested seed set. *PLoS ONE*, 6(12), e28315.
- Mallika, V., Sivakumar, K. C., & Soniya, E. V. (2011). Evolutionary implications and physicochemical analyses of selected proteins of type III polyketide synthase family. *Evolutionary Bioinformatics Online*, 7.
- Mandal, S. M., Chakraborty, D., & Dey, S. (2010). Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant Signaling & Behavior*, 5(4), 359-368.

- Mbah, A. N., Kamga, H. L., Awofolu, O. R., & Isokpehi, R. D. (2012). Drug target exploitable structural features of adenylyl cyclase activity in *Schistosoma mansoni*. *Drug Target Insights*, 6.
- McGibbon, Robert T., Beauchamp, Kyle A., Harrigan, Matthew P., Klein, C., Swails, Jason M., Hernández, Carlos X., . . . Pande, Vijay S. (2015). MDTraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophysical Journal*, 109(8), 1528-1532.
- Meng, X.-Y., Zhang, H.-X., Mezei, M., & Cui, M. (2011). Molecular docking: A powerful approach for structure-based drug discovery. *Current Computer-aided Drug Design*, 7(2), 146-157.
- 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 & Medical Modelling*, 10, 22.
- Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., & Beckstein, O. (2011). MDAAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry*, 32(10), 2319-2327.
- Moriguchi, T., Kita, M., Tomono, Y., EndoInagaki, T., & Omura, M. (1999). One type of chalcone synthase gene expressed during embryogenesis regulates the flavonoid accumulation in citrus cell cultures. *Plant and Cell Physiology*, 40(6), 651-655.
- Morita, H., Takahashi, Y., Noguchi, H., & Abe, I. (2000). Enzymatic formation of unnatural aromatic polyketides by chalcone synthase. *Biochemical and Biophysical Research Communications*, 279(1), 190-195.
- Morita, H., Kondo, S., Oguro, S., Noguchi, H., Sugio, S., Abe, I., & Kohno, T. (2007). Structural insight into chain-length control and product specificity of pentaketide chromone synthase from *Aloe arborescens*. *Chemistry & Biology*, 14(4), 359-369.
- Morita, H., Shimokawa, Y., Tanio, M., Kato, R., Noguchi, H., Sugio, S., . . . Abe, I. (2010). A structure-based mechanism for benzalacetone synthase from *Rheum palmatum*. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2), 669-673.
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins: Structure, Function, and Bioinformatics*, 12(4), 345-364.

- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785-2791.
- National Parks of Board of Singapore (2013). *Boesenbergia rotunda*. [Online image]. Retrieved on June 2, 2018 from: <https://florafaunaweb.nparks.gov.sg/Special-Pages/plant-detail.aspx?id=4849> .
- Neria, E., Fischer, S., & Karplus, M. (1996). Simulation of activation free energies in molecular systems. *Journal of Chemical Physics*, 105(5), 1902-1921.
- Nishimura, R., Tabata, K., Arakawa, M., Ito, Y., Kimura, Y., Akihisa, T., . . . Suzuki, T. (2007). Isobavachalcone, a chalcone constituent of *Angelica keiskei*, induces apoptosis in neuroblastoma. *Biological and Pharmaceutical Bulletin*, 30(10), 1878-1883.
- Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-coffee: A novel method for fast and accurate multiple sequence alignment: Edited by J. Thornton. *Journal of Molecular Biology*, 302(1), 205-217.
- Nurnadiah, R. (2017). *Isolation, characterization and overexpression of Chalcone synthase gene in suspension cultures of Boesenbergia rotunda* (Doctoral dissertation, University of Malaya). Retrieved on December 2, 2018 from <http://studentsrepo.um.edu.my/8412/>
- Ongwisepaiboon, O., & Jiraungkoorskul, W. (2017). Fingerroot, *Boesenbergia rotunda* and its aphrodisiac activity. *Pharmacognosy Reviews*, 11(21), 27-30.
- Oostenbrink, C., Villa, A., Mark, A. E., & van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, 25(13), 1656-1676.
- Orlikova, B., Tasdemir, D., Golais, F., Dicato, M., & Diederich, M. (2011). Dietary chalcones with chemopreventive and chemotherapeutic potential. *Genes & Nutrition*, 6(2), 125-147.
- Paissoni, C., Spiliotopoulos, D., Musco, G., & Spitaleri, A. (2014). GMXPBSA 2.0: A GROMACS tool to perform MM/PBSA and computational alanine scanning. *Computer Physics Communications*, 185(11), 2920-2929.

- Paissoni, C., Spiliotopoulos, D., Musco, G., & Spitaleri, A. (2015). GMXPBSA 2.1: A GROMACS tool to perform MM/PBSA and computational alanine scanning. *Computer Physics Communications*, 186, 105-107.
- Panche, A. N., Diwan, A. D., & Chandra, S. R. (2016). Flavonoids: An overview. *Journal of Nutritional Science*, 5.
- Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single-crystals - A new molecular-dynamics method. *Journal of Applied Physics*, 52(12), 7182-7190.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605-1612.
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., . . . Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16), 1781-1802.
- Pirhadi, S., Sunseri, J., & Koes, D. R. (2016). Open source molecular modeling. *Journal of Molecular Graphics & Modelling*, 69, 127-143.
- Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1), 1-19.
- Potipiranun, T., Adisakwattana, S., Worawalai, W., Ramadhan, R., & Phuwapraisirisan, P. (2018). Identification of pinocebrin as an anti-glycation agent and alpha-glucosidase inhibitor from fingerroot (*Boesenbergia rotunda*): The tentative structure(-)activity relationship towards MG-trapping activity. *Molecules*, 23(12).
- Rahman, R., Zakaria, I. I., Salleh, A. B., & Basri, M. (2012). Enzymatic Properties and mutational studies of chalcone synthase from *Physcomitrella patens*. *International Journal of Molecular Sciences*, 13(8), 9673-9691.
- Roche, D. B., & McGuffin, L. J. (2016). *In silico* identification and characterization of protein-ligand binding sites. *Computational Design of Ligand Binding Proteins*, 1-21.
- Rohini, K., & Srikumar, P. S. (2013). Insights from the docking and molecular dynamics simulation of the phosphopantetheinyl transferase (PptT) structural model from *Mycobacterium tuberculosis*. *Bioinformation*, 9(13), 685-689.

- Saini, R. D. (2017). Thermodynamics of protein-ligand interactions and their analysis. *Journal of Proteins and Proteomics*, 8(4).
- Salama, S. M., Bilgen, M., Al Rashdi, A. S., & Abdulla, M. A. (2012). Efficacy of *Boesenbergia rotunda* treatment against thioacetamide-induced liver cirrhosis in a rat model. *Evidence-Based Complementary and Alternative Medicine*, 2012, 1-12.
- Salentin, S., Haupt, V. J., Daminelli, S., & Schroeder, M. (2014). Polypharmacology rescored: Protein–ligand interaction profiles for remote binding site similarity assessment. *Progress in Biophysics and Molecular Biology*, 116(2), 174-186.
- Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), 779-815.
- Sanjari, S., Shobbar, Z. S., Ebrahimi, M., Hasanloo, T., Sadat-Noori, S. A., & Tirnaz, S. (2015). Chalcone synthase genes from milk thistle (*Silybum marianum*): Isolation and expression analysis. *Journal of Genetics*, 94(4), 611-617.
- 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.
- Schüz, R., Heller, W., & Hahlbrock, K. (1983). Substrate specificity of chalcone synthase from *Petroselinum hortense*. Formation of phloroglucinol derivatives from aliphatic substrates. *Journal of Biological Chemistry*, 258(11), 6730-6734.
- Selvaraj, C., Sakkiah, S., Tong, W., & Hong, H. (2018). Molecular dynamics simulations and applications in computational toxicology and nanotoxicology. *Food Chem Toxicol*, 112, 495-506.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., . . . Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 539-539.
- Singh, N., & Warshel, A. (2010). Absolute binding free energy calculations: On the accuracy of computational scoring of protein-ligand interactions. *Proteins*, 78(7), 1705-1723.
- Sivashanmugam, M., Raghunath, C., & Vetrivel, U. (2013). Virtual screening studies reveal linarin as a potential natural inhibitor targeting CDK4 in retinoblastoma. *Journal of Pharmacology and Pharmacotherapeutics*, 4(4), 256-264.

- Song, C., Ring, L., Hoffmann, T., Huang, F.-C., Slovin, J., & Schwab, W. (2015). Acylphloroglucinol Biosynthesis in Strawberry Fruit. *Plant Physiology*, 169(3), 1656.
- Sousa, S. F., Ribeiro, A. J. M., Coimbra, J. T. S., Neves, R. P. P., Martins, S. A., Moorthy, N. S. H. N., . . . Ramos, M. J. (2013). Protein-ligand docking in the new millennium – a retrospective of 10 years in the field. *Current Medicinal Chemistry*, 20(18), 2296-2314.
- Spiliotopoulos, D., & Caflisch, A. (2014). Molecular dynamics simulations of bromodomains reveal binding-site flexibility and multiple binding modes of the natural ligand acetyl-lysine. *Israel Journal of Chemistry*, 54(8 - 9), 1084-1092.
- Stevensen, C. (1999). JAMU: An Indonesian herbal tradition with a long past, a little known present and an uncertain future. *Complementary Therapies in Nursing and Midwifery*, 5(1), 1-3.
- Suh, D. Y., Fukuma, K., Kagami, J., Yamazaki, Y., Shibuya, M., Ebizuka, Y., & Sankawa, U. (2000). Identification of amino acid residues important in the cyclization reactions of chalcone and stilbene synthases. *Biochemical Journal*, 350(Pt 1), 229-235.
- Sukari, M. A., Mohd Sharif, N. W., Yap, A. L. C., Tang, S. W., Neoh, B. K., Rahmani, M., . . . Yusof, U. K. (2008). Chemical constituents variations of essential oils from rhizomes of four Zingiberaceae species. *Malaysian Journal of Analytical Sciences*, 12(3), 638-644.
- Sun, W., Meng, X., Liang, L., Jiang, W., Huang, Y., He, J., . . . Wang, L. (2015). Molecular and biochemical analysis of chalcone synthase from *Freesia hybrid* in flavonoid biosynthetic pathway. *PLoS ONE*, 10(3), e0119054.
- Supinya, T., Subhadhirasakul, S., Puripattanavong, J., & Tassanee, P. (2003). HIV-1 protease inhibitory substances from the rhizomes of *Boesenbergia pandurata* Holtt. *Songklanakarinn Journal of Science and Technology*, 29(4), 503-508.
- Tan, B. C., Tan, S. K., Wong, S. M., Ata, N., Rahman, N. A., & Khalid, N. (2015). Distribution of flavonoids and cyclohexenyl chalcone derivatives in conventional propagated and in vitro-derived field-grown *Boesenbergia rotunda* (L.) Mansf. *Evidence-Based Complementary and Alternative Medicine*, 2015, 451870.
- Thomas, S., & Andreas, L. (2010). Towards accurate free energy calculations in ligand protein-binding studies. *Current Medicinal Chemistry*, 17(8), 767-785.

- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), 4673-4680.
- Tropf, S., Karcher, B., Schroder, G., & Schroder, J. (1995). Reaction mechanisms of homodimeric plant polyketide synthase (stilbenes and chalcone synthase). A single active site for the condensing reaction is sufficient for synthesis of stilbenes, chalcones, and 6'-deoxychalcones. *Journal of Biological Chemistry*, 270(14), 7922-7928.
- Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, 31(2), 455-461.
- Tushar, Basak, S., Sarma, G. C., & Rangan, L. (2010). Ethnomedical uses of zingiberaceous plants of Northeast India. *Journal of Ethnopharmacology*, 132(1), 286-296.
- van der Meer, I. M., Spelt, C. E., Mol, J. N., & Stuitje, A. R. (1990). Promoter analysis of the chalcone synthase (chsA) gene of *Petunia hybrida*: A 67 bp promoter region directs flower-specific expression. *Plant Molecular Biology*, 15(1), 95-109.
- van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark, A. E., . . . Tironi, I. G. (1996). *Biomolecular simulation: The GROMOS96 manual and user guide*. Zürich: Biomos.
- van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastiris, P. L., Karaca, E., . . . Bonvin, A. M. J. J. (2016). The HADDOCK2.2 web server: User-friendly integrative modeling of biomolecular complexes. *Journal of Molecular Biology*, 428(4), 720-725.
- Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., . . . MacKerell, A. D. (2010). CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry*, 31(4), 671-690.
- Veldkamp, J. F. (2013). Nomenclatural Notes on *Boesenbergia Kuntze* (Zingiberaceae). *Philippine Journal of Science*, 142(Special Issue), 215-221.

- Verdan, A. M., Wang, H. C., García, C. R., Henry, W. P., & Brumaghim, J. L. (2011). Iron binding of 3-hydroxychromone, 5-hydroxychromone, and sulfonated morin: Implications for the antioxidant activity of flavonols with competing metal binding sites. *Journal of Inorganic Biochemistry*, 105(10), 1314-1322.
- Vieira, A. C., Marschalk, C., Biavatti, D. C., Lorscheider, C. A., Peralta, R. M., & Seixas, F. A. V. (2015). Modeling Based Structural insights into biodegradation of the herbicide diuron by laccase-1 from *Ceriporiopsis subvermispora*. *Bioinformation*, 11(5), 224-228.
- Vladilo, G., & Hassanali, A. (2018). Hydrogen Bonds and Life in the Universe. *Life*, 8(1), 1.
- Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *Journal of Molecular Graphics*, 8(1), 52-56.
- Vyas, V. K., Ukawala, R. D., Ghate, M., & Chintha, C. (2012). Homology modeling a fast tool for drug discovery: Current perspectives. *Indian Journal of Pharmaceutical Sciences*, 74(1), 1-17.
- Wang, B., Li, L., Hurley, T. D., & Meroueh, S. O. (2013). Molecular recognition in a diverse set of protein–ligand interactions studied with molecular dynamics simulations and end-point free energy calculations. *Journal of Chemical Information and Modeling*, 53(10), 2659-2670.
- Wang, C., Nguyen, P. H., Pham, K., Huynh, D., Le, T.-B. N., Wang, H., . . . Luo, R. (2016). Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis. *Journal of Computational Chemistry*, 37(27), 2436-2446.
- Wang, J., Cieplak, P., & Kollman, P. A. (2000). How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry*, 21(12), 1049-1074.
- Wani, T. A., Pandith, S. A., Gupta, A. P., Chandra, S., Sharma, N., & Lattoo, S. K. (2017). Molecular and functional characterization of two isoforms of chalcone synthase and their expression analysis in relation to flavonoid constituents in *Grewia asiatica* L. *PLoS ONE*, 12(6), e0179155.
- Wannapinpong, S., Srikulnath, K., Thongpan, A., Choowongkamon, K., & Peyachoknagul, S. (2013). Molecular cloning and characterization of the CHS gene family in turmeric (*Curcuma longa* Linn.). *Journal of Plant Biochemistry and Biotechnology*, 24(1), 25-33.

- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., . . . Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(1), 296-303.
- Weiner, P. K., & Kollman, P. A. (1981). AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *Journal of Computational Chemistry*, 2(3), 287-303.
- Wilkins, M. R., Lindskog, I., Gasteiger, E., Bairoch, A., Sanchez, J. C., Hochstrasser, D. F., & Appel, R. D. (1997). Detailed peptide characterization using PEPTIDEMASS-A World-Wide-Web-accessible tool. *Electrophoresis*, 18(3-4), 403-408.
- Wingender, R., Röhrig, H., Hörnicke, C., Wing, D., & Schell, J. (1989). Differential regulation of soybean chalcone synthase genes in plant defence, symbiosis and upon environmental stimuli. *Molecular and General Genetics*, 218(2), 315-322.
- Winkel-Shirley, B. (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology*, 126(2), 485.
- Xie, H., Li, Y., Yu, F., Xie, X., Qiu, K., & Fu, J. (2015). An investigation of molecular docking and molecular dynamic simulation on imidazopyridines as b-raf kinase inhibitors. *International Journal of Molecular Sciences*, 16(11), 27350-27361.
- Yu, D., Xu, F., Zeng, J., & Zhan, J. (2012). Type III polyketide synthases in natural product biosynthesis. *IUBMB Life*, 64(4), 285-295.
- Zaeoung, S., Plubrukarn, A., & Keawpradub, N. (2005). Cytotoxic and free radical scavenging activities of Zingiberaceous rhizomes. *Songklanakarinn Journal of Sciences and Technology*, 27(4), 799-812.
- Zhao, Y., Zeng, C., & Massiah, M. A. (2015). Molecular dynamics simulation reveals insights into the mechanism of unfolding by the a130t/v mutations within the mid1 zinc-binding bbox1 domain. *PLoS ONE*, 10(4), e0124377.
- Zhong, P., Wu, L., Qian, Y., Fang, Q., Liang, D., Wang, J., . . . Liang, G. (2015). Blockage of ROS and NF- κ B-mediated inflammation by a new chalcone L6H9 protects cardiomyocytes from hyperglycemia-induced injuries. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1852(7), 1230-1241.

Zhou, P., Huang, J., & Tian, F. (2012). Specific noncovalent interactions at protein-ligand interface: Implications for rational drug design. *Current Medicinal Chemistry*, 19(2), 226-238.

Zwanzig, R. W. (1954). High - temperature equation of state by a perturbation method. I. Nonpolar gases. *The Journal of Chemical Physics*, 22(8), 1420-1426.

University of Malaya

LIST OF PUBLICATIONS AND PAPERS PRESENTED

PUBLICATION

1. **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.

PAPER PRESENTED

1. **Sanmugavelan, R** & Teoh, T. C. (2017, Dec). *In silico* study of chalcone synthase variant 2 of *Boesenbergia rotunda*. Paper presented at the Biological Sciences Graduate Congress, Singapore.

University of Malaya

In vitro and in silico studies of chalcone synthase variant 2 in *Boesenbergia rotunda* and its substrate specificity

Ragaventhan SANMUGAVELAN, Teow Chong TEOH*, Nurnadiah ROSLAN, Zulqarnain MOHAMED
Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia

Received: 31.10.2017 • Accepted/Published Online: 29.03.2018 • Final Version: 13.06.2018

Abstract: In this study, transformation of *BrCHS var 2* into *B. rotunda* cell suspension culture, followed by chalcone synthase enzymatic assay and HPLC analysis was conducted to investigate whether the substrate specificity for *BrCHS var 2* is either cinnamoyl-CoA or *p*-coumaroyl-CoA. The HPLC profile showed an increase in the amount of pinocembrin chalcone when cinnamoyl-CoA and malonyl-CoA were added but not *p*-coumaroyl-CoA. Molecular docking was performed to explore the binding of cinnamoyl-CoA and *p*-coumaroyl-CoA to *BrCHS var 2* receptor and the docking results showed that cinnamoyl-CoA formed numerous hydrogen bonds and more negative docked energy than *p*-coumaroyl-CoA. Cinnamoyl-CoA showed good interactions with Cys 164 to initiate the subsequent formation of pinocembrin chalcone, whereas the hydroxyl group of *p*-coumaroyl-CoA formed an unfavorable interaction with Gln 161 that caused steric hindrance to subsequent formation of naringenin chalcone. Docked conformation analysis results also showed that malonyl-CoA formed hydrogen bonding with Cys 164, His 303, and Asn 336 residues in *BrCHS var 2*. The results show that cinnamoyl-CoA is the preferred substrate for *BrCHS var 2*.

Key words: Chalcone synthase, cell suspension culture, homology modelling, molecular docking

1. Introduction

Ginger was used as a food, spice, and herbal remedy over 2000 years ago. It is a monocot plant from the family Zingiberaceae that can be found widely in Southeast Asia. Common members of this family include root ginger (*Zingiber officinale*), fingerroot (*Boesenbergia rotunda*), turmeric (*Curcuma longa*), and myoga (*Zingiber mioga* Roscoe). There has been increasing interest in its rhizome as a good source of medical treatment for humans and it has been utilized as a vital source to exhibit inhibitory activities as an anticancer, antimicrobial, antiviral, and antiinflammatory agent (Sohn et al., 2005; Kiat et al., 2006; Voravuthikunchai et al., 2006; Kirana et al., 2007). In most plants, chalcone synthase (CHS; EC 2.3.1.74) is one of the key enzymes involved in the initiation of the flavonoid biosynthesis pathway. CHS is a plant-specific polyketide synthase type III (PKSIII) that forms chalcone by the condensation of one molecule of *p*-coumaroyl CoA with three molecules of malonyl-CoA to form metabolites such as an intermediate naringenin chalcone. To date, there are more than 2000 nonvolatile compounds detected using (LCMS) in fresh ginger and rhizome, but less than 100 have been structurally identified (Koo et al., 2013).

The CHS gene constitutes a multigene family in which it has irregular copy number of *CHS* in different plant species. Published studies have reported that only one *CHS* was found in *Arabidopsis thaliana* (Feinbaum and Ausubel, 1988) and eight in *Sorghum bicolor* and *Glycine max*, while *Vitis vinifera* and *Physcomitrella patens* have three and seventeen copies, respectively (Goto-Yamamoto et al., 2002; Lo et al., 2002; Tuteja et al., 2004; Koduri et al., 2010). Initial results on gene isolation from *B. rotunda* *CHS* (*BrCHS*) revealed multiple *BrCHS* variants that were identified from different parts of the *B. rotunda* plant. *BrCHS variant 2* (*BrCHS var 2*) was found to be expressed predominantly in the rhizome. As such, this variant was chosen to be introduced into *B. rotunda* cell suspension cultures for downstream HPLC analyses.

CHS has broad substrate preference toward aromatic and aliphatic CoA esters (Jez et al., 2002; Samappito et al., 2002; Abe et al., 2007). The first study on evaluating catalytic activity of the CHS enzyme was performed in cell suspension cultures of parsley (*Petroselinum hortense*) (Kreuzaler and Hahlbrock, 1975). *p*-Coumaroyl-CoA was reported to be the most preferential starter molecule for *CHS* in many plants as well as cinnamoyl-CoA, which is catalyzed at a considerable rate as compared to

* Correspondence: ttchong@um.edu.my