



DISCOVERY OF ANTIMALARIAL NATURAL PRODUCTS AGAINST MALARIAL
TARGETS USING BIOAFFINITY MASS SPECTROMETRY

Ben Yang

Submitted in fulfillment of the requirements of the degree of
Doctor of Philosophy
in the School of Natural Sciences and the Eskitis Institute for Drug
Discovery

Supervisors

Professor Ronald J. Quinn

Doctor Yun Feng

Doctor Hoan Vu

February 2016

ABSTRACT

Natural products have proven to be effective in fighting various human diseases. The identification of bioactive natural products is an efficient method of finding new drug leads. Screening of natural product extracts on therapeutic targets for the presence of active compounds is a rapid and effective strategy to identify noncovalent interactions between active ligands from natural product extracts and their therapeutic protein targets. This thesis presents an effort of exploring a novel approach to screen natural product extracts using bioaffinity mass spectrometry against malarial targets and studies the relationship between HPLC retention time and Log *P* values of natural products.

With an aim to identify noncovalent complexes and to obtain chemical information of active compounds at the screening stage as well as exploit the connection between HPLC retention time and Log *P* values of chemical constituents isolated from the active extracts, a research plan containing four major steps has been undertaken in this study. Step one includes understanding of mass spectrometry technique and its approaches in drug discovery screening, understanding of physicochemical properties of antimalarial natural products and using ChemGPS to assess drug potential of reported anti-malarial natural products. Step two involves developing optimized conditions for ESI-FTIMS screening. Step three aims to use the optimized method to screen natural product extracts, and identify noncovalent complexes formed between a natural product ligand and a protein target, obtaining accurate molecular mass of the active components. The active ligands are then isolated and purified. The binding of the pure compounds with proteins is validated to confirm the accomplishment of the research project. In step four, the isolation and structure elucidation of additional compounds from the selected extracts that showed a protein-ligand complex was performed to prove that reversed-phase HPLC is an appropriate method to capture drug-like chemical constituents (Log *P* value of no more than 5).

The thesis reports outcomes against all the aims. For aim 1, analysis of the physicochemical properties of 506 antimalarial natural products and comparison with the physicochemical properties of malaria drugs revealed major differences between the Log *P* distributions and the molecular weight distributions of these two sets. Principle component analysis of the 506 antimalarial natural products and the 30 malaria drugs using ChemGPS-NP indicated that not many antimalarial natural products were toward the cluster of low molecular weight malaria drugs.

For aim 2, the procedure to achieve the optimal ESI-MS condition for detecting protein-small molecule complex has been successfully developed using bCA II as the test target protein. Results showed that methanol can improve the intensity of proteins signals on ESI-FTMS.

For aim 3, three malarial targets were screened with natural product extracts using ESI-FTMS. It was the first time a bioaffinity MS screening method has been used to study the malarial target *PfRab11a* and *PfUCH-L3*, and it is the first time that natural product binders have been reported. A new natural product inhibitor of *PfRab11a* has been reported. In order to predict the binding sites of the active compounds, virtual docking was employed to provide valuable clues for future research.

For aim 4, including the ligands identified by ESI-FTMS, a number of natural products, including flavonoids, alkaloids, triterpenoids and so on, were isolated from the active extracts. Physicochemical property analysis indicated that most of them met Lipinski's Ro5. Their Log *P* value showed a linear relationship with HPLC retention time.

STATEMENT OF ORIGINALITY

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

(Signed)_____

Name of Student

TABLE OF CONTENTS

CHAPTER ONE INTRODUCTION	27
1.1 Malaria	27
1.2 Current clinical agents	28
1.2.1 Nucleic acid antimalarial drugs	29
1.2.2 Blood schizontocytes	31
1.2.3 Antibiotic based malaria drugs	34
1.3 Drug resistance	34
1.4 Malarial targets	36
1.4.1 <i>Pfd</i> UTPase	36
1.4.2 <i>Pf</i> Rab11a	39
1.4.3 <i>Pf</i> UCH-L3	40
1.5 Protein-ligand noncovalent interactions	42
1.6 Mass spectrometry (MS)	45
1.6.1 General	46
1.6.2 Ion sources	47
1.6.2.1 MALDI	48
1.6.2.2 ESI	49
1.6.3 Analyzers	51
1.6.3.1 TOF	51
1.6.3.2 Quadrupole	52
1.6.3.3 Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS, FT-MS)	52
1.6.3.3.1 FT-MS theory	52
1.6.3.3.2 FT-MS applications	55
1.7 Lead-like enhanced extracts (LLEE) library	56
1.8 Objectives	57

1.9	Research plan	57
CHAPTER TWO PREDICTING NATURAL PRODUCT VALUE, AN EXPLORATION OF ANTIMALARIAL DRUG SPACE		59
2.1	Introduction	59
2.2	Data and method	60
2.2.1	Data	60
2.2.2	Method	61
2.3	Results and discussion	64
2.3.1	Physicochemical properties of malaria drugs	65
2.3.2	Physicochemical properties of antimalarial natural products (506) reported in the time frame of 1990-2012	68
2.3.3	Comparison of antimalarial natural product physicochemical space vs current malaria drugs	70
2.4	Summary	72
CHAPTER THREE OPTIMIZATION FOR ELECTROSPRAY IONIZATION FOURIER TRANSFORM MASS SPECTROMETRY		74
3.1	Introduction	74
3.2	Instrument used in this thesis	75
3.2.1	ESI source	76
3.2.2	Ion transfer line	79
3.2.3	ICR cell	80
3.3	Materials and methods	80
3.3.1	Materials	80
3.3.2	Methods	81
3.4	Results and discussion	82
3.4.1	Sensitivity of protein in ammonium acetate	82
3.4.2	Influence of methanol on protein	83
3.4.3	Influence of trifluoroethanol on protein	84

3.4.4	Influence of different ligand/protein ratios	85
3.5	Summary	86
CHAPTER FOUR SCREENING LEAD-LIKE ENHANCED EXTRACTS AGAINST		87
PFDUTPASE USING ESI-FTMS		
4.1	Introduction	87
4.1.1	<i>PfdUTPase</i>	87
4.1.2	Lead-like enhanced extract library	87
4.2	Materials and methods	88
4.2.1	<i>PfdUTPase</i> and buffer exchange	88
4.2.2	A known binder to <i>PfdUTPase</i>	89
4.2.3	LLEE library establishment	90
4.2.4	Bioaffinity ESI-FTMS screening	91
4.2.5	<i>PfdUTPase</i> expression and purification	91
4.3	Results and discussion	92
4.3.1	Desalting and buffer exchange	93
4.3.2	Detection of complex formed between <i>PfdUTPase</i> and its known ligand in NP extract matrix	93
4.3.3	LLEE library and Bioaffinity ESI-FTMS screening	95
4.3.4	<i>PfdUTPase</i> expression and purification	96
4.4	Summary	97
CHAPTER FIVE SCREENING LEAD-LIKE ENHANCED EXTRACTS AGAINST PFRAB11A		99
USING ESI-FTMS		
5.1	Introduction	100
5.2	Results and discussion	101
5.2.1	Detection of protein by ESI-FTMS in physiological pH	101
5.2.2	Detection of complexes formed by protein and active components in natural product extracts	101
5.2.3	Determination of ligand molecular mass	103

5.2.4	Structure elucidation of natural products and the analysis of their physicochemical properties	105
5.2.4.1	Structure elucidation	105
5.2.4.2	The physicochemical property analysis	147
5.2.5	Determination of K_D	152
5.2.6	Determination of IC_{50}	154
5.2.7	Evaluation of possible <i>PfRab11a</i> binding sites of those natural product binders via virtual docking	155
5.3	Summary	160
5.4	Experimental	160
5.4.1	<i>PfRab11a</i>	160
5.4.2	Natural product extracts	161
5.4.3	ESI-FTMS analysis	161
5.4.4	HPLC, MS and NMR instruments	162
5.4.5	Extraction and isolation	162
5.4.5.1	Chemical constituents from the Australian plant <i>Psydrax montigena</i>	163
5.4.5.2	Chemical constituents from the Australian plant <i>Wendlandia basistaminea</i>	165
5.4.5.3	Chemical constituents from the Australian plant <i>Psydrax lamprophylla</i>	167
5.4.5.4	Chemical constituents from the Australian plant <i>Prunus persica</i>	169
5.4.5.5	Chemical constituents from the Australian plant <i>Larsenaikia ochreatea</i>	172
5.4.6	Dissociation constant (K_D) determination	174
5.4.7	Enzyme assay	175
5.4.8	Virtual docking	176
CHAPTER SIX SCREENING LEAD-LIKE ENHANCED EXTRACTS AGAINST PFUCH-L3		178

USING ESI-FTMS

6.1	Introduction	178
6.2	Results and discussion	179
6.2.1	Detection of protein by ESI-FTMS in physiological pH	179
6.2.2	Detection of complexes formed by protein and active components in natural product extracts	179
6.2.3	Determination of ligand molecular mass	181
6.2.4	Structure elucidation of natural products and the analysis of their physicochemical properties	182
6.2.4.1	Structure elucidation	182
6.2.4.2	Physicochemical property analysis	200
6.2.5	Determination of K_D	202
6.2.6	Valuation of possible <i>PfUCH-L3</i> binding sites of those natural product binders via virtual docking	203
6.3	Summary	206
6.4	Experimental	207
6.4.1	<i>PfUCH-L3</i>	207
6.4.2	Natural product extracts	207
6.4.3	ESI-FTMS analysis	207
6.4.4	HPLC, MS and NMR instruments	208
6.4.5	Extraction and isolation	208
6.4.5.1	Chemical constituents from the Australian plant <i>Prunus Rosaceae</i>	209
6.4.5.2	Isolation of target compound from the sponge <i>Ianthella</i> <i>quadrangulata</i>	211
6.4.5.3	Chemical constituents from the Australian plant <i>Canthium</i> <i>Rubiaceae</i>	212
6.4.6	Dissociation constant (K_D) determination	214
6.4.7	Virtual docking	214

DISCUSSION AND CONCLUSION	216
APPENDICES	220
REFERENCES	227

LIST OF FIGURES

Figure 1.1.1	Countries with ongoing transmission of malaria, 2013	27
Figure 1.2.1.1	Chemical structures of sulfadoxine (1), pyrimethamine (2), biguanides (proguanil) (3), cycloguanil (4), sulfalene (5) and dapsone (6).	29
Figure 1.2.1.2	Chemical structure of atovaquone (7)	30
Figure 1.2.2.1	Chemical structures of chloroquine (8), amodiaquine (9), pyronaridine (10), lumefantrine (11), quinine (12), mefloquine (13), primaquine (14), halofantrine (15) and piperaquine (16).	31
Figure 1.2.2.2	Chemical structures of artemether (17), artemisinin (18), dihydroartemisinin (19) and artesunate (20).	33
Figure 1.2.3.1	Chemical structures of tetracycline (21), doxycycline (22), and clindamycin (23)	34
Figure 1.4.1.1	<i>Pfd</i> UTPase with 5'-Diphenyl Nucleoside Inhibitors (PDB ID: 3T60)	36
Figure 1.4.1.2	Chemical structures of the securinine fragment series shown to bind <i>Pfd</i> UTPase	39
Figure 1.4.2.1	Crystal structure of <i>Pf</i> Rab11a in complex with GDP (PDB ID: 3BFK)	39
Figure 1.4.3.1	Crystal structure of <i>Pf</i> UCH-L3 (PDB ID: 2WE6)	40
Figure 1.5.1	“Lock and Key” model: The green key (ligand) has the right shape for the red lock (protein) and a ligand-protein complex can be formed. The blue key	42

	with unsuitable shape thus no complex is formed.	
Figure 1.5.2	An illustration of conformational change in protein acquired by ESI-MS. (a) Charge states for native protein. (b) Charge states for denatured protein.	45
Figure 1.6.1.1	Main components of a mass spectrometer	47
Figure 1.6.2.1.1	A schematic of MALDI	48
Figure 1.6.2.2.1	Electrospray ionization	49
Figure 1.6.2.2.2	Two ESI mechanisms: CRM and IDM	50
Figure 1.6.3.1.1	TOF Schematic	51
Figure 1.6.3.1.2	Equations for TOF	52
Figure 1.6.3.2.1	Quadrupole Mass Spectrometer Schematic	52
Figure 1.7.1	Example of a lead-like enhanced extract HPLC chromatogram	57
Figure 2.2.2.1	Step One	61
Figure 2.2.2.2	Step Two and Three	62
Figure 2.2.2.3	Step Four and Five	62
Figure 2.2.2.4	Step Six and Seven	63
Figure 2.2.2.5	Step Eight. Each dot represents one molecule.	63
Figure 2.3.1.1	Pie chart presentation of the percentage of malaria drugs compliant with Lipinski's Ro5	66
Figure 2.3.1.2	Structure of solithromycin (24)	66
Figure 2.3.1.3	Physicochemical property histograms of malaria drugs: a) MW, b) calculated Log <i>P</i> , c) HBD, and d) HBA.	67

Figure 2.3.2.1	Pie chart presentation of the percentage of antimalarial natural products compliant with Lipinski's Ro5	68
Figure 2.3.2.2	Physicochemical property histograms of antimalarial natural products: a) MW, b) calculated Log <i>P</i> , c) HBD, and d) HBA.	69
Figure 2.3.3.1	Score plot of malaria drugs.	70
Figure 2.3.3.2	Score plot of antimalarial natural products.	71
Figure 2.3.3.3	Comparison of the physicochemical space of 506 antimalarial natural products (blue dots) with physicochemical space of 30 malaria drugs and candidates (red dots).	72
Figure 3.2.1	ESI-FTICR-MS, Bruker 4.7 Tesla	75
Figure 3.2.2	A schematic cutaway view of ESI-FTICR-MS	76
Figure 3.2.1.1	Apollo ESI source diagram	77
Figure 3.2.1.2	Xmass tune page for ESI source	78
Figure 3.2.2.1	A schematic of the ion transfer line	79
Figure 3.2.2.2	Electric potential values of the ion transfer optic line. A decrease voltage value from +1 V to -2.5 kV helps to accelerate the ions. An increase voltage from -2.5 kV to +1 V makes the ions slowly go into the ICR cell.	79
Figure 3.2.3.1	A drawing of Bruker ICR cell	80

Figure 3.4.1.1	Protein intensity in 10 mM ammonium acetate	83
Figure 3.4.2.1	Influence of methanol on protein	83
Figure 3.4.2.2	A spectrum of bCA II in 10mM ammonium acetate with 20% MeOH.	84
Figure 3.4.3.1	Influence of TFE on protein.	84
Figure 3.4.4.1	Compared different ligand/protein ratios.	85
Figure 3.4.4.2	A spectrum of bCA II and its complex. For instance, 2644.48 is the signal of [bCA II] ¹¹⁺ , on its right hand side is the signal of [bCA II-ligand complex] ¹¹⁺	86
Figure 4.1.2.1	Example of a lead-like enhanced extract HPLC chromatogram	88
Figure 4.2.1.1	Buffer exchange steps. (1) Column Preparation. Excess storage buffer is allowed to flow through the column. (2) Column Equilibration. Column is equilibrated with buffer chosen by user (ammonium bicarbonate in this case). (3) Sample application. Sample is applied to column. (4) Elution. Purified sample is eluted from the column with buffer chosen by user (ammonium bicarbonate in this case).	89
Figure 4.2.2.1	Chromatogram of chosen LLEF extract.	90
Figure 4.3.1.1	MS spectrum of <i>Pfd</i> UTPase without buffer exchange	93
Figure 4.3.1.2	MS spectrum of buffer exchanged <i>Pfd</i> UTPase	93
Figure 4.3.2.1	Comparison of protein-ligand complexes formed by	94

*Pfd*UTPase with ligand in different concentrations.

Figure 4.3.4.1	Lane 1: Marker (15, 30, 68 kDa); Lane 2: fraction (fr) 20/2; Lane 3: fr 20/5; Lane 4: fr 50/2; Lane 5: fr 50/5; Lane 6: fr 100/2; Lane 7: fr 100/5; Lane 8: fr 200/2; Lane 9: fr 200/5; Lane 10: 500/2. (For instance, fr 20/2 means this is fraction 2 of buffer containing 20 mM imidazole)	97
Figure 4.3.4.2	ESI-FTMS spectrum of <i>Pfd</i> UTPase (15.2 μ M after buffer exchange) made here. m/z 2034.8: [dUTPase] ¹⁰⁺ ; m/z 2260.8: [dUTPase] ⁹⁺ .	97
Figure 5.2.1.1	Mass spectrum of <i>Pf</i> Rab11a acquired under native condition	101
Figure 5.2.2.1	Mass spectrum of <i>Pf</i> Rab11a-ligand Complex One	102
Figure 5.2.2.2	Mass spectrum of <i>Pf</i> Rab11a-ligand Complex Two	102
Figure 5.2.2.3	Mass spectrum of <i>Pf</i> Rab11a-ligand Complex Three	103
Figure 5.2.2.4	Mass spectrum of <i>Pf</i> Rab11a-ligand Complex Four	103
Figure 5.2.2.5	Mass spectrum of <i>Pf</i> Rab11a-ligand Complex Five	103
Figure 5.2.4.1.1	Chemical structure and ¹ H NMR spectrum of 29	105
Figure 5.2.4.1.2	Chemical structure and ¹³ C NMR spectrum of 29	106
Figure 5.2.4.1.3	Key COSY (in bold) and HMBC correlations (H→C) for 29	107
Figure 5.2.4.1.4	Chemical structure and ¹ H NMR spectrum of compound 33	109

Figure 5.2.4.1.5	Chemical structure and ^1H NMR spectrum of 39	110
Figure 5.2.4.1.6	Chemical structure and ^1H NMR spectrum of 45	111
Figure 5.2.4.1.7	Chemical structure and ^1H NMR spectrum of 52	111
Figure 5.2.4.1.8	Chemical structure and ^{13}C NMR spectrum of 52	112
Figure 5.2.4.1.9	Chemical structure and ^1H NMR spectrum of 27	113
Figure 5.2.4.1.10	Chemical structure and ^{13}C NMR spectrum of 27	113
Figure 5.2.4.1.11	Chemical structure and ^1H NMR spectrum of 28	114
Figure 5.2.4.1.12	Chemical structure and ^{13}C NMR spectrum of 28	114
Figure 5.2.4.1.13	Chemical structure and ^1H NMR spectrum of 30	116
Figure 5.2.4.1.14	Chemical structure and ^{13}C NMR spectrum of 30	116
Figure 5.2.4.1.15	Chemical structure and ^1H NMR spectrum of 31	117
Figure 5.2.4.1.16	Chemical structure and ^{13}C NMR spectrum of 31	118
Figure 5.2.4.1.17	Chemical structure and ^1H NMR spectrum of 32	119
Figure 5.2.4.1.18	Chemical structure and ^1H NMR spectrum of 34	120
Figure 5.2.4.1.19	Chemical structure and ^{13}C NMR spectrum of 34	120
Figure 5.2.4.1.20	Chemical structure and ^1H NMR spectrum of 35	121
Figure 5.2.4.1.21	Chemical structure and ^{13}C NMR spectrum of 35	122
Figure 5.2.4.1.22	Chemical structure and ^1H NMR spectrum of 36	123
Figure 5.2.4.1.23	Chemical structure and ^{13}C NMR spectrum of 36	123

Figure 5.2.4.1.24	Chemical structure and ^1H NMR spectrum of 37	124
Figure 5.2.4.1.25	Chemical structure and ^1H NMR spectrum of 38	125
Figure 5.2.4.1.26	Chemical structure and ^{13}C NMR spectrum of 38	126
Figure 5.2.4.1.27	Chemical structure and ^1H NMR spectrum of 40	127
Figure 5.2.4.1.28	Chemical structure and ^{13}C NMR spectrum of 40	127
Figure 5.2.4.1.29	Chemical structure and ^1H NMR spectrum of 41	128
Figure 5.2.4.1.30	Chemical structure and ^{13}C NMR spectrum of 41	128
Figure 5.2.4.1.31	Chemical structure and ^1H NMR spectrum of 42	129
Figure 5.2.4.1.32	Chemical structure and ^1H NMR spectrum of 43	130
Figure 5.2.4.1.33	Chemical structure and ^{13}C NMR spectrum of 43	130
Figure 5.2.4.1.34	Chemical structure and ^1H NMR spectrum of 44	131
Figure 5.2.4.1.35	Chemical structure and ^1H NMR spectrum of 46	132
Figure 5.2.4.1.36	Chemical structure and ^{13}C NMR spectrum of 46	133
Figure 5.2.4.1.37	Chemical structure and ^1H NMR spectrum of 47	134
Figure 5.2.4.1.38	Chemical structure and ^{13}C NMR spectrum of 47	134
Figure 5.2.4.1.39	Chemical structure and ^1H NMR spectrum of 48	135
Figure 5.2.4.1.40	Chemical structure and ^{13}C NMR spectrum of 48	136
Figure 5.2.4.1.41	Chemical structure and ^1H NMR spectrum of 49	137
Figure 5.2.4.1.42	Chemical structure and ^{13}C NMR spectrum of 49	137

Figure 5.2.4.1.43	Chemical structure and ^1H NMR spectrum of 50	138
Figure 5.2.4.1.44	Chemical structure and ^{13}C NMR spectrum of 50	139
Figure 5.2.4.1.45	Chemical structure and ^1H NMR spectrum of 51	140
Figure 5.2.4.1.46	Chemical structure and ^{13}C NMR spectrum of 51	140
Figure 5.2.4.1.47	Chemical structure and ^1H NMR spectrum of 53	141
Figure 5.2.4.1.48	Chemical structure and ^{13}C NMR spectrum of 53	142
Figure 5.2.4.1.49	Chemical structure and ^1H NMR spectrum of 54	143
Figure 5.2.4.1.50	Chemical structure and ^{13}C NMR spectrum of 54	143
Figure 5.2.4.1.51	Chemical structure and ^1H NMR spectrum of 55	144
Figure 5.2.4.1.52	Chemical structure and ^{13}C NMR spectrum of 55	145
Figure 5.2.4.1.53	Chemical structure and ^1H NMR spectrum of 56	146
Figure 5.2.4.1.54	Chemical structure and ^{13}C NMR spectrum of 56	146
Figure 5.2.4.2.1	Natural products isolated from the Australian plant <i>Psydrax montigena</i>	149
Figure 5.2.4.2.2	Natural products isolated from the Australian plant <i>Wendlandia basistaminea</i>	149
Figure 5.2.4.2.3	Natural products isolated from the Australian plant <i>Psydrax lamprophylla</i>	150
Figure 5.2.4.2.4	Natural products isolated from the Australian plant <i>Prunus persica</i>	150

Figure 5.2.4.2.5	Natural products isolated from the Australian plant <i>Larsenaikia ochreata</i>	151
Figure 5.2.4.2.6	Relationship between Log <i>P</i> and Retention Time (min)	151
Figure 5.2.5.1	Titration of 5 ligands with <i>PfRab11a</i> for determination of K_D . L ₃ (39), L ₄ (45) and L ₅ (52) didn't give curves as their high K_D values.	153
Figure 5.2.6.1	IC ₅₀ calculations	154
Figure 5.2.7.1	Ligand-protein interaction diagrams of the selected ligands generated by LIGPLOT.	157
Figure 5.2.7.2	Ligand-protein interaction diagrams of the selected ligands generated by LIGPLOT (Rab2_neg).	158
Figure 5.2.7.3	Compound 29 (Rab1), 33 (Rab2), 39 (Rab3), 45 (Rab4) and 52 (Rab5) bound to four different pockets in the <i>PfRab11a</i> (3BFK) binding site. The ligands are shown in stick model while the binding site is depicted as a surface for better visualization.	159
Figure 5.2.7.4	Compound 33 (Rab2_neg) bound to <i>PfRab11a</i> (3BFK). The ligand is shown in stick model while the binding site is depicted as a surface for better visualization.	159
Figure 5.4.5.1.1	HPLC chromatogram of <i>Psydrax montigena</i> sp.	164
Figure 5.4.5.2.1	HPLC chromatogram of <i>Wendlandia basistaminea</i> sp.	166
Figure 5.4.5.3.1	HPLC chromatogram of <i>Psydrax lamprophylla</i> sp.	168
Figure 5.4.5.4.1	HPLC chromatogram of <i>Prunus persica</i> sp.	170
Figure 5.4.5.5.1	HPLC chromatogram of <i>Larsenaikia ochreata</i> sp.	172

Figure 6.2.1.1	Mass spectrum of <i>Pf</i> UCH-L3 acquired under native condition	179
Figure 6.2.2.1	Mass spectrum of <i>Pf</i> UCH-L3-ligand Complex One	180
Figure 6.2.2.2	Mass spectrum of <i>Pf</i> UCH-L3-ligand Complex Two	180
Figure 6.2.2.3	Mass spectrum of <i>Pf</i> UCH-L3-ligand Complex Three	181
Figure 6.2.4.1.1	Chemical structure and ¹ H NMR spectrum of 58	182
Figure 6.2.4.1.2	Chemical structure and ¹ H NMR spectrum of 63	183
Figure 6.2.4.1.3	Chemical structure and ¹³ C NMR spectrum of 63	183
Figure 6.2.4.1.4	Chemical structure and ¹ H NMR spectrum of 68	184
Figure 6.2.4.1.5	Chemical structure and ¹³ C NMR spectrum of 68	185
Figure 6.2.4.1.6	Chemical structure and ¹ H NMR spectrum of 57	186
Figure 6.2.4.1.7	Chemical structure and ¹³ C NMR spectrum of 57	186
Figure 6.2.4.1.8	Chemical structure and ¹ H NMR spectrum of 59	187
Figure 6.2.4.1.9	Chemical structure and ¹³ C NMR spectrum of 59	188
Figure 6.2.4.1.10	Chemical structure and ¹ H NMR spectrum of 60	189
Figure 6.2.4.1.11	Chemical structure and ¹³ C NMR spectrum of 60	189
Figure 6.2.4.1.12	Chemical structure and ¹ H NMR spectrum of 61	190
Figure 6.2.4.1.13	Chemical structure and ¹³ C NMR spectrum of 61	191
Figure 6.2.4.1.14	Chemical structure and ¹ H NMR spectrum of 62	192
Figure 6.2.4.1.15	Chemical structure and ¹³ C NMR spectrum of 62	192

Figure 6.2.4.1.16	Chemical structure and ^1H NMR spectrum of 64	193
Figure 6.2.4.1.17	Chemical structure and ^{13}C NMR spectrum of 64	194
Figure 6.2.4.1.18	Chemical structure and ^1H NMR spectrum of 65	195
Figure 6.2.4.1.19	Chemical structure and ^{13}C NMR spectrum of 65	195
Figure 6.2.4.1.20	Chemical structure and ^1H NMR spectrum of 66	196
Figure 6.2.4.1.21	Chemical structure and ^{13}C NMR spectrum of 66	196
Figure 6.2.4.1.22	Chemical structure and ^1H NMR spectrum of 67	197
Figure 6.2.4.1.23	Chemical structure and ^{13}C NMR spectrum of 67	198
Figure 6.2.4.1.24	Chemical structure and ^1H NMR spectrum of 69	199
Figure 6.2.4.1.25	Chemical structure and ^{13}C NMR spectrum of 69	199
Figure 6.2.4.2.1	Natural products isolated from the Australian plant <i>Prunus Rosaceae</i>	201
Figure 6.2.4.2.2	Natural products isolated from the Australian plant <i>Canthium Rubiaceae</i>	201
Figure 6.2.4.2.3	Relationship between Log <i>P</i> and Retention Time (min). Each blue dot represents a natural product.	202
Figure 6.2.5.1	Titration of 3 ligands with <i>PfUCH-L3</i> for determination of K_D . Upper curve belongs to L_2 (63) and the other curve belongs to L_3 (68). L_1 (58) didn't give a curve as its high K_D value.	203
Figure 6.2.6.1	Ligand-protein interaction diagrams of the selected	205

ligands generated by LIGPLOT.

Figure 6.2.6.2	Ligands are shown in stick model while the binding site is depicted as a surface for better visualization. Structural data was from PDB structures 2WE6.	206
Figure 6.4.5.1.1	HPLC chromatogram of <i>Prunus Rosaceae</i>	210
Figure 6.4.5.2.1	HPLC chromatogram of <i>Ianthella quadrangulata</i>	212
Figure 6.4.5.3.1	HPLC chromatogram of <i>Canthium Rubiaceae</i>	213

LIST OF TABLES

Table 1.4.1.1	Structures of some <i>Pfd</i> UTPase inhibitors	37
Table 1.4.1.2	Inhibition activity of nucleosides	38
Table 1.5.1	Advantage and disadvantages of some biophysical methods to study protein-ligand interactions	43
Table 1.6.1.1	Progresses in MS	46
Table 2.3.1.1	Physicochemical properties of 31 approved malaria drugs and candidates in clinical trials.	65
Table 3.3.1.1	Structure and K_D of ethoxzolamide, an inhibitor of bCA II	81
Table 4.2.2.1	Structure of the known binder (26) to <i>Pfd</i> UTPase	89
Table 4.3.3.1	Sources of the samples in Nature Bank	95
Table 5.2.4.1.1	NMR spectroscopic data for arborside E	108
Table 5.2.4.2.1	Physicochemical properties of compounds 27-56	147
Table 5.2.5.1	Results for the determination of K_D (μ M)	153
Table 5.2.6.1	Results for the determination of IC_{50} (μ M)	154
Table 5.2.7.1	Docking scores of the ligands interacting with five different binding sites in 3BFK predicted by SiteMap.	155
Table 5.4.5.1	Taxonomy information of active extracts	162
Table 6.2.4.2.1	Physicochemical properties of compounds 57-69	200
Table 6.2.5.1	Results for the determination of K_D (μ M)	203
Table 6.2.6.1	Docking scores of the ligands interacting with five different binding sites in 2WE6 predicted by SiteMap.	204
Table 6.4.5.1	Taxonomy information of active extracts	208

ABBREVIATIONS

ACT	Artemisinin combination therapy
BAMS	Bioaffinity mass spectrometry screening
bCA II	Bovine carbonic anhydrase II
COSY	(Homonuclear chemical shift) Correlation spectroscopy
CRM	Charge residue model
dUTPase	desoxyuridine 5'-triphosphate nucleotidohydrolase
ESI	Electrospray ionization
FT-MS or FTMS	Fourier transform mass spectrometry
GC	Gas chromatography
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum correlation
ICR	Ion Cyclotron Resonance
IDM	Ion desorption model
LC	Liquid chromatography
LLEE	Lead-like enhanced extract
LLEF	Lead-like enhanced fraction
<i>m/z</i>	mass-to-charge ratio
MALDI	Matrix-Assisted Laser Desorption Ionization
MeOH	Methanol
MHz	mega hertz
NMR	Nuclear magnetic resonance
NP	Natural product
RF	Radio frequency
<i>Pf</i>	<i>Plasmodium falciparum</i>
Ro5	Rule of five
ROESY	Rotating frame Overhauser effect spectroscopy

SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel
SEC	Size-exclusion chromatography
SPE	Solid-phase extraction
TFE	Trifluoroethanol
TOF	Time of flight
UCH-L3	Ubiquitin carboxyl-terminal esterase L3

ACKNOWLEDGMENTS

Looking back on the four years of PhD research and life here in Eskitis Institute bit by bit, I cannot help a myriad of thoughts. Ph.D. is a long and fulfilling process. Without the encouragement and support from my dear supervisors, friends and family members, I was not able to get it. In occasion of the completion of this thesis, I would like to express my most sincere gratitude to those people.

This thesis is written under the guidance of my supervisors Professor Ron Quinn, Doctor Yun Jiang Feng and Doctor Hoan Vu. Many thanks to Ron for giving me the chance to start my PhD in Australia and also for his guidance and all kinds of support throughout my PhD studies. I would like to thank Yun for her efforts and valuable suggestions to my chemistry experiments and manuscripts. I would like to thank Hoan for his help with the mass spectrometry.

During the last four years, I am fortunate to get to know lots of kind people in Eskitis Institute. Thanks to all of you.

ALL PAPERS INCLUDED ARE CO-AUTHORED

Acknowledgement of Papers included in this Thesis

Included in this thesis is one paper in *Chapter Five* which is co-authored with other researchers. My contribution to this co-authored paper is outlined at the front of the relevant chapter. The bibliographic details (if published or accepted for publication)/status (if prepared or submitted for publication) for this paper including all authors, are:

Paper name: Bioaffinity Mass Spectrometry Screening.

Authors: Ben Yang, Yun Jiang Feng, Hoan Vu, Brendan McCormick, Jessica Rowley, Liliana Pedro, Gregory J. Crowther, Wesley C. Van Voorhis, Paul I. Forster, and Ronald J. Quinn.

This paper was accepted for publication by the Journal of Biomolecular Screening in November, 2015.

Appropriate acknowledgements of those who contributed to the research but did not qualify as authors are included in this paper.

(Signed) _____ (Date) _____

Name of Student

(Countersigned) _____ (Date) _____

Name of Supervisor

CHAPTER ONE INTRODUCTION

Abstract: This chapter presents the background of this project, including malaria, current clinical agents, drug resistance, malarial targets, protein-ligand noncovalent interactions, mass spectrometry and natural product lead-like enhanced extract library. Malaria is one of the most global harmful acute parasitic diseases. Due to the severe drug resistance, it is urgent to search for new antimalarial therapeutics. Plasmodium falciparum dUTPase, Rab11a and UCH-L3 are three potential malarial targets involved in this research. Natural product lead-like enhanced extracts, being rich in small molecules with high drug-like properties, can be a source of the future therapeutics. Active drug compounds can form noncovalent complexes with the therapeutic protein targets. Bioaffinity mass spectrometry, which is a speed and sensitive approach, will be employed to detect the noncovalent binding between malarial protein targets and active components from lead-like enhanced extracts. These binding compounds will be developed further for becoming antimalarial drugs.

1.1 Malaria

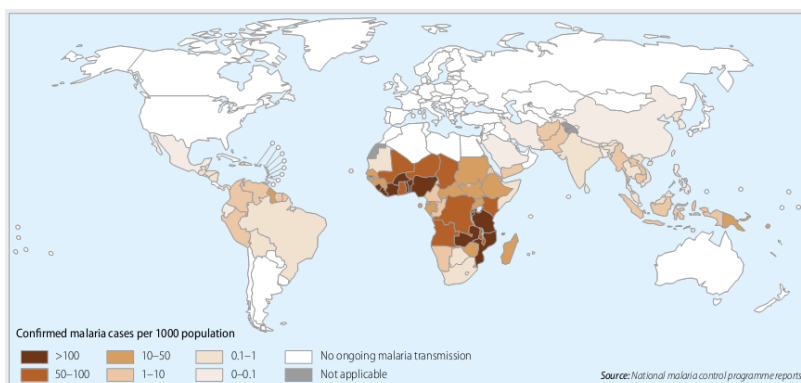


Figure 1.1.1 Countries with ongoing transmission of malaria, 2013.¹ (Permission authorized by WHO)

Malaria is one of the most harmful infectious diseases in the world. According to the World Malaria Report, 97 countries and territories had ongoing malaria transmission (Figure 1.1.1). An estimated 3.3 billion people were at risk of malaria, of whom 1.2

billion are at high risk.¹ There were an estimated 198 million cases of malaria worldwide (range 124-283 million) in 2013, and an estimated 584,000 deaths (range 367,000-755,000). 90% of all malaria deaths occur in Africa. In 2013, an estimated 437,000 African children died before their fifth birthday due to malaria. Globally, the disease caused an estimated 453,000 under-five deaths in the same year.¹

There are five species of *Plasmodium* that can cause malaria in humans, including *Plasmodium falciparum* (*P. falciparum* or *Pf*), *Plasmodium vivax* (*P. vivax*), *Plasmodium ovale*, *Plasmodium malariae* and *Plasmodium knowlesi*. Malaria due to *P. falciparum* is the most deadly, and it predominates in Africa.¹ Malaria is transmitted to humans by the bite of infected female mosquitoes, which introduces the protists via its saliva into the circulatory system, and ultimately to the liver where they mature and reproduce. The disease causes symptoms that typically include fever and headache, which in severe cases can progress to coma or death.²

The life cycle of malaria parasites is complicated. At first, malarial parasites are transmitted to human when an infected female *Anopheles* mosquito feeds.³ Secondly, infectious sporozoites move to the liver, where they develop and produce thousands of merozoites. The merozoites rapidly invade erythrocytes, grow, reproduce, exit from the cells, and infect new erythrocytes, causing many different kinds of disease symptoms.⁴ After a feeding mosquito ingests some reproductive gametocytes produced by mature parasites in its gut, they will become oocysts that develop, burst, and release sporozoites, which migrate to the mosquito's salivary glands in order to complete the cycle.⁵

1.2 Current clinical agents

Therapy and prevention for malaria depend on a lot of factors such as the type of infection (by which kind of *Plasmodium* parasites), severity of clinical cases and drug resistance. Anti-malarial drugs can be generally classified by their mode of action or chemical nucleus. However, at present, many antimalarial drugs available have not had their mode of action unequivocally elaborated.

1.2.1 Nucleic acid antimalarial drugs

The nucleic acid antimalarial drugs include folate antagonists and atovaquone.

1) Folate Antagonist Drugs

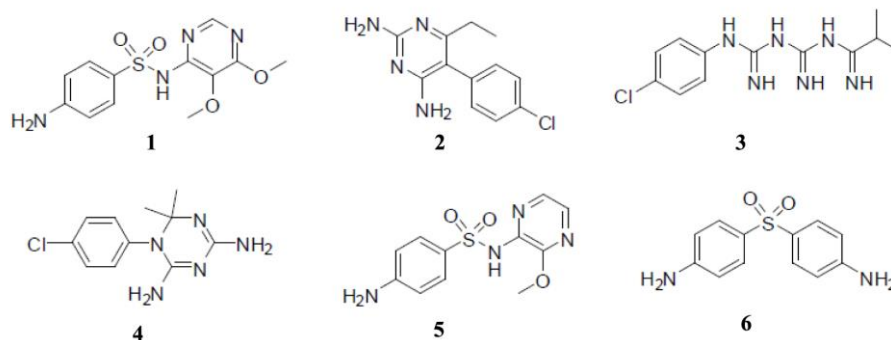


Figure 1.2.1.1 Chemical structures of sulfadoxine (1), pyrimethamine (2), biguanides (proguanil) (3), cycloguanil (4), sulfalene (5) and dapsone (6).

Folate antagonists can inhibit enzymes involved in the folate pathway and pyrimidine synthesis. In other words, they can reduce the quantity of DNA and serine as well as block the methionine formation.⁶ As shown in Figure 1.2.1.1, these kind of drugs consists of sulfadoxine (1), pyrimethamine (2), biguanides (proguanil) (3), cycloguanil (4), sulfalene (5) and dapsone (6).

Sulfalene (5) is a structural analogue to sulfadoxine (1) and also a competitive antagonist of *p*-aminobenzoic acid. Both of them are competitive inhibitors of dihydropteroate synthase, which is an important enzyme responsible for the incorporation of *p*-aminobenzoic acid during the synthetic process of folic acid.⁷

Pyrimethamine (2) currently is only used in synergistic combination with sulfadoxine (1), sulfalene (5) or dapsone (6) for prevention.⁸ By inhibiting plasmodial dihydrofolate reductase (DHFR), it blocks nucleic acids synthesis indirectly in the body of malaria parasite.⁹ The elimination half-life is approximately 100 hours.¹⁰ In the past, it was effective against all kinds of human malaria; however, in Southeast Asia, drug resistance has emerged unexpectedly.¹¹ This drug is also used in the treatment of toxoplasmosis and isosporiasis and can be used in the prevention

against *Pneumocystis carinii* pneumonia.⁷

As a biguanide compound, proguanil (**3**) is metabolized in the human body via the polymorphic cytochrome P450 enzyme CYP2C19 to the active metabolite, cycloguanil (**4**).¹² About 3% of Caucasian and African populations and 20% of asian people are “poor metabolizers” and have seriously suppressed biotransformation from proguanil (**3**) to cycloguanil (**4**).¹³ Furthermore, cycloguanil (**4**) also inhibits plasmodial dihydrofolate reductase.¹³ Proguanil (**3**) has weak antimalarial activity through an unknown mechanism of action. It also has sporontocidal activity, rendering gametocytes non-infective to the mosquito vector.⁸ Unfortunately, drug resistance to proguanil (**3**) grew rapidly, and it is not used alone for malaria therapy any more. Now it is administered as the hydrochloride salt in combination with atovaquone (**7**).⁷

2) Atovaquone (**7**)

Atovaquone (**7**) (Figure 1.2.1.2) is one of the hydroxynaphthoquinone antiparasitic drugs against all *Plasmodium* species.¹⁴ It curbs pre-erythrocytic development in human’s liver and oocyst development in the body of mosquito.¹⁵ It is currently combined with proguanil (**3**) for the treatment of malaria.⁷ Atovaquone (**7**) selectively curbs electron transport through the parasite mitochondrial cytochrome bc1 complex and collapses the mitochondrial membrane potential at much lower concentrations than those at which the mammalian system is influenced.¹⁴

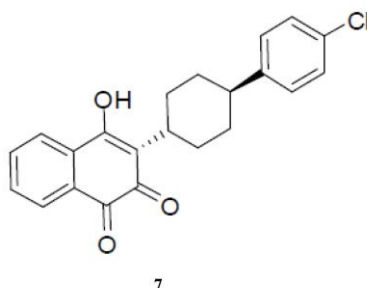


Figure 1.2.1.2 Chemical structure of atovaquone (**7**)

1.2.2 Blood schizontocycles

Blood schizontocycles are classified as drugs that are active against intra-erythrocytic parasites. This kind of drugs includes the quinoline-containing compounds and the artemisin-type compounds.

1) Quinoline-containing drugs

As shown in Figure 1.2.2.1, the quinoline containing drugs consist of chloroquine (**8**), amodiaquine (**9**), pyronaridine (**10**), lumefantrine (**11**), quinine (**12**), mefloquine (**13**) and primaquine (**14**).¹⁶

Chloroquine (**8**) is a 4-aminoquinoline that has been used widely for the treatment of malaria. It disrupts parasite haem detoxification and inhibits the growth of parasitemia. It is not active against the liver-infecting forms.¹⁷ Although it is effective against *P. vivax*, *P. ovale* and *P. malariae*, widespread drug resistance has developed in *P. falciparum*.^{18, 19}

Amodiaquine (**9**), a 4-aminoquinoline compound, has the same mode of action as that of chloroquine (**8**).⁸ Even though there is some cross drug resistance, it is still effective against some chloroquine-resistant strains of *P. falciparum*.²⁰

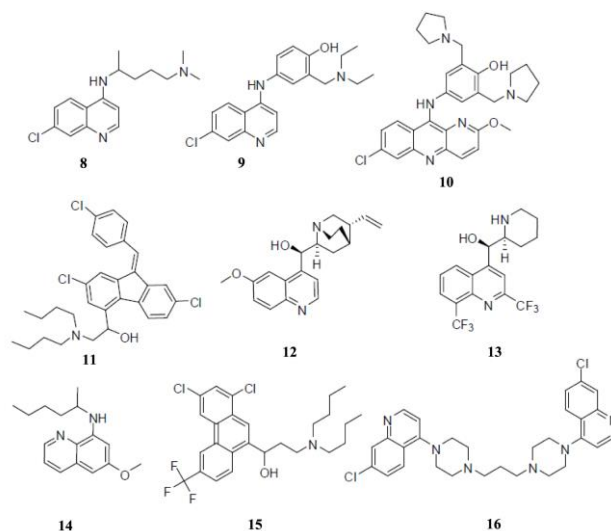


Figure 1.2.2.1 Chemical structures of chloroquine (**8**), amodiaquine (**9**), pyronaridine (**10**), lumefantrine (**11**), quinine (**12**), mefloquine (**13**), primaquine (**14**), halofantrine (**15**) and piperazine (**16**).

Pyronaridine (**10**), a 4-aminoquinoline compound, is highly effective against *P. falciparum*. It is regularly used with artesunate in combination therapy for common malaria.²¹

Lumefantrine (**11**) is a racemic chlorinated derivative. It is only effective in oral preparation combined with artemether (**17**).^{7, 8, 22}

Quinine (**12**) was isolated from the bark of the Cinchona tree. Quinine inhibits the mature trophozoite stage of *P. falciparum*⁹ and also kills the sexual stages of *P. vivax*, *P. malariae* and *P. ovale*, but not mature *P. falciparum* gametocytes. Its mechanism of action appears to be the inhibition of parasite haem detoxification in the food vacuole.⁷ It is generally used in combination with tetracycline (**21**) or doxycyclin (**22**) or clindamycin (**23**).^{23, 24}

Mefloquine (**13**), a 4-methanolquinoline, is effective against all forms of malaria.⁷ It is normally used in combination with artesunate and has been successful in the treatment of *P. vivax* malaria in chloroquine-resistant areas.²⁵

Primaquine (**14**), an 8-aminoquinoline, is effective against intrahepatic forms of all types of malarial parasite.⁹ It has blood schizontocidal and gametocytocidal activity via inhibition of protein transport.⁸ Primaquine (**14**) is also gametocytocidal against *P. falciparum* and has significant blood stages activity against *P. vivax* (and some are against asexual stages of *P. falciparum*). The drug has serious adverse effects including a life-threatening hemolysis in humans with glucose-6-phosphate-dehydrogenase (G6PD) deficiency, a genetic polymorphism particularly abundant in Africa and Asia.²⁶

Halofantrine (**15**) is an oral antimalarial, but not used for prevention. A possible mechanism of action has been shown in a crystallographic study, indicating that halofantrine binds to hematin *in vitro*.²⁷

Piperaquine (**16**), a 4-aminoquinoline derivative, has the same mode of action as chloroquine (**8**).²² It is now commonly used as a fixed dose formulation with dihydroartemisinin (**18**) against *Plasmodium* parasites.²⁵

2) Artemisinin (**18**) and its derivatives

Artemisinin (**18**), a peroxy containing sesquiterpene lactone, was first isolated from the Chinese medicinal herb *Artemisia annua* L. (sweet wormwood), also known as “qinghaosu”.²⁸ In China, *Artemisia annua* L. has been used to treat fever for thousands of years. The structure was first elucidated in 1972,¹⁷ and several semi-synthetic derivatives of artemisinin have been reported.²⁷ Artemisinin is an effective and rapidly acting blood schizontocide *in vitro*. It is active against all *Plasmodium* species with an unusually broad activity against asexual parasites, killing all stages from young rings to schizonts.²⁹ Furthermore, artemisinin is highly effective in reducing the parasite biomass 10,000-fold in a single asexual cycle, making artemisinin the most active and rapidly acting antimalarial drug.³⁰ These drugs are now used as a combination therapy (Artemisinin Combination Therapies; ACTs) to prevent the parasite resistance.³¹

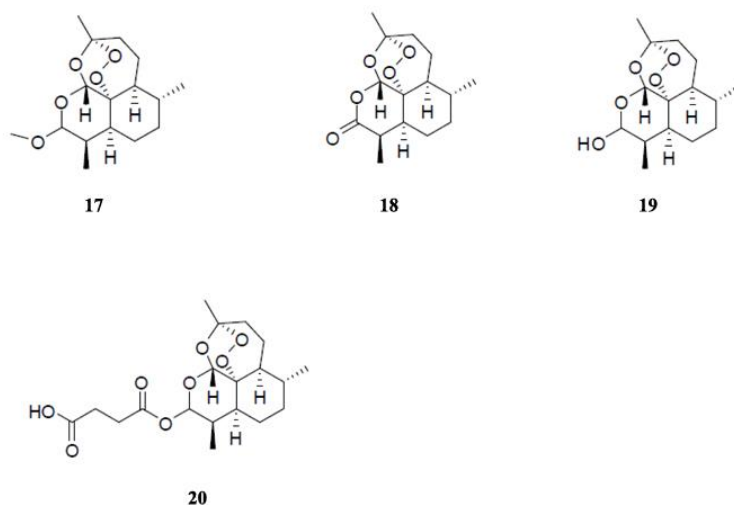


Figure 1.2.2.2 Chemical structures of artemether (**17**), artemisinin (**18**), dihydroartemisinin (**19**) and artesunate (**20**).

A number of analogues have also been reported and widely used, including artemether (**17**), dihydroartemisinin (**19**) and artesunate (**20**).²⁷ Their chemical structures are shown in Figure 1.2.2.2.

1.2.3 Antibiotic based malaria drugs

The well-known antibiotics including tetracyclines (**21**), doxycycline (**22**) and clindamycin (**23**) (Figure 1.2.3.1) have been successfully used for the treatment of malaria.⁷

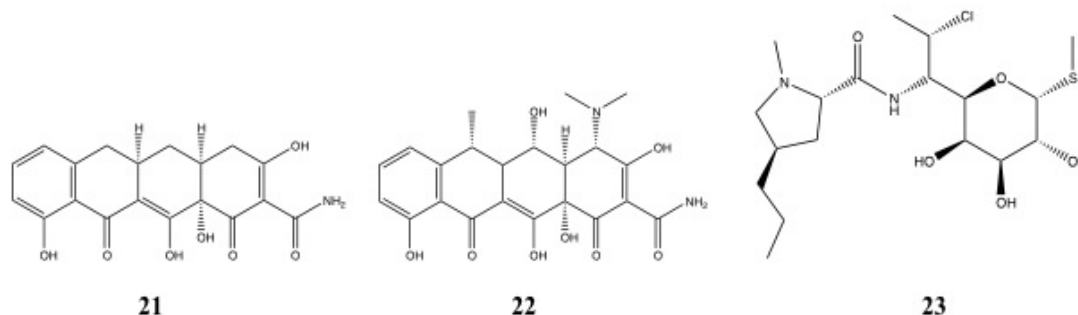


Figure 1.2.3.1 Chemical structures of tetracycline (**21**), doxycycline (**22**), and clindamycin (**23**)

Synthetic tetracyclines (**21**) are applied in malaria therapy as well as the natural products which were derived from a variety of *Streptomyces* genera.³² The hydrochloride salt or phosphate complex of tetracycline is given either orally or intravenously.⁷ Tetracyclines (**21**) inhibit binding of aminoacyl-tRNA during protein synthesis.³³

Doxycycline (**22**) is an analogue of tetracycline. It has a longer half-life (11-23 hours),³⁴ which can make dosing schedules much easier.⁷ Besides its long half-life, more reliable absorption and better safety profile in patients with renal dysfunction are the reasons it is generally preferred over tetracycline.³⁵

Clindamycin (**23**) is a lincosamide antibiotic and it acts against the early stages of protein synthesis.³⁶ It is also employed as the phosphate either via the intramuscular route or the intravenous route.⁷

1.3 Drug resistance

Although there are quite a lot of antimalarial drugs on the market, increasing drug resistance poses a growing problem in current malarial treatment. Drug resistance to

all classes of antimalarials has emerged in the last several decades. Antimalarial resistance is more common in *Plasmodium falciparum*. Resistance in *P. vivax* has appeared comparatively later and is mainly in southeast Asia.^{37,38}

In the 1950s, drug resistance to antimalarials first came into prominence.³⁹ In the 1970s and 1980s, chloroquine resistance became common and was responsible for the malaria recurrence in the tropics and in Africa.⁴⁰ Now, there are many countries where chloroquine resistance in *P. falciparum* is observed. Compared to *P. falciparum*, drug resistance to *P. vivax* is relatively insignificant. However, treatment failure with chloroquine in patients infected with *P. vivax* has been reported.⁴¹

Resistance in *P. falciparum* to the antifolate drugs (pyrimethamine-sulfadoxine) was reported in the same year as the introduction of the drugs. Antifolate resistance became widespread in the early 1980s in Southeast Asia, especially the Thai-Cambodian border where antimalarial resistance first emerged.^{42, 43}

Present knowledge indicates that drug resistance in malaria emerges through selection of parasites. Genetic mutations confer a selective advantage to some parasites over sensitive parasites in the presence of drug pressure.⁴⁴ Resistance to antifolates is known to be caused by mutations in the genes encoding the drugs therapeutic targets, dihydrofolate reductase (DHFR) and dihydropteroate synthase (DHPS).⁴⁵ In the case of quinoline-based compounds, such as chloroquine and mefloquine, resistance is thought to be dependent on the exclusion of the drug from the site of action.⁴⁶ This was thought to be mediated by mutations in two genes denoted *pfmdr1* and *pfcr1*, which encode proteins localized on the membrane of the parasitophorous food vacuole where these drugs are thought to act.⁴⁶ In recent years, health authorities have been increasingly recommending the use of ACTs in an effort to prevent the spread of antimalarial drug resistance.⁴⁷ Resistance to artemisinin derivatives has not yet been reported in human, but studies in rodent malaria models suggested that artemisinin resistance occurred in *P. falciparum*.⁴⁸

Given the critical drug resistance, there is an urgent need for new therapeutics to fight malaria.

1.4 Malarial targets

With the outstanding progress in molecular biology, many proteins with key roles in the function of both normal and abnormal cells have been identified, which has allowed the formation of specific hypotheses about how modulating the function of defined proteins that are linked to disease could be a route to new drugs. Such disease-linked proteins are commonly known as targets. Nowadays, target-based drug discovery becomes popular in the pharmaceutical industry. The target-based approach can very effectively develop novel treatments for a therapeutic target. In this study, three malarial targets were involved.

1.4.1 *Pfd*UTPase



Figure 1.4.1.1 *Pfd*UTPase with 5'-Diphenyl Nucleoside Inhibitors (PDB ID: 3T60)⁴⁹

dUTPase (Figure 1.4.1.1) is indispensable in eukaryotes as well as in prokaryotes. It catalyzes the hydrolysis of 2'-Deoxyuridine 5'-Triphosphate (dUTP) to deoxyuridine monophosphate (dUMP). It is involved in pyrimidine metabolism, which is a major route for therapeutic intervention against malaria. The enzyme is thought to be crucial to DNA integrity in two ways. It prevents the buildup of dUTP and ensures the provision of dUMP, the substrate for thymidylate synthase in the biosynthesis of deoxythymidine triphosphate (dTTP). As a result, a low dUTP/dTTP ratio is maintained, which greatly reduces uracil incorporation into DNA. Under normal circumstances, following dUTP misincorporation into DNA, uracil is excised and replaced by thymine through a repair process catalyzed by uracil-DNA glycosylase.⁴⁹

However, when dUTP levels are abnormally high, repetitive cycles of introduction and excision of uracil will take place, giving rise to DNA fragmentation and ultimately cell death.⁴⁹ In another word, inhibition of dUTPase in *Plasmodium* parasites results in the death of parasites. If we could selectively inhibit *Pfd*UTPase, we could selectively kill *Plasmodium* parasites. Therefore, selective inhibition of *Pfd*UTPase could lead to potential antimalarial drugs.

There are several different oligomeric forms of dUTPase: monomeric (in herpes virus and Epstein-Barr virus), dimeric (in trypanosomes and leishmania), and trimeric (in mammals, plasmodium, and various bacteria and viruses).⁴⁹ From the kinetic point of view, trimeric dUTPases appear extremely selective toward dUTP; they can discriminate between the four nucleobases and are also selective toward the sugar moiety.⁴⁹ Thus, other naturally occurring nucleotides such as dTTP, deoxycytosine triphosphate (dCTP), and uridine triphosphate (UTP) are very poor substrates.⁴⁹

Whittingham and her colleagues compared the crystal structures of six trimeric dUTPases (*P. falciparum*, human, equine infectious anemia virus, feline immunodeficiency virus, *E. coli* and *M. tuberculosis*). They found that dUTPases from these 6 oligomeric classes had similar 3D structures though they shared only low-level pair-wise sequence identities. These dUTPase trimers possessed five conserved sequence motifs, which formed the substrate recognition site and reaction center.⁴⁹ *Pfd*UTPase had a low sequence identity (28.4%) compared with its human ortholog. This might make it a suitable target for drug development.⁴⁹

Table 1.4.1.1 Structures of some *Pfd*UTPase inhibitors⁵⁰

No.	X	Y	R	R ₁
I	O	C	Phenyl	OH

II	NH	C	Phenyl	OH
III	O	Si	Phenyl	OH
IV	O	C	Phenyl	F
V	NH	C	Phenyl	F
VI	O	Si	Isopropyl	OH

A number of *Pfd*UTPase inhibitors were reported. Nucleoside derivatives (Table 1.4.1.1) were tested for their inhibitory activity against *Pfd*UTPase and human dUTPase.⁴⁹ The result suggested that compound **II** inhibited *Pfd*UTPase with a K_i value of 0.2 μM , in comparison to 46.3 μM for the human enzyme, giving a selectivity of the order of 200-fold (Table 1.4.1.2). In addition, these inhibitors were tested against *P. falciparum* K1 strain cultured in erythrocytes. The results suggested that compound **II** also selectively inhibited *Plasmodium* parasites with an IC_{50} value of 4.5 μM , while it had 10-fold less activity against L6 cells (Table 1.4.1.2).

Table 1.4.1.2 Inhibition activity of nucleosides⁴⁹⁻⁵⁰

No.	Inhibition of the Enzyme			Activity against Parasites		
	<i>Pf</i> K_i (μM)	Human K_i (μM)	Select.	<i>Pf</i> IC_{50} (μM)	L6-cells IC_{50} (μM)	Select.
I	1.8	17.7	9.8	6	192	32
II	0.2	46.3	232	4.5	43.5	9.7
III	2.8	909	325	1.1	ND	
IV	4.98	457	91	2.0	35.4	17.7
V	12.45	>1000	>80	5.3	30.3	5.7
VI	227.1	>1000	>4.4	>13	25.2	

ND = Not detectable.

IC_{50} values for standard drugs against *P. falciparum*: Chloroquine = 0.195 μM ; artemisinin = 0.0057 μM .

L-6 cells are rat skeletal myoblasts and compounds are assayed against these cells to give an idea of toxicity against mammalian cells.

Selectivity (Select.) values are shown in parentheses, where:

Selectivity for the enzyme is defined as K_i (human)/ K_i (*P. falciparum*, *Pf*). Selectivity for the parasite is defined as IC_{50} (L-6 cells)/ IC_{50} (*Pf*).

Research carried out in Professor Ron Quinn's group at Eskitis Institute has identified seven securinine-derivatives (Figure 1.4.1.2) which bound to *Pfd*UTPase through by

ESI-FTMS. Subsequent *in vitro* whole cell assay against *P. falciparum* suggested that the securinine-derivatives also inhibited viability of both *P. falciparum* gametocyte (sexual) and blood (asexual) stage parasites with IC₅₀ of 36.7 (±3.0) µM to 81.1 (±2.1) µM and 17.1 (± 0.9) µM to 82.4 (± 10.8) µM respectively.⁵¹

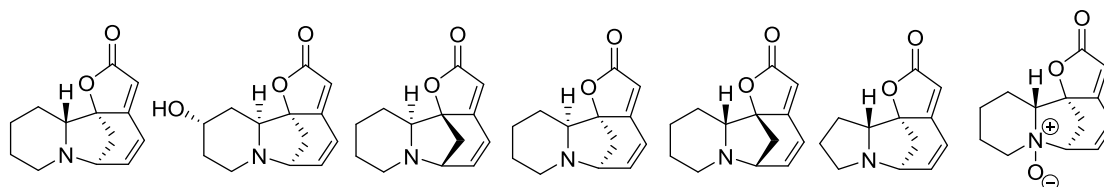


Figure 1.4.1.2 Chemical structures of the securinine fragment series shown to bind *PfdUTPase*⁵¹

Our results in combination with literature reports indicated that *PfdUTPase* might be a valid target for the discovery and development of new antimalarial drugs.

1.4.2 *PfRab11a*

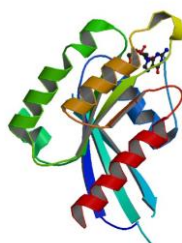


Figure 1.4.2.1 Crystal structure of *PfRab11a* in complex with GDP (PDB ID: 3BFK)⁵²

Rab GTPases, constituting the largest family of small monomeric GTPases, are key regulators of vesicular traffic in eukaryotic cells. The small G-protein Rab11a is conserved in eukaryotes and it has been shown to play an important role in regulating trafficking of certain plasma membrane receptors through recycling endosomes.⁵³ Furthermore, in animal cells, Rab11a has been demonstrated to be essential for delivering plasma membrane to the cleavage furrow, indicating a conserved function during cell division.⁵⁴⁻⁵⁵

Plasmodium parasites were reported to have a family of 11 different Rabs.⁵⁶ *Plasmodium falciparum* Rab11a was first described by Langsley *et al* in 1996 and

subsequently shown to be expressed in the asexual blood stage of parasites.⁵⁶⁻⁵⁷ In 2003, Quevillon *et al* reported that *PfRab11a* might mediate an essential recycling endosome function in *P. falciparum*.⁵⁶ Later, Agop-Nersesian *et al* demonstrated that the final step during parasite cell division involves delivery of new plasma membrane to the daughter cells, a process requiring functional Rab11a.⁵⁸ Importantly, Rab11a can be found in MTIP (Myosin-Tail-Interacting-Protein), which is a member of a 4-protein motor complex that is crucial for parasite to invade the host cell.⁵⁸ Ablation of Rab11a function results in daughter parasites having an incompletely formed IMC (Inner Membrane Complex) that leads to a block at a late stage of cell division.⁵⁸ *PfRab11a* therefore might be a potential target for anti-malarial drug discovery. To date, no efforts have been put into screening *PfRab11a* using ESI-FTMS. Meanwhile, no natural product inhibitor/binder of *PfRab11a* has been reported.

1.4.3 *PfUCH-L3*

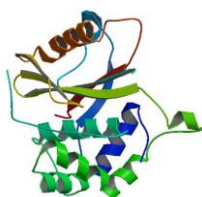


Figure 1.4.3.1 Crystal structure of *PfUCH-L3* (PDB ID: 2WE6)⁵⁹

PfUCH-L3, an ubiquitin C-terminal hydrolase (UCH), is thought to be essential for *Plasmodium falciparum* to survive.⁵⁹ Ubiquitin (Ub) is a protein present in all eukaryotes that is composed of a well-conserved sequence of 76 amino acids. This ubiquitin-proteasome system is crucial to eukaryotic survival and is involved in protein degradation as well as cell's maintenance of synthesized and degraded proteins.⁵⁹ When this balance is out of equilibrium, the cell can run into functional problems and can eventually lead to cell death.⁶⁰

PfUCH-L3 consists of a central six-stranded anti-parallel β -sheet surrounded by seven α -helices. Viewing *PfUCH-L3* in complex with an ubiquitin-based suicide substrate (UbVME) allows us to know the dual specificity of this enzyme. The β -sheet

conformation of the C-terminus of ubiquitin has backbone carbonyl and amide groups that hydrogen bonds to the *PfUCH-L3* residues. The C-terminus of Ubiquitin binds to the narrow groove, lined by the catalytic triad residues: Cys-92, His-164, and Asp-179. These three residues are all within hydrogen bonding distance of each other. Ser-12 forms an additional hydrogen bond to the backbone nitrogen of Ub Leu-73 in the Ub-binding groove of *PfUCH-L3*. All of the hydrogen bond donors and acceptors of Ub C-terminal residues 71-75 are then fulfilled.⁵⁹

An important element of the bound complex is the active site crossover loop. In unbound *PfUCH-L3*, the crossover loop is very ordered. This is due to crystal contacts that hold the loop in a distinct conformation. No significant structural changes occur upon the binding of Ubiquitin.⁶¹

This loop encounters the C-terminal of the Ubiquitin suicide substrate (UbVME) and is thought to assist in the positioning of the substrate. When the substrate is not present, the crossover loop becomes flexible and cannot be detected in the crystal structure, hence why we cannot see it on this free molecule. However, when ubiquitin is present, the crossover loop connects the two halves of the catalytic center, bringing key residues close enough to catalyze the binding of ubiquitin. The crossover loop also functions as a substrate filter by limiting the types and sizes of substrates the enzyme can hydrolyze. While the structure and function of the cross-over loop is not yet well understood, there are a few key features that are visible in the ubiquitin bound crystal structure. Leu-71 and Leu-73 point to the opposite side of the binding groove and occupy a hydrophobic binding pocket where the substrate attaches. Asp-157 is located in the crossover loop and forms a salt bridge with the side chain of Arg-74 of Ub to provide stability to the bound complex. This ubiquitin binding pocket can now be visualized.⁵⁹

The key difference between *HsUCH-L3* and *PfUCH-L3* is that the sequence of the crossover loop is not conserved. Therefore, these two enzymes do not interact with Ub in the same pattern and do not share the same conformation. While the Ub

recognition and binding mode between each enzyme and Ub is similar, the Ub interfaces of both enzymes have a few key differences. In *PfUCH-L3*, Ser-12 in the Ub-binding-groove forms an additional hydrogen bond to the backbone nitrogen of Ub Leu-73. Ub-Arg-74, a highly conserved residue and central to Ub recognition, forms a network of hydrogen bonds with *HsUCH-L3* residues. However, *PfUCH-L3* residues have a completely different conformation of hydrogen bonds. One example of this is the formation of a salt bridge to Asp-157 in *PfUCH-L3* that is not present in *HsUCH-L3*. Also, the interfaces of *PfUCH-L3* and Ub have the nonconserved residues of Thr-163 and Ser-219. These key differences allow researchers to specifically target the parasite rather than the human host cells.⁵⁹ There was no natural product inhibitor of *PfUCH-L3* reported. Hence, *PfUCH-L3* might be treated as a potential target for antimalarial drug discovery.

1.5 Protein-ligand noncovalent interactions

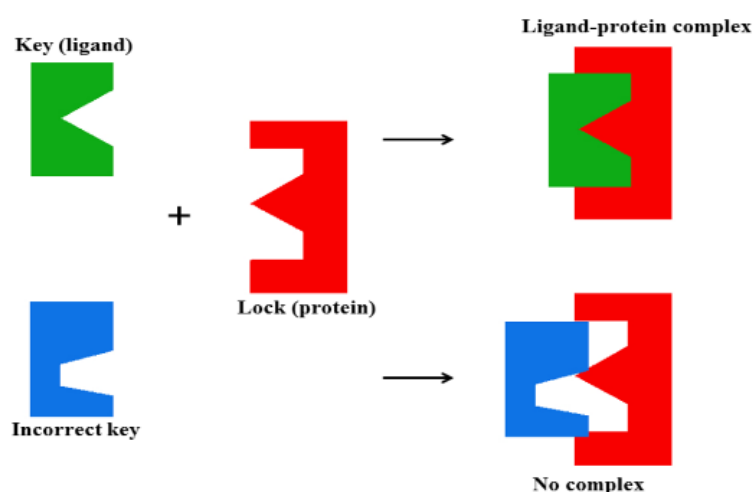


Figure 1.5.1 “Lock and Key” model: The green key (ligand) has the right shape for the red lock (protein) and a ligand-protein complex can be formed. The blue key with unsuitable shape thus no complex is formed.

Understanding protein-ligand interactions is essential in order to obtain novel therapeutic drugs to deal with diseases. Interactions between a protein and a ligand have been explained by the “lock and key” model introduced by Emil Fischer.⁶²

Based on this theory, the “key” represents the ligand and the “lock” represents the protein or some other receptor. Only those ligands with a right geometrically shape can fit into the keyhole of the lock (Figure 1.5.1).

There are two different kinds of chemical bonds between ligand and protein: the covalent bond and the noncovalent bond. The covalent bond is directionality and saturated. When a ligand and a receptor form a covalent bond, due to the solid structure and strong bond, this kind of binding is irreversible, making it less desirable to be used a drug. Meanwhile, there are many types of noncovalent interaction such as ionic bonds, hydrophobic interactions, hydrogen bonds, van der Waals forces and dipole-dipole bonds.⁶³ A noncovalent complex is formed when a noncovalent bond is created between a ligand and a protein. Today most drugs have a pharmacological effect on a therapeutic target by forming a noncovalent complex with the target.⁶⁴

Currently, there are a lot of biophysical methods available for studying protein-ligand interactions, including ESI-MS, X-ray crystallography, Nuclear Magnetic Resonance (NMR), Surface plasmon resonance (SPR), Isothermal titration calorimetry (ITC) and Spectroscopy methods. Their advantages and disadvantages are summarized in Table 1.5.1.

Table 1.5.1 Advantage and disadvantages of some biophysical methods to study protein-ligand interactions

Methods	Advantages	Disadvantages
ESI-MS	<ul style="list-style-type: none"> • Fast • Able to detect small amount of sample • Able to observe protein in native state 	<ul style="list-style-type: none"> • Little structure information • Destructive method • Not suitable for insoluble protein
X-ray crystallography	<ul style="list-style-type: none"> • Highly accurate • Determination of structure and interaction with ligands at high 	<ul style="list-style-type: none"> • Difficult to generate crystal • Time-consuming • Difficult to capture the dynamic

	atomic resolution	information of biomacromolecules
	<ul style="list-style-type: none"> • No molecular weight limit 	<ul style="list-style-type: none"> • High amount of sample required in most cases
Nuclear Magnetic Resonance (NMR)	<ul style="list-style-type: none"> • Detect biomolecules configuration under normal physiological conditions • Observe continuous dynamic information • Non-destructive method 	<ul style="list-style-type: none"> • Molecular weight limit (<30, 000 Da) • Large amount of sample required • Difficult and time-consuming to analyze the spectrum • Not suitable for insoluble protein
Surface plasmon resonance (SPR)	<ul style="list-style-type: none"> • Convenient • Fast • Suitable for turbid, opaque or colored solution • Observe continuous dynamic process 	<ul style="list-style-type: none"> • Influenced by external environment (temperature and etc) • Difficult to distinguish non-specific adsorption • Large mass ligands
Isothermal titration calorimetry (ITC)	<ul style="list-style-type: none"> • Accurate • Sample can be used in solution • Non-destructive method 	<ul style="list-style-type: none"> • Slow (30-60 min/sample) • Large amount of sample required
Spectroscopy (Fluorescence, Ultraviolet-visible, Circular dichroism)	<ul style="list-style-type: none"> • Sample can be used in solution • Low amount of sample required 	<ul style="list-style-type: none"> • Influenced by external environment (pH, temperature, etc) • Little structure information

Compared with other methods, ESI-MS provides a more rapid and sensitive choice for the identification of noncovalent interactions in mixtures.⁶⁴ Since it is a soft-ionization ion method, ESI can detect fragile noncovalent complexes.⁶⁴ ESI-MS can also distinguish whether a protein exists as its native folded state (pH=7) or remains in a denatured state (pH=1-3) in liquid-phase. Difference between native state and denatured state of one protein is shown in Figure 1.5.2. Native protein has a broad and low charge state mass spectrum (high m/z), while a narrow and high charge state spectrum (low m/z) is correlated to denatured (unfolded) protein. This is mainly because folded protein has fewer basic sites available for protonation than its unfolded conformation.⁶⁵

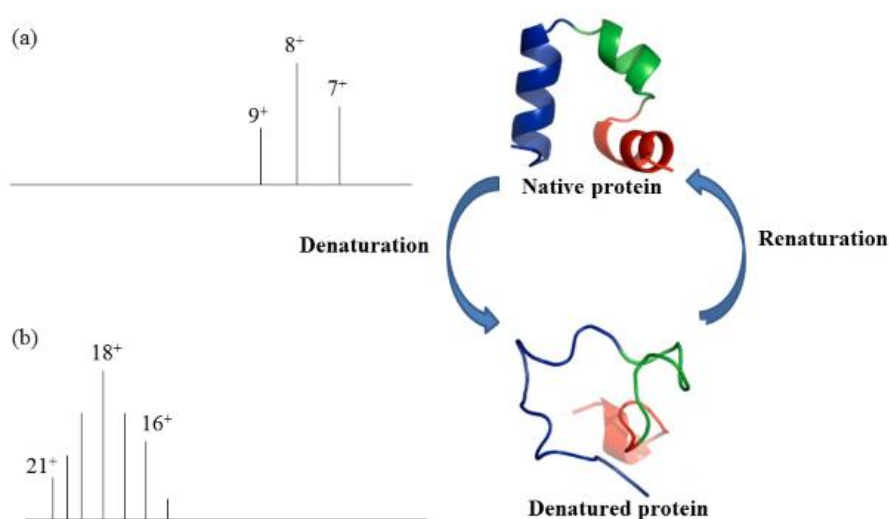


Figure 1.5.2 An illustration of conformational change in protein acquired by ESI-MS.

(a) Charge states for native protein. (b) Charge states for denatured protein.

The coupling of ESI and FT-MS has greatly advanced the capabilities of mass spectrometry for biological applications.⁶⁶ One significant application is to use the ESI-FTMS for screening, which can be referred as bioaffinity mass spectrometry screening (BAMS).

1.6 Mass spectrometry (MS)

With the purpose to get a better understanding of BAMS, an overview of related MS information is presented below.

1.6.1 General

MS is an analytical technique used to separate ions according to their mass to charge ratio (m/z). In the early of last century, this technique was used to measure masses of atoms.⁶⁷ And with its help, scientists successfully demonstrated the existence of isotopes; this discovery fueled the contemporaneous ongoing debates about the structure of the atom.⁶⁸ By the 1940s, MS was used to detect the abundances of small hydrocarbons in process streams by chemists working in the petroleum industry.⁶⁹ It was not until the 1960s that MS was applied to natural products chemistry and other chemists really began to realize that how complex molecules MS can deal with and to ponder the range of its possible applications.⁷⁰

To really recognize how the research area of MS has expanded to its present-day size, it is useful to review the great advances in the field. Some important progresses are listed in Table 1.6.1.1.

Table 1.6.1.1 Progresses in MS

Year	Progress
1897	Early MS
1949	Ion Cyclotron Resonance (ICR)
1953	Quadrupole Analyzers
1956	Identifying Organic Compounds with Mass Spectrometry
1968	Electrospray Ionization
1974	Fourier Transform Ion Cyclotron Resonance
1984	Quadrupole/Time-Of-Flight Mass Analyzer
1985	Matrix-Assisted Laser Desorption Ionization (MALDI)
1989	ESI on Biomolecules
2005	Direct Analysis in Real Time (DART)

Nowadays, MS has both qualitative and quantitative applications, which include identifying unknown compounds, determining the isotopic composition of elements in a compound, and determining the structure of a molecule by observing its fragmentation. Other applications include quantifying the amount of a compound in a sample or studying the fundamentals of gas phase ion chemistry (the chemistry of ions and neutrals in a vacuum). This technology is commonly used in analytical laboratories to study biological, chemical, or physical properties of a wide range of molecules.

Generally, an ionization source, a mass analyzer and a detector are the main components of a mass spectrometer (Figure 1.6.1.1). In order to maintain the pressure far below the atmospheric pressure, a good vacuum system is required. Sample vaporization is the first step, followed by ionization, separation according to different m/z ratio, and detection of the ions. The signal is processed into mass spectra via a computer system.

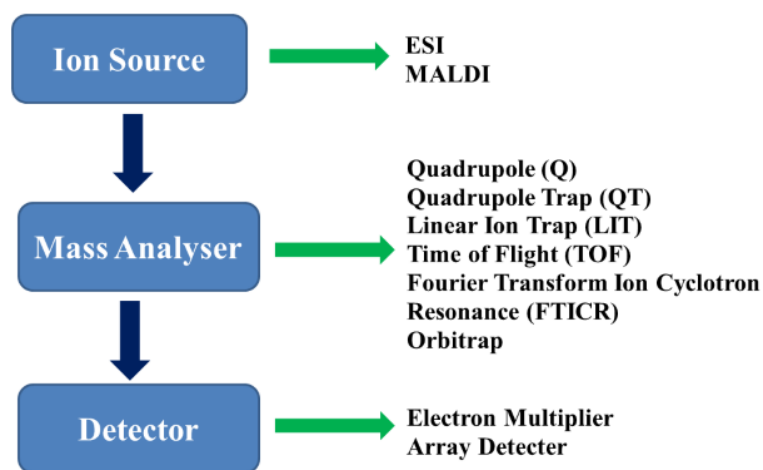


Figure 1.6.1.1 Main components of a mass spectrometer

1.6.2 Ion sources

Ionization is the process by which an electrically neutral atom, molecule, or radical loses or gains one or more electrons and becomes an ion. Ionization can occur in gases, liquids, or solids.

There are many different types of ionization, such as Electron Impact (EI), Chemical Ionization (CI), Field Ionization (FI), Field Desorption (FD), Fast Atom Bombardment (FAB), Matrix-Assisted Laser Desorption Ionization (MALDI) and ESI. MALDI and ESI are the most common used methods for biological applications.

1.6.2.1 MALDI

Ions in MALDI are assembled by pulsed-laser irradiation of a sample.⁷¹ A co-crystallization between the sample and the solid matrix take place, which is able to absorb the wavelength of light emitted by the laser.⁷² The mechanism of the ion formation is not quite clear though the formation of singly-protonated analytes is characteristic.⁷³ A notable advantage is that MALDI has a very high level of sensitivity. It can couple with time-of-flight (TOF) to preserve molecular weight information rapidly.⁷¹ MALDI has a high tolerance to salts and buffers, however its drawback is that it needs more substance than ESI.⁷⁴ In addition, the presence of a matrix causes chemical noise at m/z ratio below 500 Da.⁷⁵ As a result, samples with low MW are normally hard to analyse. MALDI schematic is shown in Figure 1.6.2.1.1.

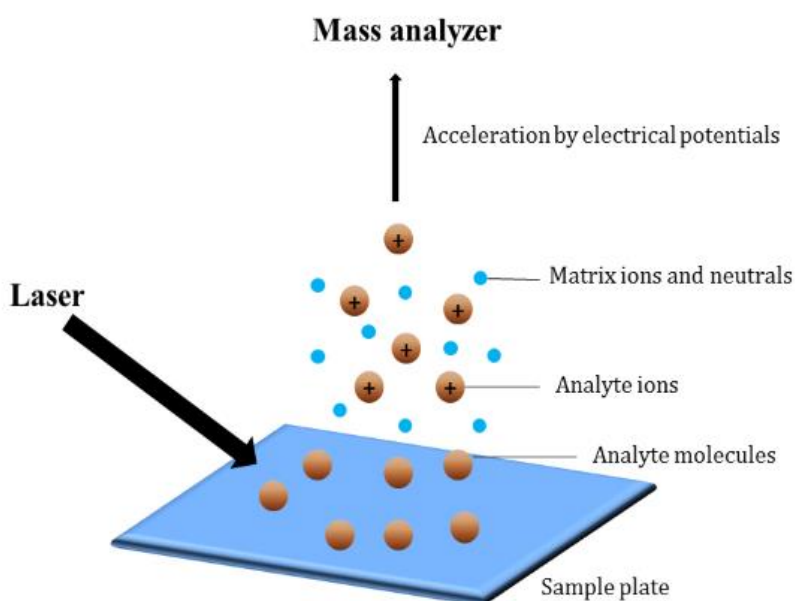


Figure 1.6.2.1.1 A schematic of MALDI

1.6.2.2 ESI

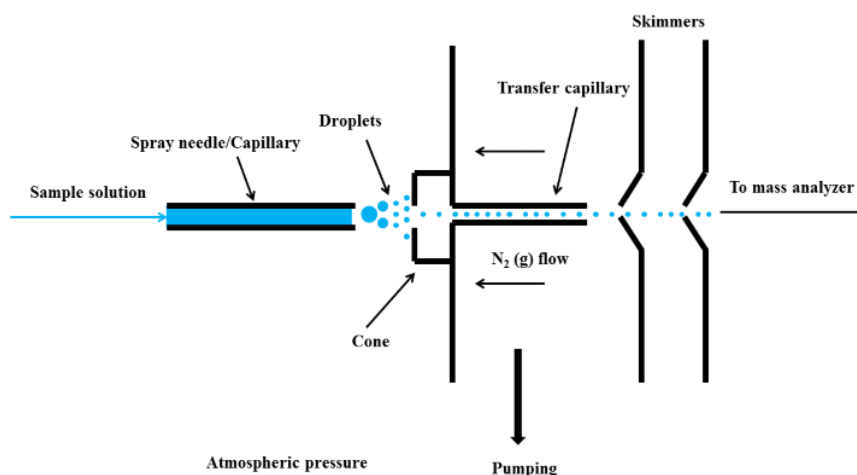


Figure 1.6.2.2.1 Electrospray ionization

In Dole and co-workers' research, gas-phase ions were generated by spraying a solution from the tip of an electrically charged capillary.⁷⁶ Based on Dole's work, Fenn and co-workers developed electrospray as an ionization technique for MS.⁷⁷

In the process of ESI (Figure 1.6.2.2.1), the analyte is injected into the source in solution either from a syringe pump or as the eluent flow from liquid chromatography (LC) at atmospheric pressure. The analyte solution flow goes through a capillary into a small-diameter needle held at high potential (typically within vary from 2.5 to 4 kV). This forces the spraying of charged droplets from the needle with a surface charge of the same polarity to the charge on the needle. The droplets are released from the needle towards the cone on the counter electrode. While the droplets move towards the cone, solvent evaporation happens. The droplet becomes smaller until it reaches the Rayleigh limit, at which the repulsive Coulombic force is equal to the surface tension force.⁷⁸ The droplets become unstable and generate smaller droplets. This Coulomb fission process is repeated until there are small highly charged droplets.

Currently, there is no consensus about the mechanism to explain how the charged droplets generate solute ions.⁷⁹ Charge residue model (CRM) by Dole et al and ion

desorption model (IDM) by Iribarne and Thomson are two famous theories on ion formation.^{76, 80}

Dole *et al* assumed that droplets become smaller with solvent evaporation. Due to the Coulomb fission process, the droplet divides into smaller ones. This process continues until one droplet contains only one charged molecule. After the last solvent evaporates, a completely desolvated ion is formed. The CRM is more plausible for macromolecules.⁸¹

Iribarne and Thomson supposed that when a droplet is highly charged, the surface electric field becomes large enough for the desorption of ions. This IDM theory is more likely to be valid for small ions.⁸²

A diagram demonstrating the CRM and IDM models of ion formation in electrospray is shown in Figure 1.6.2.2.2.

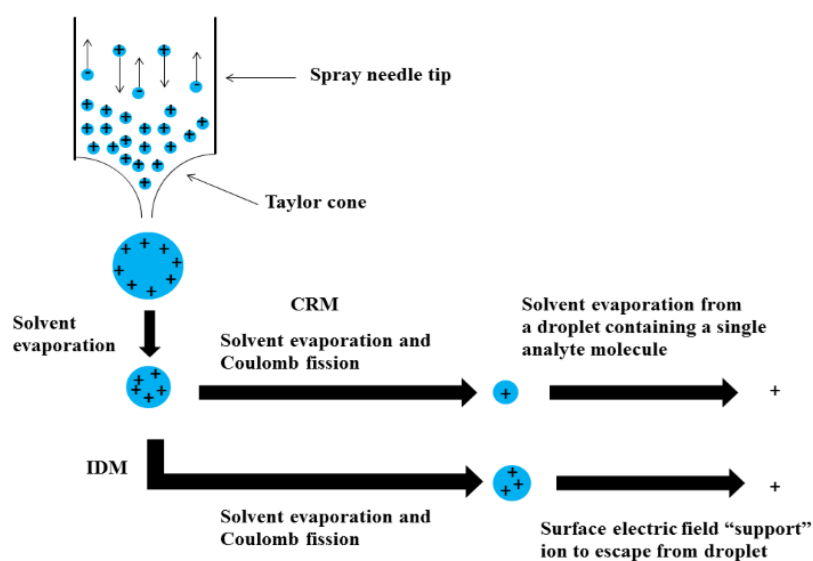


Figure 1.6.2.2.2 Two ESI mechanisms: CRM and IDM

In addition to its high sensitivity, ESI has a lot of advantages. First, ESI can be easily connected with a LC system. Secondly, multiple charged ions can be formed and it's possible to measure macromolecules. Thirdly, unlike MALDI, ions with a small m/z can be clearly observed. However, ESI has shortcomings. For instance, most samples

need to be desalted before analysis, because high salt conditions can inhibit ion formation.⁸³

1.6.3 Analyzers

1.6.3.1 TOF

Figure 1.6.3.1.1 shows the working principle of a linear TOF mass spectrometer. In order to allow the ions to fly through the flight path without hitting anything else, all the air molecules have been pumped out to create an ultra high vacuum. TOF identify molecules by measuring the time that sample molecules, all starting with the same kinetic energy, require to fly a known distance.⁸⁴ Once the sample molecules are ionized, an electrical field accelerates them all to the same energy. According to the equation (1) in Figure 1.6.3.1.2, lower m/z ions achieve higher velocities than ions with higher m/z ratio. As they all travel the same distance, and their velocity is dependent upon their mass, measuring the flight time each ion takes to fly through the time measurement region is just proportional to the square root of their mass (equation 2 in Figure 1.6.3.1.2). An advantage of TOF is that the mass range is theoretically limitless.⁸⁵

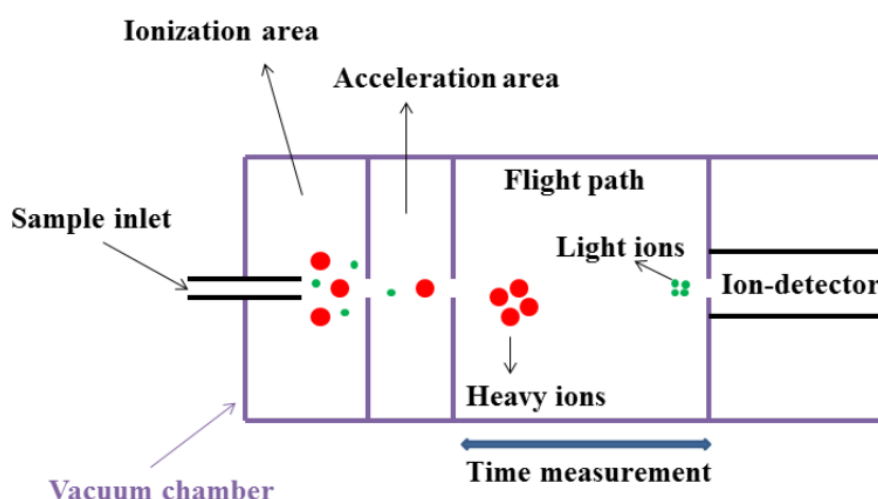


Figure 1.6.3.1.1 TOF Schematic

$$Velocity = \sqrt{\frac{2 \times energy}{mass}} \quad (1)$$

$$Flight_time = \frac{time\ measurement\ region}{velocity} = time\ measurement\ region \times \sqrt{\frac{mass}{2 \times energy}} \quad (2)$$

Figure 1.6.3.1.2 Equations for TOF

1.6.3.2 Quadrupole

The quadrupole mass analyzer has become more and more popular. Especially for gas chromatography/mass spectrometry (GC/MS) as well as liquid chromatography/mass spectrometry (LC/MS), the quadrupole was the first choice.⁸⁶
⁸⁷ An advantage of the quadrupole is the low cost of the instruments and their ease of automation.⁸⁸ And the mass separation in a quadrupole is addicted directly by the m/z ratio of the ion and is a result of ion motion in a dynamic electric field (Figure 1.6.3.2.1).⁸⁹

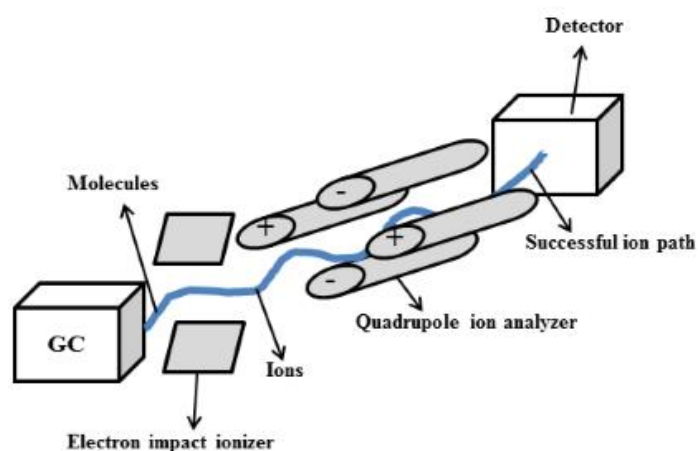


Figure 1.6.3.2.1 Quadrupole Mass Spectrometer Schematic

1.6.3.3 Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS, FT-MS)

1.6.3.3.1 FT-MS theory

The technique of ICR-MS was first published in the middle 1950s, when it was demonstrated for measurement of very small mass differences at very high

precision.⁹⁰ The technique remained a largely academic tool until the application of FT methods by Alan Marshall and Melvin Comisarow in 1970's.⁹¹ In 1978, Marshall and Comisarow built the first FT-ICR-MS mass spectrometer, and the performance of these instruments made a major leap forward and laid the basis of modern FT-MS.⁹² Coupled with superconducting magnets in 1980, FT-MS's stability and reproducibility was significantly improved. At the same time, since the magnetic field increased, the resolution, measurement accuracy and quality range of FT-MS was greatly improved. It is now one of the most sensitive methods of ion detection in existence and has almost unlimited resolution ($>10^7$).⁹²

FT-MS is based on the ICR principle. Ions generated by electrospray ionization (ESI) are trapped in the ICR cell, which is located in a homogenous magnetic field of a superconducting magnet. In the absence of an electrical field, an ion with mass (m) and charge (q) moves in circular motion as a result of the Lorentz force and the centrifugal force working on it in opposite direction. The angular frequency of this motion is given by: $\omega_c = v/r = qB/m$ (equation 1), in which ω_c is the cyclotron frequency in radians per second, B is the magnetic field strength in Tesla, and m/q is the mass to charge ratio of the ions in kg/C. The total charge on an ion can be described as $q = ze$, in which z is the net number of electronic charges on the ion and e is the charge per electron in Coulomb ($1\text{ e} = 1.6022 \times 10^{-19}\text{ C}$). Equation (1) can be rewritten as: $f_c = zeB/2\pi m = 15357zB/m$ (equation 2), in which f_c is the cyclotron frequency in Hertz, z is the charge state of the ion, and m is the mass of the ion in Da ($1\text{ Da} = 1.66054 \times 10^{-27}\text{ kg}$). Following equation (2) it is evident that ICR frequency is independent of ion velocity, or of kinetic energy. Ions of a given mass will have the same ion cyclotron frequency, regardless of the spread in the ions' kinetic energy distribution. This avoids the peak broadening and achieves ultra high mass resolution. The equation (1) is the unperturbed cyclotron frequency in which there is only a uniform magnetic field applied on the ICR cell. In reality an electrostatic DC is applied to the trapping plate of the ICR cell to prevent the ions escaping along the z axis. The

combination of the magnetic field and the electrostatic field create three motions in the cell: the axial trapping motion, the magnetron motion and the cyclotron motion, and lead to the real cyclotron frequency which is smaller than that of equation (1). $\omega_c \text{ real} = \omega_c - \omega_m$, in which ω_m is the magnetron frequency. To obtain high resolution and mass accuracy, the magnetron motion has to be eliminated or minimized. High trap potentials and ion density increase the growth of magnetron motion, whereas on-axis injection and longer injection times to increase the symmetry of the trapped ion cloud minimize magnetron motion. Coulomb interactions occur between an orbiting ion packet and other ion packets or with the electrostatic trapping plates, cause frequency shifts and peak coalescence effect (coupling of cyclotron frequencies) and affect the resolution and mass accuracy.

Here is a brief introduction about how FT-MS instrument works. In a general FT-MS instrument, the ions are produced in the source (for instance, ESI) and then pass through a series of pumping stages at increasingly high vacuum into the ion trap (also known as ICR cell). The ion trap is located inside a spatial uniform static superconducting high field magnet cooled by liquid nitrogen and liquid helium. The ions are bent into a circular motion in a plane perpendicular to the field due to the Lorentz Force, when they enter the magnetic field. At each end of the ion trap, there is a trapping plate to prevent the ions from escaping. The ion cyclotron frequency is dependent on their m/z ratio. At this stage, there is no signal observed because the radius of the motion is very small. Excitation of each individual m/z is achieved by a swept radio frequency (RF) pulse across the excitation plates of the cell. Each individual excitation frequency will couple with the ions natural motion and excite them to a higher orbit where they can generate an alternating current between the detector plates. The cyclotron frequency of the ions is the same as the frequency of this current. In addition, the number of ions is proportional to the current's intensity. The signal of this current will be recorded, processed and finally transformed into the mass spectrum. (The ions return to their natural orbit, when the RF is off

resonance for that m/z .) Deconvolution of this signal by FT methods results in the deconvoluted frequency vs. intensity spectrum, which is then converted to the mass vs. intensity spectrum (the mass spectrum). It is also usual to correct for mass errors at this stage by applying a calibration.

1.6.3.3.2 FT-MS applications

FT-MS is used in many applications, from small molecules to larger polymers, due to its high resolution and high mass accuracy.

(1) Macromolecular analysis

He *et al* managed to use FT-MS to distinguish two peptides with a molecular weight difference of 3.4 mDa.⁹³ In order to analyze the preliminary amino acid sequence, Wood *et al* determined the accurate mass of protein by FT-MS.⁹⁴ MS data obtained from Wood's research confirmed the presence of a dimer in the solution.

Yu *et al* reported the applications of FT-MS in the analysis of RNA structure as well as RNA and protein interactions.⁹⁵ Without gel electrophoresis experiments, FT-MS can directly provide clear RNA structure information, which greatly improves the working efficiency and analysis accuracy.

High polymers have large molecular weight and complex structures. Aaserud reviewed the application of mass spectrometry in the plastics industry, with particular emphasis on the large analysis range and high-resolution of MALDI-FTMS and ESI-FTMS.⁹⁶

(2) Metabolomics analysis

FT-MS can directly detect and identify metabolites without chromatography. What's more, MS/MS can play an important role in elemental composition analysis of the substance and its structure identification.⁹⁷

In 2002, Aharoni *et al* identified some small molecular metabolites from strawberry on its metabolic transition from immature to ripe fruit. And they draw out a metabolic "fingerprinting" about the ripening process of strawberry.⁹⁸

(3) Others

FT-MS has also been applied to isotopic analysis. Zhang *et al* reported the use of this method to analyse the natural abundance of the elements in a protein molecule.⁹⁹ Xiong *et al* identified two kinds of compounds from a surfactant mixture.¹⁰⁰

In addition, coupling ESI with FTMS has greatly advanced the capabilities of MS for biological applications. The high resolution of FTMS and the multiple charging of ESI make ESI-FTMS extremely useful in determining accurately the mass of large biomolecules such as proteins or protein-ligand complexes.

In this thesis, ESI-FTMS is used to screen natural product extracts against malarial targets for antimalarial natural products.

1.7 Lead-like enhanced extracts (LLEE) library

Natural products have historically been an invaluable source of therapeutic agents and medicines. Natural product extracts are rich in chemical components, which may have the possibility to form noncovalent complexes with therapeutic targets.

Eskitis LLEE library consists of over 18,000 Lead-like enhanced (LLE) extracts of plants and marine organisms from Australia, China, Malaysia and Papua New Guinea. The library was constructed by LLE extraction procedure followed by reverse-phase HPLC fractionation.¹⁰¹ In comparison with traditional natural product drug discovery, which can carry out physicochemical assessment only after natural products are isolated, purified and their structures elucidated, the new protocol in Eskitis allows front-loading of the lead-like properties at the extract and fraction stage.¹⁰²

Lead-like enhanced extraction procedure involves: solvent extraction by DCM and MeOH, HLB solid phase extractions (SPE) eluted with 90% MeOH/10% H₂O to eliminate compounds with $\log P > 5$, and HPLC fractionation using reversed-phase solvent conditions (MeOH/H₂O/0.1% TFA) and a Phenomenex C18 Monolithic HPLC column (4.6 mm \times 100 mm). Five fractions were collected per sample within the region of the chromatogram corresponding to $\log P < 5$ (Figure 1.7.1). Previous research suggested that these five fractions mostly contain compounds with drug-like properties.¹⁰² In this thesis, natural product extracts from the LLEE library

was screened against malarial target *Pfd*UTPase using ESI-FTMS.

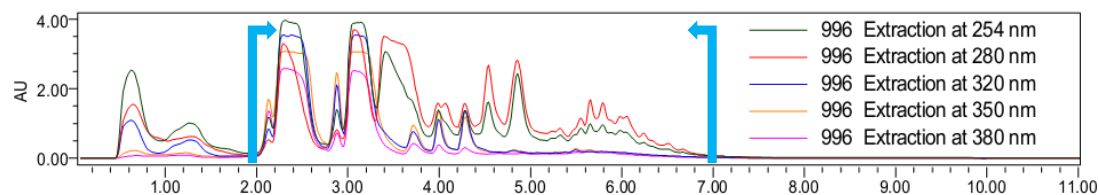


Figure 1.7.1 Example of a lead-like enhanced extract HPLC chromatogram

1.8 Objectives

Overall the project is to screen natural product extracts for ligands, which bind to potential malarial targets using ESI-FTMS.

This thesis aimed to: (1) use ChemGPS to analyze physicochemical properties of antimalarial natural products and drugs; (2) optimize ESI-FTMS screening conditions; (3) detect noncovalent complexes using ESI-FTMS and provide molecular mass of the active compounds; (4) isolate active molecules bound to the target; (5) evaluate the binding activity and inhibitory activity of the active compounds; (6) predict binding sites for the active compounds using virtual docking.

In order to prove that the LLE process has the ability to capture drug-like chemical constituents (Log *P* of no more than 5), this thesis also aimed to: (1) select samples that show a protein-ligand complex and isolate drug-like molecules; (2) evaluate drug-like properties of all the isolated natural products including the active compounds; (3) study the relationship between Log *P* and HPLC retention time.

1.9 Research plan

A research plan containing four major steps was undertaken in this study. Step one included understanding of mass spectrometry technique and its approaches in drug discovery screening (Chapter 1.6), understanding of physicochemical properties of antimalarial natural products (Chapter 2) and using ChemGPS to assess drug potential of reported anti-malarial natural products and drugs (Chapter 2). Step two involved developing optimized conditions for ESI-FTIMS screening (Chapter 3). Step

three aimed to use optimized method to screen natural product extracts, and identify noncovalent complexes formed by active natural product ligand and protein targets, obtaining accurate molecular mass of active components. The active ligands are then isolated and purified. The binding of the pure compounds with proteins is validated through titration experiments (Chapter 4, 5 and 6). In step four, Chapter 5 and Chapter 6 included the isolation and structure elucidation of additional compounds from the selected extracts that showed protein-ligand complex with the purpose to provide evidence to that the LLE process can capture drug-like molecules.

CHAPTER TWO PREDICTING NATURAL PRODUCT VALUE, AN EXPLORATION OF ANTIMALARIAL DRUG SPACE

Abstract: Malaria remains a worldwide threat to human health. Many antimalarial drugs are available on the market. However, there are very few new chemical entities being developed to solve this global problem. From 1990 to 2012, 506 antimalarial natural products were reported in the literature. Here we present a perspective based on an analysis of the drug-like properties of the reported antimalarial natural products in order to assess drug potential.

2.1 Introduction

Malaria, caused by parasites of the genus *Plasmodium*, is one of the significant infectious diseases in many tropical and subtropical regions, where more than 3 billion people are at risk of being infected with malaria or developing disease.¹ It has been responsible for the death of approximately 600,000 deaths per year, and 78% of malaria deaths occurred in children aged under 5 years.¹

Malaria can be cured by undertaking therapeutic drug treatments suggested by World Health Organization (WHO). However, due to a number of factors such as limited access and treatment in developing countries and the emergence of antimalarial drug resistance, malaria still remains a leading threat to human health.¹ In addition to the resistance-mediated decline in efficacy, current antimalarial drugs have their shortcomings, such as side effects, etc. There is an urgent call for new efficacious and affordable chemotherapeutics in order to improve patient outcomes, reduce side effects, and reduce the emergence of resistance strains.

Nature is an infinite source of active compounds. Approximately 50% of the pharmaceuticals currently available are derived from natural products.¹⁰³ Compared to synthetic molecules, nature-derived compounds have a higher chemical diversity as well as an average increase of heteroatoms. It was also revealed that natural products are more similar to the chemical space occupied by drugs than compounds

obtained from synthesis.¹⁰⁴ Most natural products are biologically active and have better ADME (adsorption, distribution, metabolism and excretion) properties compared to synthetic molecules.¹⁰⁴

Historically, natural products play a vital role in antimalarial drug discovery. By way of example, in China, during the second century BCE, the Qinghao plant (*Artemisia annua*), which is also known as the annual or sweet wormwood, was described in the medical treatise.¹⁰⁵ In 340 CE, the antifever properties of Qinghao were first described.²⁸ The active ingredient of Qinghao, known as artemisinin, was isolated by Chinese scientists in 1971.¹⁰⁶ Derivatives of this extract, known collectively as artemisinins, are today very potent and effective antimalarial drugs, especially in combination with other medicines.¹⁰⁷

A number of natural products have been reported to have activity against malaria during the last two decades and are covered by a lot of recent comprehensive reviews of antimalarial natural products.¹⁰⁸⁻¹¹¹ However, the vast majority of these natural products have not been developed for treatment of malaria. In this chapter, we review literatures and present a perspective based on an analysis of drug-like properties of antimalarial natural products.

2.2 Data and method

2.2.1 Data

In order to analyze the physicochemical properties of antimalarial natural products, first we have to find out how many natural products were reported to have antimalarial activity in the last twenty years. A literature search was launched to achieve this goal. Malaria, natural product, and anti-malaria (or antimalarial) were the key words used to search literatures in the database, including Web of Science, PubMed, Embase, etc. Thousands of literatures were found. After a filtration process by reading their abstract, several literatures¹⁰⁸⁻¹¹¹ were selected to study. Hundreds of natural products with antimalarial activity involved in these literatures were recorded and their structures were drawn in InstantJChem 3.0.4 [InstantJChem 3.0.4,

2009 ChemAxon Ltd. (<http://www.chemaxon.com>)] to calculate the physicochemical properties of each compound. This process was repeated to find malaria drugs (including promising candidates in clinical trials) and to have their physicochemical properties calculated.

2.2.2 Method

ChemGPS-NP¹¹² was used for principle component analysis of antimalarial natural products and malaria drugs (including promising candidates in clinical trials), in order to compare the distribution of malaria drugs and antimalarial natural products in physicochemical space. Figures from 2.2.2.1 to 2.2.2.5 are used to illustrate how this method works.

1. To export your project from JChem click on “Export to file wizard” icon.

Cid	Structure	Mol Weight	Formula	Name	Reference	LogP	TPSA
1		304.47	C20H...	Azorellanone	19/2012/22	3.82	37.3
2		334.49	C21H...	17-acetoxy-13a-hydr...	19/2012/24	3.47	46.5
3		348.52	C22H...	Azorellanol	19/2012/23	3.70	46.5
4		304.47	C20H...	Mulinol	20/2012/25	3.82	32.7

Figure 2.2.2.1 Step One

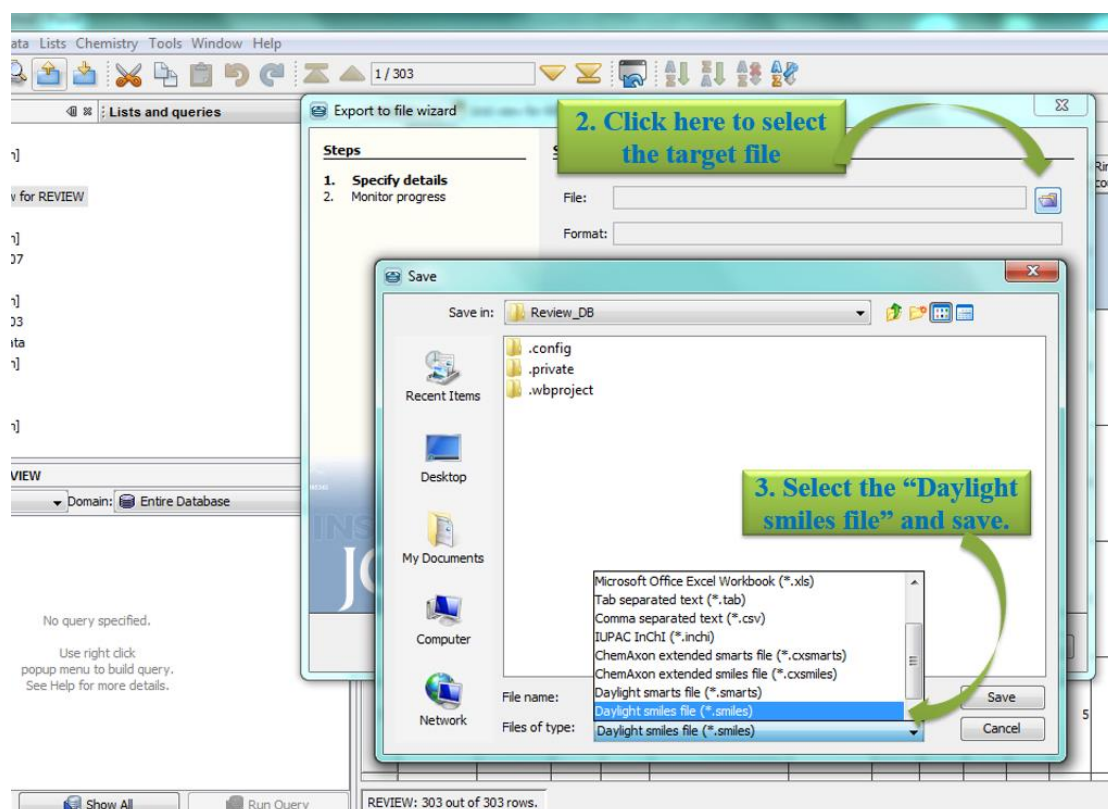


Figure 2.2.2.2 Step Two and Three

4. Go to ChemGPS-NP website. The address is:
<http://chemgps.bmc.uu.se/batchelor/queue.php?show=submit>

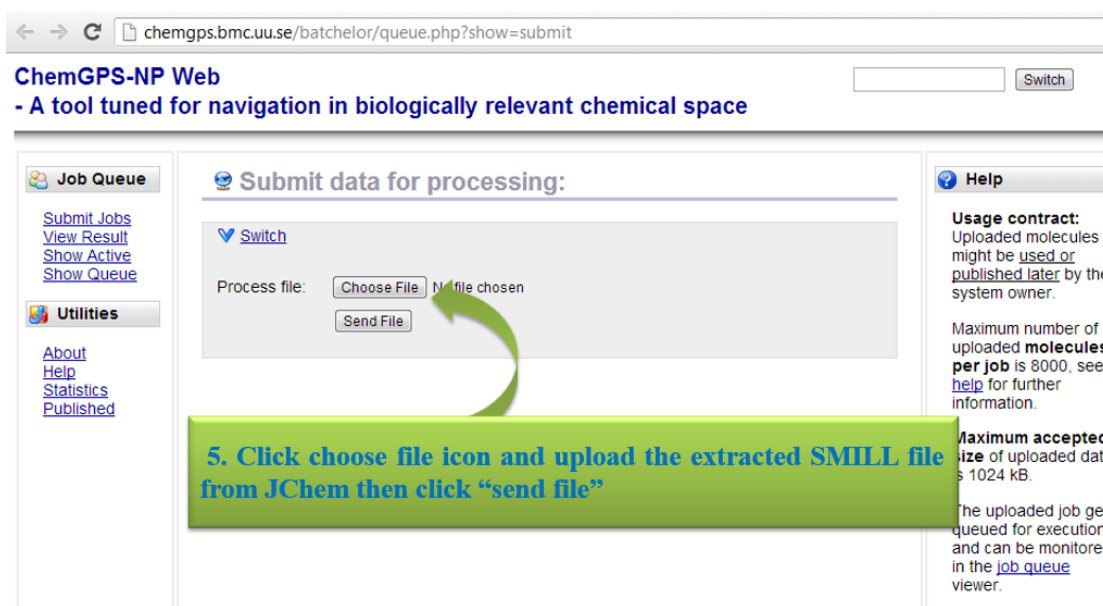


Figure 2.2.2.3 Step Four and Five

chemgps.bmc.uu.se/batchelor/queue.php?show=queue&sort=started&filter=finished&js=on

ChemGPS-NP Web

- A tool tuned for navigation in biologically relevant chemical space

Job Queue

[Submit Jobs](#)
[View Result](#)
[Show Active](#)
[Show Queue](#)

Utilities

[About](#)
[Help](#)
[Statistics](#)
[Published](#)

Job Queue:

Filter Options: Sort on: Started Show: Finished Refresh

Queued	Finished	Started	Job	Notice	Download	Delete	Publish
	✓	2013-03-29 12:19:20	Job 1675				
	✓	2013-03-29 10:45:53	Job 1674				
	✓	2013-03-29 07:18:03	Job 1672				

Delete Jobs

Help

Use the 'sort on' and 'show' droplist to filter jobs in queue.

Pressing the 'delete jobs' button will delete all jobs that are currently shown.

Simplified job control is enabled. Clicking the 'trash bin' icon on a running job will send the 'term' signal to the job.

Description of exit status icons:

- ✓ = Finished successful
- ⓘ = Finished with

6. Click "View result" to see or download the results

7. You can see or download the results here

Figure 2.2.2.4 Step Six and Seven

8. In next step you need to use a software (Matlab, Mathematica or Origin) to draw a 3D physicochemical space like the following picture, which is just an example. You can also compare the approved drugs with reported anti-XXX natural products in same graph.

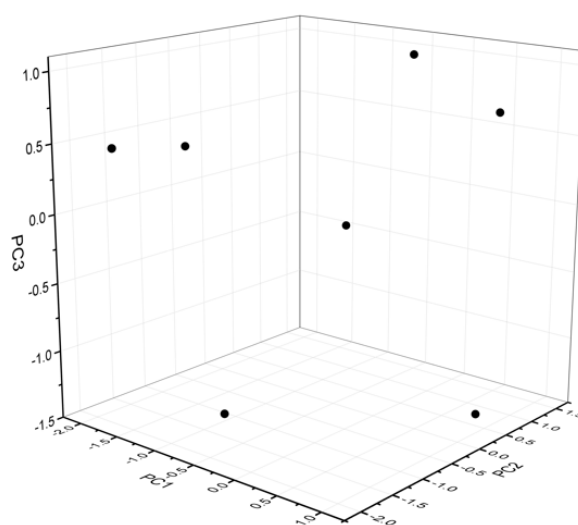


Figure 2.2.2.5 Step Eight. Each dot represents one molecule.

2.3 Results and discussion

Generally, the activity of any compound depends on its ability to permeate the cell membrane in order to reach and modify the target. By examining the physicochemical properties of the compound, it is possible to predict its cell permeability and bioavailability. Lipinski and his co-workers selected 2,245 compounds that reached phase II clinical trials from World Drug Index (WDI) and evaluated their physicochemical properties including molecular weight (MW), Log *P* (the logarithm of the partition coefficient between water and 1-octanol), hydrogen-bond acceptor number (HBA, such as number of N, O) and hydrogen-bond donor number (HBD, such as number of –NH, –OH), in order to understand which factors are responsible for compound attrition in clinical development. After their investigation based on a simple set of calculated physicochemical properties, the “rule of five” (Ro5) was released as their result.¹¹³ They concluded that for a compound to be drug-like it should have the cut-off numbers of five or multiples of five in MW, HBD, HBA and Log *P*, based on the comparison of drug-like and non-drug-like molecules. According to Lipinski’s Ro5, a drug-like compound should have a MW of 500 Da or less, equal or less than 5 HBDs, equal or less than 10 HBAs, and a partition coefficient Log *P* of 5 or less. If two or more properties are violated for a molecule, there is high probability of lack of oral activity and bioavailability.¹¹³ However, it is vital to keep in mind that even if a molecule is in line with the Ro5, there is still no guarantee that the compound is druggable. The aim of Ro5 was to guide medicinal chemists for a better selection and design of molecules, for the purpose to reduce the attrition during clinical development owing to unsatisfactory pharmacokinetics.¹¹⁴

Here, the four individual physicochemical properties MW, HBD, HBA, and Log *P* were analyzed for 31 malaria drugs (approved and candidate in clinical trials) as well as 506 antimalarial natural compounds reported between 1990 and 2012.

The following paragraphs are devoted to discussion of physicochemical properties of malaria drugs, followed by physicochemical properties of antimalarial natural products and comparison between these two sets.

2.3.1 Physicochemical properties of malaria drugs

The calculated data for physicochemical properties of approved malaria drugs and candidates in clinical trials are shown in Table 2.3.1.1.

Table 2.3.1.1 Physicochemical properties of 31 approved malaria drugs and candidates in clinical trials.

Name	MW*	Log <i>P</i> *	HBD*	HBA*	PSA*	ROTB*	Ro5*
tafenoquine	463.50	4.97	2	5	78.63	10	Pass
doxycycline	444.44	-3.29	6	9	181.62	2	Pass
artesanate	384.43	3.10	1	7	100.52	5	Pass
mefloquine	378.32	4.11	2	3	45.15	4	Pass
ferroquine	433.76	1.88	1	3	28.16	7	Pass
NITD-246	373.79	3.71	3	2	56.92	0	Pass
CDRI-97/78	500.59	4.96	1	7	100.52	10	Pass
fosmidomycin	138.06	-0.74	2	4	70.06	1	Pass
clindamycin	424.98	1.04	4	6	102.26	7	Pass
cipargamin	390.24	4.17	3	2	56.92	0	Pass
rosiglitazone	357.43	2.45	1	5	71.53	7	Pass
artefenomel	469.62	5.44	0	6	49.39	5	Pass
DSM-265	415.33	5.68	1	4	55.11	4	Pass
artenimol	284.35	2.84	1	5	57.15	0	Pass
piperaquine	535.52	5.27	0	6	38.74	6	Pass
KAF-156	411.46	3.07	2	4	76.18	4	Pass
arterolane	392.54	3.11	2	5	82.81	4	Pass
artemether	298.38	3.48	0	5	46.15	1	Pass
lumefantrine	253.73	2.38	5	5	83.79	2	Pass
atovaquone	528.94	9.19	1	2	23.47	10	Fail
proguanil	366.84	5.00	1	3	54.37	2	Pass
AQ-13	291.82	3.00	1	3	28.16	7	Pass
halofantrine	500.43	8.06	1	2	23.47	11	Fail
solithromycin	845.02	5.76	2	12	197.87	11	Fail
quinine	376.49	3.15	0	4	68.43	2	Pass
hypoestoxide	355.87	3.76	2	4	48.39	6	Pass
amodiaquine	518.06	4.20	2	7	73.75	7	Pass
pyronaridine	312.41	3.84	0	5	46.15	2	Pass

artemotil	369.47	2.38	2	5	72.48	8	Pass
bulaquin	319.88	3.93	1	3	28.16	8	Pass
chloroquine	324.42	2.51	1	4	45.59	4	Pass

* MW, molecular weight; Log *P*, partition coefficient; HBD, hydrogen bond donor; HBA, hydrogen bond acceptor; PSA, polar surface area; ROTB, rotatable bond; Ro5, rule of five.

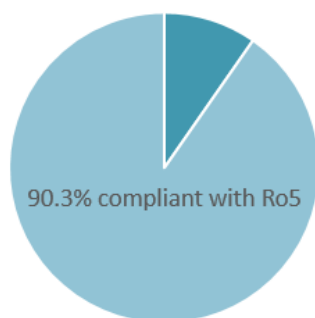


Figure 2.3.1.1 Pie chart presentation of the percentage of malaria drugs compliant with Lipinski's Ro5

The percentage of malaria drugs compliant with Lipinski's Ro5 is depicted in Figure 2.3.1.1. Approximately 90.3% of malaria drugs are in line with all of the Lipinski's Ro5 or just have one violation while 9.7 percent (three drugs), all of which were from semi-or synthesis, have two or more violations. The three drugs that violated Lipinski's Ro5 are atovaquone (**7**), halofantrine (**15**) and solithromycin (**24**) (Figure 2.3.1.2), which are all orally bioavailable drugs.

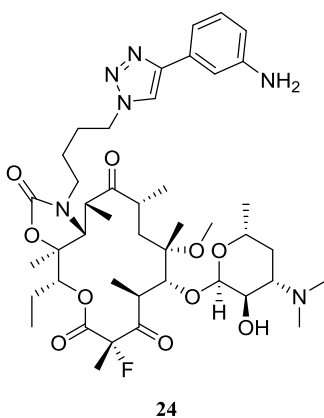


Figure 2.3.1.2 Structure of solithromycin (**24**)

As described in Chapter 1.2.1, atovaquone (**7**) is currently combined with proguanil (**3**) for the treatment and prophylaxis of *P. falciparum* malaria; halofantrine (**15**), a phenanthrene-methanol, is a highly effective treatment for malaria in many parts of the world and has been remarkably free from adverse effects.¹¹⁵ Solithromycin (**24**), a macrolide antibiotic, was developed for the potential treatment of *P. vivax* infection.¹¹⁶

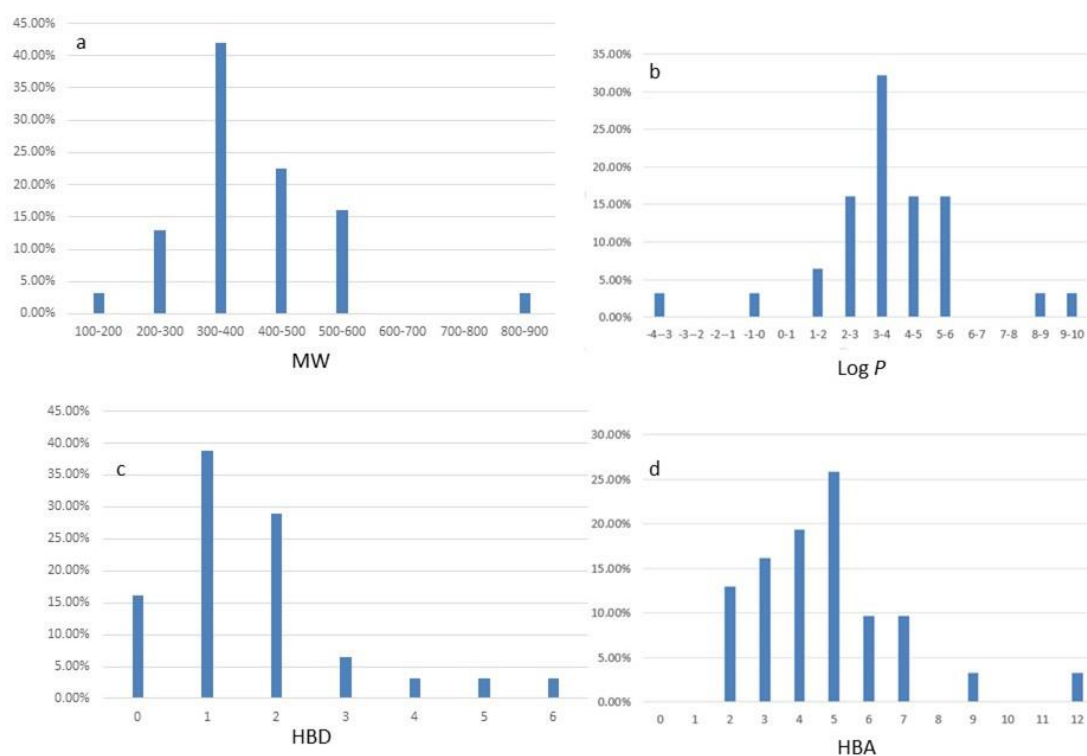


Figure 2.3.1.3 Physicochemical property histograms of malaria drugs: a) MW, b) calculated Log *P*, c) HBD, and d) HBA.

The physicochemical property histograms for MW, calculated Log *P*, HBD and HBA of these malaria drugs are depicted in Figure 2.3.1.3. Approximately 90.3% of malaria drugs have a molecular weight less than 500 Da (Figure 2.3.1.3a). The distribution of the calculated Log *P* (Figure 2.3.1.3b) revealed a wide range of variability. The majority of molecules are distributed between Log *P* of 1 to 6. The distribution of HBD (Figure 2.3.1.3c) showed a maxima at 1 and then steadily decreased, while a wide range of variability was observed in the distribution histogram of HBA (Figure

2.3.1.3d) of the malaria drugs. There was no compound that had violations from the HBD cut-off numbers 5. There was only one molecule that had violation against the HBA cut-off numbers of 10. This drug is solithromycin (**24**), which has 12 HBAs.

2.3.2 Physicochemical properties of antimalarial natural products (506) reported in the time frame of 1990-2012

A dataset consisting of 506 compounds reported in the literature during the period from 1990 to 2012 was compiled. The calculated physicochemical properties of the natural products with reported antimalarial activity is presented in the Appendix.

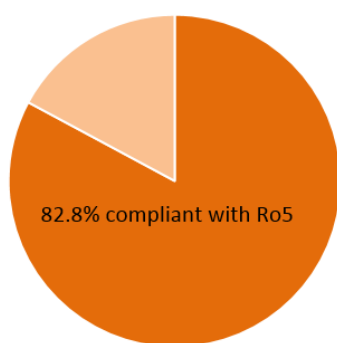


Figure 2.3.2.1 Pie chart presentation of the percentage of antimalarial natural products compliant with Lipinski's Ro5

The percentage of malaria drugs compliant with Lipinski's Ro5 is depicted in Figure 2.3.2.1. Approximately 82.8% of antimalarial natural products are in line with all of the Lipinski's Ro5 or just have one violation while 17.2 percent have two or more violations.

Histograms of the calculated physicochemical properties of antimalarial natural products are depicted in Figure 2.3.2.2. The histogram of molecular weight (Figure 2.3.2.2a) showed a peak at 300-400 Da. About 23.9% of the analyzed molecules had molecular weights more than 500 Da. The MW distribution was similar to the MW distribution of 126,140 natural products from Dictionary of Natural Products (DNP), which was previously examined for Lipinski's Ro5.¹¹⁷ The calculated Log *P* histogram of antimalarial natural products (Figure 2.3.2.2b) showed a normal distribution,

however, a shift to a higher log P value was noticed compared to other Lipinski's parameters. In contrast to malaria drugs where 9.7% showed negative calculated Log P , a higher percentage (25.9%) of the antimalarial natural products had a negative Log P . Another difference between these two sets was observed that the distribution of the calculated Log P for antimalarial natural products had a much wider range of variability. The distribution of HBD is depicted in Figure 2.3.2.2c. Similar to malaria drugs the distribution showed a peak at 1 following by a steady decrease. Only about 4.3% of these antimalarial natural products had more than 5 HBDs. The histogram of HBA (Figure 2.3.2.2d) showed a peak at 5 that included 21.7% of the molecules followed by a steady decrease. About 97.0% of these antimalarial natural products had no more than 10 HBAs.

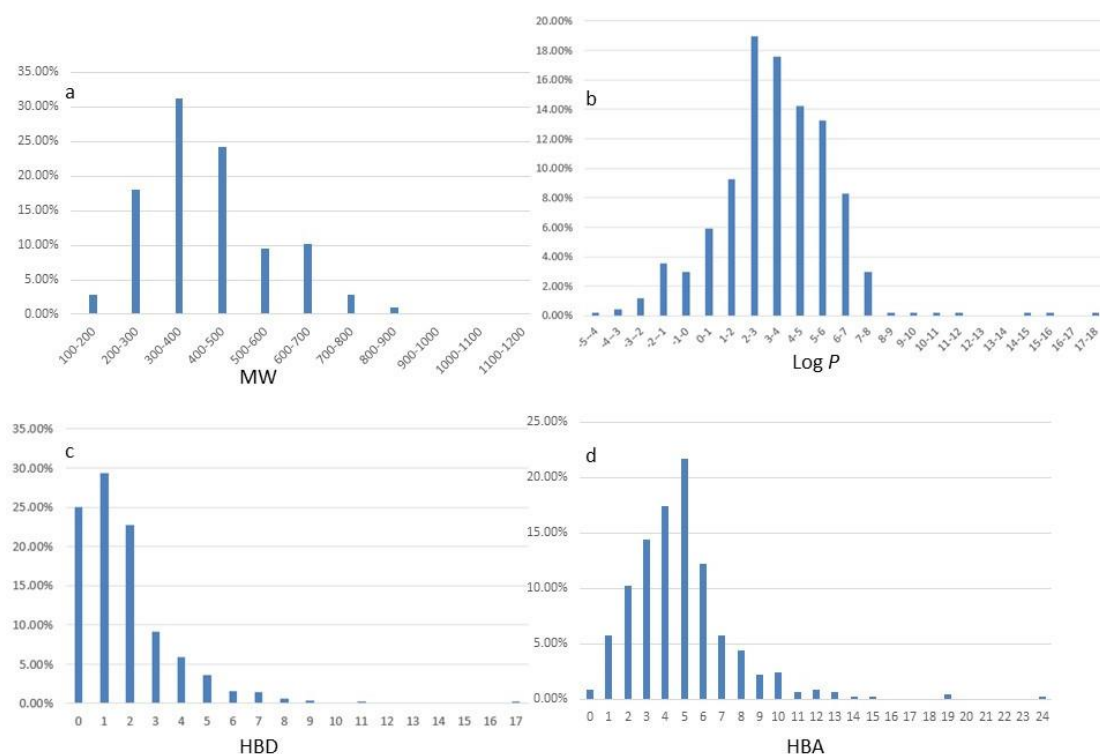


Figure 2.3.2.2 Physicochemical property histograms of antimalarial natural products: a) MW, b) calculated Log P , c) HBD, and d) HBA.

While 90.3% of malaria drugs were compliant with the Lipinski's Ro5, the proportion of antimalarial natural products was lower and about 82.8% of the molecules had no violation or just one violation from Lipinski's Ro5. The result for antimalarial natural

products is similar to two previous analyses of natural products.^{104, 117} In both researches, approximately 80% of analyzed molecules were in line with Lipinski's Ro5.

2.3.3 Comparison of antimalarial natural product physicochemical space vs current malaria drugs

With the purpose to compare the distribution of malaria drugs and antimalarial natural products in physicochemical space, ChemGPS-NP¹¹² was used for principle component analysis (PCA) of the 506 antimalarial natural products and the 31 malaria drugs (including both the marketed drugs and promising candidates in clinical trials). The first three principle components are plotted in Figures 3a and 3b. PC1 represents size, shape, and polarizability, PC2 expresses aromatic and conjugation related properties, and PC3 describes lipophilicity, polarity and H-bond capacity.¹¹²

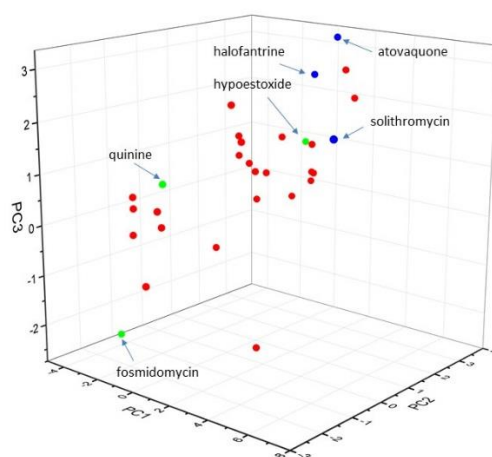


Figure 2.3.3.1 Score plot of malaria drugs.

Physicochemical space of 30 approved malaria drugs and candidates in clinical trials (except ferroquine, as it has the highest PC3 value of -168.856, which was too far away from the area focused) was demonstrated in Figure 2.3.3.1. Green dots show the three drugs, which are purely from nature. While, blue dots represents the three

drugs, which didn't obey Lipinski's Ro5. It shows that nature derived and synthetic malaria drugs occupy a broad range of physicochemical space.

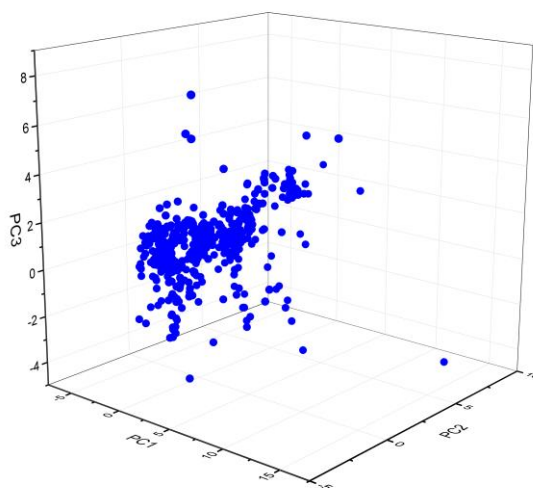


Figure 2.3.3.2 Score plot of antimalarial natural products.

Physicochemical space of 506 antimalarial natural products was demonstrated in Figure 2.3.3.2. It shows the majority of these molecules fall into drug-like physicochemical space.

In order to evaluate the 506 antimalarial natural products reported in the 1990-2012 timeframe against malaria drugs physicochemical space, their first three PCs are plotted and compared to malaria drugs in Figure 2.3.3.3. Red dots represent malaria drugs and blue dots show antimalarial natural products. Similar to the distribution of malaria drugs, the majority of antimalarial natural compounds fall into drug-like space. Compared to malaria drugs, antimalarial natural products occupy a much wider range of physicochemical space. However, there were more antimalarial natural products with two or more violations against Lipinski's parameters.

About 9.7% of malaria drugs and 23.9% antimalarial natural products have MW value more than 500 Da, which indicates that natural products are more likely to have a high MW. Log *P* plays a vital role in druggability of antimalarial natural products. About 25.8% of antimalarial natural products with Log *P* value more than 5, demonstrates the ability of lipophilic molecules to pass through the cell wall in order

to reach the target enzyme. As to malaria drugs, only three compounds had Log *P* value more than 5.

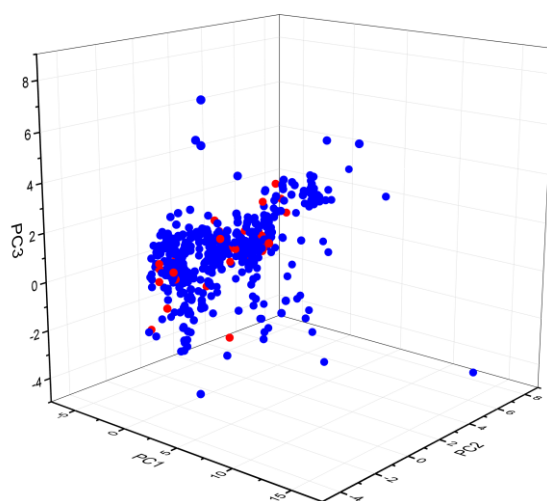


Figure 2.3.3.3 Comparison of the physicochemical space of 506 antimalarial natural products (blue dots) with physicochemical space of 30 malaria drugs and candidates (red dots).

2.4 Summary

Natural product-based malaria drugs are important components of malaria drugs and have significant efficacy in malaria treatment. Efforts of the past two decades for identification of antimalarial natural products have been very productive. About 506 natural compounds belonging to various chemical classes are reported to have antimalarial activity. Analysis of the physicochemical properties of these antimalarial natural products and comparison with the physicochemical properties of malaria drugs revealed major differences between the Log *P* distributions and the molecular weight distributions of these two sets. Antimalarial natural products showed a shift to higher Log *P* values and a high percentage (25.9%) had negative calculated Log *P*, while a small percentage of malaria drugs (9.7%) had negative Log *P* values. 76.1% of antimalarial natural products possess molecular weight no more than 500 Da whereas 90.3% of malaria drugs have a molecular weight no more than 500 Da.

Principle component analysis of the 506 antimalarial natural products and the 30 malaria drugs using ChemGPS-NP also indicated that not many antimalarial natural products were toward the cluster of the low molecular weight malaria drugs. The drug space of existing malaria drugs may help prioritization of a continuing discovery pipeline of antimalarial natural products and identification of antimalarial natural products reported in the literature.

CHAPTER THREE OPTIMIZATION FOR ELECTROSPRAY IONIZATION FOURIER TRANSFORM MASS SPECTROMETRY

Abstract: This chapter described an optimization strategy for protein detection on ESI-FTMS. Experiments were designed and employed to study the influence of methanol and trifluoroethanol as well as the ratio of ligand to protein. Results showed that the addition of methanol and trifluoroethanol at the concentration of <20% and <30%, respectively, can improve the intensity of protein signals. The proper ratio of ligand to protein (bCA II) to observe bCA II-ligand complex was 2:1.

3.1 Introduction

ESI-FTMS has proven to be a powerful tool in the study of noncovalent complexes.¹¹⁸⁻¹²⁰ It is well suited for high throughput affinity screening, and a variety of types of complexes can be analyzed. For example, RNA has been used as a target for drug discovery in an assay called multi-target affinity-specificity screening to interrogate the noncovalent interaction with components derived from a bacterial natural product library.¹²¹ There are also examples to apply FT-MS to detect protein-ligand complexes, including oligosaccharides,¹²² peptides,¹²³ and other proteins.¹²⁴

Despite its sensitivity and accuracy, bioaffinity ESI-FTMS still has technical limitations, particularly when the sample matrix is complicated as in natural product crude extracts,¹²⁵⁻¹²⁸ or when the sample preparation is intricate, such as protein with a nonvolatile buffer, which is incompatible with the ESI process.¹²⁹ As a result, further optimization is required to achieve best protein and complex desolvation in the presence of methanol or trifluoroethanol. In the initial spray solution, the protein concentration is an important factor that significantly affects the ESI-MS signal intensity. Optimization to find out a proper protein concentration to employ bioaffinity mass spectrometry screening is also required.

Proteins selected for ESI-FTMS optimization were carbonic anhydrase (CA, EC 4.2.1.1), *PfRab11a*, and *PfUCH-L3*. *PfRab11a* and *PfUCH-L3* were described in Chapter 1. CAs has been studied clinically for a long time. For instance, a role for CA inhibition as an anticancer therapy has been identified, indicating it could be a therapeutic target.¹³⁰⁻¹³¹ Poulsen has reported an application of ESI-FTMS to the direct screening of a dynamic combinatorial library against bCA II and observed bCA II-small molecule complexes.¹³² In this chapter, bCA II was used for the candidate to learn how to prepare samples for screening and how to operate the instrument.

3.2 Instrument used in this thesis

An actively shielded 4.7 Tesla superconducting magnet (Magnex Scientific Ltd., Oxford, UK) is cooled in a cryosystem by liquid helium to 4.2 Kelvin. The horizontal bore of the magnet is lined up with a metal frame supporting a vacuum system, ion sources and a gas inlet system. Data acquisition and analysis is performed using Xmass software running on a 1200 MHz Pentium III data station under Windows 2000 operating system. A picture of ESI-FTICR-MS used in this project is shown in Figure 3.2.1.



Figure 3.2.1 ESI-FTICR-MS, Bruker 4.7 Tesla

A schematic cutaway view shows different pressure regions of the ESI-FTICR-MS (Figure 3.2.2). The vacuum system has two regions, the ultra high vacuum analyzer and the source vacuum. The ultra high vacuum system is capable of sustaining a base pressure of below 4×10^{-10} mbar. The source vacuum system, which consists of two turbo-molecular pumps and a rotary pump, has a base pressure of 10^{-6} mbar.

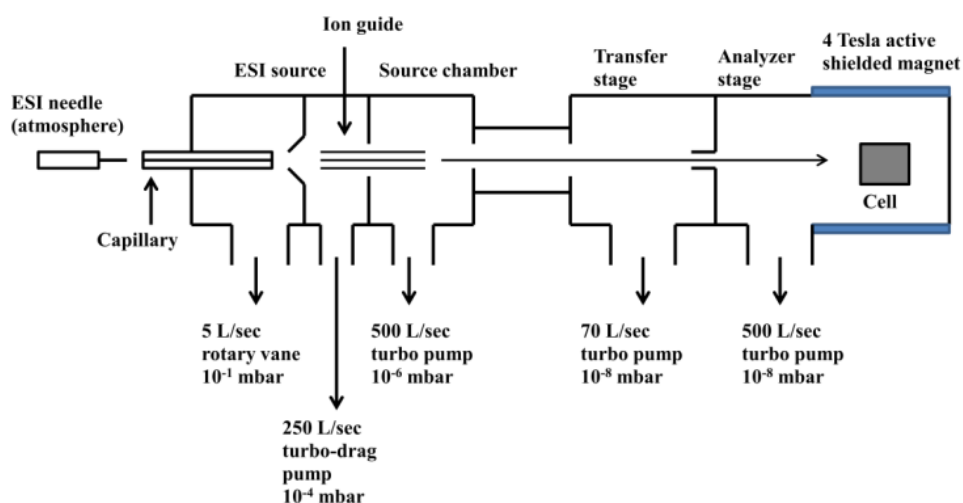


Figure 3.2.2 A schematic cutaway view of ESI-FTICR-MS (adapted from the Bruker Manual and drawn by the candidate)

Including a good ultra high vacuum system, a FT-MS instrument consists of three other main parts, which are external ion sources, ion transfer line and ICR, cell.

3.2.1 ESI source

The external ion source we have is called Apollo source (Figure 3.2.1.1). The main components of this source are: the off-axis grounded nebulizer, high voltage endplate, glass ion transport capillary, skimmer 1, pre-hexapole, skimmer 2, main hexapole, and the trapping/extracting electrode.

The distance from the off-axis and grounded nebulizer to the end plate is approximately 1.5 mm. Sample solution is injected into the spray chamber via the nebulizer. A spray of charged droplets is produced through an electrical field in the chamber, aided by a nebulizer nitrogen gas. Heated drying nitrogen gas helps the

desolvation and therefore the ionization of the droplets, as it flow counter current towards the stream of the droplets. Driven by the electric force, ions are moving towards the inlet glass capillary, which is 15 cm long, 0.5 mm in diameter and both its ends are capped with platinum. The capillary acts like a bridge between the atmosphere pressure and the vacuum. After transferred through the capillary, ions enter into the source vacuum.

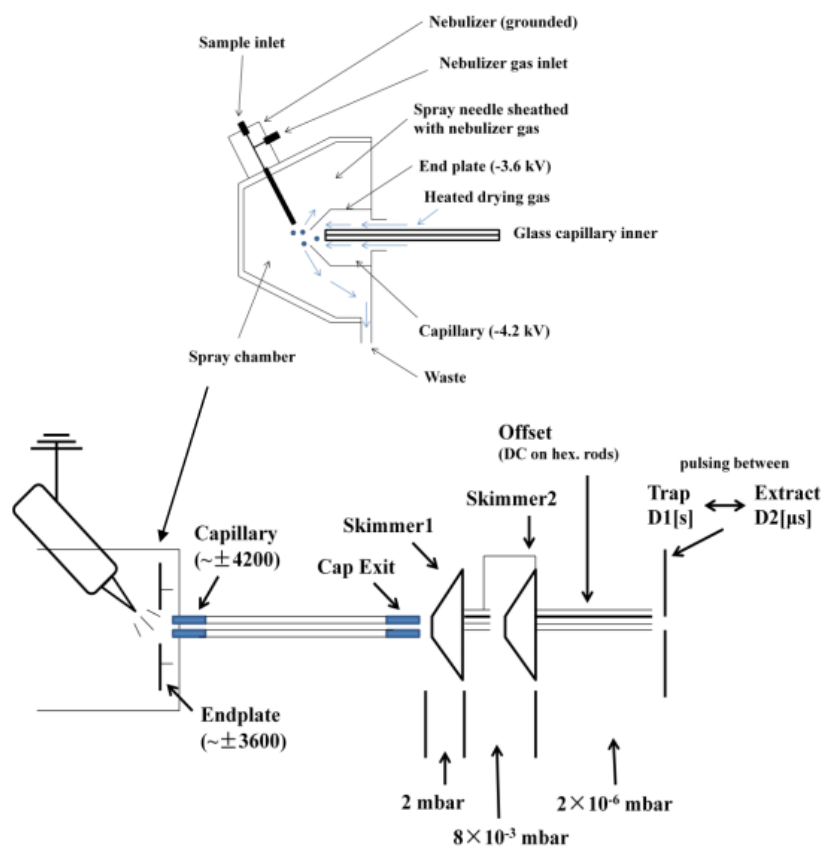


Figure 3.2.1.1 Apollo ESI source diagram (adapted from the Bruker Manual and drawn by the candidate)

Figure 3.2.1.1 demonstrates ions transfer from capillary through skimmer 1 and enter into pre-hexapole, which is biased at the same voltage as skimmer 2. Once they move through the pre-hexapole and skimmer 2, ions are accumulated in the main hexapole, which is biased at a DC offset voltage. By applying a positive voltage to the electrode for positive ions, they can be trapped between the skimmer 2 and the trap/extract electrode (the gate). The voltage of the trap/extract electrodes is

decreased after a user pre-determined time (d1) and ions come out from the hexapole. The gate continues to be open for a user-defined time (d2). At the end of the gate pulse, the voltage applied to the trap/extract electrodes is increased again and the next cycle of ion accumulation in the hexapole takes place. Because of this step, ions are produced in ion pulses from the ESI source into the ICR cell.

Increasing the potential difference between the transfer capillary exit and the skimmer 1, the collisions transform the kinetic energy of the ions into internal energy causing fragmentation of the ions. Therefore structural information can be detected. The name of this process is Collision Induced Dissociation (CID). All source parameters, which can affect the desolvation and dissociation, including flow rate, desolvation nebulising gas pressure, drying gas, end plated voltage, capillary voltage, temperature, capillary exit voltage, skimmer 1 and skimmer 2 can be adjusted on the Xmass tune page (Figure 3.2.1.2).

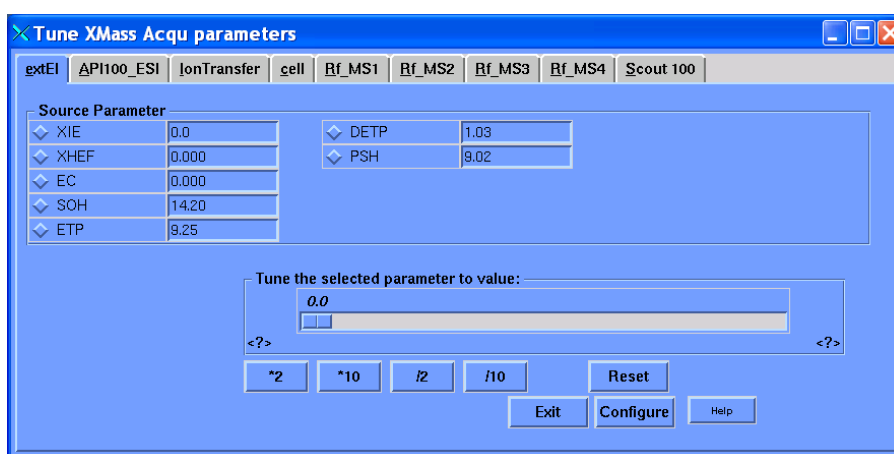


Figure 3.2.1.2 Xmass tune page for ESI source

3.2.2 Ion transfer line

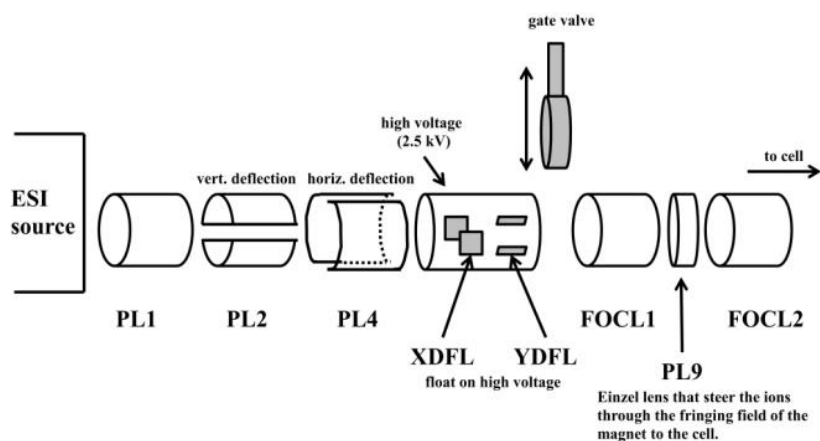


Figure 3.2.2.1 A schematic of the ion transfer line (adapted from the Bruker Manual and drawn by the candidate)

As shown in Figure 3.2.2.1, the ion transfer line includes deflection plates (PL1, PL2 and PL4), high acceleration voltage (XDFL and YDFL) and Einzel lens (FOCL1, PL9 and FOCL2). By speeding up them to high potential energy, ions generated in the ESI source are driven through various pumping circumstances. Then, ions are de-accelerated to help in ion capture within the analyzer or trapping cell.

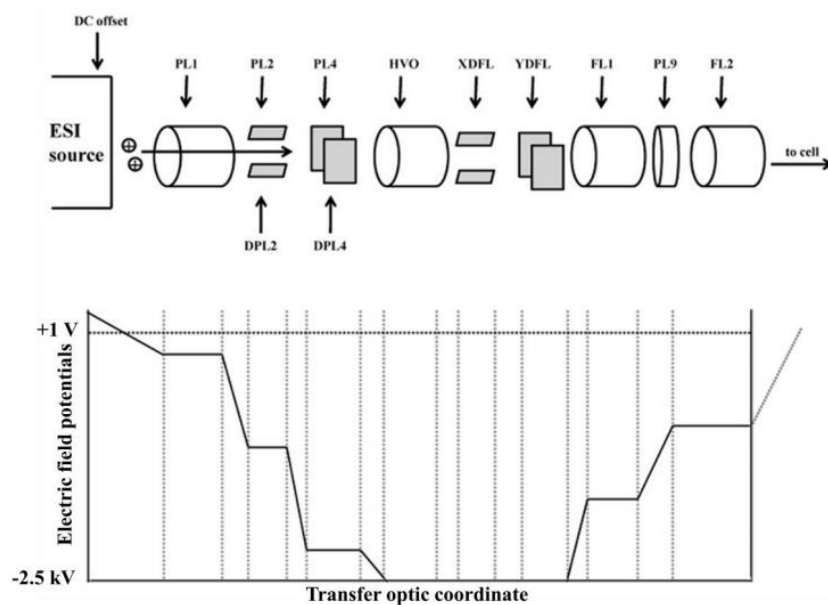


Figure 3.2.2.2 Electric potential values of the ion transfer optic line. A decrease voltage value from +1 V to -2.5 kV helps to accelerate the ions. An increase voltage from -2.5 kV to +1 V makes the ions slowly go into the ICR cell (adapted from the Bruker Manual and drawn by the candidate).

The correlation between the values of electric field potentials and the transfer optic coordinate is shown in Figure 3.2.2.2, which shows the acceleration and de-acceleration of the ions towards the analyzer cell.

3.2.3 ICR cell

Cubic or cylindrical design cells are most common in use today. Generally, a Bruker analyzer ICR cell has six electrodes, which consist of two trap electrodes with trapping voltage PV1, and PV2, two excitation electrodes and two detection electrodes. It also has two sidekick electrodes, which are used for sidekick trapping technique. Figure 3.2.3.1 is a drawing of Bruker ICR cell.

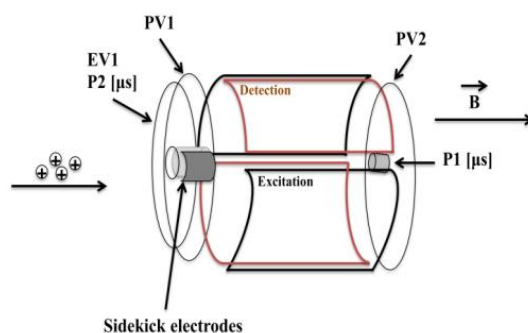


Figure 3.2.3.1 A drawing of Bruker ICR cell (adapted from Bruker Manual and drawn by the candidate).

In this thesis, ESI-FTMS is used to screen natural product extracts against malarial targets for antimalarial natural products.

3.3 Materials and methods

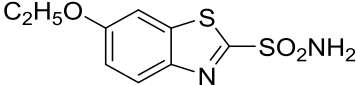
3.3.1 Materials

Bovine carbonic anhydrase II (bCA II, EC 4.2.1.1, 29089 Da) was purchased from Sigma-Aldrich, and used without further purification. It was dissolved in ammonium bicarbonate (10 mM, pH 7.0) to generate stock solution (34 μ M).

PfRab11a and *PfUCH-L3* were from Van Voorhis Group (University of Washington, WA, USA). They were dissolved in ammonium bicarbonate (10 mM, pH 7.0) to generate stock solution (20 μ M).

Ethoxzalamide (**25**) was purchased from Sigma-Aldrich (Table 3.3.1.1). It was dissolved in methanol to generate stock solutions (400 μ M).

Table 3.3.1.1 Structure and K_D of ethoxzalamide, an inhibitor of bCA II

Name	Structure	MW	K_D
Ethoxzalamide ¹³³⁻¹³⁴		258.31	0.25x10 ⁻⁹ M 0.15x10 ⁻⁹ M

3.3.2 Methods

Series of experiments were carried out as follows:

- (1) Sensitivity of protein (20 μ M, 10 μ M, 5 μ M, 2 μ M, 1 μ M) in buffer (ammonium acetate).
- (2) Influence of methanol (MeOH) (from 10% to 90%) on protein.
- (3) Influence of trifluoroethanol (TFE) (from 10% to 90%) on protein.
- (4) The ratio of ligand (ethoxzalamide for bCA II) and protein (10:1, 5:1, 2:1, 1:1).

All experiments were performed on a Bruker Apex III 4.7 Tesla external ESI source FTICR mass spectrometer. Samples were injected directly by a Cole-Parmer syringe pump with a flow rate of 2 μ L per minute. The end plate voltage was biased at 3200 V and the capillary voltage at 4500 V relative to the ESI needle during data acquisition. A nebulizing N₂ gas with a pressure of 50 psi and a counter-current drying N₂ gas with a flow of 30 L per minute were employed. The drying gas temperature was maintained at 100°C for direct infusion ESI-FTICR-MS. The capillary exit voltage was tuned at 140 V and skimmer 1 voltage at 15 V. Ions were accumulated in an external ion reservoir comprised of an re-only hexapole, a skimmer cone (skimmer 2) with a tuning voltage of 12 V, and an auxiliary gate electrode, prior to injection into the cylindrical Infinity™ analyzer cell, where they were mass analyzed. Each run took 3-5 minutes. These optimized conditions were used based on our previous studies.^{51, 135}

Mass spectra were recorded in the positive ion mode with mass range from 160-6000 m/z for broadband low-resolution acquisition. All data acquisitions were performed on a 1200 MHz Pentium III data station running Bruker's Xmass software under Windows 2000 operating system.

3.4 Results and discussion

3.4.1 Sensitivity of protein in ammonium acetate

bCA II was dissolved in ammonium acetate (10 mM, pH 7.0) to generate solutions with different concentration (20 μ M, 10 μ M, 5 μ M, 2 μ M, 1 μ M). The total volume of each solution was 100 μ l. After incubated at room temperature for 1 h, 15 μ l of each solution was injected into the ESI-FTMS for analysis. For nondenaturing ESI-MS, the buffer pH is an important factor. The buffer pH must be the same as that of physiological condition so that denaturing of protein does not occur and the formation of protein-ligand complex can take place. In this case, the pH of the ammonium acetate was at 7.0, which was in the range of physiological condition needed for the protein to maintain in its native form.

Mass spectrum of bCA II acquired under this physiological pH showed the peaks at charge states of 9+, 10+ and 11+ with the major peak at the charge state of 10+. Mass spectrum of *PfRab11a* acquired under this physiological pH showed the peaks at charge states of 9+, 10+ and 11+ with the major peak at the charge state of 10+. Mass spectrum of *PfUCH-L3* acquired under this physiological pH showed the peaks at charge states of 10+, 11+ and 12+ with the major peak at the charge state of 11+. The intensity of the major peak for each protein {[bCA II]¹⁰⁺, [*PfRab11a*]¹⁰⁺, and [*PfUCH-L3*]¹¹⁺} was plotted and shown in Figure 3.4.1.1. The result suggested that 10 μ M of bCA II, 5 μ M *PfRab11a* and 5 μ M *PfUCH-L3* gave the most intense protein signal respectively. This result indicated that an appropriate solution concentration was important for protein observation on ESI-FTMS. Either a much lower concentration or a much higher concentration should be avoided.

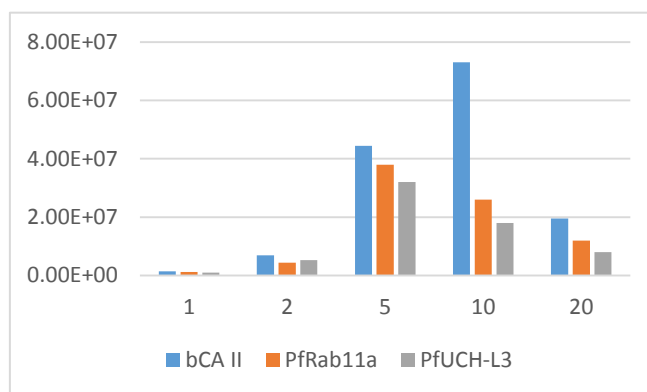


Figure 3.4.1.1 Protein intensity in 10 mM ammonium acetate.

3.4.2 Influence of methanol on protein

bCA II (10 μ M), PfRab11a (5 μ M), and PfUCH-L3 (5 μ M) were incubated with different MeOH concentrations (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90%), respectively. Samples were incubated overnight at 4°C before being analyzed by ESI-FTMS. The intensity of the major peak for each protein {[bCA II]¹⁰⁺, [PfRab11a]¹⁰⁺, and [PfUCH-L3]¹¹⁺} was plotted and shown in Figure 3.4.2.1.

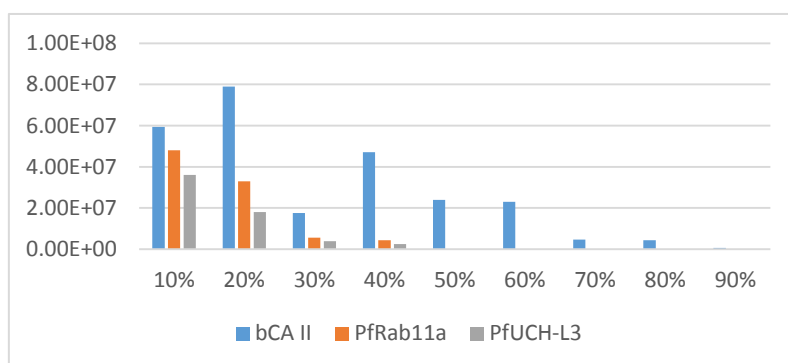


Figure 3.4.2.1 Influence of methanol on protein

Methanol (MeOH) can increase the solubility of natural products in buffer solution, thus improve the interaction between protein and ligands. However, too much methanol can cause protein denaturation.¹³⁶ The results indicated that low proportion of methanol (up to 20%) significantly increased the intensity of protein signal. And the 10-charged signal was much more stable and easy to observe. The signal was weak or even no signal when the concentration of methanol was very high,

probably because the protein was denatured.¹³⁶ Figure 3.4.2.2 is a spectrum of bCA II in 10mM ammonium acetate with 20% MeOH.

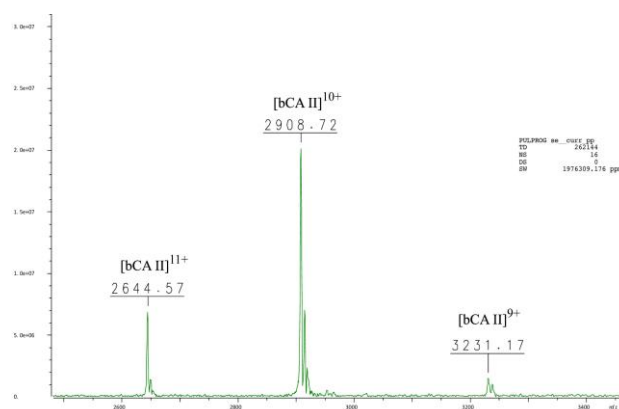


Figure 3.4.2.2 A spectrum of bCA II in 10mM ammonium acetate with 20% MeOH.

As a result, in the next stage of study, the percentage of methanol in the sample solution should be no more than 20%.

3.4.3 Influence of trifluoroethanol on protein

bCA II (10 μ M), *PfRab11a* (5 μ M), and *PfUCH-L3* (5 μ M) were incubated with different concentrations of TFE (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90%), respectively. Samples were incubated overnight at 4°C before being analyzed by ESI-FTMS. The intensity of the major peak for each protein {[bCA II]¹⁰⁺, [*PfRab11a*]¹⁰⁺, and [*PfUCH-L3*]¹¹⁺} was plotted and shown in Figure 3.4.3.1.

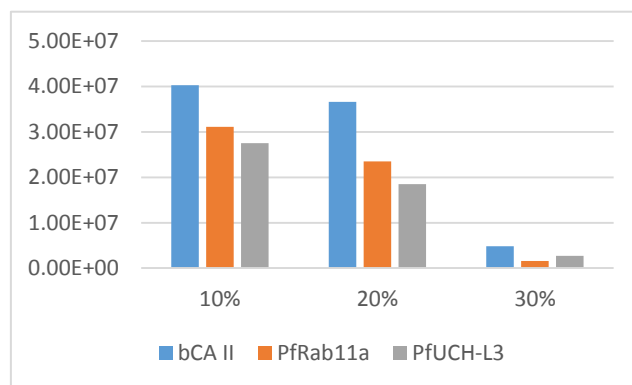


Figure 3.4.3.1 Influence of TFE on protein.

It is reported that low concentrations of trifluoroethanol (TFE) can stabilize protein tertiary structure.¹³⁷ The result suggested that there was no protein signal when the percentage of TFE reached 40%.

The results indicated that certain percentage of MeOH (no more than 20%) or TFE (no more than 30%) can improve the intensity of protein signal in FT-MS spectrum. When compared the effects of MeOH and TFE on the protein signal, MeOH gave a better result. For instance, the intensity of [bCA II]¹⁰⁺ with 20% of MeOH and 20% TFE was 7.9×10^7 and 2.8×10^7 respectively.

As a result, MeOH was used to dissolve the lead-like enhanced extracts and its percentage will be no more than 20% in the sample, which will be acquired on ESI-FTMS.

3.4.4 Influence of different ligand/protein ratios

Different ligand/protein ratios (10:1, 5:1, 2:1, and 1:1) were used to observe the influence of the ratio to the intensity of the protein. The results suggested that a 2:1 ratio of ligand to protein gave a better spectrum (Figure 3.4.4.1).

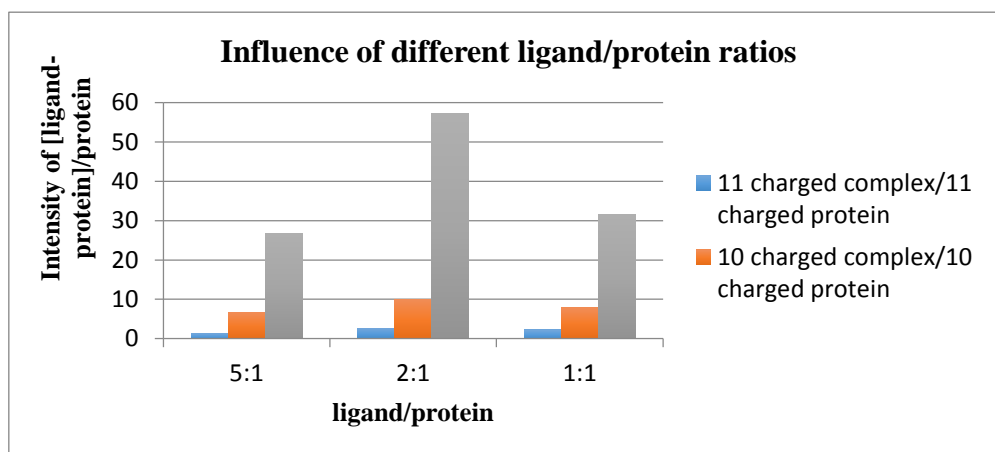


Figure 3.4.4.1 Compared different ligand/protein ratios.

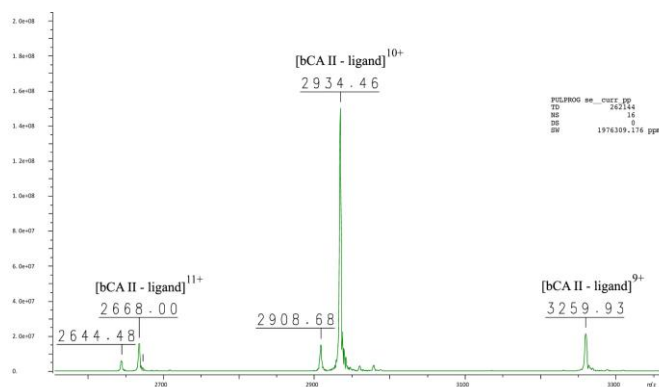


Figure 3.4.4.2 A spectrum of bCA II and its complex. For instance, 2644.48 is the signal of $[bCA II]^{11+}$, on its right hand side is the signal of $[bCA II\text{-}ligand]^{11+}$.

Figure 3.4.4.2 is an ESI-FTMS spectrum of bCA II and its complex formed with ethoxzalamide. As shown in this spectrum, six peaks were observed. The two at m/z 2644.48 and 2668.00 represented $[bCA II]^{11+}$ and $[bCA II\text{-}ligand]^{11+}$ complex respectively; peaks at m/z 2908.68 and 2934.46 represented $[bCA II]^{10+}$ and $[bCA II\text{-}ligand]^{10+}$ complex respectively; peaks at m/z 3259.93 represented $[bCA II\text{-}ligand]^{9+}$ complex. In this case, peak of $[bCA II]^{9+}$ was too weak to be observed. The molecular weight of the ligand (ethoxzalamide) was calculated based on the following equation, Molecular Weight of ethoxzalamide = $[m/z \text{ (bCA II-ethoxzalamide complex)} - m/z \text{ (free bCA II)}] \times (z) = [2668.00 - 2644.48] \times 11 = 258.72 \text{ (Da)}$.

3.5 Summary

In this chapter, an optimization strategy for protein detection on ESI-FTMS was successfully developed. According to the results, the addition of methanol and trifluoroethanol at concentration of <20% and <30%, respectively, can improve the intensity of protein signals. And, the proper ratio of ligand to protein (bCA II) to observe bCA II-ligand complex was 2:1. When compared with TFE, a percentage of MeOH (no more than 20%) can help to get higher intensity ESI-FTMS spectra. As a result, in the next stage of research, MeOH will be used to dissolve the LLE extracts and its percentage will be no more than 20% in the sample solution.

CHAPTER FOUR SCREENING LEAD-LIKE ENHANCED EXTRACTS AGAINST PFDUTPASE USING ESI-FTMS

Abstract: This chapter presents a strategy for using ESI-FTMS to detect the complex formed between PfdUTPase and its ligand in natural product extract matrix. A known PfdUTPase binder was spiked into a marine sample extract and the mixture was used to confirm the feasibility and sensitivity of the method. Then, 4,400 lead-like enhanced extracts were screened against PfdUTPase on ESI-FTMS. This chapter also describes the establishment of the lead-like enhanced extracts library and the cloning and purification of PfdUTPase.

4.1 Introduction

4.1.1 PfdUTPase

Described in Chapter 1.4.1.

4.1.2 Lead-like enhanced extract library

Eskitis lead-like enhanced extract (LLEE) library was established based on Nature bank, which consists of over 200,000 fractions derived from over 18,000 biota samples. The library was constructed by Lead-like enhanced (LLE) extraction procedure followed by reverse-phase HPLC fractionation. In comparison with traditional natural product (NP) drug discovery, which can carry out physicochemical assessment only after natural products are isolated, purified and their structures elucidated, the new protocol allows front-loading of the lead-like properties at the extract and fraction stage.¹³⁸

Lead-like enhanced extraction procedure involves solvent extraction by DCM (dichloromethane) and methanol (MeOH), solid phase extractions (SPE) eluted with 90% MeOH/10% H₂O to eliminate compounds with Log *P* > 5, and HPLC fractionation using reversed-phase solvent conditions (MeOH/H₂O/0.1% TFA) and a Phenomenex C18 Monolithic HPLC column (4.6 mm × 100 mm). Five fractions were collected per

sample within the region of the chromatogram corresponding to $\log P < 5$ (Figure 4.1.2.1). Previous work suggested that these five fractions mostly contain compounds with drug-like properties.¹³⁸ Five LLEFs were combined and dried to generate a LLE extract (LLEE). In this project, a LLEE library was built up and used for screening against *Pfd*UTPase on ESI-FT-MS.

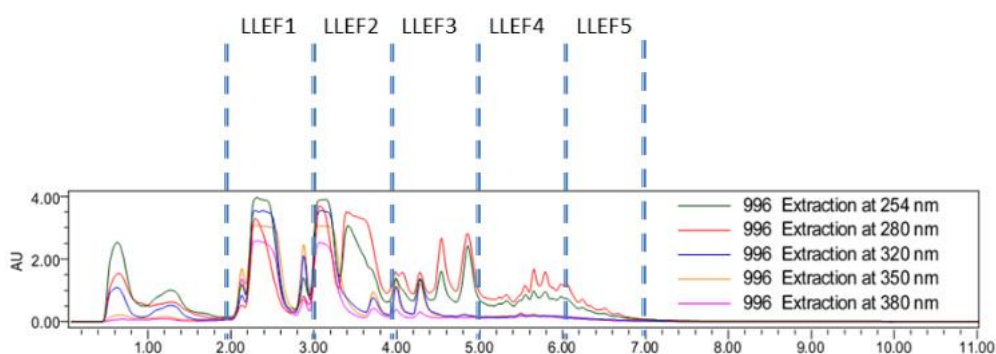


Figure 4.1.2.1 Example of a lead-like enhanced extract HPLC chromatogram

4.2 Materials and methods

4.2.1 *Pfd*UTPase and buffer exchange

*Pfd*UTPase was from Van Voorhis Group (University of Washington, WA, USA) and stored in a nonvolatile buffer (Tris, pH 8.0), which is incompatible with MS. The Tris buffer was then replaced by ammonium bicarbonate (10 mM, pH 7.0) through a buffer-exchange process.

Gel filtration uses size exclusion to separate proteins from small molecules. The separation takes place due to different access to the porous beads (gel filtration media). Small molecules permeate to the pores of the gel filtration media, while proteins are excluded from the pores and elute faster. Based on this characteristic, gel filtration is used for the desalting and buffer exchanging of proteins in non-volatile buffer. The gel column is first equilibrated with the volatile buffer. The protein is introduced to the column with volatile buffer constantly passing through. Small nonvolatile buffer molecules diffuse into the pores and elute later, whereas

the protein elute first together with the volatile buffer. In this study, Sephadex G25 column was used for desalting and buffer exchanging (Figure 4.2.1.1).

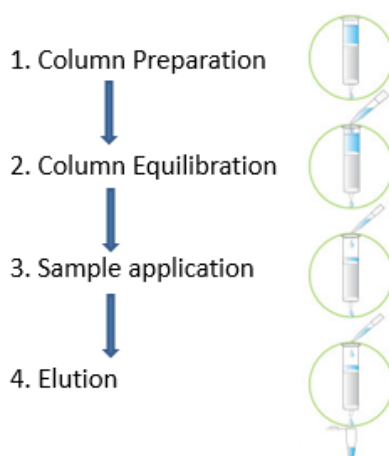


Figure 4.2.1.1 Buffer exchange steps. (1) Column Preparation. Excess storage buffer is allowed to flow through the column. (2) Column Equilibration. Column is equilibrated with buffer chosen by user (ammonium bicarbonate in this case). (3) Sample application. Sample is applied to column. (4) Elution. Purified sample is eluted from the column with buffer chosen by user (ammonium bicarbonate in this case).¹³⁵

4.2.2 A known binder to *PfdUTPase*

The ligand (Table 4.2.2.1) used in this project is a securinine derivative (**26**) according to our research. It was dissolved in MeOH to generate different concentrations (5 mM, 2.5 mM, 1.25 mM, 0.5 mM, 0.4 mM, 0.3 mM, 0.2 mM, and 0.1 mM).

Table 4.2.2.1 Structure of the known binder (**26**) to *PfdUTPase*

Structure	MW
	217.2

For the spiked experiments, one LLE extract was chosen because its HPLC chromatogram showed that it consisted of a number of chemical components (Figure 4.2.2.1).

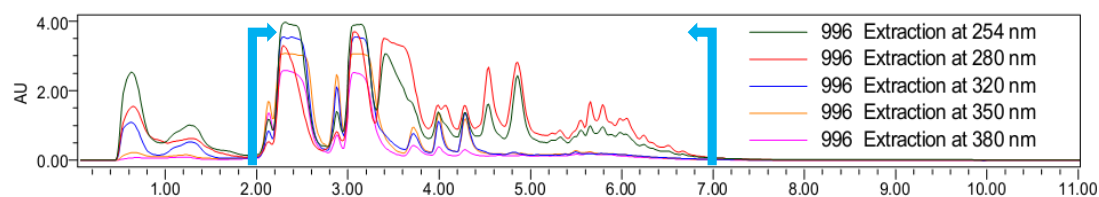


Figure 4.2.2.1 Chromatogram of chosen LLEF extract.

One set of samples for proof of concept experiments was prepared as follows:

LLE extract (5 μ L) and the ligand (1 μ L, 5 mM/2.5 mM/1.25 mM/0.5 mM/0.4 mM/0.3 mM/0.2 mM/0.1 mM) were incubated with *Pfd*UTPase (50 μ L, 18.0 μ M) in ammonium bicarbonate (10 mM, pH 7.0). The final concentrations of the ligand were 89.3 μ M, 44.6 μ M, 22.3 μ M, 8.9 μ M, 7.1 μ M, 5.4 μ M, 3.6 μ M, and 1.8 μ M.

The total volume of each sample was 56 μ L. They were incubated overnight at 4°C before being analyzed by ESI-FTMS.

4.2.3 LLEE library establishment

LLEE library has been built according to the LLE protocol as followed. A small amount (300 mg) of dried biota sample was packed into a SPE cartridge (Waters Oasis HLB, 400 mg), washed with *n*-hexane (8 mL), and then extracted with DCM (8 mL), then MeOH (8 mL). The DCM and MeOH extracts were combined, dried, and resuspended in DMSO (300 μ L). The extract (100 μ L) was fractionation by C18 analytical HPLC with solvent conditions consisting of a linear gradient from 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) to 50% H₂O (0.1% TFA)/50% MeOH (0.1% TFA) in 3 min at a flow rate of 4 mL/min, followed by a linear gradient to 100% MeOH (0.1% TFA) in 3.5 min at a flow rate of 3 mL/min. This was held at 100% MeOH (0.1% TFA) for 0.5 min at a flow rate of 3 mL/min and for further 1.0 min at a flow rate of 4 mL/min, then a linear gradient back to 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) in 1 min at a flow rate of 4 mL/min was applied. Finally, the gradient was held at 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) for 2 min at a flow rate of 4 mL/min, ready for the next injection. Total run time for each injection was 11 min, and 5 fractions were collected between 2.0 and 7.0 min, one min per fraction.

Five LLE fractions were combined, dried, resuspended in MeOH (200 μ L). 10 μ L was taken to a second 96-well V-plate for bioaffinity MS screening against *PfdUTPase*.

4.2.4 Bioaffinity ESI-FTMS screening

For the LLEE library screening, 5 μ L of LLEE sample (12.5 μ g/ μ L) were incubated with *PfdUTPase* (9.01 μ M) in ammonium bicarbonate (10 mM, pH 8.0) for direct infusion ESI-FTMS. The total volume of each sample was 55 μ L. They were incubated overnight at 4°C before being analyzed by MS. Fifty 96-well plates (each plate contains 88 samples; 4,400 samples in total) were screened. Mass spectral data was obtained in the positive ion mode on a Bruker Apex III 4.7 Tesla spectrometer, which was equipped with an ESI Apollo source. Samples were directly infused by a Cole-Parmer syringe pump with a flow rate of 120 μ L per minute. The end plate or counter electrode voltage was biased at 4,000 V and the capillary voltage at 4,400 V relative to the ESI needle. N₂ gas was used as nebulizing gas with a pressure of 40 psi and as counter-current drying N₂ gas with a flow of 20 L/min. The drying gas temperature was maintained at 125°C. The capillary exit voltage was tuned at 100 V. ESI mass spectra were recorded in the *m/z* range 50-5,000. Data acquisition and processing were performed using Xmass software.^{51, 135}

4.2.5 *PfdUTPase* expression and purification

The plasmid containing the *PfdUTPase* gene was a gift from Greg Crowther (Department of Medicine, The University of Washington). The full-length mature protein was cloned into the vector X that incorporates a HIS6-tag to give the following primary structure:

MAHHHHHHMHLKIVCLSDEVREMYKNHKTHHEGDSGLDLFIVKDEVLPKSTTFVKLGKAI
ALQYKSNYYYKCEKSENKKKDDDKSNIVNTSFLFPRSSISKTPRLANSIGLIDAGYRGEIIAALDNTS
DQEYHIKKNDKLVQLVSFTGEPLSFELVEELDETSRGEGGFGSTSNKY

The construct was transformed into the expression strain BL21(DE3)pLysS to enable over-expression of *PfdUTPase* using auto-induction method.¹³⁹ Cells were grown at

37°C for 4 hours and at 15°C for another 72 hours. Cells were harvested at 3000 rpm for 30 min in an Avanti J-E centrifuge (Beckman Coulter, USA).

Pellets were resuspended in buffer containing 20 mM Tris (pH 8.0) and 100 mM NaCl (50 mL per 20 g of pellet) and frozen at -80°C.

Cells were lysed using a thaw-freeze cycle for three times to lyse the bacterial cells.¹⁴⁰ In the freeze thaw cycle, resuspended cells were frozen at -80°C and then thawed in warm water. After the first thawing step, 100 microliter of PSMF inhibitor was added per 10 mL of lysate. The lysed cells were then sonicated to disrupt the cell membrane using a sonicator (Branson, USA) with a duty cycle of 80% for 5 min. Ultrasonics propagates in liquid media by pressure waves that alternatively expand and contract, thereby creating microbubbles or 'cavities'. Collapse of these cavities can produce extreme shear forces with the ability to disrupt membranes.¹⁴¹

Cells were then centrifuged at 40,000 rpm for 30 min at 4°C using an Optima XL 100K ultracentrifuge (Beckman, USA).

The supernatant was passed through a column prepared with freshly charged Ni-NTA resin, so that the His-tagged *Pfd*UTPase can bind to it. The resin was washed with 100 mL buffer containing 20 mM Tris (pH 8.5) and 100 mM NaCl. The protein was then eluted step-wise with the same buffer and 20 mM, 50 mM, 100 mM, 200 mM and 500 mM imidazole. A 12% SDS PAGE was performed to check for fractions that contain *Pfd*UTPase and to roughly determine the purity. Fractions containing protein were pooled and loaded onto a column with Q sepharose media for ion exchange to remove unwanted protein bands. A NaCl-gradient with buffer containing 20 mM Tris (pH 8.5) and 0.1 M NaCl was performed to further purify *Pfd*UTPase to remove unwanted protein side products. Fractions were assessed by SDS-PAGE on a 12% gel and fractions containing *Pfd*UTPase pooled and concentrated. A buffer exchange step was introduced so that the final buffer was 20 mM Tris (pH 8.5) and 150 mM NaCl.

4.3 Results and discussion

4.3.1 Desalting and buffer exchange

Size-exclusion chromatography (SEC) is a chromatographic method in which molecules in solution are separated by their size, and in some cases molecular weight.¹⁴² It is usually applied to large molecules or macromolecular complexes such as proteins and industrial polymers. Typically, when an aqueous solution is used to transport the sample through the column, the technique is known as gel-filtration chromatography. SEC is a widely used polymer characterization method because of its ability to provide good molar mass distribution results for polymers.

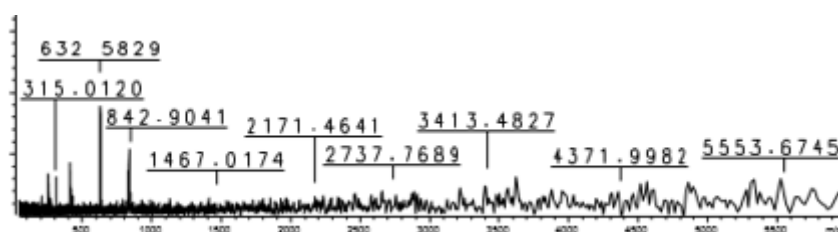


Figure 4.3.1.1 MS spectrum of *PfdUTPase* without buffer exchange

As shown in Figure 4.3.1.1, protein solution without buffer exchange was difficult to show signals in FTMS, compared with buffer exchanged protein solution (Figure 4.3.1.2)

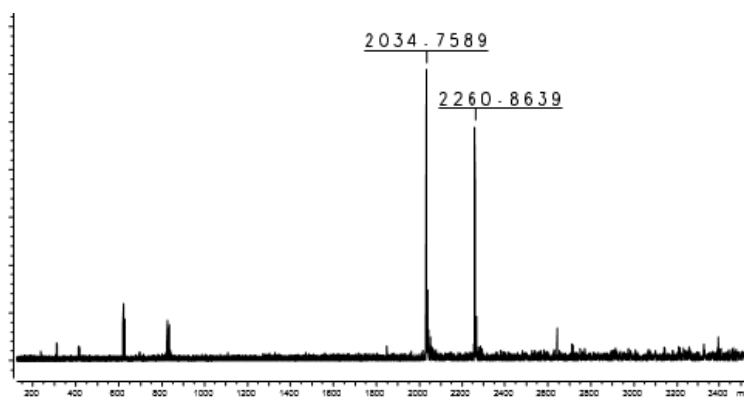


Figure 4.3.1.2 MS spectrum of buffer exchanged *PfdUTPase*

4.3.2 Detection of complex formed between *PfdUTPase* and its known ligand in NP extract matrix

All samples showed signals of unbound protein. 4 samples showed the signals of complexes. Some raw spectra of *Pfd*UTPase-LLE extract and *Pfd*UTPase-binder-LLE extract samples were presented in Figure 4.3.2.1.

In Figure 4.3.2.1, with the increased ligand concentration, the intensity of protein-ligand complexes was improved. According to the method discussed in Chapter Three, the molecular weight of the ligand was 217.2. This result confirmed the accuracy of the method and the instrument.

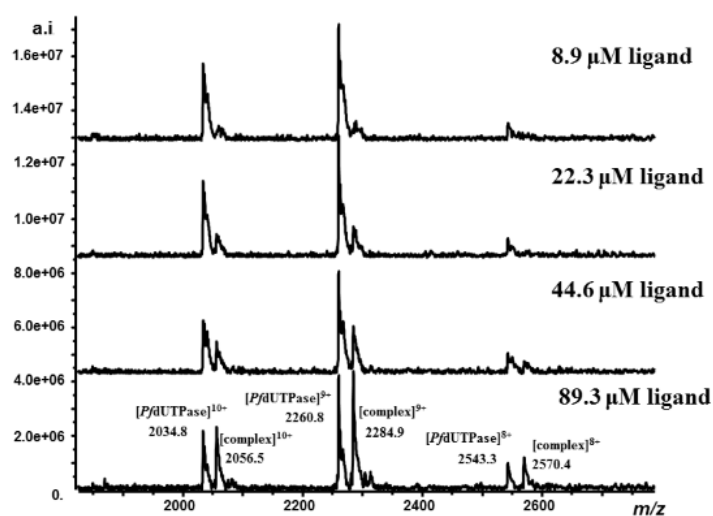


Figure 4.3.2.1 Comparison of protein-ligand complexes formed by *Pfd*UTPase with ligand in different concentrations.

The followed condition was used to test the sensitivity of the ESI-FTMS screening methodology. The concentration of the NP extract was 12.5 μg/μL and its volume used for each sample was 5 μL. The volume of ligand was 1 μL. The total sample volume was 56 μL. The ligand was titrated in a range from 89.3 μM to 1.8 μM, in the presence of protein and NP extract. The prtein-ligand complex was detected when the concentration of the ligand was increased to 8.9 μM. With this low concentration, a compound with percentage yield of 0.17% in the extract can be detected if it bound to the target protein. $[(8.9 \mu\text{M} \times 56 \times 10^{-6} \text{ L} \times 217.2 \mu\text{g}/\mu\text{mol})]/(5 \mu\text{L} \times 12.5$

$\mu\text{g}/\mu\text{L}) \times 100\%] = 0.17\%$. These results indicated that ESI-FTMS was sensitive and accurate to observe non-covalent complexes formed by protein target and its ligand.

4.3.3 LLEE library and Bioaffinity ESI-FTMS screening

Natural products have formed the basis of around half of all drugs discovered over the last two decades. Nature Bank at Eskitis Institute is a huge drug discovery source of desperately needed new medicines.

Nature Bank is a collection of over 72,000 samples from plants and marine organisms ready for screening against any disease. During the last twenty years the Eskitis Institute has acquired a significant biota collection of predominately terrestrial plants from Queensland, China and Papua New Guinea, and marine organisms from the Great Barrier Reef and Tasmania (Table 4.3.3.1). Value adding to the biota samples was achieved via a generic procedure that aligned extracts and subsequent fractions within lead- and drug-like physicochemical space based on Log *P*. The current methodology relies on initially preparing crude extracts that are then passed through an SPE cartridge containing a copolymer of divinylbenzene and *N*-vinylpyrrolidone to afford a lead-like enhanced extract. Subsequent chromatography provides simplified fractions for screening. Besides, Nature Bank is enriched with 4,500 pure natural products. These pure compounds were acquired in sub-milligram to gram quantities. In-house analysis showed that over 80% of the compound library comply with Lipinski's Rule of Five, i.e. $\text{MW} \leq 500$, $\log P \leq 5$, hydrogen bond acceptor (HBA) ≤ 10 , and hydrogen bond donor (HBD) ≤ 5 . This result validates the drug-like quality of the current Nature Bank compounds and is consistent with the published data calculated on a smaller library of 814 natural product compounds.

Table 4.3.3.1 Sources of the samples in Nature Bank

Country	Plants	Marine organism
Australia	From Queensland	From Queensland
	>20,000 plant samples	> 8,000 invertebrates
		> 3,500 species
	>8,000 plant species	From Tasmania
		> 1,200 invertebrates

	276 families	> 700 species
China	6,545 samples >2,000 plant species 183 plant families	NA
Papua New Guinea	5,743 samples >1,500 plant species 163 families	NA

In this project, 4,400 lead-like enhanced extracts were screened against *PfdUTPase* on ESI-FTMS. This was the most time-consuming and labour-requiring stage of the project. More frustrating no complexes were detected by FT-MS; therefore no binder was identified for *PfdUTPase*.

According to previous research, *PfdUTPase* was a validated malarial target and its securinine derivative binders showed antimalarial activity. Moreover, spiked experiments were successful. Failure to identify any potential binders may be due to many reasons, one being low percentage yield of any potential ligands or no binders in the LLE extracts screened.

4.3.4 *PfdUTPase* expression and purification

PfdUTPase was expressed and purified as described in Chapter 4.2.5. Figure 4.3.4.1 shows bands of eluted *PfdUTPase* from Ni-NTA resin. Most of *PfdUTPase* could be eluted in buffer containing 50 mM imidazole (Lane 5), but a band size of *PfdUTPase* could also be traced in buffer containing 500 mM imidazole (Lane 10).

After buffer exchange to ammonium bicarbonate (10 mM, pH 7.0), expressed and purified *PfdUTPase* (15.2 μ M) was injected into ESI-FTMS (Figure 4.3.4.2). The results showed that the purity of *PfdUTPase* was as good as the protein from the University of Washington. Thus, it can be used for bioaffinity FTMS screening in future research. This work played a vital role to provide enough protein material for FTMS screening.

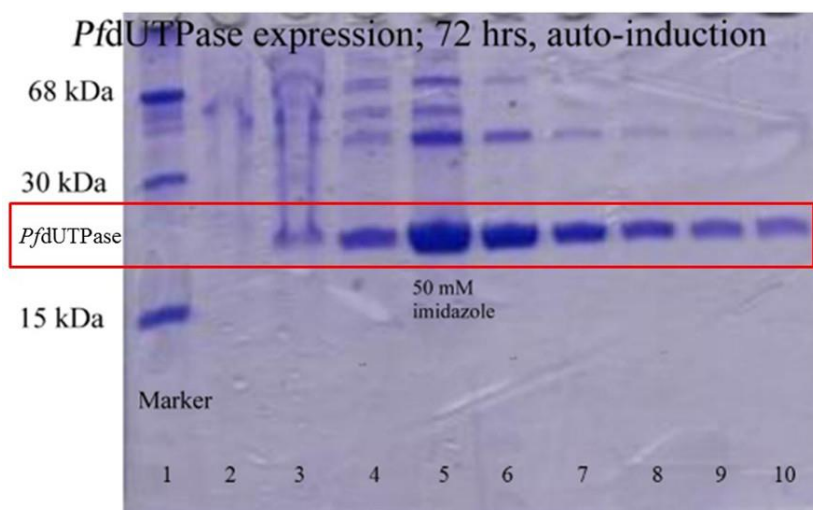


Figure 4.3.4.1 Lane 1: Marker (15, 30, 68 kDa); Lane 2: fraction (fr) 20/2; Lane 3: fr 20/5; Lane 4: fr 50/2; Lane 5: fr 50/5; Lane 6: fr 100/2; Lane 7: fr 100/5; Lane 8: fr 200/2; Lane 9: fr 200/5; Lane 10: 500/2. (For instance, fr 20/2 means this is fraction 2 of buffer containing 20 mM imidazole)

The final concentration of expressed *PfdUTPase* was 76.2 μM (5 mL).

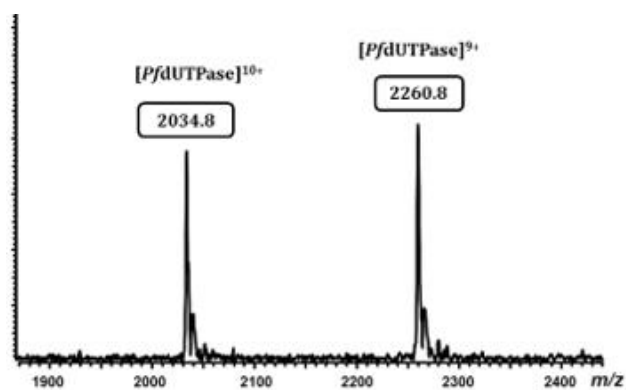


Figure 4.3.4.2 ESI-FTMS spectrum of *PfdUTPase* (15.2 μM after buffer exchange) made here. m/z 2034.8: [*dUTPase*] $^{10+}$; m/z 2260.8: [*dUTPase*] $^{9+}$.

4.4 Summary

This chapter described a strategy for using ESI-FTMS to detect the complex formed by *PfdUTPase* and its known ligand spiked in natural product extract matrixes. The results confirmed the feasibility and sensitivity of the method. 9,600 LLE extracts were prepared to establish a LLEE library according to Eskitis' in house protocol and

4,400 LLE extracts were screened against *Pfd*UTPase on ESI-FTMS. The cloning and purification of *Pfd*UTPase has been done with the purpose to keep the protein quantity. However, as there was no hit found with *Pfd*UTPase, other malarial targets would be screened in the next stage of research.

CHAPTER FIVE SCREENING LEAD-LIKE ENHANCED EXTRACTS AGAINST

PFRAB11A USING ESI-FTMS

STATEMENT OF CONTRIBUTION TO CO-AUTHORED PUBLISHED PAPER

This chapter includes a co-authored paper. The bibliographic details (if published or accepted for publication)/status (if prepared or submitted for publication) of the co-authored paper, including all authors, are:

Paper name: Bioaffinity Mass Spectrometry Screening.

Authors: Ben Yang, Yun Jiang Feng, Hoan Vu, Brendan McCormick, Jessica Rowley, Liliana Pedro, Gregory J. Crowther, Wesley C. Van Voorhis, Paul I. Forster, and Ronald J. Quinn.

This paper was published in the Journal of Biomolecular Screening in February 2016.

My contribution to the paper involved:

Bioaffinity screening of lead-like enhanced natural product extracts using ESI-FTMS; enzyme assay; the provision of the data; data analysis; manuscript writing.

(Signed) _____ (Date) _____
Name of Student

(Countersigned) _____ (Date) _____
Corresponding author of paper: Name of Corresponding Author

(Countersigned) _____ (Date) _____
Supervisor: Name of Supervisor

Abstract: This chapter presents a strategy for using ESI-FTMS to detect the complex formed between PfRab11a and its ligand in natural product extract matrix. Lead-like enhanced extracts were incubated with PfRab11a and then analyzed by ESI-FTMS. 3,520 extracts were screened against PfRab11a and five noncovalent complexes were

identified. Mass-directed isolation and purification gave five natural products (including one new natural product) as potentially active compounds. Subsequently, their binding to *PfRab11a* was confirmed by ESI-FTMS. The quantification of each ligand binding to *PfRab11a* was illustrated by measuring the dissociation constant (K_D) for the *PfRab11a*-ligand complex. The enzyme assay result also confirmed their binding activity. In order to predict the binding sites of these binders, virtual docking was carried out. This was the first time a bioaffinity MS screening method has been used to study the malarial target *PfRab11a*, and it is the first time that binders to this target and a new natural product inhibitor (arborside E) have been reported. In this chapter, contents related to arborside E has generated a manuscript, which was accepted by *Journal of Biomolecular Screening*. A number of natural products from the samples that gave positive FTMS screening result was isolated and purified based on spectroscopic data analysis to study the relationship between Log P value and HPLC retention time.

5.1 Introduction

This chapter demonstrates the ability of ESI-FTMS in detecting noncovalent complexes and identifying the active ligands in natural product extracts. The target selected for this study was *Plasmodium falciparum* Rab11a (*PfRab11a*, PF13_0119) with the molecular mass of 23 kDa.

ESI-FTMS was applied to screen lead-like enhanced extracts based on the formation of a noncovalent complex between a protein target and active constituent. 3,520 plant extracts from Nature Bank were incubated with *PfRab11a* and analyzed by ESI-FTMS. Five active extracts were identified. Five natural products isolated from these extracts were confirmed to bind to *PfRab11a*. The K_D of active compounds were measured based on titration experiments and their binding sites were predicted according to a virtual docking method. Additional compounds were isolated from the active extracts during the process; their structures were determined based on spectroscopic data analysis.

5.2 Results and discussion

5.2.1 Detection of protein by ESI-FTMS in physiological pH

PfRab11a was dissolved in ammonium acetate buffer solution (10 mM, pH 6.8) to give a protein solution (4.5 μ M). *PfRab11a* ammonium acetate solution (50 μ L) was injected into mass spectrometer to give peaks at charge states of 9+, 10+ and 11+ with the major peak at the charge state of 10+ (Figure 5.2.1.1). *PfRab11a* has the molecular weight of 23,238 Da, whilst three peaks at m/z 2583.1570, 2325.0116, and 2113.6818 were observed from MS spectrum, indicating a 9+, 10+, and 11+ charge state of the protein.

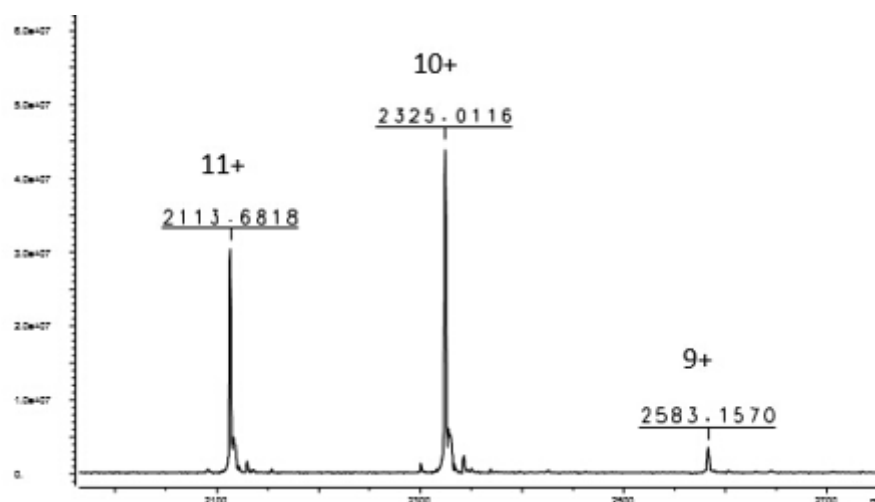


Figure 5.2.1.1 Mass spectrum of *PfRab11a* acquired under native condition

5.2.2 Detection of complexes formed by protein and active components in natural product extracts

3,520 extracts were screened against *PfRab11a* by FT-MS. Each extract (250 μ g/ μ L, 5 μ L) was incubated with *PfRab11a* (4.5 μ M, 50 μ L) at what temperature for how long. A portion of the protein-extract mixture (55 μ L) was directly infused to the ESI-FTMS. 98.6% of incubated samples showed signals of native protein or complexes.

Of the 3,520 extracts, five samples showed signals of *PfRab11a*-ligand noncovalent complexes in MS spectrum (Figure 5.2.2.1-5.2.2.5). Their mass spectra showed peaks of *PfRab11a*-ligand noncovalent complexes (Complex One to Five) with the major

peak at charge state 10+. The masses of the charge states, from 9+ to 11+, confirm the molecular mass of the protein to be $23,239 \pm 0.5$ Da were in excellent agreement for low-resolution mode with the calculated mass (23,238 Da) obtained from the amino acid composition. It was observed that the complex remained intact and there was still free protein in the assay solution. The experiments were performed in a low-resolution mode for a quick turnaround time (3-5 minutes per run).

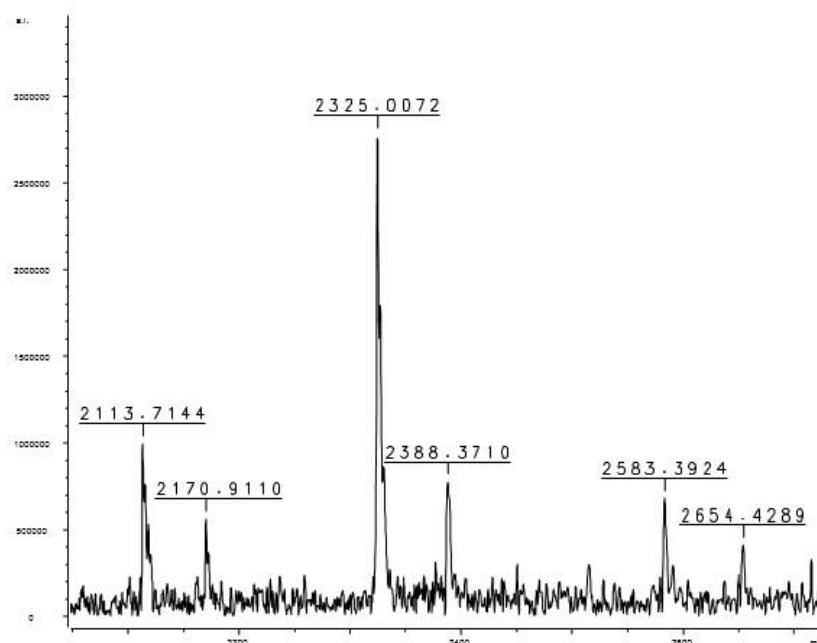


Figure 5.2.2.1 Mass spectrum of *PfRab11a*-ligand Complex One

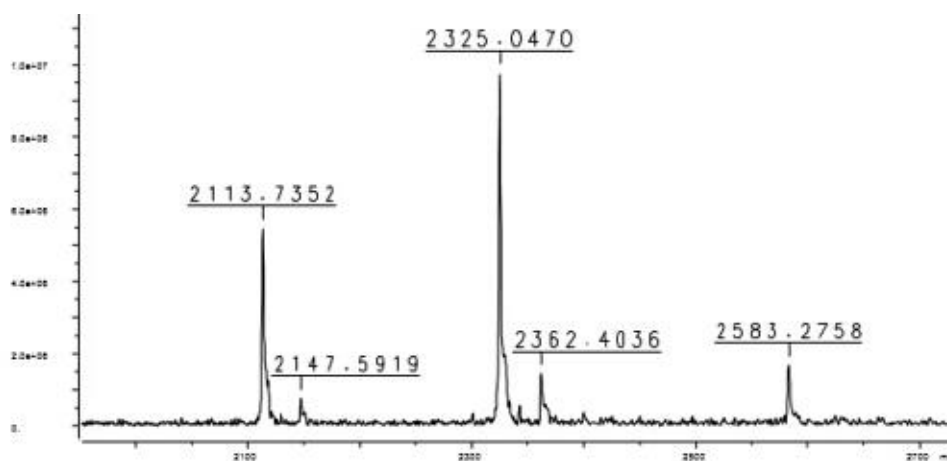


Figure 5.2.2.2 Mass spectrum of *PfRab11a*-ligand Complex Two

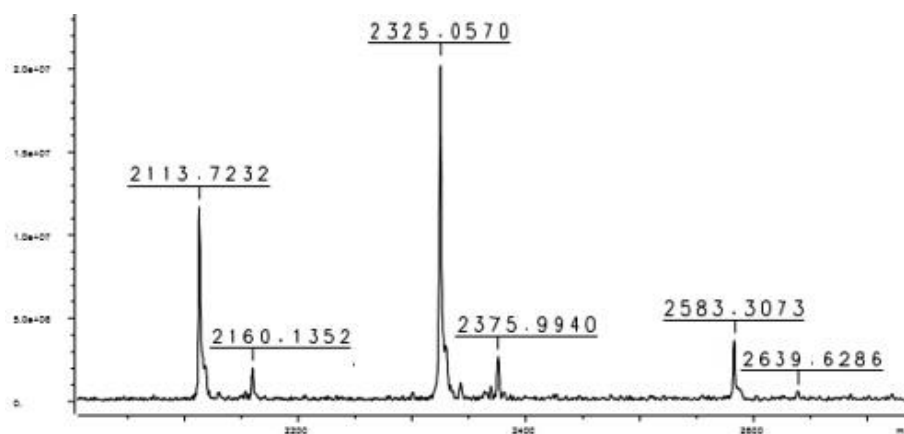


Figure 5.2.2.3 Mass spectrum of *PfRab11a*-ligand Complex Three

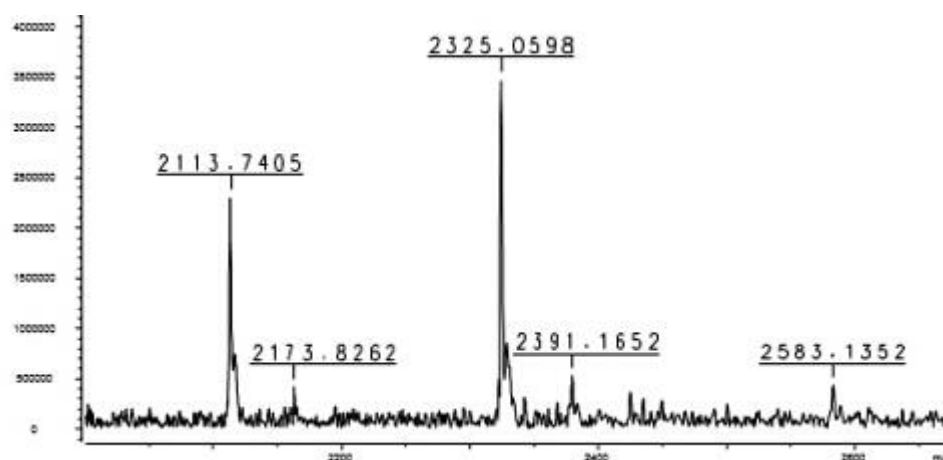


Figure 5.2.2.4 Mass spectrum of *PfRab11a*-ligand Complex Four

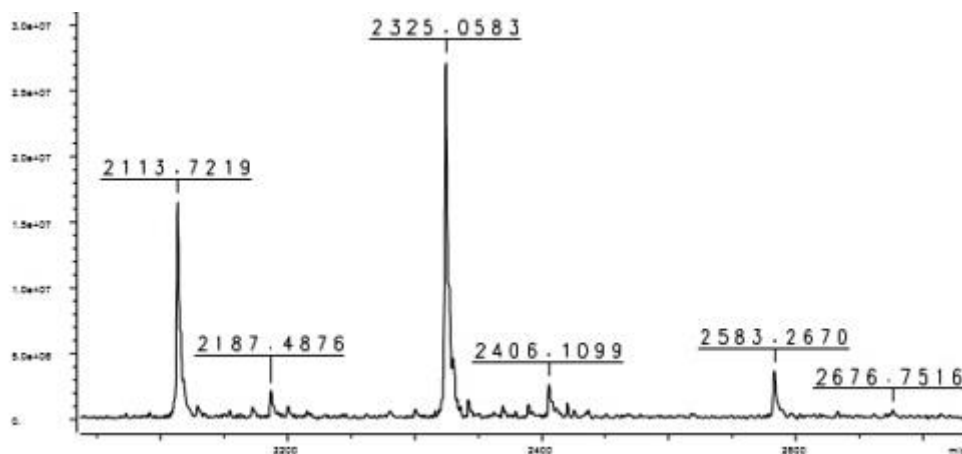


Figure 5.2.2.5 Mass spectrum of *PfRab11a*-ligand Complex Five

5.2.3 Determination of ligand molecular mass

Knowledge of molecular mass of an active component (ligand) in crude natural product extracts is vital to solve the bottleneck in the isolation process for natural

product discovery. If the molecular mass is known, the commonly used and time-consuming process of bio-guided fractionation will be shortened. Owing to the high resolution of the FTMS, the molecular mass of the ligand can be deduced from the mass-to-charge ratio (m/z) of the protein-ligand complex.

For example, in Figure 5.2.2.1, each charge state consisted of a grouping of two peaks: one peak at a lower m/z value corresponded to free *PfRab11a*, and the other peak at higher m/z corresponded to *PfRab11a*-ligand complex. The ligand mass is the difference between the average molecular mass of the complex and the average molecular mass of *PfRab11a*. Using the following equation, the molecular weight of ligand one was calculated as 629.1626 Da.¹⁴³

$$\text{MW}_{\text{Ligand 1}} = [m/z (\text{PfRab11a-natural product ligand complex}) - m/z (\text{free PfRab11a})] \times (z) = [2170.9110 - 2113.7144] \times 11 = 629.1626 \text{ Da}$$

The molecular weight of the other four natural product ligands (Ligand 2 to Ligand 5) was obtained in the same procedure as follows.

$$\text{MW}_{\text{Ligand 2}} = [m/z (\text{PfRab11a-natural product ligand complex}) - m/z (\text{free PfRab11a})] \times (z) = [2147.5191 - 2113.7352] \times 11 = 372.4237 \text{ Da}$$

$$\text{MW}_{\text{Ligand 3}} = [m/z (\text{PfRab11a-natural product ligand complex}) - m/z (\text{free PfRab11a})] \times (z) = [2160.1352 - 2113.7232] \times 11 = 510.5320 \text{ Da}$$

$$\text{MW}_{\text{Ligand 4r}} = [m/z (\text{PfRab11a-natural product ligand complex}) - m/z (\text{free PfRab11a})] \times (z) = [2173.8262 - 2113.7405] \times 11 = 660.9427 \text{ Da}$$

$$\text{MW}_{\text{Ligand 5}} = [m/z (\text{PfRab11a-natural product ligand complex}) - m/z (\text{free PfRab11a})] \times (z) = [2187.4876 - 2113.7219] \times 11 = 811.4227 \text{ Da}$$

In order to obtain sufficient quantity of the ligands for further study, mass-directed isolation and purification was undertaken to isolate these active natural products from the lead-like enhanced extracts, which led to the identification of arborside E (Ligand 1, MW 630.2; compound **29**) from the Australian plant *Psydrax montigena*, geniposidic acid (Ligand 2, MW 374.3; compound **33**) from the Australian plant *Wendlandia basistaminea*, 6-*O*-benzoylshanzhiside methyl ester (Ligand 3, MW

510.5; compound **39**) from Australian plant *Psydrax lamprophylla*, acantrifoside D (Ligand 4, MW 660.8; compound **45**) from the Australian plant *Prunus persica*, and rivaloside E (Ligand 5, MW 810.9; compound **52**) from the Australian plant *Larsenaikia ochreata*.

5.2.4 Structure elucidation of natural products and the analysis of their physicochemical properties

5.2.4.1 Structure elucidation

Ligand 1: MW 630.2

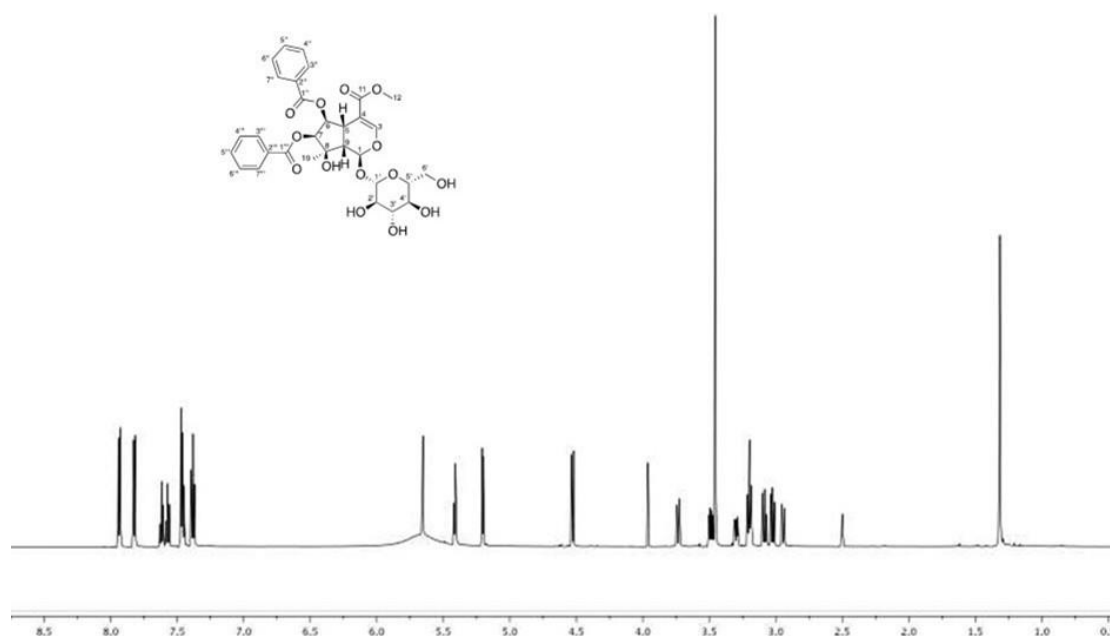


Figure 5.2.4.1.1 Chemical structure and ^1H NMR spectrum of **29**

Compound **29** was obtained as a light yellow, amorphous solid with an $[\alpha]_{\text{D}}^{28} + 10.8$ ($c=0.5$, CHCl_3). Its molecular formula was determined to be $\text{C}_{31}\text{H}_{34}\text{O}_{14}$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+$ ion at m/z 631.19419}, which was consistent with 15 degrees of unsaturation. The ^1H NMR spectrum (Figure 5.2.4.1.1) displayed characteristic signals for two mono-substituted aromatic rings [δ_{H} 7.81 (dd, $J=8.2$, 1.2 Hz, 2H), 7.56 (m, 1H), 7.37 (t, $J=1.6$, 2H)]; δ_{H} 7.92 (dd, $J=8.2$, 1.2 Hz, 2H), 7.59 (m, 1H), 7.45 (m, 2H)], a sugar moiety [δ_{H} 4.51 (d, $J=8.2$ Hz, 1H), 3.33-3.47 (m, 5H)], three oxygenated methines [δ_{H} 5.64 (d, $J=2.3$ Hz, 1H), 5.39 (t, $J=4.7$, 1H), 5.19 (d, $J=5.3$ Hz,

1H)], two other methines [δ_{H} 3.29 (m, 1H), 2.94 (dd, $J=11.2, 1.8$ Hz, 1H)], one methoxyl [δ_{H} 3.44 (s, 3H)], and one methyl [δ_{H} 1.30 (s, 3H)]. Analysis of the ^{13}C NMR (Figure 5.2.4.1.2) and HSQC (data provided in Table 5.2.4.1.1) spectra indicated a total of 31 resonances ascribed to two methyls (δ_{C} 51.4, 22.5), one methylene (δ_{C} 61.6), 23 methines [δ_{C} 151.9, 92.5, 79.1, 76.2, 47.3, 34.0; a sugar moiety (δ_{C} 98.2, 77.7, 77.0, 73.6, and 70.5); 12 aromatics (δ_{C} 133.7, 133.6, 129.9×2 , 129.8×2 , 129.5×2 , 129.0×2 , and 128.9×2], and five quaternary carbons [δ_{C} 109.4, 76.6; three carbonyls (δ_{C} 166.7, 165.5 and 165.0)].

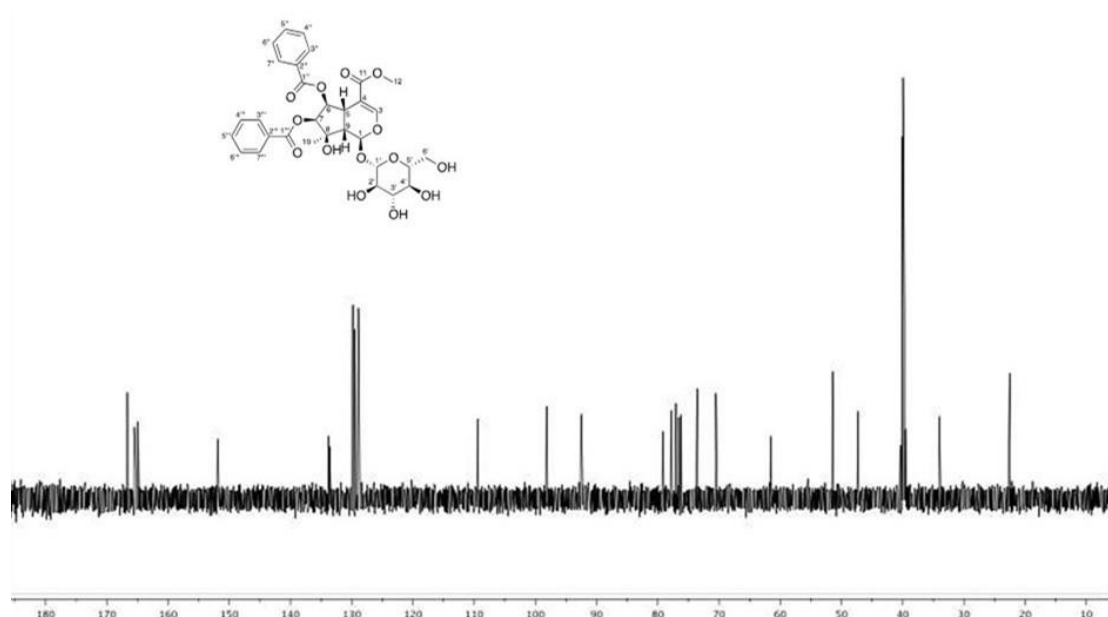


Figure 5.2.4.1.2 Chemical structure and ^{13}C NMR spectrum of **29**

Analysis of the COSY correlation data established four spin systems: two mono-substituted benzene rings (**a** and **b**), one sugar moiety (**c**), and one -CH-CH-CH-CH-CH-CH₂- moiety (**d**) (Figure 5.2.7.1.4, in bold). Further HMBC correlations from a methine doublet in **d** (δ_{H} 5.64, $J=2.3$ Hz, 1H) to an oxygenated olefinic carbon (δ_{C} 151.9), and from another methine in **d** (δ_{H} 3.29, m, 1H) to the oxygenated olefinic carbon (δ_{C} 151.9) and a quaternary carbon (δ_{C} 109.4) indicated the formation of a six-member ring system. HMBC correlations from the methine doublet (δ_{H} 5.64, $J=2.3$ Hz, 1H) and two oxygenated methines [δ_{H} 5.39 (t, $J=4.7$ Hz, 1H), 5.19 (d, $J=5.3$ Hz, 1H)] in **d** to the quaternary carbon (δ_{C} 76.6) indicated the

formation of a five-member ring system. Therefore, an iridoid nucleus was confirmed.

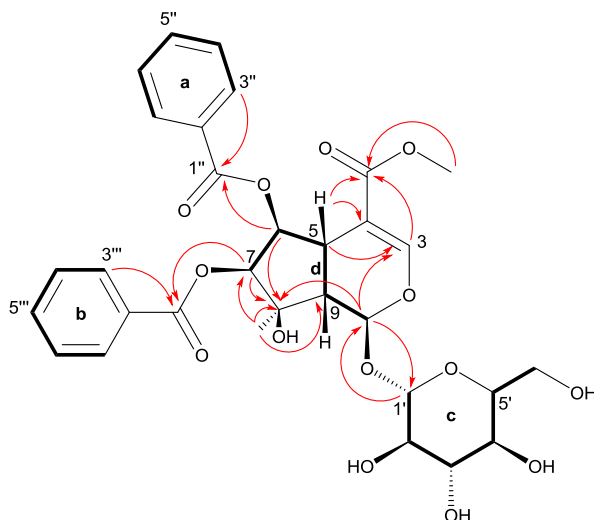


Figure 5.2.4.1.3 Key COSY (in bold) and HMBC correlations ($H \rightarrow C$) for **29**

On the basis of HMBC correlations from an aromatic proton signal in **a** [δ_H 7.81 (dd, $J=8.2, 1.2$ Hz, 2H)) to a carbonyl carbon (δ_C 165.0) and from δ_H 7.92 (dd, $J=8.2, 1.2$ Hz, 2H) in **b** to δ_C 165.5, two benzoyl moieties were established. Further HMBC correlation from a methine triplet in **d** (δ_H 5.39, $J=4.7$ Hz, 1H) to the carbonyl signal (δ_C 165.0) suggested that a benzoyl moiety was connected to the C-6 position via an ester functionality. Similarly, the second benzoyl moiety was connected to the C-7 position via an ester group by the presence of HMBC correlation from a methine doublet in **d** (δ_H 5.19, $J=5.3$ Hz, 1H) to the second carbonyl signal (δ_C 165.5).

The deshielded carbon chemical shift of anomeric carbon (δ_C 98.2), the presence of four oxymethine resonances and a deshielded methylene (δ_C 3.71/3.47, δ_C 61.6) suggested a monosaccharide. The coupling constant of the anomeric proton ($J=8.2$ Hz) and the chemical shift of the anomeric carbon (δ_C 98.2) indicated β -O-glycosidic linkage of the glucopyranoside moiety in **29**. HMBC correlations from a methine doublet in **d** (δ_H 5.64, $J=2.3$ Hz, 1H) to the anomeric carbon (δ_C 98.2), and from the anomeric proton (δ_H 4.51) to δ_C 92.5 in **d** suggested that the glucopyranoside moiety was attached to the C-1 position in **d**. HMBC correlations from H-3 (δ_H 7.45), H-5 (δ_H 3.29), and a methyl singlet (δ_H 3.44) to the carbonyl (δ_C 166.7) suggested that a

CH₃-O-C=O- moiety was connected to the C-4 position. The chemical shift of C-8 (δ_c 76.6) suggested that a hydroxyl group was present at C-8 position. A methyl was connected to C-8 position proven by the HMBC correlations from the methyl singlet (δ_H 1.30, 3H) to C-7, C-8 and C-9. Based on all the evidence, the planar structure of **29** was elucidated. All the key COSY and HMBC correlations were presented in Figure 5.2.4.1.3.

The relative configuration of **29** was inferred from the ROESY experiment (data provided in Table 5.2.4.1.1). The correlations between H-6/CH₃-10, H-7/CH₃-10, and H-1/CH₃-10 were observed clearly, suggesting that H-1, H-6, H-7 and CH₃-10 were co-facial. The presence of correlation of H-5/H-9 and the absence of correlations between H-5/H-7 and H-9/CH₃-10 indicated that these two protons were on the other face of the ring system. The relative configuration of **29** was therefore determined as shown in the structure.

Table 5.2.4.1.1 NMR spectroscopic data for arborside E*

No.	¹³ C	¹ H (mult., <i>J</i> in Hz, int.)	HMBC	ROESY
1	92.5	5.64 (d, 2.3, 1H)	3, 5, 8, 1'	9, 1'
3	151.9	7.45 (s, 1H)	1, 4, 5, 11	
4	109.4			
5	34.0	3.29 (m, 1H)	3, 4, 6, 9 4, 5, 7, 8, 9,	9
6	76.2	5.39 (t, 4.7, 1H)	1'' 5, 8, 9, 10,	7
7	79.1	5.19 (d, 5.3, 1H)	1'''	6
8	76.6			
9	47.3	2.94 (dd, 11.2, 1.8, 1H)	1, 4, 5, 8, 10	5
10	22.5	1.30 (s, 3H)	7, 8, 9	1, 6, 7
11	166.7			
12	51.4	3.44 (s, 3H)	11	
1'	98.2	4.51 (d, 8.2, 1H)	1	4'
2'	73.6	3.00 (m, 1H)	1', 3'	
3'	77.0	3.18 (m, 1H)	2', 4'	
4'	77.7	3.18 (m, 1H)	3', 5'	
5'	70.5	3.07 (m, 1H)	4', 6'	
6'	61.6	3.47 (dd, 11.7, 5.9,	4', 5'	4'

		1H)		
		3.71 (dd, 11.7, 1.8,		
		1H)	5'	4'
1''	165.0			
2''	129.9			
3''	129.5	7.81 (dd, 8.2, 1.2, 2H)	1'', 2'', 4'', 5''	4''
4''	128.9	7.37 (t, 1.6, 2H)	2'', 3'', 5''	
5''	133.6	7.56 (m, 1H)	3''	4''
6''	128.9	7.37 (t, 1.6, 2H)		
7''	129.5	7.81 (dd, 8.2, 1.2, 2H)		
1'''	165.5			
2'''	129.9			
			1''', 2''', 4''',	
3'''	129.8	7.92 (dd, 8.2, 1.2, 2H)	5'''	4'''
4'''	129.0	7.45 (m, 2H)	2''', 3''', 5'''	
5'''	133.7	7.59 (m, 1H)	3'''	4'''
6'''	129.0	7.45 (m, 2H)		
7'''	129.8	7.92 (dd, 8.2, 1.2, 2H)		

* Spectra were recorded in DMSO-*d*₆ at 600 MHz.

Ligand 2: MW 374.3

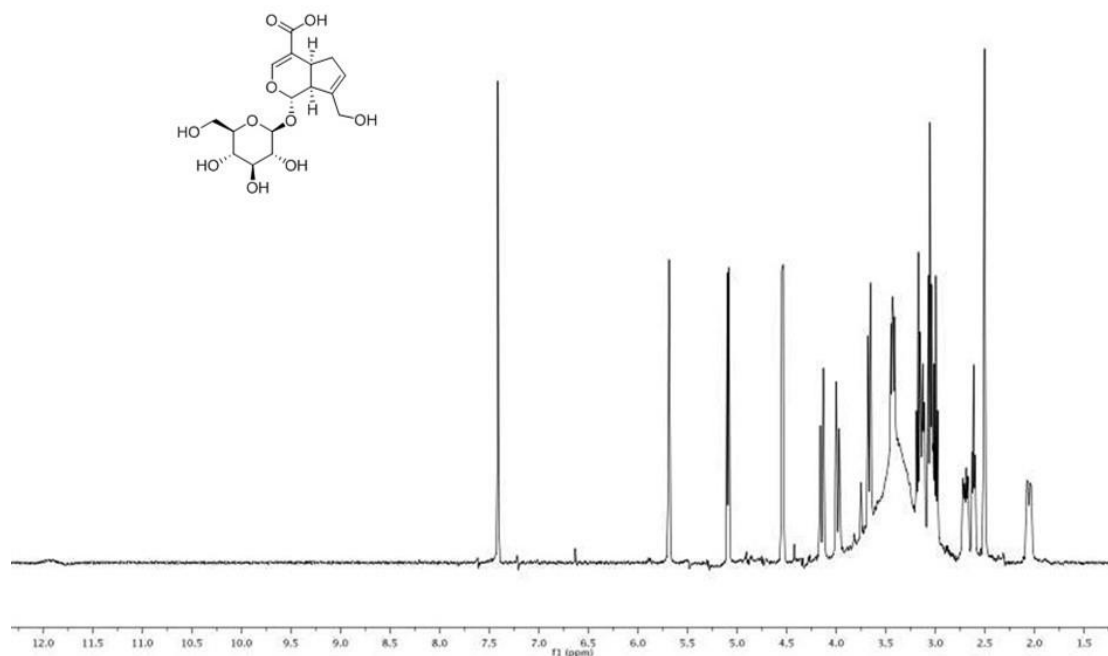


Figure 5.2.4.1.4 Chemical structure and ¹H NMR spectrum of compound **33**

Compound **33** was identified from the Australian plant *Wendlandia basistaminea*. It was obtained as yellow needles. Its molecular formula was determined to be

C₁₆H₂₂O₁₀ on the basis of (+)-HRESIMS measurements {[M-H]⁺ ion at *m/z* 373.5926}, which was consistent with 6 degrees of unsaturation. The ¹H NMR spectrum (Figure 5.2.4.1.4) displayed a sugar moiety and other signals. The ¹³C NMR spectrum showed a total of 16 resonances. Its molecular formula and genus name were searched in the DNP, and 33 hits were found. By comparing their ¹H and ¹³C NMR data, compound **33** was determined as geniposidic acid.¹⁴⁴ Geniposidic acid has antifungal, anti-inflammatory and other activities, it can be used to lower blood pressure in clinical, pain, and sedation.¹⁴⁵⁻¹⁴⁶ For instance, geniposidic acid showed inhibitory activity to the mycelial mass development of *Colletotrichum gloeosporioides* at 150 ug/mL.¹⁴⁷

Ligand 3: MW 510.5

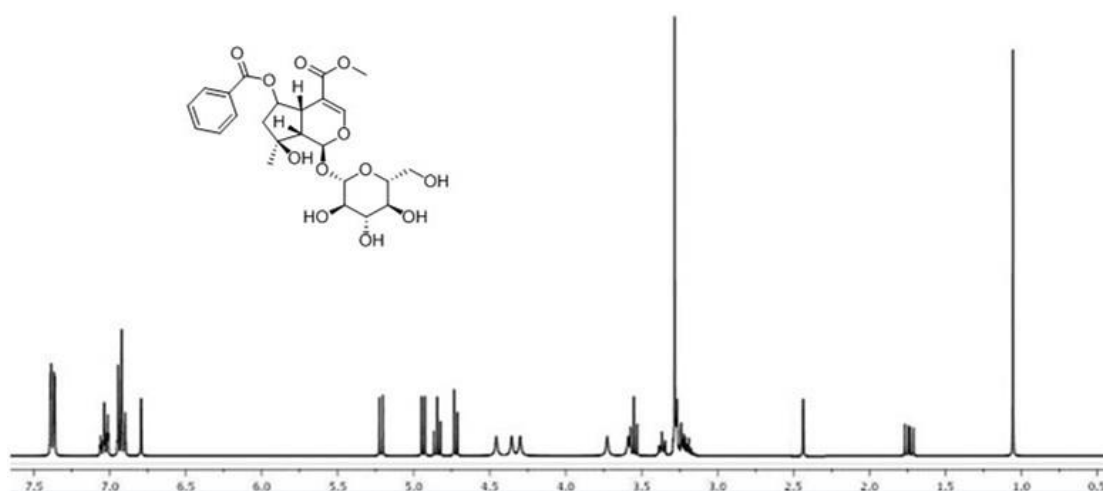


Figure 5.2.4.1.5 Chemical structure and ¹H NMR spectrum of **39**

Compound **39** was identified from the Australian plant *Psydrax lamprophylla*. It was obtained as brown powder. Its molecular formula was determined to be C₂₄H₃₀O₁₂ on the basis of (+)-HRESIMS measurements {[M+H]⁺ ion at *m/z* 511.5204}, which was consistent with 10 degrees of unsaturation. The ¹H NMR (Figure 5.2.4.1.5) displayed peaks for aromatics and sugar moieties. The ¹³C NMR spectrum showed a total of 24 resonances. Its molecular formula and genus name were searched in the DNP, and 31 hits were found. By comparing their ¹H and ¹³C NMR data, compound **39** was determined as 6-*O*-benzoylshanzhiside methyl ester.¹⁴⁸

Ligand 4: MW 660.9

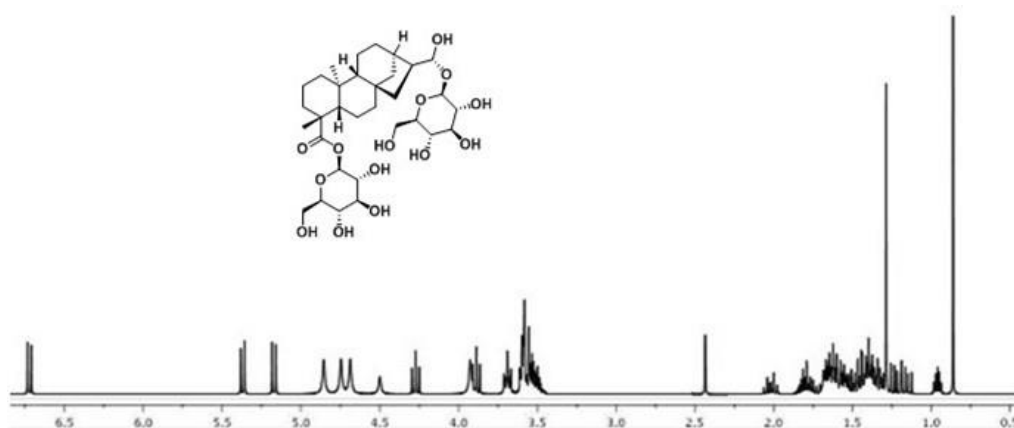


Figure 5.2.4.1.6 Chemical structure and ¹H NMR spectrum of **45**

Compound **45** was identified from the Australian plant *Prunus persica*. It was obtained as white needles. Its molecular formula was determined to be C₃₂H₅₂O₁₄ on the basis of (+)-HRESIMS measurements {[M+H]⁺ ion at *m/z* 661.7984}, which was consistent with 10 degrees of unsaturation. The ¹H NMR (Figure 5.2.4.1.6) displayed sugar moieties and other signals. The ¹³C NMR spectrum showed a total of 32 resonances. Its molecular formula and genus name were searched in the DNP, and five hits were found. By comparing their ¹H and ¹³C NMR data, compound **45** was determined as acantrifoside D.¹⁴⁹

Ligand 5: MW 810.9

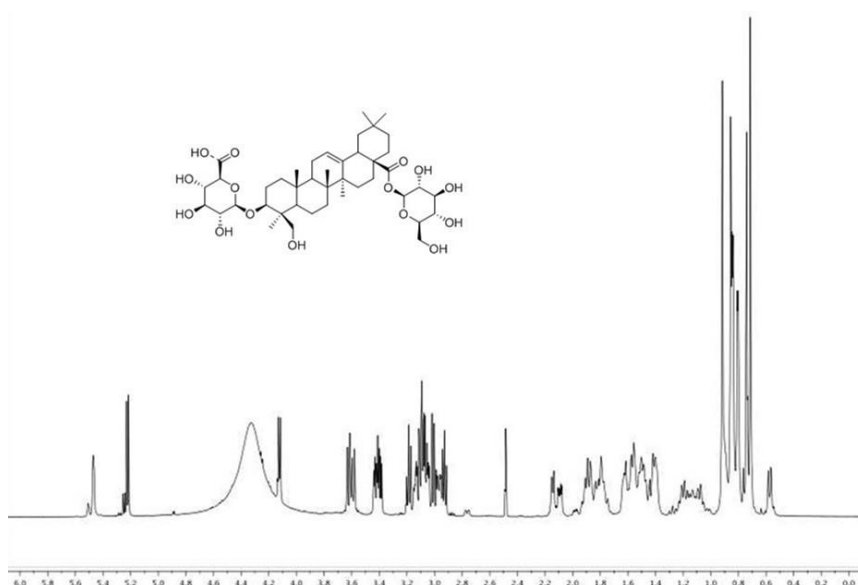


Figure 5.2.4.1.7 Chemical structure and ¹H NMR spectrum of **52**

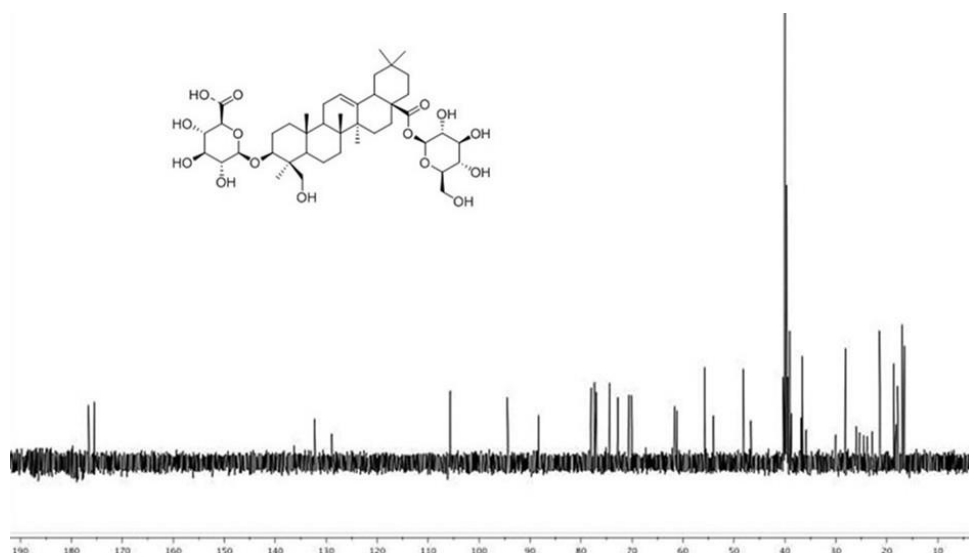


Figure 5.2.4.1.8 Chemical structure and ^{13}C NMR spectrum of **52**

Compound **52** was identified from the Australian plant *Larsenaikia ochreatea*. It was obtained as amorphous solid. Its molecular formula was determined to be $\text{C}_{42}\text{H}_{66}\text{O}_{15}$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+$ ion at m/z 811.4362}, which was consistent with 10 degrees of unsaturation. The ^1H NMR displayed sugar moieties and other signals (Figure 5.2.4.1.7). The ^{13}C NMR spectrum showed a total of 42 resonances (Figure 5.2.4.1.8). Its molecular formula and genus name were searched in the DNP, and 61 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **52** was determined as rivaloside E.¹⁵⁰

Additional compounds were also identified from the purification process. They included: compounds **27-28** and **30-32** from the Australian plant *Psydrax montigena*; compounds **34-38** from the Australian plant *Wendlandia basistaminea*; compounds **40-44** from the Australian plant *Psydrax lamprophylla*; compounds **46-50** from the Australian plant *Prunus persica*; **51** and **53-56** from the Australian plant *Larsenaikia ochreatea*.

Compound **27**

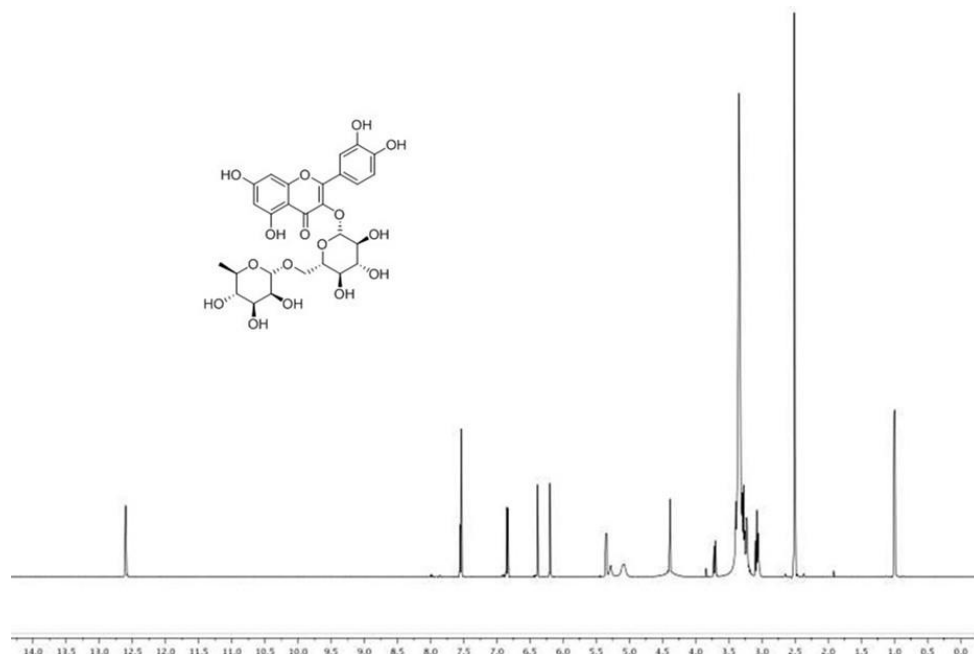


Figure 5.2.4.1.9 Chemical structure and ^1H NMR spectrum of **27**

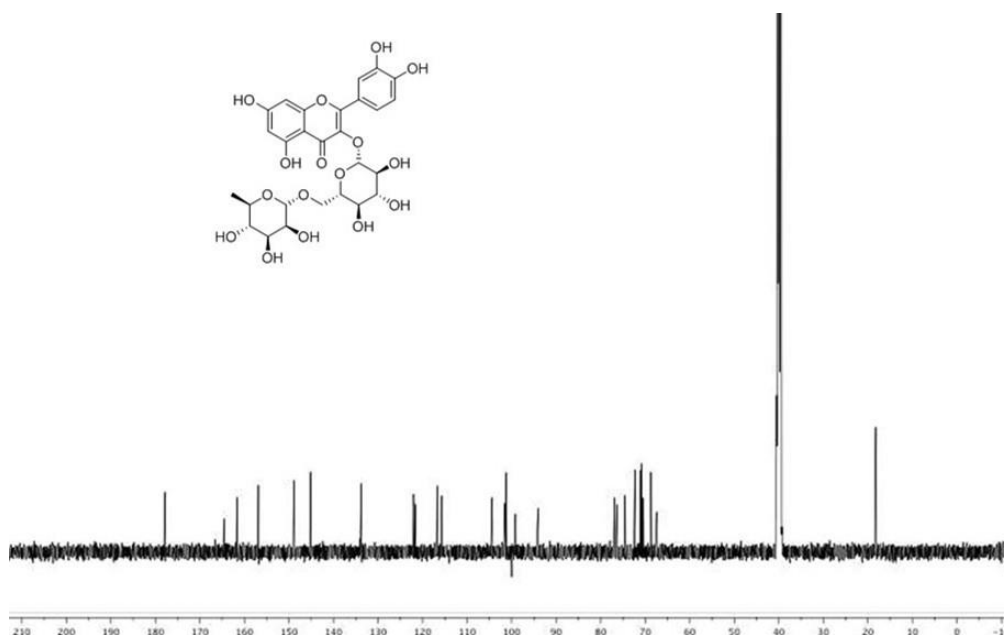


Figure 5.2.4.1.10 Chemical structure and ^{13}C NMR spectrum of **27**

Compound **27** was obtained as yellow powder with an $[\alpha]_{\text{D}}^{26} -5.1^\circ$ ($c=0.5$, CHCl_3). Its molecular formula was determined to be $\text{C}_{27}\text{H}_{30}\text{O}_{16}$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+$ ion at m/z 611.2643}, which was consistent with 13 degrees of unsaturation. ^1H NMR spectrum (Figure 5.2.4.1.9) displayed one exchangeable proton singlet (δ_{H} 12.60), five aromatics [7.56 (1H, d, $J=2.1$ Hz), 7.54 (1H, d, $J=2.3$ Hz), 6.84 (1H, d, $J=8.1$ Hz), 6.39 (1H, d, $J=2.0$ Hz), 6.20 (1H, d, $J=2.0$ Hz)], two oxygenated

methines (δ_{H} 5.35, 4.39), and sugar moieties (δ_{H} 3.07-3.40). The ^{13}C NMR spectrum of **27** (Figure 5.2.4.1.10) showed a total of 27 resonances, including 12 sp^3 hybridized and 17 sp^2 hybridized carbons. Its molecular formula and genus name were searched in the DNP, and 150 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **27** was determined as rutin. Rutin has exhibited a variety of pharmacological properties, including antimicrobial, antifungal, and anti-allergic activity.¹⁵¹

Compound **28**

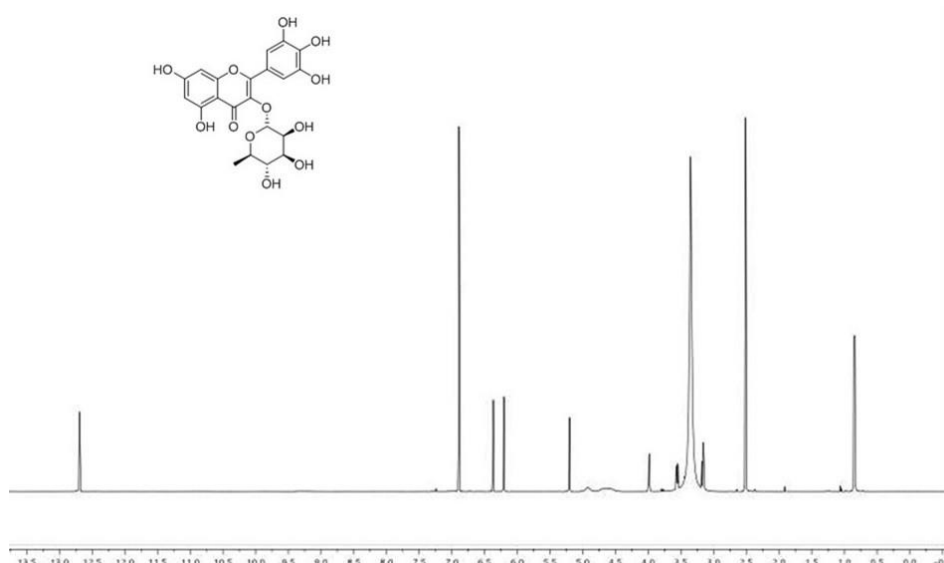


Figure 5.2.4.1.11 Chemical structure and ^1H NMR spectrum of **28**

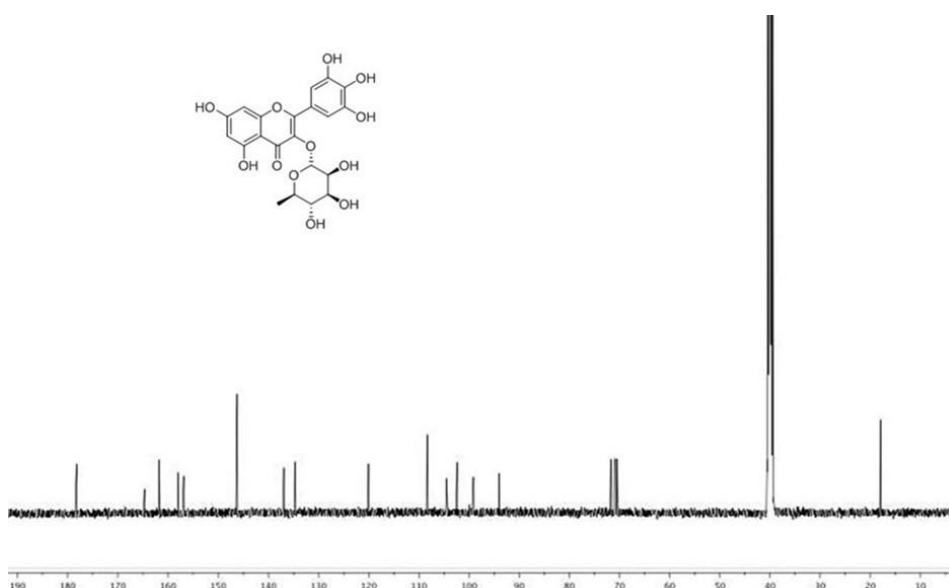


Figure 5.2.4.1.12 Chemical structure and ^{13}C NMR spectrum of **28**

Compound **28** was obtained as yellow powder with an $[\alpha]_D^{26} -126.3^\circ$ ($c=0.5$, MeOH). Its molecular formula was determined to be $C_{21}H_{20}O_{12}$ on the basis of (+)-HRESIMS measurements $\{[M+H]^+$ ion at m/z 465.0351 $\}$, which was consistent with 12 degrees of unsaturation. The 1H -NMR spectrum (Figure 5.2.4.1.11) of compound **28** exhibited a characteristic *meta*-coupled proton signal at δ_H 6.19 (1H, d, $J=2.0$ Hz) and 6.38 (1H, d, $J=2.0$ Hz) corresponding to a flavonoid A ring. Another AX coupling system was indicated by the presence of δ_H 7.36 (2H, brs) from B ring. The sugar moiety was suggested by the characteristic sugar anomeric proton signal at δ_H 5.33 (1H, d, $J=1.1$ Hz). The ^{13}C NMR spectrum of **28** (Figure 5.2.4.1.12) showed a total of 21 resonances, including 6 sp^3 hybridized and 15 sp^2 hybridized carbons. Its molecular formula and genus name were searched in the DNP, and 80 hits were found. By comparing their 1H and ^{13}C NMR data, compound **28** was determined as myricitrin.¹⁵² Chung *et al* examined the antioxidant activity of myricitrin and found out that it showed lower IC_{50} values (0.697 mM and 0.034 mM) than ascorbic acid.¹⁵³ In the study carried out by Winekenstdde *et al*, myricitrin showed strong inhibition of 5-lipoxygenase with an IC_{50} value of $7.8 \pm 0.2 \mu M$, suggesting good anti-inflammatory activity.¹⁵⁴ Qin *et al* demonstrated that myricitrin can inhibit oxidized low-density lipoprotein-induced endothelial apoptosis and prevent plaque formation at an early stage in an atherosclerotic mouse model, which results in reduced atherosclerotic plaque formation, indicating that myricitrin might be used as a drug candidate for the treatment of cardiovascular diseases.¹⁵⁵

Compound **30**

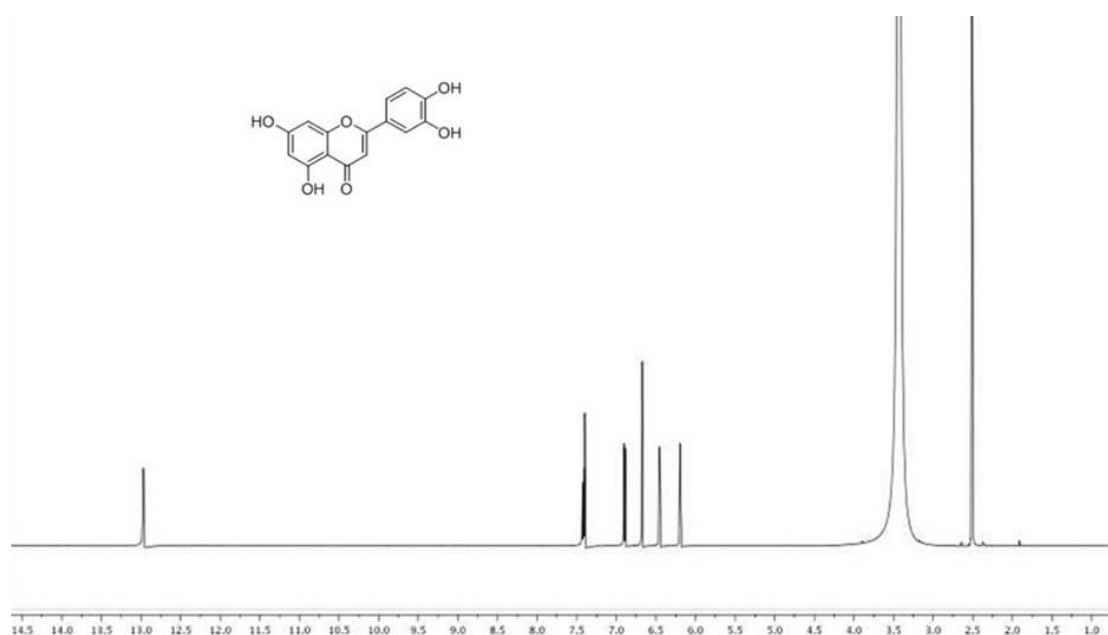


Figure 5.2.4.1.13 Chemical structure and ¹H NMR spectrum of **30**

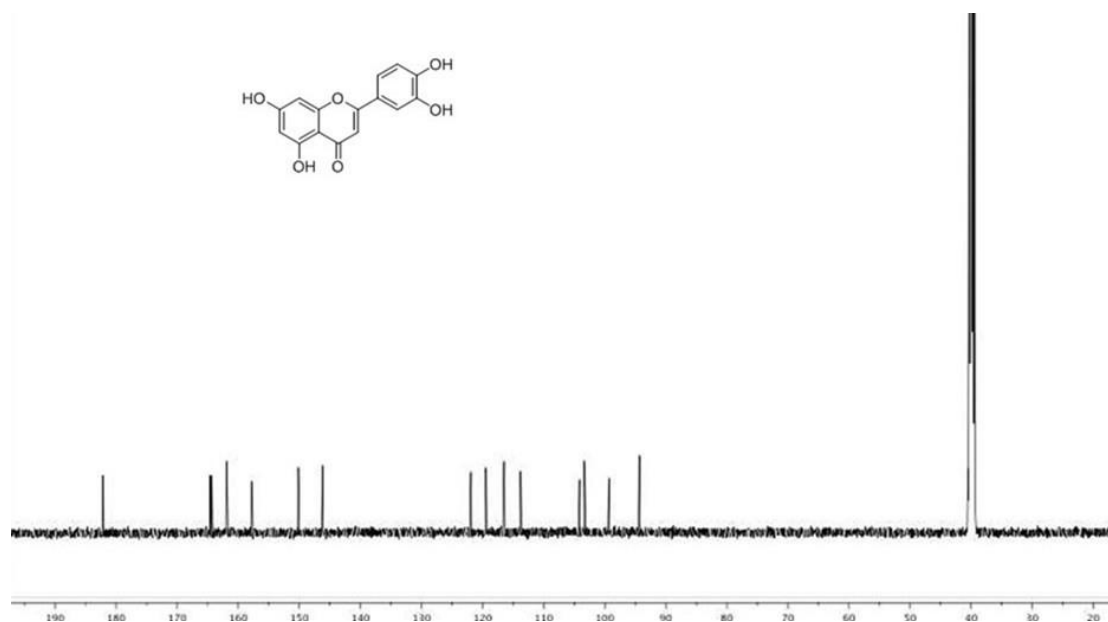


Figure 5.2.4.1.14 Chemical structure and ¹³C NMR spectrum of **30**

Compound **30** was obtained as yellow needles. Its molecular formula was determined to be C₁₅H₁₀O₆ on the basis of (+)-HRESIMS measurements {[M+H]⁺ ion at *m/z* 287.3108}, which was consistent with 11 degrees of unsaturation. The ¹H NMR spectrum of compound **30** (Figure 5.2.4.1.13) showed the presence of three *meta* coupled aromatic doublets at δ_H 6.28, 6.46 and 7.36, one *ortho* coupled aromatic doublet at δ_H 6.85, one doublet at δ_H 7.56 corresponding to an *ortho* and

meta coupled aromatic proton, and a singlet at δ_H 6.75. The ^{13}C NMR spectrum of **30** (Figure 5.2.4.1.14) showed a total of 15 sp^2 hybridized carbon resonances, six of which were oxygenated. Compound **30** was confirmed as a flavonoid from the positive characteristic chemical test with aluminum trichloride solution. Its molecular formula and genus name were searched in the DNP, and 152 hits were found. By comparing their 1H and ^{13}C NMR data, compound **30** was determined as luteolin.¹⁵⁶ Studies have shown that luteolin possessed a variety of pharmacological activities, including anti-inflammatory, antioxidant, anticancer, and antimicrobial activities.¹⁵⁷ In the study carried out by Miyazawa *et al*, luteolin showed antimutagenic activity with an IC_{50} value of 0.44 $\mu\text{mol/ml}$.¹⁵⁷

Compound **31**

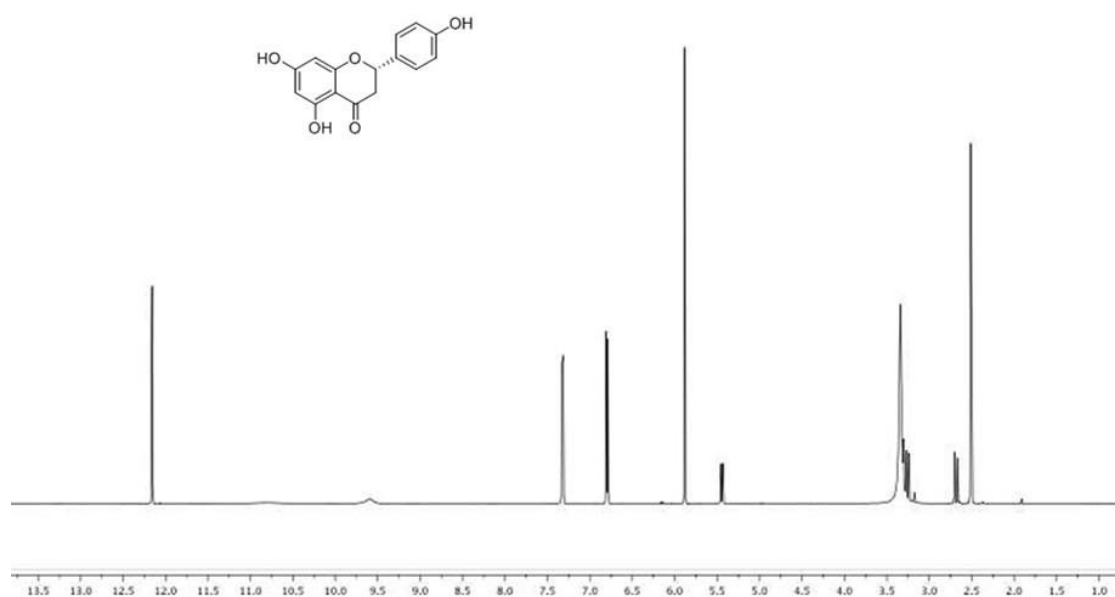


Figure 5.2.4.1.15 Chemical structure and 1H NMR spectrum of **31**

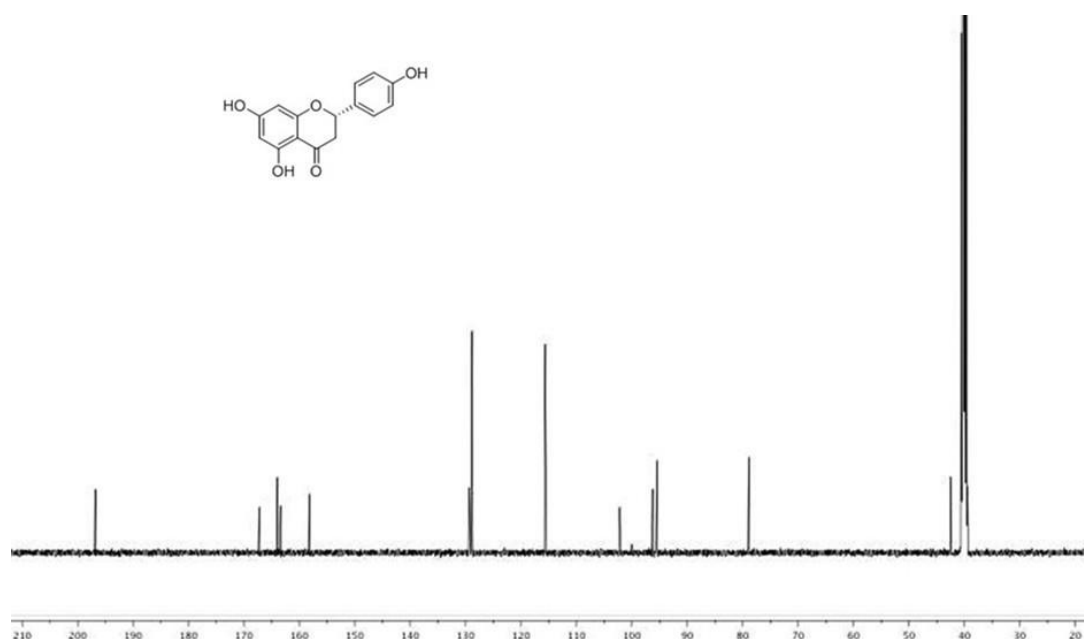


Figure 5.2.4.1.16 Chemical structure and ^{13}C NMR spectrum of **31**

Compound **31** was obtained as light brown powder with an $[\alpha]_D^{26} -20.1^\circ$ ($c=0.1$, MeOH). Its molecular formula was determined to be $\text{C}_{15}\text{H}_{12}\text{O}_5$ on the basis of (+)-HRESIMS measurements $\{[\text{M}-\text{H}]^+$ ion at m/z 273.3354}, which was consistent with 10 degrees of unsaturation. ^1H NMR spectrum of compound **31** (Figure 5.2.4.1.15) showed singlets at δ_{H} 12.18, 9.56 and 5.45 representing exchangeable protons. Aromatic peaks at δ_{H} 5.94 (1H, $J=2.0$ Hz) and 5.90 (1H, $J=1.8$ Hz) showed the pattern due to a *tetra*-substituted benzene ring. Aromatic doublets at δ_{H} 7.38 (2H, $J=8.6$ Hz) and δ_{H} 6.88 (2H, $J=8.3$ Hz) indicated another benzene ring. The ^{13}C NMR spectrum (Figure 5.2.4.1.16) showed the presence of one carboxyl at δ_{C} 196.8, one methylene at δ_{C} 42.4, one sp^3 hybridized methine at δ_{C} 78.9, and 12 sp^2 hybridized methines, four of which were oxygen-bearing. Its molecular formula and genus name were searched in the DNP, and 243 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **31** was determined as naringenin.¹⁵⁸ Naringenin showed effectiveness in the protection against oxidative damage to lipids in a dose-dependent manner and it was effective in reducing DNA damage.¹⁵⁹ Cytotoxic effects were induced in human cancer cell lines when high concentrations of naringenin were administered (50% effective concentration: 150-560 μM).¹⁶⁰

Compound **32**

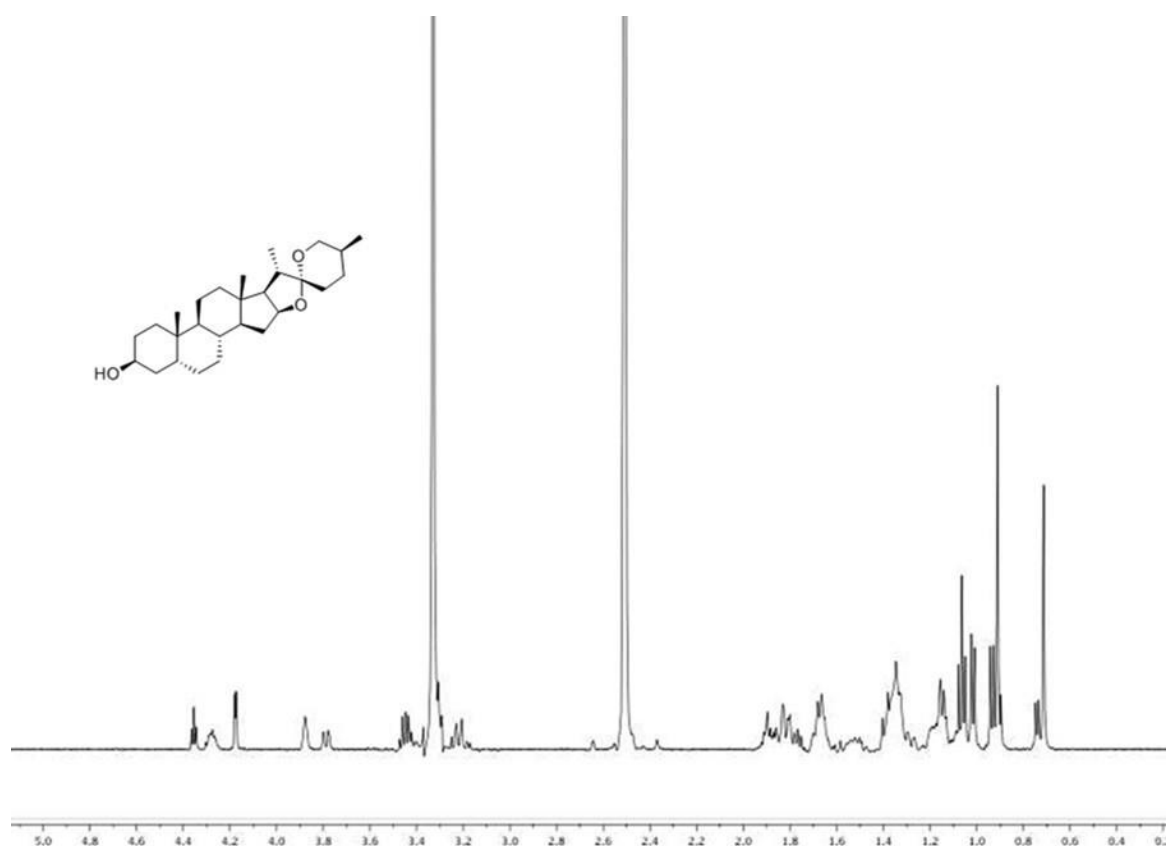


Figure 5.2.4.1.17 Chemical structure and ¹H NMR spectrum of **32**

Compound **32** was obtained as light brown powder with an $[\alpha]_D^{26} -81.1^\circ$ ($c=0.5$, CH₃Cl). Its molecular formula was determined to be C₂₇H₄₄O₃ on the basis of (+)-HRESIMS measurements {[M-3H]⁺ ion at m/z 413.4254}, which was consistent with 6 degrees of unsaturation. The ¹H NMR spectrum showed four methyls and other signals (Figure 5.2.4.1.17). The ¹³C NMR spectrum of **32** showed a total of 27 *sp*³ hybridized carbons, four of which were oxygenated. Compound **32** was confirmed as a steroidal aglycone from the positive characteristic chemical tests with Liebermann-Burchard and Molisch's reagents. Its molecular formula and genus name were searched in the DNP, and 150 hits were found. By comparing their ¹H and ¹³C NMR data, compound **32** was determined as sarsasapogenin.¹⁶¹ Sarsasapogenin derivatives displayed excellent selective cytotoxicity toward human cancer cell lines, and one of them exhibited significantly inhibitory activity against A375-S2 (IC₅₀=0.56 μM) and HT1080 (IC₅₀=0.72 μM) cells.¹⁶²

Compound **34**

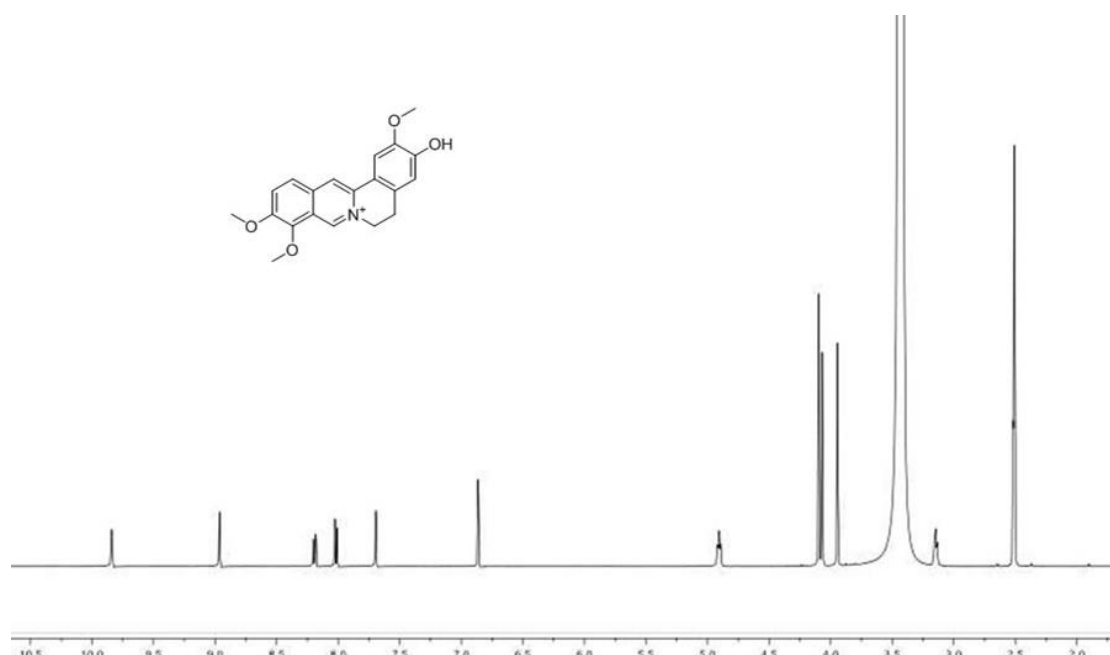


Figure 5.2.4.1.18 Chemical structure and ^1H NMR spectrum of **34**

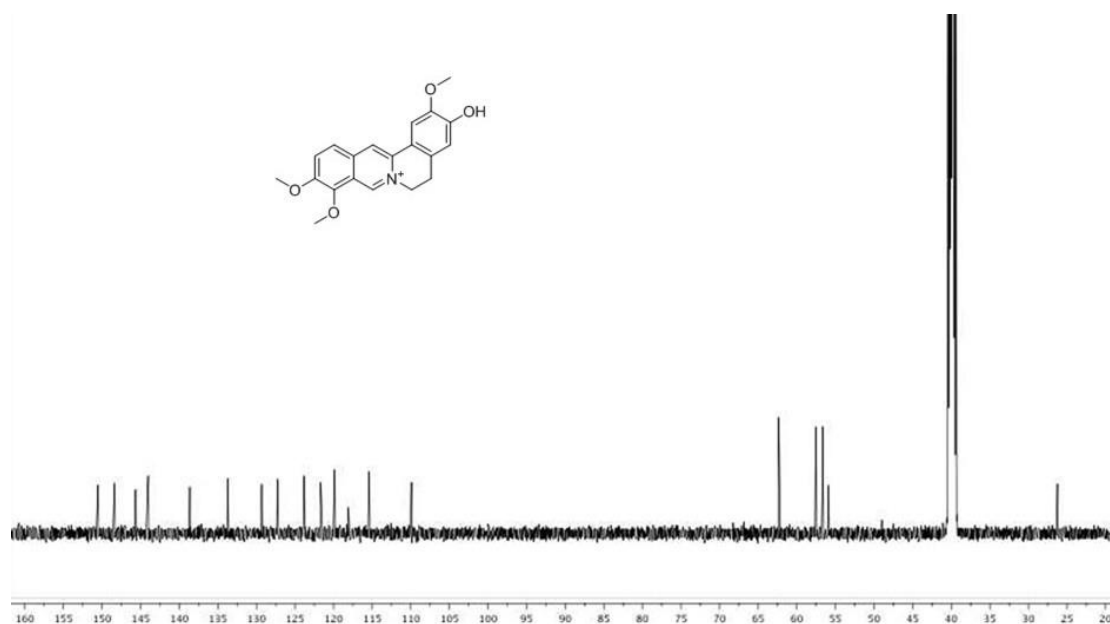


Figure 5.2.4.1.19 Chemical structure and ^{13}C NMR spectrum of **34**

Compound **34** was obtained as brown powder. Its molecular formula was determined to be $\text{C}_{20}\text{H}_{20}\text{NO}_4$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 338.5144, which was consistent with 12 degrees of unsaturation. ^1H NMR (Figure 5.2.4.1.18) displayed one exchangeable proton (δ_{H} 8.97), four aromatics [δ_{H} 8.19 (1H, dd, $J=2.3, 9.2$ Hz), 8.02 (1H, d, $J=9.2$ Hz), 7.70 (1H, s), 6.86 (1H, s)], and

other signals. The ^{13}C NMR spectrum of **34** (Figure 5.2.4.1.19) showed a total of 20 carbons, 10 of which were oxygen- or nitrogen-bearing. Compound **34** gave positive results with Dragendorff's reagents suggesting that it was an alkaloid. Its molecular formula and genus name were searched in the DNP, and 15 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **34** was determined as jatrorrhizine.¹⁶³ Jatrorrhizine was tested by Kong *et al* and showed anti-bacterial activity with IC_{50} value of 479.3 $\mu\text{g/mL}$.¹⁶⁴ It also showed strong antimicrobial activity against the 20 clinical isolates of *Propionibacterium acnes* with minimal inhibitory concentration values between 25 and 50 $\mu\text{g/mL}$.¹⁶⁵

Compound **35**

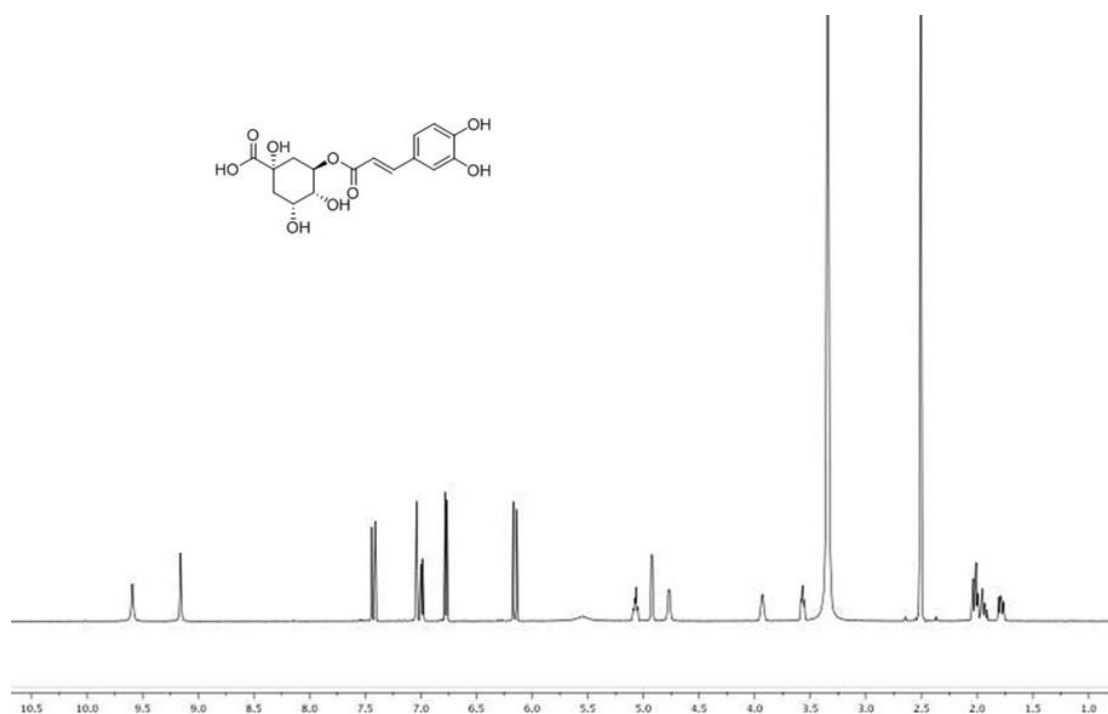


Figure 5.2.4.1.20 Chemical structure and ^1H NMR spectrum of **35**

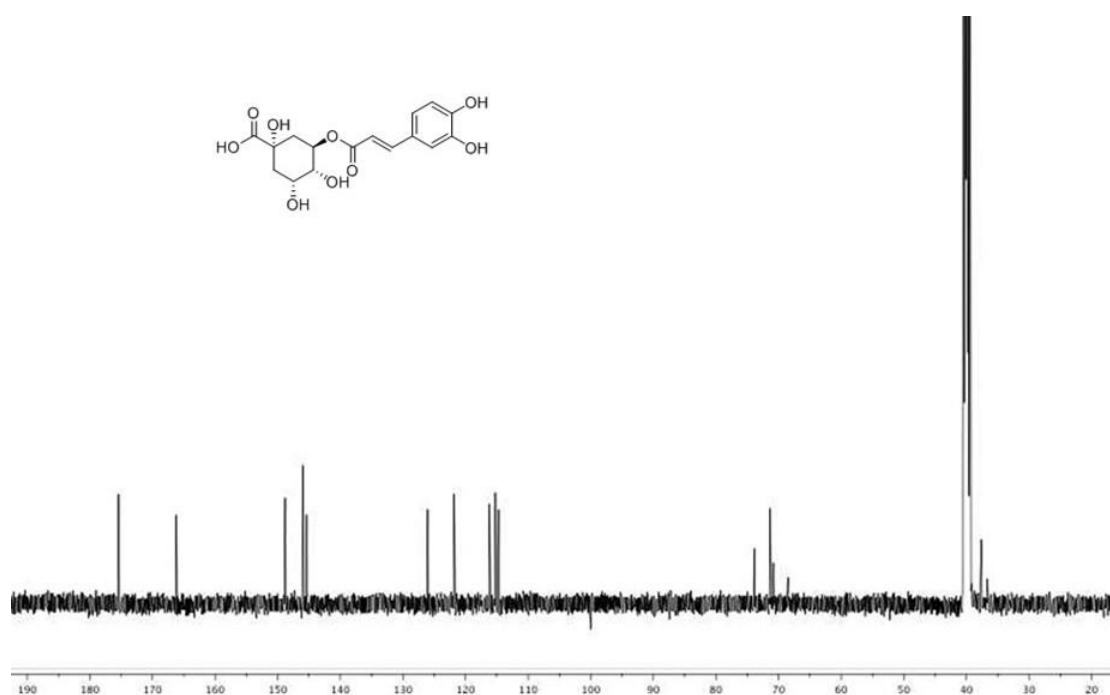


Figure 5.2.4.1.21 Chemical structure and ¹³C NMR spectrum of **35**

Compound **35** was obtained as colorless needles with an $[\alpha]_D^{26}$ -29.4° ($c=0.1$, MeOH). Its molecular formula was determined to be $C_{16}H_{18}O_9$ on the basis of (+)-HRESIMS measurements $\{[M-H]^+$ ion at m/z 355.3417}, which was consistent with 8 degrees of unsaturation. The ¹H NMR spectrum (Figure 5.2.4.1.20) exhibited signals in aliphatic and aromatic range. In the aromatic range, the ¹H NMR spectrum revealed an ABX system at δ_H 7.03 (1H, d, $J=2.0$ Hz), 6.91 (1H, dd, $J=8.2, 2.1$ Hz), 6.73 (1H, d, $J=8.0$ Hz) characteristic for 1, 3, 4-*tri*-substituted benzene. The ¹³C NMR spectrum of **35** showed a total of 20 resonances (Figure 5.2.4.1.21). Its molecular formula and genus name were searched in the DNP, and 28 hits were found. By comparing their ¹H and ¹³C NMR data, compound **35** was determined as chlorogenic acid.¹⁶⁶ Sato *et al* measured superoxide anion-scavenging activity of chlorogenic acid and its IC₅₀ value was 41.0 ± 12.1 μ M.¹⁶⁷ Chlorogenic acid showed *in vitro* antifungal activity against strains, such as *Candida albicans*, *Trichosporon beigeli* and *Malassezia furfur* with minimum inhibitory concentration values between 40 and 80 μ g/ml.¹⁶⁸ Study carried out by Guo *et al* indicated that chlorogenic acid could inhibit the inflammatory reaction in herpes simplex virus encephalitis via the suppression of

Toll-like receptor 2/Toll-like receptor 9 myeloid differentiation factor 88 signaling pathways.¹⁶⁹

Compound **36**

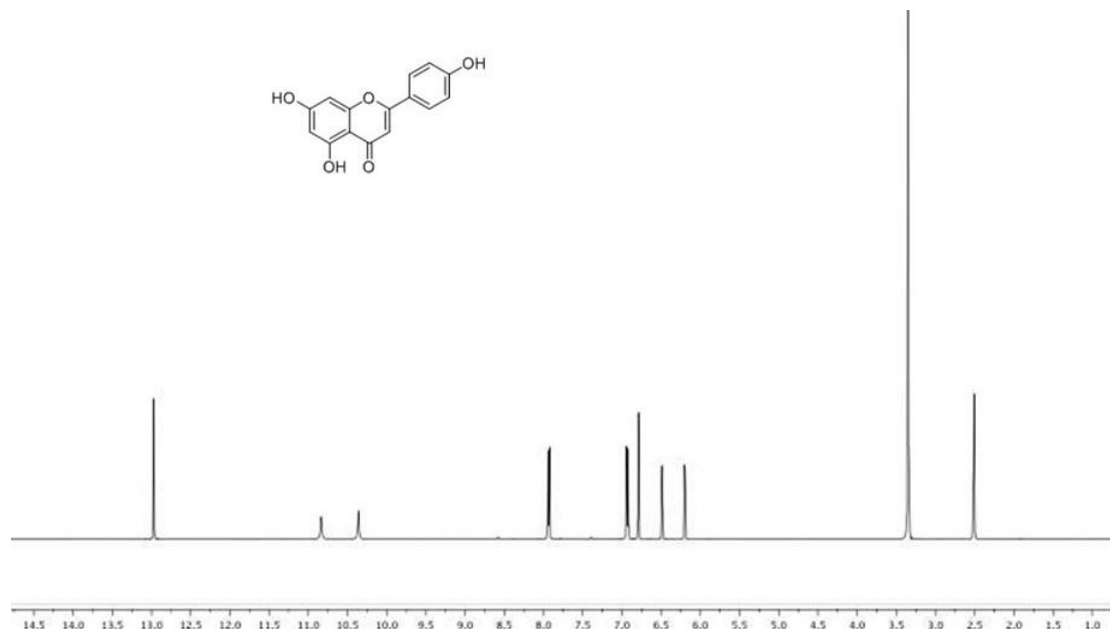


Figure 5.2.4.1.22 Chemical structure and ¹H NMR spectrum of **36**

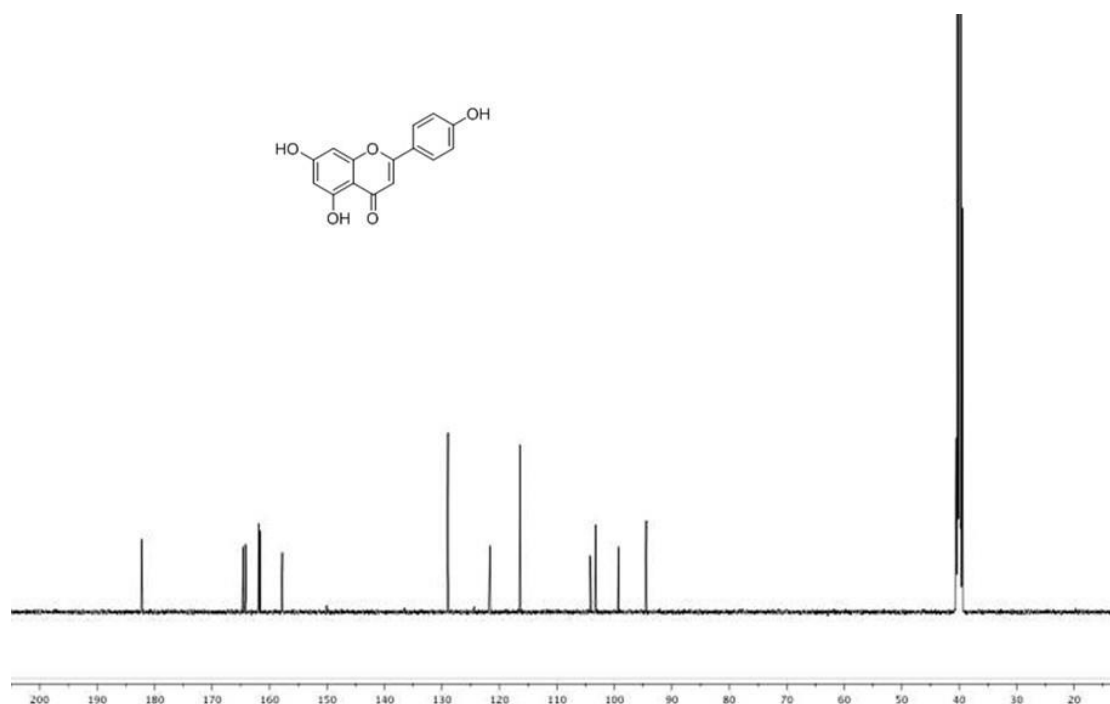


Figure 5.2.4.1.23 Chemical structure and ¹³C NMR spectrum of **36**

Compound **36** was obtained as yellow needles. Its molecular formula was determined to be C₁₅H₁₀O₅ on the basis of (+)-HRESIMS measurements {[M-H]⁺ ion at

m/z 269.3385}, which was consistent with 11 degrees of unsaturation. In the ^1H NMR spectrum (Figure 5.2.4.1.22), aromatic doublets at δ_{H} 7.68 (2H, $J=8.8$ Hz) and 6.92 (2H, $J=8.8$ Hz) indicated two aromatic protons, suggesting a *para* hydroxyl-substituted benzene ring. Singlets at δ_{H} 6.47 and 6.18 indicated an *ortho* hydroxyl-substituted benzene ring. ^{13}C NMR spectrum (Figure 5.2.4.1.23) gave presence of 15 carbons, six of which were oxygenated. Compound **36** gave positive result with aluminum trichloride solution, which confirmed it was a flavonoid. Its molecular weight and genus name were searched in the DNP, and 160 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **36** was determined as apigenin.¹⁵⁶ Apigenin was screened by Liu *et al* for its *in vitro* antibacterial activity against four bacteria (*Staphylococcus aureus*, *Bacillus subtilis*, *Escherichia coli* and *Pseudomonas aeruginosa*) and gave minimum inhibitory concentrations between 31.25 and 62.25 $\mu\text{g/ml}$.¹⁷⁰ Choudhury *et al* found out that apigenin induced apoptosis and cell death in lung epithelium cancer (A549) cells with an IC_{50} value of 93.7 ± 3.7 μM for 48 h treatment.¹⁷¹ In the research carried out by Miyazawa *et al*, apigenin suppressed the furylfuramide-induced SOS response in the *umu* test with an IC_{50} value of 0.55 $\mu\text{mol/ml}$, indicating its anti-mutagenic activity.¹⁵⁶

Compound **37**

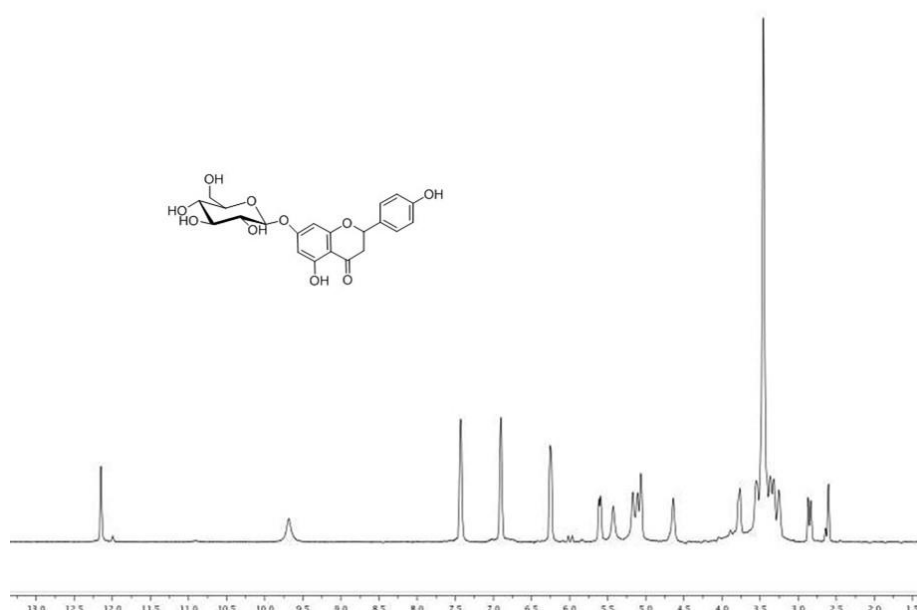


Figure 5.2.4.1.24 Chemical structure and ^1H NMR spectrum of **37**

Compound **37** was obtained as brown powder. Its molecular formula was determined to be $C_{21}H_{22}O_{10}$ on the basis of (+)-HRESIMS measurements $\{[M-H]^+$ ion at m/z 433.5624 $\}$, which was consistent with 11 degrees of unsaturation. The 1H NMR spectrum of compound **37** (Figure 5.2.4.1.24) was similar to naringenin (**31**), but with two more proton signals (δ_H 3.51 and 3.57) and peaks for a sugar moiety. The ^{13}C NMR spectrum of **37** showed a total of 21 resonances. Its molecular formula and genus name were searched in the DNP, and 79 hits were found. By comparing their 1H and ^{13}C NMR data, compound **35** was determined as naringenin 7-*O*-glucoside.¹⁷² Naringenin 7-*O*-glucoside has been tested by Rigano *et al* for its antibacterial activity against 12 bacteria strains and gave minimum inhibitory concentration between 32 and 256 $\mu g/ml$.¹⁷³

Compound **38**

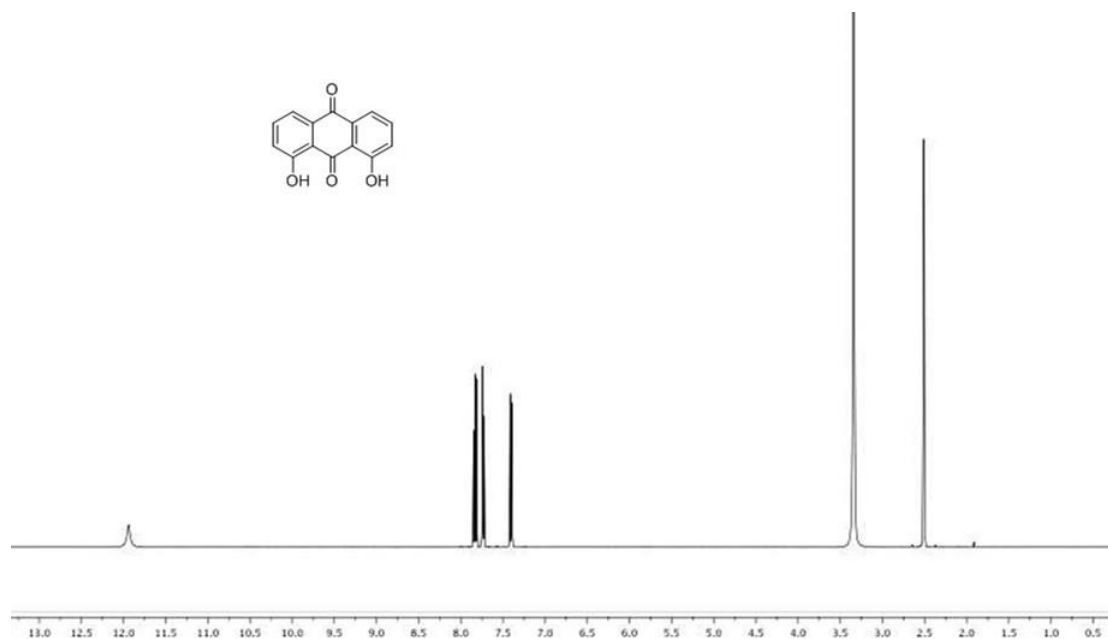


Figure 5.2.4.1.25 Chemical structure and 1H NMR spectrum of **38**

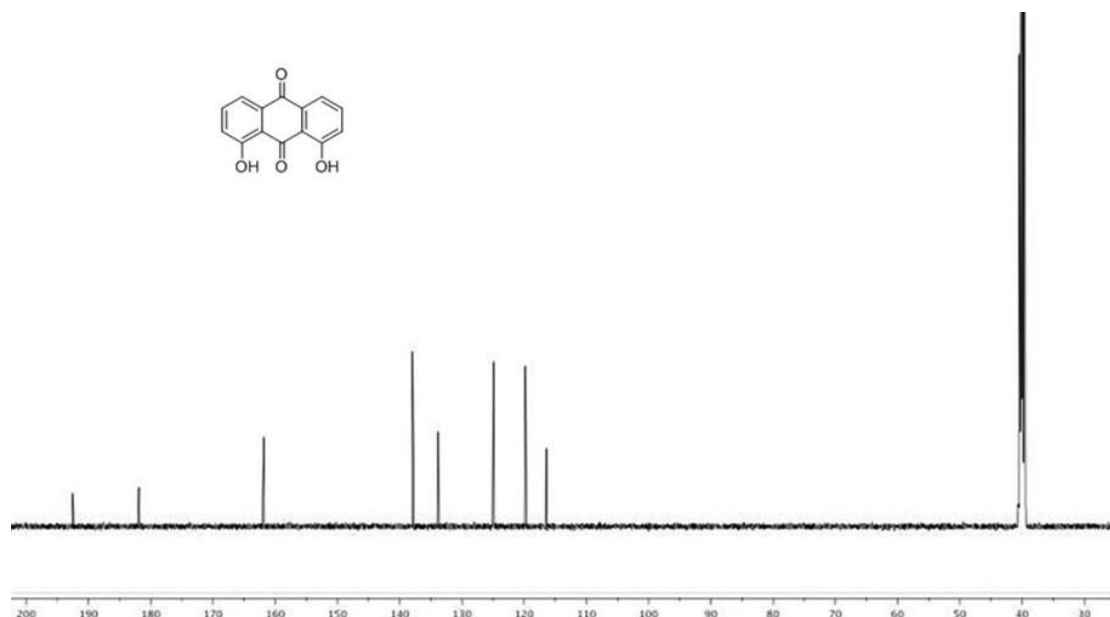


Figure 5.2.4.1.26 Chemical structure and ^{13}C NMR spectrum of **38**

Compound **38** was obtained as yellow needles. Its molecular formula was determined to be $\text{C}_{14}\text{H}_8\text{O}_4$ on the basis of (+)-HRESIMS measurements $\{[\text{M}-\text{H}]^+\}$ ion at m/z 239.2835, which was consistent with 11 degrees of unsaturation. In the ^1H NMR spectrum (Figure 5.2.4.1.25), δ_{H} 7.74 (2H, d, $J=7.7$ Hz), 7.65 (2H, dd, $J=7.7$ Hz, 8.2 Hz), 7.06 (2H, d, $J=8.2$ Hz) indicated the existence of benzene ring; δ_{H} 11.9 (2H, brs) suggested exchangeable protons. The ^{13}C NMR spectrum of **38** (Figure 5.2.4.1.26) displayed a total of 21 resonances, four of which were oxygenated. Compound **38** gave positive result with Borntrger agents, which confirmed that it was a flavonoid. Its molecular formula and genus name were searched in the DNP, and 14 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **38** was determined as 1,8-dihydroxyanthraquinone.¹⁷⁴ 1,8-dihydroxyanthraquinone can be administered to terminal cancer patients to alleviate the side effects resulting from morphine use.¹⁷⁵

Compound **40**

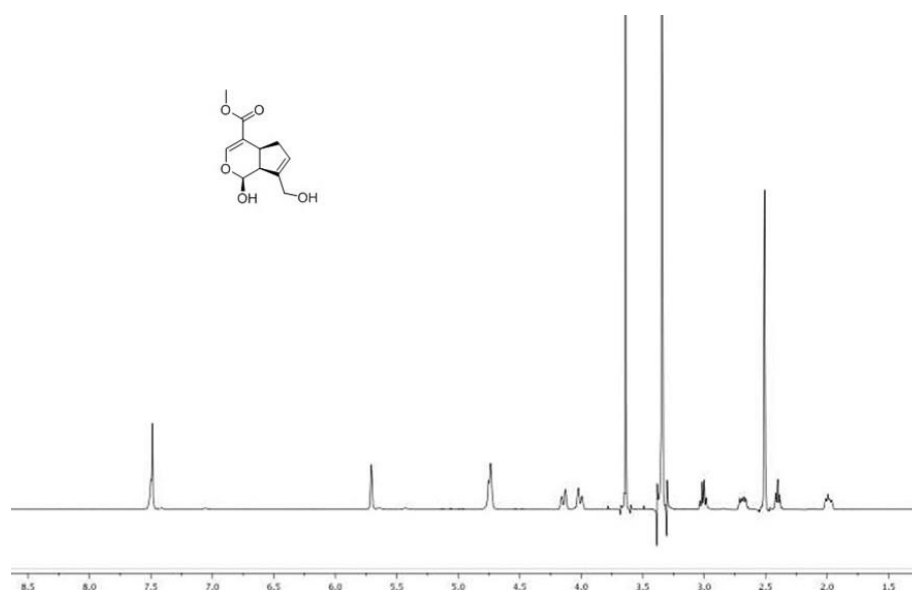


Figure 5.2.4.1.27 Chemical structure and ¹H NMR spectrum of **40**

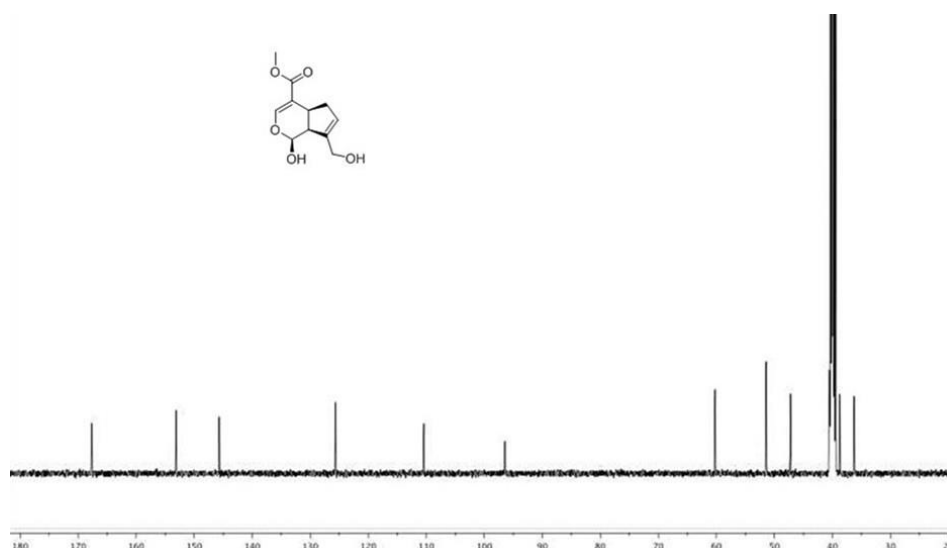


Figure 5.2.4.1.28 Chemical structure and ¹³C NMR spectrum of **40**

Compound **40** was obtained as brown powder. Its molecular formula was determined to be C₁₁H₁₄O₅ on the basis of (+)-HRESIMS measurements {[M+H]⁺ ion at *m/z* 227.2674}, which was consistent with 5 degrees of unsaturation. ¹H NMR spectrum was shown in Figure 5.2.4.1.27. The ¹³C NMR spectrum of **40** displayed a total of 11 resonances (Figure 5.2.4.1.28). Its molecular formula and genus name were searched in the DNP, and 95 hits were found. By comparing their ¹H and ¹³C NMR data, compound **40** was determined as genipin.¹⁷⁶ Genipin has been tested by

Chaipukdee *et al* against Human small cell lung cancer cell line (NCI-H187) with an IC_{50} value of 18.28 $\mu\text{g/ml}$.¹⁷⁷ Son *et al* determined that the 50% cytotoxicity dose (CD_{50}) of genipin was 70 μM , indicating that genipin caused a strong cytotoxicity in latently infected host cell Epstein-Barr virus.¹⁷⁸

Compound **41**

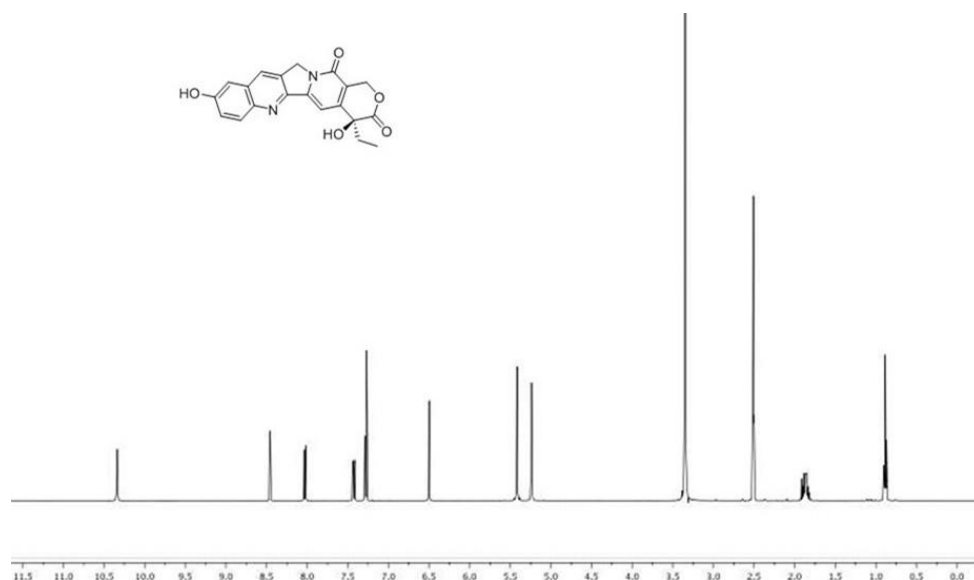


Figure 5.2.4.1.29 Chemical structure and ^1H NMR spectrum of **41**

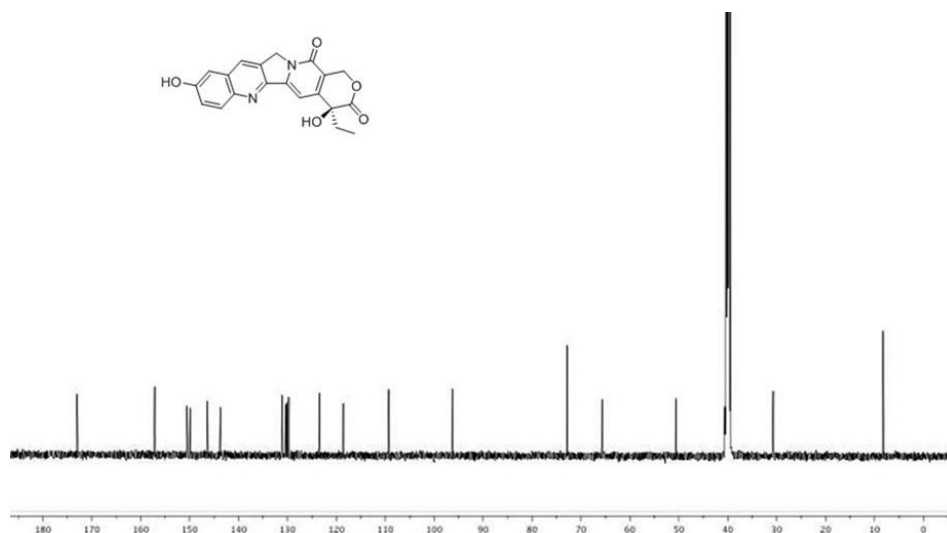


Figure 5.2.4.1.30 Chemical structure and ^{13}C NMR spectrum of **41**

Compound **41** was obtained as light yellow needles with an $[\alpha]_D^{26}$ -136.5° ($c=0.5$, CHCl_3). Its molecular formula was determined to be $\text{C}_{20}\text{H}_{16}\text{N}_2\text{O}_5$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+$ ion at m/z 365.2147}, which was consistent

with 12 degrees of unsaturation. The ^1H NMR spectrum (Figure 5.2.4.1.29) showed four aromatic signals at δ_{H} 8.45 (1H, s, H-7), 8.04 (1H, d, $J=9.3$ Hz, H-12), 7.28 (1H, d, $J=2.4$ Hz, H-9), 7.27 (1H, s, H-14), one methyl at δ_{H} 0.89 (3H), and one proton at δ_{H} 6.60 (1H, s). The ^{13}C NMR spectrum of **41** (Figure 5.2.4.1.30) displayed a total of 20 resonances, nine of which were oxygen- or nitrogen- bearing. Compound **41** was confirmed as an alkaloid by the positive results with Dragendorff's reagents. Its molecular formula and genus name were searched in the DNP, and six hits were found. By comparing their ^1H and ^{13}C NMR data, compound **41** was determined as 10-hydroxycamptothecin.¹⁷⁹ Research by Ping *et al* showed that a relatively low dose (5-20 nM) of 10-hydroxycamptothecin was able to inhibit the growth of human colon cancer.¹⁸⁰

Compound **42**

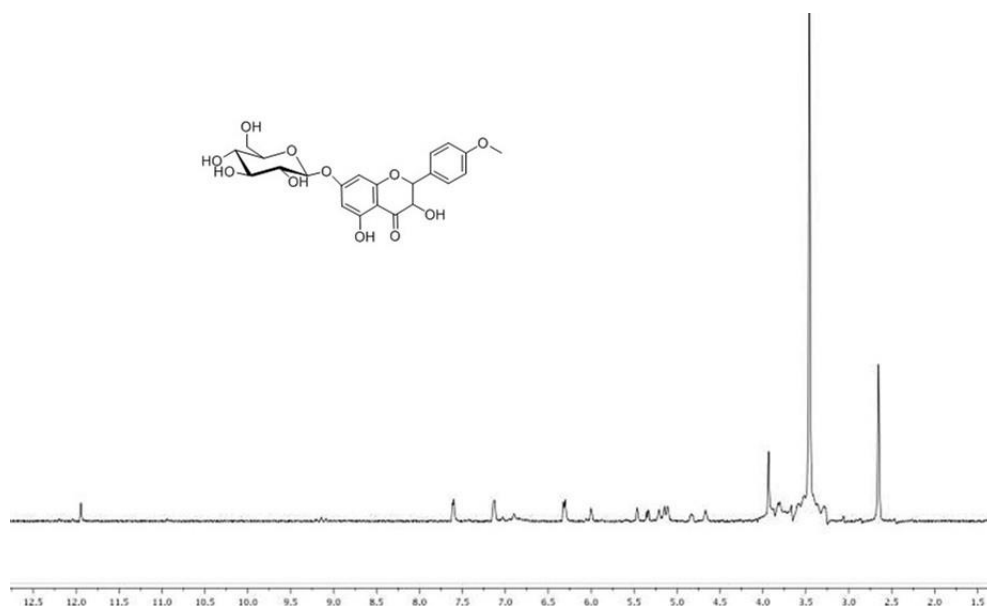


Figure 5.2.4.1.31 Chemical structure and ^1H NMR spectrum of **42**

Compound **42** was obtained as brown powder. Its molecular formula was determined to be $\text{C}_{21}\text{H}_{24}\text{O}_{11}$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 465.3716, which was consistent with 10 degrees of unsaturation. Compound **42** has similar NMR spectrum to naringenin 7-*O*-glucoside (**37**). ^1H NMR of compound **42** (Figure 5.2.4.1.31) showed one more methoxyl at δ_{H} 3.81 (3H, s). Another difference was the downfield shift of C-3 position from δ_{C} 43.1 in naringenin

7-*O*-glucoside (**37**) to δ_c 79.3 in compound **42**. The ^{13}C NMR spectrum of **42** displayed a total of 21 resonances. Its molecular formula and genus name were searched in the DNP, and 52 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **42** was determined as glucoferide.¹⁸¹⁻¹⁸²

Compound **43**

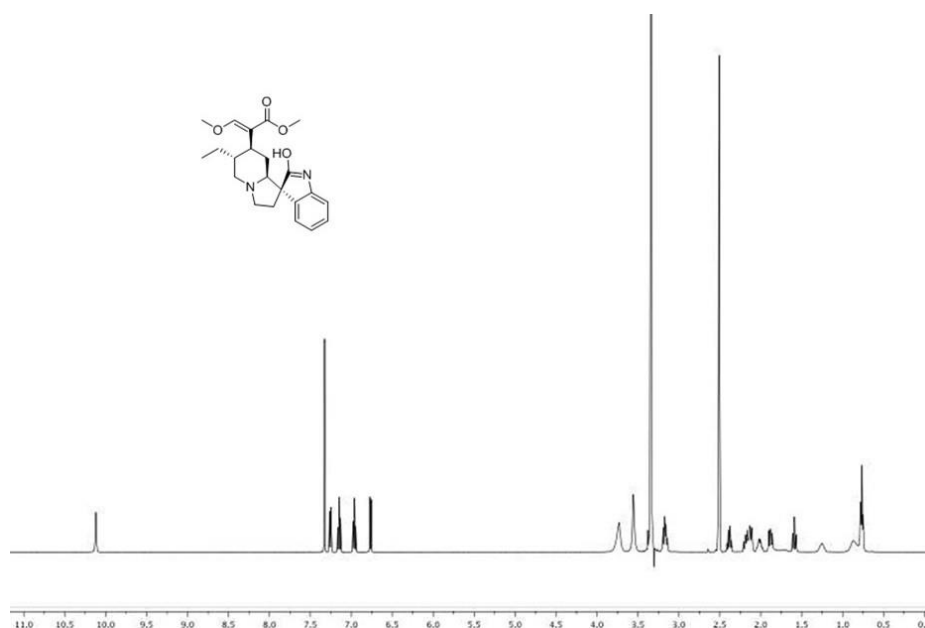


Figure 5.2.4.1.32 Chemical structure and ^1H NMR spectrum of **43**

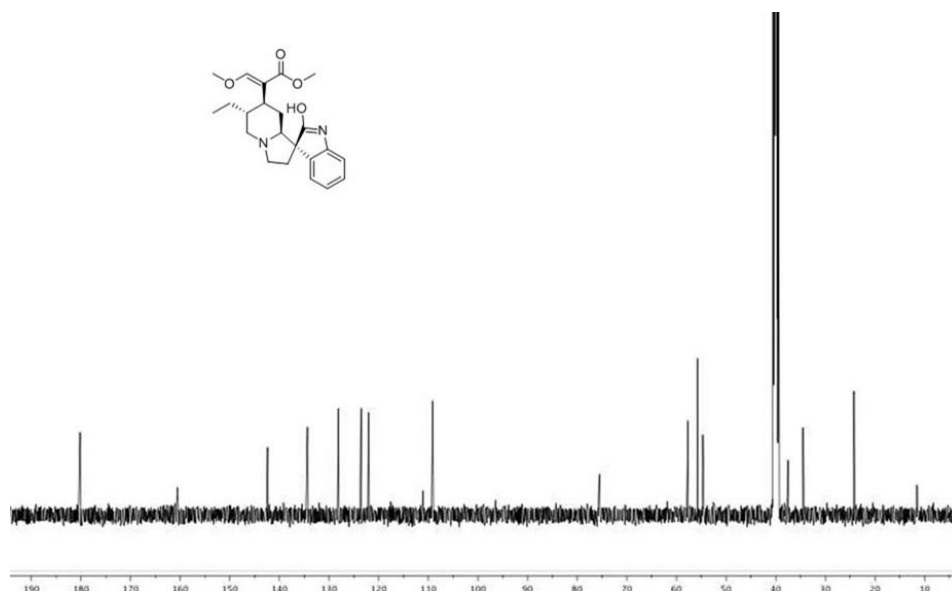


Figure 5.2.4.1.33 Chemical structure and ^{13}C NMR spectrum of **43**

Compound **43** was obtained as white needles with an $[\alpha]_D^{26}$ -18.2° ($c=2.5$, CHCl_3). Its molecular formula was determined to be $\text{C}_{22}\text{H}_{28}\text{N}_2\text{O}_4$ on the basis of (+)-HRESIMS

measurements $\{[M+H]^+$ ion at m/z 385.4185}, which was consistent with 10 degrees of unsaturation. In the ^1H NMR spectrum (Figure 5.2.4.1.32), δ_{H} 7.35 (1H, s), 7.25 (1H, d, $J=6.9$ Hz), 7.15 (1H, ddd, $J=7.7, 7.7, 1.1$ Hz), 7.05 (1H, ddd, $J=7.5, 7.5, 0.9$ Hz) indicated that there was an *ortho*-substituted benzene ring; δ_{H} 3.74 (3H, s) and δ_{H} 3.60 (3H, s) were signals from two methoxyls. The ^{13}C NMR spectrum of **43** displayed a total of 22 resonances (Figure 5.2.4.1.33). Compound **43** gave positive results with Dragendorff's reagents, which confirmed that it was an alkaloid. Its molecular formula and genus name were searched in the DNP, and 71 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **43** was determined as rhynchophylline.¹⁸³ Rhynchophylline can act on cardiovascular and central nervous system diseases, including hypertension and bradycardia.¹⁸⁴ It has shown effects on anticoagulation, inhibited vascular smooth muscle cell proliferation, and has shown to be anti-endotoxemic.¹⁸⁴

Compound **44**

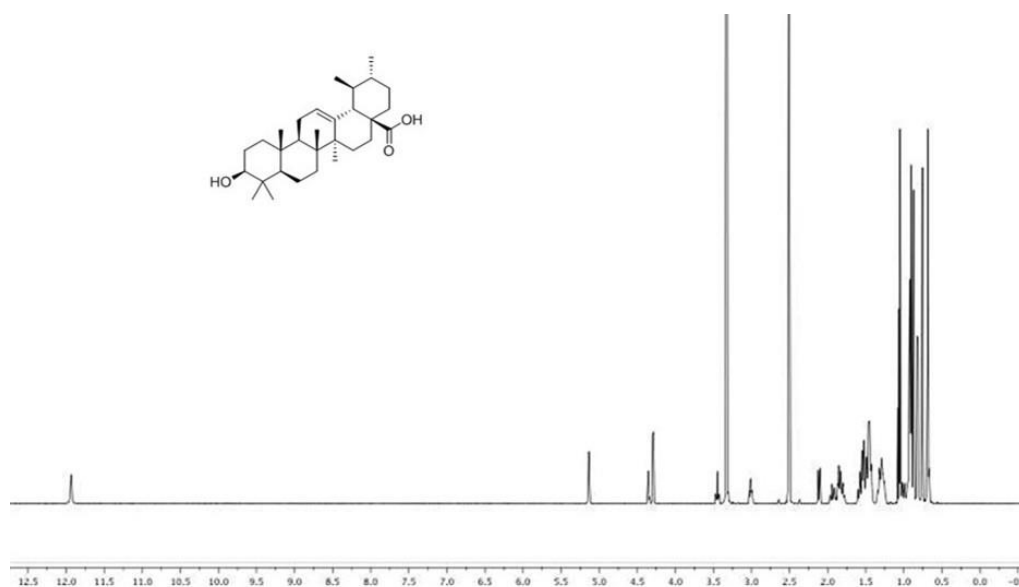


Figure 5.2.4.1.34 Chemical structure and ^1H NMR spectrum of **44**

Compound **44** was obtained as brown powder with an $[\alpha]_{\text{D}}^{26} +64.8^\circ$ ($c=0.3$, $\text{C}_5\text{H}_5\text{N}$). Its molecular formula was determined to be $\text{C}_{30}\text{H}_{48}\text{O}_3$ on the basis of (+)-HRESIMS measurements $\{[M+H]^+$ ion at m/z 455.4326}, which was consistent with 7 degrees of unsaturation. In the ^1H NMR spectrum (Figure 5.2.4.1.34), there were seven

methyl signals between δ_H 0.40 and 1.10 characteristic for pentacyclic triterpenoids, which was also confirmed by the positive result of adding 5% sulfuric acid to the solution of compound **44**. The ^{13}C NMR spectrum of **44** displayed a total of 30 resonances. By comparing its 1H and ^{13}C NMR data with references, compound **44** was determined as ursolic acid.¹⁸⁵⁻¹⁸⁶ In the study by Nascimento *et al*, ursolic acid has showed antimicrobial activity against *Staphylococcus aureus* with a minimal inhibitory concentration value of 32 $\mu g/ml$.¹⁸⁷ It has been effective against *Escherichia coli*, *Klebsiella pneumoniae* and *Shigella flexneri* with a minimal inhibitory concentration value of 64 $\mu g/ml$.¹⁸⁷ By inhibiting DPPH (2,2-diphenyl-1-picrylhydrazyl), ursolic acid strongly scavenged DPPH radical, with an IC_{50} value of $5.97 \times 10^{-2} \pm 1 \times 10^{-3} mg/ml$.¹⁸⁷

Compound **46**

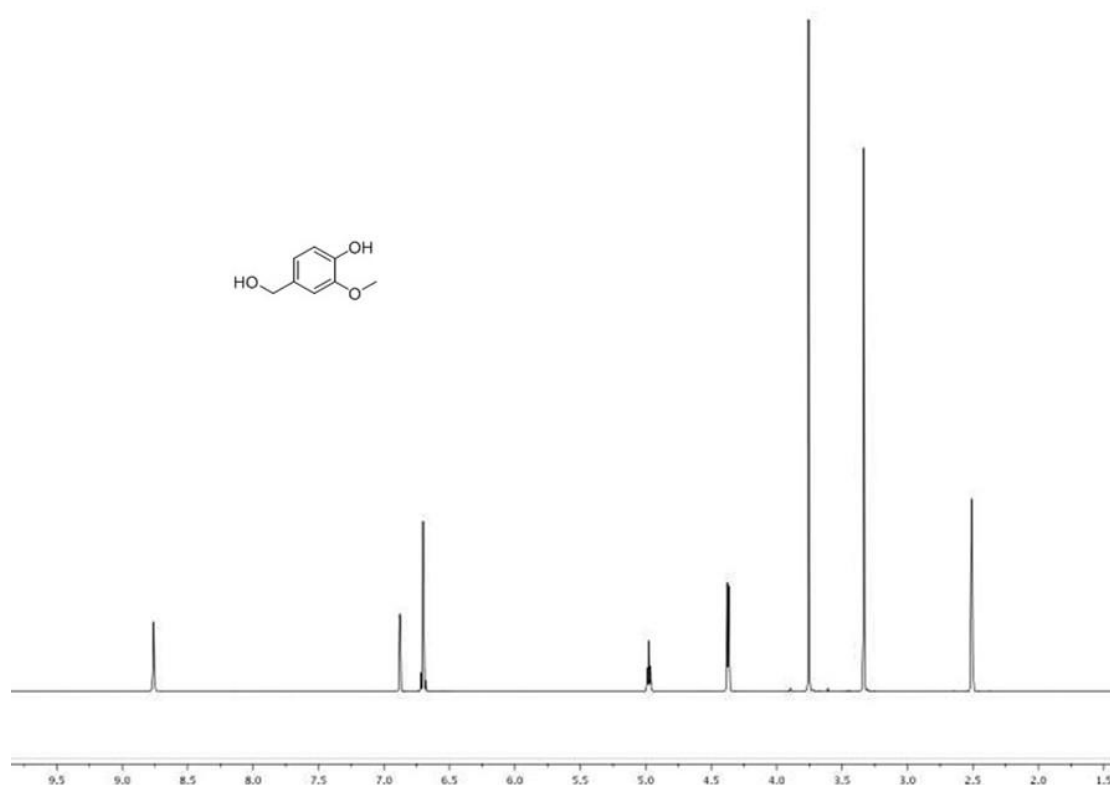


Figure 5.2.4.1.35 Chemical structure and 1H NMR spectrum of **46**

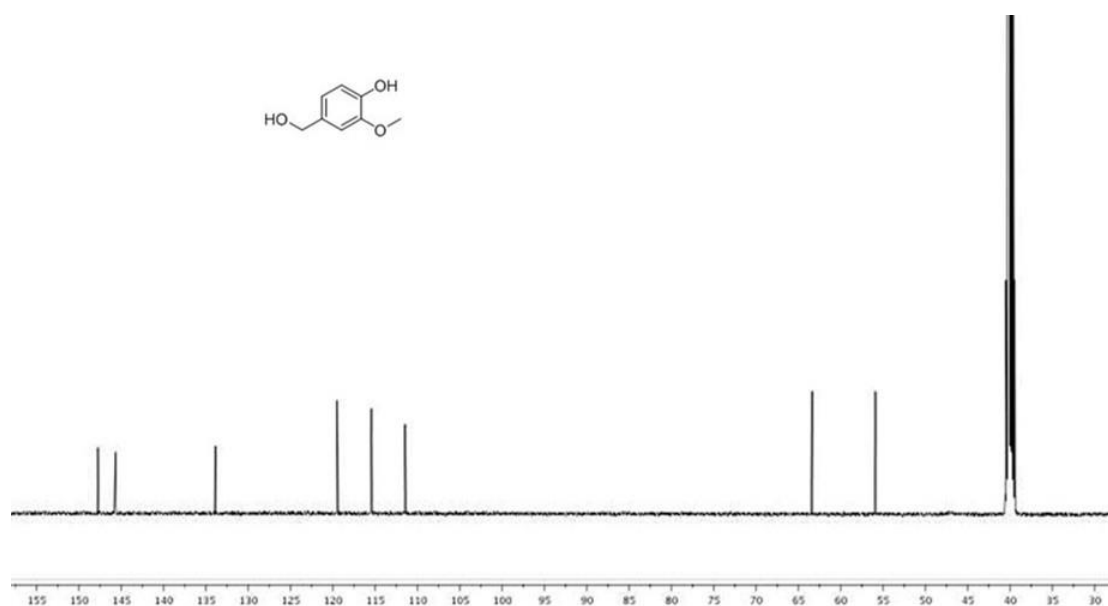


Figure 5.2.4.1.36 Chemical structure and ^{13}C NMR spectrum of **46**

Compound **46** was obtained as white needles. Its molecular formula was determined to be $\text{C}_8\text{H}_{10}\text{O}_3$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 154.1276, which was consistent with 4 degrees of unsaturation. In the ^1H NMR (Figure 5.2.4.1.35), δ_{H} 6.81 (1H, d, $J=7.4$ Hz, H-7), 6.79 (1H, d, $J=7.6$ Hz, H-4), 6.48 (1H, d, $J=7.4$ Hz, H-6) suggested a 1, 3, 4-*tri*-substituted benzene ring; δ_{H} 3.78 (3H, s) represented one methoxyl. The ^{13}C NMR spectrum of **46** displayed a total of 8 resonances (Figure 5.2.4.1.36). Its molecular formula and genus name were searched in the DNP, and 84 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **46** was determined as vanillyl alcohol.¹⁸⁸⁻¹⁸⁹ Vanillyl alcohol has been tested by Jung *et al* for anti-angiogenic activity through chick chorioallantoic membrane assay and its IC_{50} was determined to be $4.04\ \mu\text{g}/\text{egg}$.¹⁹⁰ Vanillyl alcohol at the oral doses of 3, 10, and 30 mg/kg showed an inhibition of 13.7%, 28.3%, and 34.2% in vascular permeability assay respectively, indicating that vanillyl alcohol had an acute anti-inflammatory activity.¹⁹⁰

Compound **47**

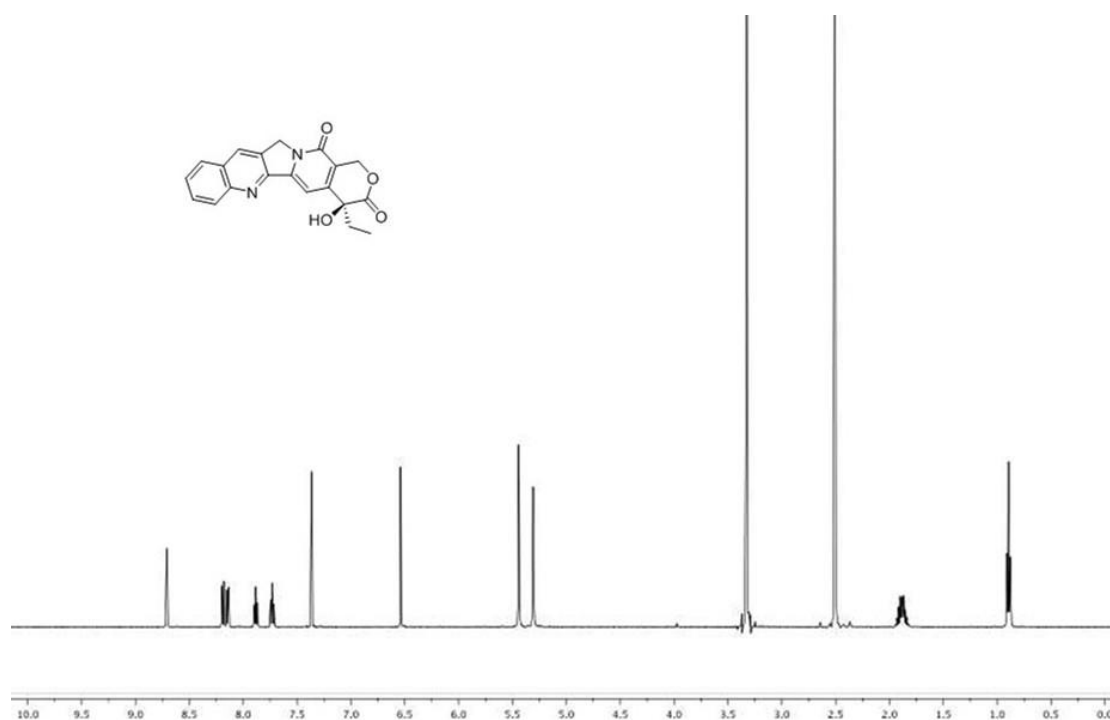


Figure 5.2.4.1.37 Chemical structure and ¹H NMR spectrum of **47**

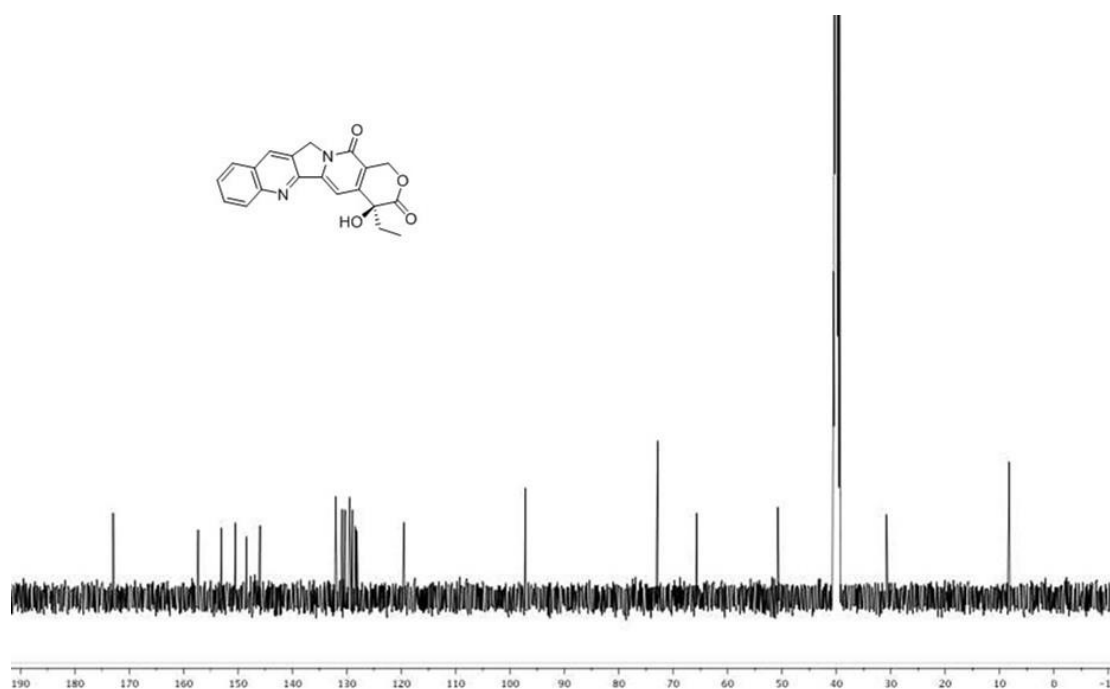


Figure 5.2.4.1.38 Chemical structure and ¹³C NMR spectrum of **47**

Compound **47** was obtained as light yellow crystal powder with an $[\alpha]_D^{26} +40.1^\circ$ ($c=0.1$, CHCl_3). Its molecular formula was determined to be $\text{C}_{20}\text{H}_{16}\text{N}_2\text{O}_4$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+$ ion at m/z 348.1439}, which was consistent

with 14 degrees of unsaturation. Compound **47** exhibited similar ^1H and ^{13}C NMR spectra (Figure 5.2.4.1.37 and Figure 5.2.4.1.38) with 10-hydroxycamptothecin (**41**). Its positive results with Dragendorff's reagents suggested that **47** was an alkaloid. Its molecular formula and genus name were searched in the DNP, and seven hits were found. By comparing their ^1H and ^{13}C NMR data, compound **47** was determined as camptothecin.¹⁹¹⁻¹⁹² Camptothecin has been reported as a strong antitumor agent and displayed nanomolar potency in cytotoxicity against many human tumor cell lines, such as HT29, LOX, SKOV3, and SKVLB, with IC_{50} values between 37 nM and 48 nM.¹⁹³ According to Luzzio *et al*, combined with TNF, camptothecin can induce apoptosis in primary mouse hepatocytes with an IC_{50} value of 13 μM .¹⁹³

Compound **48**

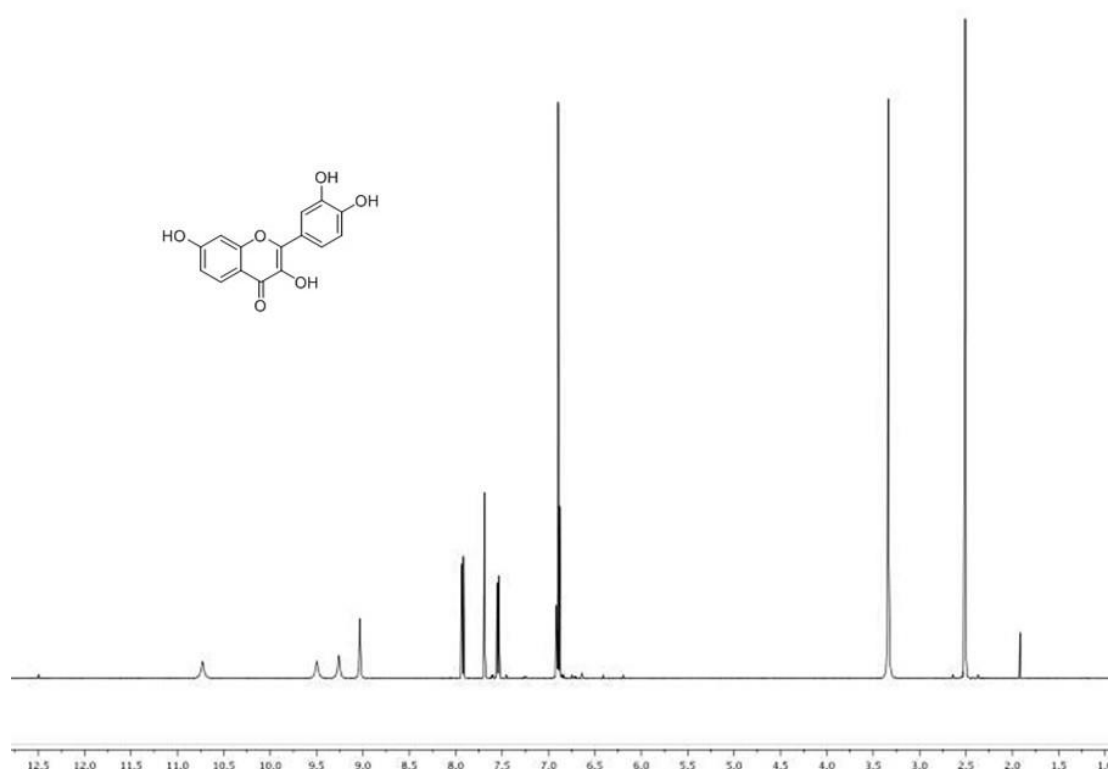


Figure 5.2.4.1.39 Chemical structure and ^1H NMR spectrum of **48**

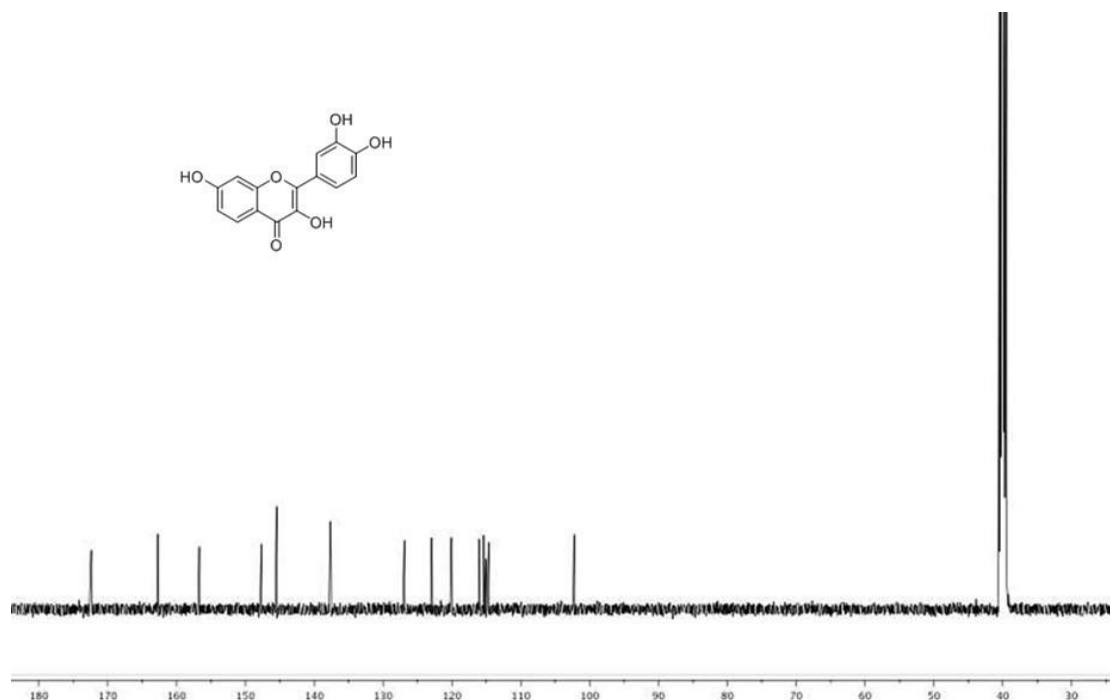


Figure 5.2.4.1.40 Chemical structure and ^{13}C NMR spectrum of **48**

Compound **48** was obtained as yellow needles. Its molecular formula was determined to be $\text{C}_{15}\text{H}_{10}\text{O}_6$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 285.3427, which was consistent with 11 degrees of unsaturation. The ^1H NMR spectrum (Figure 5.2.4.1.39) displayed four exchangeable proton singlets (δ_{H} 10.73, 9.50, 9.26, 9.03) and aromatic peaks. The ^{13}C NMR spectrum of **48** displayed a total of 15 resonances (Figure 5.2.4.1.40). The positive result with aluminum trichloride solution suggested that **48** was a flavone. Its molecular formula and genus name were searched in the DNP, and 152 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **48** was determined as fisetin.¹⁹⁴⁻¹⁹⁵ In the research by Zhuo *et al*, fisetin could inhibit cell viability of both A549 and A549-CR cells in a dose-dependent manner; IC_{50} of fisetin was 214.47 μM for A549 and 320.42 for A549-CR cells respectively, suggesting that fisetin could be a promising agent for lung cancer therapy.¹⁹⁶ The cytotoxic effect of fisetin was determined using MTT (3-(4,5-Dimethylthiazol-2-Yl)-2,5- Diphenyltetrazolium Bromide) method by Zandi *et al* and its 50% cytotoxicity (CC_{50}) value was 247 $\mu\text{g}/\text{mL}$.¹⁹⁷

Compound **49**

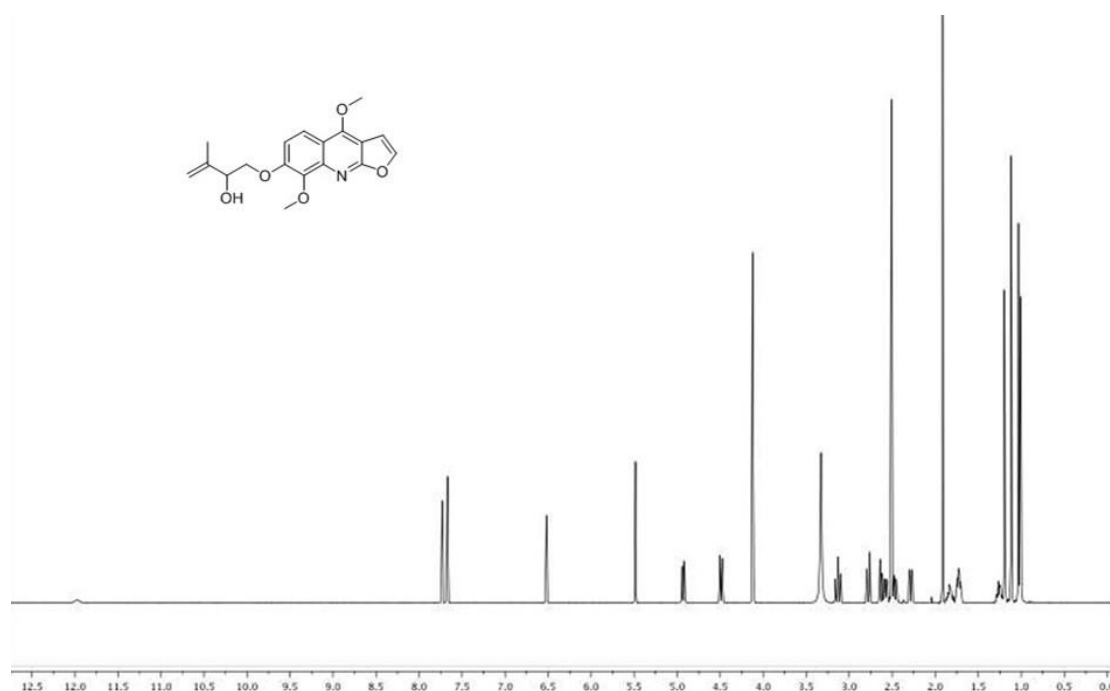


Figure 5.2.4.1.41 Chemical structure and ^1H NMR spectrum of **49**

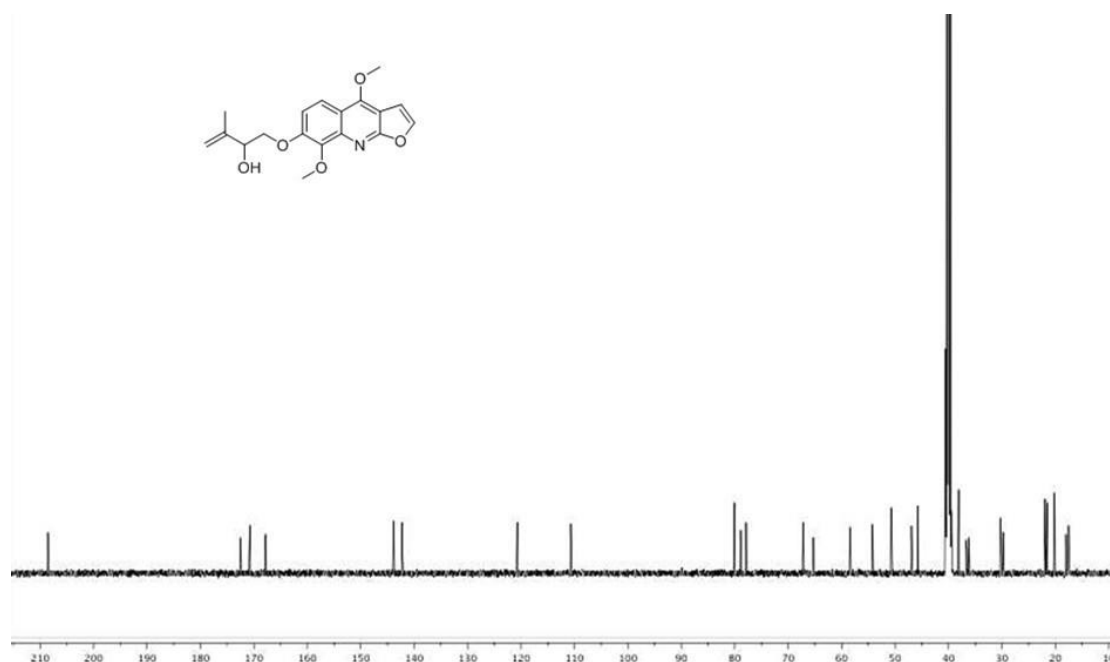


Figure 5.2.4.1.42 Chemical structure and ^{13}C NMR spectrum of **49**

Compound **49** was obtained as light brown powder. Its molecular formula was determined to be $\text{C}_{18}\text{H}_{19}\text{NO}_5$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+$ ion at m/z 329.8592}, which was consistent with 10 degrees of unsaturation. The ^1H NMR spectrum (Figure 5.2.4.1.41) displayed one exchangeable proton (δ_{H} 11.96), two methoxyls (δ_{H} 3.83, 4.16), one methyl, and other proton signals. The ^{13}C NMR spectrum of **49** (Figure 5.2.4.1.42) showed a total of 18 carbons, 10 of which were

oxygen- or nitrogen-bearing. Its molecular formula and genus name were searched in the DNP, and 58 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **49** was determined as evodine.¹⁹⁸⁻¹⁹⁹ Cai *et al* reported that the ethanol extract of *Evodia Rutaecarpa* exhibited better antinociceptive activity, which appeared to be related to the higher contents of evodine existed in the extract.¹⁹⁹

Compound **50**

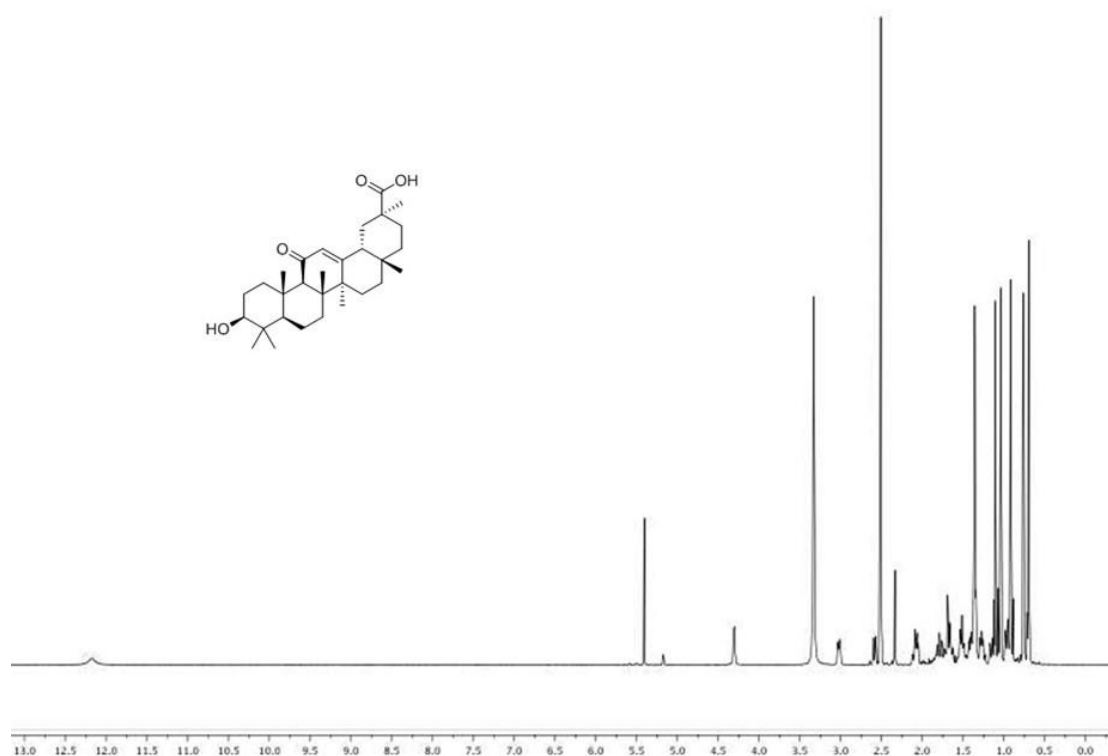


Figure 5.2.4.1.43 Chemical structure and ^1H NMR spectrum of **50**

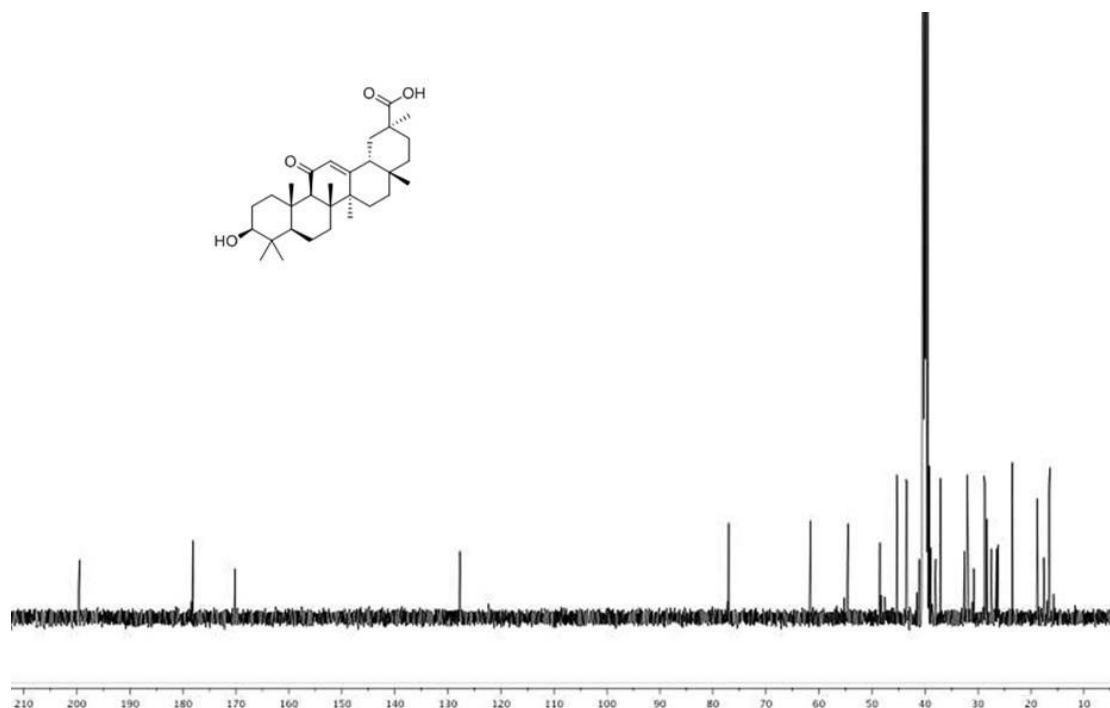


Figure 5.2.4.1.44 Chemical structure and ^{13}C NMR spectrum of **50**

Compound **50** was obtained as yellow crystal. Its molecular formula was determined to be $\text{C}_{30}\text{H}_{46}\text{O}_4$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 471.2406, which was consistent with 8 degrees of unsaturation. Compound **50** showed similar NMR spectra with ursolic acid (**44**). The ^1H NMR spectrum (Figure 5.2.4.1.43) showed seven methyls between δ_{H} 0.65 and 1.25 characteristic for pentacyclic triterpenoids, which was confirmed by the positive result of adding 5% sulfuric acid to the solution of compound **50**. The ^{13}C NMR spectrum of **50** showed a total of 30 resonances (Figure 5.2.4.1.44). There was one more keto group in **50**, indicated by the carbon downfield shift from δ_{C} 23.8 in ursolic acid (**44**) to δ_{C} 199.3 in **50** at C-11 position. The carbon chemical shift also suggested that the position of carboxyl group connected to C-17 position in ursolic acid (**44**), but connected to C-20 position in **50**. By comparing its ^1H and ^{13}C NMR data with references, compound **50** was determined as glycyrrhetic acid.²⁰⁰⁻²⁰¹ The anticancer activities of glycyrrhetic acid against MCF-7, MDA-MB-231 and hTERT-RPE1 were determined by *Li et al* through an *in vitro* MTT assay and the IC_{50} of glycyrrhetic acid were ranging from 63.41 to 84.70 μM .²⁰² Schwarz *et al* tested the antitumor activity of glycyrrhetic

acid through a sulforhodamine B assay using 15 different human tumor cell lines, and the IC₅₀ values were between 18.52 and 86.80 μ M.²⁰³

Compound **51**

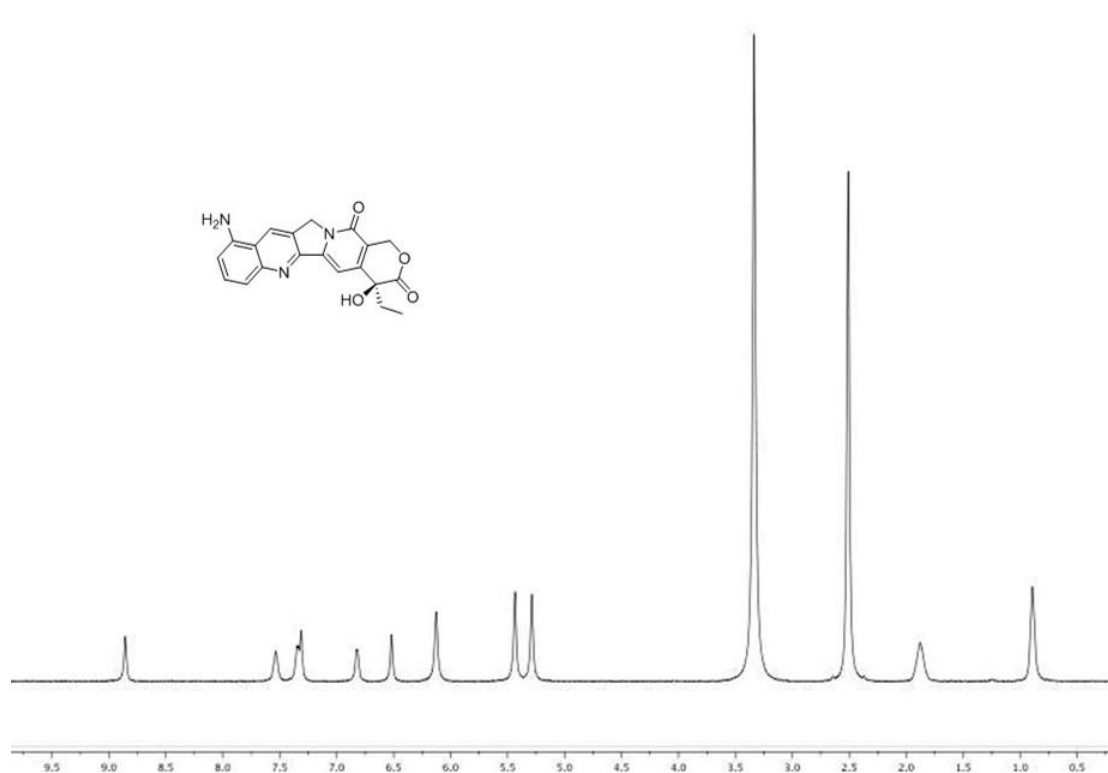


Figure 5.2.4.1.45 Chemical structure and ¹H NMR spectrum of **51**

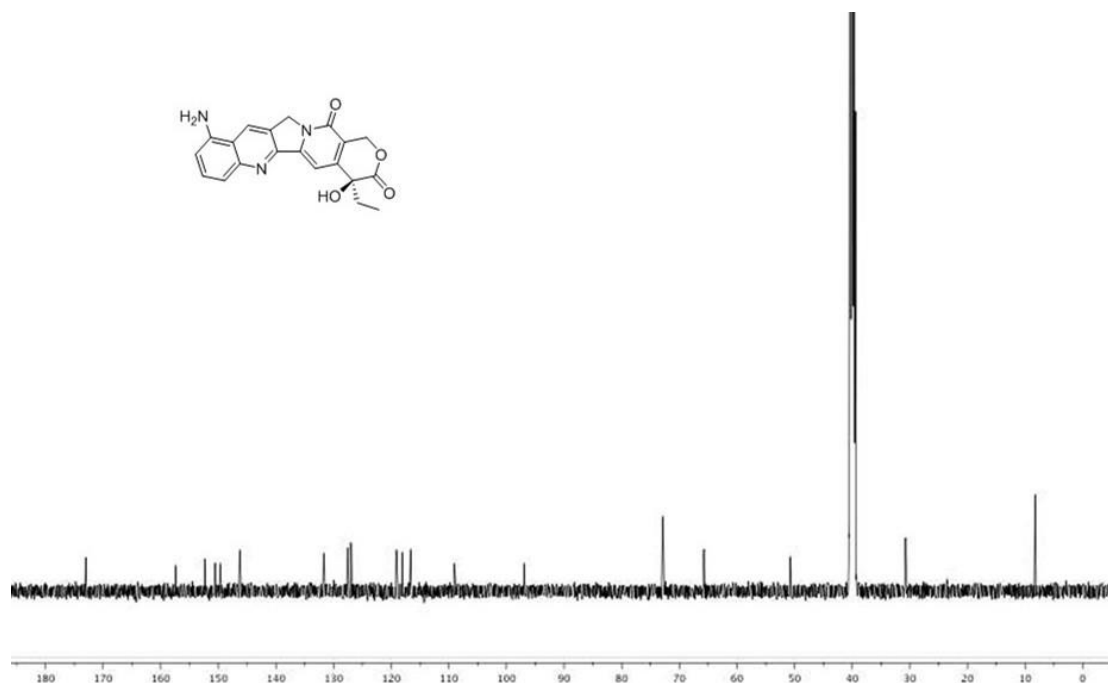


Figure 5.2.4.1.46 Chemical structure and ¹³C NMR spectrum of **51**

Compound **51** was obtained as white powder. Its molecular formula was determined to be $C_{20}H_{17}N_3O_4$ on the basis of (+)-HRESIMS measurements $\{[M+H]^+$ ion at m/z 363.7345}, which was consistent with 14 degrees of unsaturation. Compound **51** exhibited similar 1H and ^{13}C NMR spectra (Figure 5.2.4.1.45 and Figure 5.2.4.1.46) with 10-hydroxycamptothecin (**41**) and camptothecin (**47**). Its positive results with Dragendorff's reagents suggested that **51** was an alkaloid. Its molecular formula and genus name were searched in the DNP, and five hits were found. By comparing their 1H and ^{13}C NMR data, compound **51** was determined as 9-aminocamptothecin.²⁰⁴⁻²⁰⁵ According to the study by Naik *et al*, 9-aminocamptothecin showed inhibitory activity to the growth of prostate cancer in a dose-dependent manner across a 1 to 100 ng/ml concentration range.²⁰⁶

Compound **53**

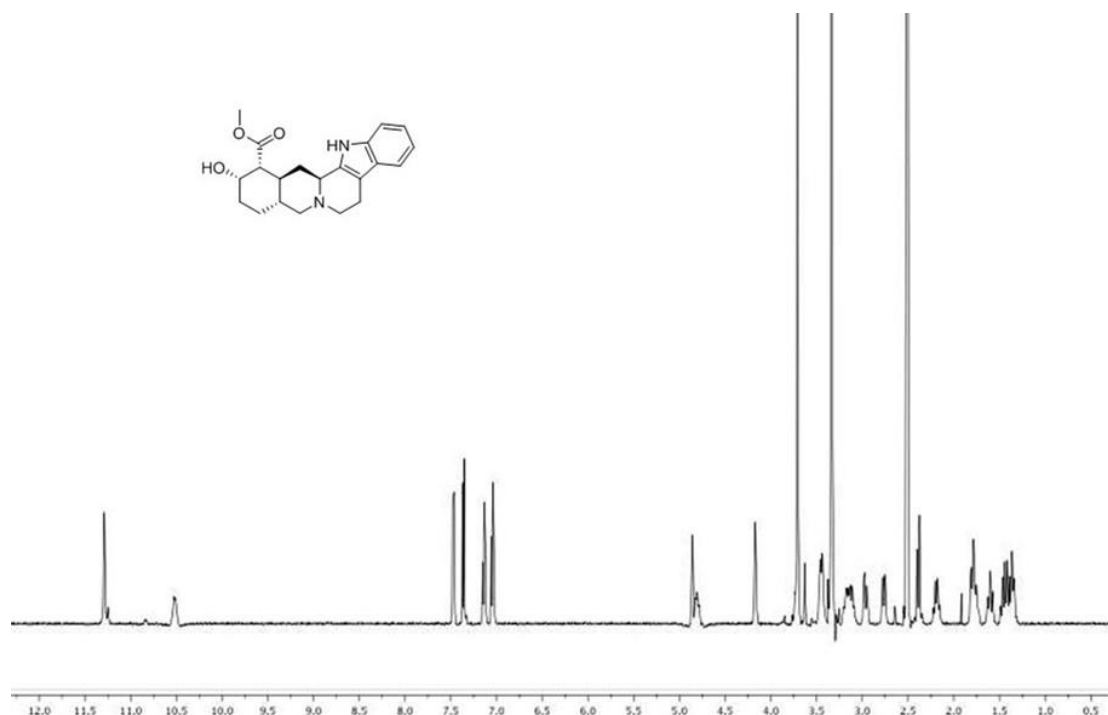


Figure 5.2.4.1.47 Chemical structure and 1H NMR spectrum of **53**

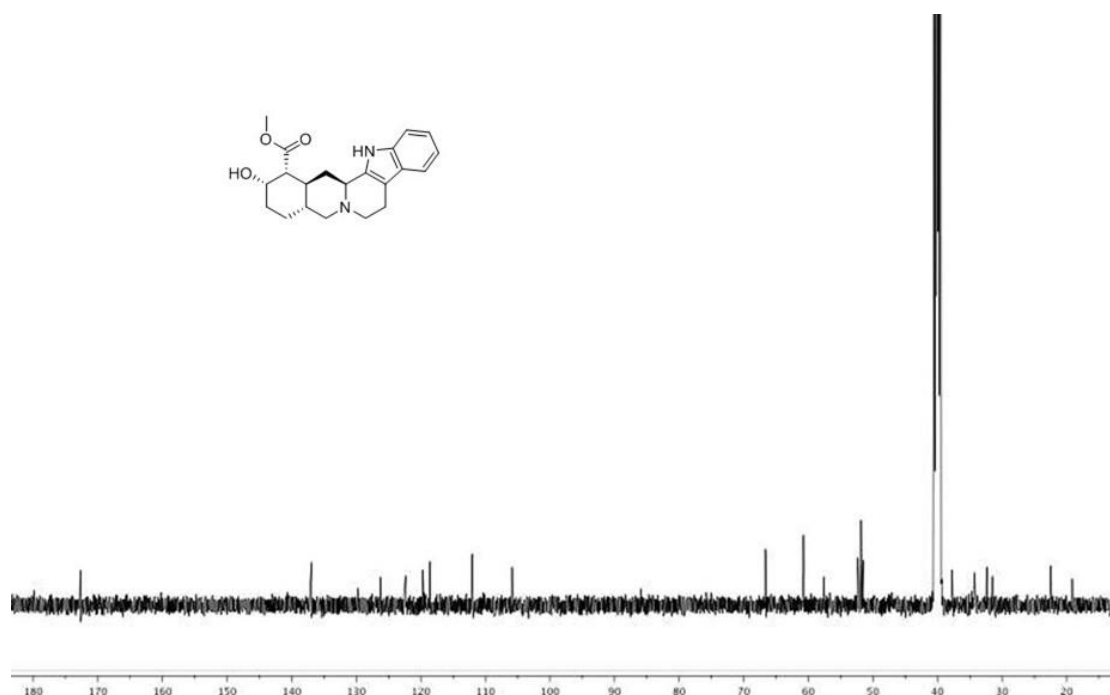


Figure 5.2.4.1.48 Chemical structure and ^{13}C NMR spectrum of **53**

Compound **53** was obtained as white powder. Its molecular formula was determined to be $\text{C}_{21}\text{H}_{26}\text{N}_2\text{O}_3$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 355.2905}, which was consistent with 10 degrees of unsaturation. Its ^1H and ^{13}C NMR were displayed in Figure 5.2.4.1.47 and Figure 5.2.4.1.48. Its molecular formula and genus name were searched in the DNP, and 109 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **53** was determined as yohimbine.²⁰⁷⁻²⁰⁸ Yohimbine has been studied as a potential treatment for erectile dysfunction and could be used as a drug lead for type 2 diabetes.²⁰⁹

Compound **54**

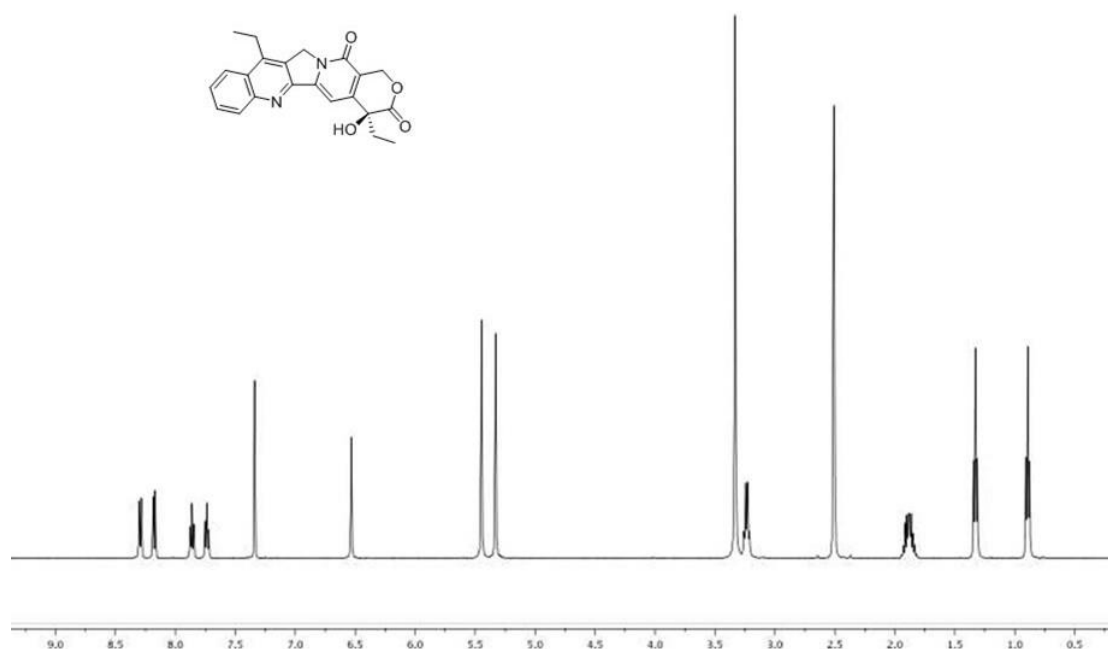


Figure 5.2.4.1.49 Chemical structure and ¹H NMR spectrum of **54**

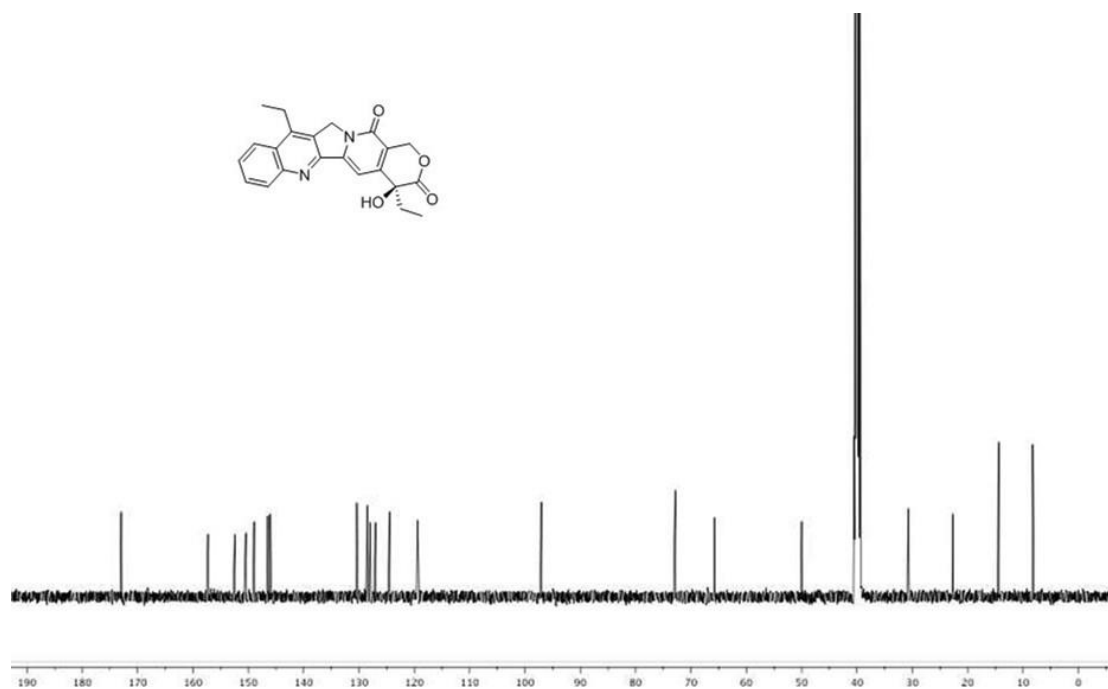


Figure 5.2.4.1.50 Chemical structure and ¹³C NMR spectrum of **54**

Compound **54** was obtained as yellow crystal powder. Its molecular formula was determined to be C₂₂H₂₀N₂O₄ on the basis of (+)-HRESIMS measurements {[M+H]⁺ ion at *m/z* 377.3264}, which was consistent with 14 degrees of unsaturation. Compound **54** has similar NMR spectra (Figure 5.2.4.1.49 and Figure 5.2.4.1.50) to camptothecin

(47). The only difference was that C-7 position in **54** was connected to an ethyl group, evidenced by one methyl signal at δ_H 1.33 (3H, t, $J=7.6$ Hz) and one ethyl signal at δ_H 3.24 (2H, q, $J=7.5$ Hz). The ^{13}C NMR spectrum of **54** showed a total of 22 carbons. By comparing its 1H and ^{13}C NMR data with references, compound **54** was determined as 7-ethylcamptothecin.²¹⁰ In the research by Nagata *et al*, 7-ethylcamptothecin had better antitumor activity than camptothecin.²¹¹ 7-ethylcamptothecin inhibited the growth of KB cells *in vitro*, 50% effective doses of 3.5 ng/ml, compared to 8.6 ng/ml of camptothecin.²¹¹

Compound **55**

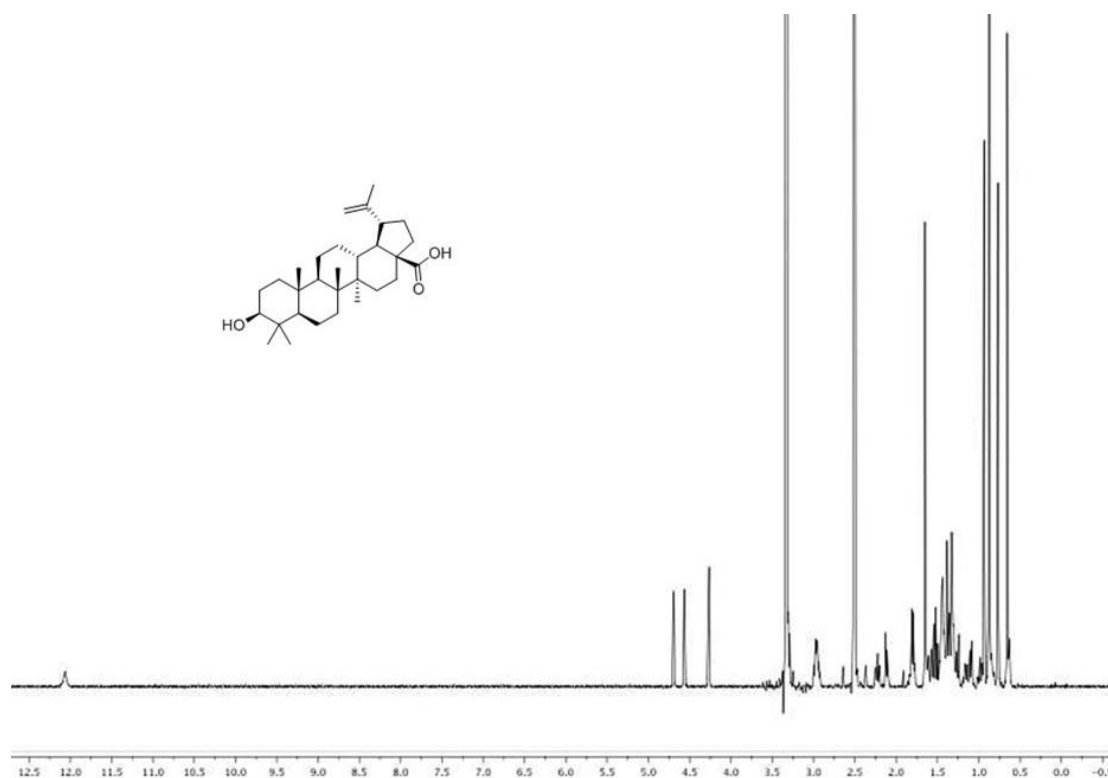


Figure 5.2.4.1.51 Chemical structure and 1H NMR spectrum of **55**

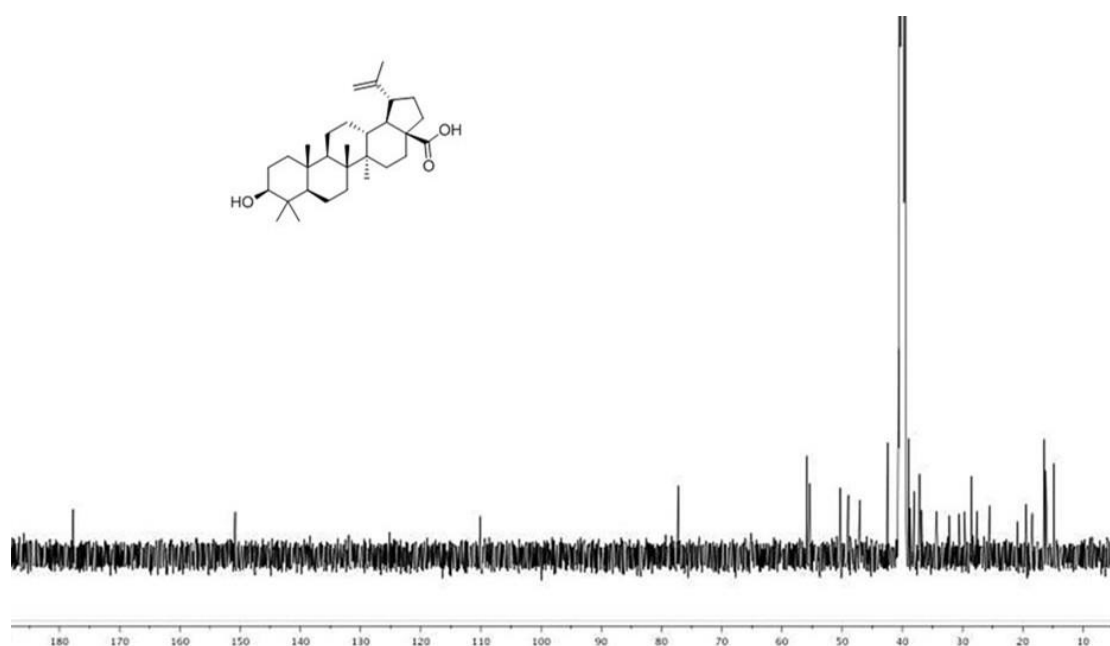


Figure 5.2.4.1.52 Chemical structure and ^{13}C NMR spectrum of **55**

Compound **55** was obtained as white powder. Its molecular formula was determined to be $\text{C}_{30}\text{H}_{48}\text{O}_3$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 455.4107, which was consistent with 7 degrees of unsaturation. Compound **55** displayed similar ^1H and ^{13}C NMR spectra (Figure 5.2.4.1.51 and Figure 5.2.4.1.52) with glycyrrhetic acid (**50**). By comparing its ^1H and ^{13}C NMR data with references, compound **55** was determined as betulinic acid.²¹²⁻²¹³ Betulinic acid was tested by Kumar *et al* and showed good anti-leukemic activity with IC_{50} of values at 13.73, 12.84, 15.27 mg/ml in human leukemic cell lines U937, HL60 and K562, respectively.²¹⁴ In the study by Fujioka *et al*, betulinic acid has showed inhibitory activity against HIV-1 replication with an EC_{50} value of 1.4 μM and inhibited uninfected H9 cell growth with an IC_{50} value of 13 μM .²¹⁵ Betulinic acid has showed *in vitro* and *in vivo* antimalarial activity. The *in vitro* anti-plasmodial IC_{50} values of betulinic acid against sensitive (T9-96) and chloroquine resistant (K1) *Plasmodium falciparum* were 25.9 $\mu\text{g}/\text{ml}$ and 19.6 $\mu\text{g}/\text{ml}$, respectively.²¹⁶ Betulinic acid was tested for *in vitro* anti-malarial activity against asexual erythrocytic stages of the *Plasmodium falciparum* parasite and showed the anti-malarial activity between of moderate to good.²¹⁷

Compound **56**

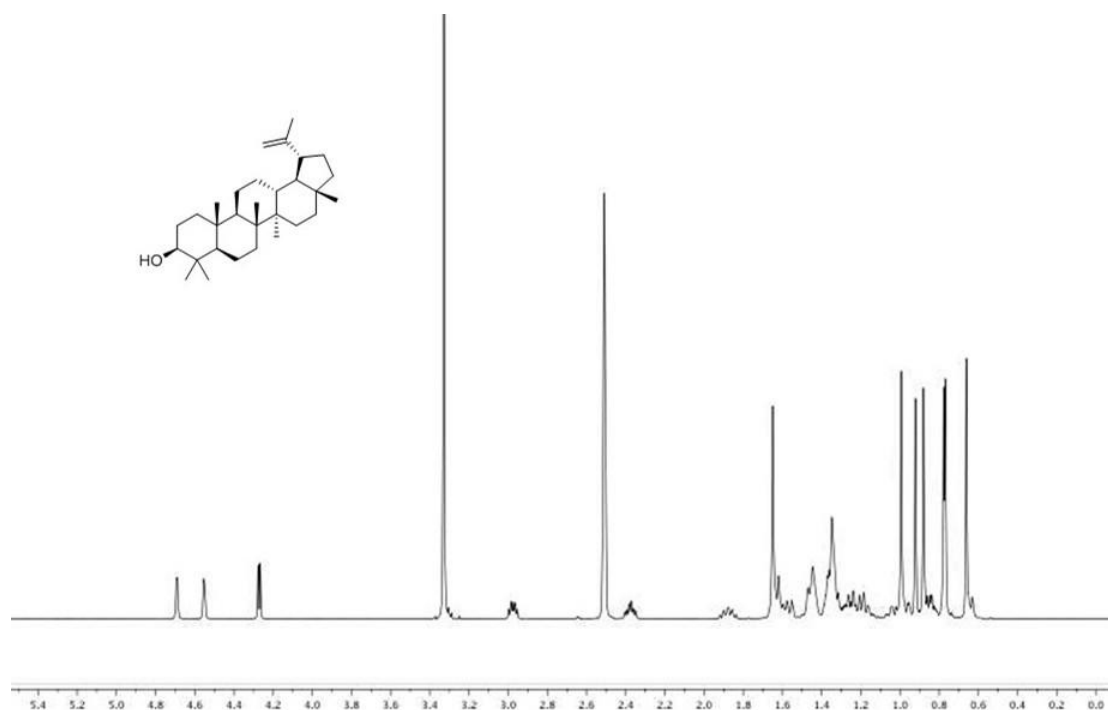


Figure 5.2.4.1.53 Chemical structure and ¹H NMR spectrum of **56**

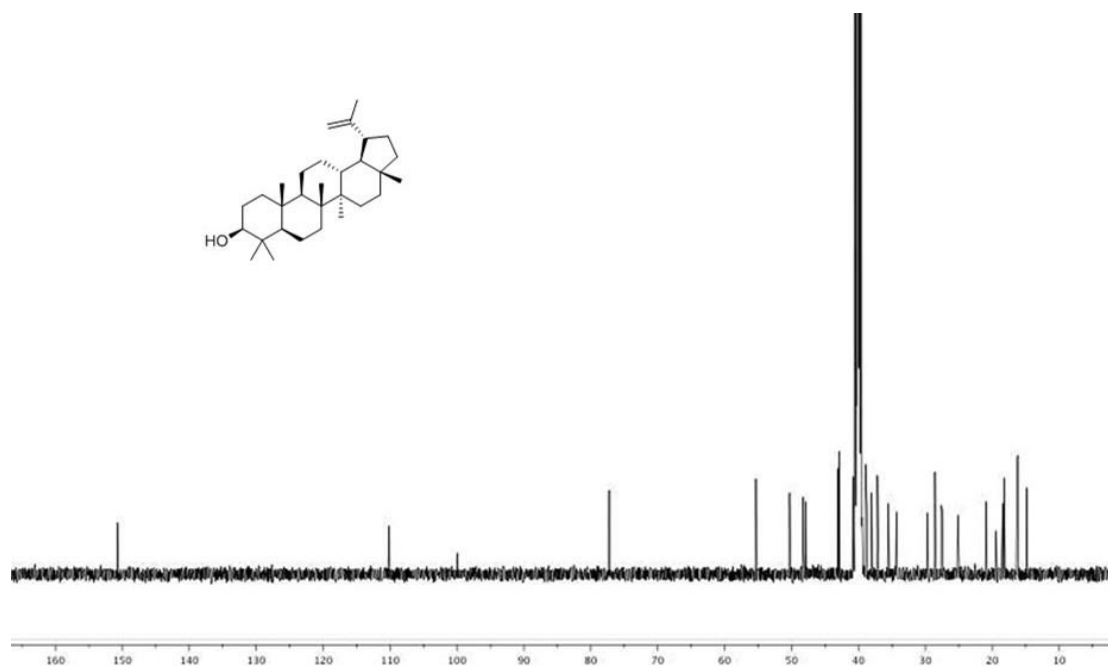


Figure 5.2.4.1.54 Chemical structure and ¹³C NMR spectrum of **56**

Compound **56** was obtained as white powder. Its molecular formula was determined to be C₃₀H₅₀O on the basis of (+)-HRESIMS measurements {[M+H]⁺ ion at *m/z* 427.3925}, which was consistent with 6 degrees of unsaturation. In the ¹H NMR

spectrum (Figure 5.2.4.1.53), there were seven methyl signals between δ_H 0.50 and 1.05 characteristic for pentacyclic triterpenoids, which was also confirmed by the positive result of adding 5% sulfuric acid to the solution of compound **56**. The ^{13}C NMR spectrum of **56** showed a total of 30 resonances (Figure 5.2.4.1.54). By comparing its 1H and ^{13}C NMR data with references, compound **56** was determined as as lupeol.²¹⁸ Lupeol displayed antiprotozoal, antimicrobial, anti-inflammatory, antitumor and chemopreventive properties.²¹⁹ For instance, in the last 10 years, lupeol and its derivatives were tested for anti-malarial activities against lots of *Plasmodium* strains and their IC₅₀ values were ranging from 1.5 to more than 391 $\mu g/ml$ *in vitro*.²¹⁹

5.2.4.2 The physicochemical property analysis

In order to investigate their drug-like properties, physicochemical properties of compounds **27-56** were calculated using the method described in Chapter Two. The results are shown in Table 5.2.4.2.1. Five out of thirty natural products violated Lipinski's Ro5, which means the molecule has no less than two violations against Ro5.¹¹³

Table 5.2.4.2.1 Physicochemical properties of compounds **27-56**

No.	MW	Log <i>P</i>	HBD	HBA	Ro5
27	610.52	-0.87	10	16	Fail
28	464.38	0.60	8	12	Fail
29	630.60	1.26	5	11	Fail
30	286.24	2.40	4	6	Pass
31	272.26	2.84	3	5	Pass
32	416.65	5.33	1	3	Pass
33	374.34	-2.59	6	10	Pass
34	374.84	-1.37	1	4	Pass
35	354.31	-0.27	6	8	Pass
36	270.24	2.71	3	5	Pass
37	434.1	3.32	1	7	Pass
38	240.21	3.61	2	4	Pass
39	510.49	-0.55	5	10	Pass
40	226.23	-0.44	2	4	Pass
41	364.36	0.92	2	5	Pass

42	464.42	2.10	2	3	Pass
43	384.48	2.61	1	4	Pass
44	456.71	6.58	2	3	Pass
45	660.75	-0.72	9	13	Fail
46	154.17	0.74	2	3	Pass
47	348.36	1.22	1	4	Pass
48	286.24	1.81	4	6	Pass
49	329.35	2.53	1	5	Pass
50	470.69	6.03	2	4	Pass
51	363.37	0.39	2	5	Pass
52	810.98	1.60	9	14	Fail
53	354.45	2.10	2	3	Pass
54	376.41	2.18	1	4	Pass
55	456.71	6.64	2	3	Pass
56	426.73	7.45	1	1	Pass

*Content in red means violation against Ro5.

Log *P* is one of the key factors when drug molecules transferred to cell membrane or protein binding sites (hydrophobic) from the aqueous phase. And, the HPLC method employed in the thesis aims to generate front-loading of both extracts and subsequent fractions with desired physicochemical properties (especially Log *P*) prior to ESI-FTMS screening. Thus, in this thesis, the relationship between Log *P* and HPLC retention time was studied. Figures from 5.2.4.2.1 to 5.2.4.2.5 showed the Log *P* value of each isolated natural product and the fraction from which it was isolated.

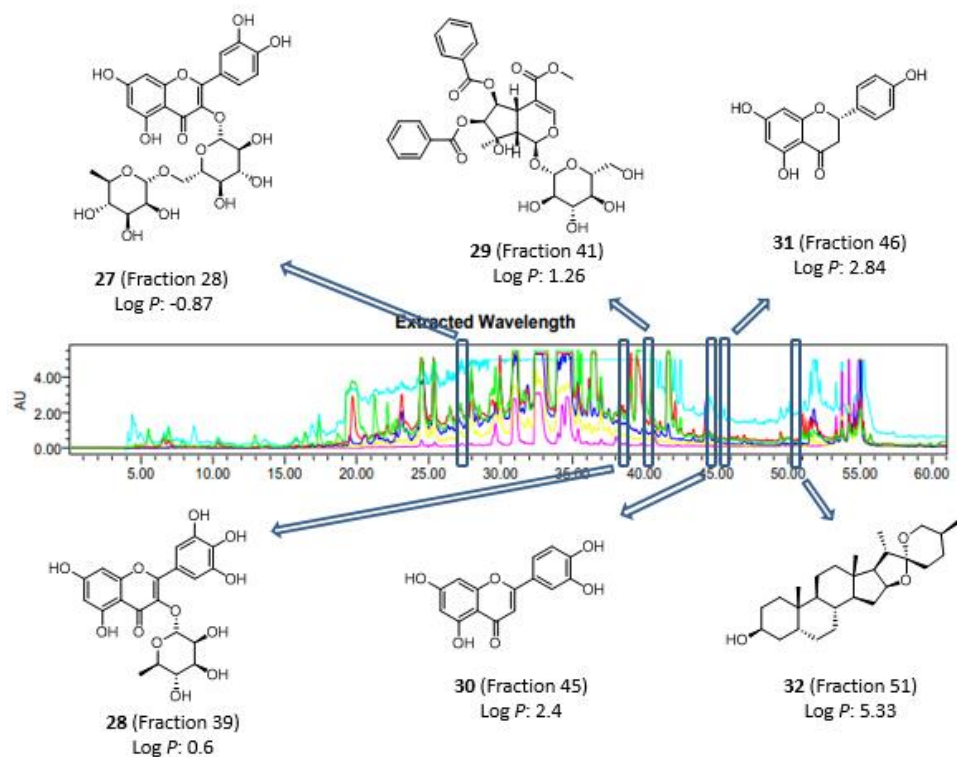


Figure 5.2.4.2.1 Natural products isolated from the Australian plant *Psydraz montigena*

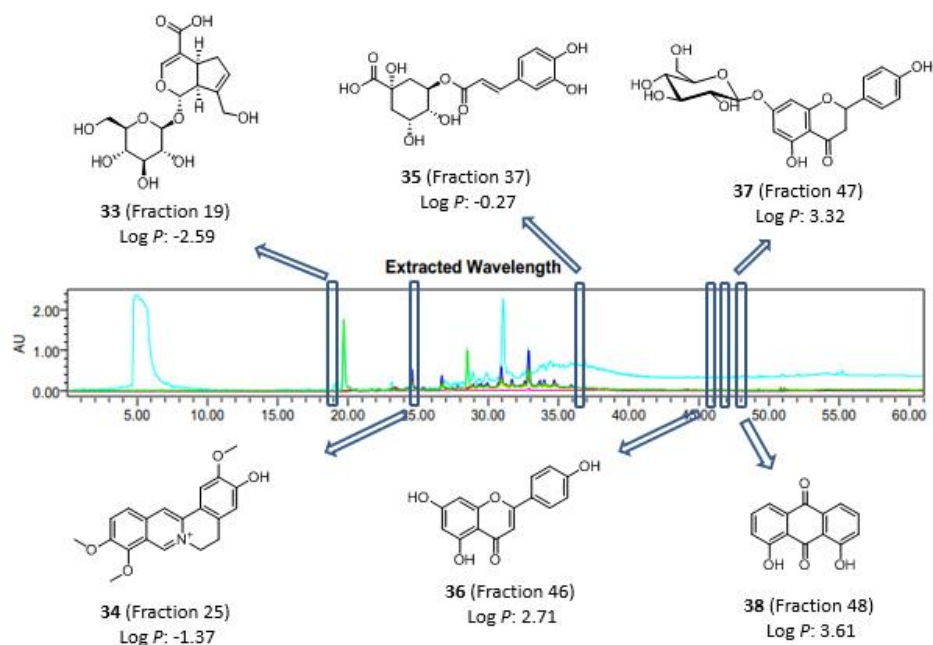


Figure 5.2.4.2.2 Natural products isolated from the Australian plant *Wendlandia basistaminea*

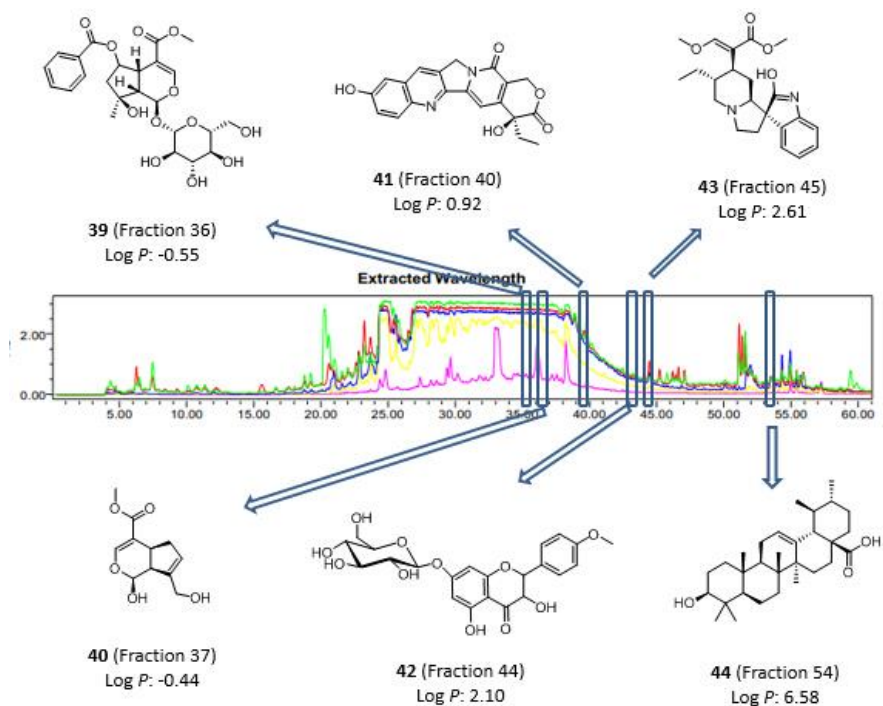


Figure 5.2.4.2.3 Natural products isolated from the Australian plant *Psyrax lamprophylla*

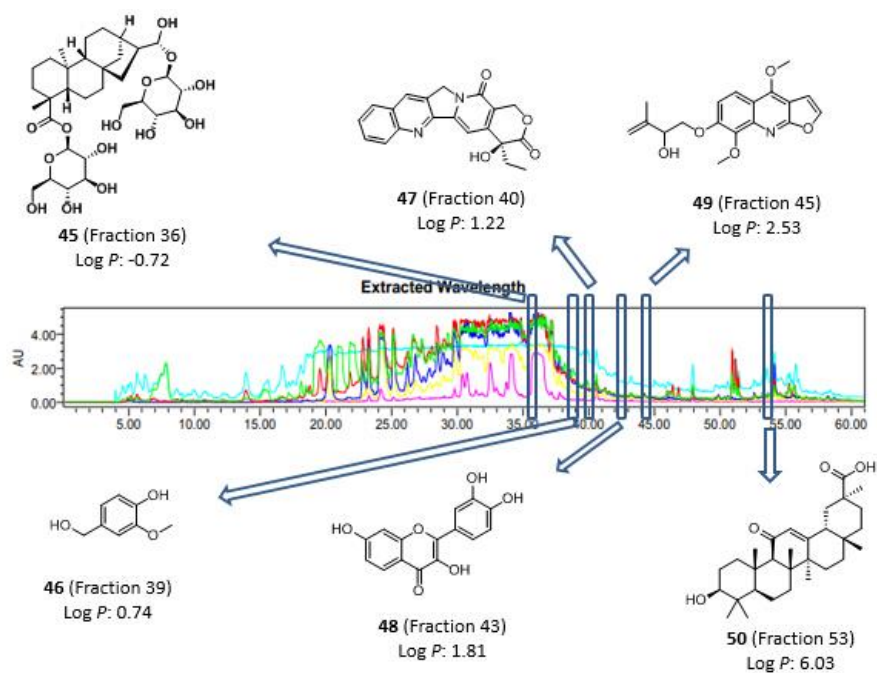


Figure 5.2.4.2.4 Natural products isolated from the Australian plant *Prunus persica*

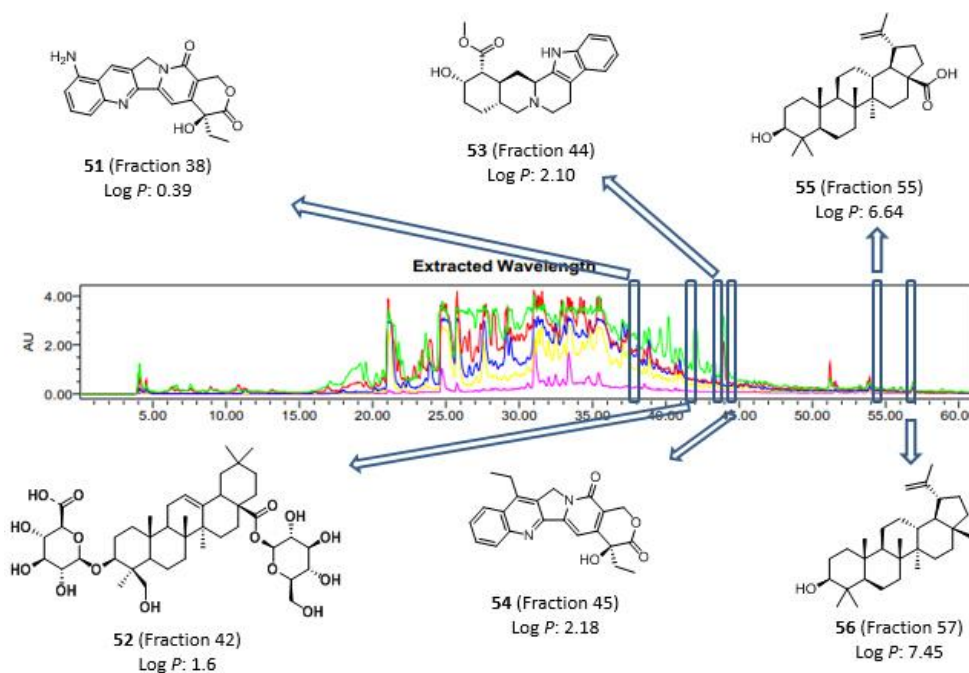


Figure 5.2.4.2.5 Natural products isolated from the Australian plant *Larsenaikia ochreata*

Lead-like enhanced extraction of selected samples was processed with solid-phase extraction (SPE) protocol that filtered on Log *P*. With the increasing of methanol in the mobile phase, the polarity of mobile phase becomes decreasing and the polarity of compounds isolated from Fraction 1 to 60 should also decrease. Whilst, their Log *P* values should increase at the same time.

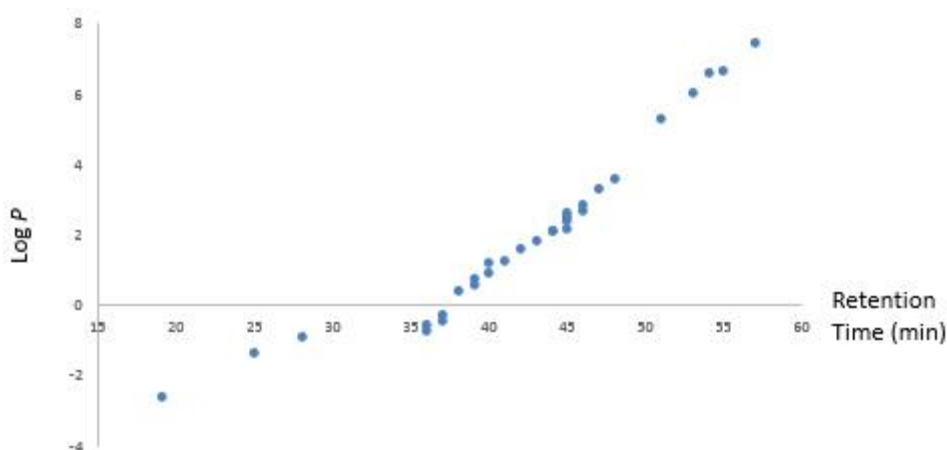


Figure 5.2.4.2.6 Relationship between Log *P* and Retention Time (min)

As is shown in Figure 5.2.4.2.6, the retention time for the isolated natural products in this chapter had a linear relationship with their Log *P* value. Similar with previous

researches, this result proved that the LLE process has the ability to capture most drug-like chemical constituents.

5.2.5 Determination of K_D

Compounds (**29**, **33**, **39**, **45** and **52**) were confirmed as active molecules, which bound to *PfRab11a*. Their molecular weight met the calculation results. In order to determine the binding affinity of these five ligands, the K_D was measured by a titration experiment monitoring the equilibrium population of free protein and ligand-protein complex using a constant concentration of protein (*PfRab11a*, 4.5 μ M in 10 mM ammonium acetate buffer, pH 6.8) and increasing concentrations of ligand (2 μ M, 3 μ M, 4 μ M, 8 μ M, 40 μ M and 50 μ M). Each protein-ligand mixture was incubated at 4 °C for 1 hr before analyzing by ESI-FTMS. The experiment was run in triplicate.

It is assumed that the observed protein ion intensity reflects the true protein concentrations. Either peak height or peak area can be used for quantitation method. The peak height intensity of the *PfRab11a* at 10+ charge state was used for the calculation of the bound *PfRab11a* because the *PfRab11a* peak height intensity at other charge states was in lower abundance. Prism software (Version 6.01) was used to plot the fraction of bound protein against the total ligand concentration of each ligand (Figure 5.2.5.1). According to Equation 6, the value of K_D of each ligand was determined. The results were shown in Table 5.2.5.1.

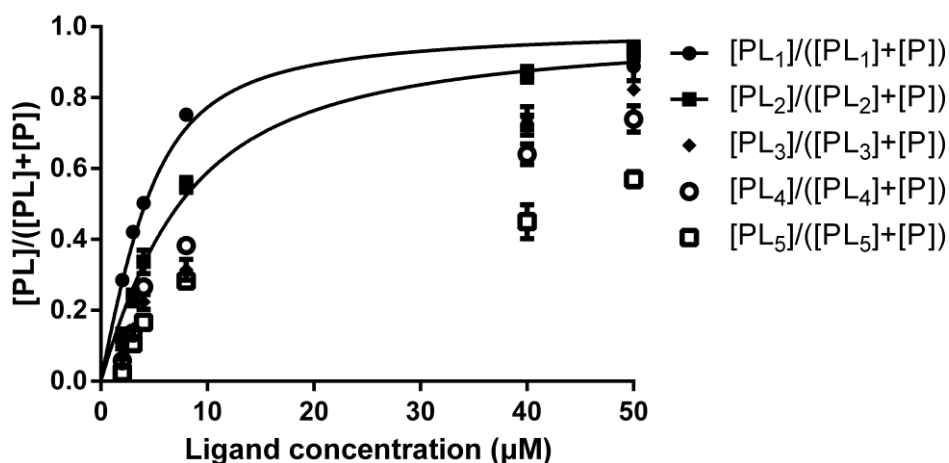


Figure 5.2.5.1 Titration of 5 ligands with *PfRab11a* for determination of K_D . L₃ (**39**), L₄ (**45**) and L₅ (**52**) didn't give curves as their high K_D values.

Table 5.2.5.1 Results for the determination of K_D (μM)

No.	Best-fit values	Standard deviation
L ₁ (29)	1.921	0.4028
L ₂ (33)	5.131	0.3485
L ₃ (39)	12.58	0.6860
L ₄ (45)	14.34	1.412
L ₅ (52)	34.07	3.548

K_D is the equilibrium dissociation constant, a ratio of K_{off}/K_{on} , between the protein and its ligand. K_D and affinity are inversely related. The K_D value relates to the concentration of ligand (the amount of ligand needed for a particular experiment) and so the lower the K_D value (lower concentration) and thus the higher the affinity of the ligand. According to the titration experiments, arborside E (**29**) had the lowest K_D value. Therefore, its binding affinity was the best among all the five *PfRab11a* natural product binders identified by ESI-FTMS. In order to distinguish a real affinity from titration-like-binding, we used the same method to measure the K_D value of arborside E (**29**) again with a lower protein concentration (*PfRab11a*, 3.0 μM), and the K_D value was 1.963 ± 0.2134 μM, which indicated a real K_D was measured.

5.2.6 Determination of IC₅₀

Nonspecific binding can be detected by a competition experiment using a known ligand to compete with binding of the first ligand to the specific site. However, there is no published ligand for *PfRab11a*. Under this circumstance, an orthogonal enzyme assay was run against *PfRab11a* using the Innova GTPase Kit to confirm the inhibitory activity of Compounds (**29**, **33**, **39**, **45** and **52**). Titration experiments were carried out with *PfRab11a* at a fixed concentration of 4.5 μM in 10 mM ammonium acetate buffer (pH 6.8) and the ligand at a varied concentration from 0.01 μM to 100 μM (0.01 μM , 1.0 μM , 1.5 μM , 4.5 μM , 10 μM , 20.0 μM , 30.0 μM , 40.0 μM , 50.0 μM , 80.0 μM and 100.0 μM). The results were showed in Figure 5.2.6.1 and Table 5.2.6.1.

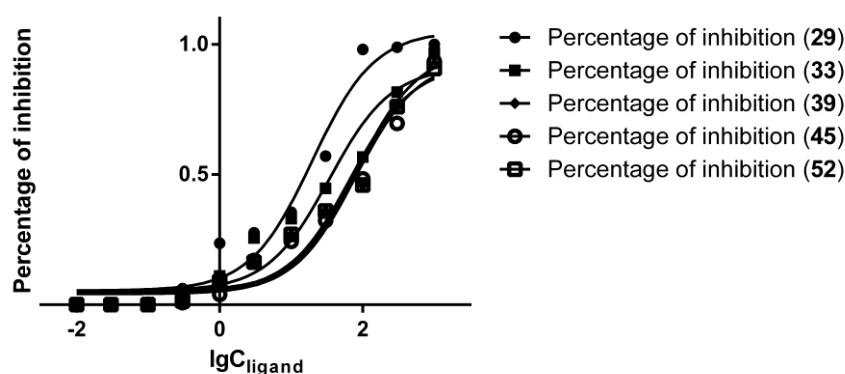


Figure 5.2.6.1 IC₅₀ calculations

Table 5.2.6.1 Results for the determination of IC₅₀ (μM)

No.	IC ₅₀	Standard deviation
29	19.73	1.174
33	34.24	0.8943
39	82.54	1.341
45	79.66	0.7761
52	67.50	0.9311

The results suggested that arborside E (**29**) had an IC₅₀ value of $19.73 \pm 1.174 \mu\text{M}$, which was the best in all the five compounds. This result also confirmed that

ESI-FTMS can be used for bioaffinity screening to identify bioactive small molecules, which can inhibit a protein target.

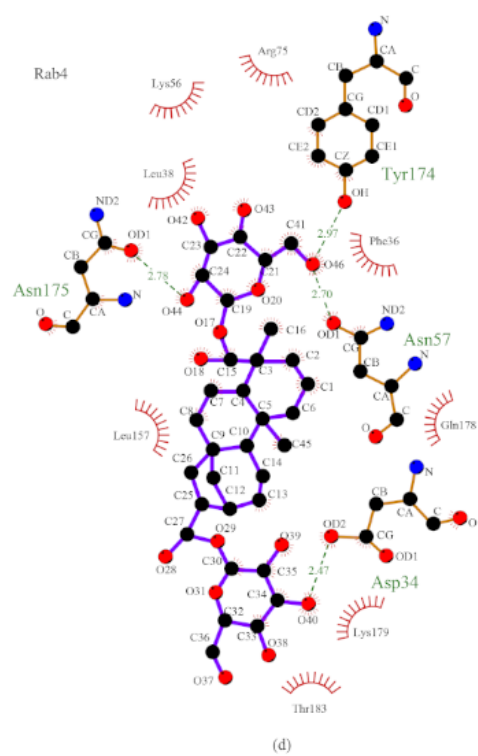
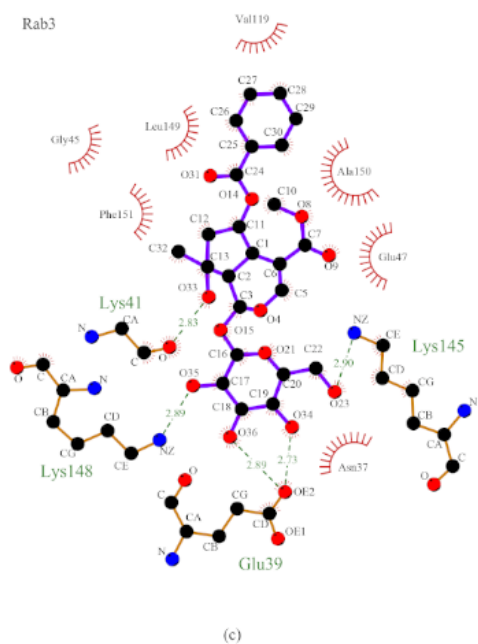
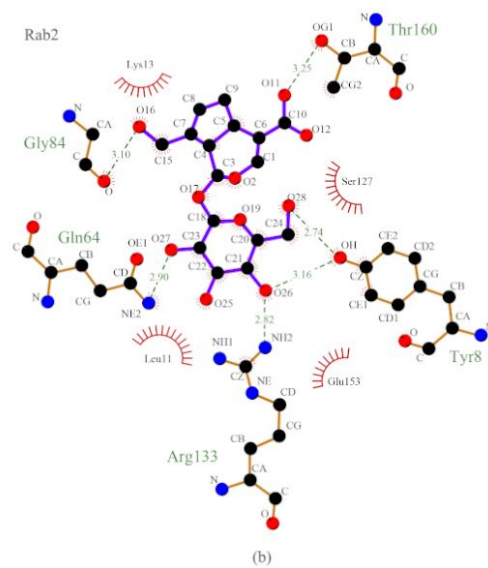
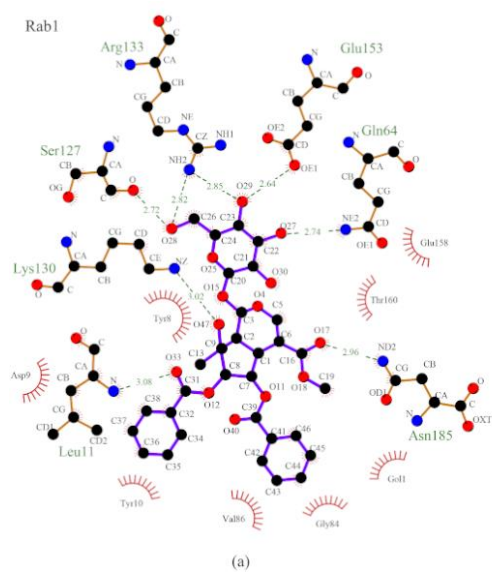
5.2.7 Evaluation of possible *PfRab11a* binding sites of those natural product binders via virtual docking

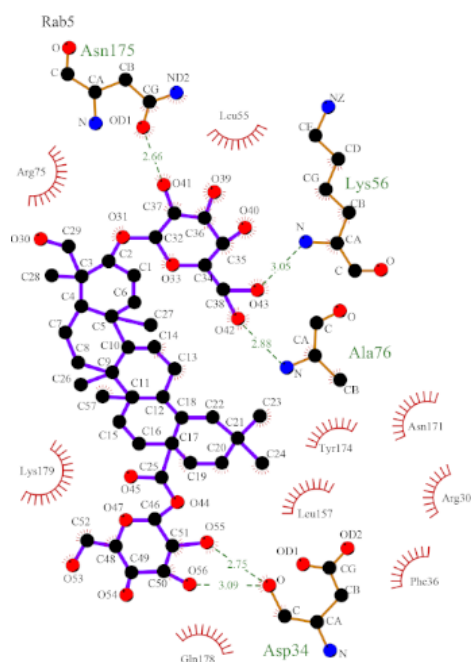
Molecular Docking. Five binding sites were obtained for the protein *PfRab11a* (3BFK) using SiteMap. The molecules obtained by ESI-FTMS screening were docked to the sitemaps respectively with Glide software to find the most appropriate binding site of the target and to investigate their ligand-protein interactions. The binding orientation and features of each molecule relative to the receptor protein were determined and scored with internal scoring function GlideScore. The docking results are given in Table 5.2.7.1. The binding site of the highest docking score ligand was selected and the binding models of them are shown in Figure 5.2.7.1 to Figure 5.2.7.4.

Table 5.2.7.1 Docking scores of the ligands interacting with five different binding sites in 3BFK predicted by SiteMap.

Docking Score (kcal/mol)	Site 1	Site 2	Site 3	Site 4	Site 5
29 (Rab1)	-5.24	-4.83	-6.21	-5.98	-6.16
33 (Rab2)	-5.38	-5.25	-5.45	-3.91	-5.99
33 (Rab2_neg)	-5.56	-5.46	-5.58	-4.76	-4.79
39 (Rab3)	-4.78	-4.15	-5.38	-5.61	-5.09
45 (Rab4)	-6.06	-6.75	-6.17	-3.71	-4.52
52 (Rab5)	-4.31	-7.72	-6.48	-5.69	-6.53

Binding Mode Analysis. The 2D representation of the selected binding modes of the molecules and their receptor hydrogen bond and hydrophobic interactions is shown in Figure 5.2.7.1 and Figure 5.2.7.2. The ligand-protein interaction diagrams were generated in LIGPLOT²²⁰. Hydrogen bond interactions and their atomic distances (in Å) are shown in dashed lines, whereas hydrophobic contacts are shown in red crescents.



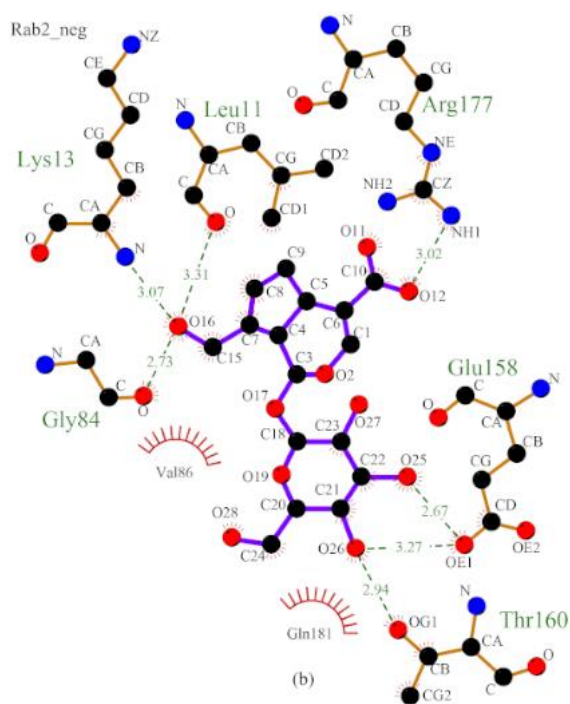


Key

The meaning of the items on the plot is as follows:

- — ● 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 5.2.7.1 Ligand-protein interaction diagrams of the selected ligands generated by LIGPLOT.



Key

The meaning of the items on the plot is as follows:



Figure 5.2.7.2 Ligand-protein interaction diagrams of the selected ligands generated by LIGPLOT (Rab2_neg).

As shown in Figure 5.2.7.1 and Figure 5.2.7.2, the ligands of **29** (Rab1), **33** (Rab2), **39** (Rab3), **45** (Rab4) and **52** (Rab5) bound to four different pockets in the 3BFK binding site. Overall, the numbers of bonded interactions and hydrophobic contacts were observed to be high, suggesting a strong binding between the ligands and the target protein.

29 (Rab1) had hydrogen bond interactions with Leu11, Gln64, Asn185, Arg133, Ser127, Lys130, and Glu153; it also had hydrophobic contacts with Gln1, Tyr8, Asp9, Tyr10, Gly84, Val86, Glu158, and Thr160. **33** (Rab2) had hydrogen bond interactions with Tyr8, Gln64, Gly84, Arg133, and Thr160; it also had hydrophobic contacts with Leu11, Lys13, Ser127, and Glu153. When the molecule of **39** (Rab3) was in negative mode, it had hydrogen bond interactions with Leu11, Lys13, Gly84, Glu158, Thr160, and Arg177; it also had hydrophobic contacts with Val86 and Gln181. **39** (Rab3) had hydrogen bond interactions with Lys41, Glu39, Lys145, and Lys148; it also had hydrophobic contacts with Asn37, Gly45, Glu47, Val119, Leu149, Ala150, and Phe151. **45** (Rab4) had hydrogen bond interactions with Asp34, Asn57, Tyr174, and Asn175; it also had hydrophobic contacts with Phe36, Leu38, Lys56, Arg75, Leu157, Gln178, Lys179, and Thr183. **52** (Rab5) had hydrogen bond interactions with Asp34, Lys56, Ala76, and Asn175; it also had hydrophobic contacts with Arg30, Phe36, Leu55, Arg75, Leu157, Asn171, Tyr174, Gln178, and Lys179.

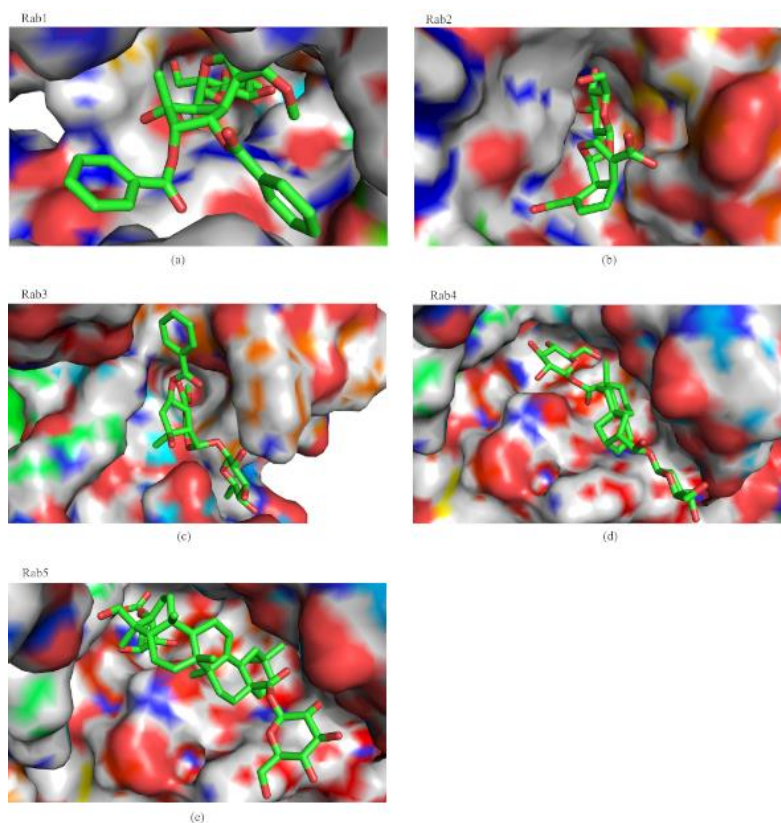


Figure 5.2.7.3 Compound **29** (Rab1), **33** (Rab2), **39** (Rab3), **45** (Rab4) and **52** (Rab5) bound to four different pockets in the *PfRab11a* (3BFK) binding site. The ligands are shown in stick model while the binding site is depicted as a surface for better visualization.

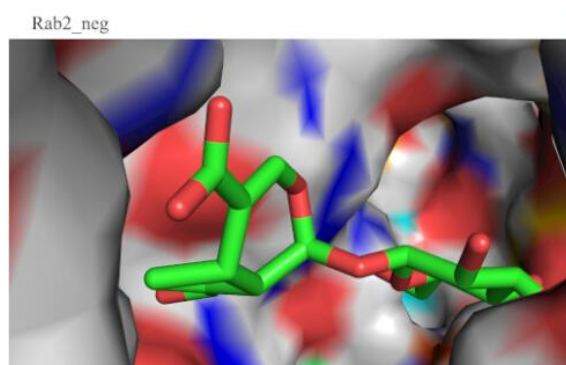


Figure 5.2.7.4 Compound **33** (Rab2_neg) bound to *PfRab11a* (3BFK). The ligand is shown in stick model while the binding site is depicted as a surface for better visualization.

5.3 Summary

ESI-FTMS was used to screen 3,520 natural product extracts against *PfRab11a*. Five extracts from Australian plants were identified to contain molecules that bound to the protein *PfRab11a*. MS-directed isolation led to the identification of a new natural product (**29**) and other four known natural products (**33**, **39**, **45** and **52**). Their binding activity and inhibitory activity against *PfRab11a* was confirmed by ESI-FTMS titration experiments and an orthogonal enzyme assay. To our knowledge, it is the first time a bioaffinity MS screening method has been used to study the malarial target *PfRab11a*, it is the first time that natural product binders (**29**, **33**, **39**, **45** and **52**) to this target have been reported, and a new natural product inhibitor reported. This work successfully demonstrates that direct bioaffinity screening using ESI-FTMS can detect a *PfRab11a*-ligand complex in a complex natural product extract. Meanwhile, a number of additional natural products were isolated from the samples that show a protein-ligand complex, and their Log *P* values and HPLC retention time showed a linear relationship, which provided evidence that the LLE process has the ability to capture drug-like chemical constituents.

5.4 Experimental

5.4.1 *PfRab11a*

PfRab11a was from Van Voorhis Group (University of Washington, WA, USA) and stored in a nonvolatile buffer (Tris, pH 8.0), which is incompatible with MS. Its sequence of amino acid is:

GDYYDYLFKIVLIGDSGVGKSNLLSRFTRDEFNLESKSTIGVEFATKSIQLKNNKIKAQIWDTAGQ
ERYRAITSAYYRGAVGALLVYDITKKNSFENIEKWLKELRDNADSNIVILLVGNKSDLKHLRVIND
NDATQYAKKEKLAFIETSALEATNVELAFHQLLNEIYNVRQKKQATKN.

The nonvolatile buffer was replaced by ammonium acetate (10 mM, pH 6.8) through a buffer-exchange process using Sephadex G25 column to generate stock solution (4.5 μ M).

5.4.2 Natural product extracts

3,520 natural product extracts (0.5 μ L, 250 μ g/ μ L each extract) randomly chosen from Nature Bank were dried and re-solubilized with 5 μ L methanol. Each extract (5 μ L, 25 μ g/ μ L) was incubated with *PfRab11a* at 4.5 μ M in ammonium acetate (10 mM, pH 6.8) for direct infusion ESI-FTMS. The total volume of each sample was 55 μ L. They were incubated for 1 hour at 4°C before being analyzed by MS.

5.4.3 ESI-FTMS analysis

All experiments were performed on a Bruker Apex III 4.7 Tesla external ESI source FTICR mass spectrometer. Samples were injected directly by a Cole-Parmer syringe pump with a flow rate of 2 μ L per minute. The end plate voltage was biased at 3200 V and the capillary voltage at 4500 V relative to the ESI needle during data acquisition. A nebulizing N₂ gas with a pressure of 50 psi and a counter-current drying N₂ gas with a flow of 30 L/min were employed. The drying gas temperature was maintained at 100°C for direct infusion ESI-FTMS. The capillary exit voltage was tuned at 140 V and skimmer 1 voltage at 15 V. Ions were accumulated in an external ion reservoir comprised of an radio frequency only hexapole, a skimmer cone (skimmer 2) with a tuning voltage of 12 V, and an auxiliary gate electrode, prior to injection into the cylindrical InfinityTM analyzer cell, where they were analyzed. Each run took 3-5 minutes. These optimized conditions were used based on our previous studies.^{51, 135}

Mass spectra were recorded in the positive ion mode with mass range 160-6,000 m/z for broadband low-resolution acquisition. Each spectrum was an average of 128 transients (scans) composed of 32,000 data points for high-resolution mode and 32 transients of 32,000 data points for low-resolution mode. All aspects of pulse sequence control and data acquisition were performed on a Pentium III (1.2 GHz) data station running Bruker's Xmass software under Windows operating system.

5.4.4 HPLC, MS and NMR instruments

HPLC was performed on a Waters 600 HPLC Controller coupled with a Waters 996 Photodiode Array Detector. LR-LC-MS was performed on a Waters 2790 Separations Module HPLC coupled with a Waters micromass ZQ spectrometer. NMR spectra were recorded at 30°C on a Varian 500 MHz spectrometer; samples were dissolved in DMSO-*d*₆.

5.4.5 Extraction and isolation

Five extracts gave positive results during FTMS screening. Their taxonomy information was shown in Table 5.4.5.1.

Table 5.4.5.1 Taxonomy information of active extracts

Plant species	Family	Location	Collectors	Collection Date	Voucher held
<i>Psydrax montigena</i>	Rubiaceae	Daintree National Park, NE Queensland	P.I.F.Forster & R.Booth, R. Jago, R.Jensen	Feb-1997	Queensland Herbarium (BRI)
<i>Wendlandia basistaminea</i>	Rubiaceae	Wooroonooran National Park, Boonjie, NE Queensland	P.I.F.Forster & M.C.Tucker	Jun-1996	Queensland Herbarium (BRI)
<i>Psydrax lamprophylla</i>	Rubiaceae	Mt Hedley, 3 km ENE of Rossville, NE Queensland	P.I.Forster & R.Booth	Apr-1999	Queensland Herbarium (BRI)
<i>Prunus persica</i>	Rosaceae	Camp Eden, Currumbin Valley, SE Queensland	P.I.Forster & G.Leiper	Dec-1997	Queensland Herbarium (BRI)

<i>Larsenaikia ochreatea</i>	Rubiaceae	Mungana National Park, Chillagoe, NE Queensland	P.I.Forster & T.Ryan	Feb-1996	Queensland Herbarium (BRI)
----------------------------------	-----------	--	-------------------------	----------	----------------------------------

5.4.5.1 Chemical constituents from the Australian plant *Psydrax montigena*

The freeze-dried and ground plant material (10 g) was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under gravity then discarded. CH₂Cl₂ (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration the biota was extracted with another volume of MeOH (250 mL) whilst being shaken at 200 rpm for 16 h. All CH₂Cl₂/MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (2.80 g). A portion of this material (1.0 g) was preadsorbed to C18-bonded silica (1 g), and then packed into a stainless steel cartridge (10 × 30 mm) that was subsequently attached to a Thermo Betasil C18 semipreparative HPLC column (150 × 21.2 mm). Isocratic HPLC conditions of 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) were employed for the first 10 min; then a linear gradient from 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9 mL/min. Sixty fractions (60 × 1 min) were collected from the start of the HPLC run (Figure 5.4.5.1.1). Looking for the ion of interest at *m/z* 629.1626 identified in the ESI-FTMS screen, fractions were analyzed by (±)-LRESIMS and NMR spectroscopy.

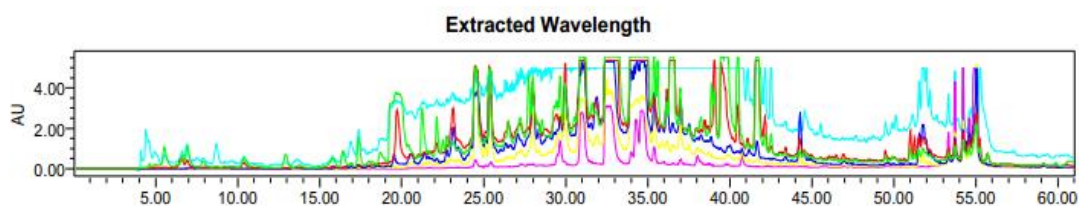


Figure 5.4.5.1.1 HPLC chromatogram of *Psyrdrax montigena*

Fraction 28 yielded compound **27** (2.4 mg). Yellow powder. LR-ESI-MS: m/z 611.2 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.60 (1H, s), 7.56 (1H, d, $J=2.1$ Hz), 7.54 (1H, d, $J=2.3$ Hz), 6.84 (1H, d, $J=8.1$ Hz), 6.39 (1H, d, $J=2.0$ Hz), 6.20 (1H, d, $J=2.0$ Hz), 5.35 (1H, m), 4.39 (1H, d, $J=1.4$ Hz), 3.71 (1H, d, $J=10.1$ Hz), 3.40 (1H, dd, $J=3.2, 1.5$ Hz), 3.30 (1H, d, $J=2.9$ Hz), 3.28 (1H, d, $J=3.1$ Hz), 3.27-3.21 (m), 1.00 (3H, d, $J=6.3$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 177.8, 164.6, 161.7, 157.0, 156.9, 148.9, 145.2, 133.7, 122.0, 121.6, 116.7, 115.7, 104.4, 101.6, 101.2, 99.1, 94.0, 76.9, 76.4, 74.5, 72.3, 71.0, 70.8, 70.4, 68.7, 67.5, 18.2.

Fraction 39 yielded compound **28** (2.5 mg). Yellow powder. LR-LC-MS: m/z 465.0 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.70 (1H, s, -OH), 6.89 (2H, s), 6.37 (1H, d, $J=2.0$ Hz), 6.20 (1H, d, $J=2.0$ Hz), 5.20 (1H, d, $J=1.4$ Hz), 3.98 (1H, dd, $J=1.5, 3.0$ Hz), 3.55 (1H, dd, $J=3.2, 9.5$ Hz), 3.16 (1H, t, $J=9.4$ Hz), 0.85 (3H, d, $J=6.1$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 178.2, 164.7, 161.8, 157.9, 156.9, 146.2, 136.9, 134.7, 120.1, 108.3, 104.5, 102.4, 99.1, 94.0, 71.7, 71.0, 70.8, 70.5, 18.0.

Fraction 41 yielded compound **29** (1.4 mg). Light yellow, amorphous solid. LR-LC-MS: m/z 629.7 $[M+H]^+$. Its NMR data and full structure elucidation is described in the following Chapter 5.3.4.1.

Fraction 45 yielded Compound **30** (2.7 mg). Yellow needles. LR-LC-MS: m/z 287.3 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.97 (1H, s, -OH), 7.43 (1H, d, $J=2.1$ Hz), 7.41 (1H, d, $J=2.3$ Hz), 6.89 (1H, d, $J=8.4$), 6.67 (1H, s, H-8), 6.45 (1H, d, $J=2.0$ Hz), 6.19 (1H, d, $J=2.0$). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 182.1, 164.6, 164.3, 161.9, 157.7, 150.2, 146.2, 121.9, 119.5, 116.5, 113.8, 104.1, 103.3, 99.3, 94.3.

Fraction 46 yielded compound **31** (2.1 mg). Light brown powder. LR-LC-MS: m/z 273.3 $[M-H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.16 (1H, s, -OH), 7.40 (2H, d, $J=6.5$ Hz), 6.91 (2H, d, $J=9.5$ Hz), 5.88 (2H, s), 5.44 (1H, dd, $J=2.9, 12.5$ Hz), 2.70 (1H, d, $J=3.1$ Hz), 2.66 (1H, d, $J=3.1$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 196.8, 167.2, 163.9, 163.4, 158.2, 129.3, 128.8, 115.6, 102.2, 96.3, 95.4, 78.9, 42.4.

Fraction 51 yielded compound **32** (1.7 mg). Light brown powder. LR-LC-MS: m/z 413.4 $[M-3H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 4.49 (1H, s), 3.56 (1H, m), 3.54 (1H, m), 1.95 (1H, m), 1.34-1.79 (m), 1.04-1.14 (m), 0.78-0.89 (each 3H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 109.3, 80.9, 71.4, 66.9, 62.8, 56.3, 54.3, 42.2, 40.5, 40.0, 37.1, 36.0, 35.6, 32.4, 31.9, 31.4, 30.8, 30.1, 29.1, 24.3, 20.9, 18.9, 16.4, 14.9, 12.8.

5.4.5.2 Chemical constituents from the Australian plant *Wendlandia basistaminea*

The freeze-dried and ground plant material (30 g) from the Australian plant *Wendlandia basistaminea* sp. was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under gravity then discarded. CH_2Cl_2 (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration the biota was extracted with another volume of MeOH (250 mL) whilst being shaken at 200 rpm for 16 h. All CH_2Cl_2 /MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (3.83 g). A portion of this material (1.0 g) was preadsorbed to C18-bonded silica (1 g), and then packed into a stainless steel cartridge (10×30 mm) that was subsequently attached to a Thermo Betasil C18 semipreparative HPLC column (150 × 21.2 mm). Isocratic HPLC conditions of 90% H_2O (0.1% TFA)/10% MeOH (0.1% TFA) were employed for the first 10 min; then a linear gradient from 90% H_2O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9 mL/min. Sixty fractions (60 × 1

min) were collected from the start of the HPLC run (Figure 5.4.5.2.1). Looking for the ion of interest at m/z 372.4237 identified in the ESI-FTMS screen, fractions were analyzed by (\pm)-LRESIMS and NMR spectroscopy.

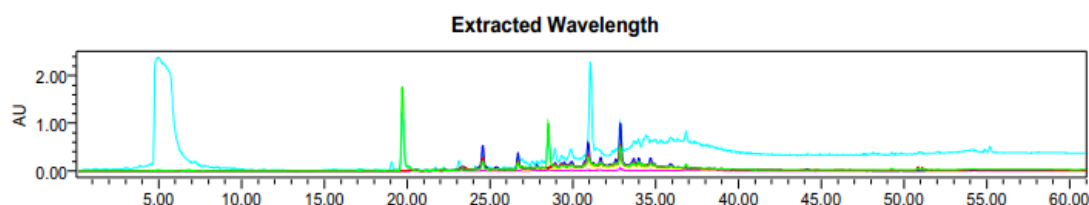


Figure 5.4.5.2.1 HPLC chromatogram of *Wendlandia basistaminea* sp.

Fraction 19 yielded compound **33** (3.7 mg). Yellow needles. LR-LC-MS: m/z 373.6 $[M-H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 11.93 (1H, brs), 7.41 (1H, s), 5.69 (1H, s), 5.09 (1H, d, $J=6.8$ Hz), 4.54 (1H, d, $J=7.8$ Hz), 4.14 (1H, d, $J=14.9$ Hz), 3.98 (1H, d, $J=14.8$ Hz), 3.67 (1H, d, $J=11.7$ Hz), 3.43 (1H, dd, $J=5.9, 11.6$ Hz), 3.17 (1H, d, $J=8.8$ Hz), 3.12 (1H, d, $J=6.0$ Hz), 3.05 (1H, t, $J=8.3$ Hz), 2.99 (1H, t, $J=8.5$ Hz), 2.70 (1H, dd, $J=8.2, 16.6$ Hz), 2.61 (1H, t, $J=7.5$ Hz), 2.06 (2H, m). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 173.0, 153.6, 147.9, 128.5, 117.2, 105.4, 100.7, 96.2, 77.7, 77.0, 73.9, 70.6, 61.5, 60.6, 50.2, 38.5, 32.7.

Fraction 25 yielded compound **34** (2.3 mg). Brown powder. LR-LC-MS: m/z 338.5 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 9.84 (1H, d, $J=2.7$ Hz), 8.96 (1H, s), 8.19 (1H, dd, $J=2.3, 9.2$ Hz), 8.02 (1H, d, $J=9.2$ Hz), 7.70 (1H, s), 6.86 (1H, s), 4.91 (2H, t, $J=6.0$ Hz), 4.10 (3H, d, $J=0.9$ Hz), 4.07 (3H, d, $J=1.1$ Hz), 3.94 (3H, s), 3.15 (2H, t, $J=6.3$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 150.6, 150.5, 148.4, 145.6, 144.0, 138.6, 133.7, 129.3, 127.2, 123.8, 121.7, 119.9, 118.1, 115.4, 109.9, 62.3, 57.5, 56.7, 55.9, 26.2.

Fraction 37 yielded compound **35** (2.0 mg). Colorless needles. LR-LC-MS: m/z 355.3 $[M-H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 9.59 (1H, s), 9.16 (1H, s), 7.42 (1H, d, $J=15.9$ Hz), 7.04 (1H, d, $J=2.1$ Hz), 6.99 (1H, dd, $J=2.1, 8.3$ Hz), 6.77 (1H, d, $J=8.2$ Hz), 6.15 (1H, d, $J=15.9$ Hz), 5.07 (1H, m), 4.92 (1H, d, $J=5.0$ Hz), 4.77 (1H, d, $J=5.8$ Hz), 3.93 (1H, brs), 3.56 (1H, m), 2.02 (3H, m), 1.95 (2H, dd, $J=3.7, 13.2$ Hz), 1.79 (1H, dd,

$J=7.7, 13.0$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 175.4, 166.2, 148.8, 146.0, 145.4, 126.0, 121.8, 116.2, 115.2, 114.7, 73.9, 71.3, 70.8, 68.5, 37.7, 36.7.

Fraction 46 yielded compound **36** (2.0 mg). Yellow needles. LR-LC-MS: m/z 269.3 $[\text{M}-\text{H}]^+$. ^1H NMR (DMSO- d_6 , 500 MHz), δ_{H} : 12.97 (1H, s), 10.84 (1H, s), 10.36 (1H, s), 7.93 (2H, d, $J=8.7$ Hz), 6.93 (2H, d, $J=8.9$ Hz), 6.79 (1H, s), 6.49 (1H, d, $J=2.0$ Hz), 6.20 (1H, d, $J=2.0$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 182.2, 164.6, 164.2, 161.9, 161.6, 157.8, 128.9, 121.6, 116.4, 104.2, 103.3, 99.3, 94.4.

Fraction 47 yielded compound **37** (2.5 mg). Brown powder. LR-LC-MS: m/z 433.6 $[\text{M}-\text{H}]^+$. ^1H NMR (DMSO- d_6 , 500 MHz), δ_{H} : 12.01 (1H, s), 7.11 (2H, dd, $J=7.2, 7.8$ Hz), 6.71 (2H, dd, $J=7.3, 8.0$ Hz), 6.20 (2H, s), 6.80 (1H, s), 5.51 (1H, m). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 196.8, 166.8, 163.2, 157.4, 130.9, 127.5, 117.0, 109.2, 101.8, 94.8, 94.1, 82.5, 81.5, 76.8, 73.4, 71.5, 62.2, 43.2.

Fraction 48 yielded compound **38** (2.6 mg). Yellow needles. LR-LC-MS: m/z 239.3 $[\text{M}-\text{H}]^+$. ^1H NMR (DMSO- d_6 , 500 MHz), δ_{H} : 11.94 (1H, brs), 7.85 (1H, s), 7.83 (1H, d, $J=0.8$ Hz), 7.81 (1H, s), 7.74 (1H, d, $J=1.1$ Hz), 7.73 (1H, d, $J=1.1$ Hz), 7.41 (1H, d, $J=1.1$ Hz), 7.40 (1H, d, $J=1.1$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 192.5, 181.9, 161.8, 137.9, 133.8, 124.9, 119.8, 116.5.

5.4.5.3 Chemical constituents from the Australian plant *Psydraz lamprophylla*

The freeze-dried and ground plant material (10 g) from the Australian plant *Psydraz lamprophylla* sp. was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under gravity then discarded. CH_2Cl_2 (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration the biota was extracted with another volume of MeOH (250 mL) whilst being shaken at 200 rpm for 16 h. All CH_2Cl_2 /MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (2.41 g). A portion of this material (1.0 g) was

preadsorbed to C18-bonded silica (1 g), and then packed into a stainless steel cartridge (10×30 mm) that was subsequently attached to a Thermo Betasil C18 semipreparative HPLC column (150 × 21.2 mm). Isocratic HPLC conditions of 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) were employed for the first 10 min; then a linear gradient from 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9 mL/min. Sixty fractions (60 × 1 min) were collected from the start of the HPLC run (Figure 5.4.5.3.1). Looking for the ion of interest at *m/z* 510.5320 identified in the ESI-FTMS screen, fractions were analyzed by (±)-LRESIMS and NMR spectroscopy.

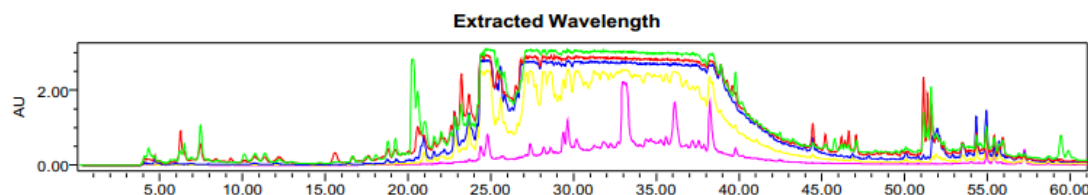


Figure 5.4.5.3.1 HPLC chromatogram of *Psydraz lamprophylla* sp.

Fraction 36 yielded compound **39** (1.3 mg). Brown powder. LR-LC-MS: *m/z* 511.5 [M+H]⁺. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_H: 8.02 (2H, m), 7.52 (1H, m), 7.38 (2H, m), 7.36 (1H, s), 5.49 (1H, m), 5.46 (1H, d, *J*=1.5 Hz), 1.35 (3H, s). ¹³C NMR (DMSO-*d*₆, 500 MHz), δ_C: 151.8, 132.7, 130.1, 129.3, 128.1, 108.6, 98.4, 93.6, 78.2, 77.6, 76.3, 76.2, 73.0, 70.0, 61.5, 50.0, 46.7, 36.9.

Fraction 37 yielded compound **40** (1.8 mg). Brown powder. LR-LC-MS: *m/z* 227.3 [M+H]⁺. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_H: 7.49 (1H, d, *J*=5.8 Hz), 5.71 (1H, s), 4.75 (1H, d, *J*=9.0 Hz), 4.14 (1H, d, *J*=14.8), 4.01 (1H, d, *J*=13.7 Hz), 3.01 (1H, m), 2.69 (1H, dd, *J*=8.3, 15.8 Hz), 2.40 (1H, t, *J*=7.8 Hz), 1.99 (1H, m). ¹³C NMR (DMSO-*d*₆, 500 MHz), δ_C: 167.6, 153.1, 145.7, 125.6, 110.4, 96.4, 62.2, 51.4, 47.2, 38.7, 36.2.

Fraction 40 yielded compound **41** (2.6 mg). Light yellow needles. LR-LC-MS: *m/z* 365.2 [M+H]⁺. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_H: 10.33 (1H, s), 8.46 (1H, s), 8.03 (1H, d, *J*=9.2 Hz), 7.43 (1H, dd, *J*=2.7, 9.2 Hz), 7.29 (1H, d, *J*=2.7 Hz), 7.27 (1H, s), 6.50 (1H,

s), 5.42 (2H, s), 5.24 (2H, s), 1.87 (2H, m), 0.88 (3H, t, $J=7.3$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 173.0, 157.3, 157.1, 150.5, 149.9, 146.4, 143.7, 131.1, 130.4, 130.2, 129.7, 123.5, 118.6, 109.2, 96.3, 72.9, 65.7, 50.6, 30.7, 8.25.

Fraction 44 yielded compound **42** (2.7 mg). Brown powder. LR-LC-MS: m/z 465.4 $[\text{M}+\text{H}]^+$. ^1H NMR (DMSO- d_6 , 500 MHz), δ_{H} : 12.01 (1H, s), 7.11 (2H, dd, $J=7.2, 7.8$ Hz), 6.71 (2H, dd, $J=7.3, 8.0$ Hz), 6.20 (2H, s), 6.80 (1H, s), 5.51 (1H, m). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 196.8, 166.8, 163.2, 157.4, 130.9, 127.5, 117.0, 109.2, 101.8, 94.8, 94.1, 82.5, 81.5, 76.8, 73.4, 71.5, 62.2, 43.2.

Fraction 45 yielded compound **43** (1.6 mg). White needles. LR-LC-MS: m/z 385.4 $[\text{M}+\text{H}]^+$. ^1H NMR (DMSO- d_6 , 500 MHz), δ_{H} : 10.12 (1H, s), 7.33 (1H, s), 7.26 (1H, d, $J=7.3$ Hz), 7.15 (1H, ddd, $J=1.2, 7.7$ Hz), 6.96 (1H, ddd, $J=1.0, 7.6$ Hz), 6.76 (1H, d, $J=7.5$ Hz), 3.73 (2H, s), 3.56 (3H, s), 3.17 (2H, m), 2.37 (1H, m), 2.15 (4H, m), 2.02 (2H, m), 1.88 (1H, dd, $J=6.9, 12.7$ Hz), 1.59 (1H, t, $J=10.8$ Hz), 0.76 (3H, t, $J=7.3$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 180.2, 160.5, 142.4, 134.3, 128.1, 123.6, 122.0, 111.1, 119.2, 75.5, 57.7, 55.8, 54.7, 40.4, 40.6, 37.5, 34.5, 24.2, 11.5.

Fraction 54 yielded compound **44** (1.5 mg). Brown powder. LR-LC-MS: m/z 455.4 $[\text{M}+\text{H}]^+$. ^1H NMR (DMSO- d_6 , 500 MHz), δ_{H} : 11.94 (1H, s), 5.14 (2H, t, $J=3.4$ Hz), 4.35 (1H, t, $J=5.0$ Hz), 4.29 (1H, d, $J=5.2$ Hz), 3.45 (1H, dd, $J=5.0, 6.9$ Hz), 3.01 (2H, m), 2.11 (1H, d, $J=11.3$ Hz), 1.94 (1H, td, $J=3.6, 13.4$ Hz), 1.85 (4H, dd, $J=5.0, 10.5$ Hz), 1.79 (1H, d, $J=9.0$ Hz), 1.49 (1H, m), 1.29 (1H, m), 1.07 (1H, d, $J=7.0$ Hz), 1.05 (3H, s), 0.93 (4H, m), 0.90 (3H, s), 0.87 (3H, s), 0.82 (3H, d, $J=6.4$ Hz), 0.76 (3H, s), 0.68 (3H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 178.7, 138.5, 125.8, 79.1, 55.5, 53.1, 48.1, 47.8, 42.3, 39.7, 39.4, 39.2, 39.0, 38.9, 37.2, 37.1, 33.3, 30.9, 28.3, 28.3, 27.0, 24.5, 23.8, 23.6, 21.5, 18.6, 17.3, 17.1, 15.9, 15.7.

5.4.5.4 Chemical constituents from the Australian plant *Prunus persica*

The freeze-dried and ground plant material (10 g) from the Australian plant *Prunus persica* sp. was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under

gravity then discarded. CH₂Cl₂ (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration the biota was extracted with another volume of MeOH (250 mL) whilst being shaken at 200 rpm for 16 h. All CH₂Cl₂/MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (2.51 g). A portion of this material (1.0 g) was preadsorbed to C18-bonded silica (1 g), and then packed into a stainless steel cartridge (10 × 30 mm) that was subsequently attached to a Thermo Betasil C18 semipreparative HPLC column (150 × 21.2 mm). Isocratic HPLC conditions of 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) were employed for the first 10 min; then a linear gradient from 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9 mL/min. Sixty fractions (60 × 1 min) were collected from the start of the HPLC run (Figure 5.4.5.4.1). Looking for the ion of interest at *m/z* 660.9427 identified in the ESI-FTMS screen, fractions were analyzed by (±)-LRESIMS and NMR spectroscopy.

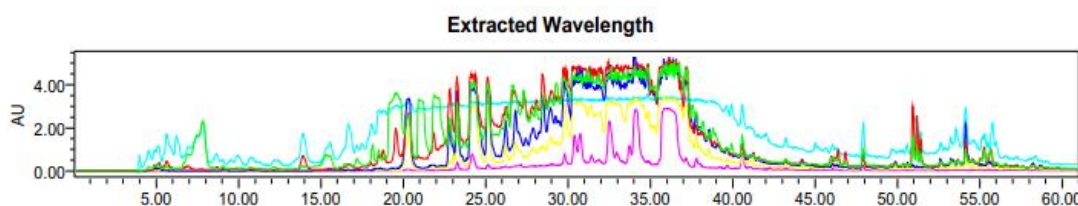


Figure 5.4.5.4.1 HPLC chromatogram of *Prunus persica* sp.

Fraction 36 yielded compound **45** (1.6 mg). White needles. LR-LC-MS: *m/z* 660.8 [M+H]⁺. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_H: 6.77 (1H, m), 5.40 (1H, m), 5.19 (1H, m), 4.29 (1H, m), 2.01 (2H, m), 1.2-1.8 (each 2H, 8 X -CH₂), 0.84 (1H, s). ¹³C NMR (DMSO-*d*₆, 500 MHz), δ_C: 176.5, 107.0, 99.6, 96.3, 82.0, 81.3, 56.6, 54.8, 44.9, 44.0, 40.5, 37.8, 28.2, 21.1, 18.6, 15.8.

Fraction 39 yielded compound **46** (2.5 mg). White needles. LR-LC-MS: m/z 154.1 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 8.76 (1H, s), 6.88 (1H, d, $J=7.4$ Hz), 6.70 (1H, d, $J=7.6$ Hz), 4.98 (1H, t, $J=5.6$ Hz), 4.37 (1H, d, $J=5.6$ Hz), 3.34 (3H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 147.8, 145.7, 133.9, 119.5, 115.5, 111.5, 63.4, 55.9.

Fraction 40 yielded compound **47** (2.4 mg). Light yellow crystal powder. LR-LC-MS: m/z 348.1 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 8.71 (1H, s), 8.19 (1H, d, $J=8.5$ Hz), 8.15 (1H, d, $J=7.6$ Hz), 7.88 (1H, ddd, $J=1.4, 6.9, 8.5$ Hz), 7.73 (1H, ddd, $J=1.1, 6.9, 7.5$), 7.37 (1H, s), 6.54 (1H, s), 5.44 (1H, s), 5.33 (1H, s), 1.88 (1H, m), 0.89 (3H, t, $J=7.4$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 172.9, 157.3, 153.0, 150.4, 148.4, 146.0, 132.0, 130.9, 130.3, 129.5, 129.0, 128.4, 128.1, 119.5, 97.2, 72.8, 65.7, 50.7, 30.8, 8.2.

Fraction 43 yielded compound **48** (1.9 mg). Yellow needles. LR-LC-MS: m/z 285.3 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 10.73 (1H, s), 9.50 (1H, s), 9.26 (1H, s), 9.03 (1H, s), 7.93 (1H, d, $J=7.5$ Hz), 7.69 (1H, d, $J=2.3$ Hz), 7.54 (1H, dd, $J=2.2, 8.5$ Hz), 6.89 (3H, m). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 172.4, 162.7, 156.7, 147.7, 145.5, 137.6, 126.9, 122.9, 120.1, 116.0, 115.4, 115.1, 114.7, 102.3.

Fraction 45 yielded compound **49** (1.5 mg). Light brown powder. LR-LC-MS: m/z 329.9 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 11.98 (1H, s), 7.73 (1H, m), 7.67 (1H, t, $J=1.7$ Hz), 6.52 (1H, d, $J=0.6, 1.8$ Hz), 5.49 (1H, s), 4.93 (1H, d, $J=13.1$ Hz), 4.49 (1H, d, $J=13.1$ Hz), 4.12 (3H, s), 3.33 (3H, s), 3.13 (1H, t, $J=15.2$ Hz), 2.78 (1H, d, $J=15.4$ Hz), 2.64 (1H, d, $J=4.0$ Hz), 2.61 (1H, d, $J=4.0$ Hz), 2.58 (1H, m), 2.47 (1H, dd, $J=3.3, 15.6$ Hz), 2.28 (1H, dd, $J=3.3, 14.9$ Hz), 1.82 (1H, m), 1.73 (1H, m), 1.26 (1H, m), 1.20 (3H, s), 1.12 (3H, s), 1.02 (d, $J=11.3$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 208.5, 172.5, 170.7, 167.8, 143.8, 142.2, 120.7, 110.7, 80.0, 78.9, 77.9, 67.2, 65.3, 58.4, 54.2, 50.8, 46.9, 45.7, 38.1, 36.7, 36.2, 30.2, 29.7, 21.9, 21.5, 20.2, 18.0, 17.5.

Fraction 43 yielded compound **50** (2.2 mg). Yellow crystal. LR-LC-MS: m/z 471.2 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.18 (1H, brs), 5.41 (1H, s), 4.31 (1H, d, $J=4.9$ Hz), 3.33 (3H, s), 3.02 (1H, m), 2.58 (1H, m), 2.34 (1H, s), 2.09-0.70 (each 3H, s).

^{13}C NMR (DMSO- d_6 , 500 MHz), δ_{C} : 199.5, 178.1, 170.1, 127.7, 77.0, 61.6, 54.5, 48.5, 45.3, 43.5, 43.4, 41.1, 40.6, 39.2, 39.0, 38.0, 37.1, 32.6, 32.0, 30.8, 28.9, 28.6, 28.3, 27.4, 26.5, 26.3, 23.5, 18.8, 17.6, 16.7, 16.5.

5.4.5.5 Chemical constituents from the Australian plant *Larsenaikia ochreata*

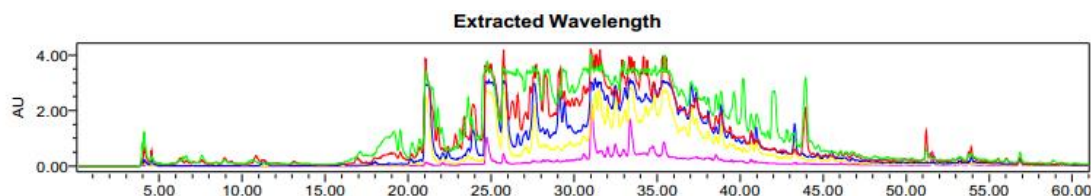


Figure 5.4.5.5.1 HPLC chromatogram of *Larsenaikia ochreata* sp.

The freeze-dried and ground plant material (10 g) from the Australian plant *Larsenaikia ochreata* sp. was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under gravity then discarded. CH_2Cl_2 (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration the biota was extracted with another volume of MeOH (250 mL) whilst being shaken at 200 rpm for 16 h. All CH_2Cl_2 /MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (2.27 g). A portion of this material (1.0 g) was preadsorbed to C18-bonded silica (1 g), and then packed into a stainless steel cartridge (10 × 30 mm) that was subsequently attached to a Thermo Betasil C18 semipreparative HPLC column (150 × 21.2 mm). Isocratic HPLC conditions of 90% H_2O (0.1% TFA)/10% MeOH (0.1% TFA) were employed for the first 10 min; then a linear gradient from 90% H_2O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9 mL/min. Sixty fractions (60 × 1 min) were collected from the start of the HPLC run (Figure 5.4.5.5.1). Looking for the

ion of interest at m/z 811.4227 identified in the ESI-FTMS screen, fractions were analyzed by (\pm)-LRESIMS and NMR spectroscopy.

Fraction 38 yielded compound **51** (1.9 mg). White powder. LR-LC-MS: m/z 363.7 [M+H]⁺. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_{H} : 8.86 (1H, s), 7.54 (1H, s), 7.32 (1H, d, $J=11.9$ Hz), 6.83 (1H, s), 6.52 (1H, s), 6.13 (1H, s), 5.44 (1H, s), 5.29 (1H, s), 1.88 (1H, s), 0.89 (1H, s). ¹³C NMR (DMSO-*d*₆, 500 MHz), δ_{C} : 173.0, 157.4, 152.3, 150.5, 149.7, 146.3, 131.7, 127.6, 127.0, 119.1, 118.0, 116.6, 109.0, 96.9, 72.9, 65.7, 50.7, 30.7, 8.2.

Fraction 42 yielded compound **52** (1.0 mg). Colorless needles. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_{H} : 6.30 (1H, d, $J=7.0$ Hz), 5.42 (1H, brt), 5.13 (1H, d, $J=7.5$ Hz), 4.35 (1H, d, $J=11.0$ Hz), 3.62 (1H, d, $J=11.0$ Hz), 3.59 (1H, dd, $J=11.0, 4.0$ Hz), 1.53 (3H, s), 1.29 (3H, s), 1.06 (3H, s), 0.93 (3H, s), 0.88 (3H, s), 0.76 (3H, s). ¹³C NMR (DMSO-*d*₆, 500 MHz), δ_{C} : 176.8, 173.1, 144.3, 123.1, 106.6, 95.6, 89.1, 78.6, 78.3, 78.2, 75.4, 73.8, 73.6, 71.0, 63.4, 62.3, 56.2, 48.0, 47.2, 46.3, 44.5, 42.3, 41.9, 40.0, 38.6, 36.8, 34.1, 33.5, 32.6, 30.8, 28.5, 27.0, 26.1, 23.8, 23.5, 24.2, 19.0, 17.4, 15.6.

Fraction 44 yielded compound **53** (2.4 mg). White powder. LR-LC-MS: m/z 355.3 [M+H]⁺. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_{H} : 7.35 (1H, d, $J=7.3$ Hz), 7.21 (1H, d, $J=7.2$ Hz), 6.93 (1H, m), 6.91 (1H, m), 4.20 (1H, m), 3.71 (3H, s), 3.62 (1H, m). ¹³C NMR (DMSO-*d*₆, 500 MHz), δ_{C} : 172.7, 136.9, 129.7, 126.2, 122.4, 119.7, 118.6, 112.0, 105.9, 66.6, 60.8, 57.6, 52.3, 51.8, 51.5, 37.7, 34.3, 32.4, 31.5, 22.4, 19.1.

Fraction 45 yielded compound **54** (2.6 mg). Yellow crystal powder. LR-LC-MS: m/z 377.3 [M+H]⁺. ¹H NMR (DMSO-*d*₆, 500 MHz), δ_{H} : 8.29 (1H, d, $J=8.2$ Hz), 8.18 (1H, d, $J=8.2$ Hz), 7.86 (1H, t, $J=7.2$ Hz), 7.74 (1H, t, $J=7.2$ Hz), 7.34 (1H, s), 6.53 (1H, s), 5.45 (2H, s), 5.33 (2H, s), 3.24 (2H, q, $J=7.5$ Hz), 1.88 (2H, m), 1.33 (3H, t, $J=7.6$ Hz), 0.89 (3H, t, $J=7.2$ Hz). ¹³C NMR (DMSO-*d*₆, 500 MHz), δ_{C} : 173.0, 157.3, 152.4, 150.5, 149.0, 146.5, 146.0, 130.4, 130.3, 128.5, 128.0, 127.0, 124.5, 119.4, 97.0, 72.8, 65.7, 50.0, 30.7, 22.7, 14.2, 8.2.

Fraction 55 yielded compound **55** (2.8 mg). White powder. LR-LC-MS: m/z 455.4 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.06 (1H, s), 4.69 (1H, d, $J=2.3$ Hz), 4.57 (1H, s), 4.27 (1H, d, $J=5.2$ Hz), 2.97 (2H, m), 2.23 (1H, t, $J=10.3$ Hz), 2.12 (1H, d, $J=9.0$ Hz), 1.81 (1H, m), 1.65 (3H, s), 1.49 (3H, m), 1.33 (3H, m), 0.94 (3H, s), 0.87 (3H, d, $J=1.7$ Hz), 0.77 (3H, s), 0.66 (3H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 177.7, 150.8, 110.1, 77.2, 55.9, 55.3, 50.4, 49.0, 47.1, 42.5, 40.7, 39.0, 38.7, 38.1, 37.2, 36.8, 34.4, 32.2, 30.6, 29.7, 28.6, 27.6, 25.5, 20.9, 19.4, 18.4, 16.4, 16.3, 16.2, 14.9.

Fraction 57 yielded compound **56** (2.1 mg). White powder. LR-LC-MS: m/z 427.4 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 4.69 (1H, d, $J=2.4$ Hz), 4.55 (1H, d, $J=1.1$ Hz), 4.27 (1H, d, $J=5.2$ Hz), 2.98 (1H, dt, $J=5.5, 10.4$ Hz), 2.38 (1H, dt, $J=5.8, 10.9$ Hz), 1.88 (1H, m), 0.66-1.65 (each 3H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 150.7, 110.1, 77.2, 55.3, 50.4, 48.2, 47.9, 43.0, 42.8, 40.8, 40.5, 39.0, 38.8, 38.0, 37.1, 35.5, 34.3, 29.7, 28.6, 27.6, 27.5, 25.1, 20.9, 19.4, 18.4, 18.2, 16.4, 16.3, 16.2, 14.8.

5.4.6 Dissociation constant (K_D) determination

For a protein with a single and specific binding site for a ligand, like *PfRab11a*, the equilibrium dissociation constant (K_D) of the protein-ligand complex is given by:

$$K_D = \frac{[P][L]}{[PL]} \quad (1)$$

Where [P] and [L] are the free concentrations of protein and ligand, respectively, and [PL] is the concentration of the binary protein-ligand complex. As the only known quantities were the total concentration of added protein and ligand, the mass conservation law was applied and the equilibrium dissociation equation rewritten as:

$$K_D = \frac{([P]_t - [PL])([L]_t - [PL])}{[PL]} \quad (2)$$

Where $[P]_t$ and $[L]_t$ were the total concentrations of protein and ligand, respectively. The equation was multiplied on both sides by [PL] and subtracted on both sides by $K_D[PL]$. After rearrangement, the following was obtained:

$$0 = [PL]^2 - ([P]_t + [L]_t + K_D)[PL] + [P]_t[L]_t \quad (3)$$

This quadratic equation would have two possible solutions. However, only one was physically meaningful:

$$[PL] = \frac{([P]_t + [L]_t + K_D) - \sqrt{([P]_t + [L]_t + K_D)^2 - 4[P]_t[L]_t}}{2} \quad (4)$$

In a typical titration experiment, the concentration of the protein is kept constant and the concentration of its ligand is increased until full saturation of the protein is achieved. To determine the K_D , the relative abundances of bound to total protein in the mass spectra were correlated to the relative equilibrium concentrations of bound to total protein in solution, as follows:

$$\frac{\sum I(PL)^{n+}/n}{\sum I(P)^{n+}/n + \sum I(PL)^{n+}/n} = \frac{[PL]}{[P]_t} \quad (5)$$

Where the total abundances of free protein and protein-ligand complex were obtained by the sum of all the respective charges states intensity peaks (I) normalized for charge state (n). By plotting experimentally observed ratios between bound and total protein ion abundances against total concentration of ligand, K_D could be obtained as a parameter of a non-linear least squares curve fitting:

$$\frac{\sum I(PL)^{n+}/n}{\sum I(P)^{n+}/n + \sum I(PL)^{n+}/n} = \frac{([P]_t + [L]_t + K_D) - \sqrt{([P]_t + [L]_t + K_D)^2 - 4[P]_t[L]_t}}{2[P]_t} \quad (6)$$

PfRab11a (4.5 μ M in 10 mM ammonium acetate buffer, pH 6.8, 50 μ L) and increasing concentrations of ligand (2 μ M, 3 μ M, 4 μ M, 8 μ M, 40 μ M and 50 μ M, 5 μ L each) was incubated at 4 °C for 1 hr before analyzed by ESI-FTMS. Triplicate values of titration were used for the calculation of K_D . Prism software (Version 6.01) was used to plot the fraction of bound protein against the total ligand concentration of each ligand.

5.4.7 Enzyme assay

PfRab11a was cloned from *P. falciparum* cDNA, produced and purified as previously described.²²¹ The protein was dissolved in ammonium acetate (10 mM, pH 6.8) to generate stock solution (4.5 μ M).

PfRab11a activity was measured with a colorimetric GTPase assay kit purchased from Innova Biosciences (BioNovus Life Science) following the manufacturer's protocol. Different concentrations of five ligands (0.01 μ M, 1.0 μ M, 1.5 μ M, 4.5 μ M, 10 μ M, 20.0 μ M, 30.0 μ M, 40.0 μ M, 50.0 μ M, 80.0 μ M and 100.0 μ M) were incubated respectively with *PfRab11a* (4.5 μ M) for 1 hr; GTP (10 μ M) was added and the mixture incubated for 30 min at 4 °C. Colorimetric buffer (Pi ColorLock Gold and Accelerator) was added and the mixture was incubated for 5 min at 4 °C. Then, stabilizer buffer was added, and the reaction mixture was incubated for 30 min. Optical density was determined at 620 nm by using an EnVision Multilabel Reader (PerkinElmer). The experiment was run in triplicate.

5.4.8 Virtual docking

All virtual docking applications were performed using Schrödinger Suite 2012 (LLC, New York, NY, USA) on a Linux platform. The following Schrödinger modules were used: Protein Preparation Wizard (Epik Version 2.2, Impact version 5.7, Prime version 2.3, Schrödinger, New York, NY, USA, 2011), LigPrep (Version 2.5, Schrödinger, New York, NY, USA, 2011), SiteMap (Version 2.5, Schrödinger, New York, NY, USA, 2011) and Glide²²²⁻²²⁴. All modules were accessed via Maestro graphical interface (Version 9.2, Schrödinger, New York, NY, USA, 2011).

Receptor Protein Preparation. The PDB is used to download the target *PfRab11a* file (3BFK). The crystal structure of *PfRab11a* was refined using the Protein Preparation Wizard in the Maestro suite from Schrödinger (Version 9.2, New York, NY, USA, 2011). Hydrogen atoms were added to the heavy atoms and all waters were deleted. Using force field OPLS 2001, minimization was carried out setting maximum heavy atom RMSD to 0.30Å.

Receptor Grid Generation. After preparation, receptor grids were generated with Glide by specifying the binding site with a 3D cubic box. SiteMap was used to estimate the location of the active site by searching regions near the protein surface, generating hydrophobic and hydrophilic contour maps of the protein, and calculating

energy potentials. Rotation of all receptor hydroxyl and thiol groups within the grid was allowed.

Ligand preparation. The small molecules identified as binders to *PfRab11a* were drawn in Maestro workspace, and these 2D structures were converted to all-atom 3D structures using embedded LigPrep script.²²⁵ Up to 32 stereoisomers were generated per ligand by determining chiralities of 3D structures, and all possible ionization states of ligands were generated at target pH of 7+/- 2. After converting to 3D, conformation search was carried out to generate conformers and search for low energy structures using OPLS 2005 force field.

Glide Docking. The docking program used was Glide. In this study, Glide-SP (standard precision) was used for docking simulations. The Glide algorithm utilizes precomputed grids generated by sitemap. No bonding constraints were given during docking calculations. Using Monte Carlo random search algorithm, ligand poses were generated for each input molecule. Binding affinity of these molecules to receptors was predicted in terms of Glide docking score. Post docking minimization was performed with OPLS 2005 force field, and one pose per ligand was saved. In the present study, we have utilized PyMol (version 1.5.0.4; Schrödinger, LLC: New York, 2012) to analyze the binding sites.

CHAPTER SIX SCREENING LEAD-LIKE ENHANCED EXTRACTS AGAINST PFUCH-L3 USING ESI-FTMS

Abstract: This chapter presents a strategy for using ESI-FTMS to detect the complex formed between PfUCH-L3 and its ligand in natural product extract matrix. Lead-like enhanced extracts were incubated with PfUCH-L3 and then analyzed by ESI-FTMS. 3,080 extracts were screened against PfUCH-L3 and three noncovalent complexes were identified. Mass-directed isolation and purification gave three natural products as active compounds. Subsequently, their binding to PfUCH-L3 was confirmed by ESI-FTMS. The quantification of each ligand binding to PfUCH-L3 was illustrated by measuring the dissociation constant (K_D) for the PfUCH-L3-ligand complex. Simultaneously, a number of natural products from the samples, which gave positive FTMS screening results, was isolated and purified based on spectroscopic data analysis to investigate the connection between Log P value and HPLC retention time.

6.1 Introduction

This chapter describes the use of ESI-FTMS to detect noncovalent complexes and identify active compounds in natural product extracts. The target selected for this study was *Plasmodium falciparum* ubiquitin carboxyl terminal hydrolase (PfUCH-L3, PF14_0576) with the molecular weight of 25 kDa.

ESI-FTMS was used to screen lead-like enhanced extracts for noncovalent complexes formed between the protein target and its ligand. 3,080 natural product extracts were incubated with PfUCH-L3 and analyzed by ESI-FTMS. Three active extracts were identified. Three natural products isolated and purified from these extracts were confirmed to bind to PfUCH-L3. The K_D values of the active compounds were measured based on the titration experiments, and their binding sites were predicted according to a virtual docking method. Additional compounds were also isolated from the active extracts during the process, their structures were determined based

on spectroscopic data analysis with the purpose to study their physicochemical properties especially the relationship between Log *P* and HPLC retention time.

6.2 Results and discussion

6.2.1 Detection of protein by ESI-FTMS in physiological pH

PfUCH-L3 was dissolved in ammonium acetate buffer solution (10 mM, pH 6.8) to give a protein solution (4.5 μ M). *PfUCH-L3* ammonium acetate solution (50 μ L) was injected into mass spectrometer to give peaks at charge states of 10+, 11+ and 12+ with the major peak at the charge state of 11+ (Figure 6.2.1.1). *PfUCH-L3* has the molecular weight of 25158.6 Da, whilst three peaks at *m/z* 2317.2169, 2527.7498, and 2780.6067 were observed from MS spectrum, indicating a 10+, 11+, and 12+ charge state of the protein.

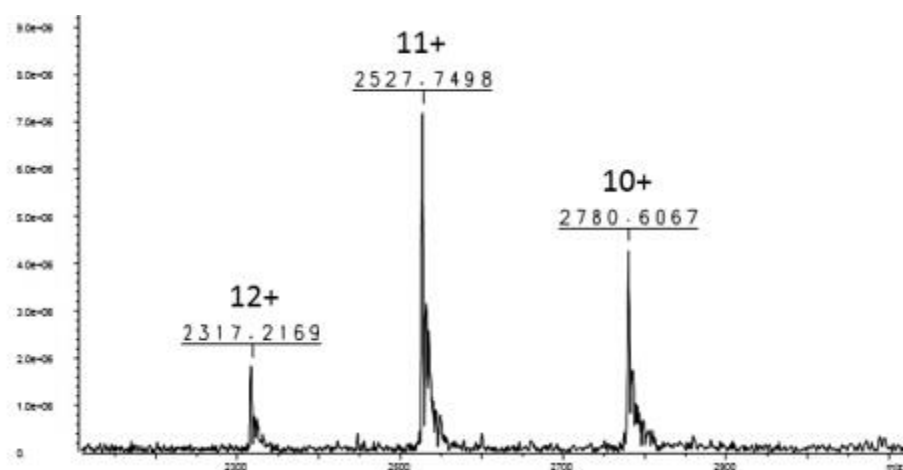


Figure 6.2.1.1 Mass spectrum of *PfUCH-L3* acquired under native condition

6.2.2 Detection of complexes formed by protein and active components in natural product extracts

3,080 extracts were screened against *PfUCH-L3* by FT-MS. Each extract (250 μ g/ μ L, 5 μ L) was incubated with *PfUCH-L3* (4.5 μ M, 50 μ L) at what temperature for how long. A portion of the protein-extract mixture (55 μ L) was directly infused to the ESI-FTMS. 98.4% of incubated samples showed signals of native protein or complexes.

Of the 3080 extracts, three samples showed signals of *Pf*UCH-L3-ligand noncovalent complexes in MS spectrum (Figure 6.2.2.1-6.2.2.3). Their mass spectra showed peaks of complexes (Complex One to Three) with the major peak at charge state 11+. The masses of the charge states, from 10+ to 12+, confirm the molecular mass of the protein to be $25,157 \pm 0.5$ Da were in highly agreement for low-resolution mode with the calculated mass (25,158 Da) obtained from the amino acid composition. It was observed that the complex remained intact and there was still free protein in the assay solution. The experiments were performed in a low-resolution mode for a quick turnaround time (3-5 minutes per run).

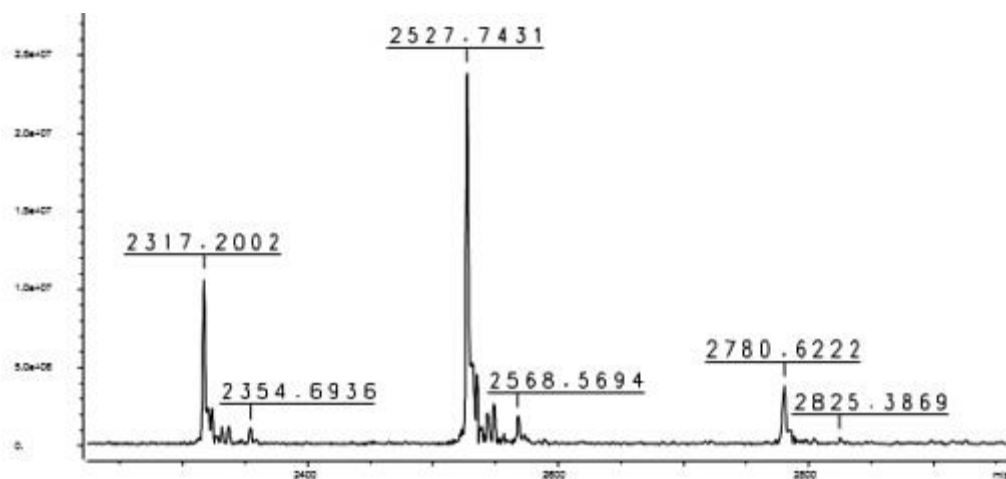


Figure 6.2.2.1 Mass spectrum of *Pf*UCH-L3-ligand Complex One

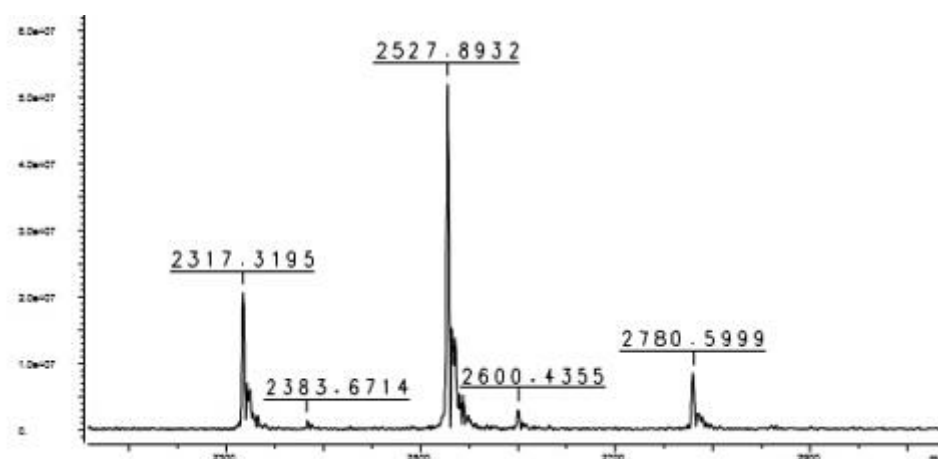


Figure 6.2.2.2 Mass spectrum of *Pf*UCH-L3-ligand Complex Two

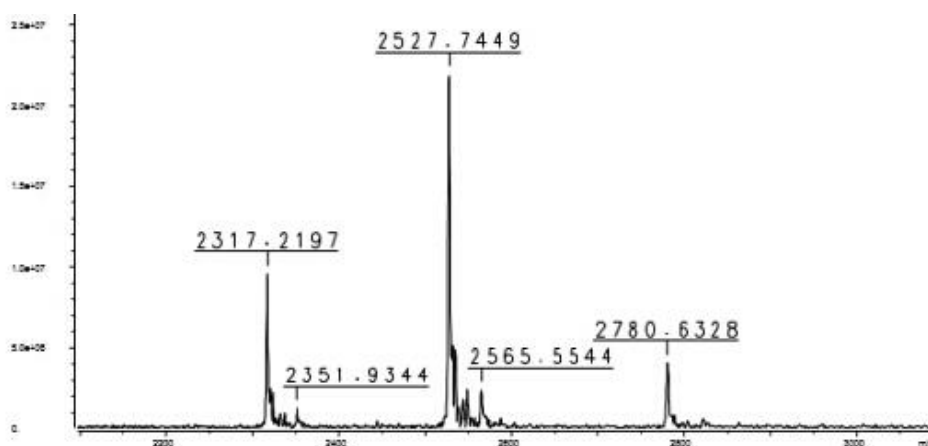


Figure 6.2.2.3 Mass spectrum of *PfUCH-L3-ligand Complex Three*

6.2.3 Determination of ligand molecular mass

The molecular mass of the ligand can be deduced from the mass-to-charge ratio (m/z) of the protein-ligand complex.

For example, in Figure 6.3.2.1, each charge state consisted of a grouping of two peaks: one peak at a lower m/z value corresponded to free *PfUCH-L3*, and the other peak at higher m/z corresponded to *PfUCH-L3-ligand complex*. The ligand mass is the difference between the average molecular mass of the complex and the average molecular mass of *PfUCH-L3*. Using the following equation, the molecular weight of ligand one was calculated as 449.9208 Da.¹⁴³

$$MW_{\text{Ligand 1}} = [m/z (\text{PfUCH-L3-natural product ligand complex}) - m/z (\text{free PfUCH-L3})] \times (z) = [2354.6936 - 2317.2002] \times 12 = 449.9208 \text{ Da}$$

The molecular weight of the other two ligands (Ligand 2 and Ligand 3) was obtained in the same procedure as follows.

$$MW_{\text{Ligand 2}} = [m/z (\text{PfUCH-L3-natural product ligand complex}) - m/z (\text{free PfUCH-L3})] \times (z) = [2383.6714 - 2317.3195] \times 12 = 796.2228 \text{ Da}$$

$$MW_{\text{Ligand 3}} = [m/z (\text{PfUCH-L3-natural product ligand complex}) - m/z (\text{free PfUCH-L3})] \times (z) = [2351.9344 - 2317.2197] \times 12 = 416.5764 \text{ Da}$$

In order to obtain sufficient quantity of the ligands for further study, mass-directed isolation and purification was undertaken to isolate these active natural products

from the lead-like enhanced extracts, which led to the identification of sinensin (Ligand 1, MW 449.9; compound **58**) from Australian plant *Prunus Rosaceae*, ianthesine E (Ligand 2, MW 796.2; compound **63**) from Australian sponge *Ianthella quadrangulata*, and gominsin A (Ligand 3, MW 416.6; compound **68**) from Australian plant *Canthium Rubiaceae*.

6.2.4 Structure elucidation of natural products and the analysis of their physicochemical properties

6.2.4.1 Structure elucidation

Ligand 1: MW 450.4

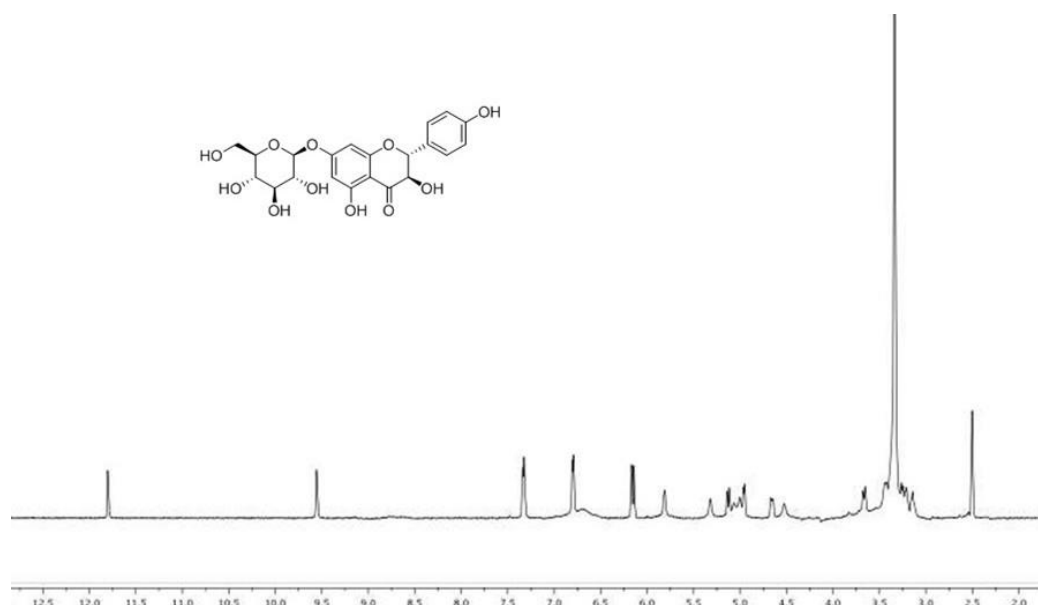


Figure 6.2.4.1.1 Chemical structure and ^1H NMR spectrum of **58**

Compound **58** was identified from the Australian plant *Prunus Rosaceae*. It was isolated as colorless solid. Its molecular formula was determined to be $\text{C}_{21}\text{H}_{22}\text{O}_{11}$ on the basis of (+)-HRESIMS measurements $\{[\text{M}-\text{H}]^+ \text{ ion at } m/z \text{ 449.4147}\}$, which was consistent with 11 degrees of unsaturation. The ^1H NMR displayed aromatic protons and a sugar moiety (Figure 6.2.4.1.1). The ^{13}C NMR spectrum of **58** showed a total of 21 resonances. Its molecular weight and genus name were searched in the DNP, and 95 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **58** was determined as sinensin.²²⁶

Ligand 2: MW 796.2

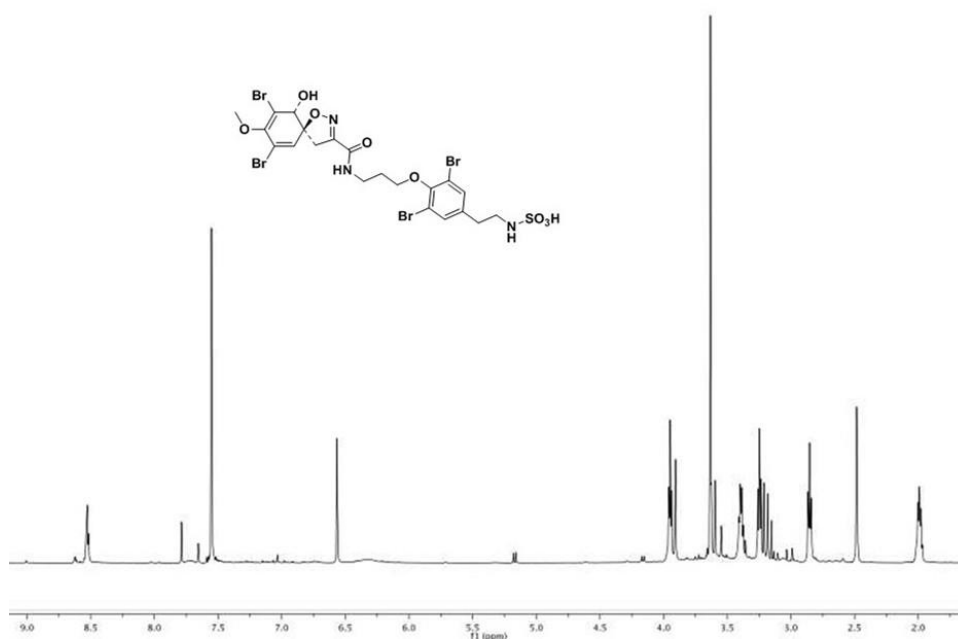


Figure 6.2.4.1.2 Chemical structure and ¹H NMR spectrum of **63**

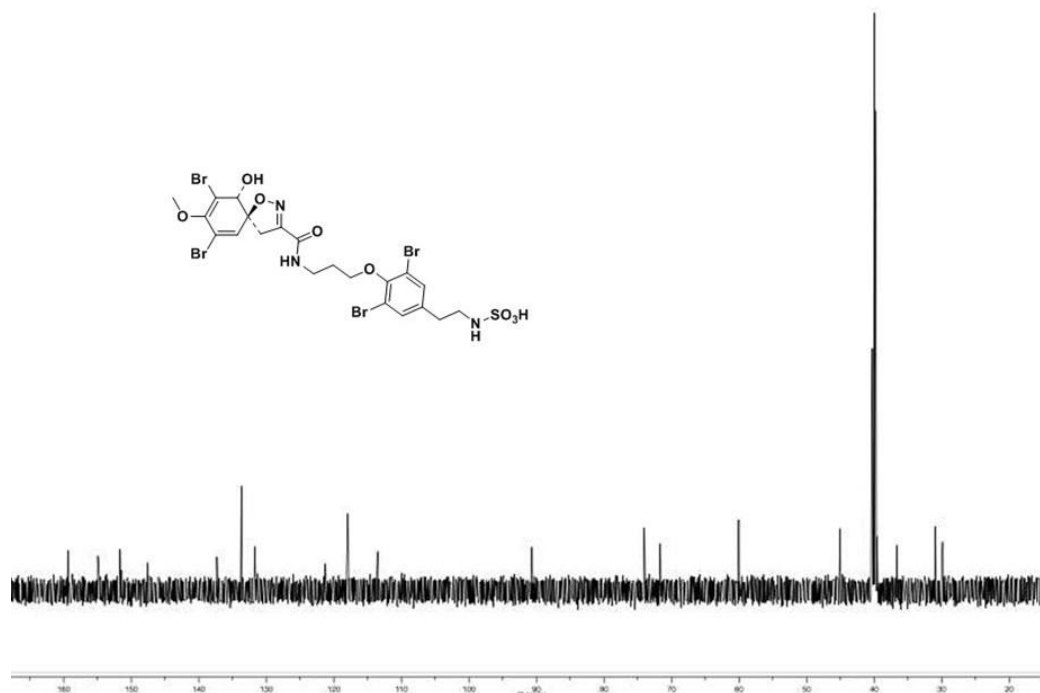


Figure 6.2.4.1.3 Chemical structure and ¹³C NMR spectrum of **63**

Compound **63** was identified from the Australian sponge *Ianthella quadrangulata*. It was obtained as yellow needle-like crystals. Its molecular formula was determined to be C₂₁H₂₃Br₄N₃O₈S on the basis of (+)-HRESIMS measurements {[M-H]⁺ ion at *m/z* 795.7819}, which was consistent with 10.5 degrees of unsaturation. The ¹H NMR spectrum (Figure 6.2.4.1.2) showed a single amide -NH shift at δ_H 8.53 integrating for

one proton, a singlet at δ_H 6.57, three doublets at δ_H 6.35, 3.21, and 3.62; a singlet at δ_H 7.50 suggested the presence of a 1,2,4,6-tetrasubstituted aromatic ring. The presence of a sulphate group was supported by a cluster center with five isotope peaks around m/z 795.7819 in the high resolution ESI mass spectrum. The ^{13}C NMR spectrum of **63** showed a total of 21 resonances (Figure 6.2.4.1.3). Its molecular formula and genus name were searched in the DNP, and one hit was found. By comparing their 1H and ^{13}C NMR data, compound **63** was determined as ianthesine E.²²⁷ Ianthesine E was reported to have binding activity towards the adenosine A₁ receptor.²²⁷

Ligand 3: MW 416.4

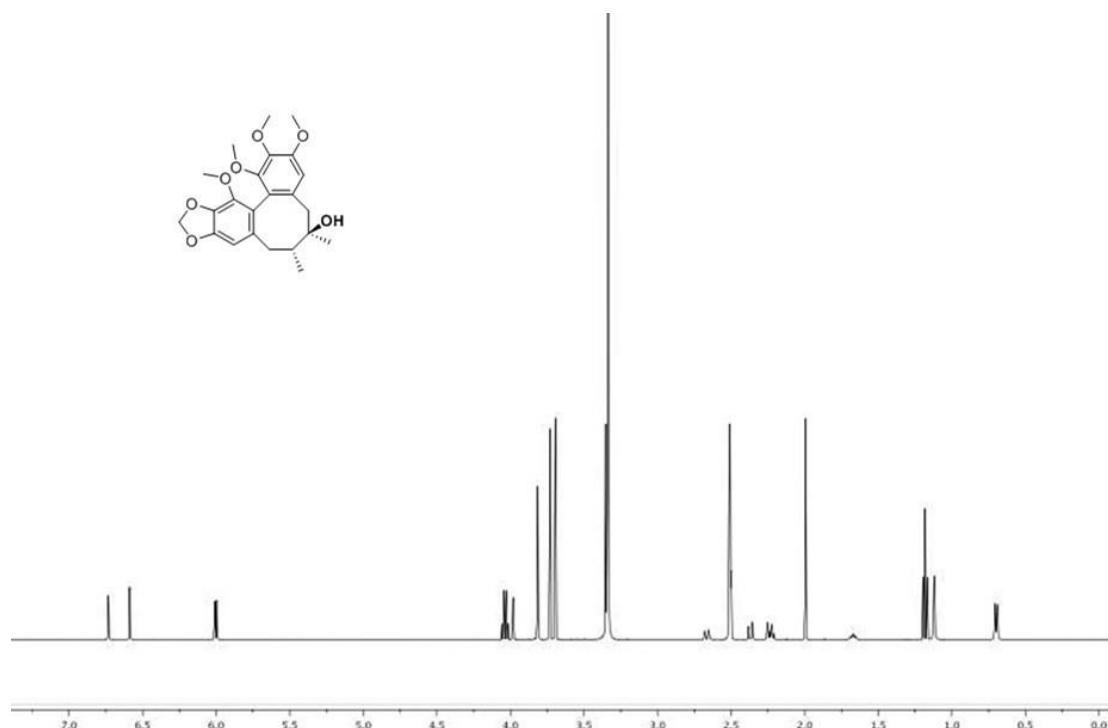


Figure 6.2.4.1.4 Chemical structure and 1H NMR spectrum of **68**

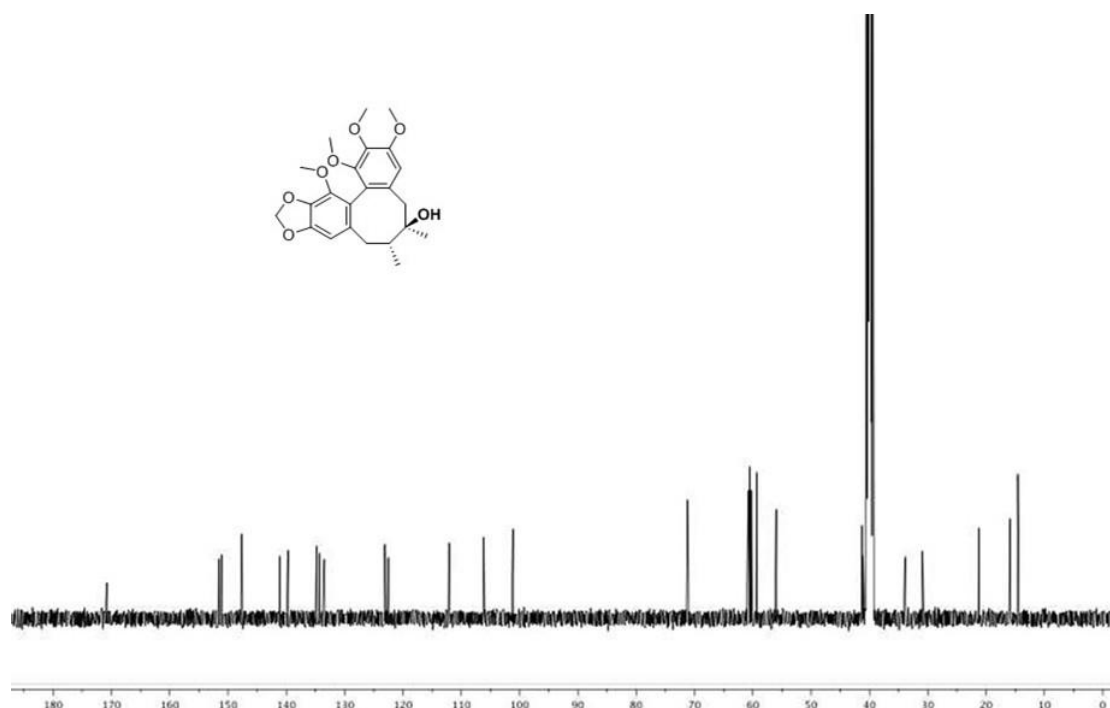


Figure 6.2.4.1.5 Chemical structure and ^{13}C NMR spectrum of **68**

Compound **68** was identified from the Australian plant *Canthium Rubiaceae*. It was obtained as colorless needles. Its molecular formula was determined to be $\text{C}_{23}\text{H}_{28}\text{O}_7$ on the basis of (+)-HRESIMS measurements $\{[\text{M}-\text{H}]^+ \text{ ion at } m/z \text{ 415.3408}\}$, which was consistent with 10 degrees of unsaturation. The ^1H NMR spectrum (Figure 6.2.4.1.4) displayed the presence of one methyl at δ_{H} 0.81 (3H, d, $J=7.5$ Hz), methoxys (δ_{H} 3.51–3.90), two aromatic protons [δ_{H} 6.47 (1H, s), 6.62 (1H, s)] and other proton signals. The ^{13}C NMR spectrum of **68** (Figure 6.2.4.1.5) showed a total of 23 resonances, including three methyls (δ_{C} 40.2, 30.1, 15.8), two methylenes (δ_{C} 40.5, 33.7), one methylene dioxy (δ_{C} 100.7), six oxygenated aromatic quaternary carbons (δ_{C} 152.1, 151.9, 147.2, 141.1, 140.6, 134.5), two aromatic carbons (δ_{C} 110.2, δ_{C} 105.8), four methoxys (δ_{C} 61.0, 60.5, 59.6, 55.9), and another three quaternary carbons (δ_{C} 124.0, 121.7, 71.6). Its molecular formula and genus name were searched in the DNP, and 80 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **68** was determined as gomisin A.^{172, 228}

Additional compounds were also identified from the purification process. They included: compounds **57** and **59-62** from the Australian plant *Prunus Rosaceae*; compounds **64-67** and **69** from the Australian plant *Canthium Rubiaceae*.

Compound **57**

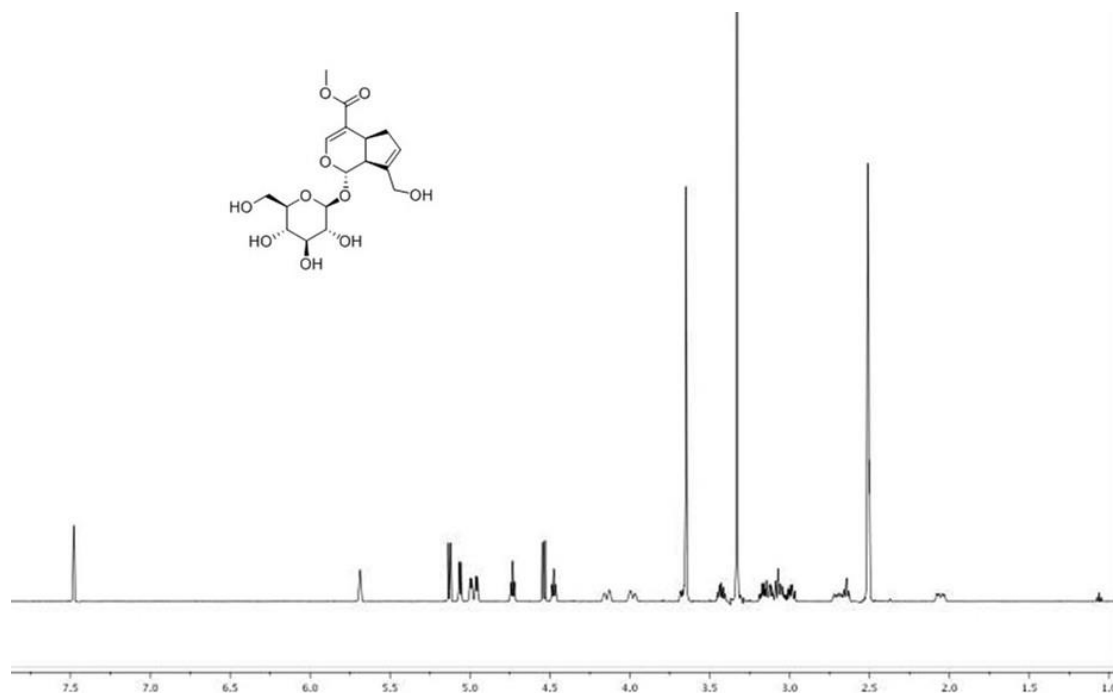


Figure 6.2.4.1.6 Chemical structure and ^1H NMR spectrum of **57**

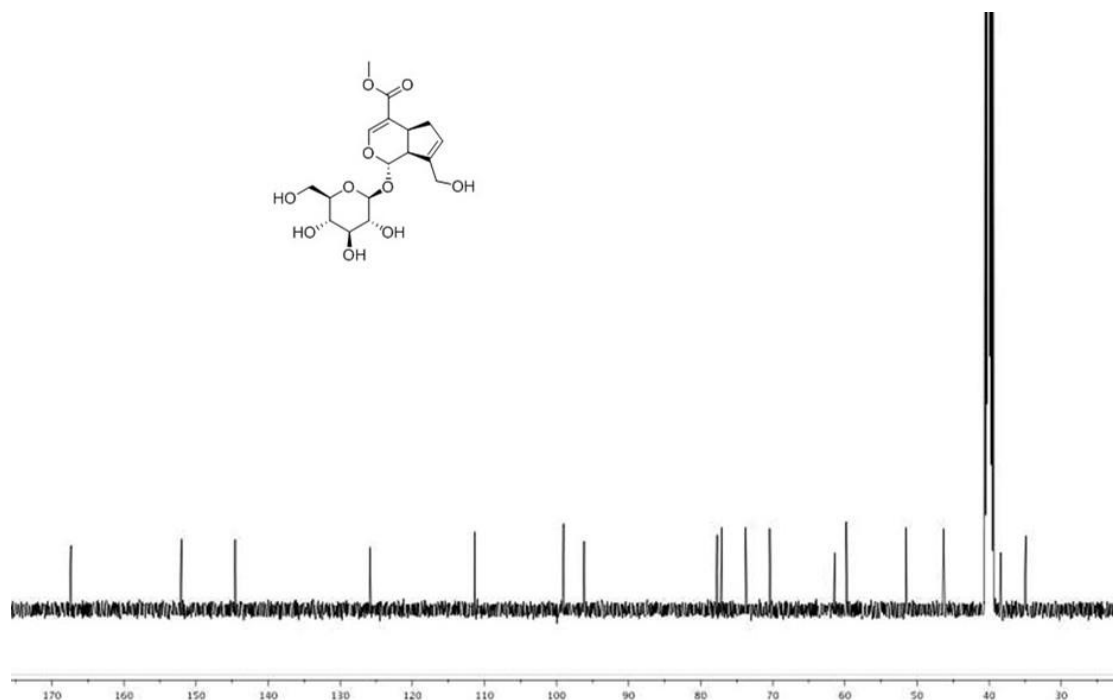


Figure 6.2.4.1.7 Chemical structure and ^{13}C NMR spectrum of **57**

Compound **57** was obtained as light brown powder. Its molecular formula was determined to be $C_{17}H_{24}O_{10}$ on the basis of (+)-HRESIMS measurements $\{[M+H]^+$ ion at m/z 389.3245}, which was consistent with 6 degrees of unsaturation. The 1H NMR spectrum (Figure 6.2.4.1.6) displayed sugar signals between δ_H 4.54 and 3.21, one methoxyl at δ_H 3.65 and other proton signals. The ^{13}C NMR spectrum of **57** (Figure 6.2.4.1.7) showed a total of 17 resonances, including two ethylenes (δ_C 152.0, 111.4; δ_C 144.6, 125.9), one α -carbonyl group at δ_C 167.4, one methoxyl at δ_C 34.9, three methines (δ_C 96.2, 46.3, 51.5), one methylene at δ_C 38.4, six sugar carbons (δ_C 99.0, 77.7, 77.1, 73.8, 70.4 and 61.4), and one hydroxymethyl group at δ_C 59.8. Its molecular formula and genus name were searched in the DNP, and 41 hits were found. By comparing their 1H and ^{13}C NMR data, compound **57** was determined as geniposide.^{144, 229} According to the study by Cai *et al*, geniposide has showed antidepressant-like effect on chronic unpredictable mild stress-induced depressive rats.²³⁰

Compound **59**

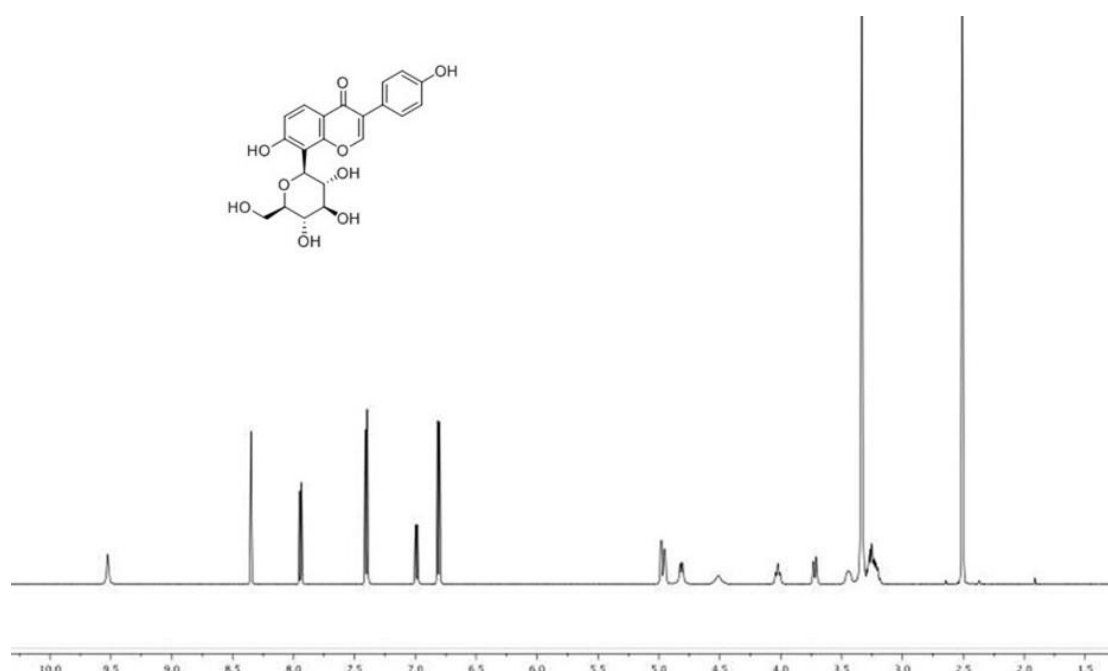


Figure 6.2.4.1.8 Chemical structure and 1H NMR spectrum of **59**

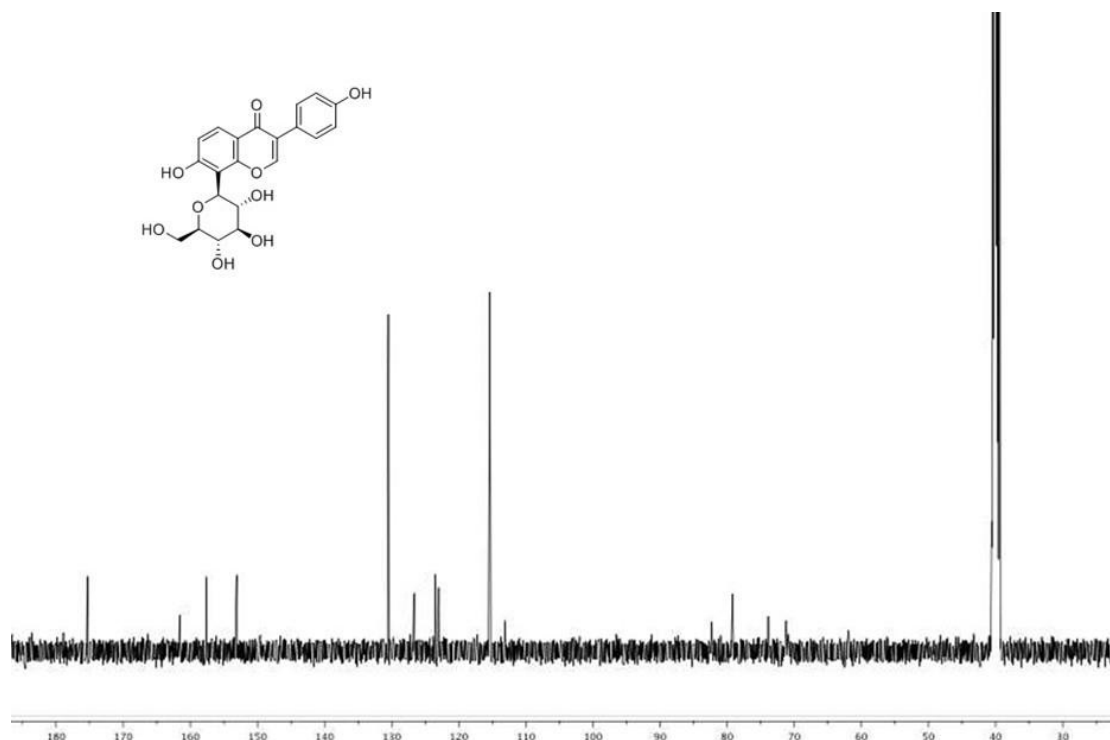


Figure 6.2.4.1.9 Chemical structure and ^{13}C NMR spectrum of **59**

Compound **59** was obtained as light yellow crystals. Its molecular formula was determined to be $\text{C}_{21}\text{H}_{20}\text{O}_9$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+$ ion at m/z 417.2063}, which was consistent with 12 degrees of unsaturation. The ^1H NMR (Figure 6.2.4.1.8) showed a sugar moiety (δ_{H} 4.81, 3.20-3.34), six aromatic signals, and other proton signals. The ^{13}C NMR spectrum of **59** showed a total of 21 resonances (Figure 6.2.4.1.9). When hydrochloric acid and magnesium was added to compound **59**, the color of solution became light yellow. This reaction confirmed that compound **59** was one of *iso*-flavonoids. Its molecular formula and genus name were searched in the DNP, and 80 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **59** was determined as puerarin.²³¹⁻²³² Puerarin has shown anti-oxidative and anti-inflammatory activities.²³³

Compound **60**

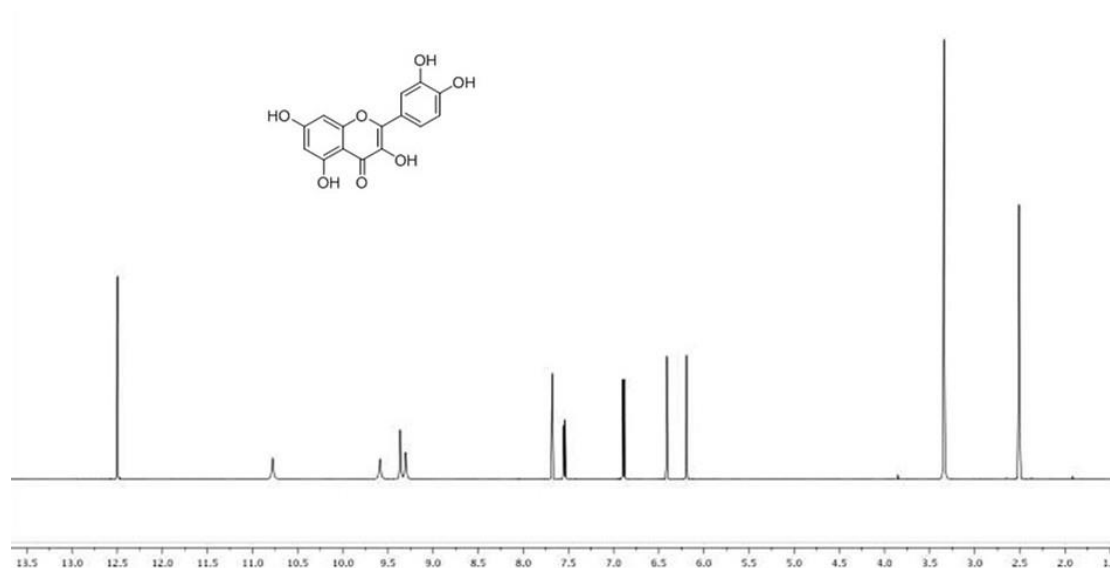


Figure 6.2.4.1.10 Chemical structure and ^1H NMR spectrum of **60**

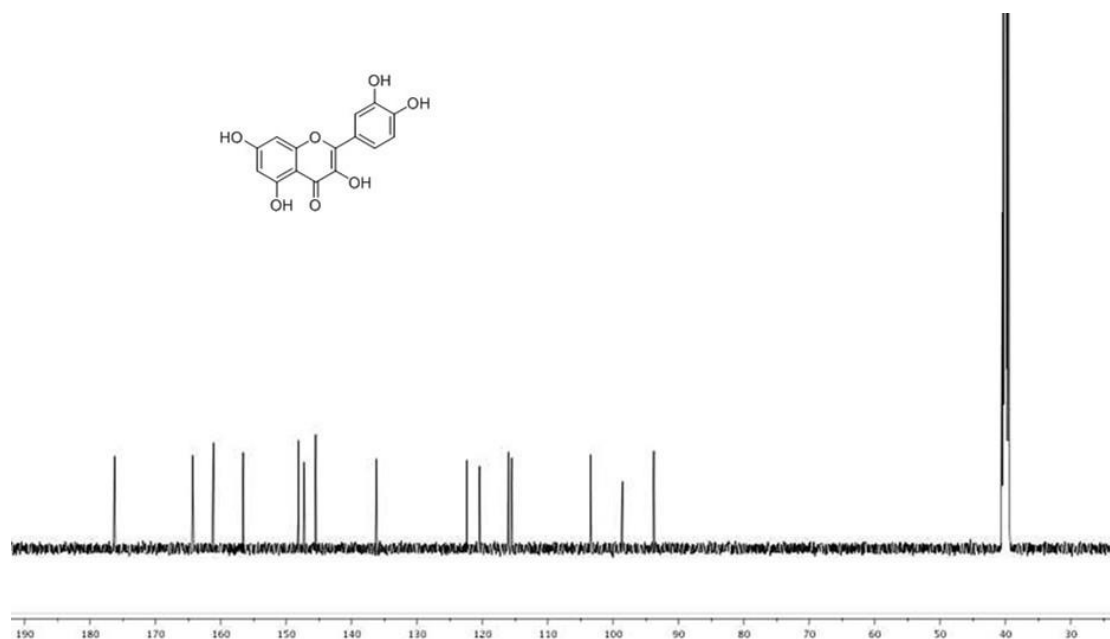


Figure 6.2.4.1.11 Chemical structure and ^{13}C NMR spectrum of **60**

Compound **60** was obtained as yellow needle-like crystals. Its molecular formula was determined to be $\text{C}_{15}\text{H}_{10}\text{O}_7$ on the basis of (+)-HRESIMS measurements $\{[\text{M}]^+ \text{ ion at } m/z \text{ 302.2085}\}$, which was consistent with 11 degrees of unsaturation. The ^1H and ^{13}C NMR spectra of compound **60** exhibited resonances due to aromatic systems. In the ^1H NMR spectrum (Figure 6.2.4.1.10), the aromatic region exhibited an ABX system at δ_{H} 7.66 (1H, d, $J=2.5$ Hz), 7.54 (1H, dd, $J=2.0, 8.5$ Hz), and 6.88 (1H, d, $J=8.5$ Hz) and a typical *meta* coupled pattern for two aromatic doublets (6.19 and 6.41, d,

$J=2.0$ Hz). The ^{13}C NMR spectrum displayed a total of 15 carbons (Figure 6.2.4.1.11). When hydrochloric acid and magnesium was added to compound **60** (dissolved in methanol), the solution became orange, which indicated that **60** was a flavonoid. Its molecular formula and genus name were searched in the DNP, and 128 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **60** was determined as quercetin.^{156, 234} Gordon *et al* investigated the antioxidant activity of quercetin and found out that the activity of quercetin increased more strongly with concentration and it was very effective at a concentration of 5×10^{-2} mol/mol phospholipid.²³⁵

Compound **61**

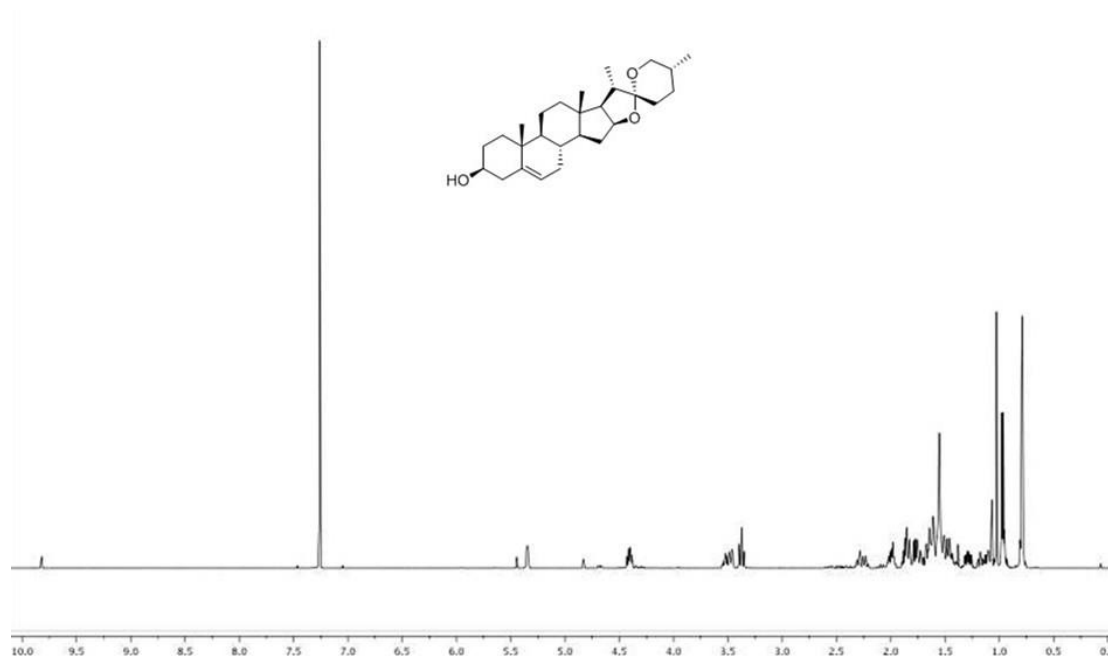


Figure 6.2.4.1.12 Chemical structure and ^1H NMR spectrum of **61**

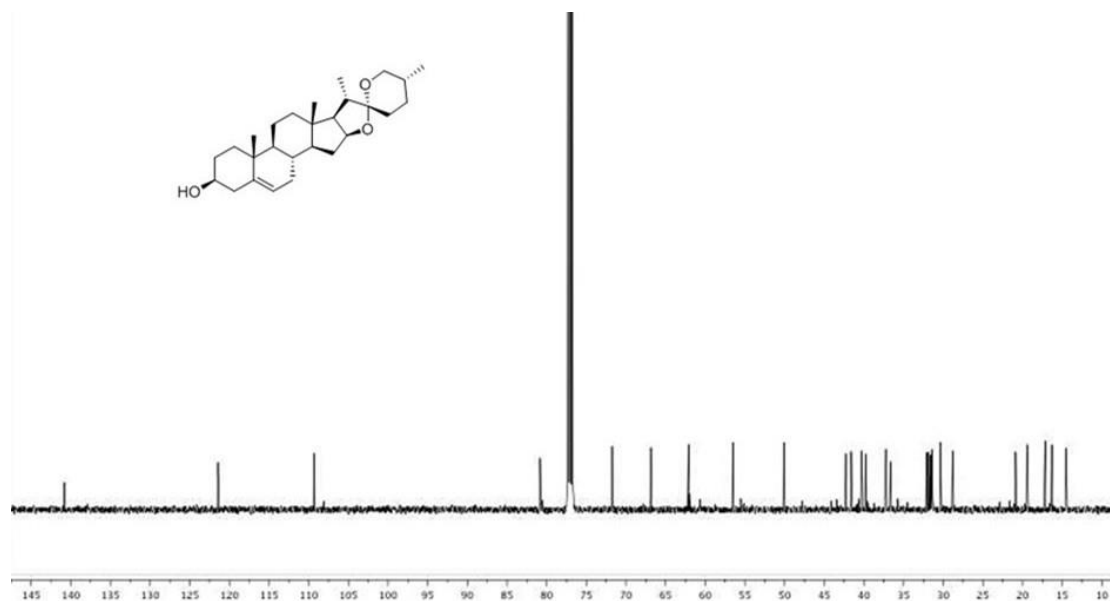


Figure 6.2.4.1.13 Chemical structure and ^{13}C NMR spectrum of **61**

Compound **61** was obtained as white powder. Its molecular formula was determined to be $\text{C}_{27}\text{H}_{42}\text{O}_3$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 413.4952, which was consistent with 7 degrees of unsaturation. The ^1H NMR spectrum displayed four methyls, and other proton signals (Figure 6.2.4.1.12). The ^{13}C NMR spectrum displayed a total of 27 resonances (Figure 6.2.4.1.13). When acetic anhydride and concentrated sulfuric acid (20:1) was added to compound **61**, the solution became green and then faded (Liebermann-Burchard reaction). This evidence indicated that compound **61** was a steroid. Its molecular formula and genus name were searched in the DNP, and 131 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **61** was determined as diosgenin.²³⁶⁻²³⁷ Diosgenin has showed anti-inflammatory activity, inhibiting the growth of fibroblast-like synoviocytes from human rheumatoid arthritis 2.²³⁸ Diosgenin has also exhibited anticancer and antitumor activities.²³⁸

Compound **62**

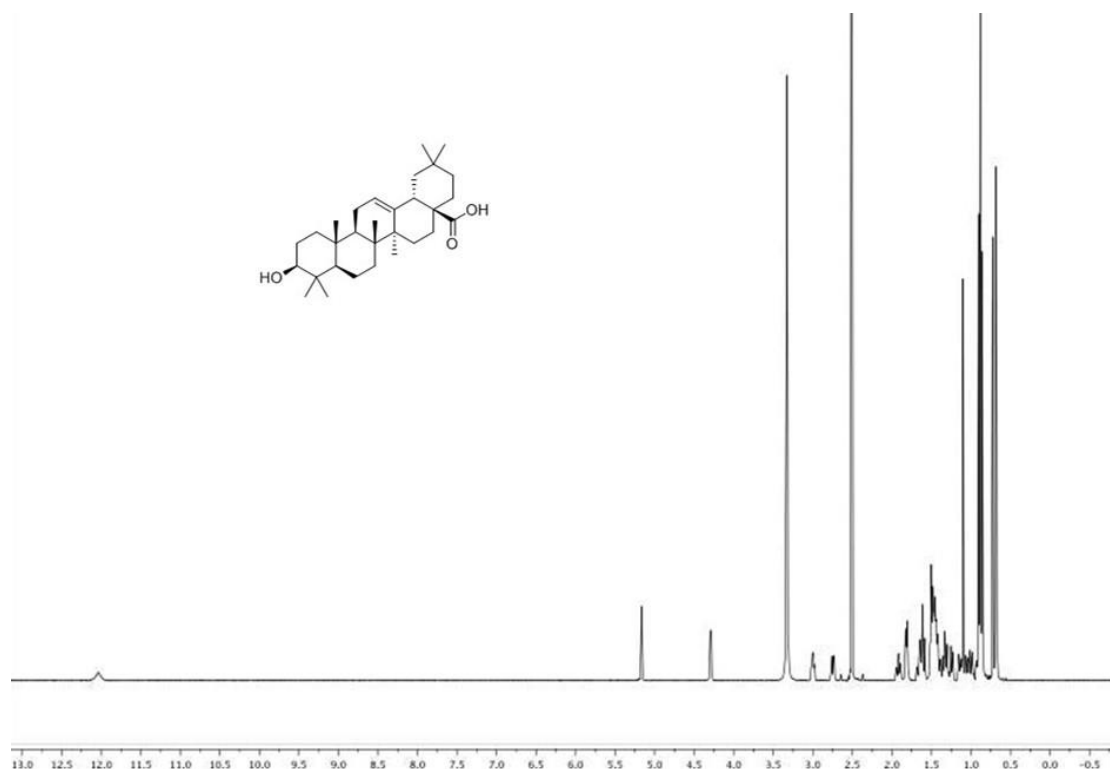


Figure 6.2.4.1.14 Chemical structure and ¹H NMR spectrum of **62**

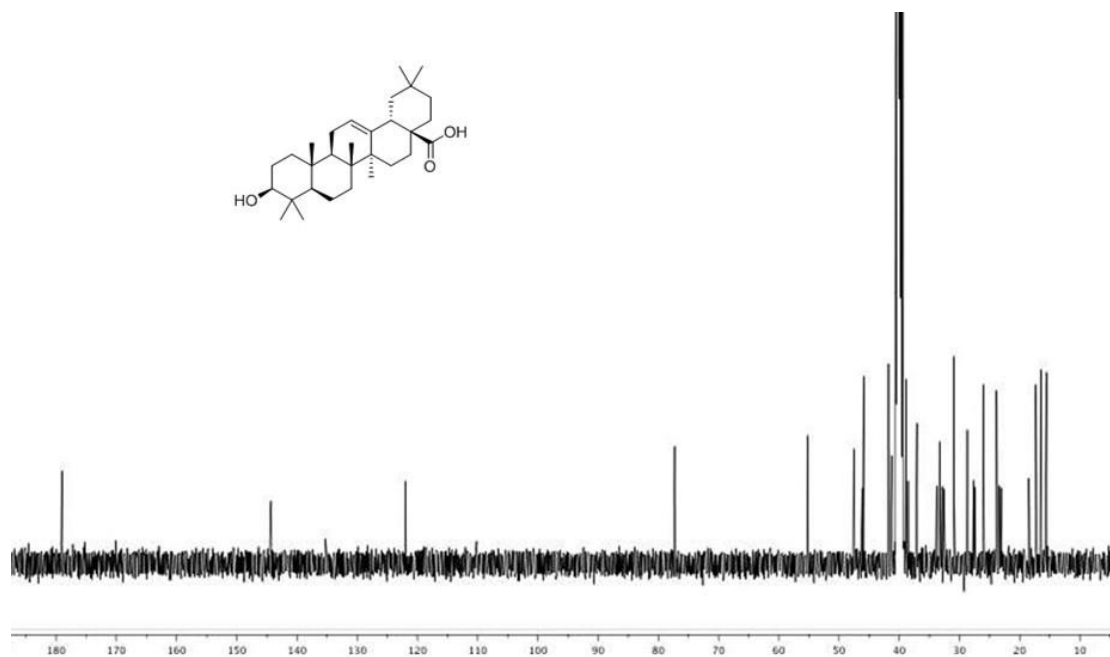


Figure 6.2.4.1.15 Chemical structure and ¹³C NMR spectrum of **62**

Compound **62** was obtained as white powder. Its molecular formula was determined to be C₃₀H₄₈O₃ on the basis of (+)-HRESIMS measurements {[M+H]⁺ ion at *m/z* 455.3145}, which was consistent with 7 degrees of unsaturation. The ¹H NMR

showed seven methyls, and other proton signals (Figure 6.2.4.1.14). The ^{13}C NMR spectrum displayed a total of 30 resonances (Figure 6.2.4.1.15). The result of Liebermann-Burchard reaction for compound **62** was positive, but the colour of solution turned dark red instead of green and then faded, so compound **62** can be classified as one of the triterpenoids. By comparing its ^1H and ^{13}C NMR data with references, compound **62** was determined as oleanic acid.²³⁹⁻²⁴⁰ Oleanic acid and its derivatives were capable of inhibiting HIV-1 protease with an IC_{50} of 4-20 $\mu\text{g}/\text{ml}$.²⁴¹ When tested against *Mycobacterium tuberculosis* streptomycin-, isoniazid-, rifampin-, and ethambutol-resistant strains, the minimum inhibitory concentration of oleanic acid was 50 $\mu\text{g}/\text{ml}$.²⁴² In the study by Steele *et al*, oleanic acid showed antimalarial effects with IC_{50} values of 88.8 $\mu\text{g}/\text{ml}$ and 70.6 $\mu\text{g}/\text{ml}$ for chloroquine-resistant and chloroquine-sensitive *Plasmodium falciparum* strains, respectively.²¹⁶

Compound **64**

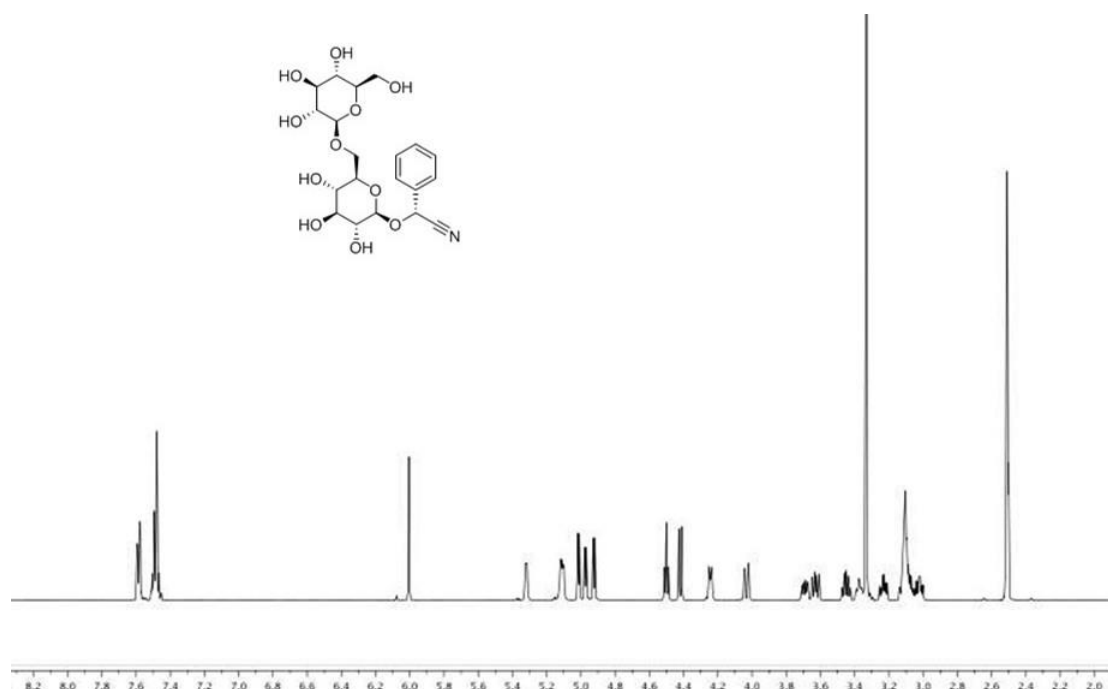


Figure 6.2.4.1.16 Chemical structure and ^1H NMR spectrum of **64**

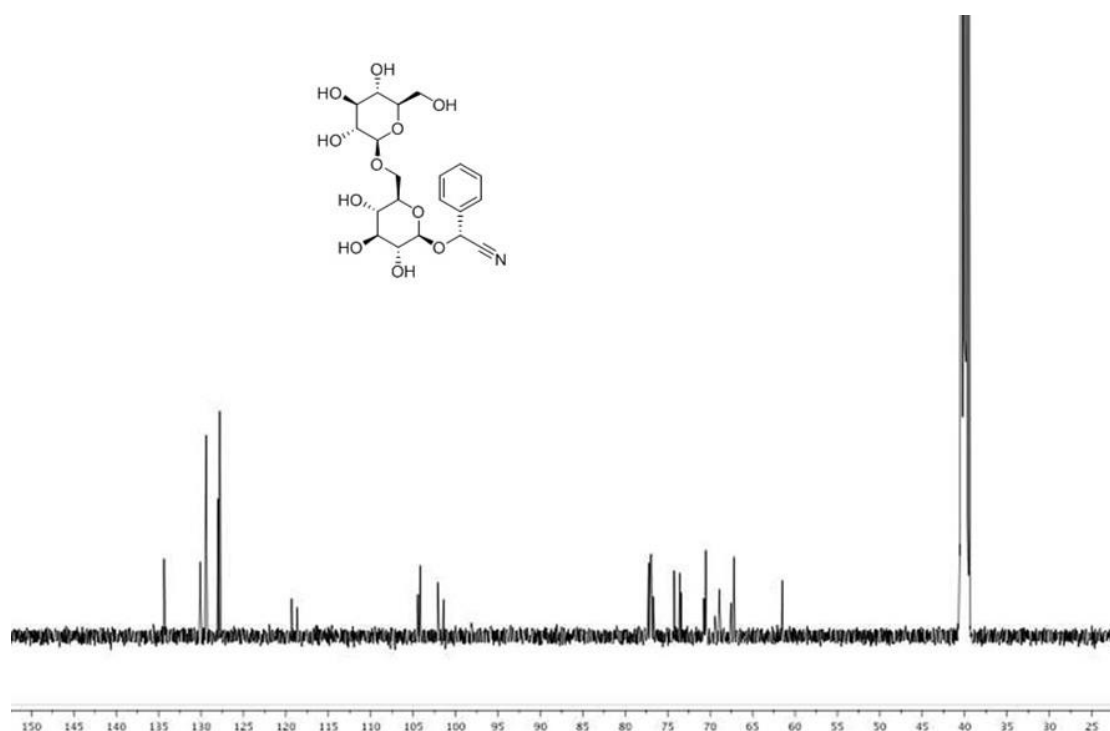


Figure 6.2.4.1.17 Chemical structure and ^{13}C NMR spectrum of **64**

Compound **64** was obtained as white crystals. Its molecular formula was determined to be $\text{C}_{20}\text{H}_{27}\text{NO}_{11}$ on the basis of (+)-HRESIMS measurements $\{[\text{M}-\text{H}]^+\}$ ion at m/z 456.4108}, which was consistent with 8 degrees of unsaturation. The ^1H NMR spectrum indicated aromatic and sugar moieties (Figure 6.2.4.1.16). The ^{13}C NMR spectrum displayed a total of 20 resonances (Figure 6.2.4.1.17). Its molecular formula and genus name were searched in the DNP, and six hits were found. By comparing their ^1H and ^{13}C NMR data, compound **64** was determined as amygdalin.²⁴³⁻²⁴⁴ Amygdalin has showed antitumor activity against cervical cancer, by inhibiting the growth of HeLa cell xenografts through a mechanism of apoptosis.²⁴⁵ According to Guo *et al*, amygdalin could have therapeutic potentials for patients suffering from fibrotic kidney diseases.²⁴⁶

Compound **65**

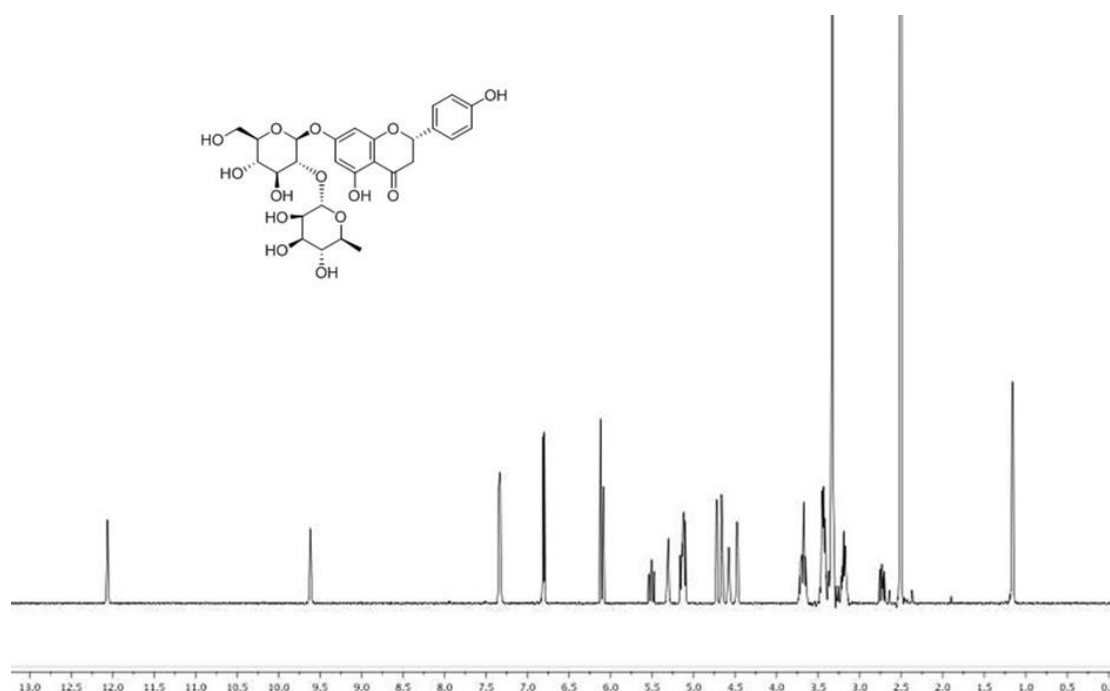


Figure 6.2.4.1.18 Chemical structure and ^1H NMR spectrum of **65**

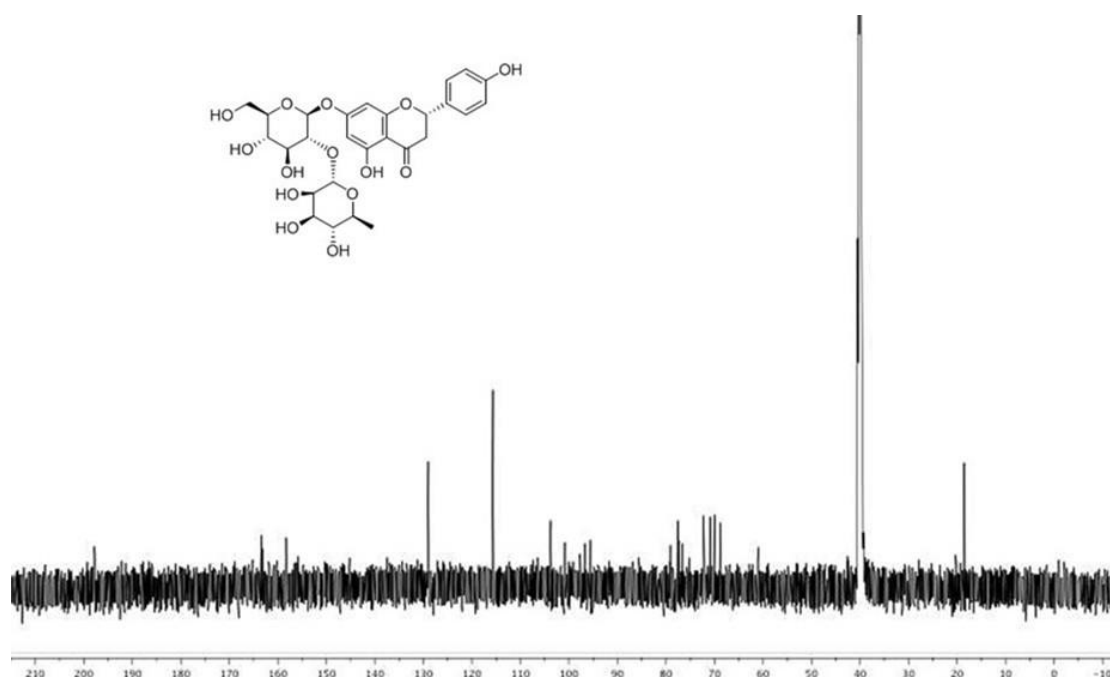


Figure 6.2.4.1.19 Chemical structure and ^{13}C NMR spectrum of **65**

Compound **65** was obtained as light yellow powder. Its molecular formula was determined to be $\text{C}_{27}\text{H}_{32}\text{O}_{14}$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 581.3146, which was consistent with 12 degrees of unsaturation. The ^1H NMR spectrum displayed aromatic and sugar moieties (Figure 6.2.4.1.18). The ^{13}C

NMR spectrum displayed a total of 27 resonances (Figure 6.2.4.1.19). Its molecular formula and genus name were searched in the DNP, and 45 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **65** was determined as naringin.²⁴⁷⁻²⁴⁸ According to Amaro *et al*, naringin has showed a 50% decrease of NO and iNOS expression levels and 80% inhibition of transcription factor NF- κ B, indicating its anti-inflammatory activity.²⁴⁹

Compound **66**

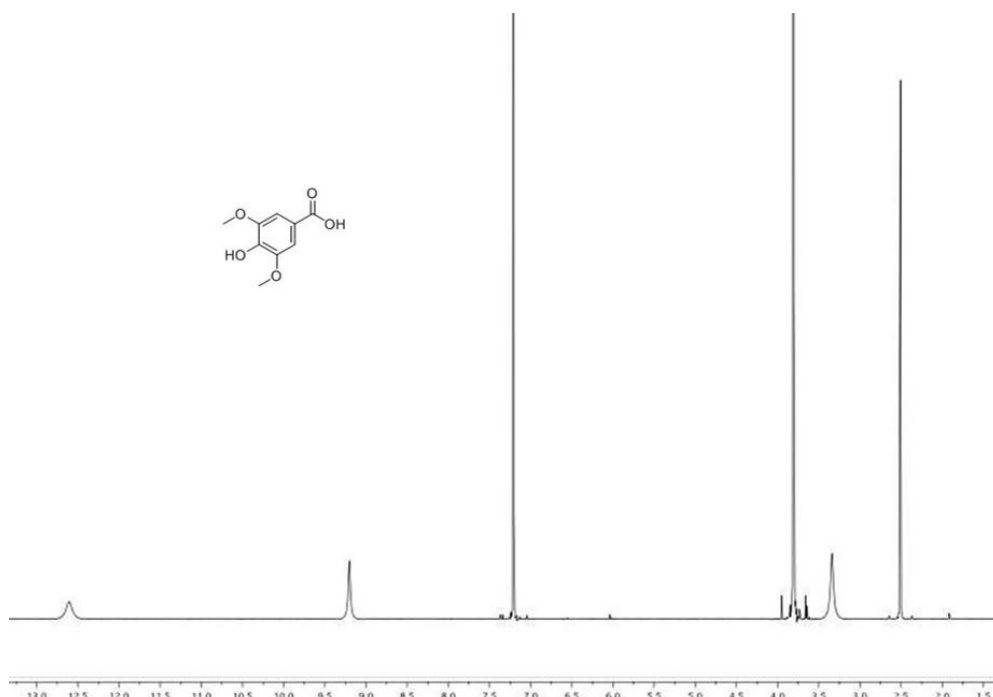


Figure 6.2.4.1.20 Chemical structure and ^1H NMR spectrum of **66**

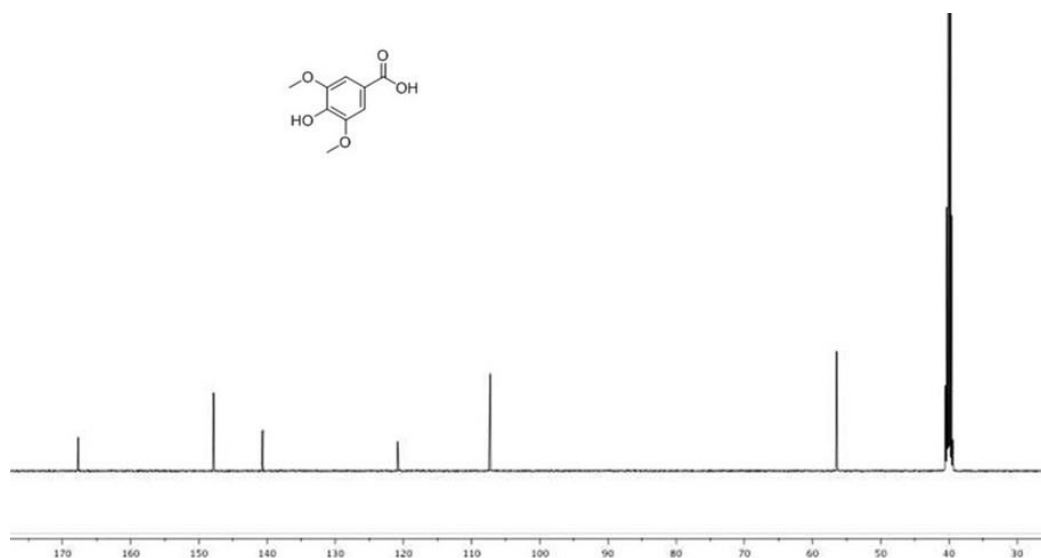


Figure 6.2.4.1.21 Chemical structure and ^{13}C NMR spectrum of **66**

Compound **66** was obtained as light yellow powder. Its molecular formula was determined to be $C_9H_{10}O_5$ on the basis of (+)-HRESIMS measurements $\{[M+H]^+$ ion at m/z 199.2410}, which was consistent with 5 degrees of unsaturation. 1H -NMR spectrum (Figure 6.2.4.1.20) showed two exchangeable proton signals [δ_H 12.60 (1H, brs), 9.20 (1H, brs)], two methoxyls at δ_H 3.81 (6H, s), and one aromatic proton at δ_H 7.21 (2H, s). The ^{13}C NMR spectrum displayed a total of 9 resonances (Figure 6.2.4.1.21). When $FeCl_3$ was added to compound **66** (dissolved in methanol), the color of solution became purple. This result confirmed that the structure of compound **66** contained phenolic hydroxyl proton. Its molecular formula and genus name were searched in the DNP, and 59 hits were found. By comparing their 1H and ^{13}C NMR data, compound **66** was determined as syringic acid.²⁵⁰⁻²⁵¹ In the study by Karthik *et al*, the apparent angiogenic effect of syringic acid was seen from 20 to 50 μM , which significantly reduced number of circulating red blood cells in ISV region compared to that of control.²⁵² Chong *et al* tested syringic acid for *in vitro* antimicrobial activity and it was found to be very fungitoxic to *Ganoderma boninense* even at concentration of 0.5 mg/ml.²⁵³ In the study by Shi *et al*, the minimum inhibitory concentration of syringic acid against all tested *Cronobacter sakazakii* strains was 5 mg/ml.²⁵⁴

Compound **67**

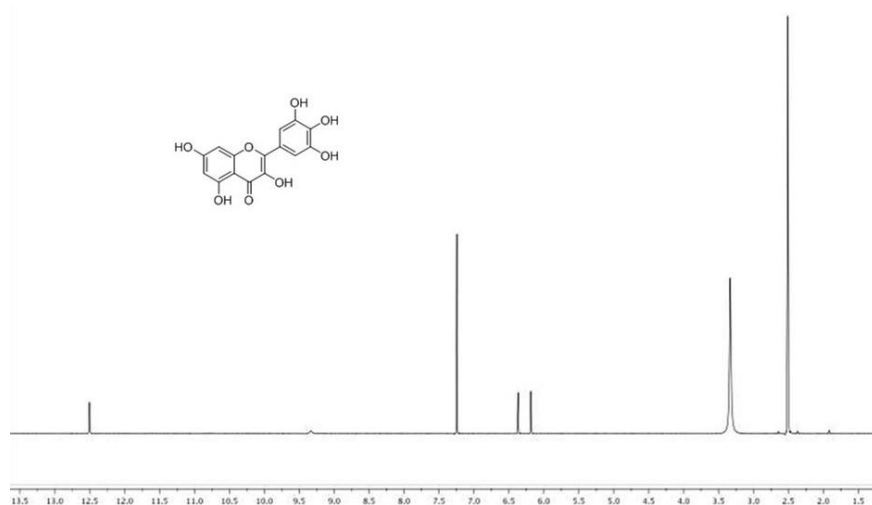


Figure 6.2.4.1.22 Chemical structure and 1H NMR spectrum of **67**

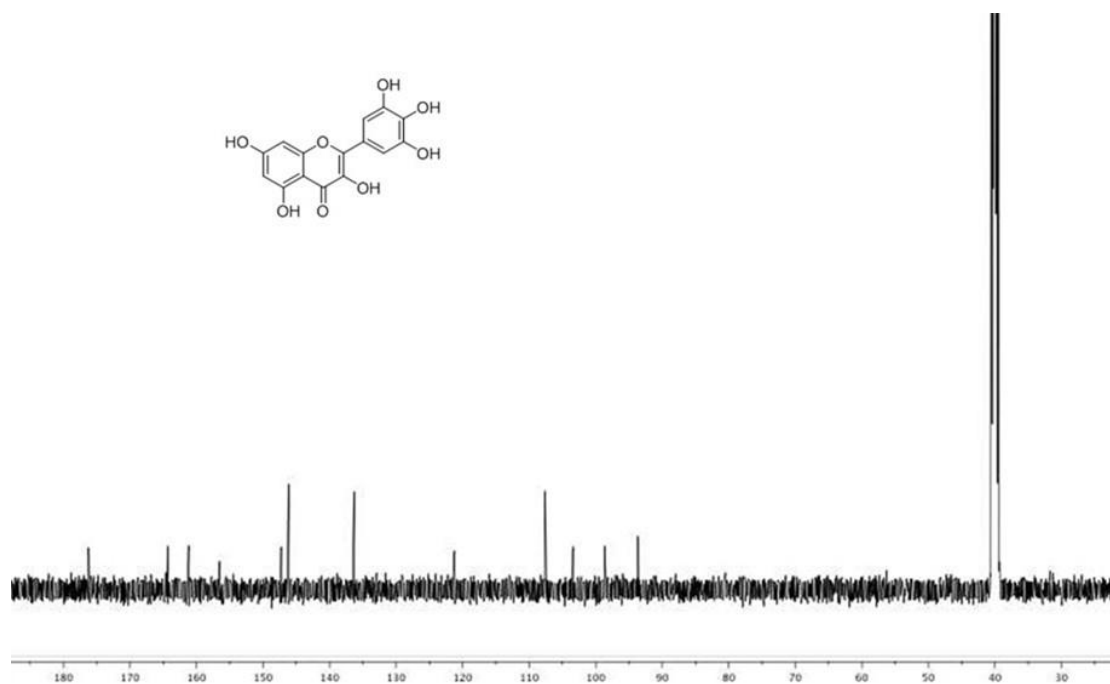


Figure 6.2.4.1.23 Chemical structure and ^{13}C NMR spectrum of **67**

Compound **67** was obtained as white powder. Its molecular formula was determined to be $\text{C}_{15}\text{H}_{10}\text{O}_8$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+\text{H}]^+\}$ ion at m/z 319.3251, which was consistent with 11 degrees of unsaturation. The ^1H -NMR spectrum indicated aromatic moieties (Figure 6.2.4.1.22). ^{13}C -NMR spectra showed a total of 15 resonances (Figure 6.2.4.1.23). When hydrochloric acid and magnesium was added to compound **67** (dissolved in methanol), the solution became orange. This reaction suggested that compound **67** was one flavonoid. Its molecular formula and genus name were searched in the DNP, and 76 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **67** was determined as myricetin.²⁵⁵⁻²⁵⁶ Myricetin has been reported to have various biological activities, including antioxidant, anticancer, anti-inflammatory, anti-amyloidogenic, antibacterial, antiviral, and antidiabetic effects.²⁵⁷

Compound **69**

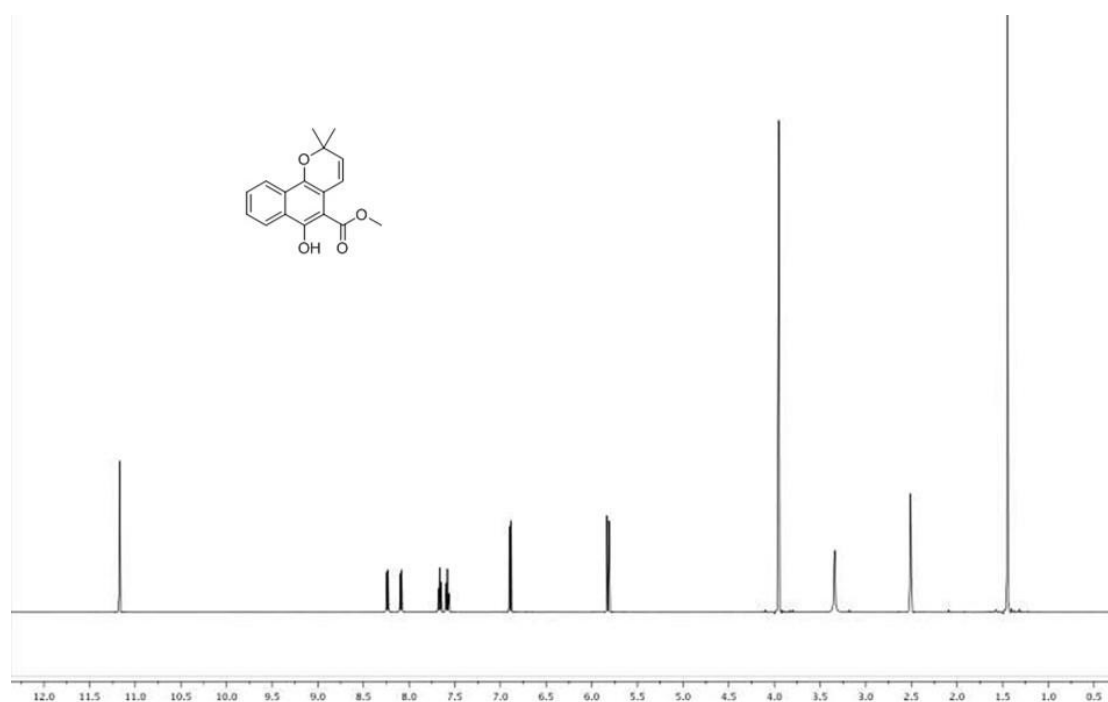


Figure 6.2.4.1.24 Chemical structure and ^1H NMR spectrum of **69**

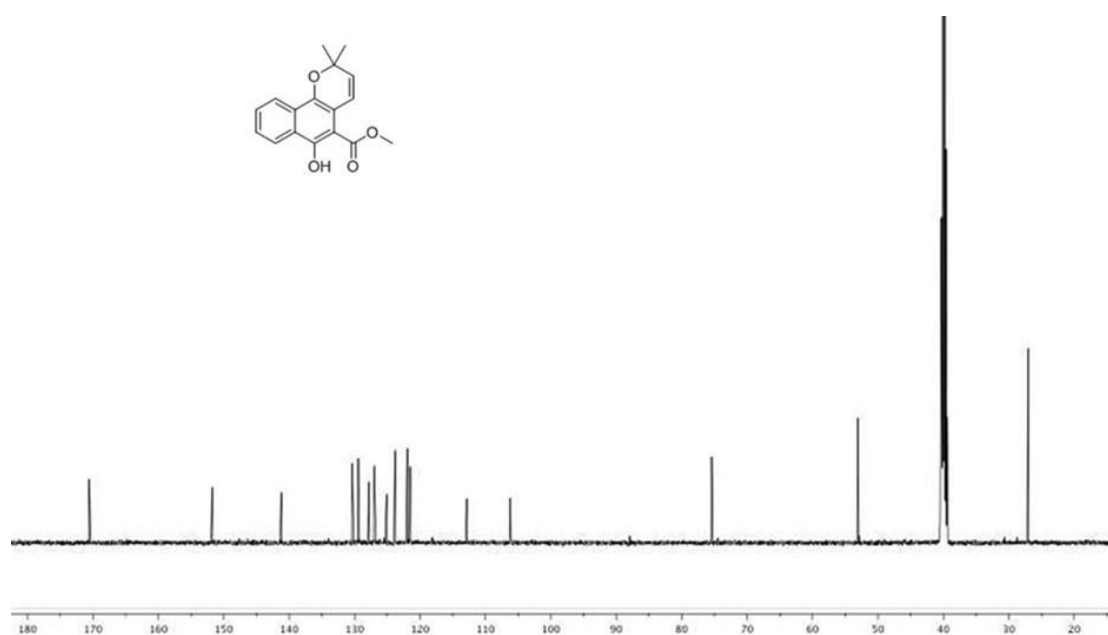


Figure 6.2.4.1.25 Chemical structure and ^{13}C NMR spectrum of **69**

Compound **69** was obtained as yellow-green flakes. Its molecular formula was determined to be $\text{C}_{17}\text{H}_{16}\text{O}_4$ on the basis of (+)-HRESIMS measurements $\{[\text{M}+2\text{H}]^+\}$ ion at m/z 283.5164, which was consistent with 10 degrees of unsaturation. In the ^1H NMR spectrum (Figure 6.2.4.1.24), four proton signals at δ_H 8.38 (1H, ddd, $J=1.0, 2.0, 5.0$ Hz), 8.10 (1H, ddd, $J=1.5, 2.0, 2.5$ Hz), 7.66 (1H, ddd, $J=1.5, 7.0, 15.0$ Hz), 7.57 (1H,

ddd, $J=1.5, 7.0, 15.0$ Hz), were in accordance with a benzene ring. The ^{13}C NMR spectrum gave a total of 17 resonances (Figure 6.2.4.1.24). Its molecular formula and genus name were searched in the DNP, and 119 hits were found. By comparing their ^1H and ^{13}C NMR data, compound **69** was determined as mollugin.²⁵⁸⁻²⁵⁹

6.2.4.2 Physicochemical property analysis

Physicochemical properties of compounds **57-69** were calculated using the method described in Chapter Two. The results are shown in Table 6.2.4.2.1.

Table 6.2.4.2.1 Physicochemical properties of compounds **57-69**

No.	MW	Log P	HBD	HBA	Ro5
57	388.37	-2.21	5	9	Pass
58	450.40	-0.15	7	11	Fail
59	416.38	-0.03	6	9	Pass
60	302.24	2.16	5	7	Pass
61	414.63	4.93	1	3	Pass
62	456.71	6.59	2	3	Pass
63	797.10	0.64	4	9	Pass
64	457.43	-2.59	7	12	Fail
65	580.54	-0.16	8	14	Fail
66	198.17	1.01	2	5	Pass
67	318.24	1.85	6	8	Pass
68	416.47	3.32	1	7	Pass
69	284.31	4.21	1	3	Pass

*Content in red means violation against Ro5.

In this study, three out of fourteen natural products violated Lipinski's Ro5. As discussed previously, Log P is an important drug-like property. The relationship between Log P and HPLC retention time was studied. Figures from 6.2.4.2.1 to 6.2.4.2.2 showed the Log P value of each isolated natural product and the fraction from which it was isolated.

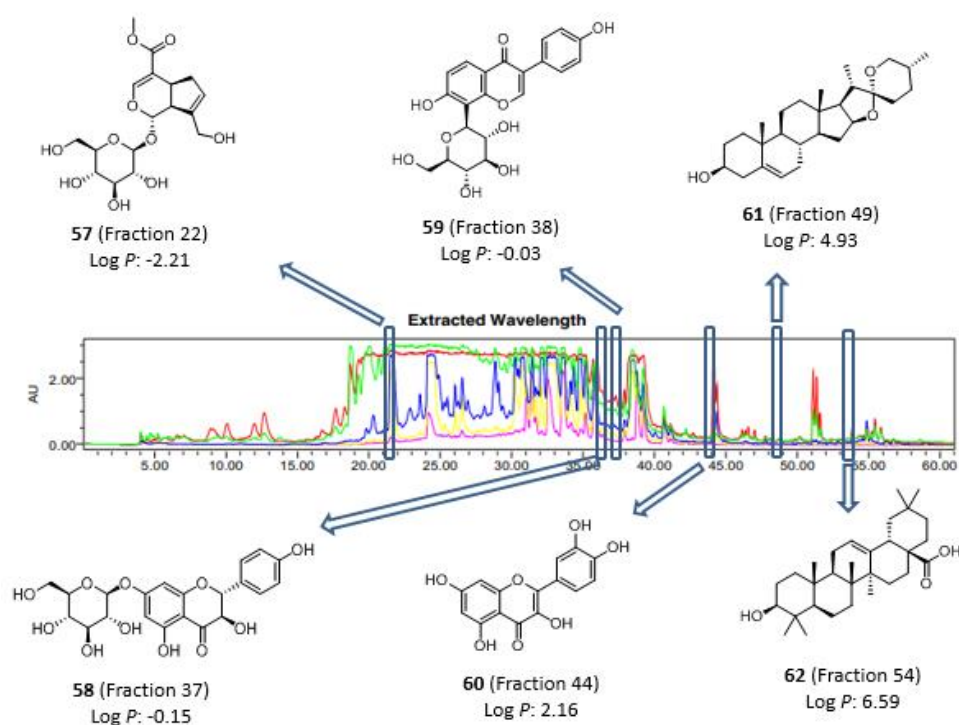


Figure 6.2.4.2.1 Natural products isolated from the Australian plant *Prunus Rosaceae*

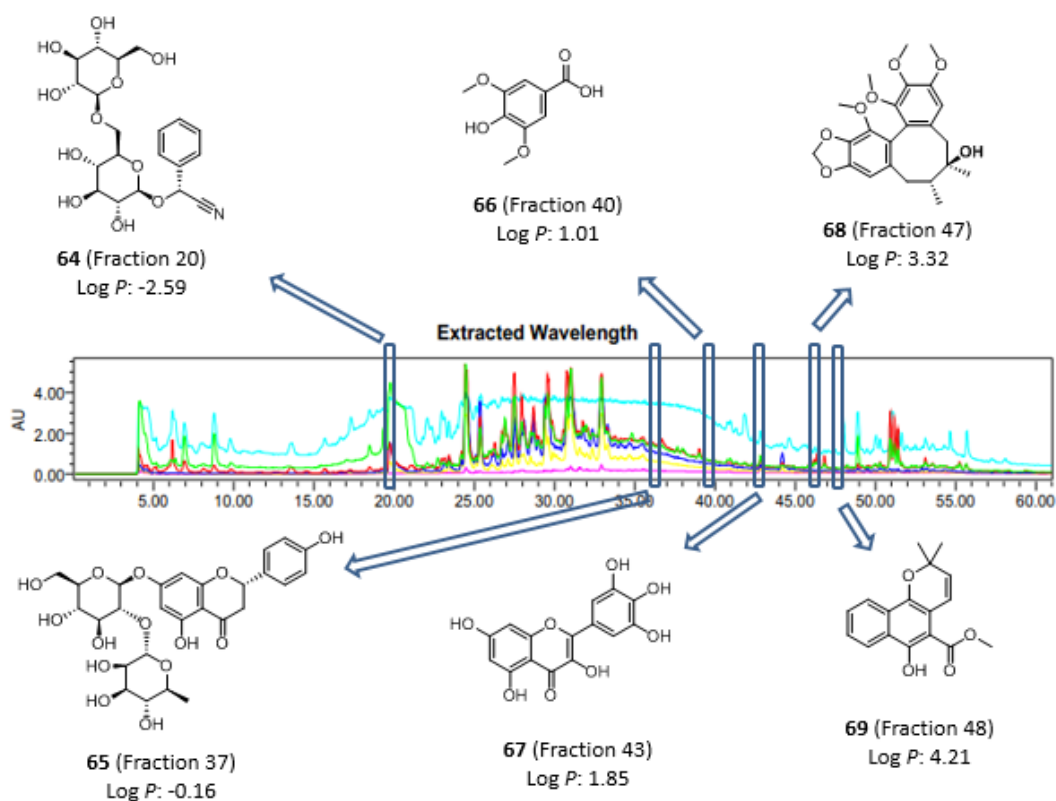


Figure 6.2.4.2.2 Natural products isolated from the Australian plant *Canthium*

Rubiaceae

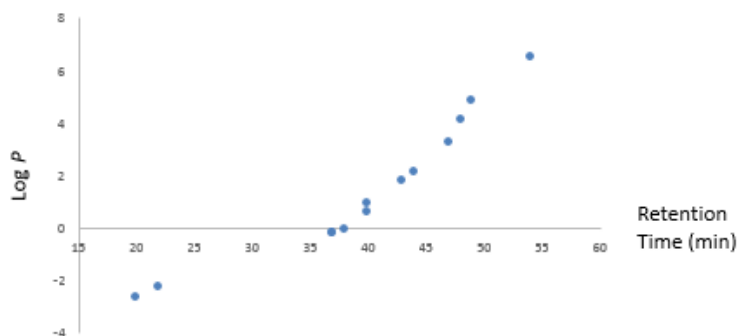


Figure 6.2.4.2.3 Relationship between Log P and Retention Time (min). Each blue dot represents a natural product.

As is shown in Figure 6.3.4.2.3, the retention time for the isolated natural products in this chapter had a linear relationship with their Log P value. This result also proved that the LLE process has the ability to capture most drug-like chemical constituents.

6.2.5 Determination of K_D

Compounds (**58**, **63**, and **68**) were confirmed as active molecules, which bound to *PfUCH-L3*. Their molecular weight met the calculation results. Their structure elucidation and physiochemical property analysis was described in the Chapter 6.3.4. To determine the binding affinity of three active compounds (**58**, **63**, and **68**), the K_D was measured by a titration experiment monitoring the equilibrium population of free protein and ligand-protein complex using a constant concentration of protein (*PfUCH-L3*, 4.5 μ M in 10 mM ammonium acetate buffer, pH 6.8) and increasing concentrations of ligand (1 μ M, 5 μ M, 15 μ M, 25 μ M, 50 μ M and 100 μ M). Each protein-ligand mixture was incubated at 4 °C for 1 hr before analyzing by ESI-FTMS. The experiment was run in triplicate.

It is assumed that the observed protein ion intensity reflects the true protein concentrations. Either peak height or peak area can be used for quantitation method. The peak height intensity of the *PfUCH-L3* at 11+ charge state was used for the calculation of the bound *PfUCH-L3* because the *PfUCH-L3* peak height intensity at other charge states was in lower abundance. Prism software (Version 6.01) was used

to plot the fraction of bound protein against the total ligand concentration of each ligand (Figure 6.2.5.1). According to Equation 6 in Chapter Five, the value of K_D of each ligand was determined. The results were shown in Table 6.2.5.1.

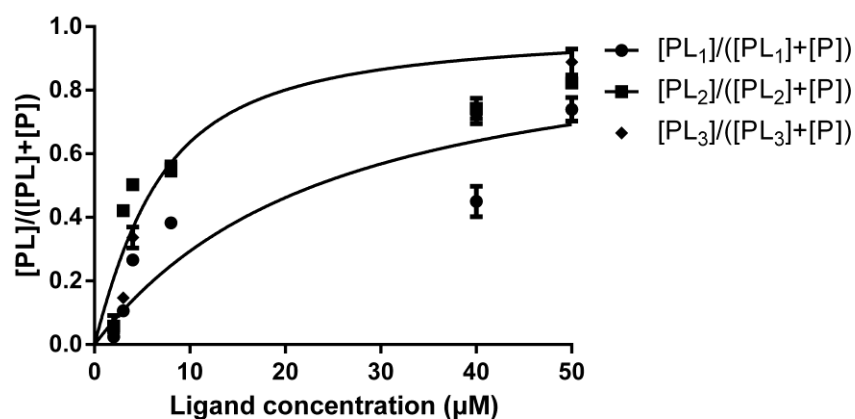


Figure 6.2.5.1 Titration of 3 ligands with *Pf*UCH-L3 for determination of K_D . Upper curve belongs to L₂ (**63**) and the other curve belongs to L₃ (**68**). L₁ (**58**) didn't give a curve as its high K_D value.

Table 6.2.5.1 Results for the determination of K_D (μM)

No.	Best-fit values	Standard deviation
L ₁ (58)	20.85	3.7480
L ₂ (63)	4.076	0.9141
L ₃ (68)	7.147	1.033

According to the titration experiments, compound **63** had the lowest K_D value. Therefore, its binding affinity was the best among all the three *Pf*UCH-L3 natural product binders identified by ESI-FTMS.

6.2.6 Valuation of possible *Pf*UCH-L3 binding sites of those natural product binders via virtual docking

Molecular Docking. Five binding sites were obtained for the protein *Pf*UCH-L3 (2WE6) using SiteMap. The small molecules identified by ESI-FTMS were docked to the

sitemaps respectively with Glide software to find the most appropriate binding site of the target and to investigate their ligand-protein interactions. The binding orientation and features of each molecule relative to the receptor protein were determined and scored with internal scoring function GlideScore. The docking results are given in Table 6.2.6.1. The binding site of the highest docking score ligand was selected and the binding models of them are shown in Figure 6.2.6.1 and Figure 6.2.6.2.

Table 6.2.6.1 Docking scores of the ligands interacting with five different binding sites in 2WE6 predicted by SiteMap.

DockingScore(kcal/mol)	site1	site2	site3	site4	site5	site6	site7
UCH1 (58)	-5.86	-6.49	-6.22	-5.94	-4.99	-5.93	-4.71
UCH2 (63)	-5.77	-6.47	-5.96	-5.93	-5.31	-5.15	-
UCH3 (68)	-3.08	-2.71	-3.44	-5.71	-	-3.11	-1.98

-: No poses were found by docking.

Binding Mode Analysis. The 2D representation of the selected binding modes of the molecules and their receptor hydrogen bond and hydrophobic interactions is shown in Figure 6.3.6.1. The ligand-protein interaction diagrams were generated in LIGPLOT.²²⁰ Hydrogen bond interactions and their atomic distances (in Å) are shown in dashed lines, whereas hydrophobic contacts are shown in red crescents.

As shown in Figure 6.2.6.1, **58** (UCH1) had hydrogen bond interactions with Asp79, Val81, Phe83, Arg105, Asn106, Asn124, Arg131, and Glu172; it also had hydrophobic contacts with Phe78, Asn80, Trp82, Gly102, and Asn103. **63** (UCH2) had hydrogen bond interactions with Arg131, Gln170, and Val81; it also had hydrophobic contacts with Phe78, Asp79, Asn80, Trp82, Gly102, Asn103, Arg105, Asn106, Asn124, Ser127, Ala128, and Ile171. **68** (UCH3) had hydrogen bond interactions with Cys149; it also had hydrophobic contacts with Glu11, Ser16, Leu19, Glu147, Phe148, Cys149, Gly150, Gln151, and Val152.

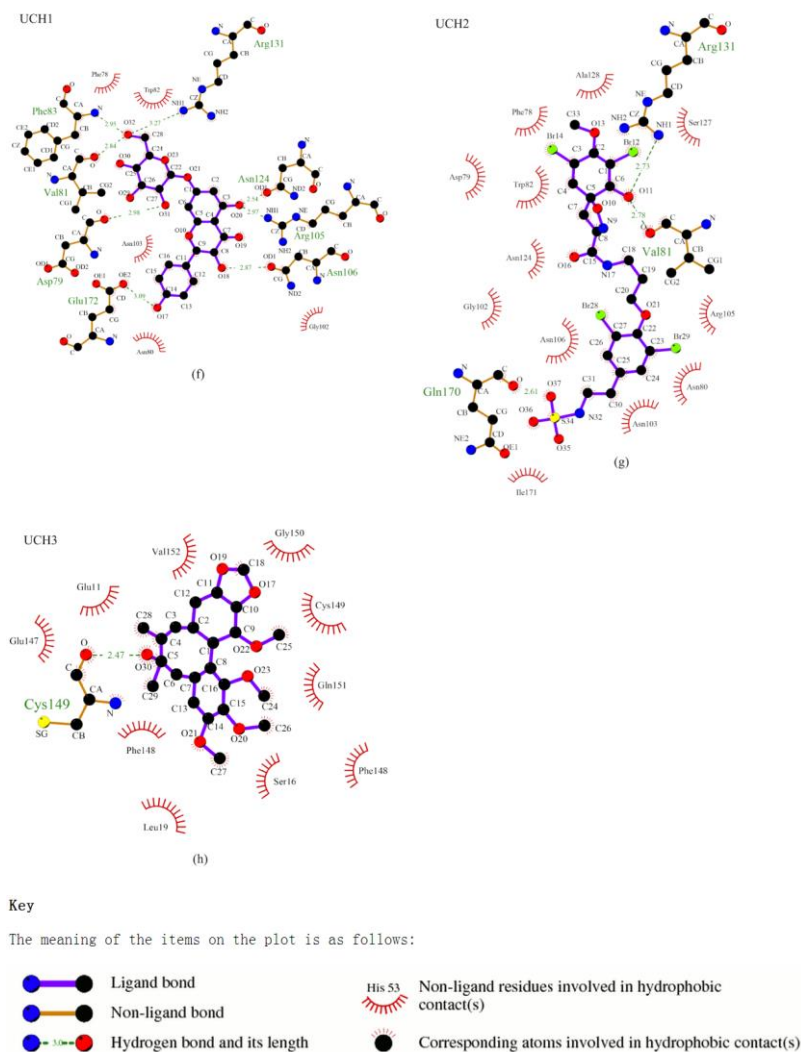


Figure 6.2.6.1 Ligand-protein interaction diagrams of the selected ligands generated by LIGPLOT.

As shown in Figure 6.2.6.2, the ligands of **58** (UCH1), **63** (UCH2), and **68** (UCH3) bound to two different pockets in the 2WE6 binding site. **58** (UCH1) and **63** (UCH2) may bind to the same site. However, **58** (UCH1) had more hydrogen bond interactions than **63** (UCH2). **63** (UCH2) had more hydrophobic contacts than **58** (UCH1). Overall, the numbers of bonded interactions and hydrophobic contacts were observed to be high, suggesting a strong binding between the ligands and the target protein. These results can be used to help future protein and ligand structural research.

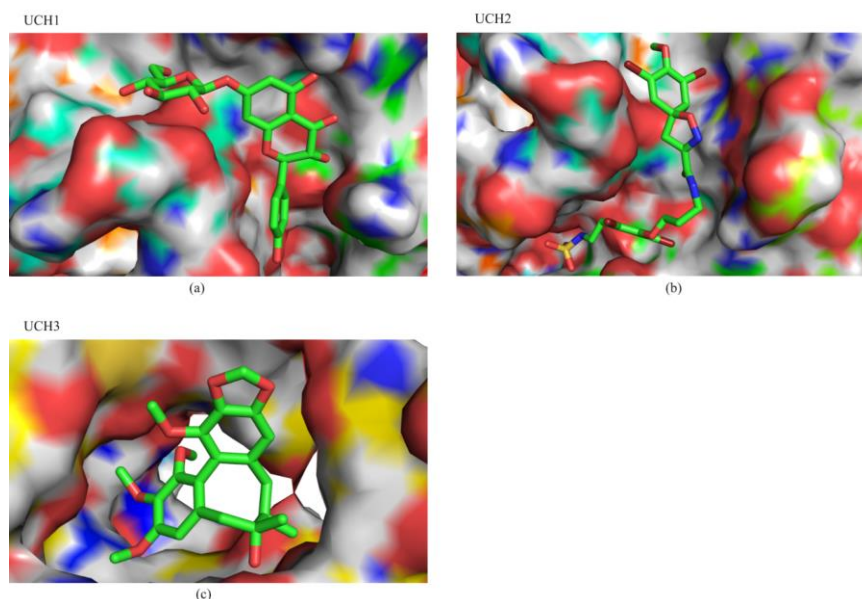


Figure 6.2.6.2 Ligands are shown in stick model while the binding site is depicted as a surface for better visualization. Structural data was from PDB structures 2WE6.

6.3 Summary

3,080 natural product extracts were screened against *PfUCH-L3* using ESI-FTMS. Three natural product extracts were identified to contain compound that bound to the protein *PfUCH-L3*. MS-directed isolation led to the identification of three known natural products (**58**, **63**, and **68**) as active compounds. Their binding activity to *PfUCH-L3* was confirmed by ESI-FTMS titration experiments. It is the first time a bioaffinity MS screening method has been used to study the malarial target *PfUCH-L3*, and it is the first time that natural product binders (**58**, **63**, and **68**) to this target have been reported. In order to predict the binding sites of the active compounds, virtual docking was employed to provide valuable clues for future research. Simultaneously, a number of additional natural products were isolated from the samples that show a protein-ligand complex. Together with three binding ligands, their Log *P* values and HPLC retention time showed a linear relationship, which proved that the LLE process has the ability to capture drug-like chemical constituents.

6.4 Experimental

6.4.1 *PfUCH-L3*

PfUCH-L3 was from Van Voorhis Group (University of Washington, WA, USA) and stored in a nonvolatile buffer (Tris, pH 8.0), which is incompatible with MS. Its sequence of amino acid is:

MAHHHHHMA KNDIWTPLES NPDSLYLYSC KLGQSKLK FV DIYGFNNDLL DMIPQPVQAV
IFLYPVNDNI VSENNTNDKH NLKENFDNVW FIKQVKIITL CNMNNILPIL YVCFNSIELK
NNKSIENLHH EFCGQVENRD DILDVDTHFI VFVQIEGKII ELDGRKDHPT VHCFTNGDNF
LYDTGKIIQD KFIEKCKDDL RFSALAVIPN DNFDII.

The nonvolatile buffer was replaced by ammonium acetate (10 mM, pH 6.8) through a buffer-exchange process using Sephadex G25 column to generate stock solution (4.5 μ M).

6.4.2 Natural product extracts

3,080 natural product extracts (0.5 μ L, 250 μ g/ μ L, each extract) randomly chosen from the Lead-like Enhanced Extract Library were dried and re-solubilized with 5 μ L methanol. Each extract (5 μ L, 25 μ g/ μ L) was incubated with 50 μ L *PfUCH-L3* at 4.5 μ M in ammonium acetate (10 mM, pH 6.8) for direct infusion ESI-FTMS. The total volume of each sample was 55 μ L. They were incubated for 1 hour at 4°C before being analyzed by FTMS.

6.4.3 ESI-FTMS analysis

All experiments were performed on a Bruker Apex III 4.7 Tesla external ESI source FTICR mass spectrometer. Samples were injected directly by a Cole-Parmer syringe pump with a flow rate of 2 μ L per minute. The end plate voltage was biased at 3200V and the capillary voltage at 4,500V relative to the ESI needle during data acquisition. A nebulizing N₂ gas with a pressure of 50 psi and a counter-current drying N₂ gas with a flow of 30 L/min were employed. The drying gas temperature was maintained at 100°C for direct infusion ESI-FTMS. The capillary exit voltage was tuned at 140 V and skimmer 1 voltage at 15 V. Ions were accumulated in an external ion reservoir

comprised of a radio frequency only hexapole, a skimmer cone (skimmer 2) with a tuning voltage of 12 V, and an auxiliary gate electrode, prior to injection into the cylindrical Infinity™ analyzer cell, where they were analyzed. Each run took 3-5 minutes. These optimized conditions were used based on our previous studies.^{51, 135} Mass spectra were recorded in the positive ion mode with mass range m/z 160-6,000 for broadband low-resolution acquisition. Each spectrum was an average of 128 transients (scans) composed of 32,000 data points for high-resolution mode and 32 transients of 32,000 data points for low-resolution mode. All aspects of pulse sequence control and data acquisition were performed on a 1200 MHz Pentium III data station running Bruker's Xmass software under Windows operating system.

6.4.4 HPLC, MS and NMR instruments

HPLC was performed on a Waters 600 HPLC Controller coupled with a Waters 996 Photodiode Array Detector. LR-LC-MS was performed on a Waters 2790 Separations Module HPLC coupled with a Waters micromass ZQ spectrometer. NMR spectra were recorded at 30°C on a Varian 500 MHz spectrometer; samples were dissolved in DMSO- d_6 or $CDCl_3$.

6.4.5 Extraction and isolation

Three extracts gave positive results during FTMS screening. Their taxonomy information was shown in Table 6.4.5.1.

Table 6.4.5.1 Taxonomy information of active extracts

Species	Family	Location	Collectors	Collection Date	Voucher held
<i>Psydrax lamprophylla</i>	Rubiaceae	Palen Creek, near Mt Lindesay, SE Queensland	P.Grimshaw	Nov-1993	Queensland Herbarium (BRI)

<i>Ianthella quadrangulata</i>	Ianthellidae	Idalia National Park, SW Queensland	P.I.Forster & T.Ryan	Jul-2004	Queensland Herbarium (BRI)
<i>Psyrax latifolia</i>	Rubiaceae	Idalia National Park, SW Queensland	P.I.Forster & T.Ryan	Mar-1993	Queensland Herbarium (BRI)

6.4.5.1 Chemical constituents from the Australian plant *Prunus Rosaceae*

The freeze-dried and ground plant material (10 g) was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under gravity then discarded. CH₂Cl₂ (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration the biota was extracted with another volume of MeOH (250 mL) whilst being shaken at 200 rpm for 16 h. All CH₂Cl₂/MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (2.80 g). A portion of this material (1.0 g) was pre-adsorbed onto C18-bonded silica (1 g), and then packed into a stainless steel cartridge (10 × 30 mm) that was subsequently attached to a Thermo Betasil C18 semipreparative HPLC column (150 × 21.2 mm). Isocratic HPLC conditions of 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) were employed for the first 10 min; then a linear gradient from 90% H₂O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9 mL/min. Sixty fractions (60 × 1 min) were collected from the start of the HPLC run (Figure 6.4.5.1.1). Looking for the ion of interest at *m/z* 449.9208 identified in the ESI-FTMS screen, fractions were analyzed by (±)-LRESIMS and NMR spectroscopy.

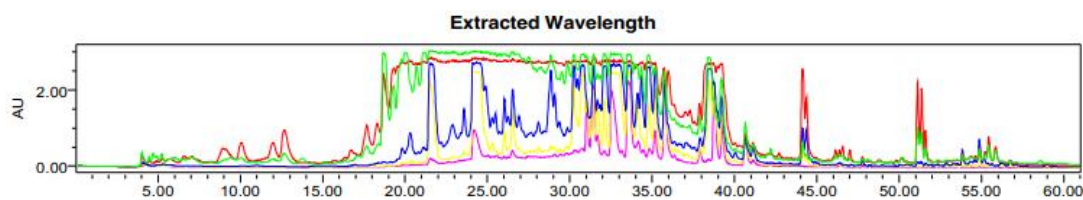


Figure 6.4.5.1.1 HPLC chromatogram of *Prunus Rosaceae*

Fraction 22 yielded compound **57** (2.6 mg). Light brown powder. LR-LC-MS: m/z 389.3 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 7.48 (1H, brs), 5.69 (1H, brs), 5.13 (1H, d, $J=7.0$ Hz), 4.54 (1H, d, $J=8.0$ Hz), 4.16 (1H, d, $J=15.0$ Hz), 3.99 (1H, d, $J=15.0$ Hz), 3.65 (3H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 167.4, 152.0, 144.6, 125.9, 111.4, 99.0, 96.2, 77.7, 77.1, 73.8, 70.4, 61.4, 59.8, 51.5, 46.3, 38.4, 34.9.

Fraction 37 yielded compound **58** (1.5 mg). Colorless solid. LR-LC-MS: m/z 449.5 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 7.35 (2H, d, $J=8.6$ Hz), 6.82 (2H, d, $J=8.6$ Hz), 6.20 (1H, s), 6.17 (1H, s), 4.97 (2H, m), 4.58 (1H, d, $J=1.8$ Hz), 3.85 (1H, s), 3.67 (1H), 3.37-3.50 (4H, m). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 199.3, 167.3, 164.2, 159.2, 130.4, 129.1, 116.2, 115.9, 101.3, 98.3, 97.0, 85.1, 78.2, 77.7, 74.6, 73.7, 71.1, 62.3.

Fraction 38 yielded compound **59** (1.9 mg). Light yellow crystals. LR-LC-MS: m/z 417.2 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 8.35 (1H, s), 7.94 (1H, d, $J=8.5$ Hz), 7.40 (2H, d, $J=9.5$ Hz), 6.99 (1H, d, $J=8.5$ Hz), 6.81 (2H, d, $J=9.5$ Hz), 4.81 (1H, d, $J=10.0$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 175.3, 161.6, 157.6, 153.1, 130.5, 126.7, 123.5, 122.9, 116.7, 115.4, 113.1, 82.3, 79.2, 73.9, 71.2, 61.9.

Fraction 44 yielded compound **60** (2.0 mg). Yellow needle-like crystals. LR-LC-MS: m/z 302.2 $[M]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.50 (1H, s), 7.66 (1H, d, $J=2.5$), 7.54 (1H, dd, $J=2.0, 8.5$ Hz), 6.88 (1H, d, $J=8.5$ Hz), 6.41 (1H, d, $J=2.0$ Hz), 6.19 (1H, d, $J=2.0$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 176.3, 164.3, 161.2, 156.6, 148.1, 147.2, 145.5, 136.2, 122.4, 120.4, 116.0, 115.5, 103.5, 98.6, 93.8.

Fraction 49 yielded compound **61** (1.8 mg). White powder. LR-LC-MS: m/z 413.5 $[M+H]^+$. 1H NMR ($CDCl_3$, 500 MHz), δ_H : 5.35 (1H, d, $J=5.2$ Hz), 4.04 (1H, m), 3.49 (2H, m), 3.37 (1H, td, $J=3.6, 10.8$ Hz), 2.28 (2H, m), 2.00 (2H, m), 1.86 (2H, m), 1.75 (2H,

m), 1.64 (2H, m), 1.55 (2H, m), 1.29 (1H, m), 1.14 (2H, m), 1.02 (2H, s), 0.97 (3H, dd, $J=4.3, 6.6\text{Hz}$), 0.79 (3H, s). ^{13}C NMR (CDCl_3 , 500 MHz), δ_{C} : 140.8, 121.4, 109.3, 80.8, 71.7, 66.9, 62.1, 56.5, 50.1, 42.3, 41.6, 40.3, 39.8, 37.2, 36.7, 32.1, 31.9, 31.6, 31.4, 31.3, 30.3, 28.8, 20.9, 19.4, 17.1, 16.3, 14.5.

Fraction 54 yielded compound **62** (2.3 mg). White needle-like crystals. LR-LC-MS: m/z 455.3 $[\text{M}+\text{H}]^+$. ^1H NMR ($\text{DMSO}-d_6$, 500 MHz), δ_{H} : 5.17 (1H, brs), 3.00 (1H, m), 2.75 (1H, dd, $J=4.5\text{ Hz}, 14.0\text{ Hz}$), 0.68-1.10 (each 3H, s, $7 \times -\text{CH}_3$). ^{13}C NMR ($\text{DMSO}-d_6$, 500 MHz), δ_{C} : 179.0, 144.3, 122.0, 77.3, 55.2, 47.5, 46.2, 45.9, 41.8, 41.3, 38.9, 38.5, 38.1, 37.1, 33.8, 33.3, 32.9, 32.6, 30.9, 28.7, 27.7, 27.4, 26.1, 23.8, 23.4, 23.1, 18.5, 17.3, 16.5, 15.6.

6.4.5.2 Isolation of target compound from the sponge *Ianthella quadrangulata*

The dried and ground marine material (10 g) from the sponge was sequentially extracted with n-hexane (250 mL, shaken at 200 rpm for 2 h), CH_2Cl_2 (250 mL, shaken at 200 rpm for 2 h), and CH_3OH (250 mL, shaken at 200 rpm for 2 h twice). All $\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ extractions were combined and dried to yield a crude extract (2.8 g). 1.0 g of the extract was subsequently pre-adsorbed onto C18-bonded silica (1.0 g), then packed into a stainless steel cartridge (10 \times 30 mm) and attached to a Thermo Betasil C18 semipreparative HPLC column (150 \times 21.2 mm). Isocratic conditions of 10% CH_3OH (0.1% TFA)/90% H_2O (0.1% TFA) were used for the first 10 min, followed by a linear gradient from 90% H_2O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% 100% CH_3OH (0.1% TFA) in 40 min, then isocratic conditions of CH_3OH (0.1% TFA) were run for 10 min. Sixty fractions (60 \times 1 min) were collected from time from 0 min to 60 min, all at a flow rate of 9 mL/min (Figure 6.4.5.2.1). Looking for the ion of interest at m/z 796.2228 identified in the ESI-FTMS screen, fractions were analyzed by (\pm)-LRESIMS and NMR spectroscopy.

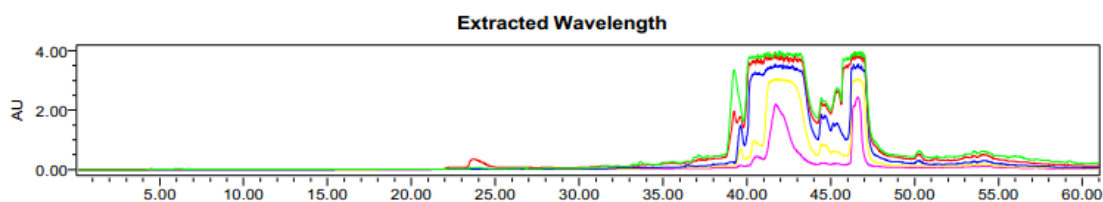


Figure 6.4.5.2.1 HPLC chromatogram of *lanthella quadrangulata*

Fraction 40 yielded compound **63** (1.1 mg). Yellow needle-like crystals. LR-LC-MS: m/z 795.7 $[M-H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 8.53 (1H, t, $J=5.8$ Hz), 7.55 (2H, s), 6.57 (1H, s), 3.95 (1H, t, $J=6.3$ Hz), 3.91 (1H, s), 3.60 (1H, s), 3.63 (1H, d, $J=4.2$ Hz), 3.39 (1H, td, $J=5.9, 8.0$ Hz), 3.25 (1H, t, $J=7.5$ Hz), 2.85 (1H, t, $J=7.5$ Hz), 1.99 (1H, m). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 159.4, 154.9, 151.6, 148.6, 135.3, 133.7, 131.7, 117.9, 113.5, 90.7, 74.0, 60.1, 45.1, 36.7, 30.9, 29.9.

6.4.5.3 Chemical constituents from the Australian plant *Canthium Rubiaceae*

The freeze-dried and ground plant material (10 g) was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under gravity then discarded. CH_2Cl_2 (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration the biota was extracted with another volume of MeOH (250 mL) whilst being shaken at 200 rpm for 16 h. All CH_2Cl_2 /MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (1.90 g). A portion of this material (1.0 g) was pre-adsorbed onto C18-bonded silica (1 g), and then packed into a stainless steel cartridge (10 × 30 mm) that was subsequently attached to a Thermo Betasil C18 semipreparative HPLC column (150 × 21.2 mm). Isocratic HPLC conditions of 90% H_2O (0.1% TFA)/10% MeOH (0.1% TFA) were employed for the first 10 min; then a linear gradient from 90% H_2O (0.1% TFA)/10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9

mL/min. Sixty fractions (60 × 1 min) were collected from the start of the HPLC run (Figure 6.4.5.3.1). Looking for the ion of interest at m/z 416.5764 identified in the ESI-FTMS screen, fractions were analyzed by (\pm)-LRESIMS and NMR spectroscopy.

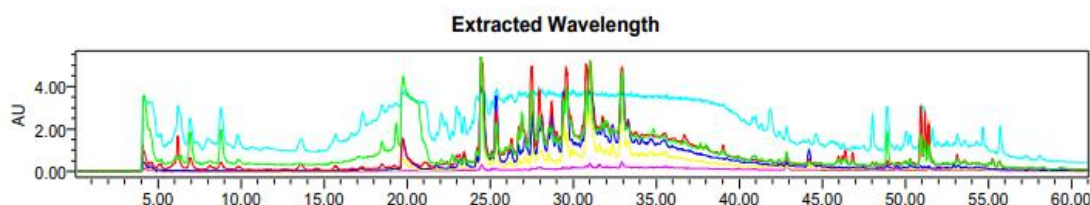


Figure 6.4.5.3.1 HPLC chromatogram of *Canthium Rubiaceae*

Fraction 20 yielded compound **64** (2.7 mg). White crystals. LR-LC-MS: m/z 456.4 $[M-H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 7.58 (2H, m), 7.51 (3H, m), 6.00 (1H, s), 4.50 (1H, d, $J=7.50$ Hz), 4.26 (1H, d, $J=7.50$ Hz), 3.00-4.04 (m). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 134.2, 130.1, 129.4, 127.8, 118.7, 104.4, 102.1, 77.3, 76.9, 76.6, 74.1, 73.4, 70.7, 70.4, 68.9, 67.3, 61.5.

Fraction 37 yielded compound **65** (1.8 mg). Light yellow powder. LR-LC-MS: m/z 581.3 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.10 (1H, s), 9.62 (1H, s), 7.34 (2H, d, $J=8.5$ Hz), 6.80 (2H, d, $J=8.5$ Hz), 6.12 (1H, d, $J=2.0$ Hz), 6.00 (1H, d, $J=2.0$ Hz), 5.50 (1H, t, $J=2.5$ Hz), 5.31 (1H, d, $J=5.0$ Hz), 5.15 (1H, m), 3.70 (2H, m), 3.20 (1H, m), 2.75 (1H, m), 1.15 (3H, d, $J=6.5$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 197.8, 163.4, 158.3, 129.0, 115.7, 103.8, 100.8, 96.8, 95.6, 77.6, 72.3, 70.0, 68.8, 60.9, 18.5.

Fraction 40 yielded compound **66** (1.6 mg). Shallow grey amorphous. LR-LC-MS: m/z 199.2 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.60 (1H, brs), 9.20 (1H, brs), 7.21 (2H, s), 3.81 (6H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 167.7, 147.9, 140.6, 120.8, 107.3, 56.4.

Fraction 43 yielded compound **67** (2.2 mg). White powder. LR-LC-MS: m/z 319.3 $[M+H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 12.5 (1H, s), 7.24 (3H, s), 6.37 (2H, d, $J=2.0$ Hz), 6.18 (2H, d, $J=2.0$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 176.2, 164.3, 161.2, 156.5, 147.2, 146.2, 136.3, 121.2, 107.6, 103.4, 98.6, 93.6.

Fraction 47 yielded compound **68** (1.4 mg). Colorless needles. LR-LC-MS: m/z 415.3 $[M-H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 6.47 (1H, s), 6.62 (1H, s), 3.51 (3H, s), 2.67 (1H, d, $J=13.5$ Hz), 2.34 (1H, d, $J=13.5$ Hz), 2.57 (1H, dd, $J=1.5, 14.1$ Hz), 2.33 (1H, dd, $J=8.1, 14.1$ Hz), 1.24 (3H, s), 0.81 (3H, d, $J=7.5$ Hz). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 152.1, 151.9, 147.2, 141.1, 140.6, 134.5, 124.0, 121.7, 110.2, 105.8, 40.5, 40.2, 33.7, 30.1, 15.8.

Fraction 48 yielded compound **69** (2.2 mg). Yellow-green flakes. LR-LC-MS: m/z 283.5 $[M+2H]^+$. 1H NMR (DMSO- d_6 , 500 MHz), δ_H : 11.2 (1H, s), 8.38 (1H, ddd, $J=5.0, 1.0, 2.0$ Hz), 8.10 (1H, ddd, $J=1.5, 2.0, 2.5$ Hz), 7.66 (1H, ddd, $J=1.5, 7.0, 15.0$ Hz), 7.57 (1H, ddd, $J=1.5, 7.0, 15.0$ Hz), 6.89 (1H, d, $J=10.0$ Hz), 5.82 (d, 1H, d, $J=10.0$ Hz), 3.95 (3H, s), 1.45 (6H, s). ^{13}C NMR (DMSO- d_6 , 500 MHz), δ_C : 170.6, 151.8, 141.2, 130.4, 129.4, 127.8, 127.0, 125.1, 123.8, 122.0, 121.5, 112.9, 106.2, 75.4, 53.1, 27.0.

6.4.6 Dissociation constant (K_D) determination

PfUCH-L3 (4.5 μ M in 10 mM ammonium acetate buffer, pH 6.8, 50 μ L) and increasing concentrations of ligand (1 μ M, 5 μ M, 15 μ M, 25 μ M, 50 μ M and 100 μ M, 5 μ L each) was incubated at 4 $^{\circ}$ C for 1 hr before analyzed by ESI-FTMS. Triplicate values of titration were used for the calculation of K_D . Prism software (Version 6.01) was used to plot the fraction of bound protein against the total ligand concentration of each ligand.

6.4.7 Virtual docking

All virtual docking applications were performed using Schrödinger Suite 2012 (LLC, New York, NY, USA) on a Linux platform. The following Schrödinger modules were used: Protein Preparation Wizard (Epik Version 2.2, Impact version 5.7, Prime version 2.3, Schrödinger, New York, NY, USA, 2011), LigPrep (Version 2.5, Schrödinger, New York, NY, USA, 2011), SiteMap (Version 2.5, Schrödinger, New York, NY, USA, 2011) and Glide²²²⁻²²⁴. All modules were accessed via Maestro graphical interface (Version 9.2, Schrödinger, New York, NY, USA, 2011).

Receptor Protein Preparation. The PDB is used to download the target *Pf*UCH-L3 file (2WE6). The crystal structure of *Pf*UCH-L3 was refined using the Protein Preparation Wizard in the Maestro suite from Schrödinger (Version 9.2, New York, NY, USA, 2011). Hydrogen atoms were added to the heavy atoms and all waters were deleted. Using force field OPLS 2001, minimization was carried out setting maximum heavy atom RMSD to 0.30Å.

Receptor Grid Generation. After preparation, receptor grids were generated with Glide by specifying the binding site with a 3D cubic box. SiteMap was used to estimate the location of the active site by searching regions near the protein surface, generating hydrophobic and hydrophilic contour maps of the protein, and calculating energy potentials. Rotation of all receptor hydroxyl and thiol groups within the grid was allowed.

Ligand preparation. The small molecules were drawn in Maestro workspace, and these 2D structures were converted to all-atom 3D structures using embedded LigPrep script.²²⁵ Up to 32 stereoisomers were generated per ligand by determining chiralities of the 3D structures, and all possible ionization states of ligands were generated at target pH of 7+/- 2. After converting to 3D, conformation search was carried out to generate conformers and search for low energy structures using OPLS 2005 force field.

Glide Docking. The docking program used was Glide. In this study, Glide-SP (standard precision) was used for docking simulations. The Glide algorithm utilizes precomputed grids generated by sitemap. No bonding constraints were given during docking calculations. Using Monte Carlo random search algorithm, ligand poses were generated for each input molecule. Binding affinity of these molecules to receptors was predicted in terms of Glide docking score. Post docking minimization was performed with OPLS 2005 force field, and one pose per ligand was saved. In the present study, we have utilized PyMol (version 1.5.0.4; Schrödinger, LLC: New York, 2012) to analyze the binding sites.

DISCUSSION AND CONCLUSION

Direct screening of natural product extracts on therapeutic targets for the presence of active compounds was employed in this thesis as a rapid and effective strategy to identify the noncovalent interactions between natural products and native (folded) proteins. Simultaneously, with the purpose to study the relationship between the Log *P* value and the HPLC retention time, additional natural products were isolated, purified and structure elucidated together with the active ligands. This thesis presents the outcomes of exploring new antimalarial natural products utilizing bioaffinity mass spectrometry against malarial targets as well as the investigation of the physicochemical properties of metabolites especially the correlation between their HPLC retention time and Log *P* value.

FTMS optimization to detect noncovalent protein-ligand complexes in the gas phase was established using bCA II as a test target protein. Results showed that certain amounts of methanol could improve the intensity of protein signals in ESI-FTMS. It was shown that under optimal conditions, ESI-FTMS could be used as a fast and efficacious method to detect noncovalent complexes under physiological conditions and to identify active molecules from natural product extracts.

Three malarial targets were screened with natural product extracts using ESI-FTMS under optimized condition. To our knowledge, it is the first time a bioaffinity MS screening method has been used to study two of these malarial targets (*PfRab11a* and *PfUCH-L3*), and it was the first time that natural product binders had been reported. Most importantly, one of those binders was a new natural product, and was confirmed as an inhibitor of *PfRab11a* by an enzyme assay. The outcomes related to *PfRab11a* generated a paper, which has been published in the Journal of Biomolecular Screening. Virtual docking was employed to predict the binding sites of the active compounds.

FTMS coupled with ESI plays a key role in the bioaffinity mass spectrometry. Most of proteins can be analyzed by ESI-FTMS. Enzymes can be generally classified into six

groups on the basis of the type of reactions that they catalyze.²⁶⁰ Our group has examined 141 soluble proteins and found native MS spectra for 112. The following table lists the screening results for each class of enzymes (unpublished).

Group	Name	No. of screened	No. of positive
EC1	Oxidoreductases	17	12
EC2	Transferases	26	17
EC3	Hydrolases	23	20
EC4	Lyases	10	8
EC5	Isomerases	10	7
EC6	Ligases	15	12
Other	NA	40	36
Total	NA	141	112

However, insoluble proteins represent a major headache. For instance, membrane-bound and membrane-associated proteins are difficult to analyze by MS, as the association with lipids impedes the isolation and solubilization of the proteins in buffers suitable for MS and the efficient generation of positively charged peptide ions by ESI.²⁶¹

The physicochemical property analysis of isolated natural products showed that their Log *P* values and HPLC retention time displayed a linear relationship, which provided evidence to that reversed-phase HPLC was an appropriate method to capture drug-like chemical constituents (Log *P* value of no more than 5).


Nowadays, high-throughput screening (HTS) has become an important discipline undertaken by academic institutions and pharmaceutical companies for identifying new lead compounds. However, it has also introduced a large number of peculiar compounds, which interfere drug screening to show false results. Time and research money can be wasted to optimize the activity of these molecules. The original article about pan assay interference compounds (PAINS) by Baell *et al* considered synthetic libraries and reported a number of substructural features, which can help to identify

'frequent hitters' in the HTS.²⁶² In general, the opportunity that a PAIN will be progressed to a useful lead molecule is overwhelmingly smaller than the opportunity it will not.²⁶³ Clear structure-activity relationship (SAR) investigations can help to the identification of PAINS.²⁶⁴ More recently, a number of publications have considered natural product PAINS. Bisson *et al* used the natural products alert (NAPRALERT) database to reveal 39 most reported natural products.²⁶⁵ Half of them have not been explained by phenomena known for synthetic libraries, and all showed a variety of bioactivities.²⁶⁵ Pouliot *et al* has reported several antifungal natural products, whose structures contained PAINS motif.²⁶⁶ Glaser *et al* has discussed the most common assay-interfering characteristics of antiprotozoal natural products, described significant examples from recent literatures, and proposed methods to handle the promiscuous molecules.²⁶⁷ Baell suggested that the context of the biological readout must be brought into consideration as a key factor to discuss natural product PAINS.²⁶⁴ According to previous studies, myricitrin (**28**), luteolin (**30**), naringenin (**31**), apigenin (**36**), 1,8-Dihydroxyanthraquinone (**38**), quercetin (**60**), and myricetin (**67**) contain the PAINS motif. More investigations (*e.g.*, SAR) should be considered if researchers would like to study bioactivities of these compounds against any therapeutic target. As to the eight natural products, which showed binding activity against two malarial targets (five for *Pf*Rab11a, another three for *Pf*UCH-L3), each molecule hasn't been screened against a second protein target. Therefore, we are not sure if these ligands are nuisance compounds or not.

This research showed that direct screening has the ability to overcome the bottleneck of bioassay-guided isolation as it allows direct isolation and purification of active molecules from natural product extracts utilizing ligand mass information obtained in the screening stage. Due to its sensitivity, speed and reliability, direct bioaffinity screening mass spectrometry has great potential to serve as an extensively utilized screening method for natural product drug discovery. Simultaneously, the in-house developed LLE procedure that involves two steps,


initial SPE to remove $\text{Log } P > 5$ constituents followed by reversed-phase HPLC, can be used in large-scale isolation. Using reversed-phase HPLC of extracts, provided the $\text{Log } P$ /retention time correlation is used during isolation.

APPENDICES

 Original Research

Bioaffinity Mass Spectrometry Screening

Ben Yang¹, Yun Jiang Feng¹, Hoan Vu¹, Brendan McCormick¹,
Jessica Rowley¹, Liliana Pedro¹, Gregory J. Crowther²,
Wesley C. Van Voorhis², Paul I. Forster³, and Ronald J. Quinn¹

Journal of Biomolecular Screening
2016, Vol. 21(2) 194–200
© 2015 Society for Laboratory
Automation and Screening
DOI: 10.1177/1087057115622605
jbx.sagepub.com


Abstract

Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR-MS or ESI-FTMS) was used to screen 192 natural product extracts and a 659-member natural product-based fragment library for bindings to a potential malaria drug target, *Plasmodium falciparum* Rab11a (PfRab11a, PF13_0119). One natural product extract and 11 fragments showed binding activity. A new natural product, arborside E, was identified from the active extract of *Psydrax montigena* as a weak binder. Its binding activity and inhibitory activity against PfRab11a were confirmed by ESI-FTMS titration experiments and an orthogonal enzyme assay.

Keywords

malaria, FTMS, PfRab11a

Introduction

Malaria is one of the most harmful infectious diseases in the world. According to the World Health Organization, there were approximately 200 million malaria cases in 2014, resulting in about 600,000 deaths. Around 3 billion people continue to be in danger of contracting malaria, mostly in Africa and Southeast Asia.¹

Natural products have been an important source for the development of new drugs to treat malaria.² The use of medicinal plants against parasitic illnesses has been recorded over thousands of years. By way of example, the bark of the Cinchona tree was the first effective treatment for malaria used by the Romans in the 17th century,³ followed by the later identification of the active constituent,⁴ Artemisinin, also known as “qinghaosu,” which is active against all *Plasmodium* species, was isolated from the Chinese medicinal herb *Artemisia annua* L. (sweet wormwood) that had been used to treat malaria for over 2000 years.⁵

Rab GTPases are key regulators of vesicular traffic in eukaryotic cells. In 2003, Langsley and his colleagues reported that *Plasmodium falciparum* Rab11a (PfRab11a, PF13_0119) might mediate an essential recycling endosome function in *P. falciparum*.⁶ Later, they demonstrated that the final step in parasite cell division involves delivery of new plasma membrane to the daughter cells, a process requiring functional Rab11a. Importantly, Rab11a can be found in myosin-tail-interacting protein (MTIP), also known as myosin light chain 1 (MLC1).⁷ MLC1 is a member of a four-protein motor complex called the glideosome that is known to be crucial for parasite invasion of host cells. Ablation of Rab11a function results in a block at a late

stage of cell division.⁷ PfRab11a therefore is a potential target for antimalarial drug discovery.

Bioaffinity mass spectrometry has been used as an effective screening method to identify small molecules that bind to proteins.⁸ The fundamental concept of affinity screening methodologies is to detect a noncovalent complex containing protein and small ligand.⁸ We have shown that a target protein added to a natural product extract can form a noncovalent complex, and we have established a method that employs electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTMS) to screen extracts.⁸ We then adapted the method to fragment-based screening and identified seven securinine derivatives, which bound to *P. falciparum* deoxyuridine 5'-triphosphate nucleotidohydrolase (PfUTPase, PF11_0282).⁹ A subsequent in vitro whole-cell assay against *P. falciparum* showed that the securinine derivatives also had inhibitory activity against the parasite.⁹

¹Eskitis Institute for Drug Discovery, Griffith University, Brisbane, Australia

²University of Washington, Seattle, WA, USA

³Queensland Herbarium, DSIT, Brisbane Botanic Gardens, Queensland, Australia

Received Jul 27, 2015, and in revised form Nov 22, 2015. Accepted for publication Nov 24, 2015.

Supplementary material for this article is available on the *Journal of Biomolecular Screening* Web site at <http://jbx.sagepub.com/supplemental>.

Corresponding Author:

Ronald J. Quinn, Eskitis Institute for Drug Discovery, Griffith University, Brisbane, Queensland 4111, Australia.

Email: r.quinn@griffith.edu.au

In this study, ESI-FTMS was used to screen 192 natural product extracts and an in-house natural product-based fragment library of 659 members for active constituents that bind to *PfRab11a*. One extract from the plant *Psydrax montigena* and 11 fragments (details in Table 1) were identified by ESI-FTMS to bind to *PfRab11a*. Mass guided fractionation and purification of the extract resulted in the identification of a new natural product, arboreside E. We here report the isolation of arboreside E, its binding affinity and inhibitory activity against *PfRab11a*, and the structures of the fragment hits.

Materials and Methods

General Experimental Procedures

Nuclear magnetic resonance (NMR) spectra were acquired at 30 °C on an Agilent (Santa Clara, CA, USA) INOVA 600 MHz spectrometer, which was equipped with a triple-resonance cold probe. Low-resolution electrospray ionization mass spectrometry (LRESIMS) spectra were measured on a Waters ZQ mass spectrometer (Milford, CT, USA). High-resolution electrospray ionization mass spectrometry (HRESIMS) spectra were recorded on a Bruker (Bremen, Germany) Apex III 4.7 Tesla external electrospray ionization (ESI) source Fourier transform ion cyclotron resonance (FTICR) mass spectrometer. A Thermo Electron (Waltham, MA, USA) C₁₈ Betasil 5-μm 143Å column (21.2 × 150 mm) and a Phenomenex (Lane Cove, Australia) C₁₈ Onyx monolithic column (4.6 × 100 mm) were used for semi-preparative high-performance liquid chromatography (HPLC) and compound purity analysis. HPLC separations were conducted on a Waters 600 pump equipped with a 996 Photodiode Array Detector and Gilson 215 Liquid Handler (Middleton, WI, USA).

All ESI-FTMS experiments were performed on a Bruker Apex III 4.7 Tesla external ESI source FTICR mass spectrometer. Samples were directly infused by a Cole-Parmer syringe pump (Chicago, IL, USA) with a flow rate of 2 μL/min. The end plate voltage was biased at 3200 V and the capillary voltage at 4500 V relative to the ESI needle during data acquisition. A nebulizing N₂ gas pressure of 50 psi and a counter-current drying N₂ gas with a flow of 30 L/min was used. The drying gas temperature was maintained at 100 °C for direct-infusion ESI-FTMS. The capillary exit voltage was tuned at 140 V and the skimmer 1 voltage at 15 V. Ions were accumulated in an external ion reservoir composed of a hexapole, a skimmer cone (skimmer 2) with a tuning voltage of 12 V, and an auxiliary gate electrode, prior to injection into the cylindrical Infinity analyzer cell (Bruker, Bremen, Germany), where they were mass analyzed. These optimized conditions were used based on our previous studies.^{8,9}

Mass spectra were recorded in the positive ion mode with mass range from 50 to 6000 *m/z* for broadband low-resolution acquisition. All data acquisitions were performed on a

1200-MHz Pentium III data station running Bruker's Xmass software under the Windows 2000 (Microsoft Corp., Redmond, WA) operating system.

ESI-FIMS Optimization

Based on previous work,^{8,9} optimization of protein concentration and the influence of methanol (MeOH) (from 10%–90%) on *PfRab11a* mass spectra were determined.

PfRab11a was dissolved in ammonium acetate (10 mM, pH 6.8) to generate solutions with different concentration (0.5 μM, 1.0 μM, 1.5 μM, 4.5 μM, and 10 μM). The total volume of each solution was 100 μL. After being incubated at 4 °C for 1 h, 50 μL of each sample was injected into the ESI-FTMS for analysis.

PfRab11a (4.5 μM) was incubated with different concentrations (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90%) of MeOH. Samples were incubated at 4 °C for 1 h and then analyzed by ESI-FTMS.

Sample Preparation and ESI-FTMS Screening

In total, 192 extracts from Nature Bank (Eskitis Institute, Brisbane, Australia) were added to each well of a V-plate microtitre plate (BioCentrix, Carlsbad, CA, USA). Each well contained lead-like enhanced extract prepared by extraction and solid-phase extraction (SPE) from 125 μg of dried biota dissolved in MeOH (5 μL) as previously described.² The *PfRab11a* (50 μL, 4.5 μM) was added into each extract. The mixtures were incubated for 1 h at 4 °C and then analyzed by ESI-FTMS.

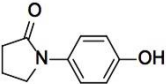
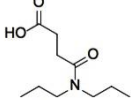
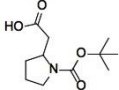

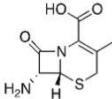
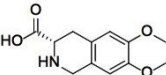
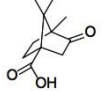
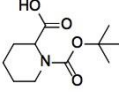
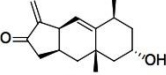
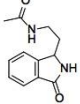

For ESI-FTMS fragment screening, 659 natural product fragments were provided as methanol solutions in 96-well plates. The blank control was prepared by mixing the protein *PfRab11a* (50 μL, 7.29 μM) with methanol (1 μL). Screening samples were prepared by mixing the protein (50 μL, 7.29 μM) with a pool of eight fragments in methanol (1 μL). The final total volume of each well was 51 μL. The concentration of each fragment in the screening assay wells was 98.03 μM. The fragment to protein ratio was ~14:1. All plates were covered and incubated at 4 °C for 1 h and then analyzed by ESI-FTMS. The acquired spectra were compared with the experiment control spectrum.

Plant Material

The plant sample (family: Rubiaceae; species: *P. montigena*) was collected at Daintree National Park, Mt Sorrow track, North Queensland, Australia, in December 1997 by Dr. P. I. F. Forster. A voucher sample has been deposited at the Queensland Herbarium, Brisbane, Australia.

The natural product fragment library had 659 members and was built from a selection of our natural product compound collection. The selection criteria for fragments were as follows: molecular weight, ≤250 Da; calculated log

Table 1. Fragment Hits of *PfRab1* Ia.

Binding Activity ^a	Structure	Name	Molecular Weight
Strong		Lepiotin C	177.2
		N,N-Dipropylsuccinamic acid	201.3
Medium		2-(1-Boc-pyrrolidin-2-yl)acetic acid	229.3
		(R)-2-((2R,5S)-5-((S)-2-hydroxybutyl)tetrahydrofuran-2-yl)propanoic acid	216.3
		Deacetoxycephalosporanic acid	214.2
		(S)-6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid	237.3
		(+)-Camphanic acid	198.2
		1-(tert-butoxycarbonyl)piperidine-2-carboxylic acid	229.3
		2-Hydroxylantolactone	248.3
Weak		N-(2-(3-oxoisindolin-1-yl)ethyl)acetamide	218.3
		1-(1-carboxyethyl)-5-oxopyrrolidine-2-carboxylic acid	201.2

^aStrong: protein complex/protein ratio >50%, medium: protein complex/protein ratio >15% and <50%, and weak: protein complex/protein ratio <15%.

P, <4; hydrogen bond donors, ≤4; hydrogen bond acceptors, ≤5; rotatable bonds, ≤6; and polar surface area, <45.

Extraction and Isolation

The freeze-dried and ground plant material (10 g) was placed into a conical flask (1 L), *n*-hexane (250 mL) was added, and the flask was shaken at 200 rpm for 2 h. The *n*-hexane extract was filtered under gravity and then discarded. CH₂Cl₂ (250 mL) was added to the defatted plant material in the conical flask and shaken at 200 rpm for 2 h. The resultant extract was filtered under gravity and set aside. MeOH (250 mL) was added, and the MeOH/plant mixture was shaken for a further 2 h at 200 rpm. Following gravity filtration, the biota was extracted with another volume of MeOH (250 mL) while being shaken at 200 rpm for 16 h. All CH₂Cl₂/MeOH extracts were combined and dried under reduced pressure to yield a dark brown solid (2.80 g). A portion of this material (1.0 g) was preadsorbed to C₁₈-bonded silica (1 g) and then packed into a stainless steel cartridge (10 × 30 mm) that was subsequently attached to a C₁₈ semipreparative HPLC column. Isocratic HPLC conditions of 90% H₂O (0.1% trifluoroacetic acid [TFA])/10% MeOH (0.1% TFA) were employed for the first 10 min, and then a linear gradient to MeOH (0.1% TFA) was run over 40 min, followed by isocratic conditions of MeOH (0.1% TFA) for a further 10 min, all at a flow rate of 9 mL/min. Sixty fractions (60 × 1 min) were collected from the start of the HPLC run. Fractions were analyzed by (±)-LRESIMS and NMR spectroscopy. Fraction 41 yielded arboreside E (1.4 mg, 0.58% dry weight).

Enzyme Assay

P/Rab11a was cloned from *P. falciparum* complementary DNA (cDNA) and produced and purified as previously described.¹⁰ The protein was dissolved in ammonium acetate (10 mM, pH 6.8) to generate stock solution (4.5 μM).

P/Rab11a activity was measured with a colorimetric GTPase assay kit purchased from Innova Biosciences (BioNovus Life Sciences, Cherrybrook, Australia) following the manufacturer's protocol. Briefly, different concentrations of arboreside E (0.01 μM, 1.0 μM, 1.5 μM, 4.5 μM, 10 μM, 20.0 μM, 30.0 μM, 40.0 μM, 50.0 μM, 80.0 μM, and 100.0 μM) were incubated with *P/Rab11a* (4.5 μM) for 1 h; GTP (10 μM) was added, and the mixture incubated for 30 min at 4 °C. Colorimetric buffer (Pi ColorLock Gold and Accelerator; Innova Biosciences) was added, and the mixture was incubated for 5 min at 4 °C. Then, stabilizer buffer was added, and the reaction mixture was incubated for 30 min. Optical density was determined at 620 nm by using an EnVision Multilabel Reader (PerkinElmer, Waltham, MA).

K_D Determination

Dissociation constant (K_D) is the equilibrium constant for the dissociation of a complex into its components. To determine

the K_D, the relative abundances of bound to total protein in the mass spectra were correlated to the relative equilibrium concentrations of bound to total protein in solution, as follows:

$$\frac{\sum I(\text{PL})^{n+} / n}{\sum I(\text{P})^{n+} / n + \sum I(\text{PL})^{n+} / n} = \frac{[\text{PL}]}{[\text{P}]}, \quad (1)$$

where the total abundances of free protein and protein-ligand complex were obtained by the sum of all the respective charges states intensity peaks (I) normalized for charge state (n). As in each spectrum, three charge states (9+, 10+, and 11+) of free protein and protein-ligand complexes can be observed. The total abundances of free protein and protein-ligand complex were obtained by the sum of their three charge states intensity peaks. By plotting experimentally observed ratios between bound and total protein ion abundances against total concentration of ligand, K_D was obtained as a parameter of a nonlinear least squares curve fitting¹¹:

$$\frac{\sum I(\text{PL})^{n+} / n}{\sum I(\text{P})^{n+} / n + \sum I(\text{PL})^{n+} / n} = \frac{([P]_t + [L]_t + K_D) - \sqrt{([P]_t + [L]_t + K_D)^2 - 4[P]_t[L]_t}}{2[P]_t} \quad (2)$$

Results and Discussion

In the initial spray solution, the protein concentration is an important factor that significantly affects the ESI-mass spectrometry (MS) signal intensity.¹² In this study, ESI-FTMS optimization suggested that 4.5 μM was the most suitable protein concentration to do bioaffinity MS screening (details in supplementary materials). MeOH can increase the solubility of natural products in buffer solution, thus appearing to improve the interaction between protein and ligands. However, too much methanol can cause protein denaturation.¹³ The results showed that methanol increased the intensity of the protein signal (details in supplementary materials).

A 659-member fragment library and 192 crude extracts from Australian plants were screened using ESI-FTMS. Eleven fragments and one extract were identified to bind to the target. Depending on the protein-ligand/protein ratio, the binding of the fragments was ranked as strong, medium, or weak (Table 1). The active extract from the plant *P. montigena* was identified to contain a constituent that bound to the protein *P/Rab11a* (Fig. 1). As shown in the ESI-FTMS spectrum, six peaks were observed. The two at *m/z* 2113.7144 and 2170.9110 represented [*P/Rab11a*]¹¹⁺ and [*P/Rab11a*-ligand]¹¹⁺ complex, respectively; two peaks at *m/z* 2325.4072 and 2388.3710 represented [*P/Rab11a*]¹⁰⁺ and [*P/Rab11a*-ligand]¹⁰⁺ complex, respectively; and two peaks at *m/z* 2583.3924 and 2653.4289 represented [*P/Rab11a*]⁹⁺ and [*P/Rab11a*-ligand]⁹⁺ complex, respectively.

The molecular weight (MW) of the bound ligand was calculated as 629.6380 Da using the following equation¹⁴:

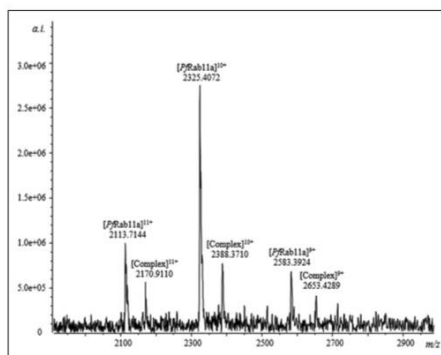


Figure 1. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTMS) spectrum of the protein-ligand complex.

$$MW = \left[\frac{m/z(\text{P/Rab11a} - \text{natural product ligand complex})}{-m/z(\text{free P/Rab11a})} \right] \times (z) =$$

$$[2388.3710 - 2325.407] \times 10 = 629.6380 \text{ (Da)}.$$

To obtain sufficient quantity of the compound for further structure identification and biological activity study, 10 g freeze dried plant sample was extracted with *n*-hexane, CH_2Cl_2 , and MeOH. (+)-LRESIMS measurement of the combined CH_2Cl_2 /MeOH extracts confirmed the presence of the target compound in the extract. The combined CH_2Cl_2 /MeOH extract was then fractionated using reversed-phase C_{18} HPLC eluting with 10% MeOH (0.1% TFA) to 100% MeOH (0.1% TFA) over 60 min. LRESIMS measurement indicated that fraction 41 contained the ion of interest at m/z 631.0 Da, which led to the isolation of a new compound, arboreside E.

Arboreside E (Fig. 2) was isolated as a light yellow, amorphous solid with a $[\alpha]_D^{25} + 10.8$ ($c = 0.5$, CHCl_3). The molecular formula was determined to be $\text{C}_{31}\text{H}_{34}\text{O}_{14}$ on the basis of (+)-HRESIMS measurements (m/z 630.1942) and NMR data (Table 2). Detailed analysis of ^1H -NMR and ^{13}C -NMR spectral data revealed that arboreside E had a similar structure to that of arboreside A.¹⁵ The only difference was the presence of a hydroxyl functionality at the C-8 position in arboreside E. This was evident by the replacement of the methyl doublet at δ_{H} 0.99 in arboreside A by a methyl singlet at δ_{H} 1.30 in arboreside E, as well as the absence of a methine quartet at δ_{H} 2.31. It was also confirmed by the downfield shift of C-8 from δ_{C} 35.6 in arboreside A to δ_{C} 76.6 in arboreside E.

To determine the binding affinity of arboreside E, the K_D was measured by a titration experiment monitoring the equilibrium population of free protein and ligand-protein complex using a constant concentration of protein (4.5 μM in 10 mM ammonium acetate buffer, pH 6.8) and increasing

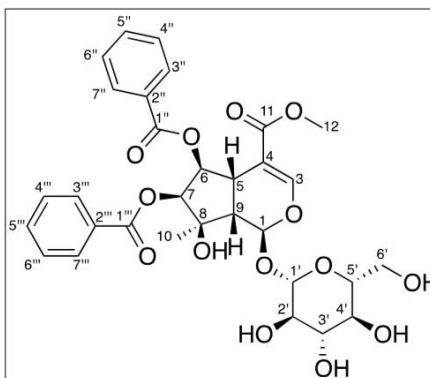


Figure 2. Structure of arboreside E.

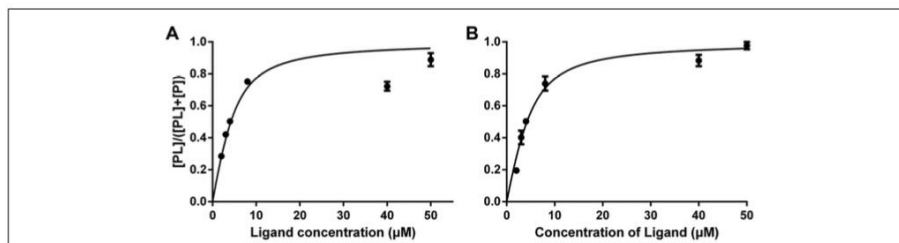
concentrations of ligand (2 μM , 3 μM , 4 μM , 8 μM , 40 μM , and 50 μM). Each protein-ligand mixture was incubated at 4 $^{\circ}\text{C}$ for 1 h before analyzing by ESI-FTMS. The experiment was run in triplicate. Prism software (GraphPad Prism, La Jolla, CA) was used to plot the fraction of bound protein against the total ligand concentration of arboreside E. According to equation (2), the value of K_D of arboreside E was determined as $1.921 \pm 0.4028 \mu\text{M}$ (Fig. 3a). To distinguish a real affinity from titration-like binding, we used the same method to measure the K_D again with a lower protein concentration (3.0 μM), and the K_D value was $1.963 \pm 0.2134 \mu\text{M}$ (Fig. 3b), which indicated a real K_D was measured.

Nonspecific binding can be detected by a competition experiment using a known ligand to compete with binding of the first ligand to the specific site. However, there is no published ligand for P/Rab11a. Under this circumstance, an orthogonal enzyme assay was run against P/Rab11a using the Innova Biosciences GTPase kit to confirm the inhibitory activity of arboreside E. Titration experiments were carried out with P/Rab11a at a fixed concentration of 4.5 μM in 10 mM ammonium acetate buffer (pH 6.8) and the ligand at a varied concentration from 0.01 μM to 100 μM (0.01 μM , 1.0 μM , 1.5 μM , 4.5 μM , 10 μM , 20.0 μM , 30.0 μM , 40.0 μM , 50.0 μM , 80.0 μM , and 100.0 μM). The result showed that arboreside E had an IC_{50} value of $19.73 \pm 1.174 \mu\text{M}$ (Fig. 4). Inhibition constant (K_i) was calculated from the IC_{50} using a web-based tool¹⁶ to give a value of $3.042 \pm 0.1767 \mu\text{M}$. The Michaelis constant (K_m) is the substrate concentration at which the reaction rate is at half-maximum and is an inverse measure of the substrate's affinity for the enzyme. According to a published method,¹⁷ K_m was calculated to have a value of 2.134 μM . All these data confirmed arboreside E's binding activity and inhibitory activity against P/Rab11a.

Table 2. Nuclear Magnetic Resonance (NMR) Spectroscopic Data for Arborside E.^a

No.	¹³ C	¹ H (multiplicity, J in Hz, integration)	HMBC	ROESY
1	92.5	5.64 (d, 2.3, 1H)	3, 5, 8, 1'	9, 1'
3	151.9	7.45 (s, 1H)	1, 4, 5, 11	
4	109.4			
5	34.0	3.29 (m, 1H)	3, 4, 6, 9	9
6	76.2	5.39 (t, 4.7, 1H)	4, 5, 7, 8, 9, 1''	7
7	79.1	5.19 (d, 5.3, 1H)	5, 8, 9, 10, 1'''	6
8	76.6			
9	47.3	2.94 (dd, 11.2, 1.8, 1H)	1, 4, 5, 8, 10	5
10	22.5	1.30 (s, 3H)	7, 8, 9	1, 6, 7
11	166.7			
12	51.4	3.44 (s, 3H)	11	
1'	98.2	4.51 (d, 8.2, 1H)	1	4'
2'	73.6	3.00 (m, 1H)	1', 3'	
3'	77.0	3.18 (m, 1H)	2', 4'	
4'	77.7	3.18 (m, 1H)	3', 5'	
5'	70.5	3.07 (m, 1H)	4', 6'	
6'	61.6	3.47 (dd, 11.7, 5.9, 1H)	4', 5'	4'
		3.71 (dd, 11.7, 1.8, 1H)	5'	4'
1''	165.0			
2''	129.9			
3''	129.5	7.81 (dd, 8.2, 1.2, 2H)	1'', 2'', 4'', 5''	4''
4''	128.9	7.37 (t, 1.6, 2H)	2'', 3'', 5''	
5''	133.6	7.56 (m, 1H)	3''	4''
6''	128.9	7.37 (t, 1.6, 2H)		
7''	129.5	7.81 (dd, 8.2, 1.2, 2H)		
1'''	165.5			
2'''	129.9			
3'''	129.8	7.92 (dd, 8.2, 1.2, 2H)	1''', 2''', 4''', 5'''	4'''
4'''	129.0	7.45 (m, 2H)	2''', 3''', 5'''	
5'''	133.7	7.59 (m, 1H)	3'''	4'''
6'''	129.0	7.45 (m, 2H)		
7'''	129.8	7.92 (dd, 8.2, 1.2, 2H)		

HMBC, heteronuclear multiple bond correlation; ROESY, rotating frame overhauser effect spectroscopy.

^aSpectra were recorded in DMSO-d₆ at 600 MHz.**Figure 3.** Titration of arborside E in *P/Rab11a* for determination of K_D .

ESI-FTMS was used to screen 192 natural product extracts and a library of 659 fragment-sized natural products against *P/Rab11a*. One extract from Australian plant *P. montigena* and 11 fragments were identified to bind to the

protein *P/Rab11a*. Screening a fragment library gave a higher hit rate (1.6%) than screening the natural product extracts (0.5%). MS-directed isolation led to the identification of a new natural product, arborside E. Its binding

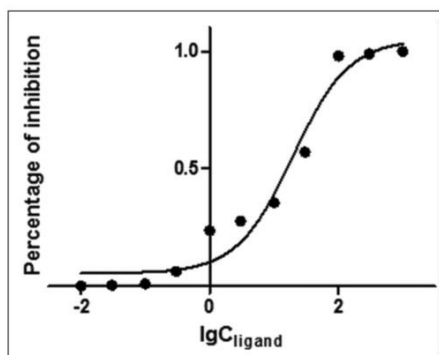


Figure 4. Enzyme assay result.

activity and inhibitory activity against *Pf*Rab11a were confirmed by ESI-FTMS titration experiments and an orthogonal enzyme assay.

To our knowledge, this is the first time a bioaffinity MS screening method has been used to study the malarial target *Pf*Rab11a, and it is the first time that binders (a compound and 11 fragments) to this target and a new natural product inhibitor have been reported.

Our study demonstrated that ESI-FTMS screening is a quick and accurate method to identify binders to the potential drug target *Pf*Rab11a. ESI-FTMS can detect a protein-ligand complex in a complex natural product extract and detect protein-fragment interactions. Bioaffinity ESI-FTMS fragment screening has several advantages. First, there is no need for labeling of protein or compounds. Second, the screening and analysis are fast. The library of 659 fragment members required only 3 h for data acquisition and another 3 h for data analysis and confirmatory testing. Third, the consumption of the protein is very small. The screen of 659 fragments consumed less than 1 mg protein. Fourth, the mass signal of the protein-ligand complex provides the molecular weight of the active binding fragment in a pooled library and the molecular weight of the binding constituent in a complex extract.

In summary, this direct bioaffinity ESI-FTMS screening has a great potential to become a widely used screening method in drug discovery owing to its sensitivity, reliability, and speed.

Acknowledgments

The authors thank the Queensland Herbarium for the collection of the plant sample.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Bruker 4.7 Tesla Apex III was partially funded by the Australian Research Council (ARC LIEF LE0237908). We acknowledge support from the Australian Research Council (ARC Discovery DP0343419, ARC Linkage Project LP120100485) and the Bill & Melinda Gates Foundation Grand Challenges Explorations Grant Phase II OPP1035218 GCE.

References

- World Health Organization. World Malaria Report 2014. http://www.who.int/malaria/publications/world_malaria_report_2014/wmr-2014-no-profiles.pdf. Accessed July 15, 2015.
- Camp, D.; Davis, R. A.; Campitelli, M.; et al. Drug-Like Properties: Guiding Principles for the Design of Natural Product Libraries. *J. Nat. Prod.* **2012**, *75*, 72–81.
- Bwire, R. Quinine: Malaria and the Quest for a Cure That Changed the World. *J. Hist. Med. All. Sci.* **2006**, *61*, 229–230.
- Schlitz, M. Antimalarial Drugs—What Is in Use and What Is in the Pipeline. *Arch. Pharm. (Weinheim)* **2008**, *341*, 149–163.
- Miller, L. H.; Su, X. Z. Artemisinin: Discovery from the Chinese Herbal Garden. *Cell* **2011**, *146*, 855–858.
- Quevillon, E.; Spielmann, T.; Brahimi, K.; et al. The *Plasmodium falciparum* Family of rab GTPases. *Gene* **2003**, *306*, 13–25.
- Agop-Nersesian, C.; Naissant, B.; Ben Rached, F.; et al. Rab11A-Controlled Assembly of the Inner Membrane Complex Is Required for Completion of Apicomplexan Cytokinesis. *PLoS Pathog.* **2009**, *5*, e1000270.
- Vu, H.; Pham, N. B.; Quinn, R. J. Direct Screening of Natural Product Extracts Using Mass Spectrometry. *J. Biomol. Screen.* **2008**, *13*, 265–275.
- Vu, H.; Roullier, C.; Campitelli, M.; et al. *Plasmodium* Gametocyte Inhibition Identified from a Natural-Product-Based Fragment Library. *ACS Chem. Biol.* **2013**, *8*, 2654–2659.
- Mehlin, C.; Boni, E.; Buckner, F. S.; et al. Heterologous Expression of Proteins from *Plasmodium falciparum*: Results from 1000 Genes. *Mol. Biochem. Parasit.* **2006**, *148*, 144–160.
- Bligh, S. W. A.; Haley, T.; Lowe, P. N. Measurement of Dissociation Constants of Inhibitors Binding to Src SH2 Domain Protein by Non-Covalent Electrospray Ionization Mass Spectrometry. *J. Mol. Recognit.* **2003**, *16*, 139–147.
- Kebarle, P.; Tang, L. From Ions in Solution to Ions in the Gas-Phase—The Mechanism of Electrospray Mass-Spectrometry. *Anal. Chem.* **1993**, *65*, A972–A986.
- Fernandez, A.; Sinanoglu, O. Denaturation of Proteins in Methanol Water Mixtures. *Biophys. Chem.* **1985**, *21*, 163–166.
- Green, M. K.; Vestling, M. M.; Johnston, M. V.; et al. Distinguishing Small Molecular Mass Differences of Proteins by Mass Spectrometry. *Anal. Biochem.* **1998**, *260*, 204–211.
- Srivastava, V.; Rathore, A.; Ali, S. M.; et al. New Benzoic Esters of Loganin and 6-Beta-Hydroxyloganin from *Nyctanthes-Arbor-Tristis*. *J. Nat. Prod.* **1990**, *53*, 303–308.
- Cer, R. Z.; Mudunuri, U.; Stephens, R.; et al. IC₅₀-to-Ki: A Web-Based Tool for Converting IC₅₀ to Ki Values for Inhibitors of Enzyme Activity and Ligand Binding. *Nucleic Acids Res.* **2009**, *37*, W441–W445.
- Lineweaver, H.; Burk, D. The Determination of Enzyme Dissociation Constants. *J. Am. Chem. Soc.* **1934**, *56*, 658–666.

REFERENCES

1. WHO. World Malaria Report. **2014**.
2. Fairhurst, R. M.; Wellems, T. E., Chapter 275. Plasmodium species (malaria). In *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases.*, Churchill Livingstone/Elsevier: Philadelphia, Pennsylvania, 2010.
3. Huff, C. G. LIFE CYCLE OF MALARIAL PARASITES *Annual reviews*. **1947**, *1*, 43-60.
4. Cowman, A. F.; Crabb, B. S. Invasion of red blood cells by malaria parasites. *Cell*. **2006**, *124* (4), 755-766.
5. euroclinix.net. <http://www.euroclinix.net/malaria-transmission.html>. (accessed March 25, 2013).
6. Nzila, A. The past, present and future of antifoates in the treatment of Plasmodium falciparum infection. *The Journal of antimicrobial chemotherapy*. **2006**, *57* (6), 1043-1054.
7. WHO *Guideline for the treatment of malaria*. 2010.
8. Murambiwa, P.; Masola, B.; Govender, T.; Mukaratirwa, S.; Musabayane, C. T. Anti-malarial drug formulations and novel delivery systems: A review. *Acta Trop*. **2011**, *118* (2), 71-79.
9. Schlitzer, M. Antimalarial drugs - what is in use and what is in the pipeline. *Arch Pharm (Weinheim)*. **2008**, *341* (3), 149-163.
10. Winstanley, P. Modern chemotherapeutic options for malaria. *The Lancet infectious diseases*. **2001**, *1* (4), 242-250.
11. Black, F.; Bygbjerg, I.; Efferso, P.; Gomme, G.; Jepsen, S.; Jensen, G. A. Fansidar resistant falciparum malaria acquired in South East Asia. *Trans R Soc Trop Med Hyg*. **1981**, *75* (5), 715-716.
12. Carrington, H. C.; Crowther, A. F.; Davey, D. G.; Levi, A. A.; Rose, F. L. A Metabolite of Paludrine with High Antimalarial Activity. *Nature*. **1951**, *168* (4286), 1080-1080.
13. Srivastava, I. K.; Vaidya, A. B. A mechanism for the synergistic antimalarial action of atovaquone and proguanil. *Antimicrob Agents Chemother*. **1999**, *43* (6), 1334-1339.
14. Mather, M. W.; Darrouzet, E.; Valkova-Valchanova, M.; Cooley, J. W.; McIntosh, M. T.; Daldal, F.; Vaidya, A. B. Uncovering the molecular mode of action of the antimalarial drug atovaquone using a bacterial system. *J Biol Chem*. **2005**, *280* (29), 27458-27465.
15. Hudson, A. T.; Dickins, M.; Ginger, C. D.; Gutteridge, W. E.; Holdich, T.; Hutchinson, D. B.; Pudney, M.; Randall, A. W.; Latter, V. S. 566C80: a potent broad spectrum anti-infective agent with activity against malaria and opportunistic infections in AIDS patients. *Drugs under experimental and clinical research*. **1991**, *17* (9), 427-435.
16. Oliaro, P. Mode of action and mechanisms of resistance for antimalarial drugs. *Pharmacology & therapeutics*. **2001**, *89* (2), 207-219.
17. Ahluwalia, V. K.; Chopra, M. *Medicinal Chemistry*. CRC Press: New South Wales, 2008.

18. Hastings, I. M. The origins of antimalarial drug resistance. *Trends in parasitology*. **2004**, 20 (11), 512-518.
19. Wellems, T. E.; Plowe, C. V. Chloroquine-resistant malaria. *The Journal of infectious diseases*. **2001**, 184 (6), 770-776.
20. Tafazoli, S.; O'Brien, P. J. Amodiaquine-induced oxidative stress in a hepatocyte inflammation model. *Toxicology*. **2009**, 256 (1-2), 101-109.
21. Kurth, F.; Pongratz, P.; Belard, S.; Mordmuller, B.; Kremsner, P. G.; Ramharter, M. In vitro activity of pyronaridine against *Plasmodium falciparum* and comparative evaluation of anti-malarial drug susceptibility assays. *Malaria journal*. **2009**, 8, 79.
22. Muregi, F. W. Antimalarial drugs and their useful therapeutic lives: rational drug design lessons from pleiotropic action of quinolines and artemisinins. *Current drug discovery technologies*. **2010**, 7 (4), 280-316.
23. Baird, J. K. Effectiveness of antimalarial drugs. *The New England journal of medicine*. **2005**, 352 (15), 1565-1577.
24. Obonyo, C. O.; Juma, E. A.; Ogutu, B. R.; Vulule, J. M.; Lau, J. Amodiaquine combined with sulfadoxine/pyrimethamine versus artemisinin-based combinations for the treatment of uncomplicated *falciparum* malaria in Africa: a meta-analysis. *Trans R Soc Trop Med Hyg*. **2007**, 101 (2), 117-126.
25. Wells, T. N.; Alonso, P. L.; Gutteridge, W. E. New medicines to improve control and contribute to the eradication of malaria. *Nature reviews. Drug discovery*. **2009**, 8 (11), 879-891.
26. Taylor, W. R.; White, N. J. Antimalarial drug toxicity: a review. *Drug safety : an international journal of medical toxicology and drug experience*. **2004**, 27 (1), 25-61.
27. Ioset, J. R. Natural products for neglected diseases: A review. *Curr Org Chem*. **2008**, 12 (8), 643-666.
28. Miller, L. H.; Su, X. Z. Artemisinin: Discovery from the Chinese Herbal Garden. *Cell*. **2011**, 146 (6), 855-858.
29. Chotivanich, K.; Sattabongkot, J.; Udomsangpetch, R.; Looareesuwan, S.; Day, N. P. J.; Coleman, R. E.; White, N. J. Transmission-blocking activities of quinine, primaquine, and artesunate. *Antimicrob Agents Ch*. **2006**, 50 (6), 1927-1930.
30. Woodrow, C. J.; Haynes, R. K.; Krishna, S. Artemisinins. *Postgrad Med J*. **2005**, 81 (952), 71-78.
31. Ogbonna, A.; Uneke, C. J. Artemisinin-based combination therapy for uncomplicated malaria in sub-Saharan Africa: the efficacy, safety, resistance and policy implementation since Abuja 2000. *Trans R Soc Trop Med Hyg*. **2008**, 102 (7), 621-627.
32. Maffii, G.; Mainardi, L. [Pharmacological study of a new antibiotic: tetracycline]. *Il Farmaco; edizione scientifica*. **1955**, 10 (4), 197-210.
33. Chopra, I.; Roberts, M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiology and molecular biology reviews : MMBR*. **2001**, 65 (2), 232-260 ; second page, table of contents.

34. Newton, P. N.; Chaulet, J. F.; Brockman, A.; Chierakul, W.; Dondorp, A.; Ruangveerayuth, R.; Looareesuwan, S.; Mounier, C.; White, N. J. Pharmacokinetics of oral doxycycline during combination treatment of severe falciparum malaria. *Antimicrob Agents Chemother.* **2005**, *49* (4), 1622-1625.
35. Ejaz, A.; Haqnawaz, K.; Hussain, Z.; Butt, R.; Awan, Z. I.; Bux, H. Treatment of uncomplicated plasmodium falciparum malaria with quinine-doxycycline combination therapy. *JPMMA. The Journal of the Pakistan Medical Association.* **2007**, *57* (10), 502-505.
36. Lell, B.; Kremsner, P. G. Clindamycin as an antimalarial drug: review of clinical trials. *Antimicrob Agents Chemother.* **2002**, *46* (8), 2315-2320.
37. Talisuna, A. O.; Bloland, P.; D'Alessandro, U. History, dynamics, and public health importance of malaria parasite resistance. *Clin Microbiol Rev.* **2004**, *17* (1), 235-254.
38. Achan, J.; Talisuna, A. O.; Erhart, A.; Yeka, A.; Tibenderana, J. K.; Baliraine, F. N.; Rosenthal, P. J.; D'Alessandro, U. Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malaria journal.* **2011**, *10*, 144.
39. Parija, S. C.; Praharaj, I. Drug resistance in malaria. *Indian journal of medical microbiology.* **2011**, *29* (3), 243-248.
40. Lehane, A. M.; McDevitt, C. A.; Kirk, K.; Fidock, D. A. Degrees of chloroquine resistance in Plasmodium - is the redox system involved? *International journal for parasitology, drugs and drug resistance.* **2012**, *2*, 47-57.
41. Baird, J. K. Resistance to therapies for infection by Plasmodium vivax. *Clin Microbiol Rev.* **2009**, *22* (3), 508-534.
42. Dondorp, A. M.; Yeung, S.; White, L.; Nguon, C.; Day, N. P.; Socheat, D.; von Seidlein, L. Artemisinin resistance: current status and scenarios for containment. *Nature reviews. Microbiology.* **2010**, *8* (4), 272-280.
43. Alker, A. P.; Lim, P.; Sem, R.; Shah, N. K.; Yi, P.; Bouth, D. M.; Tsuyuoka, R.; Maguire, J. D.; Fandeur, T.; Ariey, F.; Wongsrichanalai, C.; Meshnick, S. R. Pfmdr1 and in vivo resistance to artesunate-mefloquine in falciparum malaria on the Cambodian-Thai border. *The American journal of tropical medicine and hygiene.* **2007**, *76* (4), 641-647.
44. Crowther, G. J.; Napuli, A. J.; Gilligan, J. H.; Gagaring, K.; Borboa, R.; Francek, C.; Chen, Z.; Dagostino, E. F.; Stockmyer, J. B.; Wang, Y.; Rodenbough, P. P.; Castaneda, L. J.; Leibly, D. J.; Bhandari, J.; Gelb, M. H.; Brinker, A.; Engels, I. H.; Taylor, J.; Chatterjee, A. K.; Fantauzzi, P.; Glynn, R. J.; Van Voorhis, W. C.; Kuhen, K. L. Identification of inhibitors for putative malaria drug targets among novel antimalarial compounds. *Mol Biochem Parasit.* **2011**, *175* (1), 21-29.
45. Zhao, R.; Goldman, I. D. Resistance to antifolates. *Oncogene.* **2003**, *22* (47), 7431-7457.
46. Frederich, M.; Hayette, M. P.; Tits, M.; De Mol, P.; Angenot, L. Reversal of chloroquine and mefloquine resistance in Plasmodium falciparum by the two monoindole alkaloids, icajine and isoretuline. *Planta medica.* **2001**, *67* (6), 523-527.
47. WHO. *World malaria report 2012*; 2012.

48. Cheng, Q.; Kyle, D. E.; Gatton, M. L. Artemisinin resistance in *Plasmodium falciparum*: A process linked to dormancy? *International journal for parasitology, drugs and drug resistance*. **2012**, *2*, 249-255.
49. Whittingham, J. L.; Leal, I.; Nguyen, C.; Kasinathan, G.; Bell, E.; Jones, A. F.; Berry, C.; Benito, A.; Turkenburg, J. P.; Dodson, E. J.; Perez, L. M. R.; Wilkinson, A. J.; Johansson, N. G.; Brun, R.; Gilbert, I. H.; Pacanowska, D. G.; Wilson, K. S. dUTPase as a platform for antimalarial drug design: Structural basis for the selectivity of a class of nucleoside inhibitors. *Structure*. **2005**, *13* (2), 329-338.
50. McCarthy, O.; Musso-Buendia, A.; Kaiser, M.; Brun, R.; Ruiz-Perez, L. M.; Johansson, N. G.; Pacanowska, D. G.; Gilbert, I. H. Design, synthesis and evaluation of novel uracil acetamide derivatives as potential inhibitors of *Plasmodium falciparum* dUTP nucleotidohydrolase. *Eur J Med Chem*. **2009**, *44* (2), 678-688.
51. Vu, H.; Roullier, C.; Campitelli, M.; Trenholme, K. R.; Gardiner, D. L.; Andrews, K. T.; Skinner-Adams, T.; Crowther, G. J.; Van Voorhis, W. C.; Quinn, R. J. Plasmodium Gametocyte Inhibition Identified from a Natural-Product-Based Fragment Library. *Acs Chem Biol*. **2013**, *8* (12), 2654-2659.
52. Crystal structure of *Plasmodium falciparum* Rab11a in complex with GDP. <http://www.rcsb.org/pdb/explore/explore.do?structureId=3bfbk> (accessed Nov 12).
53. Saraste, J.; Goud, B. Functional symmetry of endomembranes. *Mol Biol Cell*. **2007**, *18* (4), 1430-1436.
54. Wilson, G. M.; Fielding, A. B.; Simon, G. C.; Yu, X. Z.; Andrews, P. D.; Hames, R. S.; Frey, A. M.; Peden, A. A.; Gould, G. W.; Prekeris, R. The FIP3-Rab11 protein complex regulates recycling endosome targeting to the cleavage furrow during late cytokinesis. *Mol Biol Cell*. **2005**, *16* (2), 849-860.
55. Fielding, A. B.; Schonteich, E.; Matheson, J.; Wilson, G.; Yu, X. Z.; Hickson, G. R. X.; Srivastava, S.; Baldwin, S. A.; Prekeris, R.; Gould, G. W. Rab11-FIP3 and FIP4 interact with Arf6 and the Exocyst to control membrane traffic in cytokinesis. *Embo J*. **2005**, *24* (19), 3389-3399.
56. Quevillon, E.; Spielmann, T.; Brahimi, K.; Chattopadhyay, D.; Yeramian, E.; Langsley, G. The *Plasmodium falciparum* family of rab GTPases. *Gene*. **2003**, *306*, 13-25.
57. Langsley, G.; Chakrabarti, D. *Plasmodium falciparum*: The small GTPase rab11. *Exp Parasitol*. **1996**, *83* (2), 250-251.
58. Agop-Nersesian, C.; Naissant, B.; Ben Rached, F.; Rauch, M.; Kretzschmar, A.; Thiberge, S.; Menard, R.; Ferguson, D. J. P.; Meissner, M.; Langsley, G. Rab11A-Controlled Assembly of the Inner Membrane Complex Is Required for Completion of Apicomplexan Cytokinesis. *Plos Pathog*. **2009**, *5* (1).
59. Artavanis-Tsakonas, K.; Weihofen, W. A.; Antos, J. M.; Coleman, B. I.; Comeaux, C. A.; Duraingh, M. T.; Gaudet, R.; Ploegh, H. L. Characterization and Structural Studies of the *Plasmodium falciparum* Ubiquitin and Nedd8 Hydrolase UCHL3. *J Biol Chem*. **2010**, *285* (9), 6857-6866.
60. Kim, J. H.; Park, K. C.; Chung, S. S.; Bang, O.; Chung, C. H. Deubiquitinating

- enzymes as cellular regulators. *J Biochem.* **2003**, *134* (1), 9-18.
61. Popp, M. W.; Artavanis-Tsakonas, K.; Ploegh, H. L. Substrate Filtering by the Active Site Crossover Loop in UCHL3 Revealed by Sortagging and Gain-of-function Mutations. *J Biol Chem.* **2009**, *284* (6), 3593-3602.
 62. Fischer, E. Influence of configuration on the action of enzymes. *Ber Dtsch Chem Ges.* **1894**, *27*, 2985-2993.
 63. Ermondi, G.; Caron, G. Recognition forces in ligand-protein complexes: blending information from different sources. *Biochemical pharmacology.* **2006**, *72* (12), 1633-1645.
 64. Hofstadler, S. A.; Sannes-Lowery, K. A. Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. *Nature reviews. Drug discovery.* **2006**, *5* (7), 585-595.
 65. Chowdhury, S. K.; Katta, V.; Chait, B. T. Probing Conformational-Changes in Proteins by Mass-Spectrometry. *J Am Chem Soc.* **1990**, *112* (24), 9012-9013.
 66. Bruce, J. E.; Anderson, G. A.; Chen, R. D.; Cheng, X. H.; Gale, D. C.; Hofstadler, S. A.; Schwartz, B. L.; Smith, R. D. Bio-Affinity Characterization Mass-Spectrometry. *Rapid Commun Mass Sp.* **1995**, *9* (8), 644-650.
 67. Thomson, J. J. *Rays of positive electricity and their application to chemical analysis.* London, Longmans, 1913.
 68. Aston, F. W. *Mass spectra and isotopes.* London, Longmans, 1933.
 69. Washburn, H. W.; Wiley, H. F.; Rock, S. M. The mass spectrometer as an analytical tool. *Ind Eng Chem.* **1943**, *15*, 541-547.
 70. Griffiths, J. A brief history of mass spectrometry. *Anal Chem.* **2008**, *80* (15), 5678-5683.
 71. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yohida, T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid communications in mass spectrometry : RCM.* **1988**, *2*, 151-153.
 72. Beavis, R. C.; Chait, B. T. Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. *Rapid communications in mass spectrometry : RCM.* **1989**, *3* (12), 432-435.
 73. Zenobi, R.; Knochenmuss, R. Ion formation in MALDI mass spectrometry. *Mass Spectrom Rev.* **1999**, *17*, 337-366.
 74. Schiller, J.; Suss, R.; Arnhold, J.; Fuchs, B.; Lessig, J.; Muller, M.; Petkovic, M.; Spalteholz, H.; Zschornig, O.; Arnold, K. Matrix-assisted laser desorption and ionization time-of-flight (MALDI-TOF) mass spectrometry in lipid and phospholipid research. *Progress in lipid research.* **2004**, *43* (5), 449-488.
 75. Richel, A.; Vanderghem, C.; Simon, M.; Wathelet, B.; Paquot, M. Evaluation of matrix-assisted laser desorption/ionization mass spectrometry for second-generation lignin analysis. *Analytical chemistry insights.* **2012**, *7*, 79-89.
 76. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Molecular beams of macroions. *J Chem Phys.* **1968**, *49* (2240-2249).
 77. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray

- ionization for mass spectrometry of large biomolecules. *Science*. **1989**, 246 (4926), 64-71.
78. Taflin, D. C.; Ward, T. L.; Davis, E. J. Electrified droplet fission and the Rayleigh limit. *Langmuir*. **1989**, 5, 376-384.
79. Kebarle, P. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *Journal of mass spectrometry : JMS*. **2000**, 35 (7), 804-817.
80. Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M. Molecular Beams of Macroions .2. *J Chem Phys*. **1970**, 52 (10), 4977.
81. Fernandez de la Mora, J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytica chimica acta*. **2000**, 406, 93-104.
82. Gamero-Castano, M.; Mora, J. F. Kinetics of small ion evaporation from the charge and mass distribution of multiply charged clusters in electrosprays. *Journal of mass spectrometry : JMS*. **2000**, 35 (7), 790-803.
83. Banerjee, S.; Mazumdar, S. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International journal of analytical chemistry*. **2012**, 2012, 282574.
84. Guilhaus, M. Principles and Instrumentation in Time-of-Flight Mass Spectrometry. *Journal of Mass Spectrometry*. **1995**, 30, 1519-1532.
85. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. An introduction to quadrupole-time-of-flight mass spectrometry. *Journal of mass spectrometry : JMS*. **2001**, 36 (8), 849-865.
86. Fiehn, O.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal Chem*. **2000**, 72 (15), 3573-3580.
87. Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal Chem*. **1985**, 57 (3), 675-679.
88. van den Heuvel, R. H.; van Duijn, E.; Mazon, H.; Synowsky, S. A.; Lorenzen, K.; Versluis, C.; Brouns, S. J.; Langridge, D.; van der Oost, J.; Hoyes, J.; Heck, A. J. Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem*. **2006**, 78 (21), 7473-7483.
89. El-Aneed, A.; Cohen, A.; Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Applied Spectroscopy Reviews*. **2009**, 44 (3), 210-230.
90. Hipple, J. A.; Sommer, H.; Thomas, H. A. A Precise Method of Determining the Faraday by Magnetic Resonance. *Phys. Rev*. **1949**, 76 (12), 1877-1878.
91. Comisarow, M. B.; Marshall, A. G. Fourier transform ion cyclotron resonance spectroscopy. *Chem Phys Lett*. **1974**, 25, 282-283.
92. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier transform ion cyclotron

- resonance mass spectrometry: A primer. *Mass Spectrom Rev.* **1998**, 17 (1), 1-35.
93. He, F.; Emmett, M. R.; Hakansson, K.; Hendrickson, C. L.; Marshall, A. G. Theoretical and experimental prospects for protein identification based solely on accurate mass measurement. *Journal of proteome research.* **2004**, 3 (1), 61-67.
 94. Dufresne, C. P.; Wood, T. D.; Hendrickson, C. L. High-resolution electrospray ionization Fourier transform mass spectrometry with infrared multiphoton dissociation of glucokinase from *Bacillus Stearotherophilus*. *J Am Soc Mass Spectrom.* **1998**, 9 (11), 1222-1225.
 95. Yu, E.; Fabris, D. Direct probing of RNA structures and RNA-protein interactions in the HIV-1 packaging signal by chemical modification and electrospray ionization fourier transform mass spectrometry. *J Mol Biol.* **2003**, 330 (2), 211-223.
 96. Aaserud, D. J.; Simonsick, W. J. Modern mass spectrometry for coatings. *Progress in Organic Coatings.* **1998**, 34, 206-213.
 97. Zhao, X.; He, K. Technique of hybrid fourier transform ion cyclotron resonance mass spectrometry and its application in the. *Modern Scientific Instruments.* **2008**, 2, 43-58.
 98. Aharoni, A.; Keizer, L. C.; Van Den Broeck, H. C.; Blanco-Portales, R.; Munoz-Blanco, J.; Bois, G.; Smit, P.; De Vos, R. C.; O'Connell, A. P. Novel insight into vascular, stress, and auxin-dependent and -independent gene expression programs in strawberry, a non-climacteric fruit. *Plant Physiol.* **2002**, 129 (3), 1019-1031.
 99. Zhang, Z.; Guan, S.; Marshall, A. G. Enhancement of the effective resolution of mass spectra of high-mass biomolecules by maximum entropy-based deconvolution to eliminate the isotopic natural abundance distribution. *J. Am. Soc. Mass Spectrom.* **1997**, 8 (6), 659-670.
 100. Xiong, S.; Pu, D.; Xin, B.; Wang, G. Determination of complicated surfactant components using fourier transform ion cyclotron resonance mass spectrometry. *Chinese Journal of Analytical Chemistry.* **2003**, 31 (4), 429-432.
 101. <http://www.nature-bank.com.au/>. (accessed October 20, 2012).
 102. Camp, D.; Davis, R. A.; Campitelli, M.; Ebdon, J.; Quinn, R. J. Drug-like properties: guiding principles for the design of natural product libraries. *Journal of natural products.* **2012**, 75 (1), 72-81.
 103. Kingston, D. G. I. Modern Natural Products Drug Discovery and Its Relevance to Biodiversity Conservation. *J Nat Prod.* **2011**, 74 (3), 496-511.
 104. Feher, M.; Schmidt, J. M. Property distributions: Differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comp Sci.* **2003**, 43 (1), 218-227.
 105. Xu, W. Y. Traditional Chinese Medicines Against Malaria. *Parasitol Res Mg.* **2014**, 6, 67-78.
 106. Faurant, C. From Bark to Weed: The History of Artemisinin. *Parasite.* **2011**, 18 (3), 215-218.
 107. Watsierah, C. A.; Ouma, C. Access to artemisinin-based combination therapy (ACT) and quinine in malaria holoendemic regions of western Kenya. *Malaria J.* **2014**,

- 13.
108. Saxena, S.; Pant, N.; Jain, D. C.; Bhakuni, R. S. Antimalarial agents from plant sources. *Curr Sci India*. **2003**, *85* (9), 1314-1329.
109. Gademann, K.; Kobylinska, J. Antimalarial Natural Products of Marine and Freshwater Origin. *Chem Rec*. **2009**, *9* (3), 187-198.
110. Kaur, K.; Jain, M.; Kaur, T.; Jain, R. Antimalarials from nature. *Bioorgan Med Chem*. **2009**, *17* (9), 3229-3256.
111. Xu, Y. J.; Pieters, L. Recent Developments in Antimalarial Natural Products Isolated from Medicinal Plants. *Mini-Rev Med Chem*. **2013**, *13* (7), 1056-1072.
112. Larsson, J.; Gottfries, J.; Muresan, S.; Backlund, A. ChemGPS-NP: Tuned for navigation in biologically relevant chemical space. *J Nat Prod*. **2007**, *70* (5), 789-794.
113. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliver Rev*. **1997**, *23* (1-3), 3-25.
114. Lipinski, C.; Hopkins, A. Navigating chemical space for biology and medicine. *Nature*. **2004**, *432* (7019), 855-861.
115. Taylor, W. R. J.; White, N. J. Antimalarial drug toxicity - A review. *Drug Safety*. **2004**, *27* (1), 25-61.
116. Wittlin, S.; Ekland, E.; Craft, J. C.; Lotharius, J.; Bathurst, I.; Fidock, D. A.; Fernandes, P. In Vitro and In Vivo Activity of Solithromycin (CEM-101) against Plasmodium Species. *Antimicrob Agents Ch*. **2012**, *56* (2), 703-707.
117. Quinn, R. J.; Carroll, A. R.; Pham, N. B.; Baron, P.; Palframan, M. E.; Suraweera, L.; Pierens, G. K.; Muresan, S. Developing a Drug-like Natural Product Library (vol 71, pg 468, 2008). *J Nat Prod*. **2008**, *71* (10), 1792-1792.
118. Bruce, J. E.; Smith, V. F.; Liu, C. L.; Randall, L. L.; Smith, R. D. The observation of chaperone-ligand noncovalent complexes with electrospray ionization mass spectrometry. *Protein Sci*. **1998**, *7* (5), 1180-1185.
119. Gao, H.; Leary, J.; Carroll, K. S.; Bertozzi, C. R.; Chen, H. Y. Noncovalent complexes of APS reductase from M-tuberculosis: Delineating a mechanistic model using ESI-FTICR MS. *J Am Soc Mass Spectr*. **2007**, *18* (2), 167-178.
120. Gao, J. M.; Wu, Q. Y. Q.; Carbeck, J.; Lei, Q. P.; Smith, R. D.; Whitesides, G. M. Probing the energetics of dissociation of carbonic anhydrase-ligand complexes in the gas phase. *Biophys J*. **1999**, *76* (6), 3253-3260.
121. Cummins, L. L.; Chen, S.; Blyn, L. B.; Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. Multitarget affinity/specificity screening of natural products: Finding and characterizing high-affinity ligands from complex mixtures by using high-performance mass spectrometry. *J Nat Prod*. **2003**, *66* (9), 1186-1190.
122. Seyfried, N. T.; Atwood, J. A.; Yongye, A.; Almond, A.; Day, A. J.; Orlando, R.; Woods, R. J. Fourier transform mass spectrometry to monitor hyaluronan-protein interactions: use of hydrogen/deuterium amide exchange. *Rapid Commun Mass Sp*. **2007**, *21* (2), 121-131.
123. Freitas, M. A.; Hendrickson, C. L.; Marshall, A. G.; Rostom, A. A.; Robinson, C.

V. Competitive binding to the oligopeptide binding protein, OppA: In-trap cleanup in an Fourier transform ion cyclotron resonance mass spectrometer. *J Am Soc Mass Spectr.* **2000**, *11* (11), 1023-1026.

124. Martinovic, S.; Berger, S. J.; Pasa-Tolic, L.; Smith, R. D. Separation and detection of intact noncovalent protein complexes from mixtures by on-line capillary isoelectric focusing-mass spectrometry. *Anal Chem.* **2000**, *72* (21), 5356-5360.

125. Gentili, A.; Perret, D.; Marchese, S.; Sergi, M.; Olmi, C.; Curini, R. Accelerated solvent extraction and confirmatory analysis of sulfonamide residues in raw meat and infant foods by liquid chromatography electrospray tandem mass spectrometry. *J Agr Food Chem.* **2004**, *52* (15), 4614-4624.

126. Iavarone, A. T.; Williams, E. R. Mechanism of charging and supercharging molecules in electrospray ionization. *J Am Chem Soc.* **2003**, *125* (8), 2319-2327.

127. Petrovic, M.; Barcelo, D. Analysis of ethoxylated nonionic surfactants and their metabolites by liquid chromatography/atmospheric pressure ionization mass spectrometry. *J Mass Spectrom.* **2001**, *36* (11), 1173-1185.

128. Teravainen, M. J.; Pakarinen, J. M. H.; Wickstrom, K.; Vainiotalo, P. Comparison of the composition of Russian and North Sea crude oils and their eight distillation fractions studied by negative-ion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry: The effect of suppression. *Energ Fuel.* **2007**, *21* (1), 266-273.

129. Mendes, M. A.; Souza, B. M.; Marques, M. R.; Palma, M. S. The effect of glycerol on signal supression during electrospray ionization analysis of proteins. *Spectrosc-Int J.* **2004**, *18* (2), 339-345.

130. Pastorekova, S.; Casini, A.; Scozzafava, A.; Vullo, D.; Pastorek, J.; Supuran, C. T. Carbonic anhydrase inhibitors: The first selective, membrane-impermeant inhibitors targeting the tumor-associated isozyme IX. *Bioorg Med Chem Lett.* **2004**, *14* (4), 869-873.

131. Scozzafava, A.; Owa, T.; Mastrolorenzo, A.; Supuran, C. T. Anticancer and antiviral sulfonamides. *Curr Med Chem.* **2003**, *10* (11), 925-953.

132. Poulsen, S. A. Direct screening of a dynamic combinatorial library using mass spectrometry. *J Am Soc Mass Spectr.* **2006**, *17* (8), 1074-1080.

133. Chen, R. F.; Kernohan, J. C. Combination of Bovine Carbonic Anhydrase with a Fluorescent Sulfonamide. *J Biol Chem.* **1967**, *242* (24), 5813.

134. Mocharla, V. P.; Colasson, B.; Lee, L. V.; Roper, S.; Sharpless, K. B.; Wong, C. H.; Kolb, H. C. In situ click chemistry: Enzyme-generated inhibitors of carbonic anhydrase II. *Angew Chem Int Edit.* **2005**, *44* (1), 116-120.

135. Vu, H.; Pham, N. B.; Quinn, R. J. Direct screening of natural product extracts using mass spectrometry. *J Biomol Screen.* **2008**, *13* (4), 265-275.

136. Fernandez, A.; Sinanoglu, O. Denaturation of Proteins in Methanol Water Mixtures. *Biophys Chem.* **1985**, *21* (3-4), 163-166.

137. Povey, J. F.; Smales, C. M.; Hassard, S. J.; Howard, M. J. Comparison of the effects of 2,2,2-trifluoroethanol on peptide and protein structure and function. *J*

Struct Biol. **2007**, *157* (2), 329-338.

138. Camp, D.; Davis, R. A.; Campitelli, M.; Ebdon, J.; Quinn, R. J. Drug-like Properties: Guiding Principles for the Design of Natural Product Libraries. *J Nat Prod.* **2012**, *75* (1), 72-81.

139. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expres Purif.* **2005**, *41* (1), 207-234.

140. Johnson, B. H.; Hecht, M. H. Recombinant Proteins Can Be Isolated from Escherichia-Coli-Cells by Repeated Cycles of Freezing and Thawing. *Bio-Technol.* **1994**, *12* (13), 1357-1360.

141. Miller, M. C.; Resnick, J. B.; Smith, B. T.; Lovett, C. M. The Bacillus subtilis dinR gene codes for the analogue of Escherichia coli LexA - Purification and characterization of the DinR protein. *J Biol Chem.* **1996**, *271* (52), 33502-33508.

142. Viode, A.; Dagany, X.; Kerleroux, M.; Dugourd, P.; Doussineau, T.; Charles, L.; Antoine, R. Coupling of size-exclusion chromatography with electrospray ionization charge-detection mass spectrometry for the characterization of synthetic polymers of ultra-high molar mass. *Rapid Commun Mass Sp.* **2016**, *30* (1), 132-136.

143. Green, M. K.; Vestling, M. M.; Johnston, M. V.; Larsen, B. S. Distinguishing small molecular mass differences of proteins by mass spectrometry. *Anal Biochem.* **1998**, *260* (2), 204-211.

144. Guvenalp, Z.; Kilic, N.; Kazaz, C.; Kaya, Y.; Demirezer, L. O. Chemical constituents of Galium tortumense. *Turk J Chem.* **2006**, *30* (4), 515-523.

145. Gao, Y.; Chen, Z. Y.; Liang, X.; Xie, C.; Chen, Y. F. Anti-atherosclerotic effect of geniposidic acid in a rabbit model and related cellular mechanisms. *Pharm Biol.* **2015**, *53* (2), 280-285.

146. Carrillo-Ocampo, D.; Bazaldua-Gomez, S.; Bonilla-Barbosa, J. R.; Aburto-Amar, R.; Rodriguez-Lopez, V. Anti-Inflammatory Activity of Iridoids and Verbascoside Isolated from Castilleja tenuiflora. *Molecules.* **2013**, *18* (10), 12109-12118.

147. Luciano, J. H. S.; Lima, M. A. S.; Silveira, E. R.; Vasconcelos, I. M.; Fernandes, G. S.; de Souza, E. B. ANTIFUNGAL IRIDOIDS, TRITERPENES AND PHENOL COMPOUNDS FROM Alibertia myrciifolia SPRUNGE EX. SCHUM. *Quim Nova.* **2010**, *33* (2), 292-294.

148. Gournelis, D.; Skaltsounis, A. L.; Tillequin, F.; Koch, M.; Pusset, J.; Labarre, S. Plants of New-Caledonia .121. Iridoids and Alkaloids from Plectronia-Odorata. *J Nat Prod.* **1989**, *52* (2), 306-316.

149. Kiem, P. V.; Cai, X. F.; Van Minh, C.; Lee, J. J.; Kim, Y. H. Kaurane-type diterpene glycoside from the stem bark of Acanthopanax trifoliatum. *Planta Med.* **2004**, *70* (3), 282-284.

150. Amimoto, K.; Yoshikawa, K.; Arihara, S. Triterpenoid Saponins of Aquifoliaceous Plants .12. Ilexosides Xlvi-Li from the Leaves of Ilex-Rotunda Thunb. *Chem Pharm Bull.* **1993**, *41* (1), 77-80.

151. Al-Dhabi, N. A.; Arasu, M. V.; Park, C. H.; Park, S. U. An up-to-Date Review of Rutin and Its Biological and Pharmacological Activities. *Excli J.* **2015**, *14*, 59-63.

152. Nicollier, G.; Thompson, A. C. Flavonoids of *Desmanthus-Illinoensis*. *J Nat Prod*. **1983**, *46* (1), 112-117.
153. Chung, S. K.; Kim, Y. C.; Takaya, Y.; Terashima, K.; Niwa, M. Novel flavonol glycoside, 7-O-methyl mearnsitrin, from *Sageretia theezans* and its antioxidant effect. *J Agr Food Chem*. **2004**, *52* (15), 4664-4668.
154. Winekenstadde, D.; Angelis, A.; Waltenberger, B.; Schwaiger, S.; Tchoumtchoua, J.; Konig, S.; Werz, O.; Aligiannis, N.; Skaltsounis, A. L.; Stuppner, H. Phytochemical Profile of the Aerial Parts of *Sedum sediforme* and Anti-inflammatory Activity of Myricitrin. *Nat Prod Commun*. **2015**, *10* (1), 83-88.
155. Qin, M.; Luo, Y.; Meng, X. B.; Wang, M.; Wang, H. W.; Song, S. Y.; Ye, J. X.; Pan, R. L.; Yao, F.; Wu, P.; Sun, G. B.; Sun, X. B. Myricitrin attenuates endothelial cell apoptosis to prevent atherosclerosis: An insight into PI3K/Akt activation and STAT3 signaling pathways. *Vasc Pharmacol*. **2015**, *70*, 23-34.
156. Miyazawa, M.; Hisama, M. Antimutagenic activity of flavonoids from *Chrysanthemum morifolium*. *Biosci Biotech Bioch*. **2003**, *67* (10), 2091-2099.
157. Kim, H. J.; Lee, W.; Yun, J. M. Luteolin Inhibits Hyperglycemia-Induced Proinflammatory Cytokine Production and Its Epigenetic Mechanism in Human Monocytes. *Phytother Res*. **2014**, *28* (9), 1383-1391.
158. Jeon, S. H.; Chun, W.; Choi, Y. J.; Kwon, Y. S. Cytotoxic constituents from the bark of *Salix hulteni*. *Archives of Pharmacal Research*. **2008**, *31* (8), 978-982.
159. Cavia-Saiz, M.; Busto, M. D.; Pilar-Izquierdo, M. C.; Ortega, N.; Perez-Mateos, M.; Muniz, P. Antioxidant properties, radical scavenging activity and biomolecule protection capacity of flavonoid naringenin and its glycoside naringin: a comparative study. *J Sci Food Agr*. **2010**, *90* (7), 1238-1244.
160. Park, J. H.; Lee, J. W.; Paik, H. D.; Cho, S. G.; Nah, S. Y.; Park, Y. S.; Han, Y. S. Cytotoxic Effects of 7-O-Butyl Naringenin on Human Breast Cancer MCF-7 Cells. *Food Sci Biotechnol*. **2010**, *19* (3), 717-724.
161. Agrawal, P. K. Determining ring-F configuration in spirostane-type steroidal sapogenins by (1)H NMR. *Indian J Chem B*. **2005**, *44* (5), 1092-1094.
162. Yin, Y.; Zhao, X. C.; Wang, S. J.; Gao, P. Y.; Li, L. Z.; Ikejima, T.; Song, S. J. Synthesis and biological evaluation of novel sarsapogenin derivatives as potential anti-tumor agents. *Steroids*. **2015**, *93*, 25-31.
163. Jung, H. A.; Yoon, N. Y.; Bae, H. J.; Min, B. S.; Choi, J. S. Inhibitory Activities of the Alkaloids from *Coptidis Rhizoma* against Aldose Reductase. *Archives of Pharmacal Research*. **2008**, *31* (11), 1405-1412.
164. Kong, W. J.; Zhao, Y. L.; Xiao, X. H.; Jin, C.; Liu, Y.; Li, Z. L. Comparison of Anti-bacterial Activity of Four Kinds of Alkaloids in *Rhizoma Coptidis* Based on Microcalorimetry. *Chinese J Chem*. **2009**, *27* (6), 1186-1190.
165. Slobodnikova, L.; Kost'alova, D.; Labudova, D.; Kotulova, D.; Kettman, V. Antimicrobial activity of *Mahonia aquifolium* crude extract and its major isolated alkaloids. *Phytother Res*. **2004**, *18* (8), 674-676.
166. Han, T.; Li, H. L.; Zhang, Q. Y.; Zheng, H. C.; Qin, L. P. New thiazinediones and

other components from *Xanthium strumarium*. *Chem Nat Compd*. **2006**, 42 (5), 567-570.

167. Sato, Y.; Itagaki, S.; Kurokawa, T.; Ogura, J.; Kobayashi, M.; Hirano, T.; Sugawara, M.; Iseki, K. In vitro and in vivo antioxidant properties of chlorogenic acid and caffeic acid. *Int J Pharmaceut*. **2011**, 403 (1-2), 136-138.

168. Sung, W. S.; Lee, D. G. Antifungal action of chlorogenic acid against pathogenic fungi, mediated by membrane disruption. *Pure Appl Chem*. **2010**, 82 (1), 219-226.

169. Guo, Y. J.; Luo, T.; Wu, F.; Mei, Y. W.; Peng, J.; Liu, H.; Li, H. R.; Zhang, S. L.; Dong, J. H.; Fang, Y.; Zhao, L. Involvement of TLR2 and TLR9 in the anti-inflammatory effects of chlorogenic acid in HSV-1-infected microglia. *Life Sci*. **2015**, 127, 12-18.

170. Liu, R.; Zhang, H. C.; Yuan, M. S.; Zhou, J.; Tu, Q.; Liu, J. J.; Wang, J. Y. Synthesis and Biological Evaluation of Apigenin Derivatives as Antibacterial and Antiproliferative Agents. *Molecules*. **2013**, 18 (9), 11496-11511.

171. Choudhury, D.; Ganguli, A.; Dastidar, D. G.; Acharya, B. R.; Das, A.; Chakrabarti, G. Apigenin shows synergistic anticancer activity with curcumin by binding at different sites of tubulin. *Biochimie*. **2013**, 95 (6), 1297-1309.

172. Blunder, M.; Pferschy-Wenzig, E. M.; Fabian, W. M. F.; Hufner, A.; Kunert, O.; Saf, R.; Schuhly, W.; Bauer, R. Derivatives of schisandrin with increased inhibitory potential on prostaglandin E-2 and leukotriene B-4 formation in vitro. *Bioorgan Med Chem*. **2010**, 18 (7), 2809-2815.

173. Rigano, D.; Formisano, C.; Basile, A.; Lavitola, A.; Senatore, F.; Rosselli, S.; Brtmo, M. Antibacterial activity of flavonoids and phenylpropanoids from *Marrubium globosum* ssp *libanoticum*. *Phytother Res*. **2007**, 21 (4), 395-397.

174. Berger, Y.; Castonguay, A. C-13 Nuclear Magnetic-Resonance Spectra of Anthraquinone, 8 Polyhydroxyanthraquinones and 8 Polymethoxyanthraquinones. *Org Magn Resonance*. **1978**, 11 (8), 375-377.

175. Bennett, M.; Cresswell, H. Factors influencing constipation in advanced cancer patients: a prospective study of opioid dose, dantron dose and physical functioning. *Palliative Med*. **2003**, 17 (5), 418-422.

176. Muzzarelli, R. A. A. Genipin-crosslinked chitosan hydrogels as biomedical and pharmaceutical aids. *Carbohydr Polym*. **2009**, 77 (1), 1-9.

177. Chaipukdee, N.; Kanokmedhakul, K.; Kanokmedhakul, S.; Lekphrom, R.; Pyne, S. G. Two new bioactive iridoids from *Rothmannia wittii*. *Fitoterapia*. **2016**, 113, 97-101.

178. Son, M.; Lee, M.; Ryu, E.; Moon, A.; Jeong, C. S.; Jung, Y. W.; Park, G. H.; Sung, G. H.; Cho, H.; Kang, H. Genipin as a novel chemical activator of EBV lytic cycle. *J Microbiol*. **2015**, 53 (2), 155-165.

179. Liu, G. S.; Dong, Q. L.; Yao, Y. S.; Yao, Z. J. Expeditious Total Syntheses of Camptothecin and 10-Hydroxycamptothecin. *Org Lett*. **2008**, 10 (23), 5393-5396.

180. Ping, Y. H.; Lee, H. C.; Lee, J. Y.; Wu, P. H.; Ho, L. K.; Chi, C. W.; Lu, M. F.; Wang, J. J. Anticancer effects of low-dose 10-hydroxycamptothecin in human colon cancer.

Oncol Rep. **2006**, *15* (5), 1273-1279.

181. Rayyan, S.; Fossen, T.; Nateland, H. S.; Andersen, O. M. Isolation and identification of flavonoids, including flavone rotamers, from the herbal drug 'Crataegi folium cum flore' (hawthorn). *Phytochem Analysis.* **2005**, *16* (5), 334-341.
182. Shimoda, K.; Kubota, N.; Taniuchi, K.; Sato, D.; Nakajima, N.; Hamada, H.; Hamada, H. Biotransformation of naringin and naringenin by cultured *Eucalyptus perriniana* cells. *Phytochemistry.* **2010**, *71* (2-3), 201-205.
183. Paradowska, K.; Wolniak, M.; Pisklak, M.; Glinski, J. A.; Davey, M. H.; Wawer, I. C-13, N-15 CPMAS NMR and GIAO DFT calculations of stereoisomeric oxindole alkaloids from Cat's Claw (*Uncaria tomentosa*). *Solid State Nucl Mag.* **2008**, *34* (4), 202-209.
184. Zhou, J. Y.; Zhou, S. W. Antihypertensive and neuroprotective activities of rhynchophylline: The role of rhynchophylline in neurotransmission and ion channel activity. *J Ethnopharmacol.* **2010**, *132* (1), 15-27.
185. Gnoatto, S. C. B.; Dassonville-Klimpt, A.; Da Nascimento, S.; Galera, P.; Boumediene, K.; Gosmann, G.; Sonnet, P.; Moslemi, S. Evaluation of ursolic acid isolated from *Ilex paraguariensis* and derivatives on aromatase inhibition. *Eur J Med Chem.* **2008**, *43* (9), 1865-1877.
186. Taketa, A. T. C.; Breitmaier, E.; Schenkel, E. P. Triterpenes and triterpenoidal glycosides from the fruits of *Ilex paraguariensis* (Mate). *J Brazil Chem Soc.* **2004**, *15* (2), 205-211.
187. do Nascimento, P. G. G.; Lemos, T. L. G.; Bizerra, A. M. C.; Arriaga, A. M. C.; Ferreira, D. A.; Santiago, G. M. P.; Braz Filho, R.; Costa, J. G. M. Antibacterial and Antioxidant Activities of Ursolic Acid and Derivatives. *Molecules.* **2014**, *19* (1), 1317-1327.
188. Zhou, Y. H.; Gao, G. C.; Li, H.; Qu, J. P. A convenient method to reduce hydroxyl-substituted aromatic carboxylic acid with NaBH(4)/Me(2)SO(4)/B(OMe)(3). *Tetrahedron Lett.* **2008**, *49* (20), 3260-3263.
189. Alencar, J.; Gosset, G.; Robin, M.; Pique, V.; Culcasi, M.; Clement, J. L.; Mercier, A.; Pietri, S. Improving the Stability and Antioxidant Properties of Sesame Oil: Water-Soluble Spray-Dried Emulsions from New Transesterified Phenolic Derivatives. *J Agr Food Chem.* **2009**, *57* (16), 7311-7323.
190. Jung, H. J.; Song, Y. S.; Lim, C. J.; Park, E. H. Anti-angiogenic, Anti-inflammatory and Anti-nociceptive Activities of Vanillyl Alcohol. *Archives of Pharmacol Research.* **2008**, *31* (10), 1275-1279.
191. Cabral, H.; Nakanishi, M.; Kumagai, M.; Jang, W. D.; Nishiyama, N.; Kataoka, K. A Photo-Activated Targeting Chemotherapy Using Glutathione Sensitive Camptothecin-Loaded Polymeric Micelles. *Pharm Res-Dord.* **2009**, *26* (1), 82-92.
192. Zhou, H. B.; Liu, G. S.; Yao, Z. J. Highly efficient and mild cascade reactions triggered by bis(triphenyl)oxodiphosphonium trifluoromethanesulfonate and a concise total synthesis of camptothecin. *Org Lett.* **2007**, *9* (10), 2003-2006.
193. Luzzio, M. J.; Besterman, J. M.; Emerson, D. L.; Evans, M. G.; Lackey, K.;

- Leitner, P. L.; McIntyre, G.; Morton, B.; Myers, P. L.; Peel, M.; Sisco, J. M.; Sternbach, D. D.; Tong, W. Q.; Truesdale, A.; Uehling, D. E.; Vuong, A.; Yates, J. Synthesis and Antitumor-Activity of Novel Water-Soluble Derivatives of Camptothecin as Specific Inhibitors of Topoisomerase-I. *J Med Chem.* **1995**, *38* (3), 395-401.
194. Hasan, A.; Sadiq, A.; Abbas, A.; Mughal, E.; Khan, K. M.; Ali, M. Isolation and synthesis of flavonols and comparison of their antioxidant activity. *Nat Prod Res.* **2010**, *24* (11), 995-1003.
195. Rasku, S.; Wahala, K. Synthesis of deuterium labeled polyhydroxy flavones and 3-flavonols. *Tetrahedron.* **2000**, *56* (6), 913-916.
196. Zhuo, W. L.; Zhang, L.; Zhu, Y.; Zhu, B.; Chen, Z. T. Fisetin, a dietary bioflavonoid, reverses acquired Cisplatin-resistance of lung adenocarcinoma cells through MAPK/Survivin/Caspase pathway. *Am J Transl Res.* **2015**, *7* (10), 2045-2052.
197. Zandi, K.; Teoh, B. T.; Sam, S. S.; Wong, P. F.; Mustafa, M. R.; AbuBakar, S. In vitro antiviral activity of Fisetin, Rutin and Naringenin against Dengue virus type-2. *J Med Plants Res.* **2011**, *5* (23), 5534-5539.
198. Zheng, Y. J.; Jiang, X.; Gao, F.; Song, J. X.; Sun, J. X.; Wang, L. X.; Sun, X. X.; Lu, Z. H.; Zhang, H. Y. Identification of plant-derived natural products as potential inhibitors of the Mycobacterium tuberculosis proteasome. *Bmc Complem Altern M.* **2014**, *14*.
199. Cai, Q. Y.; Li, W. R.; Wei, J. J.; Mi, S. Q.; Wang, N. S. Antinociceptive activity of aqueous and alcohol extract of *Evodia rutaecarpa*. *Indian J Pharm Sci.* **2014**, *76* (3), 235-239.
200. Wu, X. A.; Ying, P.; Liu, J. Y.; Shen, H. S.; Chen, Y.; He, L. Lithium Chloride-Assisted Selective Hydrolysis of Methyl Esters Under Microwave Irradiation. *Synthetic Commun.* **2009**, *39* (19), 3459-3470.
201. Rao, G. S. R. S.; Kondaiah, P.; Singh, S. K.; Ramanan, P.; Sporn, M. B. Chemical modifications of natural triterpenes-glycyrrhetic and boswellic acids: evaluation of their biological activity. *Tetrahedron.* **2008**, *64* (51), 11541-11548.
202. Li, Y.; Feng, L.; Song, Z. F.; Li, H. B.; Huai, Q. Y. Synthesis and Anticancer Activities of Glycyrrhetic Acid Derivatives. *Molecules.* **2016**, *21* (2).
203. Schwarz, S.; Csuk, R. Synthesis and antitumour activity of glycyrrhetic acid derivatives. *Bioorgan Med Chem.* **2010**, *18* (21), 7458-7474.
204. Xu, C.; Barchet, T. M.; Mager, D. E. Quantitative structure-property relationships of camptothecins in humans. *Cancer Chemoth Pharm.* **2010**, *65* (2), 325-333.
205. Xu, S.; Shang, M. Y.; Liu, G. X.; Xu, F.; Wang, X.; Shou, C. C.; Cai, S. Q. Chemical Constituents from the Rhizomes of *Smilax glabra* and Their Antimicrobial Activity. *Molecules.* **2013**, *18* (5), 5265-5287.
206. Naik, H.; Lehr, J. E.; Pienta, K. J. Inhibition of prostate cancer growth by 9-aminocamptothecin and estramustine. *Urology.* **1996**, *48* (3), 508-511.
207. Hong, B.; Cheng, W. M.; Wu, J. A.; Zhao, C. J. Screening and Identification of Many of the Compounds Present in *Rauvolfia verticillata* by Use of High-Pressure LC

- and Quadrupole TOF MS. *Chromatographia*. **2010**, 72 (9-10), 841-847.
208. Duflos, A.; Redoules, F.; Fahy, J.; Jacquesy, J. C.; Jouannetaud, M. P. Hydroxylation of yohimbine in superacidic media: One-step access to human metabolites 10 and 11-hydroxy-yohimbine. *J Nat Prod*. **2001**, 64 (2), 193-195.
 209. Rosengren, A. H.; Jokubka, R.; Tojjar, D.; Granhall, C.; Hansson, O.; Li, D. Q.; Nagaraj, V.; Reinbothe, T. M.; Tuncel, J.; Eliasson, L.; Groop, L.; Rorsman, P.; Salehi, A.; Lyssenko, V.; Luthman, H.; Renstrom, E. Overexpression of Alpha2A-Adrenergic Receptors Contributes to Type 2 Diabetes. *Science*. **2010**, 327 (5962), 217-220.
 210. Wang, X.; Wu, X. J.; Cheng, N.; Zhao, H. Q.; Gu, Z. H.; Shen, X. A. New Route for Conversion of Camptothecin to 7-Ethylcamptothecin and 7-Propylcamptothecin. *Synthetic Commun*. **2007**, 37 (1-3), 519-523.
 211. Nagata, H.; Kaneda, N.; Furuta, T.; Sawada, S.; Yokokura, T.; Miyasaka, T.; Fukada, M.; Notake, K. Action of 7-Ethylcamptothecin on Tumor-Cells and Its Disposition in Mice. *Cancer Treat Rep*. **1987**, 71 (4), 341-348.
 212. Sharma, P.; Gupta, Y. K.; Sharma, M. C.; Dobhal, M. P. Two new compounds from the stem of Nerium oleander. *Indian J Chem B*. **2010**, 49 (3), 374-378.
 213. Siddiqui, S.; Hafeez, F.; Begum, S.; Siddiqui, B. S. Oleanderol, a New Pentacyclic Triterpene from the Leaves of Nerium-Oleander. *J Nat Prod*. **1988**, 51 (2), 229-233.
 214. Kumar, D.; Mallick, S.; Vedasiromoni, J. R.; Pal, B. C. Anti-leukemic activity of Dillenia indica L. fruit extract and quantification of betulinic acid by HPLC. *Phytomedicine*. **2010**, 17 (6), 431-435.
 215. Fujioka, T.; Kashiwada, Y.; Kilkuskie, R. E.; Cosentino, L. M.; Ballas, L. M.; Jiang, J. B.; Janzen, W. P.; Chen, I. S.; Lee, K. H. Anti-Aids Agents .11. Betulinic Acid and Platanic Acid as Anti-Hiv Principles from Syzigium-Claviflorum, and the Anti-Hiv Activity of Structurally Related Triterpenoids. *J Nat Prod*. **1994**, 57 (2), 243-247.
 216. Steele, J. C. P.; Warhurst, D. C.; Kirby, G. C.; Simmonds, M. S. J. In vitro and in vivo evaluation of betulinic acid as an antimalarial. *Phytother Res*. **1999**, 13 (2), 115-119.
 217. Alakurtti, S.; Makela, T.; Koskimies, S.; Yli-Kauhaluoma, J. Pharmacological properties of the ubiquitous natural product betulin. *Eur J Pharm Sci*. **2006**, 29 (1), 1-13.
 218. Fotie, J.; Bohle, D. S.; Leimanis, M. L.; Georges, E.; Rukunga, G.; Nkengfack, A. E. Lupeol long-chain fatty acid esters with antimalarial activity from Holarrhena floribunda. *J Nat Prod*. **2006**, 69 (1), 62-67.
 219. Gallo, M. S., M. Biological Activities of Lupeol. *International Journal of Biomedical and Pharmaceutical Sciences* **2009**, 46-66.
 220. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. Ligplot - a Program to Generate Schematic Diagrams of Protein Ligand Interactions. *Protein Eng*. **1995**, 8 (2), 127-134.
 221. Mehlin, C.; Boni, E.; Buckner, F. S.; Engel, L.; Feist, T.; Gelb, M. H.; Haji, L.; Kim, D.; Liu, C.; Mueller, N.; Myler, P. J.; Reddy, J. T.; Sampson, J. N.; Subramanian, E.; Van

- Voorhis, W. C.; Worthey, E.; Zucker, F.; Hol, W. G. J. Heterologous expression of proteins from *Plasmodium falciparum*: Results from 1000 genes. *Mol Biochem Parasit.* **2006**, *148* (2), 144-160.
222. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem.* **2006**, *49* (21), 6177-6196.
223. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem.* **2004**, *47* (7), 1739-1749.
224. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem.* **2004**, *47* (7), 1750-1759.
225. Chen, I. J.; Foloppe, N. Drug-like Bioactive Structures and Conformational Coverage with the LigPrep/ConfGen Suite: Comparison to Programs MOE and Catalyst. *J Chem Inf Model.* **2010**, *50* (5), 822-839.
226. Li, C. W.; Cui, C. B. One New and Nine Known Flavonoids from *Choerospondias axillaries* and Their in Vitro Antitumor, Anti-Hypoxia and Antibacterial Activities. *Molecules.* **2014**, *19* (12), 21363-21377.
227. Kalaitzis, J. A.; Leone, P. D. A.; Hooper, J. N. A.; Quinn, R. J. Ianthesine E, a new bromotyrosine-derived metabolite from the Great Barrier Reef sponge *Pseudoceratina* sp. *Nat Prod Res.* **2008**, *22* (14), 1263-1269.
228. Kim, D. H.; Hung, T. M.; Bae, K. H.; Jung, J. W.; Lee, S.; Yoon, B. H.; Cheong, J. H.; Ko, K. H.; Ryu, J. H. Gomisins A improves scopolamine-induced memory impairment in mice. *Eur J Pharmacol.* **2006**, *542* (1-3), 129-135.
229. Zhou, T. T.; Fan, G. R.; Hong, Z. Y.; Chai, Y. F.; Wu, Y. T. Large-scale isolation and purification of geniposide from the fruit of *Gardenia jasminoides* Ellis by high-speed counter-current chromatography. *J Chromatogr A.* **2005**, *1100* (1), 76-80.
230. Cai, L.; Li, R.; Tang, W. J.; Meng, G.; Hu, X. Y.; Wu, T. N. Antidepressant-like effect of geniposide on chronic unpredictable mild stress-induced depressive rats by regulating the hypothalamus-pituitary-adrenal axis. *Eur Neuropsychopharm.* **2015**, *25* (8), 1332-1341.
231. Jun, M.; Fu, H. Y.; Hong, J.; Wan, X.; Yang, C. S.; Ho, C. T. Comparison of antioxidant activities of isoflavones from kudzu root (*Pueraria lobata* Ohwi). *J Food Sci.* **2003**, *68* (6), 2117-2122.
232. Xi, J.; Guo, R. Interactions of puerarin with micelles: pK(a) shifts and thermodynamics. *J Solution Chem.* **2008**, *37* (1), 107-118.
233. Xiao, C.; Li, J.; Dong, X. X.; He, X. J.; Niu, X. Y.; Liu, C.; Zhong, G. Y.; Bauer, R.; Yang, D. J.; Lu, A. P. Anti-oxidative and TNF- α suppressive activities of puerarin derivative (4AC) in RAW264.7 cells and collagen-induced arthritic rats. *Eur J Pharmacol.* **2011**, *666* (1-3), 242-250.

234. Olejniczak, S.; Potrzebowski, M. J. Solid state NMR studies and density functional theory (DFT) calculations of conformers of quercetin. *Org Biomol Chem.* **2004**, *2* (16), 2315-2322.
235. Gordon, M. H.; Roedig-Penman, A. Antioxidant activity of quercetin and myricetin in liposomes. *Chem Phys Lipids.* **1998**, *97* (1), 79-85.
236. Martin, R.; Schmidt, A. W.; Theumer, G.; Krause, T.; Entchev, E. V.; Kurzchalia, T. V.; Knolker, H. J. Synthesis and biological activity of the (25R)-cholesten-26-oic acids-ligands for the hormonal receptor DAF-12 in *Caenorhabditis elegans*. *Org Biomol Chem.* **2009**, *7* (5), 909-920.
237. Wu, G. W.; Gao, J. M.; Shi, X. W.; Zhang, Q.; Wei, S. P.; Ding, K. Microbial Transformations of Diosgenin by the White-Rot Basidiomycete *Coriolus versicolor*. *J Nat Prod.* **2011**, *74* (10), 2095-2101.
238. PATEL, K. G., M.; TAHILYANI, V.; PATEL, D. A review on pharmacological and analytical aspects of diosgenin: a concise report. *Nat. Prod. Bioprospect.* **2012**, *2*, 46-52.
239. Mallavadhani, U. V.; Mahapatra, A.; Raja, S. S.; Manjula, C. Antifeedant activity of some pentacyclic triterpene acids and their fatty acid ester analogues. *J Agr Food Chem.* **2003**, *51* (7), 1952-1955.
240. Acebey-Castellon, I. L.; Voutquenne-Nazabadioko, L.; Huong, D. T. M.; Roseau, N.; Bouthagane, N.; Muhammad, D.; Debar, E. L. M.; Gangloff, S. C.; Litaudon, M.; Sevenet, T.; Nguyen, V. H.; Lavaud, C. Triterpenoid Saponins from *Symplocos lancifolia*. *J Nat Prod.* **2011**, *74* (2), 163-168.
241. Nakamura, N. Inhibitory effects of some traditional medicines on proliferation of HIV-1 and its protease. *Yakugaku Zasshi.* **2004**, *124* (8), 519-529.
242. Jimenez, A.; Meckes, M.; Alvarez, V.; Torres, J.; Parra, R. Secondary metabolites from *Chamaedorea tepejilote* (Palmae) are active against *Mycobacterium tuberculosis*. *Phytother Res.* **2005**, *19* (4), 320-322.
243. Widmalm, G.; Jansson, K.; Pellijeff, G.; Sandstrom, D. Probing segmental mobility in the cyanogenic glycoside amygdalin by C-13 solid-state NMR. *J Phys Chem B.* **2003**, *107* (42), 11794-11798.
244. Liu, E. H.; Qi, L. W.; Peng, Y. B.; Cheng, X. L.; Wu, Q.; Li, P.; Li, C. Y. Rapid separation and identification of 54 major constituents in Buyang Huanwu decoction by ultra-fast HPLC system coupled with DAD-TOF/MS. *Biomed Chromatogr.* **2009**, *23* (8), 828-842.
245. Song, Z. Q.; Xu, X. H. Advanced research on anti-tumor effects of amygdalin. *J Cancer Res Ther.* **2014**, *10* (5), 3-7.
246. Guo, J. Q.; Wu, W. Z.; Sheng, M. X.; Yang, S. L.; Tan, J. M. Amygdalin inhibits renal fibrosis in chronic kidney disease. *Mol Med Rep.* **2013**, *7* (5), 1453-1457.
247. Chang, E. J.; Lee, W. J.; Cho, S. H.; Choi, S. W. Proliferative effects of flavan-3-ols and propelargonidins from rhizomes of *Drynaria fortunei* on MCF-7 and osteoblastic cells. *Arch Pharm Res.* **2003**, *26* (8), 620-630.
248. Tomeckova, V.; Guzy, J.; Kusnir, J.; Fodor, K.; Marekova, M.; Chavkova, Z.;

- Perjesi, P. Comparison of the effects of selected chalcones, dihydrochalcones and some cyclic flavonoids on mitochondrial outer membrane determined by fluorescence spectroscopy. *J Biochem Bioph Meth.* **2006**, 69 (1-2), 143-150.
249. Amaro, M. I.; Rocha, J.; Vila-Real, H.; Eduardo-Figueira, M.; Mota-Filipe, H.; Sepodes, B.; Ribeiro, M. H. Anti-inflammatory activity of naringin and the biosynthesised naringenin by naringinase immobilized in microstructured materials in a model of DSS-induced colitis in mice. *Food Res Int.* **2009**, 42 (8), 1010-1017.
250. Al Arni, S.; Drake, A. F.; Del Borghi, M.; Converti, A. Study of Aromatic Compounds Derived from Sugarcane Bagasse. Part I: Effect of pH. *Chem Eng Technol.* **2010**, 33 (6), 895-901.
251. Pan, S. D., H; Chang, W; Lin, H. Phenols from the aerial parts of *Leonurus sibiricus*. *Chinese Pharmaceutical Journal (Taipei, Taiwan).* **2006**, 58 (1), 35-40.
252. Karthik, G. A., M.; VijayaKumar, A.; Natarajapillai, S. Syringic acid exerts antiangiogenic activity by downregulation of VEGF in zebrafish embryos. *Biomedicine & Preventive Nutrition* **2014**, 203-208.
253. Chong, K. P. In Vitro Antimicrobial Activity and Fungitoxicity of Syringic Acid, Caffeic Acid and 4-hydroxybenzoic Acid against *Ganoderma Boninense*. *Journal of Agricultural Science.* **2009**, 1 (2), 15-20.
254. Shi, C.; Sun, Y.; Zheng, Z. W.; Zhang, X. R.; Song, K. K.; Jia, Z. Y.; Chen, Y. F.; Yang, M. C.; Liu, X.; Dong, R.; Xia, X. D. Antimicrobial activity of syringic acid against *Cronobacter sakazakii* and its effect on cell membrane. *Food Chem.* **2016**, 197, 100-106.
255. Elgammal, A. A.; Mansour, R. M. A. Antimicrobial Activities of Some Flavonoid Compounds. *Zbl Mikrobiol.* **1986**, 141 (7), 561-565.
256. Zhang, Y. J.; Peng, C. W.; Xu, S. Q.; Zheng, Y. Y.; Xie, Q. J. [Optimization of extraction on flavonoids from leaves of *Heritiera littoralis*]. *Zhong Yao Cai.* **2012**, 35 (4), 638-641.
257. Park, K. S.; Chong, Y.; Kim, M. K. Myricetin: biological activity related to human health. *Appl Biol Chem.* **2016**, 59 (2), 259-269.
258. Son, J. K.; Jung, S. J.; Jung, J. H.; Fang, Z.; Lee, C. S.; Seo, C. S.; Moon, D. C.; Min, B. S.; Kim, M. R.; Woo, M. H. Anticancer constituents from the roots of *Rubia cordifolia* L. *Chem Pharm Bull.* **2008**, 56 (2), 213-216.
259. Itokawa, H.; Qiao, Y. F.; Takeya, K. Anthraquinones, Naphthoquinones and Naphthohydroquinones from *Rubia-Oncotricha*. *Phytochemistry.* **1991**, 30 (2), 637-640.
260. McDonald, A. G.; Tipton, K. F. Fifty-five years of enzyme classification: advances and difficulties. *Febs J.* **2014**, 281 (2), 583-592.
261. Mirza, S. P.; Halligan, B. D.; Greene, A. S.; Olivier, M. Improved method for the analysis of membrane proteins by mass spectrometry. *Physiol Genomics.* **2007**, 30 (1), 89-94.
262. Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their

Exclusion in Bioassays. *J Med Chem.* **2010**, 53 (7), 2719-2740.

263. Baell, J. B.; Ferrins, L.; Falk, H.; Nikolakopoulos, G. PAINS: Relevance to Tool Compound Discovery and Fragment-Based Screening. *Aust J Chem.* **2013**, 66 (12), 1483-1494.

264. Baell, J. B. Feeling Nature's PAINS: Natural Products, Natural Product Drugs, and Pan Assay Interference Compounds (PAINS). *J Nat Prod.* **2016**, 79 (3), 616-628.

265. Bisson, J.; McAlpine, J. B.; Friesen, J. B.; Chen, S. N.; Graham, J.; Pauli, G. F. Can Invalid Bioactives Undermine Natural Product-Based Drug Discovery? *J Med Chem.* **2016**, 59 (5), 1671-1690.

266. Pouliot, M.; Jeanmart, S. Pan Assay Interference Compounds (PAINS) and Other Promiscuous Compounds in Antifungal Research. *J Med Chem.* **2016**, 59 (2), 497-503.

267. Glaser, J.; Holzgrabe, U. Focus on PAINS: false friends in the quest for selective anti-protozoal lead structures from Nature? *Medchemcomm.* **2016**, 7 (2), 214-223.