



If you have discovered material in AURA which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown Policy](#) and [contact the service](#) immediately

Computer-Aided Design, Synthesis and Screening of Potential Anti-Microbial Compounds

Jiangmeng Ren

Doctor of Philosophy

**Aston University
August 2009**

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Aston University
Computer-Aided Design, Synthesis and Screening of
Potential Anti-Microbial Compounds

A thesis submitted by Jiangmeng Ren for the degree of Doctor of Philosophy
August 2009

Summary:

The work presented in this thesis falls into three main categories: The design and synthesis of potential anti-tuberculosis drugs targeting a mycobacterial esterase and the enzyme dUTPase; synthesis and anti-microbial SAR studies on a set of carboxamidrazones; synthesis and anti-microbial SAR studies on a set of thiosemicarbazones.

M-TB (*Mycobacterium tuberculosis*) is associated with HIV and has killed millions of people. M-TB has developed to MDR-TB (multi-drug resistant TB), even XDR-TB (extensively drug-resistant TB). All these facts show that developing new effective anti-TB drugs is an urgent issue. In this study, homology modelling of a mycobacterial esterase and using virtual screening techniques with docking molecule libraries into the mycobacterial esterase and another mycobacterial enzyme dUTPase were used to design and develop the potential drugs which may have anti-TB properties. Six types of compounds were designed using virtual screening techniques and seven compounds from two types were successfully synthesised and isolated. Those compounds have been submitted to the TAACF for screening against *M. tuberculosis* H37Rv.

Following some previous work, four types of carboxamidrazones were synthesised and tested against *M-TB* H37Rv and other organisms such as *E. coli* and *S. aureus*. Three *N*¹-(4-alkylbenzoyl)-pyridine-2-carboxamidrazones where the alkyl group was pentyl, hexyl and heptyl had the lowest IC₉₀ values of 1.4, 1.3 and 1.5 µg/mL respectively against M-TB H37Rv. However, the cytotoxicity assay data shows that those compounds were too cytotoxic to use as drugs. Therefore, the further development of these kinds of drugs with low cytotoxicity is required.

Eleven thiosemicarbazones were also synthesised and tested against both *M-TB* H37Rv and the Gram-positive bacteria *Clostridium difficile*, *S. epidermidis*, *S. aureus* MRSA and *P. acnes*. One compound, *N*¹-Ethyl-2-(5-nitro-2-furylidene)-allylic-1-carbothioamide, was found to have strong activity against all those kinds of Gram-positive bacteria. At the same time nearly all those compounds were found to be active against *M-TB* H37Rv with the lowest IC₉₀ of 2.1 µg/mL. However, those compounds were also cytotoxic and could not be used as drugs. These results have been published in *Bioorg. Med. Chem. Lett.*, 2008, 16, 879-883.

Acknowledgments

I would like to thank Dr Dan. L. Rathbone for this constant support and guidance throughout the course of this work and for all his encouragement, patience and understanding.

Many thanks to Prof. Peter Lambert for his help with the antibacterial screening.

I would like to express my gratitude for the help, technical assistance and advice from Mike Davis and Karen Farrow.

I also would like to thanks to Chatzakis for the work on dUTPase, Chong (Winnie) for the help on synthetic work and Chauhan on biological screening work.

Thanks to everyone in the medicinal chemistry LAB who kept me going when things went pear-shaped, Dr Alexandre Mongeot, Xiaomei Hu, Mike Davis, Soumick Sikdar. Thank you also to Dr Bill Fraser, for the advice on chemistry and life in general.

Finally, I would like to thank my Mum and Dad for their support and encouragement throughout my life in the UK. Also may thanks to my wife who gives me support and encouragement during my study.

List of contents

	Page
Title Page	1
Summary	2
Acknowledgements	3
List of Contents	4
List of Tables	7
List of Figures	9
List of Schemes	11
Abbreviation	12
 Chapter 1 Molecular Modelling 	
1.1 Introduction	13
1.1.1 Background of Tuberculosis	13
1.1.2 Background of Carboxyl Esterase Rv3487c	18
1.1.3 Background of Homology Modelling	22
1.1.4 Introduction of Molecular Mechanics	26
1.1.5 Introduction of Molecular Dynamics	31
1.1.6 Introduction of Quantum Mechanics	41
1.1.7 Introduction of Docking and Scoring	45
1.1.7.1 <i>Scoring functions</i>	45
1.1.7.2 <i>Docking functions</i>	49
1.1.8 Design of Inhibitors of Mycobacterial dUTPase	51
1.2 Result and Discussion	54
1.2.1 Homology modelling	54
1.2.2 Docking Results and Discussion	74
1.2.2.1 <i>Docking potential ligands into the esterase homology models</i>	74
1.2.2.1.1 <i>Initial docking</i>	74

	Page	
1.2.2.1.2	<i>The first generation of docking</i>	92
1.2.2.1.3	<i>Further generations of docking</i>	97
1.2.2.1.3.1	<i>The second generation of BTB02557</i>	99
1.2.2.1.3.2	<i>The third and fourth generations of BTB02557</i>	100
1.2.2.1.3.3	<i>The fifth generation of BTB02557</i>	104
1.2.2.1.3.4	<i>The second and third generation of XBX00012</i>	108
1.2.2.2	<i>Docking potential ligands into dUTPase</i>	110
1.3	Conclusion	120
1.4	Recommendation	121
1.5	Experimental Parameters	122

Chapter 2 Synthesis of the designed compounds

2.1	Introduction	123
2.1.1	Designed Compounds as Potential Inhibitors of the Mycobacterial Carboxyl Esterase Rv3487c	123
2.1.1.1	<i>Intended synthetic route of Type 1A compounds</i>	124
2.1.1.2	<i>Intended synthetic route of Type 1B compounds</i>	125
2.1.1.3	<i>Intended synthetic route of Type 1C compounds</i>	125
2.1.2	Designed Compounds as Potential Inhibitors of the Mycobacterial dUTPase	127
2.1.2.1	<i>Intended synthetic route of Type 2A compounds</i>	128
2.1.2.2	<i>Intended synthetic route of Type 2B compounds</i>	129
2.1.2.3	<i>Intended synthetic route of Type 2C compounds</i>	130
2.2	Results and Discussion	132
2.2.1	Synthesis of Potential Inhibitors of the Mycobacterial Esterase Rv3487c	132
2.2.1.1	<i>Synthesis of Type 1A compounds</i>	132
2.2.1.2	<i>Synthesis of Type 1B compounds</i>	135
2.2.1.3	<i>Synthesis of Type 1C compounds</i>	136
2.2.2	Synthesis of Potential Inhibitors of the Mycobacterial dUTPase	148

	Page
2.2.2.1	<i>Synthesis of Type 2A compounds</i> 148
2.3	Conclusion 151
2.4	Recommendation 151
2.5	Experimental 153

Chapter 3 Carboxamidrazones synthesis and screening

3.1	Introduction	158
3.2	Results and Discussion	163
3.2.1	Type I compounds	163
3.2.1.1	<i>Chemistry of compounds of Type I</i>	163
3.2.1.2	<i>Microbiology screening of compounds of Type I</i>	164
3.2.2	Type II compounds	166
3.2.2.1	<i>Chemistry of compounds of Type II</i>	166
3.2.2.2	<i>Microbiology screening of compounds of Type II</i>	175
3.2.3	Type III and Type IV compounds	178
3.2.3.1	<i>Chemistry of compounds of Type III and Type IV</i>	178
3.2.3.2	<i>Microbiology screening of compounds of Type III and Type IV</i>	180
3.3	Conclusion and Recommendation	182
3.4	Experimental	183

Chapter 4 Thiosemicarbazones Active Against *C. difficile*

4.1	Introduction	216
4.2	Results and Discussion	220
4.2.1	Chemistry of Thiosemicarbazones	220
4.2.2	Microbiology Screening of Thiosemicarbazones	224
4.3	Conclusion	230
4.4	Experimental	231
	References	238

List of Tables

		Page
1.1.1	Name and structure of first-line and second-line anti-TB drugs	17
1.1.2	Comparison of steepest descents and conjugate gradients methods in different gradient change calculations under the same pc conditions.	30
1.1.3	Summarise of protein motion type and the time scales	36
1.1.4	Ligand atom type	48
1.1.5	Protein atom type	49
1.1.6	Further analogues of T2-3-9	53
1.2.1	PDB number, name and the position related to active site and alpha helix	60
1.2.2	PDB number, name and the position related to active site and alpha helix	71
1.2.3	Initial docking score for the Mybridge reduced set into T1	76
1.2.4	Average docking score of top 10 ranked compounds	82
1.2.5	First generation of designed compounds arising from Mybridge XB00012	84
1.2.6	First generation of designed compounds arising from Mybridge AW00957	86
1.2.7	First generation of designed compounds arising from Mybridge BTB02557	88
1.2.8	First generation of designed compounds arising from Mybridge BTB12775	90
1.2.9	First generation of designed compounds arising from Mybridge SCR00846	92
1.2.10	The ranked docking score by average of first generation of those selected best Mybridge compounds	94
1.2.11	Leading structures of third of the molecular design	97
1.2.12	Docking score of further generation of the molecules	99
1.2.13	Structure of the second generation of the second generation of BTB02557	99
1.2.14	The structure of third and fourth generations of BTB02557	101
1.2.15	Structure and the average docking score of BTB02557 C5_1 series	105
1.2.16	Structure and the average docking score of BTB02557 C5_3 series	106
1.2.17	The next generation of XB00012 compounds	108
1.2.18	LogP and Molecular Weight values of some chosen compounds	109
1.2.19	Structure and the docking score of selected analogues of T2-3-9	112
1.2.20	Top ranked molecules with their structure and docking score information	112

	Page	
1.2.21	The molecules that had unfavourable docking score	117
2.1.1	General structures of designed compounds as potential inhibitors of the mycobacterial carboxyl esterase Rv3487c	123
2.1.2	General structures of designed compounds as potential inhibitors of the mycobacterial dUTPase	127
2.2.1	Structures of the 12 compounds that were planned to be synthesised	138
2.2.2	Compounds were planned to be synthesised by Wittig reactions	148
3.1.1	Active compounds from Billington <i>et. al.</i>	159
3.1.2	Four types of synthesised carboxamidrazones	160
3.1.3	Previous anti-TB testing results	162
3.2.1	Structure with the yield and melting of Type I compounds	164
3.2.2	Biological testing results against <i>E. coli</i> and <i>S. aureus</i>	165
3.2.3	Testing result from TAACF against <i>M. tuberculosis</i> H37Rv	166
3.2.4	Structure with the yield and melting of Type II compounds	170
3.2.5	Biological testing results against <i>E. coli</i> and <i>S. aureus</i>	176
3.2.6	Testing result of Type II compounds against M-TB H37Rv	177
3.2.7	Structure with the yield and melting of Type III and Type IV compounds	179
3.2.8	Testing result of Type III and Type IV compounds against M-TB H37Rv	181
3.2.9	Cytotoxicity assay of Type III and Type IV compounds	181
4.1	Structure of synthesised thiosemicarbazones	219
4.2	Structure of four intermediates	220
4.3	Comparison of thiosemicarbazones' yield and melting point	222
4.4	MIC ($\mu\text{g/mL}$) of all the compounds against an initial panel of organisms	226
4.5	MIC ($\mu\text{g/mL}$) of compound 12 against an expanded panel of organisms	228
4.6	LogP value of 12 compounds	228
4.7	Testing results of thiosemicarbazones against M-TB H37Rv	228
4.8	Cytotoxicity assay of active thiosemicarbazones	229

List of Figures

	Page	
1.1.1	Schematic presentation of the α/β -hydrolase fold. β -sheet (1-8) are shown as dark arrows, α -helices (A-F) as dark columns. The relative position of the amino acids of the catalytic triad are indicated as light circles	19
1.1.2	Amino-acid sequence alignment between Rv3487c, AFestA (abbreviation: AFestA, the carboxyl esterase from <i>A. fulgidus</i>) and AAest2 (abbreviation: the esterase 2 M211S/R215L mutant from the thermophilic eubacterium <i>A. acidocaldarius</i>). The catalytic triad was identified as Ser90-Glu219-His219	21
1.1.3	Two zones of protein sequence alignment	23
1.1.4	Ramachandran plots of Ijji	25
1.1.5	Illustration of steepest descent method in a molecular potential energy surface	29
1.1.6	Illustration of the demerit of Molecular Mechanics method	31
1.1.7	(left) with a very small step, long time calculation is used. (right) with long time step, some of the atom may come together then fly away	36
1.1.8	Illustration of period boundary condition	38
1.1.9	Three ways to treat cut-off condition	39
1.1.10	An example of a molecular dynamic trajectory	41
1.1.11	The best compounds derived from Mybridge compound BTB110716	52
1.2.1	Ramachandran plot for T1	58
1.2.2	Ramachandran plot for T4	59
1.2.3	Attempted correction of T1	61
1.2.4	Attempted correction of T4	62
1.2.5	Ramachandran plot for T2	65
1.2.6	Attempted correction of T4	66
1.2.7	Ramachandran plot for T3	69
1.2.8	Ramachandran plot for T5	70
1.2.9	Five homology models	73
1.2.10	Four examples of docked structures. Only the ligand and the nearby residues are shown. The coloured web shows the calculated surface pocket. The brown thread with numbers shows the bumps between ligand and residues	80
1.2.11	Partial docked structure and the structure of XBX00012	83
1.2.12	Partial docked structure and the structure of AW00957	85
1.2.13	Partial docked structure and the structure of BTB02557	87
1.2.14	Partial docked structure and the structure of BTB12775	89
1.2.15	Partial docked structure and the structure of SCR00846	91
1.2.16	An example of docked structure of compound XBX00012 C1_5	95
1.2.17	Structure comparison of compounds BTB02557 in their docked positions	96

		Page
1.2.18	Partial docked structure of BTB02557 C3_3 into T2	102
1.2.19	Comparison of two docked structures in the fourth generation	103
1.2.20	Partial docked structure in T4 and the molecule of BTB02557 C5_101	105
1.2.21	Partial docked structure of BTB02557 C5_303 in T1	107
1.2.22	Partial docked structure of BTB02557 C5_305 in T2	107
1.2.23	Partial docked structure of BWO12	113
1.2.24	Partial docked structure of BW4	114
1.2.25	Partial docked structure of BW7	115
1.2.26	Partial docked structure of BWO8	116
1.2.27	Partial docked structure of CW1	118
2.2.1	¹ H NMR spectrum and the structure of compound 4 in d ₆ -DMSO	135
2.2.2	Suggested structure of the compound exhibiting a mass spectral ion of m/z-221	139
2.2.3	¹ H NMR (CDCl ₃) of Compounds of 11f , 11g and 11h	144
2.2.4	Predicted fragmentation of compounds 11	145
2.2.5	The ¹ H NMR spectra (CDCl ₃) and the structure of 11e before and after filtration	147
2.2.6	Structure of the intermediate if B11 and B12	150
3.2.1	Comparison of 205 with and without D ₂ O shake in DMSO	172
3.2.2	Partial H ¹ NMR spectrum of 206	174
3.2.3	Partial H ¹ NMR spectrum of 208	174
3.2.4	Possible explanation of the structure associated with (M-17) ⁺ signal	175
4.1	Bar chart of averaged 1/MIC (μg/mL) for compound P1-P12 for the various classes of organisms tested	226

List of Schemes

	Page
1.1.1	Esterase hydrolysis process 20
2.1.1	Intended synthetic route for Type 1A compound 124
2.1.2	Intended synthetic route for Type 1B compound 125
2.1.3	Intended synthetic route for Type 1C compound 126
2.1.4	Intended synthetic route for Type 2A compound 128
2.1.5	Intended synthetic route for Type 2B compound 129
2.1.6	Intended synthetic route for Type 2B compound 130
2.2.1	One of the side reaction of Type 1A compound 132
2.2.2	Intended synthesis route of Type 1A 's alternative 134
2.2.3	Prediction of the reaction between step 2 and step 3 140
3.2.1	Synthesis scheme of benzylidene-pyridine-3-carboxamidrazone-1-N-oxide 163
3.2.2	Synthesis scheme of benzylidene-anthroyl-pyridine-carboxamidrazone 166
3.2.3	Synthesis scheme of pyridine-carboxamidrazone-amides 178
4.1	Synthesis scheme of thiosemicarbazones compounds

Abbreviations

3-D	Three Dimension
AAest2	The esterase 2 M211S/R215L mutant from the thermophilic eubacterium <i>A. acidocaldarius</i>
AFestA	AFestA, the carboxyl esterase from <i>A. fulgidus</i>
AIDS	Acquired Immune Deficiency Syndrome
ALA	Alanine
AMBER	Assisted Model Building with Energy Refinement
APCI-MS	Atmospheric pressure chemical ionisation mass spectrometry
Ar	Aromatic
ARG	Arginine
ASP	Aspartic acid
BuLi	n-Butyllithium
<i>C. albicans</i>	<i>Candida albicans</i>
<i>C. difficile</i>	<i>Clostridium difficile</i>
CC50	Concentration of the drug that cause 50% of cell death
CDAD	<i>Clostridium difficile</i> associated diarrhoea
CDCl ₃	Deuterated chloroform
CYS	Cysteine
d	Doublet
D ₂ O	Heavy water, Deuterium oxide
DCM	Dichloromethane
DIEA	<i>N,N</i> -Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
dUMP	2'-deoxyuridine-5'-monophosphate
dUTP	2'-deoxyuridine-5'-triphosphate
dUTPase	deoxyUridine TriPhosphate nucleotidohydrolase
<i>E. coli</i>	<i>Escherichia coli</i>
<i>E. faecalis</i>	<i>Enterococcus faecalis</i>
<i>E. faecium</i>	<i>Enterococcus faecium</i>
EIAV	Equine Infectious Anaemia Virus
EtOAc	Ethyl Acetate
EtOH	Ethanol
FIV	Feline Immunodeficiency Virus
FQs	fluoroquinolones
GLN	Glutamine
GLU	Glutamic acid
GLY	Glycine
HGP	Human Genome Project
HIS	Histidine
HIV	Human Immunodeficiency Virus
IC50	Half maximal inhibitory concentration
IC90	90% inhibitory concentration
ILE	Isoleucine
IR	Infra-red
K ₂ CO ₃	Potassium Carbonate

KBr	Potassium Bromide
LAB	Laboratory
LEU	Leucine
LYS	Lysine
m	Multiplet
<i>M. fortuitum</i>	<i>Mycobacterium fortuitum</i>
MD	Molecular Dynamics
MDR-TB	Multi-Drug Resistant Tuberculosis
MeOH	Methanol
MET	Methionine
MIC	Minimum Inhibitory Concentration
MM	Molecular Mechanics
MNDO	Modified Neglect of Diatomic Overlap
MOPAC	Molecular Orbital PACKage
MRSA	Methicilin-resistant <i>Staphylococcus aureus</i>
MS	Mass spectrometry
M-TB	<i>Mycobacterium Tuberculosis</i>
MX	moxifloxacin
NaCl	Sodium Chloride
NMR	Nuclear Magnetic Resonance
OPLS	Optimised Potentials for Liquid Simulations
<i>P. acnes</i>	<i>Propionibacterium acnes</i>
PDB	Protein Data Bank
Ph	Phenol
PM3	Parameterised model number 3
PMF	Potential of Mean Force
PRO	Proline
ps	picosecond
q	quartet
QM	Quantum Mechanics
rf	Retention factor
RO5	Rule of Five
s	Singlet
<i>S. epidermis</i>	<i>Staphylococcus epidermidis</i>
<i>S. haemilyticus</i>	<i>Staphylococcus haemilyticus</i>
<i>S.aureus</i>	<i>Staphylococcus aureus</i>
SAR	Structure-activity Relationships
SER	Serine
SI	Selectivity Index
SIV	Simian Immunodeficiency Virus
t	Triplet
TB	Tuberculosis
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
TYR	Tyrosine
VAL	Valine
VDW	Van der Waals
WHO	World Health Organisation
XDR-TB	Extensively Drug-Resistant Tuberculosis

Chapter 1

Molecular Modelling

1.1 Introduction

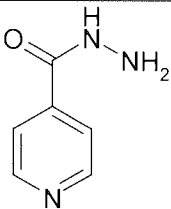
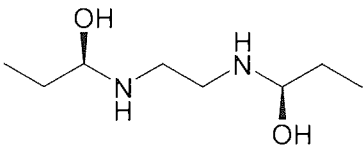
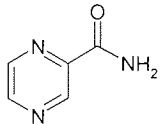
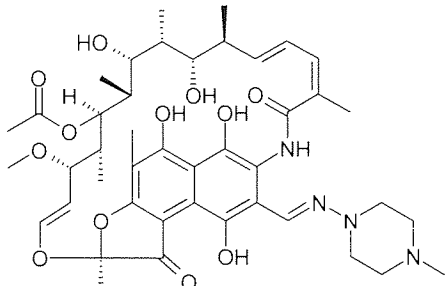
1.1.1 Background of Tuberculosis

According to the WHO (World Health Organisation) report (2008), by the end of 2005, there were around 8.8 million new cases of TB around the world. Most of these patients were in Africa which contributed for 29%, South-East Asia which accounted for 34% and the Western Pacific area which was 22%. From those numbers, it is clear that most of the TB patients live in developing countries. Among all these patients, there were around 1.6 million people who die from TB, including those co-infected with HIV/AIDS. Although the TB incidence per capita was slightly decreasing, offset by population growth, the actual number of patients is still increasing in the regions of Africa, the Eastern Mediterranean and South-East Asia.

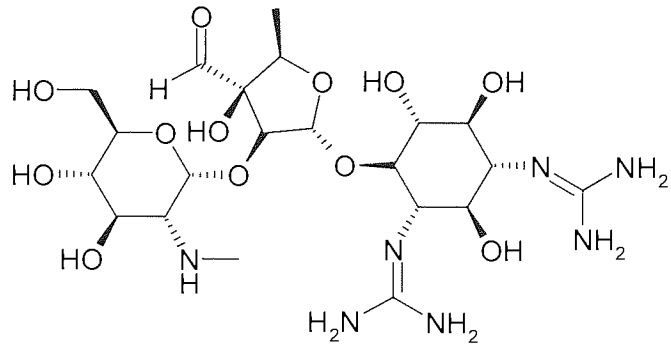
Tuberculosis is contagious and can be spread into the air by cough, sneeze, spit or talk. People nearby can be infected by only a small number of TB bacilli. Even people infected by TB bacilli may not become sick if their immune system works well. In this case, the bacilli go to a dormant or persistent phase and can hide there for years. When some people's immune systems become weaker, typically with the HIV infection, patients are more likely to get sick. At the same time, poor diet can also weaken the immune system and cause sickness. That is why TB patients mostly live in the developing countries. *M. tuberculosis* is difficult to kill and the drugs against it are so limited. This is because *M. tuberculosis* has its own unique complex cell wall so that the drugs have difficulty penetrating into the cell. (Dover 2003) According to Brennan and Nikaido (1995), the *M. tuberculosis* cell wall has unusually low

permeability because it contains a great amount of C60-C90 fatty acids which are peptidoglycan covalently linked to arabinogalactan and in turn attached to mycolic acid. Mycolic acid (MA) is one of the major components in *M. tuberculosis* cell wall. They are unique to mycobacteria and closely aligned genera and constitute 40-60% by dry weight of the cell envelope. They constitute α -alkyl- β -hydroxyl fatty acids of exceptional length and complexity and it appears either covalently associated with the cell wall or non-covalently attached in the form of the glycosylated derivative. (Korf *et. al.* 2005) These MA give the bacterial cell wall strong resistance to degradation by host enzyme, impermeability to toxic macromolecules and inactivation of reactive oxygen and nitrogen derivatives and the ability to interact with host cell receptor to facilitate uptake of the bacterium and to modulate the host immune responses.

There are usually two types of anti-TB drugs namely first-line and second-line anti-TB drugs. They are all shown in **Table 1.1.1.1**.

First-line Anti TB drugs	
Name	Structure
Isoniazid	
Ethambutol	
Pyrazinamide	
Rifampicin	

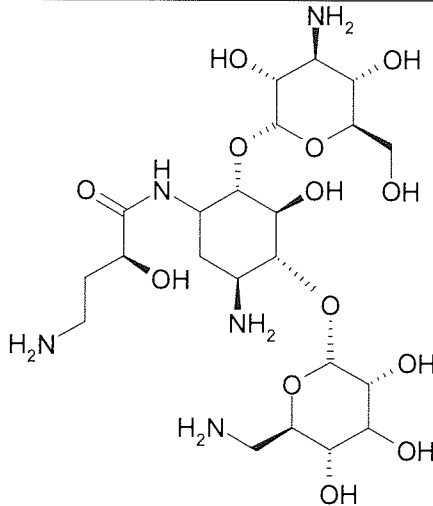
streptomycin



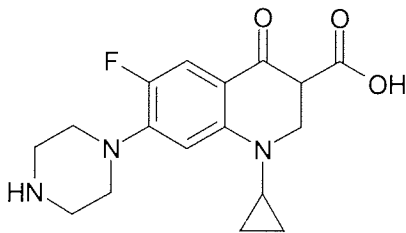
Second-line anti-TB drugs

Name	Structure
Amikacin	 <chem>CC1=CC=C(C=C1)C(=O)N2CCN(C2)C(=O)N3C(O)C(O)C(O)C3N</chem>
Ciprofloxacin	 <chem>CC12CCN(C1)CC2C3=CC=C(C=C3)C(=O)N4C(=O)C(O)C4</chem>
Moxifloxacin	 <chem>CC12CCN(C1)CC2C3=CC=C(C=C3)C(=O)N4C(=O)C(O)C4</chem>
Ethionamide	 <chem>CCc1ccc(cc1)C(=S)N</chem>

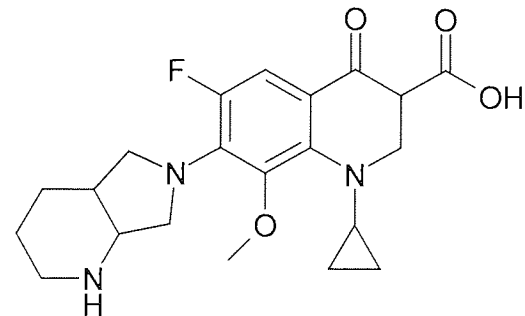
Amikacin



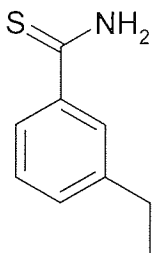
Ciprofloxacin



Moxifloxacin



Ethionamide



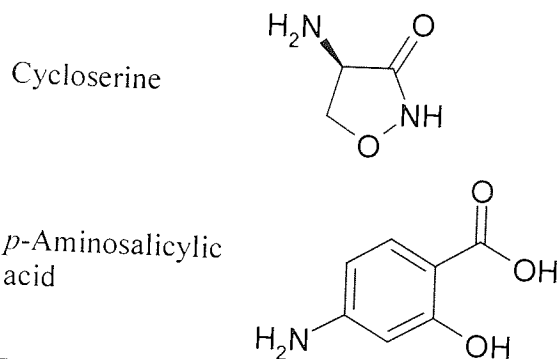


Table 1.1.1: Name and structure of first-line and second-line anti-TB drugs.

The first line anti-TB drugs include isoniazid (1952), ethambutol (1966), pyrazinamide (1970), rifampicin (1965) and streptomycin (1943). The second-line anti-TB drugs are usually less effective with high toxic side-effects or they are too expensive. There are several second-line anti-TB drugs which include amikacin (1999), ciprofloxacin (1983), moxifloxacin (1997) ethionamide (2002), cycloserine, *p*-aminosalicylic acid (1948), *etc.*

If people get sick due to TB, the standard cure course will be conducted using the first line anti-TB drugs. With the combination of those drugs, the whole treatment should take six months. However, during this long period, if the patient does not take the right dose or even not complete the whole course, Multi-drug resistant tuberculosis (MDR-TB) could be developed. Such cases always happen in those people who have weaker immune systems as well as in the poorer nations with failing tuberculosis control programmes. Using a scientific definition, if TB is resistant to isoniazid and rifampicin, those strains will be called multi-drug resistant tuberculosis (MDR-TB). Usually, if a person is ill because of MDR-TB, he or she can be treated using second-line anti-TB drugs. However, this treatment will take a longer time to cure with more side effects and much more cost than the first-line anti-TB drugs. In October 2006, WHO give a new definition of extensively drug-resistant tuberculosis (XDR-TB). Those TB can resistance to at least isoniazid and rifampicin, as well as any member of quinolone family and at least one of the second-line anti-TB drugs

include kanamycin, capreomycin or amikacin (WHO, 2006). XDR-TB is still treatable. However, the success rate was low. A study showed only 48.3% of the XDR-TB patients were successfully treated (Keshavjee S. *et. al.* 2008).

Although we have so many anti-TB drugs, it is noticeable that all of the first-line anti-TB drugs were all developed before 1970, which means that no highly effective anti-TB drugs have been found during the last 30 years. (Arcus *et.al.* 2006) Therefore, there are three major goals which need to be achieved in 21st century to combat *M. tuberculosis* (Barry 1997).

- Develop new drugs to decrease the treatment period in order to reduce the chance for *M. tuberculosis* becoming MDR-TB and kill the organism completely.
- Develop new drugs to treat the already existing drug-resistant bacilli.
- New chemotherapeutics should be developed to specifically target dormant bacilli to treat the people who are latently infected with TB.

1.1.2 Background of Carboxyl Esterase Rv3487c

Carboxyl esterases are active in catalysing the cleavage and formation of ester bonds. Many of them appear in animals, plants and microorganisms. They are thought to evolve to enable access to carbon sources or to be involved in catabolic pathways. Unlike lipases, carboxyl esterases are only involved in the hydrolysis of simple esters, which are usually shorter than C6 (e.g. ethyl acetate and short chain triglycerides). The carboxyl esterase which was used in this molecular modelling study is *Rv3487c* which is found in the organism *M. tuberculosis* (*lipF*, GenBank accession No. NC_000962). The data in Zhang *et.al.*'s (2005) study showed that *Rv3487c* had optimal activity for glycerol acetate and *p*-nitrophenyl acetate. *Rv3487c* could be considered an esterase since the preferential substrates for hydrolysis were water-soluble esters of short chain fatty acids. *Rv3487c* is involved in the pathogenicity of *M. tuberculosis* infection in the host. It is the promoter of granulomas formation

which can resist the elimination of the host's immune system (Zahrt 2003). At the same time, Saviola *et.al.* (2003) reported that a possible function of *lipF*(Rv3487c) could be to hydrolyze toxic fatty acids present in caseous necrotic debris during tuberculosis. Alternatively, it may modify the external cell wall as an adaptive response to acid damage. Therefore, if *Rv3487c* can be inhibited, granulomas should be destroyed easily and can be eliminated by immune system.

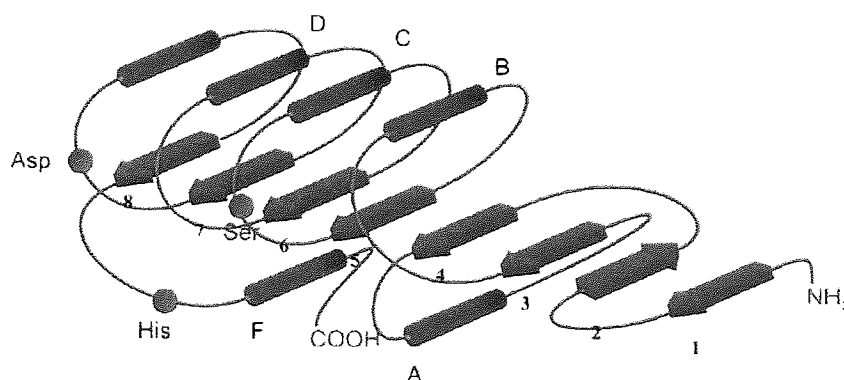
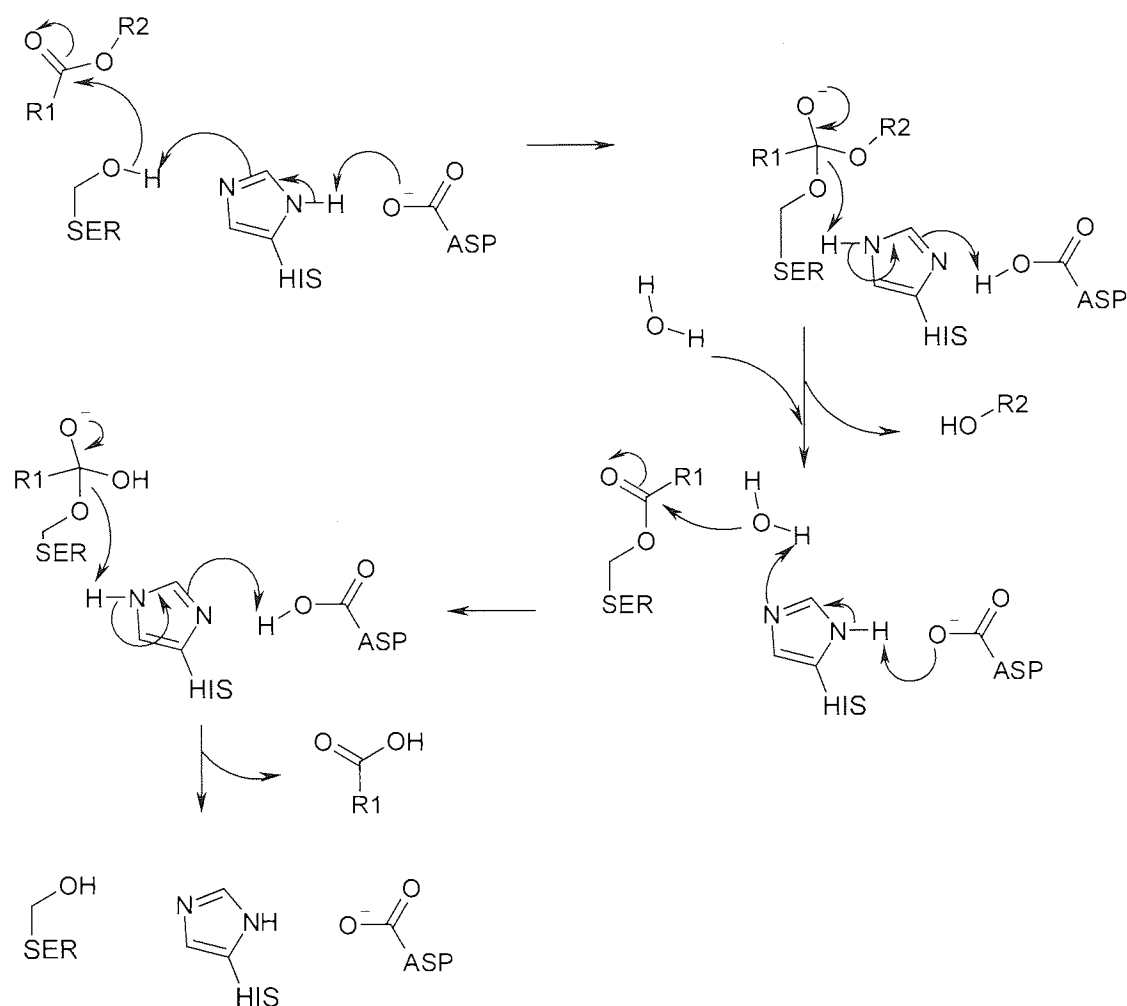


Figure 1.1.1: Schematic presentation of the α/β -hydrolase fold. β -sheet (1-8) are shown as dark arrows, α -helices (A-F) as dark columns. The relative position of the amino acids of the catalytic triad are indicated as light circles (Bornscheuer 2002)

Carboxyl esterases have similar three-dimensional structures which are defined by α -helices and β -sheets fold (**Figure 1.1.1**). The catalytic triad is usually Ser-Asp-His (sometimes Asp can be Glu) (Bornscheuer 2002). The mechanism for forming or breaking the ester bond is simple as it does not require any cofactors. The whole process is a typical four steps chymotrypsin-like hydrolysis mechanism which involves Ser-His-Asp catalytic triad (sometime can be Ser-His-Glu). **Scheme 1.1.1** shows the reaction mechanism.



Scheme 1.1.1: Ester hydrolysis process.

In the first step, the ester is bound to the active serine which is stabilized by histidine and aspartic acid or glutamic acid. The intermediate is tetrahedral and can eliminate an alcohol leaving an acyl-enzyme complex. Then, a nucleophile (water in this case) activated by the catalytic triad attaches to the acyl-enzyme complex to form a tetrahedral form again. Finally, this intermediate collapses to give an acid and a clear catalytic triad (Stryer 1995, Lowe 1992).

Zhang *et.al.* (2005) had the knowledge of the amino acid sequence of the esterase. However, they did not know the 3-D crystal structure. Therefore, an on-line BLAST search was conducted by Zhang *et.al.* And two of the proteins were identified as the

templers. One of them was carboxyl esterase (1jji) from *A. fulgidus* and the other one was Esterase 2M211S/R215L mutant (1qz3A). Esterase (1jji) and *Rv3487c* have around half their residues in common. There are 41% identical and 61.4% similarity between residues 81-168 in 1jji and residues 15-98 in *Rv3487c*. Overall, there are 25.5% identity and 51.6% similarity between residues 77 - 311 in 1jji and residues 8 - 245 in *Rv3487c*. 1qz3A was chosen to be the second chose. They showed that 26.5% identity and 46.3% similarity between residues 8-238 of *Rv3487c* and residues 69-308 of 1qz2A. In Zhang *et.al.*'s paper, the homology was built in order to identify the catalytic triad and it was found in *Rv3487c* to be Ser90-Glu189-His219. **Figure 1.1.2** shows their sequence alignments.

<i>Rv3487c</i>	8	A	ADGAGRULYLHGGAFUMCGPNHSRIUNALSGFAESPULIUDYRLIPK
AFestA	77	P	---DSPULUYHGGGFUICSIESHDLCRRIARLSNSTUUSUDYRLAPE
AAest2	69	G	UEPPYPALUYHGGGWUGDLETHDPUCRULAKDGRAUVFSUDYRLAPE
<i>Rv3487c</i>	58		HSLGMALDDCHDAYQ--WLRARGYRPE--QIULAGDSAGGYLALALAQR
AFestA	124		HKFPAAUVDYDATKWAENAELRIDPSKIFUGGDSAGGNLAAUSIMA
AAest2	119		HKFPAAUVEDAYDALQWIAERAADFHLDPARIAUGGDSAGGNLAAUTSILA
<i>Rv3487c</i>	104	Q	CDDEKPAAIUVAISPLLQLAKG-PKQDHPNIGTDANFPARAFDALAAWU
AFestA	174	R	DSGEDFIKHQILYPUUNFU--APTSLLEFG-EGLWILD--QKIMSWF
AAest2	169	K	ERGGPALAFQLLIYPSTGYDPAHPPASIEENA-E-GYLLT--GGMSLWF
<i>Rv3487c</i>	152	R	AAAAKNMUDGRPEDLYEPLDHI-ESSLPPTLIHUSGSEULLNDAQLGAG
AFestA	219	S	SEQYF-SREEDKFNPLASVIFAD-LENLPPALIITAEYDPLRDEGEVFGQ
AAest2	215	L	DQYL-NSLEELTHPWFSPULYPDLSGLPPAYIATAQYDPLRDUGKLYAE
<i>Rv3487c</i>	201	K	LAAAGUCAEURUWPGQAHLFQLATPLUPEATRSLRQIGQFIRDA
AFestA	267	M	LRRAGUEASIUVRGULHGFINYYPULKAARDAINQIAALLUFD
AAest2	264	A	LNKAGUKVEIENFEDLIHGFAQFYSLSPGATKALURIAEKLRDA

Figure 1.1.2: Amino-acid sequence alignment between *Rv3487c*, AFestA (abbreviation: AFestA, the carboxyl esterase from *A. fulgidus*) and AAest2 (abbreviation: the esterase 2 M211S/R215L mutant from the thermophilic eubacterium *A. acidocaldarius*). The catalytic triad was identified as Ser90-Glu219-His219. (Zhang *et.al.* 2005)

The two proteins and the *Rv3487c* come from same protein family together with their over 25% sequence identity make the homology modelling viable. (Rost 1999).

1.1.3 Background of Homology Modelling

In the recent years, molecular biology techniques have developed rapidly. This resulted in a faster identification, isolation and sequencing of genes such as in The Human Genome Project (HGP) so that more and more protein sequences were inferred in this way. At the same time, other kinds of protein sequencing technologies were developed. All these factors led to a massive number of protein sequences being determined. However, the determination of protein 3D structures is still a time-consuming and a high cost task of using X-ray crystallography or NMR spectroscopy. This cannot catch up with the speed that a great number of proteins have already been sequenced.

Homology modelling is one method that could generate a reasonable protein model. This also called comparative modelling. This method mainly uses a template protein to predict a target protein using its known amino acid sequence and the template protein's structure. The templates are usually homologous to the target protein. As their structure and amino acid sequence are quite similar, the target protein's amino acid sequence can be aligned with the 'template' to predict the 3D structure of the 'target'. Several papers have already reviewed these techniques (Blundell *et. al.* 1987, Sali *et. al.* 1990, Greer 1991, Fetrow & Bryant 1993) and some papers used homology modelling to create their own models (Yang *et. al.* 2004, Minh & Warfel 2004, Xiao *et. al.* 2006).

Before conducting homology modelling, it is important to find out if it is workable. **Figure 1.1.3** shows a general theory of whether a homology modelling is viable or not.

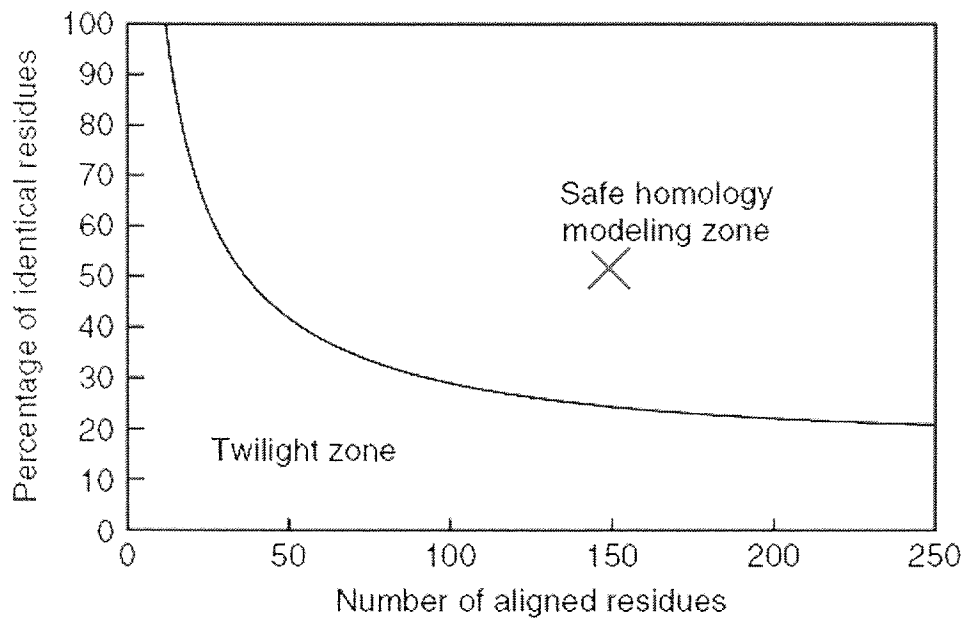


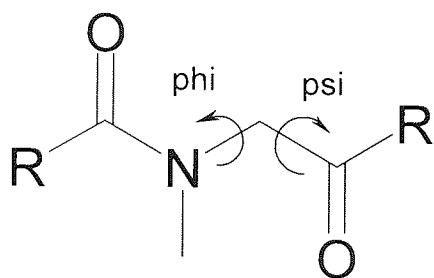
Figure 1.1.3: Two zones of protein sequence alignment. (Kridger et. al. 2003)

Figure 1.1.3 clearly shows two zones which are defined by the number of total sequence and the percentage of identical residues. With a lower number of protein sequence aligned, a higher percentage of identical residues is required. If two alignments have more than 250 residues, with a 25 percent of identical residues, the homology modelling falls into the safe side (Kridger et. al. 2003).

In general, homology modelling is usually a multi-step procedure.

- Identify a template protein determined by both 'target' and 'template' amino acid sequence.
- Sequence alignment.
- Backbone modelling.
- Loops modelling.
- Side chain modelling.
- Refine the final model.
- Validation of the model.

Several on-line database sites can conduct template identification and sequence alignment such as BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and SWISS-MODEL (<http://swissmodel.expasy.org/workspace/index.php>). By inputting the target protein's amino acid sequence, several templates with sequence alignment will be generated automatically. After the sequence is aligned, the building model should start. In this project, CaChe workspace (Version 6.1.10 Fujitsu Limited) was used to generate the structure of the target using Molecular Mechanics, Molecular Dynamics and Quantum Mechanics (MM/DM/QM). There was guidance in this programme about how to create a homology model. However, several modified approaches were used during the modelling. In this case, more than one model was built and used in the later stages. After the homology model was built, a validation method should be carried out to check the final model because all of those models were predictions and they may contain errors. A common validation method is to plot the ϕ/ψ torsion angle into Ramachandran plots (Ramachandran 1963). This is usually done by PROCHECK. **Figure 1.1.4** shows an example of ϕ/ψ torsion angle and a Ramachandran plots of Ijji, which is one of the templates used in this study.



phi-angle is the torsion angle of N-C α

psi-angle is the torsion angle of C α -C

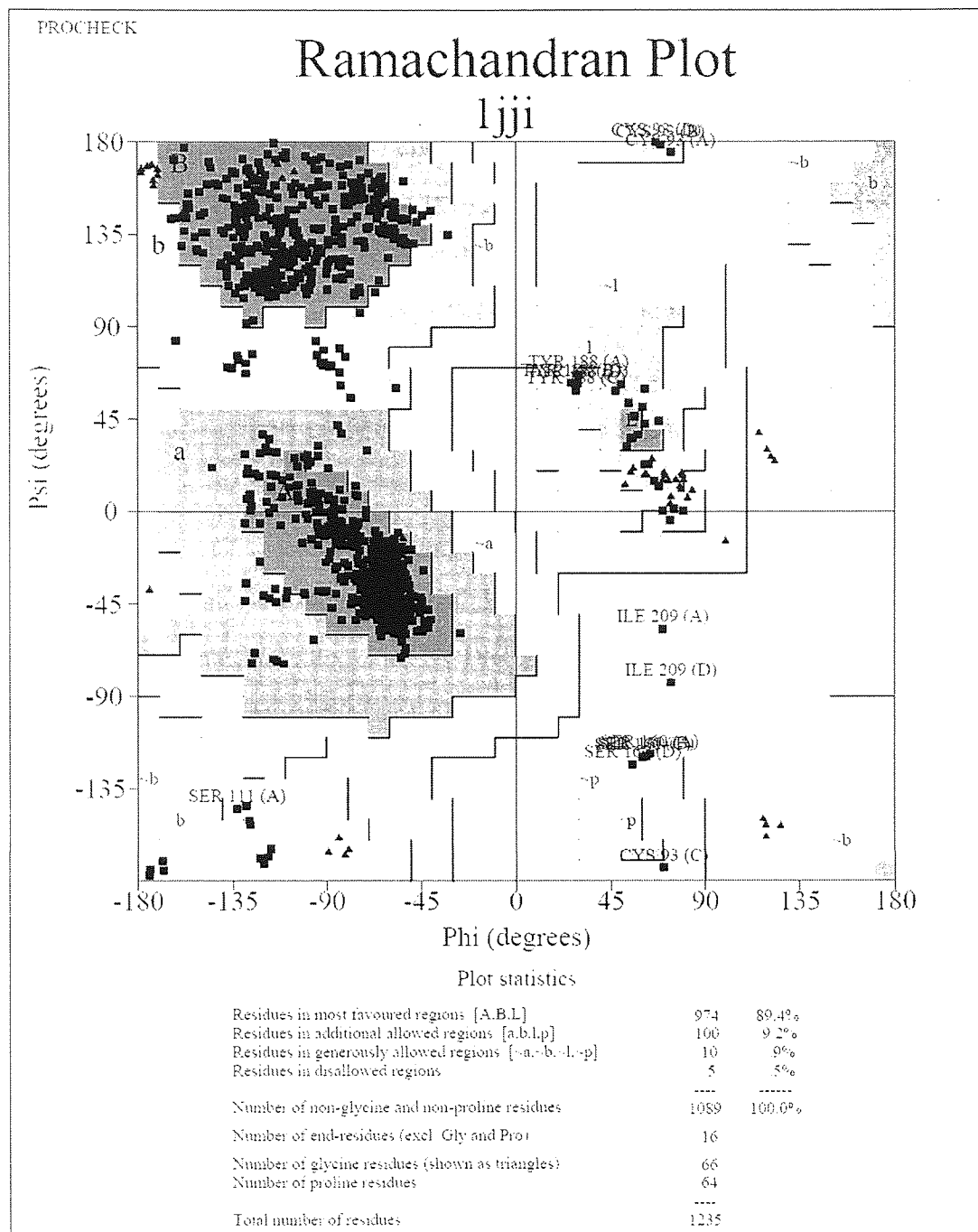


Figure 1.1.4: Ramachandran plots of 1jji.

Typically, a Ramachandran plot of a protein should show that around 90% of the residues should fall into the most favoured regions which are shown above as [A, B, L] and no residues, except glycine and proline, should be in the disallowed regions which are shown in the white area. This was based on an analysis of 118 protein structures of resolution of at least 2.0 angstroms and R-factor no greater than 20%.

The error of the homology model should relate to the initial structure and the sequence alignment. If the percentage of identical residues is greater than 90%, it should contain fewer errors and the final model should be close to a crystallographically determined structure with a few exceptions of side chains (Chothia & Lesk 1986, Sippl 1993). If the identical residues are around 25%, the final structure usually should contain large errors. (Kridger et. al. 2003)

1.1.4: Introduction of Molecular Mechanics

Molecular mechanics methods are simpler than quantum mechanics methods because they use classical Newtonian physical law equations. Although there are several molecular mechanics method such as DREIDING (Mayo S. L., Olafson B. D. & Goddard W. A. III 1990), MM2/MM3 (Allinger, N. L. and others, J. 1977; Allinger N. L., Yuh, Y. H., Lii, J-H., 1989; Bowen, J. P., and Shim, J-Y., 1998), AMBER (Weiner S. J., Kollman P. A., Case D. A. & Singh V. C. 1984), OPLS (Jorgensen W. L. & Tirado-River J. 1988), they each use similar general force field functions such as the following:

$$V = V_{bond} + V_{angles} + V_{torsions} + V_{improper} + V_{electrostatic} + V_{Van\ der\ Waals}$$

In here, V represents the potential energy of the atoms.

$$V_{bond} = \sum \frac{1}{2} k_b (b - b_o)^2 \quad 1.1$$

$$V_{angles} = \sum \frac{1}{2} k_\theta (\theta - \theta_o)^2 \quad 1.2$$

$$V_{torsions} = \sum \frac{1}{2} (1 + \cos(n\phi - \delta)) \quad 1.3$$

$$V_{improper} = \sum \frac{1}{2} k_\psi (\psi - \psi_o)^2 \quad 1.4$$

torsions

$$V_{electrostatic} = \sum \frac{q_i q_j}{4 \pi \epsilon_0 r_{ij}} \quad 1.5$$

$$V_{Van\ der\ Waals} = \sum 4 \epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \quad 1.6$$

In these functions, b_o , θ_o , δ , ψ_o are the minimum values or equilibrium values of each terms. b , θ , φ , ψ are the actual value of the atoms. The electrostatic interactions are described by Coulomb's law where q_i , q_j are the net charge of atom i and atom j and r_{ij} is the distant between them. The Van der Waals' interactions are described by Lennard-Jones 12-6 equation where the r^{-12} term stands for repulsion and the r^{-6} term stands for attraction. (Leach A. R. 2001)

In this study, two different kinds of molecular mechanics methods were used, namely MM3 force field and AMBER99 force field. Except those general functions of bond stretch, bond angle, dihedral angle, improper torsion angle, Van der Waals and electrostatics, MM3 and AMBER all add some extra function terms or definitions. For instance, in the MM3 force field, hydrogen bond, torsion stretch and bend-bend interactions were added in. (CaChe user guide, 2003) In the AMBER99 force field, four different forms were added in. Firstly, a 12-10 potential function was used to describe hydrogen bonds. Secondly, AMBER uses a Coulomb term with a distance-dependent relative permittivity to include solvent effects. Thirdly, AMBER treats carbon and attached hydrogen as a whole. The AMBER99 force field includes sulphur hydrogen bonding. (Hinchliffe 2003)

In molecular mechanics calculation, there are usually two approaches. One of them is steepest descent and other one is conjugate gradients. These two methods are called first-order minimisation methods. The steepest descent method was introduced by Wiberg K. (1965). This method was usually used for initial molecular minimization

when the starting structure was far from the local minimum. That is because by using this method, it can bring the energy down quickly. The steepest descent method works more like a downhill walk as shown in **Figure 1.1.5**. If the initial atom started from point A which is on the molecular potential energy surface, energy in point A will be calculated. After that, a small step of movement along in X, Y and Z axis will be made and each of the energy changes will be calculated. Then, the atom will be moved to new position in a direction which has the most negative energy change. Therefore, atom moves from point A to B. This process will be repeated to point C and D until it reaches the final local minimum. It is noticeable that the step size decreases each time. The step size mainly depends on the nature of potential energy surface. If the surface is relevantly flat, the step size should be big. If it is steep, the step size should be much less. However, this leads to one drawback of the steepest descent method. If the potential energy surface of a structure is too steep, the step size will be very small and calculation will always make right-angle turn in each step which leads to a much longer calculation and it would not be the best route to the minimum.

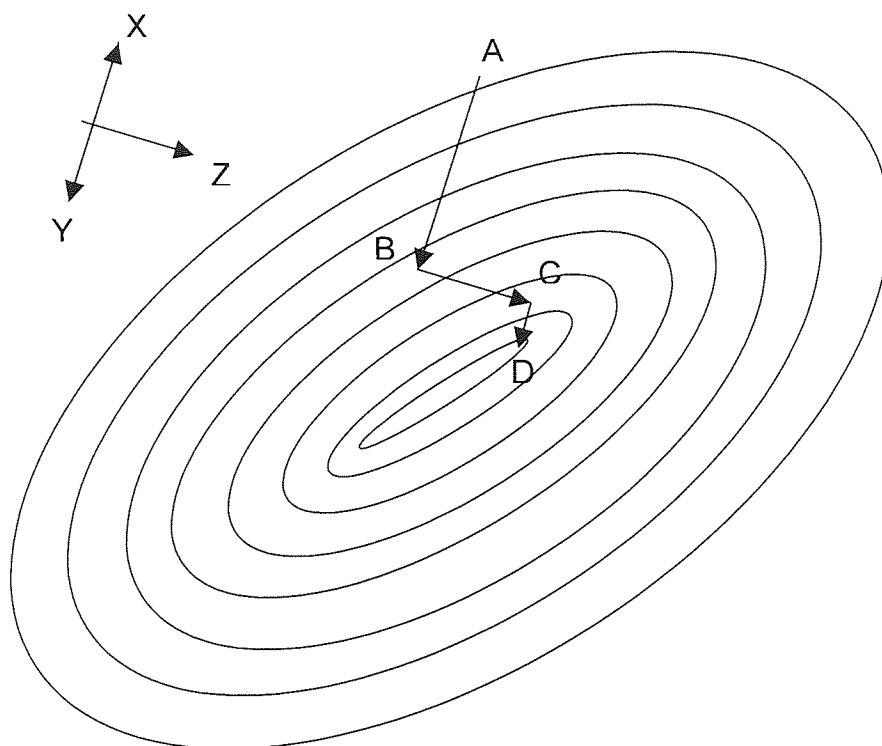


Figure 1.1.5: Illustration of steepest descent method in a molecular potential energy surface.

In 1964, Fletcher and Reeves proposed another algorithm which is called the conjugate gradient method. By using this method in narrow and gully potential energy surfaces, the direction of each turn will not be a right-angle because this method will calculate the direction of movement in each step. Therefore, the drawbacks of steepest descent method will be overcome. In this method, the direction vector V_k in point X_k will be given as:

$$V_k = -g_k + \gamma_k V_{k-1} \quad 1.8$$

In this equation, g_k is the gradient vector at point X_k , V_{k-1} is the direction vector of previous step and γ_k is a scalar which is given by:

$$\gamma_k = \frac{\mathbf{g}_k \cdot \mathbf{g}_k}{\mathbf{g}_{k-1} \cdot \mathbf{g}_{k-1}} \quad 1.9$$

Usually, both of the algorithms are used in a molecular mechanics calculation. At the beginning of the MM calculation, steepest descent will be used because the starting structure may far away from the final point. The steepest descent method can quickly bring the energy down. Then, the conjugate gradients method will be used to continue the calculation where the gradient changes are smaller. In **Table 1.1.2**, Leach (2001) lists a set of data to show the relative performance of these two algorithms.

Method	Initial refinement (AV. Gradient <1 kcal Å ⁻²)		Stringent minimisation (AV. Gradient <0.1 kcal Å ⁻²)	
	CPU time (s)	NO. of iterations	CPU time (s)	NO. of iterations
Steepest descents	67	98	1405	1893
Conjugate gradients	149	213	257	367

Table 1.1.2: Comparison of steepest descents and conjugate gradients methods in different gradient change calculations under the same PC conditions. (Leach 2001)

From **Table 1.1.2**, we can see that at initial minimisation where average gradients change at a higher rate, the steepest descent method uses less CPU time and fewer iterations. However, in a stringent minimisation, the steepest descents method works much worse. It almost uses six times more CPU time and iterations than it does in the conjugate gradients method. Therefore, in order to get a better minimised structure, these two methods are usually implemented together. Steepest descents is used at the beginning of the minimisation to let the energy gradient comes down quickly. After this, the conjugate gradients method is used to quickly get to the final point.

Overall, the purpose of molecular mechanics is to systematically move all the atoms and optimize their position until the net force on each atoms and the global force field reach to their lowest level. Basically, molecular mechanics calculations start with the input sample file. It will calculate the initial potential energy of the atoms required.

Then, it changes the atomic positions and calculates the potential energy of the adjusted structure. Usually molecular mechanics repeats this process for enough iterations so that moving the atoms between iterations will not significantly decrease the potential energy, in another word; the structure reaches its local energy minimum. However, the final structure will depend on the starting structure and sometimes, it will be trapped in the local minimum and cannot reach the global minimum. **Figure 1.1.6** demonstrates this. It is obvious that a molecule starting at conformation A should arrive at a lower minimum energy than a molecule starting at conformation B

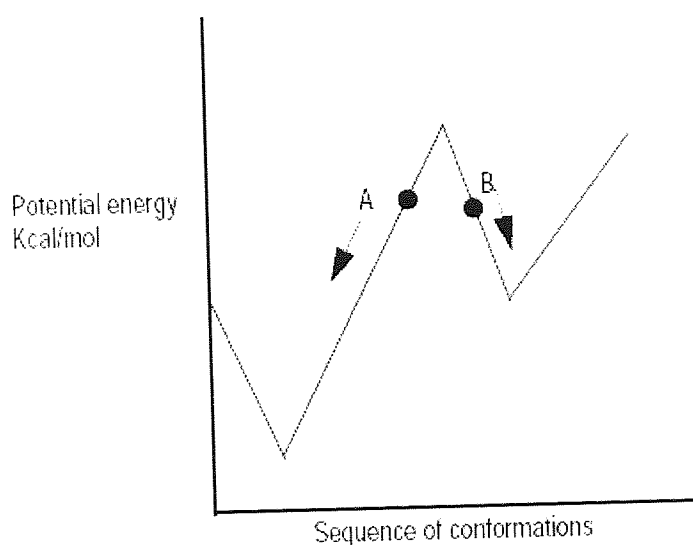


Figure 1.1.6: Illustration of the demerit of Molecular Mechanics method. (CaChe user guide, 2003)

1.1.5 Introduction of Molecular Dynamics

As shown in **Figure 1.1.6**, a structure started from point B will never reach the global minimum energy conformation if only a molecular mechanics calculation is used. That is because molecular mechanics cannot break the energy barrier. However, Molecular Dynamics (MD) can overcome this. Although molecular dynamics use the same force field as molecular mechanics, the molecular motion term with associated

temperature is added in the MD calculation using Newton's second law of motion. ($F=ma$) The result of the MD is a trajectory which describes the position and velocities of the atoms in the system with time steps. The trajectory can be achieved by solving the differential equation of Newton's law:

$$\frac{d^2 x_i}{dt^2} = \frac{F(x_i)}{m_i} \quad 2.1$$

Normally in a MD calculation, the initial structure is read from an input file. The system will give these atoms a random initial velocity by using Maxwell-Boltzmann distribution of velocity:

$$f(v) = 4 \cdot \pi \left(\frac{M}{2\pi RT} \right)^{\frac{3}{2}} \cdot v^2 \cdot \exp\left(\frac{-Mv^2}{2RT} \right) \quad 2.2$$

The force on each atom can be calculated by the force field. Now, we have the force on each atom, the mass of the atoms and their initial velocities. Therefore, the atoms' motion at the new time step $t+\Delta t$ can be calculated. At the $t+\Delta t$ time step, the new forces on the atoms will be calculated with the velocities from the previous calculation. This will result in new positions and velocities in time step $t+2\Delta t$, and so on.

There are several algorithms for integrating Newton's law of motion. Probably, the Verlet algorithms (1969) are the most commonly used method. Although there is more than one algorithm, all of them were derived from Taylor series expansions:

$$r(t + \delta t) = r(t) + \frac{\delta t \cdot v(t)}{1!} + \frac{\delta t^2 \cdot a(t)}{2!} + \frac{\delta t^3 \cdot b(t)}{3!} + \frac{\delta t^4 \cdot c(t)}{4!} + \dots \quad 2.3$$

$$v(t + \delta t) = v(t) + \frac{\delta t \cdot a(t)}{1!} + \frac{\delta t^2 \cdot b(t)}{2!} + \frac{\delta t^3 \cdot c(t)}{3!} + \dots \quad 2.4$$

$$a(t + \delta t) = a(t) + \frac{\delta t \cdot b(t)}{1!} + \frac{\delta t^2 \cdot c(t)}{2!} + \dots \quad 2.5$$

$$b(t + \delta t) = r(t) + \frac{\delta t \cdot c(t)}{1!} + \dots \quad 2.6$$

In the Verlet algorithms, a forward time step ($t+\delta t$) and a backward time step ($t-\delta t$) are used:

$$r(t + \delta t) = r(t) + \frac{\delta t \cdot v(t)}{1!} + \frac{\delta t^2 \cdot a(t)}{2!} + \dots \quad 2.7$$

$$r(t - \delta t) = r(t) - \frac{\delta t \cdot v(t)}{1!} + \frac{\delta t^2 \cdot a(t)}{2!} - \dots \quad 2.8$$

Here v is the velocity and a is the acceleration. By adding these two equations together, it will give

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \quad 2.9$$

We can see that velocity is not included in this equation. However, the velocity can be calculated from the difference between the two time steps $r(t+\delta t)$ and $r(t-\delta t)$.

$$v(t) = \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \quad 2.10$$

The Verlet algorithms are moderately simple and require less memory than other methods. However, it has its own disadvantages. Due to the lack of a velocity term in the equations, it is difficult to get the velocity value until the forward time step $r(t+\delta t)$ has been calculated. Another disadvantage is that the calculation will have a problem at the starting point when $t=0$. It is obvious that it cannot calculate the backward position $r(t-\delta t)$. Therefore, some other methods need to be used to obtain the position at $t-\delta t$.

Several other algorithms have been developed after the Verlet algorithms. One of them is called the leap-frog algorithms. (Hockney 1970) It uses the following equations:

$$r(t + \delta t) = r(t) + \delta t v\left(t + \frac{1}{2} \delta t\right) \quad 2.11$$

$$v\left(t + \frac{1}{2}\delta t\right) = v\left(t - \frac{1}{2}\delta t\right) + \delta t a(t) \quad 2.12$$

In these algorithms, velocities at time $t + \frac{1}{2}\delta t$ are calculated firstly using the velocities at time $t - \frac{1}{2}\delta t$ and the accelerations at time t . The position $r(t + \delta t)$ can then be gained using equation (3.11) and the velocities at time t can be calculated as:

$$v(t) = \frac{1}{2}\left(v\left(t + \frac{1}{2}\delta t\right) + v\left(t - \frac{1}{2}\delta t\right)\right) \quad 2.13$$

In this way, the velocities 'leap-frog' over the position $r(t)$ to give the value $t + \frac{1}{2}\delta t$.

Then the position at $t + \delta t$ leap-frogs over the velocities $t + \frac{1}{2}\delta t$. In the next round,

$v\left(t + \frac{3}{2}\delta t\right)$ leaps over $r(t + \delta t)$ and $r(t + 2\delta t)$ leaps over $v\left(t + \frac{3}{2}\delta t\right)$ and so on. Unlike the

Verlet method, leap-frog algorithms can directly calculate the velocities and it does not require much memory. However, it has its own disadvantage. The velocities and positions are not synchronised. Therefore, it is not possible to calculate the kinetic energy contribution to the total energy at the same time as the positions are defined.

In 1982, velocity Verlet algorithms (Swope *et. al.*) were developed. This method can calculate positions, velocities and acceleration at the same time.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) \quad 2.14$$

$$v(t + \delta t) = v(t) + \frac{1}{2}\delta t(a(t) + a(t + \delta t)) \quad 2.15$$

Basically, there are four steps in this method. Firstly, the new position $r(t + \delta t)$ will be calculated using initial position, velocity and acceleration at time t . Secondly, the velocity at time $t + \frac{1}{2}\delta t$ will be calculated using equation 2.16

$$v\left(t + \frac{1}{2}\delta t\right) = v(t) + \frac{1}{2}\delta t a(t) \quad 2.16$$

Thirdly, the new force at new position $r(t + \delta t)$ can be determined and thus the acceleration at time $t + \delta t$. Finally, the velocity at time $t + \delta t$ will be calculated using equation 3.17

$$v(t + \delta t) = v\left(t + \frac{1}{2}\delta t\right) + \frac{1}{2}\delta t a(t + \delta t) \quad 2.17$$

If equation 2.16 and 2.17 combine together, we will have equation 2.15. Comparing this method to Verlet algorithmic, the velocity term is available in the calculations. Therefore, the kinetic energy at the time $t+\delta t$ can be calculated with the potential energy. The velocity Verlet algorithms are straightforward simple and numerically stable. The main drawback of these algorithms is that it is computationally a little more expensive than Verlet or leap-frog algorithms because it uses four steps in the calculations.

There is another method called Beeman's algorithm (Beeman 1976) which is related to the Verlet algorithms. It uses more accurate equations to calculate the position and the velocity at time step $t+\delta t$.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{2}{3} \cdot \delta t^2 \cdot a(t) - \frac{1}{6} \delta t^2 a(t - \delta t) \quad 2.18$$

$$v(t + \delta t) = v(t) + \frac{1}{3} \cdot \delta t a(t) + \frac{5}{6} \cdot \delta t a(t) - \frac{1}{6} \delta t a(t - \delta t) \quad 2.19$$

Because these equations are more complex than those in Verlet algorithms, this method is computationally more expensive.

In all those algorithms, the selection of the time step δt will be an important parameter. If in a calculation, δt is too big, some of the atoms would move too close to each other in $t+\delta t$ time steps. This will result high energy introduced into the system and finally, these atoms will fly away which will cause the program failure. Therefore, if the time step is too big, the calculation could not be accurate. On the other hand, if δt is too small, it will take a long time to calculate one step of the atom movement and this will waste the computer power. Thus, the choose of δt should give a balance between accuracy and economy. **Figure 1.1.7** gives some illustration of what will happen in different time steps.

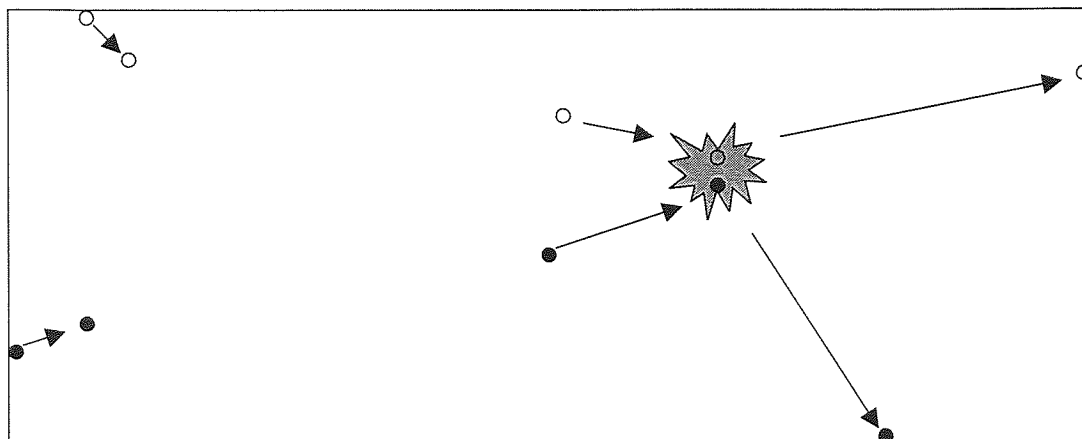


Figure 1.1.7: (left) with a very small step, long time calculation is used. (right) with long time step, some of the atom may come together then scatter.

Inside a protein, different time scales relate to the different type of motion. **Table 1.1.3** shows the different type of motion in the protein with their time scales.

Type of the motion	Functionality examples	Time and Amplitude scales
Local Motions Atomic fluctuation Side chain motion	Ligand docking flexibility, Temporal diffusion pathways	Femtoseconds (fs) to picoseconds (ps) (10^{-15} - 10^{-12} s); less than 1 Å
Medium-scale Motions Loop motion Terminal-arm motion Ridge-body motion	Active site conformation adaptation, Binding specificity	Nanoseconds (ns) to microseconds (μ s) (10^{-9} - 10^{-6} s); 1-5 Å
Large-scale Motion Domain motion Subunit motion	Hinge-bending motion Allosteric transitions	Microseconds (μ s) to milliseconds (ms) (10^{-6} - 10^{-3} s); 5-10 Å
Global motions Helix-coil transition Folding/unfolding Subunit association	Hormone activation Protein functionality	Milliseconds (ms) to hours (10^{-3} - 10^4 s); more than 10 Å

Table 1.1.3 Summary of protein motion type and the time scales. (Becker and Watanabe 2001)

From **Table 1.1.3**, we can see that the motion in the protein can vary from femtoseconds (fs) to hours. Although there is enormous difference in time scale, they

are not independent. Those motions are all coupled with each other which means that large-scale motion cannot occur without medium-scale and local motion. Usually in a molecular dynamics calculation, an appropriate time step should be chosen according to the fastest motion. In a protein, the fastest motion normally is atomic fluctuation which should be the bond stretching between hydrogen and other heavy atoms. Therefore, for a protein simulation, the time step of δt should be around 1 fs. This is because there is a general rule of choosing time step. If τ is the time scale of fastest motion in the protein, an appropriate time step δt should be

$$\frac{\tau}{\delta t} \approx 10 \quad 2.20$$

The frequency of bond stretching between hydrogen and heavy atoms is around 10 fs. Therefore, $\delta t \approx 1$ fs.

When the simulation involves conformational change of the flexible protein, the major conformational changes are caused by the change of bond angle, bond rotation and torsion angle. In this situation, bond stretching will be less important. Therefore, the integration time step will be changed. The fastest motions which are the bond stretching between hydrogen and other heavy atoms should be removed and calculated in some other way. Then, the next level of the fastest molecular motion would be greater than 10 fs and the time step will be greater than 1 fs. The most common method which applied is called SHAKE (Ryckaert, Cicotti and Berensden 1977). This algorithm constrains the bond of hydrogen and other heavy atoms. This is because that in a simulation, these bonds are always fluctuating. Therefore, the deviation of the bond length should be smaller than a certain tolerance value.

When running a MD on a cluster of molecules such as a box of water, the molecules on the surface would scatter under normal conditions. If the simulation took a long time, those water molecules would vaporise completely. In order to contain all the particles while running molecular dynamics, a periodic boundary condition should be

applied. The simplest model of period boundary condition is two-dimensional box surrounded by eight identical boxes. An example has been given in **Figure 1.1.8**.

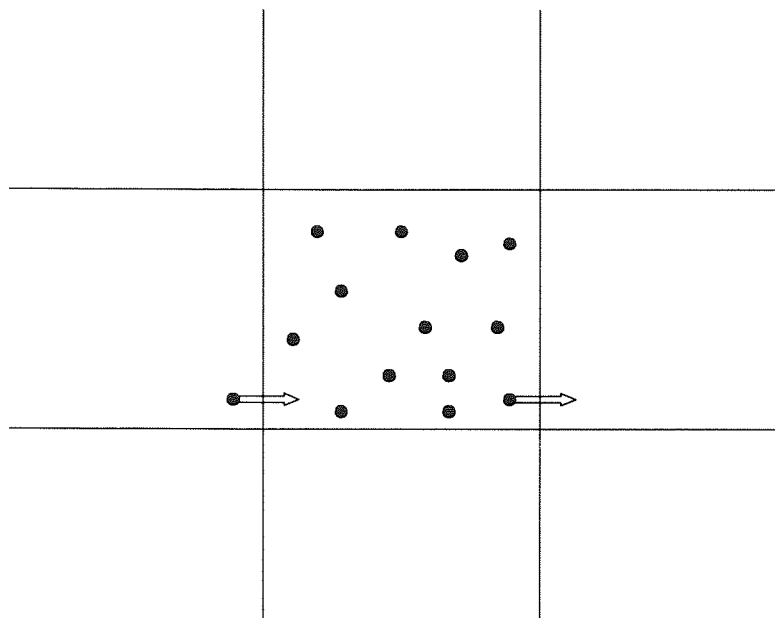


Figure 1.1.8: Illustration of period boundary condition

As shown in **Figure 1.1.8**, if a particle leaves the box during the calculation, an identical particle will enter the box from the opposite side. In this way, the total number of the molecules would not change during the simulation. Although the period boundary conditions are often used in simulation, they have some drawbacks. The first thing to consider is that it could not simulate a fluctuation for which the wavelength is greater than the length of the box. Another problem is dealing with the long-range electrostatic interactions. If the box is too small, the particle may be calculated an interaction with itself (Leach 2001). In some calculations, if the periodic boundary conditions cannot be used, there could be an alternative way by creating a boundary manually. For example, when simulating a small protein in the water, a box of water molecules can be created where the diameter is 10 \AA greater than that of the protein. The outer 5 \AA of the water molecules can act as the ‘shell’ or the ‘skin’ of whole system.

happen because the *van der Waals'* interactions approach zero rapidly when the distance between two atoms increases. By using this condition, the calculation can be speeded up and computer power could be saved. There are several ways to treat the cut-off condition. Figure 2.5 shows three ways to treat this condition.

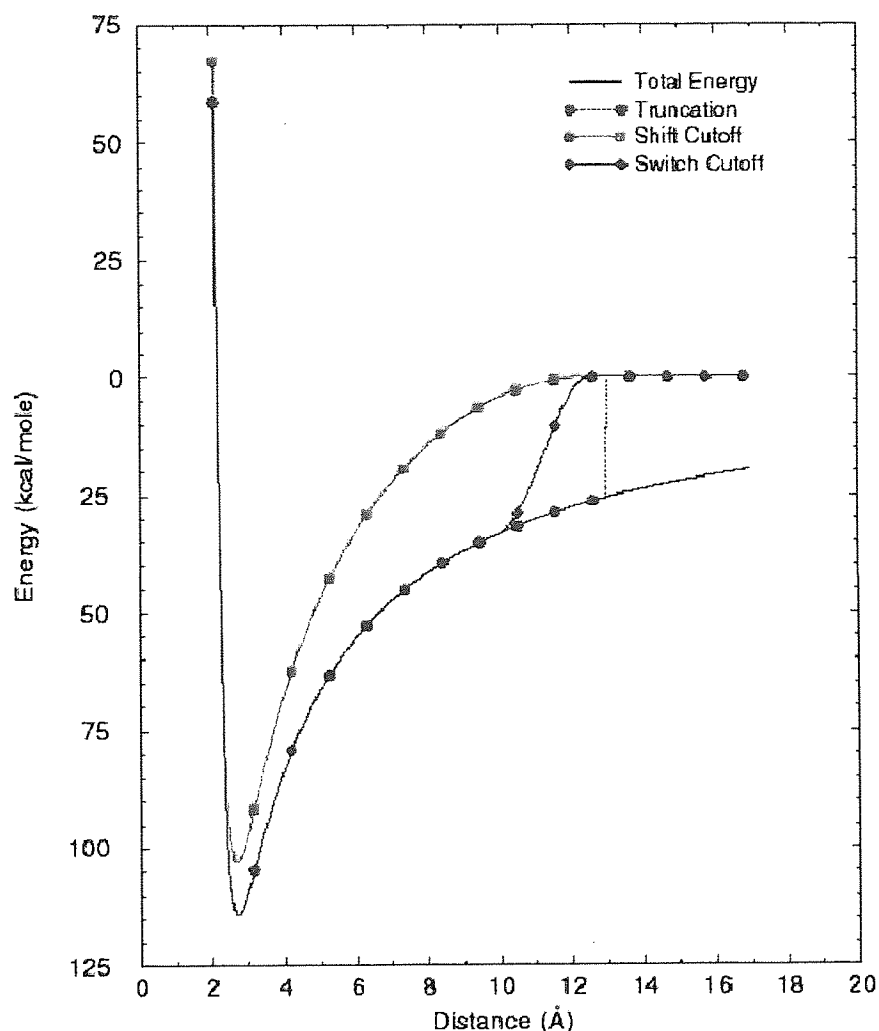


Figure 1.1.9 Three ways to treat cut-off condition (Adapted from http://www.ch.embnet.org/MD_tutorial/)

The first one is called the 'truncation' method. As shown in **Figure 1.1.9**, it uses a normal Lennard-Jones potential before the cut-off distance. When it reaches the cut-off point, the interaction suddenly becomes zero. This method is easy to use. However, it can cause problems around the cut-off point. A huge force could be

The first one is called the 'truncation' method. As shown in **Figure 1.1.9**, it uses a normal Lennard-Jones potential before the cut-off distance. When it reaches the cut-off point, the interaction suddenly becomes zero. This method is easy to use. However, it can cause problems around the cut-off point. A huge force could be introduced during the dynamics simulation if several atoms suddenly jump inside the cut-off distance. Therefore, this condition is not often used. The second one is called the 'shifted' method. Instead of using a true Lennard-Jones potential, an additional term is added so that it will reach zero at the cut-off point. The drawback for this method is that it is rarely used in 'true' simulation, especially for those calculating thermodynamic properties, as the Lennard-Jones potential has been changed. The third one is the 'switch' method. This method uses the true Lennard-Jones potential until a small distance before the cut-off point. Then, it switches to zero within this short range. This method can overcome the problem of the 'shift' method. However, it still suffers a strong force change inside the switch region (Leach 2001).

Although there were various parameters to choose in order to run a molecular dynamics calculation, the actual process of any molecular dynamics run shares a common protocol which would include the following steps (Ercolessi 1997)

- i. Prepare the input file. As the entire PDB file does not include hydrogen atoms, these missing parts need to be added on before MD as well as making sure there are no collisions.
- ii. MD parameters should be set. These mainly include the initial velocity of all the atoms, time steps, equilibration duration, MD running time duration, temperature, cut-off, etc.
- iii. Heat up the whole system. As the calculation started from 0K, the system needs to be gradually heated up to the desired temperature.
- iv. Equilibrating for a certain period in order to stabilise the whole system.
- v. Sample the MD run. After the system is equilibrated, the programme starts to output the data along with the time in order to produce dynamics trajectory

- vi. Analysis of the trajectory. After the dynamics has finished, the trajectory file can be used to analyse whether the MD process is successful. Then, the useful information can be extracted from the trajectory for further studies.

The following graph showed an example of dynamic trajectory.

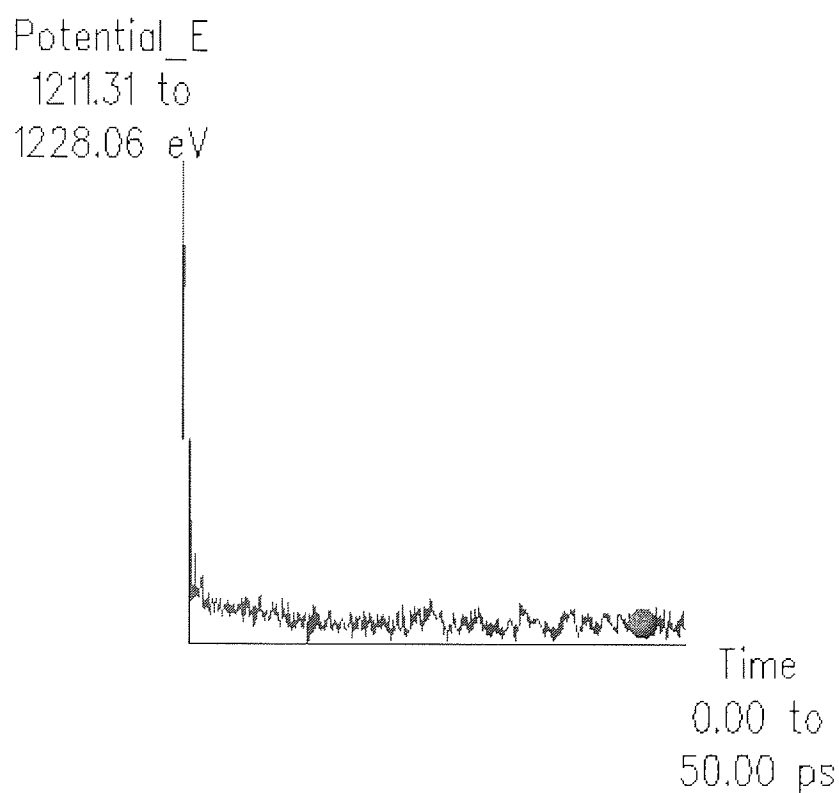


Figure 1.1.10: An example of a molecular dynamic trajectory.

It can be seen from **Figure 1.1.10** that the potential energy map started with a high-energy conformation. it quickly went down and stabilised. Any steric clashes could not be seen in this trajectory map. Therefore, this MD calculation could be considered successful.

1.1.6 Introduction of Quantum Mechanics

Different from classical Newtonian physics-based molecular mechanics method, quantum mechanics method is based on modern quantum theory which involves more complicated calculations. Usually, if MM (molecular mechanics) and QM (quantum mechanics) are applied on the same molecule, QM calculations will consume more computing time than MM calculations. This is because that the time required for MM calculations increases as the square of the number of atoms, while this relationship for QM calculation is the cube of the number of atoms (CaChe user guide, 2003). When first introduced, quantum mechanics calculation can only used on very simple systems such as H₂, He, N₂ etc. and all these calculations were done by hand. With the development of computer hardware and software, QM calculation can be applied to real molecular systems (Leach 2001).

The starting point for the quantum mechanics should refer to the time-dependent form of the Schrödinger equation.

$$\left\{ -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \gamma \right\} \Psi(r,t) = i\hbar \frac{\partial \Psi(r,t)}{\partial t} \quad 3.1$$

In equation (3.3.1), $\hbar = 1.0546 \times 10^{-34}$ is the reduced Planck's constant; m is the particle mass; $i = \sqrt{-1}$, $\Psi(r,t)$ is the wave function of position r associated with time t and the

γ is the external field. $\left\{ -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \gamma \right\}$ is the Hamiltonian operator

and it is usually abbreviate as \hat{H} . Therefore, equation 3.3.1 can be generalised to equation 3.3.2

$$\hat{H} \Psi(r,t) = i\hbar \frac{\partial \Psi(r,t)}{\partial t} \quad 3.2$$

Based on the Schrödinger equation, molecular orbital theory was derived. All the electron cloud outside the atomic nucleus is described as different orbitals which are characterised by three quantum numbers n , m and l . In here, n is the principal quantum number which defines the radial of the electron to its nucleus. l is the

azimuthal quantum number which defines the shape of the molecular orbitals. m is the magnetic quantum number which describes the angular momentum of the orbitals (Greenhow 1990). As QM calculation involves electrons, any properties associated with them can be determined. The most interesting one is the breaking and forming of the chemical bond which cannot be treated in MM and MD calculations.

The most commonly used two kinds of QM calculation are *ab initio* and semi-empirical methods. The *ab initio* calculation is more computationally expensive because it calculates everything from the beginning by only using physical constants such as:

Planck's constant $h=6.62606896 \times 10^{-34}$ Js

mass of electron $m_e=9.10538215 \times 10^{-31}$ kg

charge on an electron $|e|=-1.602176487 \times 10^{-19}$ C etc

Then, it will apply the full Hartree-Fock/Roothaan-Hall equations and calculate every integral and any other terms in the Hamiltonian. However, in semi-empirical calculations, some of the integrals and some of the terms in Hamiltonians are simplified by using parameters which derived from experimental data. In this case, fewer computational resources are needed. In this project, CaChe Workspace (Version 6.1.10 Fujitsu Limited) used MOPAC 2000 (abbreviated from Molecular Orbital PACKage, Stewart 1999), which is a semi-empirical method to conduct QM calculations. The MOPAC applied in the homology modelling process used PM3 parameterisation. This method is based on the MNDO (Modified Neglect of Diatomic Overlap, Dewar & Thiel 1997a, b) and stands for the third parameterisation of MNDO devised by Stewart (1989a,b). There are more molecular properties which can be calculated using MOPAC. These include

- Bond order
- Dipole moments
- Dynamics maps
- Ionization potentials
- Molecular orbital energies

- Optimum geometry
- Partial charges
- Potential energy maps
- Intersystem crossings
- Reaction pathways
- Transition states
- Vibration frequencies and spectra
- Structure and properties in solution

In CaChe Workspace, there is a patented method of MOPAC called MOZYME (US patent, NO. 5604686). As mentioned before, the time for a quantum mechanics run scales as the third power of the atom number. As most of the protein will contain several thousand atoms, directly applying MOPAC on such a large system becomes impractical. MOZYME uses localized molecular orbitals and alters the way that SCF (Self-Consistent Field) equation is solved. This method could give the same result in a different way. By using MOZYME, both time and the memory required are less. However, MOZYME has its own drawbacks. It can only be used for simple geometric optimization. If the calculation required vibrational frequencies or frequency-dependent polarisability, normal MOPAC should be applied. It is also suggested that when running a MOZYME calculation, it is better to break the whole process into several runs. This is because during the calculation, the geometry of the structure may change significantly. The interaction between some of the atoms may increase dramatically along with the time and the calculation may collapse with a huge gradient change. Therefore, stopping the calculation and reloading the updated information can reduce the chance of calculation crashes.

1.1.7 Introduction of Docking and Scoring

1.1.7.1 Scoring functions

After docking, scoring the final protein-ligand complexes is still considered as a major challenge (Krovat *et. al.* 2005). There are several scoring methods that have been developed since the 1990s (Tame 1999). Among them, the most accurate one should be the conformational sampling-based method, such as free energy perturbations (Kollman 1996) or linear interaction energies approximations (Aqvist *et. al.*). However, those methods consume too much computational time. Therefore, at the speed of current computers, they are not fit for screening ligand libraries. Apart from those complicated scoring functions, there are three main type of scoring methods, namely force-field-based methods, empirical methods and knowledge-based methods. Some of the examples are shown as following.

1: Force-field-based scoring functions

- AMBER (Ponder & Case 2003)
- CHARMM (Brook *et. al.* 1983)

2: Empirical scoring functions

- Ludi (Böhm 1994)
- ChemScore (Eldridge, 1997)

3: Knowledge-based scoring functions

- PMF score (Muegge and Martin, 1999)
- Drug score (Gohlke *et. al.* 2000)

In the force-field-based scoring methods, both AMBER and CHARMM are the name of the programmes. In these programmes, they usually calculate the energy of ligand-active site and ligand. Then the total energy difference before docking and after docking will be considered as docking score. Different force-field functions depend on the different force field parameters. Generally, van der Waals and electrostatic energy would be calculated in the protein-ligand interactions. The force-field-based

scoring method has its own drawbacks because at the beginning of its development, it mainly calculated the enthalpy but did not include solvation and entropic terms. At the same time, one of the force-field parameters, 'cut-off distances' which defines the effective non-bonded atom interactions may miss some of the long-range effects in a protein-ligand complex.

In the empirical scoring function, LUDI is the name of the programme and its scoring function is of empirical-based. ChemScore is the name of the scoring function which is used in GOLD program (Jones *et. al.* 1995). This kind of scoring function is based on the protein-ligand binding energy which can be calculated by a group of individual uncorrelated terms in which their coefficients are derived from regression analysis using known protein-ligand binding energy, especially the X-ray crystal structure information. In general, these kind of scoring methods calculate the binding ΔG value in the sum of some chemical and physical properties such as hydrogen bond, metal ion, VDW, electron static, solvation terms, rotatable bonds *etc.* Because of this, the coefficients are highly dependent on the structure information sets which are used in the regression analysis and fitting. One set of coefficients may not be used for another set of complexes (Kitchen *et. al.* 2004).

Finally, in knowledge-based scoring functions, PMF and DrugScore are all the name of the scoring functions. In these kinds of methods, the energy of the complex is not calculated. Instead, atom pair interaction potentials are used to represent the binding affinity. Atom pair interaction potentials have been derived from reference protein-ligand complexes. This method is simple and quick for screening large number of complexes. However, its disadvantage is that they still depend on the limited reference structure set used (Kitchen *et. al.* 2004). If a new atom pair appears and it is not represented in the reference set, the docking score cannot be calculated. Overall, all these kind of scoring functions only give approximations. Until now, no single scoring method can fully reproduce absolute structure activity profiles (Bissantz 2000 & Krovat *et. al.* 2005).

In this study, CaChe Workspace (Version 6.1.10 Fujitsu Limited) uses PMF (Potential of Mean Force) scoring method to rank the docked results. This method was developed by Muegge and Martin in 1999 and it is a knowledge based scoring method which is derived from the statistical analysis of protein and ligand complexes. They developed thirty-four ligand atom types and sixteen protein atom types from six hundred and ninety-seven protein structures from Brookhaven Protein Data Bank. All these protein and ligand atom types are shown in **Table 1.1.4** and **Table 1.1.5**. Muegge expressed the PMF score between a ligand atom type j and a protein atom type i as the following.

$$A_{ij}(r) = -k_B T \ln \left[f_{\text{Vol_corr}}^j(r) \frac{\rho_{\text{seg}}^{ij}(r)}{\rho_{\text{bulk}}^j} \right]$$

$$\text{PMF_score} = \sum_{\substack{kl \\ r < r_{\text{cut-off}}^{ij}}} A_{ij}(r)$$

In the first equation, r is the atom pair distance, k_B is the Boltzmann constant and T is the absolute temperature. Inside the brackets $\rho_{\text{seg}}^{ij}(r)/\rho_{\text{bulk}}^j$ indicate the pair correlation (radial distribution) function between a ligand atom type j and a protein atom type i in PDB data structure and $f_{\text{Vol_corr}}^j(r)$ is the ligand volume correction factor. In the second equation, $r_{\text{cut-off}}^{ij}$ is the cut-off radius for the atom pair type ij and finally they sum over all protein-ligand atom pairs kl in a database.

Atom type	Description
CF	nonpolar carbon sp^3 aliphatic
CP	polar sp^3 carbon bonded to an atom other than carbon or hydrogen
cF	nonpolar carbon aromatic
cP	polar carbon aromatic
C3	nonpolar carbon sp^2 not aromatic

CW	polar carbon sp^2 not aromatic (e.g., bonded to carbonyl oxygen)
CO	carbon bonded to a negatively charged oxygen
CN	carbon bonded to a positively charged nitrogen
C0	sp carbon
NC	positively charged nitrogen (e.g., NH_3^+ or guanidino group)
NP	planar nitrogen bonded to 2 or 3 carbons but not to a hydrogen (can occur in a nonaromatic ring)
NA	nitrogen as a hydrogen bond acceptor, not in a ring
ND	nitrogen as a hydrogen bond donor, not in a ring (e.g., amide nitrogen)
NR	planar nitrogen in a ring structure (e.g., pyridine)
N0	sp nitrogen bound to 1 carbon
NS	nitrogen bound to atoms other than carbon or hydrogen and not of type ND
OC	negatively charged oxygen (e.g., carboxylate)
OA	oxygen as hydrogen bond acceptor (e.g., keto, amide oxygen)
OE	oxygen in an ether bond
OR	oxygen in planar ring
OS	oxygen bonded to atoms other than carbon or hydrogen
OD	oxygen bonded to hydrogen, except water
P	phosphorus
SA	sulfur as hydrogen bond acceptor
SD	sulfur as hydrogen bond donor
HL	hydrogen
Zn	zinc
CL	chlorine
Mn	manganese
Mg	magnesium
F	fluorine
Fe	iron
Br	bromine
V	vanadium

Table 1.1.4 Ligand atom types

Atom type	Description
CF	nonpolar aliphatic carbon (e.g., $C\alpha$)
CP	polar aliphatic sp^2 or sp^3 carbon bonded to atoms other than carbon or hydrogen (e.g., backbone C or CR)
cF	nonpolar carbon aromatic
cP	polar carbon aromatic
CO	carbon bonded to a negatively charged oxygen
CN	carbon bonded to a positively charged nitrogen
NC	positively charged nitrogen
ND	nitrogen as a hydrogen bond donor (e.g., backbone N, TRP NE, ASN ND)
OC	nitrogen in a planar ring structure (e.g., HIS ND, HIS NE)
OA	oxygen as hydrogen bond acceptor (e.g., backbone O, ASN OD, GLN OE)
OD	oxygen as hydrogen bond donor (e.g., TYR OH, SER OG, THR OG)
OW	water oxygen
SA	sulfur as hydrogen bond acceptor (MET SD)
SD	sulfur as hydrogen bond donor (CYS SG)

Table 1.1.5 Protein atom types

In practice, Muegge and Martin pointed out that the PMF scoring function in DOCK4 is more reliable than the force-field scoring function. At the same time, they also compared it with other kinds of scoring functions and the results show that PMF performs better than LUDI and SMOG scoring in eight different protein-ligand test sets (Ewing *et. al.*1997). There was a similar scoring function which was developed by Verkhivker and co-worker (1995). Their reference set came from thirty HIV and SIV protein-ligand structures and only twelve different atom pairs were developed. Muegge and Martin emphasised that large number of atom types derived from a large protein database can enhance the power of PMF scoring. PMF scoring function can be improved by defining more atom types based on more new PDB structures. However, Muegge and Martin predicted that the number of atom pairs might not increase within a few years by waiting for some new PDB structure. There still may be other ways to enhance PMF scoring function. One of them is to recombine these atom types to improve scoring function. Another one is that some metal atoms in ligand atom type do not have sufficient data to derive their potential of mean force. Therefore, this needs a further study.

1.1.7.2 Docking functions

Docking programmes cannot function without a scoring method. Every docking programme should have at least one or more scoring functions. This is because that during the docking, each iteration needs to be ranked using a certain scoring function. Also, there are several kinds of docking algorithms present. A docking programme is a combination of docking algorithms with scoring functions.

There are mainly four types of docking algorithms.

- Fast shape matching
- Incremental construction

- Molecular dynamics
- Random search

In fast shape matching algorithms, the programme generally calculates the shape of the active site and the ligand first. Then, by giving a systematic search based on the position, orientation and the rotatable bonds in ligand or even in protein, the programme should be able to locate favourable docking orientation (Pospisil 2004). In the incremental construction methods, the shape of the active site is calculated as well. It uses a different way to treat the ligand by breaking the ligand apart along its rotatable bond into fragments. Then, it will calculate the biggest core fragment and place this part into the active site first. In the next few steps, the programme will incrementally construct the whole ligand inside the active site (Böhm 1994). In the molecular dynamics docking methods, the ligand together with the protein starts in a random or an existing state. By performing molecular dynamics calculation, a dynamics trajectory should be created. The lowest local minimum energy conformation should be as the final docked structure (Pearlman *et. al.* 1995). There are two kinds of methods including in random search docking namely Monte Carlo search and Genetic algorithm. In the Monte Carlo algorithm, the ligand starts with a random initial configuration inside the active site. Then, the initial conformation would be scored. After this, a new configuration should be generated and scored. The newly generated pose will be judged by Metropolis criterion to decide whether the new conformation is accepted. This means that if the new score is better than the old one, it will be accepted. If it is not, a Boltzman-based probability function will be applied to test this new position. The new position can only be accepted once it can pass this test. This same process will be repeated until the desired configuration is achieved (Rubinstein c1981).

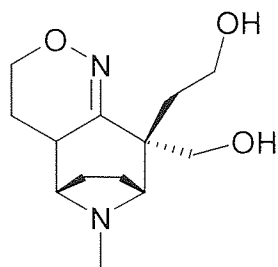
The docking algorithm used in CaChe workspace (Version 6.1.10 Fujitsu Limited) uses genetic algorithms in the docking to obtain the best docking position. This is the other kind of random search method. This method was inspired from Darwin's theory of natural selection that genetic information can crossover and mutate to produce the

next generations (Camila et. al. 2004). Therefore, all the ligand and the receptor's conformation information were represented by chromosomes. The programme would start by generating a certain number of the population containing those chromosomes. Next, a defined genetic operator including the crossover rate and the mutation rate would be given. Then, those chromosomes will mutate and give the next generation randomly. The new generations then should be scored and the best new population will go to the next iterations. The programme continues until the set number of iterations has been performed or the desired criteria have been satisfied. If the algorithms stop because the set number of iteration has been reached, then the result may not be the optimal one.

1.1.8 Design of Inhibitors of Mycobacterial dUTPase

The crystal structure of *Mycobacterium tuberculosis* dUTPase as well as its catalytic mechanism was reported by Chan *et. al.* (2004) According to Persson *et.al.* (2001), the dUTPase can be found in most prokaryotes, eukaryotes and a large majority of viruses. They reported that researchers have already published the structure of different dUTPase from *Escherichia coli* (*E. coli.*), human, feline immunodeficiency virus (FIV) and equine infectious anaemia virus (EIAV). The role of dUTPase is to degrade a toxic material 2'-deoxyuridine-5'-triphosphate (dUTP) to dUMP and pyrophosphate (PP_i). dUTP is toxic due to its being one methyl group different from dUTTP so that during DNA synthesis thymidine residues are changed to uracil residues. Finally, this will break down the cell repair system and eventually cause cell death. Therefore, dUTPase is a potential target for design anti-tuberculosis compounds.

A previous project student (Chatzakis 2006) had docked a set of 21 compounds from the Maybridge compounds collection into both *Mycobacterium tuberculosis* dUTPase (Mtb-dUTPase) and human-dUTPase (h-dUTPase). The following compound T2-3-9 was studied and has the potential to be an inhibitor for *M. tuberculosis* dUTPase.



Compound Code T 2-3-9

Figure 1.1.11: The best compound derived from Maybridge compound BTB110716 (Chatzakis 2006)

This structure shown in **Figure 1.1.11** would be difficult to synthesis owing to its many chiral centres. Therefore, following on from Chatzakis' work, further analogues of T 2-3-9 were designed and docked into *M. tuberculosis* dUTPase using Fast Docking programme in CAChe Work System Pro (Version 6.1.10 manufactured by Fujitsu Ltd). The refined protein structure also came from the previous project student. After the docking, the best structure will be chosen for the chemical synthesis and tested against *M. tuberculosis* as well as other microorganisms such as *E. coli* and *S. aureus*. There were mainly two types of T 2-3-9 analogues with several kind of variations which are shown in **Table 1.1.6**.

	General structure	Side chain
Type 2A		R = C ₁ -C ₆
Type 2B		R = C ₁ -C ₆ R' = -CH ₂ -OH, -CH ₂ -CH ₂ -OH
Type 2C		X = NH, O R' = H, -CH ₂ -OH, -CH ₂ -CH ₂ -OH R'' = C ₁ -C ₃

Table 1.1.6: Further analogues of T2-3-9.

The design of those compounds in **Table 1.1.6** were based on their ease of synthesis. The designed synthesis approach will be shown in **Chapter 2**. It can be seen from **Table 1.1.6** that two types of the compounds are described in separate columns. This was because two different starting materials were accessible, namely tropinone and 1-methyl-4-piperidone. The extra part in the tropinone was considered as a 'bridge'. However, it can be seen that nearly all the compounds in **Table 1.1.6** will have isomers, especially the Type B compounds having two or three chiral centers. In this case, each of them will have four or eight isomers. This situation made it difficult to dock all the isomers. Therefore, only one of the conformations were used in the modelling and docking. If a lower docking score together with a reasonable docking position can be found, other isomers may be docked as well to see the comparison.

1.2 Results and Discussion

1.2.1 Homology modelling

As showed previously, this homology modelling work will use 1jji.pdb and 1qz3.pdb for the templates for Rv3487c. Zhang *et. al.* (2005) already did an on-line BLAST search and gave the aligned protein sequences. Those can be seen in **Figure 1.1.2**. The crystal structures of 1jji and 1qz3 can be downloaded from protein data bank (www.rcsb.org PDB code: 1jji and 1qz3). However, CaChe workspace (Version 6.1.10 Fujitsu Limited) uses .csf file format during the calculations. Those PDB files were converted to .csf files first. 1jji contains four identical subunits, each containing an active site namely, 1jjiA, 1jjiB, 1jjiC and 1jjiD. 1jjiA was chosen as the template.

The PDB files were loaded and converted into CaChe workspace and the valence and hybridisation were corrected so that all the hydrogen atoms were added onto the structure and to make sure all the hybridisation was right. Then, a standard process, which relaxed only all the added hydrogen atoms, was conducted. This procedure included three steps:

1. By allowing the hydrogen atoms to move freely, molecular mechanics method (MM2 force field using steepest descent and conjugate gradient for 90000 iterations each) was used first to erase any steric collision.
2. A short molecular dynamics simulation was conducted (MM2 force field, 600K with equilibration period of 1ps, 1fs simulation time-step and simulation duration 20ps) to overcome the local barriers of those hydrogen atoms and the lowest conformation was chosen for the next step.
3. Molecular mechanics (MM2 force field using steepest descent and conjugate gradient for 90000 iterations each) was conducted again to refine the hydrogen position from molecular dynamic simulation.

As all the PDB file downloaded from PDB data bank do not contain hydrogen atoms, these hydrogen atoms need to be added onto the protein and to make sure that they stay in reasonable positions. Before performing molecular dynamics simulation, steric clashes must be minimised. That is why step 1 was used. During the molecular dynamics simulation, high energy was used (600K) so that all the hydrogen atoms could reach any possible position to find a global minimum. After the molecular dynamics, some of the bonds might be too long due to the high-energy input. Therefore, step 3 was used to correct those errors.

Using the protein sequence alignment in **Figure 1.1.2** from Zhang *et. al.* (2005), both amino acid sequences of IjjiA and Iqz3 were changed to the sequence of Rv3487c manually. Inserting new loops and deleting unwanted amino acids were carried out manually. After these changes, there were many gaps and collisions between backbone and side chains of protein. That is why molecular modelling methods were used to refine changed structures. Followed by the CaChe 6.1 user guide (2003) the **Pathway 1** was used to refine the changed structure of IjjiA and Iqz3.

Pathway 1:

Step 1: In this step, all the inserted loops and the gaps were refined first. Every atom was locked except those new loops and the gaps together with the nearest two residues on either side. Then, molecular mechanics MM3 was used with steepest descent and conjugate gradient for 90000 iterations each to let the loops move to fit the space. Molecular dynamics (MM3 force field, 1000K with equilibration period of 1ps, 1fs simulation time step and simulation duration 50ps) was conducted afterwards. Again, the lowest energy conformation was chosen and molecular mechanics MM3 using the same parameters was conducted to achieve a local minimum.

Step 2: In this step, the backbone and the helix were locked. All the loops and the β -sheets were free to move. Then, molecular mechanics energy

minimisation (MM3, steepest descent and steepest descent and conjugate gradient for 90000 iterations each), molecular dynamics (MM3, 1000K, equilibration period 1ps, simulation duration 100ps) and molecular mechanics energy minimisation (MM3, steepest descent and steepest descent and conjugate gradient for 90000 iterations each) were used to refine the unlocked part.

Step 3: The purpose of this step was to refine the helix of the protein. Firstly, everything except the first turn (four residues) of each helix was locked and followed by molecular mechanics energy minimisation (MM3, steepest descent and conjugate gradient for 90000 iterations each), molecular dynamics (MM3, 600K, equilibration period 1ps, simulation duration 100ps) and molecular mechanics energy minimisation (MM3, steepest descent and conjugate gradient for 90000 iterations each). Then, the second turn (next four residues) of each helix was unlocked and underwent the same procedure until all the helixes were unlocked and refined.

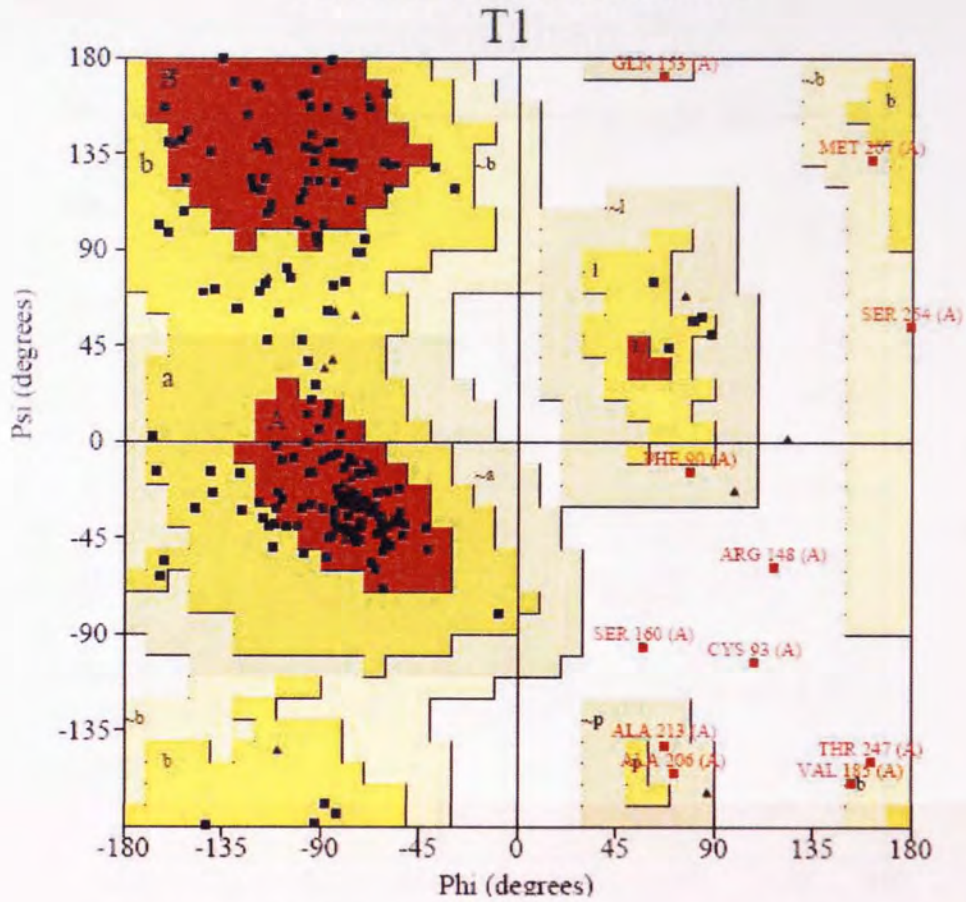
Step 4: The residues that were 8 Å around SER90-GLU189-HIS219 catalytic triad were chosen to be the active site pocket. Then, semi-empirical quantum mechanical geometry optimisation (PM3 MOZYME, under vacuum) was applied to the active site only in this step to give a better portrayal of the protein active site. The parameter 'RELSCF' was set to 10 and 'GNORM' was set from 100 down to minimum of 6 in around 10 step-by-step calculations.

Step 4 is not present in the standard procedure of the user guide. The purpose of this step is to give a special treatment of active site with a higher-level geometry optimisation. The parameter 'RELSCF' defines how easily the calculation could pass the SCF test. The default value of RELSCF was 1. However, under this condition, the calculation collapsed with a very large gradient. The calculation can run smoothly with RELSCF=10 which means it was ten times easier to pass the SCF test. Another

important parameter was GNORM which defines the termination criteria in the geometry optimisation. The default setting was GNORM=1. Nevertheless, the calculation collapsed under the default conditions. After investigating the calculated result, it was found that the gradient could easily go down to 100. At the same time, it was suggested that the whole process of MOZYME calculation would be better broken down into several runs. Therefore, the calculation of homology model from 1jji was started with GNORM=100. The second calculation was set to GNORM=50. Then, this number was reduced to 30, 20, 15, 10, 8, 7, 6 respectively. It seemed that the gradient could not go down below 5. Therefore, GNORM=6 was the minimum value of the MOZYME calculation. When the homology model from 1qz3 was under MOZYME calculation, the GNORM was also started with 100. However, the GNORM was reduced to 80 in the second run. Then, the number was reduced to 60, 40, 30, 20, 15, 10, 8, 7, 6. Although there were two more steps, the total time used was less than the first calculation of homology model from 1jji.

After **Pathway 1**, two homology models **T1** and **T4**, which came from 1jji and 1qz3 were generated respectively. Their psi/phi angles are checked by PROCHECK (Laskowski *et. al.* 1993 and Rullmann 1996) and shown in **Figure 1.2.1** and **Figure 1.2.2** as Ramachandran plots

Ramachandran Plot

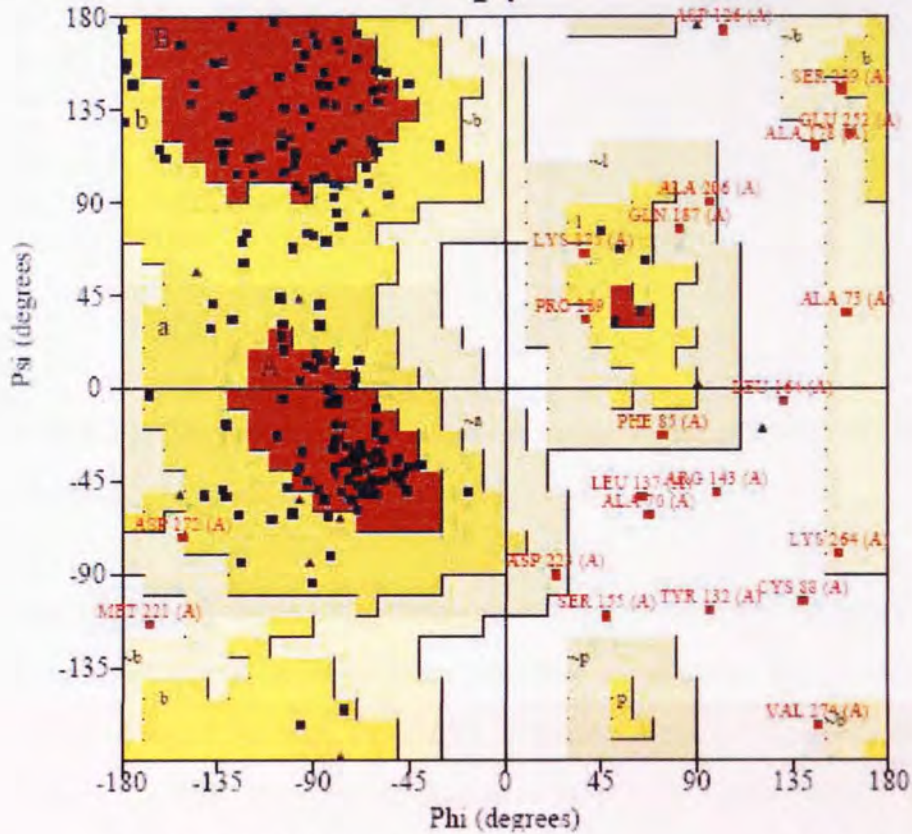


Residues in most favoured regions [A,B,L]	134	70.50%
Residues in additional allowed regions [a,b,l,p]	45	23.70%
Residues in generously allowed regions [~a,~b,~l,~p]	7	3.70%
Residues in disallowed regions	4	2.10%
Nmumber of non-glycine and non-proline residues	190	100%
Number of end-residues (excl. Gly and Pro)	7	
Number of glycine residues (shown as triangles)	19	
Nmumber of proline residues	16	
Total number of residues	232	

Figure 1.2.1: Ramachandran plot for T1

Ramachandran Plot

T4



Residues in most favoured regions [A,B,L]	121	62.70%
Residues in additional allowed regions [a,b,l,p]	51	26.40%
Residues in generously allowed regions [~a,~b,~l,~p]	10	5.20%
Residues in disallowed regions	11	5.70%
Nmumber of non-glycine and non-proline residues	193	100%
Number of end-residues (excl. Gly and Pro)	6	
Number of glycine residues (shown as triangles)	20	
Nmumber of proline residues	16	
Total number of residues	235	

Figure 1.2.2: Ramachandran plot for T4

The most important features in these two figures are the residues that fell into disallowed regions which is the white area in the Ramachandran plot. Four residues fell into disallowed regions in T1 namely CYS 93, ARG 148, SER 160 and THR 247. Only SER 160 was inside the active site. ARG 148 and SER 160 belong to the α -helix.

In T4, there were 11 residues that fell into the disallowed region. **Table 1.2.1** shows their PDB number and their relative positions.

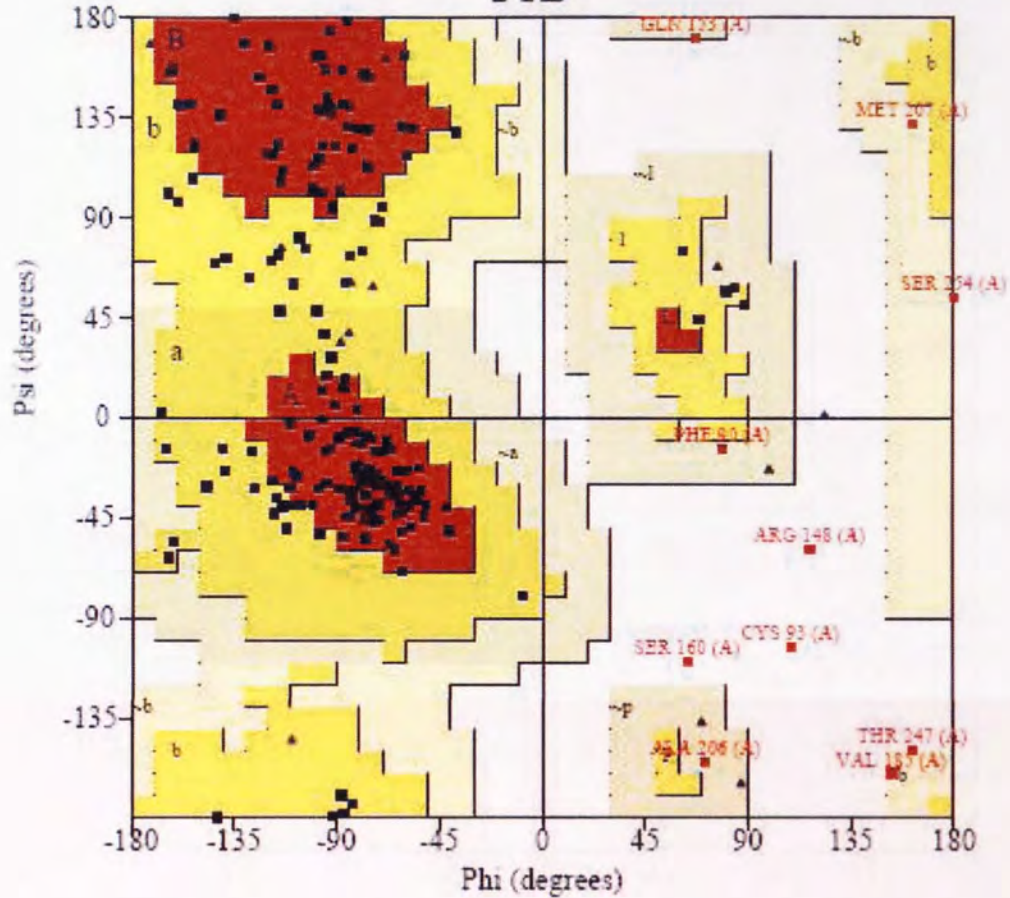
PDB NO.	Residues name	Inside or outside of active site	Inside or outside of alpha helix
70	ALA	-	-
88	CYS	+ *	-
126	ASP	-	+
132	TYR	-	+
137	LEU	-	+
143	ARG	-	-
155	SER	+	-
164	LEU	-	+
178	ALA	-	-
221	MET	-	-
274	VAL	-	-

Table 1.2.1 PDB number, name and the position related to active site and alpha helix ('+': inside, '-': outside, '*': at the edge)

From **Table 1.2.1**, it can be seen that two of the residues fell in the active site and four of them belong to the alpha helix. In order to eliminate the residues inside the disallowed region for both **T1** and **T4**, a further molecular dynamics and molecular mechanics that had the same parameters as **Step 3** was conducted with every part of the protein locked except the turn of the helix (four residues) containing those disallowed regional residues. After several MD with MM runs, CYS 88 and SER 155 in **T4** moved out of the disallowed regions. However, SER 160 in **T1** was still inside the disallowed zones. The following two figures show the Ramachandran plot result.

PROCHECK

Ramachandran Plot T1D

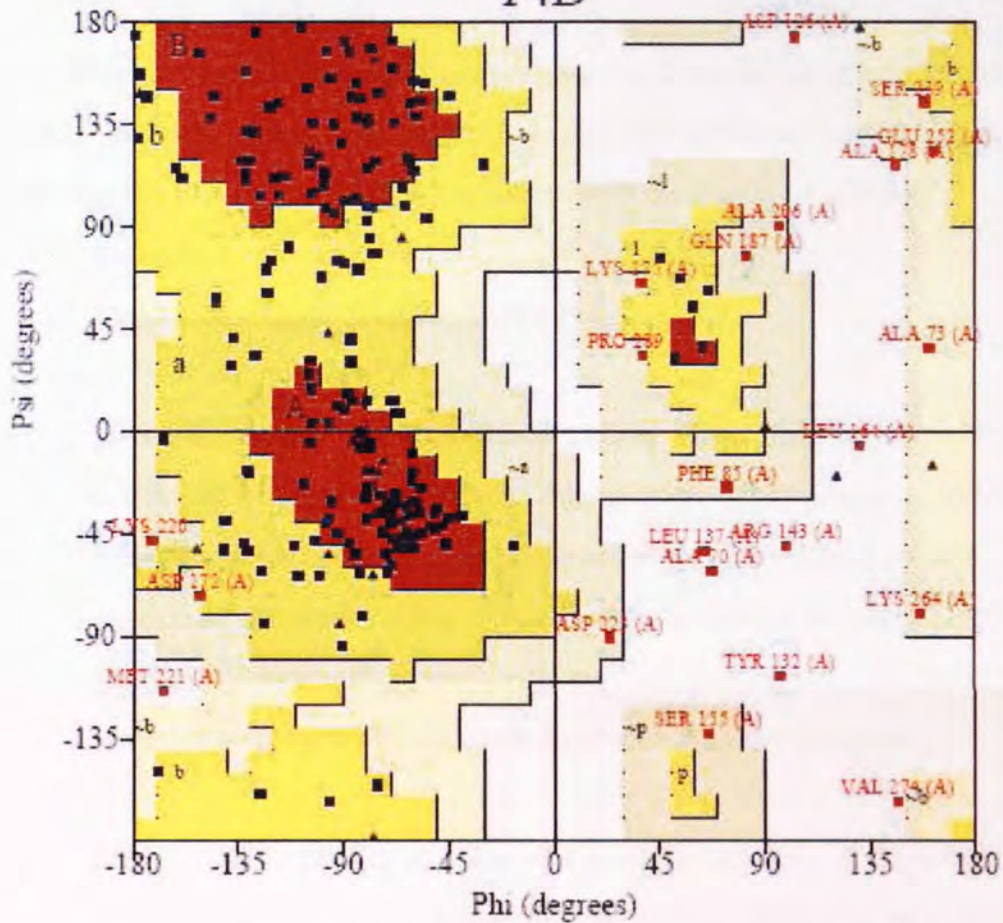


Residues in most favoured regions [A,B,L]	132	70.20%
Residues in additional allowed regions [a,b,l,p]	46	24.50%
Residues in generously allowed regions [~a,~b,~l,~p]	6	3.20%
Residues in disallowed regions	4	2.10%
Nmuber of non-glycine and non-proline residues	188	100%
Number of end-residues (excl. Gly and Pro)	12	
Number of glycine residues (shown as triangles)	19	
Nmumber of proline residues	17	
Total number of residues	236	

Figure 1.2.3: Attempted correction of T1

PROCHECK

Ramachandran Plot T4D



Residues in most favoured regions [A,B,L]	120	60.00%
Residues in additional allowed regions [a,b,l,p]	59	29.50%
Residues in generously allowed regions [~a,~b,~l,~p]	12	6.00%
Residues in disallowed regions	9	4.50%
Number of non-glycine and non-proline residues	200	100%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	20	
Number of proline residues	16	
Total number of residues	238	

Figure 1.2.4: Attempted correction of T4

It can be seen from **Figure 1.2.4** compared with **Figure 1.2.2**, CYS 88 disappeared in **Figure 1.2.4** which means that it moved into either most favoured regions or additional allowed regions. SER 155 moved from disallowed regions into generously

allowed zones shown as ~p area in **Figure 1.2.4**. If **Figure 1.2.1** and **Figure 1.2.3** are compared, although SER 160 was still inside the disallowed zones, it moved towards ~p area which is shown on the graph. Several snapshot from the dynamics trajectory all showed similar results.

In addition to these small changes, another **pathway 2** was designed in order to give a slightly different approach to the final model. The difference between these two pathways lies in step 2 and step 3. The following is the detail of **Pathway 2**.

Pathway 2:

Step 1: Same as the **Step 1** in **Pathway 1**.

Step 2: In this step, all the loops were refined. Firstly, helix, catalytic triad and unchanged residues were locked, following by MM3 energy minimisation with steepest descent and conjugate gradient for 90000 iterations each. Then, molecular dynamics MM3 (1000K with equilibration period 1ps and simulation duration 50ps) was conducted. Finally, molecular mechanics MM3 energy minimisation with the same parameters was applied again.

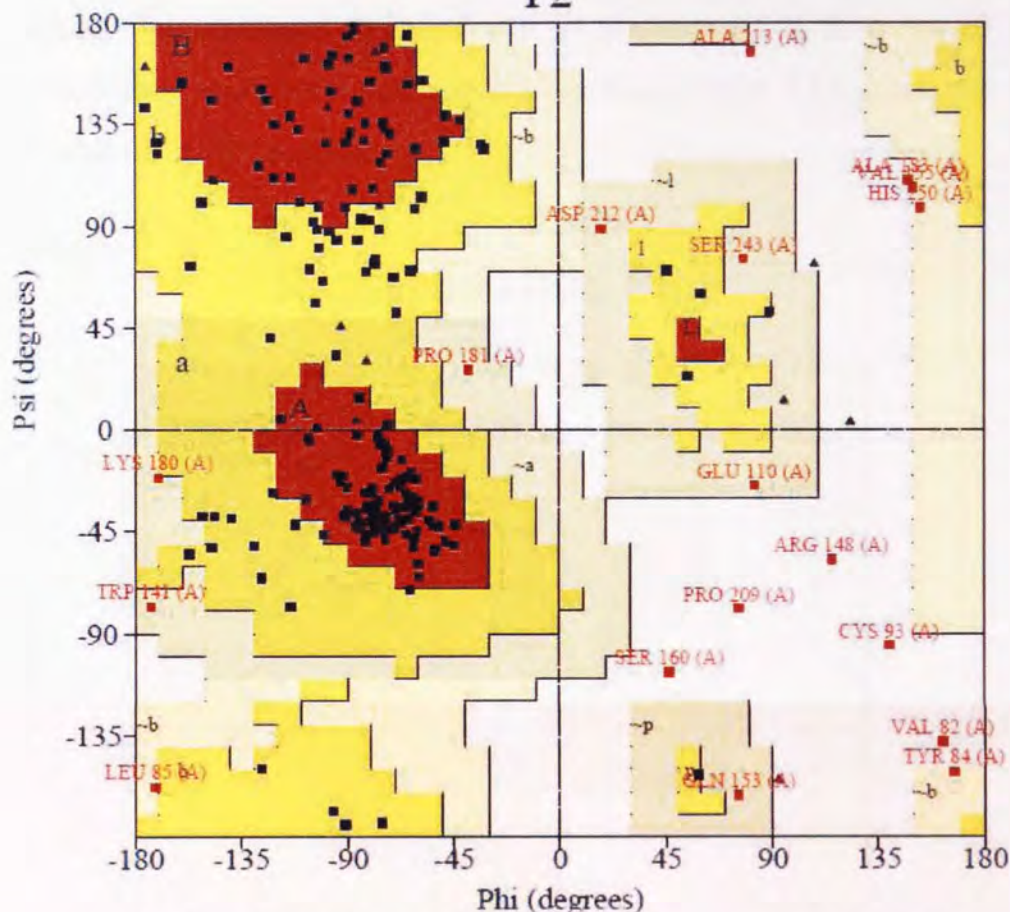
Step 3: The purpose of this step was to refine the helix of the protein. The catalytic triad was locked during this step. Firstly, everything except the first turn (four residues) of each helix was locked and followed by the molecular mechanics energy minimisation (MM3, steepest descent and conjugate gradient for 90000 iterations each), molecular dynamics (MM3, 600K, equilibration period 1ps, simulation duration 100ps), molecular mechanics energy minimisation (MM3, steepest descent and conjugate gradient for 90000 iterations each). Then, everything except the second turn (next four residues) of each helix was locked and followed by the same process until all the helixes were refined.

Step 4: Same as the **Step 4** in **Pathway 1**.

In **step 2**, unchanged residues on the beta sheet and loops were fixed while the calculation was running. In **Step 3**, each turn of the helixes was refined separately. This time, only the homology model from Ijji went through **Pathway 2** and generates **T2**. At the same time, its psi/phi angles were also checked by PROCHECK and the result is shown in **Figure 1.2.5**.

Ramachandran Plot

T2



Residues in most favoured regions [A,B,L]	126	66.30%
Residues in additional allowed regions [a,b,l,p]	48	25.30%
Residues in generously allowed regions [~a,~b,~l,~p]	9	4.70%
Residues in disallowed regions	7	3.70%
Number of non-glycine and non-proline residues	190	100%
Number of end-residues (excl. Gly and Pro)	7	
Number of glycine residues (shown as triangles)	19	
Number of proline residues	16	
Total number of residues	232	

Figure 1.2.5: Ramachandran plot for T2

By comparing **Figure 1.2.5** and **Figure 1.2.1**, the former seems to have more residues inside disallowed regions. Meanwhile, it was found that two of these residues were in the active site namely SER 160 and PRO 209. This result means that **pathway 2** is not as good as **pathway 1**. However, this structure will be used in the docking studies as a

reference. In order to move these two residues out of disallowed regions, everything except the turn involving SER 160 and the nearby two residues of PRO 209 at either side were locked and the same procedure of molecular dynamics and molecular mechanics was conducted. The result showed that PRO 209 could be moved out of the disallowed zones while SER 160 still could not. **Figure 1.2.6** shows one of the corrected PROCHECK results.

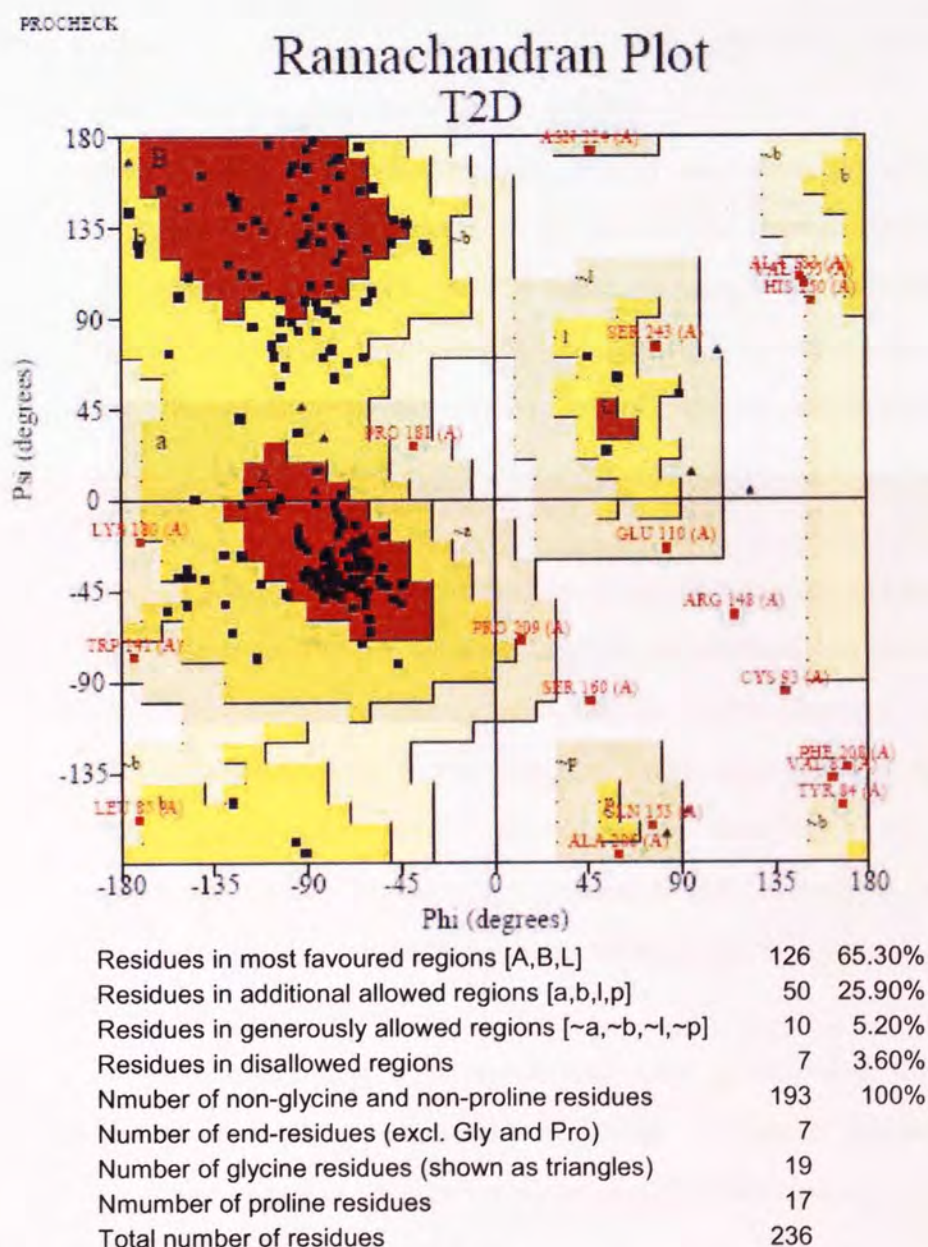


Figure 1.2.6: Attempted correction of T4.

As **Pathway 2** did not make the homology model better, 1qz3 did not use this pathway. A different pathway did not get the desired result. This might mean that the pathway that the CaChe User Guide gave should be the suitable for this programme. At the same time, a small amount of local changes toward the final homology model also were attempted, yet still some of the problems remained. A further long MD simulation might be needed in order to refine those models.

As nearly all the proteins function inside an aqueous environment, two structures from **Pathway 1** underwent a further optimisation in order to simulate the protein inside water. These processes are shown as follows:

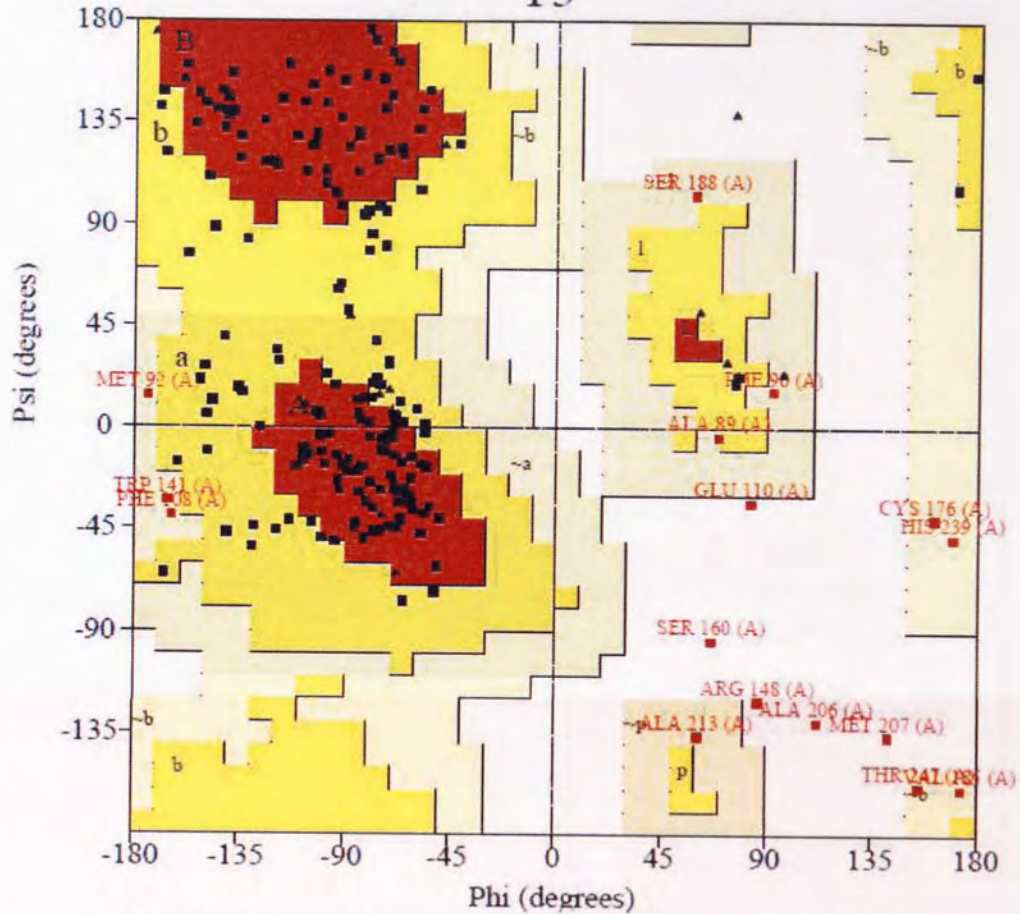
Step 1: In CaChe WorkSpace system, a water molecule sphere with a diameter 16 Å greater than the diameter of the protein was created. Then, the protein structure was transferred into this water box; any water molecule that was below 2 Å from the protein was deleted. After that, the water molecules which are 3 Å around the protein was free to move and the outside water molecule layer which is 5 Å thick was locked in order to contain all the molecule inside.

Step 2: After all these preparation, the protein was locked and the 3 Å water molecule layer was free to move, a quick optimisation was conducted using molecular mechanics energy minimisation (MM3, steepest descent and conjugate gradient for 90000 iterations each), molecular dynamics (MM3, 600K, equilibration period 1ps, simulation duration 20ps) and finally molecular mechanics energy minimisation (MM3, steepest descent and conjugate gradient for 90000 iterations each) to deal with water layer first.

Step 3: In this step, both protein and water molecules in the 3 Å layer underwent a long simulation of 500ps molecular dynamics (600K, equilibration period 1ps). The lowest energy conformation was extracted to be one of the final homology models.

The water treated pathway gave out two models namely **T3** and **T5**, which came from 1jji and 1qz3 respectively. Their PROCHECK results are shown in **Figure 1.2.7** and **Figure 1.2.8**.

Ramachandran Plot T3



Residues in most favoured regions [A,B,L]	120	63.20%
Residues in additional allowed regions [a,b,l,p]	54	28.40%
Residues in generously allowed regions [~a,~b,~l,~p]	11	5.80%
Residues in disallowed regions	5	2.60%
Nmuber of non-glycine and non-proline residues	190	100%
Number of end-residues (excl. Gly and Pro)	7	
Number of glycine residues (shown as triangles)	19	
Nmumber of proline residues	16	
Total number of residues	232	

Figure 1.2.7: Ramachandran plot for T3

Ramachandran Plot

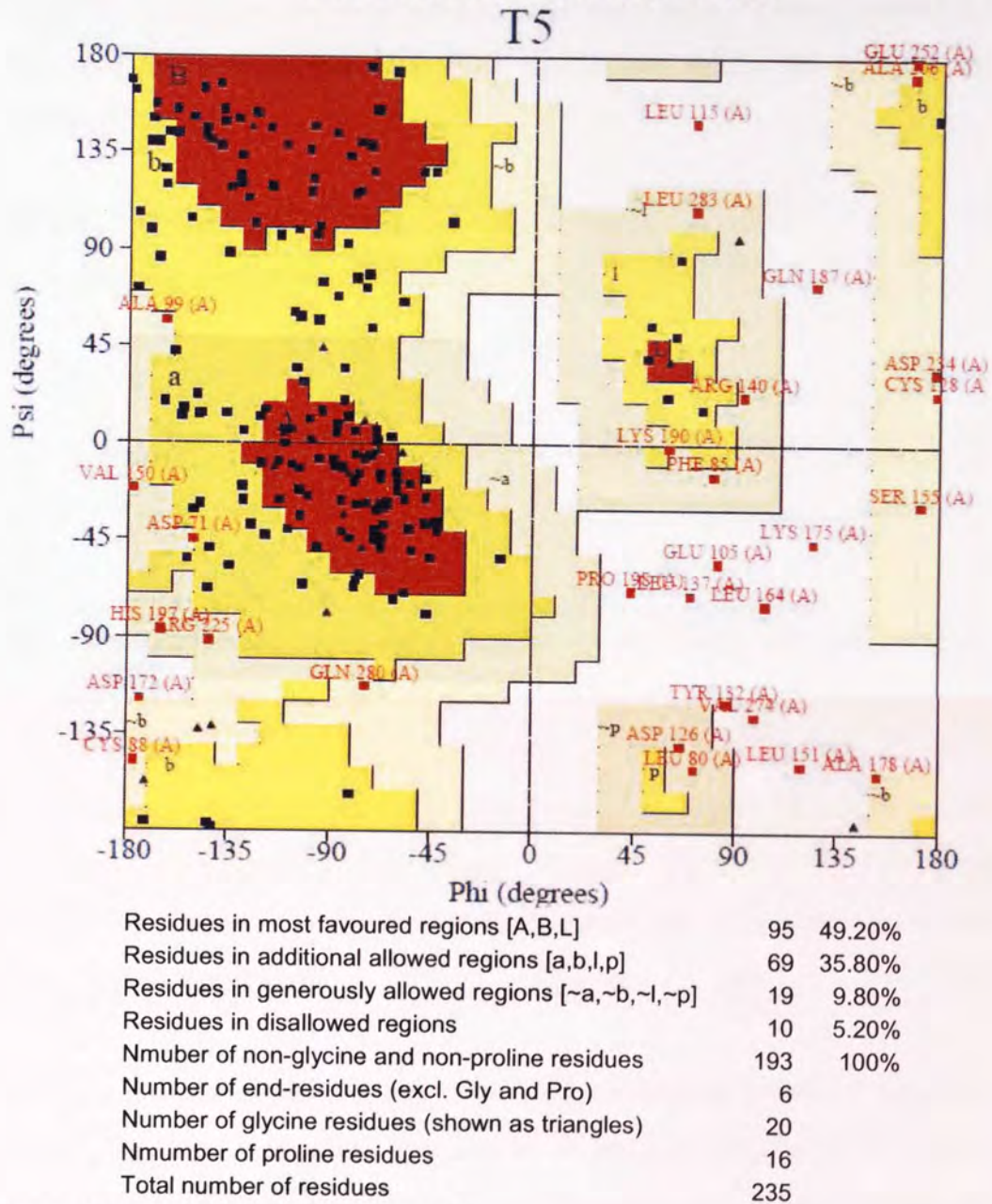


Figure 1.2.8: Ramachandran plot for T5

If the Ramachandran plots of T1, T2 and T3 were compared, T2 has the largest number of residues fell inside the disallowed regions. There was no significant difference between T1 and T3. The SER 160 in those homology models always fell

into disallowed zones. Also, a quick MD and MM calculation concentrated on SER 160 were conducted and the results showed the same as what happened in **T1** and **T2**.

T4 and **T5** unfortunately had more residues inside the disallowed zones. **Table 1.2.2** also shows the residues with their PDB number and names that fell into the disallowed regions.

PDB NO.	Residues name	Inside or outside of active site	Inside or outside of alpha helix
105	GLU	-	+
115	LEU	-	-
132	TYR	-	+
137	LEU	-	+
151	LEU	-	+
164	LEU	-	+
175	LYS	-	-
187	GLN	-	-
198	PRO	-	+
274	VAL	-	-

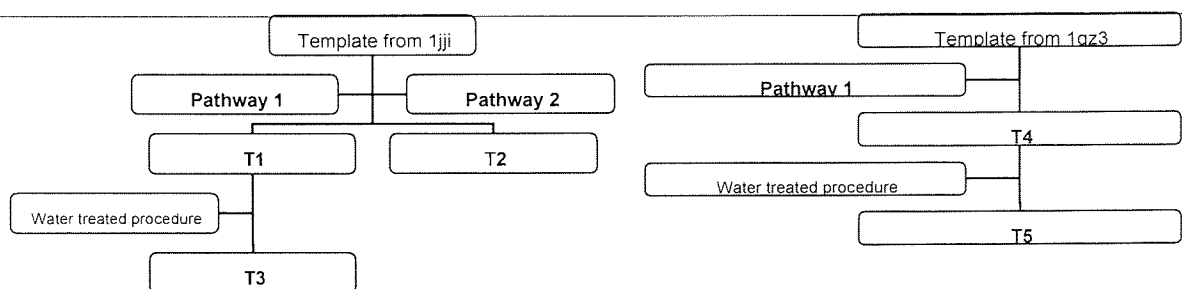
Table 1.2.2 PDB number, name and the position related to active site and alpha helix ('+': inside, '-': outside.)

By comparing **Table 1.2.1** and **Table 1.2.2**, it was clear that in **T5**, residues in the active site did not fall into disallowed regions. However, the downside of **T5** was that it had fewer residues in the most favoured regions and more inside generously allowed regions than **T4** had.

In addition to this, these five structures were given a trial run in AMBER (version 9.0) molecular dynamics calculation. The purpose of this is to see that if the AMBER calculation could fix those problems above. The result showed that although the overall percentage of residues in most favoured regions had increased and fewer residues fell into disallowed regions, the final structures had changed dramatically. This might be due to the conversion of CaChe file to PDB file and the fact that

AMBER tried to move those unfavoured residues led the overall structure changing significantly compared to these five final homology models.

Overall, all these five homology models were chosen for further docking work. As mentioned before, **T1**, **T2** and **T3** came from 1jjiA that went through **Pathway 1**, **Pathway 2** and the water treated process. **T4** and **T5** came from 1qz3 that went through **Pathway 1** and the water treated process. Their images and relations are shown in **Figure 1.2.9**.



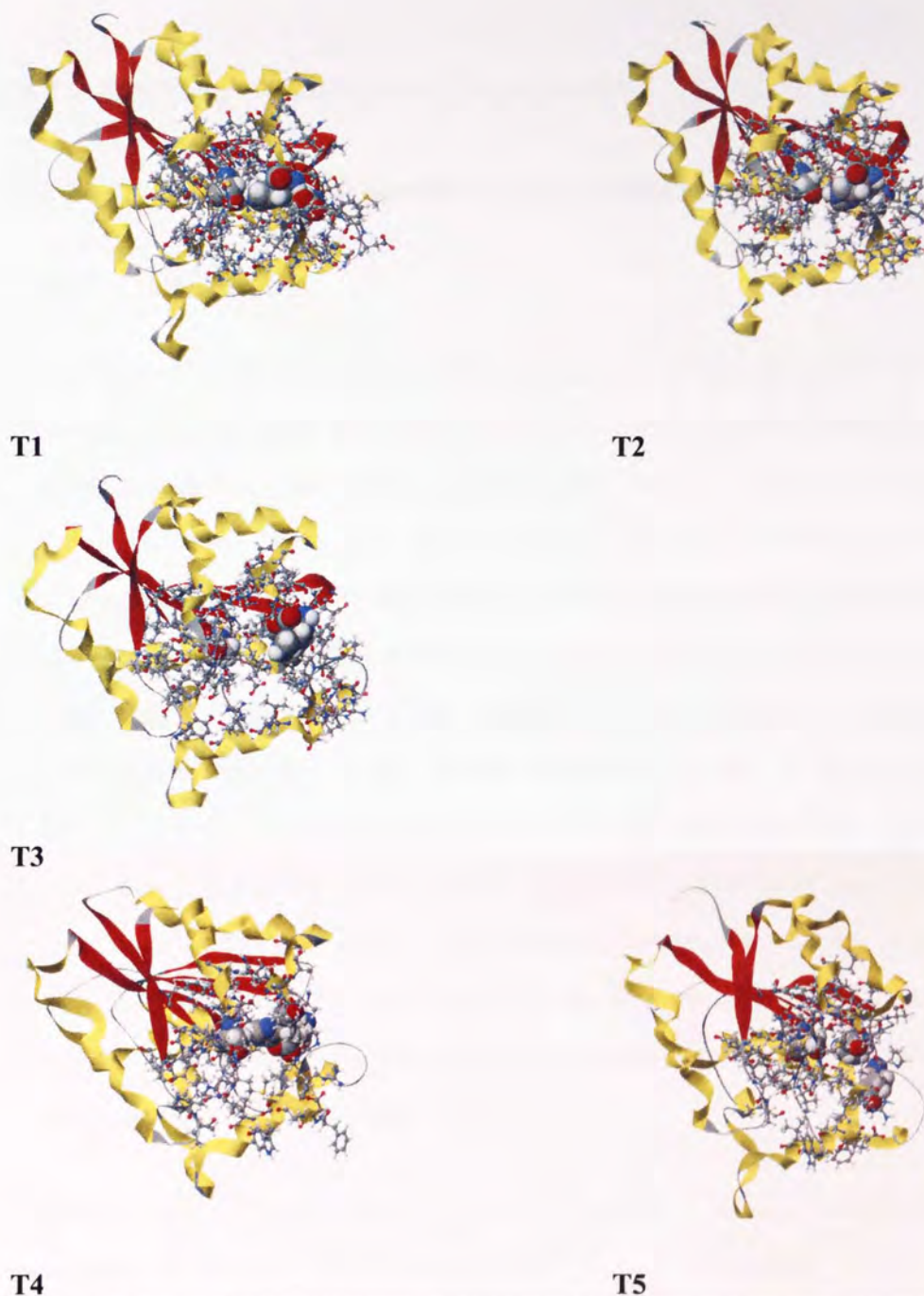


Figure 1.2.9: Five homology models.

All these five homology models are shown as flat ribbon together with all the residues that belong to the active site. The catalytic triad is shown as a space filling model. After all the modelling work had finished, the active site was defined and all the atoms were unlocked in order to use them in the docking process.

1.2.2 Docking Results and Discussion

1.2.2.1 Docking potential ligands into the esterase homology models

1.2.2.1.1: Initial docking

Under natural conditions, both protein and the drug molecules should be flexible. Therefore, flexible ligand and flexible active site conditions were chosen during the docking calculation. The ligand set used in the docking came from Maybridge's screening database. As a great number compounds (around 59,000) were in the Maybridge set, the Chem-X (1999.1 release, Oxford Molecular) programme was used to make a pharmacophore based diversity scan. The exclusion limit was set to 10%. At the end of this search, 290 compounds were chosen to represent the pharmacophore diversity of the 59,000 Maybridge set (D. L. Rathbone, Aston University). This reduced set was used due to the calculation time required for flexible ligand to flexible protein docking calculation, which takes four to ten hours per docking. For the same reason, one of the five homology models had to be chosen for the initial docking. **T1** was chosen for the first docking as it had the largest number of the residues inside the most favoured and additional allowed regions. The docking scores are shown in **Table 1.2.3**.

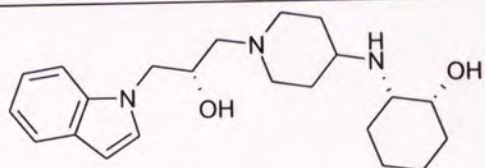
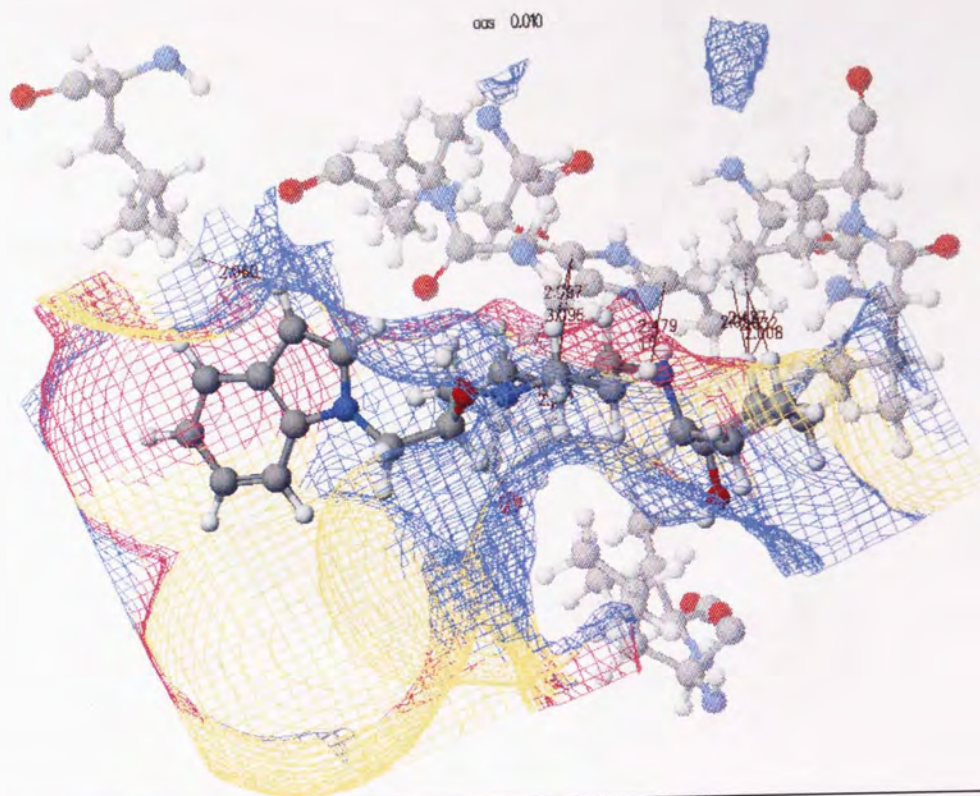
Mybridge set compound code	Docking Score (kcal/mole)	Compound code	Docking Score (kcal/mole)	Compound code	Docking Score (kcal/mole)	Compound code	Docking Score (kcal/mole)
AW 00957	-554.651	AW 00549	-347.237	S 11555	-279.685	AW 00238	-85.182
XBX 00012	-544.987	BTB 03112	-346.876	NRB 00556	-277.975	AW 00221	-84.317
BTB 02557	-519.892	AW 01206	-346.71	BTB 12462	-277.858	AC 12199	-83.319
AW 00357	-513.488	CD 01618	-346.056	EN 00191	-274.944	AC 17979	-82.878
HTS 10639	-493.051	CD 09097	-344.52	CD 00566	-273.375	KM 06366	-82.72
AW 00298	-487.397	HTS 10333	-343.237	BTB 02423	-270.285	AC 10402	-82.679
BTB 12775	-476.024	BTB 03910	-342.845	BTB 00340	-268.368	AC 10751	-81.908
AW 00648	-473.356	BTB 04277	-341.165	JFD 00728	-267.983	AW 00229	-81.762
AW 01127	-470.934	JFD 00142	-339.947	BTB 07444	-267.557	AW 00243	-81.655
SCR 00846	-465.152	BTB 04278	-337.306	BTB 04416	-266.315	AC 13432	-81.647

AW 00625	-456.951	HTS 07763	-336.924	S 11562	-265.651	AC 29314	-81.021
AW 00308	-456.459	JFD 02408	-336.006	BTB 01594	-265.269	AC 12926	-80.04
CD 04746	-452.857	AW 00264	-335.038	BTB 10718	-264.058	AW 00032	-78.992
HTS 02186	-447.783	BTB 11880	-333.253	AW 00304	-259.094	AW 00301	-78.653
HTS 03466	-445.085	AW 00277	-331.231	NRB 00588	-258.523	AC 32877	-78.402
NRB 01750	-444.477	BTB 00183	-329.639	MBE 00065	-256.392	AC 29790	-78.385
PD 00736	-441.311	AW 01132	-329.315	KM 03754	-256.23	AW 00226	-78.362
AW 00944	-440.627	JFD 03156	-328.61	BTB 11530	-253.762	AC 24175	-77.166
HTS 06808	-437.903	KM 07370	-328.259	SEW 02222	-252.954	AC 34541	-76.769
AW 00400	-432.159	DP 01439	-327.896	BTB 06949	-252.269	AC 27651	-76.726
HTS 05226	-432.029	BTB 10192	-327.449	BTB 00496	-251.668	AW 00120	-75.974
AW 00409	-429.083	DSHS 00412	-326.269	KM 04716	-251.394	AW 00111	-75.915
HTS 09268	-428.447	BTB 02809	-325.628	AW 00963	-251.001	AC 33706	-75.747
HTS 01069	-426.422	AW 00547	-325.17	DP 02022	-248.406	AW 00029	-75.079
BTB 02634	-414.171	BTB 00182	-324.957	AW 01149	-248.175	AC 11641	-73.041
RJC 00553	-413.621	BTB 07024	-322.729	BTB 02406	-246.689	AC 10033	-72.85
AW 00411	-411.903	BLT 00208	-322.011	CD 03667	-246.642	AW 00137	-71.448
BTB 14837	-407.831	CD 00461	-321.912	GK 02486	-245.708	AC 13220	-71.304
AW 00407	-407.243	GK 01849	-320.045	CD 09853	-245.031	AC 27191	-70.17
BTB 06094	-402.315	BTB 02094	-319.781	KM 03988	-242.298	AW 00038	-69.429
AW 00996	-398.321	BLT 00117	-318.583	BTB 06117	-239.568	AC 22864	-68.643
AW 00946	-397.019	SEW 05791	-317.865	BTB 15076	-238.516	AW 00030	-68.566
BTB 02635	-394.452	JA 00031	-316.928	KM 03912	-236.415	AW 00155	-67.93
AW 00382	-394.299	NRB 00464	-316.333	NRB 04495	-235.35	AW 00213	-66.935
AW 00406	-392.466	BTB 09150	-314.719	JFD 00938	-233.948	AW 00170	-66.254
DP 01062	-388.337	CD 00408	-314.626	BTB 10304	-233.103	AW 00142	-64.749
BTB 13868	-386.72	AW 00329	-312.895	BTB 11514	-232.415	AC 29726	-64.678
BTB 10232	-384.236	KM 00612	-312.548	RJF 01394	-230.346	AC 29083	-64.52
BTB 13577	-384.021	CD 01691	-312.484	AW 00467	-229.008	AW 00143	-64.288
CD 00932	-382.358	CD 00410	-312.382	AW 00970	-227.162	AC 12157	-63.029
KM 08561	-377.717	BTB 12147	-311.923	TL 00046	-226.396	AW 00136	-62.135
CD 01411	-374.938	BTB 00937	-310.511	BTB 08836	-216.322	AW 00178	-59.715
RF 02844	-371.022	CD 00383	-308.105	PD 00001	-203.474	AW 00248	-59.5
HTS 13333	-370.865	BTB 00990	-307.983	RF 02505	-201.427	AW 00098	-59.492
BTB 02419	-370.581	AW 00422	-307.141	BTB 11078	-201.25	AW 00168	-58.378
CD 01409	-369.998	HTS 13344	-305.553	BTB 00479	-200.477	AW 00107	-57.808
DP 01078	-369.639	NRB 00937	-305.55	BTB 00115	-197.285	BTB 09419	-56.78
HTS 05337	-366.819	AW 00554	-305.503	BTB 12360	-194.552	AW 01119	-56.56
BTB 11465	-366.616	GK 01779	-305.451	JFD 02789	-193.206	AW 00217	-56.304
BTB 14062	-366.131	BTB 01314	-305.157	NRB 02263	-189.634	AW 00211	-54.056
BTB 10107	-365.279	BTB 06116	-302.432	BTB 14691	-189.48	AW 00239	-51.661
AW 00593	-364.903	BTB 14865	-300.823	NRB 00934	-189.27	AW 00118	-50.19
CD 02050	-363.806	BTB 11271	-300.56	AW 01071	-187.459	AC 29286	-49.783
SEW 04564	-361.455	DP 00990	-298.992	BTB 09703	-182.545	AW 00114	-48.049
NRB 02672	-360.567	CD 06183	-298.5	BTBG 00152	-179.619	AW 00181	-47.431
HTS 13175	-360.374	DP 01350	-297.481	BTB 03367	-179.384	AC 31181	-46.945
HR 00096	-360.08	NRB 03988	-294.385	BTB 09071	-178.212	AC 40022	-46.903
AW 00256	-358.89	AW 00376	-293.397	BTB 09381	-174.168	AW 00196	-45.95
JFD 02451	-358.856	RJC 02376	-292.393	TL 00237	-166.19	AW 00216	-45.899
AW 00281	-358.254	BTB 01284	-291.696	SEW 03168	-166.058	AC 12308	-45.611
RDR 02110	-357.328	BTB 00865	-291.145	BTB 10412	-159.765	AC 22693	-45.514

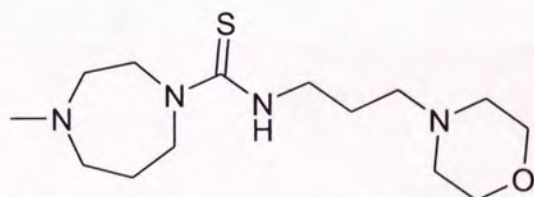
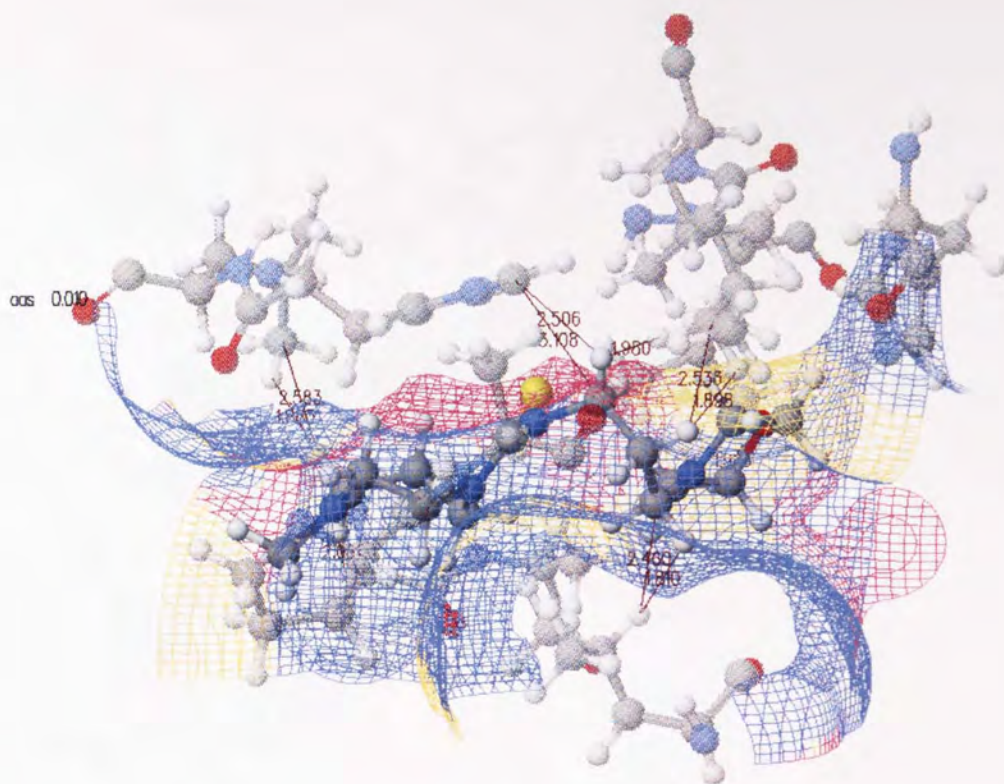
AW 00445	-357.173	GK 01846	-290.494	AW 00428	-159.56	AC 33318	-45.384
AW 00981	-356.107	BTB 10379	-290.223	DP 01824	-123.589	AC 10608	-44.904
AW 00257	-355.62	BTB 09167	-289.296	BTB 10021	-122.21	AC 10702	-44.21
JFD 01984	-354.03	BTB 14001	-285.633	AC 29806	-105.95	AC 12028	-42.522
AW 00509	-353.691	CD 11605	-284.723	BTB 05209	-89.65	AC 40100	-42.109
BTB 10720	-352.983	KM 08327	-284.702	JFD 02765	-88.95	AC 30194	-40.215
AW 00541	-351.353	CD 09096	-282.485	AW 00174	-88.523	AC 11546	-38.803
BTB 01773	-350.659	BTB 13822	-281.883	AW 00230	-88.221	AC 34762	-30.99
BTB 06499	-350.01	AW 00744	-281.531	AC 22646	-87.514	BTB 10981	-25.89
CD 08635	-348.654	BTB 10716	-281.456	AC 35726	-87.479	BTB 11055	1366.417
HTS 01502	-347.556	BTB 00016	-281.254	AW 00201	-86.903		
RJC 03401	-347.472	EN 00076	-280.406	AC 18635	-85.349		

Table 1.2.3: Initial docking score for the Maybridge reduced set into T1

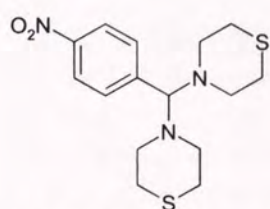
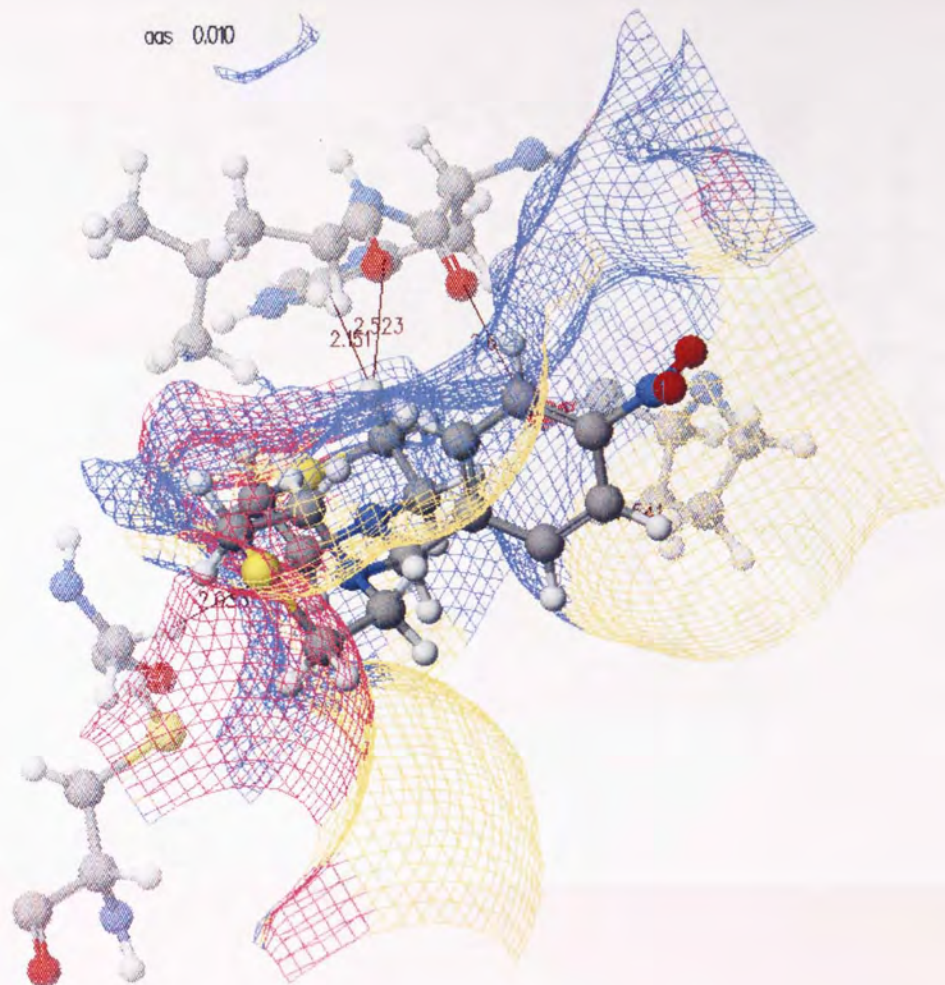
In **Table 1.2.3**, the data was sorted by the order of docking score starting from the most negative one. However, the docking score alone cannot describe a protein-ligand interaction. A close look at the docked structure is also needed. The following figures show part of the docked structures.



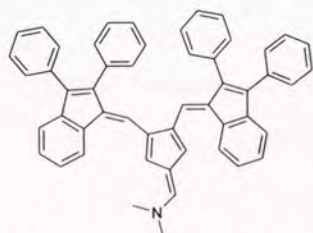
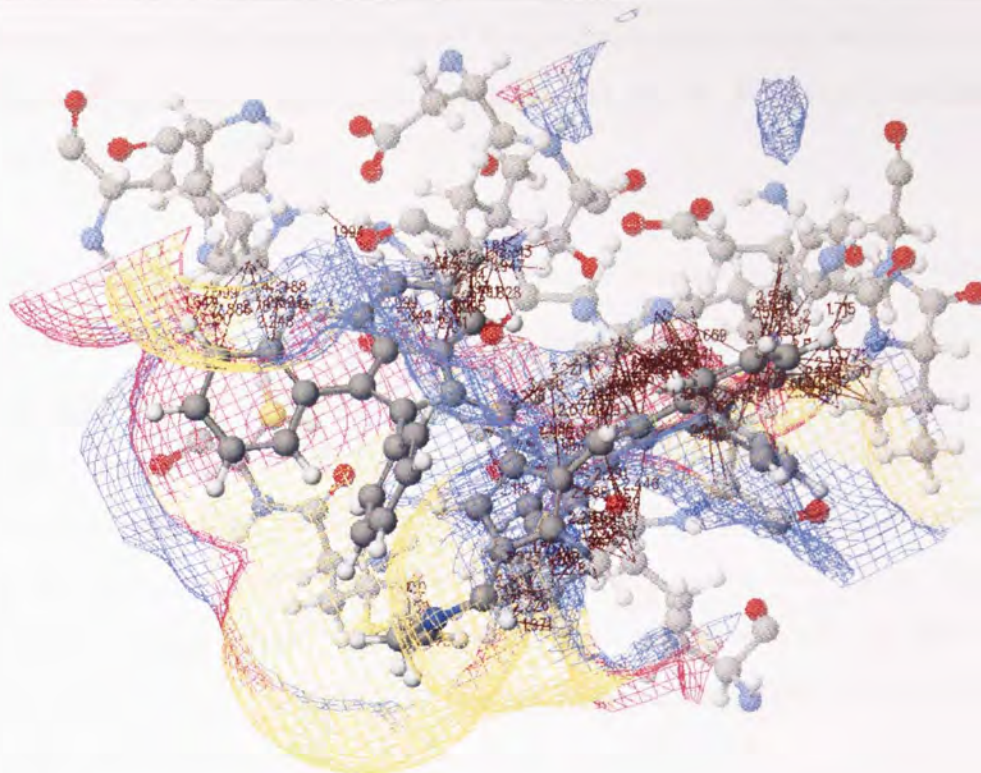
1: Part of docked structure and the molecule AW00957 (-554.651 kcal/mole)



2: Part of docked structure and the molecule BTB02557 (-544.987 kcal/mole)



3: Part of docked structure and the molecule S11055 (-279.685 kcal/mole)



4: Part of docked structure and the molecule BTB11055 (1366.417)

Figure 1.2.10: Four examples of docked structures. Only the ligand and the nearby residues are shown. The coloured web shows the calculated surface pocket. The brown thread with numbers shows the bumps between ligand and residues. (Created in the CaChe WorkSpace.)

Among those four structures, the first two came from the lowest docking score, which was below -500 kcal/mole. The third one came from the middle, which was around -300 kcal/mole. The last had a very high docking score. After the first and the second conformations in **Figure 1.2.10** were compared, it was found that, although AW00957 in **Table 1.2.3** shows the lowest docking score, it still has some bumps (steric clashes). And most importantly, part of the benzene ring crossed the surface web. At the same time, the indole substituent was pointing outside of the active site, which may mean that this part was not involved in the binding. On the other hand, in

the second graph, BTB02557 fitted quite well inside the active site. Some of the hydrogen atoms were point outside of the surface pocket, which can be accepted. Overall, by comparing their docked structure, BTB02557 should be considered as have a better binding affinity.

Besides these two best-scored molecules, two other less favoured structures were also shown in the **Figure 1.2.10**. It can be seen that the molecule of BTB 11055 had a massive structure with too many aromatic rings that was too big to fit inside the active site. A close look at the docked structure reveals nearly half of the benzene rings were pointing outside of the surface pocket and plenty of bumps existed. This molecule had the highest docking score which was above zero. In the third figure, molecule S11055 had fewer bumps. However, it only fitted part of the active site and nearly half of the molecule showed on the surface of the protein. This molecule has less contact with the protein.

A single docking may not be able to represent all the binding affinities. Therefore, docking the compounds into the same protein several times may find a better docking score and position. The top 10 ranked compounds from **Table 1.2.3** were chosen and docked into **T1** another two times in order to see if there were lower docking scores and a better docking position. The score from all these three docking is shown in **Table 1.2.4**.

Compound Code.	Docking score (kcal/mole)			
	1 st docking	2 nd docking	3 rd docking	Average
XBX00012	-544.987	-602.155	-593.292	-580.14
AW00957	-554.651	-473.156	-538.064	-521.96
BTB02557	-519.892	-506.48	-509.594	-511.99
HTS10639	-493.051	-505.448	-506.529	-501.68
BTB12775	-476.024	-478.323	-489.666	-481.34
SCR00846	-465.152	-438.859	-503.836	-469.28
AW00357	-513.488	-451.633	-297.569	-420.9
AW01127	-470.934	-354.627	-397.173	-407.58
AW00298	-487.397	-332.488	-329.506	-383.13
AW00648	-473.356	-413.559	-245.274	-377.4

Table 1.2.4: Average docking score of top 10 ranked compounds.

The data shown in **Table 1.2.4** was sorted by the average docking score. It can be seen that compounds AW00357 and AW00648 had a huge difference on the third docking. Their docked structures were checked and it was found these differences were caused by their docking position where nearly all parts of the molecules were sitting outside of the active site. This may due to calculation errors. Then, the next noticeable fact was that compound XBX00012 had the lowest total docking score and the three docking score were quite similar. By checking their docked pose, it was found that all these three structures were docked in a similar position and orientation. One of the structures is given in **Figure 1.2.11**.

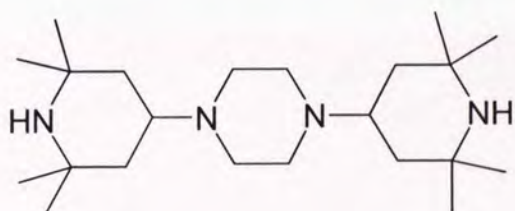
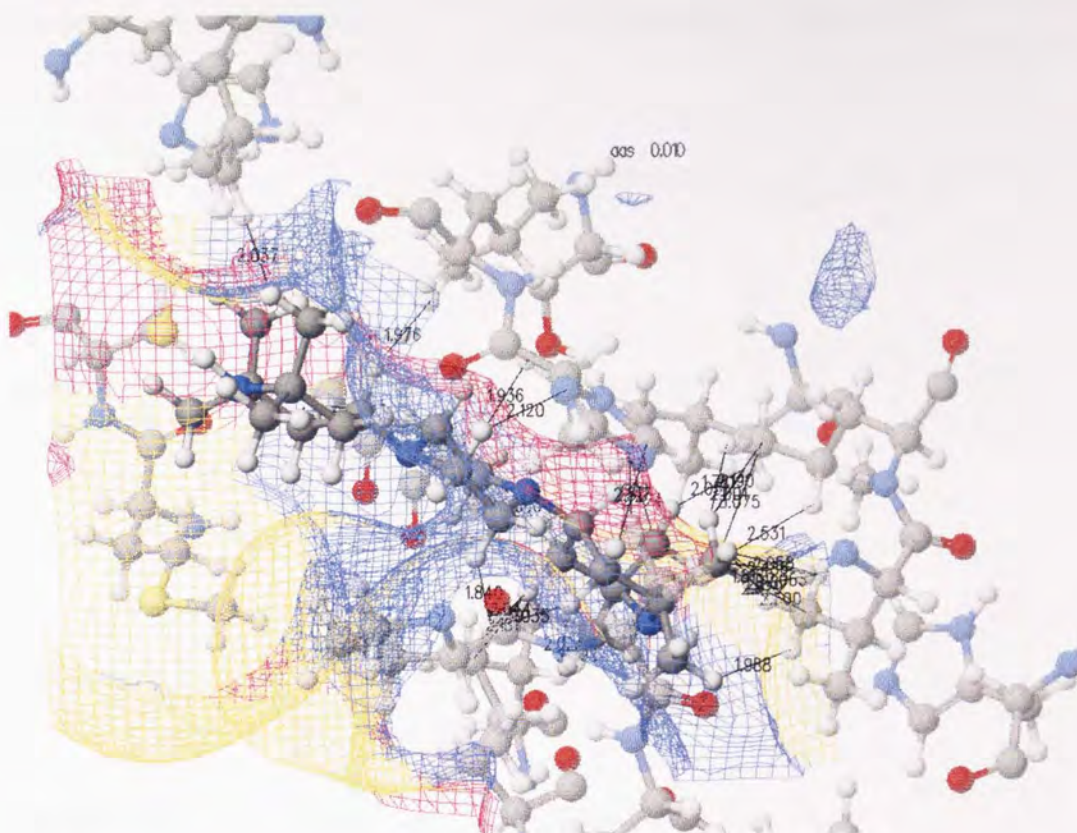


Figure 1.2.11: Partial docked structure and the structure of XBX00012.

Although the docking score was low enough, the structure analysis showed that one of the methyl groups was pointing outside of the calculated active site surface, which caused several bumps. This phenomenon could be seen in the **Figure 1.2.11**. Therefore, the methyl group could be modified in the next generation of design. It was also noticed that there was some space around the middle ring. This may mean that one or more groups could be added there. The structure in **Table 1.2.5** shows the next generations of XBX00012.

Compound code	Structure	Compound code	Structure
XBX00012 C1_1		XBX00012 C1_9	
XBX00012 C1_2		XBX00012 C1_10	
XBX00012 C1_3		XBX00012 C1_11	
XBX00012 C1_4		XBX00012 C1_12	
XBX00012 C1_5		XBX00012 C1_13	
XBX00012 C1_6		XBX00012 C1_14	
XBX00012 C1_7		XBX00012 C1_15	
XBX00012 C1_8			

Table 1.2.5: First generation of designed compounds arising from Maybridge XBX00012.

The compound AW00957 had a similar shape as the XBX00012 with three rings linked by longer chains. It was noticed that the second docking gave a slightly different docking score from other two. This was because that in the second docked structure, the indole substituent was pointing inside the active site pocket and most parts of the molecule was at the surface. This kind of the orientation should be less

favourable than the other two. A docked structure of AW00957 is given in **Figure 1.2.12**.

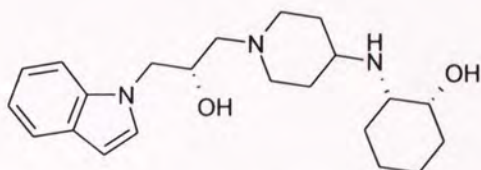
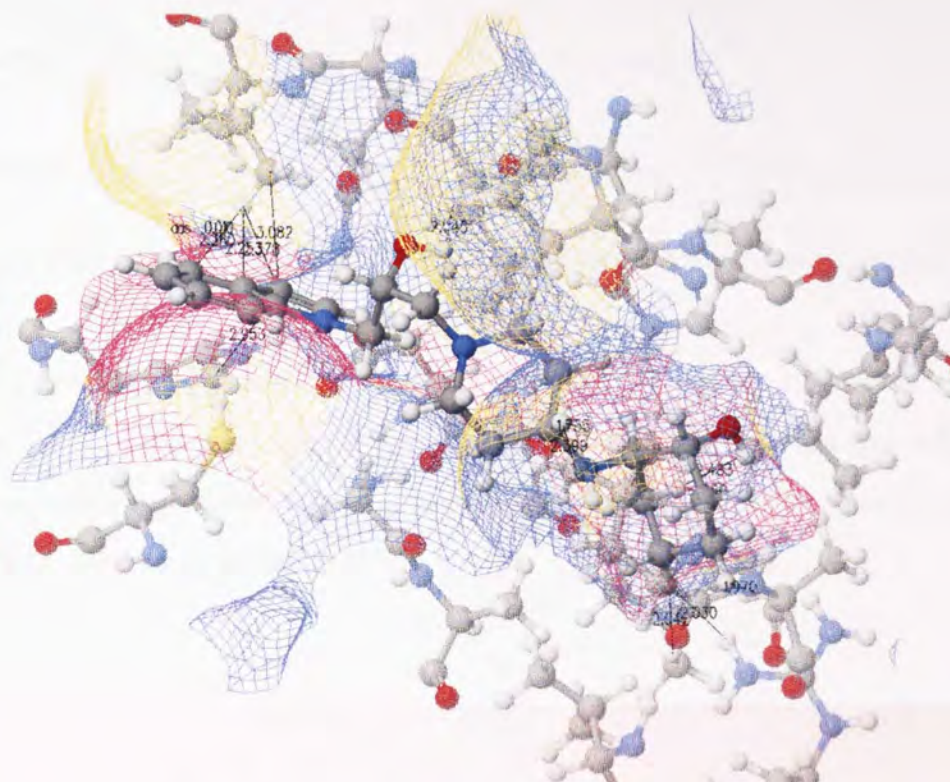


Figure 1.2.12: Partial docked structure and the structure of AW00957.

The docked complex shown in **Figure 1.2.12** should be the most favoured position for this compound. By checking the calculated active site pocket, there were some spaces available. Therefore, a little modification could be done on the cyclonehexane ring, near to the secondary amine and the piperidine ring near to the nitrogen atom side. The next generation of AW00957 compounds will be based on those modifications and the designed structures are shown in **Table 1.2.6**.

Compound code	Structure	Compounds code	Structure
AW00957 C1_1		AW00957 C1_4	
AW00957 C1_2		AW00957 C1_5	
AW00957 C1_3		AW00957 C1_6	

Table 1.2.6: First generation of designed compounds arising from Maybridge AW00957.

The three docking scores of compounds of BTB02557 showed that they were nearly the same. This was because they had the similar docking position where the morpholine substituent pointed inside the active site pocket. As shown in **Figure 1.2.13**, the docked structure did not show too many steric clashes.

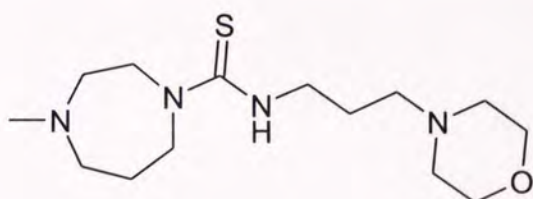
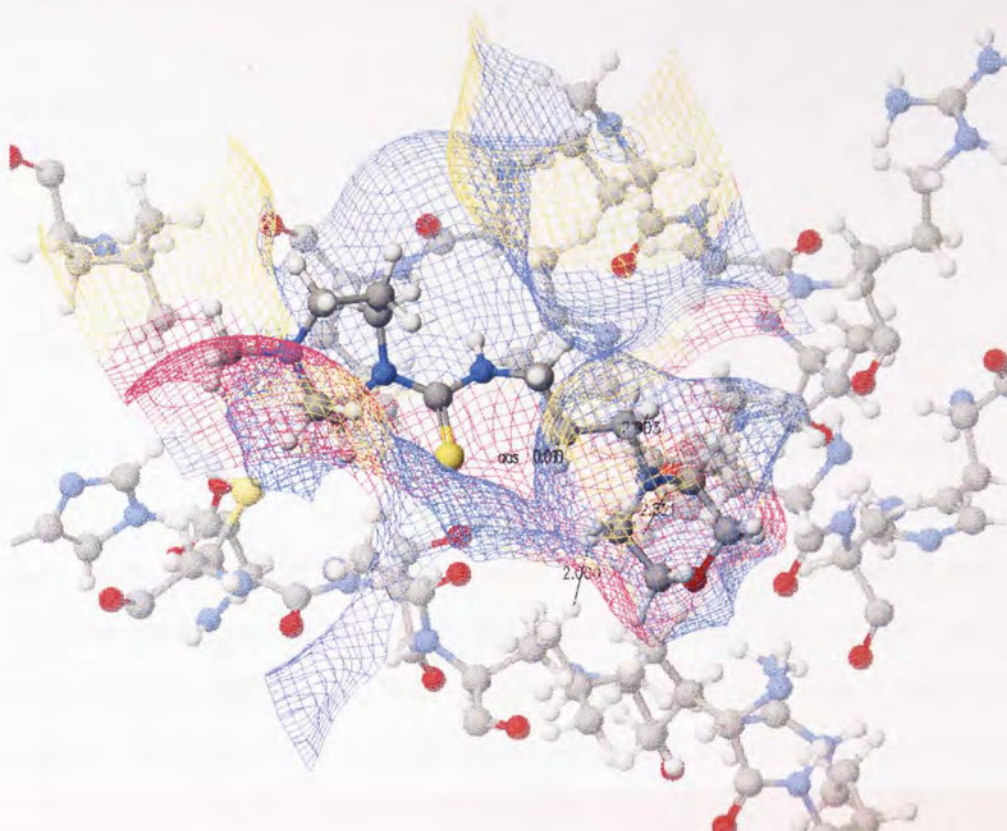


Figure 1.2.13: Partial docked structure and the structure of BTB02557.

It can be seen from **Figure 1.2.13** that although there were few bumps present, only a small part was inside the active site. There might be the possibility that the seven-member ring side together with the thio- group could point inside the active site. In this situation, they may give a better docking score. The first thing considered was reducing the size of the thio- group to a carbonyl group. Then, the modification on the morpholine substituent could be an alternative, or even making some changes on the seven-member ring substituent. These modified compounds are shown in **Table 1.2.7**.

Compound code	Structure	Compounds code	Structure
BTB02557 C1_1		BTB02557 C1_5	
BTB02557 C1_2		BTB02557 C1_6	
BTB02557 C1_3		BTB02557 C1_7	
BTB02557 C1_4		BTB02557 C1_8	

Table 1.2.7: First generation of designed compounds arising from Maybridge BTB02557.

The structure of HTS10639 was similar as XBX00012. As can be seen in **table 1.2.4** the three docking scores were more or less the same. Therefore, its next modification was similar to XBX00012. Next, the compound BTB12775 was a long molecule. Three of the docked positions were found to have a similar positions and orientations. One of the examples is given in **Figure 1.2.14**.

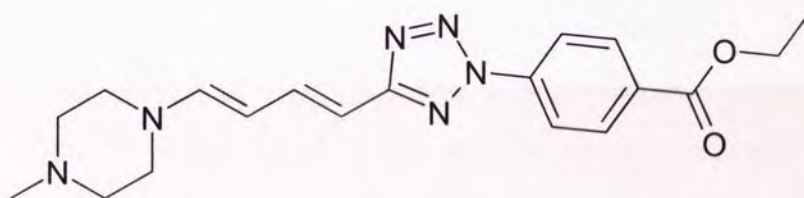
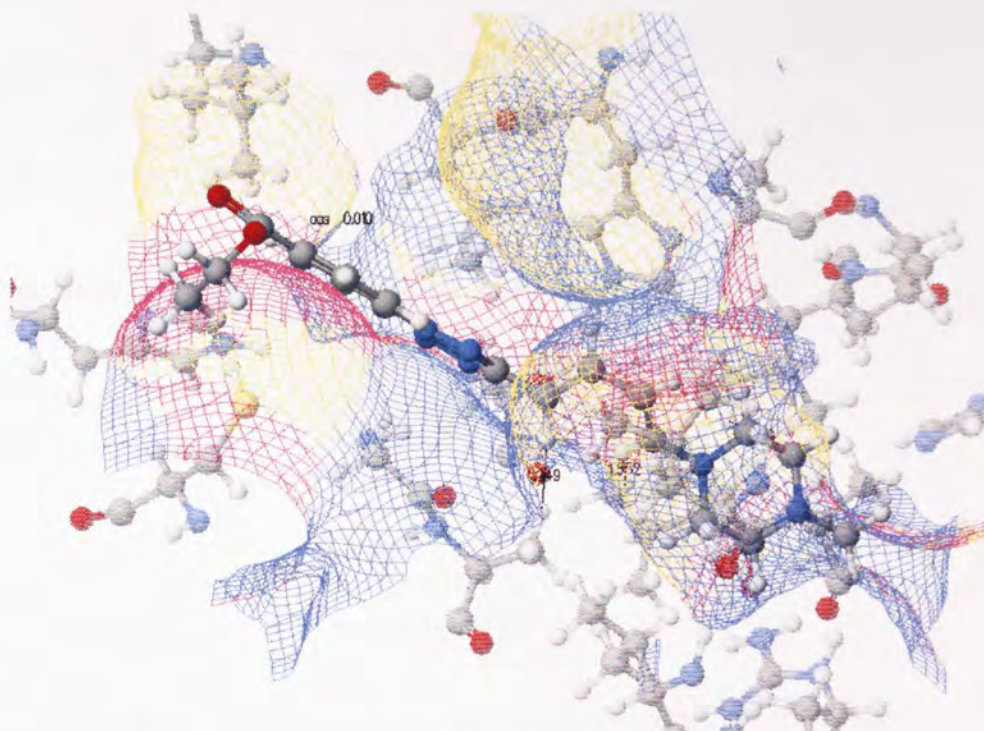


Figure 1.2.14: Partial docked structure and the structure of BTB12775.

As can be seen the molecule was very long and the docked structure showed that most of the substituent was outside of the active site. Therefore, the main modification for the next generation will be made the shorter molecules. The **Table 1.2.8** shows the modified structures.

Compound code	Structure
BTB12775 C1_1	
BTB12775 C1_2	

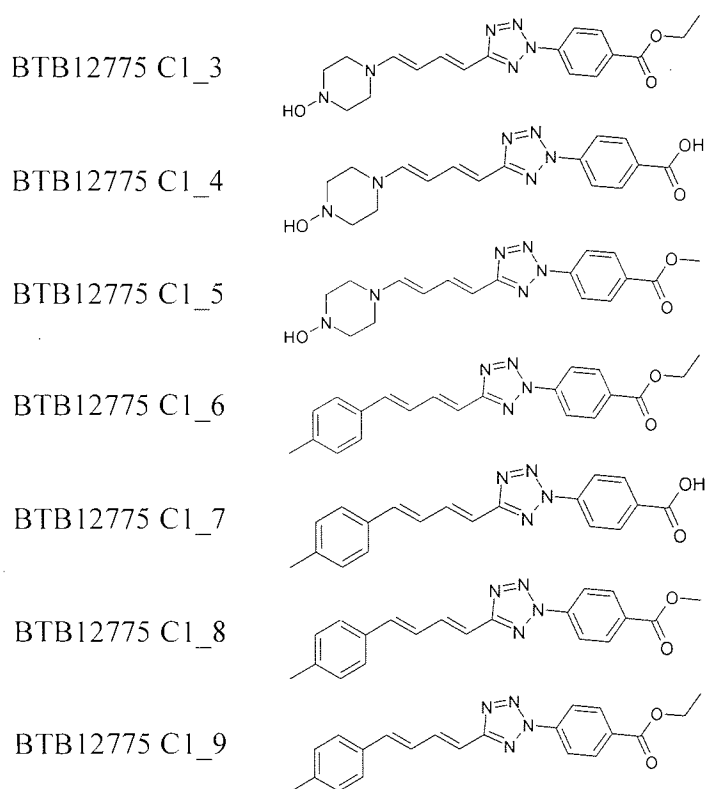


Table 1.2.8: First generation of designed compounds arising from Maybridge BTB12775.

The compound of SCR00846 was a long molecule which is quite similar to BTB12775. **Figure 1.2.15** shows one of the docked structures.

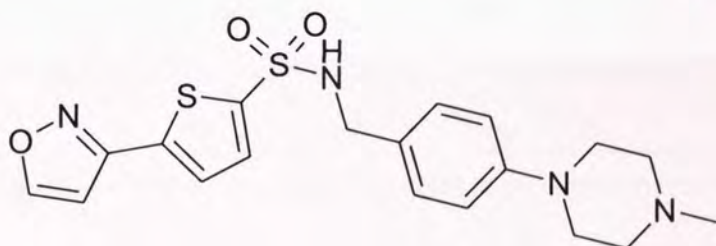
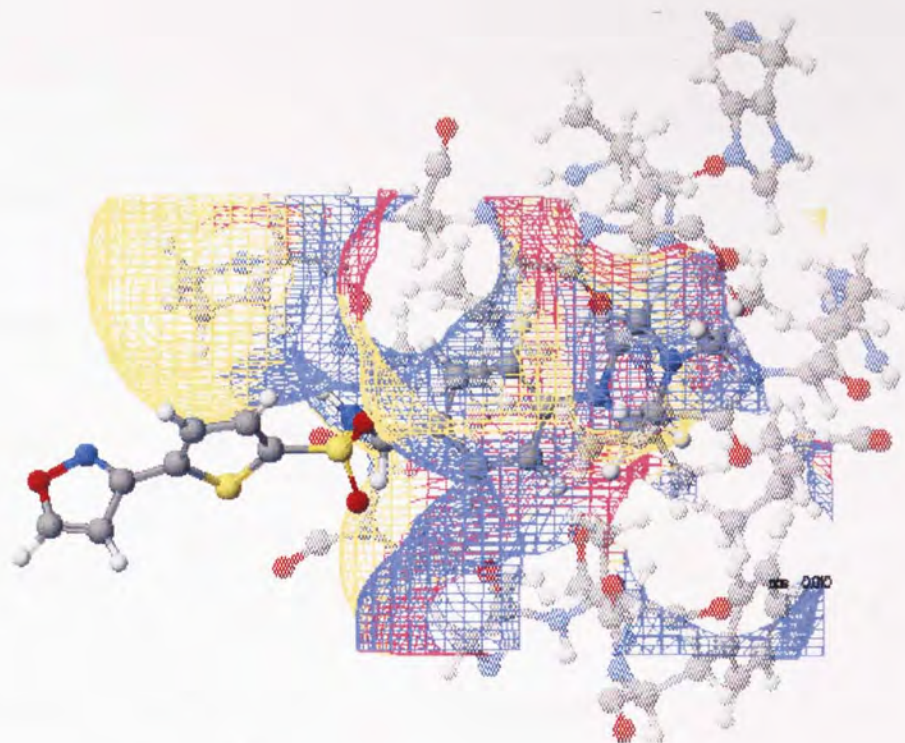


Figure 1.2.15: Partial docked structure and the structure of SCR00846

It is very clear that the isoxazole substituent falls outside the active site and all the three docked structures showed the same situation. Therefore, a replacement of isoxazole group was considered. A small change on the piperazine ring was also tried.

Table 1.2.9 shows the modified structure of SCR00846.

Compound code	Structure	Compounds code	Structure
SCR00846 C1_1		SCR00846 C1_8	
SCR00846 C1_2		SCR00846 C1_9	
SCR00846 C1_3		SCR00846 C1_10	
SCR00846 C1_4		SCR00846 C1_11	
SCR00846 C1_5		SCR00846 C1_12	
SCR00846 C1_6		SCR00846 C1_13	
SCR00846 C1_7		SCR00846 C1_14	

Table 1.2.9: First generation of designed compounds arising from Maybridge SCR00846.

Overall, all these selected compounds were found to be docked into the homology model in a linear fashion. The main difference of the docked structures was the head group might be inside or outside of the active site pocket. Therefore, the modifications for the first generation of the compounds was similar.

1.2.2.1.2: The first generation of docking

As discussed above, several changes were made to each of the selected compounds. Those modified structures were considered as the first generation. Then, all those first

generation compounds together with their parent structures were re-docked into **T1**.

The **Table 1.2.10** shows the sorted docking score.

Compound code	Docking score (kcal/mole)			
	Docking 1	Docking 2	Docking 3	Average
XBX00012C1_8	-568.259	-602.154	-500.89	-557.101
AW00957C1_3	-554.786	-574.921	-530.553	-553.42
XBX00012C1_3	-398.875	-602.343	-612.548	-537.922
XBX00012	-590.583	-415.362	-593.645	-533.197
AW00957C1_4	-579.14	-469.997	-519.117	-522.751
BTB02557C1_6	-474.891	-478.59	-614.769	-522.75
XBX00012C1_9	-386.06	-554.922	-607.378	-516.12
XBX00012C1_7	-396.571	-578.332	-561.262	-512.055
XBX00012C1_4	-370.019	-622.792	-543.217	-512.009
AW00957	-511.78	-480.942	-542.335	-511.686
BTB02557C1_1	-547.355	-544.113	-437.736	-509.735
BTB02557	-483.543	-522.586	-512.331	-506.153
XBX00012C1_12	-555.623	-369.719	-587.411	-504.251
BTB02557C1_3	-524.54	-484.745	-501.66	-503.648
BTB02557C1_5	-344.156	-522.585	-614.769	-493.837
BTB02557C1_4	-574.204	-555.821	-344.158	-491.394
AW00957C1_2	-490.4	-502.124	-461.144	-484.556
BTB12775	-489.779	-473.494	-489.577	-484.283
AW00957C1_1	-415.781	-501.06	-518.185	-478.342
BTB12775C1_2	-481.842	-488.397	-456.443	-475.561
XBX00012C1_13	-419.509	-412.248	-592.439	-474.732
XBX00012C1_2	-403.686	-395.447	-603.999	-467.711
AW00957C1_5	-498.908	-356.725	-535.81	-463.814
XBX00012C1_14	-378.624	-401.135	-608.853	-462.871
HTS10639	-396.764	-456.34	-508.188	-453.764
BTB12775C1_5	-436.106	-466.037	-456.267	-452.803
XBX00012C1_11	-412.693	-340.587	-575.239	-442.84
BTB12775C1_3	-437.847	-447.26	-442.251	-442.453
XBX00012C1_15	-413.99	-413.641	-494.158	-440.596
BTB02557C1_8	-476.306	-340.666	-480.522	-432.498
SCR00846	-449.672	-356.519	-480.713	-428.968
BTB12775C1_4	-443.832	-413.379	-412.126	-423.112
SCR00846C1_5	-420.886	-411.074	-428.397	-420.119
SCR00846C1_11	-450.644	-380.15	-419.704	-416.833
HTS10639C1_2	-442.121	-444.035	-336.905	-407.687
BTB02557C1_2	-453.692	-351.966	-406.155	-403.938
AW00957C1_6	-427.143	-329.679	-450.039	-402.287
BTB02557C1_7	-198.505	-498.655	-499.274	-398.811
SCR00846C1_3	-293.51	-455.138	-419.708	-389.452
BTB12775C1_9	-388.799	-363.779	-409.664	-387.414
SCR00846C1_7	-443.222	-281.052	-436.351	-386.875

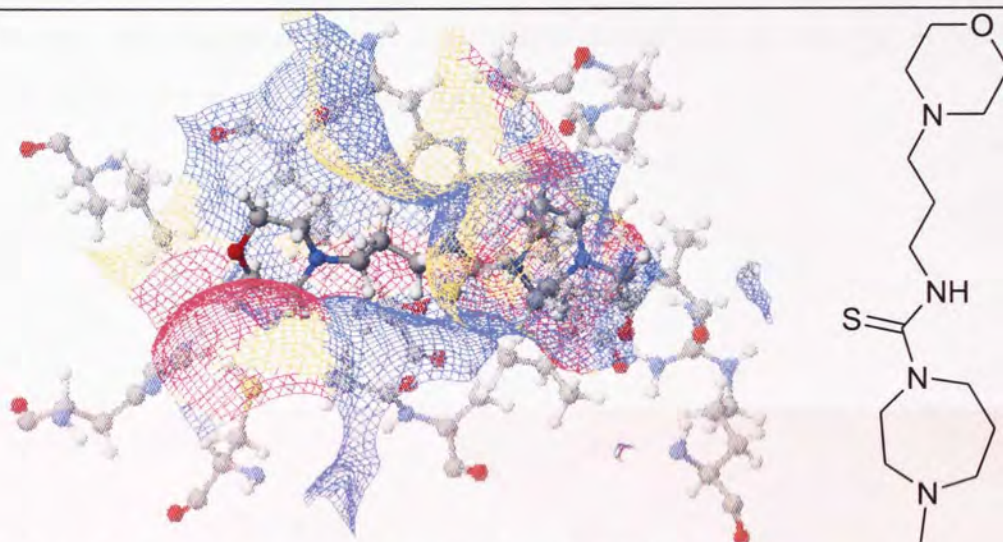
BTB12775C1_8	-375.848	-376.388	-405.122	-385.786
SCR00846C1_6	-322.03	-397.193	-435.043	-384.755
SCR00846C1_8	-426.087	-326.333	-398.276	-383.565
SCR00846C1_1	-396.848	-303.117	-438.946	-379.637
BTB12775C1_1	-242.159	-448.978	-438.369	-376.502
XBX00012C1_6	-561.515	-125.051	-426.476	-371.014
SCR00846C1_4	-434.768	-305.465	-366.332	-368.855
SCR00846C1_13	-310.5	-326.185	-463.483	-366.723
SCR00846C1_10	-305.168	-315.222	-469.953	-363.448
SCR00846C1_12	-316.918	-344.489	-423.374	-361.594
BTB12775C1_7	-362.92	-364.76	-354.901	-360.86
BTB12775C1_6	-179.504	-477.142	-407.534	-354.727
XBX00012C1_1	-356.665	-338.355	-359.935	-351.652
HTS10639C1_1	-454.391	-412.336	-173.695	-346.807
SCR00846C1_9	-328.544	-280.279	-415.129	-341.317
SCR00846C1_2	-285	-362.327	-265.648	-304.325
XBX00012C1_10	-286.984	-167.505	-452.622	-302.37
SCR00846C1_14	-294.359	-341.299	-271.404	-302.354
XBX00012C1_5	-72.409	-285.317	-234.468	-197.398

Table 1.2.10: The ranked docking score by average of first generation of those selected best Mybridge compounds.

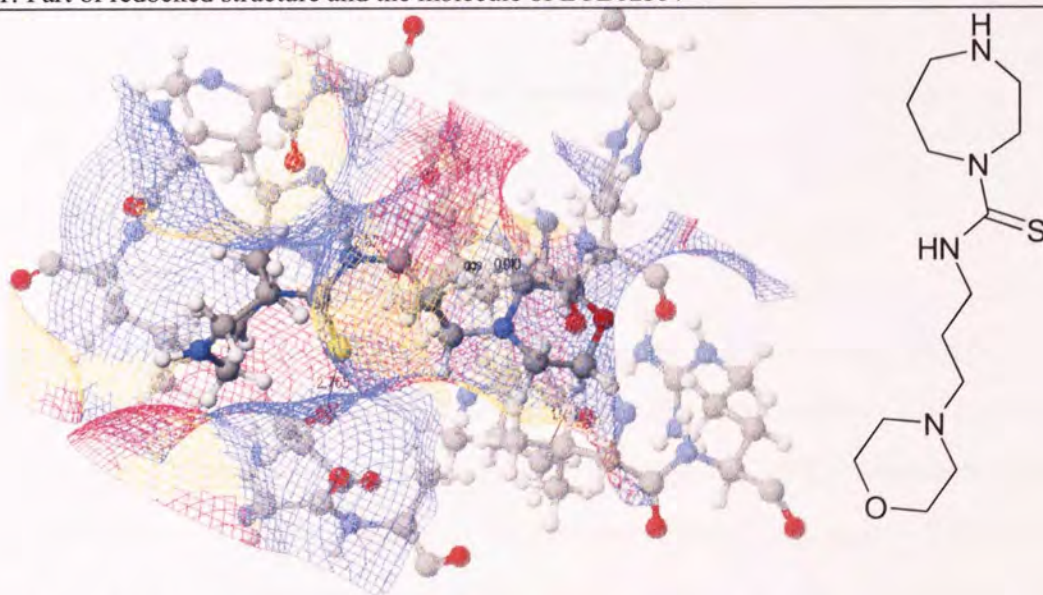
In **Table 1.2.10**, the first thing can be noticed is that the modified compounds XBX00012 and BTB02557 had the lowest docking score which should mean that they could be used as potential inhibitors in this case. The docked structures also fitted well inside the active site. Other compounds that had higher docking scores may, due to their molecule structures being a little bit longer, had part of the molecule protruding outside of the active site pocket. Therefore, these two types of the molecules were considered as the leading compounds for further docking.

It was also noticeable that some docking scores of modified compounds of XBX00012 fall at the end of the ranking table. Therefore, the docked structures of XBX00012 C1_5 and XBX00012 C1_1 were checked in order to find out why they received low docking scores. An example of them is given in **Figure 1.2.16**.

the seven-member ring part. The docked structure analysis suggested that a carbonyl group should be better than a thio-group. This is because when the steric clashes were checked, more bumps were found when the thio-group was present. This phenomenon might be due to the greater bond length of the thio-group to the carbonyl group. Another interesting phenomenon was noticed. If the methyl group on the seven-member ring was deleted or moved closer to carbonyl group, the docked position would change. **Figure 1.2.17** shows a comparison of two different structures.



1: Part of redocked structure and the molecule of BTB02557



2: Part of docked structure and the molecule of BTB02557 C1_8

Figure 1.2.17: Structure comparison of compounds BTB02557 in their docked position.

It can be seen clearly that in the first figure in **Figure 1.2.17** that the seven-member ring substituent was pointing inside the active site pocket when the methyl group was on the fourth position. However, the second graph shows a complete opposite docking pose. If the methyl group was moved, the morpholine substituent was pointing inside the pocket. The changing of methyl group gave a different docking pose together with a slight lower docking score suggest one possible modification on the third generation of molecule design.

Overall, the leading structures for the next generations of molecule design were decided and are shown in **Table 1.2.11**.

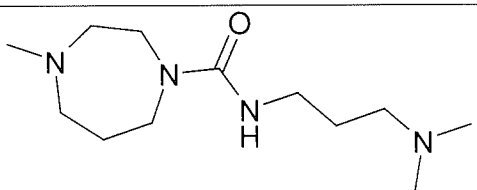
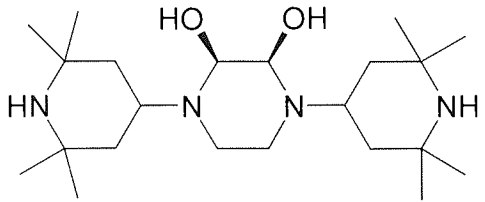
Compound Code	Structure
BTB02557C1_5	
XBX00012 C1_8	

Table 1.2.11: Leading structures of third generation of the molecule design.

1.2.2.1.3: Further generations of docking

The next generations of drug design were based on the result of the first generation of the compounds as well as the possibility of further chemical synthesis. For the series of BTB02557 like compounds, the main modifications were still concentrated on the methyl group which was attached to the seven-membered ring, carbonyl group and the *tert*-amine substituent. The main modifications for the XBX00012 like compounds were concentrated on the middle part. The newly designed compounds

were docked into all those five homology models from **T1** to **T5**. The best compounds were chosen for the chemical synthesis. The docking score of third generation together with their parent compounds are shown in **Table 1.2.12**.

Compounds Code *	T1 docking score (kcal/mole)	T2 docking score (kcal/mole)	T3 docking score (kcal/mole)	T4 docking score (kcal/mole)	T5 docking score (kcal/mole)	Average (kcal/mole)
BTB 02557	-503.3	-432.5	-546.78	-409.05	-534.378	-485.202
BTB02557 2nd	-524.31	-497.987	-598.583	-406.155	-489.593	-503.326
BTB02557 C2_1	-538.495	-484.188	-614.769	-437.736	-567.322	-528.502
BTB02557 C2_2	-484.937	-455.145	-559.908	-395.287	-495.262	-478.108
BTB02557 C2_3	-427.923	-451.165	-545.601	-410.408	-430.751	-453.17
BTB02557 C2_4	-486.765	-429.473	-461.355	-487.586	-480.318	-469.099
BTB02557 C2_5	-460.294	-371.147	-512.533	-338.884	-362.66	-409.104
BTB02557 C3_1	-491.07	-535.556	-574.026	-501.66	-526.965	-525.855
BTB02557 C3_2	-562.331	-543.679	-641.207	-344.158	-559.197	-530.114
BTB02557 C3_3	-566.053	-481.372	-659.11	-614.769	-518.911	-568.043
BTB02557 C3_4	-504.367	-458.957	-503.747	-545.268	-472.239	-496.916
BTB02557 C3_5	-498.364	-504.908	-554.58	-400.801	-516.38	-495.007
BTB02557 C3_6	-451.452	-513.501	-503.45	-411.674	-521.813	-480.378
BTB02557 C3_7	-448.854	-460.932	-490.637	-328.669	-486.61	-443.14
BTB02557 C3_8	-359.259	-442.047	-505.617	-614.769	-463.648	-477.068
BTB02557 C3_9	-477.822	-469.756	-557.661	-370.001	-461.798	-467.408
BTB02557 C3_10	-509.88	-493.337	-611.751	-408.636	-519.118	-508.544
BTB02557 C3_11	-498.071	-256.706	-529.716	-420.098	-492.715	-439.461
BTB02557 C4_1	-475.654	-463.505	-586.368	-396.837	-487.216	-481.916
BTB02557 C4_2	-432.37	-439.261	-551.798	-404.381	-509.368	-467.436
BTB02557 C4_3	-390.2	-460.369	-487.598	-381.972	-449.172	-433.862
BTB02557 C4_4	-508.373	-445.342	-540.622	-398.759	-559.521	-490.523
BTB02557 C4_5	-462.08	-474.311	-525.446	-384.075	-485.845	-466.351
BTB02557 C4_6	-522.051	-543.953	-624.737	-524.658	-524.658	-548.011
BTB02557 C4_7	-548.087	-553.373	-667.149	-445.951	-600.896	-563.091
BTB02557 C4_8	-312.488	-395.003	-497.63	-395.003	-497.63	-419.551
BTB02557 C5_1	-422.959	-381.319	-542.261	-414.317	-517.703	-455.712
BTB02557 C5_101	-466.162	-435.391	-526.918	-410.091	-471.803	-462.073
BTB02557 C5_102	-480.183	-470.163	-533.213	-408.841	-505.608	-479.602
BTB02557 C5_103	-481.206	-454.061	-599.831	-434.064	-509.064	-495.645
BTB02557 C5_104	-506.101	-459.929	-625.522	-401.914	-525.589	-503.811
BTB02557 C5_105	-549.373	-321.552	-629.313	-451.351	-576.11	-505.54
BTB02557 C5_2	-498.071	-449.017	-529.716	-420.098	-492.715	-477.923
BTB02557 C5_201	-477.702	-423.162	-502.829	-388.679	-453.109	-449.096
BTB02557 C5_202	-455.838	-472.195	-587.439	-395.303	-493.405	-480.836
BTB02557 C5_203	-488.131	-477.973	-480.523	-386.637	-503.956	-467.444
BTB02557 C5_204	-497.079	-514.751	-508.059	-425.119	-481.919	-485.385
BTB02557 C5_205	-538.473	-412.768	-593.673	-407.693	-535.323	-497.586
BTB02557 C5_3	-455.659	-453.682	-603.193	-411.934	-411.862	-467.266
BTB02557 C5_301	-458.596	-380.199	-535.492	-435.839	-540.637	-470.153
BTB02557 C5_302	-488.011	-482.846	-544.104	-476.779	-552.123	-508.773
BTB02557 C5_303	-512.17	-523.028	-548.24	-536.658	-544.691	-532.957

BTB02557 C5_304	-508.62	-507.189	-598.018	-421.732	-483.011	-503.714
BTB02557 C5_305	-376.067	-565.903	-654.457	-376.067	-508.519	-496.203
XBX 00012	-622.589	-568.69	-602.58	-428.635	-428.636	-530.226
XBX00012 C2_1	-495.76	-239.5	-401.641	-314.762	-489.576	-388.248
XBX00012 C2_2	-512.017	-377.879	-547.124	-357.284	-596.282	-478.117
XBX00012 C2_3	-420.661	389.024	-544.653	-345.675	-606.654	-305.724
XBX00012 C3_1	-452.473	-492.096	-491.031	-226.188	-500.881	-432.534
XBX00012 C3_2	-422.322	-190.47	-595.963	-249.168	-468.036	-385.192

Table 1.2.12: Docking score of further generation of the molecules. (* In Compounds Code 'Cm_n', 'm' represents the generation number.)

1.2.2.1.3.1: The second generation of BTB02557

In **Table 1.2.12**, the average score has been given in order to show a better comparison. It can also be noticed that most of these generations belonged to the BTB02557 series. This is because this compound has more substituents to be modified. Also, it can be seen that compounds of BTB02557 were docked into those five homology models twice in order to minimise any error caused by the docking programme. In the second generations of BTB02557 (from C2_1 to C2_5), the modifications involved change to the size of the ring, length of the alkyl chain between carbonyl group and morpholine substituent and thio- group changed to carbonyl group. The **Table 1.2.13** shows all the structures of the BTB02557 compounds from C2_1 to C2_5.

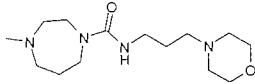
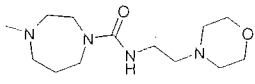
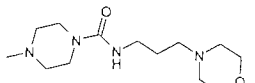
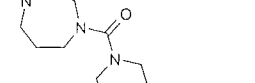
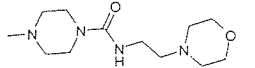
Compounds Code	Structure	Compounds Code	Structure
BTB02557 C2_1		BTB02557 C2_4	
BTB02557 C2_2		BTB02557 C2_5	
BTB02557 C2_3			

Table 1.2.13: Structure of the second generation of BTB02557.

Comparing their docking scores, only BTB02557 C2_1 has a better docking score than the parent compound. By looking at the docked structure, it was found that all the docked positions were similar in orientation with seven-member ring side pointing inside the active site. This suggests that BTB02557 C2_1 should have a better binding affinity than its parent compound. BTB02557 C2_2 and BTB02557 C2_3 had the seven-member ring replaced by the six-member ring. The docking scores were reduced. Docking position analysis shows that when the ring size was reduced, two of the docking position of BTB02557 C2_2 in **T5** and BTB02557 C2_4 in **T4** were changed with the morpholine substituent pointing inside the active site. This result may mean that the seven-member rings should have a better interaction with the active site. In the last two analogues, seven-member rings were present. However, the alkyl chain length was reduced. This modification also reduced the docking score, especially for the compound BTB02557 C2_5. The docking analysis showed that some of them were on the surface of the active site. Comparing this to the previous docking position shown in **Table 1.2.15**, hypotheses could be made here that the middle part of the active site pocket should be a long narrow tunnel. If the alkyl chain length was too short, the drug molecules could not have a proper interaction with the whole active site. Overall, in this generation of the compounds, BTB02557 C2_1 could be a potential inhibitor for the protein. The result also suggested that the ring which is close to the carbonyl group should be seven-membered and the alkyl chain length should be three.

1.2.2.1.3.2: The third and fourth generations of BTB02557

The modification of the third and fourth generation of the BTB02557 compounds (from C4_1 to C4_8) mainly focused on the morpholine substituent. All the structures of these two generations are shown in **Table 1.2.14**.

Third generation		Fourth generation	
Code	Structure	Code	Structure
BTB02557 C3_1		BTB02557 C4_1	
BTB02557 C3_2		BTB02557 C4_2	
BTB02557 C3_3		BTB02557 C4_3	
BTB02557 C3_4		BTB02557 C4_4	
BTB02557 C3_5		BTB02557 C4_5	
BTB02557 C3_6		BTB02557 C4_6	
BTB02557 C3_7		BTB02557 C4_7	
BTB02557 C3_8		BTB02557 C4_8	
BTB02557 C3_9			
BTB02557 C3_10			
BTB02557 C3_11			

Table 1.2.14: The structure of third and fourth generations of BTB02557

In the third generation of the compounds, the main modification was on the *tert* amine. Among those docking scores, BTB02557 C3_3 showed a higher docking score than its parent compound although it had a slightly lower docking score in **T2**. The **Figure 1.2.18** shows the partial docked structure.

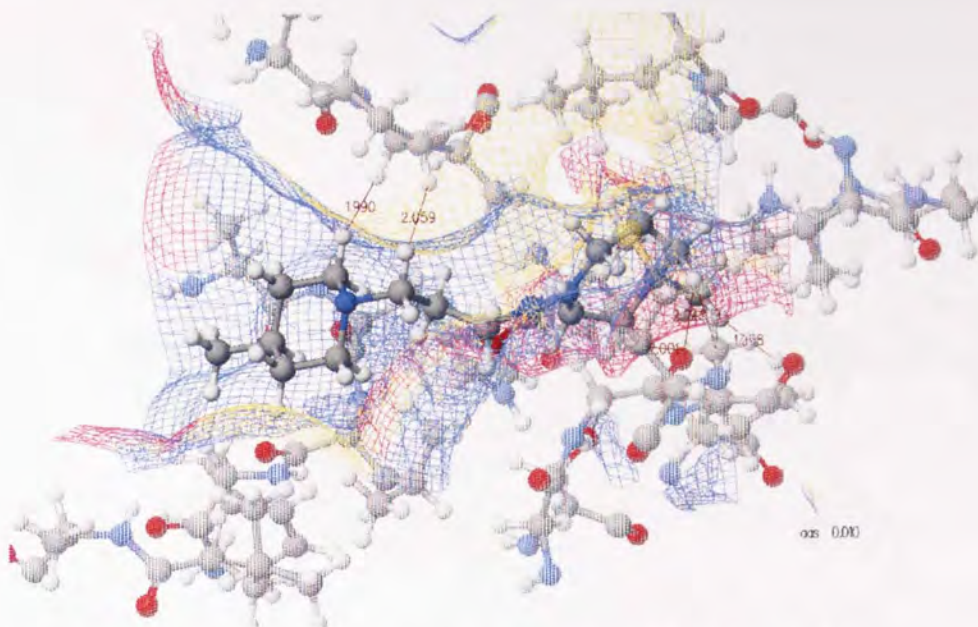
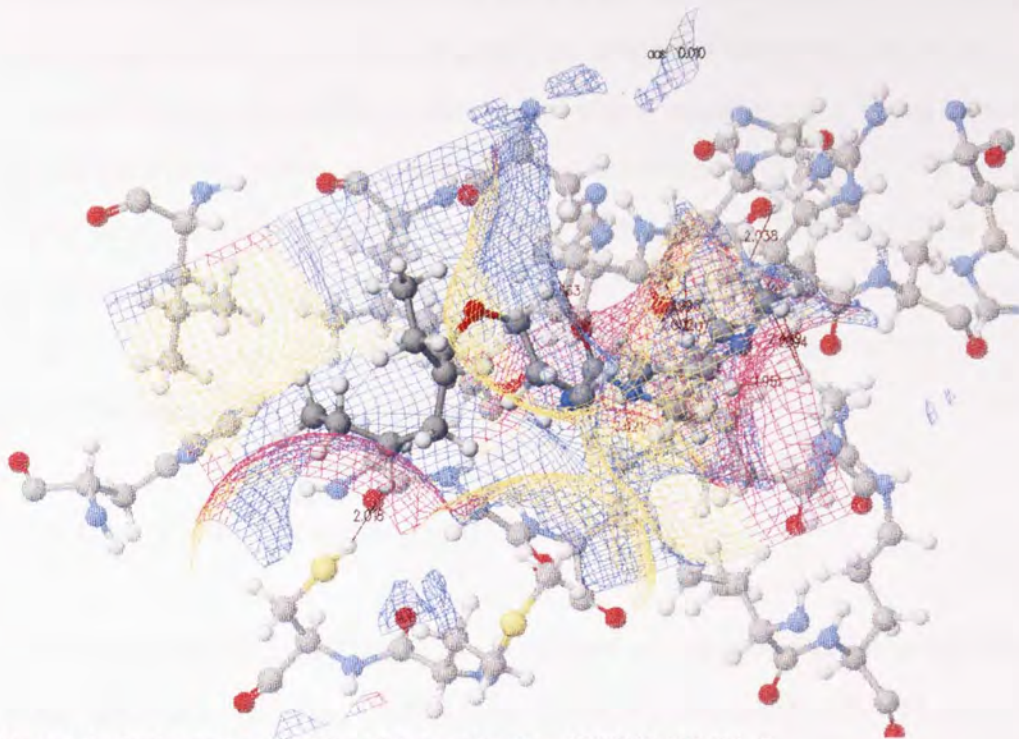


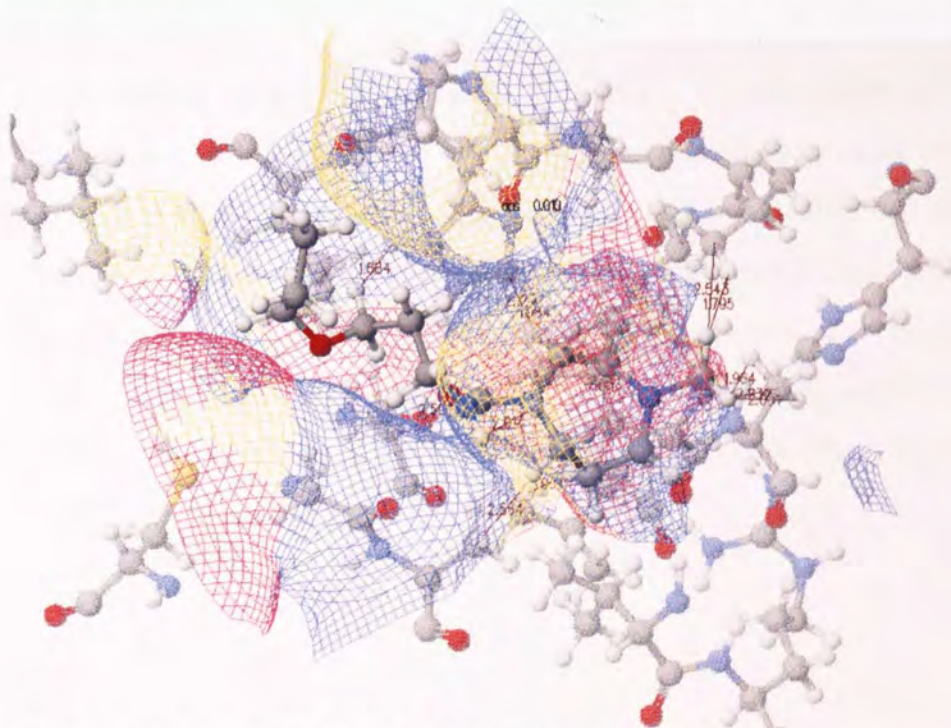
Figure 1.2.18: Partial docked structure of BTB02557 C3_3 into T2

Comparison with the first graph in **Figure 1.2.18** as well as the docked structure of docked structure in T1, T3, T4 and T5 showed that their orientation were all similar. Although the labelled bumps show that there was one bump presented on the hydrogen in the six-member ring, this molecule could be accepted as a better potential inhibitor for the homology models. Apart from this molecule, others in the third generation all showed a similar overall docking score to the parent compounds except the compound of BTB02557 C2_11. The docked structure in T4 was checked as this one had the lowest score of -256.706 kcal/mole. Its situation is similar to that shown in **Figure 1.2.18**, as all of their orientations were similar. Therefore, this docking score could be considered as calculation error and the compounds could still be considered as a potential drug molecule.

In the fourth generation, the nitrogen was replaced by oxygen. The result in **Table 1.2.12** showed that nearly all of them have a higher score than the parent compound except BTB02557 C4_6 and C4_7. **Figure 1.2.19** shows a comparison



Partial docked structure in **T2** and the molecule of BTB02557 C4_7



Partial docked structure in **T2** and the molecule of BTB02557 C4_5

Figure 1.2.19: Comparison of two docked structures in the fourth generation.

The first graphs in **Figure 1.2.19** was from the molecule whose docking score was lower than the parent compound while the second one came from one with a higher score. It can be seen that in the first graph, the long alkyl chain was very close to the calculated surface of the active site pocket which might result a strong lipophilic interaction with the nearby residues. However, compound of BTB02557 C4_5 lacked a long alkyl chain which resulted in a higher docking score. This comparison should be able to explain why BTB02557 C4_6 and C4_7 had a lower score. This generation of the molecules suggested that a large tailing substituent attached to the oxygen is needed in order to get a better score.

1.2.2.1.3.3: The fifth generation of BTB02557

The fifth generation of BTB02557 concentrated on the modification of the methyl group attached to the seven-member ring. There are three series of molecules. C5_1 series which were from BTB02557 C5_1 to C5_105 had the methyl group replaced by an allyl group. The C5_3 series from BTB02557 C5_3 to C5_305 had the methyl group replaced by a benzyl group. In the series C5_2 from BTB02557 C5_2 to C5_205, the methyl group was not changed and those compounds acted as the reference to the other two groups. The reason for this modification will be shown in **Chapter 2**. As shown previously in **Figure 1.2.17**, if the methyl group were changed, the docking position may alter. After the docked structures were checked, nearly all the molecules from the C5_1 and C5_3 series exhibited docking positions where the seven-member ring pointed outside of the active site pocket. The following **Figure 1.2.20** gives an example from C5_1 series.

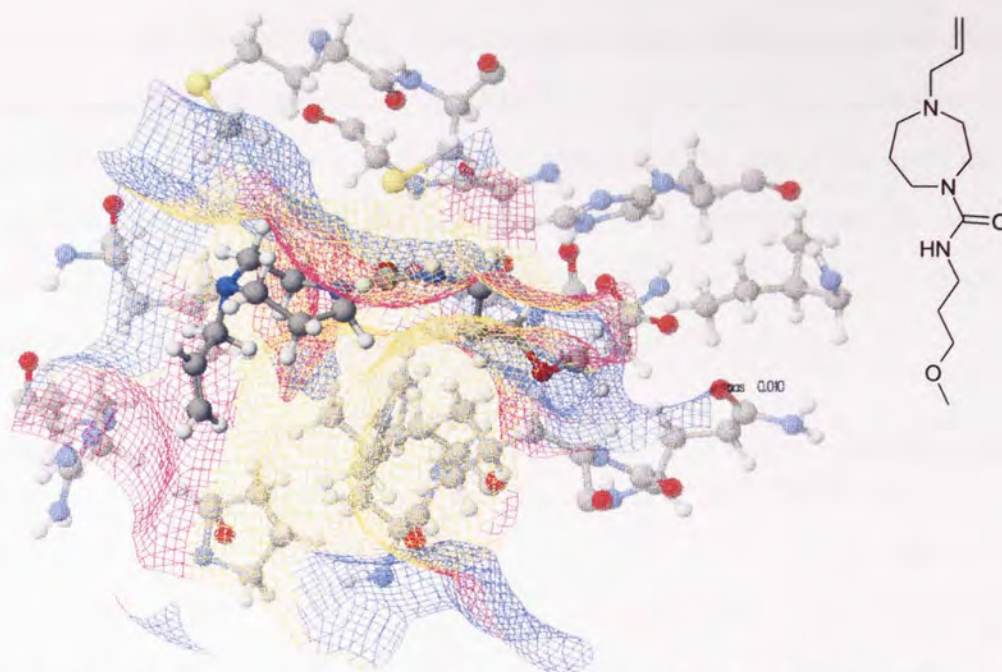


Figure 1.2.20: Partial docked structure in T4 and the molecule of BTB02557 C5_101 (-410.091 kcal/mole)

It can be seen from **Figure 1.2.20** that the methoxy group was inside the active site pocket. Analyzing the structures and the overall docking scores shows a small trend. The structure of the drug molecules and the overall docking score are shown in **Table 1.2.15**.

Compounds code	Structure of the drug molecules	Average docking score (kcal /mole)
BTB02557 C5_101		-462.073
BTB02557 C5_102		-479.602
BTB02557 C5_103		-495.645
BTB02557 C5_104		-503.811
BTB02557 C5_105		-505.54

Table 1.2.15: Structure and the average docking score of BTB02557 C5_1 series.

It was very clear that the docking score decreased as the substituent size attached to oxygen increased. The phenomena could be explained by the observation that these kind of molecules could have better binding affinities if the size of the alkyl tailing group could fit the space of the seven-member ring assuming that the docking position did not shift. However, the situation changed in the BTB02557 C5_3 series. Their structures and the average docking are shown in **Table 1.2.16**.

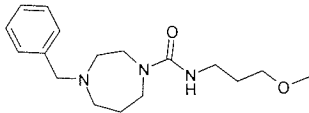
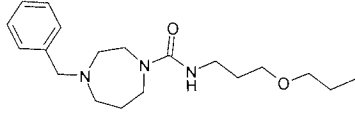
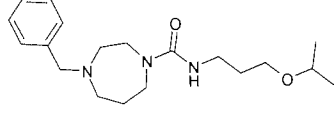
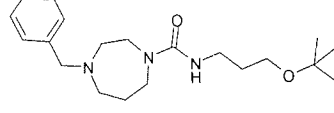
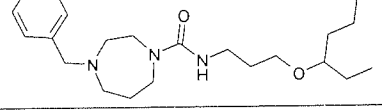
Compounds code	Structure of the drug molecules	Average docking score (kcal /mole)
BTB02557 C5_301		-470.153
BTB02557 C5_302		-508.773
BTB02557 C5_303		-532.957
BTB02557 C5_304		-503.714
BTB02557 C5_305		-496.203

Table 1.2.16: Structure and the average docking score of BTB02557 C5_3 series.

In **Table 1.2.16**, it shows that BTB02557 C5_303 had the lowest docking score. Therefore, their docked structures were analysed in order to find out the reason. The **Figure 1.2.21** and **Figure 1.2.22** shows a comparison of docked structures.

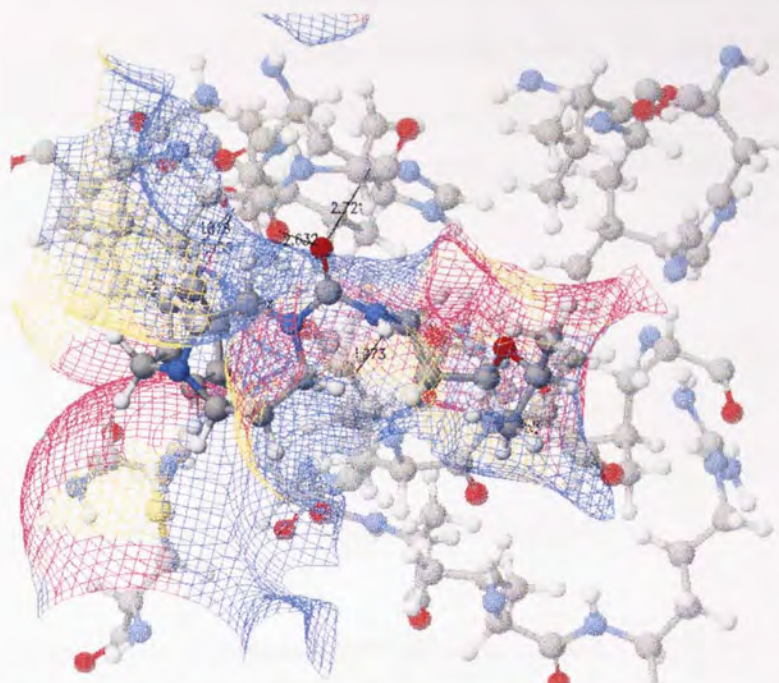


Figure 1.2.21: Partial docked structure of BTB02557 C5_303 in T1

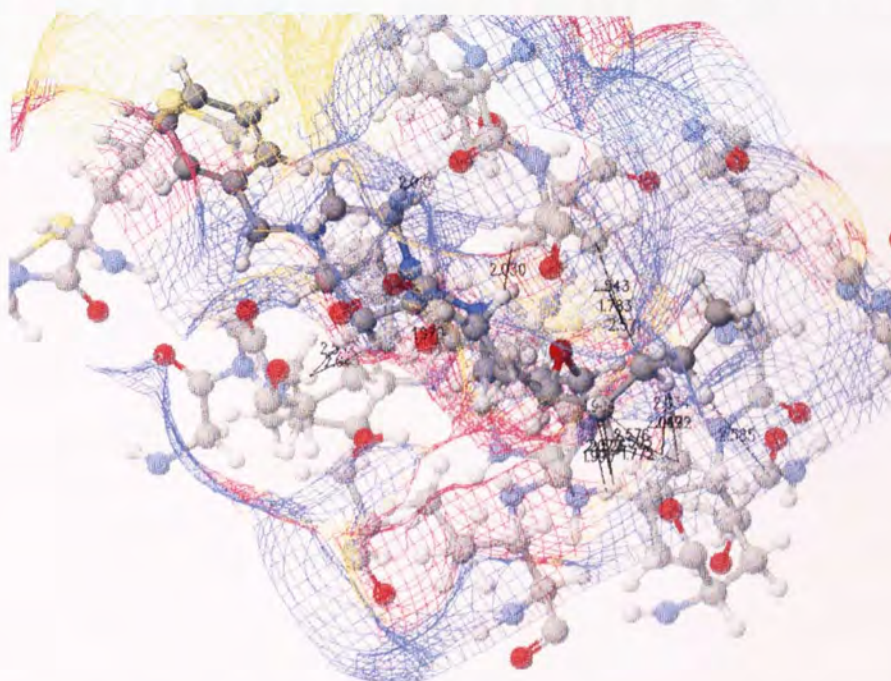


Figure 1.2.22: Partial docked structure of BTB02557 C5_305 in T2

The bumps were labelled in these two figures and **Figure 1.2.22** had more bumps which appear on the alkyl tailing group than **Figure 1.2.21** had. Based on the comparison of these two docked structures, a hypothesis could be made that although

the benzyl groups were pointing outside of the active site pocket, they could still affect the binding affinities inside the pocket. This might arise with the presence of benzyl group allowing the whole molecule shifted inside the active site pocket a little bit. Therefore, the space inside the pocket became smaller and could not fit the larger group attached to the oxygen.

Overall, in the fifth generation of BTB02557, a certain trend could be followed for the drug design. However, as all of them had a similar docking score as their parent compounds, these drug molecules could be considered as potential inhibitors for the protein and could be synthesised and screened.

1.2.2.1.3.4: The second and third generation of XBX00012

For those molecules derived from XBX00012, the next generation of the drug design did not improve the docking score. The structures of all these compounds are shown in **Table 1.2.17**.

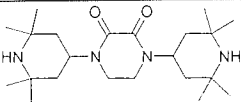
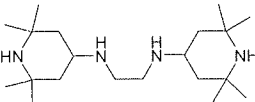
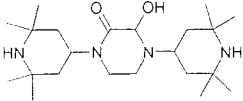
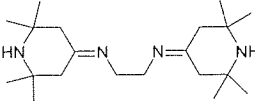
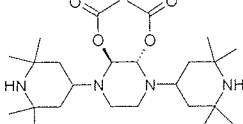
Compound code	Structure	Compound code	Structure
XBX00012 C2_1		XBX00012 C3_1	
XBX00012 C2_2		XBX00012 C3_2	
XBX00012 C2_3			

Table 1.2.17: The next generation of XBX00012 compounds.

Comparing these structures in **Table 1.2.17** and the docking with their parent compound and XBX00012 C1_2, only XBX00012 C2_2 has a closer docking score to the parent compound. At the same time, XBX00012 C3_1 shows a slightly better

docking score than others. These may mean that the presence of a hydrogen bond donor in the middle part is important for the binding affinity. However, only XBX00012 C1_2 was considered for synthesis and screening.

After all these dockings, several molecules were selected for chemical synthesis and biological screening. Their molecule weight, LogP, hydrogen bond donor and overall nitrogen and oxygen acting as hydrogen bond acceptors were also checked. These relate to the rule of five (RO5) which was proposed by Lipinski *et. al.* (1997). The molecular weight and the LogP value (the parameter set came from Ghose *et. al.* 1988) of some chosen molecules were calculated in CaChe ProjectLeader (Version 6.1.10, Fujitsu Limited) and the result are shown in **Table 1.2.18**.

Compound code	LogP	MW	Compound code	LogP	MW
BTB02557 C1_5	0.373	242	BTB02557 C5_2	-1.234	282
BTB02557 C1_6	-0.68	296	BTB02557 C5_201	-0.824	229
BTB02557 C2_1	-1.022	284	BTB02557 C5_202	-0.013	257
BTB02557 C3_1	0.042	282	BTB02557 C5_203	-0.068	257
BTB02557 C3_2	-0.878	297	BTB02557 C5_204	0.01	271
BTB02557 C3_3	0.373	296	BTB02557 C5_205	1.029	312
BTB02557 C5_1	-0.493	308	BTB02557 C5_3	0.543	358
BTB02557 C5_101	-0.083	255	BTB02557 C5_301	0.953	305
BTB02557 C5_102	0.728	283	BTB02557 C5_302	1.764	333
BTB02557 C5_103	0.672	283	BTB02557 C5_303	1.708	333
BTB02557 C5_104	0.751	297	BTB02557 C5_304	1.787	348
BTB02557 C5_105	2.402	340	BTB02557 C5_305	3.438	390
XBX00012 C1_2	0.942	396			

Table 1.2.18: LogP and Molecule Weight values of some chosen compounds

It can be seen that all the LogP values were less than 5 and all the MW values were less than 500. As all these compounds share a similar structure, it was easy to see that XBX00012 C1_2 had the highest number of hydrogen bond donors (four). Therefore, the total number of hydrogen bond donors was less than five. BTB02557 C5_2 possessed the highest number (five) of hydrogen bond acceptors. Therefore, the total number of hydrogen bond acceptors was less than ten. According to Lipininski's

rules, these numbers suggest that these compounds could be considered as pharmaceutically appropriate

1.2.2.2: Docking potential ligands into the dUTPase

In total, there were 14 compounds of Type A, 28 compounds of Type B and 36 compounds of Type C docked into the defined active site. Their docking scores and the structure information are shown in **Table 1.2.21**. In order to distinguish all those similar structures, a number with a two or three letter code was used as the compounds name. The first letter, A, B, or C mean they belong to the Type 2A, Type 2B or Type 2C respectively which was shown in **Table 1.1.6**. If the next letter is 'W', it means that the starting material will be 1-methyl-4-piperidone. If the next letter is 'WO', it means that the starting material will be tropinone.

Name	Structure information				Docking score (kcal/mole)		
	R	R'	X	R''	Score1	Score2	Average
AW1	Methyl	-	-	-	-286.537	-289.703	-288.12
AW2	Ethyl	-	-	-	-247.455	-306.936	-277.196
AW2D	Vinyl	-	-	-	-313.014	-298.141	-305.578
AW3	Propyl	-	-	-	-293.152	-239.886	-266.519
AW4	Butyl	-	-	-	-264.367	-266.077	-265.222
AW5	Pentyl	-	-	-	-300.464	-269.548	-285.006
AW6	Hexyl	-	-	-	-172.897	-265.774	-219.336
AWO1	Methyl	-	-	-	-271.794	-265.543	-268.669
AWO2	Ethyl	-	-	-	-278.988	-290.258	-284.623
AWO2D	Vinyl	-	-	-	-311.493	-335.356	-323.425
AWO3	Propyl	-	-	-	-260.369	-240.569	-250.469
AWO4	Butyl	-	-	-	-278.554	-288.45	-283.502
AWO5	Pentyl	-	-	-	-343.331	-343.56	-343.446
AWO6	Hexyl	-	-	-	-354.728	-310.12	-332.424
BW1	Methyl	Hydroxymethyl	-	-	-307.998	-327.58	-317.789
BW2	Ethyl	Hydroxymethyl	-	-	-328.984	-308.587	-318.786
BW3	Vinyl	Hydroxymethyl	-	-	-304.722	-308.595	-306.659
BW4	Propyl	Hydroxymethyl	-	-	-325.454	-351.285	-338.37
BW5	Butyl	Hydroxymethyl	-	-	-309.934	-235.046	-272.49
BW6	Pentyl	Hydroxymethyl	-	-	-343.022	-299.02	-321.021
BW7	Hexyl	Hydroxymethyl	-	-	-243.396	-407.815	-325.606
BW8	Methyl	Hydroxyethyl	-	-	-351.001	-343.091	-347.046
BW9	Ethyl	Hydroxyethyl	-	-	-318.593	-287.926	-303.26
BW10	Vinyl	Hydroxyethyl	-	-	-338.643	-297.788	-318.216
BW11	Propyl	Hydroxyethyl	-	-	-283.489	-241.324	-262.407

BW12	Butyl	Hydroxyethyl	-	-	-281.847	-346.595	-314.221
BW13	Pentyl	Hydroxyethyl	-	-	-257.722	-214.229	-235.976
BW14	Hexyl	Hydroxyethyl	-	-	-314.633	-130.87	-222.752
BWO1	Methyl	Hydroxymethyl	-	-	-312.605	-299.424	-306.015
BWO2	Ethyl	Hydroxymethyl	-	-	-290.838	-282.79	-286.814
BWO3	Vinyl	Hydroxymethyl	-	-	-299.031	-317.194	-308.113
BWO4	Propyl	Hydroxymethyl	-	-	-304.13	-310.178	-307.154
BWO5	Butyl	Hydroxymethyl	-	-	-350.394	-298.552	-324.473
BWO6	Pentyl	Hydroxymethyl	-	-	-293.202	-346.06	-319.631
BWO7	Hexyl	Hydroxymethyl	-	-	-314.498	-343.374	-328.936
BWO8	Methyl	Hydroxyethyl	-	-	-330.958	-312.525	-321.742
BWO9	Ethyl	Hydroxyethyl	-	-	-311.983	-347.432	-329.708
BWO10	Vinyl	Hydroxyethyl	-	-	-336.892	-291.886	-314.389
BWO11	Propyl	Hydroxyethyl	-	-	-330.48	-356.244	-343.362
BWO12	Butyl	Hydroxyethyl	-	-	-328.649	-345.238	-336.944
BWO13	Pentyl	Hydroxyethyl	-	-	-369.359	-283.658	-326.509
BWO14	Hexyl	Hydroxyethyl	-	-	-255.558	-309.121	-282.34
CW1	-	H	N	H	-260.39	-212.5	-236.445
CW2	-	Hydroxymethyl	N	H	-298.688	-306.737	-302.713
CW3	-	Hydroxyethyl	N	H	-334.181	-312.068	-323.125
CW4	-	H	N	Methyl	-247.907	-250.12	-249.014
CW5	-	Hydroxymethyl	N	Methyl	-284.377	-318.993	-301.685
CW6	-	Hydroxyethyl	N	Methyl	-306.318	-353.055	-329.687
CW7	-	H	N	Ethyl	-312.483	-289.58	-301.032
CW8	-	Hydroxymethyl	N	Ethyl	-269.782	-334.957	-302.37
CW9	-	Hydroxyethyl	N	Ethyl	-319.404	-335.126	-327.265
CW10	-	H	O	H	-274.543	-297.54	-286.042
CW11	-	Hydroxymethyl	O	H	-298.152	-287.094	-292.623
CW12	-	Hydroxyethyl	O	H	-314.569	-323.423	-318.996
CW13	-	H	O	Methyl	-298.175	-268.624	-283.4
CW14	-	Hydroxymethyl	O	Methyl	-297.364	-317.849	-307.607
CW15	-	Hydroxyethyl	O	Methyl	-322.722	-297.129	-309.926
CW16	-	H	O	Ethyl	-312.34	-301.23	-306.785
CW17	-	Hydroxymethyl	O	Ethyl	-297.917	-329.053	-313.485
CW18	-	Hydroxyethyl	O	Ethyl	-268.788	-314.569	-291.679
CWO1	-	H	N	H	-261.458	-277.81	-269.634
CWO2	-	Hydroxymethyl	N	H	-302.493	-301.848	-302.171
CWO3	-	Hydroxyethyl	N	H	-307.96	-314.246	-311.103
CWO4	-	H	N	Methyl	-246.532	-234.32	-240.426
CWO5	-	Hydroxymethyl	N	Methyl	-306.877	-289.573	-298.225
CWO6	-	Hydroxyethyl	N	Methyl	-324.312	-339.868	-332.09
CWO7	-	H	N	Ethyl	-261.839	-264.358	-263.099
CWO8	-	Hydroxymethyl	N	Ethyl	-290.525	-343.192	-316.859
CWO9	-	Hydroxyethyl	N	Ethyl	-330.998	-343.026	-337.012
CWO10	-	H	O	H	-237.044	-229.145	-233.095
CWO11	-	Hydroxymethyl	O	H	-275.031	-287.463	-281.247
CWO12	-	Hydroxyethyl	O	H	-316.3	-310.844	-313.572
CWO13	-	H	O	Methyl	-271.764	-250.458	-261.111
CWO14	-	Hydroxymethyl	O	Methyl	-316.31	-308.789	-312.55
CWO15	-	Hydroxyethyl	O	Methyl	-313.483	-324.633	-319.058

CWO16	-	H	O	Ethyl	-270.72	-287.569	-279.145
CWO17	-	Hydroxymethyl	O	Ethyl	-315.093	-284.666	-299.88
CWO18	-	Hydroxyethyl	O	Ethyl	-333.718	-315.615	-324.667

Table 1.2.19: Structure and the docking score of selected analogues of **T 2-3-9**.

As shown in **Table 1.2.19**, each compound selected was docked into the protein structure twice and the total score of them was given. Firstly, the compounds with an average score below -325 kcal/mole were investigated. Fourteen compounds fell into this region. Their information is shown in **Table 1.2.20**.

Name	Structure information				Docking score (kcal/mole)		
	R	R'	X	R''	Score1	Score2	Average
BW8	Methyl	Hydroxyethyl	-	-	-351.001	-343.091	-347.046
AWO5	Pentyl	-	-	-	-343.331	-343.56	-343.446
BWO11	Propyl	Hydroxyethyl	-	-	-330.48	-356.244	-343.362
BW4	Propyl	Hydroxymethyl-	-	-	-325.454	-351.285	-338.37
CWO9	-	Hydroxyethyl	N	Ethyl	-330.998	-343.026	-337.012
BWO12	Butyl	Hydroxyethyl	-	-	-328.649	-345.238	-336.944
AWO6	Hexyl	-	-	-	-354.728	-310.12	-332.424
CWO6	-	Hydroxyethyl	N	Methyl	-324.312	-339.868	-332.09
BWO9	Ethyl	Hydroxyethyl	-	-	-311.983	-347.432	-329.708
CW6	-	Hydroxyethyl	N	Methyl	-306.318	-353.055	-329.687
BWO7	Hexyl	Hydroxymethyl-	-	-	-314.498	-343.374	-328.936
CW9	-	Hydroxyethyl	N	Ethyl	-319.404	-335.126	-327.265
BWO13	Pentyl	Hydroxyethyl	-	-	-369.359	-283.658	-326.509
BW7	Hexyl	Hydroxymethyl-	-	-	-243.396	-407.815	-325.606

Table 1.2.20. Top ranked molecules with their structure and docking score information.

The docked structure of these compounds was investigated virtually. Surprisingly, most of the molecules appeared outside of the active side pocket. This should be the reason that the docking score of certain compounds gave a huge difference such as AW2, AW6, BW7, BW14, *etc.* Those docked structures all showed different docking

positions. Only two of the compounds in **Table 1.2.20** were sitting inside the active sit pocket. A comparison is given in **Figure 1.2.23** and **Figure 1.2.24**.

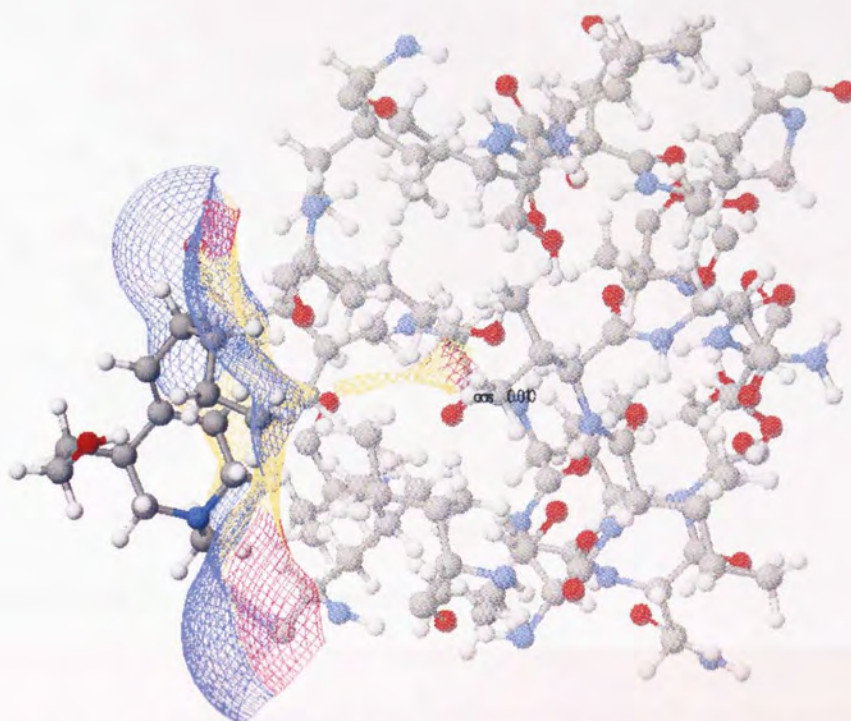


Figure 1.2.23: Partial docked structure of BWO12

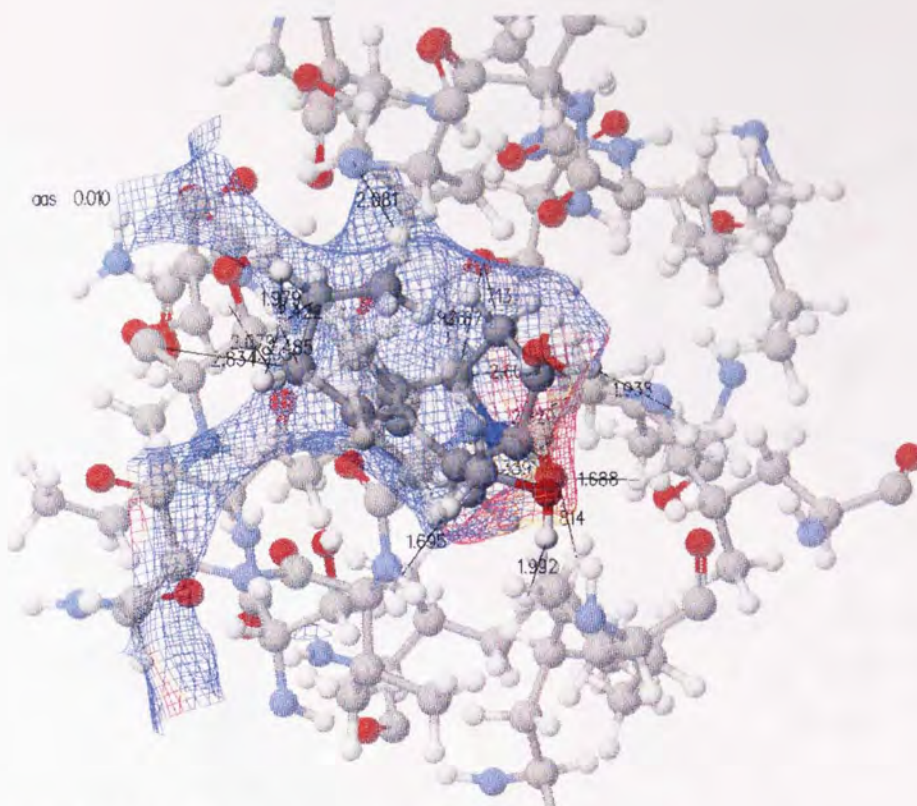


Figure 1.2.24: Partial docked structure of BW4.

It was very clear that compound BWO12 was sitting outside of the active site pocket. The lower docking score is due to the interaction between the drug molecule and the residues which lie at the entrance of the active site pocket. In **Figure 1.2.24**, compound BW4 sits inside the active site completely. Although analysis showed that some bumps were present, those bumps were caused by the hydrogen atoms which pointed outside the calculated active site pocket. Except for this, no part of the carbon, oxygen and nitrogen was outside of the surface of the pocket. This kind of the situation was acceptable. It was also noticeable that propyl group was folded up in order to fit in the small space of the pocket. In this case, it might mean that the molecules with a larger group may not be able to fit in the active site. However, there was a very interesting docking position of compound BW7. Its docked structure is shown in **Figure 1.2.25**.

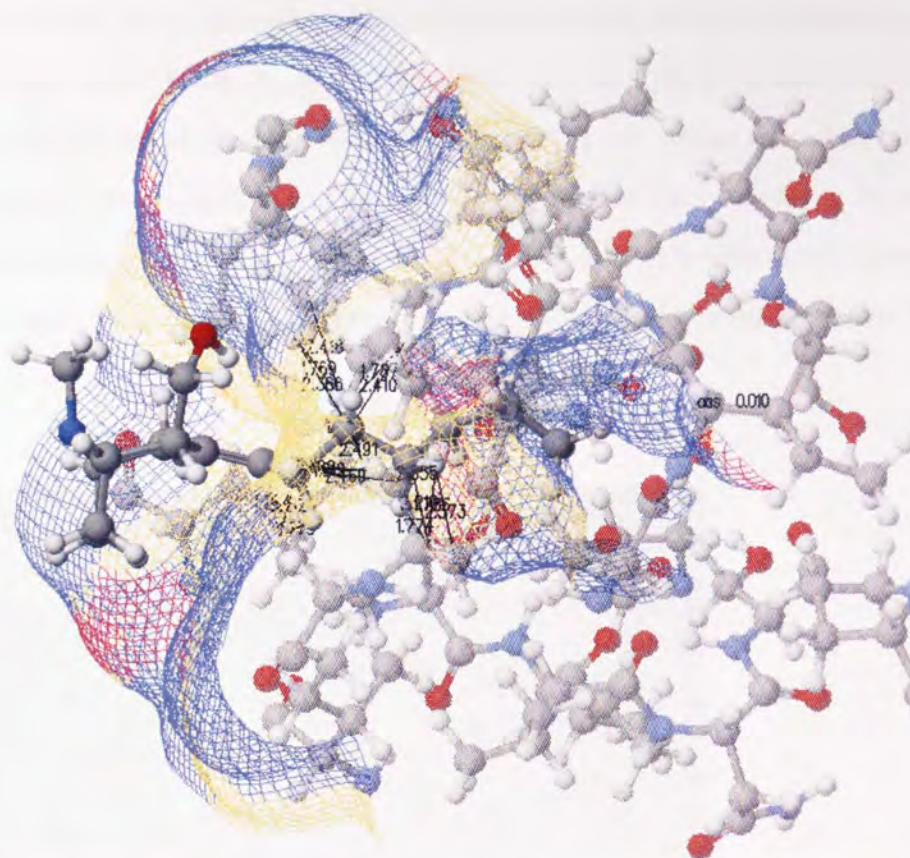


Figure 1.2.25: Partial docked structure of BW7

It can be seen from **Figure 1.2.25** that the long alkyl chain of BW7 was pointing inside the active site. The rest molecule sat at the entrance where the hydroxymethyl group might have a good interaction with the nearby residues. The other one of the docked structures was checked and it was found that the whole molecule was inside the active pocket. However, several carbon and oxygen atoms were pointing outside of the calculated active site surface and caused a lot of bumps. This should be the reason that this compound had a very different docking score. Overall, these two compounds BW4 and BW7 could be synthesized and tested.

After the top ranked structures were analysed, the molecule with a smaller group such as CW3, BWO8, CWO15, BW2, CW17, CW15, *etc.* was checked as the previous result suggested that the selected active site pocket was small and could not

accommodate large molecules. After checking several docked structures, a general trend was found. If the ligand was lipophilic such as Type A compounds, the drug molecules tended to sit at the entrance of the active site pocket even if the molecule was small. If the compound was hydrophilic, it was more likely to stay in the centre of the active site pocket and most of the Type C compounds with small groups were found inside the pocket. The **Figure 1.2.26** indicates the possible reason for this phenomenon.

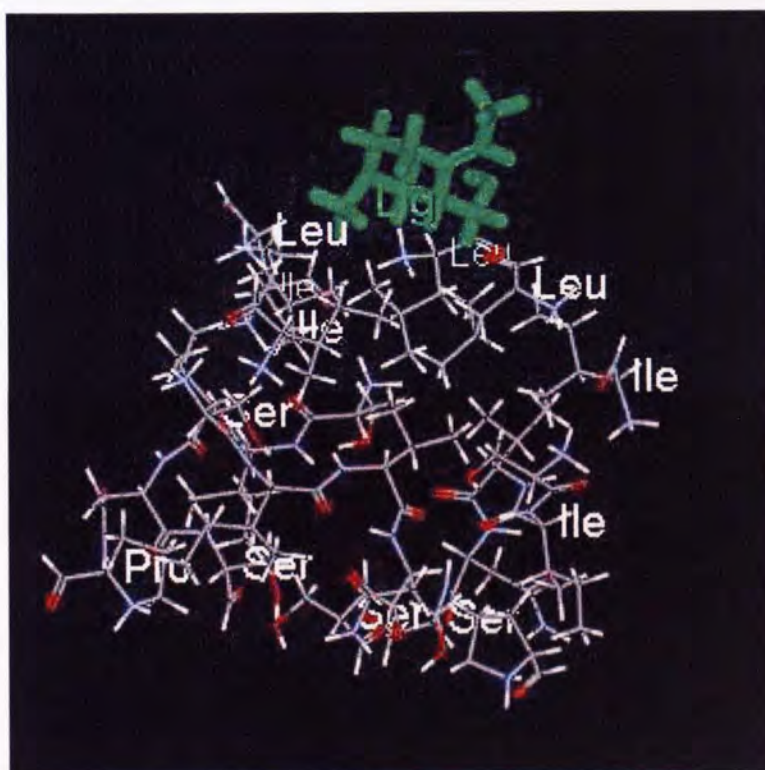


Figure 1.2.26: Partial docked structure of BWO8 (created in DS Visualizer, V. 2.01)

In the **Figure 1.2.26**, the green part is the ligand BWO8. Although it had small substituents which could fit in the active site pocket, it still remained at the entrance. The residues near to the ligand are labeled with their three-letter residue names. There were three isoleucines and three leucines. All these residues have lipophilic side chains. At the same time, the residues inside the pocket were also named. It can be seen that there were four serines and a proline are present all of which are hydrophilic.

It was assumed that lipophilic and hydrophilic interactions should be the main determinants for docking. This should explain why the most of the Type A and Type B compounds sit at the entrance and the Type C compounds were inside the active site.

The docked structures with a lower average docking score were also checked. **Table 1.2.21** showed some of the molecules that had average docking scores greater than – 250 kcal/mole.

Name	Structure information				Docking score (kcal/mole)		
	R	R'	X	R''	Score1	Score2	Average
AW6	Hexyl	-	-	-	-172.897	-265.774	-219.336
BW14	Hexyl	Hydroxyethyl	-	-	-314.633	-130.87	-222.752
CWO10	-	H	O	H	-237.044	-229.145	-233.095
BW13	Pentyl	Hydroxyethyl	-	-	-257.722	-214.229	-235.976
CW1	-	H	N	H	-260.39	-212.5	-236.445
CWO4	-	H	N	Methyl	-246.532	-234.32	-240.426
CW4	-	H	N	Methyl	-247.907	-250.12	-249.014

Table 1.2.21: The molecules that had unfavourable docking scores.

By checking the structures of those docked compounds, it was found that those molecules were either too big or too small to fit in the active site. **Figure 1.2.27** gave an example of docked structure of compound CW1.

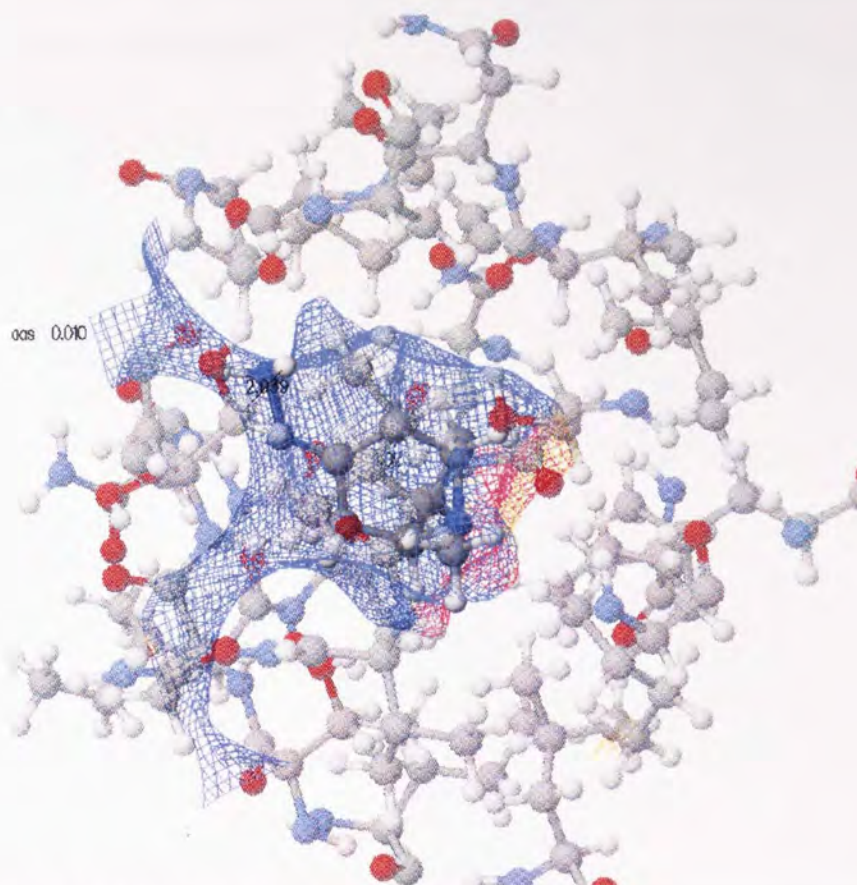


Figure 1.2.27: Partial docked structure of CW1

From this figure, it can be seen that although the molecule sits in the middle of the pocket, it was too small to make close contact with the nearby residues and there was still some free space in the pocket. There were even no bumps found around the ligand. Other compounds of CW4, CWO4 and CWO10 showed similar dockings. Therefore, the compounds in Type C will not be synthesized if the side chain is too small.

Based on the assumption made previously, no further docking should be required. The compounds selected for the synthesis were the Type A compounds. The reason for this was to see if this kind of molecule could block the entrance of the active site. Then, the short chain in Type B compounds together with BW7 were put forward as be synthetic target for testing as these kinds of compounds have both lipophilic and hydrophilic substituents. And finally, the Type C compounds with a long chain were

put forward for synthesis because this kind of compound were most likely to be inhibitors of the protein.

1.3 Conclusion

M-TB associated with HIV has led to many deaths and the number of patients is still increasing. At the same time, M-TB has developed to MDR-TB and XDR-TB. This situation was mainly caused by patients failing to take the correct dose during the long treatment period. Although there are several first-line drugs available, no new drugs have entered the clinic to treat the M-TB effectively since the 1970s. Therefore, new drugs and new chemotherapeutic methods are needed.

Rv3487c carboxyl esterase was found present in M-TB. However, only the amino acid sequence was known. As all the carboxyl esterase family shares a similar 3-D structure, its structure could be predicted by homology modelling. Zhang *et al.* (2005) identified two templates, 1jji and 1qz3A, which could be used in homology modelling as they all showed over 25% of the residues to be identical. Therefore, homology modelling using these two proteins to predict the 3-D structure of carboxyl esterase Rv3487c was conducted in this study, and molecular mechanics, molecular dynamics and quantum mechanics calculation were used to optimize the aligned homology models. Overall, five homology models were created in which three of them came from 1jji and two of them came from 1qz3A. The Ramachandran plot of those homology models was checked and only one active site residue fell into the disallowed regions. These five homology models were slight different because they all came from different protein structures or went through different homology modelling pathways. All of those five homology models were used in the docking and virtual screening.

The reduced Maybridge hitset was docked into those five homology models. Several generations of docking were conducted. The best compounds were chosen. Virtual screening techniques were used to analyze the docked structures and then, these

compounds acted as lead structures to design the next generation of compounds. Finally, three types of potential drug structures were identified for synthesis. These designed compounds were tested by the rule of five (Lipinski *et. al.* 1997) and predicted that they all should have good orally availability.

In addition to these homology models, an optimized protein structure of *Mycobacterium tuberculosis* dUTPase modified by a previous project student was also used and three types of compounds which derived from this student's lead structure were docked in to the optimized dUTPase and analysed. Two kinds of docking positions were found. This phenomenon may due to the lipophilic and hydrophilic difference between the entrance and the interior of the active site. Several compounds from these three types were chosen as synthetic targets.

1.4 Recommendation

- In the homology modelling step, different software may be used to build the homology models. This was because one of the residues belonging to the active site fell inside the disallowed region in Ramachandran plot in this study. Designing a different homology pathway may solve this problem.
- In the docking calculations, different docking programmes using different docking and scoring protocols such as AMBER 9.0 or AutoDock could be used in order to see how they dock the ligand into the active site. A better docking position may be found.

1.5 Experimental Parameters

All of the molecular modelling work was conducted in CaChe WorkSpace (Version 6.1.10, Fujitsu Limited).

Normal MM/MD/MM geometry optimisation steps:

1. MM2 or MM3 force field using steepest descent with maximum iterations of 90000 and convergence value of 0.001 kcal/mole first and then, MM2 or MM3 force field using conjugate gradient with maximum iterations of 90000 and convergence value of 0.001 kcal/mole.
2. Molecular dynamics with MM3 force field with equilibration period 1ps. The temperature and the simulation duration were changeable.
3. MM2 or MM3 force field using steepest descent with maximum iterations of 90000 and convergence value of 0.001 kcal/mole first and then, MM2 or MM3 force field using conjugate gradient with maximum iterations of 90000 and convergence value of 0.001 kcal/mole.

QM parameter key word: 'PM3 MMOK NOXYZ MOZYME NODIIS GRAPH LBFGS MOZPRP=2000 GEO-OK T=50D SAFE SHIFT + XYZ CUTOF2=6 GNORM=100 RELSCF=10 CHARGE=-5'. The 'GNORM=100 and CHARGE=-5' is variable depend on the condition

Docking parameter: Flexible ligand, Flexible side chains of active site, PMF scoring, Use grids and Amber van der Waals, Grid spacing=0.375 Å, population size=50, Crossover rate=0.8, Eltism=7, Maximum generation=5000, Mutation rate=0.2, convergence=1 kcal/mole, In the local search, Maximum iterations=300, rate=0.06.

Chapter 2

Synthesis of the designed compounds

2.1 Introduction

2.1.1: Designed Compounds as Potential Inhibitors of the Mycobacterial Carboxyl Esterase Rv3487c

In the first chapter, several types of inhibition compounds were designed for chemical synthesis. Their general structures are shown in **Table 2.1.1**.

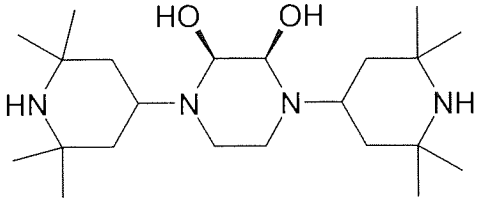
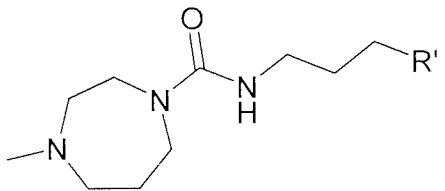
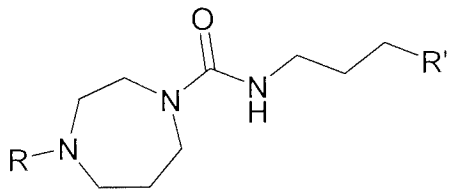
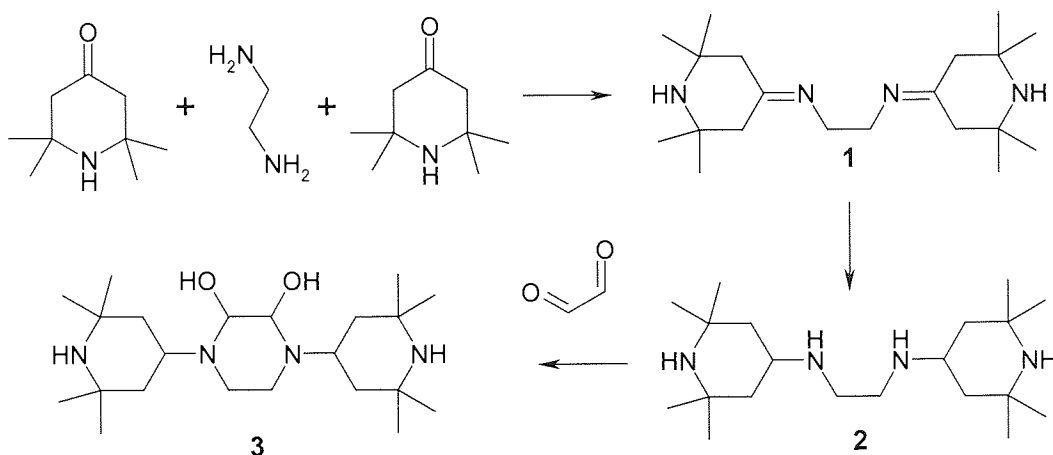
Compounds NO.	General structure
Type 1A	
Type 1B	
Type 1C	

Table 2.1.1: General structures of designed compounds as potential inhibitors of the mycobacterial carboxyl esterase Rv3487c

2.1.1.1. Intended synthetic route of Type 1A compounds

All these three types of the molecules were designed using different synthetic approaches. The intended synthetic route for **Type 1A** compounds is shown in **Scheme 2.1.1**

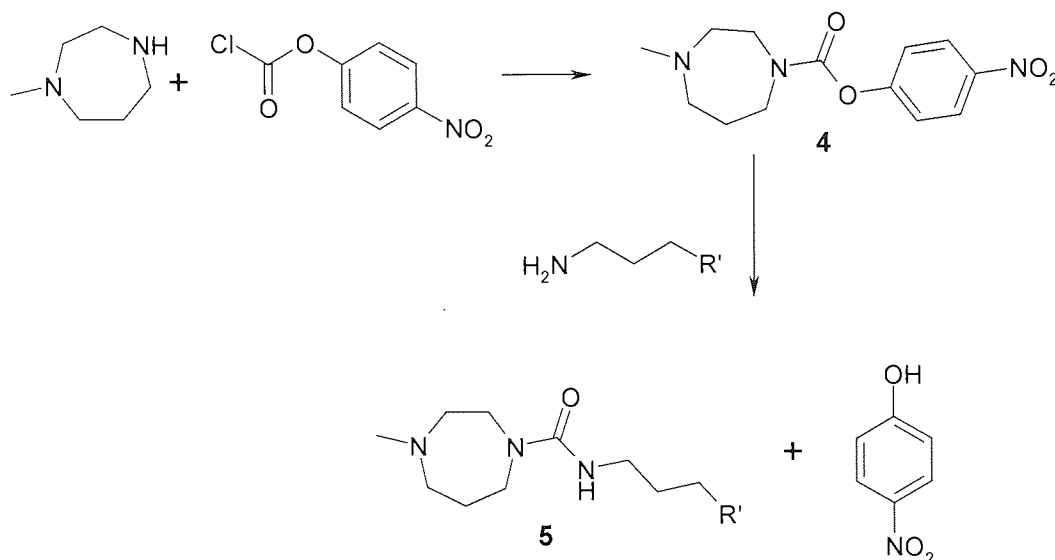


Scheme 2.1.1: Intended synthetic route for **Type 1A** compound

This is a three-step reaction. Firstly, two equivalents of 2,2,6,6-Tetramethyl-4-piperidone reacts with one portion of ethylenediamine to form N, N'-bis-(2,2,6,6-tetramethyl-piperidin-4-ylidene)-ethane-1,2-diamine (Morales *et. al.* 1984). In the second step, the two double bonds should be reduced by hydrogenation (Rylander & Nels 1979). Finally, oxalaldehyde will be added resulting in the target compound with two hydroxy groups on the same side of the ring (Willer *et. al.* 1985).

2.1.1.2. Intended synthetic route of Type 1B compounds

The intended synthetic route for **Type 1B** compounds is shown in **Scheme 2.1.2**.



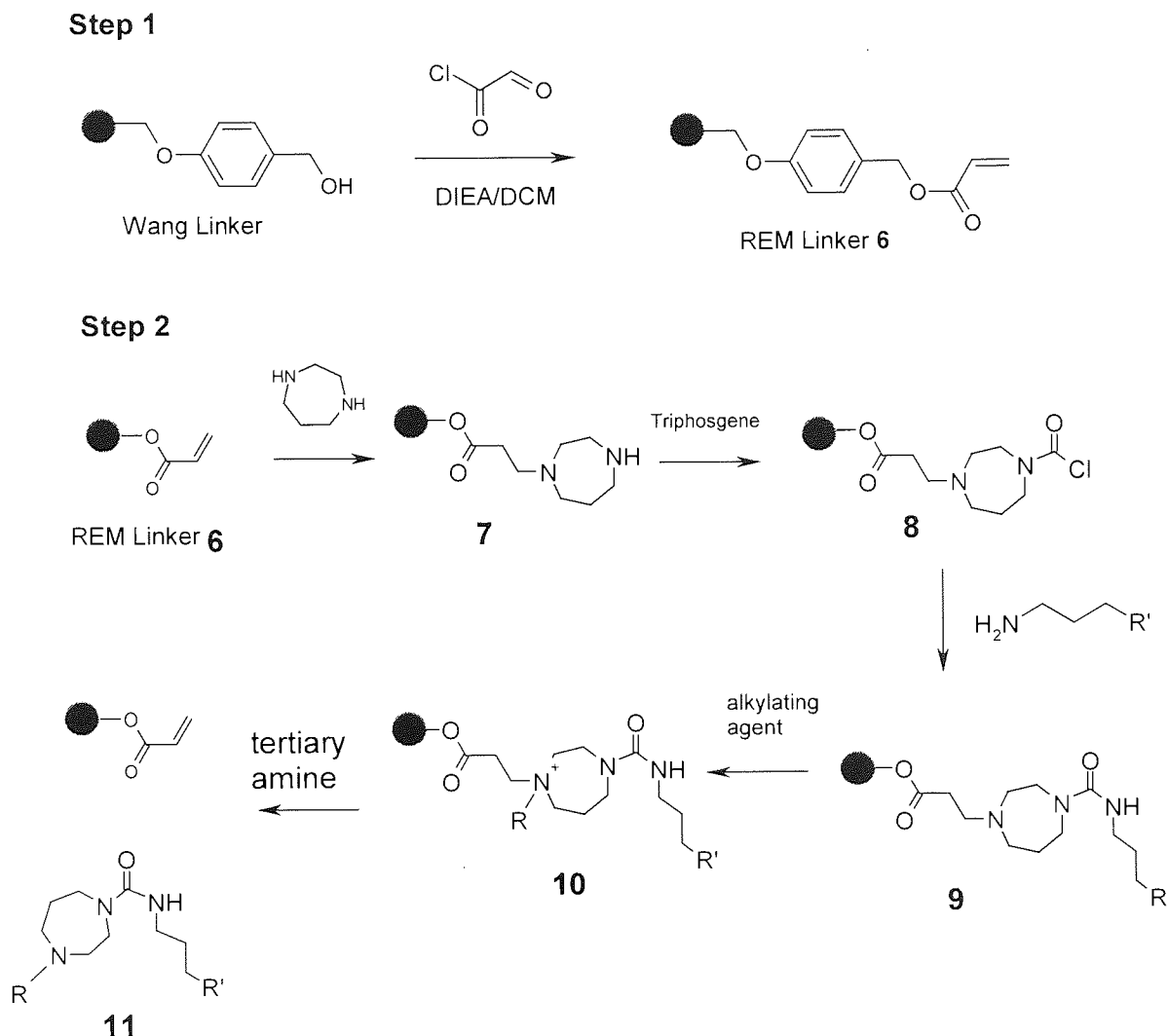
Scheme 2.1.2: Intended synthetic route for **Type 1B** compounds.

The **Scheme 2.1.2** shows a two-step reaction. The first step is the reaction of 1-methylpiperazine with 4-nitrobenzoyl chloride (Lewandowski *et. al.* 1998). In the second step, 4-nitrophenol is displaced by the chosen amine and gives the target compound **5**. (Richard 1999).

2.1.1.3. Intended synthetic route of Type 1C compounds

The synthetic route for **Type 1C** compounds was designed using solid-state chemistry and the REM linker (Morphy *et. al.* 1996). There are several advantages in using solid-state synthesis. Firstly, solid-state chemistry can use multi-step (could be more than five steps) for synthesis with a high yield which solution phase chemistry could not do. Secondly, as all the required substituents are attached on the resin, excess amount of reagents can be used and all the by products and un-reacted substances can

be eliminated by simple filtration. However, the disadvantage of this method is the reaction process is difficult to monitor. Also, the reaction will be limited by the cleavage conditions. The **Scheme 2.1.3** shows the intended synthetic route for **Type 1C** compounds.



Scheme 2.1.3: Intended synthetic route for **Type 1C** compounds.

As shown in **Scheme 2.1.3**, the sequence starts with Wang resin (Wang 1973). By adding acryloyl chloride, Wang resin is converted to REM resin **6** in **Step 1** (Morphy *et. al.* 1996). In **Step 2**, the homopiperazine is added onto **6**. (Brown *et. al.* 1996) Then, a phosgene equivalent was added to **7** and the other side of the phosgenated resin is reacted with selected amine to form **9**. (Scialdone *et. al.* 1998) After all the building blocks are added onto the resin, the elimination process is conducted by

firstly alkylating the tertiary amine of **9** and then the target compounds **11** are cleaved off the resin by adding DIEA (Morphy *et. al.* 1996). It can be seen that the REM resin returns to its original condition and could be used again which is one of the advantages of using REM resin.

2.1.2: Designed Compounds as Potential Inhibitors of the Mycobacterial dUTPase

In the first chapter, three types of further analogues of T2-3-9 with several kinds of variations were designed. Their general structures were shown in **Table 1.1.6** and here again in **Table 2.1.2**.

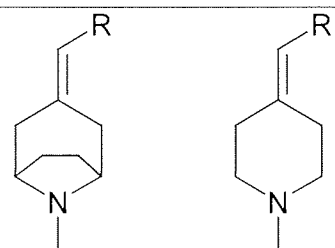
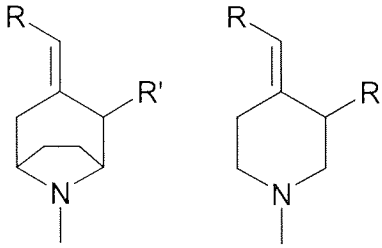
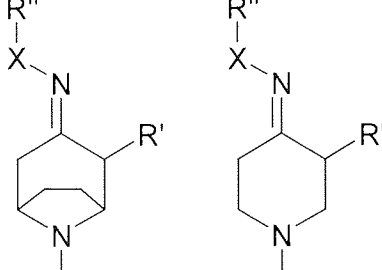
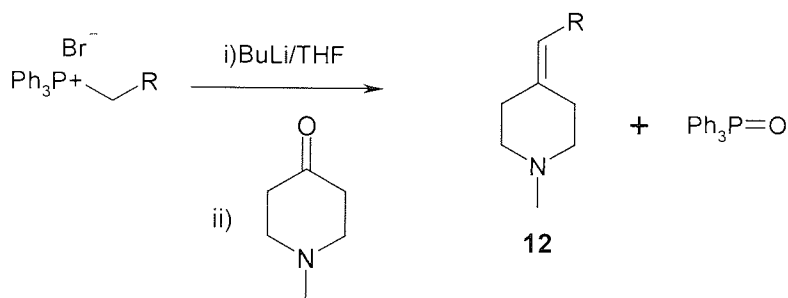
	General structure	Side chain
Type 2A		R = C ₁ -C ₆
Type 2B		R = C ₁ -C ₆ R' = -CH ₂ -OH, -CH ₂ -CH ₂ -OH
Type 2C		X = NH, O R' = -CH ₂ -OH, -CH ₂ -CH ₂ -OH R'' = C ₁ -C ₃

Table 2.1.2: General structure of designed compounds as potential inhibitors of the mycobacterial dUTPase

2.1.2.1. Intended synthetic route of Type 2A compounds

The design of those compounds in **Table 2.1.2** was based on their amenability of synthesis these compounds in chemical laboratory. **Scheme 2.1.4** shows the intended synthetic route of **Type 2A** compounds.

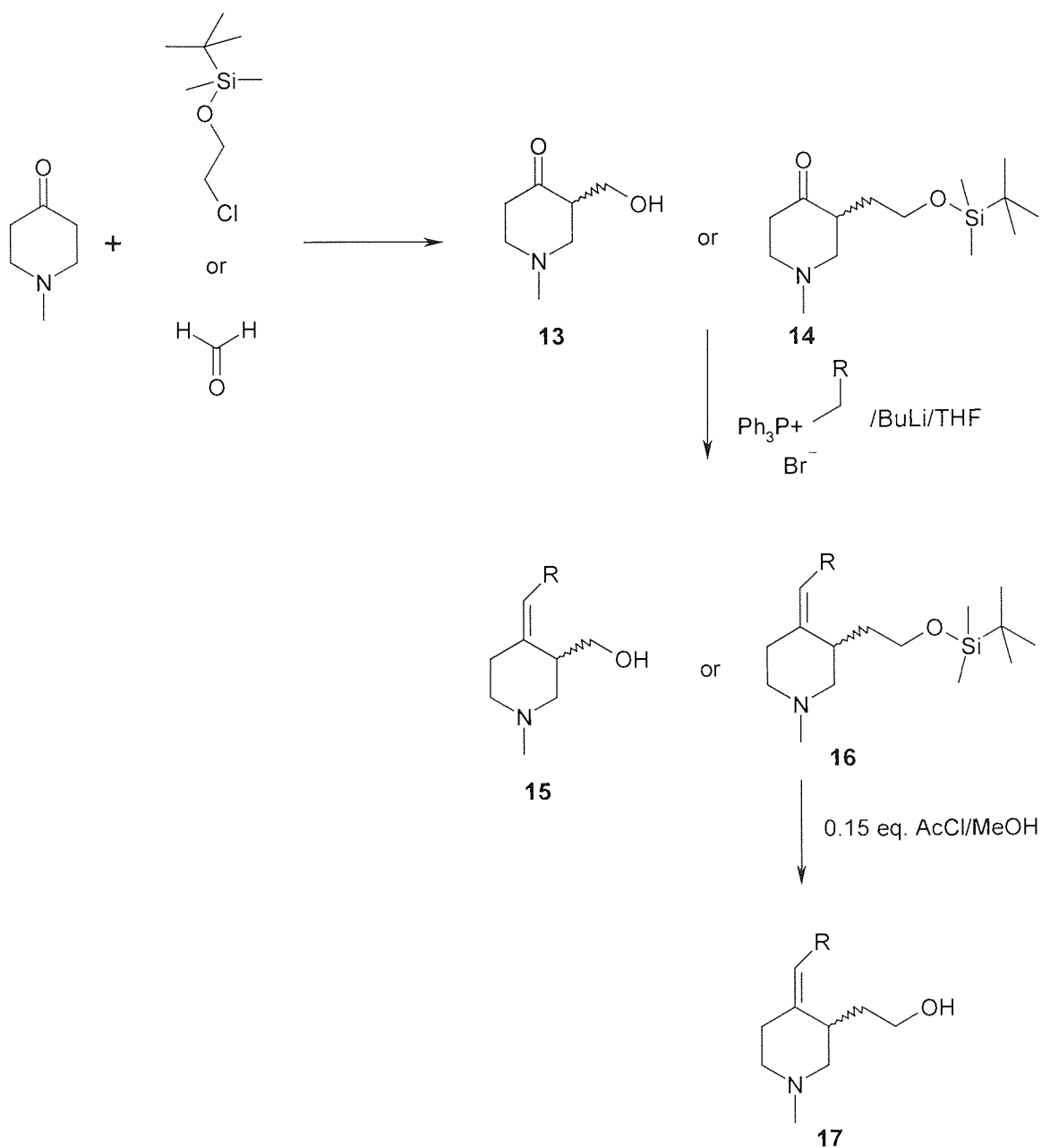


Scheme 2.1.4: Intended synthetic route of **Type 2A** compounds. (Wittig 1983)

This **Scheme 2.1.4** can be used for all three types of compounds. In **Scheme 2.1.4**, it shows that Wittig reaction starts with alkyl-triphenyl-phosphonium bromide with a strong base, BuLi in this case. They will form a phosphorus ylid (also called a phosphorane). Then by slowly adding ketone, a square type of intermediate called an oxaphosphetane should be formed. Then, the intermediate was break down and give alkene product and the by-products of triphenyl-phosphine oxide. (March &Smith 2001)

2.1.2.2. Intended synthetic route of Type 2B compounds

The intended synthetic route of **Type 2B** is shown in **Scheme 2.1.5**.



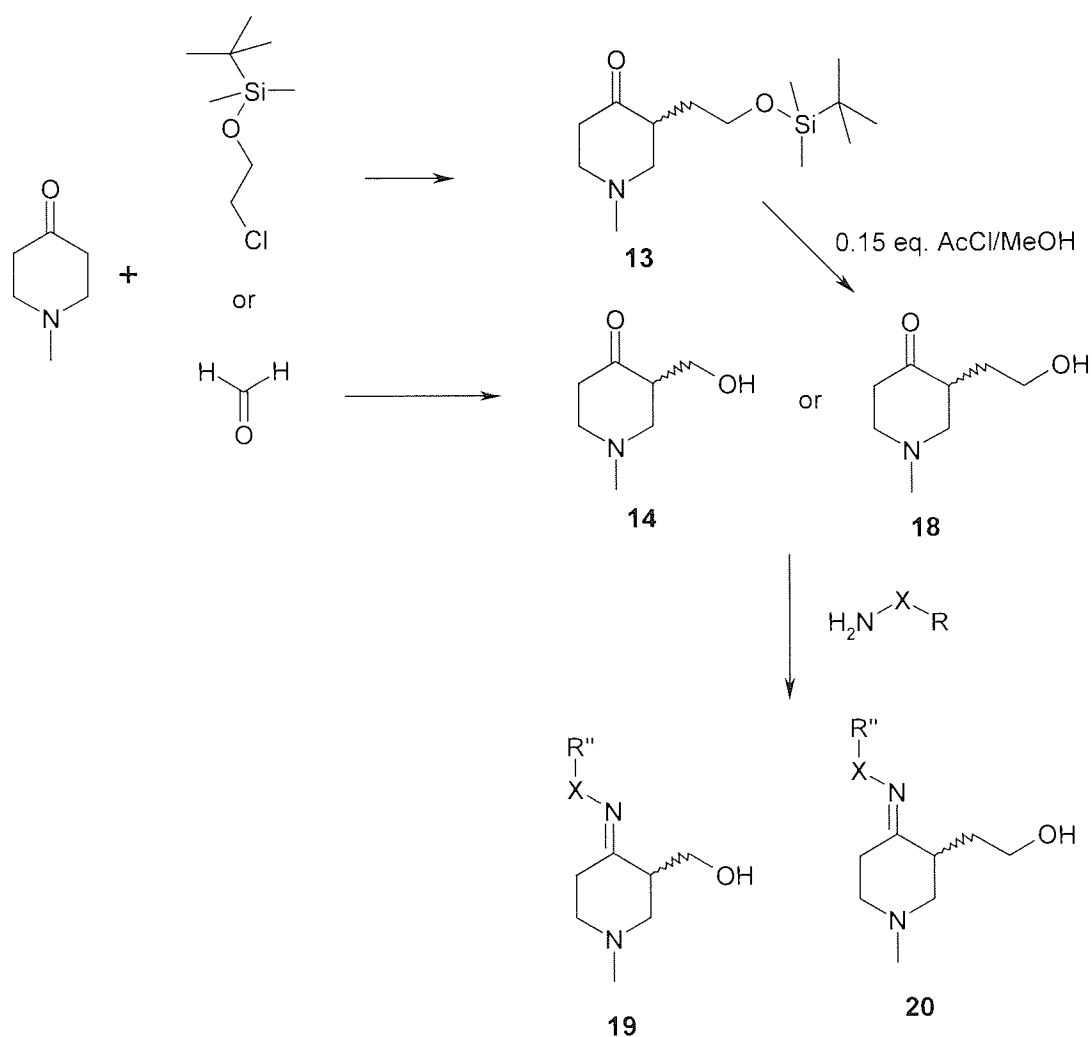
Scheme 2.1.5: Intended synthetic route of **Type 2B**

Scheme 2.1.5 shows two different kinds of alkylation processes. One of them is the hydroxymethylation of ketones by formaldehyde to give **13** (Casas et.al. 2004). The

other one is alkylation with TBDMS (tert-butyl-dimethylsilane)-protected chloroethanol to give **14** (Heslin et. al. 1988). After the alkylation and the Wittig reaction, the hydroxyl group is released by a 0.15 equivalent of acetyl chloride in dried MeOH. (Khan & Mondal 2003) The Wittig reaction is the similar as the route in **Scheme 2.1.4**.

2.1.2.3. Intended synthetic route of Type 2C compounds

The intended synthetic route of **Type 2C** is shown in **Scheme 2.1.6**.



Scheme 2.1.6: Intended synthetic route of **Type 2C**

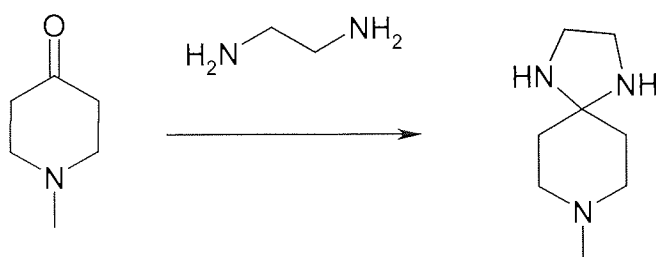
It can be seen that the alkylating process in **Scheme 2.1.6** was same as it in **Scheme 2.1.5**. The final step was adding alkyl-hydrazine or alkyl-hydroxylamine onto the carbonyl group to form the target compounds **19** and **20**. (Rathbone *et. al.* 2006)

2.2: Results and Discussion

2.2.1 Synthesis of Potential Inhibitors of the Mycobacterial Esterase Rv3487c

2.2.1.1 Synthesis of Type 1A compounds

As shown in Scheme 2.1.1, Type 1A compounds were prepared first. For the first step of the reaction, Bergmann *et. al.* (1948) showed that ethane-1,2-diamine and cyclohexanone could react at 1:1 ratio so that both amine group reacted onto the carbonyl to form a five member ring. This reaction is shown in **Scheme 2.1.1**.

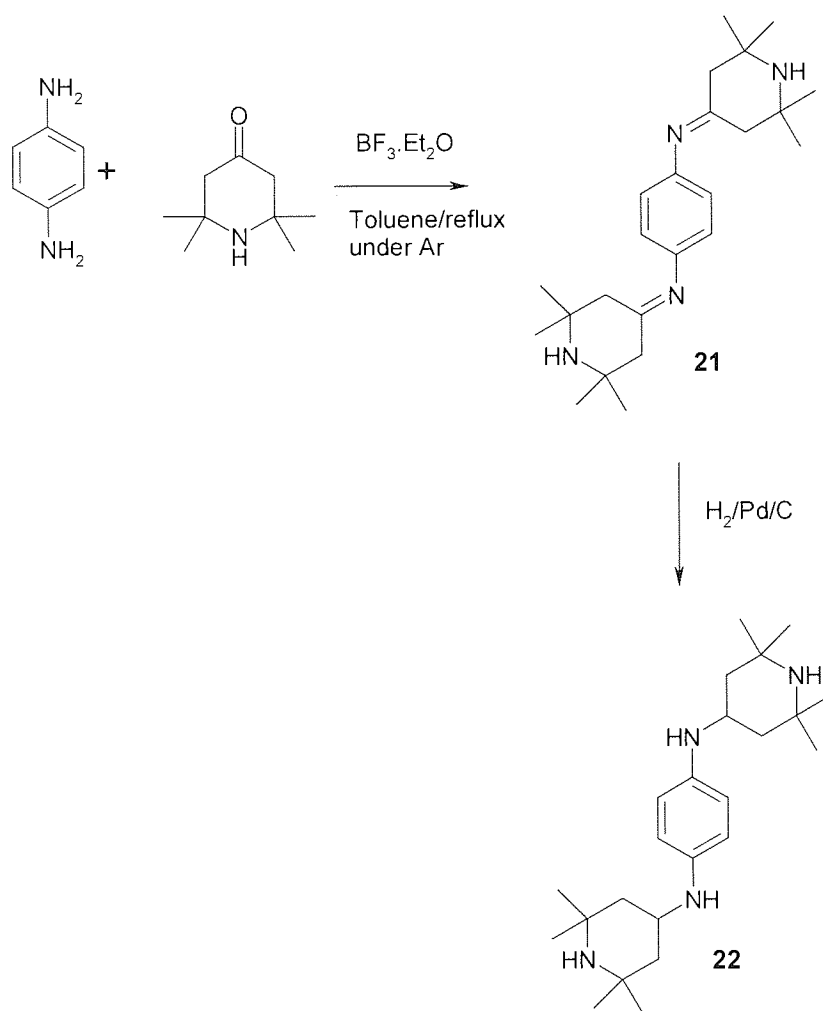


Scheme 2.2.1: One of the side reactions of **Type 1A** compound

Therefore, it was predicted that the first step of this reaction should use excess of 2,2,6,6-tetramethyl-4-piperidone as well as the slow addition of ethylenediamine into the reaction mixture. Under such conditions, the reaction of ethylenediamine should preferentially form N, N'-bis-(2,2,6,6-tetramethyl-piperidin-4-ylidene)-ethane-1,2-diimine rather than the ring form. The first reaction took place in 100°C in toluene for two hours. After the toluene was evaporated, The TLC test (MeOH:EtOAc=1:4 with and without NH₃) of the solid mixture showed that there were at least eight different components and those spots formed a stripe on the TLC plate. This made the separation by flash column nearly impossible. Also, the ring form of the by-products

might have been formed. Another reason for too many products might be that the temperature was too high. Therefore, in the next batch of the reaction, only 80°C was applied. In addition, the solvent was to isopropanol. This time, the number of spots on the TLC plate was reduced. However, even after 15 hours, the TLC of the solid products suggested that much starting material remained. Flash column chromatography was conducted using elutes 20% MeOH in EtOAc. The ¹H NMR spectrum of major part showed that it was the starting material of 2,2,6,6-tetramethyl-4-piperidone. The results suggest that the reaction did not reach completion under these conditions. In Eisch and Sanchez's paper (1985), it was suggested that the reaction of amines with ketones was more difficult than with aldehydes so that the higher temperature and a longer time are usually needed. In addition to this, a catalyst and a drying agent such as molecular sieves may be needed. (Bonnett R. & Emerson T. R. 1965) The next trial of the reaction was conducted in cyclohexane at 90°C using *p*-toluene sulphonic acid as catalyst and molecular sieves as drying agent. The reaction was stopped after twenty-four hours. After the solid products were collected and analysed by TLC, it was found that there were too many spots on the TLC plate which made it difficult for separation. When the reaction was attempted at ambient temperature, mostly starting material remained.

The main problem was thought to be that the ethane-1,2-diamine could form several products. In order to test whether this problem was caused by ethane-1,2-diamine, a compound similar to **Type 1A** was designed for the synthesis.



Scheme 2.2.2: Intended synthesis route of **Type 1A**'s alternative (Markowicz et. al. 2006)

As shown in **Scheme 2.2.2**, benzene-1,4-diamine was used instead of ethane-1,2-diamine. Under this condition, there will be no five-member ring formed. The structure of **21** and the **22** were also docked in to the five homology models in order to find out their docking score and position. However, the result showed that this compound was less favourable for the protein. In the first step of the reaction, the Dean-Stark apparatus was used to collect the water. When the reaction took place, only a small amount of water was collected in the Dean-Stark apparatus at the beginning of the reaction and the water level did not increase in the next twenty-four hours. After the reaction stopped, the NMR of solid mixture showed that there were

still starting materials. This should mean that benzene-1,4-diamine was more difficult to react with ketones.

2.2.1.2 Synthesis of Type 1B compounds.

Attempted synthesis of Type 1B compounds by solution phase was conducted using the information shown in **Scheme 2.1.2**. The types of amine can be varied. The first amine chosen was 3-(4-morpholinyl)-1-propanamine. After the first step the crude material contained compound **4** which was purified by flash chromatography. The yield of purified compounds **4** was 67%.

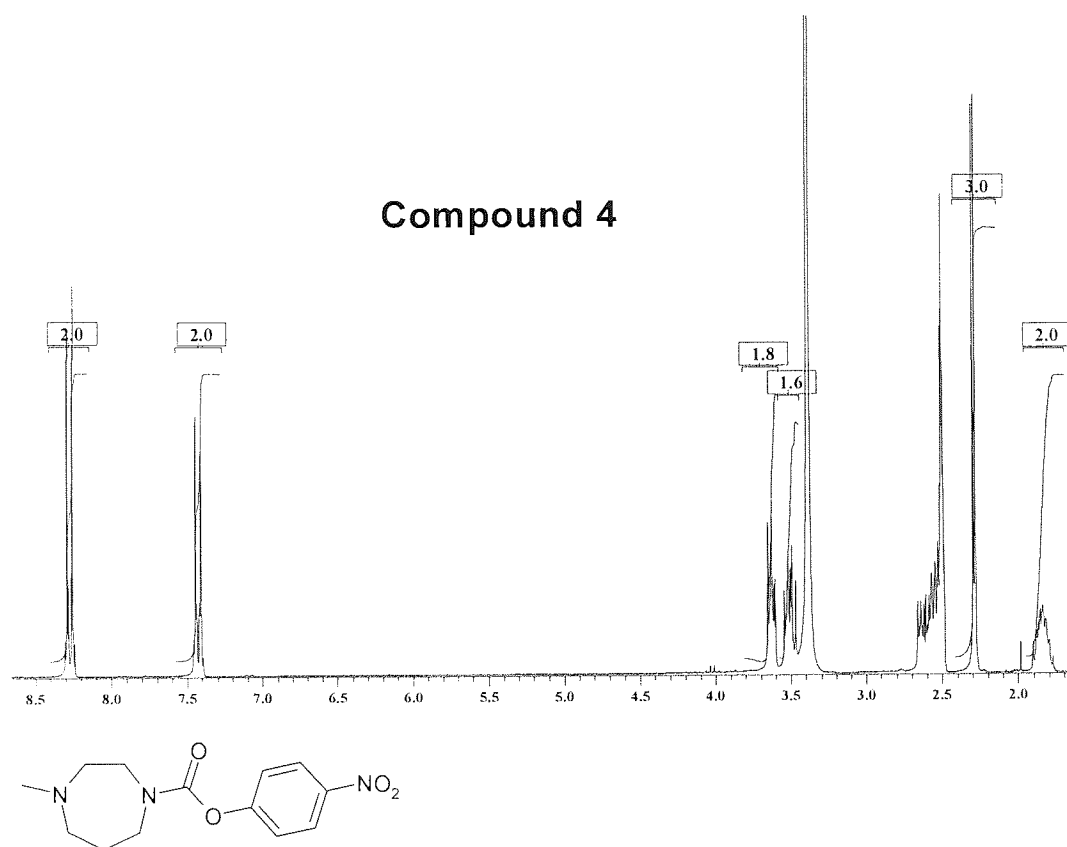


Figure 2.2.1: ^1H NMR spectrum and the structure of compound **4** in $\text{d}_6\text{-DMSO}$.

Figure 2.2.1 clearly shows that two peaks around 8ppm belong to the benzene ring. Other multiplets from 1.8 to 3.7ppm except the peak at 2.3ppm belong to the seven-member ring as they all coupled with each other. The peak at around 2.3ppm should

belong to the methyl group. It behaved as a fine doublet, which might be because of different conformations of the seven-member ring. The correct mass of the molecule was found in the mass spectrum. This was consistent with the proposed structure. During the synthetic work, it was also found that this intermediate was unstable. The ^1H NMR spectrum shows that it decomposes after two days when left under vacuum at room temperature. Therefore, this intermediate has to be used freshly after synthesis.

Coupling of the intermediate compound **4** with N-(3-aminopropyl)morpholine presented several difficulties. Owing to its water solubility, it was not possible to separate the products from the reagents and inorganic material. Flash chromatography was not successful and neither was ion exchange chromatography. Similar difficulties were encountered when 3-(dimethylamino)-1-propylamine was used as the amino component in the reaction.

2.2.1.3 Synthesis of Type 1C compounds using REM linker

As it was difficult to separate the final products, a solid phase method using REM resin was considered as an alternative route. As mentioned before, using solid-state synthesis makes the isolation of intermediates compounds more straightforward than in the solution phase. At the same time, the methyl group that attaches to the seven-member ring could be changed by using different kinds of alkylating agents. Allyl bromide was the first alkylating agent chosen for the synthesis. This was because it is a very powerful alkylating agent. The advantage of using REM resin was the elimination process would not happen without alkylation. However, this will cause trouble if 3-(dimethylamino)-1-propylamine or 3-(4-morpholinyl)-1-propanamine were to be used in the reaction. This is because the alkylating agent will add on to the amine on those substituents. Therefore, in the fourth step in **Scheme 2.1.3**, amines containing R' groups cannot be used. That is why the fifth generation of BTB02557 compounds was designed. Three kinds of alkylation agents were used in the REM

chemistry namely, allyl bromide, iodomethane and benzyl bromide. In the third step in **Scheme 2.1.3**, triphosgene was used instead of phosgene for the safety reason. The reaction was started with Wang resin (1mmole/g) using acryloyl chloride and was completed in four hours at 20°C. (Brown *et. al.* 1996) The ¹³C gel phase NMR spectrum showed that this step was successful.

Twelve different of **Type 1C** compounds were planned to be synthesised and their structures are shown in **Table 2.2.1**.

Name	R	R'
11a		
11b		
11c		
11d		
11e		
11f		
11g		
11h		
11i		
11j		
11k		
11l		

Table 2.2.1: Structures of the 12 compounds that were planned to be synthesised.

In the first trial of REM chemistry, compounds **11a** was planned to be synthesised and the reaction process was designed in the following five steps.

Step 1. REM resin **6** was swollen with a solution of homopiperazine and DMF in a flask. The mixture was agitated on a mixer for 18 hours. Then the resin was washed with DMF, DCM, methanol and dried in vacuum.

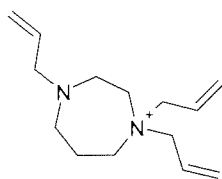
Step 2. The resin **7** was swollen with a solution of anhydrous DCM and DIEA a flask. The flask was immersed in ice-water bath. Triphosgene dissolved in anhydrous DCM was added dropwise over five minutes. After this, the flask was agitated on a mixer for 2 hours at ambient temperature. Then, the resin was washed with DCM, methanol and dried in vacuum.

Step 3. The resin **8** was swollen with a solution of DCM anhydrous and DIEA into a flask. The flask was put into ice-water bath. 1-(3-aminopropyl)-2-pyrrolidinone was added in dropwise. After this, the flask was agitated on a rotator for 2 hours at ambient temperature. Then, the resin was washed with DCM, methanol and dried in the vacuum.

Step 4. The resin **9** was swollen with a solution of DMF and allyl bromide in a flask. The flask was agitated on a mixer for 18 hours at ambient temperature. Then, the resin was washed with DCM, methanol and dried in vacuum.

Step 5. The resin **10** was swollen with a solution of DCM and DIEA in a flask. The flask was agitated on a mixer for 4 hours at ambient temperature. Then, the resin was washed with DCM, methanol, DCM. The filtrate was collected and the solvent was evaporated to give the crude oil products.

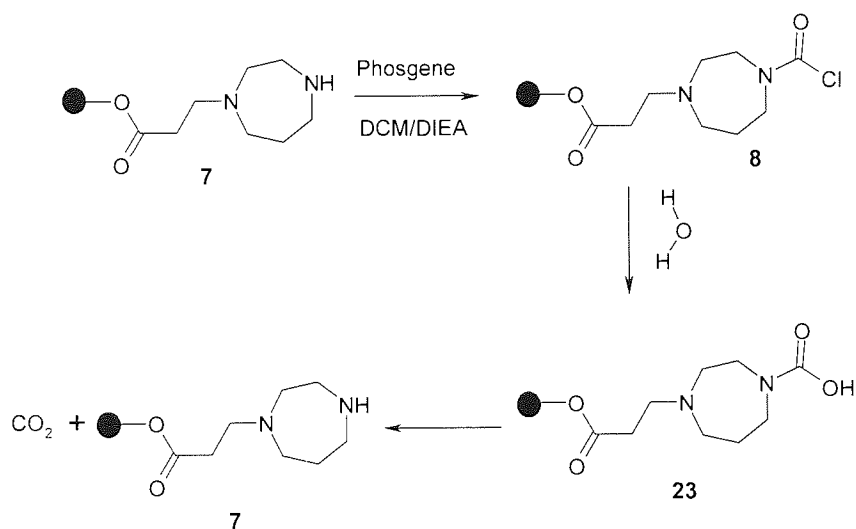
After the products were obtained, NMR and MS were conducted to see if the reactions had succeeded. The result showed no sign of the target compound. However in the MS, a large peak of 221 appeared and the molecular ion of target compound did not show on MS. By analyse the synthesitic process, it was found that 221 peak should belong to the molecular ion shown in **Figure 2.2.1**.



22

Figure 2.2.2: Suggested structure of the compound exhibiting a mass spectral ion of m/z -221

The reason for gaining this structure in the reaction may be because triphosgene did not react with the resin 7. The reaction time was increased to twenty-four hours. However, the result was the same. In this case, the triphosgene might not be a good reagent in this reaction. It was decided to use phosgene solution instead. Therefore, step 2 of the reaction was modified using five-mole equivalent of phosgene solution and every thing else was kept the same. When the phosgene solution was dropped into the reaction mixture, the colour of the mixture turned from orange to red. After every thing finished and the crude oil products were collected, once again, analysis of the crude reaction product indicated the presence of the tri-allylated compounds 22. This means that 1-(3-aminopropyl)-2-pyrrolidinone was not added onto the resin 7. **Scheme 2.2.3** shows the predicted reaction route between step 2 and step 3.



Scheme 2.2.3: Prediction of the reaction between step 2 and step 3.

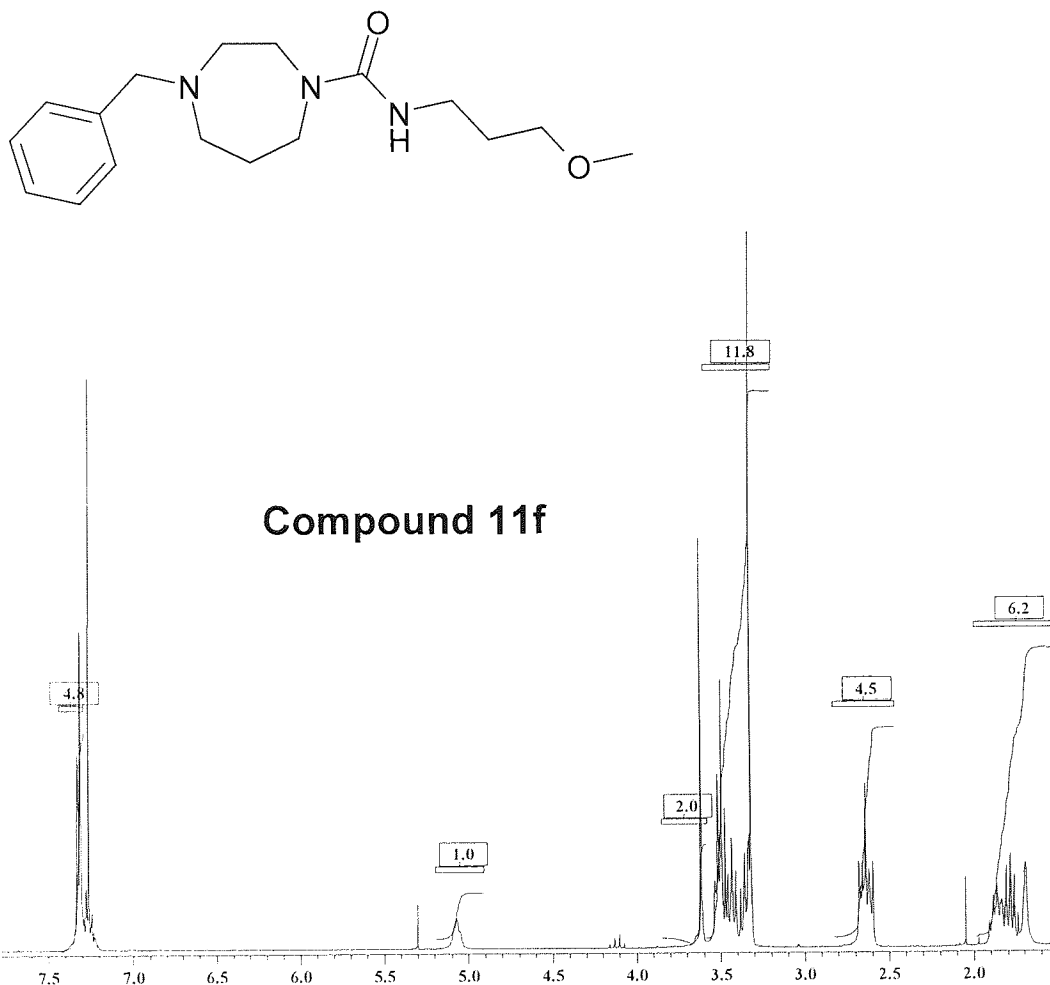
It can be seen from **Scheme 2.2.3** that after addition of phosgene, the resin may go back to its original condition by contact with the moisture in the air. Therefore, in step 3, nothing happens. The resin then was alkylated three times and eliminated to give 22.

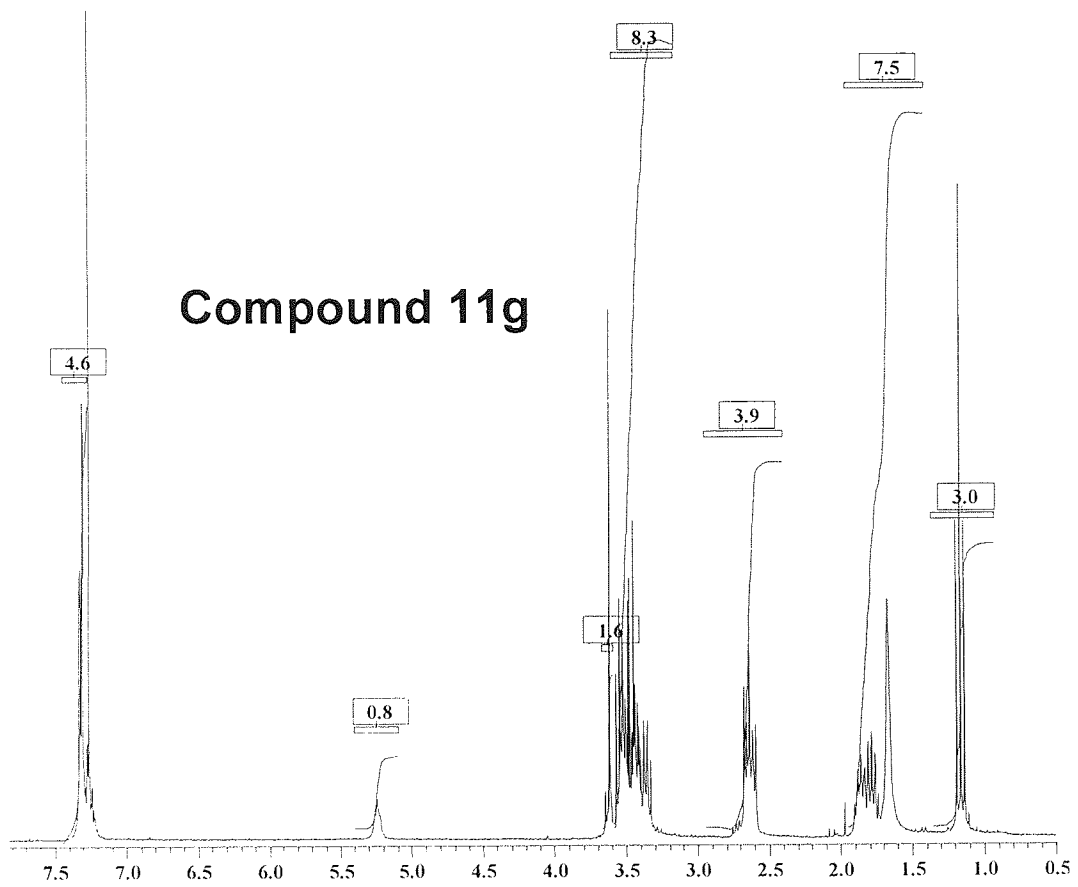
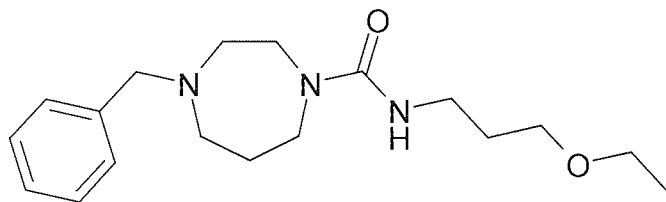
Based on these results, the reaction process was modified. Step 2 and Step 3 were combined. Five mole equivalent of phosgene was dropped into the flask. After 2 hours, 15 mole equivalent of amine was added. In this case, the phosgene will not make contact with moist air and amount of the amine should be enough for the reaction. After this new process, the molecular ion $(M+1)^+$ of the desired products could be seen in the mass spectrum and the 221 peak disappeared. This result means that the reaction was successful.

In theory, the final products should only contain the target compound, DIEA·HBr salt and a small amount of solvent and DIEA. Brown's (1997) solid phase extraction method (SPE) was implemented to remove DIEA·HBr and other impurities. A modified SPE method was used here try to get the pure compounds. After the SPE column, TLC (MeOH:EtOAc=1:4 with 2% triethylamine) indicated that there were at least three components. The NMR of the oil-like product also suggested impurities. Those impurities may be caused by the fact that in each step of the reaction, there might be some chemical trapped inside the resin and which accumulates throughout the reaction. Therefore, the Soxhlet extraction was used in order to clean the resin before the elimination step. At the beginning, the Soxhlet extraction was applied after the alkylation step. However, in the final elimination step, no product was found. When the extraction was applied before the alkylation step, the products appeared. Although some of the impurities were removed, the TLC still showed more than two tailing spots. At the same time, the ^1H NMR spectrum of the products after SPE column still suggested mixture. In this case, the current separation method could not isolate the pure compound.

As it was difficult to judge which spots on the TLC plate represented the target compound, a different alkylating agent, benzyl bromide, was used instead of allyl bromide. After the SPE column, it was very clear that four spots could be seen under UV visualisation for **11f**, **11g** and **11h** on the TLC (MeOH:EtOAc 1:4 with 2% triethylamine) plate. The TLC of compound **11e** showed at least five overlapping

spots which precluded flash chromatography. **11f** and **11g** were separated by flash chromatography with a good yield of 17.4% and 21.3% respectively. **11h** was separated by elution from the SPE column. However, this separation method gave a very low yield of 5.6%, compared to the other two products. The ^1H NMR spectra of these three compounds are shown in **Figure 2.2.2**.





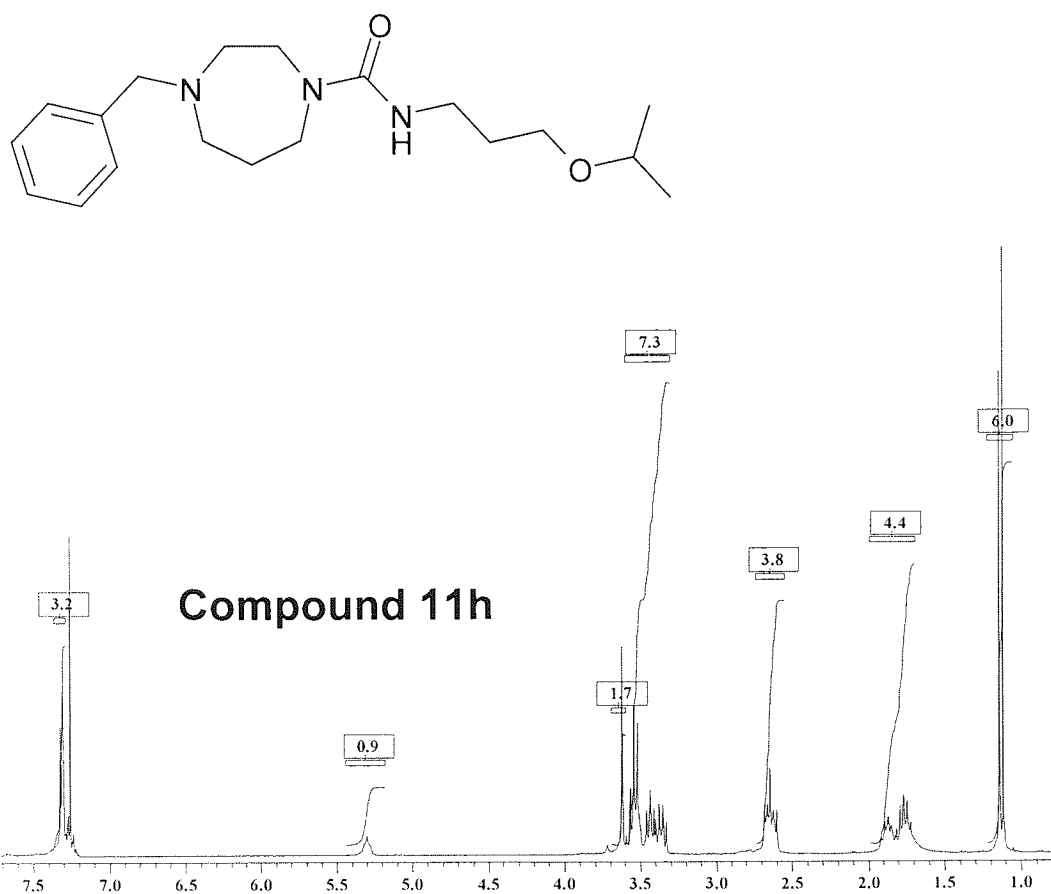


Figure 2.2.3: ^1H NMR (CDCl_3) of Compounds of **11f**, **11g** and **11h**.

When these three ^1H NMR spectra in **Figure 2.2.2** were compared and analysed, it was very clear that the peaks around 7.3 ppm belonged to the hydrogens on the benzene ring. The small triplet around 5.2 ppm should belong to the NH. This was used to confirm that the compounds had been formed. The peak in the region from 1.0 ppm to 1.7 ppm should belong to the variable of R' group shown in **Table 2.2.1**. The other peaks should all belong to all kinds of CH_2 group. The small peaks in the first ^1H NMR spectrum belong to the solvent and all those peaks disappeared after one day left in the vacuum desiccators. The mass spectra of these three products also provided the correct molecular ion. All this evidence showed the synthesis and the separation was successful.

In the MS, the peaks of 191 could be found in all the spectra of **11f**, **11g** and **11h**. The mass of the following molecular ion **24** was found to be 191. This peak should be a fragmentation ion.

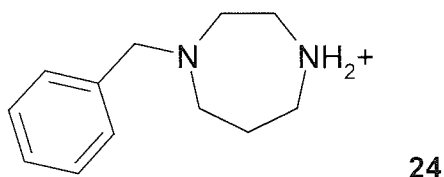
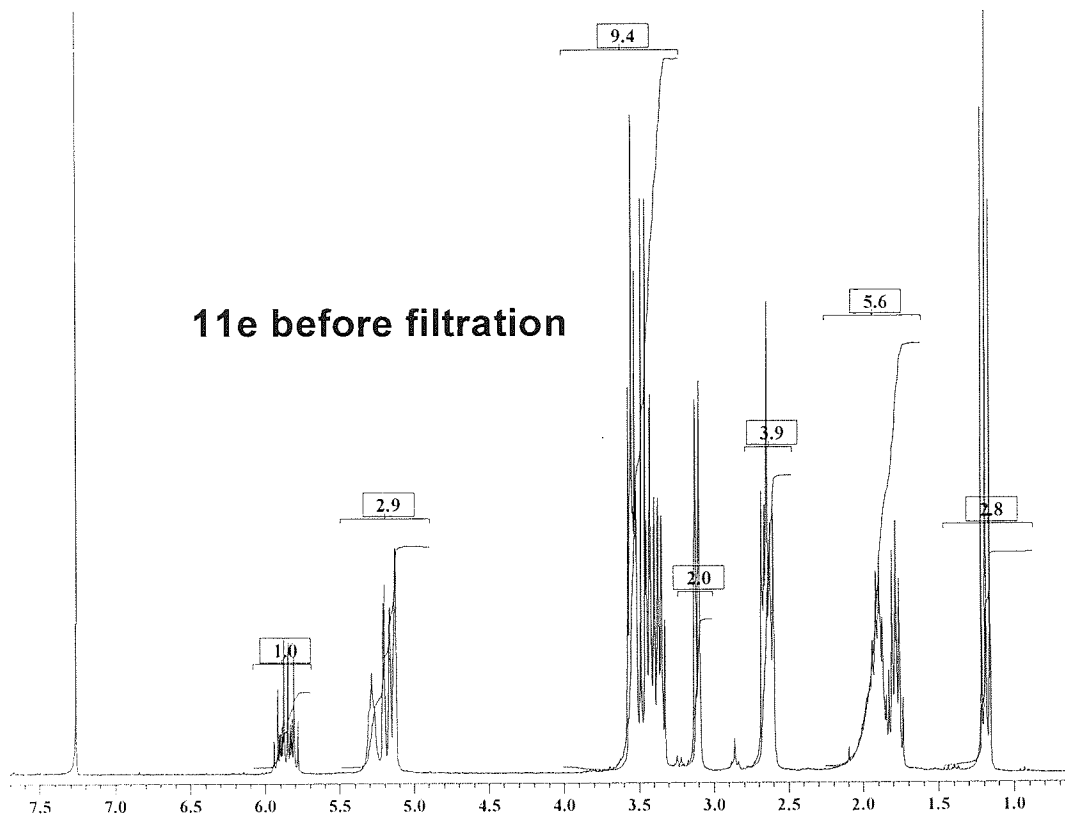
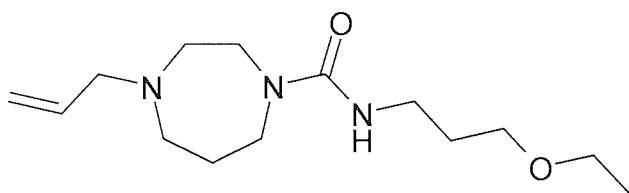


Figure 2.2.4: Predicted fragmentation of compounds **11**.

After these compounds were successfully separated, a separation of compound of **11c** was conducted using the similar method that used to separate compound **11h**. The compound **11c** was purified after elution on a SPE column. After the solvent evaporated, a yellow oil and colourless crystals appeared. That crystal was tested and was found to be the K_2CO_3 . The oil part was tested by 1H NMR and the MS and found to be the correct target compounds. In order to separate those crystals, a normal filtration was conducted. Strangely, the 1H NMR spectrum of the product before and after filtration was different. Those two spectra are shown in **Figure 2.2.5**.



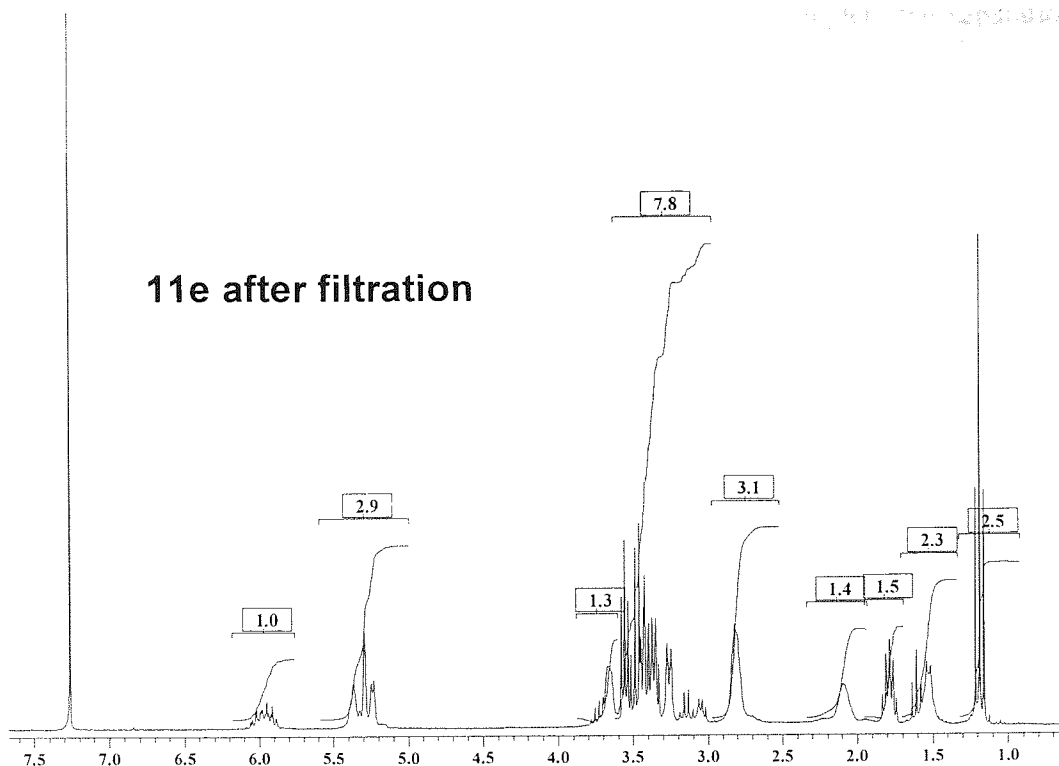


Figure 2.2.5: The ^1H NMR spectra (CDCl_3) and the structure of **11e** before and after filtration.

It can be seen in **Figure 2.2.2** that before filtration, the compound appears pure and all the peaks were in the right place when compared to the NMR of compound **11g** in **Figure 2.2.3**. However, if the two spectra in **Figure 2.2.5** are compared, it was clear that some change has happened. By analysing the second graph, it is clear that allyl polymerisation had not occurred. Higher temperature ^1H NMR at 30°C and 40°C of the product also showed the same result as in the second graph. This could mean that there was no delocalisation effect of carbonyl group. Another explanation of this might be allyl migration (Yadav & Lande 2005).

The synthesis of compounds **B31** to **B34** was also conducted. The ^1H NMR spectra and the MS all showed that the synthesis was successful. However, the separation of those compounds was still an issue. This was because the R_f value (MeOH:EtOAc=1:4 with 2% triethylamine) was below 0.1 so that the flash column

could not applied. The modified SPE extraction could not complete the separation either.

2.2.2 Synthesis of Potential Inhibitors of the Mycobacterial dUTPase

2.2.2.1 Synthesis of Type 2A compounds

It was intended to prepare **Type 2A** compounds by Wittig reactions. The compounds are shown in **Table 2.2.2**

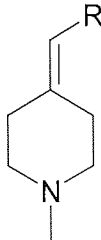
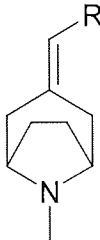
R group		
Methyl	12a	12h
Ethyl	12b	12i
Vinyl	12c	12j
Propyl	12d	12k
Butyl	12e	12l
Pentyl	12f	12m

Table 2.2.2: Compounds were planned to be synthesised by Wittig reaction.

In the first trial, **12b** was synthesised. The colour of the phosphorus ylid was red. After the 1-methyl-4-piperidone was added, the red colour slowly turned into white and the temperature probe used to monitor the temperature of the reaction mixture showed a temperature increase. This should suggest that the reaction was happening. After the reaction finished, an extraction between ether and concentrated sodium chloride solution was conducted. The solvent was evaporated and the TLC (MeOH:EtOAc=1:9 with 2% ammonia) of the biggest spot which can be seen under UV should be triphenyl-phosphine oxide. Flash chromatography was conducted using the same solvent system as used in TLC to separate the products which were dried

under vacuum. The next day, it was found that all the products had evaporated. In this case, distillation should be used to separate the products owing to their low boiling point. It was decided to use Kugelrohr distillation, as the amount of the products was small. This time, the separation was successful. Therefore, the synthesis of other compounds was attempted. When **12c** was synthesised, a white coloured gel-like product was found. It would not dissolve in water or organic solvents. This was perhaps caused by the polymerisation of the double to form a polymer. It was also found that the compounds of **12a** could not be separated from the starting material of 1-methyl-4-piperidone by distillation. Therefore, a new method should be found to purify compound **12a**.

When the starting ketone changed from 1-methyl-4-piperidone to tropinone, it was found that the Wittig reaction failed. After the ketones were added to the ylid, the colour of the reaction mixture remained red. Even after two days and a temperature of 50°C, everything was still the same. After reaction, there was no sign of the target compound using NMR and MS. All this evidence showed that the tropinone was not reacting with the phosphorus ylid. In order to gain some insight, several calculations of QM/MM/MD using CaChe WorkSpace (Version 6.1.10) were conducted. Two structures of the intermediates of compounds **25** and **26** were imported into the CaChe WorkSpace and the structures are shown in **Figure 2.2.6**.

2.3 Conclusion

The synthesis of the designed compounds may be summarized as follows. The synthesis of **Type 1A** compounds was unsuccessful. The synthesis of **Type 1B** compounds was successfully as judged by NMR spectroscopy and MS data. However, they could not be purified. **Type 1C** compounds were synthesised by solid-state chemistry using the REM linker. NMR spectroscopy and the MS data show that the reaction was successful. However, it was only possible to separated three of them by flash chromatography or SPE extraction.

The designed drugs for mycobacterium dUTPase were also synthesised. Four **Type 2A** compounds were successfully synthesised by Wittig reaction and separated by Kugelrohr distillation. The compounds **12a** could not be purified because it was thought that it had a similar boiling point to the starting material. The compounds **12h** to **12m** could not be synthesised by Wittig reaction. The molecular modelling result shows that the total energy of intermediate of **12h** to **12m** was too high. This might be the reason that the reaction would not happen.

2.4 Recommendation

- There might be a better separation method for **Type 1B** and **Type 1C** compounds.
- Whether compounds **11e** underwent allyl migration was not clear. This could be investigated.
- Synthesis and separation of **Type 2B** and **Type 2C** compounds could be conducted in future.

- All of the synthesised compounds were sent to TAACF and tested against M-TB H37Rv. After the testing resulting obtained, a further SAR study on those kinds of compounds may be required.

2.5 Experimental

Instrumentation

Proton NMR spectra were obtained on a Bruker AC 250 instrument operation at 250MHz as solutions in d_6 -DMSO or $CDCl_3$ and referenced from $\delta_{DMSO}=2.50$ ppm or $\delta_{CDCl_3}=7.26$ ppm unless otherwise stated. Infrared spectra were recorded as KBr discs or NaCl discs on a Mattson 3000 FTIR spectrophotometer. Atmospheric pressure chemical ionisation mass spectrometry (APCI-MS) was carried out on a Hewlett-Packard 5989B quadrupole instrument connected to an electrospray 59987A unit with an APCI accessory and automatic injection using a Hewlett-Packard 1100 series autosampler. Melting points were obtained using a Reichert-Jung Thermo Galen hot stage microscope and are corrected.

REM resin synthesis: Wang resin (NOVAbiochem 100-200 mesh) (10g, 1mmol/g) was swollen in a solution of DCM (anhydrous, 50mL), DIEA (16.5mL, 100mmol) and acryloyl chloride (8.2mL, 100mmole) under Ar. The reaction flask was agitated for four hours. Then, the resin was collected by filtration and washed with DCM (3 \times 20mL), methanol (3 \times 20mL) and DCM (3 \times 20mL). The resin was dried under vacuum.

^{13}C NMR in $CDCl_3$: δ 166.24 (C=O), 159.26 (PhCH-O), 131.14 (CH=), 128.55 (PhCH), 114.97 (CH₂=) 66.31 (Ph-CH₂-O), 40.62 (CH₂-CH=) ppm.

Modified process of REM Chemistry:

Step 1. REM resin **6** (2g, around 1mmole/g) was swollen with a solution of homopiperazine (2g, 20mmole) and DMF (10mL). The mixture was agitated for 18 hours. Then the resin was collected by filtration and washed with DMF (3 \times 5mL), DCM (3 \times 5mL), methanol (3 \times 5mL) and dried under vacuum.

Step 2. The resin **7** was swollen with a solution of DCM (anhydrous 10mL) and DIEA (3.3mL, 20mmole) under Ar. The reaction flask was cooled using ice-water bath. Phosgene solution (20% in toluene, 5mL, 10mmol) was added dropwise over 10 minutes. After this, the reaction flask was agitated for 2 hours at ambient temperature. Then, the selected amine (30mmol) was added dropwise over 5 minutes. The reaction flask was agitated overnight. After this, the resin was collected by filtration and extracted by Soxhlet in DCM for at least 24 hours and dried under vacuum.

Step 3. The resin **9** was swollen with a solution of DMF (10mL) and selected alkylating agents (10mmol) were added. The reaction flask was agitated for 18 hours at ambient temperature. Then, the resin was collected by

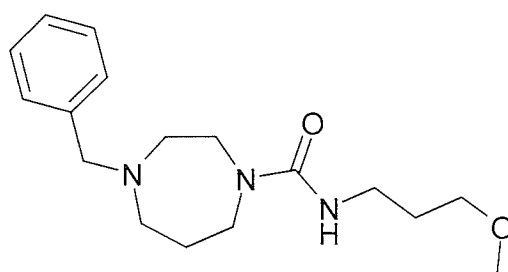
filtration and washed with DCM (3×5mL), methanol (3×5mL) and dried under vacuum.

Step 4. The resin **10** was swollen with a solution of DCM (10mL) and DIEA (3.3mL, 20mmole). The reaction flask was agitated for 4 hours at ambient temperature. Then, the resin was washed with DCM (3×5mL), methanol (3×5mL), DCM (3×5mL). The filtrate was collected and the solvent was evaporated to give the crude products as oils.

Modified SPE extraction: A column containing a mixture of silica (2g) and K₂CO₃ dry powder (0.4g) was loaded with the solution of crude oil and 1mL DCM. The column was dried under vacuum. Then, the column was washed with ethyl acetate (20mL), the product was eluted with 20% methanol in ethyl acetate in 4mL fractions and tested by TLC (MeOH:EtOAc=1:4 with 2% triethylamine).

4-Benzyl-[1,4]diazepane-1-carboxylic acid (3-methoxy-propyl)-amide [11f]

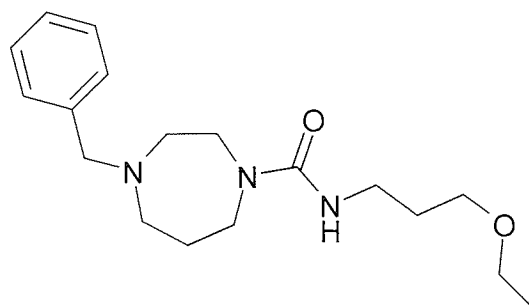
Structure



Physical state	Yellow oil
Yield	17.4% (purified by Flash Chromatography MeOH:EtOAc 1:4 with 2% triethylamine)
¹ H NMR (CDCl ₃ ; δCDCl ₃ =7.26ppm)	1.69-1.91 (overlapping m, 6H, 3×CH ₂), 2.59-2.68 (overlapping m, 4H, 2×CH ₂), 3.31-3.53 (overlapping m, 9H, CH ₃ , 3×CH ₂), 3.62 (s, 2H, CH ₂ -Ar), 5.07 (t, J=4.4Hz 1H, NH), 7.31-7.32 (overlapping m, 5H, 5×ArH) ppm
IR (NaCl disc)	3358 (NH), 2928, 2864, 2813, 1626 (C=O), 1532, 1451, 1400, 1202, 1114, 1067 cm ⁻¹
MS (APCI +ve)	m/z=306 (M+H) ⁺

4-Benzyl-[1,4]diazepane-1-carboxylic acid (3-ethoxy-propyl)-amide [11g]

Structure

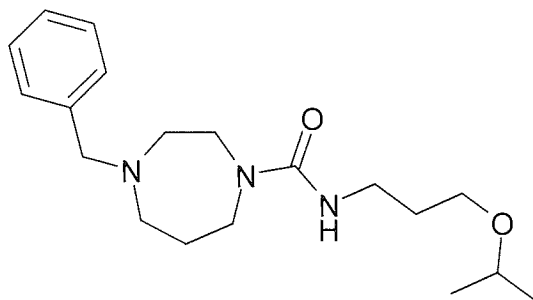


Physical state	Yellow oil
Yield	21.3% (purified by Flash Chromatography MeOH:EtOAc 1:4 with 2% triethylamine)
¹ H NMR (CDCl ₃ ;	1.16 (t, J=6.95Hz, 3H, CH ₃), 1.66-1.88 (overlapping m, 8H,

$\delta\text{CDCl}_3=7.26\text{ppm}$	4 $\times\text{CH}_2$), 2.59-2.68 (overlapping m, 4H, 2 $\times\text{CH}_2$), 3.33-3.57 (overlapping m, 9H, CH_3 , 3 $\times\text{CH}_2$), 3.62 (s, 2H, $\text{CH}_2\text{-Ar}$), 5.24 (t, $J=4.4\text{Hz}$ 1H, NH), 7.31-7.32 (overlapping m, 5H, 5 $\times\text{ArH}$) ppm
IR (NaCl disc)	3356(NH), 2966, 2929, 2863, 1622(C=O), 1531, 1451, 1398, 1374, 1350, 1210, 1108 cm^{-1}
MS (APCI +ve)	$m/z=320$ (M+H) ⁺

4-Benzyl-[1,4]diazepane-1-carboxylic acid (3-isopropoxy-propyl)-amide [11h]

Structure

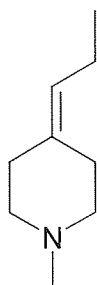


Physical state	Yellow oil
Yield	5.6% (purified by SPE extraction)
^1H NMR (CDCl_3 ; $\delta\text{CDCl}_3=7.26\text{ppm}$)	1.13 (d, $J=6.3\text{Hz}$, 6H, 2 $\times\text{CH}_3$), 1.72-1.89 (overlapping m, 7H, 3 $\times\text{CH}_2$, CH), 2.60-2.68 (overlapping m, 4H, 2 $\times\text{CH}_2$), 3.33-3.56 (overlapping m, 9H, CH_3 , 3 $\times\text{CH}_2$), 3.62 (s, 2H, $\text{CH}_2\text{-Ar}$), 5.30 (t, $J=4.4\text{Hz}$ 1H, NH), 7.30-7.32 (overlapping m, 5H, 5 $\times\text{ArH}$) ppm
IR (NaCl disc)	3355(NH), 2967, 2930, 2865, 1626(C=O), 1531, 1452, 13989, 1364, 1207, 1125, 1071 cm^{-1}
MS (APCI +ve)	$m/z=334$ (M+H) ⁺

General procedure of Wittig reaction: Alkyl-triphenylphosphonium bromide (40mmol) was suspended in THF (anhydrous 70mL) fitted with stirring bar and cooled in ice-water bath. *n*-BuLi (25mL, 1.6M in hexane) was added dropwise and carefully monitored by temperature probe. The rate of addition of the *n*-BuLi solution was controlled such that the reaction temperature remained below 30°C. It took four hours at ambient temperature. Then 1-methyl-4-piperidone (5mL, 44mmol) was added in dropwise in ice-water bath. The reaction took overnight. The reaction was stopped by quenching with brine. The products were extracted with ether and separated by Kugelrohr distillation.

1-Methyl-4-propylidene-piperidine [12b]

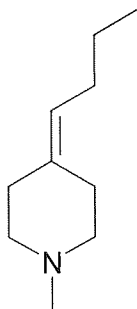
Structure



Physical state	Colourless oil
Yield	8.4%
^1H NMR (CDCl_3 ; $\delta\text{CDCl}_3=7.26\text{ppm}$)	0.94 (t, $J=7.6\text{Hz}$, 3H, $\text{CH}_3\text{-CH}_2$), 1.99, (p, $J=7.6\text{Hz}$, 2H, $\text{CH}_2\text{-CH}_3$), 2.17-2.40 (overlapping m, 11H, $\text{CH}_3\text{-N}$, $4\times\text{CH}_2$), 5.14 (t, $J=7.3\text{Hz}$ 1H, CH=) ppm
IR (NaCl disc)	2963, 2933, 2873, 2359, 2335, 1663, 1644, 1451, 1402, 1377, 1282, 1250, 1192, 1067, 1032 cm^{-1}
MS (APCI +ve)	$m/z=140$ (M+H^+)
Kugelrohr distillation condition	Reduced pressure of 50 mbar 85°C

1-Methyl-4-butylidene-piperidine [12d]

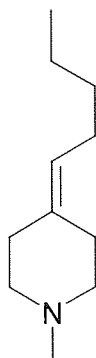
Structure



Physical state	Colourless oil
Yield	12.1%
^1H NMR (CDCl_3 ; $\delta\text{CDCl}_3=7.26\text{ppm}$)	0.88 (t, $J=7.6\text{Hz}$, 3H, $\text{CH}_3\text{-CH}_2$), 1.27-1.41 (overlapping m, 2H, $\text{CH}_2\text{-CH}_3$), 1.91-2.00, (overlapping m, 2H, $\text{CH}_2\text{-CH=}$), 2.18-2.39 (overlapping m, 11H, $\text{CH}_3\text{-N}$, $4\times\text{CH}_2$), 5.14 (t, $J=7.3\text{Hz}$ 1H, CH=) ppm
IR (NaCl disc)	2958, 2931, 2869, 2358, 2337, 1660, 1641, 1454, 1373, 1277, 1197, 1172, 1060 cm^{-1}
MS (APCI +ve)	$m/z=154$ (M+H^+)
Kugelrohr distillation condition	Reduced pressure of 14 mbar 60°C

1-Methyl-4-pentylidene-piperidine [12e]

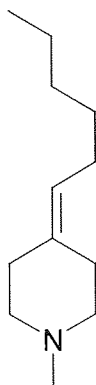
Structure



Physical state	Colourless oil
Yield	10%
^1H NMR (CDCl_3 ; $\delta\text{CDCl}_3=7.26\text{ppm}$)	0.88 (t, $J=6.7\text{Hz}$, 3H, $\text{CH}_3\text{-CH}_2$), 1.26-1.31 (overlapping m, 4H, $2\times\text{CH}_2$), 1.97 (q, $J=7.1\text{Hz}$, 2H, $\text{CH}_2\text{-CH=}$), 2.17-2.39 (overlapping m, 11H, $\text{CH}_3\text{-N}$, $4\times\text{CH}_2$), 5.13 (t, $J=7.3\text{Hz}$ 1H, CH=) ppm
IR (NaCl disc)	2957, 2932, 2778, 2357, 2338, 1670, 1459, 1448, 1366, 1276, 1196, 1136, 1262 cm^{-1}
MS (APCI +ve)	$m/z=168$ ($\text{M}+\text{H}$) $^+$
Kugelrohr distillation condition	Reduced pressure of 14 mbar 80°C

1-Methyl-4-hexylidene-piperidine [12f]

Structure



Physical state	Colourless oil
Yield	9.5%
^1H NMR (CDCl_3 ; $\delta\text{CDCl}_3=7.26\text{ppm}$)	0.88 (t, $J=7.0\text{Hz}$, 3H, $\text{CH}_3\text{-CH}_2$), 1.23-1.31 (overlapping m, 6H, $3\times\text{CH}_2$), 1.92-2.00 (overlapping m, 2H, $\text{CH}_2\text{-CH=}$), 2.18-2.40 (overlapping m, 11H, $\text{CH}_3\text{-N}$, $4\times\text{CH}_2$), 5.14 (t, $J=7.3\text{Hz}$ 1H, CH=) ppm
IR (NaCl disc)	2955, 2930, 2779, 2359, 2338, 1622, 1532, 1458, 1366, 1276, 1199, 1134, 1060 cm^{-1}
MS (APCI +ve)	$m/z=182$ ($\text{M}+\text{H}$) $^+$
Kugelrohr distillation condition	Reduced pressure of 14 mbar 120°C

Chapter 3

Carboxamidrazones Synthesis and Screening

3.1 Introduction

In a set of Mamolo's papers (1992, 1993, 1996, 1999), some pyridine-2-carboxamidrazones and pyridine-4-carboxamidrazones were synthesized and tested against *M. tuberculosis* H37Rv and *M. tuberculosis* isolated from human bronchial aspirates. Among those carboxamidrazones, some showed interesting inhibitory activity of a MIC 1-4 $\mu\text{g}/\text{mL}$ even against some of the *M. tuberculosis* organisms that were resistant to isoniazid, rifampicin and ofloxacin. In those papers, they showed that the important structural feature is the amidrazone moiety which is connected to a pyridine-based group in the ortho-position as well as hydrophobic substituents group attach to the benzylidene side which gives the whole molecule ability to penetrate the highly lipophilic mycobacterial cell wall.

In order to identify more active carboxamidrazone compounds, Billington *et.al.* (1998) applied robotic synthesis methods to the syntheses over a hundred compounds including pyridine-2-, pyrazine-2- and quinoline-2-carboxamidrazone in a short period and then did rapid primary screening on *M. fortuitum*. This organism was used because it is safer to use and much faster growing than *M. tuberculosis*. Four pyridine-2-carboxamidrazones with the MIC value less than 50 $\mu\text{g}/\text{mL}$ against *Mycobacterium fortuitum* were identified to be leading structures for the further potential anti-mycobacterial compounds library. Their structures are shown in **Table**

3.1.1

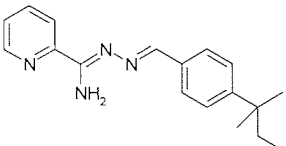
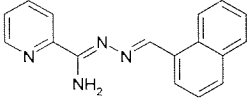
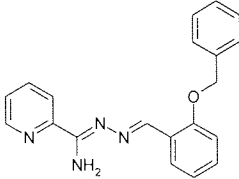
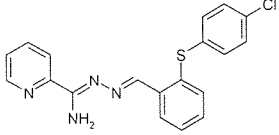
Comp.	Structure	MIC ($\mu\text{g/mL}$)
i		6.3-12.5
ii		12.5-25
iii		25-50
iv		12.5-25

Table 3.1.1: Active compounds from Billington *et.al.* (1998)

All the compounds in **Table 3.1.1** together with other closely related carboxamidrazones were tested for their toxicity using a rat liver metabolism system, human erythrocytes and mononuclear leucocytes (Coleman *et. al.* 1999, 2000, 2001). Only **iii** showed more toxic to human mononuclear leucocytes than INH.

Further work has been conducted by Rathbone *et.al.* (2006). A new set of phenolic N^1 -benzylidene-pyridinecarboxamidrazones was synthesized and tested against gram-positive bacteria including *S. aureus*, MRSA, *E. faecium*, *E. faecalis*, *S. epidemisis* and *S. haemilyticus*. Among all those compounds, N^1 -[3,5-di-(*tert*-butyl)-2-hydroxybenzylidene]-pyridine-2-carboxamidrazone was found to be highly active against all the bacteria with an MIC value of 2-4 $\mu\text{g/mL}$. In contrast, the same compound was found to be inactive against a range of gram-negative bacteria. However, the direct leucocyte toxicity test showed that this compound is extremely toxic and caused 100% cell lysis while the control substance DMSO only caused $7.4 \pm 1.1\%$ leucocytes cell death. N^1 -[3,5-di-(*tert*-butyl)-2-hydroxybenzylidene] -

pyridine-2-carbox-amidrazone is the leading compound for the next generation of carboxamidrazones.

In this study, four types of carboxamidrazones were synthesised and some of them were tested against *M. tuberculosis* H37Rv as well as other types of microorganisms such as *E. coli*, *S. aureus*. Their general structures are shown in **Table 3.1.2**.

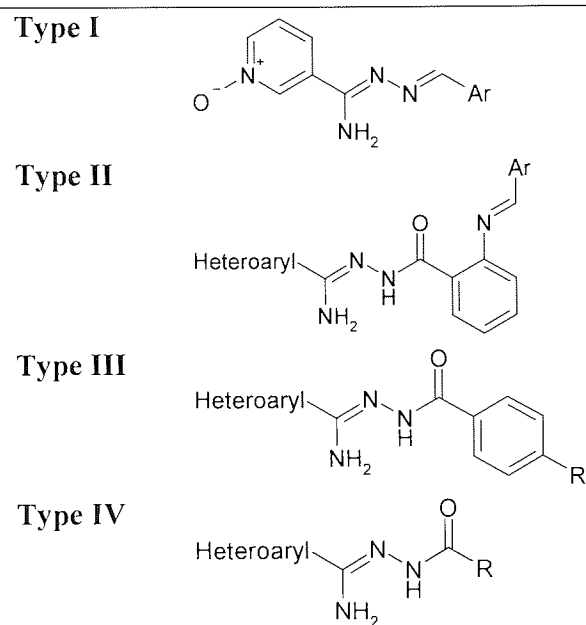
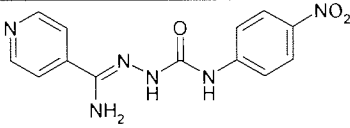
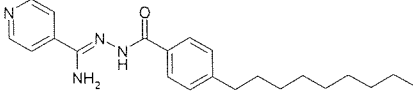
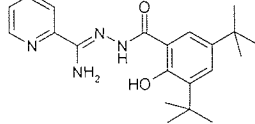
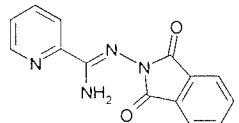
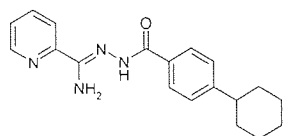
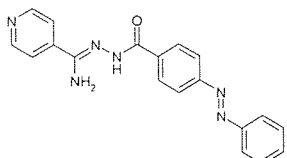
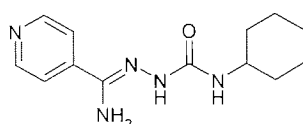
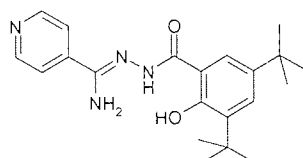


Table 3.1.2: Four types of synthesised carboxamidrazones.

Type I compounds were decided upon following the previous student Khan's (2007) work. She prepared a set of pyridylcarboxamidrazone-4-N oxides. They were tested against both *M. fortuitum* and *M. tuberculosis* H37Rv. The testing results for *M. tuberculosis* H37Rv showed the highest percentage of inhibition to be only 88%. The maximum MIC for *M. fortuitum* was only 8-16 µg/mL. Therefore, **Type I** compounds were decided to be the next target for synthesis. Some of the **Type II** compounds were based on the primary screening result from robotic synthesis. The previous robotic synthesised compounds only achieved a maximum of 90% purity. There might be some aldehyde left which will be toxic. (Coleman *et. al.* 1999) Therefore, clean

compounds should be synthesised to make sure there is no starting material left before testing against micro bacteria.

Type III, Type IV and part of **Type II** from **219** to **240** compounds were synthesised by Chong (MPham 4th year project student, 2007 Aston University) under my supervision. These compounds were selected upon after the testing results against *M. tuberculosis* H37Rv came back from TAACF (<http://www.taacf.org/>). Several active compounds were found from TAACF testing. They are shown in **Table 3.1.3**

Comp.	Structure	% inh.	MIC
1		100	<6.25
2		100	<6.25
3		100	<6.25
4		99	<6.25
5		97	<6.25
6		95	<6.25
7		94	<6.25
8		94	<6.25

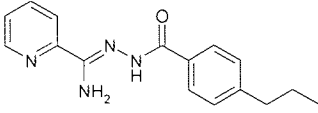
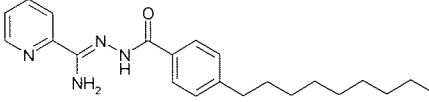
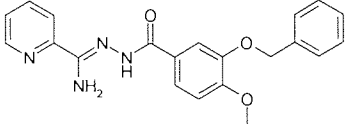
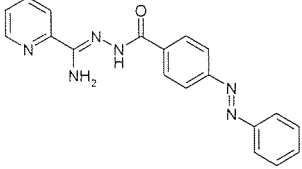
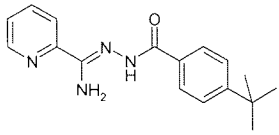
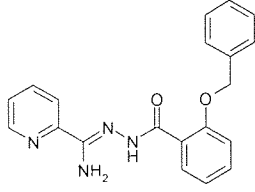
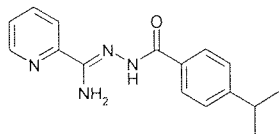
9		93	<6.25
10		93	<6.25
11		93	<6.25
12		92	<6.25
13		91	<6.25
14		90	<6.25
15		90	<6.25

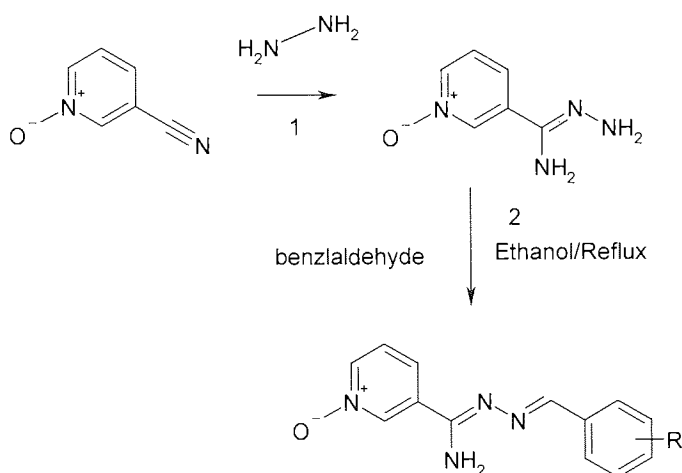
Table 3.1.3: Previous anti-TB testing results.

In **Table 3.1.3**, out of these 15 active compounds, **2, 5, 9, 10, 13, 15** possess an alkyl chain attached to the *p*-position of benzyl amide pyridine- carboxamidazone. Those were the leading compounds for **Type III** and removing the benzene ring will give the **Type IV** compounds

3.2. Results and Discussion

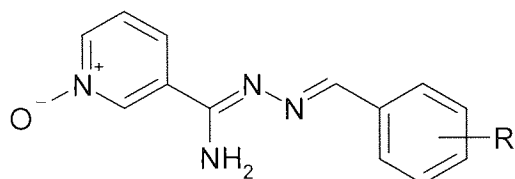
3.2.1 Type I compounds

3.2.1.1 Chemistry of compounds of Type I



Scheme 3.2.1: Synthesis scheme for benzylidene-pyridine-3-carboxamidrazone-1-N-oxide (1: Case 1965; 2: Rathbone *et.al.* 2006)

The **Type I** compounds were prepared in a two step reaction as shown in **Scheme 3.2.1**. Nearly all the compounds precipitated out after reaction mixture on cooling down and were found to be clean after filtration. All the compounds were found to have ^1H NMR spectrum, MS, IR spectrum and melting point consistent with the proposed structure. The structure of final compounds, their yield and melting point are given in **Table 3.2.1**.



Comp.	R	Yield %	Melting point °C
101	Methyl	84	230.1-232.4
102	Ethyl	70	213.3-215.5
103	Isopropyl	78	214.7-216.8
104	4- <i>tert</i> -Butyl	61	230.9-233.8
105	3,5-Di- <i>tert</i> -butyl-4-hydroxy	14	204.2-206.3
106	3,5-Di- <i>tert</i> -butyl-2-hydroxy	80	247.2-248.6
107	4-Chloro-phenylsulfanyl	57	217.6-220.7
108	3-Benzyloxy-4-methoxy	61	211.2-213.8
109	4-Benzyloxy-3-methoxy	57	198.5-201.4
110	5- <i>tert</i> -Butyl-2-hydroxy	52	206.8-208.4
111	3- <i>tert</i> -Butyl-2-hydroxy	79	239.4-241.5

Table 3.2.1: Structure with the yield and melting point of **Type I** compounds

The yield of **105** is much lower than the others. This was due to that it had to be recrystallized from ethyl acetate-petrol (60-80 °C boiling range). In the melting point data in **Table 1**, it was noticeable that from **101** to **104**, the molecular weight is increasing and their melting point should relate to their molecular weight. However, **101**, which has the lowest molecular weight has nearly the highest melting point. This may be due to the crystal packing of the molecules. Although **101** has the lowest molecular weight, it may pack well in the solid state because there is a minimal interference from the alkyl group to the intermolecular hydrogen bonding. The bulkier alkyl substituents tend to disrupt this hydrogen bonding and this results in lower melting points.

3.2.1.2 Microbiology screening of compounds of Type I

These compounds were initially tested for their biological activity against *E. coli* and *S. aureus*. This was done by S. Chauhan. (Biology placement student 2005/2006) **Table 3.2.2** shows the testing results

Compound NO.	<i>E. coli</i> ATCC 27325	<i>S. aureus</i> NCTC 6571
101	☑	☑
102	☑	☑
103	☑	☑
104	☑	☑
105	☒ 1.0 mm zone diameter	☒ 1.2 mm zone diameter
106	☑	☒ 0.8 mm zone diameter
107	☑	☑
108	☑	☑
109	☑	☑
110	☑	☑
111	☑	☑

Table 3.2.2: Biological testing results against *E. coli* and *S. aureus*.

Key: ☒ = the compound inhibited the *E. coli* or *S. aureus* growing and produced a clear circular zone

☑ = The compound was unable to inhibit the *E. coli* and *S. aureus* growth

From the **Table 3.2.2**, it is clear that only compounds **105** and **106** pass the initial screening. These two structures are quite similar to the leading structure of *N*¹-[3,5-di-(*tert*-butyl)-2-hydroxybenzylidene]-pyridine-2-carboxamidrazone. After initial screening, the MIC of compounds **105** and **106** was obtained using the same strain. The result shows that MIC of **106** was greater than 256 µg/mL. Owing to its large zone of inhibition, it was expected that **105** should be the most active compound among them. However, its MIC value was only 32-64 µg/mL which is still relatively high compared to the previous leading compound which is 2-4 µg/mL (Rathbone *et al.* 2006). The closely related compounds **105**, **106**, **110** and **111** have been synthesised. The microbiological test results tell that there is no activity and **106** only showed activity like **105**. These compounds also were sent to TAACF and tested against *M. tuberculosis* H37Rv. The testing result shows in **Table 3.2.3**.

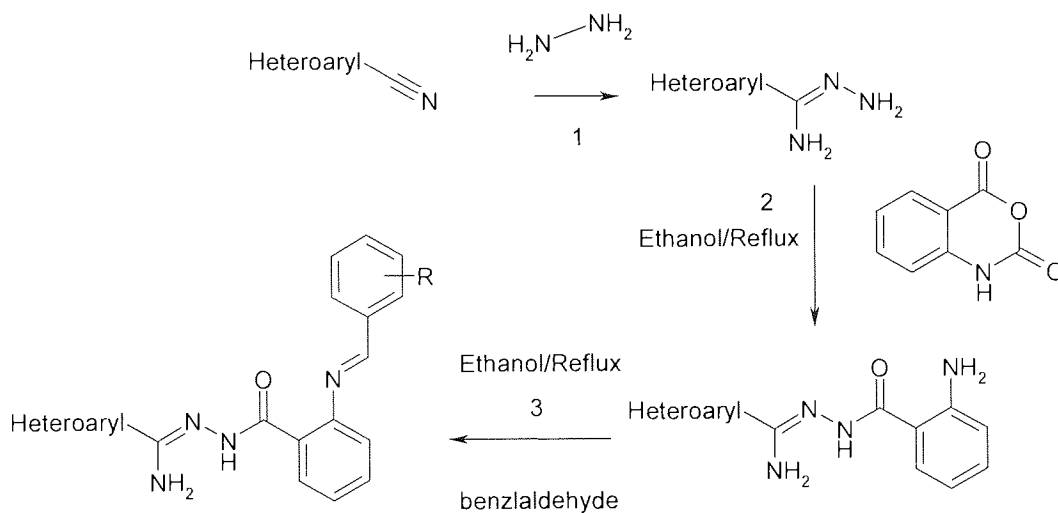
Comp.	% Inh	activity	Comp.	% Inh	activity
101	43	-	107	70	-
102	48	-	108	53	-
103	28	-	109	58	-
104	50	-	110	42	-
105	41	-	111	26	-
106	60	-			

Table 3.2.3: Testing result from TAACF against *M. tuberculosis* H37Rv.

As the TAACF considers over 90% inhibition to be active, all the above compounds are considered to be inactive against *M. tuberculosis* H37Rv. The highest one, **107**, only has 70% inhibition rate. Compared these with previous studies (Khan, 2007, Rathbone *et. al.* 2006 and Coleman *et. al.* 2003), this suggests that enough work has been done for the analogue of *N*¹-(3,5-di-tert-butyl-2-hydroxy)-phenyl-lidinepyridine-4-carbox- amidrazones.

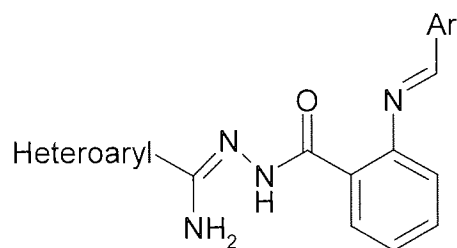
3.2.2. Type II compounds

3.2.2.1 Chemistry of compounds of Type II

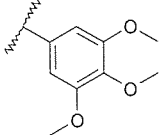
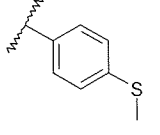
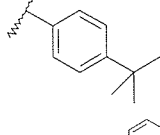
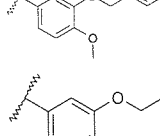
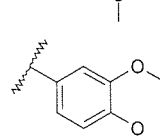
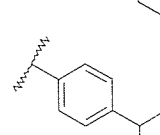
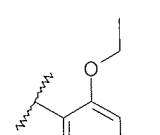
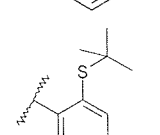
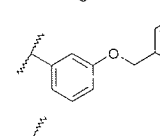
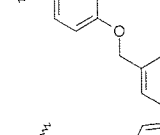
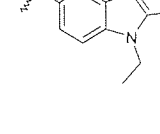



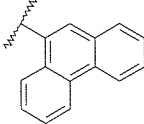
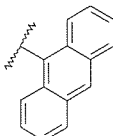
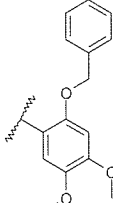
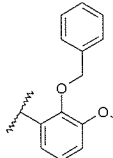
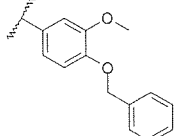
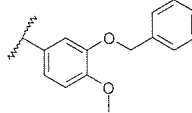
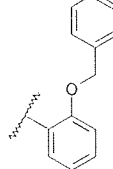
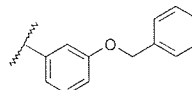
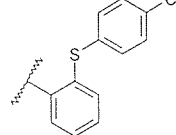
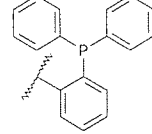
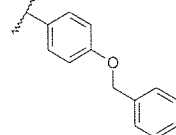
Scheme 3.2.2: Synthesis scheme of benzylidene-anthroyl-pyridine-carboxamidrazone. (1: Case 1965, 2: Clark and Wagner 1944, 3: Rathbone *et. al.* 2006)

The **Type II** compounds were prepared in a three step reaction as shown in **Scheme 3.2.2**. All the compounds precipitated out after cooling the reaction mixture and were found to be clean after filtration. All the compounds were found to have ^1H NMR spectrum, MS, IR spectrum and melting point consistent with the proposed structure. The structure of final compounds, their yield and melting point are given in **Table 3.2.4**.



Comp.	Heteroaryl	Ar	Yield %	Melting point °C
201	4-Pyridyl		70	191.9-195.9
202	4-Pyridyl		84	211.6-212.3
203	4-Pyridyl		67	225.3-226.7
204	4-Pyridyl		62	208.8-209.8
205	4-Pyridyl		91	207.2-207.6
206	2-Pyrazinyl		86	205.1-205.9
207	2-Pyrazinyl		81	215.0-216.5
208	2-Pyrazinyl		91	210.5-211.2

209	2-Pyridyl		68	200.0-202.7
210	2-Pyridyl		85	220.2-221.7
211	2-Pyridyl		71	218.2-220.1
212	2-Pyridyl		81	191.1-194.0
213	2-Pyridyl		88	222.2-225.8
214	2-Pyridyl		93	216.7-219.3
215	2-Pyridyl		61	211.0-213.6
216	2-Pyridyl		69	177.3-178.8
217	2-Pyridyl		77	222.3-225.1
218	2-Pyridyl		77	170.2-171.0
219	2-Pyridyl		81	239.8-241.4
220	2-Pyridyl		72	260.9-263.8

221	2-Pyridyl		87	250.5-252.0
222	2-Pyridyl		72	248.7-250.4
223	2-Pyridyl		71	200.7-202.2
224	2-Pyridyl		80	207.7-211.1
225	2-Pyridyl		70	201.8-203.7
226	2-Pyridyl		76	190.1-190.8
227	2-Pyridyl		80	225.2-227.2
228	2-Pyridyl		49	170.0-172.1
229	2-Pyridyl		64	200.9-207.5
230	2-Pyridyl		82	224.9-226.3
231	4-Pyridyl		79	222.8-226.5

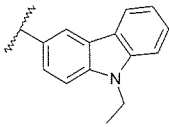
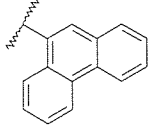
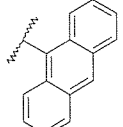
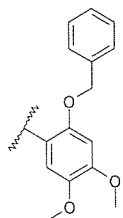
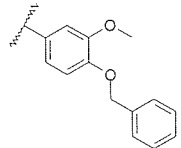
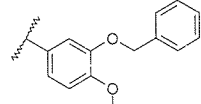
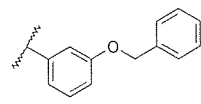
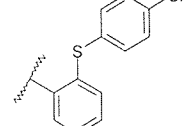
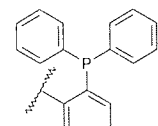
232	4- Pyridyl		68	245.0-246.2
233	4- Pyridyl		23	245.9-247.0
234	4- Pyridyl		55	249.2-250.6
235	4- Pyridyl		34	211.9-213.3
236	4- Pyridyl		80	145.7-148.6
237	4- Pyridyl		48	192.1-194.3
238	4- Pyridyl		31	201.7-202.4
239	4- Pyridyl		77	193.4-194.8
240	4- Pyridyl		76	217.1-220.1

Table 3.2.4: Structure with the yield and melting point of **Type II** compounds

If the yield information in **Table 3.2.1** and **Table 3.2.4** is compared, it can be seen that **Type II** compounds generally have a higher yield and the highest one was obtained in 93% yield. This may due to their different solubility. This can be noticed during the reaction. Compounds of **Type II** generally have a lower solubility in the

reaction solvent. Upon cooling the reaction mixture, therefore, the extent of precipitation of **type II** compounds was greater than it for **Type I**.

Although **Type II** compounds had a higher yield, if the yields information of compounds **251** to **272** are compared, the yield of compounds **251** to **262** are higher than compounds **263** to **272**. The reason for this should due to their solubility and ease of reaction. According to Rathbone *et. al.* (2006), carboxamidrazones reacting with aldehyde only takes two hours to be completed. However, the compounds from **251** to **272** took more time to react. In the chemistry laboratory, the compounds **251**, **252**, **253** and **263**, **264**, **265** were synthesised using the same condition. And the reaction took overnight. After this, ^1H NMR spectrum shows that **251**, **252** and **253** were complete. However, the solid products for **263**, **264** and **265** still contained most of the starting material which means that they need a longer time to react. In the next batch the reaction time for **263**, **264** and **265** were increased to 48 hours. This time, products formed whilst the yield was lower. This difference might due to the different solubility of the starting materials. When the chemical data of starting material was checked, it was found that under the same conditions, the yield for pyridine-2-carboxamidrazone was less than the yield for pyridine-4-carboxamidrazone which means that 4-carboxamidrazone compounds are less soluble in the ethanol. At the same time, during the reaction, it can be seen that the reaction mixture was never dissolved complete. This should be the cause of longer time reaction.

The ^1H NMR spectrum data for **Type II** compounds is very complicated because there are many aryl rings. Most of the peaks were concentrated between 6 and 8.5 ppm and overlapped with each other. At the same time, identifying the $\text{CH}=\text{N}$ peak is essential because it indicates that the required product has been formed. Therefore, a D_2O shake of **205** was conducted to identify all the peaks. The **Figure 3.2.1** shows the structure and the ^1H NMR spectrum of **205** with and without D_2O shake in DMSO.

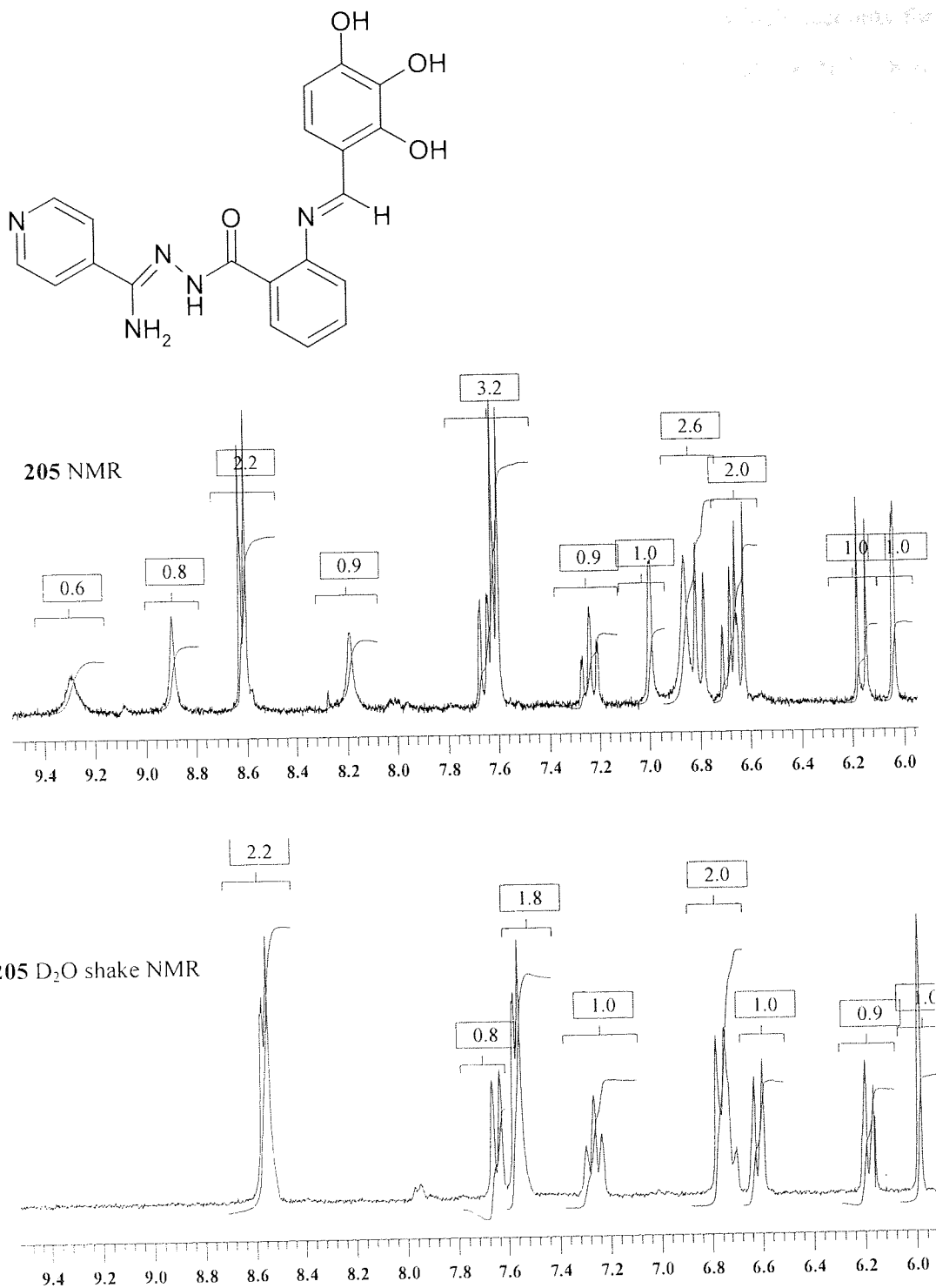


Figure 3.2.1: Comparison of **205** with and without D₂O shake in DMSO.

All the OH and NH signals do not show in a D₂O treated sample ¹H NMR spectrum. When these two ¹H NMR spectra in **Figure 3.2.1** are compared, it is easily to notice that first of all, from 8 to 9.6 ppm, three broad peaks disappeared. These three peaks

should be from the phenyl OH. At around 7.0 ppm, one singlet which accounts for one proton disappeared in D₂O shake. This indicates this peak should be NH. Then, between 6.5 and 6.9 ppm, there are two protons missing which means that the NH₂ lies in this region. The singlet around 6.0 ppm is the CH=N-Ar.

It was also noticed that the ¹H NMR spectrum chemical shift of CH=N-Ar in **Type I** compounds lies around 8.4 region. However, in most of the **Type II** compounds, ¹H NMR spectrum chemical shift of CH=N-Ar lies around 6.0 region. By comparing those structures, the major difference is the anthroyl moiety. The low chemical shift of Type II compounds should be due to the conjugation with benzyl ring on both sides while in **Type I** compounds, CH only receive electrons from one side and the other side is a strong electron withdrawing substituent.

Theoretically, the peak representing CH=N-Ar should be a singlet. This signal in all of the **Type I** and some of the **Type II** compounds was a singlet. One example is shown in **Figure 3.2.3** and the peak in 6.34 ppm represents CH=N-Ar in compound **208**. However, some of them show a small doublet such as the peaks in compound **205**, **206**, **207**, **210**, **211**, **216**, **218** and **506**. **Figure 3.2.2** is the partial ¹H NMR spectrum of compound **206** and the peak which represents CH=N-Ar in 6.26 ppm is a small doublet with a measured coupling constant of 1.22 Hz.

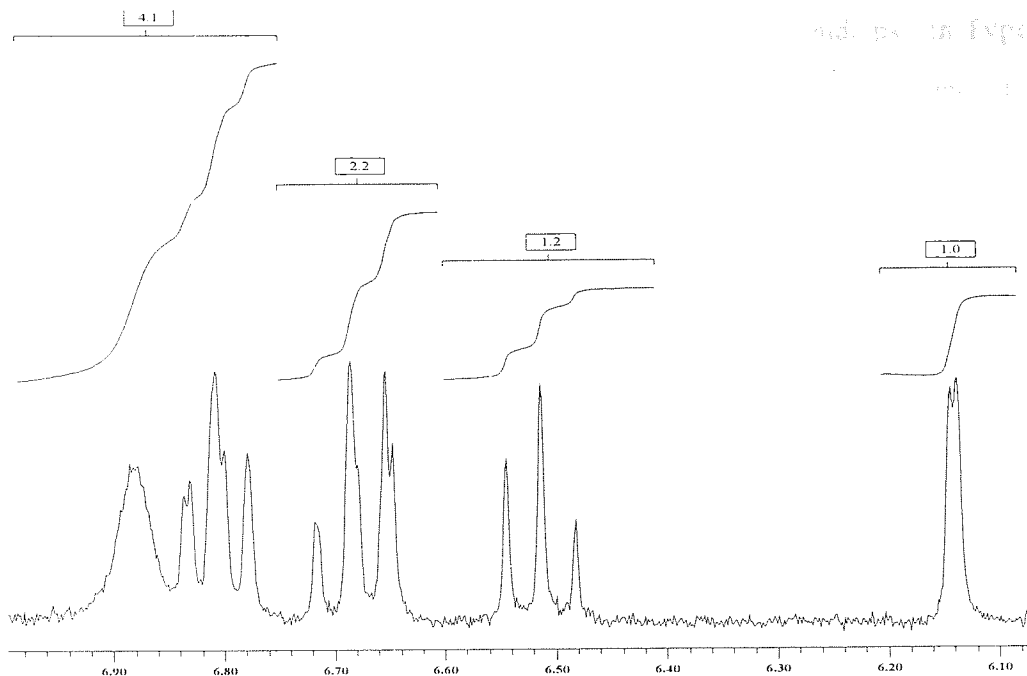


Figure 3.2.2: Partial H^1 NMR spectrum of 206

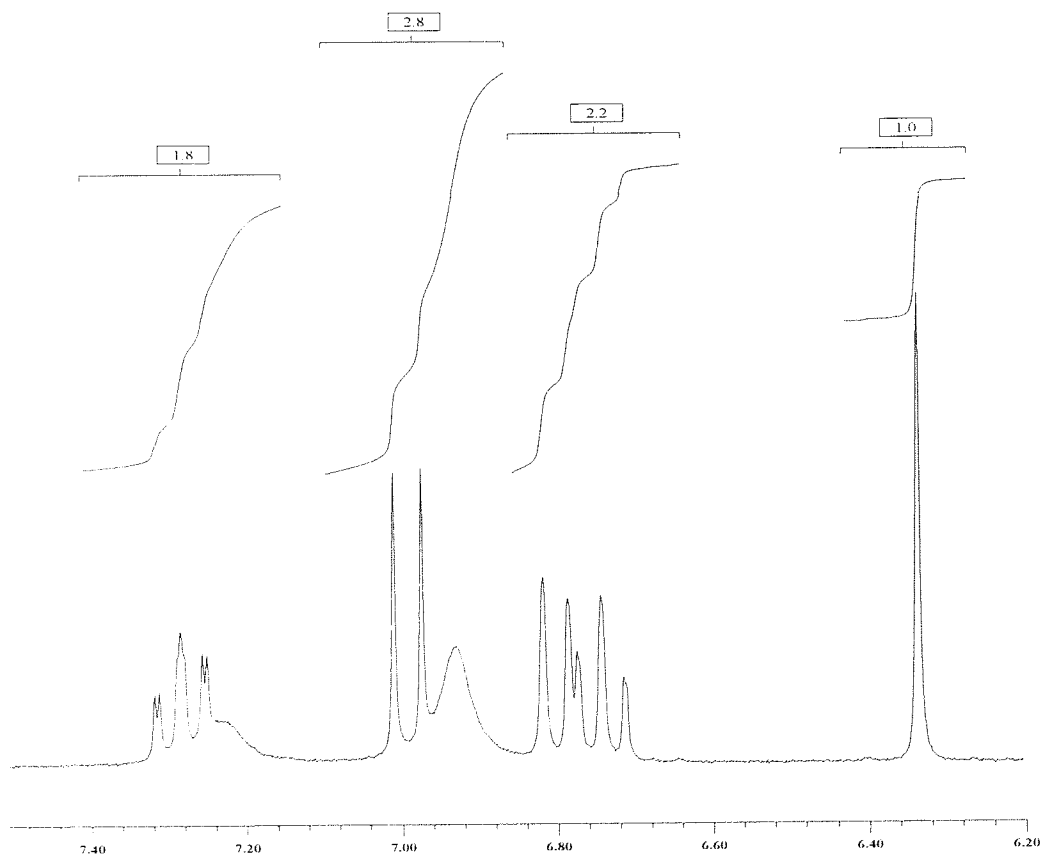


Figure 3.2.3: Partial H^1 NMR spectrum of 208.

This phenomenon cannot be explained by the delocalisation of amide part in **Type II** compounds. If assuming the CH was affected by keto-enol tautomerisation of the amide group, NH of the amide and the CH=N-Ar on anthroyl group should also be affected and should show as two singlets in NMR spectra. However, there was no such instance and all the ^1H NMR spectrum of NH was a singlet. If all the structures are compared, a general trend can be found that when the ^1H NMR spectrum of CH=N-Ar appears as a small doublet, the side groups on the benzene ring near the CH are all small, such as OH, one Cl or a small group on *p*-position. This may due to the freedom/restriction of rotation effects (Bible 1965).

All the APCI-MS +ve mass spectra of the **Type II** compounds showed the $(\text{M}+1)^+$ peak and a $(\text{M}-17)^+$ peak which indicates that the compounds may lose a water molecule. This may due be to the presence of carbonyl group and NH_2 , leading to the formation of a five membered-ring shown in **Figure 3.2.4**. Then, a water molecule could be lost in MS to give a conjugated ion. This might be the cause of $(\text{M}-17)^+$ peak.

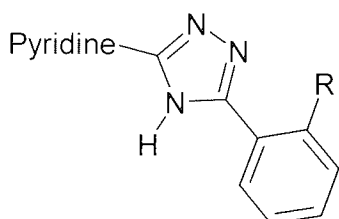


Figure 3.2.4: Possible explanation of the structure associated with $(\text{M}-17)^+$ signal.

3.2.2.2 Microbiology screening of compounds of Type II

Type II compounds **201** to **218** were tested against *E. coli* and *S. aureus*. This was done by S. Chauhan. (Biology placement student 2005/2006 Aston University) and **Table 3.2.5** shows the testing results.

Compound name	<i>E. coli</i> ATCC 27325	<i>S. aureus</i> NCTC 6571
201	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
202	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
203	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
204	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
205	<input checked="" type="checkbox"/> 1.0 cm zone diameter {precipitated within agar}	<input checked="" type="checkbox"/> 1.1 cm zone diameter {precipitated within agar}
206	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
207	<input checked="" type="checkbox"/> 3.2 cm zone diameter {overgrown zone however}	<input checked="" type="checkbox"/>
208	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
209	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
210	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
211	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
212	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
213	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
214	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
215	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
216	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
217	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
218	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Table 3.2.5: Biological testing result against *E. coli* and *S. aureus*.

Key: = the compound inhibited the *E. coli* or *S. aureus* growing and produced a clear circular zone

= The compound was unable to inhibit the *E. coli* and *S. aureus* growth

Table 3.2.5 shows that only **205** and **207** have a zone of inhibition. The size of the inhibition zone only indicates that how easily these compounds can penetrate the agar plate. After this, the MIC testing of all these compounds was conducted. However, the observed antibacterial activities were weak. Only **205** shows a MIC of 128-256 μ g/mL. All these compounds were sent to TAACF and tested against *M. tuberculosis* H37Rv. The result showed that all the compounds were inactive with an inhibition of less than 90%. However, interestingly, **212**, **214** and **218**'s percentage of inhibition were close to 90%. They are 82%, 73% and 88% respectively. By comparing their structure, they all have two or more benzene rings in the benzylidene part. Therefore, compounds **219** to **240** were designed to have more benzene rings in the benzylidene part to see if this can improve their activity.

The compounds from **219** to **240** were tested against *E. coli* and *S. aureus* and no zone of inhibition was found. They were also tested against M-TB H37Rv by TAACF. Six of them showed weak activity with a lowest IC₉₀ of 23.51 µg/mL. some results is shown in **Table 3.2.6**.

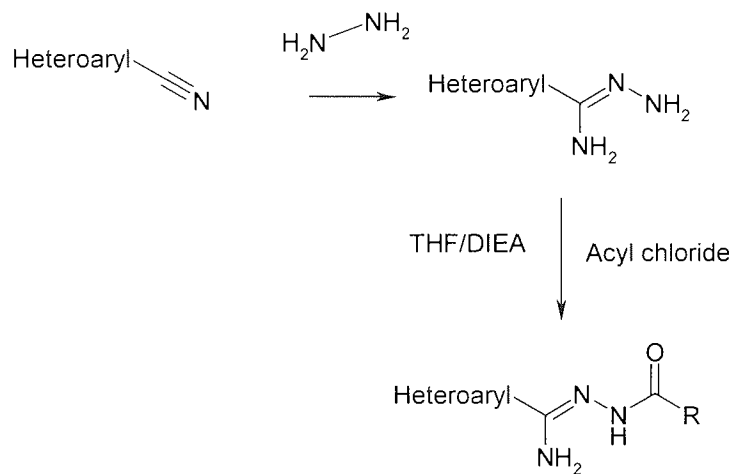
Compounds	Activity	IC ₅₀ µg/mL	IC ₉₀ µg/mL
220	Weakly Active	15.935	36.319
225	Weakly Active	26.017	42.484
230	Weakly Active	21.309	23.510
232	Weakly Active	31.505	50.546
239	Weakly Active	29.576	42.437
240	Weakly Active	17.210	24.227

Table 3.2.6: Testing result of **Type II** compounds against M-TB H37Rv

This result was quite similar to the previous testing result of compounds **201** to **218**. The most active compounds **230** and **240** have a lot of benzene rings in their structure. These results mean that simply increasing the number of aromatic ring is not appropriate. The modification on other parts of the molecule may be required.

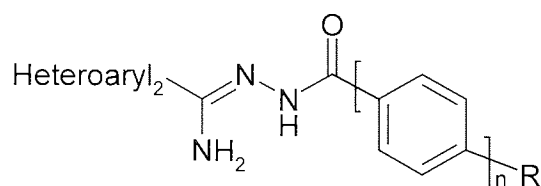
3.2.3 Type III and Type IV compounds

3.2.3.1 Chemistry of compounds of Type III and Type IV



Scheme 3.2.3: Synthesis scheme of pyridine-carboxamidrazone-amides

The **Type III** and **Type IV** compounds were prepared in a two step reaction as shown in **Scheme 3.2.3**. In the first step, the compounds precipitated out after the reaction mixture cooled down and were found to be clean after filtration. The solid products from the second step were all mixtures. Therefore, recrystallization were conducted to purify the target compounds. All the compounds were found to have ¹H NMR spectrum, MS, IR spectrum and melting point consistent with the proposed structure. The structures of final compounds, their yield and melting point are given in **Table 3.2.7**.



Comp.	Heteroaryl ₂	n	R	Yield %	Melting point °C
301	2-Pyridyl	1	Heptyl	30	215.3-216.1
302	2-Pyridyl	1	Hexyl	46	218.7-220.6
303	2-Pyridyl	1	Pentyl	45	234.1-235.9
304	2-Pyridyl	1	Decyl	32	201.2-203.2
305	2-Pyridyl	0	4-Pentenoyl	9	191.5-193.9
306	2-Pyridyl	0	10-Undecenoyl	9	120.9-124.8
307	2-Pyridyl	0	Myristoyl	33	123.0-125.3
308	2-Pyridyl	0	Palmitoyl	40	75.9-80.4
309	2-Pyridyl	0	Stearoyl	23	118.6-121.4
310	4- Pyridyl	1	Heptyl	27	182.2-182.8
311	4- Pyridyl	1	Hexyl	23	187.7-191.0
312	4- Pyridyl	1	Pentyl	38	201.2-201.8
313	4- Pyridyl	0	10-Undecenoyl	4	113.0-114.2
314	4- Pyridyl	0	Lauroyl	31	121.2-124.4
315	4- Pyridyl	0	Myristoyl	40	116.8-117.5

Table 3.2.7: Structure with the yield and melting point of **Type III** and **Type IV** compounds

First of all, it can be seen that all the yields are quite low in **Table 3.2.7**. This is because after the reaction completed and the solid products were collected, they were found to contain more than one spot based on TLC. Therefore, recrystallisations were conducted to remove all the impurities. Three different solvents including ethanol, ethyl acetate and toluene were used here for recrystallisation. Most of the products were recrystallised using ethanol acetate. **306** was cleaned using toluene and **303**, **304**, **311** and **312** were recrystallised from ethanol. It took a longer reaction time for the 4-pyridyl compounds to form which is quite similar to the situation in **Type II** compounds. It was also noticed starting material of pyridine-4-carboxamidrazones was less soluble cold THF than the pyridine-2-carboxamidrazones. Therefore, the solubility of starting material might have affected the time of reaction.

In the NMR data, the NH is formed in the final step and by identifying this peak in the ^1H NMR spectrum, it can be confirmed that the reaction was successful. In fact, all the NH peaks in ^1H NMR spectrum were found at around 10 ppm. It is interesting to notice that if the compound has a benzene ring between carbonyl group and alkyl chain, the ^1H NMR spectrum of NH was a singlet. If the compound does not have benzene ring, ^1H NMR signal of NH was two singlets and overall integral account for one proton. In this case, these two singlets represent the keto and enol form of the compound. At the same time, it means that the interconverting rate is less than the frequency of NMR machine. Otherwise, these two singlets will become one singlet if the rate is similar or greater than the frequency of NMR machine.

In **type III** and **Type IV** compounds, the length of alkyl chain will affect the melting point. Normally, longer chain length means larger molecular weight. Therefore, greater *Van der Waals* effect should apply to those compounds resulting in higher melting points. However, if the melting point of **303**, **302**, **301** and **304** are compared, there was a clear trend that when the alkyl chain increased from C5 to C10, the melting point decreased from around 235 °C to 200 °C. This phenomenon can be explained by considering conformational structure of the molecule. If the alkyl chain is long, it tends to fold into a spherical rather than linear shape. In this case, the surface area per carbon chain is decreased. This results in a lower *Van der Waals* interaction as well as low melting point. (McMurry 2002)

3.2.3.2 Microbiology screening of compounds of Type III and Type IV

All these compounds were tested against *E. coli* and *S. aureus* and no zone of inhibition was found. They were also tested against M-TB H37Rv by TAACF. Five of them were found to be active and five of them were found to be weakly active. Their testing data are shown in **Table 3.2.8**.

Active compounds			Weekly active compounds		
Comp.	IC ₅₀ µg/mL	IC ₉₀ µg/mL	Comp.	IC ₅₀ µg/mL	IC ₉₀ µg/mL
302	0.994	1.329	310	7.556	11.705
303	0.887	1.386	311	11.972	13.822
301	1.410	1.527	312	17.712	24.892
304	5.466	5.986	313	25.239	28.273
306	4.425	7.382	309	42.441	46.752

Table 3.2.8: Testing result of **Type III** and **Type IV** compounds against M-TB H37Rv

Those five active compounds share the similar feature of a 2-pyridyl group, a benzene ring connected to the carbonyl group and a long alkyl chain connected to the fourth position of the benzene ring. These compounds and the compounds **5**, **9**, **10** in **Table 3.1.3** belong to the same group. The cytotoxicity of these active compounds was tested by TAACF as well and the result is shown in **Table 3.2.9**.

Compounds	IC ₉₀ µg/mL	CC ₅₀ µg/mL	SI
302	1.329	0.428	0.322
303	1.386	1.114	0.822
301	1.527	1.003	0.656
304	5.986	1.363	0.227
306	7.382	0.565	0.076

Table 3.2.9: Cytotoxicity assay of **Type III** and **Type IV** compounds. (CC₅₀ gives the measure of mammalian cell toxicity. The selected Index SI shows the ratio of CC₅₀ and IC₉₀. SI ≥ 10 would be considered for the further screening.)

The data shown in **Table 3.2.9** means that those compounds are all too cytotoxic to use as drugs. Therefore, a further modification on **Type III** compounds would be necessary.

3.3 Conclusion and Recommendation

Overall, all four types of compounds were synthesized and their key experimental data were analysed. Certain trends were found in the same types of compounds. All eleven **Type I** compounds were tested against *E. coli*, *S. aureus* and M-TB H37Rv. Only **105** and **106** had weak activity against *E. coli* and *S. aureus* with a minimum MIC of 32-64 µg/mL. All of those eleven compounds were found inactive against M-TB H37Rv with a range of inhibition from 26% to 70%.

Similarly, **Type II** compounds were also tested against *E. coli*, *S. aureus* and M-TB H37Rv. Although compounds **205** and **207** showed zones of inhibition against *E. coli*, *S. aureus*, the MIC data showed that only **205** had a measurable MIC of 128-256 µg/mL. All of the other MICs were greater than 256 µg/mL. The testing data from TAACF shows that six of the compounds were weakly active against M-TB H37Rv. The lowest IC₉₀ was found with compounds **230** (23.51 µg/mL). This type of the compound may need further modification to increase their activities.

In **Type III** and **Type IV** compounds, five of them showed strong inhibition against M-TB H37Rv and five of them showed weak activity against M-TB H37Rv with a lowest IC₉₀ value of 1.329 µg/mL. However, the further cytotoxicity assay shows that all those strongly active compounds were too cytotoxic to be used as drug molecules. Therefore, a further modification to reduce the cytotoxicity is required.

3.4 Experimental

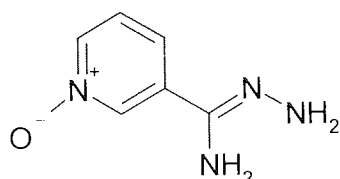
Instrumentation

Proton NMR spectra were obtained on a Bruker AC 250 instrument operation at 250MHz as solutions in d_6 -DMSO or $CDCl_3$ and referenced from $\delta_{DMSO}=2.50$ ppm or $\delta_{CDCl_3}=7.26$ ppm unless otherwise stated. Infrared spectra were recorded as KBr discs or NaCl discs on a Mattson 3000 FTIR spectrophotometer. Atmospheric pressure chemical ionisation mass spectrometry (APCI-MS) was carried out on a Hewlett-Packard 5989B quadrupole instrument connected to an electrospray 59987A unit with an APCI accessory and automatic injection using a Hewlett-Packard 1100 series autosampler. Electrospray +ve was conducted on LCT Premier Electrospray manufactured by WATERS. Melting points were obtained using a Reichert-Jung Thermo Galen hot stage microscope and are corrected.

Preparation of 3-Pyridyl Carboxamidrazone-1-N-oxide [S1]

Nicotinonitrile-1-oxide (9.72 g, 81 mmol) was suspended in ethanol (30 mL), treated with hydrazine monohydrate (20 mL) and stirred at ambient temperature for seven days. The solid material was collected by filtration, washed with ethanol (3×10 mL) and dried under vacuum to give the product as a white powder,

Structure



Physical state	White powder
Yield	5.79 g, 38 mmol, 47%
1H NMR (D6-DMSO; $\delta_{DMSO}=2.50$ ppm)	5.31 (bs, 2H, NH_2), 5.82 (bs, 2H, NH_2), 7.40 (dd, 1H, $J=8.1$, 6.4 Hz, Ar 5-H), 7.64 (ddd, 1H, $J=8.1$, 1.0, 0.9 Hz, Ar 4-H), 8.14 (ddd, 1H, $J=6.3$, 1.0, 0.9 Hz, Ar 6-H), 8.43 (m, 1H, Ar 2-H)
IR (KBr)	3408, 3284, 3174, 1666, 1589, 1564, 1489, 1428, 1390, 1309, 1231 (N-O), 1171, 1121 cm^{-1} .
MS (APCI +ve)	$m/z=153$ (M+H) $^+$, 136 (M-O) $^+$
Melting point	135.5-138.6 $^{\circ}C$ (decomposes)

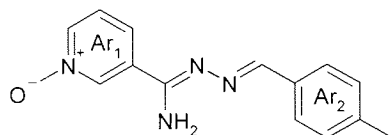
General procedure for prepares Type 1 and Type II compounds.

The starting material for Type II compounds were synthesised by the previous students and a postdoctoral work. The purity of the starting material was checked by NMR and found to be pure.

Pyridyl carboxamidrazone (0.5g) was suspended in ethanol (10mL) with aldehyde (1.1 mole equiv.) and stirred vigorously under reflux for at least 2 hours. Solid products appeared after cooling. The solid was collected by filtration, washed with ethanol (3 × 10mL) and dried under vacuum to give the final products.

***N*¹-(4-Methyl)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [101]**

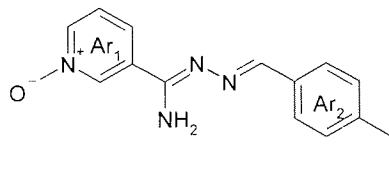
Structure



Physical state	Yellow powder
Yield	84%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	2.35 (s, 3H, CH ₃), 7.20 (bs, 2H, NH ₂), 7.25 (d, 2H, J=8.2 Hz, Ar ₂ 3-H, Ar ₂ 5-H), 7.50 (dd, 1H, J=7.9, 6.4, Hz, Ar ₁ 5-H), 7.81 (d, 2H, J=7.9 Hz, Ar ₂ 2-H, Ar ₂ 6-H), 7.84 (s, 1H, Ar ₁ 4-H), 8.31 (s, 1H, Ar ₁ 6-H), 8.42 (s, 1H, N=CH), 8.67 (s, 1H Ar ₁ 2-H)
IR (KBr)	3443, 3292, 3210, 3176, 3137, 3061, 3048, 2913, 1628, 1598, 1570, 1491, 1428, 1409, 1336, 1310, 1249 (N-O), 1177 (C-N) cm ⁻¹ .
MS (APCI +ve)	m/z=255 (M+H) ⁺ , 237 (M-OH) ⁺
Melting point	230.1-232.4 °C

***N*¹-(4-Ethyl)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [102]**

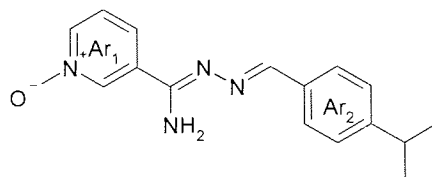
Structure



Physical state	Yellow crystalline solid
Yield	70%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.02 (t, 3H, J=7.6 Hz CH ₃), 2.65 (q, 2H, J=7.6 Hz, CH ₂) 7.20 (bs, 2H, NH ₂), 7.28 (d, 2H, J=7.9 Hz, Ar ₂ 3-H, Ar ₂ 5-H), 7.50 (dd, 1H, J=7.9, 6.7, Hz, Ar ₁ 5-H), 7.83 (d, 2H, J=8.2 Hz, Ar ₂ 2-H, Ar ₂ 6-H), 7.84 (m, 1H, Ar ₁ 4-H), 8.31 (m, 1H, Ar ₁ 6-H), 8.42 (s, 1H, N=CH), 8.68 (m, 1H Ar ₁ 2-H)
IR (KBr)	3405, 3266, 3128, 3053, 2960, 2928, 2868, 1628, 1599, 1563, 1534, 1490, 1433, 1337, 1309, 1240 (N-O), 1227, 1178(C-N) cm ⁻¹ .
MS (APCI +ve)	m/z=269 (M+H) ⁺ , 251 (M-OH) ⁺
Melting point	213.3-215.5 °C

***N*¹-(4-Isopropyl)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [103]**

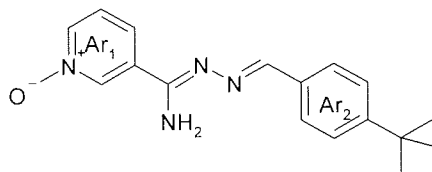
Structure



Physical state	Yellow crystalline solid
Yield	78%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.23 (d, 6H, J=6.7 Hz Me ₂), 2.93 (sept, 1H CHMe ₂) 7.19 (bs, 2H, NH ₂), 7.31 (d, 2H, J=8.2 Hz, Ar ₂ 3-H, Ar ₂ 5-H), 7.50 (dd, 1H, J=8.1, 6.5, Hz, Ar ₁ 5-H), 7.83 (d, 2H, J=8.2 Hz, Ar ₂ 2-H, Ar ₂ 6-H), 7.84 (m, 1H, Ar ₁ 4-H), 8.31 (m, 1H, Ar ₁ 6-H), 8.42 (s, 1H, N=CH), 8.68 (m, 1H Ar ₁ 2-H)
IR (KBr)	3423, 3267, 3205, 3170, 3136, 3059, 3023, 2956, 2866, 1627, 1603, 1565, 1536, 1490, 1431, 1336, 1305, 1240 (N-O), 1225, 1167(C-N) cm ⁻¹ .
MS (APCI +ve)	m/z=283 (M+H) ⁺ , 265 (M-OH) ⁺
Melting point	214.7-216.8 °C

N¹-(4-*tert*-Butyl)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [104]

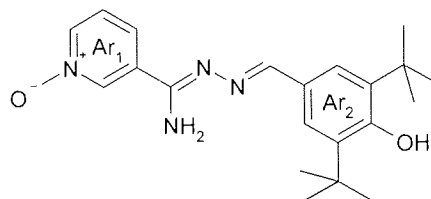
Structure



Physical state	Yellow crystalline solid
Yield	61%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.30 (s, 9H, Me ₃), 7.23 (bs, 2H, NH ₂), 7.46 (d, 2H, J=8.4 Hz, Ar ₂ 3-H, Ar ₂ 5-H), 7.51 (m, 1H, Ar ₁ 5-H), 7.81-7.87 (overlapping m, 3H, Ar ₂ 2-H, Ar ₂ 6-H, Ar ₁ 4-H), 8.32 (m, 1H, Ar ₁ 6-H), 8.42 (s, 1H, N=CH), 8.68 (t, 1H J=1.4 Hz, Ar ₁ 2-H)
IR (KBr)	3412, 3265, 3136, 3061, 2958, 2902, 2865, 1629, 1603, 1560, 1538, 1488, 1430, 1408, 1340, 1238 (N-O), 1225, 1166(C-N), 1111 cm ⁻¹ .
MS (APCI +ve)	m/z=297 (M+H) ⁺ , 279 (M-OH) ⁺
Melting point	230.9-233.8 °C

N¹-(3, 5-di-*tert*-Butyl-4-hydroxy)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [105]

Structure

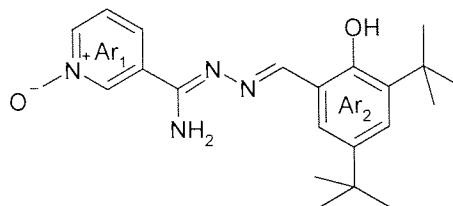


Physical state	Yellow crystalline solid
Yield	14% (recrystallised from Ethyl acetate and petrol 60-80 °C)
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.42 (s, 18H, 2×Me ₃), 7.00 (bs, 2H, NH ₂), 7.35 (bs, 1H, OH), 7.50(dd, 1H, J=8.0, 6.5, Hz, Ar ₁ 5-H), 7.62 (s, 2H, Ar ₂ 2-H, Ar ₂ 6-H) 7.83 (m, 1H, Ar ₁ 4-H), 8.34 (m, 1H, Ar ₁ 6-H), 8.36 (s, 1H, N=CH), 8.67 (m, 1H Ar ₁ 2-H)

IR (KBr)	3477, 3304, 3150, 3088, 2952, 2909, 2867, 1629, 1584, 1542, 1490, 1465, 1438, 1389, 1361, 1307, 1231 (N-O), 1230, 1202, 1173, 1131(C-N) cm^{-1} .
MS (APCI +ve)	$m/z=369$ (M+H) ⁺ , 351 (M-OH) ⁺
Melting point	202.2-204.3 °C

***N*¹-(3, 5-di-*tert*-Butyl-2-hydroxy)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [106]**

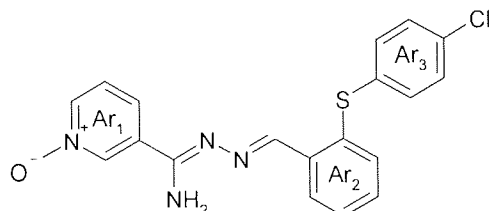
Structure



Physical state	Yellow crystalline solid
Yield	80%
¹ H NMR (D6-DMSO; δ DMSO=2.50ppm)	1.27 (s, 9H, Me ₃), 1.43 (s, 9H Me ₃), 7.25 (bs, 2H, NH ₂), 7.30 (d, 1H, J=2.4 Hz, Ar ₂ 4-H), 7.34 (d, 1H, J=2.3 Hz, Ar ₂ 6-H) 7.54 (dd, 1H, J=8.1, 6.5, Hz, Ar ₁ 5-H), 7.84 (ddd, 1H, J=8.1 1.5 1.0 Hz, Ar ₁ 4-H), 8.33 (ddd, 1H, J=6.4 1.8 0.9 Ar ₁ 6-H), 8.66 (s, 1H, N=CH), 8.67 (m, 1H Ar ₁ 2-H), 11.52 (s, 1H, OH)
IR (KBr)	3477, 3304, 3150, 3088, 2952, 2909, 2867, 1629, 1584, 1542, 1490, 1465, 1438, 1389, 1361, 1307, 1231 (N-O), 1230, 1202, 1173, 1131(C-N) cm^{-1} .
MS (APCI +ve)	$m/z=369$ (M+H) ⁺ , 351 (M-OH) ⁺
Melting point	247.2-248.6 °C

***N*¹-[2-(4-Chloroy-phenylsulfanyl)]-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [107]**

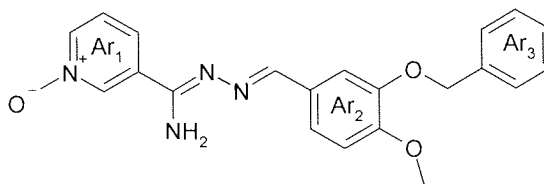
Structure



Physical state	Yellow crystalline solid
Yield	57%
¹ H NMR (D6-DMSO; δ DMSO=2.50ppm)	7.24 (d, 2H, J=8.5 Hz, Ar ₃ 3-H, Ar ₃ 5-H), 7.32-7.52 (overlapping m, 7H, Ar ₁ 5-H, Ar ₃ 2-H Ar ₃ 6-H, 2 × Ar ₂ -H, NH ₂), 7.84 (m, 1H, Ar ₁ 4-H), 8.28-8.38 (overlapping m, 2H, Ar ₁ 6-H, Ar ₂ -H), 8.66 (m, 1H Ar ₁ 2-H), 8.79 (s, 1H, N=CH),
IR (KBr)	3432, 3306, 3173, 3059, 1630, 1594, 1555, 1531, 1493, 1474, 1435, 1407, 1335, 1248 (N-O), 1229, 1166 (C-N) cm^{-1} .
MS (APCI +ve)	$m/z=383$ (M+H) ⁺ , 365 (M-OH) ⁺
Melting point	217.6-220.7 °C

***N*¹-(3-Benzyloxy-4-methoxy)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [108]**

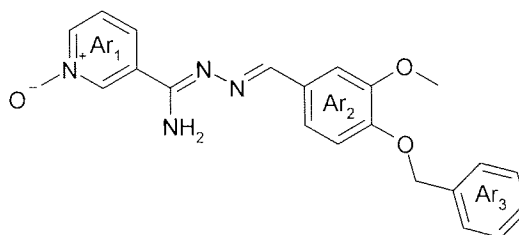
Structure



Physical state	White powder
Yield	61%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.82 (s, 3H, OMe), 5.17 (s, 2H, OCH ₂), 7.03 (d, 1H, J=8.2 Hz, Ar ₂ 5-H), 7.19 (bs, 2H, NH ₂), 7.29-7.54 (overlapping m, 7H, Ar ₁ 5-H, Ar ₂ 6-H, 5 × Ar ₃ -H), 7.81 (d, 1H, J=1.8 Hz, Ar ₂ 2-H), 7.85 (m, 1H, Ar ₁ 4-H), 8.32 (ddd, 1H, J=6.4, 1.8, 0.9 Hz, Ar ₁ 6-H), 8.36 (s, 1H, N=CH), 8.68 (m, 1H, Ar ₁ 2-H)
IR (KBr)	3441, 3237, 3106, 3072, 1618, 1596, 1546, 1508, 1486, 1428, 1385, 1351, 1323, 1262 (N-O), 1239, 1165(C-N), 1135 cm ⁻¹ .
MS (APCI +ve)	m/z=377 (M+H) ⁺ , 358 (M-OH) ⁺
Melting point	211.2-213.8 °C

***N*¹-(4-Benzyloxy-3-methoxy)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide [109]**

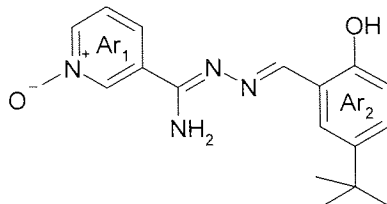
Structure



Physical state	Yellow crystalline solid
Yield	57%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.86 (s, 3H, OMe), 5.15 (s, 2H, OCH ₂), 7.09 (d, 1H, J=8.2 Hz, Ar ₂ 5-H), 7.24 (bs, 2H, NH ₂), 7.29 (dd, 1H, J=8.2, 1.5 Hz, Ar ₂ 6-H), 7.33-7.49 (overlapping m, 5H, 5 × Ar ₃ -H), 7.50 (dd, 1H, J=7.9, 6.4 Hz, Ar ₁ 5-H), 7.68 (d, 1H, J=1.5 Hz, Ar ₂ 2-H), 7.85 (m, 1H, Ar ₁ 4-H), 8.32 (m, 1H, Ar ₁ 6-H), 8.36 (s, 1H, N=CH), 8.68 (m, 1H, Ar ₁ 2-H).
IR (KBr)	3462, 3356, 3083, 3032, 1610, 1579, 1537, 1509, 1462, 1453, 1417, 1382, 1262 (N-O), 1230, 1163(C-N), 1139 cm ⁻¹ .
MS (APCI +ve)	m/z=377 (M+H) ⁺ , 359 (M-OH) ⁺
Melting point	198.5-201.4 °C

***N*¹-(2-Hydroxy-5-*tert*-butyl)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide**
[110]

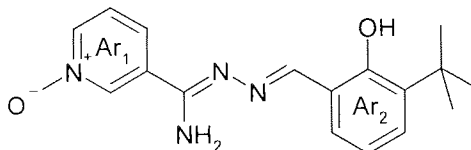
Structure



Physical state	Yellow crystalline solid
Yield	52%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.27 (s, 9H, Me ₃), 6.85 (d, 1H, J=8.5 Hz, Ar ₂ 3-H), 7.24 (bs, 2H, NH ₂), 7.33 (dd, 1H, J=8.5, 2.4 Hz, Ar ₂ 4-H), 7.51 (dd, 1H, J=8.0, 6.8 Hz, Ar ₁ 5-H), 7.66 (d, 1H, J=2.4 Hz, Ar ₂ 6-H) 7.84 (m, 1H, Ar ₁ 4-H), 8.32 (ddd, 1H, J=6.4, 1.8, 0.9 Hz, Ar ₁ 6-H), 8.67 (s, 1H, N=CH), 10.55 (bs, 1H, OH)
IR (KBr)	3413, 3308, 3122, 2961, 2864, 1652, 1624, 1604, 1585, 1564, 1547, 1422, 1401, 1361, 1341, 1302, 1283, 1264, 1239 (N-O), 1189(C-N), 1123 cm ⁻¹ .
MS (APCI +ve)	m/z=377 (M+H) ⁺ , 358 (M-OH) ⁺
Melting point	206.8-208.4 °C

***N*¹-(2-Hydroxy-3-*tert*-butyl)-benzylidene-pyridine-3-carboxamidrazone-1-N-oxide**
[111]

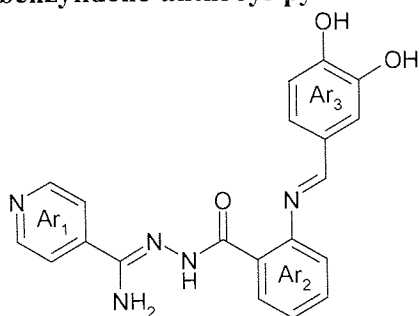
Structure



Physical state	Yellow crystalline solid
Yield	79%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.42 (s, 9H, Me ₃), 6.87 (t, 1H, J=7.6 Hz, Ar ₂ 5-H), 7.19 (dd, 1H, J=7.6, 1.5 Hz, Ar ₂ 4-H), 7.29 (bs, 2H, NH ₂), 7.34 (dd, 1H, J=7.6, 1.5 Hz, Ar ₂ 6-H), 7.52 (dd, 1H, J=7.9, 6.4 Hz, Ar ₁ 5-H), 7.84 (ddd, 1H, J=7.9, 1.5, 0.9 Hz, Ar ₁ 4-H), 8.32 (ddd, 1H, J=6.4, 1.8, 0.9 Hz, Ar ₁ 6-H), 8.65 (s, 1H, N=CH), 8.68 (m, 1H, Ar ₁ 2-H), 11.71 (s, 1H, OH)
IR (KBr)	3464, 3285, 3129, 2972, 2942, 2855, 1637, 1604, 1562, 1547, 1480, 1426, 1396, 1342, 1241 (N-O), 1196, 1162(C-N), 1115 cm ⁻¹ .
MS (APCI +ve)	m/z=313 (M+H) ⁺ , 295 (M-OH) ⁺
Melting point	239.4-241.5 °C

*N*¹-(3,4-Dihydroxy)-benzylidene-anthroyl-pyridine-4-carboxamidrazone [201]

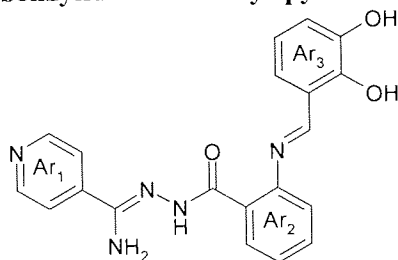
Structure



Physical state	White powder
Yield	70%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	5.73 (s, 1H, CH=N), 6.62-6.83 (overlapping m, 6H, NH ₂ , Ar ₂ -H, Ar ₂ 4H, Ar ₃ 5-H, Ar ₃ 6-H), 6.96 (d, 1H, J=1.83Hz, Ar ₃ 2-H) 7.17 (s, 1H, NH), 7.24 (m, 1H, Ar ₂ 3-H) 7.56 (dd, 2H J=1.53, 4.58Hz, Ar ₁ 2-H, Ar ₁ 6H) 7.66 (dd, 1H, J=1.22, 7.63Hz, Ar ₂ 5-H), 8.58 (dd, 2H, J=1.53, 4.58Hz, Ar ₁ 3-H, Ar ₁ 5-H) 8.77-8.85 (overlapping m, 2H, 2OH)
IR (KBr)	3527, 3504, 3399, 3290, 1642 (C=O), 1625, 1608, 1506, 1485, 1453, 1365, 1288, 1264, 1197, 1153(C-N) cm ⁻¹ .
MS (APCI +ve)	m/z=376 (M+H) ⁺ , 358 (M-OH) ⁺
Melting point	191.9-195.9 °C

*N*¹-(2,3-Dihydroxy)-benzylidene-anthroyl-pyridine-4-carboxamidrazone [202]

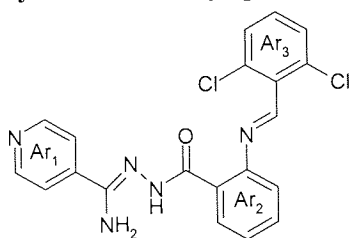
Structure



Physical state	White powder
Yield	84%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.14 (s, 1H, CH=N), 6.51 (dd, 1H, J=7.63, 7.94Hz Ar ₃ 5-H), 6.65-6.72 (overlapping m, 2H Ar ₃ 4-H, Ar ₃ 6-H), 6.78-6.89 (overlapping m, 4H, NH ₂ , Ar ₂ 2-H, Ar ₂ 4-H), 7.03 (s, 1H, NH), 7.24 (m, 1H, Ar ₂ 3-H), 7.59 (dd, 2H J=1.53, 4.58Hz, Ar ₁ 2-H, Ar ₁ 6H) 7.67 (dd, 1H, J=1.22, 7.94Hz, Ar ₂ 5-H), 8.60 (dd, 2H, J=1.53, 4.58Hz, Ar ₁ 3-H, Ar ₁ 5-H) 9.07-9.18 (overlapping m, 2H, 2OH)
IR (KBr)	3402, 3364, 1649(C=O), 1599, 1542, 1495, 1465, 1418, 1275, 1250, 1203, 1170(C-N) cm ⁻¹ .
MS (APCI +ve)	m/z=376 (M+H) ⁺ , 358 (M-OH) ⁺
Melting point	211.6-212.3 °C

***N*¹-(2,6-Dichlor)-benzylidene-anthroyl-pyridine-4-carboxamidrazone [203]**

Structure

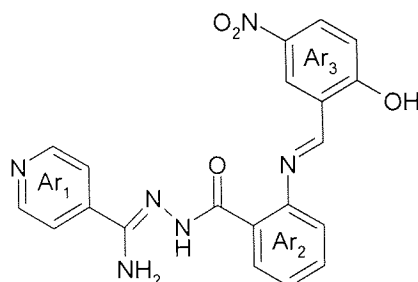


Physical state	Brown powder
Yield	67%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.61-6.69 (overlapping m, 4H, CH=N, Ar ₃ 3-H, Ar ₃ 4-H, Ar ₃ 5-H), 7.08 (d, 1H, J=0.92Hz NH), 7.25 (m, 1H Ar ₂ 3-H), 7.31-7.37 (overlapping m, 2H, Ar ₂ 2-H, Ar ₂ 4-H), 7.41-7.44 (overlapping m, 4H, NH ₂ , Ar ₁ 2-H, Ar ₁ 6H) 7.67 (dd, 1H, J=1.22, 7.94Hz, Ar ₂ 5-H), 8.56 (dd, 2H, J=1.53, 4.58Hz, Ar ₁ 3-H, Ar ₁ 5-H)
IR (KBr)	3465, 3419, 3368, 1659(C=O), 1589, 1568, 1500, 1389, 1323, 1259, 1153, 1165(C-N) cm ⁻¹ .
MS (APCI +ve)	m/z= (M+H) ⁺ , (M-OH) ⁺
Melting point	225.3-226.7 °C

***N*¹-(2-Hydroxy-5-nitro)-benzylidene-anthroyl-pyridine-4-carboxamidrazone [204]**

[204]

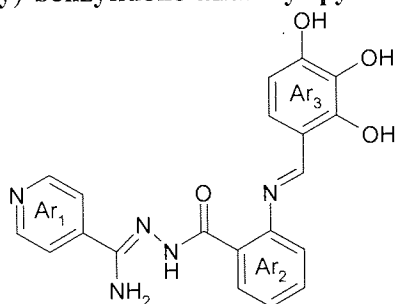
Structure



Physical state	Yellow powder
Yield	62%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.26 (s, 1H, CH=N), 6.71-6.82 (overlapping m, 2H, Ar ₂ 2-H, Ar ₂ 4-H), 6.99 (d, 1H J=8.85Hz, Ar ₃ 3-H), 7.02 (s, 2H, NH ₂), 7.22-7.32 (overlapping m, 2H, NH, Ar ₂ 3-H,) 7.53 (dd, 1H, J=1.53, 4.58Hz, Ar ₁ 2-H, Ar ₁ 6-H), 7.72 (dd, 1H, J=1.22, 7.94Hz, Ar ₂ 5-H), 8.03 (dd, 1H, J=3.05, 8.85Hz, Ar ₃ 4-H), 8.41 (d, 1H, J=3.05Hz, Ar ₃ 6-H) 8.58 (dd, 2H, J=1.53, 4.58Hz, Ar ₁ 3-H, Ar ₁ 5-H)
IR (KBr)	3413, 3320, 3183, 1628(C=O), 1498, 1439, 1415, 1392, 1338, 1291, 1154, 1084, 1164(C-N) cm ⁻¹ .
MS (APCI +ve)	m/z=405 (M+H) ⁺ , 487 (M-OH) ⁺
Melting point	208.8-209.8 °C

***N*¹-(2,3,4-Trihydroxy)-benzylidene-anthroyl-pyridine-4-carboxamidrazone [205]**

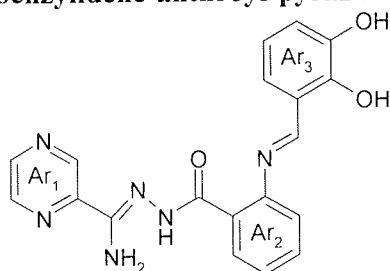
Structure



Physical state	White powder
Yield	91%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.04 (d, 1H, J=1.53Hz, CH=N), 6.16 (d, 1H, J=8.55Hz, Ar ₃ 5-H), 6.64 (d, 1H J=8.55Hz, Ar ₃ 6-H), 6.68 (m, 1H, Ar ₂ 2-H), 6.78-6.86 (overlapping m, 3H, NH ₂ , Ar ₂ 4-H), 7.00 (s, 1H, NH), 7.24(m, 1H, Ar ₂ 3-H,) 7.61 (dd, 2H, J=4.58, 1.53 Hz, Ar ₁ 2-H, Ar ₁ 6-H), 7.66 (dd, 1H, J=1.22, 7.94Hz, Ar ₂ 5-H), 8.19 (bs, 1H,Ar ₃ -OH), 8.61 (dd, 2H, J=1.53, 4.58Hz, Ar ₁ 3-H, Ar ₁ 5-H), 8.89(bs, 1H, Ar ₃ -OH), 9.29(bs, 1H, Ar ₃ -OH)
IR (KBr)	3523, 3360, 1648(C=O), 1598, 1544, 1496, 1464, 1444, 1420, 1360, 1312, 1281, 1243, 1211, 1177(C-N), 1041 cm ⁻¹ .
MS (APCI +ve)	m/z=392 (M+H) ⁺ , 374 (M-OH) ⁺
Melting point	207.2-207.6 °C

***N*¹-(2,3-Dihydroxy)-benzylidene-anthroyl-pyrazine-2-carboxamidrazone [206]**

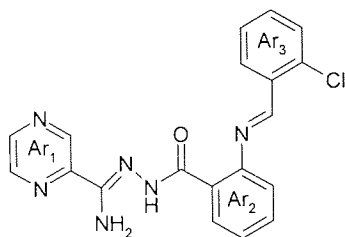
Structure



Physical state	Light yellow powder
Yield	86%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.26 (d, 1H, J=1.22Hz, CH=N), 6.52 (dd, 1H, J=7.63, 7.94Hz Ar ₃ 5-H), 6.65-6.88 (overlapping m, 6H, NH ₂ , Ar ₃ 4-H, Ar ₃ 6-H, Ar ₂ 2-H, Ar ₂ 4-H), 7.03 (s, 1H, NH), 7.24 (m, 1H, Ar ₂ 3-H), 7.69 (dd, 1H, J=1.22, 7.94Hz, Ar ₂ 5-H), 8.67 (dd, 1H, J=1.53, 2.44Hz, Ar ₁ 4-H), 8.73 (d, 1H, J=2.44Hz, Ar ₁ 3-H), 8.95 (bs, 1H, OH), 8.98 (d, 1H, J=1.53Hz, Ar ₁ 6-H), 9.23 (bs, 1H, OH)
IR (KBr)	3406, 3368, 1657(C=O), 1617, 1569, 1490, 1474, 1414, 1333, 1276, 1251, 1231, 1170(C-N), 1067 cm ⁻¹ .
MS (APCI +ve)	m/z=377 (M+H) ⁺ , 360 (M-OH) ⁺
Melting point	205.1-205.9 °C

*N*¹-(2-Chloro)-benzylidene-anthroyl-pyrazine-2-carboxamidrazone [207]

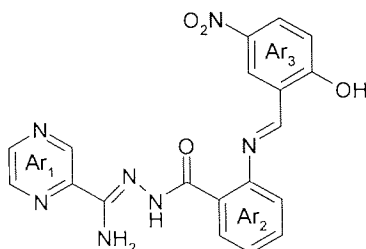
Structure



Physical state	White crystalline solid
Yield	81%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.49 (d, 1H, J=0.92Hz, CH=N), 6.72-6.79 (overlapping m, 2H, Ar ₂ 2-H, Ar ₂ 4-H), 6.85 (bs, 2H, NH ₂), 7.26-7.32 (overlapping m, 4H, NH, Ar ₂ 3-H, Ar ₃ 4-H, Ar ₃ 6-H) 7.43 (m, 1H, Ar ₃ 5-H), 7.12 (dd, 1H, J=1.22, 7.94Hz, Ar ₂ 5-H), 7.82 (m, 1H, Ar ₃ 5-H), 8.63 (dd, 1H, J=1.53, 2.44Hz, Ar ₁ 4-H), 8.70 (d, 1H, J=2.44Hz, Ar ₁ 3-H), 8.89 (d, 1H, J=1.53Hz, Ar ₁ 6-H).
IR (KBr)	3453, 3429, 3371, 1649(C=O), 1609, 1569, 1502, 1407, 1323, 1257, 1153(C-N), 1042, 1017 cm ⁻¹ .
MS (APCI +ve)	m/z= 379 (M+H) ⁺ , 361(M-OH) ⁺
Melting point	215.0-216.5 °C

*N*¹-(2-Hydroxy-5-nitroyl)-benzylidene-anthroyl-pyrazine-2-carboxamidrazone [208]

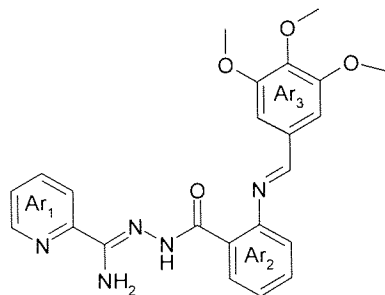
Structure



Physical state	Yellow powder
Yield	91%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.34 (s, 1H, CH=N), 6.72-6.82 (overlapping m, 2H, Ar ₂ 2-H, Ar ₂ 4-H), 6.93 (bs, 2H, NH ₂), 6.99 (d, 1H J=8.85Hz, Ar ₃ 3-H), 7.22-7.32 (overlapping m, 2H, NH, Ar ₂ 3-H,) 7.73 (dd, 1H, J=1.22, 7.94Hz, Ar ₂ 5-H), 8.03 (dd, 1H, J=3.05, 8.85Hz, Ar ₃ 4-H), 8.41 (d, 1H, J=3.05Hz, Ar ₃ 6-H), 8.65 (dd, 1H, J=1.53, 2.44Hz, Ar ₁ 4-H), 8.71 (d, 1H, J=2.44Hz, Ar ₁ 3-H), 8.96 (d, 1H, J=1.53Hz, Ar ₁ 6-H)
IR (KBr)	3431, 3326, 3076, 1631(C=O), 1610, 1590, 1571, 1522, 1499, 1413, 1343, 1284, 1219, 1158(C-N), 1087, 1019 cm ⁻¹ .
MS (APCI +ve)	m/z=406 (M+H) ⁺ , 389 (M-OH) ⁺
Melting point	210.5-211.2 °C

*N*¹-(3,4,5-Trihydroxy)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [209]

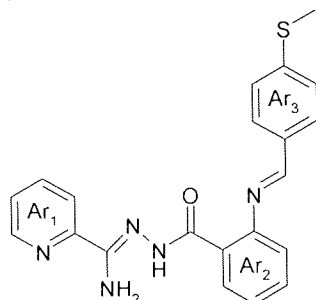
Structure



Physical state	White powder
Yield	68%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.58 (s, 3H, Ar ₃ 4OMe), 3.72 (s, 6H, Ar ₃ 3OMe, Ar ₃ 5OMe), 5.95 (s, 1H, CH=N), 6.68-6.79 (overlapping m, 4H, NH ₂ , Ar ₂ 2-H, Ar ₂ 4-H), 6.95 (s, 2H Ar ₂ 2-H Ar ₃ 6-H), 7.24-7.30 (overlapping m, 2H, NH, Ar ₂ 3-H), 7.47 (m, 1H, Ar ₁ 4-H), 7.69 (dd, 1H, J=1.22, 7.63Hz, Ar ₂ 5-H), 7.78-7.90 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.57 (d, 1H, J=4.58Hz, Ar ₁ 3-H),
IR (KBr)	3418, 3331, 3055, 3002, 2936 (CH ₃), 2834, 1641(C=O), 1613, 1565, 1503, 1461, 1422, 1368, 1325, 1299, 1238, 1127(C-N), 1004 cm ⁻¹ .
MS (APCI +ve)	m/z=434 (M+H) ⁺ , 416 (M-OH) ⁺
Melting point	200.0-202.7 °C

*N*¹-(4-Methylsulfanyl)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [210]

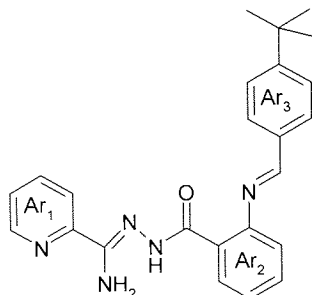
Structure



Physical state	White powder
Yield	85%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	2.41 (s, 3H, SMe), 5.91 (d, 1H, J=1.53Hz CH=N), 6.67-6.76 (overlapping m, 4H, NH ₂ , Ar ₂ 2-H, Ar ₂ 4-H), 7.17-7.31 (overlapping m, 4H, NH, Ar ₂ 3-H, Ar ₃ 3-H, Ar ₃ 5-H), 7.46-7.51 (overlapping m, 3H, Ar ₁ 4-H, Ar ₃ 3-H, Ar ₃ 5-H), 7.67 (dd, 1H, J=1.22, 7.63Hz, Ar ₂ 5-H), 7.79-7.92 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.58 (d, 1H, J=4.58Hz, Ar ₁ 3-H),
IR (KBr)	3408, 3352, 3304, 1646(C=O), 1626, 1585, 1565, 1502, 1485, 1437, 1371, 1292, 1155(C-N), 1088 cm ⁻¹ .
MS (APCI +ve)	m/z=390 (M+H) ⁺ , 372 (M-OH) ⁺
Melting point	220.2-221.7 °C

*N*¹-(4-*tert*-Butyl)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [211]

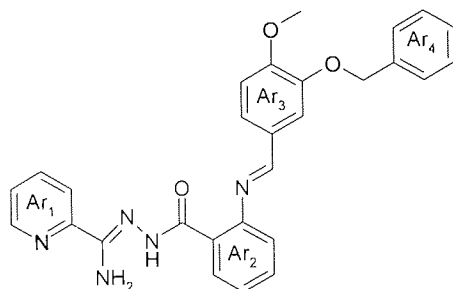
Structure



Physical state	White powder
Yield	71%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.22 (s, 9H, 3×Me), 5.89 (d, 1H, J=1.53Hz CH=N), 6.66-6.73 (overlapping m, 4H, NH ₂ , Ar ₂ 2-H, Ar ₂ 4-H), 7.21-7.36 (overlapping m, 4H, NH, Ar ₂ 3-H, Ar ₃ 3-H, Ar ₃ 5-H), 7.46-7.49 (overlapping m, 3H, Ar ₁ 4-H, Ar ₃ 3-H, Ar ₃ 5-H), 7.68 (dd, 1H, J=1.22, 7.63Hz, Ar ₂ 5-H), 7.78-7.90 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.58 (d, 1H, J=4.58Hz, Ar ₁ 3-H),
IR (KBr)	3492, 3354, 3304, 2956 (CH ₃), 2866, 1649(C=O), 1626, 1586, 1566, 1505, 1487, 1444, 1385, 1325, 1291, 1265, 1240, 1211, 1152(C-N), 1108 cm ⁻¹ .
MS (APCI +ve)	m/z=400 (M+H) ⁺ , 380,382 (M-OH) ⁺
Melting point	218.2-220.1 °C

*N*¹-(3-Benzyloxy-4-methoxy)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [212]

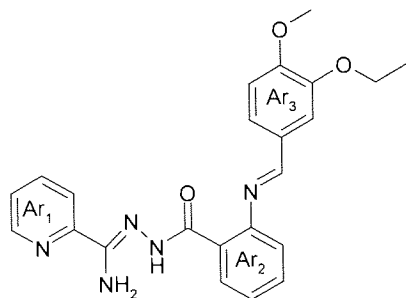
Structure



Physical state	
Yield	81%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.69 (s, 3H, Me), 4.99 (d, 2H, J=1.83Hz, CH ₂), 5.92 (s, 1H, CH=N), 6.68-7.78 (overlapping m, 4H, NH ₂ , Ar ₂ 2-H, Ar ₂ 4-H), 6.88 (d, 1H, J=8.24Hz, Ar ₃ 5-H), 7.07 (dd, 1H, J=1.83, 8.24Hz, Ar ₃ 6-H) 7.23-7.50 (overlapping m, 9H, NH, Ar ₂ 3-H, Ar ₁ 4-H, Ar ₃ 2-H, Ar ₄ 2-H, Ar ₄ 3-H, Ar ₄ 4-H, Ar ₄ 5-H, Ar ₄ 6-H), 7.69 (dd, 1H, J=1.22, 7.63Hz, Ar ₂ 5-H), 7.78-7.91 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.57 (d, 1H, J=4.58Hz, Ar ₁ 3-H),
IR (KBr)	3361, 3060, 3025, 2948 (CH ₃), 2931, 2829, 1640(C=O), 1613, 1568, 1553, 1512, 1487, 1443, 1385, 1310, 1289, 1255, 1230, 1158, 1130(C-N), 1015 cm ⁻¹ .
MS (APCI +ve)	m/z=480 (M+H) ⁺ , 462 (M-OH) ⁺
Melting point	191.1-194.0 °C

***N*¹-(3-Ethoxy-4-methoxy)-benzylidene-anthroyl-pyridine-2-carboxamidrazone**
[213]

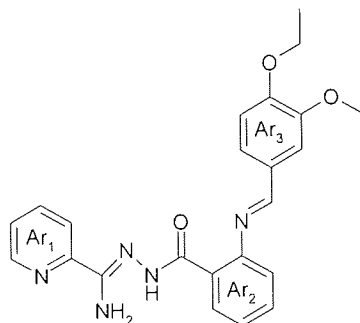
Structure



Physical state	White powder
Yield	88%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.26 (t, 3H, J=7.02Hz, CH ₃ -CH ₂), 3.68 (s, 3H, CH ₃ -O), 3.95 (q, 2H, J=7.02Hz, CH ₂ -CH ₃), 5.91 (s, 1H, CH=N), 6.67-6.86 (overlapping m, 5H, NH ₂ , Ar ₂ 2-H, Ar ₂ 4-H, Ar ₃ 5-H), 7.03 (dd, 1H, J=1.83, 8.24Hz, Ar ₃ 6-H) 7.22-7.28 (overlapping m, 3H, NH, Ar ₂ 3-H, Ar ₃ 2-H), 7.47 (m, 1H, Ar ₁ 4-H), 7.67 (dd, 1H, J=1.22, 7.63Hz, Ar ₂ 5-H), 7.78-7.92 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.57 (d, 1H, J=4.58Hz, Ar ₁ 3-H)
IR (KBr)	3508, 3395, 3303, 2947 (CH ₃), 2859, 1654(C=O), 1631, 1609, 1585, 1565, 1546, 1515, 1499, 1485, 1442, 1389, 1374, 1354, 1264, 1231, 1174, 1143(C-N), 1040, 1019 cm ⁻¹ .
MS (APCI +ve)	m/z=418 (M+H) ⁺ , 400 (M-OH) ⁺
Melting point	222.2-225.8 °C

***N*¹-(3-Methoxy-4-ethoxy)-benzylidene-anthroyl-pyridine-2-carboxamidrazone**
[214]

Structure

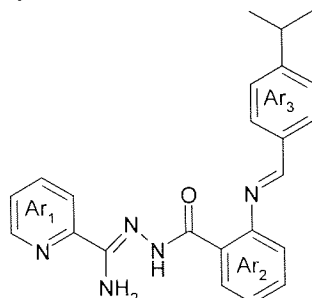


Physical state	White powder
Yield	93%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.26 (t, 3H, J=7.02Hz, CH ₃ -CH ₂), 3.69 (s, 3H, CH ₃ -O), 3.92 (q, 2H, J=7.02Hz, CH ₂ -CH ₃), 5.93 (s, 1H, CH=N), 6.68-6.84 (overlapping m, 5H, NH ₂ , Ar ₂ 2-H, Ar ₂ 4-H, Ar ₃ 5-H), 7.02 (dd, 1H, J=1.83, 8.24Hz, Ar ₃ 6-H) 7.21-7.29 (overlapping m, 3H, NH, Ar ₂ 3-H, Ar ₃ 2-H), 7.47 (ms, 1H, Ar ₁ 4-H), 7.68 (dd, 1H, J=1.22, 7.63Hz, Ar ₂ 5-H), 7.78-7.90 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.57 (d, 1H, J=4.58Hz, Ar ₁ 3-H)
IR (KBr)	3489, 3371, 3297, 2976(CH ₃), 2927, 1646(C=O), 1623, 1566, 1506, 1486, 1421, 1373, 1261, 1233, 1142(C-N), 1037 cm ⁻¹ .

MS (APCI +ve) $m/z=418 (M+H)^+$, $400 (M-OH)^+$
 Melting point 216.7-219.3 °C

***N*¹-(4-Isopropyl)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [215]**

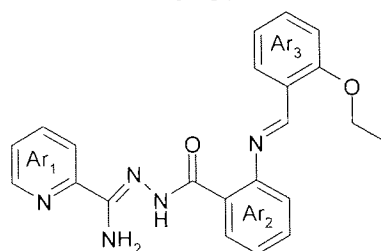
Structure



Physical state White crystalline solid
 Yield 61%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 1.14 (d, 6H, J=7.02Hz 2×Me), 2.83(sept, 1H, J=7.02Hz, CH), 5.90 (d, 1H, J=1.53Hz CH=N), 6.67-6.76 (overlapping m, 4H, NH₂, Ar₂2-H, Ar₂4-H), 7.16-7.27 (overlapping m, 3H, Ar₂3-H, Ar₃3-H, Ar₃5-H), 7.34 (d, 1H, J=0.92Hz NH) 7.45-7.50 (overlapping m, 3H, Ar₁4-H, Ar₃3-H, Ar₃5-H), 7.68 (dd, 1H, J=1.22, 7.63Hz, Ar₂5-H), 7.78-7.88 (overlapping m, 2H, Ar₁5-H, Ar₁6-H) 8.58 (d, 1H, J=4.58Hz, Ar₁3-H),
 IR (KBr) 3479, 3299, 2955 (CH₃), 1623(C=O), 1586, 1565, 1507, 1487, 1384, 1333, 1293, 1153(C-N) cm⁻¹.
 MS (APCI +ve) $m/z=386 (M+H)^+$, $368 (M-OH)^+$
 Melting point 211.0-213.6 °C

***N*¹-(2-Ethoxy)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [216]**

Structure

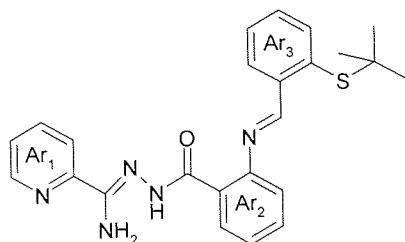


Physical state White powder
 Yield 69%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 1.33 (t, 3H, J=7.02Hz Me), 4.07(t, 1H, J=7.02Hz, CH₂), 6.23 (d, 1H, J=1.83Hz CH=N), 6.65-7.00 (overlapping m, 7H, NH₂, NH, Ar₂2-H, Ar₂4-H, Ar₃4-H, Ar₃6-H), 7.18-7.24 (overlapping m, 2H, , Ar₂3-H, Ar₃5-H), 7.41-7.50 (overlapping m, 2H, Ar₁4-H, Ar₃3-H), 7.70 (dd, 1H, J=1.22, 7.94Hz, Ar₂5-H), 7.77-7.85 (overlapping m, 2H, Ar₁5-H, Ar₁6-H) 8.58 (m, 1H, Ar₁3-H).
 IR (KBr) 3490, 3361, 3067, 2977 (CH₃), 2914, 2873, 1635(C=O), 1586, 1567, 1536, 1505, 1487, 1471, 1455, 1390, 1329, 1289, 1239, 1203, 1153(C-N), 1117, 1041 cm⁻¹.
 MS (APCI +ve) $m/z=388 (M+H)^+$, $370 (M-OH)^+$
 Melting point 177.3-178.8 °C

*N*¹-(2-*tert*-Butylsulfanyl)-benzylidene-anthroyl-pyridine-2-carboxamidrazone

[217]

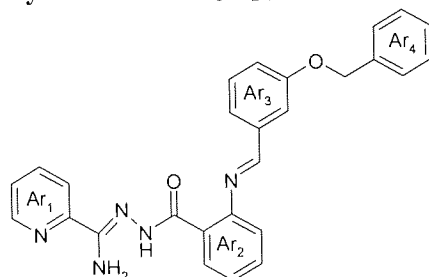
Structure



Physical state	White powder
Yield	77%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.27 (s, 9H, 3 × Me), 6.68-6.82 (overlapping m, 5H, NH ₂ , CH=N, Ar ₂ 2-H, Ar ₂ 4-H), 6.96 (s, 1H, NH), 7.20-7.54 (overlapping m, 5H, Ar ₁ 4-H, Ar ₂ 3-H, Ar ₂ 5-H, Ar ₃ 4-H, Ar ₃ 6-H), 7.69-7.77 (overlapping m, 4H, , Ar ₂ 5-H, Ar ₃ 3-H, Ar ₁ 5-H, Ar ₁ 6-H), 8.55 (d, 1H, J=4.58Hz, Ar ₁ 3-H).
IR (KBr)	3476, 3370, 3315, 3294, 3054, 2955 (CH ₃), 1647(C=O), 1628, 1587, 1565, 1486, 1438, 1368, 1161(C-N), 1049 cm ⁻¹ .
MS (APCI +ve)	m/z=432 (M+H) ⁺ , 414 (M-OH) ⁺
Melting point	222.3-225.1 °C

*N*¹-(3-Benzyloxy)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [218]

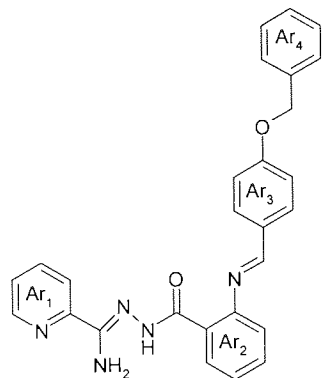
Structure



Physical state	White powder
Yield	77%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	5.04 (s, 2H, CH ₂), 5.92 (d, 1H, J=1.22 CH=N), 6.67-6.77 (overlapping m, 4H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H), 6.89 (m, 1H, Ar ₃ 4-H), 7.11-7.41 (overlapping m, 10H, 5 × Ar ₄ -H, NH, 3 × Ar ₃ -H, Ar ₂ 4-H), 7.49 (m, 1H, Ar ₁ 4-H) 7.68 (dd, 1H, J=6.71, 1.22 Hz, Ar ₂ 6-H) 7.79-7.90 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H), 8.58 (m, 1H, Ar ₁ 3-H),
IR (KBr)	3379, 3264, 3058, 2929, 1649(C=O), 1629, 1585, 1562, 1486, 1451, 1395, 1383, 1292, 1280, 1265, 1231, 1152(C-N), 1036, 1025 cm ⁻¹ .
MS (APCI +ve)	m/z=450 (M+H) ⁺ , 432 (M-OH) ⁺
Melting point	170.2-171.0 °C

*N*¹-(4-Benzyloxy)-benzylidene-anthroyl-pyridine-2-carboxamidrazone [219]

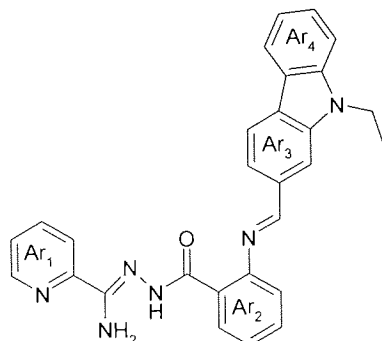
Structure



Physical state	White crystalline solid
Yield	81%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	5.04 (s, 2H, CH ₂), 5.91 (s, 1H, CH=N), 6.66-6.75 (overlapping m, 4H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H), 6.92 (d, 2H, J=8.85 Ar ₃ 3-H, Ar ₃ 5-H), 7.21-7.50 (overlapping m, 10H, 5 × Ar ₄ -H, NH, 2 × Ar ₃ -H, Ar ₂ 4-H, Ar ₁ 4-H), 7.67 (m, 1H, Ar ₂ 6-H) 7.78-7.90 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H), 8.57 (m, 1H, Ar ₁ 3-H),
IR (KBr)	3472, 3390, 3302, 3059, 3027, 2994, 1644(C=O), 1615, 1586, 1564, 1506, 1485, 1370, 1289, 1240, 1170(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 450.1930 (M+H) ⁺ , found: 450.1910
Melting point	239.8-241.4.0 °C

*N*¹-(9-Ethyl)-carbazole-anthroyl-pyridine-2-carboxamidrazone [220]

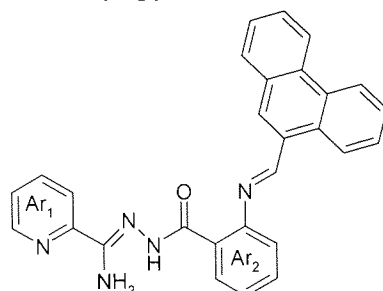
Structure



Physical state	White crystalline solid
Yield	72%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.24 (t, 3H, J=6.95Hz, CH ₃), 4.39 (q, 2H, J=6.95Hz, CH ₂), 6.18 (s, 1H, CH=N), 6.68-6.78 (overlapping m, 4H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H), 7.15-7.59 (overlapping m, 7H, 4 × Ar ₄ -H, NH, Ar ₂ 4-H, Ar ₁ 4-H), 7.70-7.76 (overlapping m, 3H, Ar ₂ 6-H, Ar ₁ 5-H, Ar ₁ 6-H), 7.84 (d, 1H, J=7.58 Hz, Ar ₃ 5-H), 8.08 (d, 1H, J=7.58, Ar ₃ 6-H) 8.35 (d, 1H, J=1.26 Hz, Ar ₃ 2-H), 8.50 (m, 1H, Ar ₁ 3-H),
IR (KBr)	3473, 3376, 3301, 3054, 2969 (CH ₃), 1647(C=O), 1625, 1547, 1510, 1487, 1465, 1359, 1222, 1147(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 461.2090 (M+H) ⁺ , found: 461.2081
Melting point	260.9-263.8 °C

*N*¹-(9-Phenanthrene)-anthroyl-pyridine-2-carboxamidrazone [221]

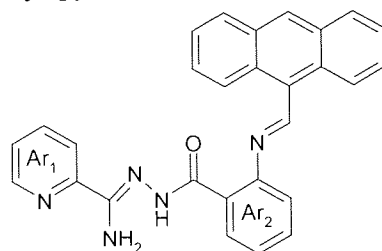
Structure



Physical state	White crystalline solid
Yield	87.2%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.73-6.82(overlapping, 4H, NH ₂ , 2Ar-H), 7.24-7.35(overlapping m, 2H, Ar ₁ =H, Ar-H), 7.40(s, 1H, CH=N), 7.51-7.70 (overlapping m, 2 × Ar ₁ -H, 4 × Ar-H), 7.80(d, 1H, J=7Hz, Ar ₂ 6-H), 7.94(d, 1H, J=7.5Hz, Ar ₁ 6-H), 8.10 (s, 1H, NH), 8.44 (d, 1H, J=5.3Hz, Ar ₁ 3-H), 8.76-8.90(overlapping m, 3H, 3 × Ar ₃ -H)ppm
IR (KBr)	3486, 3308, 3049, 2798, 1647(C=O), 1615, 1559, 1512, 1501, 1365, 1283, 1156(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 444.1824 (M+H) ⁺ , found: 444.1821
Melting point	0.7 ppm
	250.5-252.0 °C

*N*¹-(9-Anthryl)-anthroyl-pyridine-2-carboxamidrazone [222]

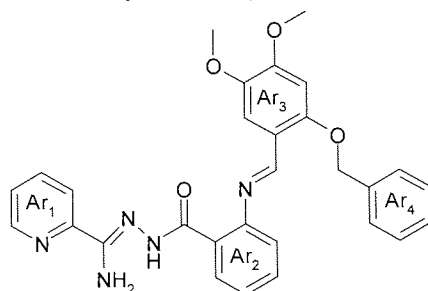
Structure



Physical state	White crystalline solid
Yield	72.3%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	6.65-6.82(overlapping m, 5H, NH ₂ , Ar ₂ 6-H, 2 × Ar-H), 7.24-7.35(overlapping m, 3H, 3 × Ar-H), 7.400(s, 1H, CH=N), 7.51-7.770 (overlapping m, 6H, 5 × Ar-H, Ar ₁ 4-H), 7.80 (d, J=7Hz, Ar ₁ 6-H), 7.94 (d, 1H, J=7.5Hz, Ar ₂ 6-H), 8.10(s, 1H, NH), 8.44(d, 1H, J=4.5Hz, Ar ₁ 3-H), 8.76-8.90(overlapping m, 3H, 3 × Ar-H)ppm
IR (KBr)	3495, 3309, 3049, 2903, 1652(C=O), 1611, 1547, 1512, 1484, 1368, 1297, 1251, 1158 (C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 444.1824 (M+H) ⁺ , found: 444.1817
Melting point	1.6 ppm
	248.7-250.4 °C

***N*¹-(2-Benzoyloxy-4,5-dimethoxy)-anthroyl-pyridine-2-carboxamidrazone [223]**

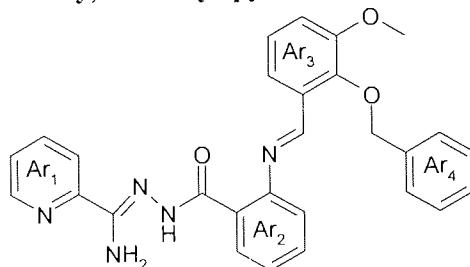
Structure



Physical state	White crystalline solid
Yield	71%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.60 (s, 3H, OCH ₃), 3.67 (s, 3H, OCH ₃), 5.14 (s, 2H, CH ₂), 6.44 (s, 1H, CH=N), 6.70-6.77 (overlapping ms, 5H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H, Ar ₃ 3-H), 6.92 (s, 1H NH), 7.25-7.50 (overlapping m, 8H, 5 × Ar ₄ -H, Ar ₁ 6-H, Ar ₂ 4-H, Ar ₃ 6-H), 7.69 (m, 1H, Ar ₂ 6-H), 7.78-7.82 (overlapping m, 2H, Ar ₁ 4-H, Ar ₁ 5-H), 8.56 (m, 1H, Ar ₁ 3-H),
IR (KBr)	3472, 3362, 3297, 3045, 2980 (CH ₃), 1644(C=O), 1608, 1561, 1524, 1484, 1428, 1400, 1378, 1350, 1314, 1301, 1272, 1208, 1199, 1185, 1148(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 510.2141 (M+H) ⁺ , found: 510.2140
Melting point	0.2 ppm 200.7-202.2 °C

***N*¹-(2-Benzoyloxy-3-methoxy)-anthroyl-pyridine-2-carboxamidrazone [224]**

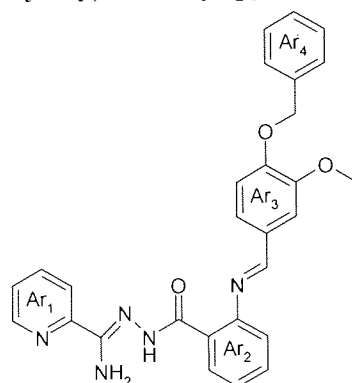
Structure



Physical state	White crystalline solid
Yield	80%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.81 (s, 3H, OCH ₃), 5.02 (ab system, 2H, J=10.74 Hz, CH ₂), 6.32 (d, 1H, J=1.26 Hz, CH=N), 6.66-6.72 (overlapping m, 3H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H), 6.96-7.01 (overlapping m, 3H, 3 × Ar ₃ -H), 7.20-7.48 (overlapping m, 8H, 5 × Ar ₄ -H, NH, Ar ₂ 4-H, Ar ₁ 4-H), 7.67 (dd, 1H, J=8.21, 1.26 Hz, Ar ₂ 6-H), 7.80-7.82 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H), 8.57 (m, 1H, Ar ₁ 3-H),
IR (KBr)	3471, 3357, 3304, 3044, 2984 (CH ₃), 1652(C=O), 1608, 1561, 1523, 1482, 1435, 1400, 1386, 1353, 1308, 1298, 1272, 1218, 1192, 1180, 1142(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 480.2036 (M+H) ⁺ , found: 480.2091
Melting point	1.5 ppm 207.7-211.1 °C

*N*¹-(3-Methoxy-4-benzyloxy)-anthroyl-pyridine-2-carboxamidrazone [225]

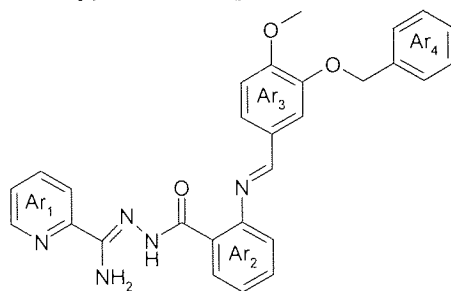
Structure



Physical state	White crystalline solid
Yield	70%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.71 (s, 3H, OCH ₃), 5.00 (s, 2H, CH ₂), 5.93 (s, 1H, CH=N), 6.67-6.77 (overlapping m, 3H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H), 6.92 (d, 1H, J=8.21 Hz, Ar ₃ 5-H), 6.99 (dd, 1H, J=8.21, 1.90 Hz, Ar ₃ 6-H), 7.22-7.37 (overlapping m, 8H, 5 × Ar ₄ -H, Ar ₃ 6-H, Ar ₂ 4-H, NH), 7.48 (ddd, 1H, J=6.95, 5.06, 1.90 Hz, Ar ₁ 4-H) 7.68 (dd, 1H, J=7.58, 1.26 Hz, Ar ₂ 6-H), 7.78-7.89 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.57 (m, 1H, Ar ₁ 3-H),
IR (KBr)	3475, 3372, 3324, 3289, 3033, 3004 (CH ₃), 2935, 2856, 1644(C=O), 1627, 1563, 1502, 1485, 1423, 1363, 1265, 1141(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 480.2049 (M+H) ⁺ , found: 480.2047
Melting point	0.4 ppm
	201.8-203.7 °C

*N*¹-(3-Benzyloxy-4-methoxy)-anthroyl-pyridine-2-carboxamidrazone [226]

Structure

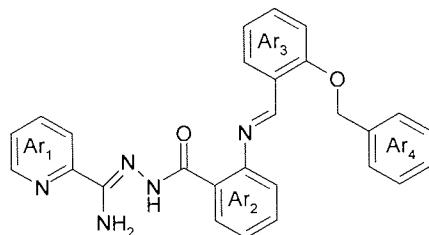


Physical state	White crystalline solid
Yield	70%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	3.68 (s, 3H, OCH ₃), 4.98 (ab system, 2H, J=12.01 Hz, CH ₂), 5.92 (s, 1H, CH=N), 6.67-6.76 (overlapping m, 3H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H), 6.87 (d, 1H, J=8.21 Hz, Ar ₃ 5-H), 7.07 (dd, 1H, J=8.21, 1.90 Hz, Ar ₃ 6-H), 7.22-7.50 (overlapping m, 9H, 5 × Ar ₄ -H, Ar ₃ 6-H, Ar ₂ 4-H, NH, Ar ₁ 4-H), 7.68 (dd, 1H, J=7.58, 1.26 Hz, Ar ₂ 6-H), 7.78-7.90 (overlapping m, 2H, Ar ₁ 5-H, Ar ₁ 6-H) 8.57 (m, 1H, Ar ₁ 3-H),
IR (KBr)	3475, 3372, 3324, 3289, 3033, 3004 (CH ₃), 2935, 2856,

1644(C=O), 1627, 1563, 1502, 1485, 1423, 1363, 1265,
1141(C-N) cm^{-1} .
Accurate MS Calc. Mass: 480.2049 (M+H)⁺, found: 480.2047
(electrospray +ve) 0.4 ppm
Melting point 201.8-203.7 °C

***N*¹-(2-Benzyloxy)-anthroyl-pyridine-2-carboxamidrazone [227]**

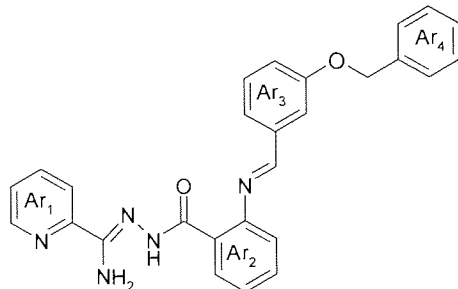
Structure



Physical state White crystalline solid
Yield 80%
1H NMR (D6-DMSO; δ DMSO=2.50ppm) 5.20 (s, 2H, CH₂), 6.37 (s, 1H, CH=N), 6.66-6.90 (overlapping m, 4H, NH₂, Ar₂3-H, Ar₂5-H, Ar₃3-H), 7.00-7.06 (overlapping m, 2H, Ar₃4-H, Ar₃5-H), 7.17-7.53 (overlapping m, 9H, 5 × Ar₄-H, Ar₃6-H, Ar₂4-H, NH, Ar₁4-H), 7.69 (dd, 1H, J=8.21, 1.26 Hz, Ar₂6-H), 7.76-7.85 (overlapping m, 2H, Ar₁5-H, Ar₁6-H) 8.57 (dd, 1H, J=6.32, 1.26 Hz, Ar₁3-H),
IR (KBr) 3483, 3382, 3305, 3060, 3033, 3008, 2936, 1655(C=O), 1607, 1567, 1552, 1486, 1453, 1357, 1330, 1292, 1230, 1158(C-N) cm^{-1} .
Accurate MS Calc. Mass: 450.1943 (M+H)⁺, found: 450.1939
(electrospray +ve) 0.9 ppm
Melting point 225.2-227.2 °C

***N*¹-(3-Benzyloxy)-anthroyl-pyridine-2-carboxamidrazone [228]**

Structure

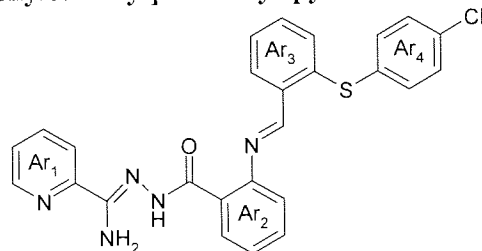


Physical state White crystalline solid
Yield 49%
1H NMR (D6-DMSO; δ DMSO=2.50ppm) 5.04 (s, 2H, CH₂), 5.92 (d, 1H, J=1.22 CH=N), 6.67-6.77 (overlapping m, 4H, NH₂, Ar₂3-H, Ar₂5-H), 6.89 (m, 1H, Ar₃4-H), 7.11-7.41 (overlapping m, 10H, 5 × Ar₄-H, NH, 3 × Ar₃-H, Ar₂4-H), 7.49 (m, 1H, Ar₁4-H) 7.68 (dd, 1H, J=6.71, 1.22 Hz, Ar₂6-H) 7.79-7.90 (overlapping m, 2H, Ar₁5-H, Ar₁6-H), 8.58 (m, 1H, Ar₁3-H),
IR (KBr) 3378, 3262, 3058, 3029, 2933, 2871, 1646(C=O), 1625, 1584, 1569, 1488, 1449, 1390, 1312, 1289, 1229, 1148(C-N) cm^{-1} .

Accurate MS Calc. Mass: 450.1930 (M+H)⁺, found: 450.1927
 (electrospray +ve) 0.7 ppm
 Melting point 170-172.1°C

N¹-[2-(4-Chloro)-phenylsulfanyl]-anthroyl-pyridine-2-carboxamidrazone [229]

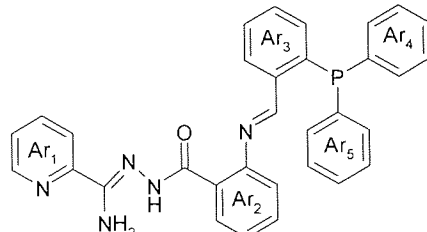
Structure



Physical state White crystalline solid
 Yield 80%
 1H NMR (D6-DMSO; δDMSO=2.50ppm) 6.59 (s, 1H, CH=N), 6.70-6.75 (overlapping m, 4H, NH₂, Ar₂3-H, Ar₂5-H), 7.17-7.38 (overlapping m, 9H, 4×Ar₄-H, 3×Ar₃-H, Ar₂4-H, NH), 7.49 (m, 1H, Ar₁4-H), 7.69-7.76 (overlapping m, 2H, Ar₂6-H, Ar₃6-H), 7.81-7.87 (overlapping m, 2H, Ar₁5-H, Ar₁6-H), 8.56 (d, 1H, J=4.42 Hz, Ar₁3-H).
 IR (KBr) 3477, 3398, 3333, 3270, 3062, 3014, 2963, 1652(C=O), 1627, 1569, 1505, 1488, 1436, 1374, 1351, 1291, 1157(C-N) cm⁻¹.
 Accurate MS Calc. Mass: 486.1155 (M+H)⁺, found: 486.1156
 (electrospray +ve) 0.2 ppm
 Melting point 200.9-207.5 °C

N¹-(2-di-Phenylphosphanyl)-anthroyl-pyridine-2-carboxamidrazone [230]

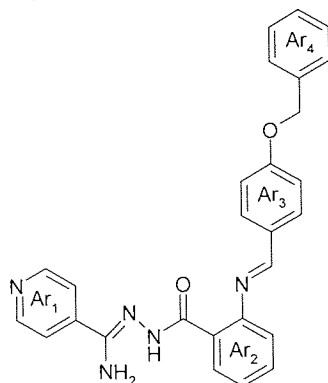
Structure



Physical state White crystalline solid
 Yield 82%
 1H NMR (D6-DMSO; δDMSO=2.50ppm) 6.70-6.76 (overlapping m, 3H, NH₂, Ar₂3-H), 6.88 (m, 1H, Ar₂5-H), 7.00-7.04 (2s, 1H, keto and enol form of NH) 7.19-7.46 (overlapping m, 17H, 5×Ar₄-H, 5×Ar₅-H, 3×Ar₃-H, 2×Ar₂-H, Ar₁4-H, Ar₂4-H, CH=N), 7.67-7.74 (overlapping m, 2H, Ar₁5-H, Ar₁6-H), 7.90 (m, 1H, Ar₃6-H), 8.52 (d, 1H, J=3.79Hz, Ar₁3-H).
 IR (KBr) 3434, 3316, 3054, 3000, 2856, 1644(C=O), 1612, 1589, 1561, 1498, 1482, 1388, 1359, 1326, 1296, 1263, 1148(C-N) cm⁻¹.
 Accurate MS Calc. Mass: 528.1953 (M+H)⁺, found: 528.1970
 (electrospray +ve) 3.2 ppm
 Melting point 224.9-226.3 °C

*N*¹-(4-Benzyloxy)-benzylidene-anthroyl-pyridine-4-carboxamidrazone [231]

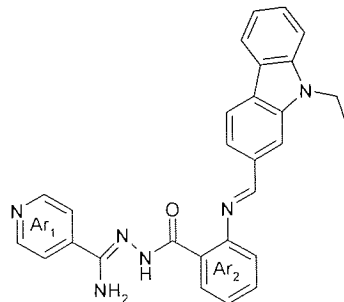
Structure



Physical state	White crystalline solid
Yield	79%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	5.06 (s, 2H, CH ₂), 5.88 (s, 1H, CH=N), 6.67-6.95 (overlapping m, 4H, NH ₂ , Ar ₂ 3-H, Ar ₂ 5-H), 6.92-6.95 (overlapping m, 2H, Ar ₃ 3-H, Ar ₃ 5-H), 7.23-7.38 (overlapping m, 7H, 5 × Ar ₄ -H, Ar ₂ 4-H, NH), 7.48-7.52 (overlapping m, 4H, Ar ₁ 2-H, Ar ₁ 6-H, Ar ₃ 2-H, Ar ₃ 6-H), 7.90 (dd, 1H, J=7.58, 1.26 Hz Ar ₂ 6-H), 8.55-8.58 (m, 2H, Ar ₁ 3-H, Ar ₁ 5-H).
IR (KBr)	3488, 3368, 3322, 3031, 2919, 2869, 2827, 1644(C=O), 1606, 1542, 1506, 1488, 1442, 1380, 1332, 1289, 1241, 1172(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 450.1930 (M+H) ⁺ , found: 450.1922
Melting point	1.8 ppm
	222.8-226.5 °C

*N*¹-(9-Ethyl)-carbazole-anthroyl-pyridine-4-carboxamidrazone [232]

Structure

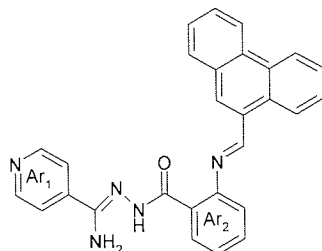


Physical state	White crystalline solid
Yield	68%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.25 (t, 3H, J=7.5Hz, CH ₃), 4.30(q, 2H, J=7.5Hz CH ₂), 6.15(s, 1H, CH=N), 6.69-6.83 (overlapping m, 4H, NH ₂ , 2 × Ar-H), 7.15-7.31 (overlapping m, 3H, 3 × Ar-H), 7.40-7.46 (overlapping m, 3H, Ar ₁ 2-H, Ar ₁ 6-H, 1 × Ar-H), 7.56 (overlapping m, 2H, 2 × Ar-H), 7.72 (d, 2H, J=7Hz, Ar ₃ 5-H, Ar ₃ 6-H), 8.08 (d, 1H, J=7.6Hz, Ar ₂ 6-H), 8.37(s, 1H, NH), 8.47(d, 2H, J=6Hz, Ar ₁ 3-H, Ar ₁ 5-H).
IR (KBr)	3494, 3397, 3284, 3054, 2982 (CH ₃), 2965, 2863, 1642(C=O),

1629, 1542, 1488, 1422, 1365, 1238, 1154(C-N) cm^{-1} .
 Accurate MS Calc. Mass: 461.2090 (M+H)⁺, found: 461.2084
 (electrospray +ve) 1.3 ppm
 Melting point 245.0-246.2 °C

***N*¹-(9-Phenanthrene)-anthroyl-pyridine-4-carboxamidrazone [233]**

Structure



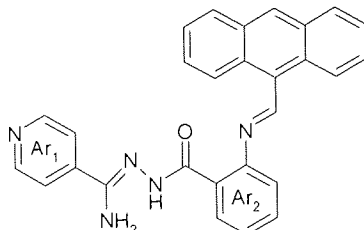
Physical state White crystalline solid
 Yield 23%
 1H NMR (D6-DMSO; $\delta_{\text{DMSO}}=2.50\text{ppm}$) 6.73-6.79 (overlapping m, 5H, NH₂, 3 × Ar-H), 7.20-7.28 (overlapping m, 3H, Ar₁2-H, Ar₁6-H, 1 × Ar-H), 7.40 (s, 1H, CH=N), 7.60-7.70 (overlapping m, 4H, 4 × Ar-H), 7.79 (d, 1H, J=7Hz, Ar₂4-H), 7.95(d, 1H, J=7.5Hz, Ar₂6-H), 8.12 (s, 1H, NH), 8.35 (d, 2H, J=6.5Hz, Ar₁3-H, Ar₁5-H), 8.79-8.93 (overlapping m, 3H, 3 × Ar-H)

IR (KBr) 3458, 3307, 3054, 2916, 1638(C=O), 1607, 1541, 1485, 1448, 1378, 1330, 1248, 1201, 1140(C-N) cm^{-1} .

Accurate MS Calc. Mass: 444.1824 (M+H)⁺, found: 444.1817
 (electrospray +ve) 1.6 ppm
 Melting point 245.9-247.0 °C

***N*¹-(9-Anthryl)-anthroyl-pyridine-4-carboxamidrazone [234]**

Structure



Physical state White crystalline solid
 Yield 55%
 1H NMR (D6-DMSO; $\delta_{\text{DMSO}}=2.50\text{ppm}$) 6.75-6.81 (overlapping m, 5H, NH₂, 3 × Ar-H), 7.22-7.24 (overlapping m, 3H, 3 × Ar-H), 7.42 (s, 1H, CH=N), 7.61-7.73 (overlapping m, 4H, Ar₁2-H, Ar₁6-H, 2 × Ar-H), 7.81(d, 1H, J=7Hz, Ar₂4-H), 7.97(dd, 1H, J=7Hz, Ar₂6-H), 8.142(s, 1H, NH), 8.37(d, 2H, J=6.5Hz, Ar₁3-H, Ar₁5-H), 8.79-8.95(overlapping m, 3H, 3 × Ar-H).

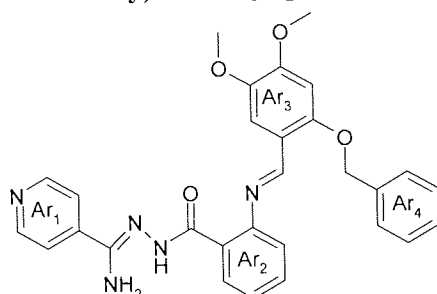
IR (KBr) 3459, 3304, 3058, 2923, 2883, 1639(C=O), 1606, 1543, 1486, 1381, 1295, 1249, 1164(C-N) cm^{-1} .

Accurate MS Calc. Mass: 444.1824 (M+H)⁺, found: 444.1801
 (electrospray +ve) 5.2 ppm

Melting point 249.2-250.6 °C

***N*¹-(2-Benzoyloxy-4,5-dimethoxy)-anthroyl-pyridine-4-carboxamidrazone [235]**

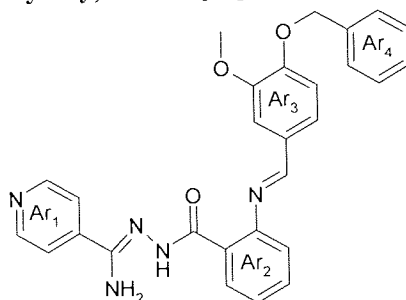
Structure



Physical state White crystalline solid
Yield 34%
1H NMR (D6-DMSO; $\delta_{\text{DMSO}}=2.50\text{ppm}$) 3.63(s, 3H, OCH₃), 3.69(s, 3H, OCH₃) 5.14 (s, 2H, CH₂), 6.42 (s, 1H, **CH=N**), 6.67-6.95 (overlapping m, 4H, NH₂, Ar₂3-H, Ar₂5-H), 6.71-6.77 (overlapping m, 2H, Ar₃3-H, Ar₃5-H), 7.23-7.40 (overlapping m, 7H, 5 × Ar₄-H, Ar₂4-H, NH), 7.48-7.52 (overlapping m, 4H, Ar₁2-H, Ar₁6-H, Ar₃2-H, Ar₃6-H), 7.68 (dd, 1H, J=7.58, 1.26 Hz Ar₂6-H), 8.55-8.58 (m, 2H, Ar₁3-H, Ar₁5-H).
IR (KBr) 3416, 3328, 3259, 3056, 3033, 3008 (CH₃), 2932, 2857, 1646(C=O), 1619, 1571, 1513, 1447, 1393, 1330, 1314, 1313, 1278, 1214, 1193(C-N) cm⁻¹.
Accurate MS Calc. Mass: 510.2141 (M+H)⁺, found: 510.2140
(electrospray +ve) 0.2 ppm
Melting point 211.9-213.3 °C

***N*¹-(3-Methoxy-4-benzyloxy)-anthroyl-pyridine-4-carboxamidrazone [236]**

Structure



Physical state White crystalline solid
Yield 80%
1H NMR (D6-DMSO; $\delta_{\text{DMSO}}=2.50\text{ppm}$) 3.73 (s, 3H, OCH₃), 5.03 (s, 2H, CH₂), 5.91 (s, 1H, **CH=N**), 6.68-7.02 (overlapping m, 6H, NH₂, Ar₂3-H, Ar₂5-H, 3 × Ar₃-H), 7.18 (s, 1H, NH), 7.23-7.37 (overlapping m, 6H, 5 × Ar₄-H, Ar₂4-H), 7.49 (dd, 2H, J=4.42 1.90 Hz, Ar₁2-H, Ar₁6-H), 7.67 (dd, 1H, J=8.21, 1.26 Hz Ar₂6-H), 8.56 (dd, 2H, J=4.42 1.90 Hz, Ar₁3-H, Ar₁5-H).
IR (KBr) 3418, 3313, 3262, 3060, 3029 (CH₃), 2963, 2928, 1658(C=O),

1612, 1544, 1498, 1477, 1374, 1353, 1267, 1228, 1189, 1139(C-N) cm^{-1} .

APCI

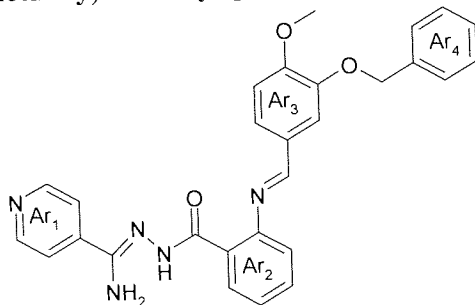
(M+1) $m/z=480$

Melting point

145.7-148.6 $^{\circ}\text{C}$

***N*¹-(3-Benzoyloxy-4-methoxy)-anthroyl-pyridine-4-carboxamidrazone [237]**

Structure



Physical state

White crystalline solid

Yield

48%

¹H NMR (D6-DMSO;

$\delta_{\text{DMSO}}=2.50\text{ppm}$)

3.70 (s, 3H, OCH₃), 4.99 (ab system, 2H, J=15.16 Hz, CH₂), 5.91 (s, 1H, CH=N), 6.68-677 (overlapping m, 2H, Ar₂3-H, Ar₂5-H), 6.87-6.90 (overlapping m, 3H, NH₂, Ar₃2-H, Ar₃5-H), 7.08 (dd, 1H, J=8.21, 1.90 Hz, Ar₃6-H), 7.21-7.52 (overlapping m, 9H, 5 × Ar₄-H, NH, Ar₂4-H, Ar₁2-H, Ar₁6-H), 7.68 (dd, 1H, J=7.58, 1.26 Hz, Ar₂6-H), 8.56 (dd, 2H, J=4.42, 1.90 Hz, Ar₁3-H, Ar₁5-H).

IR (KBr)

3461, 3326, 3062 (CH₃), 2963, 2840, 1625(C=O), 1606, 1540, 1506, 1484, 1448, 1376, 1253, 1131(C-N) cm^{-1} .

Accurate MS (electrospray +ve)

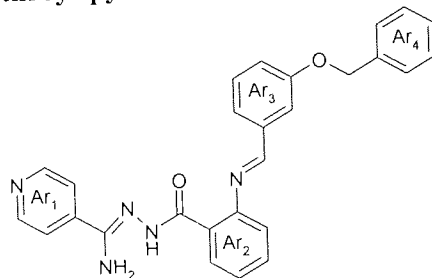
Calc. Mass: 480.2036 (M+H)⁺, found: 486.2037
0.2 ppm

Melting point

192.1-194.3 $^{\circ}\text{C}$

***N*¹-(3-Benzoyloxy)-anthroyl-pyridine-4-carboxamidrazone [238]**

Structure



Physical state

White crystalline solid

Yield

31%

¹H NMR (D6-DMSO;

$\delta_{\text{DMSO}}=2.50\text{ppm}$)

5.05 (s, 2H, CH₂), 5.91 (s, 1H, CH=N), 6.68-676 (overlapping m, 2H, Ar₂3-H, Ar₂5-H), 6.89-6.91 (overlapping m, 3H, NH₂, 2 × Ar₃-H), 7.11-7.41 (overlapping m, 9H, 5 × Ar₄-H, 2 × Ar₃-H, NH, Ar₂4-H), 7.51 (dd, 2H, J=4.42, 1.90 Hz, Ar₁2-H, Ar₁6-H), 7.68 (dd, 1H, J=7.58, 1.26 Hz, Ar₂6-H), 8.56 (dd, 2H, J=4.42, 1.90 Hz, Ar₁3-H, Ar₁5-H).

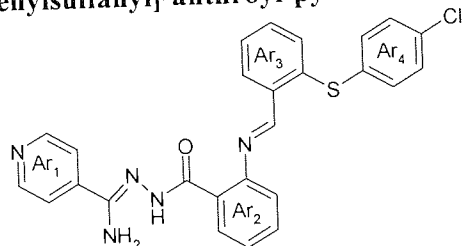
IR (KBr)

3448, 3262, 3156, 2920, 1650(C=O), 1633, 1539, 1448, 1391, 1318, 1295, 1268, 1238, 1128(C-N) cm^{-1} .

Accurate MS Calc. Mass: 450.1930 (M+H)⁺, found: 450.1948
 (electrospray +ve) 4.0 ppm
 Melting point 201.7-202.4 °C

***N*¹-[2-(4-Chloro)-phenylsulfanyl]-anthroyl-pyridine-4-carboxamidrazone [239]**

Structure

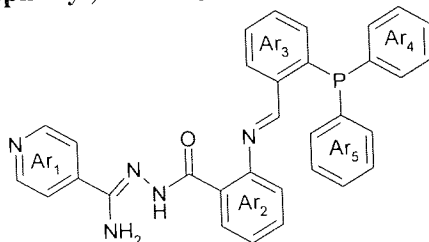


Physical state White crystalline solid
 Yield 77%
 1H NMR (D6-DMSO; δDMSO=2.50ppm) 6.57 (s, 1H, CH=N), 6.70-6.75 (overlapping m, 2H, Ar₂3-H, Ar₂5-H), 6.94 (bs, 2H, NH₂) 7.21-7.46 (overlapping m, 11H, 4 × Ar₄-H, 3 × Ar₃-H, Ar₂4-H, NH, Ar₁2-H, Ar₁6-H), 7.71 (m, 1H, Ar₃6-H), 7.88 (d, 1H, J=6.95 Hz Ar₂6-H), 8.58 (dd, 2H, J=4.42, 1.26 Hz, Ar₁3-H, Ar₁5-H).
 IR (KBr) 3487, 406, 3331, 3275, 3043, 3014, 2943, 1648(C=O), 1622, 1569, 1500, 1487, 1436, 1367, 1354, 1290, 1161(C-N) cm⁻¹.

Accurate MS Calc. Mass: 486.1155 (M+H)⁺, found: 486.1178
 (electrospray +ve) 4.7 ppm
 Melting point 193.4-194.8 °C

***N*¹-(2-Di-phenylphosphanyl)-anthroyl-pyridine-4-carboxamidrazone [240]**

Structure

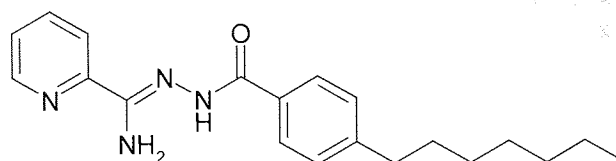


Physical state White crystalline solid
 Yield 76%
 1H NMR (D6-DMSO; δDMSO=2.50ppm) 6.71-6.76 (overlapping m, 2H, Ar₂3-H, Ar₂5-H), 6.88-6.92 (overlapping m, 3H, NH₂, Ar₃3-H), 6.96-7.00 (2s, 1H, keto and enol form of NH) 7.17-7.38 (overlapping m, 16H, 5 × Ar₄-H, 5 × Ar₅-H, 2 × Ar₃-H, 1 × Ar₂-H, Ar₁2-H, Ar₁6-H, CH=N), 7.69 (m, 1H, Ar₂6-H), 7.92 (m, 1H, Ar₃6-H), 8.49 (dd, 1H, J=4.42, 1.90 Hz, Ar₁3-H, Ar₁5-H).
 IR (KBr) 3492, 3333, 3048, 3000, 1634(C=O), 1625, 1544, 1482, 1432, 1394, 1345, 1318, 1289, 1258, 1243, 1207, 1158(C-N) cm⁻¹.

Accurate MS Calc. Mass: 528.1953 (M+H)⁺, found: 528.1950
 (electrospray +ve) 0.6 ppm
 Melting point 217.1-220.1 °C

*N*¹-(4-Heptylbenzoyl)-pyridine-2-carboxamidrazone [301]

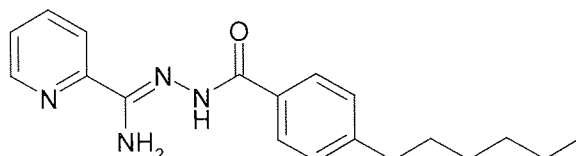
Structure



Physical state White crystalline solid
Yield 30%
1H NMR (D6-DMSO; δ DMSO=2.50ppm) 0.85 (t, 3H, J=6.9 Hz, CH₃), 1.26 (overlapping m, 8H, (CH₂)₄-CH₃), 1.59 (t, 2H, J=1.6 Hz, CH₂-(CH₂)₄), 2.63 (overlapping m, 2H, Ar-CH₂), 6.93 (s, 2H, NH₂), 7.30 (d, 2H, J=8.0 Hz, Ar3-H, Ar5-H), 7.49 (overlapping m, 1H, Py4-H), 7.79(d, 2H, J=8.0 Hz, Ar2-H, Ar6-H) 7.91 (overlapping m, 1H, Py5-H), 8.17 (m, 1H, Py6-H), 8.60 (d, 1H, H=4.3 Hz, Py3-H), 10.12 (s, 1H, NH)
IR (KBr) 3411, 3214, 3060, 2927, 1631(C=O), 1548, 1479, 1448, 1397, 1311, 1187(C-N) cm⁻¹.
MS (APCI +ve) m/z=339 (M+H)⁺.
Melting point 215.3-216.1 °C

*N*¹-(4-Hexylbenzoyl)-pyridine-2-carboxamidrazone [302]

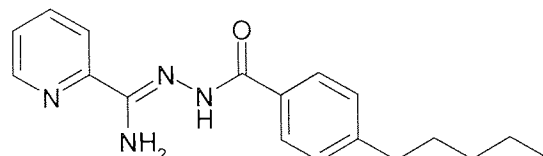
Structure



Physical state White powder
Yield 46%
1H NMR (D6-DMSO; δ DMSO=2.50ppm) 0.86 (t, 3H, J=6.8 Hz, CH₃), 1.28 (overlapping m, 6H, (CH₂)₃-CH₃), 1.59 (overlapping m, 2H, CH₂-(CH₂)₄), 2.63 (t, 2H, J=7.6 Hz, Ar-CH₂), 6.93 (s, 2H, NH₂), 7.30 (d, 2H, J=8.3 Hz, Ar3-H, Ar5-H), 7.48 (overlapping m, 1H, Py4-H), 7.79(d, 2H, J=8.3 Hz, Ar2-H, Ar6-H) 7.90 (ddd, 1H, J=8.9, 7.6, 1.3 Hz, Py5-H), 8.17 (d, 1H, J=8.2 Hz, Py6-H), 8.60 (d, 1H, H=4.3 Hz, Py3-H), 10.11 (s, 1H, NH)
IR (KBr) 3407, 3214, 3060, 2925, 2857, 1630(C=O), 1548, 1479, 1448, 1397, 1309, 1187(C-N) cm⁻¹.
Accurate MS Calc. Mass: 325.2028 (M+H)⁺, found: 325.1999
(electrospray +ve) 8.9 ppm
Melting point 218.7-220.6 °C

*N*¹-(4-Pentylbenzoyl)-pyridine-2-carboxamidrazone [303]

Structure



Physical state Light yellow crystalline solid
Yield 45%

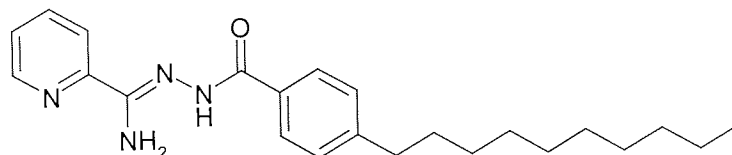
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 0.86 (t, 3H, J=6.6 Hz, CH₃), 1.26 (overlapping m, 4H, (CH₂)₂-CH₃), 1.59 (overlapping m, 2H, CH₂-(CH₂)₄), 2.63 (t, 2H, J=7.6 Hz, Ar-CH₂), 6.93 (s, 2H, NH₂), 7.28 (d, 2H, J=8.3 Hz, Ar3-H, Ar5-H), 7.48 (overlapping m, 1H, Py4-H), 7.78(d, 2H, J=8.5 Hz, Ar2-H, Ar6-H) 7.91 (ddd, 1H, J=8.9, 7.6, 1.3 Hz, Py5-H), 8.17 (d, 1H, J=8.3 Hz, Py6-H), 8.60 (d, 1H, H=5.0 Hz, Py3-H), 10.11 (s, 1H, NH)

IR (KBr) 3411, 3216, 3031, 2995, 2859, 1637(C=O), 1559, 1474, 1448, 1403, 1313, 1187(C-N) cm⁻¹.

Accurate MS (electrospray +ve) Calc. Mass: 311.1872 (M+H)⁺, found: 311.1836
 11.6 ppm
 Melting point 234.1-235.9 °C

*N*¹-(4-Decylbenzoyl)-pyridine-2-carboxamidrazone [304]

Structure



Physical state White crystalline solid
 Yield 32%

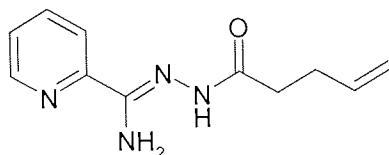
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 0.85 (t, 3H, J=6.8 Hz, CH₃), 1.26 (overlapping m, 14H, (CH₂)₇-CH₃), 1.58 (overlapping m, 2H, CH₂-(CH₂)₄), 2.63 (t, 2H, J=7.6 Hz, Ar-CH₂), 6.93 (s, 2H, NH₂), 7.30 (d, 2H, J=8.0 Hz, Ar3-H, Ar5-H), 7.48 (overlapping m, 1H, Py4-H), 7.78(d, 2H, J=7.5 Hz, Ar2-H, Ar6-H) 7.91 (ddd, 1H, J=8.9, 7.6, 1.3 Hz, Py5-H), 8.17 (d, 1H, J=8.0 Hz, Py6-H), 8.60 (d, 1H, H=4.5 Hz, Py3-H), 10.11 (s, 1H, NH)

IR (KBr) 3397, 3276, 3182, 3056, 2921, 2849, 1629(C=O), 1584, 1540, 1440, 1398, 1295, 1193(C-N) cm⁻¹.

Accurate MS (electrospray +ve) Calc. Mass: 381.2654 (M+H)⁺, found: 381.2639
 3.9 ppm
 Melting point 201.2-203.2 °C

*N*¹-(4-Pentenoyl)-pyridine-2-carboxamidrazone [305]

Structure



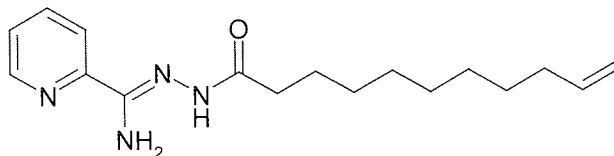
Physical state White crystalline solid
 Yield 9%

¹H NMR (D6-DMSO; δDMSO=2.50ppm) 1.13-1.27 (overlapping m, 2H, CH₂-CH), 2.28 and 2.70 (2×m, 2H, keto and enol form of CH₂-C=O), 4.95-5.10 (overlapping m, 2H, CH₂=), 5.78 (overlapping m, 1H, CH=), 6.63 (overlapping m, 2H, NH₂), 7.44 (overlapping m, 1H, Py4-H), 7.83 (overlapping m, 1H, Py5-H), 8.05 (overlapping m, 1H, Py6-H), 8.56 (overlapping m, 1H, Py3-H), 9.78, 9.91 (2 singlets, 1H, keto and enol form of NH)

IR (KBr) 3397, 3276, 3182, 3056, 2921, 2849, 1629(C=O), 1584, 1540, 1440, 1398, 1295, 1193(C-N) cm^{-1} .
 Accurate MS Calc. Mass: 219.1246 (M+H)⁺, found: 219.1213
 (electrospray +ve) 15.1 ppm
 Melting point 191.5-193.9 °C

***N*¹-(4-(10-Undecenoyl))-pyridine-2-carboxamidrazone [306]**

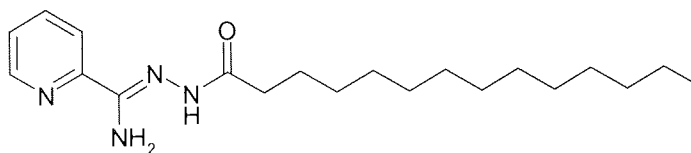
Structure



Physical state White crystalline solid
 Yield 9%
 1H NMR (D6-DMSO; $\delta_{\text{DMSO}}=2.50\text{ppm}$) 1.13-1.27 (overlapping m, 10H, (CH₂)₅-CH₂-CH=), 1.57 (overlapping m, 2H, CH₂-CH₂-C=O), 1.99(overlapping m, 2H, CH₂-CH=), 2.17-2.60 (2×m, 2H, keto and enol form of CH₂-C=O), 4.90-5.02 (overlapping m, 2H, CH₂=), 5.77 (overlapping m, 1H, CH=), 6.61 (overlapping m, 2H, NH₂), 7.44 (overlapping m, 1H, Py4-H), 7.86 (overlapping m, 1H, Py5-H), 8.05 (overlapping m, 1H, Py6-H), 8.56 (overlapping m, 1H, Py3-H), 9.73, 9.85 (2 singlets, 1H, keto and enol form of NH)
 IR (KBr) 3414, 3347, 3220, 3060, 2925, 2849, 1640(C=O), 1610, 1553, 1477, 1448, 1401, 1272, 1253, 1220, 1174(C-N) cm^{-1} .
 Accurate MS Calc. Mass: 303.2185 (M+H)⁺, found: 303.2160
 (electrospray +ve) 8.2 ppm
 Melting point 120.9-124.8 °C

***N*¹-(4-Myristoyl)-pyridine-2-carboxamidrazone [307]**

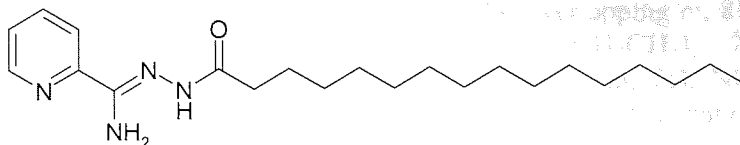
Structure



Physical state White crystalline solid
 Yield 33%
 1H NMR (D6-DMSO; $\delta_{\text{DMSO}}=2.50\text{ppm}$) 0.84 (t, 3H, J=6.5Hz, CH₃), 1.15-1.22 (overlapping m, 20H, - (CH₂)₁₀CH₃), 1.56 (overlapping m, 2H, COCH₂CH₂), 2.18 (overlapping m, 2H, COCH₂), 6.61 (overlapping m, 2H, NH₂), 7.44 (overlapping m, 1H, Py4-H), 7.86 (overlapping m, 1H, Py5-H), 8.05 (overlapping m, 1H, Py6-H), 8.56 (overlapping m, 1H, Py3-H), 9.72, 9.84 (2 singlets, 1H, keto and enol form of NH)
 IR (KBr) 3417, 3204, 3044, 2915, 2852, 1640(C=O), 1582, 1558, 1481, 1448, 1401, 1328, 1256, 1147(C-N), 1102 cm^{-1} .
 Accurate MS Calc. Mass: 347.2811 (M+H)⁺, found: 347.2791
 (electrospray +ve) 5.8 ppm
 Melting point 123.0-125.3 °C

*N*¹-(4-Palmitoyl)-pyridine-2-carboxamidrazone [308]

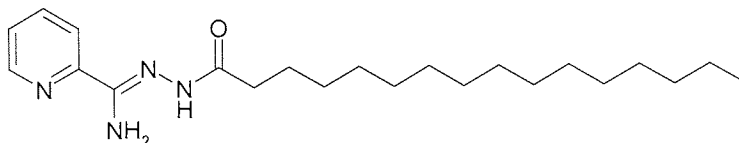
Structure



Physical state	White crystalline solid
Yield	40%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	0.84 (t, 3H, J=6.5Hz, CH ₃), 1.14-1.22 (overlapping m, 24H, - (CH ₂) ₁₂ CH ₃), 1.38 (overlapping, 2H, COCH ₂ CH ₂), 2.18 (overlapping m, 2H, COCH ₂), 6.60(overlapping m, 2H, NH ₂), 7.44 (overlapping m, 1H, Py4-H), 7.86 (overlapping m, 1H, Py5-H), 8.05 (overlapping m, 1H, Py6-H), 8.56 (overlapping m, 1H, Py3-H), 9.72,9.84 (2singlets, 1H, keto and enol from of NH)
IR (KBr)	3417, 3222, 3060, 2915, 2847, 1644(C=O), 1582, 1558, 1467, 1448, 1397, 1226, 1180(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 375.3124 (M+H) ⁺ , found: 375.3.98 6.9 ppm
Melting point	75.9-80.4 °C

*N*¹-(4-Stearoyl)-pyridine-2-carboxamidrazone [309]

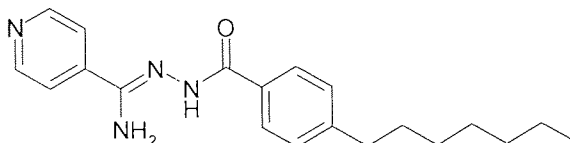
Structure



Physical state	White crystalline solid
Yield	23%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	0.85 (t, 3H, J=6.5Hz, CH ₃), 1.14-1.38 (overlapping m, 28H, - (CH ₂) ₁₄ CH ₃), 1.53 (overlapping, 2H, COCH ₂ CH ₂), 2.20 (overlapping m, 2H, COCH ₂), 6.61 (overlapping m, 2H, NH ₂), 7.44 (overlapping m, 1H, Py4-H), 7.85 (overlapping m, 1H, Py5-H), 8.06 (overlapping m, 1H, Py6-H), 8.55 (overlapping m, 1H, Py3-H), 9.73,9.84 (2singlets, 1H, keto and enol from of NH)
IR (KBr)	3418, 3222, 3058, 2920, 2847, 1644(C=O), 1584, 1558, 1477, 1447, 1397, 1226, 1181(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 403.3437 (M+H) ⁺ , found: 403.3424 3.2 ppm
Melting point	118.6-121.4 °C

*N*¹-(4-Heptylbenzoyl)-pyridine-4-carboxamidrazone [310]

Structure



Physical state	Yellow crystalline solid
----------------	--------------------------

Yield 27%

¹H NMR (D6-DMSO; δDMSO=2.50ppm) 0.85 (overlapping m, 3H, CH₃), 1.14-1.28 (overlapping m, 8H, - (CH₂)₄CH₃), 1.59 (overlapping, 2H, ArCH₂CH₂), 2.63 (overlapping m, 2H, ArCH₂), 6.90(overlapping m, 2H, NH₂), 7.30 (d, 2H, J=8.2 Hz, Ar3-H, Ar5-H), 7.72-7.84 (overlapping m, 4H, Py2-H, Py6-H, Ar2-H, Ar6-H), 8.65(d, 2H, J=5.7Hz Py3-H, Py5-H), 10.06 (s, 1H, NH)

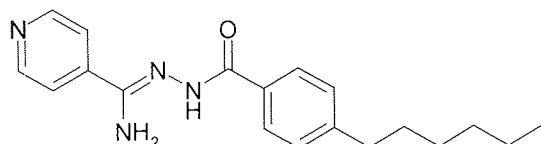
IR (KBr) 3417, 3204, 3044, 2915, 2852, 1640(C=O), 1582, 1558, 1498, 1411, 1295, 1218, 1183(C-N) cm⁻¹.

Accurate MS Calc. Mass: 339.2185 (M+H)⁺, found: 339.2170
(electrospray +ve) 4.4 ppm

Melting point 182.2-182.8 °C

*N*¹-(4-Hexylbenzoyl)-pyridine-4-carboxamidrazone [311]

Structure



Physical state Yellow crystalline solid

Yield 23%

¹H NMR (D6-DMSO; δDMSO=2.50ppm) 0.86 (overlapping m, 3H, CH₃), 1.14-1.28 (overlapping m, 6H, - (CH₂)₃CH₃), 1.59 (overlapping, 2H, ArCH₂CH₂), 2.63 (overlapping m, 2H, ArCH₂), 6.90(overlapping m, 2H, NH₂), 7.30 (d, 2H, J=8.2 Hz, Ar3-H, Ar5-H), 7.71-7.83 (overlapping m, 4H, Py2-H, Py6-H, Ar2-H, Ar6-H), 8.65 (d, 2H, J=5.7Hz Py3-H, Py5-H), 10.06 (s, 1H, NH)

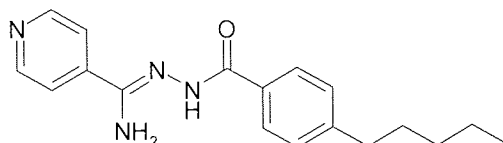
IR (KBr) 3430, 3207, 3064, 2923, 2856, 1636(C=O), 1594, 1552, 1502, 1405, 1301, 1210, 1187(C-N) cm⁻¹.

Accurate MS Calc. Mass: 325.2028 (M+H)⁺, found: 325.2000
(electrospray +ve) 8.6 ppm

Melting point 187.7-191.0 °C

*N*¹-(4-Pentylbenzoyl)-pyridine-4-carboxamidrazone [312]

Structure



Physical state Yellow crystalline solid

Yield 38%

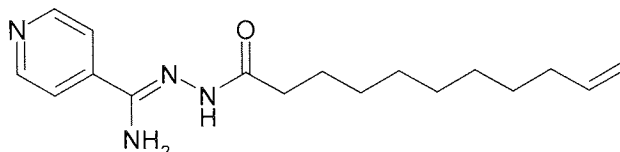
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 0.86 (overlapping m, 3H, CH₃), 1.14-1.30 (overlapping m, 4H, - (CH₂)₂CH₃), 1.58 (overlapping, 2H, ArCH₂CH₂), 2.66 (overlapping m, 2H, ArCH₂), 6.90(overlapping m, 2H, NH₂), 7.30 (d, 2H, J=8.2 Hz, Ar3-H, Ar5-H), 7.71-7.83 (overlapping m, 4H, Py2-H, Py6-H, Ar2-H, Ar6-H), 8.65 (d, 2H, J=5.7Hz Py3-H, Py5-H), 10.06 (s, 1H, NH)

IR (KBr) 3349, 3166, 3040, 2956, 2927, 1625(C=O), 1598, 1544, 1501, 1411, 1291, 1183(C-N) cm⁻¹.

Accurate MS Calc. Mass: 311.1872 (M+H)⁺, found: 311.1879
 (electrospray +ve) 2.2 ppm
 Melting point 201.2-201.8 °C

N¹-(10-Undecenoyl)-pyridine-4-carboxamidrazone [313]

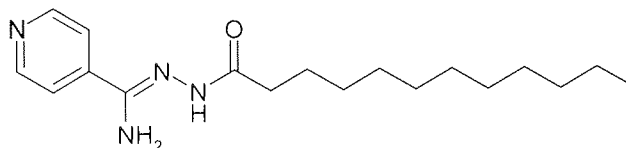
Structure



Physical state White crystalline solid
 Yield 4.3%
 1H NMR (D6-DMSO; δDMSO=2.50ppm) 1.14-1.27 (overlapping m, 10H, (CH₂)₅-CH₂-CH=), 1.43-1.56 (overlapping m, 2H, CH₂-CH₂-C=O), 2.22-2.97 2.60 (overlapping m, 4H, CH₂-CH₂-C=O), 4.91-5.02 (overlapping m, 2H, CH₂=), 5.79 (m, 1H, CH=), 6.59 (m, 2H, NH₂), 7.71-7.82 (overlapping m, 2H, Py2-H, Py6-H), 8.59-8.72 (overlapping m, 1H, Py3-H, Py5-H), 9.65, 9.81 (2 singlets, 1H, keto and enol form of NH)
 IR (KBr) 3427, 3230, 3079, 2919, 2847, 1704, 1654(C=O), 1596, 1532, 1469, 1415, 1378, 1225, 1191(C-N) cm⁻¹.
 Accurate MS Calc. Mass: 303.2185 (M+H)⁺, found: 303.2157
 (electrospray +ve) 9.2 ppm
 Melting point 113.0-114.2 °C

N¹-(Lauroyl)-pyridine-4-carboxamidrazone [314]

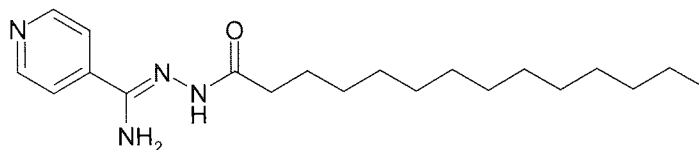
Structure



Physical state White crystalline solid
 Yield 31%
 1H NMR (D6-DMSO; δDMSO=2.50ppm) 0.83 (overlapping m, 3H, CH₃), 1.14-1.48 (overlapping m, 16H, -(CH₂)₈CH₃), 1.52 (overlapping m, 2H, COCH₂CH₂), 2.10 (overlapping m, 2H, COCH₂), 6.53(s, 2H, NH₂), 7.71-7.79 (overlapping m, 2H, Py2-H, Py6-H), 8.58-8.67 (overlapping m, 2H, Py3-H, Py5-H), 9.63, 9.80 (2 singlets, 1H, keto and enol form of NH)
 IR (KBr) 3457, 3365, 3232, 2915, 2850, 2358, 1685, 1637(C=O), 1593, 1544, 1409, 1400, 1214, 1180(C-N) cm⁻¹.
 Accurate MS Calc. Mass: 319.2498 (M+H)⁺, found: 319.2466
 (electrospray +ve) 10.0 ppm
 Melting point 121.2-124.4 °C

*N*¹-(Myristoyl)-pyridine-4-carboxamidrazone [315]

Structure



Physical state	White crystalline solid
Yield	40%
¹ H NMR (D ₆ -DMSO; δ _{DMSO} =2.50ppm)	0.84 (overlapping m, 3H, CH ₃), 1.14-1.23 (overlapping m, 20H, -(CH ₂) ₁₀ CH ₃), 1.56 (overlapping m, 2H, COCH ₂ CH ₂), 2.18 (overlapping m, 2H, COCH ₂), 6.59(s, 2H, NH ₂), 7.71-7.79 (overlapping m, 2H, Py2-H, Py6-H), 8.58-8.67 (overlapping m, 2H, Py3-H, Py5-H), 9.66,9.80 (2singlets, 1H, keto and enol from of NH)
IR (KBr)	3463, 3350, 3237, 2919, 2846, 2358, 1650(C=O), 1588, 1540, 1471, 1409, 1278, 1169(C-N) cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 347.2811 (M+H) ⁺ , found: 347.2787
Melting point	6.9 ppm
	116.8-117.5 °C

Chapter 4

Thiosemicarbazones Active

Against *C. difficile*

4.1 Introduction

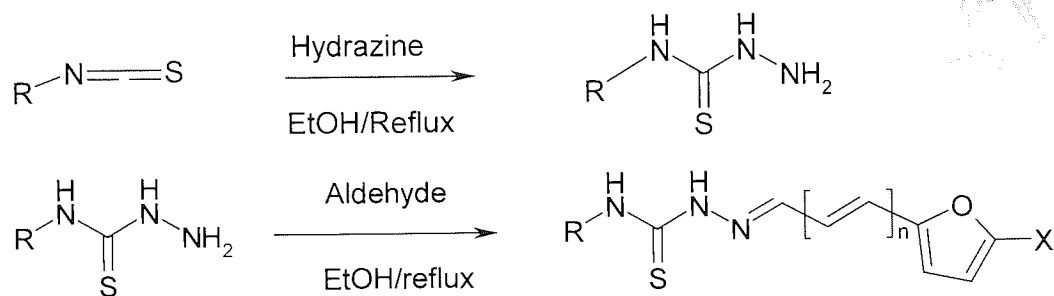
Clostridium difficile associated diarrhoea (CDAD) is a growing concern in the UK hospitals. Although CDAD is a preventable illness, the number of patients with CDAD has increased dramatically over the last decade. Among these patients, elderly people are at the greatest risk. Over 80% of *clostridium difficile* associated diseases reported are in people aged over sixty-five years old. According to a Health Protection Agency report for *clostridium difficile* infections, *MRSA bacteraemia* and *GRE bacteraemia* (July 2008), there were 45, 334 cases of *clostridium difficile* infected patients whose ages were over sixty-five in the UK from April 2007 to March 2008. In contrast, there are only 10, 059 cases of them aged under sixty-five in the same period. Therefore, control of *C. difficile* infections has become a priority in UK's hospital.

In recent years, as the antibiotics have spread, many strains of *C. difficile* have developed resistance to many common antibiotics. There are several examples of this. In the last two decades, fluoroquinolones (FQs) have been used as low risk drugs for treating CDAD (Golledge *et. al.* 1992). However, Patrizia Spigaglia *et. al.* (2008) showed that some of the strains were resistant to moxifloxacin (MX) and fluoroquinolones (FQs) in a 2005 study. At the same time in 2008, Jennifer R. O'Connor *et. al.* reported that *C. difficile* showed resistance to rifaximin which is a

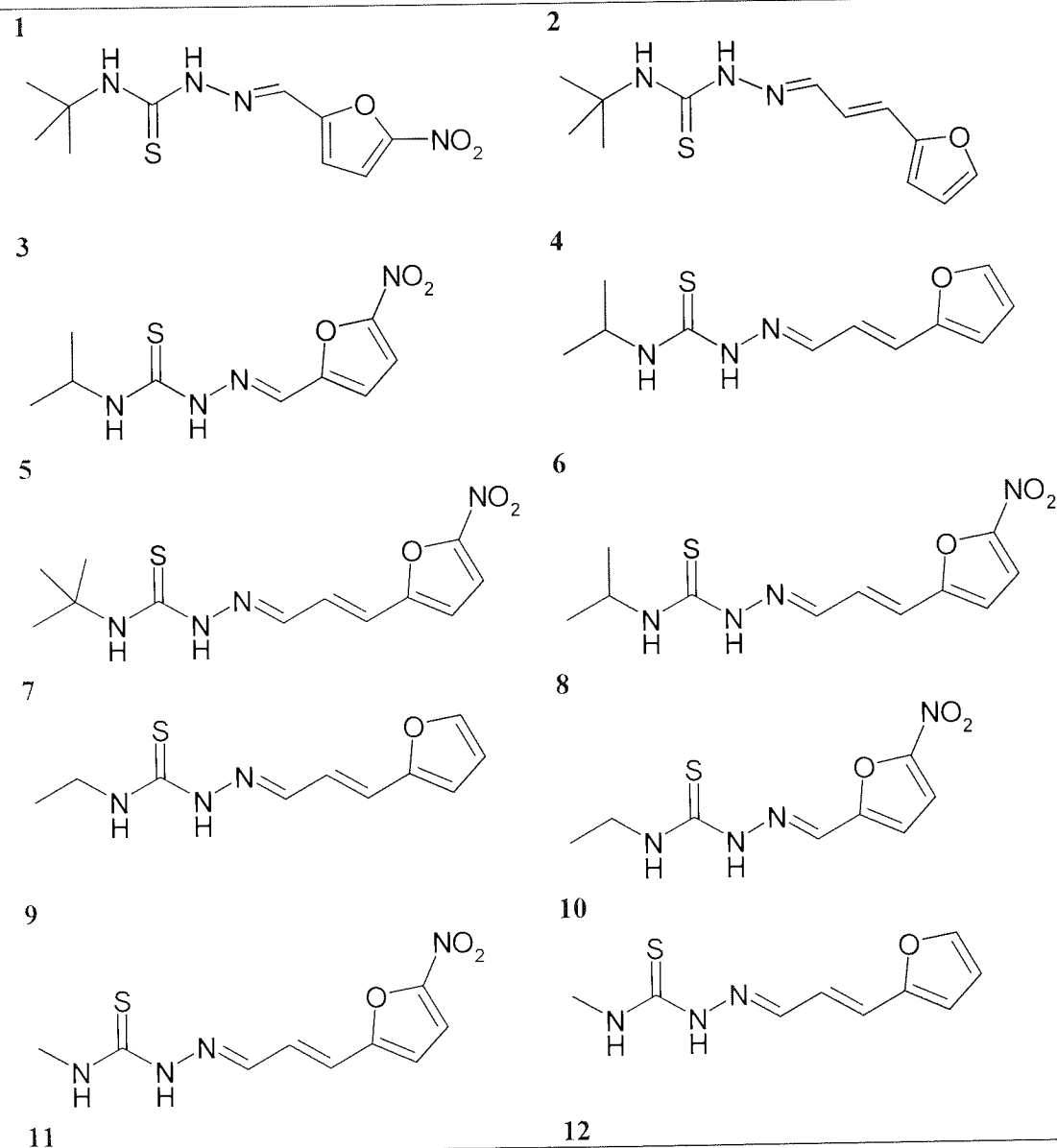
derivative of rifamycin. Rifaximin is an alternative for treating *C. difficile* infections. They found that some of the strains had a very high resistance with MIC \geq 32 μ g/mL. In early 2006, Norén *et. al.* stated that some strains of *C. difficile* had resistance to fusidic acid and metronidazole. From these examples, it is clear that normal antibiotics have been less effective in recent years. Therefore, finding novel antimicrobial agents against *Clostridium Difficile* is highly desirable.

Thiosemicarbazones have already been evaluated over the last several decades for their activities against a lot of microorganisms (Condit *et. al.* 1991). Many papers showed that they have anti-trypanosomal properties (Aguirre *et. al.* 2004, Greenbaum *et. al.* 2004, Otero *et. al.* 2006). In addition, Nardi *et.al.* (1967) also shows that thiosemicarbazones have antibacterial properties. Several papers show that the presence of a nitro group on the fifth position in the furyl ring is a requirement for their antimicrobial properties. Research demonstrated that nitro substituent contained thiosemicarbazones' biological activity is the result of a complex pathway of reduction of the nitro group to amino derivatives (Pires 2001). Greenbaum's report showed that their leading thiosemicarbazones have none or low toxicity in mice.

A set of substituted furyl thiosemicarbazones were synthesised and then tested against a panel of Gram-positive bacteria including *C. difficile*, *S. aureus* (both methicillin sensitive and resistant strains), *Staphylococcus epidermidis*, *Propionibacterium acnes*. The thiosemicarbazones were synthesised according to **Scheme 4.1.** and all these structures are shown in **Table 4.1.**



Scheme 4.1: Synthesis scheme of thiosemicarbazones compounds.



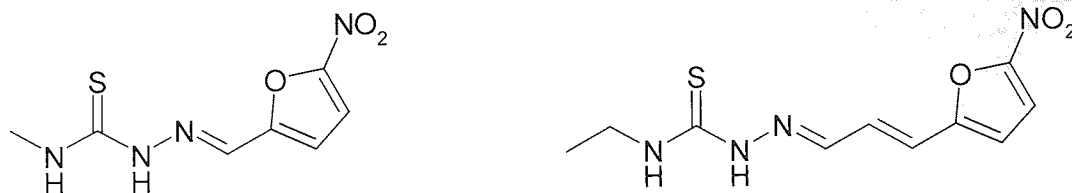


Table 4.1: Structure of synthesised thiosemicarbazones

The purpose of synthesising these compounds was to find out the effects of different sub-structure groups on anti-microbial activities. These sub-structure include the length of alkyl group at the N-4 position, the presence of nitro group at the 5-position on the furan ring and the ethylene spacer group in the furylidene portion.

4.2 Results and Discussion

4.2.1 Chemistry of Thiosemicarbazones

Overall, 11 thiosemicarbazones and 4 thiosemicarbazide intermediates were synthesized and characterized.

Because they have similar structures, they share similar physical and chemical properties and we can find some trends among these compounds. There are four intermediates shown as following:

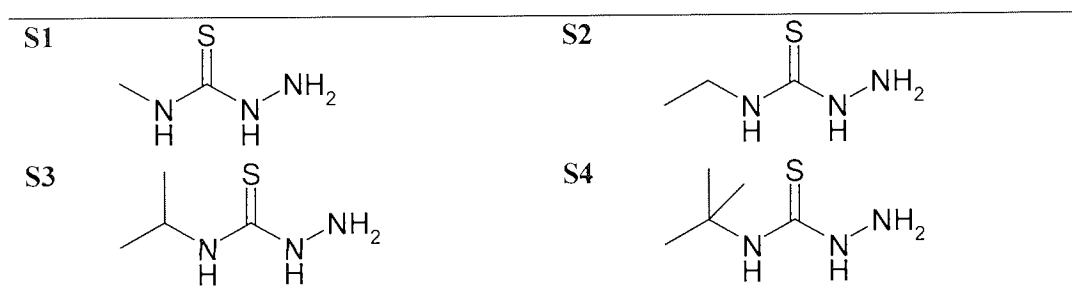
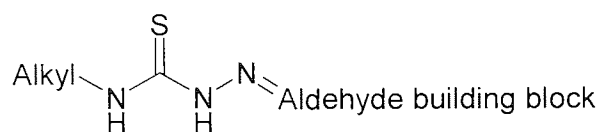


Table 4.2: Structure of four intermediates

Although they share the same preparation procedure and similar structures, some of the compounds show different properties beyond the general trend. Among those four thiosemicarbazides, the melting point for compound **S4**, **S3**, **S2** are 137.7-138.8 °C, 90.1-94.6 °C and 74.1-78.1 °C. This happens because **S4** has the greatest molecular weight and **S3** and **S2** have less molecular weight. However, the melting point of compound **S1** is tested 131.5-134.4 °C which is higher than **S3** and **S2**. This may be due to the crystal packing order of the molecules. **S1** has the lowest molecular weight and may pack well in solid state because there is a minimal interference from the alkyl group to the intermolecular hydrogen bonding. The bulkier alkyl substituents tend to disrupt this hydrogen bonding and this results in a lower melting point.

The preparation of 4 thiosemicarbazide intermediates was conducted by adding hydrazine into isothiocyanate. After 15 hours, reaction was stopped. Most of the solvent and hydrazine were removed under vacuum. Water was added and solid products appeared which was clean to use for the next step. As the solid precipitated in water after the reaction finished, the yield should be related to their solubility in water. In theory, **S4** should be the most hydrophobic due to the *tert*-butyl group and **S1** should be the most hydrophilic due to the methyl group. Theoretically, **S4** should have the greatest yield and **S1** should have the smallest yield. However, we can see that yield of **S4** is 47.1%, **S3** is 26.4%, **S2** is 78.8% and **S1** is 48.7%. The yield of **S2** and **S1** are even higher than the yield of **S4**. This happens because in the precipitation step of **S2** and **S1**, if only water (50mL) was added in, no solid appeared which means that **S2** and **S1** are easily soluble in water. In order to get solid products, a small amount of water and hexane was added to the crude material. This time, the solid precipitated and was collected. The adding of hexane decreases the solubility of **S2** and **S1**. That is why these two compounds are more hydrophilic while they still have higher yield.

The physical and chemical properties of the 11 thiosemicarbazones are more complex. However, they can be divided into three groups depending on which aldehyde was added.



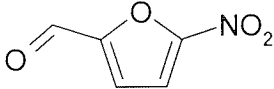
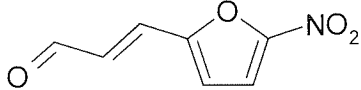
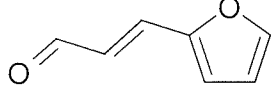
Aldehyde building block used	Comp. NO	Alkyl group	Yield (%)	Melting point (°C)
Group I 	1	<i>tert</i> -Butyl	77.3	183.1-186.3
	3	Isopropyl	72.7	192.5-194.8
	8	Ethyl	58.8	169.4-173.7
	11	Methyl	91.2	228.5-230.2
Group II 	5	<i>tert</i> -Butyl	93.4	Decomposed at 209
	6	Isopropyl	85.8	219.4-222.0
	9	Methyl	98	Decomposed at 223.5
Group III 	2	<i>tert</i> -Butyl	57.4	172.0-175.2
	4	Isopropyl	50.5	163.5-165.2
	7	Ethyl	27.2	128.7-132.8
	10	Methyl	29.5	140.5-143.7

Table 4.3: Comparison of thiosemicarbazones' yield and melting point

Firstly, the yield was compared. It is quite clear that compounds in Group II have the higher yield. Compounds in Group III have the lower yield. This is because the presence of the NO₂ group makes the compounds poorly soluble in organic solvents and most of the products precipitate from the reaction. At the same time, compounds having NO₂ group also have higher melting points and it seems that the compounds in Group II are easily decomposed at high temperature. Therefore, the presence of the NO₂ group affects the molecular properties. Comparing groups I and II, it can be seen that the length of exocyclic conjugation also affects their properties. It seems that more conjugation makes the compound less soluble in organic solvents. However, it also makes compounds prone to thermal decomposition rather than having a clean melting range.

It also can be seen that the trends in yields and melting points are more or less related. For this particular method of isolating the compounds, a higher yield means that the

compound has lower solubility and higher melting point owing to stronger intermolecular attraction. The melting point trend within each group is the same as for the thiosemicarbazide compounds. Compounds with methyl groups had higher melting points because of good crystal packing.

Some of these compounds have already been made by Aguirre *et. al.* (2004). These are compounds **8**, **9** and **11**. **11** has a higher melting point than the literature value and **9** decomposed at higher temperature than the literature value of 200 °C. In contrast, **8** melted 17 °C lower than the literature value. It should be noted here, however, that the results presented in this thesis are corrected values whereas the quoted literature provides uncorrected melting points.

Then, their NMR data were compared. In general, there are several trends to discuss. The most obvious one is that the presence of a nitro group can move the chemical shift of hydrogen to a higher ppm value. This is because the nitro group is a strong electron withdrawing group and we can see that all the compounds are highly conjugated. This makes the nitro group easier to withdraw electron from other atoms. We can compare the NMR data of **2** and **5**. The only structural difference between these two compounds is that **5** has a nitro group on the fifth position of the furyl ring. These two sets of the NMR data show that *tert*-butyl group is virtually unaffected. One of them is 1.50 ppm and the other one is 1.51 ppm. Except this, the chemical shifts of hydrogen on **5** are all higher than those on **2**. What is more, if the hydrogen is closer to the nitro group, the bigger chemical shift value difference it will have. For instance, the chemical shift of furylH-4 on **5** is 7.76 ppm. The same hydrogen on **2** is 6.57 ppm. Their difference is 1.19 ppm. Further from the nitro group, let us choose NH-N for comparison. The chemical shift difference is only 0.24 ppm. In **4** and **6**, **9** and **10**, they all show the same trend of nitro group effects. After this, we can look at the difference between coupling constants. Usually, the coupling constants on the furyl ring are between 3.8 and 4.4 Hz. The coupling constants on the allylic part are much bigger. Here we can see that one of them is 15.8 and another one is 9.5 Hz. This

happens because the distance and the orientation of those coupled hydrogens is different. The calculated distance of hydrogen on furly ring position 3-4 is 2.8 Å and they are on the same side. Two hydrogen atoms in allylic part are in a *trans* orientation and the distance is 3.0 Å. Generally, the coupling constants of hydrogen in a furyl ring is around 3.2-3.8 Hz and the *trans* state of alkenes is around 11-18 Hz. This trend is very helpful for identifying the hydrogens if the NMR data is complex. In contrast to the nitro group, alkyl chains are electron-donating groups. Nearly all the chemical shifts can be affected by an alkyl chain. Let us take **1** and **3** as examples. **1** has a *tert*-butyl group and **3** has an isopropyl group. If we compare the chemical shift of all the hydrogens except the alkyl chain, we find that all the chemical shifts in **1** are lower than those in **3**. This is because that *tert*-butyl group can donate more electron density which will push all the chemical shift to a lower ppm. Nitro groups not only affect the melting point and NMR data, it can also affect the products' colour. Usually, compounds with nitro group will have a darker colour. For example, **6** and **9** are all dark red. **11**, **8** and **5** are all orange. These compounds all have a nitro group on the furyl ring. The compounds which do not have nitro group usually have lighter colour such as **7** and **4** whose colour is lighter yellow.

4.2.2 Microbiology Screening of Thiosemicarbazones

All these thiosemicarbazones were tested against a group of Gram-positive organisms in nutrient broth. Those included *C. difficile*, *S. epidermidis*, *S. aureus*, MRSA and *P. acnes*. The minimum inhibitory concentration (MIC) was determined first. Compound **12** was found exhibit significant antimicrobial activity against all the organisms. The data is shown in **Table 3** (Vancomycin acts as reference). Nearly all the MICs were less than 8 µg/mL, except *C. difficile* O27 which is 8-16 µg/mL, *S. aureus* H1 which is 16-32 µg/mL, *P. acnes* B5 which is 16-32 µg/mL and *P. acnes* C5 which is 32-64 µg/mL.

Organisms	1	2	3	4	5	6
<i>C. difficile</i> NCTC11204	128-256	>256	>256	>256	>256	>256
<i>C. difficile</i> O27	>256	64-128	64-128	64-128	64-128	64-128
<i>C. difficile</i> Z162	128-256	>256	>256	>256	128-256	>256
<i>C. difficile</i> Z229	16-32	64-128	128-256	128-256	64-128	64-128
<i>S. epidermidis</i> NCTC11047	128-256	>256	64-128	>256	>256	64-128
<i>S. epidermidis</i> 9	>256	>256	>256	>256	>256	>256
<i>S. epidermidis</i> 4	>256	>256	>256	>256	>256	64-128
<i>S. aureus</i> ATCC6538	64-128	>256	64-128	>256	>256	>256
<i>S. aureus</i> H1	>256	>256	128-256	>256	64-128	>256
<i>S. aureus</i> K1	64-128	>256	>256	>256	32-64	>256
MRSA A1	>256	>256	>256	>256	>256	>256
MRSA C3	>256	>256	>256	>256	>256	>256
MRSA E1	>256	>256	>256	>256	>256	>256
<i>P. acnes</i> NCTC737	64-128	64-128	64-128	32-64	16-32	32-64
<i>P. acnes</i> B5	32-64	64-128	>256	>256	32-64	32-64
<i>P. acnes</i> C3	16-32	4-8	32-64	1-2	8-16	32-65
<i>P. acnes</i> C5	16-32	1-2	64-128	64-128	16-32	64-128

Continued table below

	7	8	9	10	11	12
<i>C. difficile</i> NCTC11204	>256	64-128	64-128	>256	>256	4-8
<i>C. difficile</i> O27	64-128	64-128	32-64	64-128	64-128	8-16
<i>C. difficile</i> Z162	>256	128-256	32-64	>256	>256	0.125- 0.25
<i>C. difficile</i> Z229	64-128	8-16	16-32	128-256	64-128	4-8
<i>S. epidermidis</i> NCTC11047	>256	32-64	32-64	>256	16-32	4-8
<i>S. epidermidis</i> 9	>256	64-128	64-128	>256	128-256	1-2
<i>S. epidermidis</i> 4	>256	32-64	32-64	128-256	32-64	4-8
<i>S. aureus</i> ATCC6538	>256	32-64	32-64	>256	64-128	4-8
<i>S. aureus</i> H1	>256	16-32	16-32	>256	64-128	16-32
<i>S. aureus</i> K1	>256	64-128	>256	>256	128-256	4-8
MRSA A1	>256	64-128	64-128	>256	64-128	2-4
MRSA C3	>256	64-128	64-128	>256	>256	4-8
MRSA E1	>256	64-128	64-128	>256	64-128	4-8

<i>P. acnes</i> NCTC737	16-32	16-32	16-32	32-64	16-32	4-8
<i>P. acnes</i> B5	>256	32-64	1-2	32-64	64-128	8-16
<i>P. acnes</i> C3	32-64	2-4	1-2	16-32	16-32	2-4
<i>P. acnes</i> C5	32-64	32-64	64-128	1-2	64-128	32-64

Table 4.4: MIC ($\mu\text{g/mL}$) of all the compounds against an initial panel of organisms

As can be seen from **Table 4.4** that **12** had the lowest MIC against *C. difficile* Z162 which is only 0.125-0.25 $\mu\text{g/mL}$. Besides this, nearly all of the compounds had less activity against *P. acnes* C3, e.g. **4** and **9** had a MIC value 1-2 $\mu\text{g/mL}$, **8** and **12** had a MIC value 2-4 $\mu\text{g/mL}$ and **2** had a MIC value 4-8 $\mu\text{g/mL}$. Overall, although compounds **1-11** had less activities against *C. difficile*, *S. epidermidis*, *S. aureus* and MRSA, some of them still had a good activity against *P. acnes*. **Figure 1** clearly shows their MIC difference.

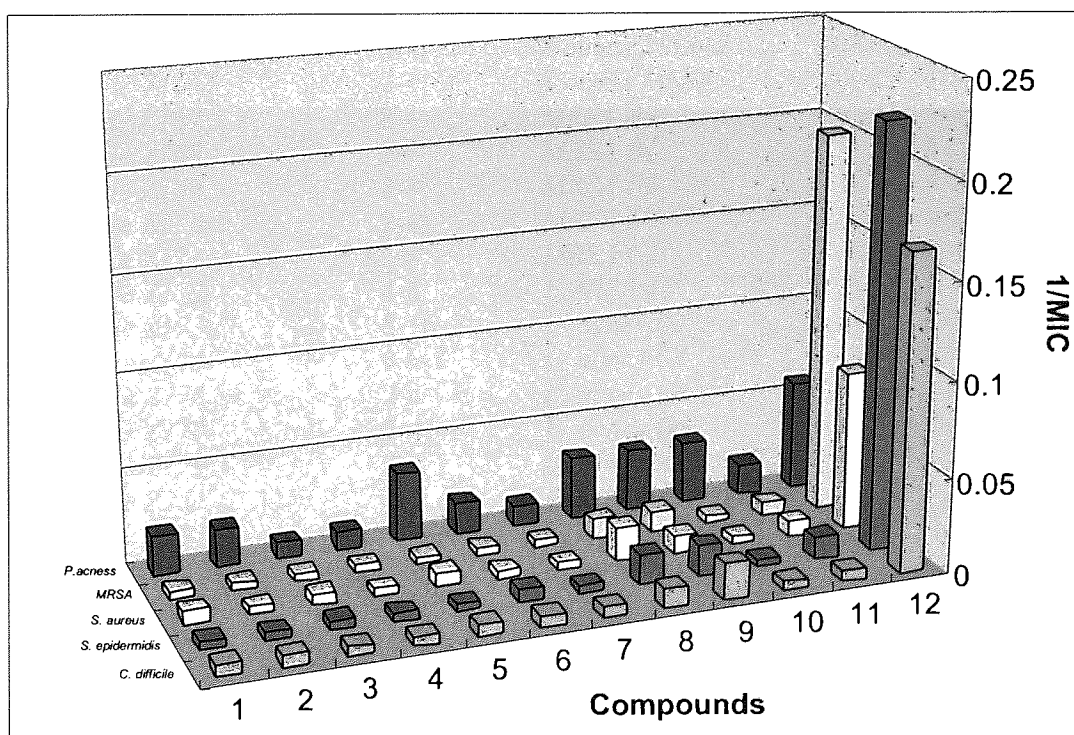


Figure 4.1: Bar chart of averaged 1/MIC ($\mu\text{g/mL}$) for compounds 1-12 for the various classes of organisms tested

It can be seen that, apart from **12**, **8** and **9**, there have no noticeable antimicrobial activities. By looking at their structures, it can be seen that all these compounds have 5-nitrofuransubstructure. It also can be seen that except **3**, **4** and **7**, other compounds are all show some activities against *P. acnes*.

After MIC, compound **12** was tested against an expanded panel of organisms. **Table 4.5** shows these results.

Organisms	MIC µg/mL	Organisms	MIC µg/mL
<i>C. difficile</i> Z16	8-16	<i>S. epidermidis</i> RP62A	4-8
<i>C. difficile</i> Z21	16-32	<i>S. epidermidis</i> 3	2-4
<i>C. difficile</i> Z66	4-8	<i>S. epidermidis</i> 7A	2-4
<i>C. difficile</i> Z150	4-8	<i>S. epidermidis</i> 8A	2-4
<i>C. difficile</i> Z158	2-4	<i>S. epidermidis</i> 10	4-8
<i>C. difficile</i> Z159	8-16	<i>S. epidermidis</i> 11	1-2
<i>C. difficile</i> Z160	8-16	<i>S. aureus</i> T1	4-8
<i>C. difficile</i> Z161	0.5-1	<i>S. aureus</i> V1	2-4
<i>C. difficile</i> Z163	2-4	<i>S. aureus</i> H1	64-128
<i>C. difficile</i> Z164	1-2	<i>S. aureus</i> MSSA2	4-8
<i>C. difficile</i> Z166	2-4	<i>S. aureus</i> MSSA3	4-8
<i>C. difficile</i> Z167	2-4	MRSA A3	4-8
<i>C. difficile</i> Z168	2-4	MRSA B1	4-8
<i>C. difficile</i> Z169	2-4	MRSA B3	4-8
<i>C. difficile</i> Z170	2-4	MRSA D1	4-8
<i>C. difficile</i> Z171	2-4	MRSA C1	2-4
<i>C. difficile</i> Z172	4-8	EMRSA 15	4-8
<i>C. difficile</i> Z173	8-16	<i>P. acnes</i> B3	6-16
<i>C. difficile</i> Z175	0.5-1	<i>P. acnes</i> C1	8-16
<i>C. difficile</i> Z176	2-4	<i>P. acnes</i> D1	4-8
<i>C. difficile</i> Z177	2-4	<i>P. acnes</i> D3	4-8
<i>C. difficile</i> Z179	2-4	<i>C. albicans</i> 1013781	128-256
<i>C. difficile</i> Z251	1-2	<i>C. albicans</i> 1012155	128-256
<i>C. difficile</i> Z387	16-32	<i>C. albicans</i> 1009116	128-256
<i>C. difficile</i> Z1568	1-2	<i>C. albicans</i> 1011377	128-256
<i>C. difficile</i> Z1577	0.5-1	<i>C. albicans</i> 1015291	128-256
<i>C. difficile</i> Z1578	1-2	<i>Acinetobacter</i> spp. 1	128-256
<i>C. difficile</i> Z1579	2-4	<i>Acinetobacter</i> spp. 2	128-256
<i>C. difficile</i> Z1580	1-2	<i>Acinetobacter</i> spp. 3	128-256
<i>C. difficile</i> Z1591	1-2	<i>Acinetobacter</i> spp. 4	128-256
<i>S. epidermidis</i> NCTC 9865	2-4	<i>Acinetobacter</i> spp. 5	128-256

Table 4.5: MIC (µg/mL) of compound **12** against an expanded panel of organisms.

From **Table 4.5**, it was clear that **12** has good activity against various kinds of Gram-positive organisms. However, there was no activity against Gram-negative microorganisms such as *Acinetobacter spp.* or yeast such as *C. albicans*. The mechanism of action compound **12** is still unknown at the present time. By calculating the LogP value of all 12 compounds (in **Table 4.6**), it was found that, except **5**, **7**, **10** and **11**, other compounds' LogP values are very similar to compound **12**. The cellular penetrations of these compounds should be similar which means that all of them should not have a cellular penetration problem. Thus, their activities should be due to their structural difference. At the same time, Aguirre *et. al.* (2004) tested the *in vivo* toxicity of compounds **8** and **12**. This indicated these compounds are nontoxic.

Compound	cLogP	compound	cLogP	compound	cLogP
1	2.9	5	3.3	9	2.5
2	2.5	6	3.2	10	1.7
3	2.8	7	2.1	11	2.1
4	2.5	8	2.4	12	2.8

Table 4.6: LogP value of 12 compounds (done by CaChe WorkSystem Version 6.1.10, Fujitsu Ltd)

These compounds were also tested against M-TB H37Rv by TAACF. Eight of them were found to be active against M-TB H37Rv. **Table 4.7** shows these testing result. Their cytotoxicity was also tested and shown in **Table 4.8**.

Compounds	IC ₅₀ µg/mL	IC ₉₀ µg/mL	Compounds	IC ₅₀ µg/mL	IC ₉₀ µg/mL
6	1.608	2.127	9	3.025	3.285
7	2.790	2.982	5	3.230	3.807
2	2.827	3.075	4	5.672	6.028
10	2.928	3.222	11	5.783	6.337

Table 4.7: Testing result of thiosemicarbazones against M-TB H37Rv

Compounds	IC ₉₀ µg/mL	CC ₅₀ µg/mL	SI
6	2.127	3.517	1.653
7	2.982	17.647	5.917
2	3.075	11.202	3.642
10	3.222	23.577	7.317
9	3.285	3.173	0.965
5	3.807	3.680	0.966
4	6.028	16.751	2.778
11	6.337	4.149	0.654

Tab 4.8 Cytotoxicity assay of active thiosemicarbazones. (CC₅₀ gives the measure of mammalian cell toxicity. The selected Index SI shows the ratio of CC₅₀ and IC₉₀. SI ≥ 10 may be considered for the further screening.)

These two tables show that although these thiosemicarbazones were active against M-TB H37Rv, they were too cytotoxic to act as drug molecules. Cocco *et.al* (2002) and Sriram *et.al.* (2007) all published some other kind of thiosemicarbazones which were active against M-TB H37Rv and had low cytotoxicity. Therefore, a further modification on the molecule structure in order to reduce the cytotoxicity may be considered as further work.

4.3 Conclusion

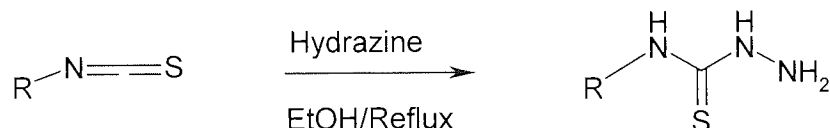
It can be seen that the structures of the thiosemicarbazones are similar, therefore, their chemical properties such as ^1H NMR chemical shift and melting point are closely related. Some general trends depend on these factors. However, there are some exceptions, for example those compounds with a methyl group have the lowest molecular weight but have a higher melting point. These compounds may pack well in the solid state because there is a minimal interference from the alkyl group to hinder intermolecular hydrogen bonding and the bulkier alkyl substituents tend to disrupt this hydrogen bonding and result in a lower melting point. At the same time, it is clear that the presence of a nitro group on the furan ring increased the yield and melting point significantly. All these compounds were tested against gram-positive bacteria including *C. difficile*, *S. epidermidis*, *S. aureus*, MRSA and *P. acnes*. Compound **12** N^1 -(ethyl)-2-(5-nitro-2-furyl-allylic)-hydrazine-1-carbothioamide was found to have the lowest MIC value against all these bacteria. This compound was also tested against an expanded panel organisms and found to be active against most of the organisms with some lower MIC values of $0.5\text{-}1.0\ \mu\text{g mL}^{-1}$. The mechanism of action of the compound is still unknown. However, the calculated LogP value showed that their cellular penetrations are similar which means that all of them should not have a cellular penetration problem. Eight of these thiosemicarbazones were found to be active against M-TB H37Rv. However, they were too cytotoxic to be used as drugs.

4.4 Experimental

Instrumentation

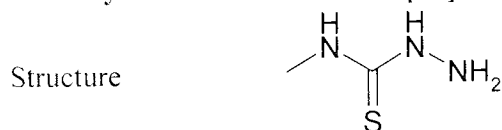
Proton NMR spectra were obtained on a Bruker AC 250 instrument operating at 250MHz as solutions in d_6 -DMSO or $CDCl_3$ and referenced from $\delta_{DMSO}=2.50$ ppm or $\delta_{CDCl_3}=7.26$ ppm unless otherwise stated. Infrared spectra were recorded as KBr discs or NaCl discs on a Mattson 3000 FTIR spectrophotometer. Atmospheric pressure chemical ionisation mass spectrometry (APCI-MS) was carried out on a Hewlett-Packard 5989B quadrupole instrument connected to an electrospray 59987A unit with an APCI accessory and automatic injection using a Hewlett-Packard 1100 series autosampler. Electrospray +ve was conducted on LCT Premier Electrospray manufactured by WATERS. Melting points were obtained using a Reichert-Jung Thermo Galen hot stage microscope and are corrected.

Preparation and characterization of thiosemicarbazide



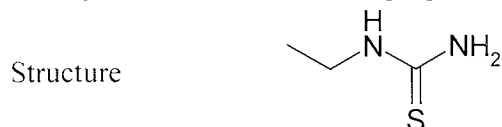
General procedure: the isothiocyanate (5g) dissolved in ethanol (20mL) was added dropwise into a solution of hydrazine monohydrate (30mL) in ethanol (30mL) which was stirred vigorously under reflux for 15hours. Most of the solvent and hydrazine were removed under vacuum. Water (50mL) was added and a solid appeared. The solid product was collected by filtration and washed with water and then dried under vacuum to give the thiosemicarbazide. This was used in the next step without purification.

4-Methyl-3-thiosemicarbazide [S1]



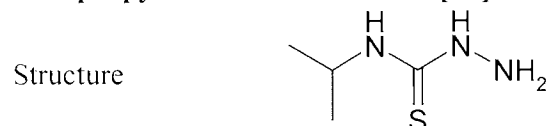
Physical state	White crystalline solid
Yield	3.5 g, 33.3 mmol, 48.7%
^1H NMR (D_6 -DMSO; $\delta_{DMSO}=2.50$ ppm)	2.88 (d, $J=4.4$ Hz, 3H, Me), 4.42 (s, 2H, NH_2), 7.83 (bs, 1H, NHMe), 8.59 (s, 1H, NHNH_2)
IR (KBr)	3285, 3236, 3140, 2956, 2932, 1613, 1637, 1563, 1464, 1266 (C=S), 1167, 1082, 1054, 1026 cm^{-1}
MS (APCI +ve)	106 (M+H) $^+$
Melting point	131.5-134.4 $^\circ\text{C}$

4-Ethyl-3-thiosemicarbazide [S2]



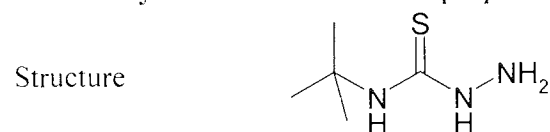
Physical state	Light yellow powder
Yield	5.39g, 45.3mmol, 78.8%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.06 (t, J=7.1Hz, 3H, Me), 3.45 (ms, 2H, CH ₂), 4.24 (s, 2H, NH ₂), 7.83 (t, J=5.1 Hz, 1H, NHCH ₂), 8.56 (s, 1H, NHHNH ₂)
IR (KBr)	3341, 3280, 3192, 2972, 1631, 1540, 1499, 1469, 1272 (C=S), 1161 cm ⁻¹
MS (APCI +ve)	120 (M+H) ⁺ 103 (M-NH ₃) ⁺
Melting point	74.1-78.1 °C

4-Isopropyl-3-thiosemicarbazide [S3]



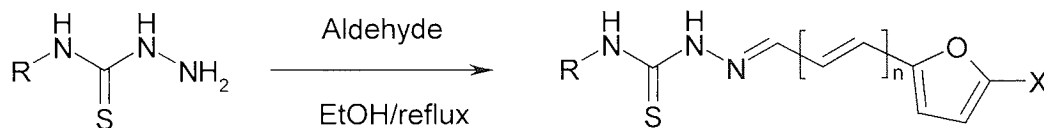
Physical state	Light yellow powder
Yield	1.74g, 13.1mmol, 26.4%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.12 (d, J=7.0 Hz, 6H, Me ₂), 4.44 (overlapping, 3H, NH ₂ & CH), 7.49 (d, J=8.2 Hz, 1H, NHCH), 8.56 (s, 1H, NHC=S)
IR (KBr)	3284, 3257, 3206, 2969, 2926, 1635, 1541, 1471, 1458, 1252 (C=S), 1174, 1063 cm ⁻¹
MS (APCI +ve)	134 (M+H), 117(M-NH ₃ ,)
Melting point	90.1-94.6 °C

4-tert-Butyl-3-thiosemicarbazide [S4]



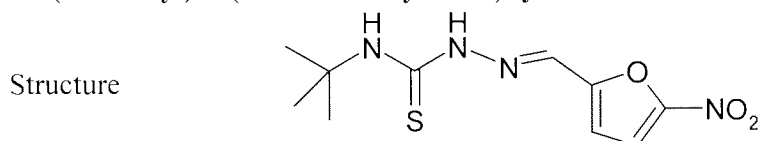
Physical state	White crystal
Yield	3.01g, 20.5mmol, 47.1%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.45 (s, 9H, Me ₃), 4.51 (s, 2H, NH ₂), 7.49 (s, 1H, NHC), 8.46 (s, 1H, NHC=S)
IR (KBr)	3339, 3311, 3200, 2961, 1625, 1531, 1498, 1399, 1354, 1272, 1227 (C=S) cm ⁻¹
MS (APCI +ve)	148 (M+H) ⁺
Melting point	137.7-138.8 °C

Preparation and characterization of thiosemicarbazones



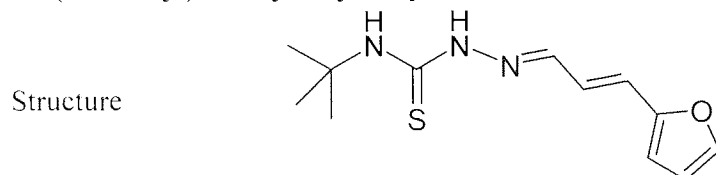
General procedure: A suspension of thiosemicarbazide (0.5g) and aldehyde (1.1 equivalents) in ethanol (15mL) was stirred and heated to reflux for 2 hours. The solid product was collected by filtration and washed with ethanol. Then, it was dried under vacuum to give the final product.

*N*¹-(*tert*-Butyl)-2-(5-nitro-2-furylidene)hydrazine-1-carbothioamide [1]



Physical state	Light yellow powder
Yield	0.71g, 2.6mmol, 77.3%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.53 (s, 9H, 3 × CH ₃), 7.31 (d, J=3.8 Hz, 1H, Fur 3-H), 7.58 (s, 1H, CH=N), 7.80 (d, J=4.4 Hz, 1H, fur 4-H), 7.99 (s, 1H, NH ₂), 11.84 (s, 1H, NH=N)
IR (KBr)	3436, 3318, 3125, 2975, 1520, 1482, 1348, 1251 (C=S), 1172, 1017 cm ⁻¹
Accurate MS (electrospray +ve)	Calc. Mass: 271.0865 (M+H) ⁺ , found: 271.0844
Melting point	183.1-186.3 °C

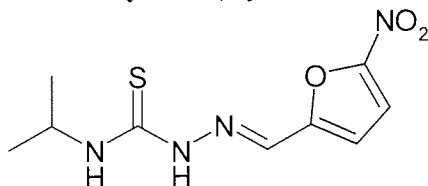
*N*¹-(*tert*-Butyl)-2-furyl-allylic-hydrazine-1-carbothioamide [2]



Physical state	Yellow crystalline solid
Yield	0.49g, 2.0mmol, 57.4%
¹ H NMR (D6-DMSO; δDMSO=2.50ppm)	1.50 (s, 9H, 3 × CH ₃), 6.57 (dd, J=3.2, 1.90 Hz, 1H, Fur 4-H), 6.65 (overlapping, 2H, Fur 3-H & CH=CH-Fur), 6.87 (d, J=15.8 Hz, 1H, CH-Fur), 7.45 (s, 1H, NH ₂), 7.75 (d, J=1.9 Hz, 1H, Fur 5-H), 7.82 (d, J=9.5 Hz, 1H, CH=N), 11.32 (s, 1H, NH=N)
IR (KBr)	3339, 3127, 2981, 2962, 1621, 1525, 1263 (C=S), 1181, 1079 cm ⁻¹
Accurate MS (electrospray +ve)	Calc. Mass: 274.0990 (M+Na) ⁺ , found: 274.1015
Melting point	172.0-175.2 °C

***N*¹-(Isopropyl)-2-(5-nitro-2-furylidene)hydrazine-1-carbothioamide [3]**

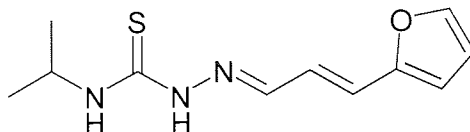
Structure



Physical state Yellow powder
Yield 0.70g, 2.7mmol, 72.7%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 1.22 (d, J=7.0 Hz 6H, 2 × CH₃), 4.50 (ms, 1H, CHMe₂), 7.41 (d, J=4.4, 1H, Fur 3-H), 7.84 (d, J=3.8 Hz, 1H, Fur4-H), 7.99 (s, 1H, CH=N), 8.13(d, J=8.2 Hz, 1H, NHCHMe₂), 11.88 (s, 1H, NH-N=)
IR (KBr) 3357, 3155, 3091, 2967, 1539, 1507, 1474, 1341, 1291, 1236 (C=S), 1213, 1148 cm⁻¹.
Accurate MS (electrospray +ve) Calc. Mass: 279.0528 (M+Na)⁺, found: 279.0506
Melting point 192.5-194.8 °C

***N*¹-(Isopropyl)-2-furyl-allylic-hydrazine-1-carbothioamide [4]**

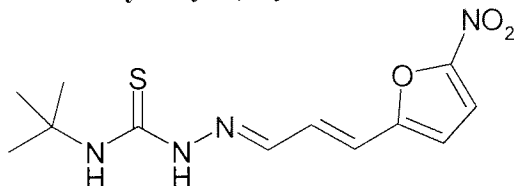
Structure



Physical state Grey powder
Yield 0.45g, 1.9mmol, 50.5%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 1.17 (d, J=7.0 Hz 6H, 2 × CH₃), 4.46 (ms, 1H, CHMe₂), 6.57 (dd, J=3.2, 1.9, 1H, Fur 4-H), 6.66 (overlapping, 2H, Fur3-H & CH=CH-Fur), 6.87(d, J=15.8 Hz 1H, CH-Fur), 7.75(d, J=1.9 Hz, 1H, Fur5-H), 7.82 (d, J=9.5 Hz 1H, CH=N), 7.93 (d, J=8.1 Hz, NHCHMe₂), 11.36 (s, 1H, NH-N=)
IR (KBr) 3369, 3137, 2974, 1627, 1530, 1512, 1287, 1246 (C=S), 1173, 1073 cm⁻¹.
Accurate MS (electrospray +ve) Calc. Mass: 260.0834 (M+Na)⁺, found: 260.0847
Melting point 163.5-165.2 °C

***N*¹-(*tert*-Butyl)-2-(5-nitro-2-furyl-allylic)-hydrazine-1-carbothioamide [5]**

Structure



Physical state Orange powder
Yield 0.94g, 3.2mmol, 93.4%
¹H NMR (D6-

DMSO; δ DMSO=2.50ppm)	H), 7.08 (d, J=3.8, 1H, CH=N), 7.57 (s, 1H, NHCMe ₃), 7.76(d, J=4.4 Hz 1H, Fur4-H), 7.87(dd, J=5.1 3.2 Hz, 1H, CH=CH-Fur), 11.56 (s, 1H, NH-N=)
IR (KBr)	3311, 3142, 2972, 1528, 1508, 1465, 1417, 1389, 1354, 1259, 1241 (C=S), 1215, 1198, 1171, 1137, 1020 cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 297.1021 (M+H) ⁺ , found: 297.1013
Melting point	Decomposed at 209.5 °C

N¹-(Isopropyl)-2-(5-nitro-2-furyl-allylic)-hydrazine-1-carbothioamide [6]

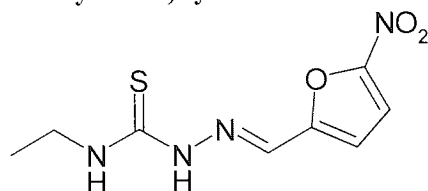
Structure	
Physical state	Dark red powder
Yield	0.91g, 3.2mmol, 85.8%
¹ H NMR (D6-DMSO; δ DMSO=2.50ppm)	1.18 (d, J=7.0 Hz 6H, 2 × CH ₃), 4.47 (ms, 1H, CHMe ₂) 7.04 (overlapping, 3H, CH-Fur, Fur3-H & CH=N), 7.75 (d, J=3.8 Hz, 1H, Fur4-H), 7.86(dd, J=7.0 1.9 Hz, 1H, CH=CH-Fur), 8.15 (d, J=8.9 Hz, 1H, NHCHMe ₂), 11.62 (s, 1H, NH-N=)
IR (KBr)	3357, 3319, 3122, 2971, 1530, 1518, 1464, 1357, 1346, 1255, 1230 (C=S), 1190 cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 283.0865 (M+H) ⁺ , found: 283.0845
Melting point	219.4-222.0 °C

N¹-(Ethyl)-2-furyl-allylic-hydrazine-1-carbothioamide [7]

Structure	
Physical state	Light yellow powder
Yield	0.26g, 1.2mmol, 27.2%
¹ H NMR (D6-DMSO; δ DMSO=2.50ppm)	1.10 (t, J=7.0 Hz 3H, Me), 3.53 (ms, 2H, CH ₂), 6.57 (dd, J=3.2, 1.9, 1H, Fur 4-H), 6.66 (overlapping, 2H, Fur3-H & CH=CH-Fur), 6.87(d, J=16.4 Hz 1H, CH-Fur), 7.74(d, J=1.9 Hz, 1H, Fur5-H), 7.82 (d, J=9.5 Hz 1H, CH=N), 8.31 (t, J=5.7 Hz, NHEt), 11.36 (s, 1H, NH-N=)
IR (KBr)	3360, 3131, 2974, 2355, 2334, 1628, 1533, 1514, 1234 (C=S), 1074 cm ⁻¹ .
Accurate MS (electrospray +ve)	Calc. Mass: 246.0677 (M+Na) ⁺ , found: 246.0719
Melting point	128.7-132.8 °C

***N*¹-(Ethyl)-2-(5-nitro-2-furylidene)hydrazine-1-carbothioamide [8]**

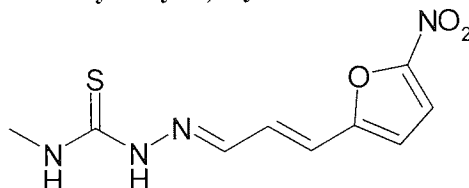
Structure



Physical state Orange powder
Yield 0.61g, 2.5mmol, 58.8%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 1.14 (t, J=7.3 Hz 3H, Me), 3.59 (ms, 2H, CH₂), 7.34 (d, J=3.8, 1H, Fur 3-H), 7.83 (d, J=3.8 Hz, 1H, Fur4-H), 7.98 (s, 1H, CH=N), 8.61(d, J=5.7 Hz, 1H, NH₂Et), 11.88 (s, 1H, NH-N=)
IR (KBr) 3435, 3358, 3335, 3151, 2970, 2356, 2337, 1545, 1472, 1353, 1246 (C=S), 1223, 1104, 1015 cm⁻¹.
Accurate MS (electrospray +ve) Calc. Mass: 265.0371 (M+Na)⁺, found: 265.0431
Melting point 169.4-173.7 °C (Literature value: 186.0-188.0 °C Aguirre *et al.* 2004)

***N*¹-(Methyl)-2-(5-nitro-2-furyl-allylic)-hydrazine-1-carbothioamide [9]**

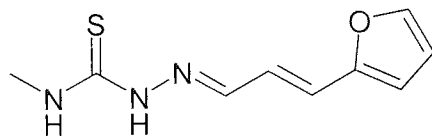
Structure



Physical state Dark red powder
Yield 1.21 g, 4.8mmol, 98%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 2.95 (d, J=5.1 Hz 3H, Me), 6.97 (overlapping, 3H, CH-Fur, Fur3-H & CH=N), 7.70 (d, J=3.8 Hz, 1H, Fur4-H), 7.84(t, J=4.4 Hz, 1H, CH=CH-Fur), 8.54 (q, J=4.4 Hz, 1H, NHMe), 11.62 (s, 1H, NH-N=)
IR (KBr) 3330, 3124, 2977, 2357, 2339, 1547, 1514, 1459, 1348, 1256, 1235 (C=S), 1088, 1015 cm⁻¹.
MS (APCI +ve) 255 (M+H)⁺
Melting point Decomposed at 223.5 °C (Literature value: 198.0-200.0 °C Aguirre *et al.* 2004)

***N*¹-(Methyl)-2-furyl-allylic-hydrazine-1-carbothioamide [10]**

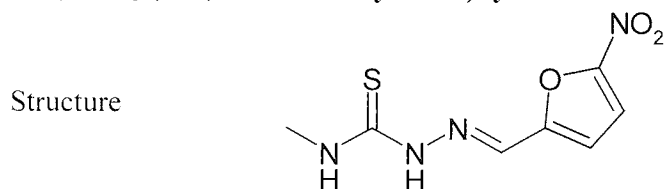
Structure



Physical state yellow powder
Yield 0.30g, 1.4mmol, 29.5%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 2.94 (d, J=5.1 Hz 3H, Me), 6.65 (dd, J=3.2, 1.9, 1H, Fur 4-H), 6.63 (overlapping, 2H, Fur3-H & CH=CH-Fur), 6.85(d, J=16.4 Hz 1H, CH-Fur), 7.70(d, J=1.9 Hz, 1H, Fur5-H), 7.79 (d, J=9.5 Hz 1H, CH=N), 8.33 (q, J=4.4 Hz, NHMe), 11.33 (s, 1H, NH-

IR (KBr) N=) 3350, 3299, 3266, 3159, 3108, 2998, 2939, 2356, 2338, 1622, 1545, 1519, 1468, 1252 (C=S), 1083, 1013 cm^{-1} .
 Accurate MS (electrospray +ve) Calc. Mass: 232.0521 (M+Na)⁺, found: 232.0554
 Melting point 7.1 ppm
 140.5-143.7 °C

***N*¹-(Methyl)-2-(5-nitro-2-furylidene)hydrazine-1-carbothioamide [11]**



Physical state Orange powder
 Yield 1.01g, 4.4mmol, 91.2%
¹H NMR (D6-DMSO; δDMSO=2.50ppm) 2.99 (d, J=4.4 Hz 3H, Me), 7.26 (d, J=3.8, 1H, Fur 3-H), 7.76 (d, J=3.8 Hz, 1H, Fur4-H), 7.95 (s, 1H, CH=N), 8.62(d, J=4.4 Hz, 1H, NHMe), 11.88 (s, 1H, NH-N=)
 IR (KBr) 3350, 3299, 3266, 3159, 3108, 2998, 2939, 2356, 2338, 1622, 1545, 1519, 1468, 1252 (C=S), 1083, 1013 cm^{-1} .
 Accurate MS (electrospray +ve) Calc. Mass: 251.0215 (M+Na)⁺, found: 251.0190
 Melting point 9.9 ppm
 228.5-230.2 °C (Literature value: 210.0-212.0 °C Aguirre *et.al.* 2004)

References

- Aguirre G., Boiani L., Cerecetto H., Fernandez M., Gonzalez M., Denicola A., Otero L., Gambino D., Rigol C., Olea-Azar C. & Faundez M. (2004) In vitro activity and mechanism of action against the protozoan parasite *Trypanosoma cruzi* of 5-nitrofuryl containing thiosemicarbazones. *Bioorg. Med. Chem.* 12: 4885-4893
- Allinger N. L. (1977) Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *J. Am. Chem. Soc.* 99: 8127-8134
- Allinger N. L., Yuh Y. H., and Lii J. H. (1989) Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* 111: 8551-8566
- Aqvist J., Medina C. & Samuelsson J. E. (1994) A new method for predicting binding affinity in computer-aid drug design, *Protein Eng.* 7: 385-391
- Banfi E., Mamolo M. G., Vio L. and Predominato E. (1993) *In-Vitro* antimycobacterial activity of new synthetic amidrazone derivatives. *J. Chemothe.* 5, 164-167
- Becker O. M., MacKerrell A. D., Jr., Roux B., Watanabe M., eds, Dekker Marcel (2001) Computational Chemistry and Biophysics: Chapter 3 "Dynamics Methods" pp39-67
- Beeman D. (1976) Some Multistep Methods for Use in Molecular Dynamics Calculations. *J. Comput. Phys* 20:130-139
- Bergmann E., Herman D. & Zimkin E. (1948) The condensation of 2,4-diamino-4-methylpentane with carbonyl compounds *J. Org. Chem.* 13: 353-356
- Bible R. H. (1965) *Interpretation of NMR spectra*. New York: Plenum Press
- Billington D.C., Coleman M. D., Ibiabuo J., Lambert P. S., Rathbone D. L. and Tims K. J. (1998) Synthesis and antimycobacterial activity of some heteroarylcarbox-amidrazone derivatives *Drug Des. Discov.* 15, 269-275
- Billington D. C., Lowe P. R., Rathbone D. L., Schwalbe C. H. (2000) a new amidrazone derivative with antimycobacterial activity. *Acta Cryst.* C56
- Bissantz C., Folkers G., Rognan D. (2000) Protein-Based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations *J. Med. Chem.* 43: 4759-4767

- Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E., & Thornton, J. M. (1987) Knowledge-Based Prediction of Protein Structures and the Design of Novel Molecules. *Nature* 326: 347-352.
- Böhm H. J. (1994) The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J. Comput. Aided Mol. Des.* 8: 243-256
- Böhm H. J. (1994) On the use of LUDI to search the fine chemicals directory for ligands of proteins of known three-dimensional structure *J. Comput. Aided Mol. Des.* 8: 623-632
- Bonnett R. & Emerson T. R. (1965) *J. Chem. Soc* 4508 (adapted from Smith M. B. & March J. 2001 *Adv. Org. chem.*. Wiley: New York
- Bornscheuer U. T. (2002) Microbial carboxyl esterase: classification, properties and application in biocatalysts *FEMS Microbiol. Rev.* 26 73-81
- Bowen, J. P., and Shim, J-Y., (1998) Molecular mechanics studies of acyl halides: I. Molecular structures and conformational analysis. *J. Comput. Chem.* 19: 1370-1386.
- Brook B. R., Bruccoleri R. E., Olafson B. D., States D. J., Swaminathan S. & Karplus M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4: 187-217
- Brown A. R., Rees D. C. Rankovic Z. & Morphy J. R. (1996) Synthesis of tertiary amines using a polystyrene (REM) resin. *J. Am. Chem. Soc.* 119: 3288-3295
- CaChes 6.1 user guide (2003) Fujitsu Limited.
- Camila S., de Magalhães, Hélio J. C., Barbosa & Dardenne L. E. (2004) A genetic algorithm for the ligand-protein docking problem. *Genet. Mol. Biol.* 27:605-610
- Case F. H. (1965) The preparation of hydrazideines and as-triazines related to substituted 2-cyanopyridine. *J. Org. Chem.* 30: 931-933
- Case J., Sundén H. & Córdova A. (2004) Direct organocatalytic asymmetric α -hydroxymethylation of ketones and aldehydes *Tetrahedron Lett.* 45: 6117-6119
- Chatzakis M. (2006) MPharm Programme student, Aston University
- Chauhan S. (2006) Biology placement student 2005/2006, Aston University
- Chong S. K. (2007) MPharm Programme student, Aston University

Chothia C., Lesk AM. (1986) The relation between the divergence of sequence and structure in proteins. *EMBO J.* 5: 823-836

Clark R. H. and E. C. Wagner (1944) Isatoic anhydride I. Reactions with primary and secondary amines and with some amides. *J. Org. Chem.* 9: 55-67

Cocco M. T., Congiu C. Onnis V., Pekkerabo M. L. & De Logu A. (2002) Synthesis and antimycobacterial activity of new *S*-alkylthiosemicarbazone derivatives *Bioorg. Med. Chem.* 10: 501-506

Coleman M. D., Rathbone D. L., Abberley L., Lambert P. A. & Bilington D. C. (1999) Preliminary in vitro toxicological evaluation of a series of 2-pyridylcarboxamidrazone candidate anti-tuberculosis compounds. *Environ. Toxicol. Pharmacol.* 7: 59-65

Coleman M. D., Tims K. J. & Rathbone D. L. (2003) The use of computational QSAR analysis in the toxicological evaluation of a series of 2-pyridylcarboxamidrazone candidate anti-tuberculosis compounds. *Environ. Toxicol. Pharmacol.* 14: 33-42

Dewar M. J. S. & Thiel W. (1997a) ground states of molecular. 38. the MNDO method Approximations and parameters. *J. Am. Chem. Soc.* 99:4899-4907

Dewar M. J. S. & Thiel W. (1997b) ground states of molecular. 39. MNDO result for molecular containing hydrogen, carbon, nitrogen and oxygen. *J. Am. Chem. Soc.* 99:4907-4917

Eisch J. J. & Sanchez R. (1985) Selective, oxophilic imination of ketones with bis(dichloroaluminum) phenylimide *J. Org. Chem.* 51: 1848-1852

Eldridge M., Murray C. W., Auton T. A., Paolini G. V. & Lee R. P. (1997) Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* 11: 425-445

Ercolessi F. (1997) *A molecular dynamics primer*. Trieste: ICTP

Ewing T. & Kuntz I. D. (1997) Critical evaluation of search algorithms for automated molecular docking and database screening. *J. Comput. Chem.* 18: 1175-1189

Fetrow, J. S. & Bryant, S. H. (1993) New Programs for Protein Tertiary Structure Prediction. *Bio/Technology* 11: 479-484.

Fletcher R. and Reeves C. M. (1964) Function minimization by conjugate gradients *The comput. J.* 7(2): 149-154

- Ghose A.K., Pritchett A. & Crippen GM. (1988) Atomic physicochemical parameters for 3-dimensional structure directed quantitative structure-activity-relationships. 3 modelling hydrophobic interactions. *J. Comput. Chem.* 9: 80-90
- Greenhow R. C. (1990) *Introductory quantum mechanics*. Bristol: Hilger
- Greer, J. (1991) Comparative Modeling of Homologous Proteins. *Meth. Enzymol.* 202: 239-252.
- Gohlke H., Hendlich M. & Klebe G. (2000) Knowledge-based scoring function to predict protein-ligand interactions. *J. Mol. Bio.* 295: 337-356
- Heslin, Julie C., Moody & Christopher J. (1988) *J. Chem. Soc., Perkin Trans.* 1: 1972-1999
- Hinchliffe A. (2006) *Molecular modelling for beginners* Sussex: Wiley
- Hockney R. W. (1970) The Potential Calculation and Some Applications. *Methods in comp. Phys.* 9:136-211
- Jones G., Willett P. & Glen R. C. (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* 245: 43-53
- Jorgensen W. L. and Tirado-Rives J. (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin *J. Am. Chem. Soc.* 110: 1657 – 1666
- Khan A. T. & Mindal E. (2003) *Synlett.* 694-698
- Khan N. (2007) PhD thesis, Aston University
- Kitchen D. B., Decornez H., Furr J. R. & Bajorath J. (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Review Drug Dis.* 3: 935-949
- Kollman P. A. (1996) Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. *Acc. Chem. Res.* 29: 461-469
- Kridger E., Nabuurs S. B. & Vriend G. (2003) *Structure Bioinformatics* Wiley-Liss Inc.
- Krovat E. M., Steindl T. & Langer T. (2005) Recent advances in docking and scoring. *Curr. Comput. Aided Drug Des.* 1: 93-102

Leach A. R. (2001) *Molecular modelling: principles and applications* Essex: Prentice Hall

Lewandowski K., Murer P., Svec F. & Fréchet J. M. J. (1998) the design of chiral separation media using monodisperse functionalised macroporous beads: effects of polymer matrix, tether, and linkage chemistry. *Anal. Chem.* 70: 1629-1638

Lipinski C. A., Lombardo F., Dominy B. W. & Feeney P. J. (1997) experimental and computational approaches to estimate solubility and permeability in drug deicoverly and development settings. *Adv. Drug Deliv. Rev.* 23: 3-25

Lowe M. E. (1992). The catalytic site residues and interfacial binding of human pancreatic lipase. *J. Biol. Chem.* 267 (24): 17069-73

Mamolo M. G., Vio, L., Banfi E., Predominanto M., Fabris C. and Asaro F. (1992) Synthesis and antimycobacterial activity of some 2-pyridinecarboxyamidrazone derivatives. *Farmaco.* 47, 1055-1066

Mamolo M. G., Vio L., Banfi E., Predominanto M., Fabris C. and Asaro F. (1993) Synthesis and antimycobacterial activity of some 4-pyridinecarboxyamidrazone derivatives. *Farmaco.* 48, 529-538

Mamolo M. G. and Vio L. (1996) Synthesis and antimycobacterial activity of some indole derivatives of pyridine-2-carboxamidrazone and quinoline-2-carboxamidrazone. *Farmaco.* 51, 65-70

Mamolo M. G., Falagiani V., Vio L. and Banfi E. (1999) Synthesis and antimycobacterial activity of some N¹-[1-[3-aryl-1-(pyridine-2-, 3- or 4-yl)-3-oxo]propyl]-2-pyridinecarboxamidrazones. *Farmaco.* 54, 761-767

March J. & Smith M. B. (2001) *Advanced organic chemistry: reactions, mechanism and structure.* Wiley: New York

Markowicz S. W., Figlus M., Lejkowski M., Karolak-Wojciechowska J., Dzierzawska-Majewska A. & Verpoore F. (2006) Enantimerically pure α -pinene derivatives from material of 65% enantiomeric purity. Part 2: C₂-symmetric N,N'-3-(2 α -hydroxy) pinane diimines and diamine *Tetrahedron: Asymmetry* 17, 434-448

Mayo S. L., Olafson B. D. & Goddard W. A. (1990) DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* 94(26):8897-8909

McMurry J. (2002) *Fundamentals of organic Chemistry* Pacific Grove: London

- Minh T. D. & Warfel J. (2004) Homology modelling of botulinum neurotoxins C-G and type A antibodies *Dimensions spring*
- Morales H. R., Juárez M. P., Cuéllar L., Mendoza L., Fernández H. & Contreras R. (1984) Borane-tetrahydrofuran as a useful reagent in the N-monoalkylation of amines and aminoalcohols by carbonyl compounds. *Synthetic Communication* 14: 1213-1219
- Muegge I. & Martin Y. C. (1999) A general and fast scoring function for protein-ligand interactions: a simplified potential approach *J. Med. Chem.* 42: 791-804
- Morphy J. R., Rankovic Z. & Rees D. C. (1996) A novel linker strategy for solid-phase synthesis. *Tetrahedron Lett.* 37: 3209-3212
- Pearlman D. A., Case D. A., Caldwell J. W., Ross W. R., Cheatham T. E., DeBole S., Ferguson D., Seibel G. & Kollman P. (1995) AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculation to elucidate the structures and energies of molecules. *Comput. Phys. Commun.* 91: 1-41
- Ponder J. W. & Case D. A. (2003) Force field for protein simulation. *Adv. Protein Chem.* 66: 27-85
- Pospisil P. & Folkers G. (2004) Making the best account of molecular docking in drug design. *J. Pharm. Sci.* 29: 81-92
- Ramachandran G. N., Ramakrishnan C. & Sasisekharan V. (1963): Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* 7, 95-99
- Rachard L. (1999) Comprehensive organic transformations: a guide to functional group preparations *Wiley-VCH*: New York
- Ramachandran G. N., Ramakrishnan C. & Sasisekharan V. (1963) Stereochemistry of polypeptide chain configuration *J. Mol. Biol.* 7: 95-99
- Rathbone D. L., K. J. Parker, M. D. Coleman, P. A. Lambert and D. C. Billington (2006) Discovery of a potent phenolic *N*¹-benzylidene-pyridinecarboxamidrazone selective against Gram-positive bacteria *Bioorg. Med. Chem. Lett.* 16: 879-883
- Rost, B., Twilight zone of protein sequence alignments, *Protein Eng.*, 1999, 12:85-94
- Rubinstein, Reuven Y. (c1981) *Simulation and the Monte Carlo method*. New York: Wiley
- Rylander P. & Nels P. (1979) Catalytic hydrogenation in organic synthesis. *Academic Press*: New York

Ryckaert J. P., Cicotti G., Berensden H. J. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics Simulation of *n*-alkanes. *J. Comput. Phys.* 23:327-341

Sali, A., Overington, J. P., Johnson, M. S., & Blundell, T. L. (1990) From Comparisons of Protein Sequences and Structures to Protein Modelling and Design. *Trends Biochem. Sci.* 15: 235-240.

Saviola B., S. C. Woolwine, W. R. Bishai (2003) Isolation of acid-inducible genes of *Mycobacterium tuberculosis* with the use of recombinase-based in vivo expression technology. *Infect. Immun.* 71(3): 1379-1388

Scialdone M. A., Shuey S. W., Soper P., Hamuro Y. & Burns D. M. (1998) Phosgenated *p*-nitrophenyl(polystyrene)ketoxime or phoxime resin. A new resin for the solid-phase synthesis of ureas via thermolytic cleavage of oxime-carbamate. *J. Org. Chem.* 63:4802-2807

Sippl MJ. (1993) Recongnition of errors in three dimensional structures of proteins. *Proteins* 17: 355-362

Sriram D., Yogeewari P., Dhakla P., Senthikumar P. & Banerhee D. (2007) *N*-Hydroxythiosemicarbazones: synthesis and in vitro antitubercular activity *Bioorg. Med. Chem. Lett.* 17: 1888-1891

Stewart J. J. P (1989a) Optimisation of parameters for semi-empirical methods I. *J. Comput. Chem.* 10: 209-220

Stewart J. J. P (1989b) Optimisation of parameters for semi-empirical methods II. *J. Comput. Chem.* 10: 221-264

Stewart J. J. P. (1999) *MOPAC 2000 Fujitsu Limited, Tokyo, Japan*

Stryer L. (1995) *Biochemistry* (4th ed.) New York: Freeman

Swope W. C., Anderson H. C., Berens P. H., Wilson K. R. (1982) A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules Application to Small Water Clusters. *J. Chem. Phys.* 76:637-649

Tame J. R. H. (1999) Scoring functions: a view from the bench, *J. Comput. Aided Mol. Des.* 13: 99-108

- Verkhivker G., Appelt K., Freer S. T. & Villafranca J. E. (1995) Empirical free energy calculations of ligand-protein crystallographic complexes. I, Knowledge-based ligand-protein interaction potentials applied to the prediction of human immunodeficiency virus I protease binding affinity. *Protein Eng.* 8: 677-691
- Verlet L. (1967) Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* 159 98-103
- Wang S. J. (1973) *p*-alkoxybenzyl alcohol resin and *p*-alkoxybenzylxycarbonylhydrazide resin for solid phase synthesis of protected peptide fragments *J. Am. Chem. Soc.* 95, 1328-1333
- Weiner S. J., Kollman P. A., Case D. A., Singh U. C., Ghio C., Alagona G., Profeta S., and Weiner P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins *J. Am. Chem. Soc.* 106: 765 – 784
- Wiberg K. B. (1965) A Scheme for Strain Energy Minimization. Application to the Cycloalkanes *J. Am. Chem. Soc.* 87: 1070 - 1078
- Willer R. L., Moor D. W., Lowe-Ma C. K. & Vanderah J. (1985) New chemistry from the reaction of N,N'-disubstituted ethylenediamines with glyoxal: synthesis of 2 imidazolidinecarboxaldehydes and 1,4,6,9-tetraalkyl-1,4,6,9-tetraaza-5,10-dioxaperhydroanthracenes *J. Org. Chem.* 50: 2368-2372
- Wittig G., Bestmann H. J. (1983) *Wittig Chemistry: dedicated to Professor Dr. G. Wittig Springer-Verlag: Berlin; New York*
- Xiao J. F., Guo Y., Chu F. & Sun P. (2006) Computational study of human phosphomannose isomerase: Insights from homology modelling and molecular dynamics simulation of enzyme bound substrate. *J. Mol. Graph. Model.* 25: 289-295
- Yadav G. D. & Lande S. V. (2005) Selective claisen rearrangement of allyl-,4-*di-tert*-butylphenyl ether to 6-allyl-2,4-*di-tert*-butylphenil catalysed by heteropolyacid supported on hexagonal mesoporous silica. *J. Mole. Cyt. A: Chemical* 243:31-39
- Yang H., Ahn Joong-Hoon, Ibrahim R. K., Lee S. & Lim Y. (2004) the three-dimensional structure of *Arabidopsis thaliana* O-methyltransferase predicted by homology-based modelling *J. Mol. Graph. Model.* 23: 77-87
- Zahrt T. C. (2003) Molecular mechanisms regulating persistent *Mycobacterium tuberculosis* infection. *Microbes and Infection* 5: 159-167

Zhang M., J. Wang, Z. Li, J. Xie, Y. Yang Y. Zhong H, Wang (2005) Expression and characterization of the carboxyl esterase Rv3487c from *Mycobacterium tuberculosis*. *Protein. Expr. Purif.* 42: 59-66