

# DESIGN AND SYNTHESIS OF POTENTIAL ANTIMICROBIAL AGENTS

SOUMICK SIKDAR

Doctor of Philosophy

ASTON UNIVERSITY

July 2010

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****DESIGN AND SYNTHESIS OF POTENTIAL ANTIMICROBIAL AGENTS**

A thesis submitted by Soumick Sikdar B.Pharm, for the degree of Doctor of Philosophy  
June 2010

**Abstract:**

Chorismate mutase is one of the essential enzymes in the shikimate pathway and is key to the survival of the organism *Mycobacterium tuberculosis*. The x-ray crystal structure of this enzyme from *Mycobacterium tuberculosis* was manipulated to prepare an initial set of *in silico* protein models of the active site. Known inhibitors of the enzyme were docked into the active site using the flexible ligand / flexible active site side chains approach implemented in *CAChe Worksystem* (Fujitsu Ltd). The resulting complexes were refined by molecular dynamics studies in explicit water using Amber 9. This yielded a further set of protein models that were used for additional rounds of ligand docking. A binding hypothesis was established for the enzyme and this was used to screen a database of commercially available drug-like compounds. From these results new potential ligands were designed that fitted appropriately into the active site and matched the functional groups and binding motifs found therein. Some of these compounds and close analogues were then synthesized and submitted for biological evaluation.

As a separate part of this thesis, analogues of very active anti-tuberculosis pyridylcarboxamidrazone were also prepared. This was carried out by the addition and the deletion of the substitutions from the lead compound thereby preparing heteroaryl carboxamidrazone derivatives and related compounds. All these compounds were initially evaluated for biological activity against various gram positive organisms and then sent to the TAACF (USA) for screening against *Mycobacterium tuberculosis*. Some of the new compounds proved to be at least as potent as the original lead compound but less toxic.

**Keywords:**

*Mycobacterium tuberculosis*, chorismate mutase, molecular modeling, molecular dynamics, docking studies, heteroaryl carboxamidrazone

## ACKNOWLEDGEMENTS

I offer countless thanks to Almighty God who has given me the ability, determination, courage and will to complete this thesis. I offer my respect to my great grandfather Late Nakul Chandra Sikdar and my grandfather Shri Shibdas Sikdar who taught me the meaning and importance of life.

I offer sincere thanks to my supervisor Dr. Daniel Rathbone for being kind, patient and an inspiring source through often difficult moments in writing my thesis. I acknowledge his motivations and overall support that kept me enthusiastic throughout the entire period of this work and I am grateful to have him.

Thanks to the School of Life and Health Sciences for the provision of the funding in this project.

I would like to thank Dr. Peter Lambert for arranging the microbiological testing of all the compounds which were prepared in this thesis. Many thanks to Dr Fraser in clarifying my doubts associated with reaction mechanisms.

I would like to express my gratitude for the help and technical assistances given to me by Dr. Mike Davis and Karen Farrow.

Many thanks to final year project students Jasjeet Singh and Pritpal Singh for their assistance in preparing some of the carboxamidrazone. I would like to thank Ren, Mei, Alexandra and Preeti for helping me to find the way through the complicated procedure of the computer packages used in this thesis.

Finally I would like to thank everybody in my family, my dad, my mum, my sister for their support and encouragement through my time at university, and last but by no means least, my wife Munmun, whose patience and understanding went beyond the call of duty.

## LIST OF CONTENTS

	<b>Page</b>
<b>TITLE PAGE</b>	<b>1</b>
<b>ABSTRACTS</b>	<b>2</b>
<b>ACKNOWLEDGEMENTS</b>	<b>3</b>
<b>LIST OF CONTENTS</b>	<b>4</b>
<b>LIST OF FIGURES</b>	<b>12</b>
<b>LIST OF TABLES</b>	<b>16</b>
<b>LIST OF SCHEMES</b>	<b>18</b>
<b>LIST OF GRAPHS</b>	<b>19</b>
<b>LIST OF EQUATIONS</b>	<b>21</b>
<b>LIST OF ABBREVIATIONS</b>	<b>23</b>
<b>CHAPTER 1</b>	<b>25</b>
<b>INTRODUCTION</b>	<b>25</b>
<b>1.1 Introduction to Tuberculosis:</b>	<b>25</b>
1.1.1 Background:	26
1.1.2 Signs and symptoms in tuberculosis:	26
1.1.3 Aetiology:	27
1.1.4 Pathophysiology:	27
1.1.5 Diagnosis and Treatment:	28
1.1.5.1 Diagnosis of TB:	28
1.1.5.2 Isoniazid (INH):	30
1.1.5.3 Ethambutol (EMB):	30
1.1.5.4 Pyrazinamide (PZA):	31
1.1.5.5 Rifampacin (RMP):	31
1.1.5.6 Streptomycin (STM):	31
1.1.5.7 Polypeptides:	31
1.1.5.8 Fluoroquinolones:	32
1.1.5.9 Thioamides	32
1.1.5.10. Para amino salicylic acid class of drugs.	32
1.1.6 Newer Drugs:	32
1.1.7 Treatment regimen:	33
1.1.8 Drug resistance:	34
<b>1.2 Biosynthetic pathway in bacteria.</b>	<b>35</b>
1.2.1 Shikimic acid pathway:	35
1.2.2 Chorismate mutase:	36
1.2.3 Transition State Analogue (TSA):	37
<b>1.3. Drug design methods:</b>	<b>37</b>
1.3.1 Protein:	38
1.3.2. Protein crystal structures	39
<b>1.4 Introduction to Molecular Modelling:</b>	<b>42</b>
<b>1.5 Molecular Mechanics:</b>	<b>42</b>
1.5.1 Bond stretching:	43

1.5.2	Bond bending:	44
1.5.3	Torsional energy:	45
1.5.4	Non bonded interactions:	46
1.5.4.1	Van der Waals interactions:	46
1.5.4.2.	Electrostatic interactions:	47
<b>1.6</b>	<b>Quantum mechanics:</b>	<b>48</b>
1.6.1	Energy calculation using ab initio quantum mechanical methods:	49
1.6.2	Energy Calculations using semi empirical methods:	50
<b>1.7</b>	<b>Geometry optimization:</b>	<b>50</b>
1.7.1	Geometry optimization by molecular mechanic methods.	51
1.7.1.1	Minimisation:	52
1.7.1.2	Molecular dynamics:	52
1.7.2	Geometry optimization by semi-empirical methods.	53
1.7.3	Geometry optimization by ab initio methods.	53
<b>1.8</b>	<b>Background and development of the Amber Force field.</b>	<b>54</b>
1.8.1.	Amber Project:	56
<b>1.9.</b>	<b>Selection of Algorithm:</b>	<b>56</b>
1.9.1	Verlet algorithm:	57
1.9.2	Leap-frog algorithm	57
1.9.3	Velocity Verlet method	58
1.9.3	Beeman's algorithm:	59
<b>1.10.</b>	<b>Starting, running, measuring and analysing simulation program:</b>	<b>59</b>
1.10.1	Starting:	59
1.10.2	Running:	60
1.10.3	Measuring and analysing:	60
<b>1.11.</b>	<b>Validation of models using Ramachandran Plot:</b>	<b>61</b>
<b>1.12.</b>	<b>Docking studies:</b>	<b>63</b>
1.13.1.	Docking	63
1.13.2.	Scoring Methods:	66
	 <b>CHAPTER 2</b>	 <b>68</b>
	<b>OPTIMIZATION OF AN INSILICO MODEL FOR MYCOBACTERIAL CHORISMATE MUTASE: THE COMPUTATIONAL APPROACH</b>	<b>68</b>
<b>2.1</b>	<b>Aim:</b>	<b>68</b>
<b>2.2</b>	<b>Background:</b>	<b>68</b>
<b>2.3</b>	<b>Crystal structure of chorismate mutase from Mycobacterium tuberculosis:</b>	<b>68</b>
2.3.1	Background of chorismate mutase:	68
2.3.2	Structure of Mycobacterium Chorismate Mutase:	69
<b>2.4</b>	<b>Active site analysis:</b>	<b>71</b>
<b>2.5</b>	<b>The Ligands:</b>	<b>73</b>
<b>2.6</b>	<b>Computational approach:</b>	<b>77</b>
2.6.1	Preparing the initial protein structure:	77
2.6.2	Preparing the initial Ligand structures:	78
2.6.3	Generating the model:	78
2.6.4	Phase I Study:	78
2.6.5	Phase II study:	83
2.6.6	Phase III study:	83
<b>2.7</b>	<b>Analyzing the models:</b>	<b>85</b>
<b>2.7.1</b>	<b>Validation of models:</b>	<b>85</b>
2.7.2	Ramachandran plot:	85
2.7.3	Analysis of the parameters.	89

2.7.4	RMSD analysis:	100
<b>2.8</b>	<b>Docking studies:</b>	<b>103</b>
2.8.1	Docking scores results:	104
<b>2.9</b>	<b>Identifying the interactions:</b>	<b>116</b>
<b>2.10</b>	<b>Discussion:</b>	<b>120</b>
2.10.1	Model Hypothesis:	120
2.10.2	Molecular modelling results and discussions:	121
2.10.3	Selecting the targeted model:	127
2.10.4	Docking results and discussion:	127
2.10.5	Structural activity studies of Ligand 1:	128
2.10.6	Fragment based approach for Ligand 1:	130
<b>2.10.7</b>	<b>Structural activity studies of Ligand 2:</b>	<b>131</b>
2.10.8	Docking studies of the commercially available ligands from Maybridge:	132
<b>2.11</b>	<b>Conclusions:</b>	<b>136</b>
	<b>CHAPTER 3</b>	<b>137</b>
	<b>SUBSTITUTED 3-NITRO 5-SULPHAMOYL BENZOIC ACIDS: DERIVATIVES AND RELATED CHEMISTRY</b>	<b>137</b>
<b>3.1</b>	<b>Aim:</b>	<b>137</b>
<b>3.2</b>	<b>Background:</b>	<b>137</b>
<b>3.3</b>	<b>Substituted 3-sulphamoyl-5-nitro benzoic acid and related compounds:</b>	<b>138</b>
<b>3.4</b>	<b>Screening synthesisable compounds:</b>	<b>141</b>
<b>3.5</b>	<b>Varying the benzoic acids moiety:</b>	<b>145</b>
<b>3.6</b>	<b>Varying the phenylamines:</b>	<b>147</b>
<b>3.7</b>	<b>Varying the benzylhydrazine moiety:</b>	<b>147</b>
<b>3.8</b>	<b>Reactivity of the substituted benzoic acids:</b>	<b>148</b>
<b>3.9</b>	<b>Reactivity of the phenylamines:</b>	<b>150</b>
<b>3.10</b>	<b>Reactivity of the benzylhydrazines:</b>	<b>151</b>
<b>3.11</b>	<b>Reactivity of the benzoyl hydrazide:</b>	<b>152</b>
<b>3.12</b>	<b>Microbiological studies:</b>	<b>153</b>
3.12.1	Dose Response screening:	153
<b>3.13</b>	<b>Discussion:</b>	<b>154</b>
3.13.1	Attempted synthesis of Compound <b>3.25</b> and other related compounds:	156
3.13.2	Attempted synthesis of Compound <b>3.20</b> and related compounds:	157
3.13.3	Attempted synthesis of 3-methoxy-2-hydroxy benzyl amine Compound <b>3.40</b> :	157
3.13.4	Attempted synthesis of Compound <b>3.29</b> :	158
3.13.5	Structural Activity Relationship studies:	159
<b>3.14</b>	<b>Conclusion:</b>	<b>163</b>
	<b>CHAPTER 4</b>	<b>164</b>
	<b>MATERIALS AND METHODS</b>	<b>164</b>
4.1.	Chemicals:	164
4.2.	Instrumentation:	164
4.3.1.1	Preparation of 4-chloro-3-chlorosulfonylbenzoic acid: Compound 3.30.	164
4.3.1.2	Preparation of 4-chloro-3-chlorosulfonyl-5-nitrobenzoic acid: Compound 3.31.	165
4.3.1.3	Preparation of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid: Compound 3.27.	166
4.4.1	Preparation of 4-(3,4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.2.	167
4.4.1.1.	Preparation of DIEA salt of 4-(3, 4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	167

4.4.1.2	Preparation of BOC substituted 4-(3,4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid :	167
4.4.1.3.	Preparation of 4-(3,4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid :	168
4.5.1.	Preparation of 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid: Compound 3.5.	168
4.5.1.1	Preparation of DIEA salt of 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	168
4.5.1.2	Preparation of BOC substituted 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	169
4.5.1.3.	Preparation of 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid :	169
4.6.1.	Preparation of 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid: Compound 3.4.	170
4.6.1.1	Preparation of DIEA salt of 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	170
4.6.1.2	Preparation of BOC substituted 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	170
4.6.1.3	Preparation of 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	171
4.7.1	Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.3.	171
4.7.1.1	Preparation of DIEA salt of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	171
4.7.1.2	Preparation of BOC substituted 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	172
4.7.1.3	Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	172
4.8.1	Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.7.	173
4.8.1.1	Preparation of DIEA salt of 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	173
4.8.1.2	Preparation of BOC substituted 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	173
4.8.1.3	Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	174
4.9.1	Preparation of 4-(3-methoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.6	174
4.9.1.1	Preparation of DIEA salt of 4-(3-methoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	174
4.9.1.2.	Preparation of BOC substituted 4-(3-methoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	175
4.9.1.3	Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:	175
4.10.1.	Preparation of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid: Compound 3.8.	176
4.10.1.1	Preparation of DIEA salt of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:	176
4.10.1.2	Preparation of BOC substituted 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:	176
4.10.1.3	Preparation of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:	177
4.11.1	Preparation of 4-(3-methoxyphenethylamino)-3, 5-dinitro benzoic acid: Compound 3.12.	177
4.11.1.1	Preparation of DIEA salt of 4-(3-methoxyphenethylamino)-3,5-dinitro benzoic acid:	177
4.11.1.2	Preparation of BOC substituted 4-(3-methoxyphenethylamino)-3,5-dinitro benzoic acid:	178
4.11.1.3	Preparation of 4-(3-methoxyphenethylamino)-3,5-dinitro benzoic acid:	178
4.12.1.	Preparation of 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid : Compound 3.9.	179
4.12.1.1.	Preparation of DIEA salt of 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid:	179
4.12.1.2.	Preparation of BOC substituted 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid:	179
4.12.1.3.	Preparation of 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid:	180
4.13.1.	Preparation of 4-phenethylamino-3,5-dinitro benzoic acid : Compound 3.11.	180
4.13.1.1.	Preparation of DIEA salt of 4-phenethylamino-3,5-dinitro benzoic acid:	180
4.13.1.2.	Preparation of BOC substituted 4-phenethylamine-3,5-dinitro benzoic acid:	181
4.13.1.3.	Preparation of 4-phenethylamine-3,5-dinitro benzoic acid:	181

4.14.1	Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3,5-dinitro benzoic acid : Compound 3.13	182
4.14.1.1.	Preparation of DIEA salt of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:	182
4.14.1.2.	Preparation of BOC substituted 4-(3,4,6-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:	182
4.14.1.3.	Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:	183
4.15.1	Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3,5-dinitro benzoic acid : Compound 3.9.	183
4.15.1.1.	Preparation of DIEA salt of 4-(3,4,5-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:	183
4.15.1.2.	Preparation of BOC substituted 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:	184
4.15.1.3.	Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:	184
4.16.1.	Preparation of 3-methoxy-4-hydroxybenzylhydrazine: Compound 3.33.	184
4.16.1.1	Preparation of 3-methoxy-4-hydroxybenzylidenehydrazine:	184
4.16.1.2	Preparation of 3-methoxy-4-hydroxybenzylhydrazine:	185
4.17.1.	Preparation of 3,4-dimethoxybenzylhydrazine: Compound 3.32.	185
4.17.1.1	Preparation of 3,4-dimethoxybenzylidenehydrazine:	186
4.17.1.2	Preparation of 3,4-dimethoxybenzylhydrazine:	186
4.18.1.	Preparation of 3,4-dihydroxybenzylhydrazine: Compound 3.24.	186
4.18.1.1	Preparation of 3,4-dihydroxybenzylidenehydrazine:	186
4.19.1.	Preparation of 2-methoxy-4-((2-(4-methylbenzylidene) hydrazinyl) methyl) phenol: Compound 3.36.	187
4.20.1.	Preparation of 4-amino-3-(N-methylsulphamoyl)-5-nitrobenzoic acid: Compound 3.35.	188

## **CHAPTER 5** **189**

### **HETEROARYLCARBOXAMIDRAZONES: DERIVATIVES AND RELATED CHEMISTRY**

		<b>189</b>
<b>5.1</b>	<b>Aim:</b>	<b>189</b>
<b>5.2</b>	<b>Background:</b>	<b>189</b>
<b>5.3</b>	<b>Structural activity relationship studies:</b>	<b>190</b>
<b>5.4</b>	<b>Optimisation of the lead compound:</b>	<b>191</b>
5.4.1	Optimisation of lead compound following Compound 5.1:	191
<b>5.5</b>	<b>Varying the aldehyde building blocks in the synthesis of benzylidene carboxamidrazones:</b>	<b>199</b>
<b>5.6</b>	<b>Varying the aryl acid building blocks in the synthesis of carboxamidrazone amides:</b>	<b>201</b>
<b>5.7</b>	<b>Varying the Pyridyl-amine moiety in the synthesis of analogues of pyridine carboxamidrazones:</b>	<b>203</b>
<b>5.8</b>	<b>Reactivity of the pyridyl-amine moiety.</b>	<b>204</b>
<b>5.9</b>	<b>Stability of the Pyridyl-amines:</b>	<b>207</b>
<b>5.10</b>	<b>Microbiological studies:</b>	<b>208</b>
5.10.1	Initial Biological Screening:	208
<b>5.10.2</b>	<b>TAACF screening:</b>	<b>211</b>
5.10.2.1	Primary Screen (Dose Response):	211
5.10.2.2	Secondary Screen:	211
<b>5.11</b>	<b>Results and Discussion:</b>	<b>215</b>
5.11.1	Microbiological studies:	220
5.11.2	Toxicology:	222
5.11.3	Structural activity relationship studies:	223
<b>5.12</b>	<b>Conclusion:</b>	<b>226</b>

<b>CHAPTER 6</b>	<b>227</b>
<b>MATERIALS AND METHODS</b>	<b>227</b>
<b>6.1. Chemicals:</b>	<b>227</b>
<b>6.2. Instrumentation:</b>	<b>227</b>
<b>6.3.1. Preparation of Pyridine-4-methylhydrazone: Compound 5.87</b>	<b>227</b>
6.3.2. Preparation of 4-[3,5-di-(tert-butyl)-2-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1,3-butadiene: Compound 5.17	228
6.3.3. Preparation of 4-[3, 5-di-(tert-butyl)-4-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1,3-butadiene: Compound 5.18	229
6.3.4. Preparation of 4-[3,5-di-iodo-2-OH-benzyl]-1-(4-pyridyl)-2,3-diaza-1,3-butadiene. Compound 5.19	229
6.3.5. Preparation of 4- [5-tert-butyl-2-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1,3-butadiene Compound 5.22	230
6.3.6. Preparation of 4-[3-tert-butyl-2-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1, 3-butadiene Compound 5.21	230
6.3.7. Preparation of 4-[3,5-di-bromo-2-OH-benzyl]-1-(4-pyridyl)-2,3-diaza-1,3-butadiene: Compound 5.20	231
<b>6.4.1 Preparation of 4-hydrazino pyridine: Compound 5.84.</b>	<b>231</b>
6.4.2. Preparation of 3,5-di-tert butyl 2-OH benzylidene 4-hydrazinopyridine Compound 5.03	232
6.4.3. Preparation of 3,5-di-tert butyl 4-OH benzylidene 4-hydrazinopyridine. Compound 5.04	233
6.4.4. Preparation of 3,5-di-iodo 2-OH benzylidene 4-hydrazinopyridine Compound 5.05	233
6.4.5. Preparation of 3,5-di-bromo 2-OH benzylidene 4-hydrazinopyridine Compound 5.06.	234
6.4.6. Preparation of 3,5-di-tert-butyl benzylidene 4-hydrazinopyridine: Compound 5.09.	234
6.4.7. Preparation of 3-tert-butyl 2-OH benzylidene 4-hydrazinopyridine: Compound 5.07	235
6.4.8. Preparation of 5-tert-butyl 2-OH benzylidene 4-hydrazinopyridine Compound 5.08	235
6.5.1. Preparation of 4-[3, 5-di-(tert-butyl)-2-OH-benzylidene]-amino-pyridine: Compound 5.29.	236
6.5.2. Preparation of 4-[3, 5-di-iodo-2-OH-benzylidene]-amino-pyridine: Compound 5.31	236
6.5.3. Preparation of 4-[3, 5-di-(tert-butyl)-4-OH-benzylidene]-amino-pyridine: Compound 5.30	237
6.5.4. Preparation of 4-[3, 5-di-bromo-2-OH-benzylidene]-amino-pyridine: Compound 5.32	238
6.6.1. Preparation of 3,5-di-tert-butyl 2-OH benzylidene 2-hydrazinopyridine. Compound 5.10	238
6.6.2. Preparation of 3,5-di-tert-butyl-4-OH benzylidene-4-hydrazinopyridine: Compound 5.11.	239
6.6.3. Preparation of 3,5-diiodo-2-OH benzylidene-4-hydrazinopyridine Compound 5.12.	240
6.6.4. Preparation of 3,5-dibromo-2-OH benzylidene-2-hydrazinopyridine: Compound 5. 13.	240
6.6.5. Preparation of 3-tert-butyl-2-OH benzylidene-2-hydrazinopyridine: Compound 5.14.	241
6.6.6. Preparation of 5-tert-butyl-2-OH benzylidene-2-hydrazinopyridine: Compound 5.15.	241
6.6.7. Preparation of 3,5-di-tert-butyl-benzylidene-2-hydrazinopyridine: Compound 5.16.	242
<b>6.7.1. Preparation of Pyridine-2-methylhydrazone: Compound 5.88.</b>	<b>242</b>
6.7.2. Preparation of 4-[3,5-di-(tert-butyl)-2-OH-benzyl]-1-(2-pyridyl)-2, 3-diaza-1,3-butadiene: Compound 5.23.	243
6.7.3. Preparation of 4-[3,5-di-(tert-butyl)-4-OH-benzyl]-1-(2-pyridyl)-2,3-diaza-1,3-butadiene. Compound 5.24.	243
6.7.4. Preparation of 4-[3-(tert-butyl)-2-OH-benzyl]-1-(2-pyridyl)-2,3-diaza-1,3-butadiene: Compound 5.25.	244
<b>6.8.1 Preparation of Pyridine-3-methylhydrazone: Compound 5.86.</b>	<b>244</b>
6.8.2. Preparation of 4-[3, 5-di-(tert-butyl)-2-OH-benzyl]-1-(3-pyridyl)-2, 3-diaza-1, 3-butadiene. Compound 5.26	245
6.8.3. Preparation of 4-[3,5-di-iodo-2-OH-benzyl]-1-(3-pyridyl)-2,3-diaza-1,3-butadiene: Compound 5.27	246
6.8.4. Preparation of 4-[3,5-di-bromo-2-OH-benzyl]-1-(3-pyridyl)-2,3-diaza-1,3-butadiene: Compound 5.28.	246
6.9.1. Preparation of 3,5-diiodo-2-OH benzylidene-3-aminopyridine: Compound 5.35.	247
6.9.2. Preparation of 3,5-di-(tert-Butyl)-2-OH benzylidene-3-aminopyridine: Compound 5.33.	247

6.9.3.	Preparation of 3,5-dibromo-2-OH benzylidene-3-aminopyridine: Compound 5.36.	248
6.9.4.	Preparation of 5-(tert-Butyl)-2-OH benzylidene-3-aminopyridine: Compound 5.37.	248
6.9.5.	Preparation of 3,5-di-(tert-Butyl)-4-OH benzylidene-3-aminopyridine: Compound 5.34.	249
6.9.6.	Preparation of 3-bromo-5-Chloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.94.	249
6.9.7.	Preparation of 3,5-dichloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.95.	250
6.9.8.	Preparation of 3,5-difluoro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.98.	250
6.9.9.	Preparation of 5-Iodo-2-hydroxy benzylidene-3-aminopyridine: Compound 5.99.	251
6.9.10.	Preparation of 3-chloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.97.	251
6.9.11.	Preparation of 2-bromo-2-hydroxy benzylidene-3-aminopyridine: Compound 5.100.	252
6.9.12.	Preparation of 5-chloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.96.	253
6.9.13.	Preparation of 5-bromo-2-hydroxy benzylidene-3-aminopyridine: Compound 5.101.	253
<b>6.10.1.</b>	<b>Preparation of pyridine-2-carboxamidrazone: Compound 5.89.</b>	<b>254</b>
6.10.2.	Preparation of N1-(2-furoylacryl)-pyridine-2-carboxamidrazone: Compound 5.38.	254
6.10.3.	Preparation of N1-2-(5-bromofuroyl)-pyridine-2-carboxamidrazone: Compound 5.41.	255
6.10.4.	Preparation of N1-2-(4,5-dibromofuroyl)-pyridine-2-carboxamidrazone: Compound 5.42.	256
6.10.5.	Preparation of N1-2-[3-(2-furoyl)-propyl]-pyridine-2-carboxamidrazone: Compound 5.44.	256
6.10.6.	Preparation of N1-3-(2,5-dimethylfuroyl)-pyridine-2-carboxamidrazone: Compound 5.43.	257
6.10.7.	Preparation of N1-3-furoyl-pyridine-2-carboxamidrazone: Compound 5.39.	257
6.10.8.	Preparation of N1-2-benzofuroyl-pyridine-2-carboxamidrazone: Compound 5.45.	258
6.10.9.	Preparation of N1-2-[5-(4-Nitrophenyl)furoyl]-pyridine-2-carboxamidrazone: Compound 5.40.	258
6.10.10.	Preparation of N1-(thiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.51.	259
6.10.11.	Preparation of N1-(3-bromothiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.54.	259
6.10.12.	Preparation of N1-(5-bromothiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.57.	260
6.10.13.	Preparation of N1-(5-chlorothiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.60.	260
6.10.14.	Preparation of N1-(2-bromobenzoyl)-pyridine-2-carboxamidrazone: Compound 5.71.	261
6.10.15.	Preparation of N1-(4-bromobenzoyl)-pyridine-2-carboxamidrazone: Compound 5.63.	261
6.10.16.	Preparation of N1-(5-[4-bromophenyl]-furan-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.66.	262
6.10.17.	Preparation of N1-(5-[2-trifluoromethylphenyl]-furan-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.69.	262
<b>6.11.1.</b>	<b>Preparation of pyridine-4-carboxamidrazone: Compound 5.91.</b>	<b>263</b>
6.11.2.	Preparation of N1-(5-chlorothiophen-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.59.	263
6.11.3.	Preparation of N1-(3-bromothiophen-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.53.	264
6.11.4.	Preparation of N1-(5-bromothiophen-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.56.	265
6.11.5.	Preparation of N1-(4-bromobenzoyl)-pyridine-4-carboxamidrazone: Compound 5.62.	265
6.11.6.	Preparation of N1-(thiophen-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.50.	266
6.11.7.	Preparation of N1-(5-[4-bromophenyl]-furan-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.65.	266
6.11.8.	Preparation of N1-(5-[2-trifluoromethylphenyl]-furan-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.68.	267
<b>6.12.1.</b>	<b>Preparation of pyridine-3-carboxamidrazone: Compound 5.90.</b>	<b>267</b>
6.12.2.	Preparation of N1-(3-bromothiophen-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.55.	268

---

6.12.3.	Preparation of N1-(5-[4-bromophenyl]-furan-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.67.	269
6.12.4.	Preparation of N1-(5-[2-trifluoromethylphenyl]-furan-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.70.	269
6.12.5.	Preparation of N1-(4-bromobenzoyl)-pyridine-3-carboxamidrazone: Compound 5.64.	270
6.12.6.	Preparation of N1-(thiophen-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.52.	271
6.12.7.	Preparation of N1-(5-bromothiophen-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.58.	271
6.12.8.	Preparation of N1-(5-chlorothiophen-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.61.	272
	<b>APPENDIX I</b>	<b>273</b>
	<b>APPENDIX II</b>	<b>274</b>
	<b>REFERENCES:</b>	<b>276</b>

## LIST OF FIGURES

	<b>Page</b>
Figure 1.1: Showing the distribution of TB cases worldwide. The scale indicates TB cases per 100,000 populations. (WHO, 2010)	25
Figure 1.2 : The mechanism of TB infection within the host (Russell. D, 2001).	28
Figure 1.3: The mechanism by which the drugs act on the bacilli. (Image obtained from <a href="http://www3.niaid.nih.gov/topics/tuberculosis/WhatIsTB/ScientificIllustrations">www3.niaid.nih.gov/topics/tuberculosis/WhatIsTB/ScientificIllustrations</a> )	30
Figure 1.4: The treatment regimen in TB patients.	34
Figure 1.5: Showing the transitional state conformation (A) and the transitional state analogue (B).	37
Figure 1.6: Peptide bond between two amino acids.	38
Figure 1.7: Showing the resonance within a peptide unit.	38
Figure 1.8: Illustrates the different structure of a protein structure. (Image obtained from <a href="http://commons.wikimedia.org/wiki/File:Main_protein_structure_levels_en">commons.wikimedia.org/wiki/File:Main_protein_structure_levels_en</a> )	39
Figure 1.9: Crystal structure of an enzyme 2FP2.pdb (chorismate mutase from <i>M. tuberculosis</i> )	40
Figure 1.10: Explaining the crystal structure data from protein data bank file.	41
Figure 1.11: Showing the distribution of the potential energy in a molecule. ( <a href="http://en.wikipedia.org/wiki/File:MM_PEF.png">http://en.wikipedia.org/wiki/File:MM_PEF.png</a> )	42
Figure 1.12: The variation of energy due to torsion angle variation.	45
Figure 1.13: The Amber program suite (Amber Manual. 2005).	56
Figure 1.14: Stepwise view of the Verlet integration algorithm and its variants. At each algorithm the dark and the light grey cells indicate the initial and calculating variables, respectively. The numbers in the cells represent the orders in the calculation procedures. The arrows point from the data used in the calculation of the variable that is being calculated each step (Becker.O, <i>et al.</i> , 2001).	58
Figure 1.15: The phi-angle ( $\phi$ ) and the psi-angle ( $\psi$ ) in a peptide unit. (Adapted from <a href="http://wiki.cmbi.ru.nl/index.php/File:Phipsi.jpg">wiki.cmbi.ru.nl/index.php/File:Phipsi.jpg</a> )	61
Figure 1.16: Example of a Ramachandran plot.	62
Figure 1.17: Showing the binding of the ligand to the receptor.	63
Figure 1.18: The steps involved in the docking process. The targeted interactions were first identified. This is then followed by identifying the best possible fit.	64
Figure 1.19: The traditional genetic algorithm (Goldberg 1989) designed to generate the initial population $P(t)$ through a random selection process. (Chromosomes represent the conformation of the molecules). The next algorithm cycles through rounds of selection, recombination and recombination until the terminal conditions are met. (The terminal conditions represent the internal energy)	65
Figure 2.1: The X-ray crystal structure of the Chorismate mutase (2FP2.pdb) determined at 1.6 Å.	70

---

Figure 2.2: Showing the ligand bound chain B. H1, H2, H3, H4, H5 and H6 represent the helices of the chain (2FP2.pdb).	71
Figure 2.3: The active site within the TSA-2FP2 complex. The above conformation is the snap shot obtained in 'CACHe' directly from the crystal structure 2FP2.pdb. The active site was defined by selecting the amino acids within the 5 Å region of the TSA.	72
Figure 2.4: The conversion of prephenate from chorismate through the transitional state.	73
Figure 2.5: The structure of the Transitional state analogue (TSA).	73
Figure 2.6: Ramachandran plots of the TSA-Protein complexes, where A is the plot from X-ray crystal structure, B is the plot following a MD study at 1120ps, C is the plot after a MD study at 2120ps and D is the plot after 12330ps.	86
Figure 2.7: Showing the Ramachandran plot of the ligand protein complex following the MD from Phase 2 studies.	88
Figure 2.8: Showing the Ligand 1 and A, B and C points to the targets where the variations have been made to design various ligands.	105
Figure 2.9: Two dimensional view (using the software 'Ligand Scout' )of the ligand interacting with the residues within the active site model obtained at 1120ps following a MD study of the TSA-protein complex (scheme 2.2).	117
Figure 2.10: The 3D view using 'DS visualiser 2', of the ligand interacting with the residues within the active site of the target model obtained at 1120ps following a MD study of the TSA-protein complex (scheme 2.2).	119
Figure 2.11: Showing the interaction of the ligands in the active site of the target model (obtained at 1120ps following the MD study using the TSA-protein complex (scheme 2.2)).	123
Figure 2.12: Showing the opening of the active site following the MD studies at 1120ps with the ligand-protein complexes.	124
Figure 2.13: The folding pattern following a long MD of the docked Ligand 1-protein complex.	126
Figure 2.14: The structure of the Ligand 1	128
Figure 2.15: The interactions produced when the Ligand 1 was fragmented and docked into the active site and the fragments were docked separately as Ligand 88 and Ligand 89.	130
Figure 2.16: The structure of Ligand 2	131
Figure 2.17: Showing the orientation of the Ligand 81 (fragment of Ligand 2) and Ligand 82 (a modification of Ligand 84) within the active site.	132
Figure 2.18: The interactions produced by some of the commercially available compounds when docked into the active site of the model obtained following MD study of the TSA-protein complex at 1120ps (Scheme 2.2).	135
Figure 3.1: Structure of 4-(3,4-dimethoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid (Compound 3.1)	137
Figure 3.2: Target compounds judged to have potential activity against chorismate mutase from <i>M. tuberculosis</i> .	140
Figure 3.3: The three main categories of the compounds used in this investigation.	141

Figure 3.4: The general synthesis of the substituted 3-sulphamoyl-5-nitro benzoic acid derivatives related to Type A compounds shown in Table 3.1.	142
Figure 3.5: The general synthesis of the substituted 3-sulphamoyl-5-nitro benzoic acid derivatives related to Type B compounds shown in Table 3.1, where R and R1 are the substituents within the ring.	143
Figure 3.6: The general synthesis of the substituted 3-sulphamoyl-5-nitro benzoic acid derivatives related to Type C compounds shown in Table 3.1.	144
Figure 3.7: The different types of benzoic acids used in this study.	145
Figure 3.8: The stages involved in the synthesis of 3-sulphamoyl-5-nitro 4-chloro benzoic acid (Compound 3.27).	146
Figure 3.9: The synthesis of 3-(N-methyl sulphonyl)-5-nitro-4-chloro benzoic acid from 3-chlorosulphamoyl-5-nitro 4-chloro benzoic acid.	146
Figure 3.10: The list of various phenylamines used in the study.	147
Figure 3.11: The stages involved in the synthesis of the benzylidene-hydrazine from the aldehyde.	148
Figure 3.12: The list of benzyl hydrazines used to prepare the analogues of Type C and Type D.	148
Figure 3.13: The presence of the electronegative groups deactivating the ring pulling the electrons thereby making the carbon bearing the Cl atom an easy target for amines.	149
Figure 3.14: Amination of Compound 3.27.	149
Figure 3.15: The reaction between the benzyl hydrazine and the aldehyde yielding corresponding imides.	152
Figure 3.16: The various benzoyl hydrazides used in this investigation.	152
Figure 3.17: Synthesis of the substituted sulphamoyl benzoic acids	154
Figure 3.18: Showing the interaction of the compound 3.1 within the active site.	160
Figure 5.1: Compound 5.1 represents N1-[3, 5-di-( <i>tert</i> -butyl)-2-hydroxybenzylidene] - pyridine-4-carboxamidrazone with MIC 2-4 µg/mL. Compound 5.2 is pyridine-2-carboxamidrazone-N1-[3, 5-di-( <i>tert</i> -butyl)-2-hydroxybenzoyl] amide with MIC <6.25 µg/mL	189
Figure 5.2: Analogues following Compound 5.1, which were designed by modifying the benzylidene group, pyridyl group and the junction between benzylidene group and pyridyl groups.	191
Figure 5.3: List of compounds following the analogues of the Compound 5.2.	195
Figure 5.4: Aldehydes used in the synthesis of the analogues of Compound 5.1.	200
Figure 5.5: Showing the list of acids used to synthesise the analogues of Compound Y.	202
Figure 5.6: The various pyridyl-carboxamidrazones used in the preparation of the analogues of Compound 5.2	203
Figure 5.7: The various amino-pyridines and pyridyl carboxamidrazones used to synthesise the analogues of Compound 5.1 and Compound 5.2	204
Figure 5.8: The resonance forms of 4-aminopyridine.	205
Figure 5.9: The various resonance forms of 3-aminopyridine.	205
Figure 5.10: The movement of the lone pair of electrons on the pyridyl-methylhydrazine and the hydrazinopyridine.	206

---

Figure 5.11: Showing the stability of the compound 5.89 due to the formation of the intra-molecular hydrogen bonding.	208
Figure 5.12: Attempted synthesis of compounds containing an acetoxy phenyl substituent.	215
Figure 5.13: Showing the schematic conversion of dibromo-furoic acid into the corresponding acid in presence of oxalyl chloride and subsequent reaction to provide the pyridine-2-carboxamidrazone amide.	216
Figure 5.14: Synthesis of the imidazole complex and subsequent intra-molecular hydrogen bonding for a 2-hydroxy salicylic acid derivative.	217
Figure 5.15: Illustrating the large bromine atom causing hindrance to the carboxylic group toward the attack of the amine	217
Figure 5.16: The reaction pathway for amide synthesis involving acid chloride.	218
Figure 5.17: Showing the keto-enol form of the compound 5.45.	219
Figure 5.18: The NMR spectrum of compound 5.45 at room temperature	219
Figure 5.19: The NMR spectrum of compound 5.45 at 50oC	220
Figure 5.20: Showing the analogue of Compound A and Compound D	224

## LIST OF TABLES

	<b>Page</b>
Table 1.1: The first line and the second line anti TB drugs.	29
Table 2.1: The broad classification of chorismate mutase (Babu. <i>et al.</i> , 2004).	69
Table 2.2: Published structures of biologically active ligands against MtCM (Agrawal, Kumar <i>et al.</i> 2007)	74
Table 2.3: List of all the inactive ligands following the biological evaluation by Himanshu <i>et al.</i> (Agrawal, Kumar <i>et al.</i> 2007).	75
Table 2.4: The table above shows the list of ligands used in the current study. Ligand 1 to Ligand 4 denotes the active ligands whereas Ligand 7 to Ligand 13 denotes the inactive ligands.	76
Table 2.5: Showing the calculation of the binding free energy when the ligand binds to the protein. The more negative is the energy, the more stable is the predicted binding.	84
Table 2.6: Showing the distribution of the % of residues in Ramachandran plot of TSA-Protein complex.	87
Table 2.7: Showing the distribution of the residues in Ramachandran plot of the protein ligand complexes obtained in the Phase 2 studies at 2240ps.	89
Table 2.8: Showing the docking score of the ligands.	104
Table 2.9: Ligands designed as analogue of Ligand 1.	108
Table 2.10: Analogues of ligand 1 containing a pyridyl moiety.	108
Table 2.11: Showing the various ligands designed as analogue of Ligand 2.	109
Table 2.12: Showing the list of the fragments used to study the interactions within the active site.	109
Table 2.13: Showing the list of analogues of Ligand 1 where the folding pattern of the ligands was fixed.	110
Table 2.14: The docking scores of the designed ligands carried out on the target protein model.	113
Table 2.15: The docking scores following the docking studies carried out with the Maybridge compounds using the target protein model.	116
Table 2.16: The RMSD values of the ligand-protein complexes from the X-ray crystal structure 2FP2.pdb.	123
Table 2.17: Showing the ranking order of the ligands as determined by the model. A denotes that the ligands follow the model hypothesis in all aspects, B denotes that the ligands requires further modification and Ligands ranking C are rejected	130
Table 2.18: Showing the list of the Maybridge compounds with rank	134
Table 3.1: Compounds of Type A.	143
Table 3.2: Analogues of compounds of Type B.	144
Table 3.3: Analogues of compounds of Type C.	145
Table 3.4: The biological activity of the compounds against <i>M .tuberculosis</i> .	153

---

Table 3.5:	The lipinski's calculation of the molecular properties. LogP was calculated using Fijitsu CAChe 6.1	160
Table 5.1:	The table above shows the list of 4-pyridyl and 2-pyridyl hydrazine substituted analogues of Compound 5.1.	192
Table 5.2:	List of 4-pyridyl, 3-pyridyl and 2-pyridyl methylhydrazone substituted analogues of Compound 5.1.	194
Table 5.3:	The list of 4-aminopyridine and 3-aminopyridine substituted analogues of Compound 5.1.	195
Table 5.4:	List of analogues of the compound D prepared by varying the action of acids shown in Figure 5.3b with the pyridine-3-carboxamidrazone, pyridine-2-carboxamidrazone and pyridine-4-carboxamidrazone from Figure 5.5.	198
Table 5.5:	Showing the list of analogues of Compound E.	198
Table 5. 6:	The table above shows the list of the compounds prepared following Compound F	199
Table 5. 7:	The zone of inhibition of the compounds versus the microorganisms (measured in mm as diameter) on a blood agar plate with 6mm diameter wells. The – sign in the 'Zone of Inhibition' column indicates that the compounds did not show any zone of inhibition. The + indicates, growth of the organism and – indicates no growth of the organism	209
Table 5. 8:	The zone of inhibition of the compounds versus the microorganisms (measured in mm as diameter) on a blood agar plate with 6mm diameter wells. The (–) sign indicates that the compounds did not show any zone of inhibition.	210
Table 5.9:	The IC50 and IC90 values of the compounds when tested against <i>Mycobacterium tuberculosis</i> at the TAACF.	213
Table 5. 10:	The selectivity index of the compounds (CC50/ IC90).	214
Table 5.11:	Showing the percentage viability of the mammalian cells at various concentrations.	222

**LIST OF SCHEMES**

	<b>Page</b>
Scheme 1.1: Showing the shikimic acid pathway (Nissen. S, et al., 2005)	36
Scheme 2.1: Schematic representation of the model design.	78
Scheme 2.2: Schematic representation of the Phase I study. The docking studies were performed three times and the average docking scores were calculated.	79
Scheme 2.3: Schematic representation of the Phase II study	83
Scheme 2.4: The schematic representation of the Phase III study.	84
Scheme 3.1: Synthesis of the DIEA salt of the substituted sulphamoyl benzoic acids	155
Scheme 3.2: The synthesis of the BOC-substitute of compound A.	155
Scheme 3.3: Showing the synthesis of substituted benzoic acid derivatives.	156
Scheme 3.4: Synthesis of the Compound 3.25.	156
Scheme 3. 5: Showing the synthesis of Compound 3.20.	157
Scheme 3.6: Attempted synthesis of 3-methoxy-2-hydroxy benzyl amine.	158
Scheme 3.7: Showing synthesis of compound 3.35.	158
Scheme 3.8: Showing structural similarity between compound 3.11 and compound 3.1	161
Scheme 5.1: Synthesis of hydrazinopyridine followed by the condensation with aldehyde.	192
Scheme 5.2: The synthesis of pyridyl-methylhydrazone.	193
Scheme 5.3: The reaction between aldehydes and the 4-aminopyridine where R-NH <sub>2</sub> represent the 4-aminopyridine and 3-aminopyridine.	194
Scheme 5.4: General method for the synthesis of pyridyl-carboxamidrazone which is then reacted with the acid to give the substituted benzylidene carboxamidrazones.	196
Scheme 5.5: The scheme above illustrates synthesis of the analogues of Compound E.	198
Scheme 5.6: Synthesis of the analogues of Compound F	199
Scheme 5.7: Formation of the imidazole-aryl complex followed by reaction with the amine to form the corresponding amide(de Figueiredo. <i>et al.</i> 2006)	206

## LIST OF GRAPHS

	<b>Page</b>
Graph 1.1: An example of energy change due to bond stretching variation	44
Graph 1.2: The change of energy due to bond bending.	45
Graph 1.3: The change of energy due to van der Waals interaction.	46
Graph 1.4 : Showing the flow of energy (potential energy) level during simulation.	51
Graph 2.1: Variation in temperature through the MD simulation performed on the TSA-protein complex over 123300ps.	89
Graph 2.2: Variation in pressure through the MD simulation performed on the TSA-protein complex over 123300ps	90
Graph 2.3: Variation in total potential energy (EPTOT) and total kinetic energy (EKTOT) through the MD simulation performed on the TSA-protein complex over 123300ps.	90
Graph 2.4: Variation in total potential energy (EPTOT) and total kinetic energy (EKTOT) through the MD simulation performed on the TSA-protein complex over 1120ps.	91
Graph 2.5: Variation in EPTOT and EKTOT through the MD simulation performed on the ligand1-protein complex over 1120ps.	91
Graph 2.6: Variation in pressure through the MD simulation performed on the Ligand1-protein complex over 1120ps.	92
Graph 2.7: Variation in temperature through the MD simulation performed on the Ligand1-protein complex over 1120ps.	92
Graph 2.8: Variation in EPTOT and EKTOT through the MD simulation performed on the Ligand2-protein complex at 1120ps.	93
Graph 2.9: Variation in temperature through the MD simulation performed on the Ligand 2 -protein complex over 1120ps.	93
Graph 2.10: Variation in pressure through the MD simulation performed on the Ligand2-protein complex over 1120ps.	94
Graph 2.11: Variation in EPTOT and EKTOT through the MD simulation performed on the Ligand4-protein complex at 1120ps.	94
Graph 2.12: Variation in pressure through the MD simulation performed on the Ligand 4-protein complex over 1120ps.	95
Graph 2.13: Variation in temperature through the MD simulation performed on the Ligand 4 -protein complex over 1120ps.	95
Graph 2.14: Variation in EPTOT and EKTOT through the MD simulation performed on the on the Ligand7-protein complex at 1120ps.	96
Graph 2.15: Variation in temperature through the MD simulation performed on the Ligand 7 -protein complex over 1120ps.	96

---

Graph 2.16: Variation in pressure through the MD simulation performed on the Ligand 7-protein complex over 1120ps.	97
Graph 2.17: Variation in EPTOT and EKTOT through the MD simulation performed on the on the Ligand 8-protein complex at 1120ps.	97
Graph 2.18: Variation in pressure through the MD simulation performed on the Ligand 8-protein complex over 1120ps.	98
Graph 2.19: Variation in temperature through the MD simulation performed on the Ligand 8 -protein complex over 1120ps.	98
Graph 2.20: Variation in EPTOT and EKTOT through the MD simulation performed on the on the Ligand 9-protein complex at 1120ps.	99
Graph 2.21: Variation in pressure through the MD simulation performed on the Ligand 9-protein complex over 1120ps.	99
Graph 2.22: Variation in temperature through the MD simulation performed on the Ligand 9 -protein complex over 1120ps.	100
Graph 2.23: RMSD changes of the TSA-protein complex during MD study at 12330ps.	100
Graph 2.24: RMSD changes of the Ligand1-protein complex during MD study at 1120ps during the Phase I study.	101
Graph 2.25: RMSD changes of the Ligand2-protein complex during MD study at 1120ps during the Phase II study	101
Graph 2.26: RMSD changes of the Ligand4-protein complex during MD study at 1120ps during the Phase II study	102
Graph 2.27: RMSD changes of the Ligand7-protein complex during MD study at 1120ps during the Phase II study	102
Graph 2.28: RMSD changes of the Ligand8-protein complex during MD study at 1120ps during the Phase II study	103
Graph 2.29: RMSD changes of the Ligand9-protein complex during MD study at 1120ps during the Phase II study.	103
Graph 2.30: The variation of the bond length throughout the molecular dynamics studies of Ligand 1-protein complex at further 2120ps (Phase2).	125
Graph 2.31: Showing the variation of the bond length throughout the molecular dynamics studies of the TSA-protein complex at further 2120ps (Phase 2).	126
Graph 5.1: The graph above showing the distribution of the selectivity index. The lower the selectivity index, the more toxic is the product relative to the antimicrobial activity.	214
Graph 5.2: The distribution of the potency of the compounds towards <i>M. tuberculosis</i> in a concentration at which compound 5.3 exhibits 100% potency.	221
Graph 5.3: The graph above showing the distribution of the percentage of cell viability of the mammalian cells versus the concentration of the compound.	223

## LIST OF EQUATIONS

	<b>PAGE</b>
$E_{tot} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{elst} + \dots$	EQUATION 1.1 43
$E = \frac{kx^2}{2}$	EQUATION 1.2 43
$E_{str} = \sum \frac{1}{2} k_b (b - b_0)^2$	EQUATION 1.3 44
$E_{bend} = \sum \frac{1}{2} k_\theta (\theta - \theta_0)^2$	EQUATION 1.4 44
$E_{tors} = \sum \frac{1}{2} K_\phi [1 + \cos(N\phi - \phi_0)]$	EQUATION 1.5 46
$E_{VDW} = \sum \left( \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right)$	EQUATION 1.6 46
$E_{elst} = \sum \frac{Q_i Q_j}{\epsilon_0 r_{ij}}$	EQUATION 1.7 47
$V(r) = \sum \frac{1}{2} k_b (b - b_0)^2 + \sum \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum \frac{1}{2} k_\phi [1 + \cos(\phi_N - \phi_0)] + \sum \left[ \frac{Q_i Q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right]$	EQUATION 1.8 47
$E\Psi = -\frac{\hbar^2}{2m} \left( \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} \right) + V\Psi$	EQUATION 1.9 48
$H\Psi = E\Psi = (U + K)\Psi$	EQUATION 1.10 48
$H = -\frac{\hbar^2}{2m} \nabla^2 + V$	EQUATION 1.11 48
$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$	EQUATION 1.12 48
$V = \frac{Ze^2}{4\pi\epsilon_0 r}$	EQUATION 1.13 48
$H(R_1, \dots, R_N) \Psi(R_1, \dots, R_N) = E\Psi(R_1, \dots, R_N)$	EQUATION 1.14 49
$H = -\frac{1}{2} \Delta^2 - \frac{Z}{r}$	EQUATION 1.15 49
$H(R_1, \dots, R_N, R_1, \dots, R_N) = H_N(R_1, \dots, R_N) + H_E(R_1, \dots, R_N, R_1, \dots, R_N)$	EQUATION 1.16 49
$H(R_1, \dots, R_N, R_1, \dots, R_N) = H_N(R_1, \dots, R_N) + H_E(R_1, \dots, R_N, R_1, \dots, R_N)$	EQUATION 1.17 49
$F_i = M_i A_i$	EQUATION 1.18 52
$F_i = -\nabla_i V$	EQUATION 1.19 52
$-\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2}$	EQUATION 1.20 52
$F_i = -\nabla_{R_i} V(R_1, \dots, R_N)$	EQUATION 1.21 53
$F_i = -\frac{\delta E}{\delta R_i}$	EQUATION 1.22 54
$\frac{d^2 X_i}{dt^2} = \frac{F_{Xi}}{m_i}$	EQUATION 1.23 56
$R(t + \Delta T) = R(T) + \Delta T V + \frac{1}{2} \Delta T^2 A(T) + \dots$	EQUATION 1.24 57

$V(t + \Delta T) = V(T) + \Delta TA + \frac{1}{2}\Delta T^2 B(T) + \dots$	EQUATION 1.25	57
$A(t + \Delta T) = A(T) + \Delta TB + \frac{1}{2}\Delta T^2 C(T) + \dots$	EQUATION 1.26	57
$X(t + \Delta T) = 2X(T) - X(T - \Delta T) + A(T) \Delta T^2 + \dots$	EQUATION 1.27	57
$v\left(t + \frac{1}{2}\delta t\right) = v\left(t - \frac{1}{2}\delta t\right) + a(t)\delta t$	EQUATION 1.28	57
$x(t + \delta t) = x(t) + v\left(t + \frac{1}{2}\delta t\right)\delta t$	EQUATION 1.29	57
$x(t + \delta t) = x(t) + v(t)\delta t + \frac{1}{2}\delta t^2 a(t).$	EQUATION 1.30	58
$v(t + \frac{1}{2}\delta t) = v(t) + \frac{1}{2}\delta t a(t)$	EQUATION 1.31	58
$v(t + \delta t) = v(t + \frac{1}{2}\delta t) + \frac{1}{2}\delta t a(t + \delta t).$	EQUATION 1.32	58
$P(v)dv = \left(\frac{m}{2\pi k_B T}\right)^{1/2} \exp\left[\frac{-mv^2}{2k_B T}\right] dv$	EQUATION 1.33	59
$T(t) = \frac{1}{k_B N_{dof}} \sum_{i=1}^{N_{dof}} m_i  v_i ^2$	EQUATION 1.34	59
$V(t) = \sum_i \sum_{j>i} \phi( r_i(t) - r_j(t) )$	EQUATION 1.35	60
$K(t) = \frac{1}{2} \sum_i m_i [v_i(t)]^2$	EQUATION 1.36	60
$E = V + K$	EQUATION 1.37	60
$K = \frac{3}{2} N k_B T$	EQUATION 1.38	61
$\Delta G = -RTL N K_A$	EQUATION 1.39	63
$K_A = \frac{[E][I]}{[EI]}$	EQUATION 1.40	63
$\Delta G = \Delta H - T\Delta S = -RTL n K$	EQUATION 1.41	66
$E_{POT} = E_{BOND} + E_{ANGLE} + E_{TORSION} + E_{OOP} + E_{ELECTROSTATICS} + E_{VDW} + E_{CONSTRAINT}$	EQUATION 1.42	66
$H_F(\text{BINDING}) = \Delta H_F(\text{COMPLEX}) - \Delta H_F(\text{ENZYME}) - \Delta H_F(\text{LIGAND}).$	EQUATION 2.43	84

**LIST OF ABBREVIATIONS**

ATP	Adenosine tri phosphate
APCI	Atmospheric pressure chemical ionization
AIDS	Acquired immunodeficiency disease
BP	Boiling point
bs	broad singlet
CC	Cytotoxic concentration
CDI	carbonyl-di-imidazole
CM	Chorismate mutase
CNDO	Complete Neglect of Differential Overlap
d	Doublet
DCM	Dichloromethane
DOT	Directly observed therapy short course
DIEA	Di-ethyl amine
DMAP	4-Dimethylaminopyridine
DMF	Dimethyl formamide
DMSO	Di-methyl sulfoxide
DNA	Deoxyribonucleic acids
ES	Electro spray
EcCM	<i>Escheria coli</i> Chorismate mutase
EMB	Ethambutol
EPTOT	Energy potential total
EKTOT	Energy kinetic total
GA	Genetic algorithm
HETATM	Hetero atom
HF	Hartree Fock
IC	Inhibitory concentration
INH	Isoniazid
INDO	Intermediate Neglect Differential Overlap
IR	Infrared
m	multiplet
MABA	Microplate Alamar Blue Assay
MD	Molecular dynamics
MDR	Multi drug resistant
MINDO	Modified Intermediate Neglect Differential Overlap
Mp	Melting point
MMFF	Molecular Mechanic force field

---

MOPAC	Molecular orbital package
MS	Mass spectroscopy
MtCM	<i>Mycobacterium tuberculosis</i> chorismate mutase
MW	Molecular weight
NDDO	Neglect of Diatomic Differential Overlap
NHS	National health services
NMR	Nuclear magnetic resonance
PDB	Protein data bank
PES	Potential energy surface
Ph	Phenyl
ps	pico seconds
Pyr	Pyridyl
PZA	Pyrazinamide
RMP	Rifampicin
RMSD	Root mean square deviation
Rf	Retention factor.
SAR	Structural activity relationship
ScCM	<i>Saccharomyces cerevisiae</i> Chorismate mutase
SI	Selectivity Index
STM	Streptomycin
t	Triplet
TAACF	Tuberculosis Antimicrobial Acquisition and coordinating facility
TB	Tuberculosis
TER	Terminal
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TFA	Trifuroic acid
TSA	Transitional state analogue
WHO	World health organization
ZDO	Zero differential overlap

## CHAPTER 1

### INTRODUCTION

#### 1.1 Introduction to Tuberculosis:

Tuberculosis (TB) is one of the infectious diseases caused by *Mycobacterium tuberculosis* and is the leading cause of death in the recent world. More than 25% of the death during the early 19<sup>th</sup> century was reported from tuberculosis. Although the death rates fall down until the late 20<sup>th</sup> century with the development of new anti-TB drugs, drug resistant TB became another important issue in most of the TB epidemic countries. Today about one third of the world's population is infected, of these almost 15 million have the active disease at all times. In the 21<sup>st</sup> century about 9.2 million people have been estimated with TB cases. Of these about 3 millions were reported from the African and South East Asia and about 2 million from the West Pacific region. India and China has been reported the most affected region (TB epidemic region). In US, the cases have declined by 10 fold since 1953. In 2007 about 13,299 cases were reported, half of these cases were reported from foreign-born people born outside US. Blacks accounted for 45% of the cases among the US-born people (Merck Manual 2010).



Figure 1.1: Showing the distribution of TB cases worldwide. The scale indicates TB cases per 100,000 populations. (WHO, 2010)

### 1.1.1 Background:

Tuberculosis has been known under a variety of names in the past. The term 'Tuberculosis' is only known from during the first half of the 19<sup>th</sup> Century. The disease has been known by other names such as the phthisis, scrofula, tabes, bronchitis, inflammation of lungs, hectic fever, gastric fever and lupus. It was also known as the great white plague and consumption. Although it is still under research, how this organism originated, the current study shows that TB was detected 18,000 years ago (Rothschild. *et al.*, 2001). Perhaps this organism originated from cattle and then was transferred into the humans (Pearce-Duvet. 2006). That this organism has evolved gradually and has mutated is clear. It was only in 1882, Dr Robert Koch isolated the tubercle bacillus and established TB as an infectious disease. It has descended from another organism called *Mycobacterium bovis* (Ernst. *et al.*, 2007). This organism was responsible in causing infection in human being and perhaps it was only after 1930 when cow milk was regularly pasteurised before they were used by human being. Later with the development of the live attenuated strain of *Mycobacterium bovis*, which is commonly known today as the BCG vaccine, the spread of TB went down. The first child to be immunised with this vaccine was in 1921, whose mother died of TB and today this vaccine is used to immunise many children worldwide.

### 1.1.2 Signs and symptoms in tuberculosis:

Tuberculosis is a pulmonary disease and therefore the symptoms will not appear until the infection reaches the lungs. Unlike the other bacteria this organism is very slow growing and symptoms precipitate years after the individual was affected by the disease. Pulmonary tuberculosis is the condition when the pulmonary system is affected by this disease and this result in persistent cough which may be associated with bloody phlegm, breathlessness, weight loss, lack of appetite and fever. Sometimes the infection can spread from lungs into the other organs such as lymph glands, bones, digestive system, reproductive system and finally into the central nervous system. Lymph nodes being closer to the lungs are usually affected first causing lymph node swelling, pain around the neck region and fluid discharge. Skeletal tuberculosis is associated with pain in bone and joints followed by fracture thereby resulting in loss of movement. T.B in the digestive systems may cause abdominal pain, diarrhoea and rectal bleeding. If the genitourinary tract is affected it may result in blood in urine, burning sensation during urination, groin pain and frequent urge to pass urine during night. Finally if the central nervous system is affected, it may result in vomiting, stiff neck,

changes in mental states, confusion and seizures. Extra-pulmonary tuberculosis is often common with patients associated with HIV or immune-deficient diseases. (NHS)

### 1.1.3 Aetiology:

TB is caused by the inhalation of the air borne particles or droplets containing *Mycobacterium tuberculosis*. These aerosolized infectious particles are released through coughing or sneezing by an infected candidate who has active pulmonary tuberculosis. Droplets can remain suspended in the air for several hours, thereby increasing the chance of spreading the disease. Patients with pulmonary TB infect about 7 close contacts on average. Inhabitants of the same poverty or institutions are at risk where a single patient is affected with active pulmonary TB. Medical practitioners and laboratory scientists who are in close contacts with the organism are also at high risk. Patients with HIV are at high risk of getting affected by this disease because cell-mediated immune system is essential for defence against TB. Other immune suppressive illness includes diabetes, anti-cancer therapy, etc. Elderly people are more susceptible to TB and this is partly due to the weaker immune system.

### 1.1.4 Pathophysiology:

The tubercle bacilli, once inside the host body, trigger a series of events from the primary infection that gradually progress into the latent stage. The primary infection develops once the organism in the form of droplets from air enters into the respiratory tract. In the lung the bacteria are engulfed by the macrophages which trigger an inflammatory response and enter into the blood stream (Figure 1.2). These cells are the building blocks for the tubercle that defines the disease. The tubercle or the granuloma consists of a kernel of infected macrophages, surrounded by the foamy giant cells and macrophages. At this stage the host does not transmit the infection, however if the immune system changes or if the host is affected with HIV-co-infection, the centre of the granuloma undergoes caseation and the infectious bacilli go into the airways. This leads to development of a productive cough where the aerosol spreads out infectious bacilli (Russell. D, 2001).



Figure 1.2 : The mechanism of TB infection within the host (Russell. D, 2001).

In the latent stage the organism can spread to any part of the body infecting the organs. 50-80% of the people who develop active TB reactivate within 1-2 years. TB damages tissues through Type IV hypersensitivity reaction producing necrosis of the organs.

### *1.1.5 Diagnosis and Treatment:*

#### *1.1.5.1 Diagnosis of TB:*

TB is diagnosed with chest X-ray, showing an abnormality in the mid and lower lung region; sometimes this is accompanied by enlarged lymph gland. A Sputum test to detect the bacilli is the only test that confirms TB infection. TB can be diagnosed by the analysis of interferon (INF) gamma in the blood sample from a TB infected patient. In an affected patient the IFN-gamma production takes place in response to the organism in the blood. Other pathological test includes tuberculin test. This is often done in patients who have no other symptoms.

Until the first half of the 20<sup>th</sup> century no effective treatment for TB was available. People were terrified by the disease. The first antibiotic to fight against TB became available only in 1946. Later in 1952 isoniazid (INH) was discovered as another potential anti-TB drug. Still today Isoniazid (INH), Ethambutol (EMB), Pyrazinamide (PZA), Rifampicin (RMP) and streptomycin (STM) are considered as a first line drug of treatment for TB. There are five classes of other second line drugs used for the treatment of tuberculosis (Table 1.1). The second line drugs are less potent and more toxic and therefore are only used either in conjunction with the first line drugs or when the first line drugs fail to treat the infection. However with the spreading of the drug resistant TB, tuberculosis is a concern to the current world. Also with the increase in the growth of the HIV and AIDS, treatment for TB infection is getting more complicated and difficult.

First Line drugs in TB treatment	Second line drugs use in TB treatment
•Ethambutol (EMB)	•Aminoglycosides. Amikacin (AMK), Kanamycin (KM)
•Isoniazid (INH)	•Polypeptides. Viomycin, capreomycin.
•Pyrazinamide (PZA)	•Fluroquinolones Ciprofloxacin (CIP), levofloxacin
•Rifampicin (RMP)	•Thioamides Ethionamide
•Streptomycin (STM)	•Para amino salicylic acids.

Table 1.1: The first line and the second line anti TB drugs.

All these drugs act on the bacilli through different pathways. Based on the mechanism of action these drugs were classified as the drugs which act by inhibiting the cell wall synthesis of the organism (TB), drugs which act by inhibiting the ATP synthesis and drugs which inhibit the RNA synthesis as shown in the figure 1.3



Figure 1.3: The mechanism by which the drugs act on the bacilli. (Image obtained from [www3.niaid.nih.gov/topics/tuberculosis/WhatIsTB/ScientificIllustrations](http://www3.niaid.nih.gov/topics/tuberculosis/WhatIsTB/ScientificIllustrations))

#### 1.1.5.2 Isoniazid (INH):

Isoniazid or isonicotinyldiazine (INH) acid was one of the first anti-TB drugs used to treat the infection. This drug was discovered in 1912, and was only used to combat against TB infection in 1951. Isoniazid was found to have carcinogenic property; however it was very useful against TB. The mechanism by which the drug acts is still under debate. It was thought that this drug acts by inhibiting the cell wall synthesis (Figure 1.3). Isoniazid is a prodrug and is converted into the active form by the bacterial Kat G catalase-peroxidase enzyme (Suarez J, *et al.* 2009). This activated form of INH binds with the NADH to form a complex thus inhibiting the fatty acid synthesis, which is essential for the bacterial cell wall synthesis. The bacteria tend to develop resistance against INH and for this reason; it is always used in combination to other anti-TB drugs.

#### 1.1.5.3 Ethambutol (EMB):

Similar to INH, ethambutol acts by inhibiting the cell wall synthesis by obstructing the building of the arabinoside residues. These arabinoside residues are an important element as they bind to the

peptidoglycan via the galactan providing support to the cell wall. Inhibiting the binding of arabinoside residues to the peptidoglycan stops the formation of the cell wall thereby increasing permeability of the cell wall.

#### 1.1.5.4 Pyrazinamide (PZA):

Pyrazinamide was discovered in 1951. Interestingly this drug was found to work well when used in combination with other anti-TB drugs. Perhaps it accelerates the activity of INH and RMP. The drug is only active in acidic conditions (Zhang. Y, *et al.*, 2003) and is converted into the active form, pyrazinoic acid by the enzyme pyrazinamidase. The accumulation of this pyrazinoic acid disrupts the membrane potential thereby causing cell lysis. This drug was also reported to act by inhibiting the fatty acid synthase I, which is required for the synthesis of fatty acid in bacteria (Zimhony. *et al.*, 2000).

#### 1.1.5.5 Rifampicin (RMP):

Rifampicin is a semi synthetic antibiotic which was discovered in 1957 and became clinically available in 1966. This drug inhibits the synthesis of the DNA-dependant RNA polymerase by binding to the beta subunits in the bacterial cell. As a result it prevents the transcription of RNA and subsequent translation of protein. Human RNA polymerase does not bind to rifampicin and therefore is not inhibited. Resistance to rifampicin results due to the glycosylation and phosphorylation of it by the host cell.

#### 1.1.5.6 Streptomycin (STM):

Streptomycin was discovered in 1943 and was first used clinically in 1947. Streptomycin belongs to the class of the Aminoglycosides. This drug binds to the 30S subunit of the bacterial ribosome (Sharma. D, *et al.*, 2007) and thereby interferes with the binding of the formyl-methionyl-tRNA which is essential for protein synthesis leading to the death of the organism. This drug was however found to produce ototoxicity in human beings. The organism also tends to develop resistance very quickly to this drug. This led to the development of new generation aminoglycosides to combat against the resistant species. There are various aminoglycosides which are currently available such as amikacin, kanamycin, vancomycin, etc.

#### 1.1.5.7 Polypeptides:

The polypeptides which are commonly used in the treatment of TB are viomycin and capreomycin. As the name suggests, these agents are members of the nonribosomal peptide

antibiotics exhibiting anti-tubercular activity. These groups are called as the tuberactinomycin. Viomycin is the first member of the tuberactinomycin family to be identified as a potential anti-TB drug. This drug targets the bacterial ribosomes by binding to the RNA and causing disruption to the bacterial protein biosynthesis. Viomycin was found to be toxic and later it was replaced by a less toxic antibiotic of the same class called capreomycin.

#### *1.1.5.8 Fluoroquinolones:*

The first generation of fluoroquinolones were introduced in 1962 for the treatment of urinary tract infection. Today fluoroquinolones are used as broad spectrum antibiotics. These agents act by inhibiting bacterial DNA gyrase and topoisomerase IV enzyme, thereby inhibiting DNA replication and transcription. Temafloxacin, an analogue of ciprofloxacin is widely used in TB infection. Levofloxacin is equally effective as ethambutol, against the organism but lower than INH and RMP (King. J, *et al.* 1999). Other drugs which are used in TB infections are Ciprofloxacin, Moxifloxacin, etc.

#### *1.1.5.9 Thioamides*

Thioamides are used in both pulmonary and extrapulmonary TB infections. The commonly used thioamides are the ethionamide (ETA) and prothionamide. Ethionamide is structurally similar to pyrazinamide and INH. It is thought that this drug is a prodrug and undergoes intracellular modification in a similar fashion as in INH (Tommaso. A., *et al* 2002). This inhibits the synthesis of mycolic acid, which is essential for the cell wall synthesis. Similar to INH, ethionamide also inhibits the enzyme enoyl-aryl carrier protein reductase inhA (Banerjee. *et al*, 1994).

#### *1.1.5.10. Para amino salicylic acid class of drugs.*

Para amino salicylic acid as oral drugs to treat TB were first introduced in 1944. This drug is used to treat multiple drug resistance tuberculosis and is always used in combinations with other anti-TB drugs. This drug acts by inhibiting the nucleic acid synthesis in the organism. Para amino salicylic acid was also used in the treatment of anti-inflammatory bowel diseases, particularly in ulcerative colitis and Crohn's disease.

#### *1.1.6 Newer Drugs:*

Although currently there are third line treatments available in TB infections, none of these drugs were as potent as the first line treatment and at the same time these drugs were also observed with some toxicity. Thus there is an urgent need to develop a new class of anti-TB

drug. In the current decades research in TB is getting more attention and this is mainly due to the gradual increase in the number of TB cases, particularly for the multi-drug resistance TB. Within the last six decades another new class of TB drugs which are unrelated to INH or RMP has emerged into research. These new agents belong to the new class of anti TB drugs called the amidrazones; these compounds are in active stage of development and have shown some high level of toxicity. These compounds have emerged from a group of hydrazines which were found to have some anti-tubercular activity (Bertrand *et al.* 1956). Later further investigations were implemented to develop the new class of compounds called carboxamidrazone (Mamalo *et al.*, 1992, 1993, 1996). This compound was thought to have a wide range of activity from antibacterial to antifungal (Vio, *et al.*, 1989). Research into this new class of compounds was also carried in the Aston University laboratory and several N<sup>1</sup>-benzylideneheteroaryl-carboxamidrazones have shown significant anti-microbial activity (Billington *et al.*, 1998, 1998, 2000). Some were found to be effective against some strains of *Mycobacterium tuberculosis* which were resistant to isoniazid and rifampicin. Further work to develop these compounds was carried out in this thesis and is discussed in the chapter 5.

#### 1.1.7 Treatment regimen:

The treatment regimen is very important in treating the TB affected patients. Patients with TB positive should initiate first line treatment (INH, RMP, PZA, EMB) for the first two months. If the smear culture is still positive, the treatment is continued with RMP and INH daily for 4 months to complete a 6 month regimen. If the chest X-ray shows the cavity, the treatment is continued for further 3 months; this completes a total of 9 months regimen. The flow diagram below shows the treatment regimen in TB patient with concomitant HIV condition.

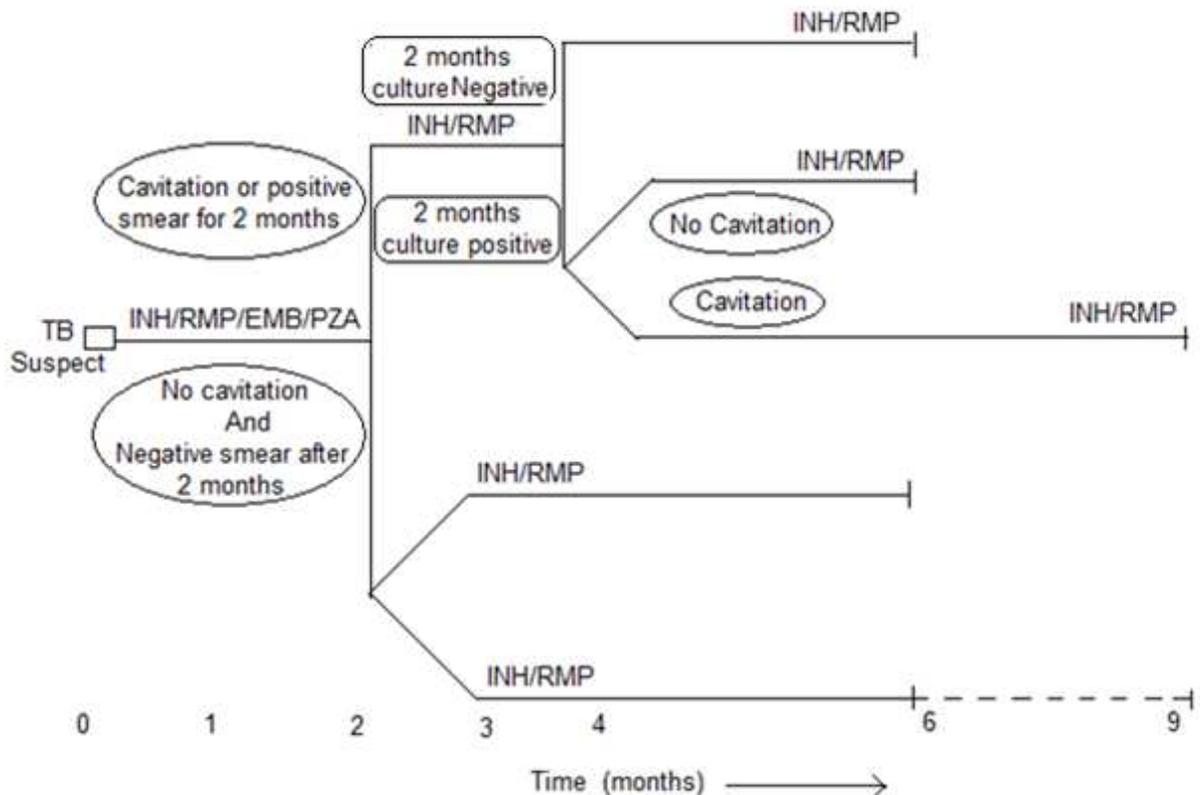


Figure 1.4: The treatment regimen in TB patients.

Unlike other anti-bacterial treatment, anti-TB drug treatments are far more complicated. This attempt causes many patients to stop the treatment without consulting with the health professionals as soon as they start feeling better. To overcome this complication directly observed therapy short course (DOTS) was introduced. While DOTS involves health workers watching their patients take their medication, FDC (Fixed-Dose Drug Combinations) became more useful in TB treatment. FDC became more beneficial in decreasing the risk of multi drug resistant tuberculosis. FDC involves an advanced method of drug delivery where two to four anti-TB drugs were administered in a fixed dose in a tablet.

#### 1.1.8 Drug resistance:

Despite innovative ideas to improve the treatment procedure, patients on anti-TB treatment regimens still develop resistance to the drug. Drug resistance due to drug happens due to inappropriate treatment or inadequate treatment, frequent inadvertently skipping the course of treatment and very often due to the underlying HIV infection followed by neglect in treatment (Zignol, *et al.*, 2006). As a result the bacterium undergoes mutation and develops resistant strains which fail to respond to the treatment. Resistance to INH is brought about by

the inactivation of the enzyme KatG which is essential for the activation of the INH when administered to the TB infected patients. Since the INH is not activated, the drug fails to exert its toxic effect in the organism. The chance of resistance is higher particularly to the less effective or the second line drugs, medium chances are for the INH and EMB and less in case of RMP (Shimao. 1987; Crofton.1970). The probability of resistance is directly proportional to the bacterial load. A bacillary load of  $10^9$  will contain several mutants resistant to one tubercular drug (Grange. *et al.*, 1990). When the organisms containing both the resistant strain and non-resistant strains are exposed to the drugs, the non-resistant strains are killed and the drug resistant strains flourish. When these forms of the strains are exposed to the second line treatment, mutants resistant to the new drugs develop. This form of TB does not respond to the standard 6 month treatment with the first line drugs and may take up to two years with the second line treatments which are less potent to the organism, more toxic and more susceptible to resistance.

## **1.2 Biosynthetic pathway in bacteria.**

### *1.2.1 Shikimic acid pathway:*

Unlike animals, bacteria do not derive their amino acids from other sources, rather all the amino acids are synthesised by the bacteria. The shikimic acid pathway (shikimate pathway) is one of the bio-synthetic routes in microorganisms to synthesise these aromatic amino acids such as L-phenylalanine, L-tryptophan, L-tyrosine and other essential amino acids. These amino acids are essential for the synthesis of folic acid in microorganisms which is necessary for their survival. This pathway is only found in plants, bacteria, fungi and some parasites but is absent in vertebrates or mammals (Haslam. *et al.*, 1993). The absence of these pathways in mammals has revealed a potential area in drug discovery.



Scheme 1.1: Showing the shikimic acid pathway (Nissen. S, et al., 2005)

The Shikimic acid pathway is an important area of research to discover antibacterial agents. Although several enzymes have been established as potential targets, perhaps the most interesting one of them was the chorismate mutase. Chorismate mutase catalyse the conversion of the chorismate to prephenate and this is a key step for the biosynthesis of several amino acids, alkaloids, etc (Scheme 1.1)

### 1.2.2 *Chorismate mutase:*

Chorismate mutase is one of the important targets to block the metabolism in bacteria, fungi or plants. It sits as a branch point enzyme in shikimic acid pathway. As mentioned this enzyme is important for the microorganism and is involved in the first step of the biosynthesis

of tyrosine and phenylalanine. Therefore blocking of such enzyme will ultimately block the metabolism in the micro organism thereby resulting in the death of micro-organism. This enzyme is however absent in humans; for these reason chorismate mutase has become a vital target in discovery of antibiotics, fungicides and herbicides. The conversion of the chorismate to prephenate takes place by Claisen rearrangement with the formation of a chair shaped intermediate. This intermediate and its conformation are of particular interest and several analogues to this intermediate have been designed as a potential target to develop anti bacterial agents.

### 1.2.3 Transition State Analogue (TSA):

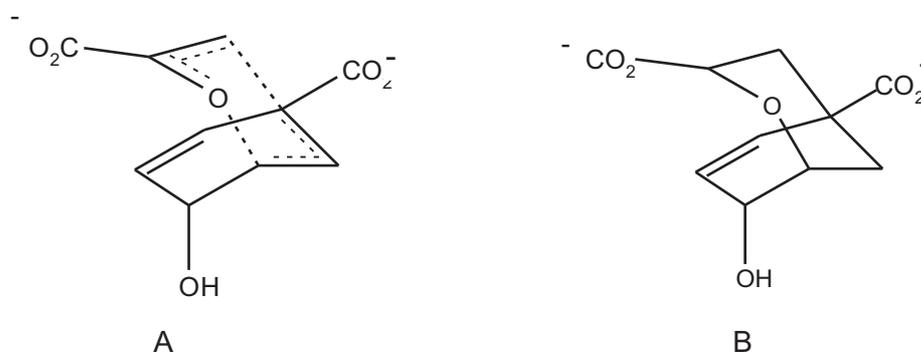


Figure1.5: Showing the transitional state conformation (A) and the transitional state analogue (B).

The figure above shows the transitional state conformation during the conversion of chorismate to prephenate. Although several inhibitors have been designed to block the conversion of chorismate to prephenate, the transitional state analogue (the endo-oxabicyclic dicarboxylic acid) was found the best inhibitor for the enzyme. (Mandal. *et al.*, 2003). TSA is a bicyclic structure and the chair form of the TSA is vital for its inhibitory activity and that the rearrangement of the chorismate to prephenate proceeds via chair like conformation (figure1) has been confirmed with various analogues of the bicarboxylic acid analogue (Andrews, P. R, *et al.* 1977).

### 1.3. Drug design methods:

The term drug design refers to the design of a chemical compound that would bind to the biological target generating the desired biological outcome. It is therefore essential that the drugs are designed based on the knowledge of the biological targets (Madsen. 2002). There are two different approaches for drug design; it can be either ligand based drug design or

structure based drug design. Ligand based drug design involves the preparation of a pharmacophore model of the binding site such that the structural characteristic would facilitate the binding of the ligand (Osman F. 2000). Structure based drug design on the other hand utilises the information from the biological active site to design ligand that would fit appropriately into the active site (Leach. A, 2007). This method thus relies on the protein crystal structures of the biological targets.

### 1.3.1 Protein:

Proteins are the polymers of amino acid. Again as the name suggests amino acids are the units of protein which contains an amino group and an acid group connected to an alpha carbon to which is attached a side chain R group. All amino acids have distinct R groups (Figure 1.6).

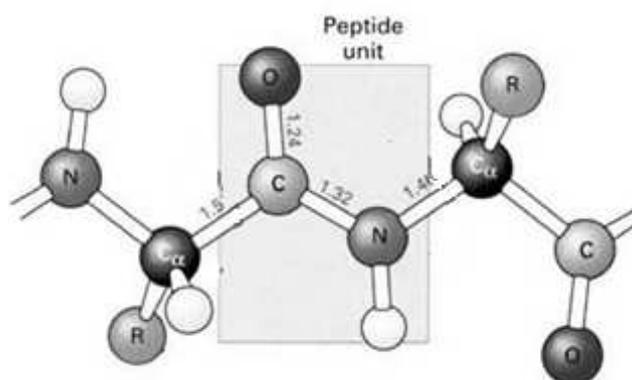


Figure 1.6: Peptide bond between two amino acids.

As shown in the figure 1.6 above, peptide bonds are the linkage between the carbonyl carbon of one amino acid to the amino nitrogen of the other. The amino acids are referred to by either a three letter code or a single letter code. The peptide bond has double bond characteristic (40%) due to resonance. The peptide bonds are therefore generally planar and this reduces the degree of freedom of the polypeptide during folding (Figure 1.7).

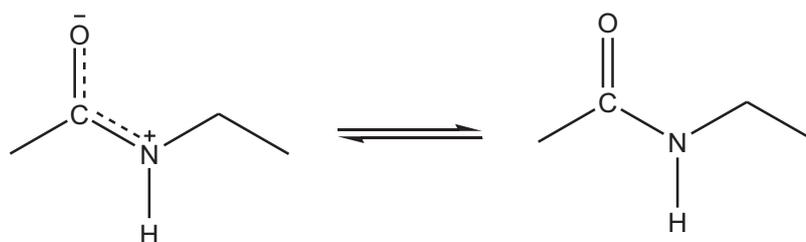


Figure 1.7 Showing the resonance within a peptide unit.

The amino acid sequence of a protein is called the primary structure. The different structure of a protein structure is illustrated in figure 1.8 below. The secondary structures in protein structure are the  $\alpha$ -helices,  $\beta$ -sheets and the  $\beta$ -turns. The secondary structure is the regular arrangement of amino acid residues in a segment of the protein structure where each residue is spatially related to its neighbour in the same way. The secondary structures are formed by the non-covalent interactions in protein such as hydrogen bonding. The tertiary structure of a protein is folded into a compacity using a variety of loops and turns. The tertiary structure of a protein is characterised by weak bonding interactions and sometimes by covalent bonds such as disulphide cross-links. The quaternary structure is formed by the non-covalent interactions between subunits of multi-subunit proteins or large protein assemblies (eg: haemoglobin contains four polypeptide chains held by non-covalent bonds).



Figure 1.8: Illustrates the different structure of a protein structure. (Image obtained from [commons.wikimedia.org/wiki/File:Main\\_protein\\_structure\\_levels\\_en](https://commons.wikimedia.org/wiki/File:Main_protein_structure_levels_en))

### 1.3.2. *Protein crystal structures*

The three dimensional co-ordinates of protein structures can be obtained following x-ray diffraction on the crystal form or NMR spectroscopy or if the experimental structure of the target is not available it can be created using homology model. A crystallographer can study the angle and the intensity of diffracted beams and can determine the arrangement of the atoms within the crystal as well as the chemical bonding and various other informations. The higher is the resolution of these x-ray diffracted beams, the better is the information about the structure.

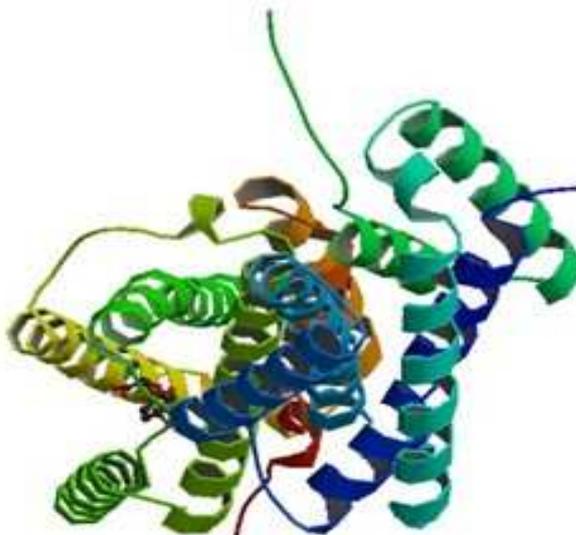


Figure 1.9: Crystal structure of an enzyme 2FP2.pdb. (chorismate mutase from *M. tuberculosis*)

The functional information of a crystal structure can be represented in the form of Cartesian coordinates (as in  $x,y,z$ ) of all the atoms present are in the form of internal coordinates. Within the internal coordinates the positions of the atoms are described as relative to the other atoms in the system. The internal coordinates are useful for a small and single molecule whereas the Cartesian coordinates are useful in defining a collection of discrete molecules. The localization of the hydrogen atoms within the protein structure are difficult to observe by x-ray crystallography. This is because the x-rays are scattered by the electron density around the atom. Each hydrogen atom contains only one electron and their influence on x-ray scattering is low and they are normally disregarded. Therefore the hydrogen positions are appointed on the basis of collected knowledge on standard bond length and angles. It is important to note that the crystalline state geometry of a molecule is based on the influence of the crystal packing force and therefore bond lengths and bond angles can differ from theoretical standard values. The energy level of the structure is thus affected by the neighbouring molecules in the crystal unit and sometimes by the solvent units. Therefore study of the global energy minima of the crystal structure is essential while using the x-ray crystal structure.

	Atom number	Atom type	Amino acids	Chain	Residue number	Cartesian coordinates (X,Y,Z)			Occupancy fraction	Temperature factor	Atom
Strand A	ATOM 1335	N	ALA	A	199	3.998	39.167	67.156	1.00	9.78	N
	ATOM 1336	CA	ALA	A	199	4.227	40.459	67.781	1.00	10.65	C
	ATOM 1337	C	ALA	A	199	4.962	40.208	69.084	1.00	11.31	C
	ATOM 1338	O	ALA	A	199	5.320	39.040	69.330	1.00	10.72	O
	ATOM 1339	CB	ALA	A	199	2.877	41.180	68.055	1.00	12.02	C
Termination of strand A	TER										
Strand B	ATOM 1342	N	SER	B	37	4.895	18.474	7.161	1.00	22.20	N
	ATOM 1343	CA	SER	B	37	4.977	16.974	7.049	1.00	19.97	C
	ATOM 1344	C	SER	B	37	3.604	16.323	6.927	1.00	18.26	C
	ATOM 1345	O	SER	B	37	2.605	16.742	7.521	1.00	17.61	O
Ligand	TER										
	HETATM 2663	C1	TSA	X	500	-4.796	5.651	29.633	1.00	11.71	C
	HETATM 2664	C2	TSA	X	500	-5.652	4.433	29.982	1.00	13.22	C

Figure 1.10: Explaining the crystal structure data from protein data bank file.

The figure 1.10 above represents a sample how the information from the amino acids are coded in a protein crystal structure file. The crystal structure of the protein may have more than one strand and this is usually defined as A,B,C etc. The amino acids from the proteins are coded with 'ATOM' and the ligands are coded with 'HETATM'. The amino acids are encoded as 3 letter code and termination of each strand is represented as TER. The atom number is the number of the atom in the whole protein. The atom type represents the atom type of the residue (N=amide, CA= alpha carbon, C= carbonyl C, O=carbonyl O). The amino acids are represented as a three letter code. The occupancy fraction represents the unit cells that contain the atom in this particular location and is usually 1.00. The temperature factor is an indication of uncertainty in this atom's position due to disorder or thermal vibration. The protein files of the crystal structure may also contains useful information such as missing residues within the data, disulphide linkage and these are added while using such crystal structure.

#### 1.4 Introduction to Molecular Modelling:

Molecular modelling is a modern technique widely used in chemistry and biology to design experiments and evaluate experimental results. Research on molecular models drew more attention after 1953 when Watson and Crick described the structure of DNA. Today molecular modelling utilises computers as a powerful tool and it requires sophisticated software and more skills to operate. Molecular modelling is used to visualise the x-ray crystal data of protein, simulate the protein and ligands, refine protein structure using molecular dynamics, evaluate the protein structure models and dock the potential structure into the enzyme active sites thereby calculating their binding affinities. Therefore the study of the potential energy in a molecular system is important while performing molecular modelling studies. Generally molecular modelling can be categorised into molecular mechanics and quantum mechanics.

#### 1.5 Molecular Mechanics:

Molecular mechanics is based on the simple empirical interaction of the atoms and molecules using classical Newton's Law. Molecular mechanics consider atoms as spheres and bonds as springs; a brief overview is shown below (Figure 1.11). These methods are used to calculate the forces acting on the atoms in molecular system and these forces are used to calculate the potential energy of the molecular system.



Figure 1.11: Showing the distribution of the potential energy in a molecule. ([http://en.wikipedia.org/wiki/File:MM\\_PEF.png](http://en.wikipedia.org/wiki/File:MM_PEF.png))

The potential energy of a molecular system in molecular mechanics is calculated using force fields. The four main components on which force field depends are the bond stretching, angle-bending, bond rotation (torsion) and the non-bonded interactions which are the van der Waals and electrostatic. The total energy of the system can be represented by the following equation:

$$E_{tot} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{elec} + \dots \quad \text{Equation 1.1}$$

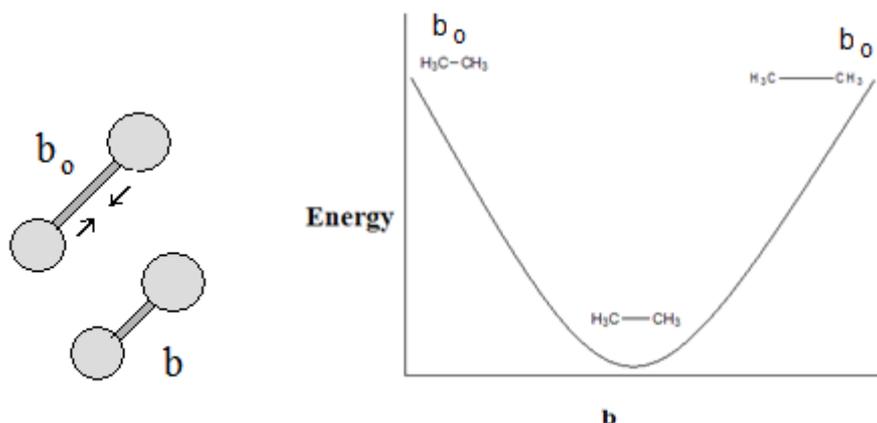
Where  $E_{tot}$  is the total energy of the molecule,  $E_{str}$  is the bond stretching energy,  $E_{bend}$  is the energy due to bond bending,  $E_{tors}$  is the torsional energy term,  $E_{vdw}$  is the van der Waals energy and  $E_{elec}$  is the electrostatic energy. Molecular mechanics thus enables the calculation of the total change in the energy system from the reference or initial energy level. This therefore involves the measurement of the energies from bond stretching, angle bending, bond rotation, and the other non-bonded interactions which is known as force fields.

#### 1.5.1 Bond stretching:

Since molecular mechanics typically treats atoms as spheres, and bonds as springs, the mathematical expression from Hooke's Law can be used to describe the ability of bonds to stretch, bend, and twist. According to Hooke's Law:

$$E = \frac{kx^2}{2} \quad \text{Equation 1.2}$$

Where,  $E$  stands for the potential energy of a deformation,  $x$  is the displacement of the spheres and  $k$  is the force constant. The same applies to all the bonded interaction as discussed earlier. The bond stretching can be represented as.



Graph 1.1: An example of energy change due to bond stretching variation

$$E_{str} = \sum \frac{1}{2} k_b (b - b_0)^2 \quad \text{Equation 1.3}$$

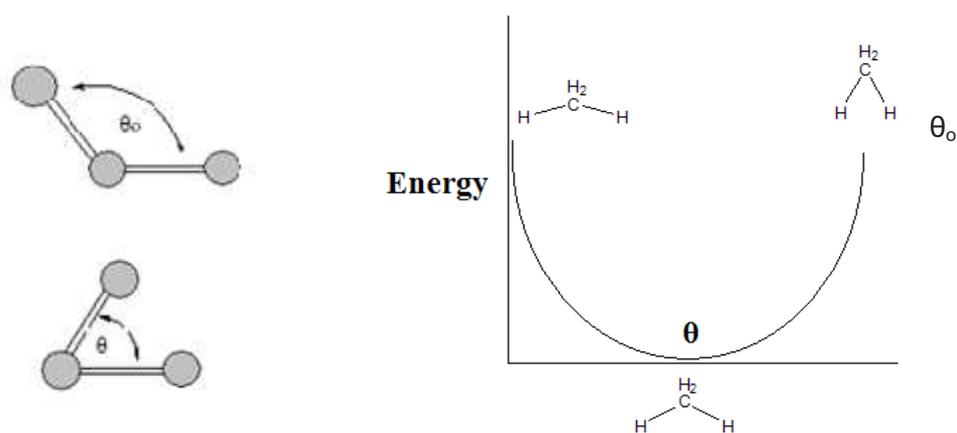
Where  $k_b$  is the bond stretching force constant,  $b$  is the actual bond length and  $b_0$  is the unstrained bond length.

### 1.5.2 Bond bending:

The energy due to angle bending  $E_{bend}$  is the energy required to bend the angle to and away from the actual angle. It can be represented as:

$$E_{bend} = \sum \frac{1}{2} k_\theta (\theta - \theta_0)^2 \quad \text{Equation 1.4}$$

Where  $K_\theta$  is the angle-bending force constant that is determined empirically,  $\theta_0$  is the equilibrium value, in other words the natural bond angle for  $\theta$  is the actual bond angle in the molecule.



Graph 1.2: The change of energy due to bond bending.

If the optimal bond angle for H-Csp<sup>2</sup>-Csp<sup>3</sup> is 122°, then any change in angle ( $\theta - \theta^0$ ), either wider or narrower, will increase the energy of the molecule.

### 1.5.3 Torsional energy:

The torsional energy represents the amount of energy that must be added to or subtracted from the Stretching Energy + Bending Energy + Non-Bonded Interaction Energy terms to make the total energy agree with experiment or rigorous quantum mechanical calculation for a model dihedral angle.

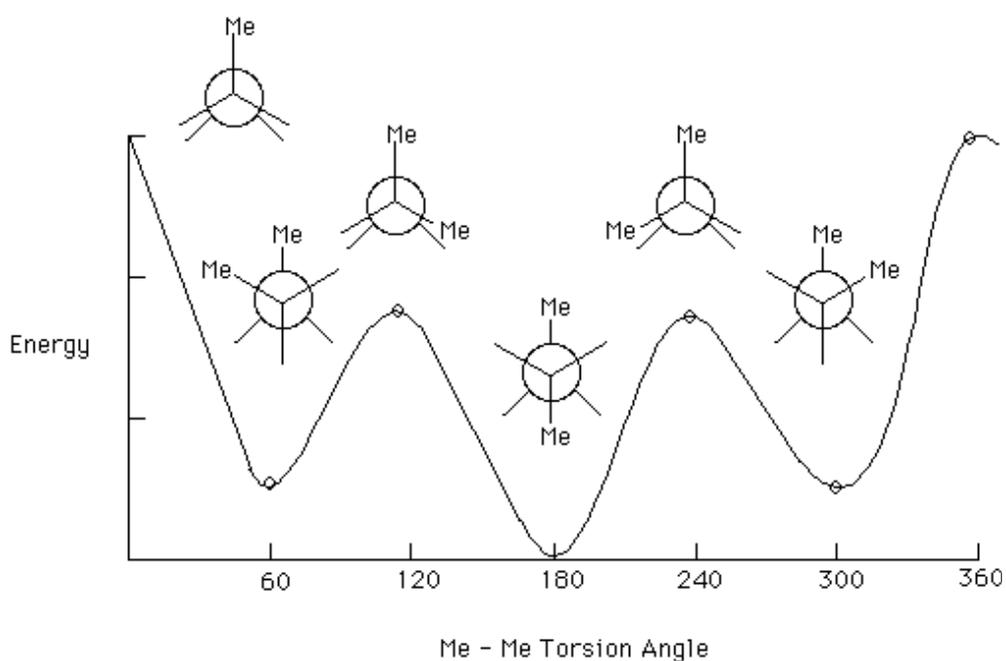


Figure 1.12: The variation of energy due to torsion angle variation.

The energy due to torsional angle can be represented as

$$E_{tors} = \sum \frac{1}{2} k_{\phi} [1 + \cos(n\phi - \phi_0)] \quad \text{Equation 1.5}$$

Where  $k_{\phi}$  is the torsional barrier,  $\phi$  is the actual torsional angle;  $n$  is the periodicity or the number of energy minima within one full cycle.  $\phi_0$  is the reference torsional angle.

#### 1.5.4 Non bonded interactions:

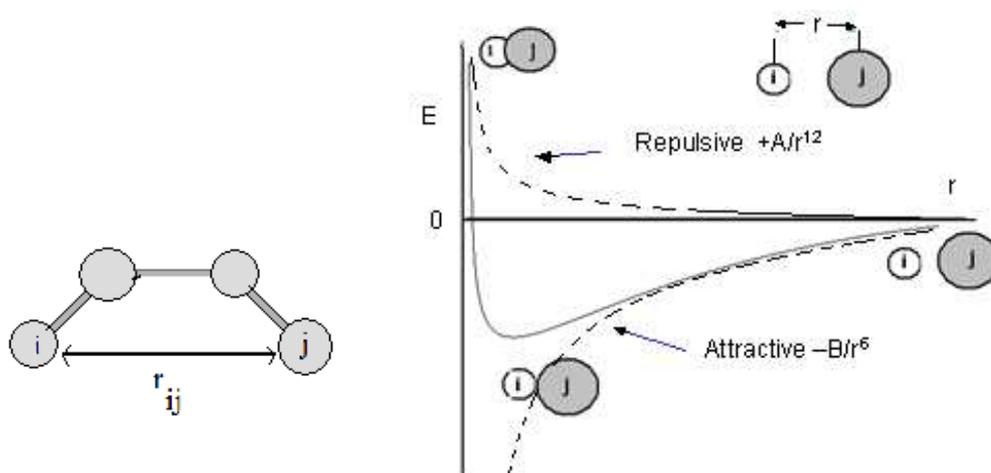
The non-bonded energy accounts for repulsion, van der Waals attraction, and electrostatic interactions.

##### 1.5.4.1 Van der Waals interactions:

The Van der Waals interactions are modelled using the Leonard-Jones potential. Van der Waals attraction occurs at short range, and rapidly dies off as the interacting atoms move apart by a few Angstroms. Repulsion occurs when the distance between interacting atoms becomes even slightly less than the sum of their contact radii. The Van der Waals interactions can be represented as:

$$E_{vdw} = \sum \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \quad \text{Equation 1.6}$$

Where  $A_{ij}$  is the repulsive term coefficient,  $B_{ij}$  is the attractive term coefficient and  $r_{ij}$  is the distance between the atoms  $i$  and  $j$ .



Graph 1.3: The change of energy due to van der Waals interaction.

#### 1.5.4.2. Electrostatic interactions:

The electrostatic energy is a function of the charge on the non-bonded atoms, their inter-atomic distance, and a molecular dielectric expression that accounts for the attenuation of electrostatic interaction by the environment. The electrostatic contribution is modelled using a Coulombic potential. It is represented as.

$$E_{ele} = \sum \frac{Q_i Q_j}{\epsilon r_{ij}} \quad \text{Equation 1.7}$$

Where  $\epsilon$  is the dielectric constant,  $Q_i$  and  $Q_j$  are the atomic charges of interacting atoms and  $r$  is the inter-atomic distance.

In summary the potential energy in the force fields can therefore be represented as

$$V(r) = \sum \frac{1}{2} k_b (b - b_0)^2 + \sum \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum \frac{1}{2} k_\phi [1 + \cos(\phi_N - \phi_0)] + \sum \left[ \frac{Q_i Q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right] \quad \text{Equation 1.8}$$

Where  $V(r)$  is the potential energy, which is the function of the position ( $r$ ) of the atoms (Ponder, J, *et al.*, 2003). Force field thus can be defined as an energy system of a molecular system which changes as the bonds are rotated and it also depends on the non-bonded interactions of the molecular system. Force fields are empirical, that is there is no correct form for a force field. If one form is shown to perform better than another, it is likely that form will be favoured. The functional form employed in molecular mechanics is often a compromise between the accuracy and computational efficiency. As the performance of the computers increase the models became more sophisticated.

In case of proteins the number of the non-bonded interactions increases with the square of the interacting sites, thus the number of the interacting sites can be reduced by considering the group as an atom (pseudo-atom) e.g. methyl group can be considered as a single atom. This is what referred to as *united atom force field* (Yang, *et al.*, 2006) and it saves a significant computational effort in carrying out calculation.

## 1.6 Quantum mechanics:

Quantum mechanical calculation of a molecular system involves the molecular orbital calculation and offers the in-depth description of the molecule's chemical behaviour. Quantum mechanical simulation is based on the law of quantum mechanics. Quantum mechanics states that energy and other related properties of molecules can be solved by using Schrodinger equation.

$$E\Psi = -\frac{\hbar^2}{2m} \left( \frac{\partial^2\Psi}{\partial X^2} + \frac{\partial^2\Psi}{\partial Y^2} + \frac{\partial^2\Psi}{\partial Z^2} \right) + V\Psi \quad \text{Equation 1.9}$$

Where,

$\hbar^2$  is the square of the reduced Planck constant  $\frac{h}{2\pi}$ .

$m$  is the mass of the particle

$\Psi$  is the particle's wave function

X,Y,Z are the co-ordinates (3 dimensional variables)

$\frac{\partial^2}{\partial X^2}$  is the second derivative with respect to dimension X.

V is the potential of the particle at X,Y,Z co-ordinates.

E is the energy of the particle.

In the above equation a single particle of mass  $m$  is moving through a space given by a position vector  $r$  and time  $t$  under the influence of an external field  $V$ .  $\Psi$  is the wave function of the particle and it is from this function the properties of the particles can be derived. The above equation can be written as

$$H\Psi = E\Psi = (U + K) \Psi \quad \text{Equation 1.10}$$

Where,

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V \quad \text{Equation 1.11}$$

$$\nabla^2 = \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2} \quad \text{Equation 1.12}$$

$$V = \frac{ze^2}{4\pi\epsilon_0 r} \quad \text{Equation 1.13}$$

In the above equation  $H$  is the Hamiltonian and  $\Psi$  is the wave function representing the distribution of the electrons around the molecule. The total energy  $E$ , contains the potential energy  $U$  and Kinetic energy  $K$ . The above equation can also be written as

$$H(\mathbf{r}_1, \dots, \mathbf{r}_n) \Psi(\mathbf{r}_1, \dots, \mathbf{r}_n) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_n) \quad \text{Equation 1.14}$$

Where, N is the number of particles (electron and proton) with r position vector. In an atom that contains a single electron, the potential energy depends upon the distance between the electron and the nucleus and thus the Hamiltonian equation is as follows.

$$H = -\frac{1}{2}\Delta^2 - \frac{Z}{r} \quad \text{Equation 1.15}$$

Where Z is the nuclear charge and r is the distance of the electron from nucleus. Unfortunately, the Schrodinger equation can only be used to solve for very small molecules such as hydrogen and helium. Therefore approximations must be introduced to make this method applicable for the systems containing more than one atom. The first approximation attempts to differentiate nuclei and electrons where the nuclei are considered much heavier than electrons and move much more slowly so that molecular systems can be viewed as electrons moving in a field of fixed nuclei (the Born-Oppenheimer approximation). This allows the separation of Hamiltonian into nuclear  $H_n$  and electronic parts  $H_e$ .

$$H(\mathbf{r}_1 \dots, \mathbf{r}_n, \mathbf{R}_1 \dots, \mathbf{R}_n) = H_n(\mathbf{R}_1 \dots, \mathbf{R}_n) + H_e(\mathbf{r}_1 \dots, \mathbf{r}_n, \mathbf{R}_1 \dots, \mathbf{R}_n) \quad \text{Equation 1.16}$$

Where  $r_1$  are the electronic and  $R_1$  are the nuclear coordinate.  $H_e$  depends on the nuclear coordinates.

$$H(\mathbf{r}_1 \dots, \mathbf{r}_n, \mathbf{R}_1 \dots, \mathbf{R}_n) = H_n(\mathbf{R}_1 \dots, \mathbf{R}_n) + H_e(\mathbf{r}_1 \dots, \mathbf{r}_n, \mathbf{R}_1 \dots, \mathbf{R}_n) \quad \text{Equation 1.17}$$

The separation of the Hamiltonian allows the separation of the Schrödinger equation. The Hartree-Fock method finds its application in the solution to the Schrödinger equation. The Hartree-Fock method is also known as the self-consistent field method (SCF). The ab initio method utilises Hartree-Fock (HF) method in which the calculations are simplified by considering the electrons in an average field. This results in the calculation of the Hamiltonian for each electron independently using a new term for its interaction with the other electrons. Thus in order to search the global minima by conformational search, it requires large computer time and this method is limited to compounds containing between ten to twenty atoms.

### 1.6.1 Energy calculation using ab initio quantum mechanical methods:

Ab initio methods do not include any empirical or semi-empirical parameters in their equations and derived directly from theoretical principles, with no inclusion of experimental data. The ab initio method was introduced to model small molecules, lately this method

became available for large molecules (Kenny. B, 1993). The simplest type of *ab initio* electronic calculation is the HF method in which the electron-electron repulsion is not taken into account and the average effect is included. The obtained approximate energies are expressed in terms of the system's wave function and are equal or greater than the exact energies. Several versions of the *ab initio* methods are available commercially such as the Gaussian series (Pople. J, 1950).

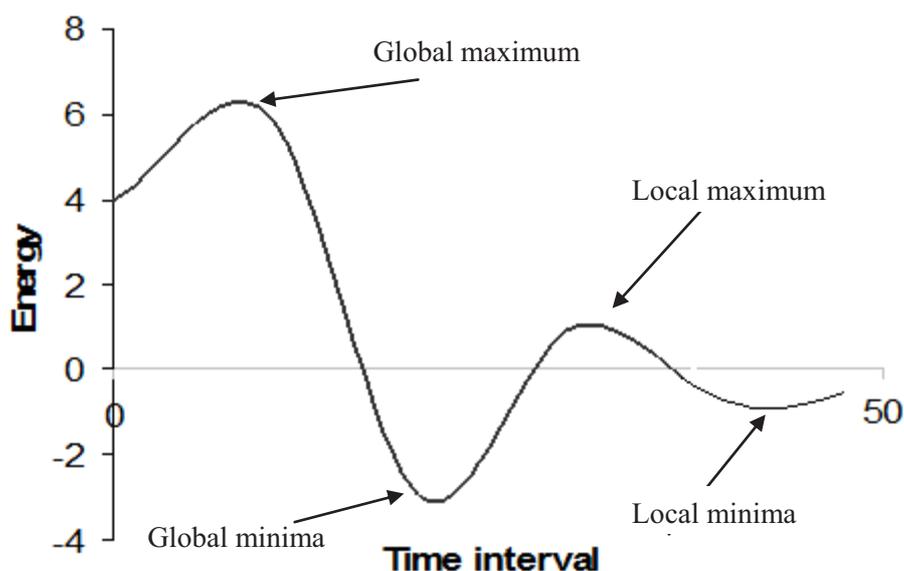
### 1.6.2 Energy Calculations using semi empirical methods:

The semi empirical calculations are based on the same or related quantum mechanical principles; with certain approximations for the majority of the electron-electron interactions to simplify the computations or to include some empirical parameters based on the experimental data. There are various types of the semi empirical methods which differ in their approach to approximation. This can be categorised based on the two research groups, the Zero Differential Overlap (ZDO) approach, the Complete Neglect Differential Overlap (CNDO) approach (Pople *et al.* 1965), intermediate neglect of differential overlap approach (INDO) and the Neglect of Diatomic Differential Overlap approach (NDDO) was carried out by John Pople and co-workers. These methods are rarely used and are modified by Michael Dewar and co-workers to make the calculation more realistic and requiring significantly less amount of time. This includes the modified neglect differential approach (MNDO), Modified Intermediate Neglect Differential Overlap approach (MINDO), (AMI), (PM3) and (SAM1). These calculations are widely used in MOPAC and AMPAC packages.

While the semi empirical methods require less computer resources when compared to the *ab initio* methods, they still are computer intensive. The major drawback of this method is that its application is limited to systems for which appropriate parameters have been developed.

## 1.7 Geometry optimization:

The geometry of a molecule determines its physical and chemical properties. It is therefore important to have the correct geometry of the molecule before performing any computation experiments. The arrangement of the atoms in the molecules, perhaps the electrons around the atoms determines the energy level of the molecule. Therefore the geometry of the molecular system has direct impact on the energy system of the molecule. One way to observe the effect of potential energy surface is to calculate the potential energy surface (PES).



Graph 1.4 : Showing the flow of energy (potential energy) level during simulation.

As shown in the graph above, the global maximum is the point on the potential energy surface where the molecule is at its highest value. At local maxima the potential energy of the molecule is at its highest value at that particular conformation. The local minimum is the point on the PES where the energy of the molecule is at its lowest value at that particular conformation. At global minima the molecule is at its lowest energy level in the entire PES. The saddle point is the point on the PES where the potential energy is maximum in one direction and minimum on the other. This is of particular interest to the chemist as they sometimes produce transitional states. The geometry generated from the saddle point may produce a structure similar to transitional state when the geometry is optimized. At each and every point on this PES the molecule has its own conformation; it is the purpose of the geometry optimization to locate the minima based on the geometry of the molecule. This can be performed by:

#### 1.7.1 Geometry optimization by molecular mechanic methods.

The molecular mechanical approach to optimize the geometry is carried out by minimization and molecular dynamics simulation method.

### 1.7.1.1 Minimisation:

In molecular modelling and drug designing every NMR or crystal structure downloaded are just the snapshots of one conformation. Minimisation is carried out to identify the corresponding local minimum on the potential energy surfaces that describes the relative energies of the conformations and geometries of molecules within a particular system. In this investigation the system is carried through a series of steps until it is no longer possible to decrease the energy gradient any further within the specified criteria. The end point yields a snapshot representing the most energetically favourable state. It is important to note that the minimisation of the system depends on the starting conformation of the system. This can be resolved by the additional use of Molecular dynamics (MD).

### 1.7.1.2 Molecular dynamics:

Molecular dynamics (MD) involves the study of the motion of the atoms in the complex protein. In drug designing molecular dynamics studies are used to refine the geometry of the proteins and to study their motions. A good molecular dynamics study for a system depends on the time length chosen for its simulation; this is because during these simulations the molecules inherit various conformations. The final result is the best possible conformation and invariably depends on how many best conformations did the molecules pass through with minimum energetic values. Molecular dynamics follows the classical Newton's Law:

$$F_i = m_i a_i \quad \text{Equation 1.18}$$

Where  $F$  is the force,  $m$ , is the mass and  $a$ , is the acceleration. The force can also be expressed as the gradient of potential energy,

$$F_i = - \nabla_i V \quad \text{Equation 1.19}$$

Combining these two equation gives:

$$- \frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \quad \text{Equation 1.20}$$

Where,  $V$  is the potential energy of the system,  $t$  is the time and  $r$  represents the Cartesian coordinates of the atom. In more pictorial term the atoms will move, bumping into each other and if the system is fluid it will oscillate in waves in concert with their neighbours, perhaps evaporating away from the surface if there is free surface, in other words it will do everything what atoms in real substance would do. In molecular dynamics the atoms interact with each other creating a force, which once again acts on the neighbouring atoms and the atoms move further because of such interaction; as these atoms move their relative position and force changes as well (Ercolessi. 1997). Thus the velocity and position of a new atom is updated to a new velocity and position. These simulations are the representation of the

*potential energy function* of the molecule from its atomic co-ordinates, which can be represented as,

$$F_i = -\nabla_{r_i} V(r_1, \dots, r_N) \quad \text{Equation 1.21}$$

As mentioned earlier;  $F$  is the force and  $V(r_1, \dots, r_N)$  represent the pairwise interaction. The lower energy states of the potential energy function are populated at thermal equilibrium and forces on the individual atoms are related to the gradient of this function; such functions are commonly referred to as force fields.

### 1.7.2 *Geometry optimization by semi-empirical methods.*

MOPAC (Molecular orbital Package) was used in this thesis as a semi-empirical method of optimization. MOPAC follows the Schrödinger equation 1.10. It finds the minimum-energy geometries which are near to the stationary geometries. Stationary geometries are those for which the energy does not change and the net force on each atom is zero. The minimum geometry is obtained by calculating the forces acting on the system, changing the geometry so as to lower the total energy. (Cache 2002). The default geometry optimizer in MOPAC uses Baker's EigenFollowing method. If this is *not* wanted, for example, if there is a need to reduce memory demands, then the Broyden Fletcher Goldfarb Shanno (BFGS) method can be used.

MOZYME is copyrighted software by Fujitsu. The geometry optimization carried out in this thesis was performed in MOZYME which was build on MOPAC methods. The methods in MOZYME were similar to MOPAC with the advantage of its use to study larger systems. This happens because the calculation time in MOPAC increases as the cube of the number of atoms, MOZYME alters the way in which the electronic structures are calculated and uses the localised molecular orbital. Therefore big molecules like proteins require MOZYME for its geometry optimisation and calculation of polarizability. However the calculations that need high precision were not possible to perform in MOZYME (Cache 2002).

### 1.7.3 *Geometry optimization by ab initio methods.*

As discussed (Equation 1.17 and 1.18) the motion of the electron and nuclei can be separated using the Born-Oppenheimer approximation. The initial geometry can be calculated by performing the single point energy calculation. The forces on the nuclei can be

calculated from the wave function using the Hellmann-Feynman theorem (Feynman 1939). The force on a nucleus  $I$  with position  $R_i$  is:

$$\mathbf{F}_I = - \frac{\delta E}{\delta R_i} \quad \text{Equation 1.22}$$

Where,  $E$  is the energy. This force can be used to find the ground state position of the atoms. As the force move toward minima, the motion for the nuclei will move toward equilibrium structure (Cheung.D, 2002).

## 1.8 Background and development of the Amber Force field.

Although a large number of protein force fields were used over the years based on equation 1.10, research interest with protein force field has only started after the development of the Monte Carlo simulation of proteins. The most popular of them were the ECEPP potentials from Scheraga and co-workers (Momaney *et al.*, 1975; Némethy *et al.*, 1983), and the consistent force field (CFF) developed by the Lifson group (Lifson and Warshel, 1969; Hagler *et al.*, 1974; Hagler and Lifson, 1974; Niketic and Rasmussen, 1977). All these force fields provided enough experience to develop more sophisticated force fields and by the 1980 research for the development of new generation of force field started; this force field is known today as the Amber force field.

During the earlier stage of this development, computers which were available had a limited power and therefore hydrogen atoms were excluded from the study. Later this hydrogen bonding were found important and were included in the force field (Weiner *et al.*, 1984) by incorporating the new concept of united atoms where the hydrogen atoms attached to a single carbon were considered as a single entity. This force field was incorporated into the Amber molecular mechanics package, which was also at the early stage of development (Weiner and Kollman *et al.*, 1981). This resulted in significant increase in calculation time. Unfortunately it was difficult with united atom models to make comparisons between computed and observed vibrational frequencies and therefore an extension of this model was published in 1986 (Weiner *et al.*, 1986). With the increase in speed of the computers a new force field was developed in 1994 known as *ff94* force field (Cornell *et al.*, 1995). Later a more serious attempt was made to develop parameters for the molecules other than the protein (Fox and Kollman, 1998). This led to the development of a key program in Amber called antechamber. The parameters from 1984 and 1986 involved the use of gas phase simulation; later the empirical water models were developed such as TIP3P and incorporated into the force field.

*Amber 11* is the latest version of Amber project with some little updates in force fields from Amber 10 and Amber 9. The work carried out in the current thesis has used Amber 9. The idea on the desire to generate simple “generic” force field models has been utilised widely in simulation of protein and nucleic acids, using simpler functions to describe bonds, angles and dihedral terms and attempts to employ as few empirical parameters as possible. Developing the torsional parameters in the Amber force field was difficult. The implementation of restrained ESP-fit method has restricted the use of torsional terms in Amber force field. The energy profile for rotation about a bond depends upon the atom types of the two atoms that comprise the central bond and not upon the terminal atoms (Leach, A., 2001). Interestingly *ff94* force field has used more than one term to describe the torsional terms (Verma, *et al.*, 2001). Amber force field has improved the reproduction of interaction energies and free energies of solvation (Cornwell *et al.*, 1995; Wang *et al.*, 2000). Also further implementation of scaling factor for the charges alongside and thereby implementing *ff02* and *ff02EP* force fields managed to produce better representation of the H-bond interactions.

Another force field developed in 1996 by Halgren and co worker called MMFF94 force field was also very popular in the development of force fields (Halgren, 1996abcd; Halgren and Nachbar, 1996; Halgren 1999ab). MMFF94 is similar to MM3 and achieves similar accuracy for small drug like ligands although it is applicable for the biological protein complexes. Halgren pioneered a novel way to more accurately model van der Waals interactions by using buffered formation in developing MMFF94 force field. The force field designed for the MMFF approach was restricted to certain types of uses and parameterised accordingly. The MMFF94 was designed specifically for MD simulation whereas the MMFF94s (static variant) performs better in minimisation. Of all the force fields developed, the most successful force fields so far are the MMFF94 and Amber force field.

The Amber project is based on these two force fields used in an overlapping manner. The Amber force field developed by Kollman and co-workers (Cornell *et al.*, 1995, Pearlman *et al.*, case *et al.*, 2002) and the MMFF94/MMFF94s force field described by Halgren (Halgren, 1996a; Halgren, 1999a). The Amber model is simpler in functional content than MMFF, but requires more computational input if the system carries ligands. The MMFF approach has been used to describe the structural geometries and charges derived from quantum-mechanical calculations on a large set of diverse small molecule species. In this way the complexity for this program was increased with distinct advantages when applied to wide chemical functionalities and features.

### 1.8.1. Amber Project:

Amber (Assisted model building with Energy refinement) is a computer based software mainly applied to nucleic acids and proteins and is best for simulating biological systems. Amber program is a collective name for a suite of programs to carry out molecular dynamics simulation. All these programs are clustered in 3 main areas (i.e preparatory program, simulation programs and analysis programs) as shown below.



Figure 1.13: The Amber program suite (Amber Manual. 2005).

### 1.9. Selection of Algorithm:

In molecular dynamics the configuration of the system is obtained by integrating the Newton's law. The result is the trajectory which is obtained by solving the differentiating Equation 1.19.

$$\frac{d^2 X_i}{dt^2} = \frac{F_{Xi}}{m_i} \quad \text{Equation 1. 23}$$

The equation above shows the motion of a particle of mass  $m_i$  along one coordinate  $X_i$  with  $F_{xi}$  is the force on the particle in that direction. Thus the particles move together colliding with

each other and thus changing position. The position of the particle  $X$  thus changes by  $V_i \delta t$ , where  $V_i$  is the constant velocity and  $\delta t$  is the time between collisions. In more realistic models the forces from each particle change as the particle changes its position. Thus the acceleration of the particles at time  $t$  can be calculated from the force; this can then be used to calculate the new position and velocities at time  $t+\delta t$ . Similarly the velocities at time  $t+\delta t$  was determined at  $t+2\delta t$ . There are several algorithms for integrating the equation of motion which are commonly used in molecular dynamics calculation. All algorithms can be derived by the Taylor series expansion:

$$r(t + \delta t) = r(t) + \delta t v + \frac{1}{2} \delta t^2 a(t) + \dots \quad \text{Equation 1.24}$$

$$v(t + \delta t) = v(t) + \delta t a + \frac{1}{2} \delta t^2 b(t) + \dots \quad \text{Equation 1.25}$$

$$a(t + \delta t) = a(t) + \delta t b + \frac{1}{2} \delta t^2 c(t) + \dots \quad \text{Equation 1.26}$$

### 1.9.1 Verlet algorithm:

The most widely used of these are the Verlet algorithm (Verlet *et al.*, 1967). The Verlet algorithm uses the position and acceleration at time  $t$  from the previous steps ( $t-\delta t$ ) to calculate the new position at  $t+\delta t$ . Knowing  $x(t)$ ,  $x(t - \delta t)$ ,  $a(t)$ . Thus the velocities in Verlet algorithm can be represented as:

$$x(t + \delta t) = 2x(t) - x(t - \delta t) + a(t) \delta t^2 + \dots \quad \text{Equation 1.27}$$

Then  $a(t + \delta t)$  the new velocities can be calculated from  $x(t + \delta t)$  via force field. Similarly the velocities at time  $(t + \delta t)$ , can be determined by the equation 1.27. Unfortunately the lack of the explicit velocity term in the equation makes it difficult to obtain the velocities. That is the velocities are not available unless the positions are computed at the next step.

### 1.9.2 Leap-frog algorithm

The leap-frog algorithm (Hockney. *et al.*, 1970) has the advantages over Verlet algorithm; it explicitly includes the velocity and does not require the calculation of the differences of large numbers. Knowing  $x(t)$ ,  $v(t - \frac{1}{2} \delta t)$ ,  $a(t)$ . The leap-frog algorithm is expressed as:

$$v(t + \frac{1}{2} \delta t) = v(t - \frac{1}{2} \delta t) + a(t) \delta t \quad \text{Equation 1.28}$$

$$x(t + \delta t) = x(t) + v(t + \frac{1}{2} \delta t) \delta t \quad \text{Equation 1.29}$$

Then  $a(t + \delta t)$  the new velocities can be calculated from  $x(t + \delta t)$  via force field and the velocities at time  $(t + \delta t)$ , can be determined by the equation 1.28 and 1. 29.

### 1.9.3 Velocity Verlet method

The Velocity Verlet method gives position, velocity and acceleration at the same time in three steps. Knowing  $x(t)$ ,  $v(t)$ ,  $a(t)$ . To calculate the new velocities the positions at  $t+\delta t$  and at time  $t$  were first calculated using the equation

$$x(t + \delta t) = x(t) + v(t)\delta t + \frac{1}{2} \delta t^2 a(t). \quad \text{Equation 1.30}$$

The velocities at time  $t+1/2\delta t$  are then determined using:

$$v(t + \frac{1}{2}\delta t) = v(t) + \frac{1}{2} \delta t a(t) \quad \text{Equation 1.31}$$

The new forces are then calculated from the current position giving  $a(t+ \delta t)$ . The velocities at time  $(t+\delta t)$  using the equation below.

$$v(t + \delta t) = v(t + \frac{1}{2}\delta t) + \frac{1}{2} \delta t a(t + \delta t). \quad \text{Equation 1.32}$$

The velocities at time  $(t + \delta t)$  thus can be determined by the equation 1.30, 1.31 and 1.32.

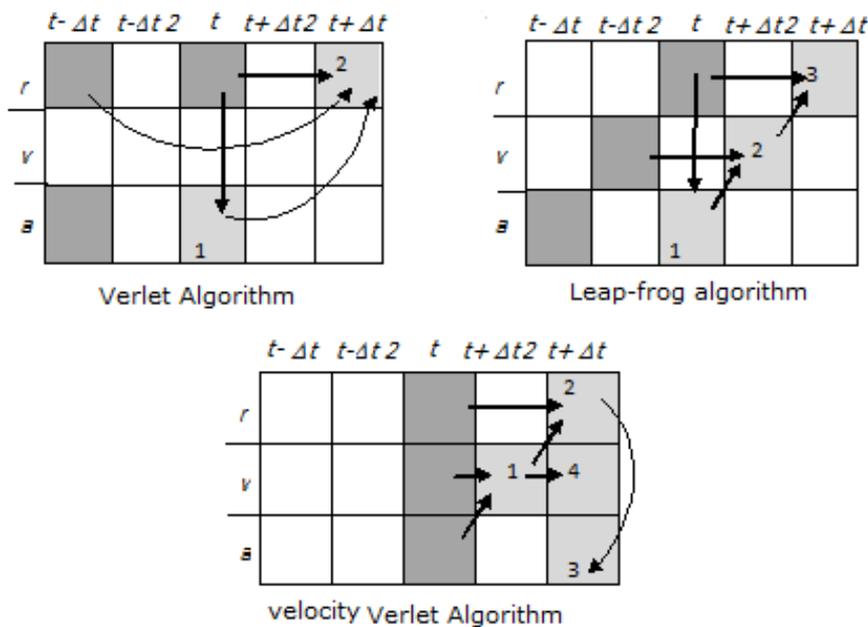


Figure 1.14: Stepwise view of the verlet integration algorithm and its variants. At each algorithm the dark and the light grey cells indicate the initial and calculating variables, respectively. The numbers in the cells represent the orders in the calculation procedures. The arrows point from the data used in the calculation of the variable that is being calculated each step (Becker.O, *et al.*, 2001).

### 1.9.3 Beeman's algorithm:

The Beeman's algorithm (Beeman 1976) is another algorithm which is related to the Verlet algorithm. It uses a more accurate expression for the velocity and therefore it produce better energy conservation because the kinetic energy is calculated directly from the velocities. These expressions are more complex and computationally more expensive to perform.

## 1.10. Starting, running, measuring and analysing simulation program:

### 1.10.1 Starting:

In order to start a simulation, it requires Newton's equation (1.10) of motion with two initial values which are typically a set of coordinates ( $r_0$ ) and a set of initial velocities ( $V_0$ ). The initial coordinates are usually obtained mainly from the X-ray crystallography and NMR experiments. This information is sometimes not enough and therefore it undergoes further preparation. The hydrogen atoms in particular are not determined by the X-ray Crystallography and therefore they need to be assigned. Sometimes there are missing residues within which also need to be added; these often happen when there are loops within the structure. These loops are not very well identified during the experiments and thus not well captured by NMR or X-ray diffraction. At other times they reflect the part of the molecules which were removed intentionally to facilitate the crystallisation process.

The initial velocities were not obtained from the experimental information. The only relevant information available about the atomic velocities was the system's temperature  $T$ . The initial velocities were randomly assigned taking from the standard Maxwell distribution at a certain temperature  $T$  as shown below.

$$P(v) dv = \left( \frac{m}{2\pi k_B T} \right)^{1/2} \exp \left[ \frac{-mv^2}{2k_B T} \right] dv \quad \text{Equation 1.33}$$

$$T(t) = \frac{1}{k_B N_{dof}} \sum_{i=1}^{N_{dof}} m_i |v_i|^2 \quad \text{Equation 1.34}$$

Velocities are initially assigned at low temperature which is then gradually increased until the system achieves the desired temperature. While doing this the system will have a small linear momentum, corresponding to a translational motion of the whole system. This is not very appropriate and thus the velocity of each particle was subtracted from this component to operate in a zero total momentum condition.

### 1.10.2 Running:

The first stage of the molecular dynamics simulation is the equilibration phase. This is carried out to bring the system to equilibrium. During this stage various parameters are closely observed and once the parameters achieve the stable conformation, the production phase is started. It is usually desired to run a MD at a specific temperature. Thus the velocity of the system is carefully scaled so that the temperature gradually reaches the desired temperature. Thus the same applies when the pressure of the system needs to be adjusted so as to run the MD at a particular pressure. Once the equilibrium is obtained, the next stage is the production phase. During the production stage the simulation of the protein is carried for the desired time length. During the production phase thermodynamic parameters are calculated and stored; careful monitoring of these properties during the simulation can show if the simulation was consistent or whether or not the simulation need to be restarted. The positions were also stored in regular time intervals from which other properties can be calculated once the simulation is completed.

### 1.10.3 Measuring and analysing:

When a protein complex is set for MD, it is important to check the root mean square deviation of the back bone of the protein to the initial structure. Too much deviation means that the protein has undergone too much change in the amino acid residues orientation whereas little deviation suggests the protein requires longer MD. In an ideal condition following the MD study the RMSD value should stay around 2 (Oliviero, *et al.*, 2003). The other most commonly measured physical properties after a MD study are as follows:

The average potential energy  $V$  is required to verify energy conservation during the MD study. For the two body interacting.

$$V(t) = \sum_i \sum_{j>i} \phi(|r_i(t) - r_j(t)|) \quad \text{Equation 1.35}$$

The Kinetic energy can be calculated using the equation:

$$K(t) = \frac{1}{2} \sum_i m_i [v_i(t)]^2 \quad \text{Equation 1.36}$$

The total energy  $E$  is calculated as the sum of the potential energy and the kinetic energy in the system.

$$E = V + K \quad \text{Equation 1.37}$$

The temperature  $T$ , is directly related to kinetic energy (using the Equation 1.40 below) and thus can be obtained from the average kinetic energy  $K$ .

$$K = \frac{3}{2} Nk_B T \quad \text{Equation 1.38}$$

A calorie curve using all these information after a MD study is helpful in understanding the overall change in these parameters throughout the simulation.

### 1.11. Validation of models using Ramachandran Plot:

A Ramachandran plot is the graphical representation in which the dihedral angle of rotation about the alpha-carbon to carbonyl-carbon bond (*Figure 1.15*) in polypeptides is plotted against the dihedral angle of rotation about the alpha-carbon to nitrogen bond. It is a 2-d plot of the  $\phi$ -  $\psi$  torsion angles of the protein back bone. It shows the possible conformations of  $\phi$  and  $\psi$  angles for a polypeptide. Certain combinations of  $\phi$  and  $\psi$  angles are 'forbidden' because they result in steric clashes between atoms or angle strains. The remaining plot corresponds to allowed conformation. Hence Ramachandran plots draw particular attention towards the unrealistic conformations in models. A Ramachandran plot is of particular importance to study the protein structure after a long molecular dynamics study. The number of the unfavourable clashes due to the improper torsion can be identified by the number of residues in the unfavourable region on the Ramachandran plot.



Figure 1.15: The phi-angle ( $\phi$ ) and the psi-angle ( $\psi$ ) in a peptide unit. (Adapted from [wiki.cmbi.ru.nl/index.php/File:Phipsi.jpg](http://wiki.cmbi.ru.nl/index.php/File:Phipsi.jpg))

As shown in the above figure 1.15, in the polypeptide the main chain N-C $\alpha$  and C $\alpha$ -C bonds are relatively free to rotate. As mentioned these rotations are represented by the torsion angles phi ( $\phi$ ) and psi ( $\psi$ ), respectively. GN Ramachandran used computer models of small polypeptides to systematically vary and with the objective of finding stable conformations

(Ramachandran. G., *et al.*, 1963). The structures were analysed and the conformations were monitored for close contacts between the atoms and the van der Waals interactions were studied. Thus the angles which result in the steric hindrance or collisions correspond to disallowed conformation of the polypeptide backbone.

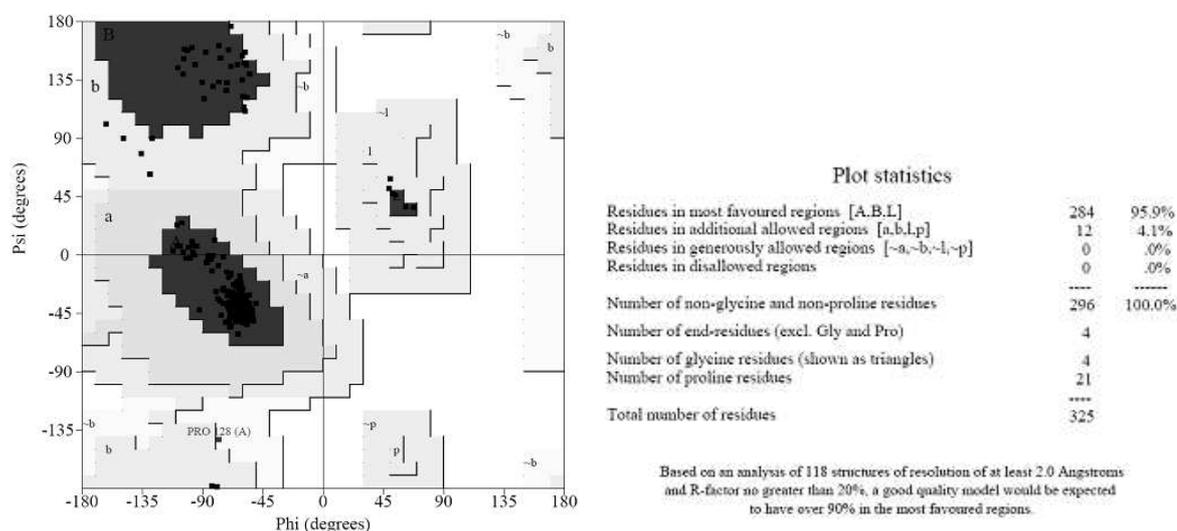


Figure 1.16: Example of a Ramachandran plot.

The diagram above is a representation of a Ramachandran plot where the white regions correspond to the residues with high steric hindrance and therefore are disallowed in a protein model. Disallowed regions generally involve steric hindrance between the side chain C methylene group and main chain atoms. Glycine residues lack a side chain and therefore can adopt phi and psi angles in all four quadrants of the Ramachandran plot. The dark region corresponds to the conformations of the residues with no steric clashes and is highly desired in a good quality model. The light grey areas shows the conformations of residues with little steric hindrance, these conformations are tolerable by the model.

Procheck is one of the standalone computer packages in addition to the other web based packages (PDBSUM and WEBMOL) to calculate the torsional angles in the form of Ramachandran plot (Laskowski., *et al.*, 2001). Unlike the other molecular dynamics program, Procheck checks the detailed stereochemical quality of the protein structure (Laskowski *et al.*, 1993). It utilises the coordinate information of the atoms from a single snapshot of the molecular structure in the form of PDB file (Bernstein *et al.*, 1977). The output or the results are available in the form of the pdf file readable in the Adobe system.

### 1.12. Docking studies:

Docking is a process by which the ligand or the inhibitor binds to the receptor site and they fit together appropriately in 3D space. Docking studies involves the posing, ranking and scoring. Posing involves the determination of the conformation and orientation of a ligand to fit appropriately into the active site. It is often a rough measurement of the ligand into the active site. Ranking enables more advanced study to estimate the binding free energy when the ligand binds to the active site. Both posing and ranking involve scoring. Scoring is a relative term and determined the placement of the ligand towards its ability to bind at the active site.

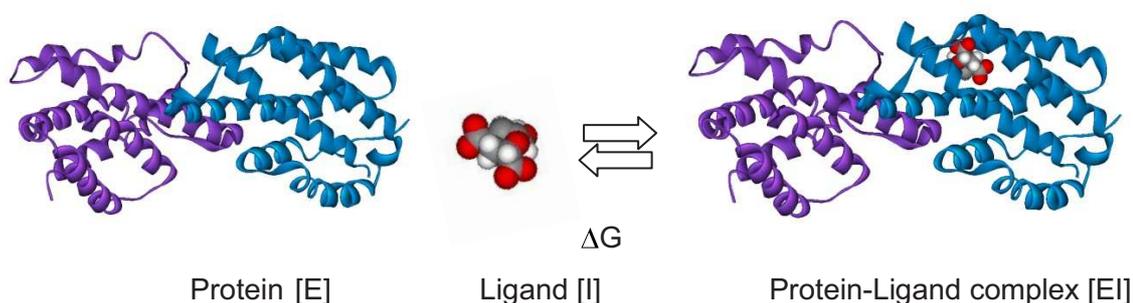


Figure 1.17: Showing the binding of the ligand to the receptor.

As shown in the above figure, binding of the ligand into the active site generates binding free energy  $\Delta G$ . The Gibbs free energy is then related to the binding constant. The  $\Delta G$  is calculated using the following equation:

$$\Delta G = -RT \ln K_A \quad \text{Equation 1.39}$$

$$K_A = \frac{[E][I]}{[EI]} \quad \text{Equation 1.40}$$

Prediction of the correct posing do not require the  $K_A$  information during the binding; this information is however important while scoring the ligands (Douglas. *et al.*, 2004).

#### 1.13.1. Docking

Docking is usually carried out to predict the binding mode of protein inhibitors. There are different types of algorithm available which are used to treat ligand flexibility and to some extent the protein flexibility. This is carried out using various search techniques. The ligand flexibility has been classified into three basic categories (Brooijmans, N, *et al.*, 2003) which are systematic search, random search and molecular dynamics search methods.

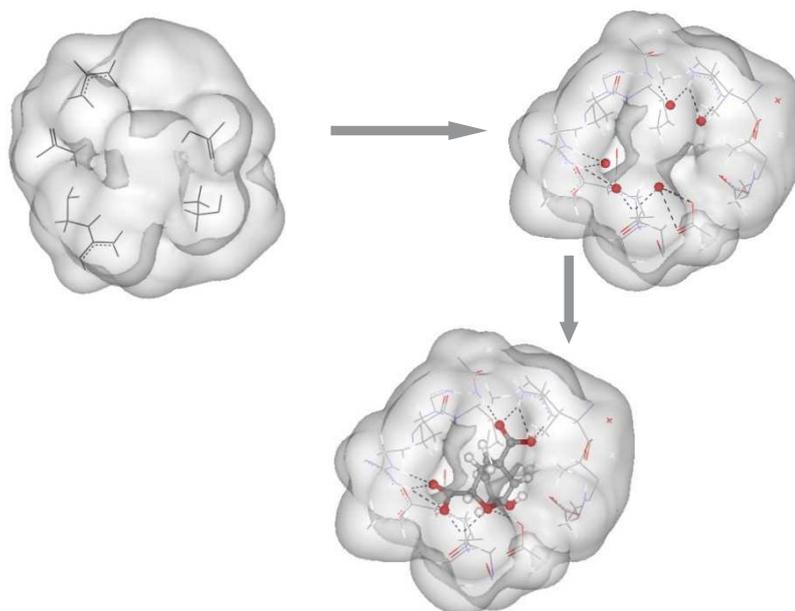


Figure 1.18: The steps involved in the docking process. The targeted interactions were first identified. This is then followed by identifying the best possible fit.

The systematic search method or the incremental method finds the best possible fit of the ligand by identifying a set of interactions at the receptor site. This is then followed by fragmenting the ligands and then placing them into the active site and finally linking these fragments. The software which utilise this algorithms are DOCK and FlexX

Molecular dynamics simulation technique can also be used to search for the best possible fit. However, these procedures may not be able to achieve the global minima and there often results the conformation in local minima (Brooijmans *et al.* 2003). Also molecular dynamics simulations involve longer time length and therefore often unable to cross high-energy barriers within feasible simulation time periods. Thus MD simulations are rarely used as stand-alone search techniques.

The random search algorithm operates by producing random changes. They have been classified into two widely used approaches i.e the Monte Carlo and the Genetic algorithm. In Monte Carlo algorithm, an initial set of configuration of ligands are created in the active site consisting of random conformation, translation and rotation. The initial configuration is then scored and a further new generation of conformation is generated and scored again. If the new solution scores better they are accepted immediately. If the solution is not a new minimum, Boltzmann based probability applies. If the solution passes the probability function test, it is accepted otherwise rejected.

Genetic algorithm (GA) is an optimization technique derived from the evolutionary theory (Holland, 1975). GA is widely used in finding the global minimum energy conformation of a molecule (Goldberg 1989). Today it is widely used in protein-ligand docking, molecular design, QSAR mapping and pharmacophore mapping (Cark. *et al.*, 1996; Jones 1998; Judson 1997). The first step involved in GA is to create a population of possible solutions. This corresponds to the set of randomly generated conformations of the molecule. The initial population is most easily obtained by randomly generating the conformation. The torsional angle values were then assigned and then the internal energy of the molecules was calculated by the molecular mechanics studies. The new population is thus generated and this became the starting point for the generation of the new population for the next cycle (Leach, A 2001). The software such as Gold, Autodock, Cache etc uses genetic algorithm to rank the ligands.

```

procedure GA
begin
  t=0;
  initialize P(t);
  evaluate structure in P(t);
  while termination condition not satisfied do
  begin
    t = t+1;
    select P(t) from P(t-1);
    recombine structures in P(t);
    evaluate structures in P(t);
  end
end.

```

Figure 1.19: The traditional genetic algorithm (Goldberg 1989) designed to generate the initial population  $P(t)$  through a random selection process. (Chromosomes represent the conformation of the molecules). The next algorithm cycles through rounds of selection, recombination and recombination until the terminal conditions are met. (The terminal condition represents the internal energy)

While considering the [EI] (Figure 1.17) the steric hindrance, electrostatic force, hydrogen bonding and the strain on the inhibitor or the enzyme are of particular importance. This is accomplished by maximizing favourable interactions and minimizing unfavourable interactions between the two molecules. It is important to note that both the structure of the receptor and the ligand changes during the binding process in a real biological environment. In this investigation we carried out the docking studies under flexible condition of the receptor and the ligand. Another important feature about the docking studies is the ligand placement algorithm to enumerate and test possible poses for the ligand in the protein's active site. In this study we use 'Project Leader' that follows *genetic algorithm* based approach.

The genetic algorithm represents the states of the degrees of freedom of the ligand as a string of digits, and this string is referred to as a gene. A population of different genes is generated at random, and each is scored using the energy function. Genes are selected to form the next population based on their score, with better scoring genes more likely to be selected. A gene may be selected more than once, and some may not be selected at all. Pairs of the selected genes are allowed to cross over with each other. In this process, a segment of the gene is selected and the values in this range are exchanged between them. Hence the values are expected to be different each time the studies are carried out under similar conditions.

### 1.13.2. Scoring Methods:

As discussed earlier that algorithm is required to search the conformation of the ligand within the active site. It is also important to rank these conformations and this is perhaps the most critical side of docking studies. Scoring methods are generally the reflexions between the correct and incorrect conformation and therefore designing a reliable scoring function is important. Although the free energy simulation method can be used to predict the protein-ligand interactions (Kollman. P., *et al.*, 199; Simonson, T., *et al.*, 2002), these techniques require expensive resources. The use of entropic effect makes this scoring function simpler and hence can be used in evaluation of large number of protein ligand complexes. The binding of the ligand to the protein to form the ligand-protein complex takes place in the aqueous environment. The transfer from the solute state with complete separation of both reactants to the complexed state involves either enthalpic and entropic effects (Bohm, *et al.*, 1996) that contribute to the Gibbs free energy of binding. Thus the free energy can be written as:

$$\Delta G = \Delta H - T\Delta S = -RTL\ln K \quad \text{Equation 1.41}$$

In the above equation  $\Delta G$  is the free energy left once the ligand binds to the protein. It is defined as the mutual compensation of the enthalpy and the entropy.  $\Delta H$  is the enthalpy and counts for the hydrogen bonds, polar interaction, van der Waals interaction, etc.  $\Delta S$  is the entropy and counts for loss of degree of freedom, gain of vibrational modes, and loss of solvation / protein structure at the binding site.

Currently three types of scoring functions are in use which are force fields based scoring, empirical based and knowledge based scoring.

The force field based scorings are carried out by the sum of the internal energies such as the energy from the bond, angle, torsional (dihedral) angle, out of plane (improper) torsional, etc and the energies due to ligand receptor interactions. These can be represented as shown below (Douglas. *et al.*, 2004).

$$E_{\text{pot}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{oop}} + E_{\text{electrostatics}} + E_{\text{vdW}} + E_{\text{constraint}} \quad \text{Equation 1.42}$$

The drawback of the method is that this method does not include solvation and entropic terms. Also a cut off distance is generally maintained for the treatment of non bonded interactions making the calculations even more complicated. The hydrogen bondings are often used in different ways depending of the geometry of interactions (Verdonk. *et al*, 2003).

The empirical scoring functions reproduce experimental data such as binding energies, conformation a sum of several parameterised functions. The binding energies can be approximated by a sum of individual uncorrelated terms derived from experimental data (X-ray structural information) (Douglas.*et al.* 2004). Although the complex empirical scoring functions can address solvation and desolvation effects, however the terms provide only incomplete descriptions of these effects on protein-ligand binding (Wang. R.,*et al.* 1998)

The knowledge based scoring or the PMF (Potential mean function) scoring of the protein ligand complexes are modelled using relatively simple interaction pair potential. The PMF scoring is the interaction potential of each atom pair in two molecules (eg. Ligand and protein) to approximate the free energy of each pairwise interactions as a function of their inter-atomic distance. This structural information is stored in databases of protein-ligand complexes to derive quick calculation of the atom pair interaction potentials (Muegge, I,*et al.*, 2000). The total energy for the protein-ligand complex is calculated by taking the sum of the PMF values of all interacting atoms. The major advantage of using the knowledge based scoring system is the computational simplicity and therefore it becomes possible to screen a large database of compounds (Douglas. B,*et al.*, 2004). Cache workstation (*Fujitsu*) utilise PMF based scoring which has been used in this study. One disadvantage is that it is essentially based on the explicitly coded information in a limited set of protein-ligand complexes.

## CHAPTER 2

### OPTIMIZATION OF AN INSILICO MODEL FOR MYCOBACTERIAL CHORISMATE MUTASE: THE COMPUTATIONAL APPROACH

#### 2.1 Aim:

The aim of this work is to design a protein model for the enzyme mycobacterial chorismate mutase (MtCM) that can be used for ligand screening.

#### 2.2 Background:

The crystal structure of the enzyme chorismate mutase (MtCM) from *Mycobacterium tuberculosis* (encoded by Rv1885c) was isolated and successfully determined at 1.6 Å (Okvist, Dey *et al.* 2006). Apart from the new topology of the dimeric structure of the enzyme, it has also revealed the entry of the ligand and its binding at the active site. This enzyme has followed simple Michaelis-Menten Kinetics ( $k_{cat}$  of 50 s<sup>-1</sup> and  $k_m$  of 180 μM) and provided a high quality model. This enzyme is involved in the pathogenesis of tuberculosis (Sasso, Ramakrishnan *et al.* 2005). In an assay to estimate the chorismate mutase activity of the enzyme from *M. tuberculosis*, over 15000 compounds were used and screen against the MtCM chorismate mutase by docking studies. Only 15 compounds have shown satisfactory results in the model and 4 compounds had good inhibitory response when tested against MtCM chorismate mutase (Agrawal, Kumar *et al.* 2007). This crystal structure of MtCM along with the known biologically active inhibitors of MtCM has opened an opportunity in designing inhibitor molecules against the deadly pathogen.

#### 2.3 Crystal structure of chorismate mutase from *Mycobacterium tuberculosis*:

##### 2.3.1 Background of chorismate mutase:



Table 2.1: The broad classification of chorismate mutase (Babu. *et al.*, 2004).

Chorismate mutase has been classified into three categories, the AroH, the prokaryotic AroQ class and the eukaryotic AroQ class. The tertiary structural folds of the AroQ class and the AroH class are completely unrelated, whereas the two types of AroQ can be structurally superimposed and the topology of the four-helix forming the active site is conserved (Lee, Stewart *et al.* 1995). Size exclusion chromatography revealed that the *Mycobacterium tuberculosis* chorismate mutase is a dimer (figure 2.2), and extensive biochemical and biophysical characterisation of the recombinant protein suggest that it is similar to the AroQ class of chorismate mutase (Qamra, Prakash *et al.* 2005). The annotated genome sequence of *M. tuberculosis* shows the presence of two reading frames Rv1885c and Rv0948c. Okvist and his co-worker in their investigation chose the secreted mycobacterial chorismate mutase encoded by frame Rv1885c in strain H37Rv and determined its crystal structure both in the unliganded form and in complex with a transition state analogue (Okvist, Dey *et al.* 2006). The PDB entries for these crystallographic models are 2FP1 and 2FP2.

### 2.3.2 Structure of Mycobacterium Chorismate Mutase:

The structure of the mycobacterium chorismate mutase has been revealed from the X-ray crystal structure determined at 1.6 Å (Okvist, Dey *et al.* 2006).

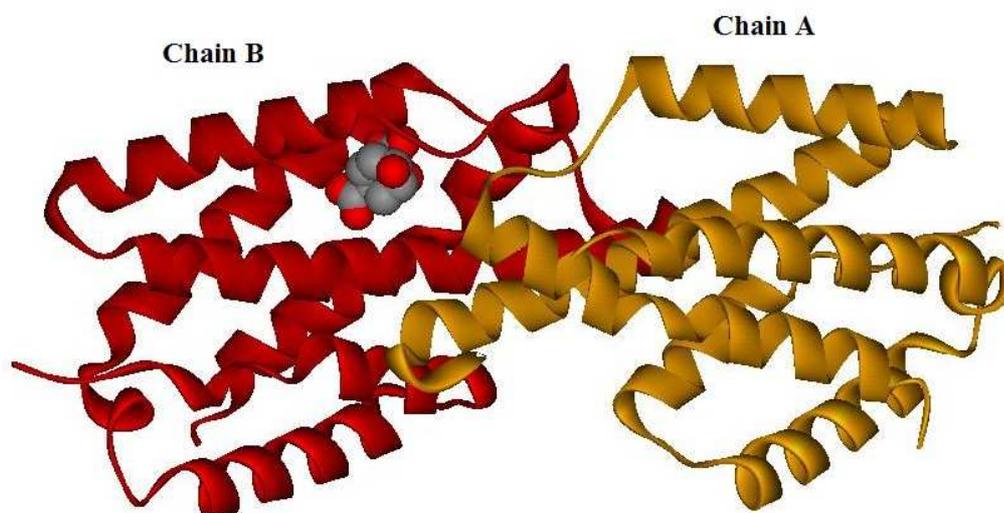


Figure 2.1: The X-ray crystal structure of the Chorismate mutase (2FP2.pdb) determined at 1.6 Å.

The MtCM is a dimeric structure (*Figure 2.1*) and this is similar to the *E. coli* chorismate mutase (EcCM) and *Saccharomyces cerevisiae* chorismate mutase (ScCM) and the active sites are located at the interface of the two protomers with the formation of an allosteric binding site. The interesting feature about the mycobacterium chorismate mutase is that the active site is located only in the B-chain of the dimer chain and there is no such allosteric regulation associated from the other chain (Okvist, Dey *et al.* 2006).

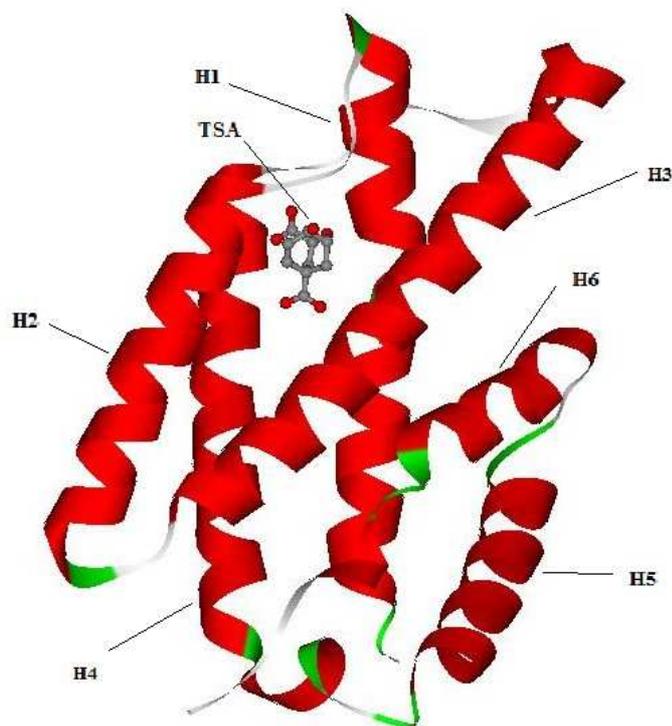


Figure 2.2: Showing the ligand bound chain B. H1, H2, H3, H4, H5 and H6 represent the helices of the chain (2FP2.pdb).

The chain consists of 6  $\alpha$ -helices which are arranged in groups of three. There is a disulphide bond between Cys160 of helix H5 and Cys193 of helix H6 thereby forming excellent hydrophobic packing in this part of the structure. Overall the structure of MtCM resembles the EcCM; the only difference is the H1, which is much longer in EcCM and the H3 which is one turn longer in MtCM (Okvist, Dey *et al.* 2006).

#### 2.4 Active site analysis:

The active site of the enzyme was identified by comparing the chorismate mutase from yeast and *E.coli* (Xue, Y *et al.* 1995). The active site of the mycobacterial chorismate mutase (2FP2.pdb) is very similar to the active sites of the other class of the AroQ chorismate mutase particularly to EcCM and ScCM (Okvist, Dey *et al.* 2006).



## 2.5 The Ligands:

A number of fundamental strategies were developed to design the best inhibitor for chorismate mutase; the principal strategy relies on enhancing the binding strength of the inhibitor at the binding site (Wolfenden, R 1976). All the previous approaches were based on the assumption that the isomerisation of chorismate to prephenate takes place through a transitional state or high energy intermediate resembling a bicyclic species and this is the structure that should be mimicked (Bartlett *et al.*, 1988).

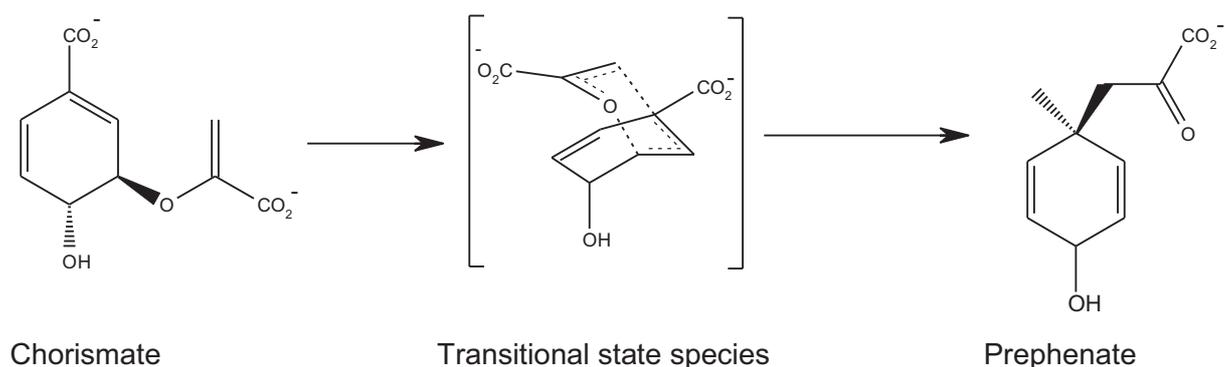


Figure 2.4: The conversion of prephenate from chorismate through the transitional state.

A number of the bicyclic diacids (Bartlett *et al.*, 1988), and a series of aza inhibitors (Mark, E.*et al.*, 2004) has been designed and screened for the chorismate activity. None of these compounds have shown sufficient inhibitory effect when compared to the endo oxabicyclic dicarboxylic acid (**TSA**).

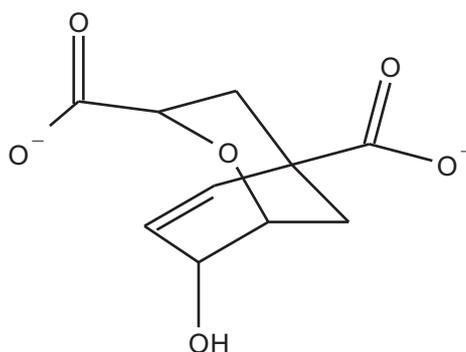


Figure 2.5: The structure of the Transitional state analogue (TSA).

**TSA** has by far been the best mimic of the geometry of the transition state of the chorismate mutase catalysed reaction and is currently the best inhibitor for the broad range of chorismate mutases (Mandal and Hilvert 2003) including the *Mycobacterium tuberculosis* chorismate mutase (Sasso, Ramakrishnan *et al.* 2005). Therefore the work in this thesis has used **TSA** as an active ligand throughout the investigation. All the other ligands used in the current investigation were following the biological evaluation of the ligands carried out by Himanshu *et. al* (Agrawal, Kumar *et al.* 2007). Out of all the fifteen compounds tested by these authors for the chorismate mutase inhibitory activity against the *Mycobacterium tuberculosis*, four compounds have shown biological activity (Table 2.2).



Table 2.2: Published structures of biologically active ligands against MtCM (Agrawal, Kumar *et al.* 2007)

The remaining eleven compounds (Table 2.3) evaluated for biological activities against MtCM were found to be inactive. Unfortunately there was not enough information available regarding seven of these eleven compounds particularly with regard to the configuration of their chiral centre and therefore they have been omitted from the current study.



Table 2.3: List of all the inactive ligands following the biological evaluation by Himanshu *et al.* (Agrawal, Kumar *et al.* 2007).

One active Ligand, Ligand **3** also has three stereo centres and the configurations of the asymmetric centres have not been provided. Therefore in the current study to prepare a computer model of the X-ray crystal structure (2FP2.pdb) seven compounds have been used (Table 2.4).

SL No	Structures	$K_i$ (nM)
Ligand <b>1</b>		5.7±1.2
Ligand <b>2</b>		17.71±3.35
Ligand <b>4</b>		28.8±4.1
Ligand <b>7</b>		Inactive
Ligand <b>8</b>		Inactive
Ligand <b>9</b>		Inactive
Ligand <b>13</b>		Inactive

Table 2. 4: The table above shows the list of ligands used in the current study. Ligand 1 to Ligand 4 denotes the active ligands whereas Ligand 7 to Ligand 13 denotes the inactive ligands.

## 2.6 Computational approach:

The computational approach was carried out by preparing the initial protein structure, preparing the ligands, setting up and running molecular dynamics, analyzing the protein models and docking studies.

### 2.6.1 Preparing the initial protein structure:

In the current investigation the X-ray crystal structure, 2FP2.pdb, of the Mycobacterial chorismate mutase (Okvist *et al.*, 2006) was downloaded from the Protein Data Bank through the World Wide Web (www.pdb.org). The initial structure was prepared by removing all the extra information and the water molecules (HOH). This results in the preparation of a pdb file with only Cartesian coordinate's information of the residues from chain A, Chain B and the ligand (**TSA**).

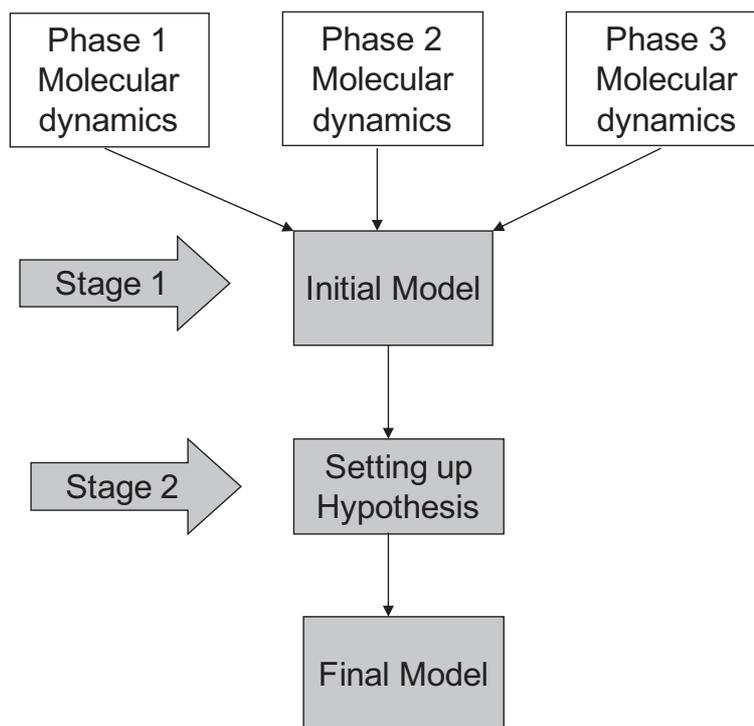
There were also two disulfide bonds between the cystine residues (CYS) 160 and 193 in both Chain A and B. The cystine residues CYS were therefore manipulated to CYX; this was to facilitate Amber program to identify the disulfide bonds in this residues.

```
SSBOND 1 CYS A 160 CYS A 193          1555 1555 2.05
SSBOND 2 CYS B 160 CYS B 193          1555 1555 2.04
```

The glutamic acid (GLU 109) in the active site was protonated (Okvist *et al.*, 2006), this was not identified during the diffraction and therefore it was converted to GLH, to be identified as protonated by Amber. The next step in the preparatory step involves the addition of the missing residues from both the chain A and B as shown below.

```
REMARK 465          M RES C SSSEQI
REMARK 465          ASP  A    34
REMARK 465          GLY  A    35
REMARK 465          PRO  A   125
REMARK 465          GLU  A   126
REMARK 465          ASP  B    34
REMARK 465          GLY  B    35
REMARK 465          THR  B    36
```

This process was carried out in *Fujitsu* Cache. Once the missing residues were added, the protein with the ligand was re-sequenced and saved as pdb file.



Scheme 2.1: Schematic representation of the model design.

#### 2.6.2 Preparing the initial Ligand structures:

All the ligands from table 2.4 were prepared in *Fujitsu* CAChe workstation by drawing the structure and beautifying the ligands. These ligands were then set for conformation search using MM2 parameters and the geometry was then optimized using MOPAC PM3 parameters. The final structures were then saved as pdb files.

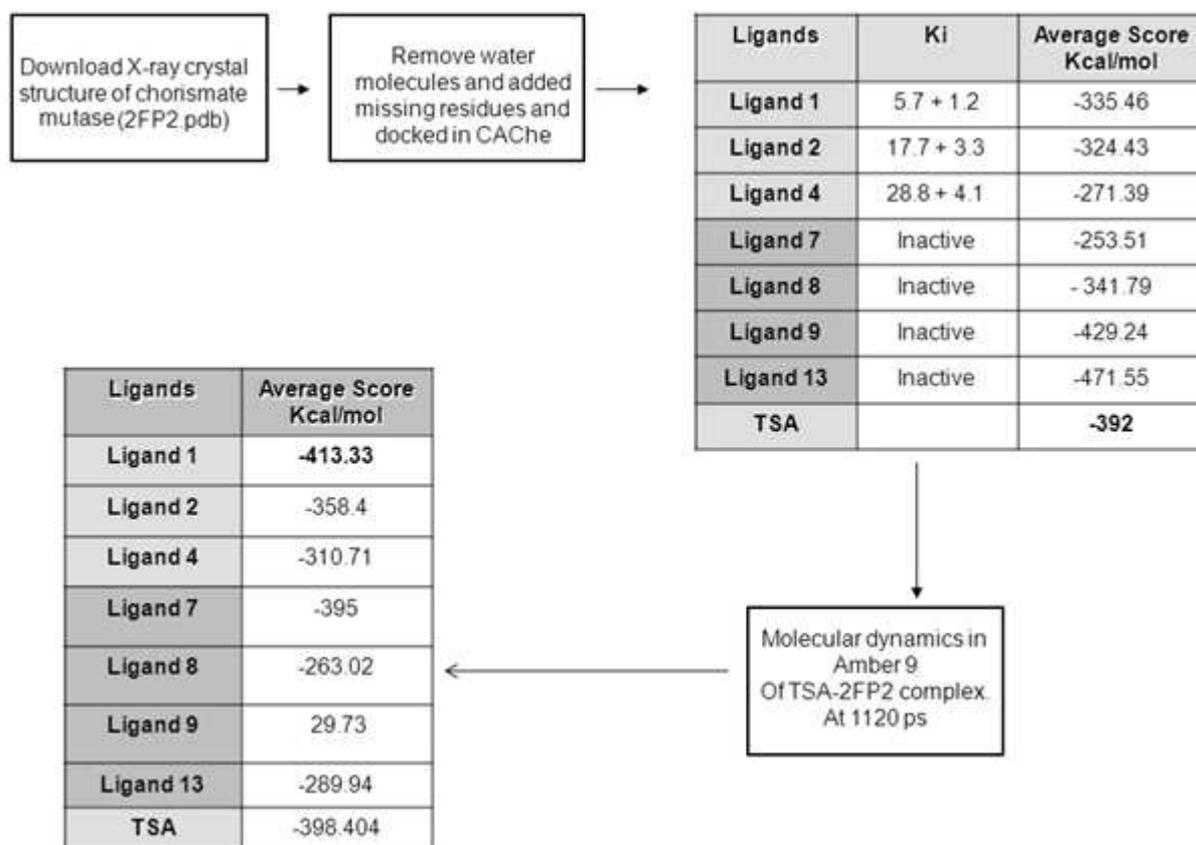
#### 2.6.3 Generating the model:

The model generation step involves various phases of study as shown in Scheme 2.1. A brief overview of the entire stages is also outlined in Appendix I. In this step the active site was first identified in the protein-ligand complex and this was carried out by identifying the neighbouring residues within the 5 Å regions. The ligand (**TSA**) was then removed and the protein was used for docking studies with a series of ligands (Table 2.4) with known biological activity.

#### 2.6.4 Phase I Study:

In the Phase I study as shown in Scheme 2.2, docking studies were performed at least three times with the ligands from table 2.4 and the average scores were considered. The active

site was further refined by molecular dynamics study carried out at 1120ps on the **TSA**-protein complex in Amber 9.



Scheme 2.2: Schematic representation of the Phase I study. The docking studies were performed three times and the average docking scores were calculated.

The pdb file (2FP2**TSA**.pdb) was prepared from the Cache file (2FP2**TSA**.csf). The ligand (**TSA**.pdb) was copied from the pdb file and the Antechamber program was used to identify the **TSA**.pdb as a ligand using the following command.

```
$AMBERHOME/exe/antechamber -i TSA.pdb -fi pdb -o TSA.prepin -fo prepi -c bcc -s 2 -nc -2
$AMBERHOME/exe/parmchk -i TSA.prepin -f prepi -o TSA.frmod
```

The Xleap was then loaded, this creates a new system or working station for the molecular dynamics to run in Amber 9 using the ff99 force field.

```
$AMBERHOME/exe/xleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.ff99
```

Before loading the protein-ligand complex for the molecular dynamics process, it was important to parameterise the ligand and this was carried out by loading the TSA.prepin and the TSA.frcmod file using the following commands.

```
source leaprc.gaff
loadamberprep TSA.prepin
loadamberparams TSA.frcmod
```

Finally the protein-ligand complex was loaded using the following command.

```
B=loadpdb 2FP2TSA.pdb
```

Where B was the name of the complex. Once everything was loaded, the system was then checked with the command 'check B' to see if the parameters were correctly loaded. The sodium ions and water molecules were then added and the system was then saved using the following commands.

```
addions B Na+ 0
solvateoct B TIP3PBOX 8.0
Saveamberparm B wat.prmtop wat.inpcrd
Output file: wat.prmtop and wat.inpcrd
```

The water molecules were explicitly used in the system by generating a box of water molecules within the 8.0 Å region of the protein (TIP3PBOX 8.0). Once the protein and ligands were prepared in the preparatory steps, the next step involves the molecular dynamics simulation studies on the protein-ligand complex.

The simulation method was performed in three steps. The initial minimisation method was carried out using the following parameters.

```
$AMBERHOME/exe/sander -O -i min.in -o wat_min.out -p wat.prmtop -c wat.inpcrd -r
wat_min.crd &
```

```
Minimisation of complex (imin)
&cntrl
imin=1, maxcyc=2000, ncyc=1000,
cut=16, ntb=1, igb=0, ntr=0
/
END
END
```

```
Input file: imin.in, wat.prmtop, wat.inpcrd
Output file: wat_min.out, wat_min.crd
```

This performed (imin=1) 2000 steps (maxcyc), using a nonbonded cut off at 16 Å (cut), a generalised Born solvent model (igb=0), no periodic boundaries (ntb=1) and no position restrain (ntr=0).

The next step was the molecular dynamics (MD1) study in constant volume which was carried out by gradually increasing the temperature to 300K over 20ps. This was carried out using the following command.

```
nohup mpirun -np 2 $AMBERHOME/exe/sander.MPI -O -i md1.in -o wat_md1.out -p
wat.prmtop -c wat_min.crd -r wat_md1.restrt -x wat_md1.mdcrd &
```

MD heating 0 to 300K over 20 ps at CONSTANT VOLUME, no shake

```
&cntrl
  imin = 0, (minimisation turned off)
  irest = 0, (we are not restarting out MD)
  ntx = 1, (first stage of MD),
  ntb = 1, (constant volume periodic boundaries),
  cut = 16, (use cutoff at 16 Å)
  ntr = 0, (no positional restraints),
  ntc = 1, (no shake),
  ntf = 1, (no shake),
  igb = 0 (generalised Born solvent model)
  tempi = 0.0, (start temp),
  temp0 = 300.0, (end temp),
  ntt = 3,
  gamma_ln = 1.0, (langevin dynamic should be used to control temp),
  nstlim = 20000, dt = 0.001
  ntpc = 500, ntwx = 500, ntwr = 1000 (restart job after 1000 step)
/
END
END
```

```
Input file      :      md1.in, wat.prmtop,
Output file     :      wat_md1.out, wat_min.crd, wat_md1.restrt, wat_md1.mdcrd
```

Once the molecular dynamics trajectory was achieved at 300K and the system reached equilibrium, the next molecular dynamics simulation (MD2) was carried out at constant pressure for initial 100ps and thereafter the step was repeated until the system achieved the desired conformations. The following commands were used for MD2 simulation.

```
nohup $AMBERHOME/exe/sander -O -i md2.in -o wat_md2.out -p wat.prmtop -c
wat_md1.restrt -r wat_md2.restrt -x wat_md2.mdcrd &
```

```
MD run const pressure (md2.in)
&cntrl
  imin = 0,
  irest = 1,
  ntx = 7,
  ntb = 2,
  cut = 16,
  ntr = 0,
  ntc = 1,
  ntf = 1,
  igb = 0
  ntp = 1
  tempi = 300.0,
  temp0 = 300.0,
  ntt = 3,
  gamma_ln = 1.0,
  nstlim = 100000, dt = 0.001
  ntp = 1000, ntwx = 1000, ntwr = 1000
/
END
END
```

```
Input file:   md2.in, wat.prmtop, md1.restrt
Output file:  wat_md2.out, wat_md2.restrt, wat_md2.mdcrd
```

The MD2 was carried out for initial 100ps and then for another 1000ps. Finally the protein-ligand simulated complex at 1120ps was minimised and the output file (wat\_md2\_min.crd) was converted to the wat\_md2\_min.pdb.pdb file using the following commands

Minimisation of the structure

```
nohup mpirun -np 2 $AMBERHOME/exe/sander.MPI -O -i min2.in -o
wat_md2_min.out -p wat.prmtop -c wat_md2.restrt -r wat_md2_min.crd &
```

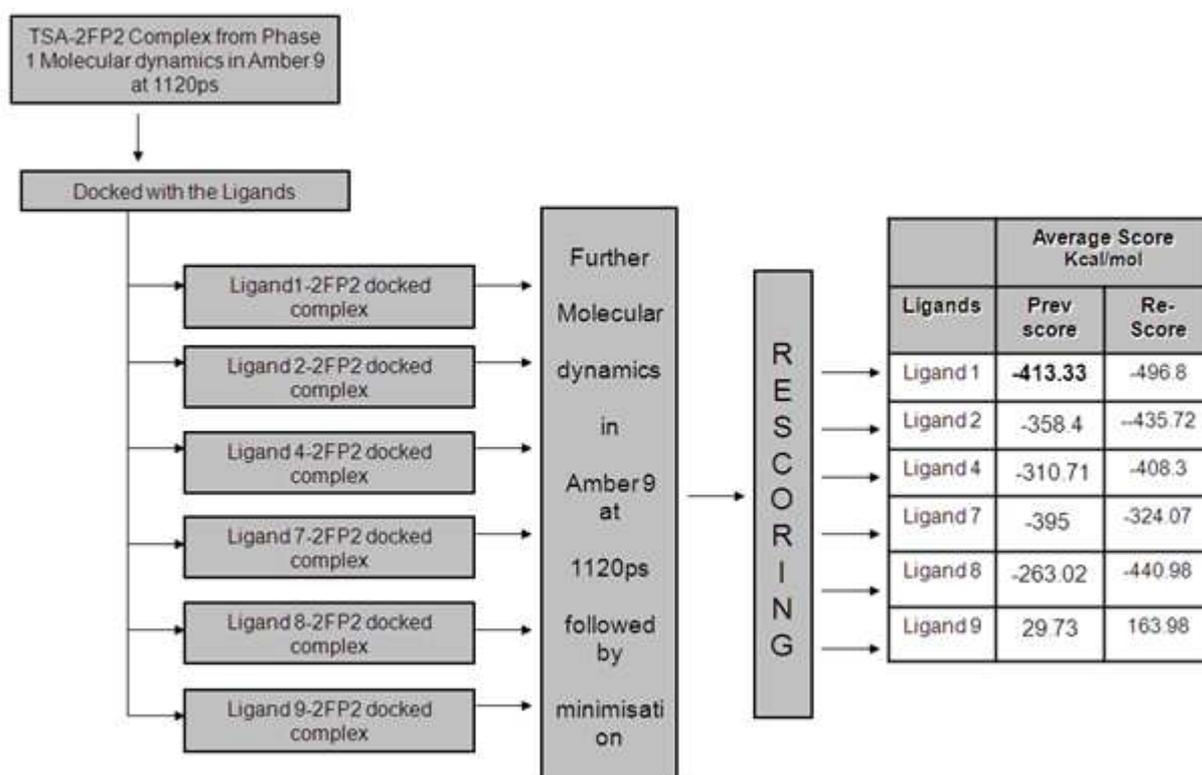
convert the output file (wat\_md2\_min.crd) into pdb.

```
$AMBERHOME/exe/ambpdb -p wat.prmtop < wat_md2_min.crd > wat_md2_min.pdb
```

The final pdb structure was then tailored by removing the water molecules and then visualised in Cache workstation. The active site was once again identified and the protein was used for docking studies (Scheme 2.2).

### 2.6.5 Phase II study:

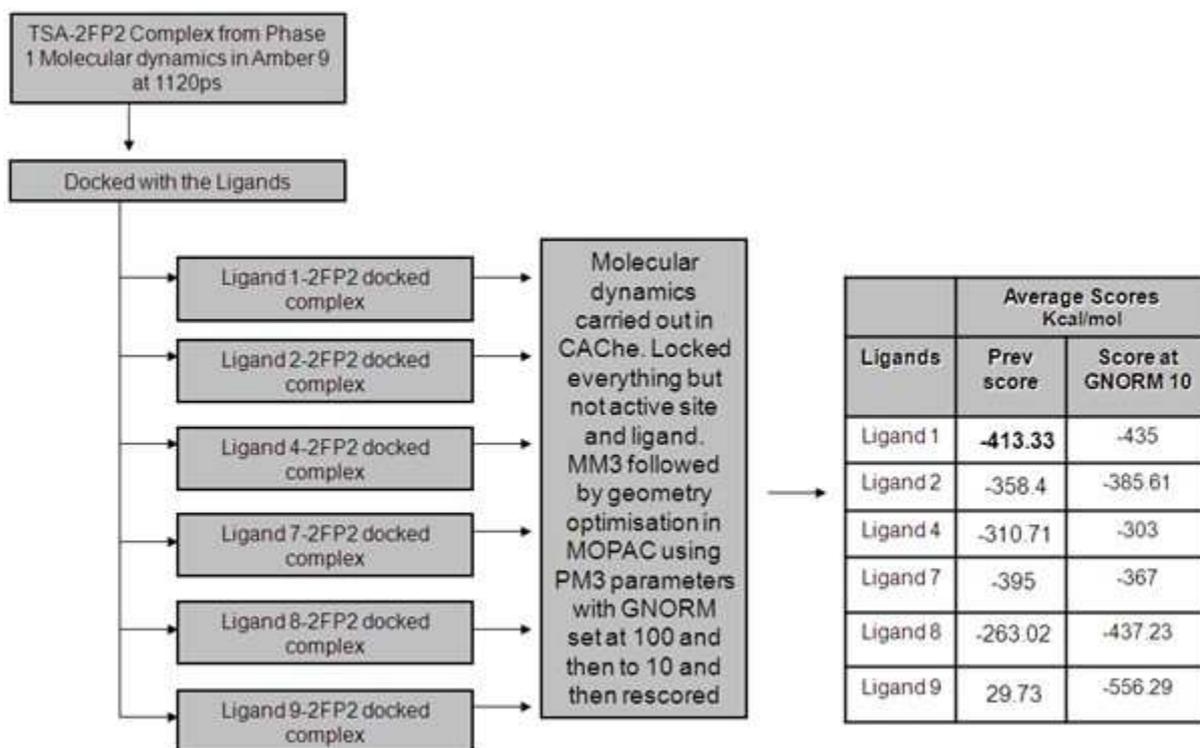
The schematic representation of the Phase II is given in Scheme 2.3. The TSA-ligand complex model following the Phase I molecular dynamics study at 1220 ps in *Amber9* was set for docking with the inactive and active inhibitors in Phase II. The docked complexes were then set for further molecular dynamics in Amber 9 at 1220 ps. These complexes were then rescored in CAChe.



Scheme 2.3: Schematic representation of the Phase II study

### 2.6.6 Phase III study:

The schematic representation of the Phase III has been illustrated in the Scheme 2.4. Phase III study was also performed as an extension from Phase I study. The ligand docked complexes were set for geometry optimization studies in '*CAChe Workstation Pro*'. The geometry of the docked complexes was optimised in MOPAC using PM3 parameters. To obtain reasonable geometry optimization results in reasonable time, the optimization was controlled by gradually decreasing the gradient norm (GNORM). The GNORM was first set to 100 (very crude optimization), and then to 10 (precise optimization) and 1 until a very precise optimization was achieved. Unfortunately GNORM 1 was not possible to obtain and therefore all the complexes were optimized at GNORM 10. These complexes were then re-scored.



Scheme 2. 4: The schematic representation of the Phase III study.

The  $\Delta H_f$  (heat of formation) of the complexes and the independent enzyme and receptor were calculated and  $\Delta H_f$  of the binding was calculated using the following equation (2.1).

$$\Delta H_f(\text{binding}) = \Delta H_f(\text{complex}) - \Delta H_f(\text{enzyme}) - \Delta H_f(\text{ligand}). \quad \text{Equation 2.43}$$

Ligands	Score GNORM 10 Kcal/mol	$\Delta H_f$ (complex) Kcal/mol	$\Delta H_f$ (ligand) Kcal/mol	$\Delta H_f$ (enzyme) Kcal/mol	$\Delta H_f$ (binding) Kcal/mol
Ligand 1	-435	-23344.70	-225.38	-22909.79	-209.53
Ligand 2	-385.61	-23185.8	-109.25	-22909.79	-166.76
Ligand 4	-303	-23131.2	-69.16	-22909.79	-152.25
Ligand 7	-367	-23104.5	-48.08	-22909.79	-146.63
Ligand 8	-437.23	-23151.9	-91.867	-22909.79	-150.243
Ligand 9	-556.29	-23120.6	28.25	-22909.79	-239.06

Table 2.5: Showing the calculation of the binding free energy when the ligand binds to the protein. The more negative is the energy, the more stable is the predicted binding.

## 2.7 Analyzing the models:

The protein models following the molecular dynamics simulations were analyzed by plotting the changes in the temperature, pressure, kinetic energy, etc. This information was obtained by extracting the data from the trajectory 'out' files.

The protein models were validated using a Ramachandran plot. This was carried out using the windows based software called Procheck (Ramachandran *et al.*, 1963). The pdb files were used directly into the software to get the output files readable in Adobe Reader.

### 2.7.1 Validation of models:

To understand whether the models had withstood the rigorous condition during the molecular dynamics studies, validations of the models are essential. Validations of these models were carried out using Ramachandran plot (Ramachandran, G.R 1963). Essentially the errors that can be introduced in a protein model during their development and refinement may vary tremendously and these produce a significant change in the final model. (Kleywegt, G. *et al.* 1995 and Brändén, C. *et al.* 1990). Therefore it was important to study the topology of the final protein structure with the help of Ramachandran plots before proceeding for further investigation.

### 2.7.2 Ramachandran plot:

Ramachandran plot is a two dimensional plot of the  $\phi$ - $\psi$  torsion angles of the polypeptides (Ramachandran, G.R 1963). The shading region on the plot represents the various  $\phi$ - $\psi$  torsional combinations (Morris *et al.* 1992) the dark region corresponds to the 'Core' region and represents the most favourable combinations. In an ideal situation and in a good protein model these regions should cover 90% of the residues. The proteins which were subjected to molecular dynamics studies were probed with Ramachandran plot with the help of the software 'Procheck' (Ramachandran, G.R 1963). The Figure 2.6 below shows the Ramachandran plot obtained at various intervals of the targeted model. The Figure 2.7 shows the Ramchandran plot of the various Ligand-protein complexes following a MD study as shown in Scheme 2.3.

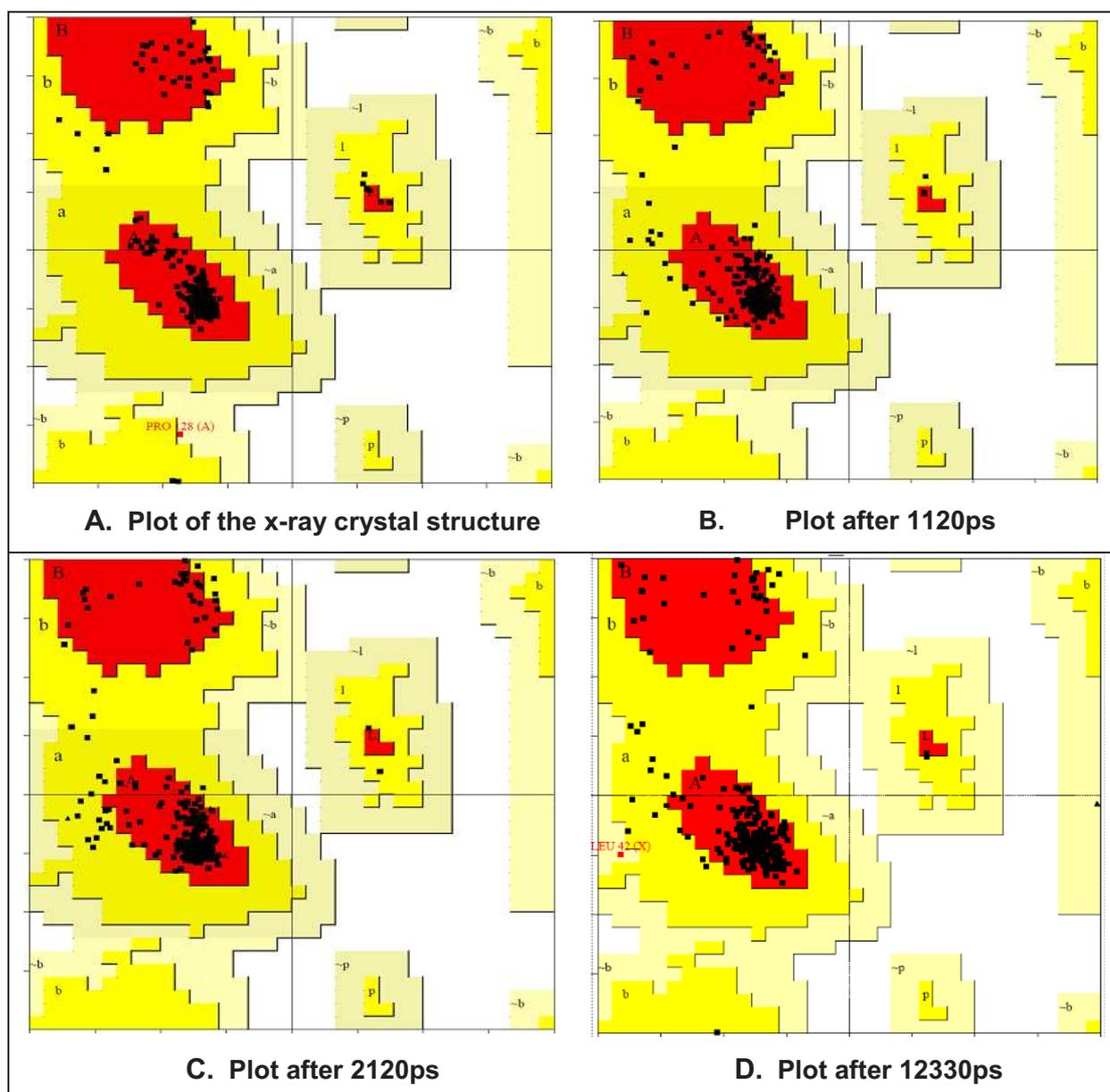


Figure 2.6: Ramachandran plots of the **TSA-Protein** complexes, where A is the plot from X-ray crystal structure, B is the plot following a MD study at 1120ps, C is the plot after a MD study at 2120ps and D is the plot after 12330ps.

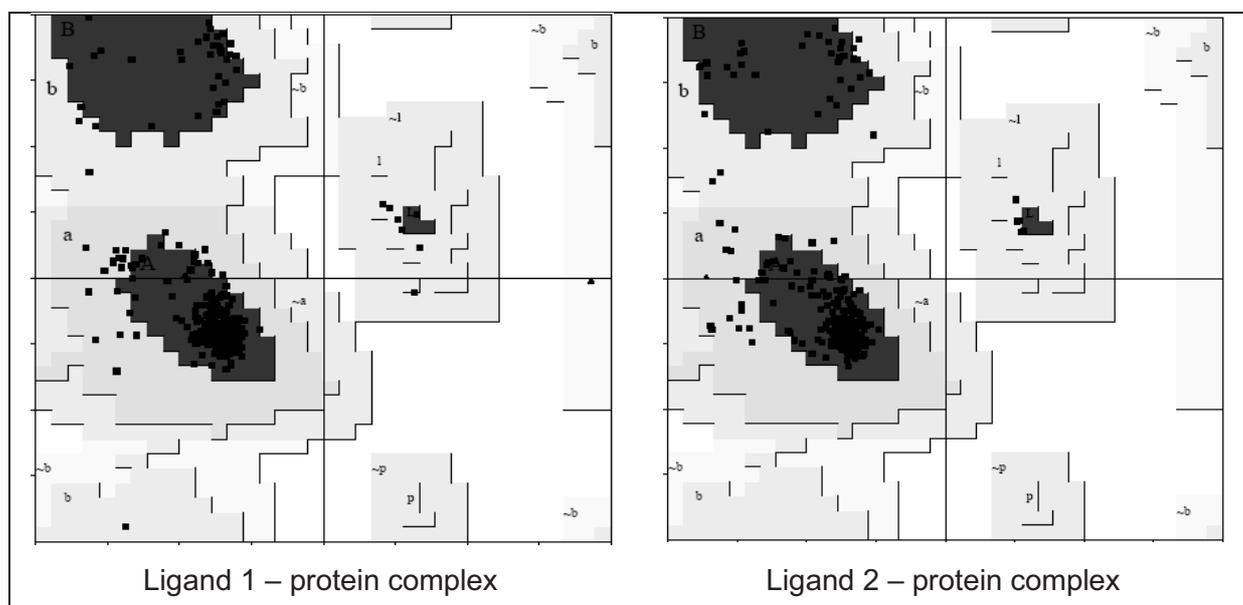
While it is important to know the percentage of the residues in the favoured region, it is also important to analyse the number of residues in the dis-allowed region in the Ramachandran plot.

Time interval in (ps)	Resdues in most allowed region (%)	Residues in additional allowed region. (%)	Residues in generously allowed region (%)	Residues in disallowed region. (%)
0	95.9	4.1	0	0
1120	93.2	6.8	0	0
2120	90.9	9.1	0	0
12330	91.6	8.0	0.4	0

Table 2.6: Showing the distribution of the % of residues in Ramachandran plot of **TSA-Protein** complex.

At various time intervals during the molecular dynamics study the Ramachandran plots were carried out to identify any unwanted interactions within the model. In an attempt to carry out the dynamics study beyond 2120 ps, the Ramachandran plot has shown some residues in the unwanted regions.

The protein-ligand complexes obtained following the MD studies from Phase 2 studies were also validated using the Ramachandran plot and each of these plots were shown in the table below.



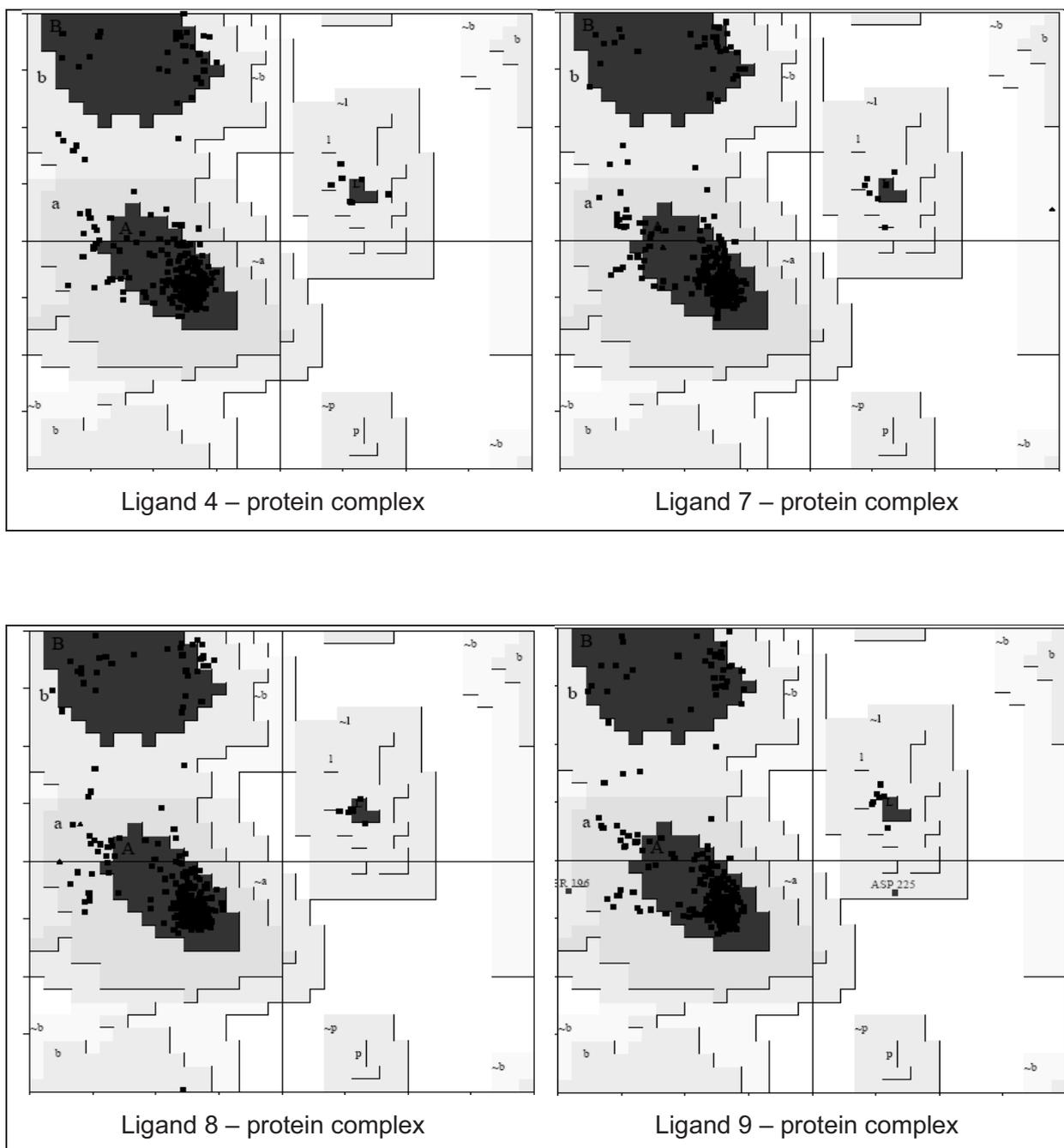


Figure 2.7: Showing the Ramachandran plot of the ligand protein complex following the MD from Phase 2 studies.

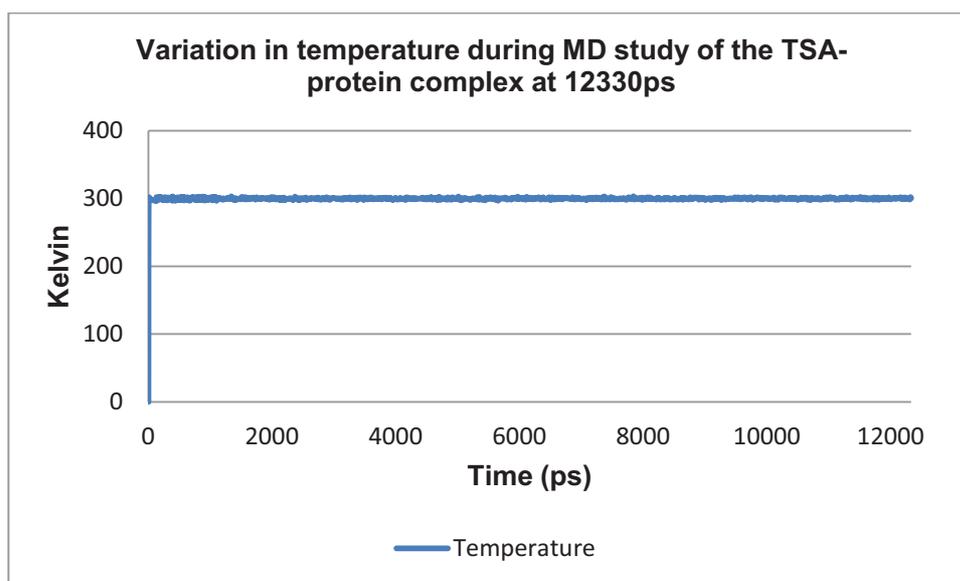
MD of Protein complexes at 2240 ps	Resdues in most allowed region (%)	Residues in additional allowed region. (%)	Residues in generously allowed region (%)	Residues in disallowed region. (%)
Ligand1-Protein	87.8	12.2	0	0
Ligand2-Protein	91.6	8.4	0	0

MD of Protein complexes at 2240 ps	Resdues in most allowed region (%)	Residues in additional allowed region. (%)	Residues in generously allowed region (%)	Residues in disallowed region. (%)
Ligand4-Protein	86.5	13.5	0	0
Ligand7-Protein	85.8	14.2	0	0
Ligand8-Protein	88.5	11.5	0	0
Ligand9-Protein	86.1	13.2	0.7	0

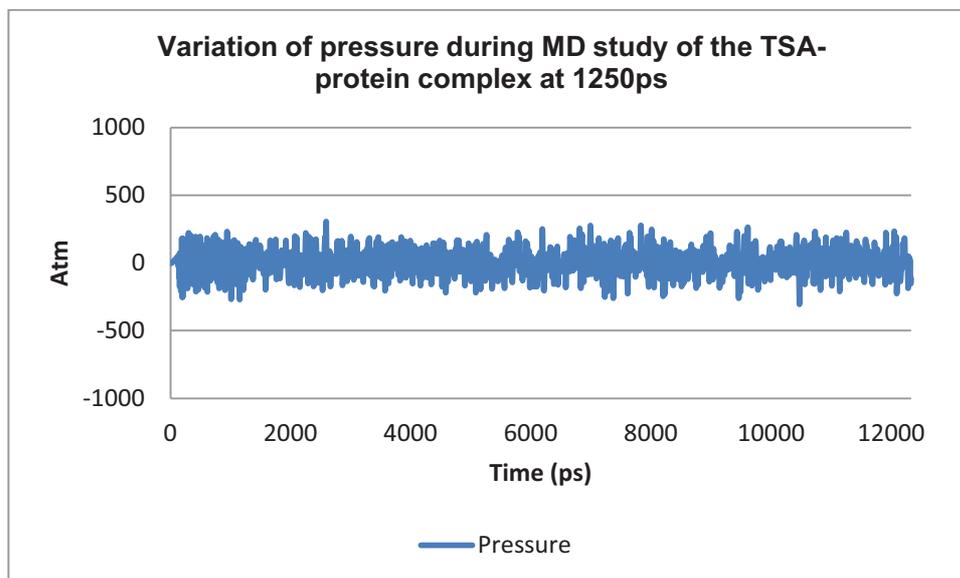
Table 2.7: Showing the distribution of the residues in Ramachandran plot of the protein ligand complexes obtained in the Phase 2 studies at 2240ps.

### 2.7.3 Analysis of the parameters.

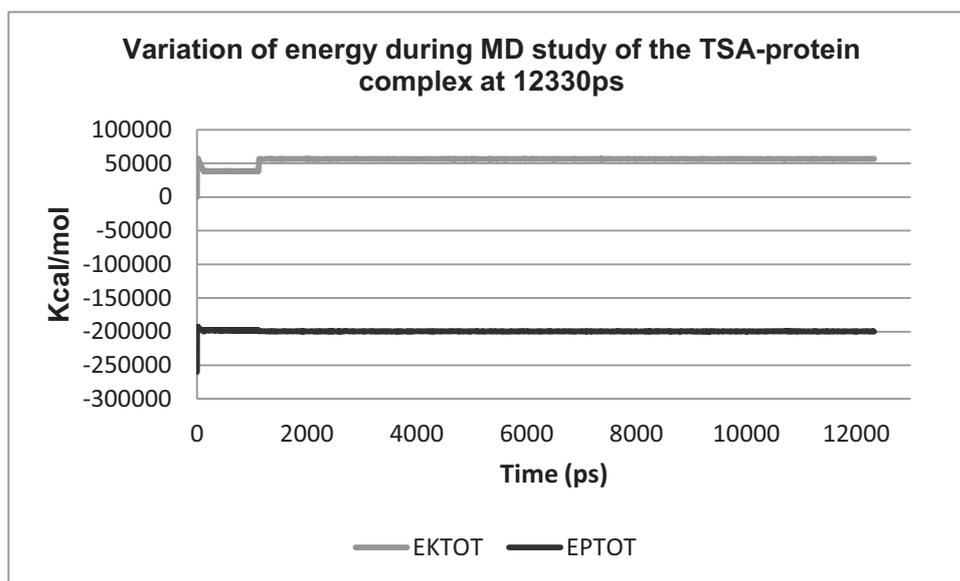
While the Ramachandran plot provides information on a particular conformational snapshot following the molecular dynamics studies, it is also important to get a clear picture of the protein model during the entire period of the molecular dynamics studies within the specified conditions; the changes in the model in presence of the various parameters need to be studied. The Graphs 2.1 to 2.4 below shows that the temperature, pressure, total kinetic energy (EKTOT) and total potential energy (EPTOT) from 0 to 2120 ps of the study maintained consistency and there was no unusual behaviour of the residues during the study.



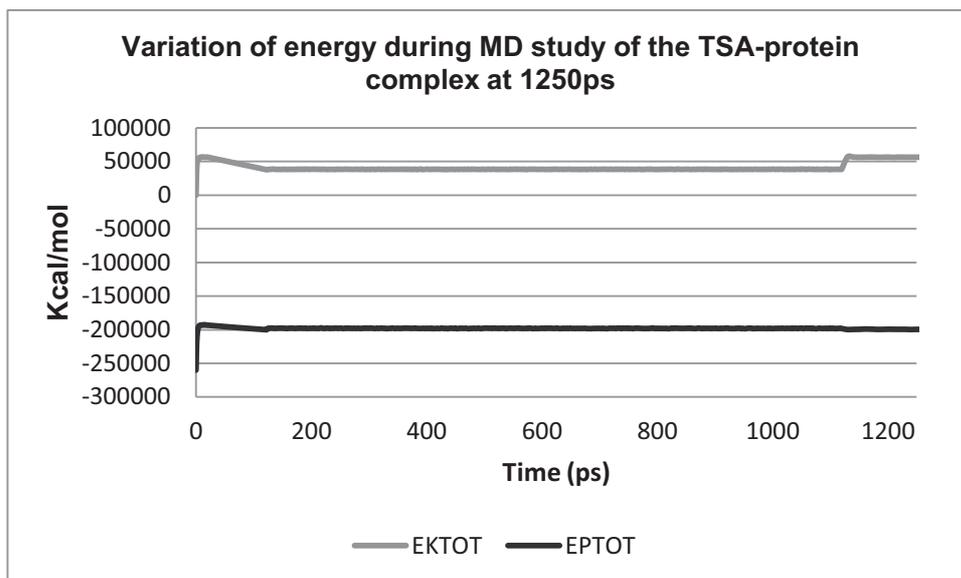
Graph 2.1: Variation in temperature through the MD simulation performed on the TSA-protein complex over 123300ps.



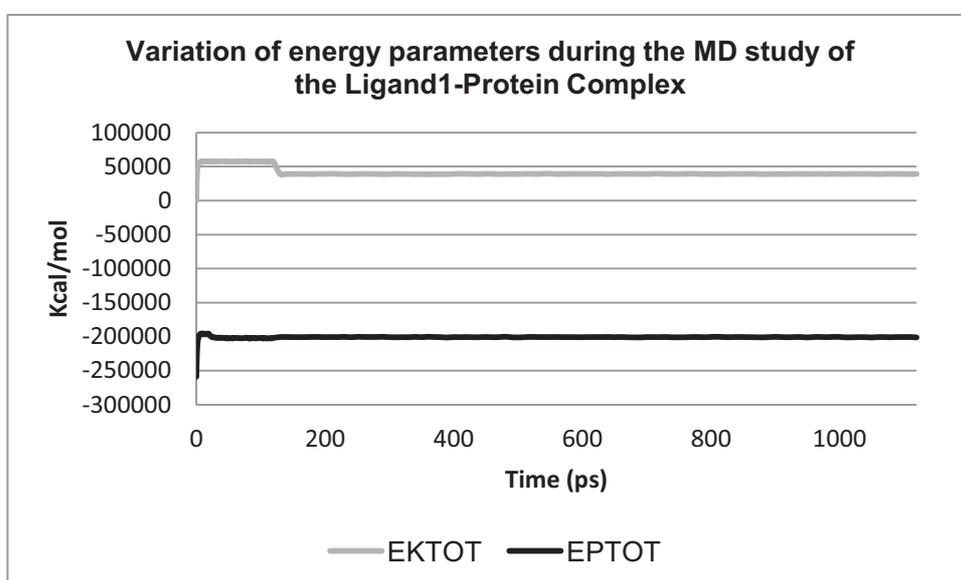
Graph 2.2: Variation in pressure through the MD simulation performed on the **TSA**-protein complex over 123300ps



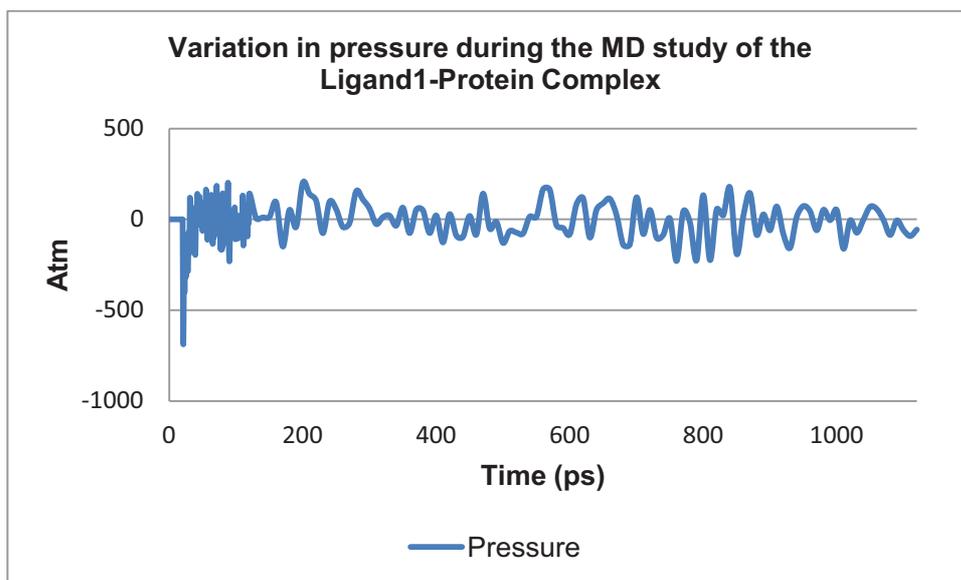
Graph 2.3: Variation in total potential energy (EPTOT) and total kinetic energy (EKTOT) through the MD simulation performed on the **TSA**-protein complex over 123300ps.



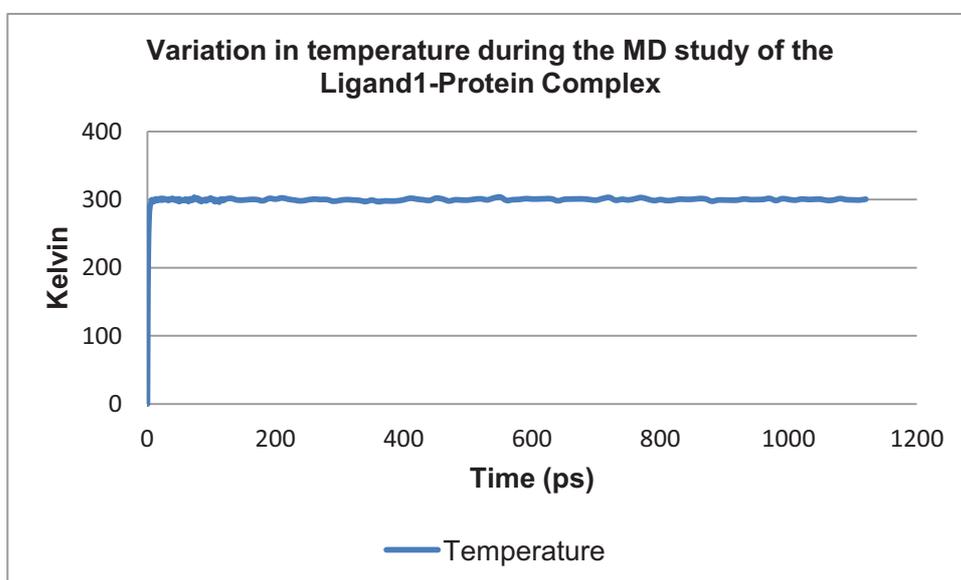
Graph 2. 4: Variation in total potential energy (EPTOT) and total kinetic energy (EKTOT) through the MD simulation performed on the **TSA**-protein complex over 1200ps.



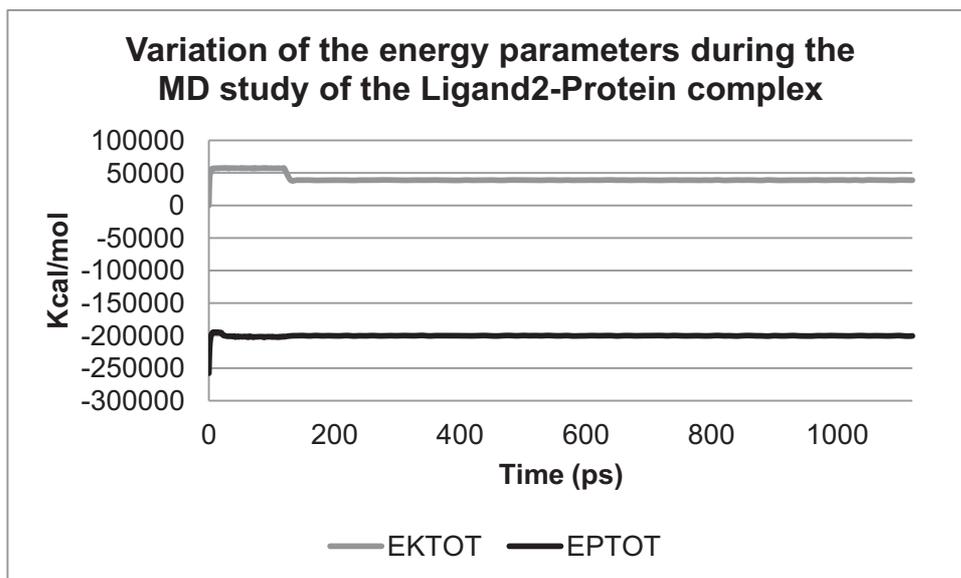
Graph 2.5: Variation in EPTOT and EKTOT through the MD simulation performed on the ligand1-protein complex over 1120ps.



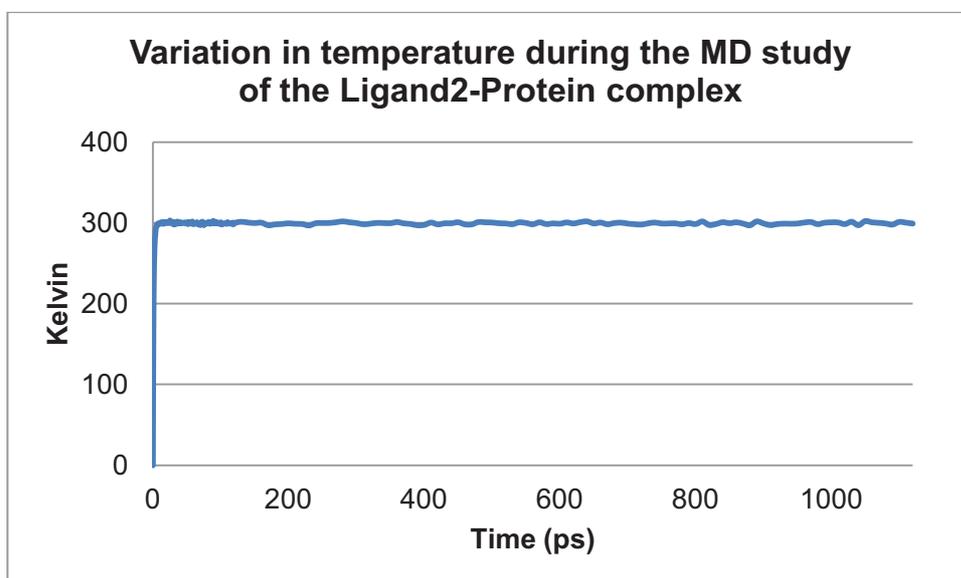
Graph 2.6: Variation in pressure through the MD simulation performed on the Ligand1-protein complex over 1120ps.



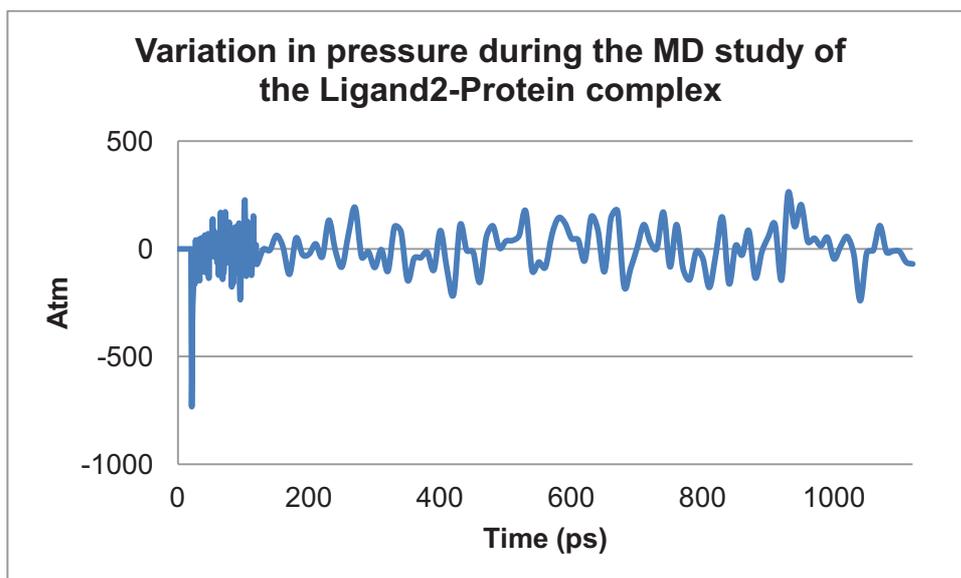
Graph 2.7: Variation in temperature through the MD simulation performed on the **Ligand1**-protein complex over 1120ps.



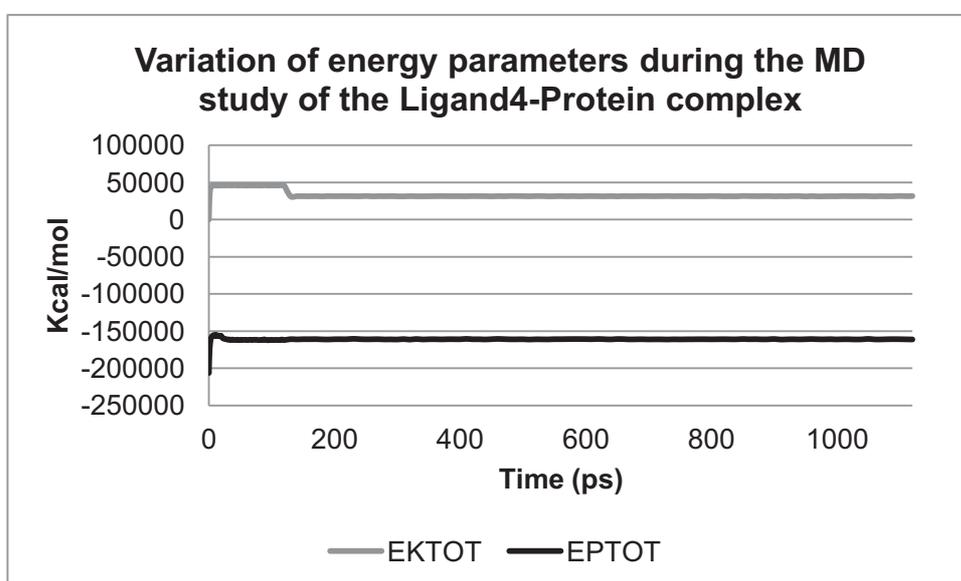
Graph 2.8: Variation in EPTOT and EKTOT through the MD simulation performed on the Ligand2-protein complex at 1120ps.



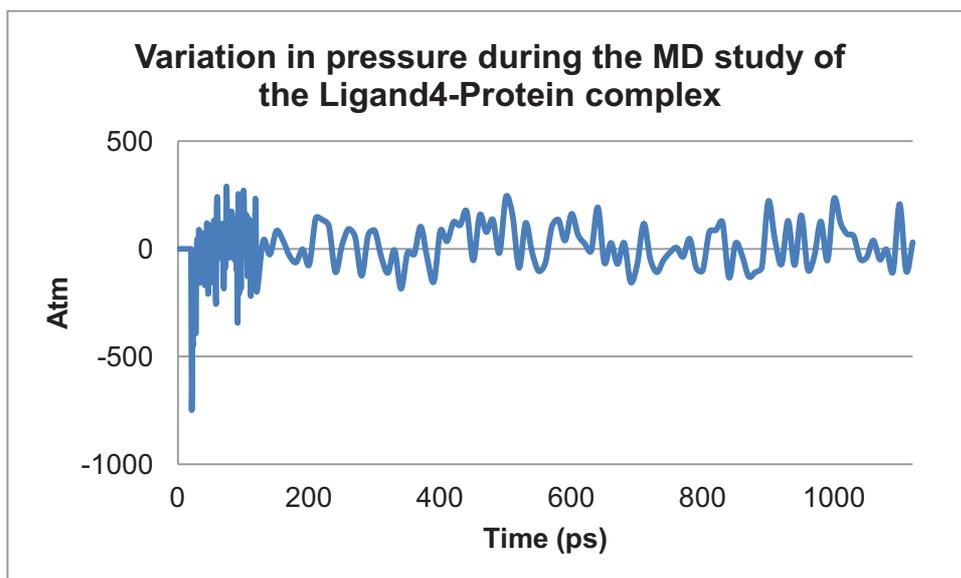
Graph 2.9: Variation in temperature through the MD simulation performed on the Ligand 2 - protein complex over 1120ps.



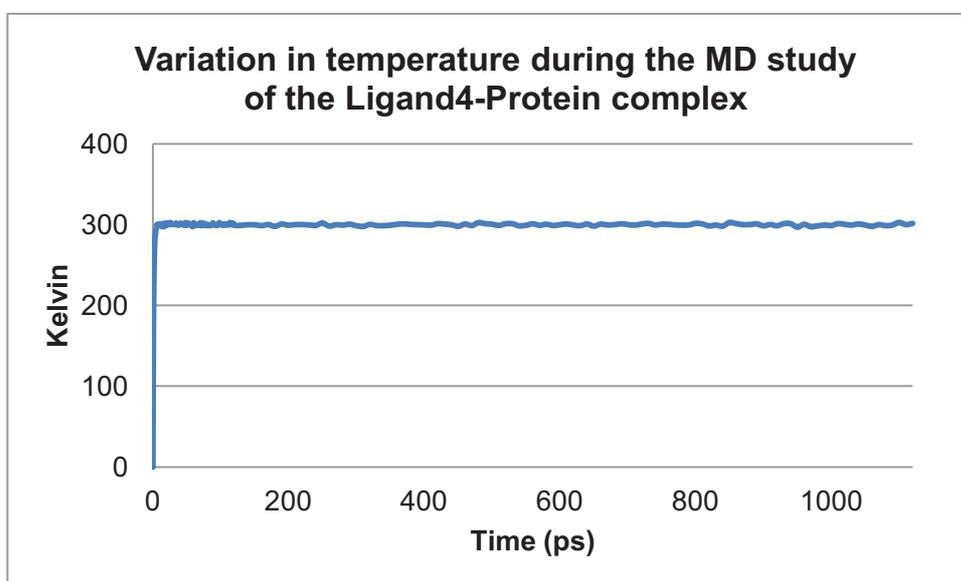
Graph 2.10: Variation in pressure through the MD simulation performed on the Ligand2-protein complex over 1120ps.



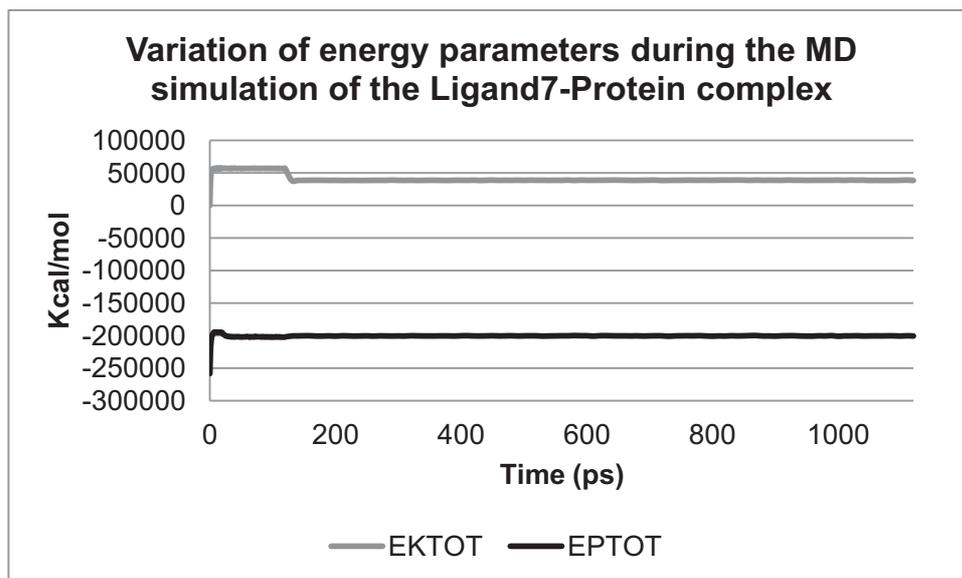
Graph 2.11: Variation in EPTOT and EKTOT through the MD simulation performed on the Ligand4-protein complex at 1120ps.



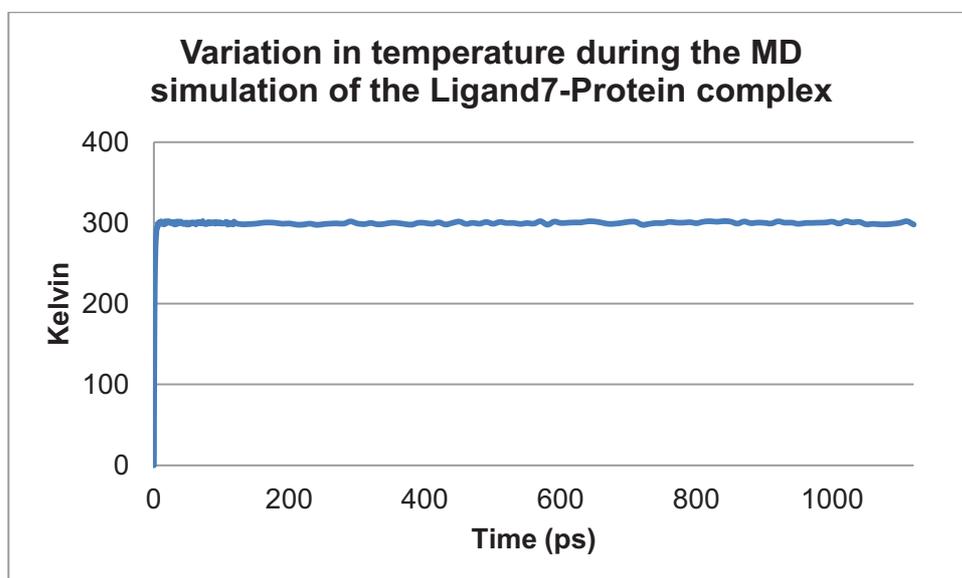
Graph 2.12: Variation in pressure through the MD simulation performed on the Ligand 4-protein complex over 1120ps.



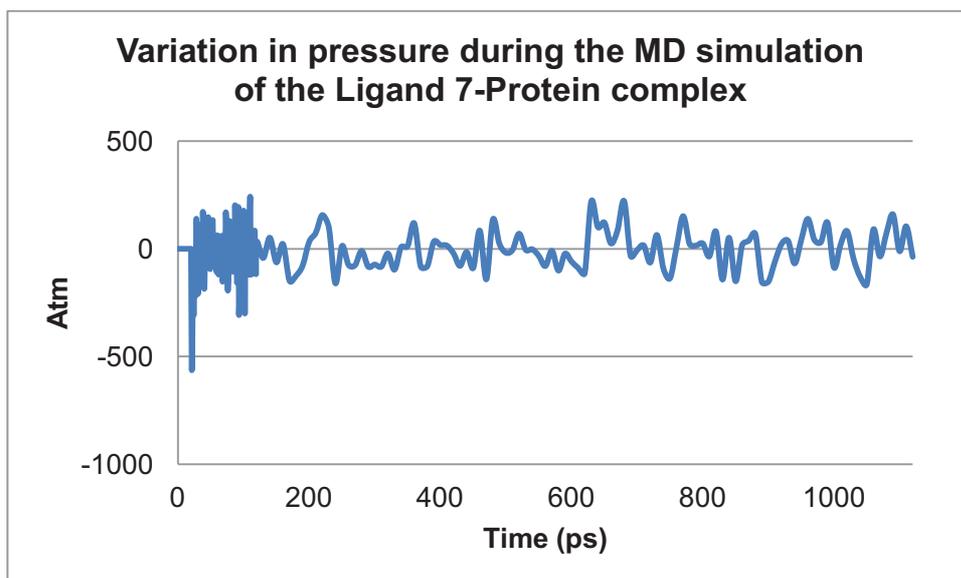
Graph 2.13: Variation in temperature through the MD simulation performed on the Ligand 4-protein complex over 1120ps.



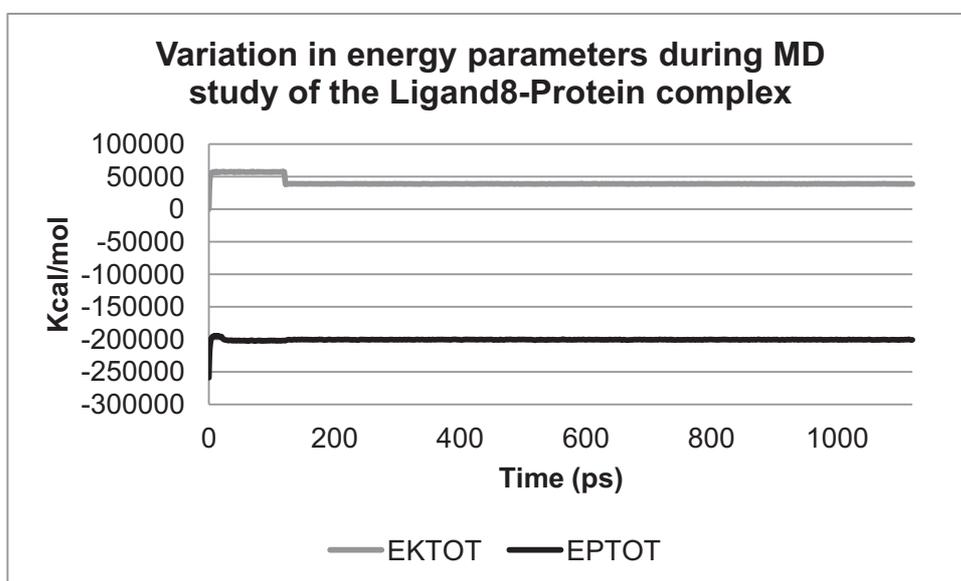
Graph 2.14: Variation in EPTOT and EKTOT through the MD simulation performed on the on the Ligand7-protein complex at 1120ps.



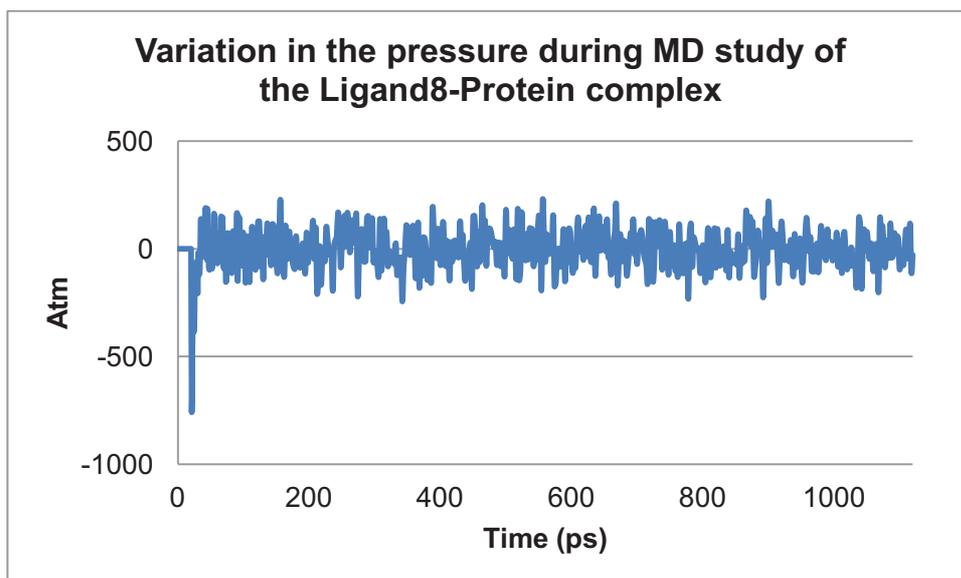
Graph 2.15: Variation in temperature through the MD simulation performed on the Ligand 7 - protein complex over 1120ps.



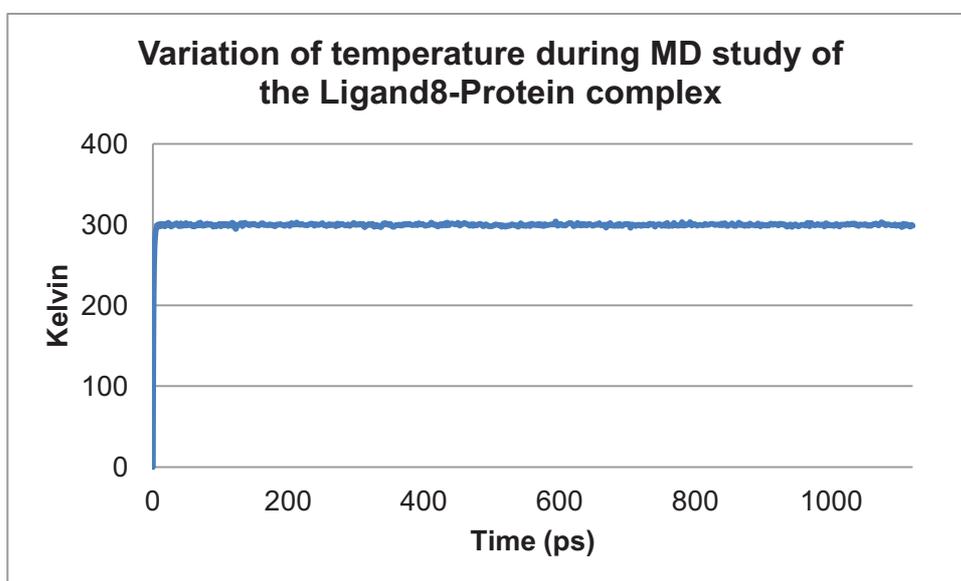
Graph 2.16: Variation in pressure through the MD simulation performed on the Ligand 7-protein complex over 1120ps.



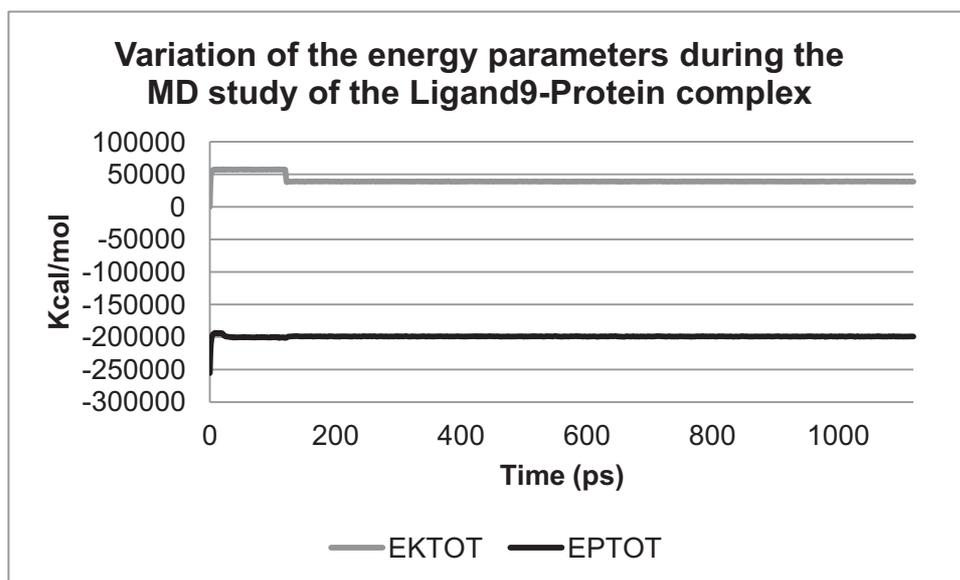
Graph 2.17: Variation in EPTOT and EKTOT through the MD simulation performed on the on the Ligand 8-protein complex at 1120ps.



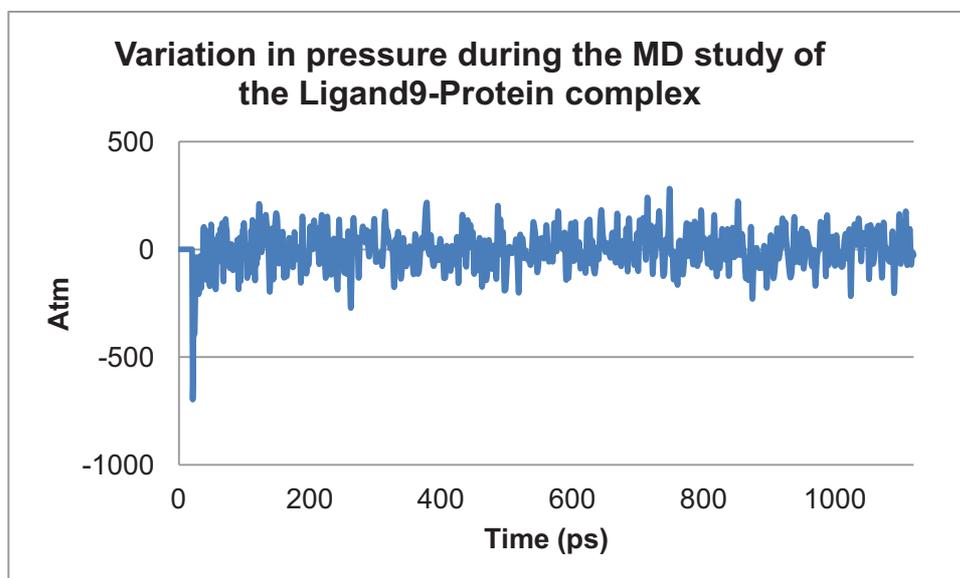
Graph 2.18: Variation in pressure through the MD simulation performed on the Ligand **8**-protein complex over 1120ps.



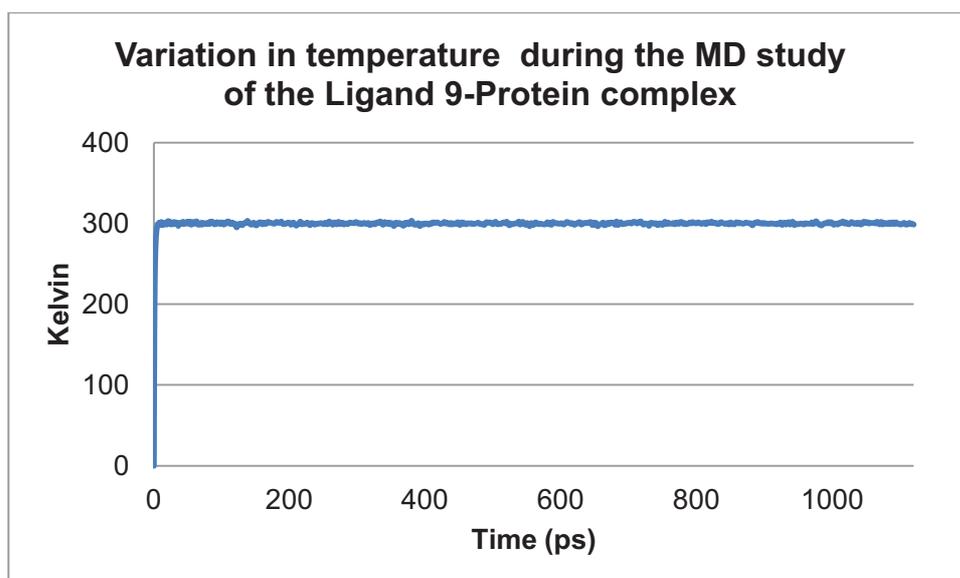
Graph 2.19: Variation in temperature through the MD simulation performed on the Ligand **8**-protein complex over 1120ps.



Graph 2.20: Variation in EPTOT and EKTOT through the MD simulation performed on the on the Ligand 9-protein complex at 1120ps.



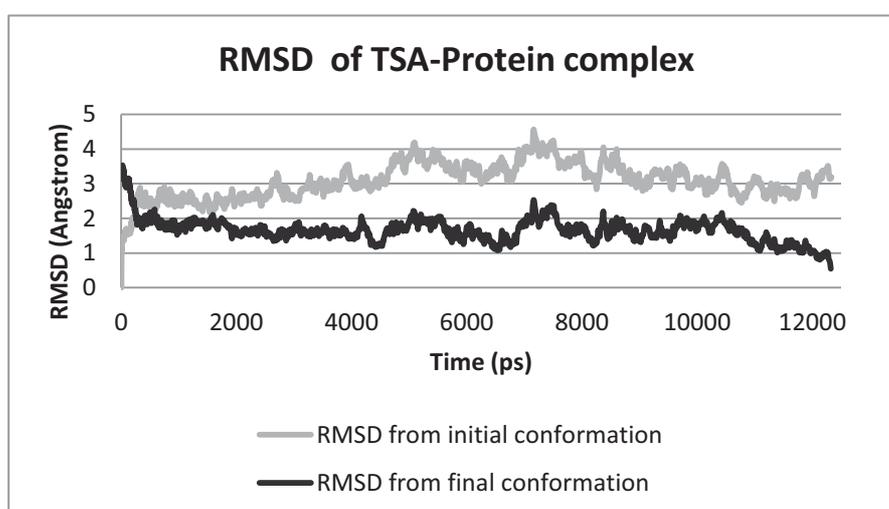
Graph 2.21: Variation in pressure through the MD simulation performed on the Ligand 9-protein complex over 1120ps.



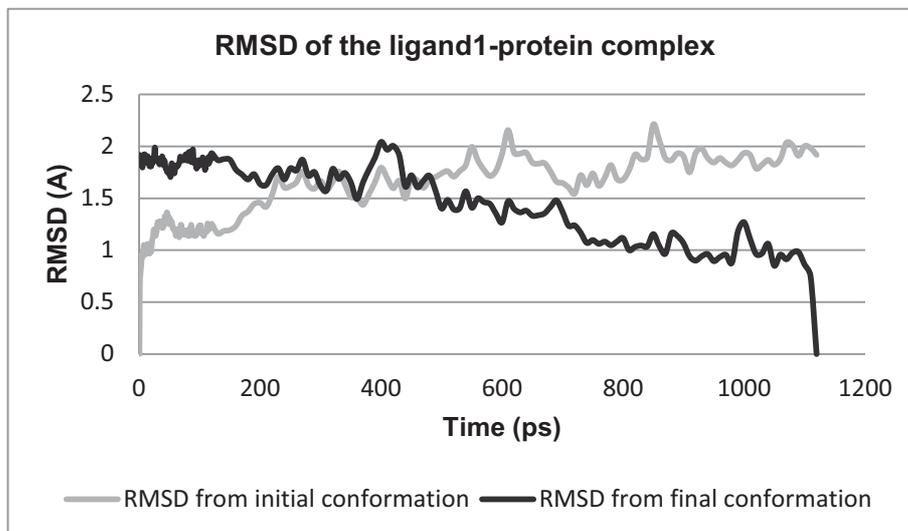
Graph 2.22: Variation in temperature through the MD simulation performed on the Ligand 9 - protein complex over 1120ps.

#### 2.7.4 RMSD analysis:

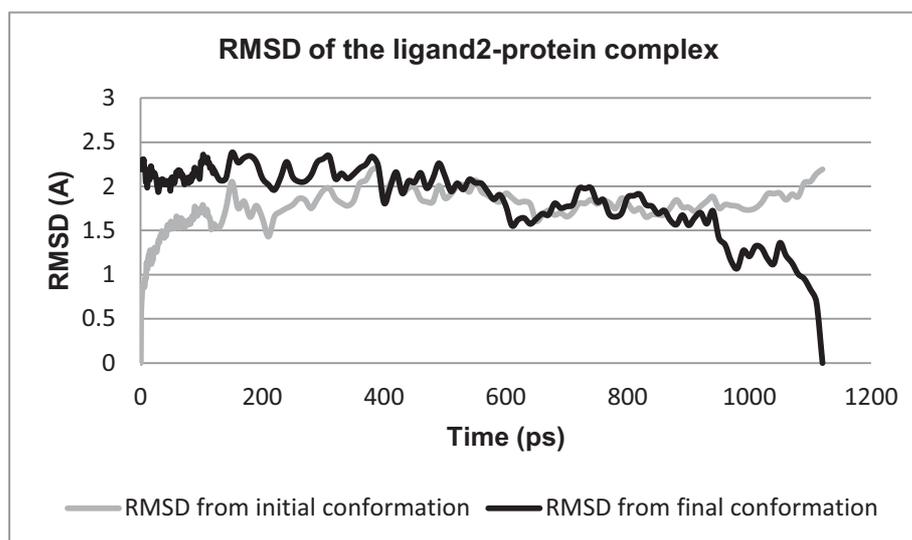
Another key estimator of the structural similarity is the RMSD (Root mean square distance) calculation (Carugo, O *et al.*, 2003). The RMSD of the final conformation in the model was also calculated from the zero-th conformation of the model structure and the average distance between the atoms was calculated. Since the initial crystal structure was determined at a very high resolution of 1.6 Å, it should have smaller RMSD value (Carugo, O. *et al.*, 2003). It was observed that the RMSD values were increased when the molecular dynamics studies were performed at more than 1120ps (Graph 2.2.)



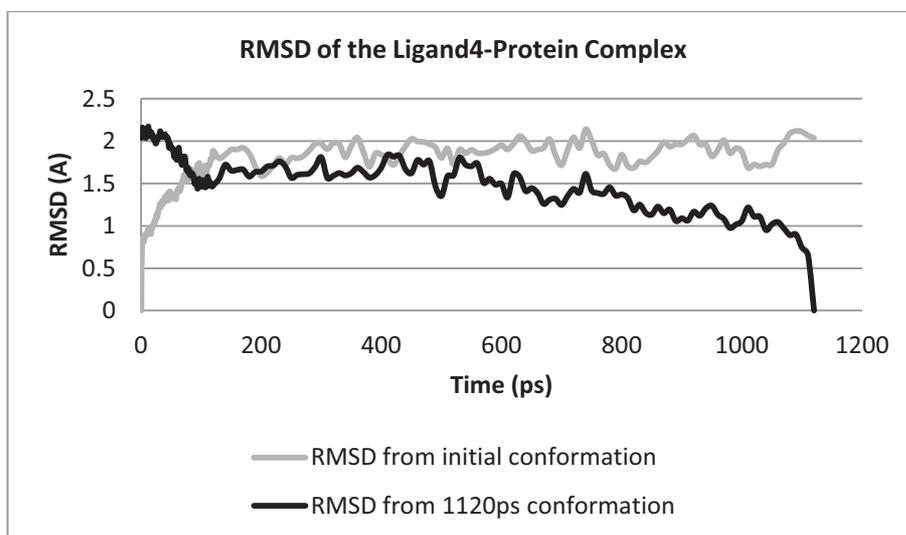
Graph 2.23: RMSD changes of the **TSA**-protein complex during MD study at 12330ps.



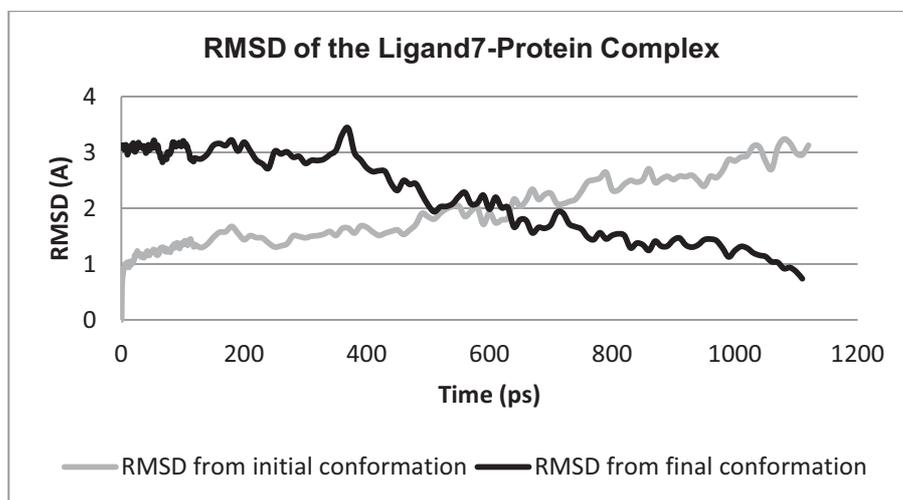
Graph 2.24: RMSD changes of the Ligand1-protein complex during MD study at 1120ps during the Phase I study.



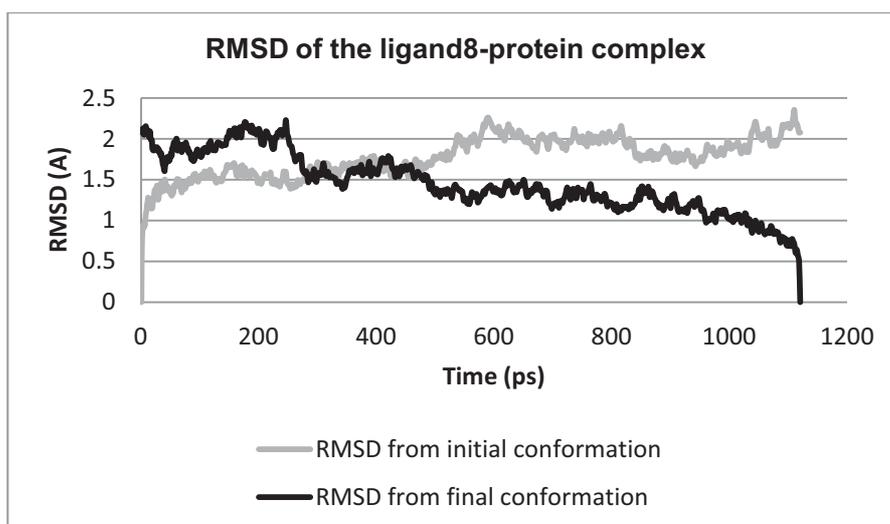
Graph 2.25: RMSD changes of the Ligand2-protein complex during MD study at 1120ps during the Phase II study



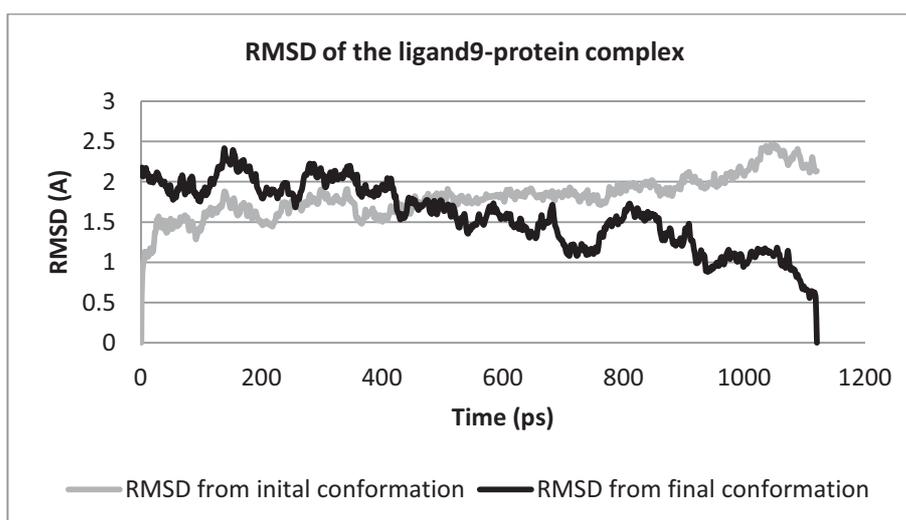
Graph 2.26: RMSD changes of the Ligand4-protein complex during MD study at 1120ps during the Phase II study



Graph 2.27: RMSD changes of the Ligand7-protein complex during MD study at 1120ps during the Phase II study



Graph 2.28: RMSD changes of the Ligand8-protein complex during MD study at 1120ps during the Phase II study



Graph 2.29: RMSD changes of the Ligand9-protein complex during MD study at 1120ps during the Phase II study.

## 2.8 Docking studies:

The docking studies carried out in this investigation were performed using the active site module in *Fujitsu* CAChe. The ligands and the proteins were used from separate files and docking studies were performed using the flexible ligand and flexible active site options. The docking program utilises the genetic algorithm based approach using Amber van der Waals based calculation in the population size of 50, cross over rate 0.800000 and maximum generation 3000. The active site bounding box was defined by 20 Å regions (X,Y and Z). All

the docking studies were carried out at least three times and the average scores were considered.

Following a series of molecular modelling studies as shown in Scheme 2.2 of the Phase 1 study, the ligands from table 2.4 were docked into the active site of the protein model obtained following a MD study on the **TSA**-protein complex at 1120ps. A brief overview of the steps employed to establish the targeted model using a flow diagram is shown in Appendix I. In the aim to establish the target model, attempts have also been made to perform MD study on the other biologically active ligand-protein complexes. These docking scores were laid out in Appendix II. Docking studies in the active site of the target model with each of the ligands were carried out at least three times and the docking scores collected were shown in Table 2.8.

Ligands	$K_i$ (nM)	PMF Score1 (Kcal/mol)	PMF Score 2 (Kcal/mol)	PMF Score 3 (Kcal/mol)	Average score (Kcal/mol)
Ligand 1	5.7±1.2	-330.19	-496.8	-413	-413.33
Ligand 2	17.71±3.35	-257.73	-415.12	-402.35	-358.4
Ligand 4	28.8±4.1	-215.41	-384.65	-332.08	-310.71
Ligand 7	Inactive	-422.8	-358	-405	-395
Ligand 8	Inactive	-330.45	-262.77	-195.86	-263.02
Ligand 9	Inactive	120.60	-20.90	-10.5	29.73
Ligand 13	Inactive	-309.553	-296.49	-263.79	-289.94
<b>TSA</b>		-398.4	-386.2	-374.08	-386.2

Table 2.8: Showing the docking score of the ligands.

### 2.8.1 Docking scores results:

Scoring was carried out to rank the placement of the ligands to one another. In an ideal situation the most active ligand would score most negative while the less active would score more positive. The overall idea for the computer model was to produce docking scores that will reflect the possibility of the biological activity of the ligands used in the docking study to the maximum extent. All the ligands were designed by studying the interactions within the active site.

Okvist, *et al.*, in the earlier research have successfully illustrated the entry of the **TSA** (Transitional state analogue) into the active site and also have determined the interaction of the **TSA** within the active site. As shown in Figure 2.3, within the active site Lys60, Arg49,

Arg 72 Glu106 and Glu109 were involved in the most interactions with the ligands. Ligand 1 has been so far reported to be the most active compound against the MtCM (Agrawal, Kumar *et al.* 2007) after **TSA**. In the current model, Ligand 1 has also successfully produced docking scores which were much negative than other active ligands including **TSA**. Therefore in the current investigation Ligand 1 was manipulated with several substituent (Figure 2.8) depending on the availability of space within the active sites.

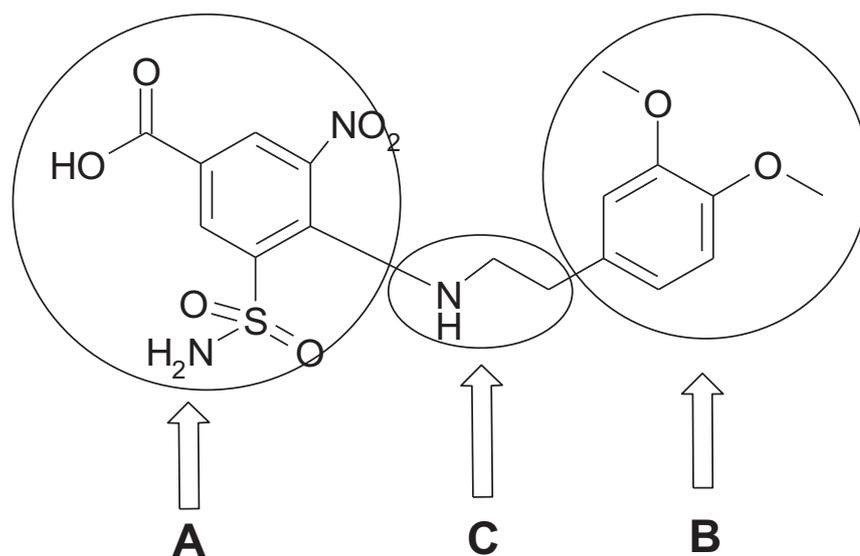
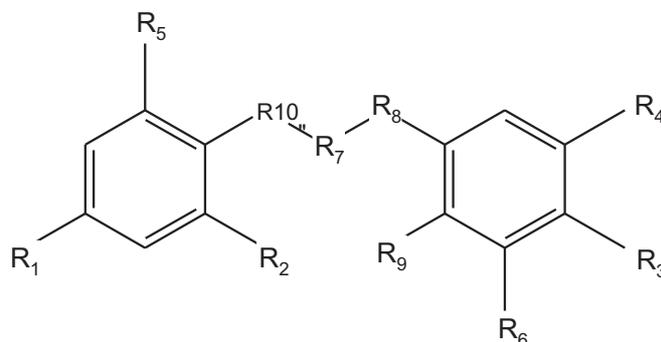


Figure 2.8: Showing the Ligand 1 and A, B and C points to the targets where the variations have been made to design various ligands.

In the aim to design inhibitors for the model the A, B and C part of the Ligand 1 was modified, as shown in Figure 2.5. The nitro and nitroso groups can serve as close analogues of carboxylic substrates for enzymes and also that the nitro group is more electronegative than the carboxylic group may serve as a better analogue of Ligand 1 (Alston, T. 1983). Therefore in this study the carboxylic group was modified with a nitro group. The sulphonamide group is an electron withdrawing species and therefore was also attempted to substitute with the nitro group. Interestingly S-(-)-dinitrobiphenic acid has shown interesting chorismate mutase activity from *E.coli*, *B.subtilis* and *S.cerevisiea* (Husain, A 1999). Therefore in this work the ligand was modified with two nitro groups. The sulphonamide chain was also increased by incorporating some methyl groups, such as N-methyl sulphonamide. The linkage between the two chains was also modified by substituting with amino or amide groups, however the distance between the two rings was left untouched. Since the methoxy groups have shown good intermolecular and intramolecular hydrogen bonding, in this work the ring B of ligand 1 (Figure 2.8) was also substituted with hydroxyl groups. At the same time these fragments

were also used separately and docked into the active site. These new ligands were docked at least three times, scored and their orientations at the active sites were studied.

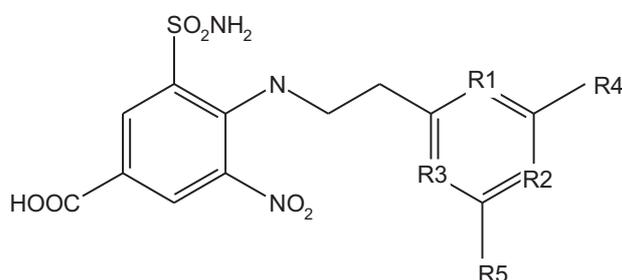


Ligands	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Ligand 16	NO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 17	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 18	CO <sub>2</sub>	NO <sub>2</sub>	CH <sub>2</sub> OH	CH <sub>2</sub> OH	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 19	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	CH <sub>2</sub> OH	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 20	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	CO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 21	CO <sub>2</sub>	CO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 22	CO <sub>2</sub>	NO <sub>2</sub>	OH	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 23	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 24	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	OH	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 25	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	NH <sub>2</sub>	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 26	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 27	NO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 28	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>2</sub> CH <sub>3</sub>	H	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>2</sub> CH <sub>3</sub>	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 29	NO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 30	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 31	CO <sub>2</sub>	CO <sub>2</sub> CH <sub>3</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 32	CO <sub>2</sub>	CH <sub>2</sub> NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 33	CO <sub>2</sub>	CO <sub>2</sub> CH <sub>3</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 34	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NHCH <sub>3</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 35	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>		H	NH

Ligands	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Ligand 36	CO <sub>2</sub>	NO <sub>2</sub>	OH	H	SO <sub>2</sub> NH <sub>2</sub>	OH	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 37	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NCH <sub>3</sub>
Ligand 38	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 39	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 40	CO <sub>2</sub>	NO <sub>2</sub>	H	H	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	N	CH <sub>2</sub>	OCH <sub>3</sub>	NH
Ligand 41	NO <sub>2</sub>	CO <sub>2</sub>	OH	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 42	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 43	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	OH	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 44	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 45	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 46	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	CO	CH <sub>2</sub>	H	NH
Ligand 47	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	N	CO	H	NH
Ligand 48	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	CO	CH <sub>2</sub>	H	NH
Ligand 49	NO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	CO	CH <sub>2</sub>	H	NH
Ligand 50	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 51	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	NO <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 52	CO <sub>2</sub>	NO <sub>2</sub>	OH	H	NO <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 53	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	OH	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 54	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	OH	NH
Ligand 55	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	NO <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH
Ligand 56	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	OH	NH
Ligand 57	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	CO	CH <sub>2</sub>	H	NH
Ligand 58	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CO	CH <sub>2</sub>	H	NH
Ligand 59	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	N	CO	H	NH
Ligand 60	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	H	N	CO	H	NH
Ligand 61	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	NO <sub>2</sub>	H	N	CO	H	NH
Ligand 62	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	N	CO	H	NH
Ligand 63	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 64	CO <sub>2</sub>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	NO <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH

Ligands	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Ligand 65	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 66	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 67	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	NO <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 68	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	N	CH <sub>2</sub>	H	NH
Ligand 69	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	NO <sub>2</sub>	H	N	CO	H	NH
Ligand 70	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	N	CO	H	NH
Ligand 71	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	NO <sub>2</sub>	H	N	CO	H	NH
Ligand 72	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	NO <sub>2</sub>	H	CO	CH <sub>2</sub>	H	NH
Ligand 73	CO <sub>2</sub>	NO <sub>2</sub>	H	OCH <sub>3</sub>	SO <sub>2</sub> NH <sub>2</sub>	H	CO	CH <sub>2</sub>	H	NH
Ligand 74	CO <sub>2</sub>	NO <sub>2</sub>	OH	OH	SO <sub>2</sub> NH <sub>2</sub>	H	CO	CH <sub>2</sub>	H	NH
Ligand 75	CO <sub>2</sub>	NO <sub>2</sub>	H	H	SO <sub>2</sub> NH <sub>2</sub>	H	CH <sub>2</sub>	CH <sub>2</sub>	H	NH

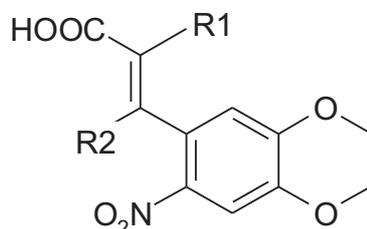
Table 2.9: Ligands designed as analogue of Ligand 1.



Ligands	R1	R2	R3	R4	R5
Ligand 76	CH	N	CH	OCH <sub>3</sub>	H
Ligand 77	N	CH	CH	OCH <sub>3</sub>	H
Ligand 78	CH	N	CH	OCH <sub>3</sub>	OCH <sub>3</sub>
Ligand 79	CH	CH	N	H	OCH <sub>3</sub>
Ligand 80	CH	N	CH	H	H

Table 2. 10: Analogues of ligand 1 containing a pyridyl moiety.

Ligand **2** was the next potentially active mycobacterial chorismate mutase inhibitor. Interestingly the ligand carries structural similarity to ligand 1. This ligand was also modified and some of the analogues designed were illustrated in the table below.



Ligands	R1	R2
Ligand <b>81</b>	H	H
Ligand <b>82</b>	H	tert-butyl
Ligand <b>83</b>	tert-butyl	H

Table 2.11: Showing the various ligands designed as analogue of Ligand **2**.

Ligand **1** contains two ring systems and therefore it was also essential to study the interactions of the substituents more closely. The ligand was fragmented and the fragments were also modified (Table 2.12) to study the interactions more closely.

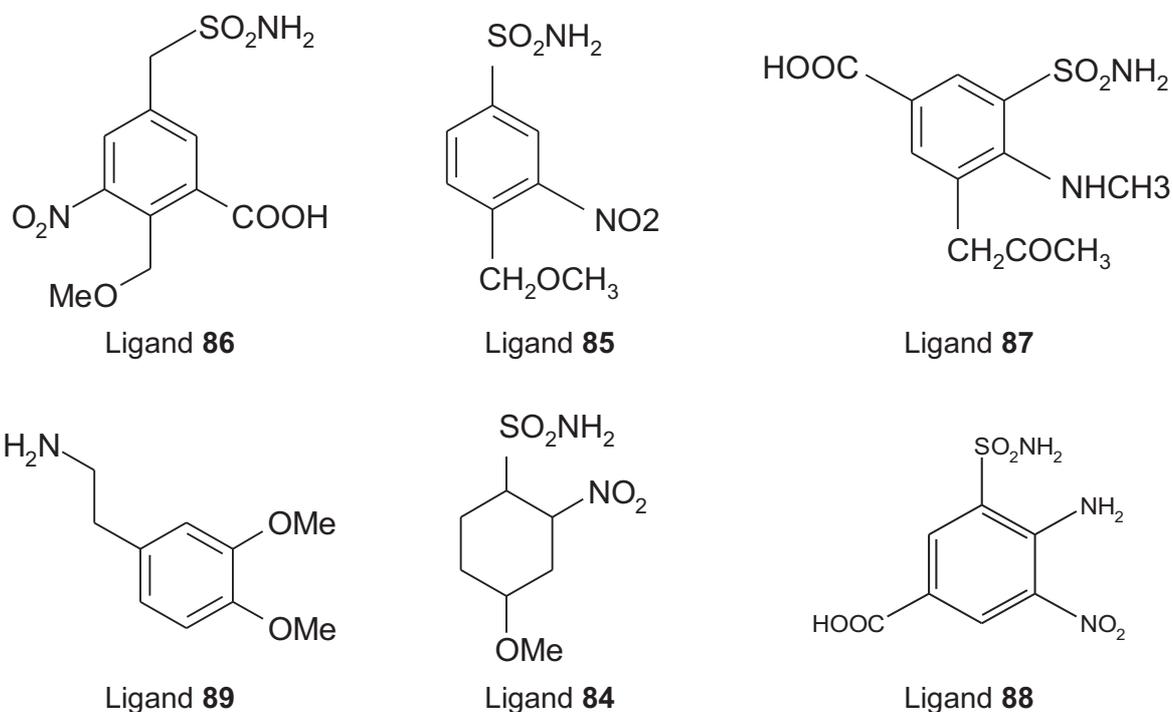
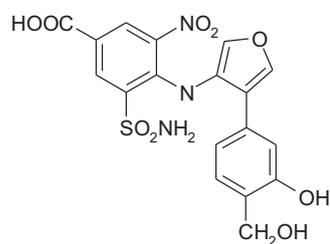
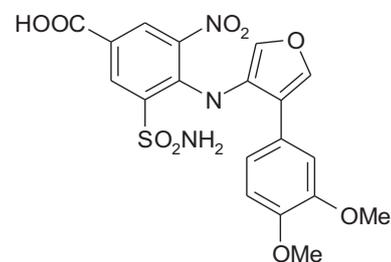


Table 2.12: Showing the list of the fragments used to study the interactions within the active site.

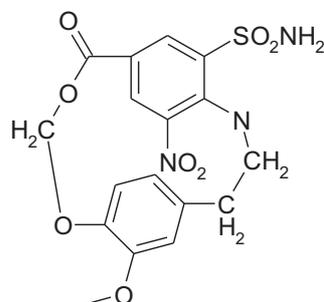
The ligand 1 when docked into the active site was observed in a folding fashion. Thus to facilitate the interactions within the active site, the folded pattern of the ligand was locked and the following ligands were produced.



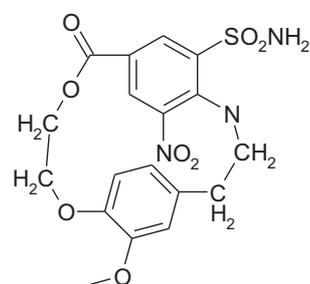
Ligand 90



Ligand 91



Ligand 93



Ligand 92

Table 2.13: Showing the list of analogues of Ligand 1 where the folding pattern of the ligands was fixed.

All the ligands shown in the table 2.9, 2.10, 2.11, 2.12 and 2.13 were docked into the active site and the docking scores were obtained as shown in the table 2.14 below.

Compounds	Dock 1 Kcal/mol	Dock 2 Kcal/mol	Dock 3 Kcal/mol	Average Score Kcal/mol
TSA	-402	-367	-385	-385
Ligand 1	-533	-336	-501	-456
Ligand 2	-464	-450	-348	-421
Ligand 4	-407	-422	-405	-411
Ligand 7	-347	-367	-348	-354
Ligand 8	-350	-455	-384	-396
Ligand 9	-303	-164	-432	-299
Ligand 16	-438	-472	-328	-413

<b>Compounds</b>	<b>Dock 1 Kcal/mol</b>	<b>Dock 2 Kcal/mol</b>	<b>Dock 3 Kcal/mol</b>	<b>Average Score Kcal/mol</b>
Ligand 17	-456	-496	-494	-482
Ligand 18	-478	-400	-205	-361
Ligand 19	-361	-428	-359	-383
Ligand 20	-298	-483	-523	-435
Ligand 21	-429	-405	-381	-405
Ligand 22	-512	-423	-502	-479
Ligand 23	-507	-248	-531	-429
Ligand 24	-384	-539	-486	-470
Ligand 25	-477	-358	-490	-442
Ligand 26	-570	-390	-356	-439
Ligand 27	-420	-349	-495	-421
Ligand 28	-469	64	-266	-224
Ligand 29	-337	-458	-499	-431
Ligand 30	-276	-509	42	-248
Ligand 31	0	-432	-450	-294
Ligand 32	-409	-390	-453	-417
Ligand 33	-409	-394	-316	-373
Ligand 34	-403	-459	-331	-398
Ligand 35	-376	-345	-384	-368
Ligand 36	-437	-437	-287	-387
Ligand 37	-412	-412	-412	-412
Ligand 38	-461	-552	-476	-496
Ligand 39	-454	-222	-492	-389
Ligand 40	-508	-302	-484	-431
Ligand 41	-537	-458	-390	-462
Ligand 42	-263	-472	-471	-402
Ligand 43	-538	-293	-469	-433
Ligand 44	-391	-414	-452	-419
Ligand 45	-275	-359	-431	-355
Ligand 46	-396	-224	-431	-350
Ligand 47	-439	-481	-417	-446
Ligand 48	-439	-447	-414	-433
Ligand 49	-442	-464	-423	-443

<b>Compounds</b>	<b>Dock 1 Kcal/mol</b>	<b>Dock 2 Kcal/mol</b>	<b>Dock 3 Kcal/mol</b>	<b>Average Score Kcal/mol</b>
Ligand 50	-507	-422	-486	-472
Ligand 51	-498	-91	-496	-362
Ligand 52	-485	-486	-437	-469
Ligand 53	-407	-497	-498	-467
Ligand 54	-435	-408	-448	-430
Ligand 55	-307	-340	-445	-364
Ligand 56	-478	-445	-525	-483
Ligand 57	-523	-380	-285	-396
Ligand 58	-223	-446	-409	-359
Ligand 59	-469	-436	-416	-440
Ligand 60	-406	-395	-491	-431
Ligand 61	-513	-129	-452	-365
Ligand 62	-420	-371	-410	-400
Ligand 63	-440	-412	-474	-442
Ligand 64	-488	-384	-512	-461
Ligand 65	-256	-425	-392	-358
Ligand 66	-301	-400	-432	-378
Ligand 67	-456	-446	-345	-416
Ligand 68	-267	-401	-397	-355
Ligand 69	-347	-502	-474	-441
Ligand 70	-492	-475	-452	-473
Ligand 71	-498	-442	-448	-463
Ligand 72	-370	-280	-446	-365
Ligand 73	-428	-436	-461	-442
Ligand 74	-455	-473	-474	-467
Ligand 75	-439	-447	-461	-449
Ligand 76	-305	-422	-429	-385
Ligand 77	-447	-515	-392	-451
Ligand 78	-334	-457	-488	-426
Ligand 79	-440	-226	-369	-345
Ligand 80	-437	-437	-287	-387
Ligand 81	-427	-431	-424	-427
Ligand 82	-233	-413	-432	-359

<b>Compounds</b>	<b>Dock 1 Kcal/mol</b>	<b>Dock 2 Kcal/mol</b>	<b>Dock 3 Kcal/mol</b>	<b>Average Score Kcal/mol</b>
Ligand <b>83</b>	-438	-492	-497	-476
Ligand <b>84</b>	-502	-479	-484	-488
Ligand <b>85</b>	-247	-403	-361	-337
Ligand <b>86</b>	-445	-444	-396	-428
Ligand <b>87</b>	-431	-194	-437	-354
Ligand <b>88</b>	-194	-354	-336	-295
Ligand <b>89</b>	-263	-472	-471	-402
Ligand <b>90</b>	-420	-347	-392	-386
Ligand <b>91</b>	-403	-382	-396	-394
Ligand <b>92</b>	-380	-314	-242	-312
Ligand <b>93</b>	-415	-351	-390	-385

Table 2.14: The docking scores of the designed ligands carried out on the target protein model.

A set of the original 56,000 Maybridge compounds was selected using a pharmacophore diversity based approach (Rathbone, personal communication). All these 260 ligands were isolated which had diverse structures and were used for docking studies within the active site of the model. The average score of three consecutive docking scores were shown in the Table 2.15 below.

<b>Compound</b>	<b>PMF score Kcal/mol</b>	<b>Compound</b>	<b>PMF score Kcal/mol</b>
Ligand <b>94</b>	903.809	Ligand <b>211</b>	-503.106
Ligand <b>95</b>	-289.859	Ligand <b>212</b>	-432.588
Ligand <b>96</b>	-370.758	Ligand <b>213</b>	-387.079
Ligand <b>97</b>	-404.065	Ligand <b>214</b>	-211.099
Ligand <b>98</b>	-423.719	Ligand <b>215</b>	-217.48
Ligand <b>99</b>	-352.42	Ligand <b>216</b>	-275.017
Ligand <b>100</b>	-433.188	Ligand <b>218</b>	-313.248
Ligand <b>101</b>	-366.871	Ligand <b>219</b>	-66.191
Ligand <b>102</b>	-368.085	Ligand <b>220</b>	-250.647
Ligand <b>103</b>	-388.601	Ligand <b>221</b>	-234.052
Ligand <b>104</b>	103.283	Ligand <b>222</b>	346.73
Ligand <b>105</b>	-188.97	Ligand <b>223</b>	-207.314

<b>Compound</b>	<b>PMF score Kcal/mol</b>	<b>Compound</b>	<b>PMF score Kcal/mol</b>
Ligand 106	-333.87	Ligand 224	-216.269
Ligand 107	5885.056	Ligand 225	59.399
Ligand 108	-264.047	Ligand 226	-247.22
Ligand 109	-307.483	Ligand 227	-453.834
Ligand 110	4856.811	Ligand 229	-273.724
Ligand 111	-300.456	Ligand 230	456.214
Ligand 112	-382.258	Ligand 231	-328.143
Ligand 113	-646.085	Ligand 232	-332.942
Ligand 114	-426.009	Ligand 233	-480.885
Ligand 115	-442.108	Ligand 234	-297.012
Ligand 116	-413.204	Ligand 235	-189.643
Ligand 117	-237.543	Ligand 236	-405.088
Ligand 118	-349.447	Ligand 237	-170.258
Ligand 119	-264.683	Ligand 238	-396.21
Ligand 120	517.59	Ligand 239	-353.808
Ligand 121	-450.52	Ligand 240	-110.18
Ligand 122	-303.779	Ligand 241	-317.626
Ligand 123	-527.327	Ligand 242	-330.445
Ligand 124	-631.255	Ligand 243	-206.029
Ligand 125	-546.48	Ligand 244	-341.858
Ligand 126	-230.249	Ligand 245	-462.645
Ligand 127	-488.645	Ligand 246	-157.58
Ligand 128	-296.453	Ligand 247	-264.888
Ligand 129	-162.865	Ligand 248	-384.475
Ligand 130	120.337	Ligand 249	-396.179
Ligand 131	-377.085	Ligand 250	60.086
Ligand 133	-464.22	Ligand 251	-411.724
Ligand 134	233.1	Ligand 252	-293.691
Ligand 135	-276.439	Ligand 254	-471.139
Ligand 136	-53.831	Ligand 255	-534.7
Ligand 137	-451.355	Ligand 256	-307.348
Ligand 138	-246.807	Ligand 257	-162.041
Ligand 139	-108.349	Ligand 258	-339.091
Ligand 140	-211.143	Ligand 259	-199.996
Ligand 141	-262.27	Ligand 260	-139.407
Ligand 142	-577.53	Ligand 261	-399.476
Ligand 143	-410.703	Ligand 262	-407.999

<b>Compound</b>	<b>PMF score Kcal/mol</b>	<b>Compound</b>	<b>PMF score Kcal/mol</b>
Ligand 144	-367.269	Ligand 263	-357.301
Ligand 145	-337.978	Ligand 264	-425.224
Ligand 146	-262.69	Ligand 265	-147.716
Ligand 147	-236.638	Ligand 266	-518.084
Ligand 148	-318.823	Ligand 267	-200.616
Ligand 149	89.489	Ligand 268	-153.544
Ligand 150	-399.989	Ligand 269	-189.727
Ligand 151	-508.814	Ligand 271	-217.007
Ligand 152	-448.471	Ligand 272	-446.237
Ligand 153	-315.915	Ligand 273	-453.142
Ligand 154	-262.519	Ligand 274	-266.539
Ligand 155	-316.276	Ligand 276	384.478
Ligand 157	-256.367	Ligand 277	-484.746
Ligand 158	-353.222	Ligand 278	-247.428
Ligand 159	-402.188	Ligand 279	-172.754
Ligand 160	-501.993	Ligand 280	-230.871
Ligand 162	-485.757	Ligand 281	-217.327
Ligand 163	-270.047	Ligand 282	-308.769
Ligand 164	-415.77	Ligand 283	-202.429
Ligand 165	-465.685	Ligand 284	-339.646
Ligand 166	-291.715	Ligand 285	-351.383
Ligand 167	-423.912	Ligand 286	-326.243
Ligand 168	-393.311	Ligand 287	-276.71
Ligand 169	-441.17	Ligand 289	-434.703
Ligand 171	-243.313	Ligand 290	-186.222
Ligand 172	-401.559	Ligand 291	-252.54
Ligand 173	-234.007	Ligand 292	-406.386
Ligand 174	-374.444	Ligand 293	-333.132
Ligand 175	2057.602	Ligand 295	-300.264
Ligand 176	-234.631	Ligand 296	-438.455
Ligand 177	-289.372	Ligand 297	-391.53
Ligand 178	-295.651	Ligand 298	-238.05
Ligand 179	-331.766	Ligand 299	-380.009
Ligand 180	-342.5	Ligand 300	-339.46
Ligand 181	-346.86	Ligand 301	-356.298
Ligand 183	-530.58	Ligand 302	-394.023
Ligand 184	-246.405	Ligand 303	-408.315

Compound	PMF score Kcal/mol	Compound	PMF score Kcal/mol
Ligand 185	874.544	Ligand 304	-284.57
Ligand 186	-107.315	Ligand 305	-290.941
Ligand 187	-163.381	Ligand 306	-238.199
Ligand 188	-375.039	Ligand 309	-240.694
Ligand 189	-460.787	Ligand 310	-183.745
Ligand 190	-462.195	Ligand 313	-320.875
Ligand 191	339.762	Ligand 314	-282.372
Ligand 192	-318.697	Ligand 315	-166.965
Ligand 193	-292.64	Ligand 316	-184.27
Ligand 194	-508.736	Ligand 317	-261.56
Ligand 195	-279.02	Ligand 318	-285.319
Ligand 196	-312.582	Ligand 319	-268.762
Ligand 197	-510.686	Ligand 320	-228.665
Ligand 198	-279.421	Ligand 321	-352.863
Ligand 199	-378.74	Ligand 322	-287.753
Ligand 200	-398.113	Ligand 323	-273.305
Ligand 201	-631.172	Ligand 324	3373.438
Ligand 202	-183.187	Ligand 325	-287.861
Ligand 204	-467.787	Ligand 326	-455.688
Ligand 205	-526.051	Ligand 327	-365.032
Ligand 206	-387.601	Ligand 328	Not found
Ligand 207	-232.137	Ligand 329	-373.639
Ligand 208	-500.701	Ligand 330	-455.089
Ligand 209	-323.651	Ligand 331	-361.732
Ligand 210	-371.257	Ligand 332	-347.454

Table 2.15: The docking scores following the docking studies carried out with the Maybridge compounds using the target protein model.

## 2.9 Identifying the interactions:

The orientation of the ligand within the active site was carefully studied and it was observed that in majority of the docking studies particularly the target model (obtained at 1120ps following the MD study on the **TSA**-protein complex) Arg 260, Lys 186 and Glu 235 interacted with the ligand in a similar fashion forming a triangular space within the receptor site (Figure 2.10a-d). **TSA**, when docked within this protein model, exhibited a similar pattern of orientation as found initially within the X-ray crystal structure (Figure 2.9). Lys 186

undergoes a hydrogen bonding with all of the active ligands. This interaction was also found within the X-ray crystal structure containing the **TSA** (Lys60 as in Figure 2.3). Therefore hydrogen bonding with Lys186 was deemed essential in the active site.

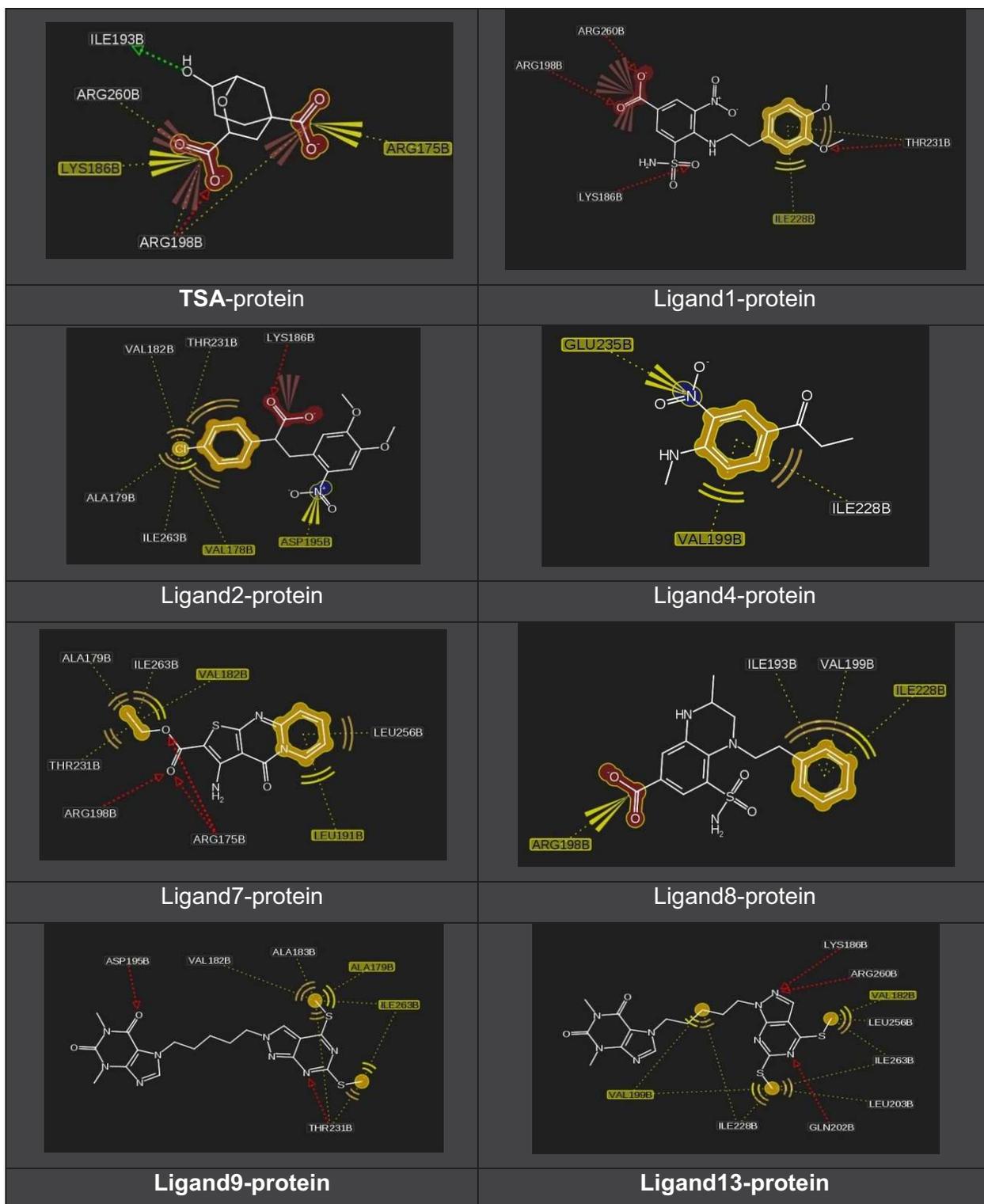
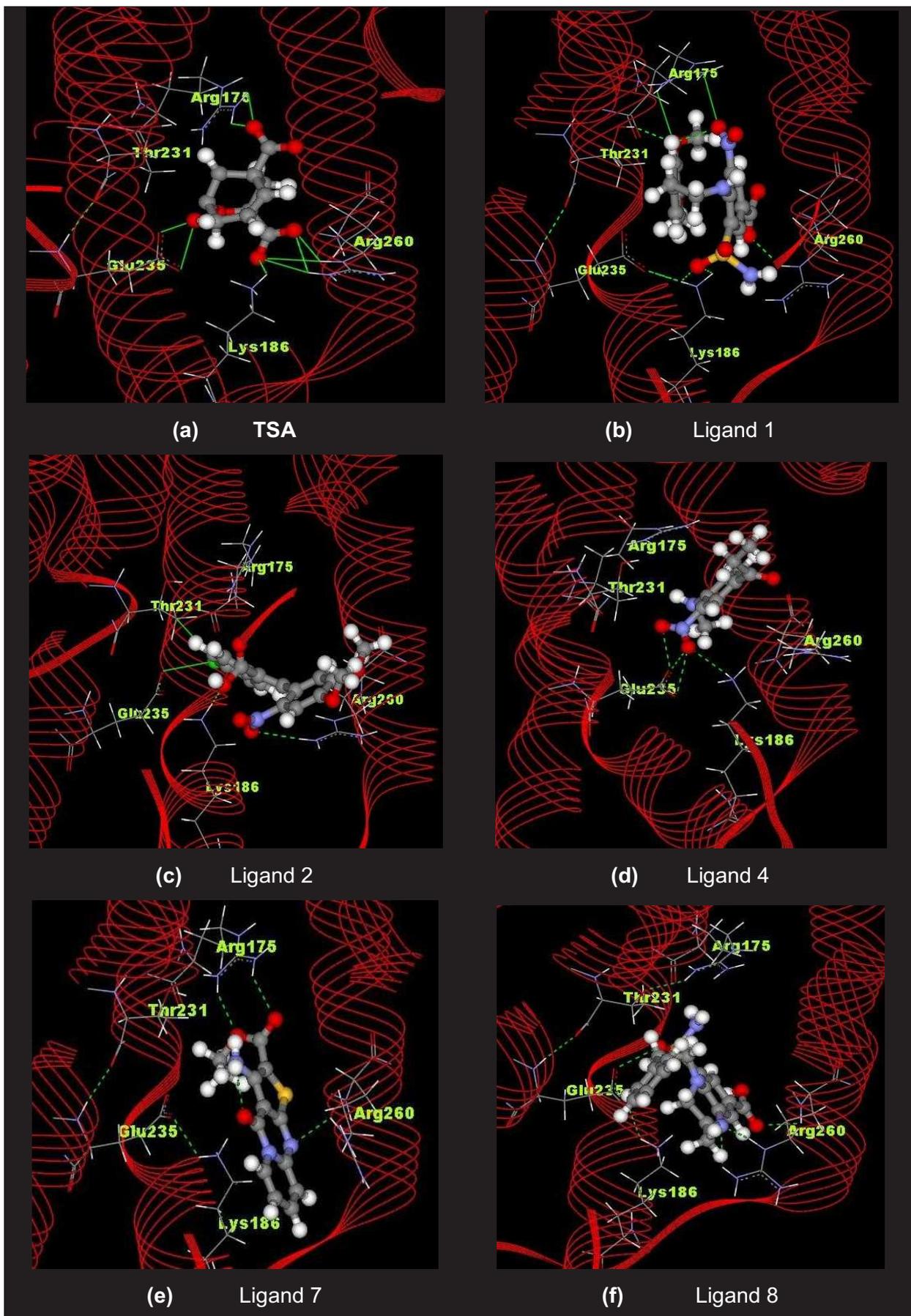


Figure 2.9: Two dimensional view (using the software 'Ligand Scout') of the ligand interacting with the residues within the active site model obtained at 1120ps following a MD study of the TSA-protein complex (scheme 2.2).



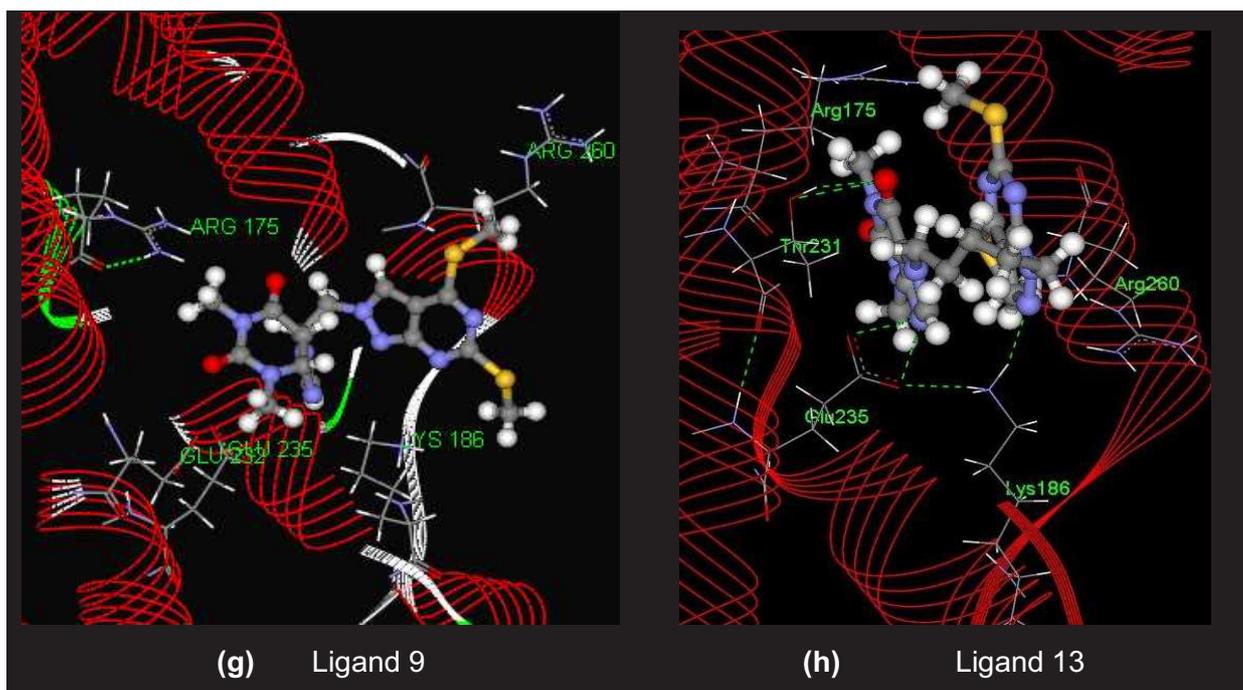


Figure 2.10: The 3D view using 'DS visualiser 2', of the ligand interacting with the residues within the active site of the target model obtained at 1120ps following a MD study of the TSA-protein complex (scheme 2.2).

No analogues of **TSA** exhibited equivalent interactions as observed with **TSA** (Hediger, M. *et al.*, 2004). In contrast Ligand **1** in this study exhibited interactions with the protein of a similar nature to that seen with **TSA**. Ligand **1** is an open chain ligand and has shown interaction within the active site of this model through a folding pattern. This pattern of folding followed by the interactions was mediated by Arg 175 with the nitro and the oxygen of methoxy group of Ligand **1** (Figure 2.10b).

An exception to this was the Ligand **2**. This ligand had a chlorine atom attached to the phenyl ring and this ligand remained unfolded upon docking. On repeated docking of this ligand, it was observed that 2 out of three times the chlorine stayed within the active site leaving the methoxy benzyl part of the molecule toward the surface. The chlorine part of the molecule tended to stay within the active site forming a straight conformation as shown in Figure 2.10c. The chlorine within the active site interacted with Glu235 and Thr231 whereas the nitro formed hydrogen bonds with Arg260.

Most of the inactive ligands had shown less negative docking scores in comparison to the active ligands. Also the interactions with Arg 260, Lys 186 and Glu 235 which were preserved in the active ligands were found absent with the inactive ligands. Hydrogen bonding with Lys186 was absent in most of the inactive ligands (Figure 2.7e-h). Also the folding pattern was absent with these ligands although the ligands has more than one ring

system. One exception to these is the Ligand **13**. It was difficult to fit this ligand into this model as the ligand was a large molecule with a large interacting surface and plenty of interacting groups. Therefore in the current investigation Ligand **13** was excluded from the study. In fact the study has rejected any large molecule. Ligands which are small in size including ligands containing more than one ring system were included in the study.

## 2.10 Discussion:

Chorismate mutase is known to act as a dimer (Figure 2.1) (Okvist *et al*, 2003) and perhaps the surface binding of one monomer to another has some influence on the three dimensional shape of the monomer and this may affect the three dimensional geometry of the active site. Thus despite the presence of the active site in the Chain B, both the chains were incorporated in the entire study on this protein structure. The target protein model (MtCM) was prepared through molecular dynamics studies carried out with the protein **TSA** complex over 1120ps as shown in Scheme 2.1 and Scheme 2.2 (Appendix I). This particular model has retained all the interesting interactions of the **TSA** (ligand) with the neighbouring residues, which were essential in the active site (Okvist *et al*, 2003). This model was therefore docked with some active and inactive ligands to set up a series of hypotheses for the model, these hypotheses were then used to design further ligands with possible biological activity against the organism.

### 2.10.1 Model Hypothesis:

- ❖ The ligand should interact with the residues Arg 260, Lys 186 and Glu 235, forming a triangular space.
- ❖ Lys 186 hydrogen bonds with the carbonyl oxygen atom of the ligand.
- ❖ Thr 231 hydrogen bonds to the nitro group as in Ac1, suggesting that the Thr 231 may be an essential residue for drug design.
- ❖ Lys 186 and Glu 235 should form a hydrogen bond together and interact with the ligand's functional groups.
- ❖ The distance between Lys186 and Arg260 should not be more than 4 Å.
- ❖ A ligand with more than two ring systems connected by a chain should fold in upon itself and this is mainly mediated by Arg 175, thereby resembling the structure of **TSA**.
- ❖ The connection between the two rings of the ligand should have 3 atom systems.
- ❖ If Cl is attached to an active ligand having two ring systems, it is unlikely that the ligand will fold, and the Cl tends to remain inside the receptor site thereby producing a triangular sitting pattern with Arg260, Lys186 and Glu235.

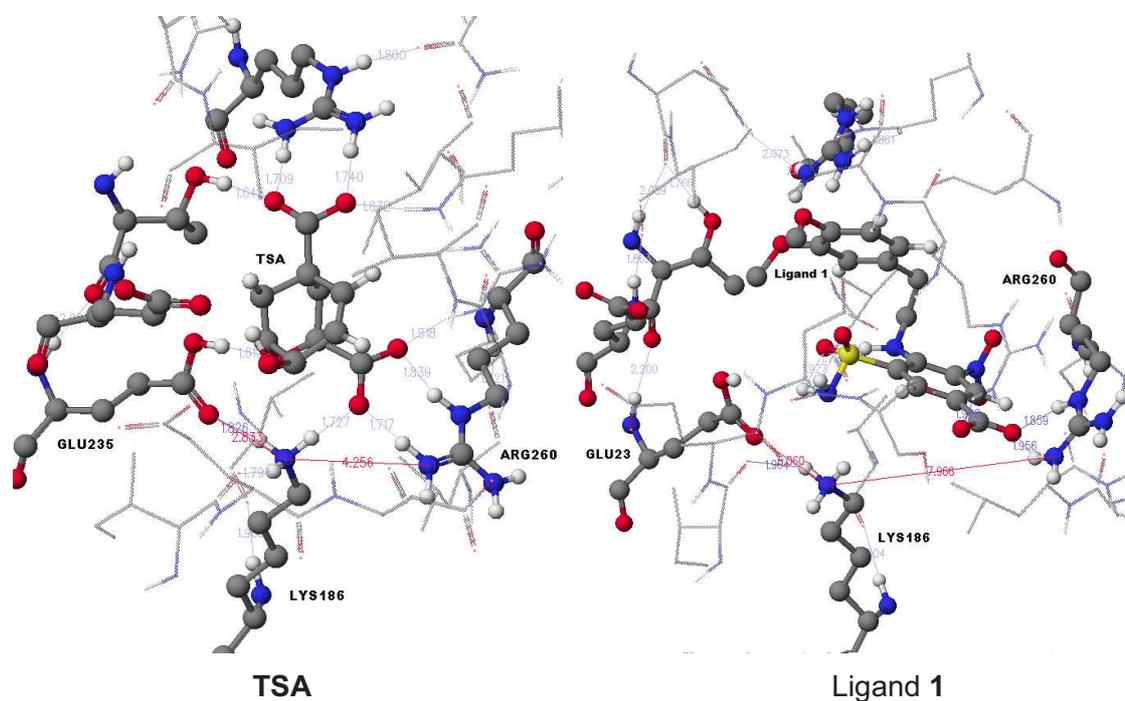
### 2.10.2 Molecular modelling results and discussions:

The protein crystal structures obtained as a pdb file from the Protein Data Bank were simple snapshots and contain co-ordinate information of only one conformation of the protein structure which may or may not be in the lowest energy conformation. Molecular dynamics was therefore essential to study the global minima or the local minima of the protein structure. When MD study was performed at 1120ps on the protein structure (2FP2.pdb) the active site shrank; this resulted in interfering with the entrance of the ligands into the active site during docking study. Thus when MD was performed directly on the downloaded protein-**TSA** complex, at 1120ps, the space available within the active site even became smaller. This has also prevented the entrance of the ligands (shown in table 2.4) into the active site during docking studies. This was because; **TSA** is bicyclic structure and was packed within the active site, running a MD simulation on this construction shrank the active site. However docking the ligand (**TSA**) into the active site enabled the active site to open up, following a MD simulation on this complex produced a model which successfully facilitated the entry of the ligands (Table 2.4) into the active site. Thus the ligands were docked into the active site of the protein (Scheme 2.2). Although the inactive ligands have shown more negative docking scores, these ligands were unable to interact appropriately with the target residues in a similar pattern as observed by the **TSA** in the active site. Thus the protein structure together with the **TSA** docked complex was set for molecular dynamics studies at 1120ps to obtain the lowest energy conformation of the active site in the protein complex. This model was then docked with the ligands shown in Table 2.4 and the interactions produced by the individual ligands were studied carefully. The docking scores obtained from Ligand **1**, Ligand **2** and Ligand **4** were more negative than the inactive ligands (scheme 2.2). The interactions produced by the **TSA** with the active site residues, were also similar to the biologically active ligands (Ligand **1**, Ligand **2** and Ligand **4**) as shown in Figure 2.11 and Figure 2.12. Although the model generated by the **TSA**-protein complex through the molecular dynamics studies at 1120ps has created a stable and sensible model which has preserved the essential interactions when ligands were docked into the active site, it was difficult to establish the model as a final protein model which can screen possible biologically active ligands.

Therefore further investigations were carried out with molecular dynamics study for another 1120ps on the ligand-protein complexes of the active ligands (Scheme 2.3) to get the most stable conformation of these ligands within the active site. These were performed essentially to study the kind of interactions produced by these ligands within the active site. Each of these models was then re-scored and the interactions at this stage were studied (scheme

2.3). The conformations of the biologically active ligands in the active site were found to have the similar interactions and conformation as was observed in Scheme 2.2. This suggests that the model produced in Scheme 2.2 generated stable conformation of these ligands.

Another key aspect that was considered while developing the model was the bond lengths. In an ideal condition within the active site Lys186 should hydrogen bond to Arg235 while interacting with the ligand. The distance between the Glu260 and the Lys186 should be around 2 Å and the distance between the Arg235 and the Lys186 should be less than 5 Å (Kameda, T *et al.* 2003). This pattern of the interaction was observed within the active site of the crystal structure (Figure 2.3a) containing the **TSA**. Interestingly when the molecular dynamics studies on the protein-ligand complexes of the Ligand 1, Ligand 2 and Ligand 4 were carried out for 1120ps (Scheme 2.3), distance between the Glu235, Lys186 and Arg260 in the active site of the protein was lost. This could be due to the large size of the ligand producing steric hindrance within the active site destroying the wanted interactions. Ligand 4 on the other hand although was smaller in size. Because it was even smaller than the **TSA**, it could preserve the interactions within the active site.



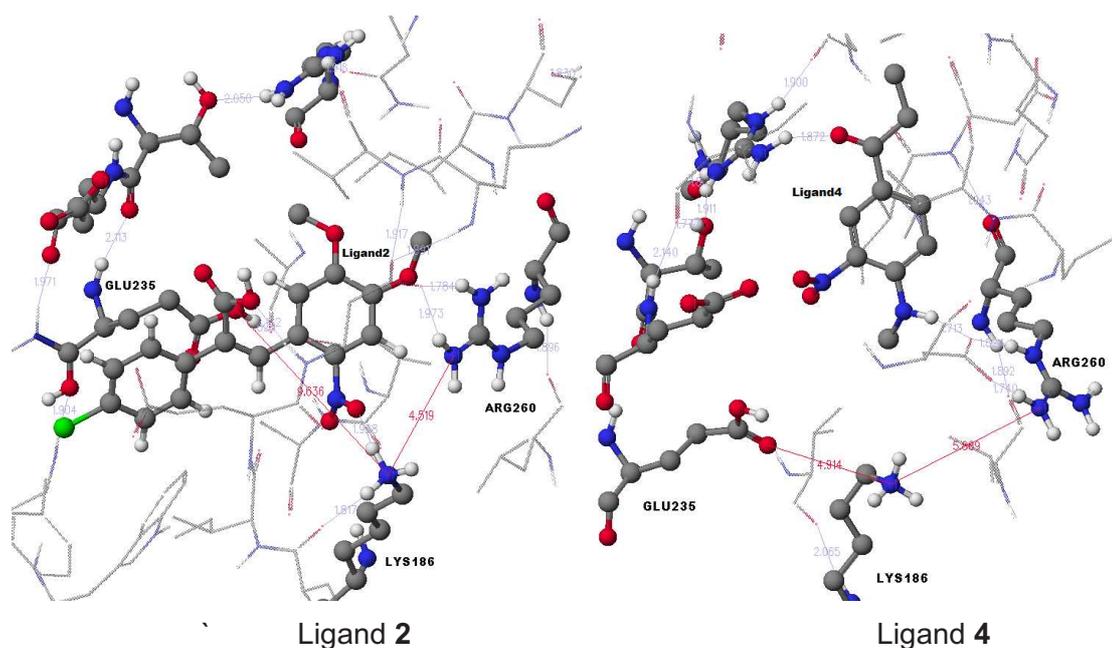


Figure 2.11: Showing the interaction of the ligands in the active site of the target model (obtained at 1120ps following the MD study using the **TSA**-protein complex (scheme 2.2)).

The RMSD value evaluation was another important aspect in identifying the correct model. The RMSD values of the ligand-protein complexes with respect to the x-ray crystal structure following the molecular dynamics studies carried out at 1120ps (table 2.16) were also investigated. The RMSD value of **TSA**-protein from the initial crystal structure at zero-th conformation to the final conformation at 1120 ps was found to be 2.5 Å. This was the least when compared with the other ligand bound protein complexes. Similarly when the RMSD of the most biologically active compound, Ligand 1 was calculated, it was found to be 3.1 Å. Molecular dynamics studies on the Ligand 1- protein complex has resulted in significant opening of the active site (Figure 2.12), interfering with the key interacting residues. Therefore **TSA**-protein complex following the molecular dynamics studies at 1120ps was designated as the final model for evaluating biological active ligands.

Ligand protein complex	RMSD from the x-ray crystal structure
<b>TSA</b> - protein complex	2.5 Å
Ligand 1- protein complex	3.1 Å
Ligand 2- protein complex	2.7 Å
Ligand 4- protein complex	3.4 Å

Table 2.16: The RMSD values of the ligand-protein complexes from the X-ray crystal structure 2FP2.pdb.

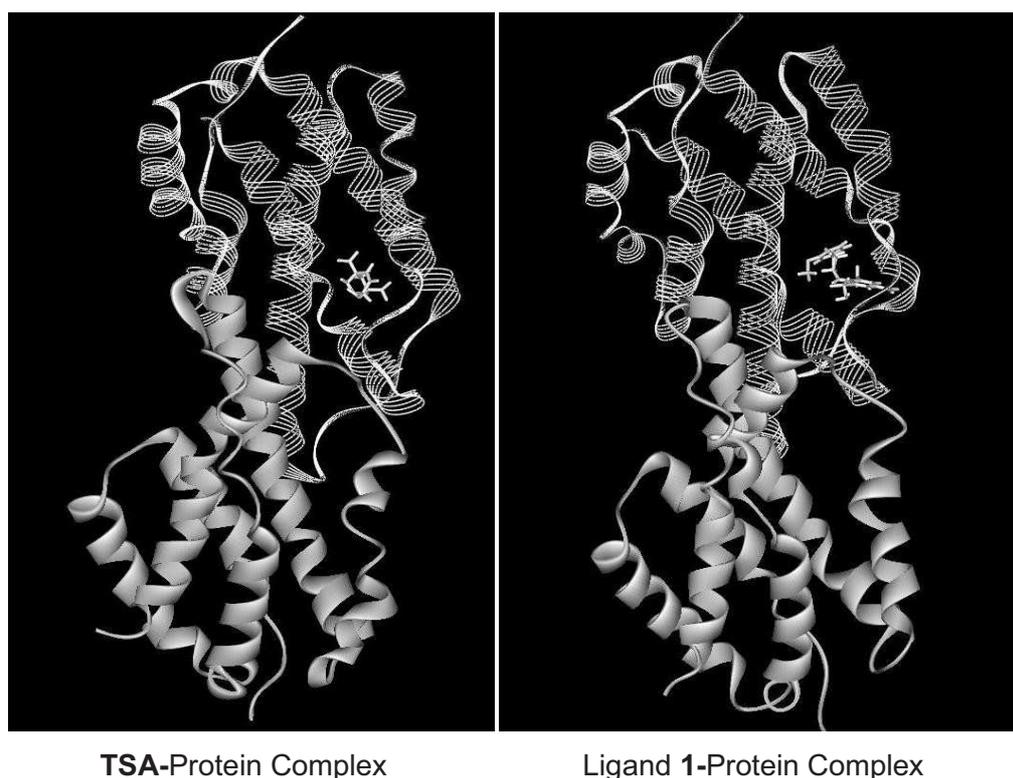


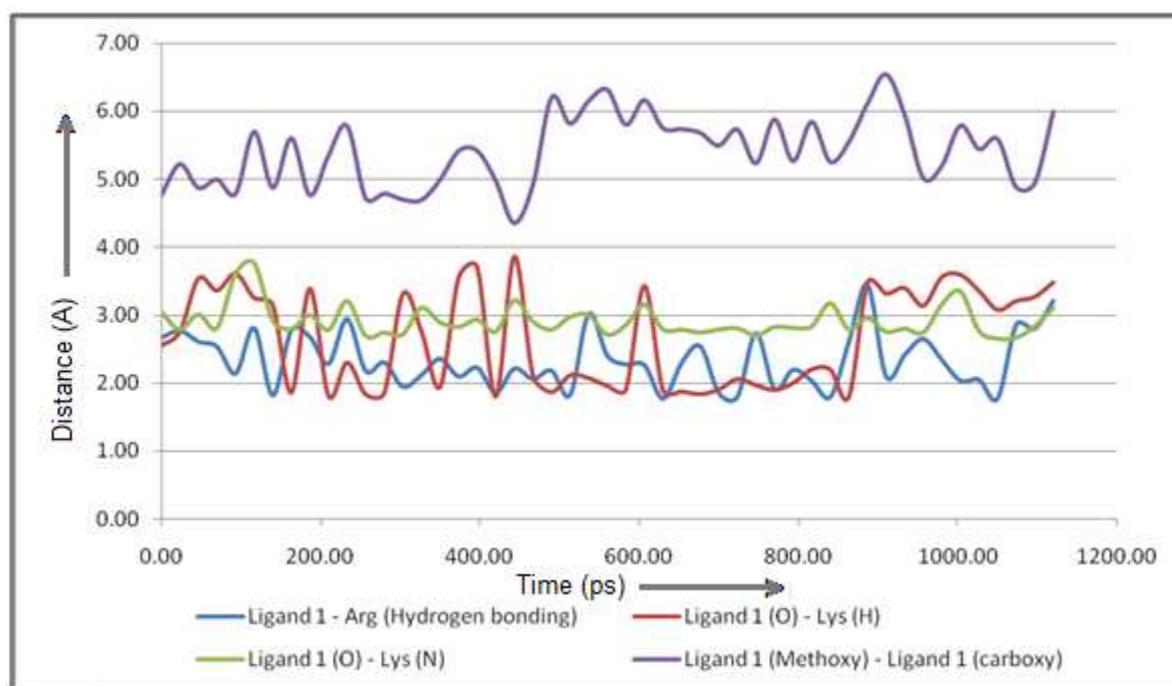
Figure 2.12: Showing the opening of the active site following the MD studies at 1120ps with the ligand-protein complexes.

The RMSD evaluation was also carried out on the MD study of the protein-TSA complex carried out at 12330ps. The RMSD was compared with the conformation obtained at 1ps as well as to the final structure as shown in graph 2.10 and was observed that the RMSD stays below 2 Å during the 1120ps of the MD study. The RMSD then gradually rises above 4 Å until the MD was performed to 8000ps after which the RMSD falls down below 3 Å. A close observation of the RMSD plot obtained after a MD study at 1120ps suggests that the RMSD was below 2 Å. The docking scores of all the ligands obtained in the phase 1 study were in accordance with the biological activity of the compounds in the organism, except ligand 7. This ligand has shown a more negative score; however when the RMSD of this ligand protein complex was evaluated (phase 2), it was found that the RMSD gradually increased beyond 3 Å and the key interactions were lost in the final confirmation.

Another interesting observation within the active site following the MD studies of the Ligand 1-protein complex was that the Ligand 1 took up a folded conformation, and this perhaps is the lowest energy conformation of the Ligand 1. When the ligand1-protein complex from phase I was set for MD study at 1120ps (scheme 2.3; Appendices I), the final structure was

found to have the similar interactions. To have a clear picture of this study the bond lengths of the key interactions were identified and the change of these bonds was observed during the MD study at 1120ps of the Ligand1-protein complex carried out in phase II study as shown in graph 2.30.

The interactions produced by the GLU 235, LYS 186 and ARG 260 were important in this model and therefore these residues were closely monitored throughout the studies. The distances between these residues and the interacting groups of the ligands (Figure 9 and Figure 2.10) were also closely monitored and the closest distance between the residues was graphically studied (Graph 2.16). The consistency over the folding pattern of the Ligand 1 (Figure 2.10 b) was also observed by monitoring the methoxy-carbonyl bond distance.



Graph 2.30: The variation of the bond length throughout the molecular dynamics studies of Ligand 1-protein complex at further 2120ps (Phase2).

It was observed that the ligand maintains the folded conformation consistently throughout the molecular dynamic studies while it maintains the interactions with the interacting residues of the model. This observation supports the fact that this folding pattern was the stable conformation of the ligand within the active site and this is mediated by the interaction of the residues caused therein.



### 2.10.3 Selecting the targeted model:

Following the MD studies of the **TSA**-protein complex, various frames were collected at various time intervals. The plot of the total kinetic energy and the total potential energy vs time interval (Graph 2.1) has shown no significant difference, however a close analysis of the plot at 1120ps (Graph 2.2) shows that there is a increase in the kinetic energy after 1120ps of the MD study. The Ramachandran plot showing the residues in the most allowed region went down from 93.5% (1120ps) to 90% (2120ps) and even less. In the frame selected at 4750ps the Ramachandran plot has shown 88.6% of the residues in the most favoured region. In an ideal model 90% or more residues should be in the most favoured region within the Ramachandran plot (Ramachandran *et al.* 1963). The orientation of the residues within the active site after MD at 1120 ps (Figure 2.11a) was also studied and compared with the orientation of the active site residues as obtained from the crystal structure (Figure 2.3). It was observed that the residues retained all interactions as observed by Okvist *et al.*, within the crystal structure of the protein, suggesting that the protein model obtained following a MD studies at 1120ps is the closest biological representation of the enzyme. Also the distances between the LYS-GLH residues and the LYS-ARG residues were altered when MD was performed at beyond 1120ps (Graph 2.31). When the active sites, obtained from the molecular dynamics studies at 2120ps and 4790ps, were carefully observed, it was found that the active sites were condensed. Although the interactions of the residues with the **TSA** were improved, it has significantly condensed the active site. These have interfered with the entrance of the ligands into the active site during the docking studies. It was observed that most of these ligands remained outside the active site. The frames obtained beyond 1120 ps following MD studies were excluded from the model. The model obtained from frame 1120ps of the MD studies has preserved all the key interactions within the active site (Okvist, M., R., *et al.* 2006). All these in together supported the frame obtained at 1120ps following the MD studies of the **TSA**-protein complex as a target model to design inhibitor compounds. A flow diagram illustrating the development of the model is shown in Appendix 1.

### 2.10.4 Docking results and discussion:

Using the targeted model obtained following a MD study on the **TSA** protein complex at 1120ps (Scheme 2.2) docking studies were carried out using the ligand (Table 2.4) with known biological activity (Table 2.8). The active sites were designed in accordance with the important interactions produced by the **TSA** (Okvist *et al.*, 2006). The ligands were docked into the active site and the docking scores were recorded (Table 2.10, Table 2.11, Table 2.12, Table 2.13 and Table 2.14). In an ideal situation the most potent ligand should score more negative and the less potent should have less negative or positive scores. Interestingly the docking score on its own was difficult to judge the possible biological activity of the ligand

due to the diversity on the ligand structures. **TSA** was a small bicyclic structure ligand and all other ligands were structurally different to **TSA**. Also only three biologically active ligands were available in this study and this was too few to calibrate the model, which were also very different from each other. Therefore in this study much emphasis has been given on the interactions produced by the key residues. Hence a series of hypotheses were established for the model and these hypotheses were used along with the docking scores to screen possible biologically active ligands. Ligand **1** has been the most active inhibitor against MtCM reported so far. Therefore in this work much effort has been given to manipulate the Ligand **1** with various substitutions to design inhibitor ligands that however consider the model hypotheses.

#### 2.10.5 Structural activity studies of Ligand 1:

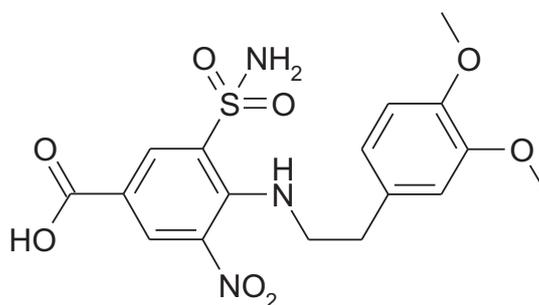


Figure 2.14: The structure of the Ligand **1**

On changing the sulphonamide group to an amide group as in Ligand **20** the interaction of the ligand to form a hydrogen bonding with the Lys 186 was lost. The interaction with the Arg 260 and Glu 236 was also changed. Again on changing the carboxylic group into nitro group as in Ligand **16**, **27** and **29** the folding pattern of the ligand was hindered and the docking score went less negative. Changing the nitro group into carboxyl group as in Ligand **21**, apart from losing the folding pattern, the hydrogen bonding with the Lys 186 was also lost. On changing the methoxy group into hydroxyl Ligand **17**, the ligand still maintained interactions with the residues thereby folding up. The distance between the Lys186 and the sulphonamide moiety increased. Perhaps the presence of a bulky group such as the methoxy groups facilitate the ligand to move toward the Lys186 forming a hydrogen bonding. The nitro groups permit the rational design of numerous potent and selective inhibitors of enzyme and are good analogues of the carbonyl groups (Alston, T. 1983). Thus on replacing the Sulphonamide part of the ligand with the nitro as in Ligand **45**, **50 – 57**, **60**, **61**, **64**, **66**, **67**, **69**, **71** and **72** preserves all the desirable interactions within the active site. This is perhaps due to the fact that although the Sulphonamide group was large compared to the nitro, both of them had similar electronic effect. On replacing the methoxy group with the

hydrogen, the ligand still followed the model hypothesis and preserved the important interactions. Incorporating one more methoxy group into the benzyl ring as in Ligand **30**, prevents the moiety to completely get within the active site. Substitution of nitro group to the carboxylic group reduces the inhibitory property of the ligand, however with the methoxy group in 3<sup>rd</sup> and 5<sup>th</sup> position on the ring B (Figure 2.8) as in Ligand **26** and the hydroxyl group on the 3<sup>rd</sup> and 4<sup>th</sup> position as in Ligand **29** moderately retains the inhibitory property of the ligand. This is probably due to the intra-molecular hydrogen bonding facilitating the folding fashion of the ligand. Another interesting observation was Ligand **75**. This ligand was devoid of any substitution on ring B (Figure 2.8), and however this ligand had folded within the active site. Replacing the ethylamine linkage with a methyl hydrazinyl linkage as in Ligand **37-40** and Ligand **64-69**, some of the ligands particularly the one with di-methoxy substitution (Ligand **38**) had a good docking score and was also successful in preserving some interaction within the active site. This ligand was also observed in adopting a partially folded conformation within the active site. Also substituting a -N-N-C=O group in place of the ethyl amine linkage as in Ligand **62**, this led to the adoption of a partially folded conformation and a less negative docking score.

Based on the docking score (Table 2.14) and the hypothesis as framed in section 2.10, all the ligands laid out in Table 2.9, Table 2.10, Table 2.11, Table 2.12 and Table 2.13, were ranked as shown in the table below.

Ligand	Ranking	Ligand	Ranking	Ligand	Ranking
Ligand <b>17</b>	A	Ligand <b>62</b>	A	Ligand <b>83</b>	B
Ligand <b>23</b>	A	Ligand <b>63</b>	A	Ligand <b>81</b>	B
Ligand <b>24</b>	A	Ligand <b>64</b>	A	Ligand <b>82</b>	B
Ligand <b>27</b>	A	Ligand <b>65</b>	A	Ligand <b>86</b>	C
Ligand <b>34</b>	A	Ligand <b>66</b>	A	Ligand <b>87</b>	C
Ligand <b>43</b>	A	Ligand <b>67</b>	A	Ligand <b>88</b>	C
Ligand <b>44</b>	A	Ligand <b>68</b>	A	Ligand <b>89</b>	C
Ligand <b>45</b>	A	Ligand <b>69</b>	A	Ligand <b>80</b>	C
Ligand <b>50</b>	A	Ligand <b>70</b>	A	Ligand <b>76</b>	C
Ligand <b>51</b>	A	Ligand <b>71</b>	A	Ligand <b>77</b>	C
Ligand <b>52</b>	A	Ligand <b>72</b>	A	Ligand <b>90</b>	C
Ligand <b>53</b>	A	Ligand <b>73</b>	A	Ligand <b>91</b>	C
Ligand <b>54</b>	A	Ligand <b>74</b>	A	Ligand <b>92</b>	C
Ligand <b>55</b>	A	Ligand <b>75</b>	A	Ligand <b>93</b>	C
Ligand <b>56</b>	A	Ligand <b>61</b>	A	Ligand <b>84</b>	C

Ligand	Ranking	Ligand	Ranking	Ligand	Ranking
Ligand 57	A	Ligand 78	B	Ligand 85	C
Ligand 59	A	Ligand 79	B		
Ligand 60	A	Ligand 58	B		

Table 2.17: Showing the ranking order of the ligands as determined by the model. A denotes that the ligands follow the model hypothesis in all aspects, B denotes that the ligands requires further modification and Ligands ranking C are rejected

### 2.10.6 Fragment based approach for Ligand 1:

In an attempt to study the interactions of the Ligand 1 within the active site a fragment based approach was also employed. The fragments of Ligand 1 were docked separately and then the orientation and interactions of the fragments (Ligand 88 and Ligand 89) were studied at the active site.

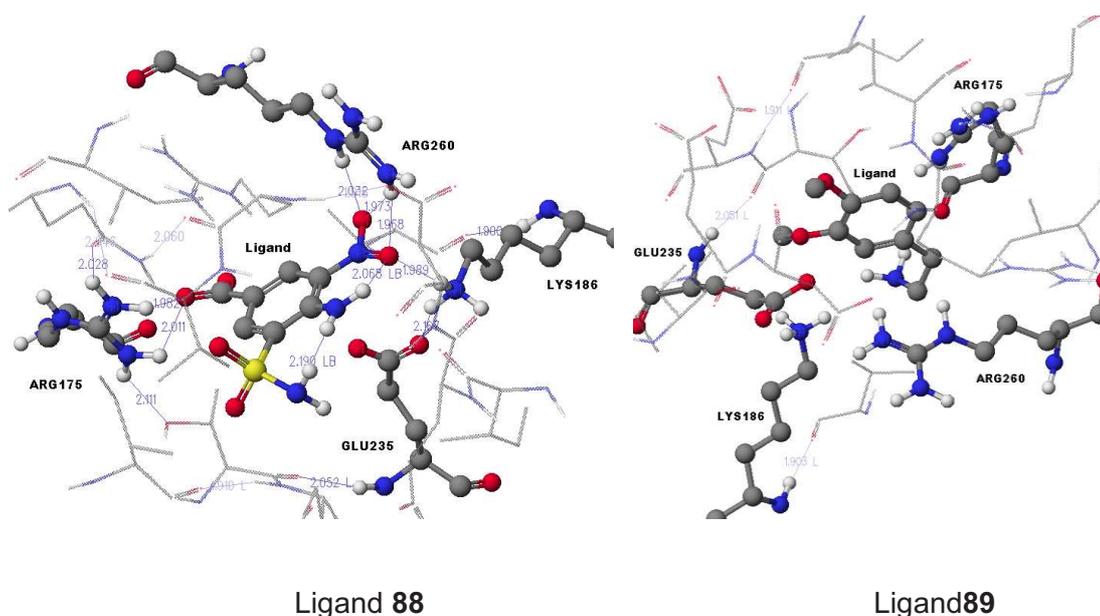


Figure 2.15: The interactions produced when the Ligand 1 was fragmented and docked into the active site and the fragments were docked separately as Ligand 88 and Ligand 89.

It was observed that 5-nitro-3-sulphomoyl-4-amino benzoic acid (Ligand 88) exhibited the most interactions within the active site, forming hydrogen bonds with Arg 175, Arg 260, Lys 186 and Glu 235. When Ligand 88 and 89 were used for docking study, the scores were found more negative than the **TSA**. Both the ligands however left plenty of space within the active site pocket. An attempt was also made to design various analogues of ligand 88 by modifying the side chains of the ligand (Ligand 84, 85, 86 and 87) as shown in Table 2.12 This was essentially to increase the interaction with the neighbouring residues. On increasing

the side chains as in Ligand **87**, the ligand tried to adopt a folded conformation. The methoxy groups in Ligand **1** have shown good interaction within the active site (Figure 2.10.b), thus in an approach to incorporate the methoxy group in the Ligand **88**, Ligand **86** was used in the study. These ligands have tried to take up a folding conformation and therefore to increase the flexibility saturated ring systems, containing the key interacting groups were used as in Ligand **84**. This ligand took up a highly strained conformation within the active site. In all the ligands used in this investigation the nitro and the sulphonamide groups produced interesting interaction with the Arg, Lys and Glu residues at the active site. Ligand **88** has shown the most successful docking score and the best interaction with the residues when compared with the other analogues of Ligand **88** shown in Table 2.12. These therefore have shown that modification in this fragment of the ligand was limited and the sulphonamide, nitro and the carboxylic in those positions of the ring were important for the ligand within the active site.

Ligand **89** did not show any interesting interaction with the residues in the active site (Figure 2.15). Neither dimethoxy phenethylamine part of Ligand **1** has shown any key interaction with the active site residues of the target model, apart from facilitating the folding conformation of Ligand **1**. Therefore the 5-nitro-3-sulphomoyl benzoic acid probably forms the lipophilic part of the Ligand **1** creating lipophilic contact with the Arg 260, Lys 175 and Glu 235 whereas the dimethoxy phenethylamine makes the hydrophobic contacts with the residues (Figure 2.14). Hence dimethoxy phenethylamine is essential in designing analogues of Ligand **1** and therefore in this model attempt has been made to create various substitution on the dimethoxyphenyl part of the Ligand **1**; these had produced ligands which were similar in properties.

#### 2.10.7 Structural activity studies of Ligand 2:

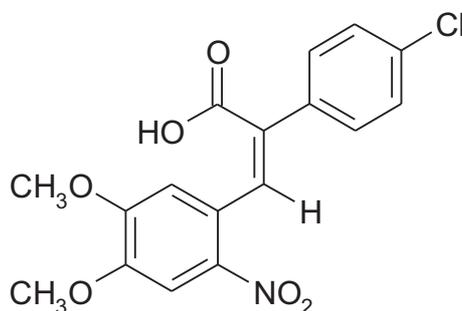


Figure 2.16: The structure of Ligand **2**

Agrawal *et al.*, in their research have observed that chlorophenyl part of the moiety of the Ligand **2** remained outside the active site, on docking this ligand into their model (Agrawal, Kumar *et al.* 2007). In the attempt to dock Ligand **2** in the active site, three out of five times the chlorophenyl part of the moiety remained within the active site. The nitro group of the ligand interacts with the Arg 260 in a similar fashion as was observed by Agarwal *et al.* The chlorine within the active site has shown some interaction with Thr 231 and Glu 235 as shown in the Figure 2.10(c). On docking the Ligand **2** into the active site, the ligand did not fold within the active site, however on removing the benzyl chloride part of the ligand, it has shown some folding pattern within the active site resembling the **TSA** like structure (Figure 2.18).

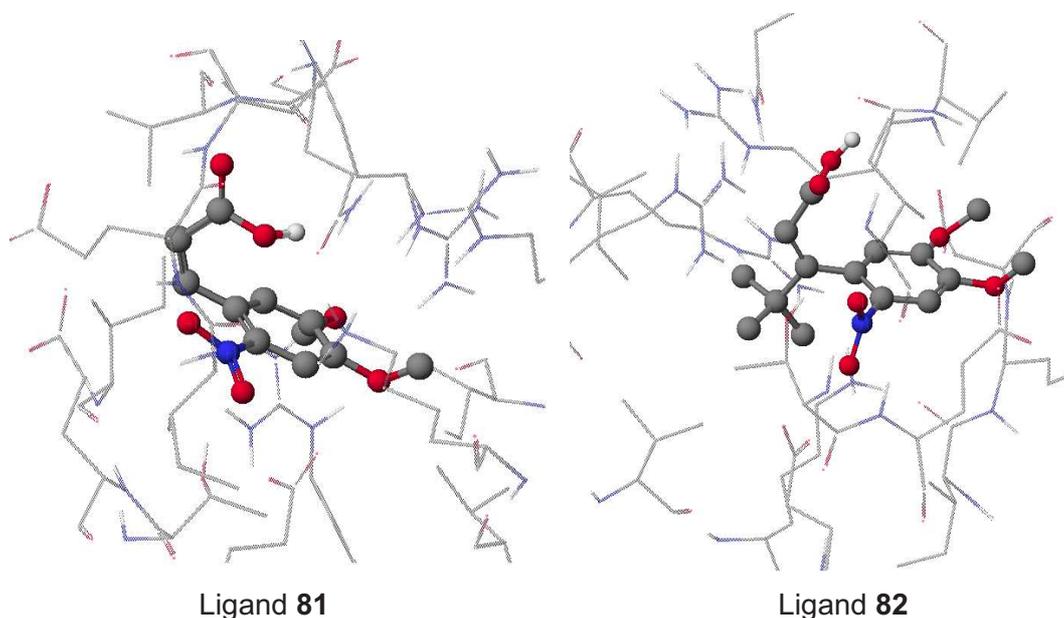
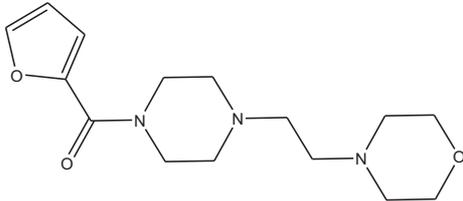
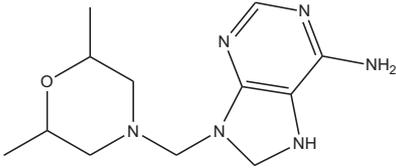
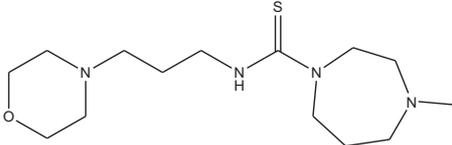
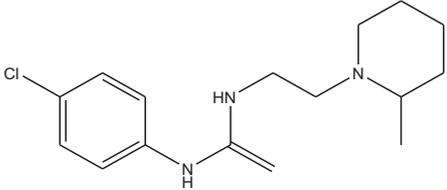
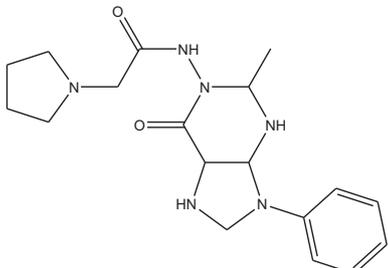
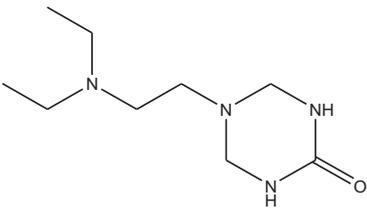
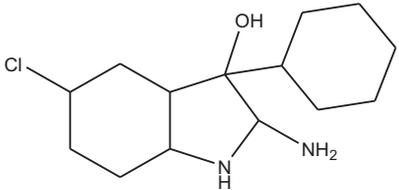


Figure 2.17: Showing the orientation of the Ligand **81** (fragment of Ligand **2**) and Ligand **82** (a modification of Ligand **84**) within the active site.

Introducing a bulky group into the moiety the ligand (Ligand **81**, **82** and **83**) has facilitated the folding pattern of the small ligand to lean the carboxylic group toward the ring (Figure 2.17).

#### 2.10.8 Docking studies of the commercially available ligands from Maybridge:

A library of ligands from the Maybridge Company was screened; a set of 260 ligands were selected for docking studies in this investigation. All these ligands have been docked three times and the average score have been shown in Table 2.15. The docked complexes were then investigated for interesting interactions. Of all the ligands used, the following ligands shown in Table 2.18 have shown good score (more negative). However, only Ligand **113** and Ligand **124** were ranked well by the model hypothesis.

SL No	Ligands	PMF Score Kcal/mole	Rank
Ligand 113		-646.08	A
Ligand 124		-631.25	A
Ligand 201		-631.72	A
Ligand 142		-577.53	A
Ligand 125		-546.48	C
Ligand 183		-530.58	C
Ligand 123		-543.32	C

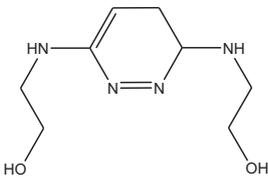
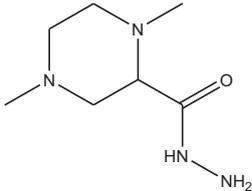
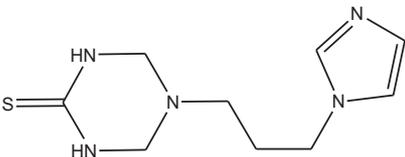
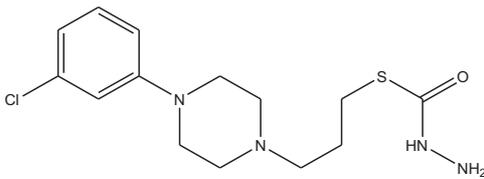
SL No	Ligands	PMF Score Kcal/mole	Rank
Ligand 160		-501.99	C
Ligand 194		-508.73	C
Ligand 197		-510.68	C
Ligand 204		-462	C

Table 2.18: Showing the list of the Maybridge compounds with rank

While screening these databases of ligands, the ligands were first ranked by docking scores and then the ligands were observed for their ability to get into the active site. Despite the more negative docking score some of these ligands were found to be on the surface of the active site. This is due to the presence of plenty of interacting groups on these ligands producing more negative score. These ligands were also excluded from the study. The final set of 11 ligands (Table 2.18) was screened using the model hypothesis and ranked accordingly. Only Ligand **113** and **124** were ranked well by the model apart from the docking scores, these ligands have shown good interactions with the Lys186, Arg260 and Glu235. These ligands have produced more negative scores and had also shown good interaction with the Lys186, Arg260 and Glu235 in a similar fashion as **TSA** and Ligand **1** (Figure 2.4). These ligands were also observed with folding up toward itself within the active site. This suggests that the folding property of the ligands within the active site of the model is not just a feature observed with the ligands designed in the study (Table 2.9, 2.10, 2.11, 2.12 and

2.13). Perhaps it is a key feature that should not be ruled out while considering this model to design inhibitor compounds.

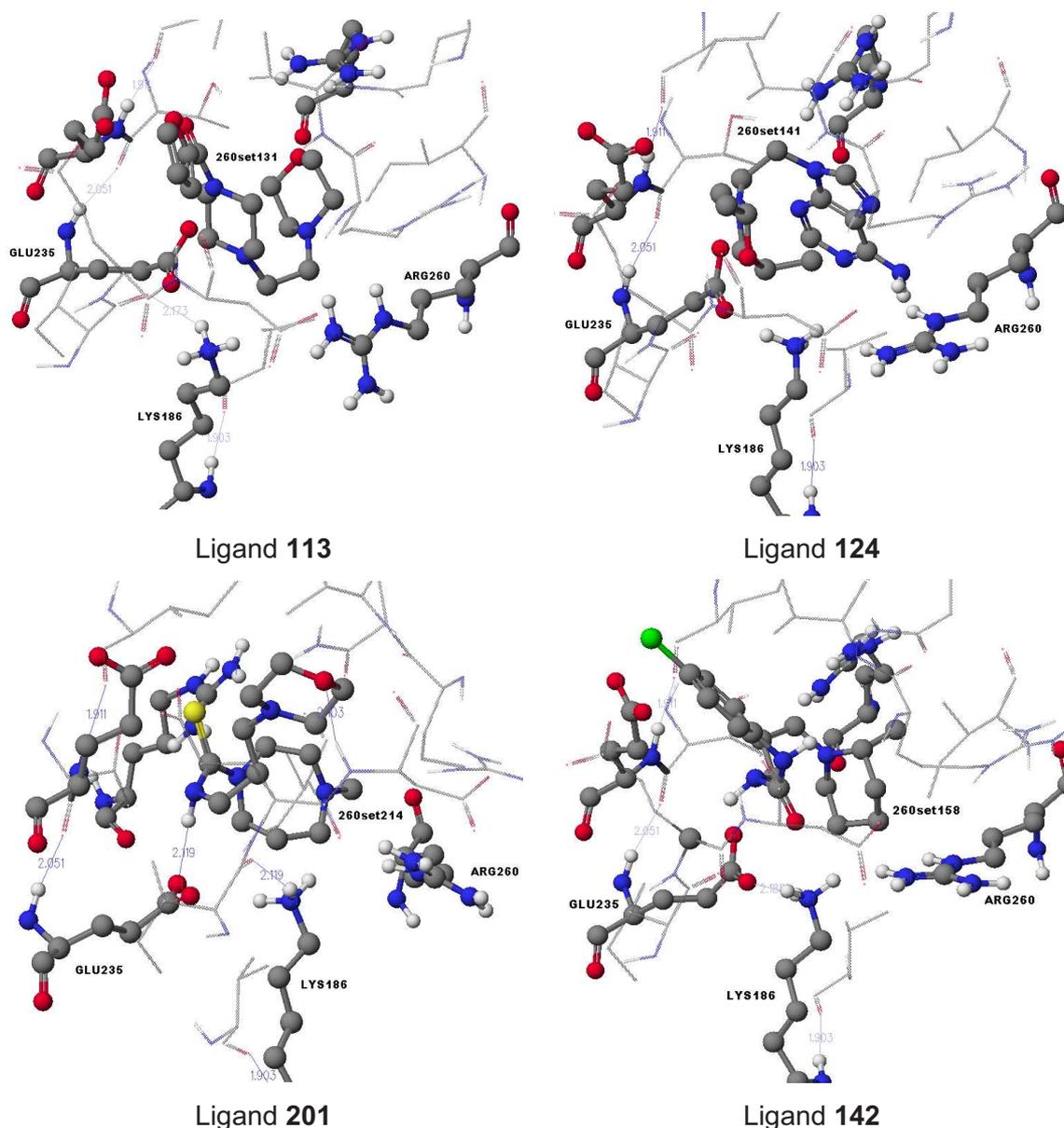


Figure 2.18: The interactions produced by some of the commercially available compounds when docked into the active site of the model obtained following MD study of the **TSA**-protein complex at 1120ps (Scheme 2.2).

Most of these ligands (Table 2.18) have more than one ring system with the saturated heterocyclic ring system. These ligands were relatively large, when compared to the structure of **TSA**, and therefore have a large surface area in general. However when these ligands were docked within the active site, they folded themselves to form a highly strained

boat conformation within the active site. This pattern of orientation favoured the interaction of the ligands within the active site (figure 2.18). The ligand 125, **183**, **123**, **120**, **194**, **197** and **204**, all went successfully into the active site on docking studies. However these ligands were unable to interact with the key residues in the active site.

### **2.11 Conclusions:**

A protein model of the enzyme chorismate mutase from *Mycobacterium tuberculosis* was prepared through a combination of molecular modelling and docking studies. A range of ligands with known biological activities were used to study the interactions and a final set of hypotheses were established specific to the model. A library of ligands was designed which were then screened and ranked based on the model hypothesis. The ligands which were ranked as good ligands were then considered for synthesis and biological screening. This model was also used to screen a database of commercially available ligands represented by a smaller set chosen by a pharmacophore based diversity analysis. Ligands **113** and **124** in particular showed interesting interactions within the active site and therefore these scaffolds represent important areas for future research to discover inhibitor compounds against the mycobacterium chorismate mutase.

## CHAPTER 3

SUBSTITUTED 3-NITRO 5-SULPHAMOYL BENZOIC ACIDS:  
DERIVATIVES AND RELATED CHEMISTRY**3.1 Aim:**

The aim of this work is to prepare a selection of the compounds which have ranked well following the docking studies in chapter 2 and to screen for biological activity particularly against *Mycobacterium tuberculosis*.

**3.2 Background:**

The protein model appropriate for the enzyme from chorismate mutase *M. tuberculosis* was prepared, which can selectively rank ligands toward the best possible biological activity. The 4-(3,4-dimethoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid which was found to be very active against the enzyme of the pathogen (Agrawal, Kumar *et al.*, 2007) has been used throughout the studies to identify the interactions with the residues in the model. Using the model over 350 ligands have been screened by docking studies and a library of compounds has been generated that fit appropriately within the active site. Ligands which are similar to the 3-nitro 5-sulphamoyl benzoic acids have shown the interesting interaction and have also scored well in the docking studies. These compounds were clustered as well ranked by the model and were selected for further synthesis.

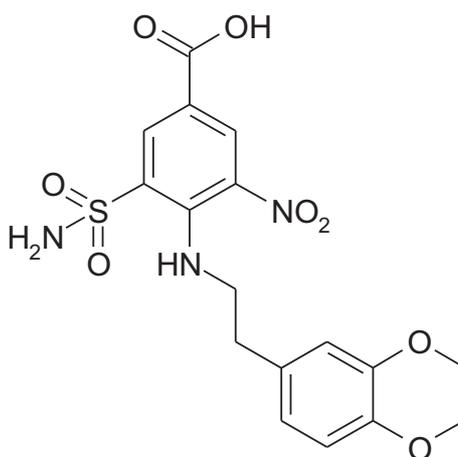
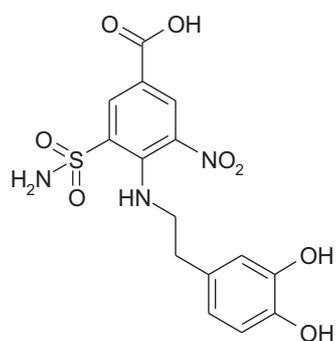
Compound **3.1**

Figure 3.1: Structure of 4-(3,4-dimethoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid (Compound **3.1**)

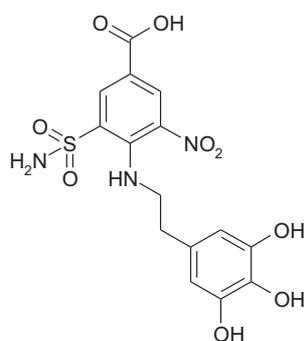
Interestingly Compound **3.1** (4-(3,4-dimethoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid) which has diuretic properties (Kansal, Kar *et al.* 1979) was found to be very active against the chorismate mutase from *Mycobacterium tuberculosis* (Agrawal, Kumar *et al.*, 2007). Although Kansal *et al.*, reported the diuretic activity of Compound **3.1**, they did not provide detailed information supporting the synthesis and characterisation of this compound. Therefore the work in this thesis has developed the preparation, isolation and characterisation of substituted 3-nitro, 5-sulphamoyl benzoic acids and related compounds.

### 3.3 Substituted 3-sulphamoyl-5-nitro benzoic acid and related compounds:

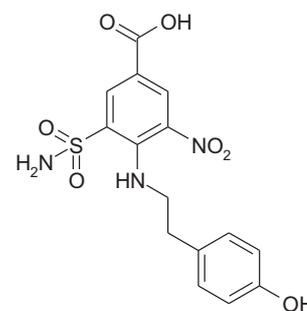
The target compounds arising from the in silico screenings are set out in Figure 3.2.



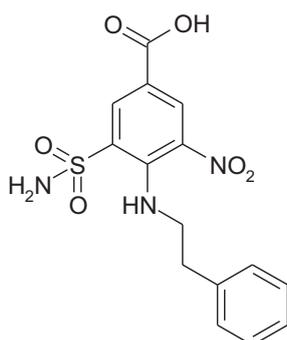
Compound **3.2**



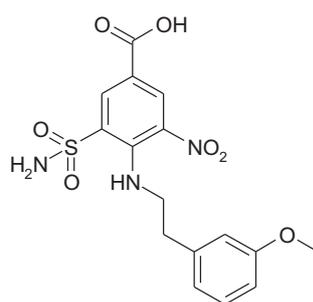
Compound **3.3**



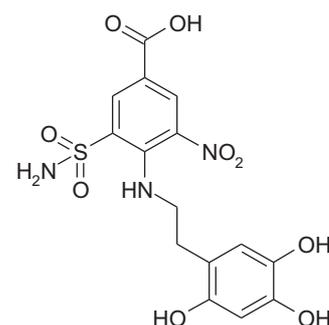
Compound **3.4**



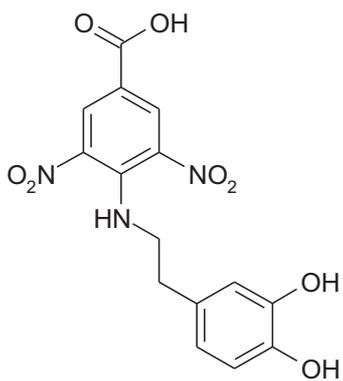
Compound **3.5**



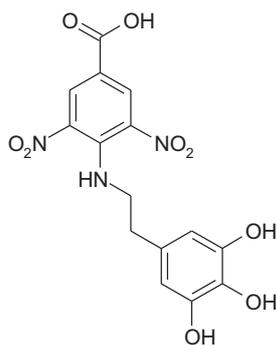
Compound **3.6**



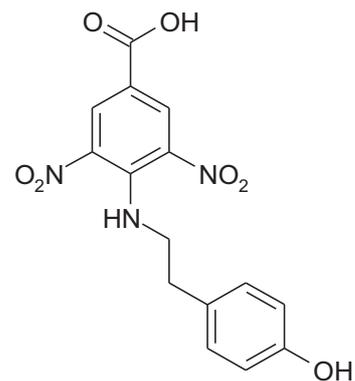
Compound **3.7**



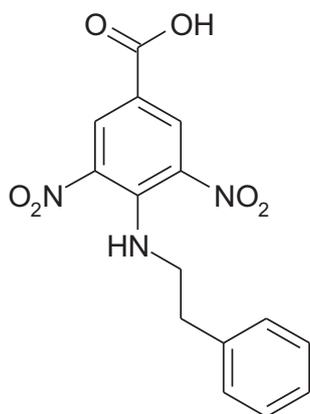
Compound 3.8



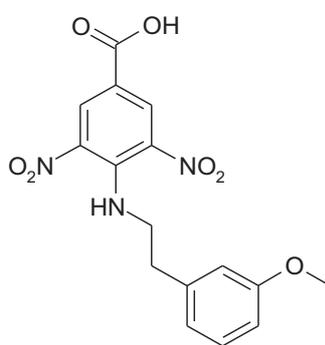
Compound 3.9



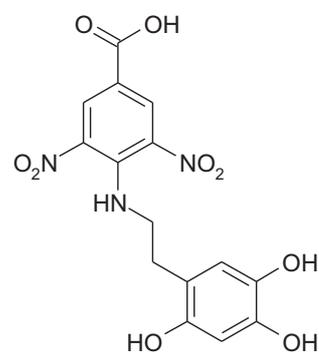
Compound 3.10



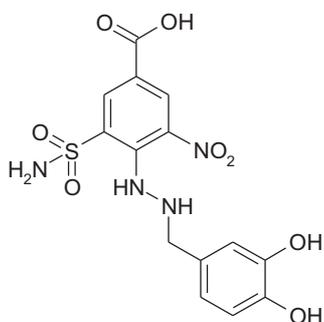
Compound 3.11



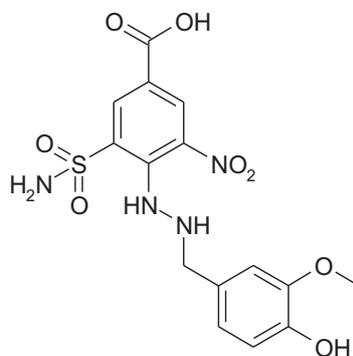
Compound 3.12



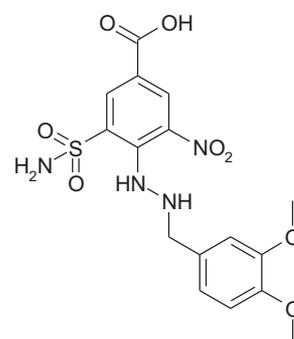
Compound 3.13



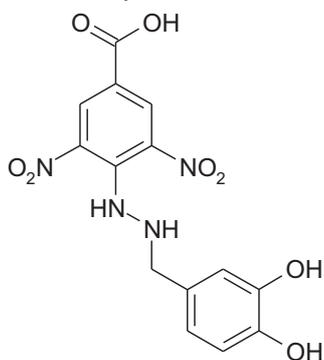
Compound 3.14



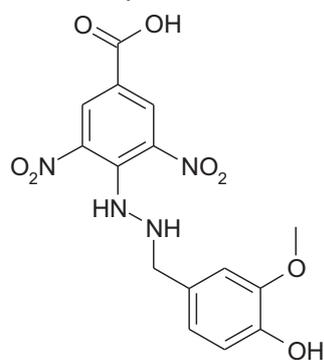
Compound 3.15



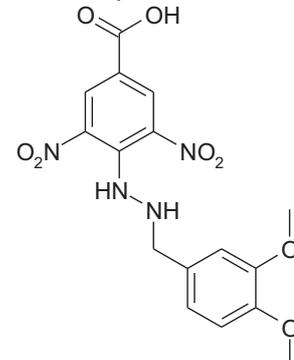
Compound 3.16



Compound 3.17



Compound 3.18



Compound 3.19

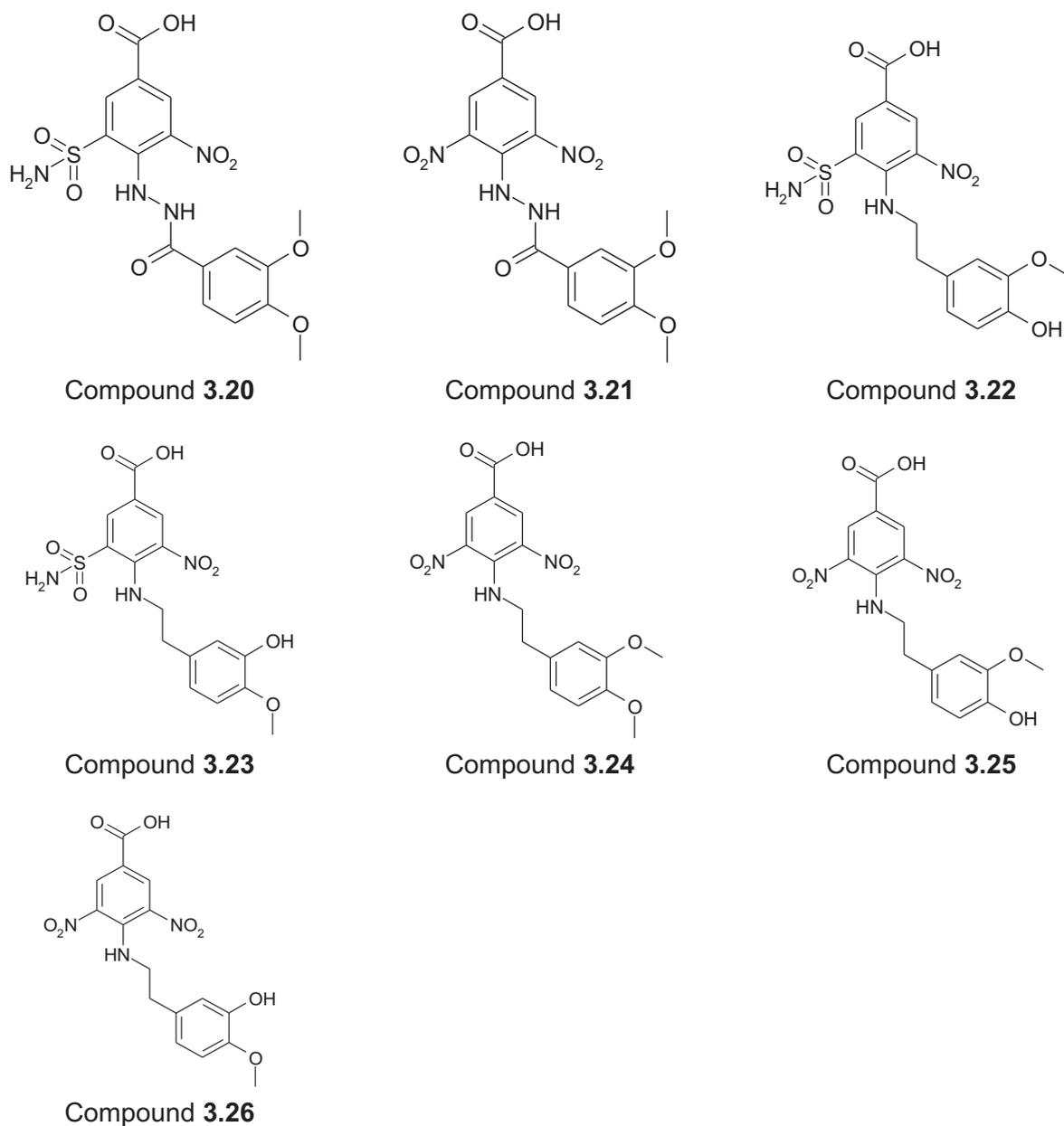


Figure 3.2: Target compounds judged to have potential activity against chorismate mutase from *M. tuberculosis*.

All the compounds shown in the Figure 3.2 have been clustered into three main analogues of the biologically active ligand Type A, Type B and Type C.

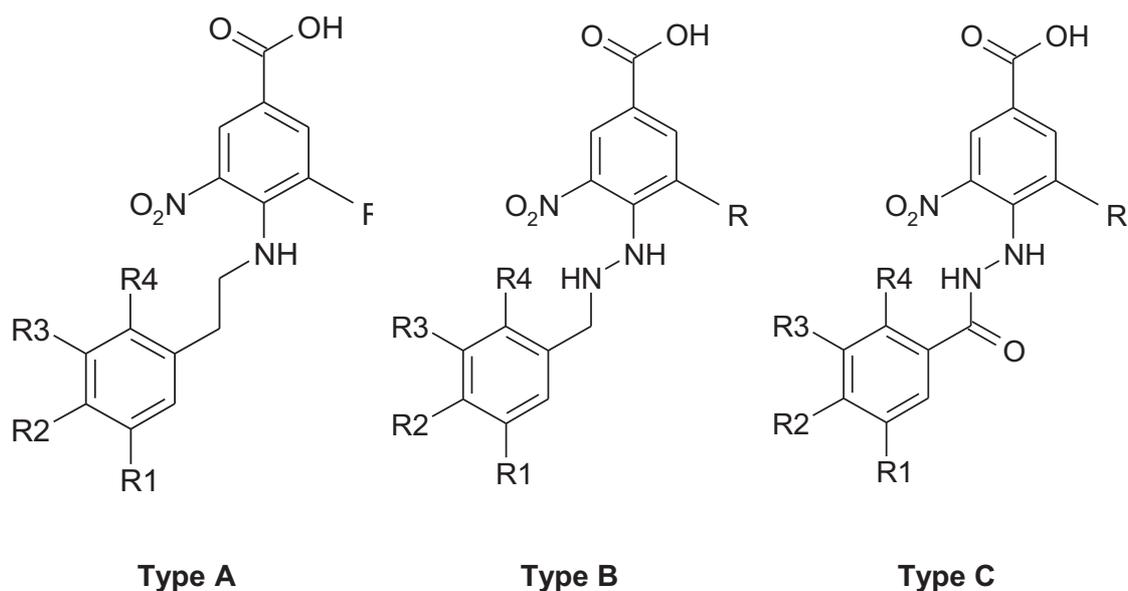


Figure 3.3: The three main categories of the compounds used in this investigation.

#### 3.4 Screening synthesisable compounds:

Unfortunately not all the compounds following Figure 3.2 were possible to synthesise. This was due to the potential use of starting materials which were classed as restricted drugs under Schedule 2 Class A by the UK Home Office. Therefore only the compounds which were free of using any controlled drugs in their synthesis were selected for synthesising and so in this study the synthesis of Compound **3.22**, Compound **3.23**, Compound **3.24**, Compound **3.25** and compound **3.26** has been omitted. The remainder analogues of the compounds of Type A (Figure 3.3) were prepared by treating the amines with the corresponding acids (Jackman *et al.*, 1962). These compounds were prepared by heating the phenethyl amine with the corresponding benzoic acid in presence of DIEA in butanol at reflux as shown in Figure 3.4 (Vlaskina *et al.*, 2004). The final compound was then isolated by first converting it into the BOC substituted derivatives (Pope. *et al.*, 1988). It was then partitioned between the ammonium chloride and ethylacetate layer. The product was then obtained from the organic phase and then BOC was removed by a strong acid such as trifluoroacetic acid (England.E, A, *et al.*, 2004) or hydrochloric acid in methanol (Williams. R, *et al.*, 2003).

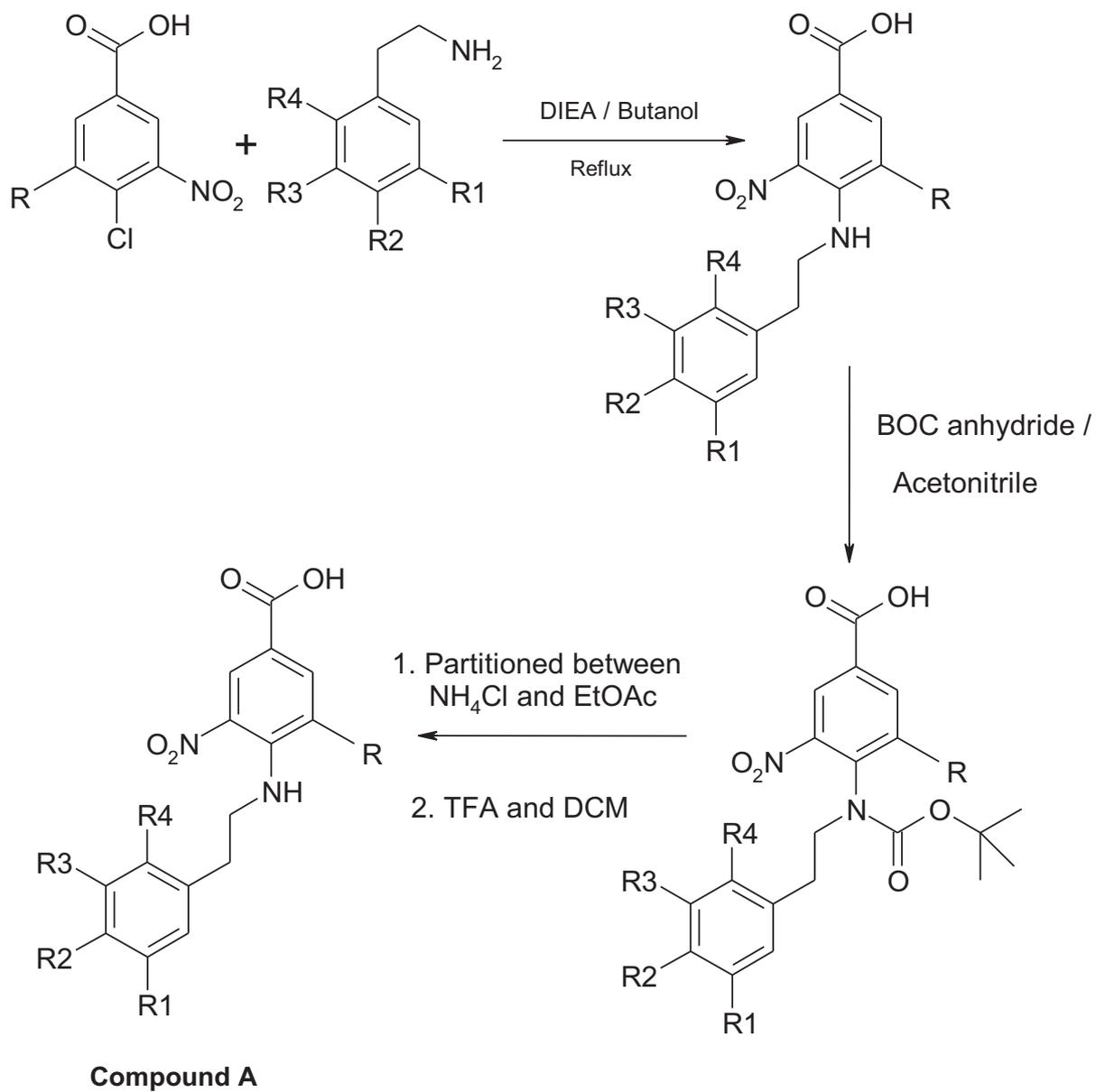


Figure 3.4: The general synthesis of the substituted 3-sulphamoyl-5-nitro benzoic acid derivatives related to Type A compounds shown in Table 3.1.

Compound	R	R1	R2	R3	R4	Yield (%)
3.2	SO <sub>2</sub> NH <sub>2</sub>	OH	OH	H	H	44
3.3	SO <sub>2</sub> NH <sub>2</sub>	OH	OH	OH	H	48
3.4	SO <sub>2</sub> NH <sub>2</sub>	H	OH	H	H	72
3.5	SO <sub>2</sub> NH <sub>2</sub>	H	H	H	H	46
3.6	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	H	H	H	40
3.7	SO <sub>2</sub> NH <sub>2</sub>	OH	OH	H	OH	48
3.8	NO <sub>2</sub>	OH	OH	H	H	34
3.9	NO <sub>2</sub>	OH	OH	OH	H	32
3.10	NO <sub>2</sub>	H	OH	H	H	61
3.11	NO <sub>2</sub>	H	H	H	H	59
3.12	NO <sub>2</sub>	OCH <sub>3</sub>	H	H	H	30
3.13	NO <sub>2</sub>	OH	OH	H	OH	10

Table 3.1: Compounds of Type A.

Similarly the analogues of the compounds from Type B were prepared by adding the acid and DIEA mixture in butanol to the hot benzyl-hydrazine solution in butanol and heating at reflux overnight (Figure 3.5).

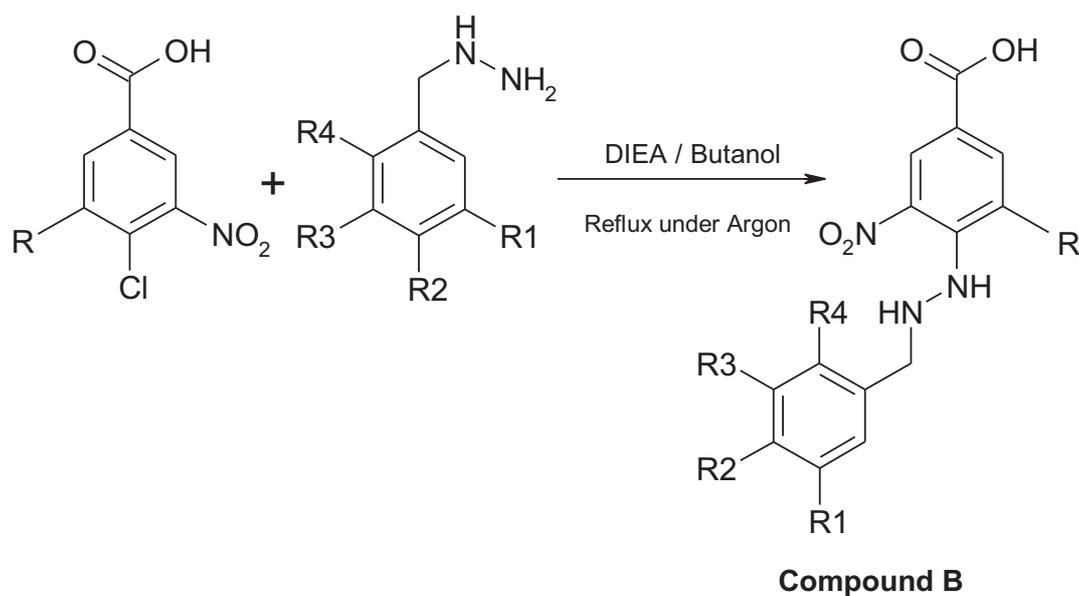


Figure 3.5: The general synthesis of the substituted 3-sulphamoyl-5-nitro benzoic acid derivatives related to Type B compounds shown in Table 3.1, where R and R1 are the substituents within the ring.

Compound	R	R1	R2	R3	R4
<b>3.14</b>	SO <sub>2</sub> NH <sub>2</sub>	OH	OH	H	H
<b>3.15</b>	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	OH	H	H
<b>3.16</b>	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H
<b>3.17</b>	NO <sub>2</sub>	OH	OH	H	H
<b>3.18</b>	NO <sub>2</sub>	OCH <sub>3</sub>	OH	H	H
<b>3.19</b>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H

Table 3.2: Analogues of compounds of Type B.

Analogues of compound Type C were also prepared similar to the analogues of Type A and Type B, however these reactions requires anhydrous conditions and therefore were carried under an argon atmosphere.

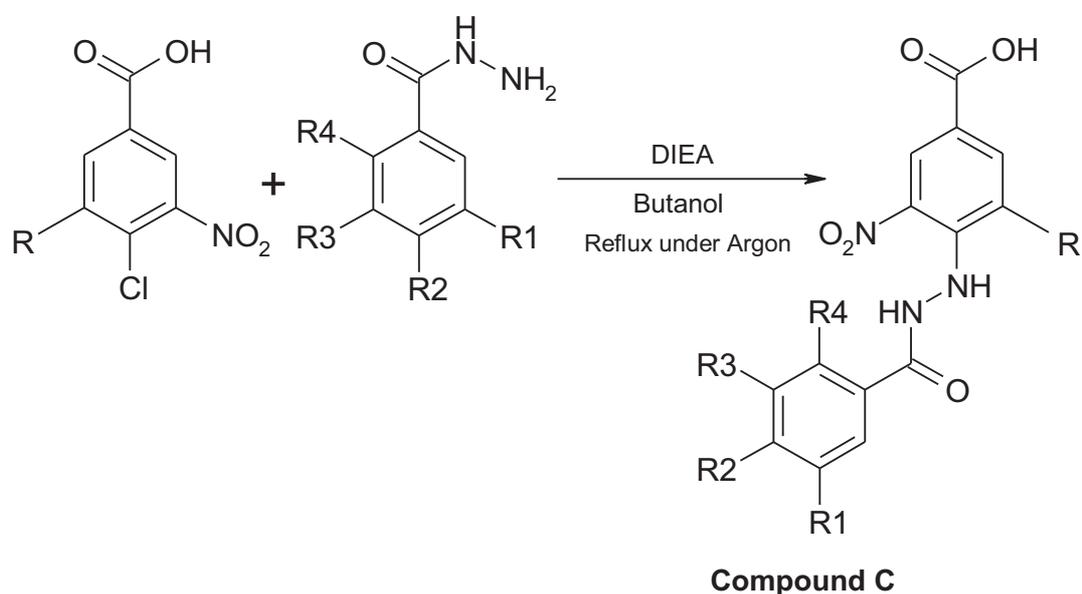


Figure 3.6: The general synthesis of the substituted 3-sulphamoyl-5-nitro benzoic acid derivatives related to Type C compounds shown in Table 3.1.

Compound	R	R1	R2	R3	R4
<b>3.20</b>	SO <sub>2</sub> NH <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H
<b>3.21</b>	NO <sub>2</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H

Table 3.3: Analogues of compounds of Type C.

### 3.5 Varying the benzoic acids moiety:

The benzoic acid part was modified in various positions. The sulphonamide group in particular was manipulated to nitro and N-methyl sulphonamide group and these modifications have shown some interesting interactions within the active site of the protein model.

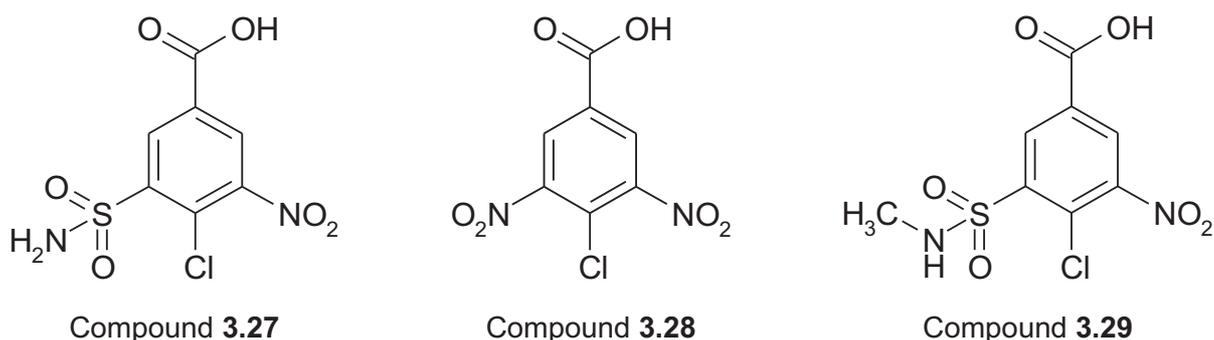


Figure 3.7: The different types of benzoic acids used in this study.

Only the 3,5-*di*-nitro-4-chlorobenzoic acid (Compound **3.28**) was commercially available. The 3-sulphamoyl-5-nitro-4-chlorobenzoic acid (Compound **3.27**) was prepared by the chlorosulfonation (Masaki, Mizuno *et al.*, 2003) of the 4-chlorobenzoic acid obtained from Sigma-Aldrich using chlorosulfonic acid in concentrated sulphuric acid to prepare the 3-chlorosulphonyl-4-chloro benzoic acid (Compound **3.30**) at 140 °C. It was then set for nitration (Palumbo, Napolitano *et al.*, 2002) using fuming nitric acid at 90 °C to prepare chlorosulphonyl-5-nitro 4-chlorobenzoic acid (Compound **3.31**) which was then finally treated with concentrated ammonia solution ( Ghadam A, K, *et al.*, 2005) to give Compound **3.27** (Kansal, Kar *et al.*, 1979). The final compound was collected by removing the excess

ammonia using a high vacuum pump and thereby adjusting the pH to 6. The schematic representation of these stages has been illustrated in the Figure 3.8 below.

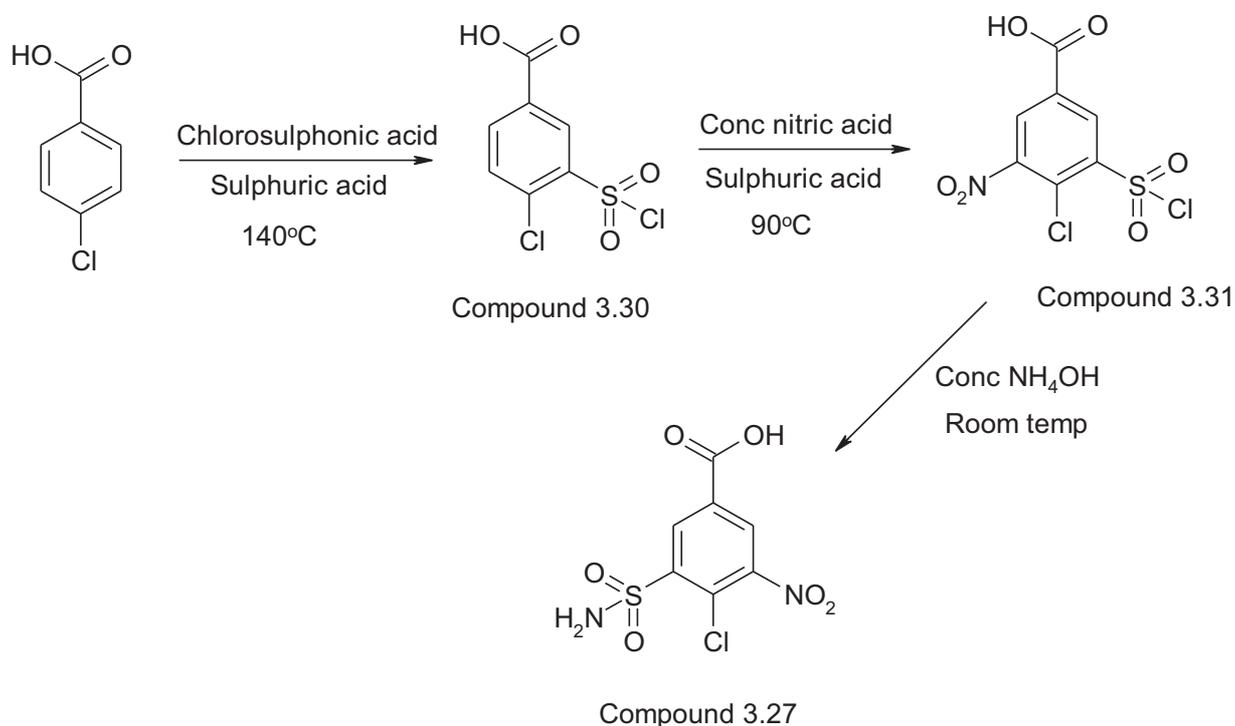


Figure 3.8: The stages involved in the synthesis of 3-sulphamoyl-5-nitro 4-chloro benzoic acid (Compound 3.27).

Compound **3.31** was used as an intermediate for the synthesis of the analogues of Type A, Type B and Type C. Compound **3.29** was prepared by treating 3-chlorosulphamoyl-5-nitro 4-chloro benzoic with N-methylamine.

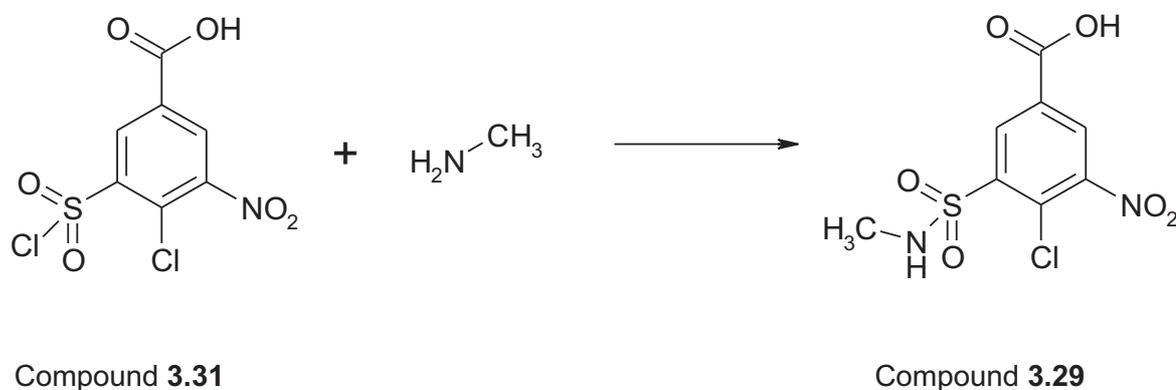


Figure 3.9: The synthesis of 3-(N-methyl sulphonyl)-5-nitro-4-chloro benzoic acid from 3-chlorosulphamoyl-5-nitro 4-chloro benzoic acid.

### 3.6 Varying the phenylamines:

All the phenylamines used in this section were commercially obtained from Sigma. The phenethylamines used in the experiment are given in Figure 3.10.

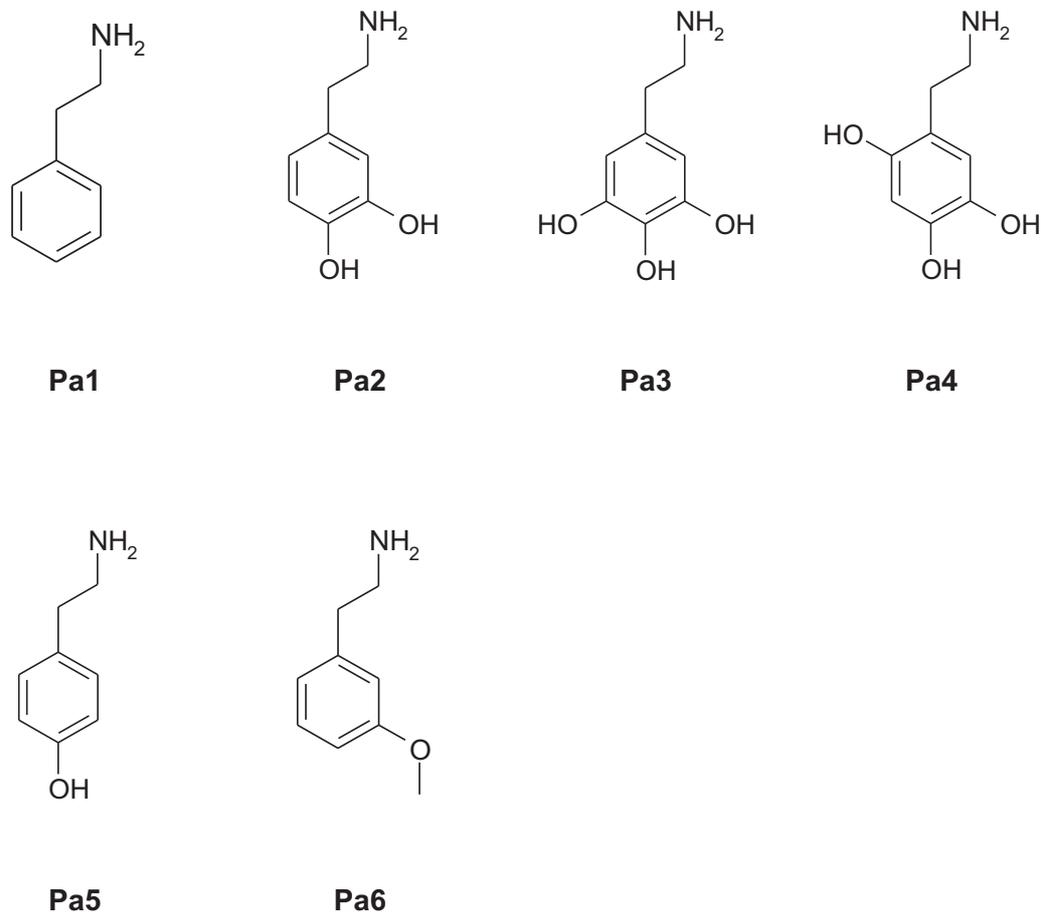


Figure 3.10: The list of various phenylamines used in the study.

### 3.7 Varying the benzylhydrazine moiety:

All the benzyl-hydrazines used in this experiment have been prepared by slowly adding the aldehydes into a mixture of excess hydrazine in ethanol heated at 60 °C with constant stirring (Szmant *et al.*, 1959). The resulting hydrazones were hydrogenated using palladium charcoal at 110 °C and 10 psi for 24 hours (Daeniker *et al.*, 1973) as shown in Figure 3.5. The benzyl-hydrazines thus obtained were used in the synthesis of the analogues of Type C and Type D.

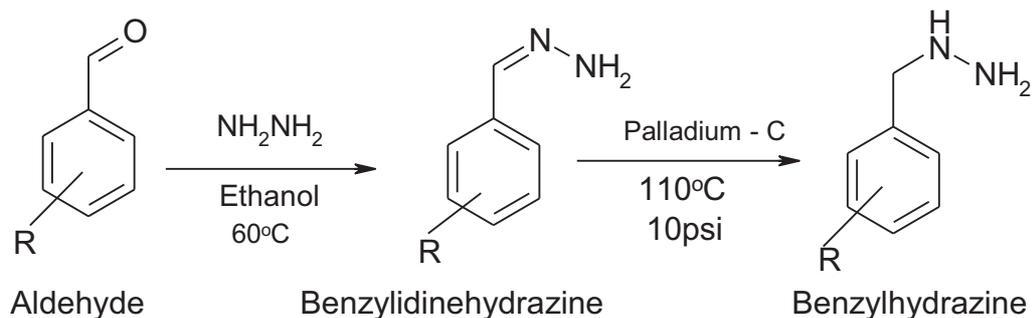


Figure 3.11: The stages involved in the synthesis of the benzylidene-hydrazone from the aldehyde.

The benzyl-hydrazines produced in this study are shown in Figure 3.11.

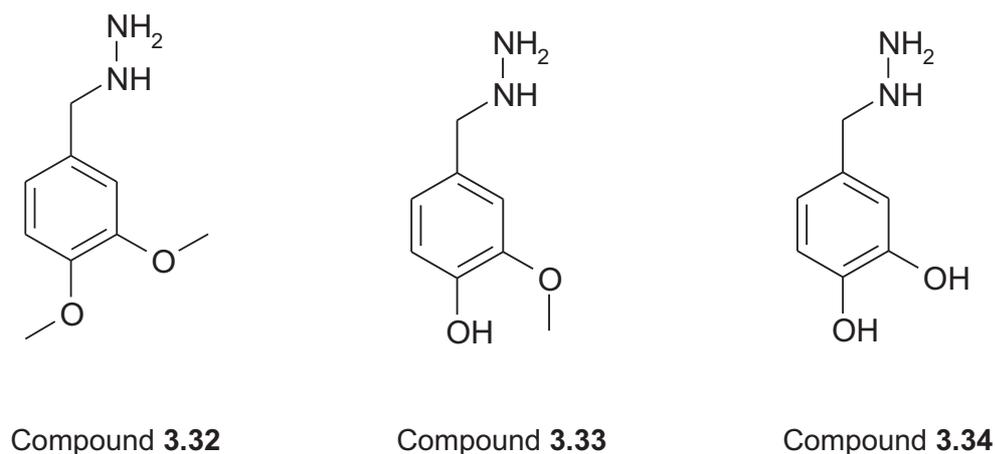


Figure 3.12: The list of benzyl hydrazines used to prepare the analogues of Type C and Type D.

### 3.8 Reactivity of the substituted benzoic acids:

The benzoic acids used in this study had a wide range of reactivity. Compound **3.28** and Compound **3.27** have similar extent of reactivity. The nitro group and the sulphonamide groups are both electronegative in nature although the extent of electro-negativity is more in case of nitro and therefore the Cl attached to the acid departs more easily in case of the di-nitro benzoic acid than in the 5-nitro, 3-sulphamoyl benzoic acid (Figure 3.13).

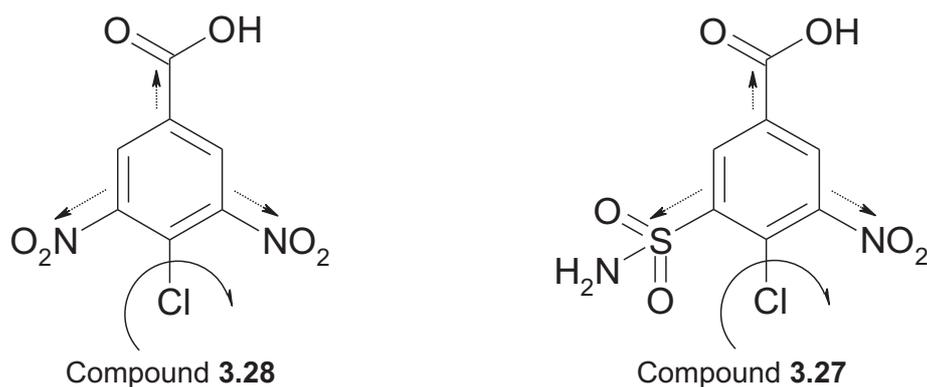


Figure 3.13: The presence of the electronegative groups deactivating the ring pulling the electrons thereby making the carbon bearing the Cl atom an easy target for amines.

Apart from using the 3,5-di-nitro benzoic acid and the 5-nitro-3-sulphamoyl benzoic acid in the current research, attempts have been made to synthesise the corresponding analogues of the Type A compounds using the 5-nitro-3-(N-methyl)sulphonyl, 4-chloro benzoic acid (Compound **3.29**). In an attempt to synthesise the compound **3.29** using compound **3.27**, the corresponding compound was found to be mixture of compound **3.29** (mono addition) and compound **3.35** (di-addition) as shown in Figure 3.14 below.

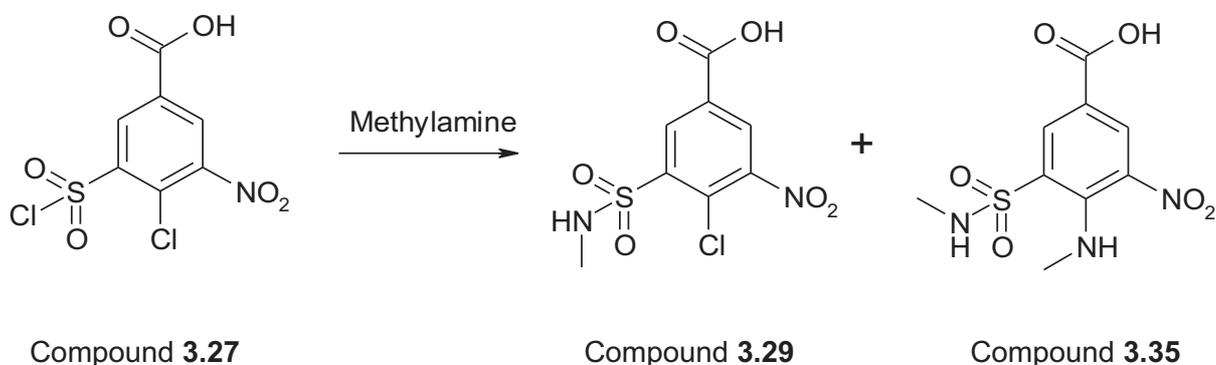


Figure 3.14: Amination of Compound **3.27**.

In the synthesis of Compound **3.27** (Figure 3.8), the final compound was obtained by simple amination of Compound **3.31** using concentrated ammonia solution. Compound **3.27** was collected as pure amorphous solid. Interestingly this reaction was not possible when N-methyl amine was used as an amine for the amination in order to prepare Compound **3.29**.

In the synthesis of Compound **3.27**, 2 M ammonia in methanol was used and the final compound was found to contain the starting material as impurities. On using concentrated

ammonia (34% ammonia solution) the final compound was obtained pure. However this was not the case with the synthesis of Compound **3.28**. On reacting with 40% methylamine solution the final compound was found to be a di-addition giving 3-(N-methyl)-sulphonyl-4-(N-methyl) benzoic acid (Compound **3.35** as in Figure 3.13). Using three equivalent of 2M methylamine in methanol after 24 hour the corresponding compound was obtained as mixture of mono (Compound **3.29**) and di-addition (Compound **3.35**) of the starting material. On reducing the reaction time to 1 hour and maintaining the temperature at 3°C, it was observed that the di-addition happens within 1 hour of the reaction as observed under NMR. This could be due to the fact that methylamine is more reactive compare to the ammonia and therefore acts much faster to give the 3-(N-methyl)sulphonyl-4-(N-methyl) benzoic acid.

Due to the presence of the carboxylic group and the methyl amine group in Compound **3.35**, this compound behaves like an amino acid and the presence of these properties had made difficulties in isolating the compound from aqueous solvent. Too much acidification made the compound soluble in the aqueous layer and therefore the pH was carefully maintained at 6 to precipitate the compound.

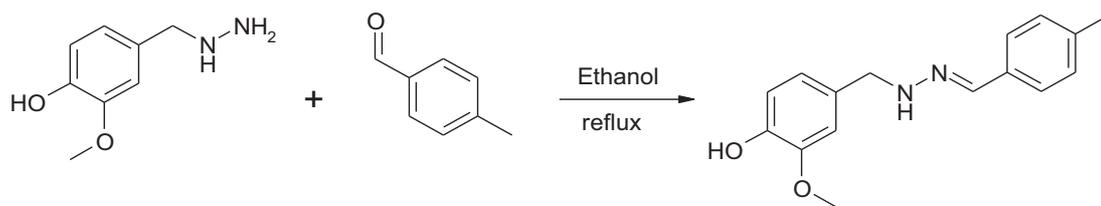
### 3.9 Reactivity of the phenylamines:

The phenethylamines used in the experiment have wide range of reactivity. Apart from the **Pa1**, **Pa5** and **Pa6** as shown in Figure 3.10 all the phenethylamines requires anhydrous condition. It was observed that on reaction with the acids in presence of the DIEA, the reaction mixture was converted into a black sticky mass, the NMR and TLC suggest plenty of impurities which were difficult to remove. **Pa3** and **Pa4** are very unstable in presence of oxygen and get converted into black sticky mass immediately on exposure to oxygen. It in fact reacts readily with the oxygen present within the apparatus. Hence Compound **3.2** was obtained as solid yellow mass on synthesis, on storage this compound get converted first into a sticky brown mass and then into a semi solid mass. Compound **3.3** on the other hand when prepared was obtained as solid which immediately converts into semi solid mass over 48 hours and on storage it converts into a black residue. This may be due to the presence of the adjacent OH groups which triggers oxidation of the moiety which is often observed with the ascorbic acids. The phenylamine **Pa2** had two adjacent OH groups and is relatively stable toward oxidation when compared with the amine **Pa3** and **Pa4** (three OH groups), however the final product has been found to absorb moisture and convert into a sticky brown mass.

### 3.10 Reactivity of the benzylhydrazines:

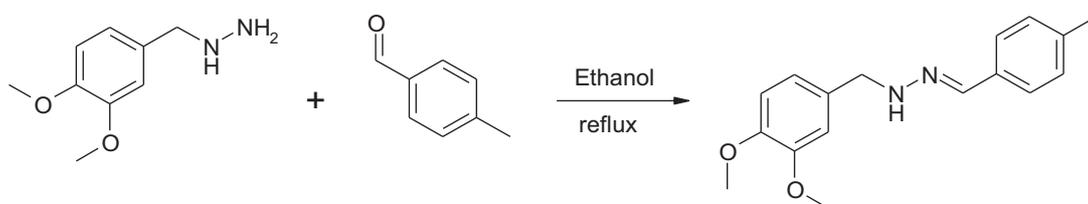
The benzyl hydrazines Compound **3.32**, Compound **3.33** and Compound **3.34** used in this work were obtained in two steps. The first step includes the action of the aldehyde on the hydrazine; then this was hydrogenated using palladium on charcoal (Figure 3.5). The reaction was successful to prepare the corresponding benzyl hydrazines Compound **3.32** and Compound **3.33**, however synthesising the benzyl hydrazine Compound **3.34** was unsuccessful. Compound **3.32** and Compound **3.33** were more reactive due to the presence of the electron donating methoxy group. Compound **3.32** is more reactive when compared with Compound **3.33** due to the presence of two electron donating methoxy groups. Compound **3.34** was the least reactive of all the benzyl hydrazines and perhaps that was the reason why the aldehydes although successfully reacted with the hydrazine in the first step to form the benzylidene hydrazine however failed in the next step to reduce the imide. Instead the reaction gave plenty of impurities which were difficult to remove.

The benzyl hydrazines were thought to react similarly to the phenyl amines used earlier in this study. Interestingly when these benzyl hydrazines Compound **3.32** and Compound **3.33** when treated with Compound **3.27** and Compound **3.28** (Figure 3.5) the final compounds were found to contain plenty of impurities which were difficult to separate. The phenylamines required higher reaction temperature and perhaps this reaction condition was too high for benzyl hydrazines. Therefore when the reaction temperatures were reduced the reaction contents were found to be the mixture of the starting materials. To further investigate the reactivity of the hydrazinyl moiety of Compound **3.32** and Compound **3.33**, each of them was set for qualitative investigation by treating with 4-methyl benzaldehyde in boiling ethanol. The corresponding imides were confirmed by NMR and a single spot on a TLC plate. Therefore this suggests that Compound **3.32** and Compound **3.33** are reactive but when reacted with Compound **3.27** and Compound **3.28** the final compounds were found to contain impurities which were difficult to remove.



Compound 3.33

Compound 3.36



Compound 3.32

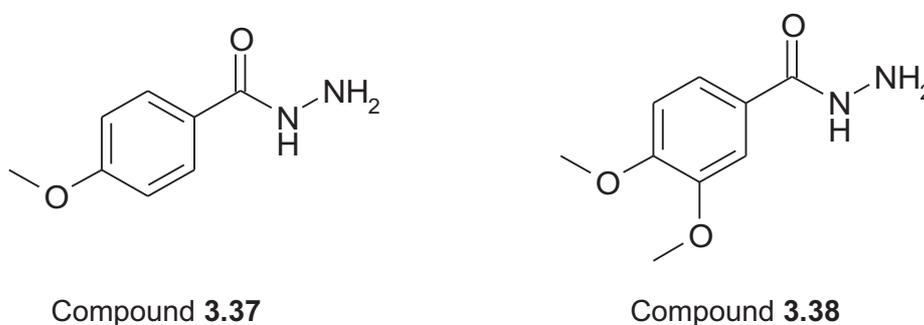
Compound 3.37

Figure 3.15: The reaction between the benzyl hydrazine and the aldehyde yielding corresponding imides.

The impurities which were observed following the reaction between Compound 3.32 and Compound 3.33 and the Compound 3.27 and Compound 3.28 were perhaps due to the presence of two reactive nitrogens in the benzyl hydrazines which contributes to the complex mixture.

### 3.11 Reactivity of the benzoyl hydrazide:

In this study the benzoyl hydrazides were found to be relatively less reactive when compared with the benzyl hydrazines used in this investigation. All the benzoyl hydrazides used in this investigation was obtained commercially.



Compound 3.37

Compound 3.38

Figure 3.16: The various benzoyl hydrazides used in this investigation.

When Compound **3.37** reacted with the aryl chlorides Compound **3.27** and Compound **3.28** under the similar condition to that with the phenylamines (Figure 3.6), the final compound was found to contain many impurities when judged with TLC. On reaction of the Compound **3.38** with the acids the resultant products were found to be a mixture of two starting materials. The presence of the electron withdrawing carboxyl group appears to render compound **3.37** and Compound **3.37** insufficiently reactive to be able to react with Compound **3.27** and Compound **3.28**.

### 3.12 Microbiological studies:

Of the 14 compounds which were synthesised, 8 were sent to the TAACF for screening against the pathogenic organism. The screening was conducted in two designed stages. The primary screening involves the determination of the 90% inhibitory concentration (IC<sub>90</sub>). The following screening or the secondary screening involves the determination of the mammalian cell cytotoxicity (CC<sub>50</sub>).

#### 3.12.1 Dose Response screening:

The dose response screen was conducted against *Mycobacterium tuberculosis* H37Rv (ATCC 27294) in BACTEC 12B medium using the Microplate Alamar Blue Assay (MABA). (Collins, L.A *et al.* 1997) Compounds were tested in ten 2-fold dilutions, typically from 100µg/mL to 0.19µg/mL. The IC<sub>90</sub> is defined as the concentration effecting a reduction in fluorescence of 90% relative to controls. This value was determined from the dose-response curve using a curve-fitting program. Any IC<sub>90</sub> value of <100µg/mL was considered "Weakly Active" for antitubercular activity and values with >100µg/mL were considered 'Inactive'.

Compound	Assay	Activity	IC <sub>50</sub> (µg/mL)	IC <sub>90</sub> (µg/mL)
<b>3.11</b>	MABA	Weakly Active	65.600	97.120
<b>3.2</b>	MABA	Inactive	>100	>100
<b>3.3</b>	MABA	Inactive	>100	>100
<b>3.4</b>	MABA	Inactive	>100	>100
<b>3.5</b>	MABA	Inactive	>100	>100
<b>3.8</b>	MABA	Inactive	>100	>100
<b>3.10</b>	MABA	Inactive	>100	>100
<b>3.12</b>	MABA	Inactive	>100	>100

Table 3.4: The biological activity of the compounds against *M. tuberculosis*.

### 3.13 Discussion:

In an attempt to synthesise the substituted sulphamoyl benzoic acids, the 4-chloro-3-nitro-5-sulphamoyl benzoic acid (Compound **3.27**) was treated directly with the amines (Figure 3.18) (Kanasal, *et al.* 1979)

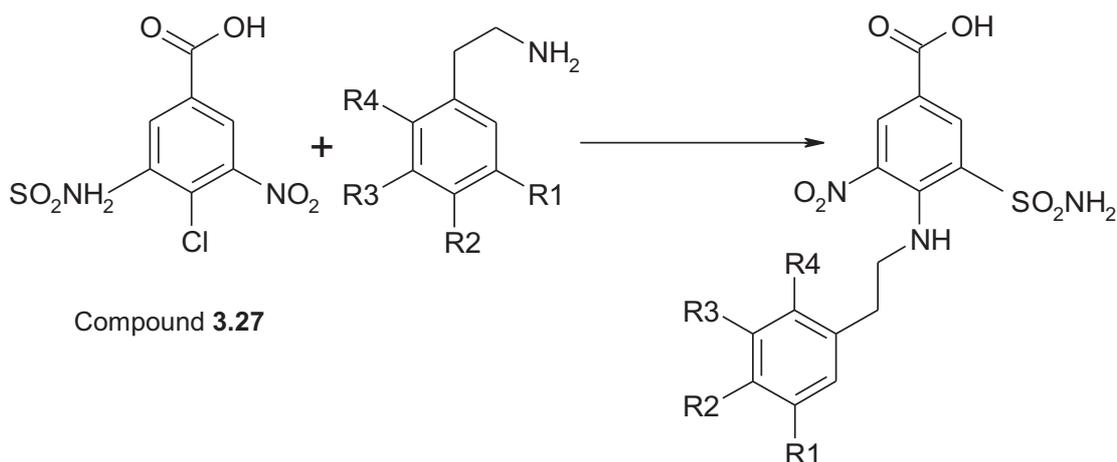
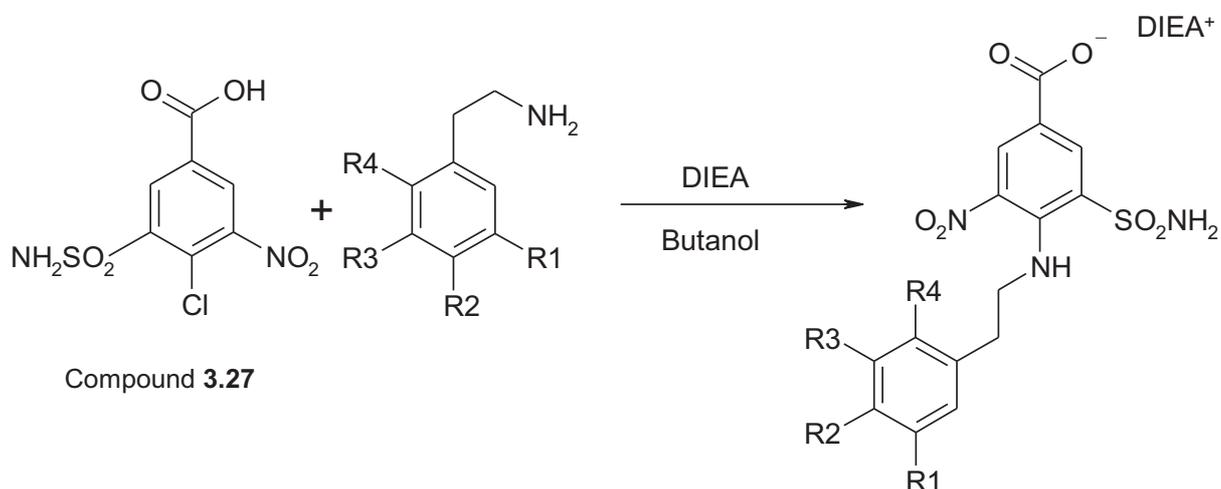


Figure 3.17: Synthesis of the substituted sulphamoyl benzoic acids

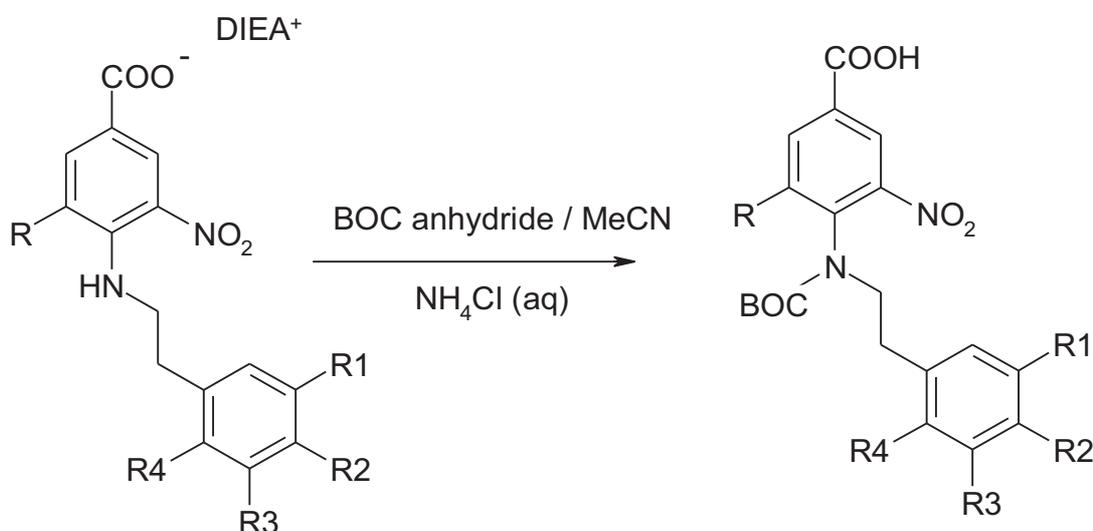
Although Kansal *et al.*, in the earlier research have successfully synthesised Compound **3.1** they did not illustrate the steps involved in the synthesis, isolation or purification of the compound. In this investigation the work has laid out steps to synthesise and isolate the substituted sulphamoyl benzoic acid derivatives Compound **3.2** to Compound **3.13** (Figure 3.2). The benzoic acid moiety was not very reactive due to the presence of the sulphonamide and the nitro group and therefore in an attempt to substitute the chlorine with the amine Compound **3.27** was treated with the amine at reflux in toluene (high boiling point solvent) in presence of pyridine (Barlow *et al.*, 1951 and Louis *et al.*, 1984); the resultant product was found to contain a mixture of the starting materials. Although 3-nitro-4-chloro-benzoic acid ethyl ester successfully reacted with the propylamine in presence of high boiling point liquid dimethylformamide giving up the corresponding 3-nitro-4-propylamino-benzoic acid ethyl ester (Özden *et al*, 2005) it was not possible in this case to synthesise Compound **3.2** under the similar condition. Using a solvent with further higher boiling point such as dimethylformamide (Özden. *et al*, 2005), propanol or butanol delivered the target compound but in an impure form. Also the resultant compound had turned into a black sticky mass; this

could possibly due to the pyridine (bp 115°C) which is susceptible under such conditions. Pyridine was not an effective solvent and n-butanol was found to be relatively better solvent under such reaction conditions (Ramage *et al.*, 1952). Incorporating DIEA (N,N-diisopropylethylamine) in place of pyridine the reaction was found successful (Ka Young *et al.*, 2004).



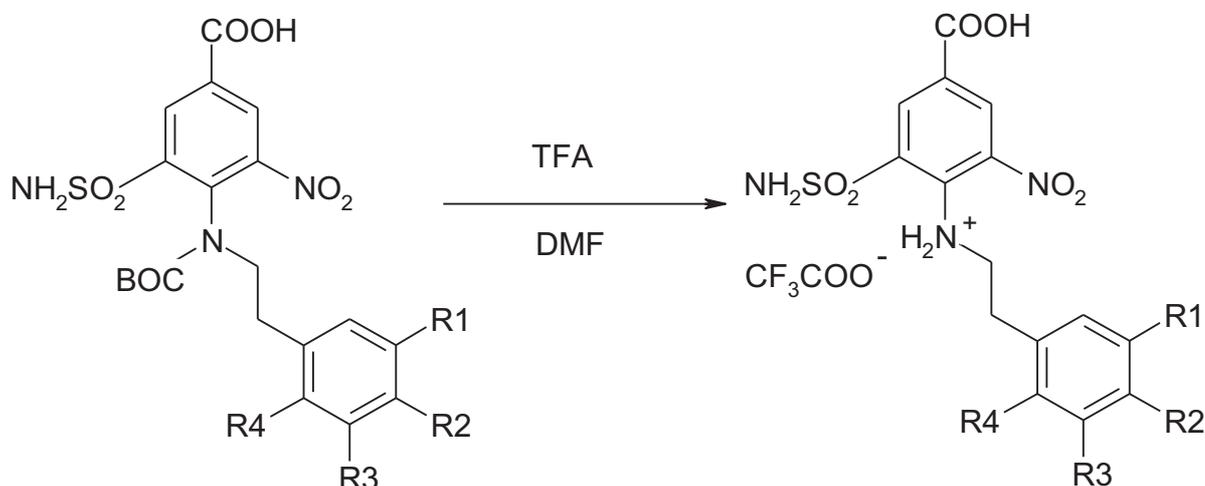
Scheme 3.1: Synthesis of the DIEA salt of the substituted sulphamoyl benzoic acids

The analogues of compound type A thus obtained were found difficult to isolate from the impurities. It was observed that the products were equally soluble into both the organic and the aqueous layer and therefore the yield of the final product was found to be very low. This could be due to the fact that the final compounds are amino acids. The basic amino group was therefore blocked with a BOC group as shown in the scheme 3.2 below using BOC anhydride (Barry. *et al.*, 1988).



Scheme 3.2: The synthesis of the BOC-substitute of compound A.

Incorporating the BOC group increases the lipophilicity of the compound and therefore it was easily extracted from the aqueous phase into the organic solvent when partitioned between ethyl acetate and  $\text{NH}_4\text{Cl}$  solution. The final compound was then generated by removing the BOC group in presence of strong acid.

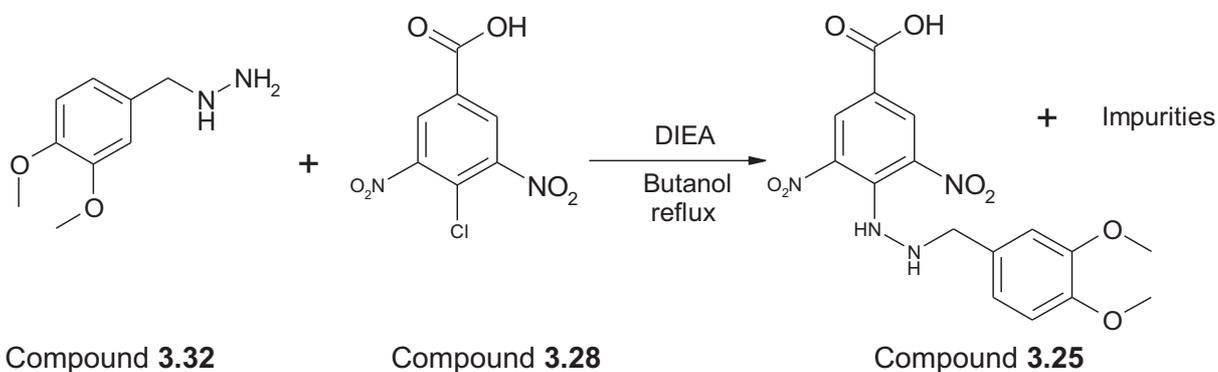


Scheme 3.3: Showing the synthesis of substituted benzoic acid derivatives.

Following the incomplete details in the literature it was not possible to prepare and isolate the target compounds. In this study the isolation was only possible via the BOC derivative. The BOC group was then removed by treating with strong acids (Shendage. D, *et al.*, 2004) to collect the title compound.

### 3.13.1 Attempted synthesis of Compound 3.25 and other related compounds:

In an attempt to synthesise Compound 3.25, Compound 3.32 was first obtained by treating the aldehyde with a large excess of hydrazine; the resultant compound was then reduced using palladium catalyst as shown in Figure 3.11.

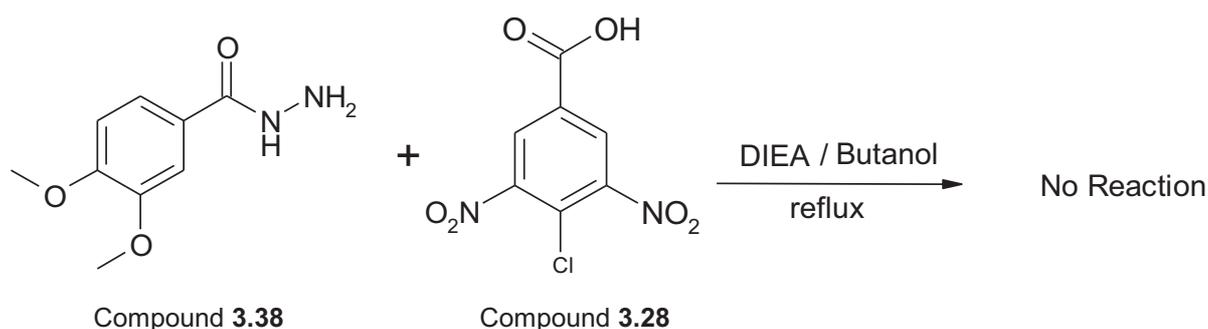


Scheme 3.4: Synthesis of the Compound 3.25.

The benzylhydrazine was then heated at reflux with Compound **3.28** in presence of Hunig's base in butanol. The  $^1\text{H}$  NMR spectrum of the amide product and the TLC indicated a complex mixture. Recrystallisation of the final product was not possible. The reaction was also carried out in reduced boiling point solvents such as toluene and ethanol and also at room temperature, but this gave no improvement. Again the presence of two potentially reactive nitrogens in the benzyl hydrazines and the derivative products (Compound **3.32**, **3.33** and **3.34**) might have been the source of some of the observed impurities.

### 3.13.2 Attempted synthesis of Compound **3.20** and related compounds:

In an attempt to prepare the Compound **3.20**, Compound **3.38** was heated at reflux with the Compound **3.28** in Hunig's base dissolved in butanol. It was observed that the resultant compounds were the mixture of two starting materials.

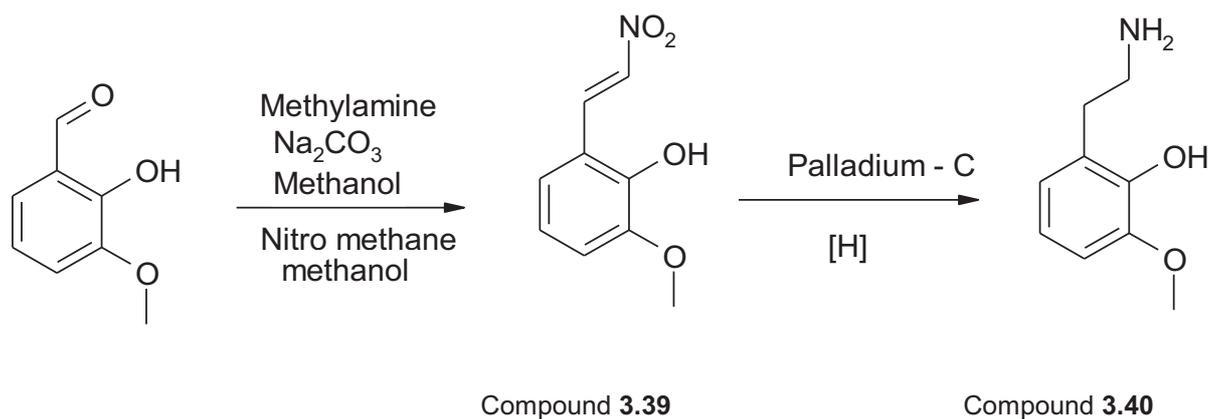


Scheme 3. 5: Showing the synthesis of Compound **3.20**.

It appears that the benzoyl hydrazide is not sufficiently nucleophilic under these conditions.

### 3.13.3 Attempted synthesis of 3-methoxy-2-hydroxy benzyl amine Compound **3.40**:

The 3-methoxy substitution on the benzyl ring of Compound **3.1** has been observed to make good interaction within the active site model and therefore an attempt has been made to synthesise the 3-methoxy-2-hydroxy benzyl substituted analogue of the compound Type A as in Figure 3.3. The aldehyde was first treated with a mixture of methylamine in methanol and  $\text{Na}_2\text{CO}_3$  and then with another mixture of nitromethane in methanol (Ma, Wu *et al.* 2004). The corresponding compound was then reduced using palladium (Daeniker, *et al.*, 1973) as shown in Figure 3.26.

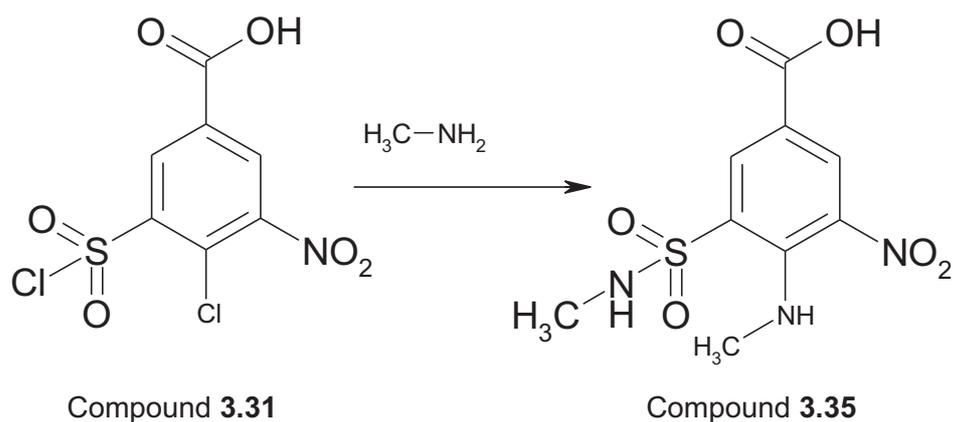


Scheme 3.6: Attempted synthesis of 3-methoxy-2-hydroxy benzyl amine.

The TLC and <sup>1</sup>H-NMR analysis indicated the presence of a complex mixture.

#### 3.13.4 Attempted synthesis of Compound **3.29**:

The methyl group attached to the sulphonamide has shown some good interaction within the active site. Therefore in an attempt to synthesise Compound **3.29**, Compound **3.31** was treated with 40% methylamine in water. The resultant compound was found to contain the double addition product of the methylamine as shown in Scheme 3.7.



Scheme 3. 7: Showing synthesis of compound **3.35**.

It was observed that on reaction of the Compound **3.31** with 2M methylamine in methanol solution, the corresponding product is a mixture of single and double addition product of the amine. On further analysis of the rate and duration of the reaction within 1<sup>st</sup>, 2<sup>nd</sup> and 24

hours, it was observed that the double addition starts immediately within 1<sup>st</sup> hour on adding the amine. This reaction was however found successful with the concentrated ammonia solution giving only the mono addition product. This could possibly be due to the fact that the methylamines are more reactive than ammonia. The presence of the electron donating methyl group pumps electrons into the amino group making the amino groups more reactive. Interestingly the amino group within the phenethylamines also carries the similar properties; these amines however required high temperature (120°C) to react with Compound **3.31**.

#### 3.13.5 Structural Activity Relationship studies:

Out of the 14 compounds synthesised, 8 compounds have been tested for biological activities against the pathogen organism. Only one compound (Compound **3.11**) has shown activity against the organism *M. tuberculosis*. The IC<sub>50</sub> value of this compound was found to be 65.6 µg/mL and the IC<sub>90</sub> value of 97.12 µg/mL; all other compounds were found inactive showing IC<sub>50</sub> values more than 100 µg/mL. In the earlier research carried out by Agrawal *et al.*, compound **3.1** was found to be very active (Agrawal *et al.*, 2007) with K<sub>i</sub> value of 5.7±1.2. In this investigation the active and the inactive compounds from the studies carried out by Agrawal *et al.*, were employed and a protein model was designed that ranks the compounds toward their possible antibacterial activity. The compound **3.2**, which was one of the closest analogues of the most potent Compound **3.1** was expected to show biological activity according to the chorismate mutase model designed in chapter 2 (Figure 3.19). Interestingly both the compounds have the nitro, carboxylic and the sulphonamide substituents which have shown similar interaction within the model. This compound only lacks the methoxy groups which were replaced with the hydroxyl groups. The hydrogen bondings which were however observed with compound **3.1** were also observed with the compound **3.2**, however it lacks efficacy within the organism, perhaps because of its pharmaceutical properties.

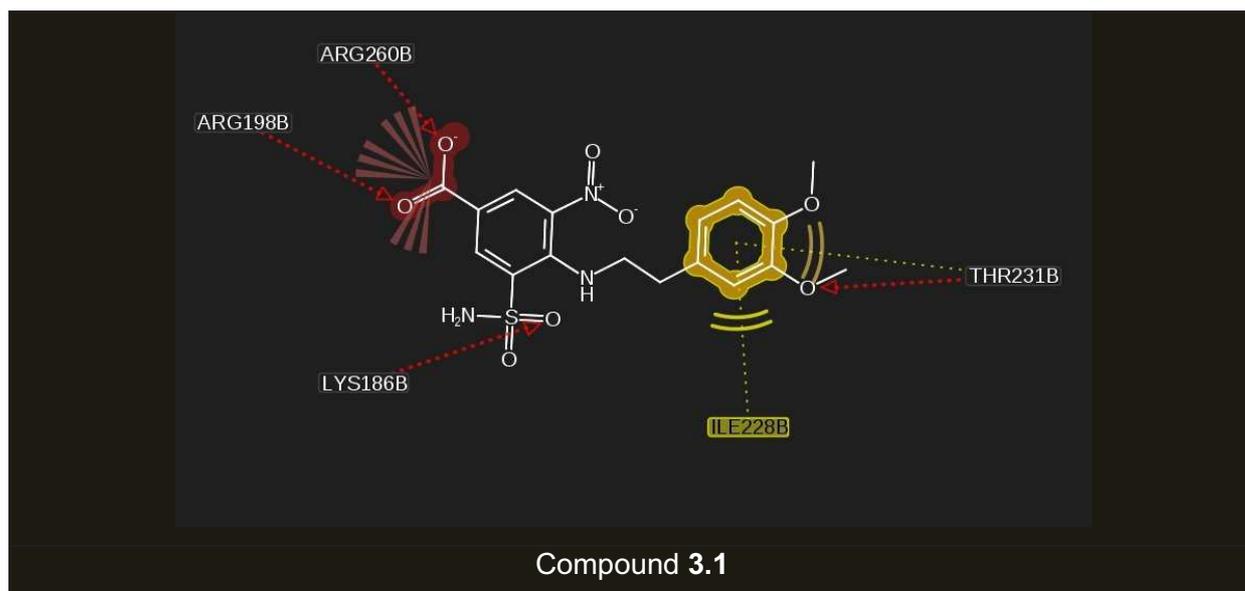


Figure 3.18: Showing the interaction of the compound **3.1** within the active site.

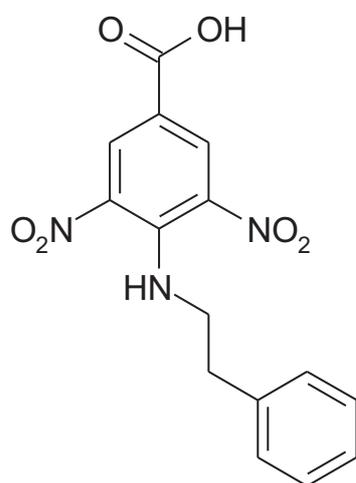
The methoxy groups of the Compound **3.1** have shown some key interaction (Figure 3.19) with the residues in the active site of the model. These interactions were improved by substituting the methoxy groups of Compound **3.1** with the hydroxyl groups as in Compound **3.2**, Compound **3.3** and Compound **3.4**. Unfortunately all these compounds lack efficacy when tested for biological activity. Interestingly all these compounds were the analogues of Compound **3.1** containing the same substituents, except the methoxy groups.

Compound	IC50 ( $\mu\text{g/mL}$ )	Mol weight	Calc LogP	Hydrogen bond acceptor	Hydrogen bond donor
<b>3.11</b>	65.600	331	3.28	9	2
<b>3.2</b>	>100	397	1.25	11	6
<b>3.3</b>	>100	413	0.98	12	7
<b>3.4</b>	>100	381	1.54	10	5
<b>3.5</b>	>100	365	2.01	9	4
<b>3.8</b>	>100	363	2.31	11	4
<b>3.10</b>	>100	347	2.80	10	3
<b>3.12</b>	>100	361	3.31	10	2

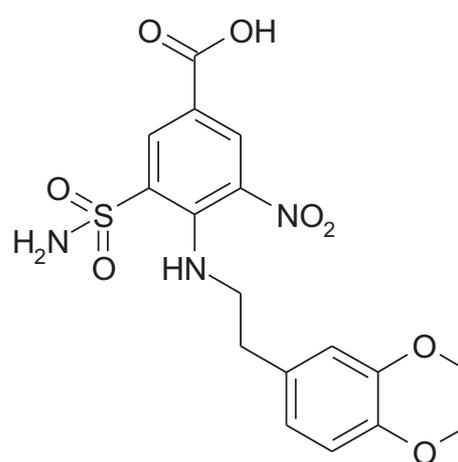
Table 3.5: The lipinski's calculation of the molecular properties. LogP was calculated using Fijitsu CAChe 6.1

All the compounds in this chapter of the thesis were isolated as TFA salts. This gives an additional weight of 114 to all the compounds shown in table 3.5. The Compound **3.1** had 11 hydrogen bond acceptors and 4 hydrogen bond donors and perhaps was isolated as a free molecule and thus falls within Lipinski's rule of 5 (Lipinski *et al.*, 2001). Interestingly Compound **3.2**, Compound **3.3** and Compound **3.4** all have more than 5 hydrogen bond donors and molecular weight more than 500 (as TFA salts form) therefore do not obey Lipinski's rule of 5. Although the lack of efficacy of these compounds particularly Compound **3.2** and Compound **3.3** can be supported at this stage by Lipinski's rule, however this requires attention when considering these compounds for further development. Compound **3.11** on the other hand stays well within the Lipinski's rule.

The Compound **3.12** in particular had the 3-methoxy substitution on the phenylamine moiety and successfully showed good interactions with the active site residues in the protein model of the enzyme, however showed no biological activity when tested in the organism. The Compound **3.11** on the other hand was found to be active with IC<sub>50</sub> value of 65.6 µg/mL. This compound had a nitro substitution instead of sulphonamide and was also devoid of any substitution on the phenethylamine moiety. Interestingly this compound had 2 hydrogen bond donors and obeys the Lipinski's rule. The nitro substitution however carries almost similar electronic effect to the sulphonamide as in Compound **3.1**.



Compound **3.11**  
IC<sub>50</sub> 65.6 µg/mL



Compound **3.1**  
K<sub>i</sub> value of 5.7±1.2 nM

Scheme 3.8: Showing structural similarity between compound **3.11** and compound **3.1**

These di-nitro substituted benzoic acid derivatives also have shown good interaction within the chorismate mutase model and these compounds were ranked well in the model. The 3,4-dimethoxy substituted phenylamines, 3-methoxy- 4-hydroxyphenylamine and the 4-methoxy-3-hydroxy substituted phenylamines have shown interesting interactions within the model but these compounds were not possible to synthesise due to the restriction of laboratory use of these compounds ([www.homeoffice.gov.uk/documents/cdlist.pdf](http://www.homeoffice.gov.uk/documents/cdlist.pdf)) .

The protein model prepared in the previous chapter (Chapter 2) was based on the potential interactions between the active site residues and the Compound **3.1** which has shown significant activity against mycobacterial chorismate mutase (Aggarwal *et al.*, 2007). Compounds **3.2** to **3.7** of figure 3.2 have the same structure with only modification on the methoxy groups and according to the model these compounds were expected to show good activity in the organism (T.B). Unfortunately none of these compounds have shown biological activity when tested against the organism. Compound **3.11** was only biologically active and this compound was different to compound **3.1** with the substitutions on both the rings. Thus Compound **3.1** together with Compound **3.11** left an interesting area of further research to discovery of chorismate mutase inhibitor from *Mycobacterium tuberculosis*.

Due to the limitation on the structural information in the database of the compounds (Aggarwal *et al.*, 2007), the selection of the biologically active and inactive compounds became very small. Only 3 active compounds were available to extract the structural information of the ligand within the active site. All of these 3 compounds were also structurally very different from each other. Despite this limited information it was still possible to prepare a model which has ranked these ligands according to their biological activity. Although a library of 23 compounds was designed and was attempted for synthesis, it was only possible to prepare 12 compounds. This is due to the restriction in the use of the various phenylamines and thus the library prepared was small. The information on the synthesis of the Compound **3.1** was also not outlined by the authors in the previous research and in this work much emphasis was given to establish the steps involved in the preparation and isolation of all the analogues of Compound **3.1**. The work in this thesis revealed another active compound (Compound **3.11**) and some inactive compounds which along with the other library (Aggarwal *et al.*, 2007) opened a vast area to carry out further research to discover inhibitors of chorismate mutase from *Mycobacterium tuberculosis*.

### **3.14 Conclusion:**

A library of the substituted sulphamoyl benzoic acid derivatives which were ranked well in the model was prepared. All these compounds have been successfully isolated, purified and characterised. These compounds have been screened for biological activity against the pathogenic organism *Mycobacterium tuberculosis*. Only one compound has shown activity against the pathogenic organism and all the rest were found inactive.

## CHAPTER 4

### MATERIALS AND METHODS

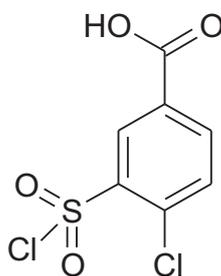
#### 4.1. Chemicals:

All chemicals and solvents were used as supplied. 4-chlorobenzoic acid, fuming nitric acid, chlorosulphonic acid, ammonia solution, phenethylamines, hydrazine monohydrate and all the aldehydes were purchased from Sigma-Aldrich. Organic solvents were purchased from Fisher Chemicals. Thin layer chromatography plates on aluminium sheets (20 X 20) of silica gel F<sub>254</sub> were obtained from Merck.

#### 4.2. Instrumentation:

Proton NMR spectra were obtained on a Bruker AC 250 instrument operating at 250 MHz as solution in d<sub>6</sub>-DMSO and referenced from  $\delta$ DMSO = 2.50 ppm unless otherwise stated. Infrared spectra were recorded as KBr disc on a Mattson 3000 FTIR spectrophotometer. (APCI-MS) was carried out on a Hewlett-Packard 5989B quadrupole instrument connected to an electrospray 59987A unit with an APCI accessory and automatic injecting using a Hewlett-Packard 1100 series autosampler. Accurate MS was carried out on TOF mass spectra (ESI mode) measured on a Waters LCT Premier Mass Spectrometer.

#### 4.3.1.1 Preparation of 4-chloro-3-chlorosulfonylbenzoic acid: Compound 3.30.



11ml or 0.165 mol of chlorosulfonic acid was taken in a three neck flask. The temperature inside the flask was maintained at 15°C. 5g or 0.0319 mol of 4-chlorobenzoic acid was added gradually with constant stirring. The temperature within the flask was then maintained at 125°C and was left to react for 24 hours. The reaction mixture was then cooled and added dropwise into 100ml of ice water with constant stirring. The white product precipitated out which was then filtered and freeze dried for 48 hours to obtain white amorphous powder.

Yield was 6.06g, 24.047 mmol, 75%.

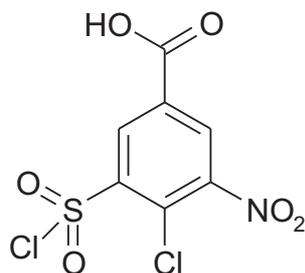
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.42 (d, 1H,  $J$  = 1.896 Hz, Benzyl-6H); 7.85 (dd, 1H,  $J$  = 1.896, 8.214 Hz, Benzyl-2H); 7.5 (d, 1H,  $J$  = 8.214 Hz, Benzyl-2H); 8.47 (s, 1H, OH) ppm.

MS (APCI -)  $m/z$  = found 201, 215, 235, 237, 249, 251, (M - H) $^+$ .

IR (KBr)  $\nu$  = 3404-3087, 2900 (-COOH), 2645, 2500, 2198, 1928, 1833, 1735, 1598, 1478, 1410 (SO<sub>2</sub>Cl), 1307  $\text{cm}^{-1}$

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.75 (single spot).

#### 4.3.1.2 Preparation of 4-chloro-3-chlorosulfonyl-5-nitrobenzoic acid: Compound 3.31.



6.06g or 0.0237 mol of 4-chloro-3-sulfonylbenzoic acid was dissolved in 98% concentrated sulphuric acid was taken in a three neck flask. The temperature was maintained within  $-15^\circ\text{C}$  to  $-10^\circ\text{C}$ . 33ml of nitrating mixture prepared from 9ml of fuming nitric acid in 24ml of concentrated sulphuric acid was added drop wise into the flask with continuous stirring. The temperature was then raised slowly to  $105^\circ\text{C}$  when the suspension dissolves to a solution and was left to react overnight. The reaction mixture was cooled and added drop wise into 300ml of ice water, it was then filtered, washed and collected as white amorphous powder.

Yield = 7.07g, 23.64mmol, 98.6 %.

Mp: 153.3-156.6  $^\circ\text{C}$

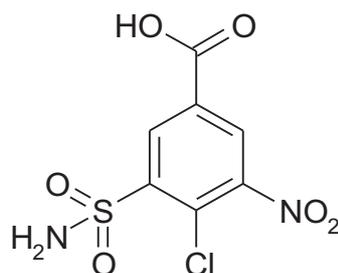
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.61 (d, 1H,  $J$  = 1.896 Hz, Benzyl-6H); 8.4 (d, 1H,  $J$  = 2.527 HZ, Benzyl-2H); 4.5 (bs, 1H, OH) ppm.

MS (APCI +)  $m/z$  = found 202, 204 (M - SO<sub>2</sub>Cl) $^-$ , 246 (M - OH, Cl) $^-$ , 298 (M - H) $^-$ .

IR (KBr)  $\nu$  = 3384-3087, 2894 (-COOH), 2645, 2500, 2198, 1958, 1847, 1698, 1598, 1544 (-NO<sub>2</sub>), 1420- 1150 (SO<sub>2</sub>Cl and NO<sub>2</sub>),  $\text{cm}^{-1}$

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.187 (single spot).

#### 4.3.1.3 Preparation of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid: Compound 3.27.



7g or 23.300 mmol of 4-chloro-3-chlorosulfonyl-5-nitro-benzoic acid was dissolved in 80ml of 2M concentrated ammonia in methanol solution. The suspension slowly turns yellow and the reaction mixture was left to react for 24 hours. The reaction mixture was then transferred dropwise into acidified water and the pH was adjusted to just below 6. The product was then extracted in ethyl acetate, evaporated, dried and collected as yellow amorphous powder.

Yield: 3.78g or 13.404mmol, 58%

Mp: 121.6-125.7 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 8.69 (d, 1H, J= 1.86 Hz, Ph-6H); 8.55 (d, 1H, J= 1.86 Hz, Ph-2H); 8.00 (bs, 1H, OH); 7.55 (bs, 2H, NH<sub>2</sub>) ppm.

MS (APCI -) m/z = found 201 (M – SO<sub>2</sub> NH<sub>2</sub>)<sup>-</sup>, 235 (M – NO<sub>2</sub>)<sup>-</sup>, 245 (M – Cl)<sup>-</sup>, 279 (M - H)<sup>-</sup>.

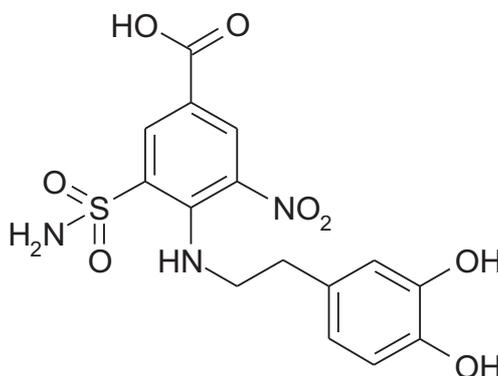
IR (KBr) ν= 3575, 3300-2790, 2315, 2295, 2100, 1890, 1810, 1665, 1610, 1550 (-NO<sub>2</sub>), 1475, 1335 (-SO<sub>2</sub>NH<sub>2</sub>) cm<sup>-1</sup>

MS (-ES) = 278.95 (M - H)<sup>-</sup> and 280.95 (M - H)<sup>-</sup>

Accurate MS (-ES) = Calculated 278.9479 (m-H)<sup>-</sup> and 280.9451 (m-H)<sup>-</sup> found 278.9478 (m-H)<sup>-</sup> and 280.9453 (m-H)<sup>-</sup>, 0.4 ppm.

TLC carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.58 (single spot tailing).

#### 4.4.1 Preparation of 4-(3,4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.2.



##### 4.4.1.1. Preparation of DIEA salt of 4-(3,4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

0.159g or 0.567 mmol of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid was dissolved in 10ml of butanol and 0.35g of diisopropylethylamine was added. 0.1g or 0.529 mmol of dopamine was then added into the reaction mixture and left reacted under reflux for 24 hours under argon. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.17g or 0.426 mmol, 75%

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.8 (s, 1H, OH); 9.65 (s, 1H, OH); 8.55 (d, 1H,  $J$  = 1.86 Hz, Ph-2H); 8.35 (d, 1H,  $J$  = 1.86 Hz, Ph-6H); 7.9 (bs, 2H,  $\text{NH}_2$ ); 7.00 (s, 1H, NH); 6.8 (d, 1H,  $J$  = 8.214 Hz, Ph-2H); 6.8 (d, 1H,  $J$  = 1.896 Hz, Ph-5H); 6.65 (dd, 1H,  $J$  = 1.896, 8.214 Hz, Ph-6H); 3.15 (dd,  $J$  = Hz, 2H,  $\text{CH}_2$ ); 2.78 (dd,  $J$  = Hz, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) shows long tailing from baseline.

##### 4.4.1.2 Preparation of BOC substituted 4-(3,4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid :

The product was then dissolved in 15ml acetonitrile and 0.12g or 0.549 mmol of BOC anhydride was added and left reacted for 48 hours under argon. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.14g

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.8 (s, 1H, OH); 9.65 (s, 1H, OH); 8.55 (d, 1H,  $J$  = 1.86 Hz, Ph-2H); 8.35 (d, 1H,  $J$  = 1.86 Hz, Ph-6H); 8.00 (bs, 2H,  $\text{NH}_2$ ); 7.00 (s, 1H, NH); 6.7 (d, 1H,  $J$  = 8.214 Hz, Ph-2H); 6.7 (d, 1H,  $J$  = 1.896 Hz, Ph-5H); 6.5 (dd, 1H,  $J$  = 1.896, 8.214 Hz, Ph-6H); 3.15 (dd,  $J$  = Hz, 2H,  $\text{CH}_2$ ); 2.78 (dd,  $J$  = Hz, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, BOC) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.72 (tailing).

**4.4 1.3. Preparation of 4-(3,4-dihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid :**

The BOC substituted product was then dissolved in 5ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 4 hours. The final product was obtained by evaporating the reaction mixture to obtain the dark yellow colour product.

Yield: 0.1g, 0.251 mmol, 44%

Mp: 63.2 – 68.1 °C.

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.8 (bs, 2H, Phenyl-OH); 8.44 (d, 1H,  $J$ = 1.896 Hz, Ph-6H); 8.32 (d, 1H,  $J$ = 1.896 Hz, Ph-2H); 7.94 (s, 2H,  $\text{NH}_2$ ); 7.05 (t, 1H, NH); 6.62 (d, 1H,  $J$ = 8.214 Hz, Ph-2H); 6.6 (d, 1H,  $J$ = 1.896 Hz, Ph-5H); 6.45 (dd, 1H,  $J$ = 1.896, 8.214 Hz, Ph-6H); 3.064 (m, 2H,  $\text{CH}_2$ ); 2.72 (m, 2H,  $\text{CH}_2$ ) ppm.

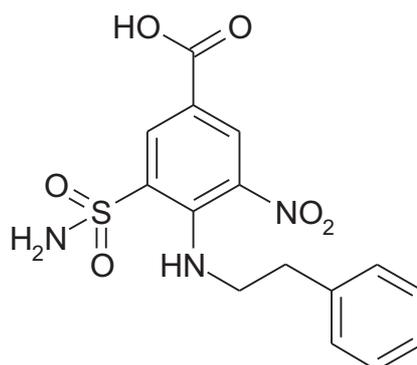
IR (KBr)  $\nu$ = 3357, 3202, 3090, 2931, 2855, 2637, 2366, 1856, 1657, 1614, 1575 ( $-\text{NO}_2$ ), 1410 ( $-\text{OH}$ ), 1325 ( $-\text{SO}_2\text{NH}_2$ )  $\text{cm}^{-1}$

MS (ES  $-$ )  $m/z$  = found 396.05 (M - H)  $^-$ .

Accurate MS (-ES) = Calculated 396.0502 (m-H)  $^-$  found 396.0505 (m-H)  $^-$  0.8 ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.44 (single spot small tailing).

**4.5.1. Preparation of 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid: Compound 3.5.**



**4.5.1.1 Preparation of DIEA salt of 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid:**

0.15g or 0.567 mmol of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid was dissolved in n-10ml of butanol and 0.3g or 2.325 mmol of diisopropylethylamine was added. 0.075g or 0.619 mmol of phenethylamine was then added into the reaction mixture and left reacted under reflux for 16 hours. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.2g 0.549 mmol, 96%

$^1\text{H}$  NMR (250 MHz  $\text{d}_6\text{-DMSO}$ ): 8.44 (d, 1H,  $J = 1.896$  Hz, Ph-6H); 8.32 (d, 1H,  $J = 1.896$  Hz, Ph-2H); 8.00 (bs, 2H,  $\text{NH}_2$ ); 7.2 (m, 5H, Ph-H); 4.4 (s, 1H, NH); 3.1 (m, 2H,  $\text{CH}_2$ ); 2.8 (dd,  $J = 2.52, 8.84$  Hz, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.75 (single spot).

#### 4.5.1.2 Preparation of BOC substituted 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

The product was then dissolved in 15ml acetonitrile and 0.12g or 0.549mmol of BOC anhydride was added and left reacted for 48 hours. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.338g,

$^1\text{H}$  NMR (250 MHz  $\text{d}_6\text{-DMSO}$ ): 8.44 (d, 1H,  $J = 1.896$  Hz, Ph-6H); 8.32 (d, 1H,  $J = 1.896$  Hz, Ph-2H); 7.8 (bs, 2H,  $\text{NH}_2$ ); 7.2 (m, 5H, Ph-H); 7.1 (s, 1H, NH); 6.9 (t, 1H, NH); 3.15 (dd,  $J = 1.89, 8.21$  Hz, 2H,  $\text{CH}_2$ ); 2.6 (dd,  $J = 2.52, 8.84$  Hz, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, BOC) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.73 (single spot).

#### 4.5.1.3. Preparation of 4-(phenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

The BOC substituted product was then dissolved in 5ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 4 hours. The final product was obtained by evaporating the reaction mixture to obtain the yellow colour product (yield = 0.24g or 0.657mmol), which was recrystallised from dichloromethane to give bright yellow product.

Yield = 0.09g or 0.246 mmol or 46%

Mp: 157.2 – 163.0  $^{\circ}\text{C}$

$^1\text{H}$  NMR (250 MHz  $\text{d}_6\text{-DMSO}$ ): 13.15 (bs, 1H, OH); 8.46 (d, 1H,  $J = 1.896$  Hz, Ph-6H); 8.33 (d, 1H,  $J = 1.896$  Hz, Ph-2H); 7.9 (s, 2H,  $\text{NH}_2$ ); 7.28 (m, 5H, Ph-H); 7.1 (s, 1H, NH); 3.16 (m, 2H,  $\text{CH}_2$ ); 2.8 (dd,  $J = 2.52, 8.84$  Hz, 2H,  $\text{CH}_2$ ); ppm.

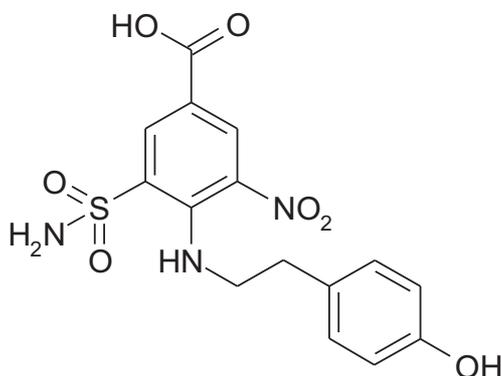
IR (KBr)  $\nu = 3378, 3347, 3272, 3083, 3025, 2958, 2658, 2624, 2529, 2282, 1955, 1688, 1610, 1550$  ( $-\text{NO}_2$ ), 1407, 1320 ( $-\text{SO}_2\text{NH}_2$ )  $\text{cm}^{-1}$

MS (ES  $-$ )  $m/z =$  found 364 (M - H) $^-$ .

Accurate MS (ES  $-$ ) = Calculated 364.0603 (m-H) $^-$  found 364.0601 (m-H) $^-$  0.5 ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.62 (single spot).

#### 4.6.1. Preparation of 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid: Compound 3.4.



##### 4.6.1.1 Preparation of DIEA salt of 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

0.151g or 0.567 mmol of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid was dissolved in 10ml of butanol and 0.3g or 2.325 mmol of diisopropylethylamine was added. 0.1g or 0.578 mmol of 4-hydroxyphenethylamine was then added into the reaction mixture and left reacted under reflux for 16 hours. The solvent was then evaporated to obtain a thick mass.

Yield = 0.19g 0.490mmol, 86%

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.45 (s, 1H, OH); 8.48 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.29 (d, 1H,  $J$  = 1.896 Hz, Ph-2H); 8.00 (bs, 2H,  $\text{NH}_2$ ); 7.06 (dd, 2H,  $J$  = 8.214, 1.896 Hz, Ph-H); 6.9 (s, 1H, OH); 6.67 (dd, 2H,  $J$  = 1.896, 8.214 Hz, Ph-H); 4.4 (s, 1H, NH); 3.064 (m, 2H,  $\text{CH}_2$ ); 2.72 (m, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.66 (long tailing)

##### 4.6.1.2 Preparation of BOC substituted 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

The product was then dissolved in 15ml acetonitrile and 0.12g or 0.549mmol of BOC anhydride was added and left reacted for 48 hours. The solvent was then evaporated and the product was dissolved and ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.121g or 0.244mmol

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.3 (bs, 1H, OH); 9.2 (bs, 1H, OH); 8.48 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.29 (d, 1H,  $J$  = 1.896 Hz, Ph-2H); 8.00 (s, 2H,  $\text{NH}_2$ ); 7.06 (dd, 2H,  $J$  = 8.214, 1.896 Hz, Ph-H); 6.9 (s, 1H, OH); 6.67 (dd, 2H,  $J$  = 1.896, 8.214 Hz, Ph-H); 3.064 (m, 2H,  $\text{CH}_2$ ); 2.72 (m, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, BOC) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.49 (single spot).

#### 4.6.1.3 Preparation of 4-(4-hydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

0.121g or 0.244mmol of the BOC substituted product was then dissolved in 5ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 4 hours. The final product was obtained by evaporating the reaction mixture to obtain the brown colour product, which was then recrystallised from dichloromethane.

Yield = 0.147g or 0.385 mmol or 72%

Mp: 160.2 – 163.5 °C

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 13.1 (bs, 1H, NH); 9.3 (s, 1H, OH); 9.2 (s, 1H, OH); 8.46 (d, 1H,  $J$ = 1.896 Hz, Ph-6H); 8.3 (d, 1H,  $J$ = 1.896 Hz, Ph-2H); 8.00 (s, 2H,  $\text{NH}_2$ ); 7.08 (dd, 2H,  $J$ = 8.214, 1.896 Hz, Ph-H); 6.67 (dd, 2H,  $J$ = 1.896, 8.214 Hz, Ph-H); 3.093 (bd, 2H,  $\text{CH}_2$ ); 2.72 (m, 2H,  $\text{CH}_2$ ); ppm.

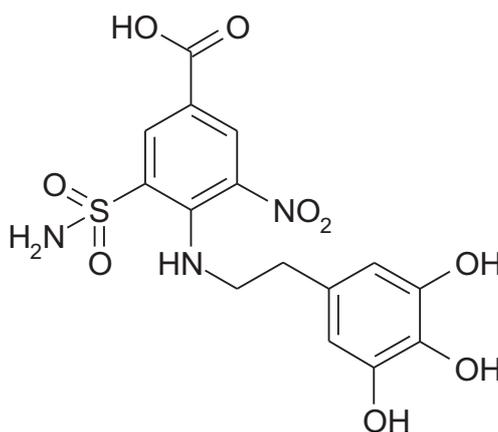
TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.238 (single spot).

IR (KBr)  $\nu$  = 3459, 3366, 3347, 3318, 3200, 3083, 2666, 2537, 1872, 1679, 1610, 1520(- $\text{NO}_2$ ), 1410 (-OH), 1407, 1330 (- $\text{SO}_2\text{NH}_2$ )  $\text{cm}^{-1}$

MS (ES -)  $m/z$  = found 380.05 (M - H) $^-$ .

Accurate MS (-ES) = Calculated 380.0552 (m-H) $^-$  found 380.0551 (m-H) $^-$  0.3 ppm.

#### 4.7.1 Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.3.



##### 4.7.1.1 Preparation of DIEA salt of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

0.14g or 0.537 mmol of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid was dissolved in 15ml of butanol and 0.3g or 2.325 mmol of diisopropylethylamine was added. 0.095g or 0.463 mmol of 3,4,5-trihydroxyphenethylamine was then added into the reaction mixture and left reacted

under reflux for 16 hours under argon. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.2g, 0.485mmol, 90.4%

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.5 (bs, 2H, OH); 8.8 (bs, 1H, OH); 8.44 (d, 1H, J= 2.527 Hz, Ph-6H); 8.32 (d, 1H, J= 2.527 Hz, Ph-2H); 7.9 (bs, 2H,  $\text{NH}_2$ ); 6.9 (bs, 1H, OH); 6.13 (s, 2H, Ph-H); 4.4 (bs, 1H, NH); 3.064 (m, 2H,  $\text{CH}_2$ ); 2.62 (m, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.25 (tailing).

#### 4.7.1.2 Preparation of BOC substituted 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

The product was then dissolved in 25ml acetonitrile and 0.12g or 0.549mmol of BOC anhydride was added and left reacted for 24 hours under argon. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the yellowish brown BOC substituted product.

Yield = 0.133g.

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.8 (s, 3H, OH); 8.44 (d, 1H, J= 1.896 Hz, Ph-6H); 8.32 (d, 1H, J= 2.527 Hz, Ph-2H); 7.9 (bs, 2H,  $\text{NH}_2$ ); 6.9 (bs, 1H, OH); 6.12 (s, 2H, Ph-H); 4.4 (t, 1H, NH); 3.034 (m, 2H,  $\text{CH}_2$ ); 2.62 (m, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, BOC); ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.125 (tailing).

#### 4.7.1.3 Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours under argon. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product.

Yield = 0.1g, 0.234 mmol or 48%

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.8 (s, 1H, OH); 8.44 (d, 1H, J= 1.896 Hz, Ph-6H); 8.32 (d, 1H, J= 2.527 Hz, Ph-2H); 7.9 (s, 2H,  $\text{NH}_2$ ); 7.3 (s, 1H, OH); 7.1 (s, 1H, OH); 6.9 (s, 1H, OH); 6.12 (s, 2H, Ph-H); 4.4 (bs, 1H, NH); 3.034 (m, 2H,  $\text{CH}_2$ ); 2.62 (m, 2H,  $\text{CH}_2$ ); ppm.

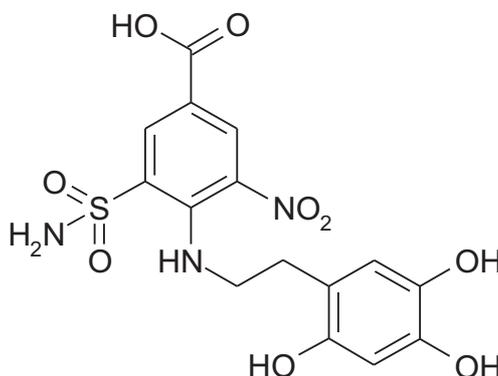
TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.17 (tailing).

MS (ES  $-$ )  $m/z$  = found 412.05 (M - H) $^-$

MS (ES  $+$ )  $m/z$  = found 414.06 (M + H) $^+$

Accurate MS (-ES) = Calculated 412.0451 (m-H) $^-$  found 412.0449 (m-H) $^-$  0.5 ppm.

#### 4.8.1 Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.7.



##### 4.8.1.1 Preparation of DIEA salt of 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

0.14g or 0.537 mmol of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid was dissolved in 15ml of butanol and 0.3g or 2.325 mmol of diisopropylethylamine was added. 0.1g or 0.487 mmol of 3,4,5-trihydroxyphenethylamine was then added into the reaction mixture and left reacted under reflux for 16 hours under argon. The solvent was then evaporated to obtain thick brown oil.

Yield = 0.26g, 0.631mmol 100%

$^1\text{H}$  NMR (250 MHz  $\text{d}_6$ -DMSO): 8.8 (s, 1H, OH); 8.44 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.32 (d, 1H,  $J$  = 2.527 Hz, Ph-2H); 7.9 (s, 2H,  $\text{NH}_2$ ); 7.0 (bs, 3H, OH); 6.12 (s, 1H, Ph-H); 5.9 (s, 1H, Ph-H); (4.4 (bs, 1H, NH); 3.034 (bt, 2H,  $\text{CH}_2$ ); 2.62 (dd,  $J$ =8.214 Hz, 2H,  $\text{CH}_2$ ); ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.25 (tailing).

##### 4.8.1.2 Preparation of BOC substituted 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

The product was then dissolved in 25ml acetonitrile and 0.125g or 0.571mmol of BOC anhydride was added and left reacted for 24 hours under argon. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the yellowish brown BOC substituted product.

Yield = 0.12g

$^1\text{H}$  NMR (250 MHz  $\text{d}_6$ -DMSO): 8.8 (s, 1H, OH); 8.44 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.32 (d, 1H,  $J$  = 2.527 Hz, Ph-2H); 7.9 (s, 2H,  $\text{NH}_2$ ); 7.0 (bs, 3H, OH); 6.12 (s, 1H, Ph-H); 5.9 (s, 1H, Ph-H); (4.4 (bs, 1H, NH); 3.034 (bt, 2H,  $\text{CH}_2$ ); 2.62 (dd,  $J$ =8.214 Hz, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, BOC); ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.25 (tailing).

#### 4.8.1.3 Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours under argon. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product.

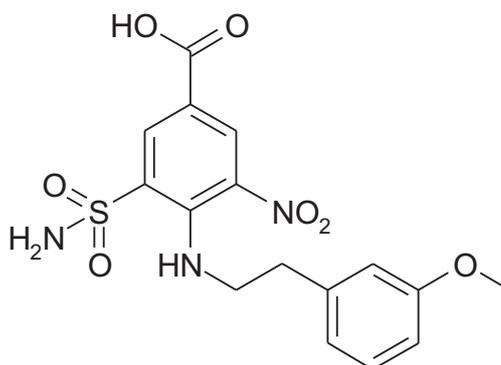
Yield = 0.101 g or 0.234 mmol or 48%

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.8 (s, 1H, OH); 8.44 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.32 (d, 1H,  $J$  = 2.527 Hz, Ph-2H); 7.9 (s, 2H,  $\text{NH}_2$ ); 7.0 (bs, 3H, OH); 6.12 (s, 1H, Ph-H); 5.9 (s, 1H, Ph-H); (4.4 (bs, 1H, NH); 3.034 (bt, 2H,  $\text{CH}_2$ ); 2.62 (dd,  $J$  = 8.214 Hz, 2H,  $\text{CH}_2$ ); ppm.

MS (ES  $-$ )  $m/z$  = found 412. ( $M - H$ ) $^-$

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.47 (tailing).

#### 4.9.1 Preparation of 4-(3-methoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid : Compound 3.6



##### 4.9.1.1 Preparation of DIEA salt of 4-(3-methoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:

0.14g or 0.5 mmol of 4-chloro-3-sulphamoyl-5-nitrobenzoic acid was dissolved in 10ml of butanol and 0.3g or 2.325 mmol of diisopropylethylamine was added. 0.079g or 0.523 mmol of 3-methoxyphenethylamine was then added into the reaction mixture and left reacted under reflux for 36 hours under argon. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.16g, 0.405mmol, 81%

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.44 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.32 (d, 1H,  $J$  = 1.896 Hz, Ph-2H); 8.00 (bs, 2H,  $\text{NH}_2$ ); 7.2 (dd, 1H,  $J$  = 8.214 Hz, Ph-4H); 6.8 (d, 1H,  $J$  = 1.896 Hz, Ph-5H); 6.8 (dd, 2H,  $J$  = 1.896, 8.214 Hz, Ph-6H); 4.4 (bs, 1H, NH); 3.064 (m, 2H,  $\text{CH}_2$ ); 2.72 (dd,  $J$  = 2.52, 8.84 Hz, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.14 (single spot).

*4.9.1.2. Preparation of BOC substituted 4-(3-methoxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:*

The product was then dissolved in 25ml acetonitrile and 0.12g or 0.549mmol of BOC anhydride was added and left reacted for 16 hours under argon. The solvent was then evaporated under high vacuum pump and the product was partitioned between ammonium hydrochloride solution and ethyl acetate. The organic layer was then evaporated to obtain the BOC substituted product as a yellow sticky mass.

Yield = 0.11g

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 8.44 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.32 (d, 1H,  $J$  = 1.896 Hz, Ph-2H); 8.00 (s, 2H,  $\text{NH}_2$ ); 7.2 (dd, 1H,  $J$  = 8.214 Hz, Ph-4H); 6.8 (d, 1H,  $J$  = 1.896 Hz, Ph-5H); 6.8 (dd, 2H,  $J$  = 1.896, 8.214 Hz, Ph-6H); 3.064 (m, 2H,  $\text{CH}_2$ ); 2.72 (dd,  $J$  = 2.52, 8.84 Hz, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, BOC); ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.15 (single spot).

*4.9.1.3 Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3-sulphamoyl-5-nitro benzoic acid:*

The BOC substituted product was then dissolved in 10ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours. The final product was obtained by evaporating the reaction mixture to obtain the brown colour product. It was then recrystallised from dichloromethane and dried.

Yield = 0.08g or 0.202 mmol or 40%

Mp: 195.1 – 199.8  $^{\circ}\text{C}$

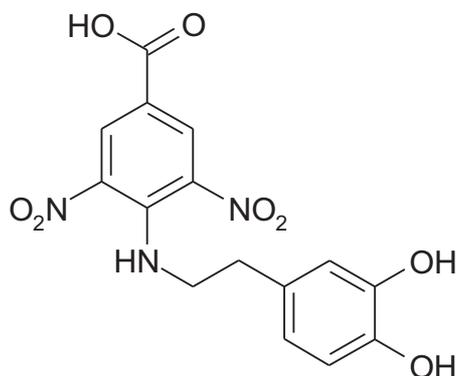
TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.44 (single spot).

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 13.1 (bs, 1H, OH); 8.44 (d, 1H,  $J$  = 1.896 Hz, Ph-6H); 8.32 (d, 1H,  $J$  = 1.896 Hz, Ph-2H); 8.00 (bs, 2H,  $\text{NH}_2$ ); 7.2 (dd, 1H,  $J$  = 8.214 Hz, Ph-4H); 6.8 (d, 1H,  $J$  = 1.896 Hz, Ph-5H); 6.8 (dd, 2H,  $J$  = 1.896, 8.214 Hz, Ph-6H); 3.064 (m, 2H,  $\text{CH}_2$ ); 2.72 (dd,  $J$  = 2.52, 8.84 Hz, 2H,  $\text{CH}_2$ ) ppm.

IR (KBr)  $\nu$  = 3433, 3266, 3087, 2957, 2834 ( $-\text{OCH}_3$ ), 2622, 2541, 2300, 1938, 1695, 1610, 1520, 1410 ( $-\text{NO}_2$ ), 1407, 1325 ( $-\text{SO}_2\text{NH}_2$ )  $\text{cm}^{-1}$

MS (ES  $-$ )  $m/z$  = found 394 (M - H) $^-$ .

#### 4.10.1. Preparation of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid: Compound 3.8.



##### 4.10.1.1 Preparation of DIEA salt of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:

0.25g or 1.016 mmol of 3, 5-dinitro-4-chloro benzoic acid was dissolved in 25ml of butanol and 0.33g or 2.557 mmol of di-isopropylethylamine was added. 0.185g or 0.978 mmol of dopamine was then added into the reaction mixture and left reacted under reflux for 16 hours under argon. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.14g, 386mmol

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.61 (s, 2H, Ph-H ); 8.45 (t, 1H, NH ); 6.69 (d, 1H, J= 7.58 Hz, Ph-2H); 6.62 (d, 1H, J= 1.896 Hz, Ph-5H); 6.45 (dd, 1H, J= 1.896, 7.58 Hz, Ph-6H); 3.13 (m, 2H,  $\text{CH}_2$ ); 2.77 (m, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.83 (single spot).

##### 4.10.1.2 Preparation of BOC substituted 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:

The product was then dissolved in 10ml acetonitrile and 0.158g or 0.722mmol of BOC anhydride was added and left reacted for 48 hours under argon. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.171g

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.61 (s, 2H, Ph-H ); 8.45 (t, 1H, NH ); 6.69 (d, 1H, J= 7.58 Hz, Ph-2H); 6.62 (d, 1H, J= 1.896 Hz, Ph-5H); 6.45 (dd, 1H, J= 1.896, 7.58 Hz, Ph-6H); 3.13 (m, 2H,  $\text{CH}_2$ ); 2.77 (m, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, *tert*-Butyl, BOC) ppm.

IR (KBr)  $\nu$  = 3416, 3316, 3171, 2954, 2282, 1868, 1690, 1621, 1595, 1520, 1410, 1407, 1320  $\text{cm}^{-1}$

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.44 (tailing).

#### 4.10.1.3 Preparation of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours under argon. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product.

Yield = 0.12g or 0.330 mmol or 34%

Mp: 206.8 – 210.1 °C.

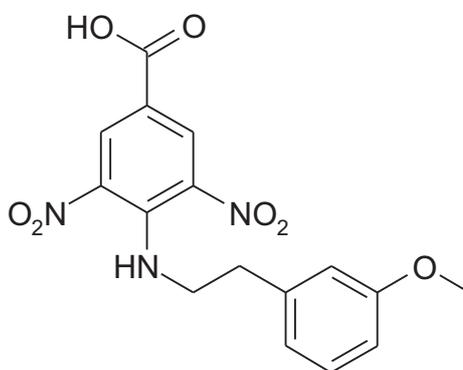
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.61 (s, 2H, Ph-H); 8.45 (t, 1H, NH); 6.69 (d, 1H,  $J$  = 7.58 Hz, Ph-2H); 6.62 (d, 1H,  $J$  = 1.896 Hz, Ph-5H); 6.45 (dd, 1H,  $J$  = 1.896, 7.58 Hz, Ph-6H); 3.13 (m, 2H,  $\text{CH}_2$ ); 2.77 (m, 2H,  $\text{CH}_2$ ) ppm.

MS (ES<sup>-</sup>)  $m/z$  = found 362 (M - H)<sup>-</sup>.

IR (KBr)  $\nu$  = 3418, 3305, 3198, 2965, 2641, 2287, 1872, 1690, 1621, 1595, 1520, 1410, 1407, 1320  $\text{cm}^{-1}$

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0..75 (long tailing).

#### 4.11.1 Preparation of 4-(3-methoxyphenethylamino)-3, 5-dinitro benzoic acid: Compound 3.12.



##### 4.11.1.1 Preparation of DIEA salt of 4-(3-methoxyphenethylamino)-3,5-dinitro benzoic acid:

0.25g or 1.016 mmol of 3, 5-dinitro-4-chloro benzoic acid was dissolved in 20ml of butanol and 0.33g or 2.557 mmol of di-isopropylethylamine was added. 0.159g or 1.052 mmol of 3-methoxyphenethylamine was then added into the reaction mixture and left reacted under reflux for 18 hours. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.861g

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.57 (s, 2H, Ph-H ); 8.33 (t, 1H, NH ); 7.23 (ddd, 1H, J= 1.26, 7.58, 8.21 Hz Ph-H); 6.79 (overlapping d, 3H, Ph-H); 3.71(s, 3H, methoxy); 3.25 (m, 2H,  $\text{CH}_2$ ); 2.93 (m, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.2 (single spot).

#### 4.11.1.2 Preparation of BOC substituted 4-(3-methoxyphenethylamino)-3,5-dinitro benzoic acid:

The product was then dissolved in 15ml acetonitrile and 0.12g or 0.549mmol of BOC anhydride was added and left reacted for 48 hours. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.171g

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.57 (s, 2H, Ph-H ); 8.33 (t, 1H, NH ); 7.23 (ddd, 1H, J= 1.26, 7.58, 8.21 Hz Ph-H); 6.79 (overlapping d, 3H, Ph-H); 3.71(s, 3H, methoxy); 3.25 (m, 2H,  $\text{CH}_2$ ); 2.93 (m, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, *tert*-Butyl, BOC) ppm.

IR (KBr)  $\nu$  = 3416, 3316, 3171, 2954, 2282, 1868, 1690, 1621, 1595, 1520, 1410, 1407, 1320  $\text{cm}^{-1}$

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.2 (single spot).

#### 4.11.1.3 Preparation of 4-(3-methoxyphenethylamino)-3,5-dinitro benzoic acid:

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product which was then re-crystallized from ether.

Yield = 0.107 or 0.277 mmol or 30%

Mp: 142.9 – 144.4  $^{\circ}\text{C}$ .

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.57 (s, 2H, Ph-H); 8.33 (t, 1H, NH); 7.23 (ddd, 1H, J= 1.26, 7.58, 8.21 Hz Ph-H); 6.79 (overlapping d, 3H, Ph-H); 3.71(s, 3H, methoxy); 3.25 (m, 2H,  $\text{CH}_2$ ); 2.93 (m, 2H,  $\text{CH}_2$ ) ppm.

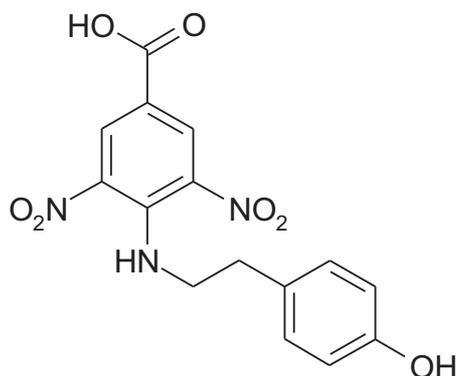
MS (ES  $-$ )  $m/z$  = found 361 (M - H)  $^-$ .

Accurate MS (-ES) = Calculated 362.0624 (m-H)  $^-$  and 362.0626 (m-H)  $^-$  found

IR (KBr)  $\nu$  = 3418, 3305, 3198, 3090, 2965, 2658, 2287, 1872, 1690, 1621, 1595, 1520, 1410, 1407, 1320  $\text{cm}^{-1}$

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.40 (single spot).

#### 4.12.1. Preparation of 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid : Compound 3.9.



##### 4.12.1.1. Preparation of DIEA salt of 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid:

0.246g or 1.00 mmol of 3,5-dinitro-4-chloro benzoic acid was dissolved in 25ml of butanol and 0.33g or 2.881 mmol of di-isopropylethylamine was added. 0.171g or 0.984 mmol of 4-hydroxyphenethylamine was then added into the reaction mixture and left reacted under reflux for 18 hours. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.37g

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 9.3 (s, 1H, OH); 8.58 (s, 2H, Ph-H); 8.34 (t, 1H, NH); 7.02 (d, 2H, J= 8.214 Hz Ph-H); 6.71 (d, 2H, J= 8.214 Hz Ph-H); 3.25 (m, 2H,  $\text{CH}_2$ ); 2.87 (m, 2H,  $\text{CH}_2$ ) ppm.

MS (ES<sup>-</sup>)  $m/z$  = found 361 (M - H)<sup>-</sup>.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.52 (tailing).

##### 4.12.1.2. Preparation of BOC substituted 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid:

The product was then dissolved in 25ml acetonitrile and 0.24g or 1.098 mmol of BOC anhydride was added and left reacted for 48 hours. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.43g

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 9.3 (s, 1H, OH); 8.58 (s, 2H, Ph-H); 8.34 (t, 1H, NH); 7.2 (bs, 2H,  $\text{NH}_2$ ); 7.02 (d, 2H, J= 8.214 Hz Ph-H); 6.71 (d, 2H, J= 8.214 Hz Ph-H); 3.25 (m, 2H,  $\text{CH}_2$ ); 2.81 (m, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, *tert*-Butyl, BOC) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.54 (single spot).

#### 4.12.1.3. Preparation of 4-(4-hydroxyphenethylamino)-3,5-dinitro benzoic acid:

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product which was then recrystallised from DCM.

Yield = 0.15g or 0.432 mmol or 61%

Mp: 219.7 – 222.6 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 9.3 (s, 1H, OH); 8.58 (s, 2H, Ph-H); 8.34 (t, 1H, NH ); 7.02 (d, 2H, J= 8.214 Hz Ph-H); 6.71 (d, 2H, J= 8.214 Hz Ph-H); 3.25 (m, 2H, CH<sub>2</sub>); 2.87 (m, 2H, CH<sub>2</sub>) ppm.

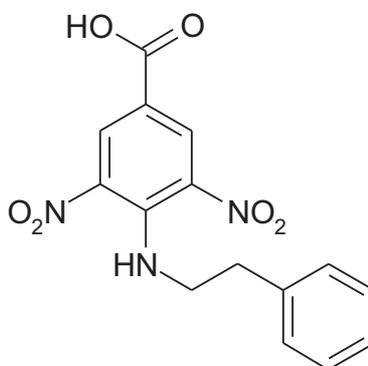
IR (KBr) ν = 3490, 3324, 3114, 2923, 2641, 2535, 2350, 2285, 1891, 1789, 1683, 1610, 1520, 1410, 1407, 1320 cm<sup>-1</sup>

MS (ES -) m/z = found 346 (M - H)<sup>-</sup> . .

Accurate MS (-ES) = Calculated 346.0675 (m-H)<sup>-</sup> and 346.0674 (m-H)<sup>-</sup> found.

TLC carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.62 (tailing).

#### 4.13.1. Preparation of 4-phenethylamino-3,5-dinitro benzoic acid : Compound 3.11.



##### 4.13.1.1. Preparation of DIEA salt of 4-phenethylamino-3,5-dinitro benzoic acid:

0.247g or 1.01 mmol of 3, 5-dinitro-4-chloro benzoic acid was dissolved in 25ml of butanol and 0.35g or 3.055 mmol of di-isopropylethylamine was added. 0.128g or 1.057 mmol of phenethylamine was then added into the reaction mixture and left reacted under reflux for 18 hours. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.73g

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 9.3 (s, 1H, OH); 8.58 (s, 2H, Ph-H ); 8.34 (t, 1H, NH ); 7.33 (m, 5H, Ph-H); 3.22 (m, 2H, CH<sub>2</sub>); 2.96 (m, 2H, CH<sub>2</sub>) ppm.

TLC carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.59 (single spot).

#### 4.13.1.2. Preparation of BOC substituted 4-phenethylamine-3,5-dinitro benzoic acid:

The product was then dissolved in 25ml acetonitrile and 0.24g or 1.098 mmol of BOC anhydride was added and left reacted for 48 hours. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.562g.

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.3 (s, 1H, OH); 8.58 (s, 2H, Ph-H); 8.34 (t, 1H, NH); 7.33 (m, 5H, Ph-H); 3.22 (m, 2H,  $\text{CH}_2$ ); 2.96 (m, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, *tert*-Butyl, BOC); ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.54 (single spot).

#### 4.13.1.3. Preparation of 4-phenethylamine-3,5-dinitro benzoic acid:

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 18 hours. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product which was then recrystallised from ether.

Yield = 0.145g or 0.438 mmol or 59%

M.p: 186.8 – 190.1  $^{\circ}\text{C}$ .

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.3 (s, 1H, OH); 8.58 (s, 2H, Ph-H); 8.34 (t, 1H, NH); 7.33 (m, 5H, Ph-H); 3.22 (m, 2H,  $\text{CH}_2$ ); 2.96 (m, 2H,  $\text{CH}_2$ ); ppm.

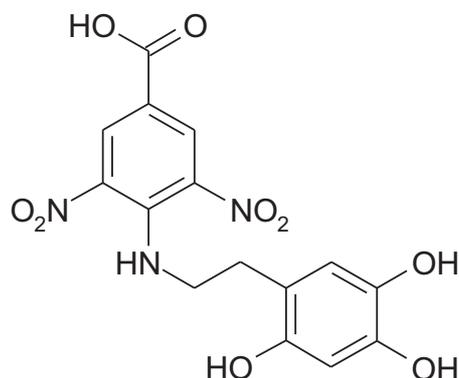
IR (KBr)  $\nu$  = 3420, 3316, 3104, 2923, 2874, 2628, 2531, 2350, 2296, 1958, 1858, 1677, 1610, 1520, 1410, 1407, 1320  $\text{cm}^{-1}$

MS (ES<sup>-</sup>)  $m/z$  = found 330 ( $\text{M} - \text{H}$ )<sup>-</sup>.

Accurate MS (-ES) = Calculated 330.0726 ( $\text{m} - \text{H}$ )<sup>-</sup> and 330.0728 ( $\text{m} - \text{H}$ )<sup>-</sup> found

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.5 (single spot).

#### 4.14.1 Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3,5-dinitro benzoic acid : Compound 3.13



##### 4.14.1.1. Preparation of DIEA salt of 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:

0.155g or 0.629 mmol of 3, 5-dinitro-4-chlorobenzoic acid was dissolved in 10ml of butanol and 0.3g or 2.325 mmol of diisopropylethylamine was added. 0.075g or 0.619 mmol of 3,4,6-trihydroxyphenethylamine was then added into the reaction mixture and left reacted under reflux for 16 hours under argon. The solvent was then evaporated to obtain thick brown oil.

Yield = 0.12g

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.9 (bs, 1H, OH); 8.58 (s, 2H, Ph-H); 8.3 (t, 1H, NH); 7.0 (bs, 3H, OH); 6.12 (s, 1H, Ph-H); 5.9 (s, 1H, Ph-H); 3.034 (bt, 2H,  $\text{CH}_2$ ); 2.72 (dd,  $J=8.214$  Hz, 2H,  $\text{CH}_2$ ) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.2 (tailing).

##### 4.14.1.2. Preparation of BOC substituted 4-(3,4,6-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:

The product was then dissolved in 15ml acetonitrile and 0.12g or 0.549mmol of BOC anhydride was added and left reacted for 48 hours under argon. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.05g

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.9 (bs, 1H, OH); 8.58 (s, 2H, Ph-H); 7.0 (bs, 3H, OH); 6.12 (s, 1H, Ph-H); 5.9 (s, 1H, Ph-H); 3.034 (bt, 2H,  $\text{CH}_2$ ); 2.72 (dd,  $J=8.214$  Hz, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, BOC) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.4 (tailing).

#### 14.1.3. Preparation of 4-(3,4,6-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours under argon. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product.

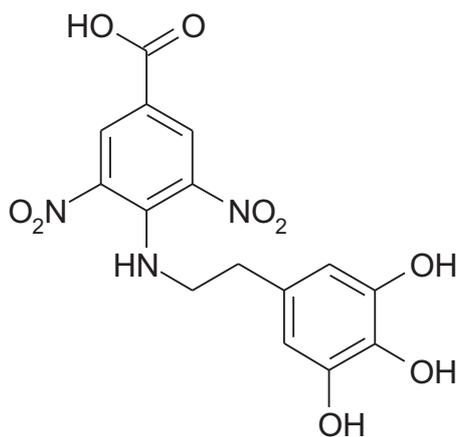
Yield = 0.02g or 10%

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 8.9 (bs, 1H, OH); 8.58 (s, 2H, Ph-H); 8.3 (t, 1H, NH); 7.0 (bs, 3H, OH); 6.12 (s, 1H, Ph-H); 5.9 (s, 1H, Ph-H); 3.034 (m, 2H,  $\text{CH}_2$ ); 2.72 (m,  $J=8.214$  Hz, 2H,  $\text{CH}_2$ ) ppm.

MS (ES<sup>-</sup>)  $m/z$  = found 378 (M - H)<sup>-</sup>.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.4 (long tailing).

#### 4.15.1 Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3,5-dinitro benzoic acid : Compound 3.9.



##### 4.15.1.1. Preparation of DIEA salt of 4-(3,4,5-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:

0.13 g or 0.528 mmol of 3, 5-dinitro-4-chlorobenzoic acid was dissolved in 10ml of butanol and 0.195g or 1.511 mmol of diisopropylethylamine was added. 0.1g or 0.487 mmol of 5-hydroxy dopamine was then added into the reaction mixture and left for reflux under argon for 16 hours under argon. The solvent was then evaporated to obtain thick yellow oil.

Yield = 0.09 g

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 9.2 (s, 1H, OH); 8.58 (s, 2H, 2H, Ph-H); 7.1 (bs, 3H, OH); 6.12 (s, 2H, Ph-H); 4.4 (bs, 1H, NH); 3.034 (m, 2H,  $\text{CH}_2$ ); 2.62 (m, 2H,  $\text{CH}_2$ ); ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.5 (long tailing).

4.15.1.2. *Preparation of BOC substituted 4-(3,4-dihydroxyphenethylamino)-3,5-dinitro benzoic acid:*

The product was then dissolved in 15ml acetonitrile and 0.12g or 0.549mmol of BOC anhydride was added and left reacted for 48 hours under argon. The solvent was then evaporated and the product was dissolved in ammonium chloride solution and then extracted into ethyl acetate to obtain the BOC substituted compound as semisolid product.

Yield = 0.04g

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 9.2 (s, 1H, OH); 8.58 (s, 2H, 2H, Ph-H); 7.1 (bs, 3H, OH); 6.12 (s, 2H, Ph-H); 4.4 (bs, 1H, NH); 3.034 (m, 2H,  $\text{CH}_2$ ); 2.62 (m, 2H,  $\text{CH}_2$ ); 1.5 (s, 9H, *tert*-Butyl, BOC) ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.5 (long tailing).

4.15.1.3. *Preparation of 4-(3,4,5-trihydroxyphenethylamino)-3,5-dinitro benzoic acid:*

The BOC substituted product was then dissolved in 15ml dichloromethane and 5ml trifluoroacetic acid was added and left reacted for 12 hours under argon. The final product was obtained by evaporating the reaction mixture to obtain the yellowish brown colour product.

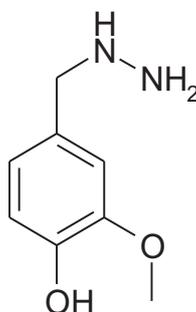
Yield = 0.065g or 32%

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 9.2 (s, 1H, OH); 8.58 (s, 2H, 2H, Ph-H); 7.1 (bs, 3H, OH); 6.12 (s, 2H, Ph-H); 4.4 (bs, 1H, NH); 3.034 (m, 2H,  $\text{CH}_2$ ); 2.62 (m, 2H,  $\text{CH}_2$ ); ppm.

MS (ES  $-$ )  $m/z$  = found 378 (M - H)  $^-$ .

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.5 (long tailing).

**4.16.1. Preparation of 3-methoxy-4-hydroxybenzylhydrazine: Compound 3.33.**



4.16.1.1 *Preparation of 3-methoxy-4-hydroxybenzylidenehydrazine:*

2g or 0.013 mol of 3-methoxy-4-hydroxybenzaldehyde was added drop wise to 9 ml of hydrazine under reflux. The reaction mixture was left reacted for 24 hours and all the solvent

was then evaporated and product was collected and recrystallised from dichloromethane. The final product was obtained as yellow amorphous powder.

Yield = 1.51 g or 70%.

$^1\text{H}$  NMR (250 MHz  $\text{d}_6\text{-DMSO}$ ): 7.61 (s, 1H,  $-\text{CH}=\text{N}-$ ); 7.11 (d, 1H,  $J= 1.269$  Hz, Ph-H); 6.8 (dd, 1H,  $J= 1.89, 8.21$  Hz, Ph-H); 6.74 (d, 1H,  $J= 7.58$  Hz, Ph-H); 6.43 (bs, 2H,  $\text{NH}_2$ ); 3.77 (s, 6H, Methoxy); ppm.

MS (ES +)  $m/z$  = found 167 ( $\text{M} + \text{H}$ ) $^+$ .

TLC carried out in ether ( $R_f$ ) = 0.9 (single spot).

#### 4.16.1.2 Preparation of 3-methoxy-4-hydroxybenzylhydrazine:

1.5g or 9.03 mmol of 3-methoxy-4-hydroxybenzylidenehydrazine was dissolved in 50 ml of methanol and 1gm of 10% Palladium charcoal was added under argon and left stirring in a hydrogenator at room temperature under 100 psi for 24 hours. The reaction mixture was then filtered through a Celite bed and the filtrate was evaporated to obtain the product as bluish green colour thick oil.

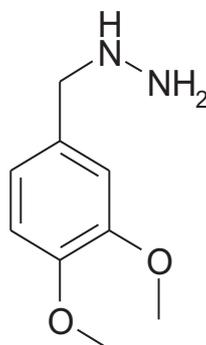
Yield = 1.18g or 78%.

$^1\text{H}$  NMR (250 MHz  $\text{d}_6\text{-DMSO}$ ): 8.66 (s, 1H, OH); 6.72 (s, 1H, Ph-H); 6.66 (d,  $J= 7.58$  Hz, 1H, Ph-H); 6.55 (d,  $J= 8.21$  Hz, 1H, Ph-H); 3.72 (s, 3H,  $\text{CH}_3$ ); 2.18 (s, 2H,  $\text{CH}_2$ ); 2.08 (s, 2H,  $\text{NH}_2$ ) ppm.

MS (ES +)  $m/z$  = found 137 ( $\text{M} - \text{NH}$ ) $^+$ , 169 ( $\text{M} + \text{H}$ ) $^+$ .

TLC carried out in ether ( $R_f$ ) = 0.77 (single spot).

#### 4.17.1. Preparation of 3,4-dimethoxybenzylhydrazine: Compound 3.32.



#### 4.17.1.1 Preparation of 3,4-dimethoxybenzylidenehydrazine:

2.01g or 1.20 mmol of 3,4-dimethoxybenzaldehyde was added drop wise to 8 ml of hydrazine under reflux. The reaction mixture was left reacted for 24 hours and all the solvent was then evaporated and product was collected as pale white amorphous powder.

Yield = 1.45 g or 80%.

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 8.65 (s, 1H, -CH=N-); 7.49 (d, 1H,  $J$ = 1.89 Hz, Ph-H); 7.4 (dd, 1H,  $J$ = 1.89, 8.84 Hz, Ph-H); 7.09 (d, 1H,  $J$ = 8.24 Hz, Ph-H); 3.83 (s, 6H, Methoxy); ppm.

MS (ES +)  $m/z$  = found 181 (M + H) $^+$ .

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.77 (single spot).

#### 4.17.1.2 Preparation of 3,4-dimethoxybenzylhydrazine:

1.03g or 5.72 mmol of 3,4-dimethoxy benzylidenehydrazine was dissolved in 50 ml of methanol and 1gm of 10% Palladium charcoal was added under argon and left stirring in a hydrogenator at room temperature under 100 psi for 24 hours. The reaction mixture was then filtered through a Celite bed and the filtrate was evaporated to obtain the product as dark yellowish thick oil.

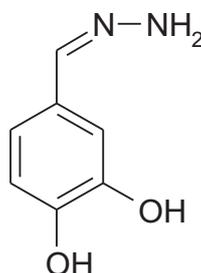
Yield = 0.8g or 65%.

$^1\text{H NMR}$  (250 MHz  $d_6$ -DMSO): 7.29 (d, 1H,  $J$ = 1.89 Hz, Ph-H); 7.09 (dd, 1H,  $J$ = 1.89, 8.84 Hz, Ph-H); 7.02 (d, 1H,  $J$ = 8.24 Hz, Ph-H); 3.83 (s, 6H, Methoxy); 3.78 (s, 2H, -CH<sub>2</sub>); ppm.

MS (ES +)  $m/z$  = found 181 (M - H) $^+$ , 151 (M - NH) $^+$ .

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.54 (single spot).

#### 4.18.1. Preparation of 3,4-dihydroxybenzylidenehydrazine: Compound 3.24.



#### 4.18.1.1 Preparation of 3,4-dihydroxybenzylidenehydrazine:

1.05g or 0.131 mmol of 3,4-dihydroxybenzaldehyde was added drop wise to 9 ml of hydrazine under reflux. The reaction mixture was left reacted for 24 hours and all the solvent was then evaporated and product was collected and recrystallised from dichloromethane. The final product was obtained as yellow amorphous powder.

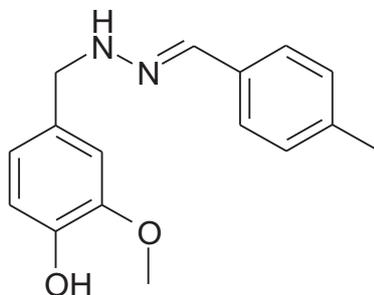
Yield = 0.8 g or 0.1925 mmol 68%.

$^1\text{H}$  NMR (250 MHz  $\text{d}_6$ -DMSO): 9.1 (bs, 2H, OH); 7.6 (s, 1H, -CH=N-); 7.4 (d, 1H,  $J$ = 1.269 Hz, Ph-H); 7.1 (dd, 1H,  $J$ = 1.89, 8.21 Hz, Ph-H); 7.04 (d, 1H,  $J$ = 7.58 Hz, Ph-H); 2.13 (bs, 2H,  $\text{NH}_2$ ); ppm.

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.65 (single spot).

MS (ES -)  $m/z$  = found 151 ( $\text{M} - \text{H}$ ) $^-$ .

#### 4.19.1. Preparation of 2-methoxy-4-((2-(4-methylbenzylidene) hydrazinyl) methyl) phenol: Compound 3.36.



0.15g of 3-methoxy-4-hydroxybenzylidenehydrazine was dissolved in 25ml toluene and 0.109g of 4-methylbenzaldehyde was added and left reacted under reflux for 18 hours. The final product was then isolated by removing the solvent by evaporation followed by re-crystallisation from ether.

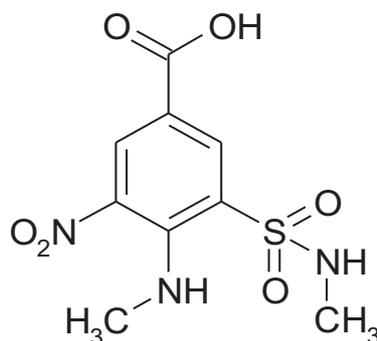
Yield: 0.11g or 76%

$^1\text{H}$  NMR (250 MHz  $\text{d}_6$ -DMSO): 8.91 (s, 1H, OH); 7.3 (d, 2H, Ph); 7.22 (s, 1H, HC=NH); 7.12 (d, 2H,  $J$ = 1.2, 1.9 Hz, Ph H); 6.83 (s, 1H, Ph-H); 6.71 (s, 2H, Ph-H); 4.32 (s, 2H,  $\text{CH}_2$ ); 3.69 (s, 3H,  $\text{CH}_3$ ); 2.27 (s, 3H,  $\text{CH}_3$ ); ppm.

MS (ES +)  $m/z$  = found 271 ( $\text{M} + \text{H}$ ) $^+$ .

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.77 (single spot).

#### 4.20.1. Preparation of 4-amino-3-(N-methylsulphamoyl)-5-nitrobenzoic acid: Compound 3.35.



4g of 4-chloro-3-chlorosulfonyl-5-nitrobenzoic acid was added slowly into 60ml of 40% methylamine solution at a temperature below 0°C and then slowly increased to room temperature and left reacted for 12 hours with continuous stirring. The final product was then collected by evaporating all the methylamine thereafter filtering the yellow solid product under vacuum.

Yield = 2.2g or 86%

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 8.35 (d, 1H, J= 2.52 Hz, Ph-H); 8.31 (d, 1H, J= 1.89 Hz, Ph-H); 8.03 (t, 1H, NH<sub>2</sub>); 7.06 (t, 1H, NH<sub>2</sub>); 7.62 (s, 1H, NH); 2.73 (d, 3H, J= 5.68 Hz, methyl), 2.45 (d, J= 5.05Hz, methyl); ppm.

IR (KBr)  $\nu$  = 3528, 3378, 2539 (-NH<sub>2</sub>), 2350, 2206, 1993, 1876, 1725, 1858, 1677, 1610-1520 (-NO<sub>2</sub>), 1477-1085 (-SO<sub>2</sub>NH<sub>2</sub>) cm<sup>-1</sup>

MS (ES -) m/z = found 288 (M - H)<sup>-</sup>.

TLC carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.55 (single spot).

## CHAPTER 5

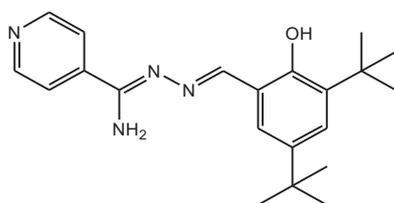
HETEROARYLCARBOXAMIDRAZONES:  
DERIVATIVES AND RELATED CHEMISTRY

## 5.1 Aim:

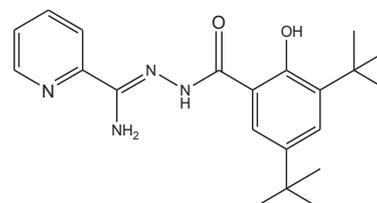
The aim of this work is to generate a library of compounds of the class heteroaryl carboxamidrazones and to screen for antimicrobial activity, particularly for anti-tubercular activity.

## 5.2 Background:

A series of  $N^1$ -arylidene-pyridine carboxyamidrazone derivatives was prepared and tested against for anti-mycobacterial activity. The 4-pyridyl (Mamalo, *et al.*, 1993) and 2-pyridyl (Mamalo, *et al.*, 1992) derivatives in particular have shown some interesting results. To further investigate the activities of these compounds a library of related benzylideneheteroaryl carboxamidrazones was prepared and tested for their antibacterial activity against various organisms. One compound in particular,  $N^1$ -[3, 5-di-(*tert*-butyl)-2-hydroxybenzylidene] - pyridine-4-carboxamidrazone (Figure 5.1 Compound 5.1), gave an intriguing antimicrobial activity profile. The compound was found to be very active against *Mycobacterium tuberculosis*; the minimum inhibitory concentration (MIC) was found to be 2-4  $\mu\text{g/mL}$  (Rathbone, *et al.*, 2006). It has also exhibited strong potency against many gram positive organisms. Unfortunately this compound was also found to be very toxic against human leucocytes causing 100% cell lysis (Coleman, M., *et al.*, 1999).



Compound 5.1



Compound 5.2

Figure 5.1: Compound 5.1 represents  $N^1$ -[3, 5-di-(*tert*-butyl)-2-hydroxybenzylidene] - pyridine-4-carboxamidrazone with MIC 2-4  $\mu\text{g/mL}$ . Compound 5.2 is pyridine-2-carboxamidrazone- $N^1$ -[3, 5-di-(*tert*-butyl)-2-hydroxybenzoyl] amide with MIC <6.25  $\mu\text{g/mL}$

Following the studies done earlier related to compound **5.1**, another compound Pyridine-2-carboxamidrazone N<sup>1</sup>-[3, 5-di-(*tert*-butyl)-2-hydroxybenzoyl] amide, (Fig 1 Compound **5.2**) exhibited potential activity against *Mycobacterium tuberculosis*. The minimum inhibitory concentration (MIC) was found to be <6.25 µg/mL causing a 100% inhibition (Begum N, *et al.*, 2007). Hence this heteroaryl carboxamidrazone and related compounds represent a useful area in the discovery of new antibacterial agents.

### 5.3 Structural activity relationship studies:

The 4-pyridyl group of compound **5.1** was substituted for 2-pyridyl and 3-pyridyl, resulting in loss of activity (Tims. K, *et al.*, 2002). Further replacing the 4-pyridyl group with pyrazinyl or 2-quinolyl abolished the antimicrobial activity (Rathbone, *et al.*, 2006). Moving the hydroxyl group from 2-position to 4-position (isomer of Compound 5.1) deactivated the compound giving an MIC > 256 µg/mL. The activity was also lost by replacing the OH at the 2 position with methyl ether; however modifying the OH into methyl ester preserved antibacterial activity MIC 4-8 µg/mL (Rathbone *et al.* 2006) though the ester may simply be acting as a prodrug for the phenol. Reducing the size of the *tert*-butyl group to methyl group again abolished the antimicrobial activity. Interestingly substituting an iodine atom instead of the *tert*-butyl group preserved the antibacterial activity, but four times less potent than Compound **5.1** (Tims. *et al.*, 2002). Moderate activity was observed by polyhydroxylated compounds and compounds with a nitro-group para to the hydroxyl group (Rathbone, *et al.* 2006). Further investigation was also made at the 3-pyridyl derivatives with N-oxide derivatives with no success (Tims. *et al.*, 2002). These suggest that the polar substitution on the benzylidene group is perhaps key for its antibacterial activity. On the other hand the 4-pyridyl moiety of the Compound **5.1** was so far the best group in the class of compounds with anti-tubercular activity. In another investigation to replace the benzene ring with a 5-nitro furan ring, it was observed that the resultant compound has been found to be very active against *Mycobacterium fortuitum* with MIC 0.1-2 µg/mL (Begum. N, 2007). In the current research efforts have been made to modify the linkages between the 4-pyridyl group and the benzylidene group so as to design compounds with similar or more potency toward its anti-tubercular effect and less toxic to human leucocytes.

## 5.4 Optimisation of the lead compound:

### 5.4.1 Optimisation of lead compound following Compound 5.1:

In an attempt to optimise the lead compound (Figure 5.1 i.e, Compound 5.1), the analogues of the following compounds shown below (Figure 5.2) were synthesised. Analogues of Compound A, Compound B and Compound C were synthesised using the aldehydes shown in Figure 5.4.

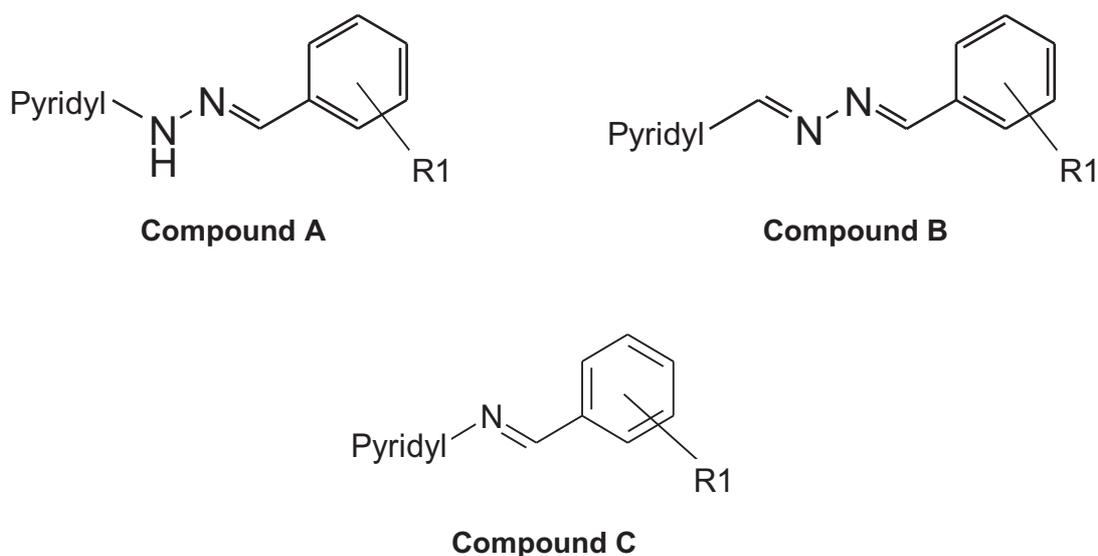
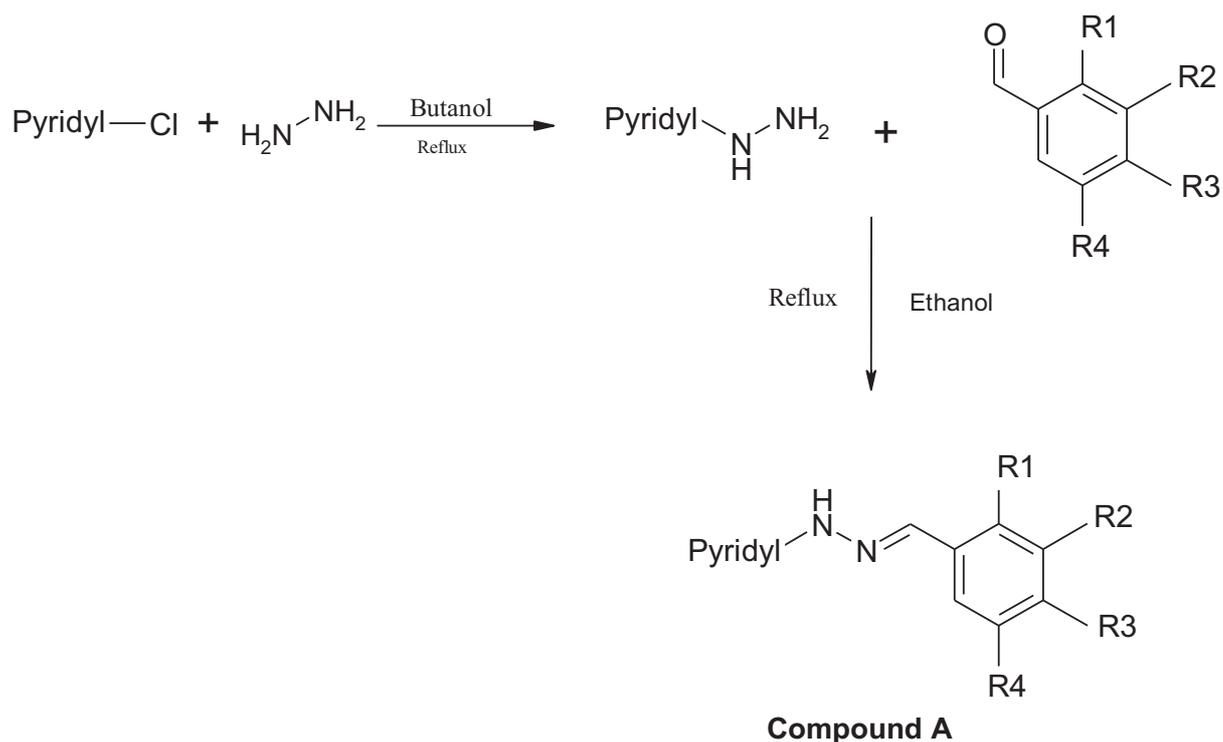


Figure 5.2: Analogues following Compound 5.1, which were designed by modifying the benzylidene group, pyridyl group and the junction between benzylidene group and pyridyl groups.

The analogues of Compound A were prepared by boiling the corresponding hydrazinopyridine with the aldehydes in ethanol at reflux (Scheme 5.1). The 4-hydrazinopyridines used in this study were prepared by boiling 4-chloropyridine with an appropriate quantity of hydrazine monohydrate in butanol (Isin, de Jonge *et al.* 2001) (Scheme 5.1). The 2-hydrazinopyridine was obtained from Sigma-Aldrich and was used directly in the reaction.

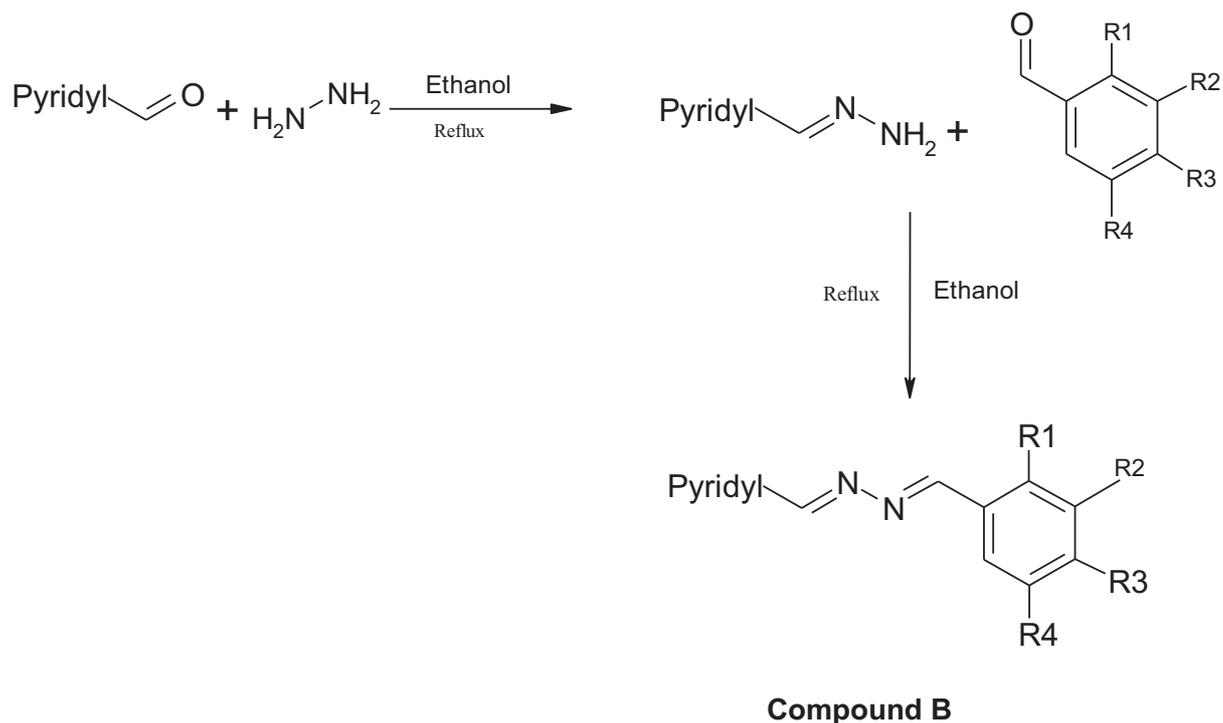


Scheme 5.1: Synthesis of hydraziropyridine followed by the condensation with aldehyde.

Compound A	Pyridyl	R1	R2	R3	R4	Yield (%)
5.3	4-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	53
5.4	4-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	69
5.5	4-pyridyl	OH	I	H	I	82
5.6	4-pyridyl	OH	Br	H	Br	63
5.7	4-pyridyl	OH	<i>tert</i> -butyl	H	H	31
5.8	4-pyridyl	OH	H	H	<i>tert</i> -butyl	37
5.9	4-pyridyl	H	<i>tert</i> -butyl	H	<i>tert</i> -butyl	77
5.10	3-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	51
5.11	3-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	17
5.12	3-pyridyl	OH	I	H	I	63
5.13	3-pyridyl	OH	Br	H	Br	54
5.14	3-pyridyl	OH	<i>tert</i> -butyl	H	H	10
5.15	3-pyridyl	OH	H	H	<i>tert</i> -butyl	29
5.16	3-pyridyl	H	<i>tert</i> -butyl	H	<i>tert</i> -butyl	33

Table 5.1: The table above shows the list of 4-pyridyl and 2-pyridyl hydrazine substituted analogues of Compound 5.1.

The analogues of Compound B were prepared by boiling the pyridyl-methylhydrazone with aldehydes at reflux (Sandbhor *et al.*, 2003). The 4-pyridyl-methylhydrazone was prepared by treating pyridine-4-carboxaldehyde with an excess hydrazine monohydrate in ethanol at 60°C as shown in Scheme 5.2 (Szmant, *et al.*, 1959). The 4-pyridyl-methylhydrazone, 3-pyridyl-methylhydrazone and 2-pyridyl-methylhydrazone were prepared similarly as shown in Scheme 5.2.



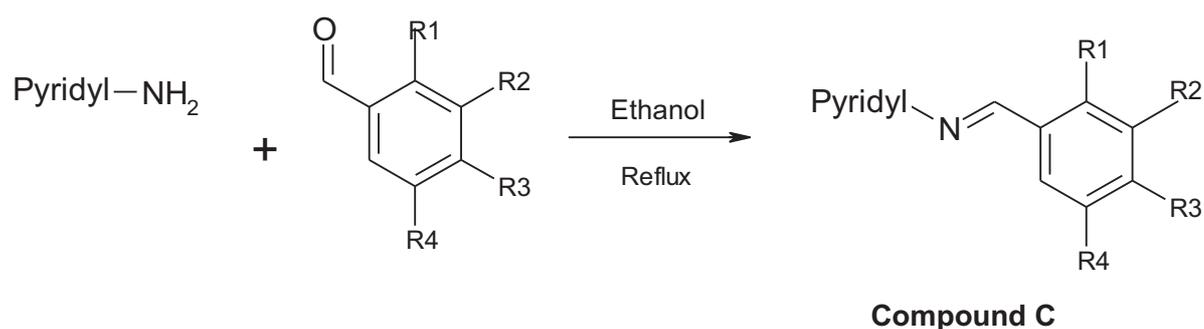
Scheme 5.2: The synthesis of pyridyl-methylhydrazone.

Compound B	Pyridyl	R1	R2	R3	R4	Yield (%)
5.17	4-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	60
5.18	4-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	65
5.19	4-pyridyl	OH	I	H	I	78
5.20	4-pyridyl	OH	Br	H	Br	62
5.21	4-pyridyl	OH	<i>tert</i> -butyl	H	H	43
5.22	4-pyridyl	OH	H	H	<i>tert</i> -butyl	65
5.23	2-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	42

Compound B	Pyridyl	R1	R2	R3	R4	Yield (%)
5.24	2-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	44
5.25	2-pyridyl	OH	<i>tert</i> -butyl	H	H	73
5.26	3-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	12
5.27	3-pyridyl	OH	I	H	I	27
5.28	3-pyridyl	OH	Br	H	Br	22

Table 5.2: List of 4-pyridyl, 3-pyridyl and 2-pyridyl methylhydrazone substituted analogues of Compound 5.1.

The analogues of Compound C were prepared by boiling the aminopyridines with the aldehydes as shown in Scheme 5. 3. The 4-aminopyridine and 3-aminopyridine were obtained commercially and used directly into the study.

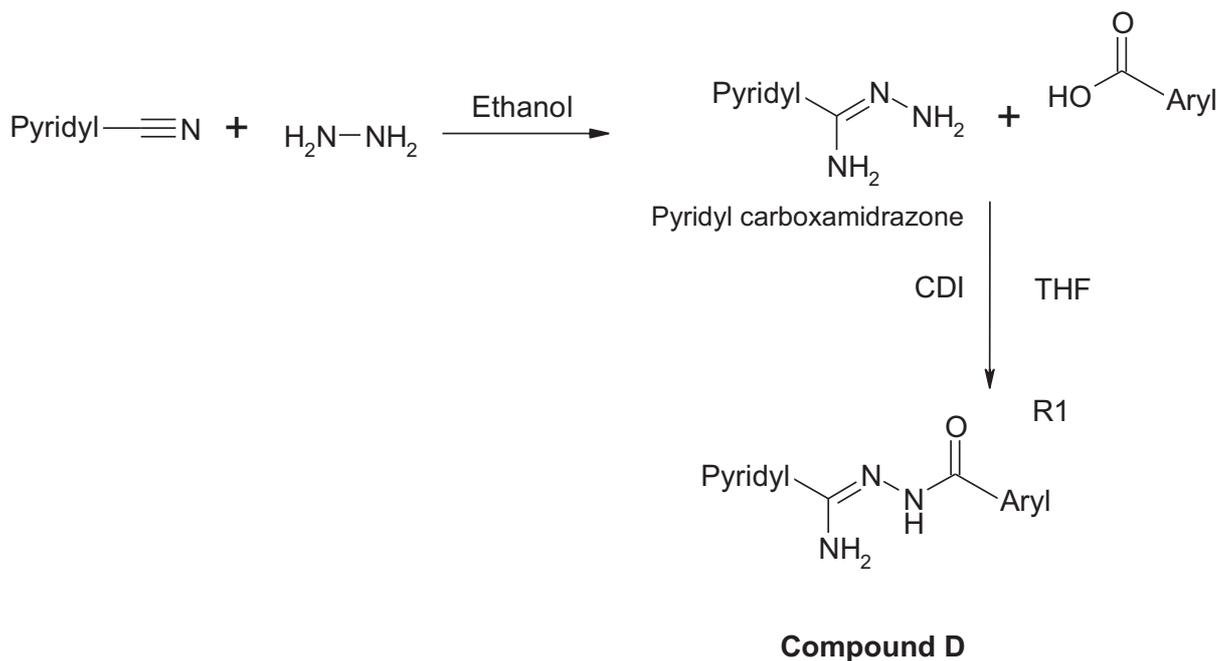


Scheme 5.3: The reaction between aldehydes and the amino pyridine where R-NH<sub>2</sub> represent the 4-aminopyridine and 3-aminopyridine.

Compound C	Pyridyl	R1	R2	R3	R4	Yield (%)
5.29	4-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	63
5.30	4-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	76
5.31	4-pyridyl	OH	I	H	I	44
5.32	4-pyridyl	OH	Br	H	Br	72
5.33	3-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	23
5.34	3-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	15
5.35	3-pyridyl	OH	I	H	I	18
5.36	3-pyridyl	OH	Br	H	Br	35



The analogues of **compound D** were prepared by treating a mixture of the acid and CDI in anhydrous THF to give the analogues of **Compound D** Scheme 5. 4 (Begum N. *et al.*, 2007).



Scheme 5.4: General method for the synthesis of pyridyl-carboxamidrazones which is then reacted with the acid to give the substituted benzylidine carboxamidrazones.

The pyridyl-4-carboxamidrazones, pyridyl-2-carboxamidrazones and pyridyl-3-carboxamidrazones were prepared in a similar way by treating the corresponding cyanopyridine with hydrazine monohydrate at room temperature with constant stirring as shown in Scheme 5. 4 above (Tanaka, Motoyama *et al.* 1994).

Compound D	Pyridyl	Aryl	Yield (%)
5.38	2-pyridyl		59
5.39	2-pyridyl		74
5.40	2-pyridyl		35
5.41	2-pyridyl		60
5.42	2-pyridyl		69

Compound D	Pyridyl	Aryl	Yield (%)
5.43	2-pyridyl		60
5.44	2-pyridyl		66
5.45	2-pyridyl		64
5.46	2-pyridyl		45
5.47	2-pyridyl		63
5.48	2-pyridyl		69
5.49	2-pyridyl		60
5.50	4-pyridyl		16
5.51	2-pyridyl		39
5.52	3-pyridyl		23
5.53	4-pyridyl		48
5.54	2-pyridyl		19
5.55	3-pyridyl		20
5.56	4-pyridyl		18
5.57	2-pyridyl		39
5.58	3-pyridyl		42
5.59	4-pyridyl		15
5.60	2-pyridyl		46
5.61	3-pyridyl		22
5.62	4-pyridyl		37
5.63	2-pyridyl		66
5.64	3-pyridyl		76
5.65	4-pyridyl		43
5.66	2-pyridyl		10
5.67	3-pyridyl		96

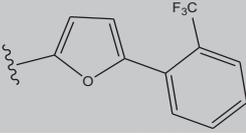
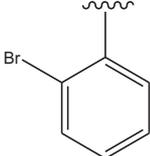
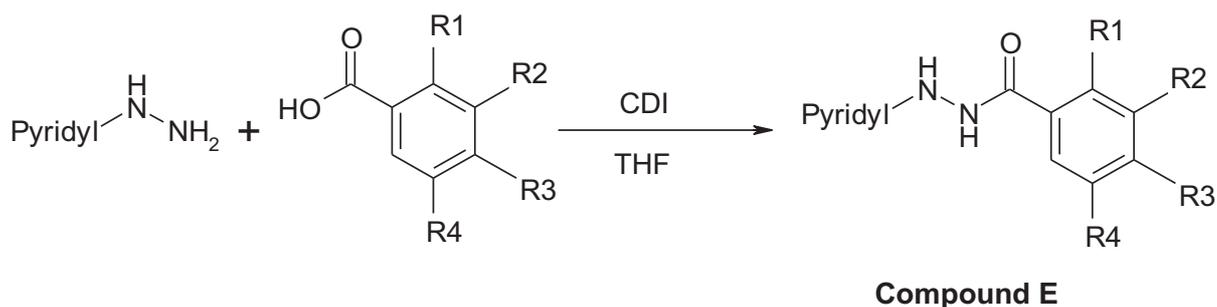
Compound D	Pyridyl	Aryl	Yield (%)
5.68	4-pyridyl		16
5.69	2-pyridyl		66
5.70	3-pyridyl		74
5.71	4-pyridyl		14

Table 5.4: List of analogues of the **compound D** prepared by varying the action of acids shown in Figure 5.3b with the pyridine-3-carboxamidrazone, pyridine-2-carboxamidrazone and pyridine-4-carboxamidrazone from Figure 5.5.

The analogues of Compound E were prepared on treating the hydrazinopyridine to the mixture of CDI and acid as shown in the Scheme 5.5 (Sandbhor *et al.*, 2003).

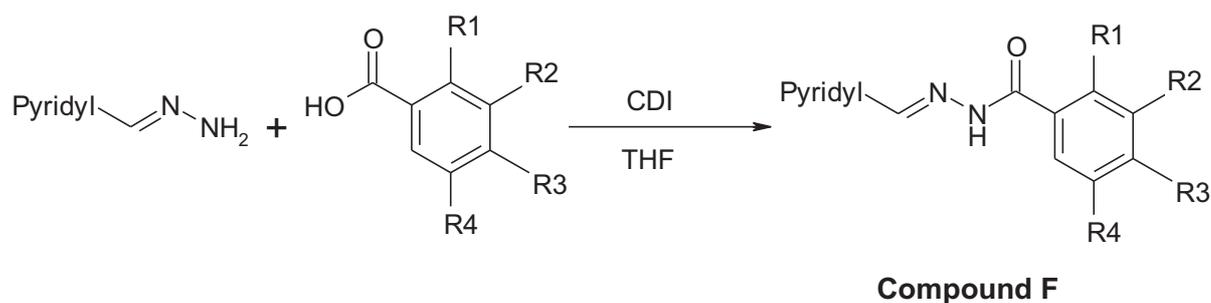


Scheme 5.5: The scheme above illustrates synthesis of the analogues of Compound E.

Compound E	Pyridyl	R1	R2	R3	R4	Yield (%)
5.72	4-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	73
5.73	4-pyridyl	OH	Br	H	Br	78
5.74	4-pyridyl	OH	I	H	I	15
5.75	3-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	68
5.76	3-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	55
5.92	2-pyridyl	OH	<i>tert</i> -butyl	H	<i>tert</i> -butyl	73
5.93	2-pyridyl	H	<i>tert</i> -butyl	OH	<i>tert</i> -butyl	65

Table 5.5: Showing the list of analogues of Compound E.

Similarly the analogues of compound F were prepared by treating the amines with the corresponding acids as shown in Scheme 5.5.

Scheme 5.6: Synthesis of the analogues of **Compound F**

Compound F	Pyridyl	R1	R2	R3	R4	Yield (%)
5.77	4-pyridyl	OH	Br	H	Br	82

Table 5.6: The table above shows the list of the compounds prepared following Compound F

### 5.5 Varying the aldehyde building blocks in the synthesis of benzylidene carboxamidrazones:

In the earlier research carried out by Rathbone *et al.*, a library of benzylidene heteroaryl carboxamidrazones was synthesised and tested for anti-tubercular activity (Rathbone, Parker *et al.* 2006). As a follow-up of the existing research the aldehydes building blocks selected in the current research were chosen to modify Compound **5.1** in the hope of increasing the biological activity and at the same time to decreasing the toxicity. The figure 5.4 below shows the list of all the aldehydes used in optimisation of the Compound **5.1**. The tert-butyl groups on the 3<sup>rd</sup> and 5<sup>th</sup> position and OH on the 2<sup>nd</sup> position on the benzyl ring have successfully shown good anti-bacterial activity in the past (Weidner-Wells, *et al.* 2001). In the recent years the benzylidene heteroaryl carboxamidrazones with such substitutions have shown anti-tubercular activity (Rathbone, Parker *et al.* 2006). The presence of iodine, chlorine or bromine at 3<sup>rd</sup> and 5<sup>th</sup> position has also shown some anti-tubercular activity (Gershon, *et al.*, 1962). It is important to note that tert-butyl group is an electron donating group, whereas iodine is a weak electron withdrawing group. Both these groups on the 3<sup>rd</sup> and 5<sup>th</sup> position on the benzylidene portion of the benzylidene carboxamidrazones have proven good anti-tubercular activity and thereby leaving a vast area into anti-tubercular drug discovery (Rathbone, Parker *et al.* 2006). Therefore in the current research apart from using the similar aldehydes such as **aa**, **ab** (isomer of **aa**) **ba** and **ac** attempt has also been made to substitute the iodine, a weak electron withdrawing group, with a range of halogens from somewhat more electronegative to highly electronegative. Therefore aldehydes with bromine, as in **ad**, chlorine as in **ae** and fluorine as in **ag** on the 3<sup>rd</sup> and 5<sup>th</sup> positions were used in this study. An

attempt was also made to incorporate both bromine and chlorine together in the same moiety **af**. Also the toxicity profile of the di-*tert*-butyl-2-hydroxy aldehyde derived compounds suggests 32% human leukocyte cell death (Coleman, Rathbone *et al.* 1999);(Coleman, Rathbone *et al.* 2001), hence the di-*tert* butyl group has also been substituted with mono-*tert*-butyl group as in **bc** and **bd**. The OH group has also been substituted with H, as in **be**. Apart from using the disubstituted halogens attempt has also been made to incorporate one halogen at a time to understand the preference of the halogens in any particular position within the aryl moiety and hence the aldehydes **bf**, **bg**, **bh**, **bi** and **bj** were used in the current investigation.

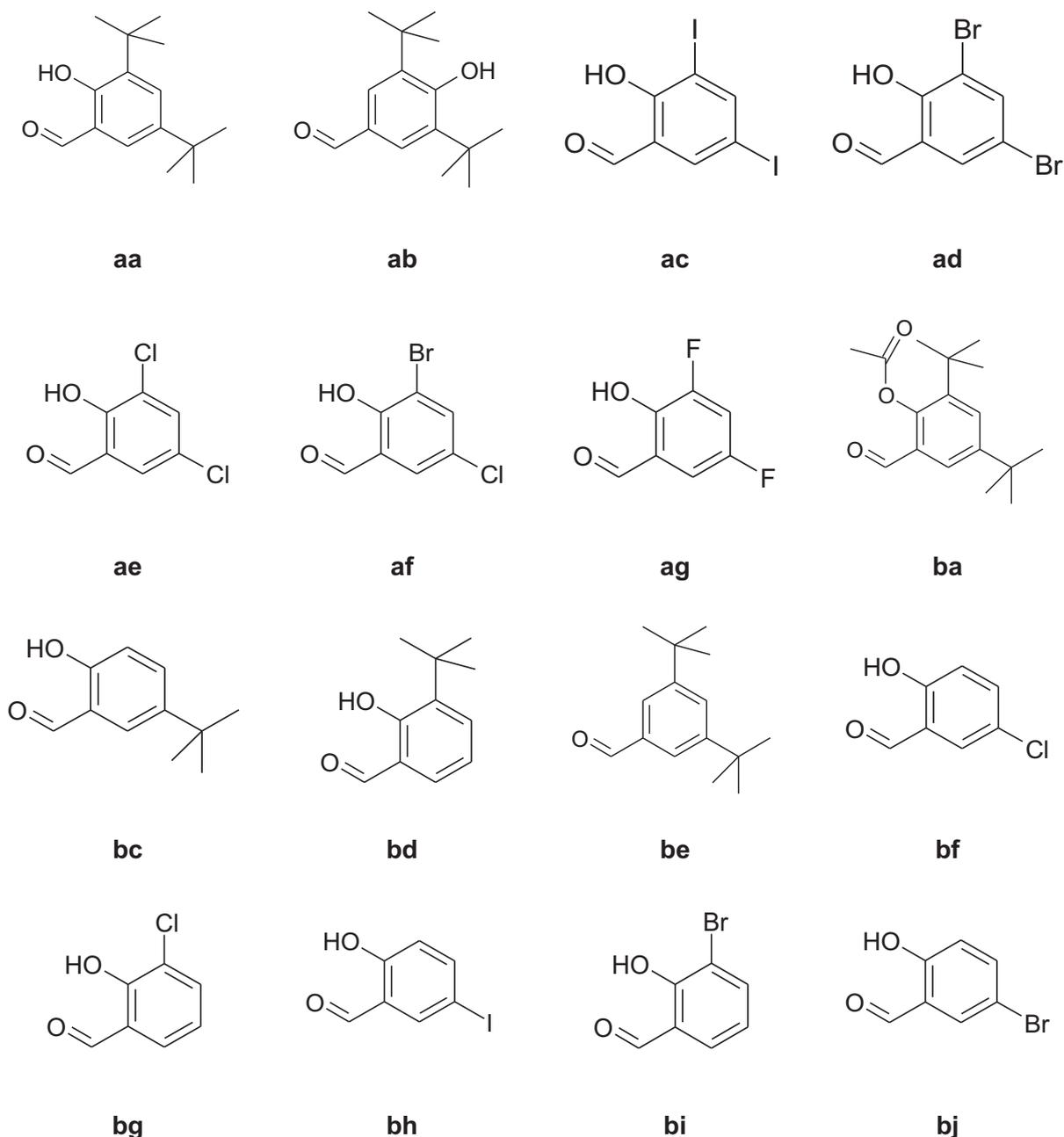
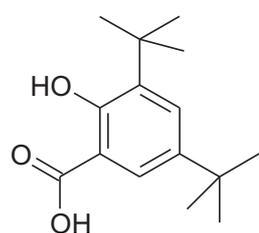
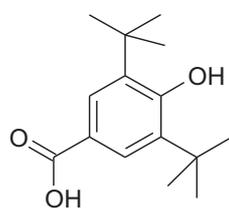
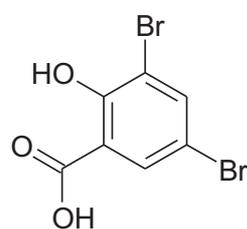
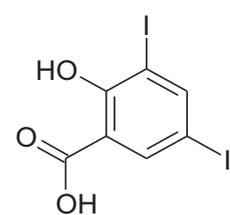
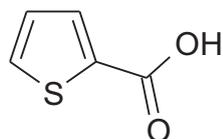
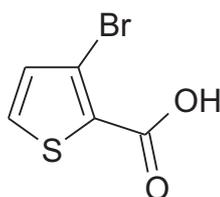
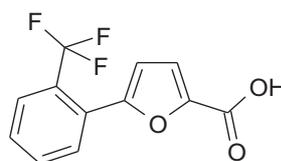
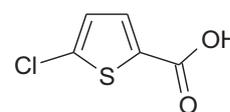
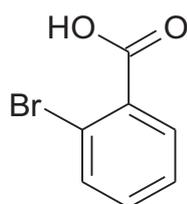
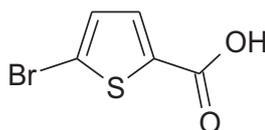
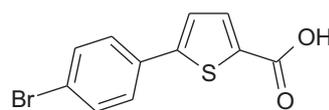
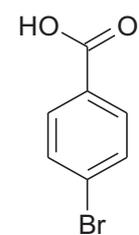
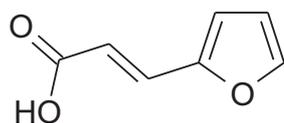
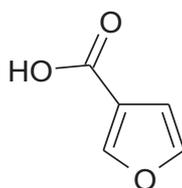
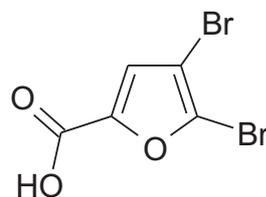
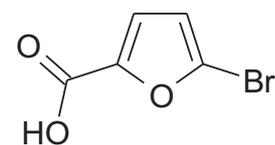


Figure 5.4: Aldehydes used in the synthesis of the analogues of Compound 5.1.

Most of the aldehydes were commercially available, apart from the **ba** which was synthesised in the laboratory (Rathbone, *et al.* 2006).

### 5.6 Varying the aryl acid building blocks in the synthesis of carboxamidrazones amides:

Apart from using various aldehydes, some acids were also used to synthesise analogues of Compound D, Compound E and Compound F.

**da****db****dc****dd****ea****eb****ec****ed****ee****ef****eg****eh****fa****fb****fc****fd**

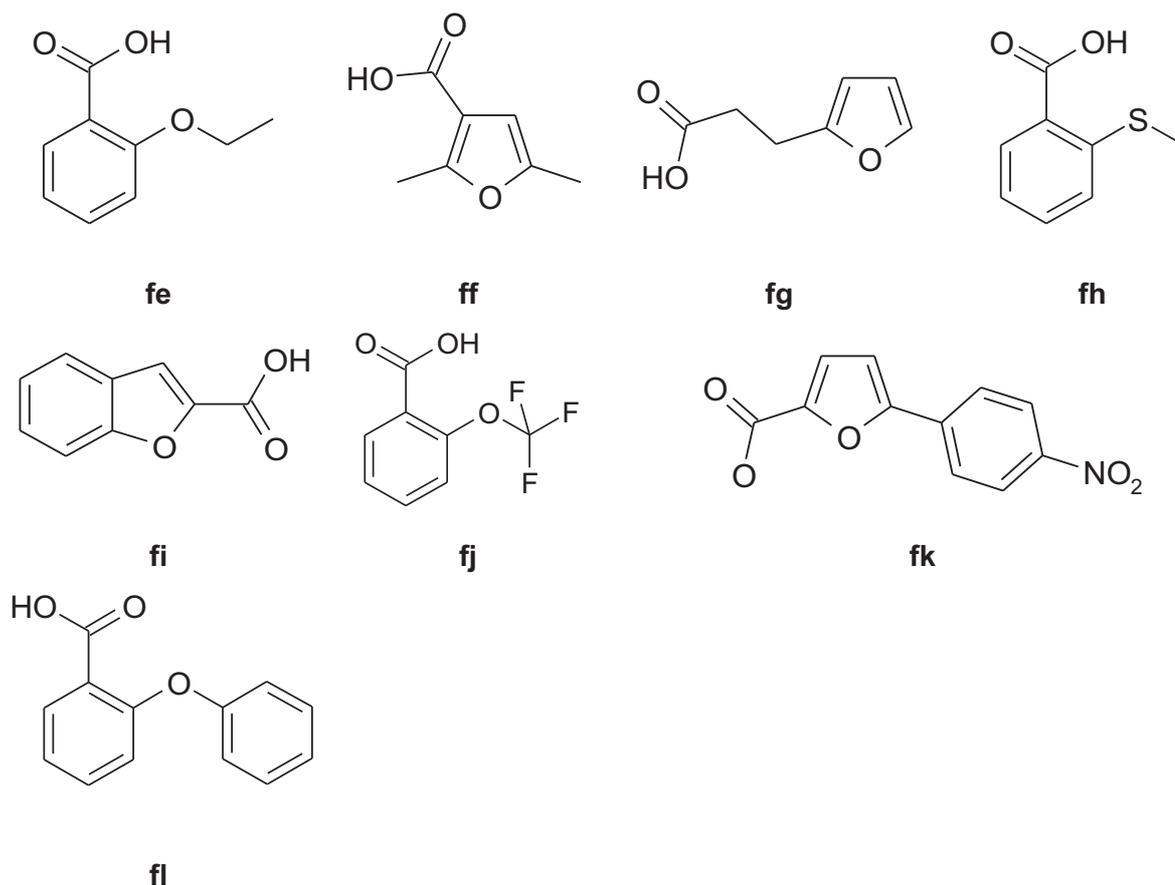


Figure 5.5: Showing the list of acids used to synthesise the analogues of Compound Y.

In the earlier research it was observed that pyridine-4-carboxamidrazone N<sup>1</sup>-[3, 5-di-(*tert*-butyl)-2-hydroxybenzoyl] amide was active against *M.tuberculosis* with a MIC of 6.25 µg/mL causing 94% inhibition (Begum. N, *et al.*, 2007). In the current research much interest has been gathered on the amide linkage of the moiety and hence synthesising analogues of Compound D. The 3, 5-di-(*tert*-butyl) part of the moiety has shown anti-microbial activity with various amidrazone moiety in the previous research (Rathbone, Parker *et al.* 2006); (Begum. N *et al.*, 2007; Tims. K *et al.*, 2002) hence **da** was used once again in the current research. Interestingly presence of iodine on the 3<sup>rd</sup> and 5<sup>th</sup> position of the aryl moiety has also shown but little activity (Rathbone, *et al.*, 2006) in the organism (*Mycobacterium tuberculosis*). Therefore in this investigation the acids containing iodine **dd**, and bromine **dc** in 3<sup>rd</sup> and 5<sup>th</sup> position of the benzyl moiety were also employed. Bromine being more electronegative than iodine was used to see the effect of stronger electronegative group in the benzyl moiety.

### 5.7 Varying the Pyridyl-amine moiety in the synthesis of analogues of pyridine carboxamidrazones:

In the earlier research, a variety of amidrazones has been used to synthesise a library of compounds (Tims. K, *et al.*, 2002). All these amidrazones (Figure 5.4) are prepared by the action of hydrazine monohydrate upon the appropriate cyano compounds (Tanaka, Motoyama *et al.* 1994).

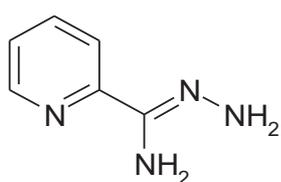
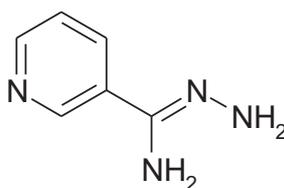
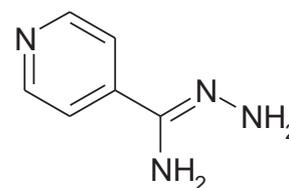
Compound **5.78**Compound **5.79**Compound **5.80**

Figure 5.6: The various pyridyl-carboxamidrazones used in the preparation of the analogues of Compound **5.2**

All these carboxamidrazones were selected based on their known antimicrobial activity. For example the aldehyde derivatives of compound 5.80 had known antimycobacterial activity (Rathbone, *et al.*, 2006). Compound **5.79** and compound **5.78** were also used because of their structural similarity to compound 5.80. In particular compound 5.78 had a structural similarity to pyrazinamide, a first line anti-tubercular drug. Compound **5.80** derivatives were found to have the best anti-tubercular activity. There was no biological activity observed in other derivatives of carboxamidrazone, therefore all other derivatives from the previous research have been omitted in the current research. The toxicity profile of the N<sup>1</sup>-[3, 5-di-(*tert*-butyl)-2-hydroxybenzylidene] - pyridine-4-carboxamidrazone, suggests 100% human leucocytes cell death. Hence in the current research work the pyridyl moiety within the carboxamidrazones has been modified (Figure 5.7) and reacted with the various aldehydes and acid derivatives (Figure 5.4 and Figure 5.5).

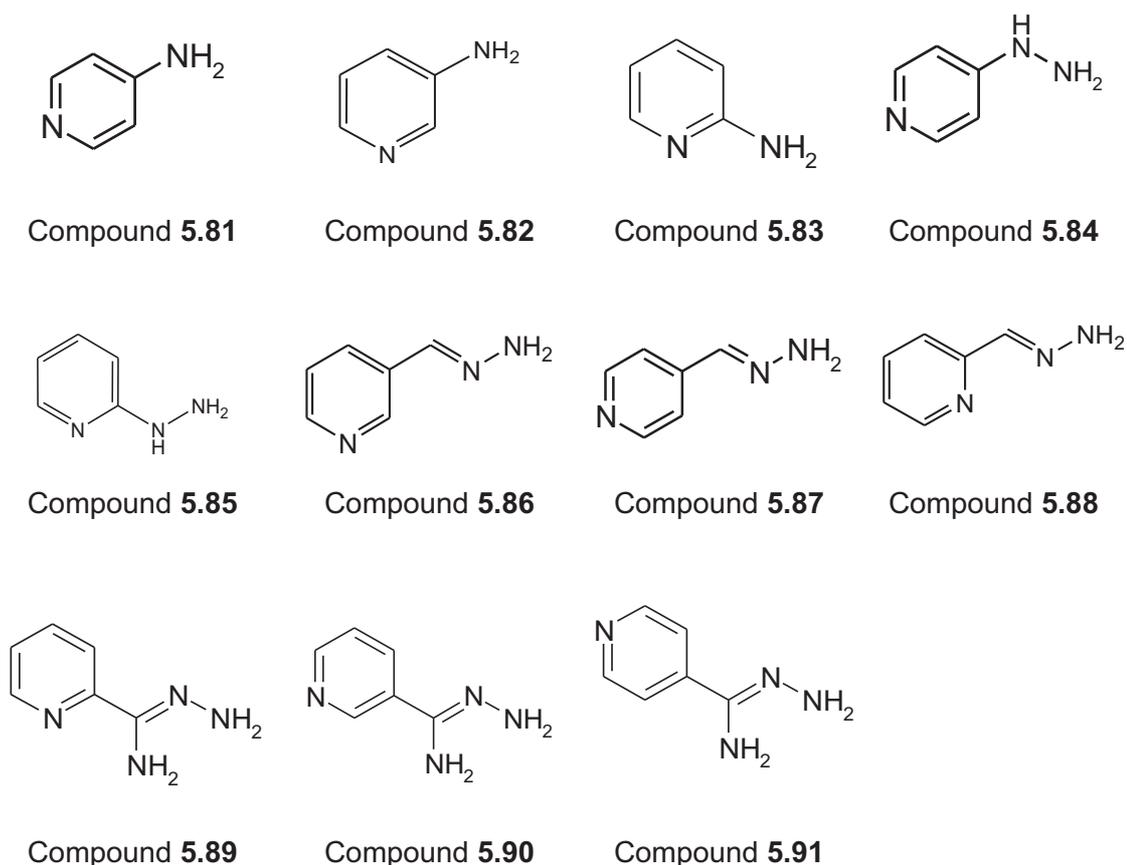


Figure 5.7: The various amino-pyridines and pyridyl carboxamidrazones used to synthesise the analogues of Compound **5.1** and Compound **5.2**

### 5.8 Reactivity of the pyridyl-amine moiety.

All the pyridyl amines listed in the Figure 5.7 appear to react with the aldehydes or the acids in different ways. The pyridyl-methylhydrazones and the hydrazinopyridines were found to be more reactive compared to the aminopyridines. The pyridyl-methylhydrazines and hydrazinopyridines reacted successfully with aldehydes (Figure 5.4) when boiled with ethanol at reflux, to synthesise the corresponding hydrazones, whereas the amino pyridines failed to follow the similar chemistry. In contrast it was necessary to melt a mixture of 4-aminopyridine and the acid (**da**) at 170-210 °C to react with the aldehydes to synthesise the corresponding amides (Matyk. J, Waisser. K, *et al.*, 2005; Waisser. K, Kunes. J, *et al.*, 2004). The reactivity of compound 5.81 was different from the compound 5.82. In case of 4-aminopyridine the lone pair of electrons is shared by the aromatic ring system and the pyridyl nitrogen as shown in Figure 5.8.

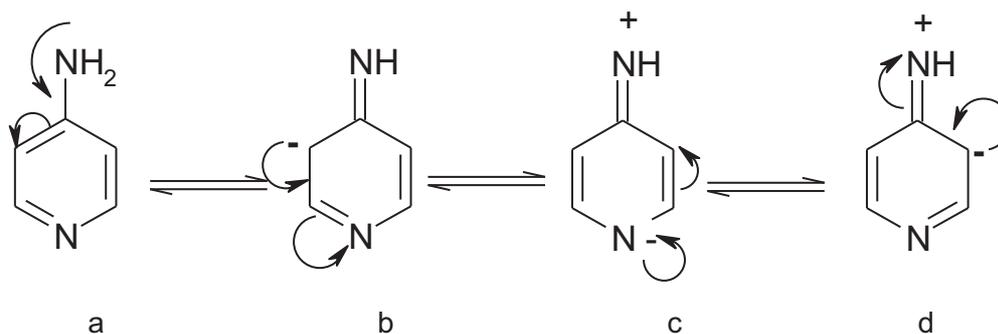


Figure 5.8: The resonance forms of 4-aminopyridine.

It was different in the case of 3-aminopyridine which was found to be more reactive than 4-aminopyridine. The negative charge of the resonance structures was only on the carbons (Figure 5.9). The lone pair on the terminal NH<sub>2</sub> is more available and less in the ring in case of the 3-amino pyridine and is therefore more reactive than 4-aminopyridine (Alam *et al.* 2000).

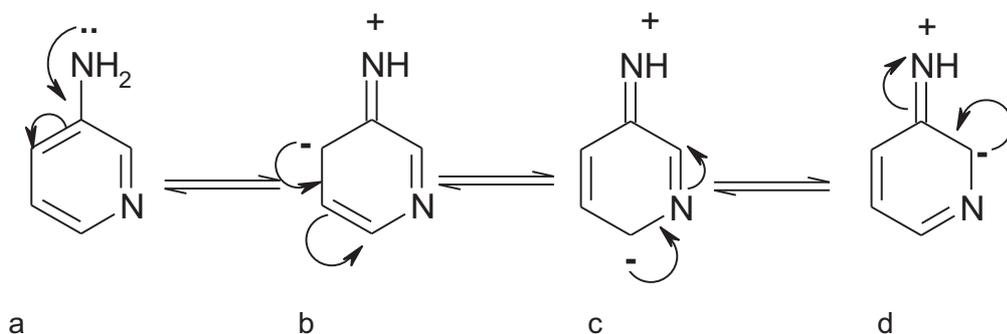


Figure 5.9: The various resonance forms of 3-aminopyridine.

On the other hand the lone pair of electrons on the terminal NH<sub>2</sub> is more easily available in case of pyridyl-methylhydrazones and hydrazinopyridines compared to the 2-amino and 3-amino pyridines as shown in Figure 5.10. In fact the hydrazinopyridines are more reactive than the pyridyl-methylhydrazones.

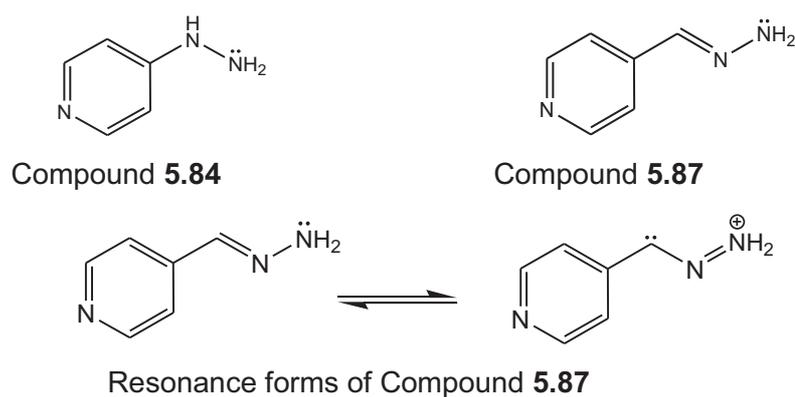
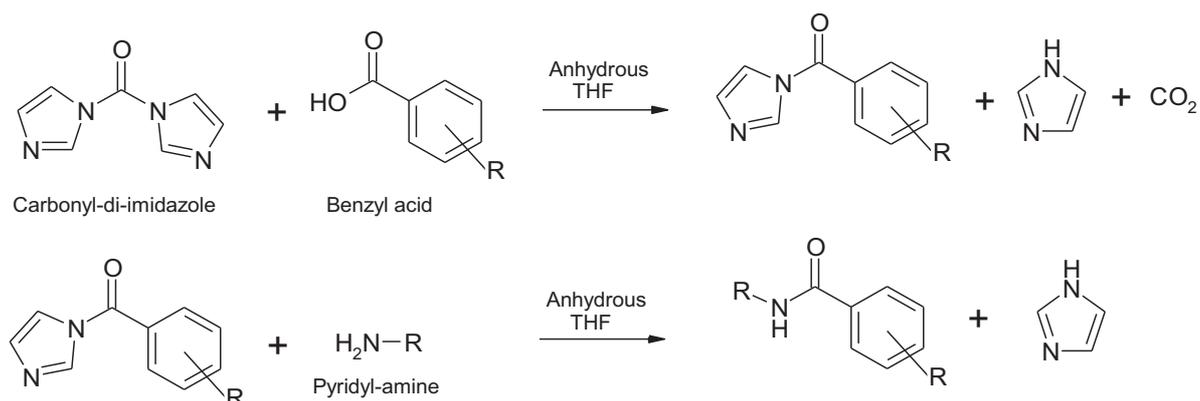


Figure 5.10: The movement of the lone pair of electrons on the pyridyl-methylhydrazine and the hydrazinopyridine.

The electron density on the terminal  $\text{NH}_2$  was found to be more in case of the compound **5.84**, than the compound **5.87**. This is because the lone pair of electrons on the  $\text{NH}_2$  of compound **5.87** gives rise to two resonance forms and therefore the electrons were not easily available (Figure 5.10). On reaction of the amines with the CDI activated acids **da**, **db**, **dc** and **dd** only the hydrazinopyridines yielded the target amides. The reaction was carried out under anhydrous condition in presence of carbonyl-di-imidazole (de Figueiredo. *et al.* 2006).



Although **da** and **dd** had electron donating and weak electron withdrawing groups, surprisingly 4-pyridyl-methylhydrazine was only successful in synthesising the corresponding amide of the acid **dc** and not **da** or **dd**. This is due to the formation of the stable acid imidazole intermediate complex with **da**. The more electron withdrawing are the substituents on the acid the more reactive is the acid-imidazole complex. Iodine in **dd** had weakly electron withdrawing property when compared with bromine **dc**, unfortunately compound 5.87 was only successful with **dc**. Perhaps compound 5.87 requires strong electron withdrawing substituents containing acid-imidazole complex to form the corresponding imide.

### 5.9 Stability of the Pyridyl-amines:

The pyridyl-amines used within this study have varying stability. Compound **5.82** and compound **5.81** are stable towards storage. Compound **5.86**, **5.87** and **5.88** were relatively unstable when compared with the amino-pyridines. Compound **5.87** was obtained as yellow oil which even on storage changes to dark colour oil. This could be due to the presence of a highly conjugated system on the moiety. Compound **5.86** and compound **5.88** were relatively stable and this is perhaps because they lack such conjugation property. The hydrazinopyridines in particular are very unstable and absorb moisture very quickly. Compound **5.84** was usually pale white solid which rapidly changes to pink on exposure to moisture and progressively converts to dark brown solid and then into semi-solid dark brown product.

All the pyridyl-carboxamidrazones have shown various stability parameters. Compound **5.90** was found to be the most unstable compound when compared with the other pyridyl carboxamidrazones (Tims, K., 2005). Compound **5.90** was very sensitive to heat and readily decomposed to a brown mass on slight increase in temperature. Compound **5.91** on the other hand is more stable compared to the compound **5.90**, although it even decomposed to a yellow and then brown mass on slight increase in temperature. Compound **5.89** was the most stable compound of all the carboxamidrazones used which could again be due to the possibility of intra-molecular hydrogen bonding as shown in figure 5.11 (Billington, D *et al.*, 2000).

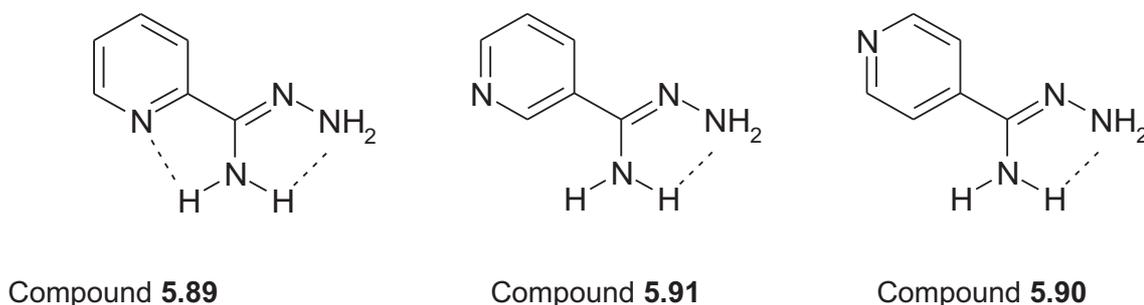


Figure 5.11: Showing the stability of the compound 5.89 due to the formation of the intra-molecular hydrogen bonding.

### 5.10 Microbiological studies:

All of these compounds were tested first for biological activity against *S.aureus*, *E.coli* and *M.fortuitum* within the Aston University laboratory. *S.aureus* and *E.coli* were used as a model to study the biological activity of these compounds against other gram positive and gram negative organisms. *Mycobacterium fortuitum* was used to screen the biological activity of these compounds against the class of *Mycobacterium*. These compounds were sent for antimicrobial activity against *Mycobacterium tuberculosis* by the Tuberculosis Antimicrobial Acquisition and Coordinating Facility ([www.TAACF.org](http://www.TAACF.org)) in U.S.

#### 5.10.1 Initial Biological Screening:

These compounds were tested for zone of inhibition on agar plate against the microorganisms. The test solutions were prepared in DMSO and 20 $\mu$ l of the compounds were placed in the wells prepared with a 6mm diameter glass cutter. The zone of inhibition was observed and the diameters of the zones were measured in mm (Table 5.7) after incubation at 37 $^{\circ}$ C for 72 hours. These compounds were also tested for the minimum inhibitory concentration (MIC). 32 $\mu$ g of the compounds were added to 25ml medium broth with 100 $\mu$ l inoculums. 1ml of this sample and 3.2 $\mu$ l of stock compound was then added to each of the test tubes and left incubating at 37 $^{\circ}$ C for 72 hours. Interestingly the compounds with larger zone of inhibition did not necessarily inhibit the organisms and did not even produce any inhibition. This suggests that the compounds had different permeability pattern which was measured as zone of inhibition.

Compound	Zone of inhibition <i>S.aureus</i>	Zone of inhibition <i>E.coli</i>	Zone of inhibition <i>M.fortuitum</i>	Growth of organism <i>M.fortuitum</i>
5.38	-	-	-	+
5.40	-	-	-	+
5.44	-	-	-	+
5.41	10	21	-	+
5.43	-	-	-	+
5.46	11	-	-	+
5.48	-	-	-	+
5.39	-	-	-	+
5.47	-	-	-	+
5.42	12	15	10	+
5.49	-	-	-	+
5.45	-	-	-	+
5.29	-	-	-	+
5.31	27	15	12	+
5.17	12	-	-	+
5.18	-	-	-	+
5.19	13	-	-	+
5.22	10	-	8	-
5.21	-	-	-	+
5.3	16	-	12	-
5.4	30	-	16	-
5.5	15	-	9	+
5.20	9	-	-	+
5.6	10	-	10	+
5.30	-	-	-	+
5.9	15	-	15	-
5.32	28	15	14	+
5.7	18	8	13	-
5.72	-	-	9	-
5.73	-	-	10	+
5.77	-	-	-	+
5.8	24	12	17	-
DMSO	-	-	-	+

Table 5.7: The zone of inhibition of the compounds versus the microorganisms (measured in mm as diameter) on a blood agar plate with 6mm diameter wells. The – sign in the ‘Zone of Inhibition’ column indicates that the compounds did not show any zone of inhibition. The + indicates, growth of the organism and – indicates no growth of the organism

With the help of the biological data further set of compounds were designed and the table 5.8 below shows the results of the further sets of the compounds.

Compounds	Zone of inhibition	Zone of inhibition
	<i>S.aureus</i>	<i>E.coli</i>
5.51	14	20
5.54	-	16
5.57	-	-
5.60	-	-
5.71	-	-
5.63	-	-
5.66	6	-
5.69	-	-
5.52	11	11
5.55	8	-
5.58	-	-
5.61	-	-
5.64	-	-
5.67	-	-
5.70	-	-
5.50	-	-
5.53	-	-
5.56	-	-
5.59	-	-
5.62	-	-
5.65	10	-
5.68	-	-
5.10	-	-
5.11	-	-
5.12	-	-
5.13	7	-
5.14	5	-
5.15	7	-
5.16	-	-
5.23	-	-
5.24	-	-
5.25	-	-
5.26	3	-
5.27	11	-
5.28	9	-
5.33	12	-
5.34	6	-
5.35	23	-
5.36	26	15
5.37	-	-
5.75	-	-
5.76	-	-

Table 5.8: The zone of inhibition of the compounds versus the microorganisms (measured in mm as diameter) on a blood agar plate with 6mm diameter wells. The (-) sign indicates that the compounds did not show any zone of inhibition.

### 5.10.2 TAACF screening:

All these compounds were sent to TAACF ([www.TAACF.org](http://www.TAACF.org)) for screening against the pathogenic organism. The screening was conducted in two designed stages. The primary screening involves the determination of the 90% inhibitory concentration (IC90). The following screening or the secondary screening involves the determination of the mammalian cell cytotoxicity (CC50).

#### 5.10.2.1 Primary Screen (Dose Response):

The initial screen was conducted against *Mycobacterium tuberculosis* H37Rv as discussed in chapter 3 (Section 3.12.1) and the data are shown in table 5.9 below. Any IC90 value of  $\leq 10\mu\text{g/mL}$  is considered "Active" and IC90 value  $< 100\mu\text{g/mL}$  is considered "Weakly Active" for antitubercular activity.

#### 5.10.2.2 Secondary Screen:

The VERO cell cytotoxicity assay was done in parallel with the TB Dose Response assay. After 72 hours exposure, viability was assessed using Promega's Cell Titer Glo Luminescent Cell Viability Assay, a homogeneous method of determining the number of viable cells in culture based on quantitation of the ATP present. Cytotoxicity was determined from the dose-response curve as the CC50 using a curve fitting program. Ultimately, the CC50 was divided by the IC90 to calculate an SI (Selectivity Index) value. SI values of  $\geq 10$  were considered for further testing.

Compound	Assay	Activity	IC50 ( $\mu\text{g/mL}$ )	IC90 ( $\mu\text{g/mL}$ )
5.3	MABA	Active	0.69	0.81
5.7	MABA	Active	1.38	1.53
5.9	MABA	Active	1.47	1.66
5.63	MABA	Active	1.63	1.86
5.57	MABA	Active	1.83	2.46
5.66	MABA	Active	2.10	2.57
5.4	MABA	Active	2.66	2.94
5.8	MABA	Active	2.90	3.15

Compound	Assay	Activity	IC50 (µg/mL)	IC90 (µg/mL)
5.60	MABA	Active	2.98	3.28
5.10	MABA	Active	3.00	3.32
5.14	MABA	Active	4.56	4.62
5.69	MABA	Active	3.82	5.00
5.42	MABA	Active	2.15	5.25
5.5	MABA	Active	5.54	6.17
5.47	MABA	Weakly Active	5.88	10.44
5.6	MABA	Weakly Active	9.12	11.22
5.15	MABA	Weakly Active	8.99	11.37
5.13	MABA	Weakly Active	6.21	11.81
5.24	MABA	Weakly Active	10.76	11.84
5.12	MABA	Weakly Active	11.22	12.60
5.72	MABA	Weakly Active	11.32	12.85
5.21	MABA	Weakly Active	11.56	13.20
5.45	MABA	Weakly Active	14.91	18.95
5.41	MABA	Weakly Active	6.27	20.54
5.43	MABA	Weakly Active	15.35	21.38
5.51	MABA	Weakly Active	9.70	23.08
5.25	MABA	Weakly Active	22.46	24.13
5.71	MABA	Weakly Active	23.74	26.91
5.68	MABA	Weakly Active	24.20	27.22
5.33	MABA	Weakly Active	26.43	55.72
5.94	MABA	Weakly Active	52.68	60.66
5.93	MABA	Weakly Active	43.60	62.20
5.35	MABA	Weakly Active	56.24	65.22
5.95	MABA	Weakly Active	63.05	68.66
5.48	MABA	Weakly Active	52.02	72.41
5.31	MABA	Weakly Active	59.28	72.80
5.38	MABA	Weakly Active	54.46	75.70
5.36	MABA	Weakly Active	64.32	77.51
5.32	MABA	Weakly Active	60.19	80.95
5.100	MABA	Weakly Active	66.59	88.33
5.73	MABA	Weakly Active	48.91	88.69
5.22	MABA	Weakly Active	75.96	89.51
5.62	MABA	Weakly Active	42.94	89.79
5.55	MABA	Weakly Active	76.69	92.66
5.28	MABA	Weakly Active	99.19	>100
5.54	MABA	Weakly Active	8.04	>100
5.59	MABA	Weakly Active	>100	>100
5.53	MABA	Weakly Active	>100	>100
5.58	MABA	Weakly Active	91.94	>100
5.61	MABA	Weakly Active	75.27	>100
5.40	MABA	Weakly Active	89.23	>100
5.44	MABA	Weakly Active	72.69	>100

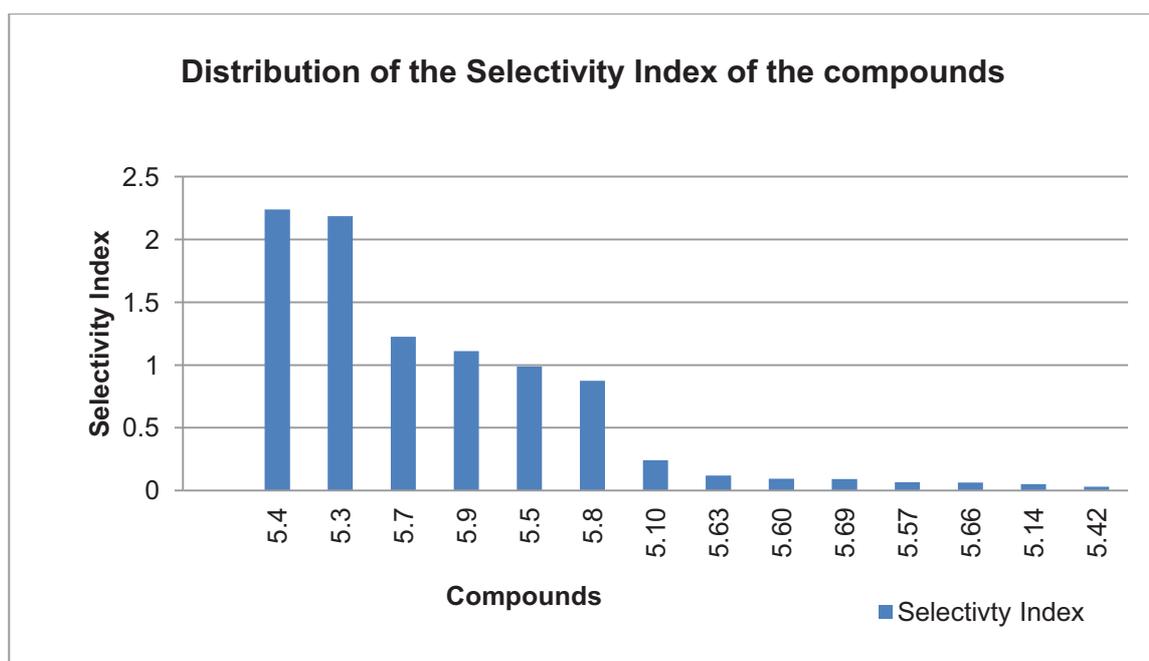
Compound	Assay	Activity	IC50 (µg/mL)	IC90 (µg/mL)
5.39	MABA	Weakly Active	47.15	>100
5.92	MABA	Weakly Active	28.08	>100
5.99	MABA	Weakly Active	95.14	>100
5.97	MABA	Weakly Active	77.38	>100
5.23	MABA	Inactive	>100	>100
5.11	MABA	Inactive	>100	>100
5.24	MABA	Inactive	>100	>100
5.26	MABA	Inactive	>100	>100
5.27	MABA	Inactive	>100	>100
5.37	MABA	Inactive	>100	>100
5.16	MABA	Inactive	>100	>100
5.34	MABA	Inactive	>100	>100
5.56	MABA	Inactive	>100	>100
5.50	MABA	Inactive	>100	>100
5.65	MABA	Inactive	>100	>100
5.67	MABA	Inactive	>100	>100
5.70	MABA	Inactive	>100	>100
5.64	MABA	Inactive	>100	>100
5.55	MABA	Inactive	>100	>100
5.102	MABA	Inactive	>100	>100
5.29	MABA	Inactive	>100	>100
5.17	MABA	Inactive	>100	>100
5.18	MABA	Inactive	>100	>100
5.19	MABA	Inactive	>100	>100
5.20	MABA	Inactive	>100	>100
5.30	MABA	Inactive	>100	>100
5.77	MABA	Inactive	>100	>100
5.74	MABA	Inactive	>100	>100
5.98	MABA	Inactive	>100	>100
5.96	MABA	Inactive	>100	>100
5.101	MABA	Inactive	>100	>100

Table 5.9: The IC50 and IC90 values of the compounds when tested against *Mycobacterium tuberculosis* at the TAACF.

Compound	MABA: H37Rv Data IC90 ( $\mu\text{g/mL}$ )	CTG: Vero Cell CC50 ( $\mu\text{g/mL}$ )	Selectivity Index
5.4	2.94	6.58	2.24
5.3	0.81	1.77	2.19
5.7	1.53	1.88	1.23
5.9	1.66	1.84	1.11
5.5	6.17	6.10	0.99
5.8	3.15	2.75	0.87
5.10	3.32	0.80	0.24
5.63	1.86	0.22	0.12
5.60	3.28	0.31	0.09
5.69	5.00	0.45	0.09
5.57	2.46	0.17	0.07
5.66	2.57	0.16	0.06
5.14	4.62	0.24	0.05
5.42	5.25	0.17	0.03

Table 5.10: The selectivity index of the compounds (CC50/ IC90).

The Selectivity indexes of the compounds were determined as the ratio between the concentration of the drug at which it inhibits 50% of the mammalian cells and the concentration at which it inhibits 90% of the bacterial cells. SI determines the relative potency of the drug to the antimicrobial activity.



Graph 5.1: The graph above showing the distribution of the selectivity index. The lower the selectivity index, the more toxic is the product relative to the antimicrobial activity.

### 5.11 Results and Discussion:

All the analogues of Compound A, Compound B, and Compound C were prepared by the reaction of the aldehydes shown in Figure 5.4 with the amino building blocks shown in Figure 5.7 in presence of ethanol as a solvent. The selection of temperature and duration of the reaction were based on the reactivity of the amines. Most of the imines were prepared by boiling a mixture of pyridyl-amine and aldehyde in ethanol at reflux. Only the analogues of Compound C were obtained by melting the amino-pyridine with the aldehyde and thereafter re-crystallizing the compound from appropriate solvent (Matyk. J, Waisser. K, *et al.*, 2005 ; Waisser. K, Kunes. J, *et al.*, 2004). As discussed in section 5.8, that those pyridylamines were less reactive when compared with all the amines used in this thesis, and therefore Compound C required higher temperature to react with the aldehydes. Interestingly the aldehyde **ba** which was successfully synthesised in laboratory was unsuccessful in preparing the final compound when reacted with the pyridyl-amines. The acetate group which was incorporated into the aldehyde was absent in the final compound. This could be that the acetate is susceptible to hydrolysis by adventitious water or perhaps to amidation by the amine used in the amine forming reaction (Figure 5.12).

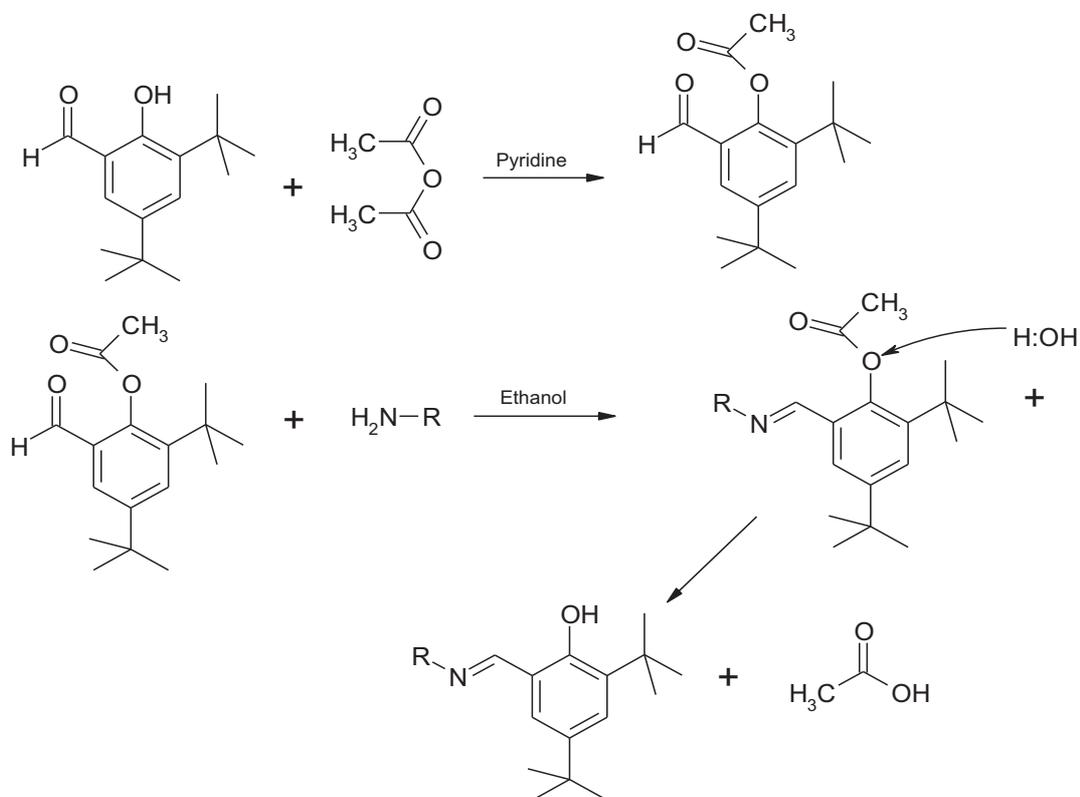


Figure 5.12: Attempted synthesis of compounds containing an acetoxy phenyl substituent.

Most of the analogues of Compound D were prepared by treating the pyridyl-amines (Figure 5.7) to a mixture of carbonyl-di-imidazole in acid shown in Figure 5.5 under anhydrous condition at room temperature as shown in Scheme 5. 4 (Begum *et al.*, 2007). Some analogues of compound D failed to respond such conditions, for example 4,5-dibromo-2-furoic acid did not react with CDI and instead it was required to convert the acid to the corresponding acid chloride using oxalyl chloride (Adams. R, *et al.*,. 1920), which was then made to react with the amine (Figure 5.13). The presence of an electronegative group made the ring highly electron-deficient and therefore the acid become less reactive toward CDI.

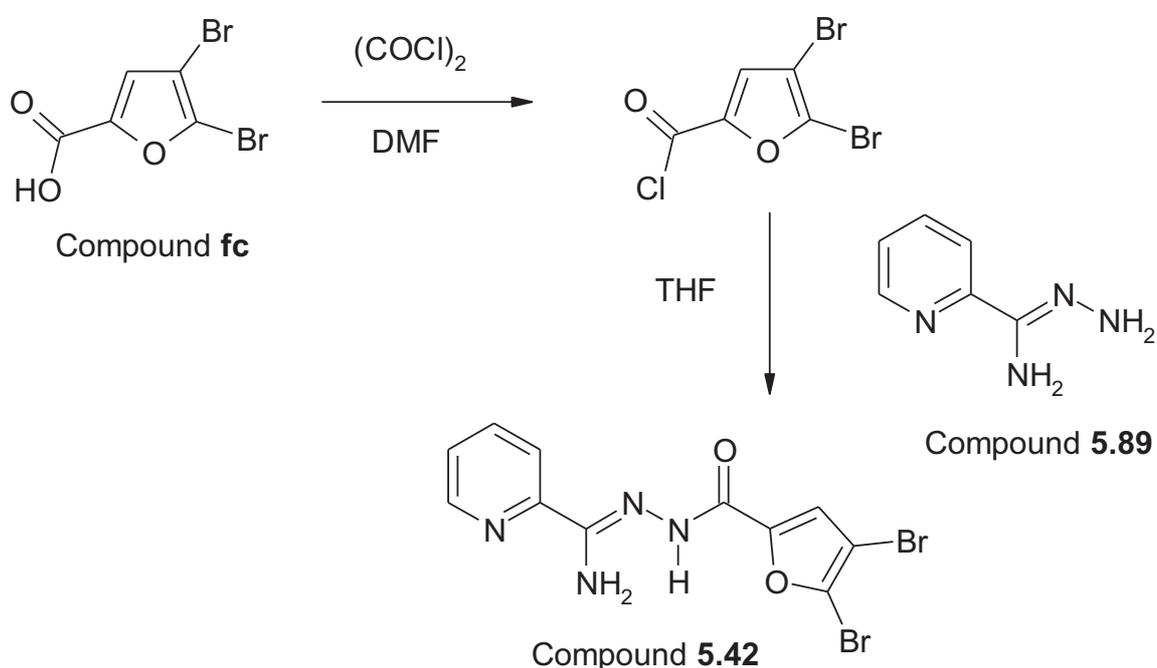


Figure 5.13: Showing the schematic conversion of dibromo-furoic acid into the corresponding acid in presence of oxalyl chloride and subsequent reaction to provide the pyridine-2-carboxamidrazone amide.

The CDI also was found unsuccessful in the approach to form the acid-imidazole complex with some of the benzoic acids containing the OH group in the 2<sup>nd</sup> position. The presence of this OH group perhaps interferes with the formation of the acid-imidazole complex. Perhaps this is due to the formation of an intra-molecular hydrogen bonding that stabilizes the acid-imidazole complex as shown in Figure 5.14.

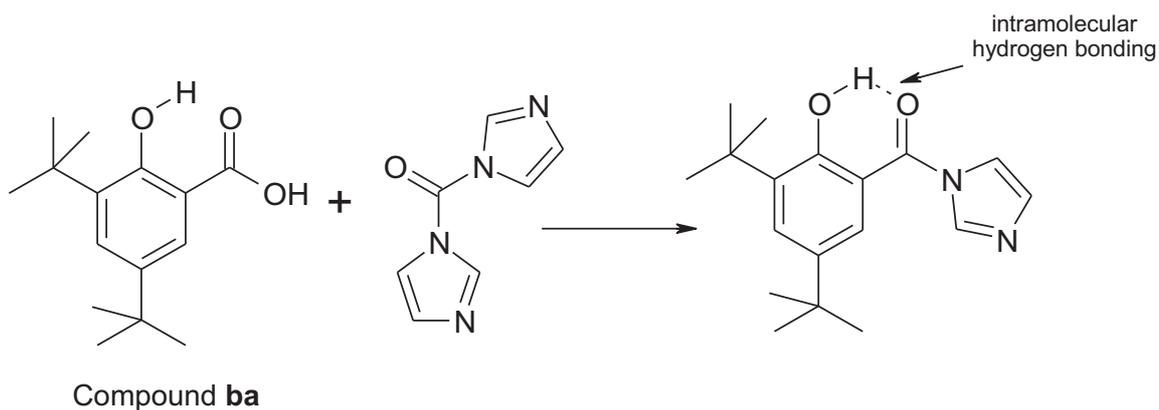


Figure 5.14: Synthesis of the imidazole complex and subsequent intra-molecular hydrogen bonding for a 2-hydroxy salicylic acid derivative.

Thus compound **5.81** became unsuccessful on reacting with acid-imidazole complex of the acid **da**, to produce the final compound. The presence of the electron donating di-tertiary butyl groups was also the other reason in stabilizing the intermediate complex (Figure 5.14). Also converting the acid **dd** into an acid chloride was also not possible. Hence to further investigate the reactivity of compound **5.81**, 2-iodo-benzoic acid was used and converted to acid chloride. This compound then became successful in giving the final compound.

All the acids used in the study reacted in a similar fashion with the pyridyl carboxamidrazones except the 2-bromo benzoic acid **ee**. Interestingly this reaction was successful with 4-bromo benzoic acid, **eh**. This is probably because of the large size of the bromine atom at the 2<sup>nd</sup> position hindering the attack of the carboxamidrazone (Figure 5.16).

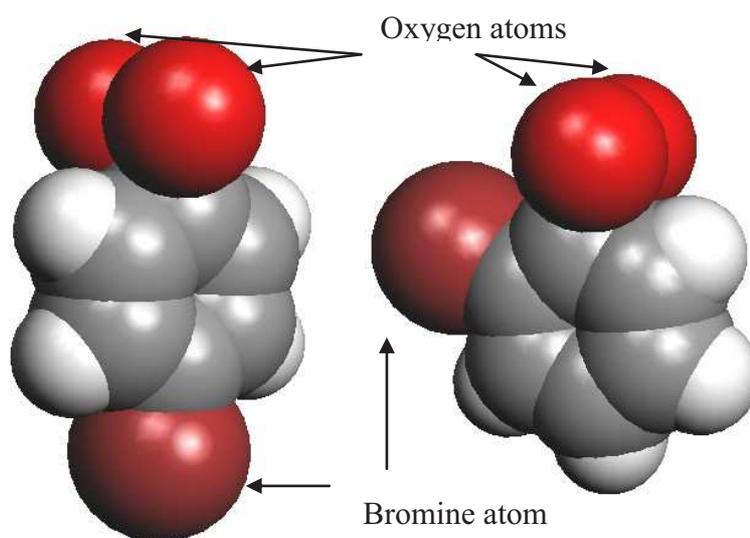


Figure 5.15: Illustrating the large bromine atom causing hindrance to the carboxylic group toward the attack of the amine

The variation within the reactivity of the acids was mainly observed within the pyridyl amines. In terms of the reactivity of the amines used in the study, it was observed that the 4-amino pyridine was the most weakly acting amine. Compound **5.81** failed to react with the acids catalysed by the CDI and even on converting the acids into more reactive acid chlorides, compound **5.81** did not respond to react with the acid chloride. However, in presence of the dimethyl amino pyridine (DMAP) and N,N-diisopropylethylamine (Hunig's base) (Konstantinova *et al.*, 2001) the reaction became successful in producing the corresponding amide as shown in Figure 5.17.

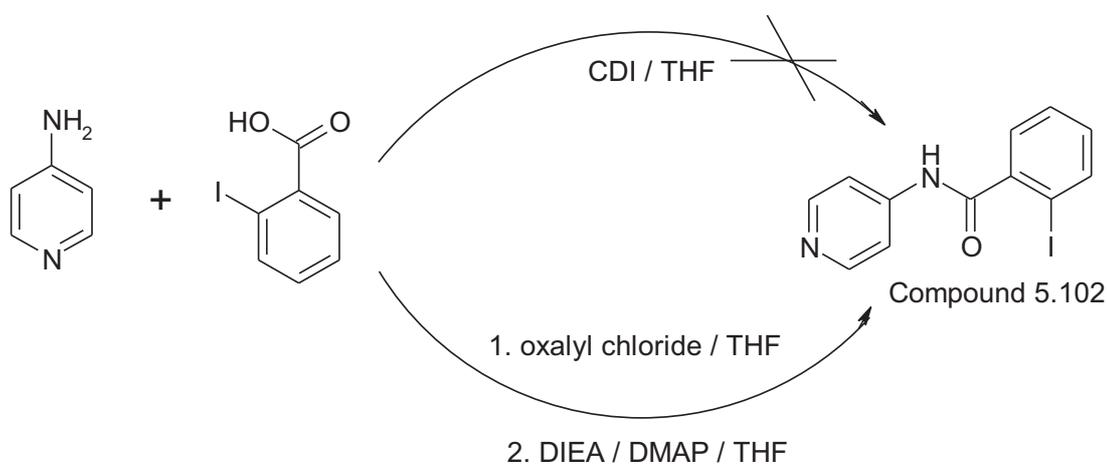


Figure 5.16: The reaction pathway for amide synthesis involving acid chloride.

The 3-Amino pyridine was relatively more reactive than 4-amino pyridine and successfully reacted with most aldehydes on boiling with ethanol at reflux. Thus 4-amino pyridine became unsuccessful in the synthesis of the 4 pyridyl analogues of the compound **5.94**, **5.95**, **5.96**, **5.97**, **5.98**, **5.99** and **5.100** when boiled in ethanol at reflux with the corresponding aldehydes. The final products were observed with two spots on the TLC plate corresponding to the starting material. This procedure was however successful in preparing the 3-pyridyl analogues. The hydrazino pyridines were more reactive than the amino pyridines used in the study. The methyl-hydrazino pyridines are less reactive than the hydrazino pyridines but they are relatively more reactive than the amino pyridines. It was observed that the hydrazino pyridines reacted with all the aldehydes listed in the Figure 5.4 to give the corresponding imides and in this study were the most reactive pyridyl amines.

The aldehyde **be** was the most weakly acting aldehyde used in the investigation. It is the most electron rich aldehyde in the set of aldehydes shown in Figure 5.4. This aldehyde only lacks the hydroxyl group at the 2<sup>nd</sup> position of the aldehyde **aa**. Interestingly **aa** was found

successful with all the pyridyl carboxamidrazones used in the current investigation and in the earlier research carried out by Rathbone *et al.*, whereas **be** became unsuccessful with most amines except the 2-hydrazino pyridine. This could be because of the intra molecular hydrogen bonding with the OH to the carbonyl oxygen in **da**, assists the attack of the amine.

The NMR spectra obtained of most of the compounds were fairly simple and were obtained in DMSO. Some compounds underwent tautomerism and this was observed on comparing the NMR spectrum obtained at room temperature and at high temperature (50°C). One compound **5.45** in particular has shown a significant change in NMR spectrum when compared at different temperatures.

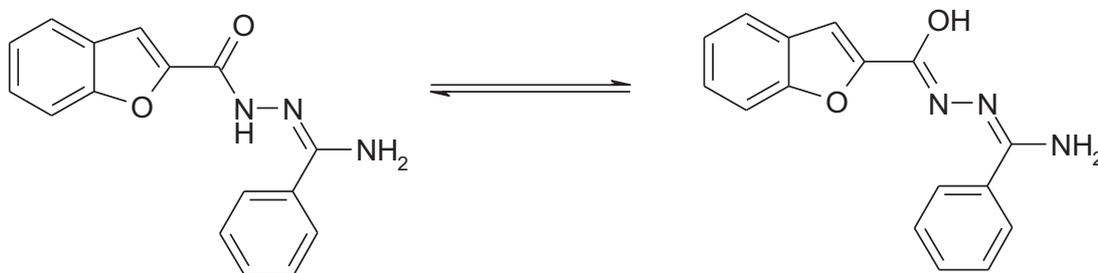


Figure 5.17: Showing the keto-enol form of the compound **5.45**.

At room temperature compound **5.45** exists in both forms as shown in Figure 5.18 and therefore there is appearance of the OH and the NH as shown in the Figure 5.19. When the same spectrum was obtained at 50°C there was the disappearance of the NH peak showing the compound in one form.

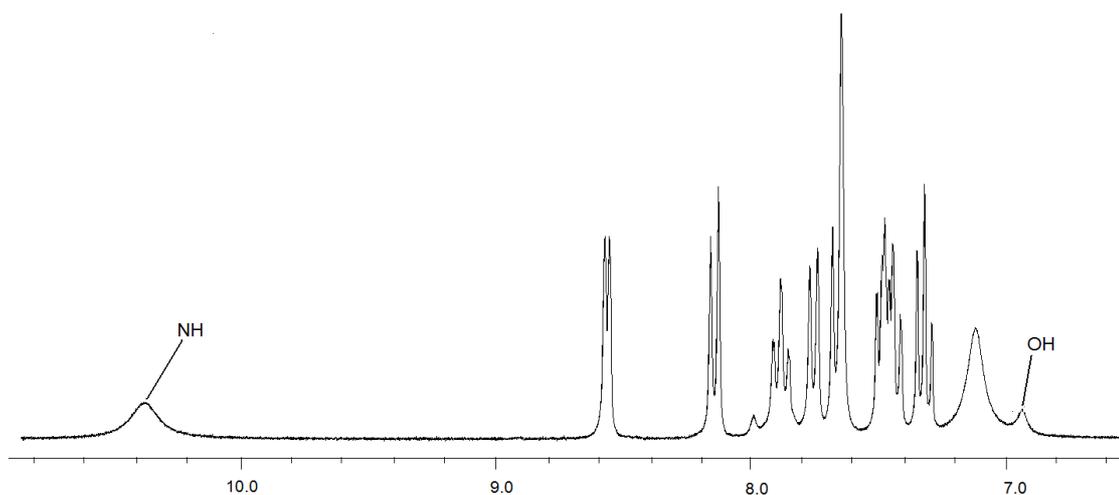


Figure 5.18: The NMR spectrum of compound **5.45** at room temperature

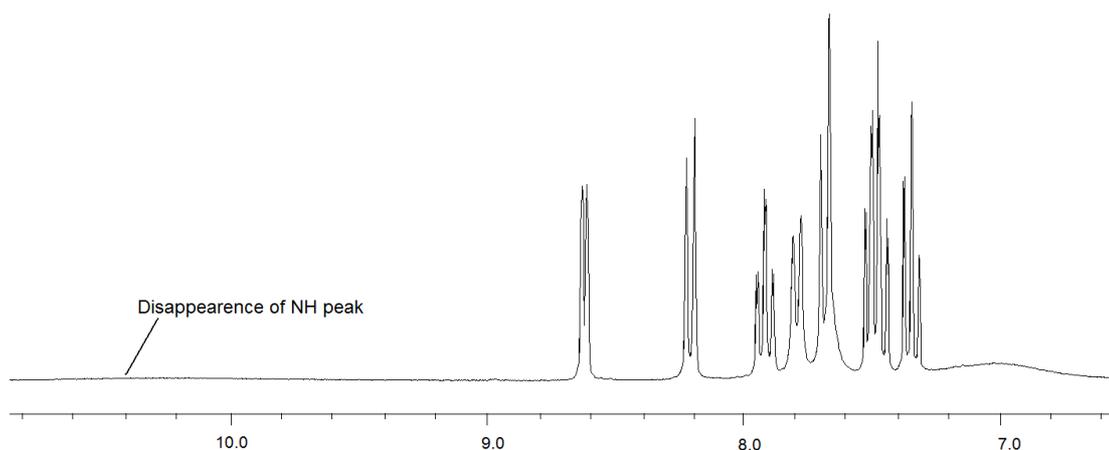
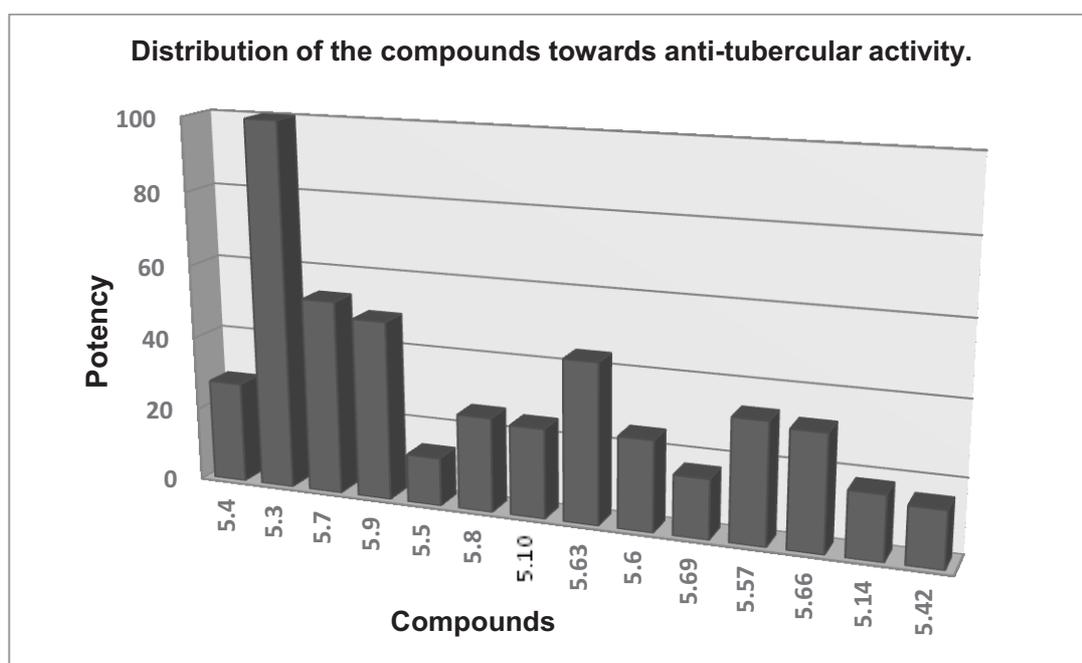


Figure 5.19: The NMR spectrum of compound **5.45** at 50°C

#### 5.11.1 Microbiological studies:

An initial set of the compounds was tested against various organisms as shown in Table 5.7. It was found that most of the compounds that successfully inhibit the growth of *S.aureus* have also successfully inhibited the growth of the *M. fortuitum*. However these compounds did not necessarily inhibit the growth of *E. coli*. This suggests that the compounds had specific activity toward certain organisms. Interestingly the compounds which successfully inhibit the growth of *M. fortuitum* in agar medium did not necessarily inhibit the growth of this organism in solution. Therefore the differences in the permeability pattern of the compounds in agar medium were also an important parameter while considering the zone of inhibition values. Only compounds **5.31** and **5.32** exhibited a very good zone of inhibition when tested against the organisms; unfortunately they were unable to inhibit the growth of *M. fortuitum* when tested in solution. On the other hand the zone of inhibition was little or none in case with the compounds **5.8**, **5.72**, **5.7**, **5.9**, **5.4** and **5.22** but these compounds have successfully inhibited the growth of *M. fortuitum* when tested in solution. Therefore the large zones of inhibition of the compound **5.31** and **5.32** were perhaps due to the high permeability of these compounds in agar plate. Unlike other organisms *M. fortuitum* behaves badly on agar, however is much easier to assay in broth (Tims. K, *et al.*, 2005). Compound **5.1** was very active against *Mycobacterium tuberculosis*; the di-iodo substituted 4-pyridylcarboxamidrazone has also shown activity against the MRSA strains of microorganism with MIC values 10-20 µg/mL (Rathbone, *et al.* 2006). Interestingly compound **5.1** had the di-tert butyl groups which are electron donating, whereas the iodo groups are weakly electron withdrawing leaving a wide area of research into drug designing. Therefore compound **5.1** was further modified by incorporating a wide range of halogens and at the same time the pyridyl-amine moiety was also modified in the study (Figure 5.7). Compound **5.35** and

compound **5.36** have shown very good inhibitory effect against *S.aureus* and *E.coli* in agar plate, unfortunately these compounds did not show any inhibition when tested against *M. tuberculosis*. Compounds **5.3**, **5.4**, **5.5**, **5.7**, **5.8**, **5.9**, **5.10**, **5.14**, **5.42**, **5.57**, **5.60**, **5.63**, **5.66** and **5.69** were found to be very active against *M. tuberculosis* (Graph 5.2). Compound **5.3** in particular has shown the highest potency (in 0.69 $\mu$ g/mL) against *M. tuberculosis*. None of the other compounds have shown similar activity. Compounds **5.4**, **5.7**, **5.8** and **5.9** were also found potent against TB organism (killing 90% of the cells) at a concentration between 0.5-3  $\mu$ g/mL. However all these compounds have only exhibited 50% or less potency when compared with compound **5.3**. All these compounds have also inhibited the growth of gram positive organisms *M. fortuitum* and *S. aureus* (Table 5.7) but were inactive against the gram negative organism *E. coli* except compound **5.7**, which has shown some activity against *E. coli*. Although compounds **5.5** and **5.42** were found active against *M.fortuitum* in the agar medium, they failed to inhibit the growth of the organism when tested in solution.



Graph 5.2: The distribution of the potency of the compounds towards *M. tuberculosis* in a concentration at which compound **5.3** exhibits 100% potency.

Compound **5.57**, **5.60**, **5.63** and **5.66** (analogues of compound 5.2) were also found active with IC<sub>50</sub> value less than 3  $\mu$ g/mL. Interestingly none of these compounds have shown activity against any other organisms on the agar plate, except compounds **5.66** and **5.14** which have shown some activity against *S.aureus* (Table 5.8). Unfortunately all these compounds (Table 5.10) were observed with high toxicity profile representing a selectivity index  $\leq 1$  (Table 5.10). Only compound **5.3** and **5.4** were observed with SI more than 2. But

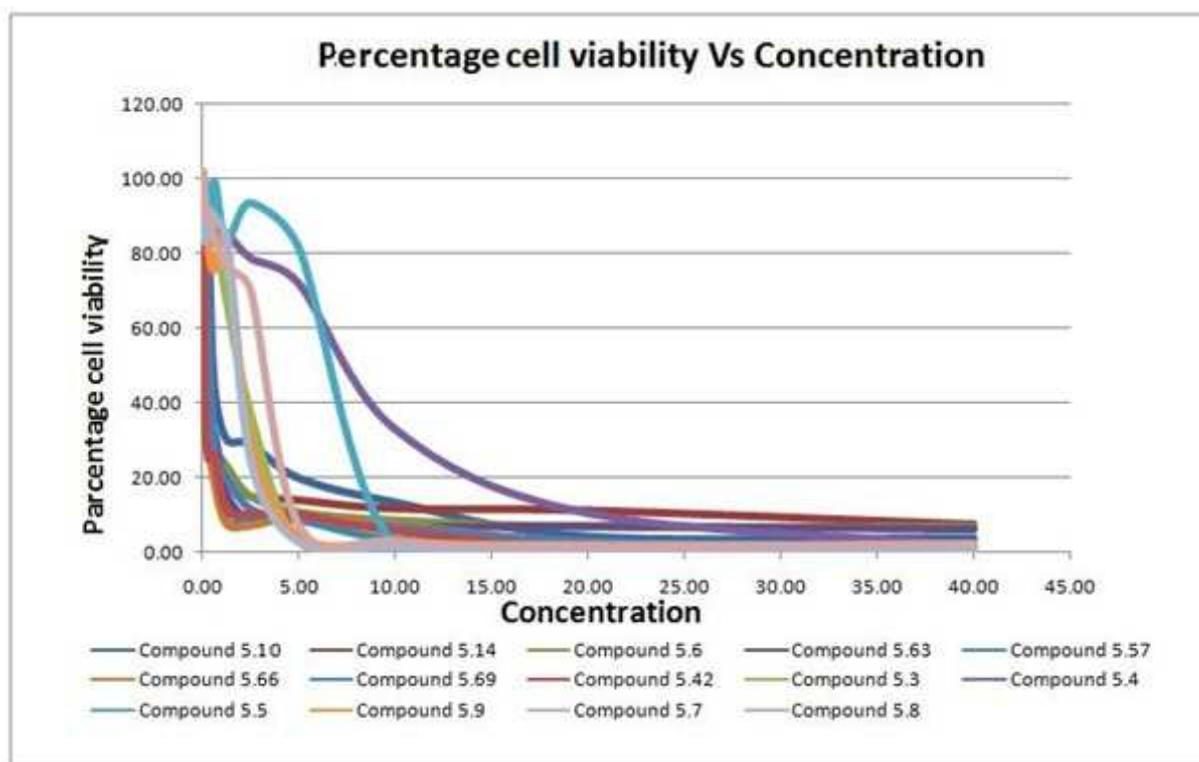
compound **5.3** was reported as the most potent compound isolated in this work showing activity at less than 1 µg/mL concentration.

#### 5.11.2 Toxicology:

All of the 14 compounds (Table 5.9) which have shown activity against the TB organism were also studied for their toxicity on mammalian cells (Graph 3). The concentrations at which these compounds kill the parent cells were obtained (Table 5.11). Most of these compounds have shown high level of toxicity causing almost 90% cell death at a concentration less than 20 µg/mL. Particularly compound **5.14**, **5.63**, **5.57**, **5.42** and **5.66** were found extremely toxic killing 50% of the parent cells at a concentration as low as 0.16 µg/mL.

Compounds	Viability of the mammalian cells in concentration (µg/mL)									
	0.08	0.16	0.31	0.63	1.25	2.50	5.00	10.00	20.00	40.00
5.3	95.42	94.07	90.58	87.64	67.91	37.53	3.15	1.71	1.62	1.58
5.4	90.29	84.83	86.77	85.51	85.70	78.91	72.15	32.92	10.29	1.58
5.5	91.52	98.86	85.09	99.10	83.37	93.42	81.49	2.05	2.08	1.58
5.7	96.54	83.47	86.66	89.78	84.07	24.72	2.69	1.25	0.69	0.68
5.8	101.90	89.74	92.04	84.49	76.40	70.41	7.36	2.96	1.62	2.26
5.9	95.31	91.79	81.39	75.62	84.53	30.50	4.20	1.94	1.74	1.36
5.10	95.09	86.89	82.96	44.38	30.23	29.02	19.73	13.21	4.05	3.62
5.14	91.18	50.40	31.95	28.54	19.77	14.85	13.89	11.62	11.10	7.36
5.42	89.51	35.80	25.00	25.62	11.16	10.20	10.04	5.69	1.27	0.91
5.57	88.39	31.13	32.96	24.49	8.26	9.64	8.76	3.53	1.74	1.92
5.60	87.50	76.05	34.30	28.09	23.37	15.19	10.97	8.77	7.06	7.02
5.63	90.18	50.51	29.26	32.02	15.81	8.96	8.76	6.83	7.06	6.23
5.66	82.92	39.68	29.48	18.88	7.79	7.26	9.46	7.52	2.89	1.47
5.69	89.84	92.70	66.03	28.09	20.12	11.79	8.64	6.83	3.01	1.92

Table 5.11: Showing the percentage viability of the mammalian cells at various concentrations.



Graph 5.3: The graph above showing the distribution of the percentage of cell viability of the mammalian cells versus the concentration of the compound.

The toxicological study of these compounds shows that the 4-hydrazino pyridine analogues of compound A were less toxic compared to any other compounds. Compounds **5.4** and **5.5** which were highly active against the gram positive organism (TB), had least toxicity profile killing less than 30% of the mammalian cells at a concentration as high as 5  $\mu\text{g/mL}$ . This concentration was found to be extremely toxic in all other compounds killing as high as 97% of the mammalian cells. The compound **5.5** however has IC<sub>90</sub> value of 6.17  $\mu\text{g/mL}$ , and compound **5.4** has 2.9  $\mu\text{g/mL}$ . This suggests that the compound **5.5** is relatively toxic to mammalian cells in the concentration at which it produces anti-bacterial activity. Compound **5.4** on the other hand was relatively toxic against the organism causing 90% death at a concentration which only kills about 25% of the mammalian cells.

#### 5.11.3 Structural activity relationship studies:

Out of the 102 compounds prepared only 14 compounds have shown activity and another 42 have shown weak activity against the pathogenic *Mycobacterium tuberculosis* (Table 5.9). Of

these 14 compounds, 8 compounds were the analogues of Compound A, and 5 of the rest were the analogues of compound D.

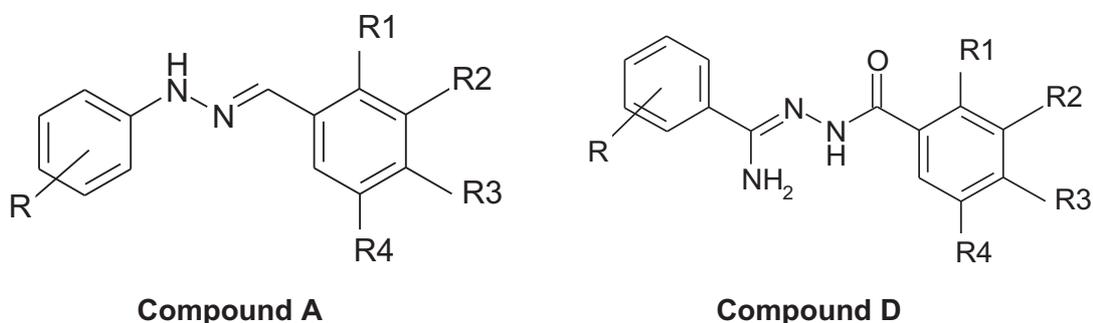


Figure 5.20: Showing the analogue of Compound A and Compound D

Of these 8 compounds which have shown potential activities against *M.tuberculosis*, 6 compounds were the 4-hydrazino pyridine derivative of Compound A. Compounds **5.3**, **5.4**, **5.9** and **5.10** had the tertiary-butyl substitution on the R2 and R3 position and were particularly active against gram positive organisms and inactive against gram negative organisms. Removing a tertiary butyl group either from R2 or R3 position, makes the compounds active against gram negative organisms as well. These suggest that the di-tert-butyl substitution is essential to make the compound specifically active against gram positive organisms. 7 out of the 8 compounds had the OH on the R1, except Compound **5.4**. Replacing the OH from the R1 to R3 (Scheme 5.1) as in Compound **5.4** the activity goes down by three fold. Compound **5.3** was the most potent compound reported in this investigation; changing the 4-pyridyl with the 3-pyridyl moiety as in compound **5.10** the activity drops by 5 fold. Although the tertiary butyl group is essential on both R2 and R4, it was observed that substitution on the R2 was more important than the R4 for its anti-tubercular activity. Thus compound **5.7** was observed as the next most potent compound in this study. Interestingly the 3-pyridyl analogue of compound **5.7** has also shown activity within the organism. This suggests that the presence of the OH on the R1 and tert-butyl on the R2 is essential within the active site. The presence of 4-pyridyl is more favourable than the 3-pyridyl within the active site. Iodine has shown some activity in the earlier studies and hence in this study a wide range of aldehydes containing halogen substitutions at R2 and R4 was used, but none of the halogen containing derivatives has shown any biological activity except iodine, compound **5.5**.

Compound D has also shown some interesting activity in the organism; these compounds were the analogues of the Compound **5.2** (Begum. *et al.*, 2007). Of these 6 compounds which have shown potential activity within the organism, 5 were the analogues of 2-pyridyl carboxamidrazone. This suggests that the presence of 2-pyridyl carboxamidrazone moiety was essential for the activity. The presence of the halogen on the 2<sup>nd</sup> position of the aryl ring of compound **5.89** has shown some interesting results in the active site (Mamalo, *et al.*, 1992). All these 6 compounds had electro-negative halogen substitution suggesting the presence of the electronegative substitution is favourable within the active site. Unfortunately all these compounds were found to be highly toxic with the selectivity index less than 1.

Compounds A, B, C, D, E and F were designed as analogues of compounds **5.1** and Compound **5.2**. This was essentially with the idea to change the various substitutions throughout the moiety to reduce the toxicity and increase the potency of Compound **5.1**, compound **5.2**. All the analogues of Compound D were reported cytotoxic (Graph 5.2 and Graph 5.3). Interestingly all these compounds had the common pyridyl carboxamidrazone moiety. In the current research the pyridyl carboxamidrazone moiety was observed as the toxicity reason of the compound and was also observed with compound **5.2** in the earlier research. This is perhaps that the pyridyl carboxamidrazone derivatives were metabolised into pyridyl derivatives in the bacteria (Tims. *et al.*, 2005). This aromatic ring of pyridine may as well act as an oxidising agent for the nicotinamide adenosine dinucleotide phosphate in the amino-oxidative metabolism in the bacteria (Ignatenko. *et al.*, 1978). The di-tert-butyl aryl moiety in Compound **5.1** was probably important to produce the biological activity. Although compounds **5.3**, **5.7** and **5.9** were found active, they had a small selectivity index just over 1. This was found among most of the ant-TB drugs used as second line treatment. While considering these two compounds as a possible anti-tubercular agent, the risk-benefit analysis must be considered. Thus any of these compounds could be used as an anti-tubercular agent if the beneficial effect of these compounds exceeds the possible risk associated with them. Another interesting feature about these agents was that they were specifically active against all the gram positive bacteria, but not gram negative. Thus these agents may also be considered as a useful compound in screening gram positive bacteria.

### 5.12 Conclusion:

A library of 102 compounds have been generated following the lead compounds N<sup>1</sup>-[3, 5-di-(*tert*-butyl)-2-hydroxybenzylidene] - pyridine-4-carboxamidrazone and N<sup>1</sup>-[3, 5-di-(*tert*-butyl)-2-hydroxybenzoyl] amide. All these compounds have been successfully characterised and tested against various gram positive and gram negative organisms including *Mycobacterium tuberculosis*. Of only 14 compounds which have shown potential activity against the pathogen *Mycobacterium tuberculosis*, 6 were found less toxic and rest all were found to be extremely toxic killing 80% of the mammalian cells at a concentration lower than 5 µg/mL. Compound **5.3** was identified as the most potent anti-tubercular agent killing 90% of the organisms at concentration 0.6µg/mL.

## CHAPTER 6

### MATERIALS AND METHODS

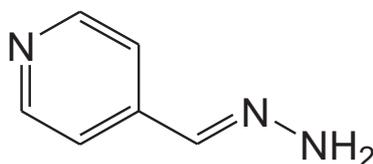
#### 6.1. Chemicals:

All chemicals and solvents were used as supplied. Hydrazine, 2-cyano pyridine, furoyl acrylic acid, N,N'-carbonyl di-imidazole, DMSO, 4-amino pyridine, and all the aldehydes were purchased from Sigma-Aldrich. Organic solvents were purchased from Fisher Chemicals. Thin layer chromatography plates on aluminium sheets (20 X 20) of silica gel F<sub>254</sub> were obtained from Merck.

#### 6.2. Instrumentation:

Proton NMR spectra were obtained on a Bruker AC 250 instrument operating at 250 MHz as solution in d<sub>6</sub>-DMSO and referenced from  $\delta$ DMSO = 2.50 ppm unless otherwise stated. Infrared spectra were recorded as KBr disc on a Mattson 3000 FTIR spectrophotometer. (APCI-MS) was carried out on a Hewlett-Packard 5989B quadrupole instrument connected to an electrospray 59987A unit with an APCI accessory and automatic injecting using a Hewlett-Packard 1100 series autosampler. Accurate MS was carried out on TOF mass spectra (ESI mode) measured on a Waters LCT Premier Mass Spectrometer.

#### 6.3.1. Preparation of Pyridine-4-methylhydrazone: Compound 5.87



An appropriate pyridine-4-carboxaldehyde (11.0287gm or 9.7mL or 90.65mmol) was dissolved in ethanol (7mL) and added dropwise through a pressure generated dropping funnel into a mixture of hydrazine monohydrate ( 9.5mL or 190mmol) in ethanol (5mL) stirred at 60°C for 24 hours. The product was collected by evaporating the excess hydrazine under high vacuum pump to give a yellow oil compound.

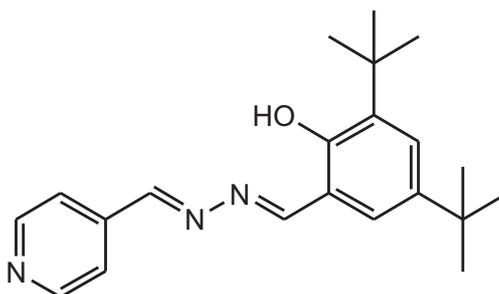
Yield was 10.7234 g or 87.178 mmol or 97%.

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 7.39 (dd, 2H,  $J = 1.2, 1.9$  Hz, Py-3,5H); 7.45 (s, 2H,  $\text{NH}_2$ ); 7.62 (s, 1H, NH); 8.47 (dd, 2H,  $J = 1.9, 1.9$  Hz, Py-2,6H) ppm.

MS (APCI +)  $m/z = \text{found } 122 (\text{M} + \text{H})^+$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.25 (single spot).

### 6.3.2. Preparation of 4-[3, 5-di(*tert*-butyl)-2-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1, 3-butadiene: Compound 5.17



3, 5-di(*tert*-butyl)-2-hydroxybenzaldehyde (3.01gm or 12.844mmol) was dissolved in ethanol (5mL). Pyridine-4-carboamidrazone (1.1gm or 9.09 mmol) was added and the reaction was stirred at 90 °C for 36 hours. The solvent was removed by filtration under vacuum and the title compound was collected as yellow amorphous solid, which was dried under vacuum.

Yield was 1.7229g. or 5.082 mmol or 60%

Melting point: 161-179.5 °C

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 1.29 (s, 9H, *tert*-butyl); 1.409 (s, 9H, *tert*-butyl); 7.43 (d, 1H,  $J = 2.5$  Hz, Ph-4H); 7.52 (d, 1H,  $J = 2.5$  Hz, Ph-6H); 7.8 (dd, 2H,  $J = 1.2, 1.2$  Hz, Py-3,5H); 8.76 (dd, 2H,  $J = 1.9, 1.9$  Hz, Py-2,6H); 8.91 (s, 1H, CH); 9.06 (m, 1H, CH); 12.2 (s, 1H, OH); ppm.

$^{13}\text{C}$  NMR (63 MHz  $d_6$ -DMSO): 29.21 ( $\text{CMe}_3$ ); 31.2 ( $\text{CMe}_3$ ); 33.93 ( $\text{CMe}_3$ ); 34.65 ( $\text{CMe}_3$ ); 116.65 (C1-Ph); 121.95 (C3 and C5 Py), 127.93 (C6 Ph); 135.77 (C4- Ph); 140.41 (C3- Ph), 140.9 (C5-Ph); 150.54 (C2 and C6 Py); 156.33 (C imide); 160.93 (C- imide); 167.98 (C2- Ph); ppm.

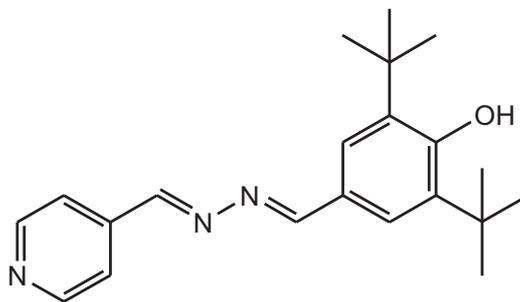
IR (KBr)  $\nu = 3073, 2948, 2869, 1955, 1850, 1796, 1755, 1677, 1420, 1390$  (-C=N-), 1433, 1385 (*tert*-butyl), 1175 (-OH)  $\text{cm}^{-1}$

MS (APCI +ve)  $m/z = \text{found } 338 (\text{M} + \text{H})^+$ .

Accurate MS (+ES) = Calculated 338.2232(m+H) $^+$  and 339.2263(m+H) $^+$  found 338.2235(m+H) $^+$  and 339.2275 (m+H) $^+$ , 0.9 ppm.

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.72 (single spot).

### 6.3.3. Preparation of 4-[3, 5-di-(*tert*-butyl)-4-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1, 3-butadiene: Compound 5.18



Yield = 0.38g or 1.127 mmol or 65%

Melting point: 143.3 – 146.9 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.41 (s, 18H, di-*tert*-butyl); 7.69 (s, 2H, Ph-H); 7.77 (dd, 2H, J= 1.422, 4.423 Hz, Py-3,5H); 8.665 (s, 1H, CH); 8.71 (dd, 2H, J= 1.422, 4.458 Hz, Py-2,6H); 8.726 (s, 1H, CH); ppm.

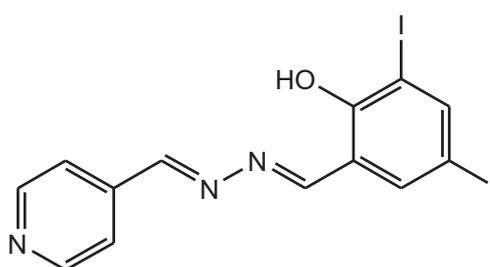
IR (KBr)  $\nu$ = 3625, 3347, 3081, 2960, 2601, 1940, 1627, 1602, 1420, 1382 (-C=N-), 1433, 1385 (*tert*-butyl), 1125 (-OH) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 230, 234, 336 (M - H)<sup>+</sup>.

Accurate MS (+ES) = Calculated 338.2232(m+H)<sup>+</sup>, found 338.2233(m+H)<sup>+</sup>, 0.3 ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.78 (single spot).

### 6.3.4. Preparation of 4-[3, 5-di-iodo-2-OH-benzyl]-1-(4-pyridyl)-2,3-diaza-1,3-butadiene: Compound 5.19



Yield = 0.25g or 0.523mmol or 78%

Melting point: 225.3 – 228.9 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 7.77 (dd, 2H, J= 1.896, 4.423 Hz, Py-3,5H); 8.05 (d, 1H, J= 1.896 Hz, Ph-H); 8.17 (d, 1H, J= 2.527 Hz, Ph-H); 8.75 (dd, 2H, J= 1.264, 4.423 Hz, Py-2,6H); 8.926 (s, 1H, CH); 8.93 (s, 1H, CH); 12.5 (bs, 1H, OH); ppm.

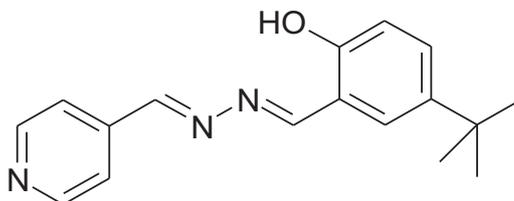
IR (KBr)  $\nu$ = 3070, 3011, 1918, 1870, 1729, 1663, 1621, 1450 (-C=N-), 1273, 1182 (-OH) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 372, 476 (M - H)<sup>+</sup>, 477 (M + H)<sup>+</sup>.

Accurate MS (+ES) = Calculated 477.8913(m+H)<sup>+</sup>, found 477.8921(m+H)<sup>+</sup>, 1.7 ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.62 (single spot).

### 6.3.5. Preparation of 4-[5-*tert*-butyl-2-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1, 3-butadiene: Compound 5.22



Yield = 0.26g or 0.918 mmol or 65%

Melting point: 204.3 – 208.9 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.275 (s, 9H, *tert*-butyl); 6.92 (d, 1H, J=8.214, Ph-H); 7.46 (t, 1H, J=8.214, 2.527 Hz, Ph-H); 7.74 (d, 1H, J=2.57 Hz, Ph-H); 7.78 (dd, 2H, J= 1.896, 4.423 Hz, Py-3,5H); 8.73 (dd, 2H, J= 1.264, 4.423 Hz, Py-2,6H); 8.83 (s, 1H, CH); 9.005 (s, 1H, CH); 10.92 (bs, 1H, OH); ppm.

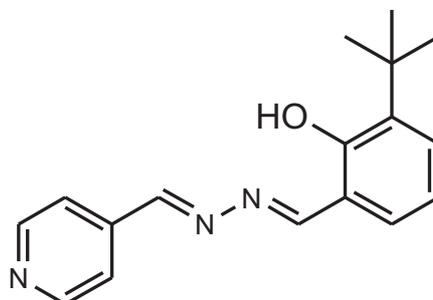
MS (APCI -ve) m/z = found 280 (M - H)<sup>+</sup>, 281 (M)<sup>+</sup>, 313 (M + MeO H)<sup>+</sup>.

IR (KBr) ν = 3041, 2959, 2869, 1937, 1899, 1760, 1621, 1567(-C=N-), 1492, 1407-1370 (*tert*-butyl), 1150 (-OH) cm<sup>-1</sup>

Accurate MS (+ES) = Calculated 282.1528 (m+H)<sup>+</sup> and 283.1637 (m+H)<sup>+</sup> found 282.1613 (m+H)<sup>+</sup>, 283.1645 (m+H)<sup>+</sup>, 2.5 ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.7.

### 6.3.6. Preparation of 4-[3-*tert*-butyl-2-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1, 3-butadiene Compound 5.21



Yellow amorphous.

Yield = 0.45g or 1.590mmol or 43%

Melting point: 232.2 – 235.6 °C

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 1.409 (s, 9H, *tert*-butyl); 6.95 (t, 1H,  $J=7.582$ , 8.214 Hz, Ph-H); 7.42 (dd, 1H,  $J=8.214$ , 1.264 Hz, Ph-H); 7.49 (dd, 1H,  $J=7.582$ , 1.264 Hz, Ph-H); 7.8 (dd, 2H,  $J=1.264$ , 4.423 Hz, Py-3,5H); 8.75 (dd, 2H,  $J=1.896$ , 4.423 Hz, Py-2,6H); 8.92 (s, 1H, CH); 9.005 (s, 1H, CH); 9.04 (bs, 1H, OH); 12.370 (s, 1H, CH); ppm.

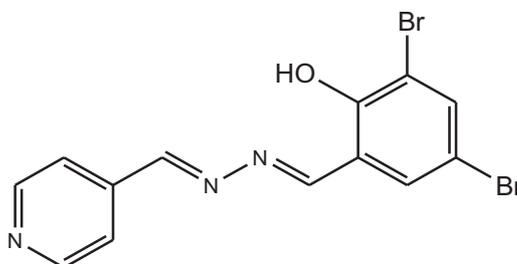
IR (KBr)  $\nu=3041$ , 2959, 2869, 1937, 1899, 1760, 1692, 1621, 1567 (-C=N-), 1490, 1412-1350 (*tert*-butyl), 1175 (-OH)  $\text{cm}^{-1}$

MS (APCI -ve)  $m/z$  = found 280 (M - H) $^+$ , 281 (M) $^+$ , 351.

Accurate MS (+ES) = Calculated 282.1606 (m+H) $^+$  found 282.1609 (m+H) $^+$ , 1.1 ppm.

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.65 (single spot).

### 6.3.7. Preparation of 4-[3, 5-di-bromo-2-OH-benzyl]-1-(4-pyridyl)-2, 3-diaza-1, 3-butadiene: Compound 5.20



Yellow amorphous.

Yield = 0.57g or 2.072 mmol or 62%

Melting point: 179.1 – 183.6  $^{\circ}\text{C}$

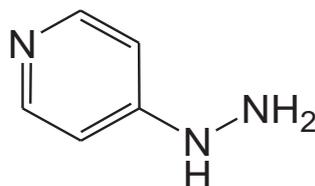
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 7.8 (dd, 2H,  $J=1.896$ , 4.423 Hz, Py-3,5H); 7.95 (d, 1H,  $J=2.527$ , Ph-H); 7.98 (d, 1H,  $J=1.896$  Hz, Ph-H); 8.76 (dd, 2H,  $J=1.896$ , 4.423 Hz, Py-2,6H); 8.946 (s, 1H, CH); 9.015 (s, 1H, CH); 12.3 (bs, 1H, OH); ppm.

IR (KBr)  $\nu=3424$ , 3046, 2987, 2869, 2358, 1939, 1627(-C=N-), 1480, 1432, 1175 (-OH) 697 (-C-Br)  $\text{cm}^{-1}$

MS (APCI -ve)  $m/z$  = found 276, 324, 382 (M - H) $^+$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.5 (single spot).

### 6.4. 1. Preparation of 4-hydrazino pyridine: Compound 5.84.



4-chloropyridine (4.93g or 43.628 mmol) was added to a mixture of hydrazine (33ml or 20eq) and n-butanol (18ml). The reaction mixture was refluxed at 100°C for 24 hours with constant stirring. The excess hydrazine was removed under high vacuum evaporation and the white solid residue was dissolved in 1 M NaOH. The final product was then obtained by extracting with ethyl acetate from the NaOH solution and then evaporating the ethyl acetate under vacuum. The final product is a white solid which on exposure to air changes to violet colour.

Yield = 3.5g or 31.531mmol or 72.27%

Melting point: 241.3 – 243.7°C

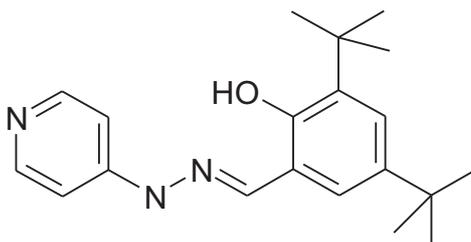
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 7.39 (dd, 2H, J= 1.2, 1.9 Hz, Py-3,5H); 7.45 (s, 2H, NH<sub>2</sub>); 7.62 (s, 1H, NH); 8.47 (dd, 2H, J= 1.9, 1.9 Hz, Py-2,6H) ppm.

MS (APCI -ve) m/z = found 120 (M - H)<sup>+</sup>.

Accurate MS (ES +ve) m/z = found 110 (M + H)<sup>+</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.57.

#### 6.4.2. Preparation of 3,5-di-*tert* butyl 2-OH benzylidene 4-hydrazinopyridine Compound 5.03



4-hydrazinopyridine (0.3g) and 3,5 di-*tert* butyl 2-OH benzaldehyde (0.69g) was dissolved in 7ml ethanol and was refluxed for 18 hours with constant stirring. The final product was then filtered washed and then recrystallised from ethanol to give a yellow amorphous compound.

Yield 0.32g. 0.984 mmol or 36%

Melting point: 191.1 – 193.9 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.29 (s, 9H, *tert*-butyl); 1.409 (s, 9H, *tert*-butyl); 4.49 (s, 1H, NH); 6.96 (dd, 2H, J=5.05, 1.896 Hz, Py-H); 7.25 (d, 2H, J=4.423 Hz, Ph-H); 8.22 (s, 1H, CH); 8.3 (dd, 1H, J=5.05, 1.896 Hz, Py-H); 11.01 (s, 1H, OH); 11.3 (s, 1H, OH); ppm.

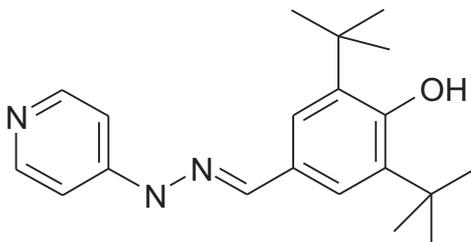
IR (KBr)  $\nu$ = 3619 (-NH), 3418, 3156, 2954, 2873, 2491, 1914, 1596 (-C=N-), 1520, 1350 (*tert*-butyl), 1300, 1140 (-OH) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 232, 233, 296, 324 (M - H)<sup>+</sup>.

Accurate MS (ES -ve) m/z = Calculated 325.2154, found 324.2108 (M - H)<sup>+</sup>. PPM=9.9

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.55 tailing.

#### 6.4.3. Preparation of 3,5-di-*tert* butyl 4-OH benzylidene 4-hydrazinopyridine. Compound 5.04



Pale yellow amorphous.

Yield = 0.73g 2.274 mmol or 69%

Melting point: 208.3 – 210.7 °C

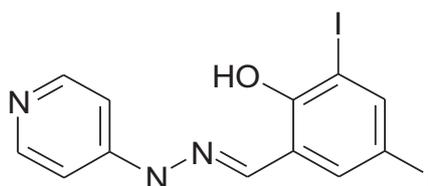
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.476 (s, 18H, di-*tert*-butyl); 4.39 (s, 1H, NH); 6.8 (dd, 2H, J=5.05, 1.896 Hz, Py-H); 7.5 (s, 2H, Ph-H); 7.93 (s, 1H, CH); 8.25 (dd, 2H, J=5.05, 1.896 Hz, Py-H); 10.652 (s, 1H, OH); ppm.

IR (KBr)  $\nu$ = 3627 (-NH), 3313, 2952, 2875, 2481, 1920, 1600 (-C=N-), 1530, 1360 (*tert*-butyl), 1300, 1120 (-OH) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 230, 323 (M - 2H)<sup>+</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.63 (single spot).

#### 6.4.4. Preparation of 3,5-di-iodo 2-OH benzylidene 4-hydrazinopyridine Compound 5.05



Yellow amorphous.

Yield = 0.782g or 1.558mmol or 82%

Melting point: 222.1 -215.3 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 6.86 (dd, 2H, J=5.05, 1.896 Hz, Py-H); 7.87 (dd, 1H, J=1.896, 2.527, 4.423 Hz, Ph-H); 7.97 (dd, 1H, J=2.527, 1.896 Hz, Ph-H); 8.1 (s, 1H, CH); 8.26 (dd, 2H, J=1.896, 5.05 Hz, Py-H); ppm.

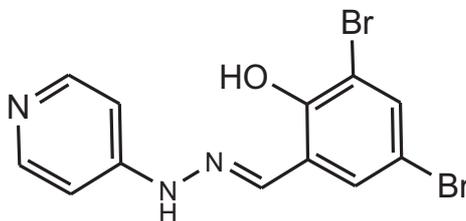
IR (KBr)  $\nu$ = 3600 (-NH), 3245, 3041, 2491, 2366, 2080, 1632 (-C=N-), 1586, 1430, 1187 (-OH) cm<sup>-1</sup>

MS (APCI +ve) m/z = found 466 (M + H)<sup>+</sup>.

Accurate MS (ES+ve) m/z = Calculated 464.8835, found 465.8904(M + H)<sup>+</sup>, 1.9 ppm

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.67 (single spot).

#### 6.4.5. Preparation of 3,5-di-bromo 2-OH benzylidene 4-hydrazinopyridine Compound 5.06.



Bright yellow powder.

Yield = 0.63g or 1.7073 mmol or 62.6%

Melting point: 239.3 – 242.7 °C

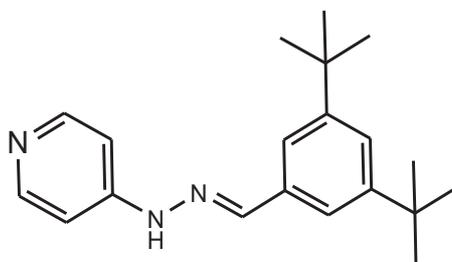
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 6.88 (d, 2H,  $J=6.318$  Hz, Py-H); 7.73 (d, 1H,  $J=2.527$  Hz, Ph-H); 7.8 (d, 1H,  $J=2.527$  Hz, Ph-H); 8.1 (s, 1H, CH); 8.24 (d, 2H,  $J=6.318$  Hz, Py-H); ppm.

IR (KBr)  $\nu$  = 3436 (-NH), 3199, 3085, 2965, 2084, 1976, 1637 (-C=N-), 1520, 1495, 1430, 1235, 1220, 1150 (-OH) 710 (C-Br)  $\text{cm}^{-1}$

MS (APCI -ve)  $m/z$  = found 277, 278, 292, 324, 370 ( $M - H$ ) $^+$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.55 (single spot).

#### 6.4.6. Preparation of 3,5-di-*tert*-butyl benzylidene 4-hydrazinopyridine: Compound 5.09.



White solid.

Yield = 0.3g or 0.964mmol or 77%

Melting point: 167.1 – 169.5 °C

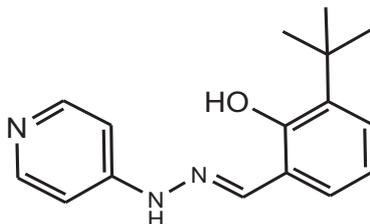
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 1.307 (s, 18H, di-*tert*-butyl); 6.94 (d, 2H,  $J=6.318$  Hz, Py-H); 7.388 (t, 1H,  $J=1.896, 1.246$  Hz, Ph-H); 7.5 (d, 2H,  $J=1.896$  Hz, Ph-H); 7.95 (s, 1H, CH); 8.2 (d, 2H,  $J=6.318$  Hz, Py-H); 10.847 (s, 1H); ppm.

IR (KBr)  $\nu$  = 3438 (-NH), 3210, 3149, 2950, 2867, 2495, 1922, 1596 (-C=N-), 1495, 1375 (*tert*-butyl), 1220, 1110  $\text{cm}^{-1}$

MS (APCI +ve)  $m/z$  = found 213, 241, 310 ( $M + H$ ) $^+$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.82 (tailing).

#### 6.4.7. Preparation of 3-*tert*-butyl 2-OH benzylidene 4-hydrazinopyridine: Compound 5.07



Pale white solid.

Yield = 0.2g 0.743mmol or 31%

Melting point: 176.1 – 178.2 °C

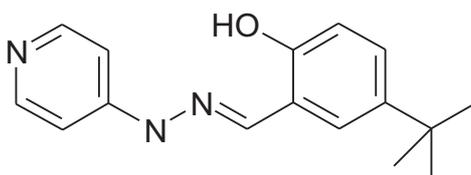
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.407 (s, 9H, *tert*-butyl); 6.8 (d, 2H, J=6.318 Hz, Py-H); 6.88 (t, 1H, J=7.582 Hz, Ph-H); 7.2 (ms, 2H, Ph-H); 8.19 (s, 1H, CH); 8.27 (d, 2H, J=6.318 Hz, Py-H); 11.1 (bs, 1H, OH); 11.46 (s, 1H, OH); ppm.

IR (KBr)  $\nu$ = 3212, 3154, 2998, 2501, 1920, 1859, 1586 (-C=N-), 1500, 1425, 1367 (*tert*-butyl), 1300, 1200, 1135 (-OH) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 176, 268 (M - H)<sup>+</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.68 tailing.

#### 6.4.8. Preparation of 5-*tert*-butyl 2-OH benzylidene 4-hydrazinopyridine Compound 5.08



Pale brown

Yield = 0.27g or 1.003 mmol or 37%

Melting point: 174.2 – 177.1 °C

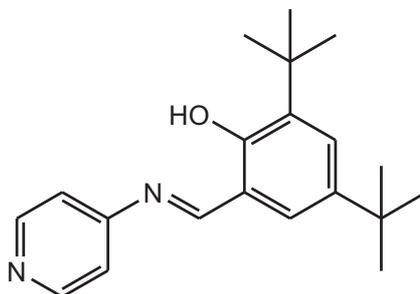
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.28 (s, 9H, *tert*-butyl); 6.83 (d, 1H, J=8.214 Hz, Ph-H); 6.88 (d, 2H, J=6.318 Hz, Py-H); 7.25 (dd, 1H, J=2.527, 8.214 Hz, Ph-H); 7.65 (d, 2H, J=2.527 Hz Ph-H); 8.22 (d, 2H, J=6.318 Hz, Py-H); 8.268 (s, 1H, CH); 10.91 (s, 1H, OH); ppm.

IR (KBr)  $\nu$ = 3434 (-NH), 3230, 3139, 2967, 2869, 1922, 1602 (-C=N-), 1495, 1367 (*tert*-butyl), 1220, 1135 (-OH) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 220,268 (M - H)<sup>+</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.72 (short tailing).

6.5.1. Preparation of 4-[3,5-di-(*tert*-butyl)-2-OH-benzylidene]-amino-pyridine: Compound 5.29.



Red crytraline

Yield = 0.51g or 1.645mmol or 63%

Melting point: 118.2 – 120.6 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.29 (s, 9H, Ph); 1.41 (s, 9H, Ph); 7.42 (dd, 2H, J= 1.264, 4.423 Hz, Py-3H,5H); 7.45 (d, 1H, J= 2.527 Hz, Ph-6H); 7.53 (d, 2H, J= 1.896 Hz, Ph-4H); 8.61 (dd, 2H, J= 1.264, 4.423 Hz, Py-2,6H); 9.05 (s, 1H, CH); 13.466 (s, 1H, OH); ppm.

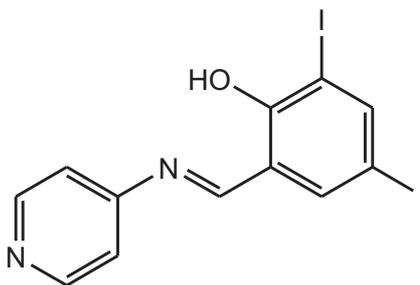
IR (KBr)  $\nu$ = 3438, 3077, 2964, 2908, 1926, 1801, 1619, 1575 (-C=N-), 1440, 1385, 1375 (*tert*-butyl), 1210, 1175 (-OH) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 309 (M - H)<sup>+</sup>.

MS (APCI +ve) m/z = found 311 (M + H)<sup>+</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.65.

6.5.2. Preparation of 4-[3,5-di-iodo-2-OH-benzylidene]-amino-pyridine: Compound 5.31



Yellow solid

Yield = 0.2g or 0.444mmol or 44%

Melting point: 179.1 – 182.6 °C

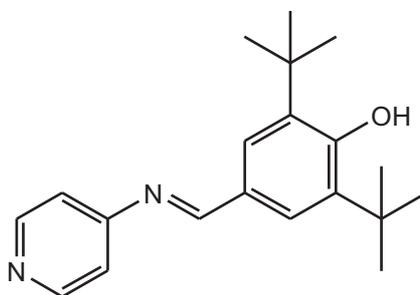
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 7.42 (dd, 2H,  $J$ = 1.264, 4.423 Hz, Py-3,5H); 8.02 (d, 1H,  $J$ = 2.527 Hz, Ph-6H); 8.201 (d, 2H,  $J$ = 2.527 Hz, Ph-4H); 8.64 (dd, 2H,  $J$ = 1.264, 4.423 Hz, Py-2,6H); 8.94 (s, 1H, CH); 13.88 (bs, 1H, OH); ppm.

IR (KBr)  $\nu$ = 3432, 3016, 2959, 2815, 2599, 1920, 1870, 1735, 1614, 1581 (-C=N-), 1482, 1436, 1160 (-OH)  $\text{cm}^{-1}$

MS (APCI -ve)  $m/z$  = found 373, 419, 449 (M - H) $^+$  and 450 (M) $^+$ ,

TLC carried out in ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.68.

### 6.5.3. Preparation of 4-[3,5-di-(*tert*-butyl)-4-OH-benzylidene]-amino-pyridine: Compound 5.30



Pale yellow solid.

Yield = 0.5g or 1.612mmol or 76%

Melting point: 141.1 – 143.3  $^{\circ}\text{C}$

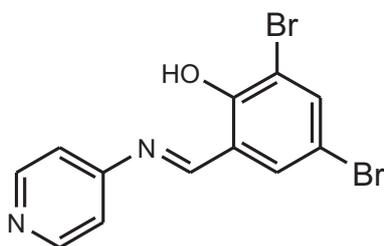
$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 1.315 (s, 18H, di-*tert* Ph); 7.3 (bd, 2H,  $J$ = 3.791 Hz, Py-3,5H); 7.7 (bs, 1H,  $J$ = 2.527 Hz, Ph-6H); 8.15 (s, 2H,  $J$ = 1.896 Hz, Ph-4H); 8.45 (dd, 2H,  $J$ = 1.264, 5.05 Hz, Py-2,6H); 10.4 (s, 1H, HC=N); ppm.

IR (KBr)  $\nu$ = 3627, 3286, 3071, 2955, 2869, 2599, 1924, 1824, 1735, 1631, 1575 (-C=N-), 1420, 1360, 1310 (*tert*-butyl), 1250, 1175 (-OH)  $\text{cm}^{-1}$

MS (APCI -ve)  $m/z$  = found 233, 279, 309 (M - H) $^+$ , 341 and 367.

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.71.

#### 6.5.4. Preparation of 4-[3, 5-di-bromo-2-OH-benzylidene]-amino-pyridine: Compound 5.32



State = Red amorphous.

Yield = 0.212g or 0.595mmol or 72%

Melting point: 167.3 – 170.7 °C

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 7.43 (dd, 1H, J= 1.896, 5.05 Hz, Py-3,5H); 7.92 (d, 1H, J= 2.527 Hz, Ph-6H); 8.01 (d, 2H, J= 2.527 Hz, Ph-4H); 8.65 (dd, 2H, J= 1.896, 5.05 Hz, Py-2,6H); 9.027 (s, 1H, HC=NH); ppm.

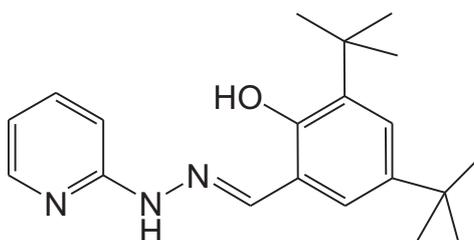
IR (KBr)  $\nu$  = 3441, 3274, 3090, 2971, 2830, 1766, 1634, 1577 (-C=N-), 1450, 1190 (-OH), 700 (Br) cm<sup>-1</sup>

MS (APCI -ve) m/z = found 355 (M - H)<sup>+</sup>.

MS (APCI +ve) m/z = found 357 (M + H)<sup>+</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.78.

#### 6.6.1. Preparation of 3,5-di-tert-butyl 2-OH benzylidene 2-hydrazinopyridine. Compound 5.10



2-hydrazinopyridine (0.1g or 0.916 mmol) and 3,5 di-*tert* butyl 2-OH benzaldehyde (0.239g or 1.02 mmol) was dissolved in 5ml ethanol and was refluxed for 18 hours with constant stirring. The final product was then filtered washed and then recrystallised from ethanol to give a yellow amorphous compound.

Yield is 0.153 g or 0.47mmol or 51 %

Mp: 178.0-182.8 °C and 187.2-189.6 °C

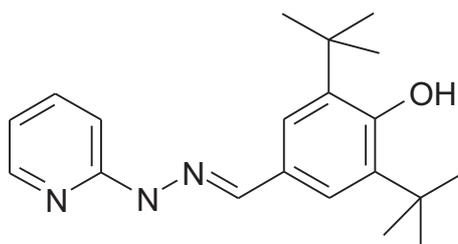
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.29 (s, 9H, CMe<sub>3</sub>); 1.42 (s, 9H, CMe<sub>3</sub>); 7.16 (d, 1H, *J*= 2.5 Hz, Ph-H4); 6.81-6.88 (overlapping m, 2H, Py-H5 and H3); 7.24 (d, 1H, *J*= 2.5 Hz, Ph-6H); 7.72 (ddd, 1H, *J*= 7.58, 5.05, 1.89 Hz, Py-H4); 8.20 (dd, 1H, *J*= 5.05, 1.26 Hz, Py-H6); 8.24 (s, 1H, N=CH); 11.07 (s, 1H, NH); 11.73 (s, 1H, OH); ppm

IR (KBr)  $\nu$ = 3463 (OH), 3213 (NH), 3069, 2957, 2865, 1593, 1541, 1458, 1396, 1366, 1317, 1291, 1241 cm<sup>-1</sup>

MS (+ES) *m/z*= 325 (M<sup>+</sup>), 326 (M+H)<sup>+</sup>, 327 (M+2H)<sup>+</sup>, 349 (M+H+Na)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.53

### 6.6.2. Preparation of 3,5-di-*tert*-butyl-4-OH benzylidene-4-hydrazinopyridine: Compound 5.11.



Yield: 0.021g or 0.065mmol or 7 %

Mp: 207.8-213.9 °C.

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 1.43 (s, 18H, di- *tert*-butyl ); 6.7 (overlapping m, H, Py-H5); 7.16 (dd, 1H, *J*= 8.21, 1.26 Hz, Py-H3); 7.4 (s, 2H, Ph-6H and Ph-H4); 10.55 (s, 1H, OH); 7.63 (overlapping, m, 1H, Py-H4); 7.95 (s, 1H, N=CH); 8.08 (dd, 1H, *J*= 5.05, 1.26 Hz, Py-H6); ppm

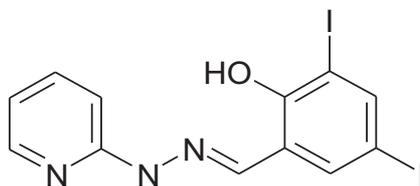
<sup>1</sup>H NMR (CDCl<sub>3</sub>;  $\delta$ CHCl<sub>3</sub>= 7.29 ppm): 1.48 (s, 18H, C Me<sub>3</sub> C3' and C Me<sub>3</sub> C5'), 6.77 (m, 1H, H3), 7.38 (d, 1H, *J*= 8.2 Hz, H5), 7.53 (s, 2H, H2' and H6'), 7.65 (m, 1H, H4), 7.76 (s, 1H, N=CH), 8.12 (m, 1H, H2), 8.44 (s, 1H, NH or OH) ppm

IR (KBr)  $\nu$ = 3623, 3435, 3187 (NH), 3063, 2953, 1601, 1542, 1463, 1443, 1320, 1233 cm<sup>-1</sup>

MS (+ES) *m/z*= 326 (M+H)<sup>+</sup>, 327 (M+2H)<sup>+</sup>, 391 (M+2H+2 CH<sub>3</sub>OH)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.72

### 6.6.3. Preparation of 3,5-diiodo-2-OH benzylidine-4-hydrazinopyridine Compound 5.12.



Yield: 0.272 g, 0.58mmol, 63 %

Mp: 211.1-212.3 °C

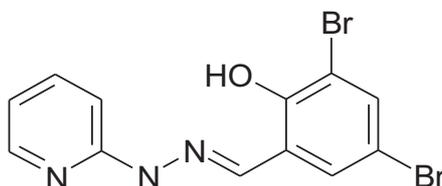
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 6.92-6.84 (overlapping m, 2H, Py-H5 and Py-H3); 7.70 (overlapping, m, 1H, Py-H4); 7.78 (d, 1H, *J* = 1.89 Hz, Ph-H6); 7.95 (d, 1H, *J* = 1.89 Hz, Ph-H4); 8.10 (s, 1H, N=CH); 8.22 (m, 1H, Py-H6); 11.3 (s, 1H, NH); 12.38 (s, 1H, OH); ppm

IR (KBr)  $\nu$  = 3437 (OH), 3168 (NH), 3063, 1602, 1596, 1545, 1449, 1313, 1227, 1147 cm<sup>-1</sup>

MS (+ES) *m/z* = 466 (M+H)<sup>+</sup>, 467 (M+2H)<sup>+</sup>, 536 (M+2H+3Na)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.65

### 6.6.4. Preparation of 3,5-dibromo-2-OH benzylidine-2-hydrazinopyridine: Compound 5.13.



Yield: 0.093 g, 0.2 5mmol, 54 %

Melting point: 242.8-245.8 °C

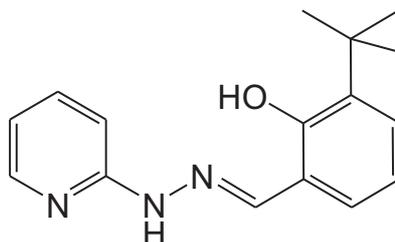
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 12.04 (bs, 1H, OH); 11.3 (s, 1H, NH); 8.21-8.22 (overlapping, m, 1H, Py-H6); 8.20 (s, 1H, N=CH); 7.70 – 7.74 (overlapping, m, 2H, Ph-H4 and Py-H4); 7.67 (d, 1H, *J* = 1.89 Hz, Ph-H6); 6.96 (d, *J* = 8.2 Hz, 1H, Py-H3); 6.88 (overlapping m, 1H, Py-H5); ppm

IR (KBr)  $\nu$  = 3463 (OH), 3180 (NH), 3069, 1596, 1544, 1439, 1343, 1314, 1248, 1149 cm<sup>-1</sup>

MS (+ES) *m/z* = 370/372/374 (M+H)<sup>+</sup> (Isotopes)

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.58

6.6.5. **Preparation of 3-*tert*-butyl-2-OH benzylidene-2-hydrazinopyridine: Compound 5.14.**



Yield: 0.002 g, 0.007 mmol,

Mp: 170.3-174 °C

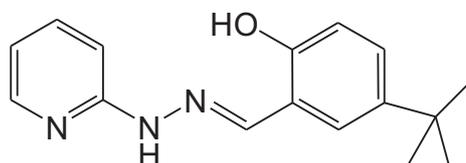
<sup>1</sup>H NMR: 1.42 (s, 9H, *tert*-butyl), 6.81-6.90 (overlapping m, 3H, H5,H3 and H5'), 7.18-7.24 (overlapping m, 2H, H4' and H6'), 7.72 (m, 1H, H4), 8.2 (m, 1H, H6), 8.25 (s, 1H, N=CH), 11.03 (s, 1H, NH), 11.89 (s, 1H, OH) ppm

IR (KBr)  $\nu$  = 3440 (OH), 3193 (NH), 3067, 2955, 1599, 1582, 1538, 1496, 1483, 1439, 1317  $\text{cm}^{-1}$

MS (+ES)  $m/z$  = 270 (M+H)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.64

6.6.6. **Preparation of 5-*tert*-butyl-2-OH benzylidene-2-hydrazinopyridine: Compound 5.15.**



Yield: 0.068 g, 0.25 mmol, 29 %

Mp: 173.2-176.7 °C

<sup>1</sup>H NMR ( $d_6$ - DMSO): 1.29 (s, 9H, *tert*-butyl), 6.75-6.84 (overlapping m, 2H, H3 and H5), 7.02 (d, 1H,  $J$  = 8.8 Hz, H3'), 7.24 (dd, 1H,  $J$  = 8.2, 2.5 Hz, H4'), 7.54 (d, 1H,  $J$  = 2.5 Hz, H6'), 7.67 (m, 1H, H4), 8.14 (dd, 1H,  $J$  = 5.1, 1.3 Hz, H6), 8.29 (s, 1H, N=CH), 10.89 (s, 1H, NH) ppm

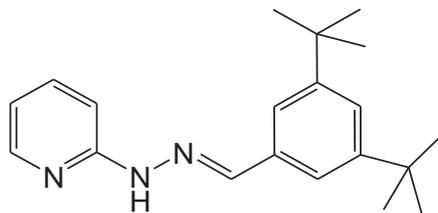
<sup>1</sup>H NMR ( $\text{CDCl}_3$ ;  $\delta\text{CHCl}_3$  = 7.26 ppm): 1.31 (s, 9H, C Me<sub>3</sub>), 6.84 (m, 1H, H3), 6.94 (d, 1H,  $J$  = 8.2 Hz, H5), 7.07 (d, 1H,  $J$  = 8.8 Hz, H3'), 7.17 (d, 1H,  $J$  = 2.5 Hz, H4'), 7.65 (m, 1H, H4), 7.94 (s, 1H, N=CH), 8.18 (dd, 1H,  $J$  = 1.3, 5.0 Hz, H2), 8.38 (s, 1H, NH or OH), 10.48 (br s, 1H, NH or OH) ppm

IR (KBr)  $\nu$  = 3457 (OH), 3193 (NH), 3109, 2957, 1597, 1534, 1494, 1442, 1366, 1307, 1255  $\text{cm}^{-1}$

MS (+ES)  $m/z$  = 270 (M+H)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.68.

### 6.6.7. Preparation of 3,5-di-*tert*-butyl-benzylidene-2-hydrazinopyridine: Compound 5.16.



Yield: 0.047 g, 0.15 mmol, 33 %

Mp: 165.8-168.5 °C

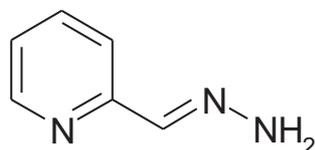
<sup>1</sup>H NMR (d<sub>6</sub>- DMSO): 1.33 (s, 18H, di- *tert*-butyl ); 6.76 (m, 1H, H5), 7.23 (d, 1H, *J*= 8.2 Hz, H3), 7.39 (m, 1H, Ph-H4); 7.48 (d, 2H, *J*= 1.89 Hz, Ph-H2 and Ph-H6); 7.66 (m, 1H, Py-H4); 8.03 (s, 1H, N=CH); 8.11 (dd, 1H, *J*= 5.05, 1.3 Hz, H6); 10.79 (s, 1H, NH) ppm

IR (KBr)  $\nu$ = 3193 (NH), 3061, 2954, 1602, 1585, 1538, 1462, 1442, 1363, 1316 cm<sup>-1</sup>

MS (+ES) *m/z*= 310 (M+H)<sup>+</sup>, 311 (M+2H)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.75.

### 6.7.1. Preparation of Pyridine-2-methylhydrazone: Compound 5.88.



An appropriate pyridine-2-carboxaldehyde (5ml) was dissolved in ethanol (5mL) and added dropwise through a pressure generated dropping funnel into a mixture of hydrazine monohydrate (9mL) in ethanol (5mL) stirred at 60°C for 24 hours. The product was collected by evaporating the excess hydrazine under high vacuum pump to give a yellow oil compound.

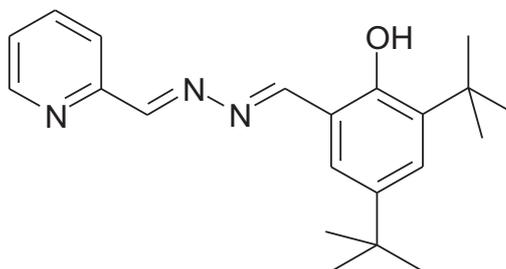
Yield was 10.7234 g, 98.37 mmol.

Melting point:

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 7.1 (ddd, 1H, *J*= 5.05, 3.2, 1.26 Hz, Py-H5), 7.4 (s, 2H, NH<sub>2</sub>); 7.6 (dd, 1H, *J*= 1.89, 7.58 Hz, Py-H4), 7.8 (d, 1H, *J*= 7.58 Hz, Py-H3), 7.9 (s, 1H, Py-CH=N), 8.50 (dd, 1H, *J*= 1.26, 5.05 Hz, Py-H6) ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.48.

**6.7.2. Preparation of 4-[3,5-di-(*tert*-butyl)-2-OH-benzyl]-1-(2-pyridyl)-2, 3-diaza-1, 3-butadiene: Compound 5.23.**



3, 5-Di-(*tert*-butyl)-2-hydroxybenzaldehyde (0.24gm or 1.02mmol) was dissolved in ethanol (5mL). Pyridine-2-carboxamidrazone (0.125gm or 1.03mmol) was added and the reaction was stirred at reflux for 18 hours. It was then cooled and the solvent was removed by filtration under vacuum. The title compound was collected as yellow amorphous solid, which was dried under vacuum.

Yield: 0.124 g, 0.37 mmol, 42 %

Melting point: 131.2-132.9 °C

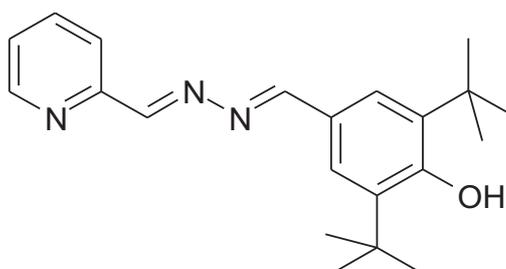
<sup>1</sup>H NMR (d<sub>6</sub>- DMSO): 1.31 (s, 9H, *tert*-butyl); 1.44 (s, 9H, *tert*-butyl); 7.43-7.58 (overlapping m, 3H, Py-H5, Ph-H4 and Ph-H6); 8.03 (dd, 1H, *J*= 7.6, 1.89 Hz, Py-H4); 8.12 (d, 1H, *J*= 7.6 Hz, Py-H3); 8.74-8.77 (overlapping m, 2H, Ph-CH=N and Py-H6); 9.08 (d, 1H, *J*= 13.9 Hz, Py-CH=N); 12.12 (br s, 1H, OH) ppm

IR (KBr)  $\nu$ = 3432 (OH), 2954, 1625, 1600, 1592, 1562, 1473, 1443, 1390, 1363 cm<sup>-1</sup>

MS (+ES) *m/z*= 338 (M+H)<sup>+</sup>, 339 (M+2H)<sup>+</sup>, 392 (M+H+3NH<sub>3</sub>)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.75

**6.7.3. Preparation of 4-[3,5-di-(*tert*-butyl)-4-OH-benzyl]-1-(2-pyridyl)-2,3-diaza-1,3-butadiene. Compound 5.24.**



Yield: 0.096 g, 0.28 mmol, 44 %

Melting point: 114.6-116.7 °C

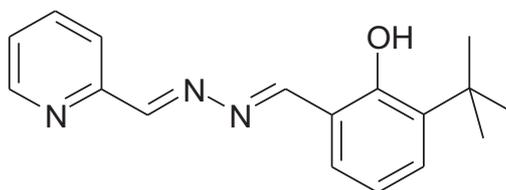
$^1\text{H}$  NMR ( $d_6$ - DMSO): 1.43 (s, 18H, di- *tert*-butyl); Ph-C3 and Ph-C5); 7.50 (ddd, 1H,  $J$ = 5.1, 3.2, 1.26 Hz, Py-H5); 7.70 (s, 2H, Ph-H2 and Ph-H6); 7.94 (dd, 1H,  $J$ = 7.58, 1.7 Hz, Py-H4); 8.08 (d, 1H,  $J$ = 8.2 Hz, Py-H3); 8.56 (s, 1H, Ph-CH=N); 8.65 (s, 1H, Py-CH=N); 8.70 (m, 1H, H6) ppm

IR (KBr)  $\nu$ = 3492 (OH), 3067, 2941, 2862, 1625, 1592, 1555, 1475, 1439, 1412, 1367, 1300  $\text{cm}^{-1}$

MS (+ES)  $m/z$ = 338 ( $\text{M}+\text{H}$ )<sup>+</sup>, 339 ( $\text{M}+2\text{H}$ )<sup>+</sup>, 401 ( $\text{M}+2\text{CH}_3\text{OH}$ )<sup>+</sup>, 402 ( $\text{M}+\text{H}+2\text{CH}_3\text{OH}$ )<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.76

#### 6.7.4. Preparation of 4-[3-(*tert*-butyl)-2-OH-benzyl]-1-(2-pyridyl)-2,3-diaza-1,3-butadiene: Compound 5.25.



Yield: 0.027g, 0.096mmol, 12%

Melting point: 264.1-269.9 °C

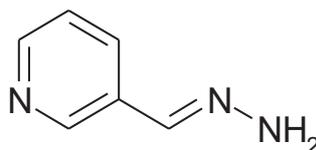
$^1\text{H}$  NMR: 1.43 (s, 9H, *tert*-butyl), 6.96 (t, 1H,  $J$ = 7.6 Hz, H5'), 7.41-7.58 (overlapping m, 3H, H5, H4' and H6'), 7.84 (td, 1H,  $J$ = 7.2, 1.5 Hz, H4), 8.12 (d, 1H,  $J$ = 7.6 Hz, H3), 8.75 (m, 1H, H6), 8.79 (s, 1H, Ar-CH=N), 9.05 (s, 1H, Py-CH=N), 12.29 (s, 1H, OH) ppm

IR (KBr)  $\nu$ = 3454 (OH), 3059, 2958, 2866, 1611, 1489, 1462, 1429, 1396, 1197  $\text{cm}^{-1}$

MS (+ES)  $m/z$ = 282 ( $\text{M}+\text{H}$ )<sup>+</sup>, 283 ( $\text{M}+2\text{H}$ )<sup>+</sup>, 353 ( $\text{M}+\text{H}+3\text{Na}$ )<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.83

#### 6.8. 1. Preparation of Pyridine-3-methylhydrazone: Compound 5.86.



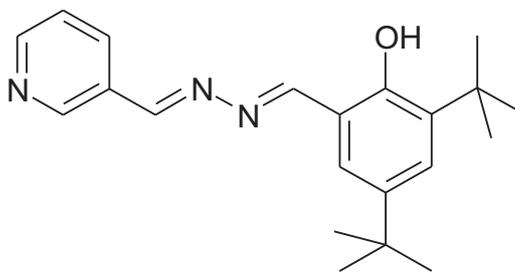
An appropriate pyridine-3-carboxaldehyde (3.54g or 32.477 mmol) was dissolved in ethanol (10mL) and added dropwise through a pressure generated dropping funnel into a mixture of

hydrazine monohydrate (9.5 mL) in ethanol (5mL) stirred at 70°C for 24 hours. The product was collected by evaporating the excess hydrazine under high vacuum pump to give a yellow oil compound.

Yield was 4.01g, 32.338mmol or 99.5%

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 7.1 (s, 2H, NH<sub>2</sub>); 7.3 (overlapping m, 1H, Py-H); 7.7 (s, 1H, Py-CH=N); 7.8 (dd, *J* = 1.89, 7.58 Hz Py-H); 8.4 (dd, 1H, *J* = 1.89, 5.05 Hz, Py-H); 8.7 (dd, 1H, *J* = 1.89 Hz, Py-H); ppm

#### 6.8.2. Preparation of 4-[3,5-di-(*tert*-butyl)-2-OH-benzyl]-1-(3-pyridyl)-2, 3-diaza-1, 3-butadiene. Compound 5.26



Yield: 0.074 g, 0.22 mmol, 27 %

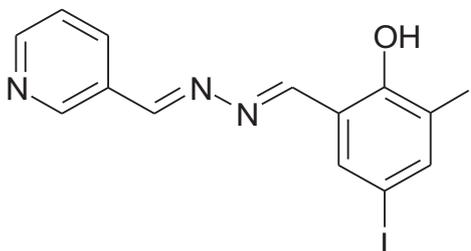
Melting point: 141.8-147.1 °C

<sup>1</sup>H NMR: 1.31 (s, 9H, *tert*-butyl), 1.44 (s, 9H, *tert*-butyl), 7.42-7.51 (overlapping m, 2H, H4' and H6'), 7.58 (dd, 1H, *J* = 8.2, 5.1 Hz, H5), 8.29 (m, 1H, H4), 8.73 (dd, 1H, *J* = 5.1, 1.9 Hz, H6), 8.95 (s, 1H, Ar-CH=N), 9.03 (m, 1H, Py-CH=N), 9.10 (s, 1H, H2), 12.24 (s, 1H, OH) ppm  
IR (KBr)  $\nu$  = 3426 (OH), 2954, 2875, 1625, 1595, 1578, 1552, 1466, 1439, 1416, 1393, 1317 cm<sup>-1</sup>

MS (+ES) *m/z* = 338 (M+H)<sup>+</sup>, 339 (M+2H)<sup>+</sup>, 379 (M+H+CH<sub>3</sub>CN)<sup>+</sup>

T TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.72

**6.8.3. Preparation of 4-[3,5-di-iodo-2-OH-benzyl]-1-(3-pyridyl)-2,3-diaza-1, 3-butadiene: Compound 5.27**



Yield: 0.041 g, 0.087 mmol, 22 %

Melting point: 215.8-218.5 °C

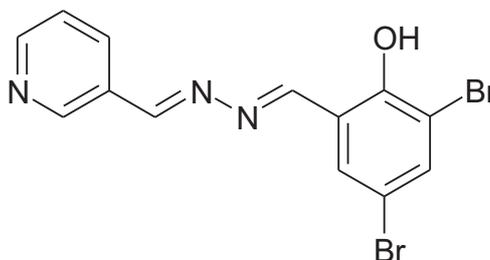
<sup>1</sup>H NMR: 7.58 (dd, 1H, *J*= 4.8, 7.9 Hz, H5), 8.04 (d, 1H, *J*= 2.5 Hz, H6'), 8.17 (d, 1H, *J*= 1.9 Hz, H4'), 8.27 (m, 1H, H4), 8.74 (dd, 1H, *J*= 1.6, 4.8 Hz, H6), 8.92 (s, 1H, Ar-CH=N), 8.99-9.02 (overlapping m, 2H, Py-CH=N and H2), 12.62 (br s, 1H, OH) ppm

IR (KBr)  $\nu$ = 3439 (OH), 3034, 1625, 1578, 1476, 1416, 1370, 1310, 1270 cm<sup>-1</sup>

MS (+ES) *m/z*= 478 (M+H)<sup>+</sup>, 519 (M+H+CH<sub>3</sub>CN)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.65

**6.8.4. Preparation of 4-[3,5-di-bromo-2-OH-benzyl]-1-(3-pyridyl)-2,3-diaza-1,3-butadiene: Compound 5.28.**



Yield: 0.069 g, 0.18 mmol, 42 %

Melting point: 178.8-181.1 °C

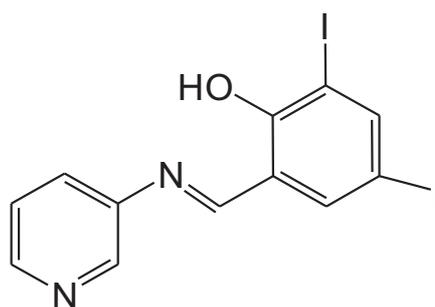
<sup>1</sup>H NMR: 7.58 (dd, 1H, *J*=7.9, 4.7Hz, H5), 7.94-7.98 (overlapping m, 2H, H4' and H6'), 8.28 (m, 1H, H4), 8.75 (dd, 1H, *J*= 4.8, 1.6 Hz, H6), 8.99-9.03 (overlapping m, 2H, Ar-CH=N and H2), 9.07 (s, 1H, Py-CH=N), 12.32 (br s, 1H, OH) ppm

IR (KBr)  $\nu$ = 3431 (OH), 3068, 1625, 1592, 1562, 1443, 1376, 1310 cm<sup>-1</sup>

MS (+ES) *m/z*= 382/384/386 (M+H)<sup>+</sup>, 423/425/427 (M+H+CH<sub>3</sub>CN)<sup>+</sup> (Isotopes)

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.58.

## 6.9.1. Preparation of 3,5-diiodo-2-OH benzylidene-3-aminopyridine: Compound 5.35.



Yield: 0.085 g, 0.19 mmol, 18 %

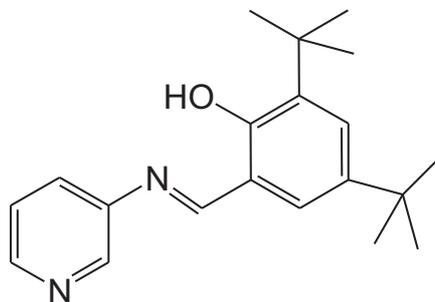
Melting point: 175.7-177.4 °C

<sup>1</sup>H NMR: 7.54 (dd, 1H, *J* = 5.1, 8.2 Hz, H5), 7.93 (ddd, 1H, *J* = 1.3, 2.5, 5.7 Hz, H4), 8.02 (d, 1H, *J* = 1.9 Hz, H6'), 8.20 (d, 1H, *J* = 1.9 Hz, H4'), 8.56 (dd, 1H, *J* = 1.6, 5.1 Hz, H6), 8.72 (d, 1H, *J* = 2.5 Hz, H2), 8.98 (s, 1H, N=CH), 14.14 (br s, 1H, OH) ppm

IR (KBr)  $\nu$  = 3439 (OH), 3054, 1611, 1572, 1429, 1270, 1157 cm<sup>-1</sup>

MS (+ES) *m/z* = 451 (M+H)<sup>+</sup>, 492 (M+H+CH<sub>3</sub>CN)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.66

6.9.2. Preparation of 3,5-di-(*tert*-Butyl)-2-OH benzylidene-3-aminopyridine: Compound 5.33.

Yield: 0.079 g, 0.25 mmol, 23%

Melting point: 114.6-116.7 °C

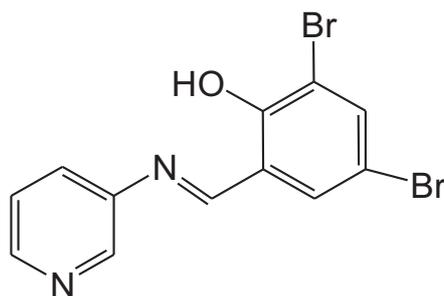
<sup>1</sup>H NMR: 1.31 (s, 9H, *tert*-butyl), 1.43 (s, 9H, *tert*-butyl), 7.44 -7.54 (overlapping m, 3H, H6', H4' and H5), 7.92 (ddd, 1H, *J* = 1.3, 2.5, 5.7 Hz, H4), 8.52 (dd, 1H, *J* = 1.6, 5.1 Hz, H6), 8.68 (d, 1H, *J* = 2.5 Hz, H2), 9.07 (s, 1H, N=CH), 13.61 (s, 1H, OH) ppm

IR (KBr)  $\nu$  = 3431 (OH), 3026, 2953, 2861, 1616, 1569, 1475, 1436, 1363, 1204 cm<sup>-1</sup>

MS (+ES) *m/z* = 311 (M+H)<sup>+</sup>, 353 (M+2H+ CH<sub>3</sub>CN)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.66.

## 6.9.3. Preparation of 3,5-dibromo-2-OH benzylidene-3-aminopyridine: Compound 5.36.



Yield: 0.066 g, 0.19 mmol, 35 %

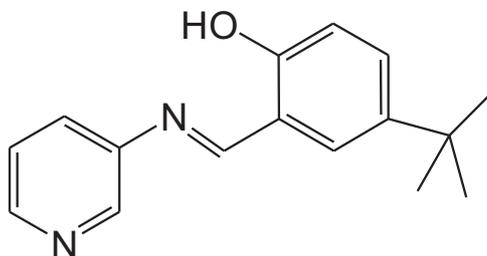
Melting point: 163.3-164.8 °C

<sup>1</sup>H NMR: 7.55 (dd, 1H, *J* = 5.7, 8.2 Hz, H5), 7.91-8.00 (overlapping m, 3H, H4, H4' and H6'), 8.57 (dd, 1H, *J* = 1.6, 4.7 Hz, H6), 8.72 (d, 1H, *J* = 2.5 Hz, H2), 9.07 (s, 1H, N=CH), 14.00 (br s, 1H, OH) ppm

IR (KBr)  $\nu$  = 3445 (OH), 3061, 1612, 1579, 1555, 1483, 1449, 1364  $\text{cm}^{-1}$

MS (+ES) *m/z* = 355/357/359 (M+H)<sup>+</sup>, 396/398/400 (M+H+CH<sub>3</sub>CN)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.61

6.9.4. Preparation of 5-(*tert*-Butyl)-2-OH benzylidene-3-aminopyridine: Compound 5.37.

Yield: 0.032 g, 0.13 mmol, 13 %

Melting point: 119.8-120.9 °C

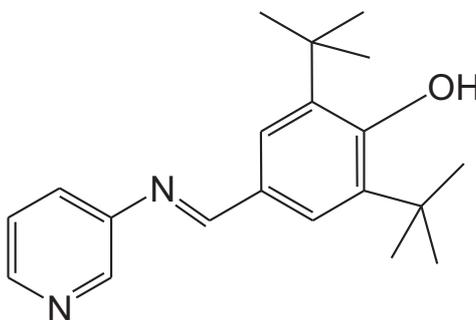
<sup>1</sup>H NMR: 1.31 (s, 9H, *tert*-butyl), 6.94 (d, 1H, *J* = 8.2 Hz, H3'), 7.48-7.54 (overlapping m, 2H, H4' and H5), 7.71 (d, 1H, *J* = 2.5 Hz, H6'), 7.86 (ddd, 1H, *J* = 1.3, 2.5, 5.7 Hz, H4), 8.51 (dd, 1H, *J* = 1.6, 4.8 Hz, H6), 8.63 (d, 1H, *J* = 2.5 Hz, H2), 9.03 (s, 1H, N=CH), 12.41 (br s, 1H, OH) ppm

IR (KBr)  $\nu$  = 3432 (OH), 3034, 2961, 2868, 1618, 1562, 1485, 1388, 1360, 1346  $\text{cm}^{-1}$

MS (+ES) *m/z* = 255 (M+H)<sup>+</sup>, 296 (M+H+CH<sub>3</sub>CN)<sup>+</sup>, 297 (M+2H+CH<sub>3</sub>CN)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.57

6.9.5. **Preparation of 3,5-di-(*tert*-Butyl)-4-OH benzylidene-3-aminopyridine: Compound 5.34.**



Yield: 0.024 g, 0.08 mmol, 15 %

Melting point: 135.8-138.7 °C

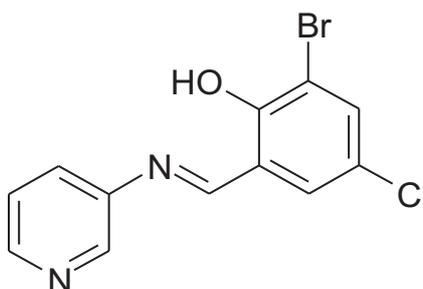
<sup>1</sup>H NMR: 1.39 (s, 18H, di-*tert*-butyl), 7.42 (dd, 1H, *J*= 8.2, 4.7 Hz, H5), 7.70 (br overlapping m, 3H, H2',H6' and H4), 8.37 (d, 1H, *J*= 4.4 Hz, H2), 8.53 (br overlapping s, 2H, H6 and N=CH) ppm

IR (KBr)  $\nu$ = 3379 (OH), 3034, 2961, 2868, 1625, 1565, 1479, 1433, 1396, 1360, 1300 cm<sup>-1</sup>

MS (+ES) *m/z*= 311 (M+H)<sup>+</sup>, 312 (M+2H)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.69.

6.9.6. **Preparation of 3-bromo-5-chloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.94.**



Yield: 0.43 g, 1.38 mmol, 45 %

Melting point: 135.2-137.5 °C

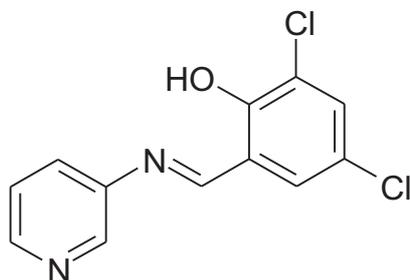
<sup>1</sup>H NMR: 14.0 (bs, 1H, OH); 9.07 (s, 1H, CH=NH), 8.73 (d, 1H, *J*= 2.52 Hz Py-2H); 8.57 (dd, 1H, *J*= 1.89, 4.42 Hz Py-6H); 7.95 (ddd, 1H, *J*= 1.89, 4.42, 8.21 Hz Py-4H); 7.91 (d, 1H, *J*= 2.57 Hz Ph-4H); 7.79 (d, 1H, *J*= 2.57 Hz Ph-6H); 7.55 (dd, 1H, *J*= 8.21, 4.42 Hz Py-5H) ppm

IR (KBr)  $\nu$ = 3440 (OH), 3067, 3046, 2896, 2830, 1930, 1787, 1611 (-CH=N-), 1565, 1479, 1433, 1396, 1360, 1300 cm<sup>-1</sup>

MS (+ES) *m/z*= 311 (M+H)<sup>+</sup>, 312 (M+2H)<sup>+</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.71.

6.9.7. **Preparation of 3,5-dichloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.95.**



Yield: 0.426 g, 1.59 mmol, 45 %

Melting point: 144.7-148.3 °C

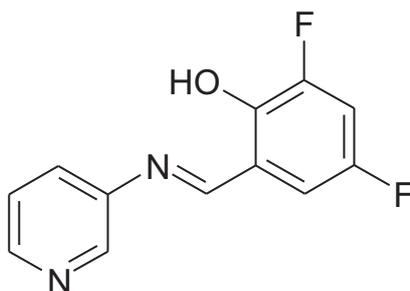
<sup>1</sup>H NMR: 14.0 (bs, 1H, OH); 9.07 (s, 1H, CH=NH), 8.73 (d, 1H, *J* = 2.52 Hz Py-2H); 8.57 (dd, 1H, *J* = 1.89, 4.42 Hz Py-6H); 7.95 (ddd, 1H, *J* = 1.89, 4.42, 8.21 Hz Py-4H); 7.80 (d, 1H, *J* = 2.57 Hz Ph-4H); 7.76 (d, 1H, *J* = 2.57 Hz Ph-6H); 7.55 (dd, 1H, *J* = 8.21, 4.42 Hz Py-5H) ppm  
 IR (KBr)  $\nu$  = 3447 (OH), 3077, 3054, 2840, 2830, 1932, 1741, 1615 (-CH=N-), 1563, 1479, 1433, 1396, 1360, 1300  $\text{cm}^{-1}$

MS (+ES) *m/z* = 267 (M+H)<sup>+</sup>, 269 (M+2H)<sup>+</sup>

Accurate MS (ES +ve) *m/z*: Calculated 267.0092, found 267.0105 (M + H)<sup>+</sup>; 4.9 ppm.

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.7.

6.9.8. **Preparation of 3,5-di-fluoro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.98.**



Yield: 0.115 g, 0.491 mmol, 45 %

Melting point: 121.2-125.1 °C

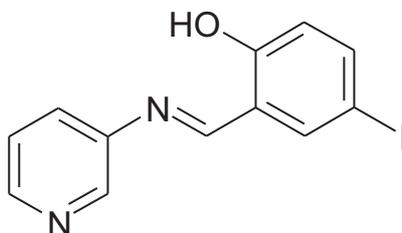
<sup>1</sup>H NMR: 12.7 (bs, 1H, OH); 9.03 (s, 1H, CH=NH), 8.67 (d, 1H, *J* = 2.52 Hz Py-2H); 8.52 (dd, 1H, *J* = 1.89, 4.42 Hz Py-6H); 7.90 (ddd, 1H, *J* = 1.89, 4.42, 8.21 Hz Py-4H); 7.57 (d, 1H, *J* = 2.57 Hz Ph-4H); 7.52 (d, 1H, *J* = 2.57 Hz Ph-6H); 7.42 (dd, 1H, *J* = 8.21, 4.42 Hz Py-5H) ppm  
 IR (KBr)  $\nu$  = 3413 (OH), 3063, 2734, 2350, 1698, 1619 (-CH=N-), 1598, 1476, 1433, 1396, 1360, 1300  $\text{cm}^{-1}$

MS (+ES) *m/z* = 235 (M+H)<sup>+</sup>, 236 (M+2H)<sup>+</sup>

Accurate MS (ES +ve) m/z: Calculated 235.0694, found 235.0690 (M + H)<sup>+</sup>; 3 ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.6.

#### 6.9.9. Preparation of 5-Iodo-2-hydroxy benzylidene-3-aminopyridine: Compound 5.99.



Yield: 0.174 g, 0.537 mmol, 45 %

Melting point: 158.2-160.5 °C

<sup>1</sup>H NMR: 12.7 (bs, 1H, OH); 8.95 (s, 1H, CH=NH), 8.63 (d, 1H, J= 2.52 Hz Py-2H); 8.52 (dd, 1H, J= 1.89, 4.42 Hz Py-6H); 8.05 (d, 1H, J= 2.52 Hz Ph-H); 7.86 (ddd, 1H, J= 1.89, 4.42, 8.21 Hz Py-4H); 7.73 (dd, 1H, J= 8.84, 1.89 Hz Ph-H); 7.52 (dd, 1H, J= 8.21, 4.42 Hz Py-5H); 6.85 (d, 1H, J= 8.21 Hz Ph-H); ppm

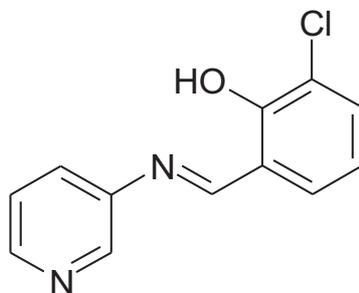
IR (KBr) ν = 3424 (OH), 3060, 2963, 2348, 1958, 1766, 1675, 1612 (-CH=N-), 1554, 1479, 1433, 1396, 1360, 1300 cm<sup>-1</sup>

MS (+ES) m/z = 324 (M+H)<sup>+</sup>, 325 (M+2H)<sup>+</sup>

Accurate MS (ES +ve) m/z: Calculated 324.9838, found 324.9832 (M + H)<sup>+</sup> 1.8 ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.5.

#### 6.9.10. Preparation of 3-chloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.97.



Yield: 0.207 g, 0.892 mmol, 45 %

Melting point: 54.2-58.5 °C

<sup>1</sup>H NMR: 13.85 (s, 1H, OH); 9.12 (s, 1H, CH=NH), 8.72 (d, 1H, *J* = 2.52 Hz Py-2H); 8.55 (dd, 1H, *J* = 1.89, 4.42 Hz Py-6H); 7.96 (ddd, 1H, *J* = 1.89, 4.42, 8.21 Hz Py-4H); 7.73 (dd, 1H, *J* = 8.84, 1.89 Hz Ph-H); 7.68-7.62 (m, 2H, Ph-H); 7.52 (dd, 1H, *J* = 8.21, 4.42 Hz Py-5H); 7.04 (t, 1H, *J* = 7.58, 8.21 Hz Ph-H); ppm

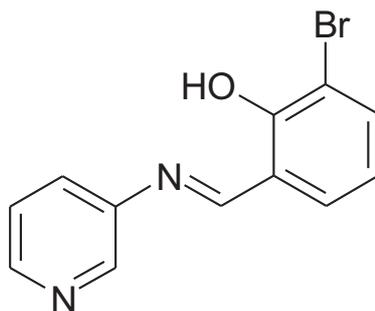
IR (KBr)  $\nu$  = 3426 (OH), 3052, 3025, 2855, 2572, 2350, 2281, 1912, 1685, 1607 (-CH=N-), 1554, 1479, 1433, 1396, 1360, 1300 cm<sup>-1</sup>

MS (+ES) *m/z* = 233 (M+H)<sup>+</sup>, 334 (M+2H)<sup>+</sup>

Accurate MS (ES +ve) *m/z*: Calculated 233.0482, found 233.0480 (M + H)<sup>+</sup> 0.9 ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.75.

#### 6.9.11. Preparation of 2-bromo-2-hydroxy benzylidene-3-aminopyridine: Compound 5.100.



Yield: 0.201 g, 0.725 mmol, 45 %

Melting point: 67.6-73.2 °C

<sup>1</sup>H NMR: 13.97 (s, 1H, OH); 9.09 (s, 1H, CH=NH), 8.73 (d, 1H, *J* = 2.52 Hz Py-2H); 8.56 (dd, 1H, *J* = 1.89, 4.42 Hz Py-6H); 7.98 (ddd, 1H, *J* = 1.89, 4.42, 8.21 Hz Py-4H); 7.76 (dd, 1H, *J* = 8.21, 1.26 Hz Ph-H); 7.75 (d, 1H, *J* = 1.89, 7.58 Ph-H); 7.56 (dd, 1H, *J* = 8.21, 4.42 Hz Py-5H); 6.99 (t, 1H, *J* = 7.58, 8.21 Hz Ph-H); ppm

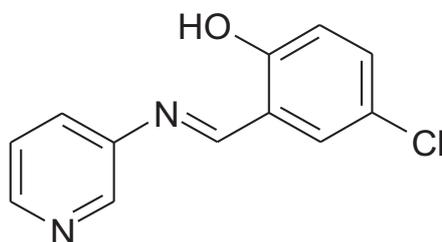
IR (KBr)  $\nu$  = 3426 (OH), 3058, 2854, 2439, 1965, 1918, 1612 (-CH=N-), 1554, 1479, 1433, 1396, 1360, 1300 cm<sup>-1</sup>

MS (+ES) *m/z* = 276 (M+H)<sup>+</sup>, 277 (M+2H)<sup>+</sup>

Accurate MS (ES +ve) *m/z*: Calculated 276.9976, found 276.9974 (M + H)<sup>+</sup> 0.7 ppm.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.71.

6.9.12. **Preparation of 5-chloro-2-hydroxy benzylidene-3-aminopyridine: Compound 5.96.**



Yield: 0.115 g, 0.495 mmol, 45 %

Melting point: 130.2-131.8 °C

<sup>1</sup>H NMR: 12.7 (bs, 1H, OH); 8.91 (s, 1H, CH=NH), 8.63 (d, 1H, *J* = 2.52 Hz Py-2H); 8.52 (dd, 1H, *J* = 1.89, 4.42 Hz Py-6H); 7.86 (ddd, 1H, *J* = 1.89, 4.42, 8.21 Hz Py-4H); 7.78 (d, 1H, *J* = 2.52 Hz Ph-H); 7.53-7.45 (m, 2H, Py-5H, Ph-H); 6.85 (d, 1H, *J* = 8.84 Hz Ph-H); ppm

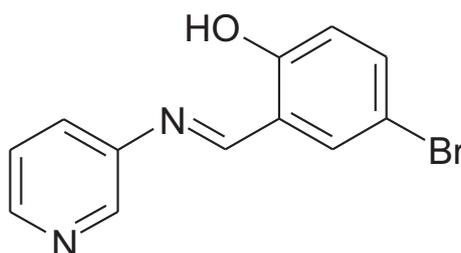
IR (KBr)  $\nu$  = 3447 (OH), 3068, 2622, 1965, 1899, 1756, 1613 (-CH=N-), 1554, 1479, 1396, 1360, 1300  $\text{cm}^{-1}$

MS (+ES) *m/z* = 233 (M+H)<sup>+</sup>, 235 (M+2H)<sup>+</sup>

Accurate MS (ES +ve) *m/z*: Calculated 233.0482, found 233.0481 (M + H)<sup>+</sup> 0.4 ppm.

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.7.

6.9.13. **Preparation of 5-bromo-2-hydroxy benzylidene-3-aminopyridine: Compound 5.101.**



Yield: 0.5 g, 1.805 mmol, 45 %

Melting point: 128.6-132.1 °C

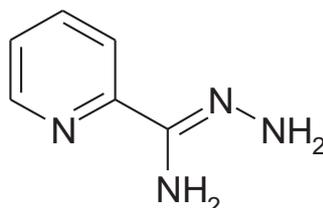
<sup>1</sup>H NMR: 12.5 (bs, 1H, OH); 8.95 (s, 1H, CH=NH), 8.63 (d, 1H, *J* = 2.52 Hz Py-2H); 8.52 (dd, 1H, *J* = 1.89, 4.42 Hz Py-6H); 7.91 (d, 1H, *J* = 3.159 Hz Ph-H); 7.86 (ddd, 1H, *J* = 1.89, 4.42, 8.21 Hz Py-4H); 7.61 (dd, 1H, *J* = 8.84, 2.52 Hz Ph-H); 7.52 (dd, 1H, *J* = 8.21, 4.42 Hz Py-5H); 6.9 (d, 1H, *J* = 8.84 Hz Ph-H); ppm

IR (KBr)  $\nu$  = 3472 (OH), 3058, 2854, 2394, 1958, 1926, 1760, 1613 (-CH=N-), 1554, 1479, 1433, 1396, 1360, 1300  $\text{cm}^{-1}$

MS (+ES)  $m/z = 278 (M+H)^+$

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.68.

#### 6.10.1. Preparation of pyridine-2-carboxamidrazone: Compound 5.89.



2- cyanopyridine (30.103g, 288.4mmol) was suspended in ethanol (90mL) and treated with hydrazine monohydrate (60mL) and stirred at ambient temperature for two days. The resulting precipitate was collected by filtration under vacuum and washed with petroleum ether (40-60°C) to give a white crystalline compound.

Yield was 31.664g or 234mmol or 83%.

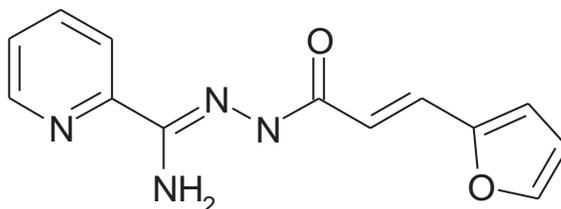
Melting point:

$^1\text{H}$  NMR (250 MHz  $d_6$ -DMSO): 5.24(bs, 2H, NH<sub>2</sub>); 5.71 (s, 2H, NH<sub>2</sub>); 7.31 (td, 1H, J=4.9, 1.2 Hz, Pyr-H4); 7.74 (td, 1H, J=6.92, 1.8 Hz, Pyr-H5); 7.91 (m, 1H, Pyr-H3); 8.48 (m, 1H, Pyr-H6) ppm.

MS (APCI +ve)  $m/z = 137 (m+H)^+$ .

TLC, carried out in ethyl acetate: methanol (2:1) ( $R_f$ ) = 0.1 (single spot).

#### 6.10.2. Preparation of N<sup>1</sup>-(2-furoylacryl)-pyridine-2-carboxamidrazone: Compound 5.38.



Anhydrous tetrahydrofuran (8mL) was added into a mixture of furoylacrylic acid (1.078g or 1.054 mmol) and carbonyl *di*-imidazole (1.30gm or 1.065mmol). The mixture was stirred at room temperature for 1 hr and then pyridine 2-carboxamidrazone (0.9824g or 7.2 mmol) dissolved in dry tetrahydrofuran (10mL) was added and the reaction was stirred at room

temperature for 18 hrs. The reaction mixture was poured dropwise into saturated sodium bicarbonate solution (75ml). The resulting precipitate was filtered dried and recrystallised from ethanol to give a brownish yellow amorphous solid.

Yield was 0.275g 1.074 mmol or 15%

Melting Point: 162.3 – 168.6 °C.

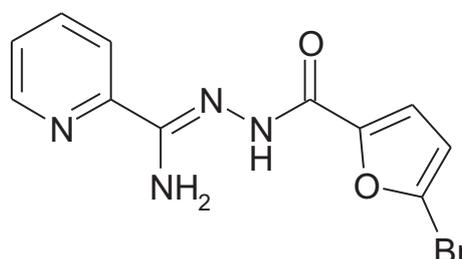
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 6.57 (d, 2H, J= 15.796 Hz, NH<sub>2</sub>); 6.6 (dd, 1H, furyl-H); 6.7 (s, 1H, NH); 6.8 (d, 1H, J=3.478 Hz, furyl-H); 7.35 (d, 1H, J= 15.638 Hz, CH); 7.44 (d, 1H, furyl-H); 7.5 (ddd, 1H, Py-5H); 7.9 (ddd, 1H, Py-4H); 8.1 (d, 1H, J= 8.056 Hz, Py-3H); 8.6 (d, 1H, J= 4.42 Hz, Py-6H); 10.2 (s, 1H, NH) ppm.

MS (+ electro spray): 257 (M + H)<sup>+</sup>; 279 (M+ Na)<sup>+</sup>

IR (KBr)  $\nu$  = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

TLC, carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.55 (single spot).

### 6.10.3. Preparation of N<sup>1</sup>-2-(5-bromofuroyl)-pyridine-2-carboxamidrazone: Compound 5.41.



Yield = 0.52g or 1.68mmol or 40%

Melting Point: 207.1 – 209.3 °C

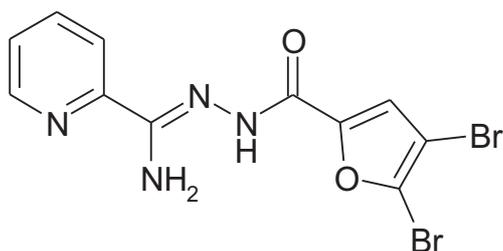
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 6.8 (d, 1H, furyl-H); 7.0 (s, 2H, NH<sub>2</sub>); 7.3 (d, 1H, J=3.478 Hz, furyl-H); 7.5 (ddd, 1H, Py-5H); 7.9 (ddd, 1H, Py-4H); 8.1 (d, 1H, J= 8.056 Hz, Py-3H); 8.6 (d, 1H, J= 4.42 Hz, Py-6H); 10.2 (s, 1H, NH) ppm.

MS (+APCI): 309, 311 (M + H)<sup>+</sup>; 292 (M- NH<sub>2</sub>)<sup>+</sup>

IR (KBr)  $\nu$  = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

TLC, carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.55 (single spot).

**6.10.4. Preparation of N<sup>1</sup>-2-(4,5-dibromofuroyl)-pyridine-2-carboxamidrazone: Compound 5.42.**



Yield = 0.201g or .515mmol or 35%

Melting Point: 237.3 – 239 °C.

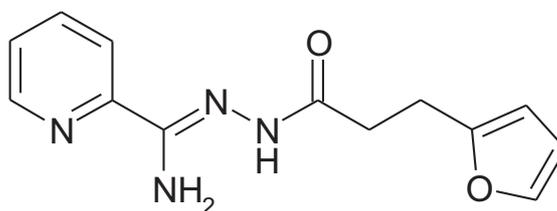
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 6.85 (bs, 2H, NH<sub>2</sub>); 7.45 (s, 1H, furyl-H); 7.5 (ddd, 1H, Py-5H); 7.9 (ddd, 1H, Py-4H); 8.1 (d, 1H, J = 8.056 Hz, Py-3H); 8.6 (d, 1H, J = 4.42 Hz, Py-6H); 10.2 (bs, 1H, NH) ppm.

MS (+ APCI): 387, 389, 391 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

TLC, carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.73 (single spot).

**6.10.5. Preparation of N<sup>1</sup>-2-[3-(2-furoyl)-propyl]-pyridine-2-carboxamidrazone: Compound 5.44.**



Yield = .426g or 1.77mmol or 65.5%

Melting Point: 177.8 – 179.7 °C.

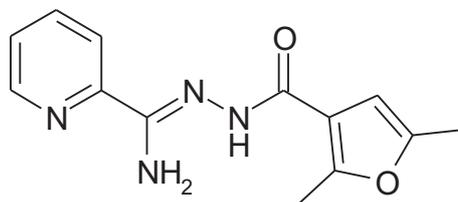
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 2.5 (d, 2H, CH<sub>2</sub>); 2.9 (d, 2H, CH<sub>2</sub>); 6.18 (d, 1H, J = 15.796 Hz, furyl-H); 6.3 (dd, 1H, furyl-H); 6.7 (ss, 2H, NH<sub>2</sub>); 7.5 (ddd, 1H, Py-5H); 7.55 (d, 1H, furyl-H); 7.9 (ddd, 1H, Py-4H); 8.1 (dd, 1H, J = 8.056 Hz, Py-3H); 8.6 (dd, 1H, J = 4.42 Hz, Py-6H); 10.2 (s, 1H, NH) ppm.

MS (+ APCI): 259 (M + H)<sup>+</sup>; 241

IR (KBr)  $\nu$  = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

TLC, carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.62 (single spot).

**6.10.6. Preparation of N<sup>1</sup>-3-(2,5-dimethylfuroyl)-pyridine-2-carboxamidrazone: Compound 5.43.**



Yield = 1.2g or 4.651mmol or 60%

Melting Point: 242.3 – 246.8 °C.

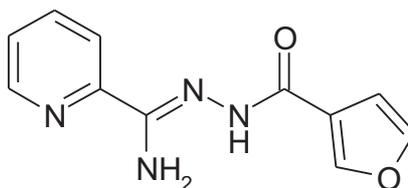
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 2.2 (s, 3H, CH<sub>3</sub>); 2.4 (s, 3H, CH<sub>3</sub>); 6.6 (s, 2H, NH<sub>2</sub>); 6.9 (s, 1H, furyl-H); 7.5 (ddd, 1H, Py-5H); 7.9 (ddd, 1H, Py-4H); 8.1 (d, 1H, J= 8.056 Hz, Py-3H); 8.6 (d, 1H, J= 4.42 Hz, Py-6H); 10.2 (s, 1H, NH) ppm.

MS (+APCI): 259 (M + H)<sup>+</sup>; 241 (M- NH<sub>2</sub>)<sup>+</sup>

IR (KBr) ν = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

TLC, carried out in ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.33 (single spot).

**6.10.7. Preparation of N<sup>1</sup>-3-furoyl-pyridine-2-carboxamidrazone: Compound 5.39.**



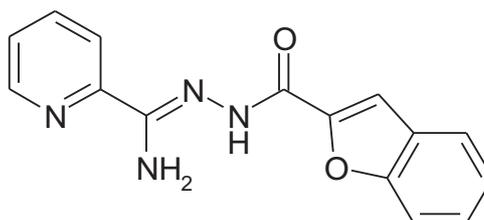
Yield = 0.975g or 4.239mmol or 64%

Melting Point: 185.3 – 187.8 °C.

<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 6.8 (s, 2H, NH<sub>2</sub>); 6.9 (d, 1H, J=3.478 Hz, furyl-H); 7.5 (ddd, 1H, Py-5H); 7.8 (d, 1H, furyl-H); 7.9 (ddd, 1H, Py-4H); 8.1 (d, 1H, J= 8.056 Hz, Py-3H); 8.3 (d, 1H, furyl-H); 8.6 (d, 1H, J= 4.42 Hz, Py-6H); 10.2 (s, H, NH) ppm.

MS (+ APCI): 213, 214 (M- NH<sub>2</sub>), 231 (M + H)<sup>+</sup>;

IR (KBr) ν = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

**6.10.8. Preparation of N<sup>1</sup>-2-benzofuroyl-pyridine-2-carboxamidrazone: Compound 5.45.**

Yield = .136g or 4.857mmol or 63.47%

Melting Point: 215.6 - 220.3°C.

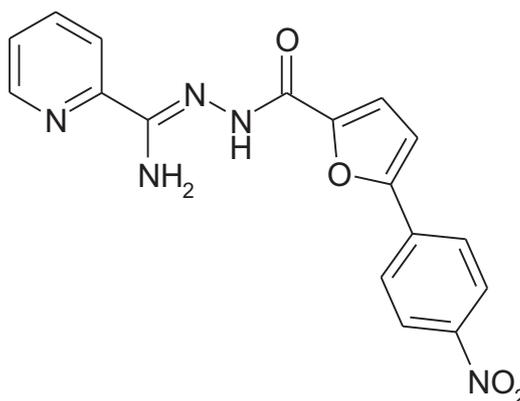
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 7.1 (s, 2H, NH<sub>2</sub>); 6.6 (dd, 1H, furyl-H); 7.4 (dd, 1H, Ph-H); 7.5 (ddd, 1H, Ph-H); 7.5 (ddd, 1H, Py-5H); 7.7 (s, 1H, furyl-H); 7.7 (d, 1H, Ph-H); 7.8 (d, 1H, Ph-H); 7.9 (ddd, 1H, Py-4H); 8.2 (d, 1H, J = 8.056 Hz, Py-3H); 8.6 (d, 1H, J = 4.42 Hz, Py-6H); 10.2 (s, 2H, NH<sub>2</sub>) ppm.

MS (+ electro spray): 281 (M + H)<sup>+</sup>;

Accurate MS (+ES): Calculated (M + H)<sup>+</sup>281.1039; found (M + H)<sup>+</sup> 281.1007; 11.4 ppm

IR (KBr)  $\nu$  = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.58.

**6.10.9. Preparation of N<sup>1</sup>-2-[5-(4-Nitrophenyl)furoyl]-pyridine-2-carboxamidrazone: Compound 5.40.**

Yield = 0.111g or 0.318mmol or 20%

Melting Point: 233.7 – 236.2°C.

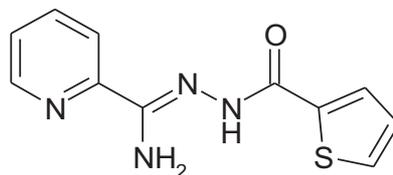
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 7.1 (s, 2H, NH<sub>2</sub>); 7.4 (d, 1H, furyl-H); 7.5 (d, 1H, Ph-H); 7.5 (ddd, 1H, Py-5H); 7.9 (ddd, 1H, Py-4H); 8.2 (d, 1H, J = 8.056 Hz, Py-3H); 8.2 (d, 2H, Ph-H); 8.4 (d, 2H, Ph-H); 8.6 (d, 1H, J = 4.42 Hz, Py-6H); 10.3 (s, 2H, NH) ppm.

MS (- APCI): 350 (M - H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3400 (NH<sub>2</sub>), 3310 (NH<sub>2</sub>), 3150, 2380 (C=O), 1680, 1620, 1550, 1540, 1500, 1470, 1440, 1390, 1225 cm<sup>-1</sup>

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = tailing on base line.

**6.10.10. Preparation of N<sup>1</sup>-(Thiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.51.**



Yield 0.370 g, 1.50 mmol, 39%.

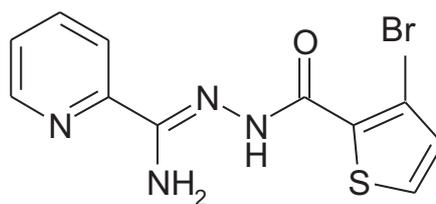
Melting Point: 214.9 - 217.8 °C

<sup>1</sup>H-NMR: 6.87 (bs, 2H, NH<sub>2</sub>), 7.19 (t, 1H, *J*=4.2 Hz, thiophenyl – H4), 7.50 (q, 1H, *J*=5.1 Hz, Py – H5), 7.78 - 8.31 (overlapping m, 4H, Py – H3, H4 and thiophenyl – H3, H5), 8.61(d, 1H, *J*=4.4 Hz, Py – H6), 10.33 (bs, 1H, CONH); MS (+ES) *m/z* = 269 (M + Na)<sup>+</sup>, 247 (M + H)<sup>+</sup>, 230 (M – NH<sub>2</sub>);

IR (KBr)  $\nu$  = 3413 (NH<sub>2</sub>), 3321(CONH), 3284, 3105, 3072 (Ar-C-H), 3027, 2848, 1654 (CO), 1617 (C=N), 1544 (Ar-C=C), 1475 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.38.

**6.10.11. Preparation of N<sup>1</sup>-(3-Bromothiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.54.**



Yield 0.230 g, 0.708 mmol, 48%.

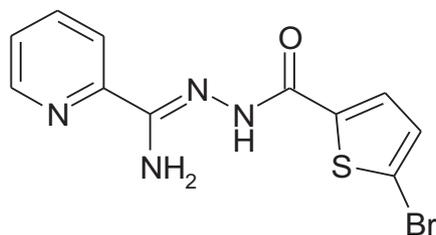
Melting Point: 183.6 – 186.4 °C

<sup>1</sup>H-NMR: 6.90 (bs, 2H, NH<sub>2</sub>), 7.20 (q, 1H, *J*=5.1 Hz, thiophenyl – H4 or H5), 7.48 (t, 1H, *J*=5.7 Hz, Py – H5), 7.75 (d, 1H, *J*=4.4 Hz, thiophenyl – H4 or H5), 7.90 (t, 1H, *J*=7.6 Hz, Py – H4), 8.13 (t, 1H, *J*=8.2 Hz, Py – H3), 8.60 (d, 1H, *J*=4.4 Hz, Py – H6), 10.40 (bs, 1H, CONH) ppm; MS (+ES) *m/z* = 349 (M + Na)<sup>+</sup>, 325 & 327 (M + H)<sup>+</sup>, 310 (M – NH<sub>2</sub>);

IR (KBr)  $\nu$  = 3439 (NH<sub>2</sub>), 3353 (CONH), 3233, 3169, 3013 (Ar-C-H), 2917, 1668 (CO), 1628 (C=N), 1585(Ar-C=C), 1566, 1492 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.52.

**6.10.12. Preparation of N<sup>1</sup>-(5-Bromothiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.57.**



Yield 0.091 g, 0.280 mmol, 39%.

Melting Point: 215.5 – 218.1 °C

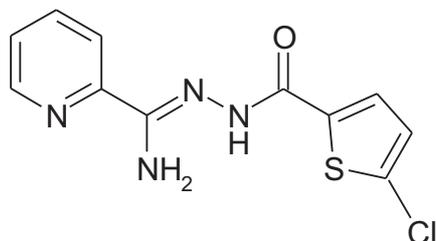
<sup>1</sup>H-NMR: 6.88 (m, 2H, NH<sub>2</sub>), 7.33 (d, 1H, *J*=4.4 Hz, thiophenyl – H<sub>4</sub>), 7.51 (t, 1H, *J*=5.1 Hz, Py – H<sub>5</sub>), 7.81 (d, 1H, *J*=4.4 Hz, thiophenyl – H<sub>3</sub>), 8.02 (t, 1H, *J*=7.6 Hz, Py – H<sub>4</sub>), 8.20 (m, 1H, Py – H<sub>3</sub>), 8.63 (d, 1H, *J*=4.4 Hz, Py – H<sub>6</sub>), 10.45 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 325 & 327 (*M* + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3431 (NH<sub>2</sub>), 3353 (CONH), 3319, 3254, 3168, 3005 (Ar-C-H), 2906, 1685 (CO), 1636 (C=N), 1583 (Ar-C=C), 1564, 1488 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.38.

**6.10.13. Preparation of N<sup>1</sup>-(5-Chlorothiophen-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.60.**



Yield = 0.198 g, 0.706 mmol, 46%.

Melting Point: 235.0 – 236.2 °C

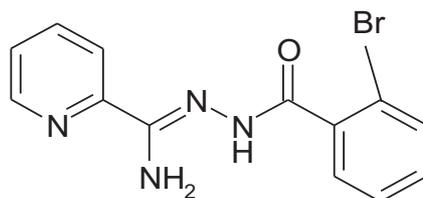
<sup>1</sup>H-NMR: 6.97 (bs, 2H, NH<sub>2</sub>), 7.25 (d, 1H, *J*=4.4 Hz, thiophenyl – H<sub>4</sub>), 7.51 (t, 1H, *J*=7.0 Hz, Py – H<sub>5</sub>), 7.86 (d, 1H, *J*=4.4 Hz, thiophenyl – H<sub>3</sub>), 8.02 (t, 1H, *J*=8.2 Hz, Py – H<sub>4</sub>), 8.21 (d, 1H, *J*=7.6 Hz, Py – H<sub>3</sub>), 8.63 (d, 1H, *J*=4.4 Hz, Py – H<sub>6</sub>), 10.50 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 281 & 283 (*M* + H)<sup>+</sup>, 231 (*M* – [NH<sub>2</sub> + Cl]);

IR (KBr)  $\nu$  = 3419 (NH<sub>2</sub>), 3382 (CONH), 3213, 3008 (Ar-C-H), 2908, 1688 (CO), 1638 (C=N), 1565 (Ar-C=C), 1432 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.32.

**6.10.14. Preparation of N<sup>1</sup>-(2-bromobenzoyl)-pyridine-2-carboxamidrazone: Compound 5.71.**



Yield 0.070 g, 0.219 mmol, 14%.

Melting Point: 206.5 – 209.4 °C

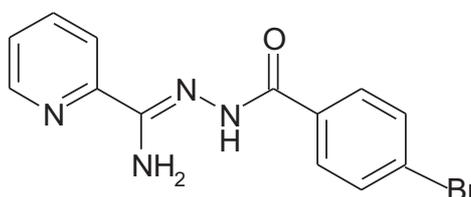
<sup>1</sup>H-NMR: 6.78 (bs, 2H, NH<sub>2</sub>), 7.37 – 7.53 (overlapping m, 4H, Ph – H<sub>3</sub>, H<sub>4</sub>, H<sub>5</sub> and Py – H<sub>5</sub>), 7.69 (dd, 1H, *J*=7.6 Hz, 1.3 Hz), 7.91 (ddd, 1H, *J*=7.6, 5.7, 1.9 Hz, Py – H<sub>4</sub>), 8.17 (dd, 1H, *J*=8.2, 1.3 Hz, Py – H<sub>3</sub>), 8.60 (dd, 1H, *J*=5.1, 1.9 Hz, Py – H<sub>4</sub>), 10.38 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 341 & 343 (M + Na)<sup>+</sup>, 319 & 321 (M + H)<sup>+</sup>, 302 (M – NH<sub>2</sub>);

IR (KBr)  $\nu$  = 3412 (NH<sub>2</sub>), 3339 (CONH), 3202, 3005 (Ar-C-H), 2841, 1664 (CO), 1633 (C=N), 1596 (Ar-C=C), 1577, 1544, 1475 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.42.

**6.10.15. Preparation of N<sup>1</sup>-(4-bromobenzoyl)-pyridine-2-carboxamidrazone: Compound 5.63.**



Yield 0.323 g, 1.013 mmol, 66%.

Melting Point: 209.9 – 212.0 °C;

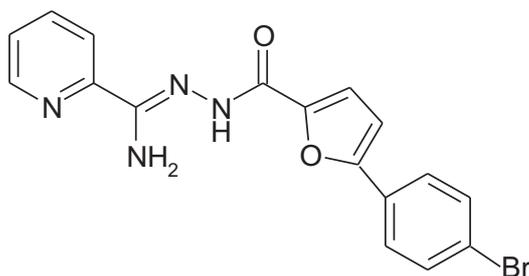
<sup>1</sup>H-NMR: 6.92 (bs, 2H, NH<sub>2</sub>), 7.50 (t, 1H, *J*=6.3 Hz), 7.68 (d, 2H, *J*=8.2 Hz, Ph – H<sub>3</sub>,H<sub>5</sub>), 7.80 – 7.93 (overlapping m, 3H, Ph – H<sub>2</sub>, H<sub>6</sub> and Py – H<sub>4</sub>), 8.17 (d, 1H, *J*=8.2 Hz, Py – H<sub>3</sub>), 8.61 (d, 1H, *J*=4.4 Hz, Py – H<sub>6</sub>), 10.32 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 319 & 321 (M + H)<sup>+</sup>, 302 (M – NH<sub>2</sub>), 207 (M – [NH<sub>2</sub> + O + Br]);

IR (KBr)  $\nu$  = 3425 (NH<sub>2</sub>), 3328 (CONH), 3307, 3188, 3057 (Ar-C-H), 1659 (CO), 1636 (C=N), 1583, 1547, 1469 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.21.

**6.10.16. Preparation of N<sup>1</sup>-(5-[4-bromophenyl]-furan-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.66.**



Yield 0.016 g, 0.042 mmol, 10%.

Melting Point: 209.7 – 212.8 °C;

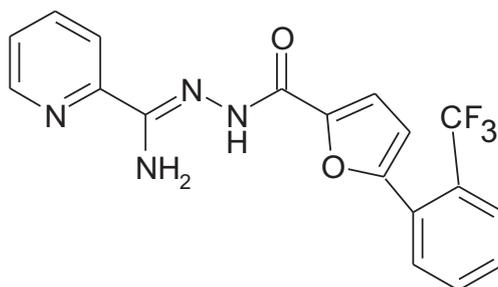
<sup>1</sup>H-NMR: 6.98 (bs, 2H, NH<sub>2</sub>), 7.19 (d, 1H, *J*=3.2 Hz, furanyl – H4), 7.28 (d, 1H, *J*=3.2 Hz, furanyl – H3) 7.51(ddd, 1H, *J*=6.3, 5.7, 1.3 Hz, Py – H5), 7.67 (d, 2H, *J*=8.2 Hz, phenyl – H3,H5), 7.86 – 7.95 (overlapping m, 3H, phenyl – H2,H6 and Py – H4), 8.18 (d, 1H, *J*=8.2 Hz, Py – H3), 8.62 (d, 1H, *J*=4.4 Hz, Py – H6), 10.27 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 407 (M + Na)<sup>+</sup>, 385 & 387 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3420 (NH<sub>2</sub>), 3286 (CONH), 3172, 3044 (Ar-C-H), 2998, 1634 (CO), 1587 (Ar-C-H), 1560, 1534, 1467 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.26.

**6.10.17. Preparation of N<sup>1</sup>-(5-[2-trifluoromethylphenyl]-furan-2-oyl)-pyridine-2-carboxamidrazone: Compound 5.69.**



Yield 0.100 g, 0.267 mmol, 66%.

Melting Point: 92.2 – 94.4 °C;

<sup>1</sup>H-NMR: 6.91 (bs, 2H, NH<sub>2</sub>), 7.32 (bs, 1H, furanyl – H4), 7.48 (ddd, 1H, *J*=6.3, 5.1, 1.3 Hz, Py – H5), 7.66 (t, 1H, *J*=7.6 Hz, phenyl – H4), 7.79 – 7.98 (overlapping m, 5H, furanyl – H3, phenyl – H3,H4,H6 and Py – H4), 8.17 (d, 1H, *J*=7.6 Hz, Py – H3), 8.60 (d, 1H, *J*=4.4 Hz, Py – H6), 10.18 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 397 (M + Na)<sup>+</sup>, 375 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3439 (NH<sub>2</sub>), 3319 (CONH), 3134, 3067 (Ar-C-H), 1635 (CO), 1602 (C=N), 1568 (Ar-C-H), 1546, 1476 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.22.

#### 6.11.1. Preparation of pyridine-4-carboxamidrazone: Compound 5.91.



4-cyanopyridine (30.103g, 288.4mmol) was suspended in ethanol (90mL) and treated with hydrazine monohydrate (60mL) and stirred at ambient temperature for two days. The resulting precipitate was collected by filtration under vacuum and washed with petroleum ether (40-60°C) to give a white crystalline compound.

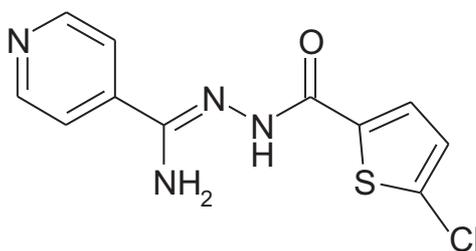
Yield was 33.450g or 244 mmol or 75%

<sup>1</sup>H-NMR: 5.33 (bs, 2H, NH<sub>2</sub>); 5.71 (bs, 2H, NH<sub>2</sub>); 7.63 (d, 2H, J=6.3 Hz, Pyr-H5 and H3); 8.51 (d, 2H, J=6.3 Hz, Pyr-H2 and H6) ppm.

MS (APCI +ve) m/z = 137 (M+H)<sup>+</sup>

TLC, carried out in ethyl acetate: methanol (2:1) (R<sub>f</sub>) = 0.18.

#### 6.11.2. Preparation of N<sup>1</sup>-(5-chlorothiophen-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.59.



5-Chloro-2-thiophenecarboxylic acid (0.283 g, 1.74 mmol) and CDI (0.283 g, 1.75 mmol) were mixed in dry THF (6mL) under argon and stirred for 20 minutes at room temperature. A solution of pyridine-4-carboxamidrazone (0.208 g, 1.53 mmol) in dry THF (4mL) was added dropwise over 5 minutes to the mixture of acid-CDI under argon and stirred for one day at room temperature. After 24 hours, the mixture was added dropwise over 10 minutes to saturated sodium bicarbonate solution (30 mL) with stirring. The solid material was collected

by filtration and washed with very small quantity of distilled water followed by ethyl acetate (10 mL). A pale yellow coloured powder was obtained which was recrystallised in hot methanol. The final product obtained was yellow coloured powder which was dried under vacuum for 24 hours. Yield 0.063 g, 0.225 mmol, 15%.

Melting Point: 216.4 – 218.4 °C;

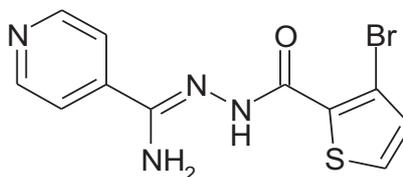
$^1\text{H-NMR}$ : 6.91 (bs, 2H,  $\text{NH}_2$ ), 7.21 (d, 1H,  $J=3.8$  Hz, thiophenyl – H4), 7.78 – 7.85 (overlapping m, 3H, Py – H3, H5 and thiophenyl – H3), 8.71 (d, 2H,  $J=5.7$  Hz, Py – H2, H6), 10.42 (bs, 1H, CONH) ppm;

MS (+ES)  $m/z$  = 303 & 305 ( $\text{M} + \text{Na}$ )<sup>+</sup>, 281 & 283 ( $\text{M} + \text{H}$ )<sup>+</sup>, 264 ( $\text{M} - \text{NH}_2$ );

IR (KBr)  $\nu$  = 3419 ( $\text{NH}_2$ ), 3386 (CONH), 3213, 3014 (Ar-C-H), 2928, 1701 (CO), 1648 (C=N), 1595 (Ar-C=C), 1486  $\text{cm}^{-1}$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.57.

### 6.11.3. Preparation of $\text{N}^1$ -(3-bromothiophen-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.53.



Yield 0.109 g, 0.335 mmol, 23%.

Melting Point: 162.2 – 164.6 °C;

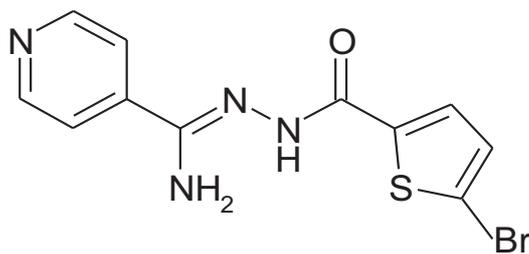
$^1\text{H-NMR}$ : 6.85 (bs, 2H,  $\text{NH}_2$ ), 7.18 (s, 1H, thiophenyl – H4 or H5), 7.72 – 7.92 (overlapping m, 3H, Py – H3, H5 and thiophenyl – H4 or H5), 8.64 (d, 2H,  $J=5,7$  Hz, Py – H2, H6), 10.29 (bs, 1H, CONH) ppm;

MS (+ES)  $m/z$  = 325 & 327 ( $\text{M} + \text{H}$ )<sup>+</sup>;

IR (KBr)  $\nu$  = 3419 ( $\text{NH}_2$ ), 3373 (CONH), 3107, 3028 (Ar-C-H), 2921, 1705 (CO), 1618 (N=C), 1595 (Ar-C=C), 1539, 1493  $\text{cm}^{-1}$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.35.

6.11.4. **Preparation of N<sup>1</sup>-(5-bromothiophen-2-oyl)-pyridine-4-carboxamidrazone:  
Compound 5.56.**



Yield 0.018 g, 0.055 mmol, 8%.

Melting Point: 205.7 – 208.2 °C

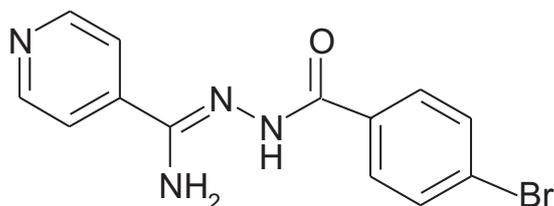
<sup>1</sup>H NMR (250 MHz d<sub>6</sub>-DMSO): 6.8 (s, 2H, NH<sub>2</sub>), 7.33 (d, 1H, *J*=4.4 Hz, thiophenyl – H4), 7.8 (overlapping m, 3H, Py – H3, H5 and thiophenyl – H4 or H5), 8.7 (d, 2H, *J*=5,7 Hz, Py – H2, H6), 10.5 (bs, 1H, NH) ppm;

MS (+ES) *m/z* = 347 (M + Na)<sup>+</sup>, 325 & 327 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3419 (NH<sub>2</sub>), 3339 (CONH), 3227, 2994(Ar-C-H), 2908, 1705 (CO), 1645 (C=N), 1599 (Ar-C=C), 1544, 1489 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.45 (single spot).

6.11.5. **Preparation of N<sup>1</sup>-(4-bromobenzoyl)-pyridine-4-carboxamidrazone:  
Compound 5.62.**



Yield 0.181 g, 0.567 mmol, 37%.

Melting Point: 269.2 – 270.1 °C;

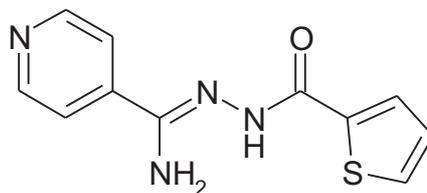
<sup>1</sup>H-NMR: 6.94 (bs, 2H, NH<sub>2</sub>), 7.67 (d, 2H, *J*=8.2 Hz, Ph – H3,H5), 7.78 (d, 2H, *J*=5.7 Hz, Py – H3,H5), 7.86 (d, 2H, *J*=8.2 Hz, Ph – H2,H6), 8.64 (d, 2H, *J*=6.3 Hz, Py – H2,H6), 10.25 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 341 & 343 (M + Na)<sup>+</sup>, 319 & 321 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3412 (NH<sub>2</sub>), 3359 (CONH), 3207, 3040 (Ar-C-H), 2826, 1649 (CO), 1592 (Ar-C=C), 1536, 1479 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.27.

**6.11.6. Preparation of N<sup>1</sup>-(thiophen-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.50.**



Yield 0.056 g, 0.228 mmol, 6%.

Melting Point: 194.8 – 196.9 °C;

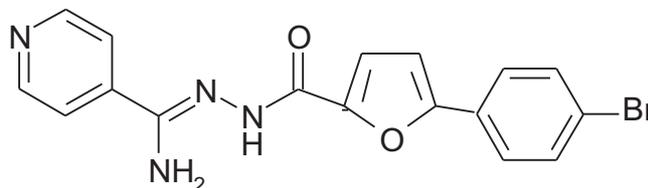
<sup>1</sup>H-NMR: 6.89 (bs, 2H, NH<sub>2</sub>), 7.16 (bs, 1H, thiophenyl – H4), 7.75 – 7.87 (overlapping m, 3H, Py – H3,H5 and thiophenyl – H5), 8.03 (bs, 1H, thiophenyl – H3), 10.25 (bs, 1H, CONH) ppm;

MS (+ES) m/z = 269 (M + Na)<sup>+</sup>, 247 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3346 (NH<sub>2</sub>), 3160, 3120, 3061 (Ar-C-H), 2994, 2829, 1665 (CO), 1625 (C=N), 1582 (Ar-C=C), 1536 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.36.

**6.11.7. Preparation of N<sup>1</sup>-(5-[4-bromophenyl]-furan-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.65.**



Yield 0.093 g, 0.242 mmol, 43%.

Melting Point: 206.3 – 208.7 °C;

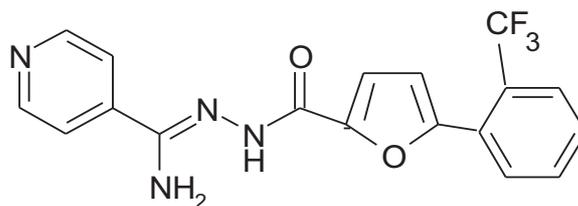
<sup>1</sup>H-NMR: 6.85 (bs, 2H, NH<sub>2</sub>), 7.15 – 7.21 (overlapping m, furanyl – H3,H4), 7.67 (d, 2H, J=8.2 Hz, phenyl – H3,H5), 7.80 – 7.86 (overlapping m, 4H, phenyl – H2, H6 and Py – H3,H5), 8.63 (d, 2H, J=5.7 Hz, Py – H2,H6), 10.26 (bs, 1H, CONH) ppm;

MS (+ES) m/z = 409 (M + Na)<sup>+</sup>, 385 & 387 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3372 (NH<sub>2</sub>), 3228 (CONH), 3123, 3037 (Ar-C-H), 2919, 1672 (CO), 1643 (N=H), 1580 (Ar-C-H), 1541 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.30.

6.11.8. **Preparation of N<sup>1</sup>-(5-[2-trifluoromethylphenyl]-furan-2-oyl)-pyridine-4-carboxamidrazone: Compound 5.68.**



Yield 0.025 g, 0.067 mmol, 16%.

Melting Point: 193.4 – 194.8 °C;

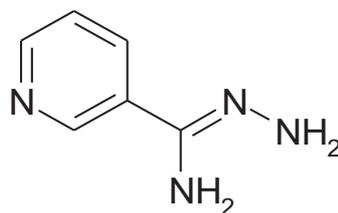
<sup>1</sup>H-NMR: 6.91 (bs, 2H, NH<sub>2</sub>), 7.31 (bs, 1H, furanyl – H4), 7.63 – 7.99 (overlapping m, 7H, furanyl – H3, phenyl – H3,H4,H5,H6 and Py – H3,H5), 8.64 (d, 2H, *J*=5.7 Hz, Py – H2,H6), 10.18 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 397 (M + Na)<sup>+</sup>, 375 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3400 (NH<sub>2</sub>), 3205 (CONH), 3071, 3024 (Ar-C-H), 1667 (CO), 1631 (C=N), 1600 (Ar-C-H), 1567, 1544 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.30.

6.12.1. **Preparation of pyridine-3-carboxamidrazone: Compound 5.90.**



3- cyanopyridine (30.103g, 288.4mmol) was suspended in ethanol (90mL) and treated with hydrazine monohydrate (60mL) and stirred at ambient temperature for two days. The resulting precipitate was collected by filtration under vacuum and washed with petroleum ether (40-60°C) to give a white crystalline compound.

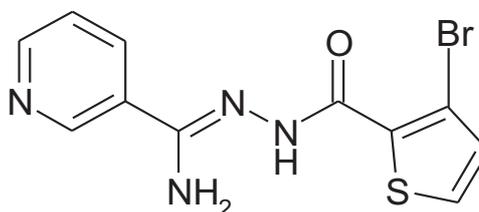
Yield was 30.214g or 222.160 mmol or 77%

<sup>1</sup>H-NMR: 4.26 (bs, 2H, NH<sub>2</sub>); 5.73 (bs, 2H, NH<sub>2</sub>); 7.34 (m, 1H, Pyr-H5); 8.01 (dt, 1H, *J*= 2.0, 3.7 Hz, Pyr-H4); 8.48 (d, 1H, *J*= 4.7Hz, Pyr-H6); 8.86 (d, H, *J*=2.1 Hz, Pyr-H2) ppm.

MS (APCI +ve) *m/z* = 137 (M+H)<sup>+</sup>

TLC, carried out in ethyl acetate: methanol (2:1) (*R<sub>f</sub>*) = 0.10 (single spot).

6.12.2. **Preparation of N<sup>1</sup>-(3-bromothiophen-2-oyl)-pyridine-3-carboxamidrazone:  
Compound 5.55.**



3-Bromo-2-thiophenecarboxylic acid (0.345 g, 1.67 mmol) and CDI (0.278 g, 1.72 mmol) were mixed in dry THF (6mL) under argon and stirred for 30 minutes at room temperature. A solution of pyridine-3-carboxamidrazone (0.208 g, 1.53 mmol) in dry THF (4mL) was added dropwise over 5 minutes to the mixture of acid-CDI under argon and stirred for two days at room temperature. After 48 hours, the mixture was added dropwise over 10 minutes to saturated sodium bicarbonate solution (30 mL) with stirring. The solid material was collected by filtration and washed with very small quantity of distilled water followed by ethyl acetate (10 mL). A pale yellow coloured powder was obtained which was dried under vacuum for 24 hours.

Yield 0.099 g, 0.305 mmol, 20%.

Melting Point: 169.2 – 171.2 °C;

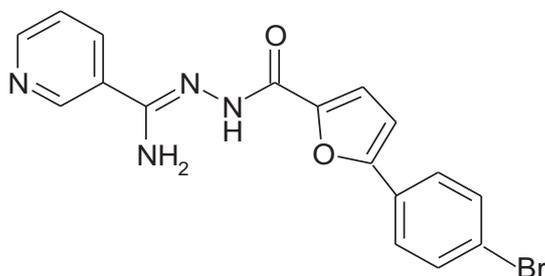
<sup>1</sup>H-NMR: 6.85 (bs, 2H, NH<sub>2</sub>), 7.03 (d, 1H, *J*=5.1 Hz, thiophenyl – H4 or H5), 7.34 – 7.48 (overlapping m, 2H, Py – H4 and thiophenyl – H4 or H5), 8.14 (d, 1H, *J*=7.6 Hz, Py – H6), 8.55 (d, 1H, *J*=3.8 Hz, Py – H5), 8.98 (bs, 1H, Py – H2), 10.24 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 347 & 349 (M + Na)<sup>+</sup>, 325 & 327 (M + H)<sup>+</sup> 244 (M – Br)<sup>+</sup>;

IR (KBr)  $\nu$  = 3376 (NH<sub>2</sub>), 3326 (CONH), 3160, 3045 (Ar-C-H), 2949, 1688 (CO), 1648 (N=C), 1618, 1580 (Ar-C=C), 1539, 1484 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.28.

6.12.3. Preparation of N<sup>1</sup>-(5-[4-bromophenyl]-furan-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.67.



Yield 0.220 g, 0.571 mmol, 96%.

Melting Point: >350 °C.

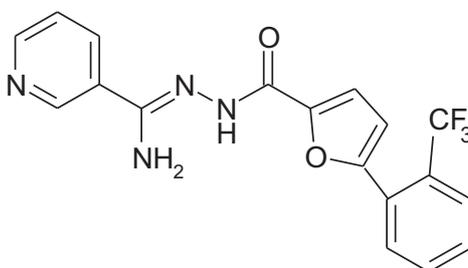
<sup>1</sup>H-NMR: 6.91 (bs, 2H, NH<sub>2</sub>), 7.17 (d, 1H, *J*=3.2 Hz, furanyl – H4), 7.22 (bs, 1H, furanyl – H3), 7.47 (q, 1H, *J*=7.6 Hz, Py – H4), 7.68 (d, 2H, *J*=8.2 Hz, phenyl – H3,H5), 7.88 (d, 2H, *J*=8.2 Hz, phenyl – H2,H6), 8.19 (d, 1H, *J*=7.6 Hz, Py – H6), 8.63 (d, 1H, *J*=5.1 Hz, Py – H5), 9.01 (d, 1H, *J*=1.9 Hz, Py – H2), 10.18 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 407 (M + Na)<sup>+</sup>, 385 & 387 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3381 (NH<sub>2</sub>), 3204 (CONH), 3049 (Ar-C-H), 2912, 1693 (CO), 1657 (N=H), 1606 (Ar-C-H), 1541 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (*R<sub>f</sub>*) = 0.23.

6.12.4. Preparation of N<sup>1</sup>-(5-[2-trifluoromethylphenyl]-furan-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.70.



Yield 0.160 g, 0.428 mmol, 74%.

Melting Point: >350 °C;

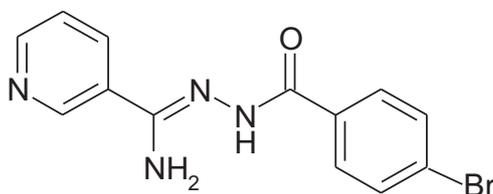
$^1\text{H-NMR}$ : 6.9 (bs, 2H,  $\text{NH}_2$ ), 7.31 (bs, 1H, furanyl – H4), 7.47 (q, 1H,  $J=7.6$  Hz, Py – H4), 7.63 – 7.99 (overlapping m, 5H, furanyl – H3, phenyl – H3,H4,H5,H6), 8.19 (d, 1H,  $J=7.6$  Hz, Py – H6), 8.63 (d, 1H,  $J=5.1$  Hz, Py – H5), 9.01 (d, 1H,  $J=1.9$  Hz, Py – H2), 10.18 (bs, 1H, NH) ppm;

MS (+ES)  $m/z = 397$  ( $\text{M} + \text{Na}$ )<sup>+</sup>, 375 ( $\text{M} + \text{H}$ )<sup>+</sup>;

IR (KBr)  $\nu = 3385$  ( $\text{NH}_2$ ), 3279 (CONH), 3135, 2924, 1695 (CO), 1652 (C=N), 1603 (Ar-C-H), 1511, 1456  $\text{cm}^{-1}$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.24.

#### 6.12.5. Preparation of N<sup>1</sup>-(4-bromobenzoyl)-pyridine-3-carboxamidrazone: Compound 5.64.



Yield 0.387 g, 1.21 mmol, 76%.

Melting Point: 198.9 – 201.2 °C;

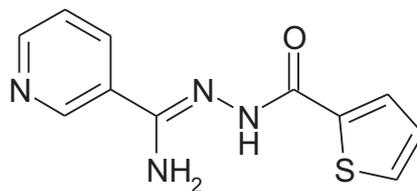
$^1\text{H-NMR}$ : 6.88 (bs, 2H,  $\text{NH}_2$ ), 7.47 (q, 1H,  $J=8.2$  Hz, Py – H4), 7.66 (d, 2H,  $J=8.2$  Hz, Ph – H3,H5), 7.86 (d, 2H,  $J=8.2$  Hz, Ph – H2,H6), 8.18 (d, 1H,  $J=7.6$  Hz, Py – H6), 8.63 (d, 1H,  $J=5.1$  Hz, Py – H5), 9.01 (d, 1H,  $J=1.9$  Hz, Py – H2), 10.24 (bs, 1H, CONH) ppm;

MS (+ES)  $m/z = 341$  & 343 ( $\text{M} + \text{Na}$ )<sup>+</sup>, 319 & 321 ( $\text{M} + \text{H}$ )<sup>+</sup>, 302 ( $\text{M} - \text{NH}_2$ );

IR (KBr)  $\nu = 3381$  ( $\text{NH}_2$ ), 3325 (CONH), 3141, 3040 (Ar-C-H), 2990, 1689 (CO), 1659 (C=N), 1606 (Ar-C=C), 1544, 1479  $\text{cm}^{-1}$ .

TLC: ethyl acetate: methanol (4:1) ( $R_f$ ) = 0.19.

6.12.6. Preparation of **N<sup>1</sup>-(thiophen-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.52.**



Yield 0.070 g, 0.285 mmol, 19%.

Melting Point: 191.5– 194.0°C;

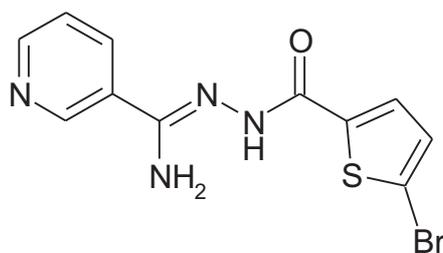
<sup>1</sup>H-NMR: 6.85 (bs, 2H, NH<sub>2</sub>), 7.17 (t, 1H, *J*=5.1 Hz, thiophenyl – H<sub>4</sub>), 7.50 (q, 1H, *J*=7.6 Hz, Py – H<sub>4</sub>), 7.88 (d, 1H, *J*=3.8 Hz, thiophenyl – H<sub>5</sub>), 8.01 (d, 1H, *J*=2.5 Hz, thiophenyl – H<sub>3</sub>), 8.20 (d, 1H, *J*=8.2 Hz, Py – H<sub>6</sub>), 8.64 (dd, 1H, *J*=5.1, 4.4 Hz, Py – H<sub>5</sub>), 9.05 (bs, 1H, Py – H<sub>2</sub>), 10.24 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 269 (M + Na)<sup>+</sup>, 247 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3439 (NH<sub>2</sub>), 3379 (CONH), 3153, 3001 (Ar-C-H), 2822, 1685 (CO), 1652 (C=N), 1546 (Ar-C=C) cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.27.

6.12.7. Preparation of **N<sup>1</sup>-(5-bromothiophen-2-oyl)-pyridine-3-carboxamidrazone: Compound 5.58.**



Yield 0.149 g, 0.458 mmol, 42%.

Melting Point: 190.7 – 193.4 °C;

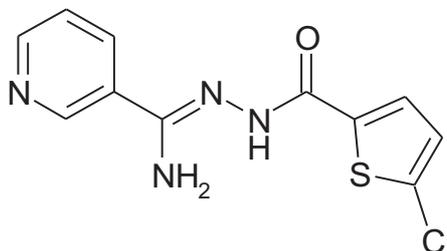
<sup>1</sup>H-NMR: 6.91 (bs, 2H, NH<sub>2</sub>), 7.30 (d, 1H, *J*=4.4 Hz, thiophenyl – H<sub>4</sub>), 7.48 (q, 1H, *J*=8.2 Hz, Py – H<sub>4</sub>), 7.78 (d, 1H, *J*=4.4 Hz, thiophenyl – H<sub>3</sub>), 8.18 (d, 1H, *J*=8.2 Hz, Py – H<sub>6</sub>), 8.66 (dd, 1H, *J*=5.1, 4.4 Hz, Py – H<sub>5</sub>), 9.04 (d, 1H, *J*=1.9 Hz, Py – H<sub>2</sub>), 10.45 (bs, 1H, CONH) ppm;

MS (+ES) *m/z* = 347 & 349 (M + Na)<sup>+</sup>, 325 & 327 (M + H)<sup>+</sup>;

IR (KBr)  $\nu$  = 3418 (NH<sub>2</sub>), 3373 (CONH), 3215, 3031(Ar-C-H), 2926, 1705 (CO), 1636 (C=N), 1521 (Ar-C=C), 1544, 1478 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.38

6.12.8. Preparation of N<sup>1</sup>-(5-chlorothiophen-2-oyl)-pyridine-3-carboxamidrazone:  
Compound 5.61.



Yield 0.094 g, 0.335 mmol, 22%.

Melting Point: 190.9 – 193.7 °C;

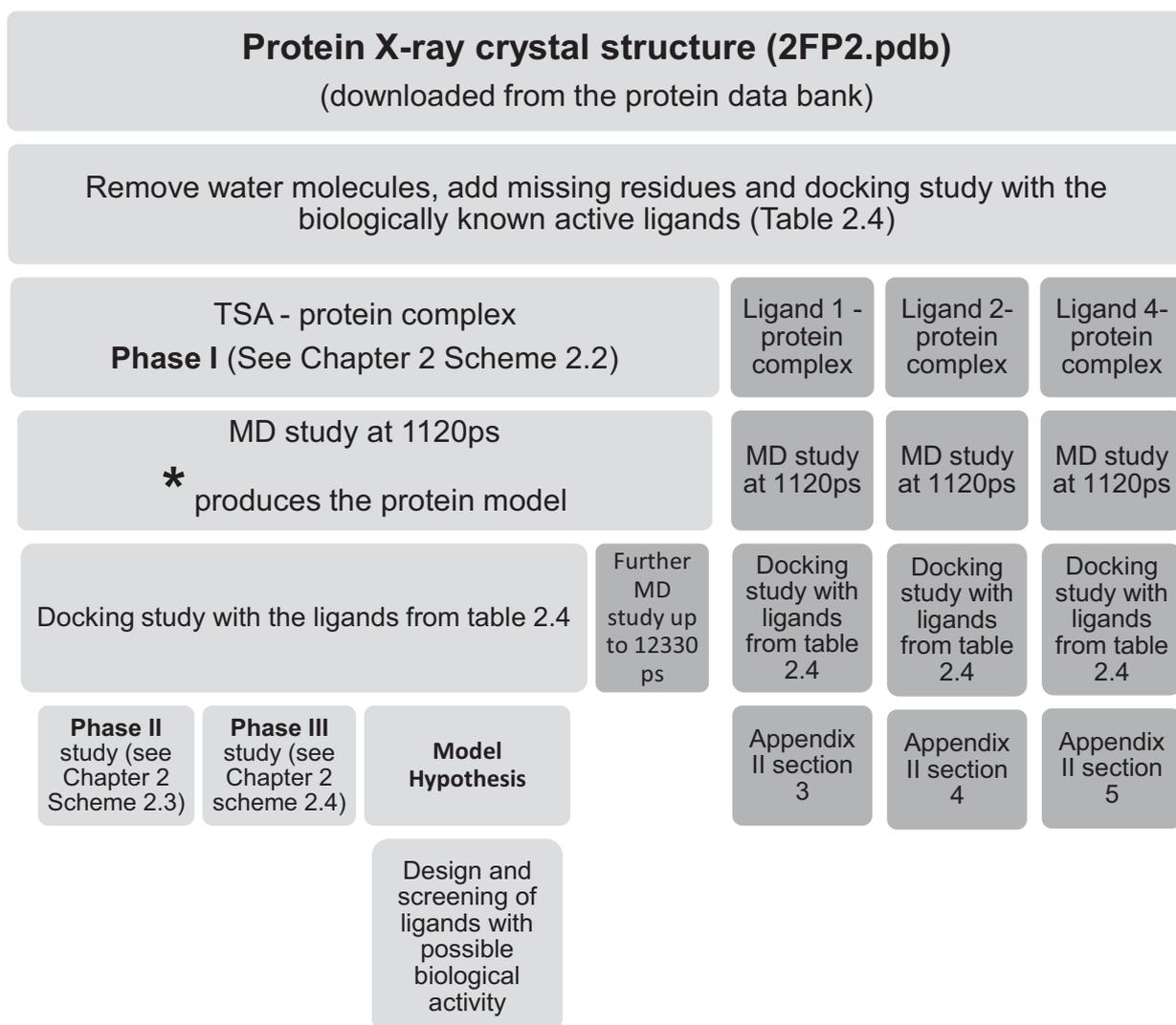
<sup>1</sup>H-NMR: 6.92 (bs, 2H, NH<sub>2</sub>), 7.21 (d, 1H, *J*=4.4 Hz, thiophenyl – H4), 7.53 (q, 1H, *J*=8.2 Hz, Py – H4), 7.84 (d, 1H, *J*=4.4 Hz, thiophenyl – H3), 8.18 (d, 1H, *J*=7.6 Hz, Py – H6), 8.66 (dd, 1H, *J*=5.1, 4.4 Hz, Py – H5), 9.04 (d, 1H, *J*=1.9 Hz, Py – H2), 10.44 (bs, 1H, CONH) ppm

MS (+ES) *m/z* = 303 (M + Na)<sup>+</sup>, 281 & 283 (M + H)<sup>+</sup>, 264 (M – NH<sub>2</sub>);

IR (KBr)  $\nu$  = 3421 (NH<sub>2</sub>), 3386 (CONH), 3227, 2994 (Ar-C-H), 2848, 1691 (CO), 1658 (C=N), 1625 (Ar-C=C), 1546 cm<sup>-1</sup>.

TLC: ethyl acetate: methanol (4:1) (R<sub>f</sub>) = 0.45 (single spot).

## APPENDIX I



■ The grey regions illustrate the steps involved in the development of the targeted protein model which was used in the screening of the ligands with possible biological activity.

\* Indicates the particular step which produces the target model.

■ The dark regions indicate the additional steps acquired during the model generation steps, however were not required to prepare the targeted protein model.

## APPENDIX II

1. Table below showing the docking score when the ligands were docked into the active site of the protein 2FP2.pdb without performing any MD study

Ligand	Docking Score Kcal/mol	Docking Score Kcal/mol	Docking Score Kcal/mol	Average Score Kcal/mol
Ligand 1	-325.467	-335.467	-312.45	-324.46
Ligand 2	-310.43	-330.23	-315.56	-318.74
Ligand 4	-294.397	-285.72	-245.55	-275.22
TSA	-410.74	-403.74	-412.87	-409.12

2. Table below showing the docking score obtained following the MD study over 220ps on the docked Ligand1-protein complex obtained from the Scheme 2.2.

Ligand	Score 1 Kcal/mol	Score 2 Kcal/mol	Score 3 Kcal/mol	Average docking score Kcal/mol
Ligand 1	-575.24	-543.27	-575.13	-564.55
Ligand 2	-246.06	-243.22	-249.37	-246.22
Ligand 4	-400.30	-236.18	-385.72	-340.73
Ligand 7	-245.88	-324.68	-241.90	-270.82
Ligand 8	-592.50	-351.24	-527.20	-490.31
Ligand 9	-667.56	-635.33	-601.40	-634.76
Ligand 13	-613.879	-596.421	-639.341	-616.55
TSA	-402.26	-408.49	-406.30	-405.68

3. Table below showing the docking score obtained following the MD study over 1120ps on the docked Ligand1-protein complex obtained from the Scheme 2.2.

Ligand	Docking Score 1 Kcal/mol	Docking score 2 Kcal/mol	Docking score 3 Kcal/mol	Average docking score Kcal/mol
Ligand 1	-481.15	-496.2	-478.36	-485.24
Ligand 2	-421.58	-430.79	-421.33	-424.57
Ligand 4	-340.82	-346.65	-337.82	-341.76
Ligand 7	-406.17	-407.62	-396.33	-403.37
Ligand 8	-556.78	-372.47	-477.36	-468.87
Ligand 9	-646.44	-639.15	-639.25	-641.61

4. Table below showing the docking score obtained following the MD study over 1120ps on the docked Ligand2-protein complex obtained from the Scheme 2.2.

Ligand	Docking Score 1 Kcal/mol	Docking score 2 Kcal/mol	Docking score 3 Kcal/mol	Average docking score Kcal/mol
Ligand 1	-326.39	-336.138	-332.26	-331.60
Ligand 2	-308.46	-358.504	-413.162	-360.04
Ligand 4	-245.1	-241.064	-239.304	-241.82
Ligand 7	-265.427	-269.406	-249.752	-261.53
Ligand 8	-478.15	-313.195	-305.804	-365.72
Ligand 9	-437.063	-302.06	-277.112	-338.75
Ligand 13	-411.596	-452.935	-372.057	-412.20
TSA	-306.45	-310.49	-303.30	-306.75

5. Table below showing the docking score obtained following the MD study over 1120ps on the docked Ligand4-protein complex obtained from the Scheme 2.2.

Ligand	Docking Score 1 Kcal/mol	Docking score 2 Kcal/mol	Docking score 3 Kcal/mol	Average docking score Kcal/mol
Ligand 1	-355.43	-346.18	-332.26	-344.62
Ligand 2	-320.61	-338.41	-413.162	-357.39
Ligand 4	-305.16	-252.04	-279.34	-278.85
Ligand 7	-365.47	-269.406	-249.752	-294.88
Ligand 8	-518.15	-513.152	-505.41	-512.24
Ligand 9	-491.03	-392.07	-485.28	-456.13
Ligand 13	-434.26	-462.935	-392.67	-429.96
TSA	-341.45	-270.19	-289.53	-300.39

**REFERENCES:**

Adams. R, Ulich. (1920). The use of oxalyl chloride and bromide for producing acid chlorides, acid bromides or acid anhydrides III. J. Am. Chem. Soc., 42 (3), 599–611

Agrawal. H, Kumar. A, *et al.*, (2007). Ligand based virtual screening and biological evaluation of inhibitors of chorismate mutase (Rv1885c) from *Mycobacterium tuberculosis* H37Rv. *Bioorganic & Medicinal Chemistry Letters* 17(11): 3053-3058.

Alam. T, Tarannum. H, *et al.* (2000). Interaction of 2-amino-, 3-amino-, and 4-aminopyridines with chromium and manganese ferrocyanides. *Journal of Colloid and Interface Science* 224(1): 133-139.

Alston. T, Porter. A, Bright. H, (1983) Enzyme inhibition by nitro and nitroso compounds. *Accounts of Chemical Research* 16: 418-424.

Amber 9 (2006) User's manual.

Andrews. P, Cain. E, Rizzardo. E, Smith G, (1977) Rearrangement of chorismate to prephenate. Use of chorismate mutase inhibitors to define the transition state structure. *Biochemistry* 16: 4848-4852.

Bandara, M. B. K., H. Zhu, *et al.* (2006). Salicylic acid reduces the production of several potential virulence factors of *Pseudomonas aeruginosa* associated with microbial keratitis. *Investigative Ophthalmology & Visual Science* 47(10): 4453-4460.

Banerjee. A, Dubnau. E, Quemard. A, Balasubramanian V, Um, K. S., Wilson T., Collins D., de Lisle G., Jacobs W. R. (1994) inhA, a Gene Encoding a Target for Isoniazid and Theionamide in *Mycobacterium tuberculosis*. *Science* 263:227–230.

Bange. F, Brown A, Jacobs. W, (1996) Jr Leucine auxotrophy restricts the growth of *Mycobacterium bovis* BCG in macrophages. *Infect. Immun.* 64, 1794–1799

Barlow. R (1951) Inhibitors of flavoprotein enzyme. *Journal of Chem Society.* 2225.

- Bartlett. P. A., Nakagawa, Y., Johnson, C. R., Reich, S. H., Luis, A. (1988) Chorismate mutase inhibitors: synthesis and evaluation of some potential transition-state analogues. *Journal of Organic Chemistry* 53: 3195-3210.
- Becker. O, Mackerell. A, Roux. B, Watanabe. M, (2001) *Computational Biochemistry and Biophysics*. Published by Marcel Dekker, NY.
- Beeman, D. (1976). Multi step methods for use in Molecular dynamics calculations. *J. Computational Physics*. 20: 130-139
- Begum. N, (2007), PhD Thesis, Aston University, Birmingham.
- Bernstein, F, Koetzle T, Williams G, Meyer. E, Brice M, Rogers. J, Kennard O, Shimanouchi. T, Tasumi. M, (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *Journal of Mol. Biol.*, 112, 535-542.
- Bertrand. J, Dobritz. C, Beerens. H, (1956) Chemistry of Amdrazones. *Bull. Soc. Pharm. Lille*, 1.39: Chem. Absr. 51.1168
- Billington, D. et al., (1998) Automated synthesis and antimycobacterial activity of a series of 2-heteroarylcarboxamidrazones and related compounds. *J. Pharm. Pharmacol.* 50, 262
- Billington, D. Coleman, M. et al (1998) Synthesis and antimycobacterial activity of some heterocarboxamidrazone derivatives. *Drug Design and Discovery* 1998, 15, 269-275.
- Billington, D. Lowe, P. Rathbone, D. Schwalbe, C. (2000) A new amidrazone derivative with antimycobacterial activity. *Acta Cryst.* C56, e211-e212
- Bohm, Klebe (1996), What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs?, *Angew. Chem. Int. Ed. Engl* 35 (22) pp. 2566–2587
- Brändén. C, & Jones, T.A. (1990). Between objectivity and subjectivity. *Nature* 343, 687–689
- Brooijmans, N. Kuntz, I. D. (2003). Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* 32, 335–373

CAChe user guide for 6.1 version 2000-2003

Cark, D. E (1996) Evolutionary algorithm in Computer aided molecular design. *J. Comp-aided molecular design*. 10 337-358.

Carugo. O, Cemazar. M, et al. (2003). Vicinal disulfide turns. *Protein Eng* 16: 637–639

Cheung. D (2002) Structures and Properties of Liquid Crystals and Related Molecules from Computer Simulation, PhD thesis. University of Durham.

Coleman, M. D, Rathbone. D, et al. (1999). Preliminary in vitro toxicological evaluation of a series of 2-pyridylcarboxamidrazone candidate anti-tuberculosis compounds. *Environmental Toxicology and Pharmacology* 7(1): 59-65.

Coleman, M. D, Rathbone. D, et al. (2000). Preliminary in vitro toxicological evaluation of a series of 2-pyridylcarboxamidrazone candidate anti-tuberculosis compounds: II. *Environmental Toxicology and Pharmacology* 8(3): 167-172.

Coleman, M. D, Rathbone. D., et al. (2001). Preliminary in vitro toxicological evaluation of a series of 2-pyridylcarboxamidrazone candidate anti-tuberculosis compounds III. *Environmental Toxicology and Pharmacology* 9(3): 99-102.

Coleman, M. D, Rathbone. D, et al. (2003), The use of computational QSAR analysis in the toxicological evaluation of a series of 2-pyridylcarboxamidrazone candidate anti-tuberculosis compounds. *Environ. Tox. Pharmacol.* 14, 33-42.

Collins, L. A. and Franzblau, S. G. (1997). Microplate Alamar Blue Assay versus BACTEC 460 System for High-Throughput Screening of Compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*. *Antimicrob. Agents Chemother.* 41:1004-1009.

Crofton J. (1970). The assessment and treatment of drug-resistance problems in tuberculosis. *Journal of the Irish Medical Association* 1970; 63: 75.

Daeniker. H, and Grob. C, (1973) 3-Quinuclidinone, hydrochloride *Organic Syntheses, Coll. Vol. 5*, p.989

de Araújo-Filho, J. A. , Vasconcelos Jr, A. C., de Sousa, E. M., Kipnis, A., Ribeiro, E., Junqueira-Kipnis, A. P. (2008) Cellular responses to MPT-51, GlcB and ESAT-6 among MDR-TB and active tuberculosis patients in Brazil. *Tuberculosis* 88: 474-481.

de Figueiredo, R. M., R. Frohlich, *et al.* (2006). N,N '-carbonyldiimidazole-mediated cyclization of amino alcohols to substituted azetidines and other N-heterocycles. *Journal of Organic Chemistry* 71(11): 4147-4154.

Dinakaran, M., Senthilkumar, P., Yogeewari, P., Sriram, D. (2009) Antitubercular activities of novel benzothiazolo naphthyridone carboxylic acid derivatives endowed with high activity toward multi-drug resistant tuberculosis. *Biomedicine and Pharmacotherapy* 63: 11-18.

Douglas B, *et al.*, (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Drug discovery*. 3. 935

EMBO reports 6, 1, 46–51 (2005) doi:10.1038/sj.embor.7400317. Published online: 17 December 2004

Englund. E, Gopi. H, Appella. D, (2004). An Efficient Synthesis of a Probe for Protein Function: 2,3-Diaminopropionic Acid with Orthogonal Protecting Groups . *Org. Lett.* 6: 213–215.

Ercolessi, F. (1997), A molecular dynamics primer. Spring College in Computational Physics, ICTP, Trieste.

Ernst JD, Trevejo-Nuñez G, Banaiee N (2007). Genomics and the evolution, pathogenesis, and diagnosis of tuberculosis . *J. Clin. Invest.* 117 (7): 1738–45.

Feit. P, Bruun. H, Nielsen. C, (1970) Aminobenzoic acid diuretics. 1. 4-Halogeno-5-sulfamylmetanilic acid derivatives *J. Med. Chem.*, 13 (6), pp 1071–1075

Feynman. R, (1939) Forces in Molecules. *Phys. Rev.* 56, 340

Frieden, T.R. Sterling, S.S. Munsiff, C.J. Watt and C. Dye, (2003), Tuberculosis, *Lancet* 362, pp. 887–899

Gershon. H, Armegiani. R, (1962) Antimicrobial Activity of 8-Quinolins, Salicylic Acids, Hydroxynaphthoic Acids, and Salts of Selected Quinolins with Selected Hydroxy-Acids. *Appl Environ Microbiol* 10(4) 348-353.

Ginsburg, A. S., J. H. Grosset, and W. R. Bishai. (2003). Fluoroquinolones, tuberculosis, and resistance. *Lancet Infect. Dis.* 3:432-442.

Gokhale, N., S. Padhye, *et al.* (2001). The crystal structure of first copper(II) complex of a pyridine-2-carboxamidrazone - a potential antitumor agent. *Inorganic Chemistry Communications* 4(1): 26-29.

Goldberg, D. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Longman publishing, San Mateo, CA.

Gottlieb. E, Kotlyar. V, Nudelman. A , (1997) NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities *Hugo J. Org. Chem.* 62, 7512-7515

Grange JM. (1990). Drug-resistance and tuberculosis elimination. *Bulletin International Union Against Tuberculosis and Lung Diseases* 65, 57.

Hagler, A., Euler, E., and Lifson, S. (1974). Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals *J. Am. Chem. Soc.* 96, 5319–5327.

Hagler, A. and Lifson, S. (1974). Energy functions for peptides and proteins. II. Amide hydrogen bond and calculation of amide crystal properties *J. Am. Chem. Soc.* 96, 5327–5335.

Halgren, T. A. (1992). The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *J. Am. Chem. Soc.* 114, 7827-7843.

Halgren, T. A. (1996a). Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94, *Journal. Computation. Chem.* 17, 490–519.

Halgren. T, (1996b) Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions. *J. Comput. Chem.* 17, 520–552.

Halgren, T. A. (1996c). Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94. *J. Comput. Chem.* 17, 553–586.

Halgren, T. A. (1996d). Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data, and Empirical Rules. *J. Comput. Chem.* 17, 616–641.

Halgren, T. A, *et al.*, (1996). Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94, *J. Comp. Chem.* 5 & 6 587-615

Haslam, E, (1993) *Shikimic Acid: Metabolism and Metabolites*, Wiley, New York.

Hediger, M. E. (2004) Design, synthesis, and evaluation of aza inhibitors of chorismate mutase. *Bioorganic and Medicinal Chemistry* 12: 4995–5010.

Hegedus, L, Holden, M, (1990). *cis-N-tosyl-3-Methyl-2-azabicyclo[3.3.0]oct-3-ene*. *Organic Synthesis, Coll. Vol. 7*, p.501.

Higashi, M, Hayashi, S, Kato, S, (2007) Transition state determination of enzyme reaction on freeenergy surface: Application to chorismate mutase *Chemical Physics Letters* 437 293–297

Hockney, R. W, (1970). "The potential calculation and some applications," in *Methods in Computational Physics*. Vol. 9, 135-211.

Holland, J. (1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI.

Husain, A, Galopin, C. C., Zhang, S., Pohnert, G., Ganem, B. (1999) *S-(-)-Dinitrobiphenic acid: A selective inhibitor of Escherichia coli chorismate mutase based on prephenate mimicry*. *Journal of American Chemical Society* 121: 2647-2648.

Ignatenko, M. A. and V. K. Promonenkov (1978). *Biological-Activity of Halogeno-Pyridines and Their Derivatives - Literature Survey*. *Khimiko-Farmatsevticheskii Zhurnal* 12(8): 20-25.

- Isin, E. M., M. de Jonge, *et al.* (2001). Studies on synthetic approaches to 1H-and 2H-indazolyl derivatives. Abstracts of Papers of the American Chemical Society 222: U84-U84.
- Jackman, G.B., Petrow, V., Stephenson, O. and Wild, A.M. (1962) 'Diuretic agents. VI. Some sulphamoylbenzoic acids'; *J Pharm Pharmacol* , 14, 679-686z
- Kameda. T, Zhao, C. Ashida, J. *et al.* (2003) Determination of distance of intra-molecular hydrogen bonding in (Ala–Gly)<sub>15</sub> with silk I form after removal of the effect of MAS frequency in REDOR experiment. *Journal of Magnetic resonance* Vol 160, Issue 2, Pages 91-96
- Kansal, V. K., K. Kar, *et al.* (1979). Diuretic Agents - Synthesis of 1, 2-Disubstituted 7-Sulphamoylbenzimidazole-5-Carboxylic Acids. *Indian Journal of Chemistry Section B-Organic Chemistry Including Medicinal Chemistry* 18(1): 88-90.
- Kast. P, Asif-Ullah, M. *et al.* (1996). Exploring the active site of chorismate mutase by combinatorial mutagenesis and selection: the importance of electrostatic catalysis. *Proc Natl Acad Sci USA*, Vol. 93, No. 10 (May 14, 1996), pp. 5043-5048
- Kenny B. Lipkowitz *et al.*, (1993) *Ab Initio Calculations on Large Molecules: Methodology and Applications*, Jerzy Cioslowski, in *Reviews in Computational Chemistry*, VCH Publishers, 1993, pages 1-33
- Khodarahmi. G, Chem. P, Hakimelahi. G, *et al.* (2005). Synthesis and Primary Cytotoxic Screening of Some 3- Sulfonamide Substituted Benzamido-benzimidazolones, *Iranian Journal of Pharmaceutical Research* 1: 43-56
- King. J, (1999) Russian jailhouse genomics. Enlisting a genome to fight MDR tuberculosis. *Modern Drug Discov.* Jan/Feb, 32-45
- Kleywegt. G, & Jones, T.A. (1995). Where freedom is given, liberties are taken. *Structure* 3, 535–540.
- Kollman, P. A. (1993) Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* 93, 2395–2417.
- Konstantinova, L. S., O. A. Rakitin, *et al.* (2001). New reactions of Hunig's base with S<sub>2</sub>Cl<sub>2</sub>: formation of monocyclic 1,2-dithioles. *Mendeleev Communications*(5): 165-166.

Kulkarni, S. Zou, M. *et al.* (2009) Structure-Activity Relationships Comparing N-(6-Methylpyridin-yl)-Substituted Aryl Amides to 2-Methyl-6-(substituted-arylethynyl)pyridines or 2-Methyl-4-(substituted-arylethynyl)thiazoles as Novel Metabotropic Glutamate Receptor Subtype 5 Antagonists *J. Med. Chem.* Vol 52 pp 3563-3575.

Laskowski, R (2001). Summary and analysis of PDB structure. *Nucleic acids Res.* 29, 221-222.

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

Leach. A, (2001) *Molecular modeling principles and application*. Published by Pearson Education EMA

Leach. A.; Harren Jhoti (2007). *Structure-based Drug Discovery*. Springer; 1 edition

Lifson, S. and Warshel, A. (1969). Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and *n*-Alkane Molecules *J. Chem. Phys.* 49, 5116–5129.

Lipinski. C, Lombardo. F, Dominy. B, (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*. Vol 46. 3-26

List of drugs currently controlled under the misuse of drugs legislation, UK Home office, 03/09/09. [www.homeoffice.gov.uk/documents/cdlist.pdf](http://www.homeoffice.gov.uk/documents/cdlist.pdf)

Lee, A. Y., J. D. Stewart, *et al.* (1995). New Insight into the Catalytic Mechanism of Chorismate Mutases from Structural Studies. *Chemistry & Biology* 2(4): 195-203.

Lee. K, Lee. C, Na. J, Kim. J, (2005) Alkynylation of N-tosylimines with aryl acetylenes promoted by ZnBr<sub>2</sub> and N,N-diisopropylethylamine in acetonitrile *Tetrahedron Letters*, Volume 46, Issue 1, 3 January 2005, Pages 69-74

Ma. D. W. W., W. G. Wu, *et al.* (2004). Tetrahydroisoquinoline based sulfonamide hydroxamates as potent matrix metalloproteinase inhibitors. *Bioorganic & Medicinal Chemistry Letters* 14(1): 47-50.

Madanbabu. M, (2004) PhD thesis submitted to the University of Cambridge

Madsen, U, Krosggaard-Larsen, Povl; Liljefors, Tommy (2002). *Textbook of Drug Design and Discovery*. Washington, DC: Taylor & Francis

Mamalo. M. G., Vio, L., Banfi, E., Predominato, M., Fabris, C., Asaro, F. (1993) Synthesis and antimycobacterial activity of some 2-pyridinecarboxamidrazone derivatives. *Farmaco* 1992, 47 (7), 1055-1066.

Mamalo. M. G., Vio, L., Banfi, E., Predominato, M., Fabris, C., Asaro, F. (1993) Synthesis and antimycobacterial activity of some 4-pyridinecarboxamidrazone derivatives. *Farmaco*, 48 (4), 529-538.

Mamalo, M. G., Vio, L., Banfi, E. (1996) Synthesis and antimycobacterial activity of some 2-pyridinecarboxamidrazone derivatives. *Farmaco*, 51 (1), 65-70.

Mamalo. M, Falagiani. V., *et al.*, (2001) Synthesis and antimycobacterial activity of [5-(pyridin-2-yl)-1,3,4-thiadiazol-2-ylthio]acetic acid arylidene-hydrazide derivatives II *Farmaco* Vol 56 (2001) 587–592

Mandal, A. and D. Hilvert (2003). Charge optimization increases the potency and selectivity of a chorismate mutase inhibitor. *Journal of the American Chemical Society* 125(19): 5598-5599.

Masaki, H, Y. Mizuno, *et al.* (2003). Structure-activity relationship of benzo[b]thiophene-2-sulfonamide derivatives as novel human chymase inhibitors. *Bioorganic & Medicinal Chemistry Letters* 13(22): 4085-4088.

Matyk. J, Waisser. K, Drazkova. K, Kunes. J, Klimesova. V, (2005) Heterocyclic isosteres of antimycobacterial salicylanilides. *Farmaco* Vol 60, 399-408.

Momany, F. A., McGuire, R. F., Burgess, A. W., and Scheraga, H. A. (1975). Energy parameters in polypeptides VII, geometric parameters, partial charges, non-bonded

interactions, hydrogen bond interactions and intrinsic torsional potentials for naturally occurring amino acids. *J. Phys. Chem.* Vol 79, 2361–2381.

Morris A. L, MacArthur M. W, Hutchinson E. G & Thornton J. M (1992). Stereochemical quality of protein structure coordinates. *Proteins*, Vol 12, 345-364

Muegge, I, *et al.*, (2000). A knowledge-based scoring function for protein-ligand interactions: Probing the reference state. *Perspectives in Drug Discovery and Design*. Vol 20, Number 1.

Muegge. I., Martin. Y. C, (1999) A General and Fast Scoring Function for Protein–Ligand Interactions: A Simplified Potential Approach *J. Med. Chem.* Vol 42, 791.

National Health Services, (2010) [www.nhs.uk/conditions/tuberculosis](http://www.nhs.uk/conditions/tuberculosis) 25-Mar 10:45 am

Némethy. G., Pottle. M. S., and Scheraga, H. A. (1983). Energy parameters in polypeptides, 9. updating of geometrical parameters, non-bonding interactions and hydrogen bonding interactions for naturally occurring amino acids. *J. Phys. Chem* Vol 87, 1883–1887.

Niketic, S. R. and Rasmussen, K. (1977). *The Consistent Force Field: A Documentation*. Springer-Verlag, New York.

Nissen. S, Namuth. D (2005). Inhibitors of aromatic amino acid biosynthesis. *Croptechology* 2005. (Obtained from <http://croptechology.unl.edu/viewLesson.cgi?LessonID=959117477>) 16-Jul-2010 12:45pm

Okvist, M., R. Dey, *et al.* (2006). 1.6 A crystal structure of the secreted chorismate mutase from *Mycobacterium tuberculosis*: Novel fold topology revealed. *Journal of Molecular Biology* 357(5): 1483-1499.

Oliviero. C, (2003). RMSD versus resolution. *J. Appl. Cryst.* 2003. Vol 36, 125-128

Osman. F. (2000). *Pharmacophore Perception, Development, and use in Drug Design*. La Jolla, 537pp. Calif: International University Line.

Özden. S, Atabey. D, Yıldız. S, Göker. H, (2005) Synthesis and potent antimicrobial activity of some novel methyl or ethyl 1H-benzimidazole-5-carboxylates derivatives carrying amide or amidine groups *Bioorganic & Medicinal Chemistry*, Vol 13 (5) 1587-1597

Palumbo, A., A. Napolitano, *et al.* (2002). Nitrocatechols versus nitro catecholamines as novel competitive inhibitors of neuronal nitric oxide synthase: Lack of the aminoethyl side chain determines loss of tetrahydrobiopterin-antagonizing properties. *Bioorganic & Medicinal Chemistry Letters* 12(1): 13-16.

Pearce-Duvet. J (2006). The origin of human pathogens: evaluating the role of agriculture and domestic animals in the evolution of human disease . *Biol Rev Camb Philos Soc* 81 (3): 369–82.

Ponder. J, Case. D., (2003) Force fields for protein simulations. *Advances in protein chemistry*, 66, 27-85.

Pope. B, Yamamoto. Y, *et al.* (1988), Di-tert-butyl dicarbonate. 1988. *Organic Syntheses, Coll. Vol. 6*, p.418.

Pople. J, Lennard-Jones. J, (1950) The Molecular Orbital Theory of Chemical Valency: IV. The Significance of Equivalent Orbitals, *Proc. Roy. Soc. A* 202 166

Qamra, R., Prakash. P, *et al.* (2005). Crystallization and preliminary X-ray crystallographic studies of Mycobacterium tuberculosis chorismate mutase. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* 61: 473-475.

Ramachandran. G. N, Ramakrishnan. C, Sasisekharan. V., (1963) Stereochemistry of polypeptide chain configurations. *Journal of Mol Biol Vol* 7 95-99.

Ramage. G, Trappe. G (1952) Tetrahydroquinoxaline. A new route from o-amino-N-2-hydroxyethylanilines. *Journal of Chem Soc.* 4406.

Rathbone. D. L., Parker, *et al.* (2006). Discovery of a potent phenolic N-1-benzylidene-pyridinecarboxamidrazone selective against Gram-positive bacteria. *Bioorganic & Medicinal Chemistry Letters* 16(4): 879-883.

Richon. A, (1994). An Introduction to Molecular Modeling. *Mathematech*, 1, 83

Rothschild B, Martin L, Lev G, Bercovier H, Bar-Gal G, Greenblatt C, Donoghue H, Spigelman M, Brittain D (2001). Mycobacterium tuberculosis complex DNA from an extinct bison dated 17,000 years before the present. Clin Infect Dis 33 (3): 305–11.

Russell. D, (2001) Mycobacterium tuberculosis: here today, and here tomorrow Nature Reviews Molecular Cell Biology 2, 569-586.

Sandbhor , U. Padhye, P. Billington, D. Rathbone, D. Franzblau, S. Anson, C. Powell, A. (2002) Metal complexes of carboxamidrazone analogs as antitubercular agents 1. Synthesis, X-ray crystal-structures, spectroscopic properties and antimycobacterial activity against Mycobacterium tuberculosis H37Rv Journal of Inorganic Biochemistry 90 (2002) 127–136

Sasso, S., C. Ramakrishnan, *et al.* (2005). Characterization of the secreted chorismate mutase from the pathogen Mycobacterium tuberculosis. Febs Journal 272(2): 375-389.

Schnappauf, G., N. Strater, *et al.* (1997). A glutamate residue in the catalytic center of the yeast chorismate mutase restricts enzyme activity to acidic conditions. Proceedings of the National Academy of Sciences of the United States of America 94(16): 8491-8496.

Sharma D, *et al.*, (2007). Mutational analysis of S12 protein and implications for the accuracy of decoding by the ribosome . Journal of Molecular Biology 374 (4): 1065–1076

Shendage. D, Fröhlich. R, Haufe. G, (2004). Highly Efficient Stereoconservative Amidation and Deamidation of  $\alpha$ -Amino Acids . Org. Lett. 6: 3675–3678.

Shimao T. (1987) Drug resistance in tuberculosis control. Tubercle; 68(suppl):5-15.

Simonson, T., Archontis, G. & Karplus, (2002) M. Free energy simulations come of age: protein–ligand recognition. Acc Chem. Res.35, 430–437.

Stark, G. R.; Bartlett, P. (1984). Design and use of potent, specific enzyme inhibitors Pharmacol. Therap. 23, 45-78.

Suarez J, Ranguelova K, Jarzecki AA, *et al.* (2009). An oxyferrous heme/protein-based radical intermediate is catalytically competent in the catalase reaction of Mycobacterium

tuberculosis catalase-peroxidase (KatG). *The Journal of Biological Chemistry* 284 (11): 7017–29

Swope, W.C., *et al.* (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.

Szmant. H, Harmuth. C. (1959) Hydrazones of nitrophenyl and pyridyl aldehydes and ketones. New preparation of o-nitrobenzophenone differences in behaviour of o-nitrobenzophenone and o-nitrophenyl mesityl ketone. *Journal of American chemical society* 81 (4), 962-966.

Tanaka, A., Y. Motoyama, *et al.* (1994). Studies on Antiplatelet Agents .4. A Series of 2-Substituted 4,5-Bis(4-Methoxyphenyl)Pyrimidines as Novel Antiplatelet Agents. *Chemical & Pharmaceutical Bulletin* 42(9): 1828-1834.

Tims. K, (2002), PhD Thesis, Aston University, Birmingham

Tommaso. A., *et al.*, (2002) The antituberculosis drug ethionamide is activated by a flavoprotein monooxygenase. *Journal of biological chemistry* 277, 12824-12829

Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. (2003) Improved protein–ligand docking using GOLD. *Proteins* 52, 609–623.

Verlet. L, *et al.*, (1967), Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* 159, 98

Verma, C. K. (2001) Molecular mechanical field force. *Biochemistry* 218

Vlaskina, A. V. and V. P. Perevalov (2004). New synthesis of substituted benzimidazoles by reduction of 4-alkylamino-3,5-dinitrobenzoic acids esters using SnCl<sub>2</sub>. *Khimiya Geterotsiklicheskikh Soedinenii* (4): 620-622.

Waisser. K, Drazkova. K, Kunes. J, Klimesova. V, Kaustova. J, (2004) Antimycobacterial N-pyridinylsalicylamides, isosteres of salicylamides. *Farmaco*, 59, 615-625.

- Wang, R, *et al.*, (1998) SCORE: A new empirical method for estimating the binding affinity of a Protein-Ligand Complex. *J.Mol.Model.*, 4 (12), 379-394.
- Weidner-Wells, M. A., Ohemeng, K. A., *et al.* (2001). Amidino benzimidazole inhibitors of bacterial two-component systems. *Bioorganic & Medicinal Chemistry Letters* 11(12): 1545-1548.
- Weiner, P. K. and Kollman, P. A. (1981). AMBER: Assisted Model Building with Energy Refinement. A general program for modeling molecules and their interactions. *J. Comput. Chem.* 2, 287–303.
- Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106, 765–784.
- Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986). An All Atom Force Field for Simulations of Proteins and Nucleic Acids *J. Comput. Chem.* 7, 230–252.
- Williams, D. E. (1988). Representation of the molecular electrostatic potential by atomic multipole and bond dipole models. *J. Comput. Chem.* 9, 745-763.
- Williams, R, Sinclair, P, DeMong, D, *et al* (2003), 4-Morpholinecarboxylic acid, 6-oxo-2,3-diphenyl-, 1,1-dimethylethyl ester, (2S,3R)-, *Org. Synth.* 80: 18,
- Wolfenden, R, (1976) Transition-state analogues: *Annu. Rev Biophys.Bioeng.*,5,271-306.
- World Health organisation, (2010) [www.who.int/research/en](http://www.who.int/research/en) 25-Mar-2010 12:15 pm.
- Yang L, Tan CH, *et al.*, (2006), New-generation Amber united-atom force field. *J Phys Chem B.*110 (26):13166-76.
- Zhang, Y, Mitchison D (2003). The curious characteristics of pyrazinamide: a review . *Int. J. Tuberc. Lung Dis.* 7 (1): 6–21
- Zignol, M, *et al.*,(2006) Global incidence of multidrug-resistant tuberculosis, *J Infect Dis* 194 pp. 479–485.

Zimhony O, Cox JS, Welch JT, Vilchèze C, Jacobs WR (2000). Pyrazinamide inhibits the eukaryotic-like fatty acid synthetase I (FASI) of *Mycobacterium tuberculosis* (abstract). *Nature Medicine* 6 (9): 1043–47