

# **The design and synthesis of bacterial RNAP inhibitors**

Andrea Sartini

Submitted in accordance with the requirements for the degree of  
Doctor of Philosophy

The University of Leeds

School of Chemistry

September 2014

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Andrea Sartini to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

© 2014 The University of Leeds and Andrea Sartini

## Acknowledgements

I would like to thank my supervisors Prof Colin Fishwick and Prof Peter Johnson for their constant support and belief in me. I am also very grateful to the University Research Scholarship and to the UK for giving me the economical possibility to achieve the dream of a PhD and for offering me a brighter future.

Special thanks to my biological collaborators Prof Ian Chopra, Dr Alex O'Neill and in particular to Dr Rachel Trowbridge and Liam Sharkey who run all the biological assays.

Thanks to all the past and present members of Fishwick/Johnson group, Marco, Anil, Donna, Ian, Craig, Fraser, Martin, Katie, Ricky and Sarah.

A really big thank you goes to Dr Jayakanth Kankanala, the most brilliant chemist and generous person I ever meet in my life; he was my point of reference and guardian angel during my PhD studies. I will never forget our lunches at the Parkinson building where we used to write organic chemistry mechanisms, design molecules and to exchange advice on organic chemistry preparations on napkins in front of a tasty pasta dish or a sandwich.

The technical support offered by Simon Barrett (NMR), Martin Huscroft (HPLC), Tanya Marinko Covell (mass spectrometry and micro analysis) and Ian Blakely (micro analysis) was invaluable throughout my PhD.

A whole-hearted thanks to all the members of Chemistry Stores, in particular to Francis who smiled to me every single day of my PhD.

I feel me very lucky to have met so many nice people and friends at the school of chemistry. The list is going to be endless and I am going to apologize in advance if I missed someone. A thousand thanks to the Italian chemist community (Valeria, Giorgia, Francesco, Roberta, Carlo), Anglo-German-Italian friends (Phil and Keeran) and the glorious British and international friends (Elghareeb, Reem, the 'four' James, Steven, Chris, Silvia, Irene, Dave, Jeff, Charles, Raj, Panchami, Alun, Dan, Kath, Ludwig, Jayapal.....sorry but I am loosing the count and memory at the moment, please visit the website

<http://www.chem.leeds.ac.uk/people.html> for a comprehensive list with no exceptions!).

My days in Leeds were also blessed by having the possibility to meet the 'Italiani nello Yorkshire' community, a small Italy inside Yorkshire with a lot of fantastic people where a particular big thank goes to Alessandro, Simona, Paola, Massimo, Roberta and Carlo for coping with the occasional stress of my PhD. Thanks also to Francesco, one of my oldest friends who gave his support and blessings from Italy.

Last but not least, my 'holy' wife Tatiana, my parents, grandmother Silvana and my family in law deserve more than a thank you for the huge love, care, affection, motivation and their blessings. There are no proper words in this world to express my love and gratitude; you gave me happiness and made my dreams come true. My heart is full of love for my little niece Arianna who paid me a visit in UK twice a year since she was born.

Finally, I hope to apply my knowledge, experience and qualifications for contributing to public health benefit and disseminating chemistry knowledge.

## Abstract

The increasing occurrence of multi drug resistant bacterial infections has underlined the urgent need for the development of new and more potent antibiotics. Bacterial RNA polymerase (RNAP) is a validated target that allows efficacy and selective toxicity. Unfortunately the efficacy of clinically used RNAP inhibitors (rifamycins) is threatened by the emergence of bacterial resistance. The recent identification of the binding region of myxopyronin (Myx), a natural product antibiotic, coupled with the availability of crystal structural data, offers the possibility of using a ligand- and structure-based design approach for the identification of new drug-like RNA polymerase inhibitors.

Validation studies of different docking algorithms were performed on the available X-ray co-crystal structures of *Thermus thermophilus* RNAP and several ligand- and structure-based virtual screening protocols on the Myx binding region were applied to identify hits for SAR-based exploration.

A combined ligand- and structure-based protocol identified two compounds which showed selective inhibitory activity towards *E. coli* RNAP in the micromolar range of concentration. A SAR expansion programme based upon one of these active molecules was conducted via the synthesis of a chemical library of acylhydrazones in order to identify the factors determining potency and to validate the putative binding mode. Similarity-based virtual screening and docking studies were also applied to explore the close chemical space of the initial hit and to prioritize the synthesis of analogs. Some of the designed compounds showed better inhibitory activity than the initial hit and moreover, one derivative possessed moderate activity towards *S. aureus* SH1000.

A structure-based virtual screening protocol was then conducted on an in-house chemical database, applying a preliminary library filtering approach based on physico-chemical descriptors. The identification of a hit active in the micromolar range of concentration underlined the predictive power of this approach and offered useful ideas for future RNAP inhibitor design.

# Table of Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Table of contents</b>	<b>vi</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Figures</b>	<b>xiii</b>
<b>Abbreviations</b>	<b>xvii</b>
<b>1. Introduction</b>	<b>1</b>
1.1 Antibiotics	1
1.2 Bacterial multi drug resistance	1
1.2.1 Current situation	1
1.2.2 Molecular mechanisms of resistance	2
1.2.3 Antibiotics: timeline, challenges and discovery strategy	3
1.2.4 Antibiotics in the pipeline and the future of antibacterial discovery	5
1.2.5 Requirements for an optimal antibacterial target and the problem of bacterial cell penetration	6
1.3 Bacterial DNA-dependent RNA polymerase	9
1.3.1 Structure	9
1.3.2 Catalytic reaction mechanism	11
1.4 RNAP inhibitors and their mechanism of action	14
1.4.1 Rifamycins	14
1.4.1.1 Chemistry and biological activity	14
1.4.1.2 Mechanism of action	18
1.4.1.3 Structure activity relationships	21
1.4.1.4 Resistance to rifamycins	22
1.4.2 Lipiarmycin	24
1.4.3 GE23077	25
1.4.4 Streptolydigin	26
1.4.5 Sorangicin	27
1.4.6 Microcin J25	28
1.4.7 Capistruin	29
1.4.8 'Switch region' inhibitors	30
1.4.8.1 Myxopyronins	31

1.4.8.2	Corallopyronins	35
1.4.8.3	Ripostatins	36
1.4.8.4	Synthetic inhibitors of the 'switch region'	36
1.4.8.4.1	Pyridyl-Benzamides	36
1.4.8.4.2	Myxopyronin based synthetic hybrids	37
1.4.8.4.3	Squaramides	38
1.4.9	Tagetitoxin	39
1.4.10	Other synthetic RNAP inhibitors	39
1.4.10.1	CBR703 and analogs	39
1.4.10.2	Ureidothiophenes	41
1.4.11	Peptides as protein-protein interaction inhibitors	42
1.4.12	Small molecules as protein-protein interaction inhibitors	42
1.5	Thesis aims and objectives	44
1.6	References	46
<b>2.</b>	<b>Computational Methodologies</b>	<b>54</b>
2.1	Virtual Screening approaches	54
2.2	Docking	54
2.2.1	Challenges in docking	55
2.2.2	Search algorithms	55
2.2.3	Scoring functions	56
2.2.4	Main limitations of scoring functions	56
2.3	Glide (Grid based ligand docking with energetic)	57
2.4	F.R.E.D. (Fast Rigid Exhaustive Docking)	58
2.5	Autodock	59
2.6	eHiTS (electronic high-throughput screening)	60
2.7	GOLD (Genetic Optimization for Ligand Docking)	60
2.8	DOCK Blaster	62
2.9	Ligand-based virtual screening	63
2.9.1	ROCS (Rapid Overlay of Chemical Structures)	64
2.10	References	65
<b>3.</b>	<b>Critical assessment of the available docking software</b>	<b>68</b>
3.1	Protein and ligand preparation protocol	68
3.2	Docking protocol used with Glide	69
3.3	Docking protocol used with eHiTS	70
3.4	Docking protocol used with Autodock	70

3.5	Docking protocol used with F.R.E.D.	71
3.6	Docking protocol used with GOLD	71
3.7	Docking protocol used with DOCK Blaster	72
3.8	Critical evaluation of results	72
3.9	Conclusions	81
3.10	References	83
<b>4.</b>	<b>Design of small molecule libraries as putative bacterial RNAP inhibitors</b>	<b>85</b>
4.1	Reported inhibitors of bacterial RNAP	85
4.2	<i>In silico</i> docking studies of reported inhibitors	86
4.3	General docking protocol	86
4.4	Design of analogs of compound 4.1	88
4.4.1	Design of the library and docking studies	90
4.4.2	Putative binding mode of compound 4.20	96
4.4.3	Putative binding mode of compound 4.30	97
4.5	Synthesis of the sulphonamide-based library	97
4.6	Synthesis of the urea-based library	99
4.7	Synthesis of the thiourea analog	101
4.8	Biological evaluation of small molecule inhibitors	101
4.9	Synthesis and biological evaluation of the sulphonamide hit	103
4.10	Conclusions	104
4.11	References	105
<b>5.</b>	<b>Scaffold-hopping and ligand-based virtual screening to assist the design and synthesis of small molecule inhibitors of the myxopyronin binding region</b>	<b>106</b>
5.1	Physico-Chemical limitations of myxopyronin A	106
5.1.1	Scaffold-hopping strategy	107
5.1.2	Ligand-based virtual screening protocol	111
5.1.3	General docking protocol	112
5.1.4	Results	113
5.2	Synthesis	121
5.3	Biological results	124
5.4	Conclusions	125

5.5	References	126
<b>6.</b>	<b>Critical reassessment of a previously reported computational study and development of a combined ligand and structure-based virtual screening protocol</b>	<b>128</b>
6.1	Limitations of a previously reported computational study	128
6.2	Critical reassessment of the protocol	129
6.2.1	General ligand-based virtual screening protocol	131
6.2.2	General docking protocol	133
6.2.3	vHTS results	134
6.2.4	Biological results	140
6.2.5	Synthesis of the acylhydrazone hit	141
6.2.6	Putative binding mode of the acylhydrazone hit	142
6.2.7	General analogue design strategy	144
6.2.8	Similarity searching on a focused library of acylhydrazones	145
6.2.9	Biological results	149
6.2.10	Design and synthesis of analogues with partial or total removal of benzyloxy groups	151
6.2.11	Biological results	152
6.2.12	SAR analysis	153
6.2.13	Design and synthesis of a new acylhydrazone library with improved solubility	155
6.2.14	Biological results	158
6.2.15	SAR analysis	160
6.2.16	Attempted evaluation of IC <sub>50</sub> values and antibacterial activity for selected compounds	161
6.3	Synthesis of intermediate 6.93	162
6.4	Synthesis of the hydrazides	163
6.5	Synthesis of the acylhydrazone library	163
6.6	Conclusions	164
6.7	References	166
<b>7.</b>	<b>Developing of a novel structure based virtual screening protocol</b>	<b>168</b>
7.1	Virtual screening strategy	168
7.1.1	Database filtering protocol	169

7.1.2	Docking protocol	170
7.1.3	vHTS results	171
7.1.4	Biological evaluation of the vHTS hits	174
7.1.5	Substructure query-based search and docking of structural analogues of compound 7.18	175
7.1.6	Biological results	178
7.1.7	SAR analysis of the purchased analogues	179
7.1.8	Putative binding mode of the identified hit	180
7.2	Conclusions	184
7.3	References	185
<b>8.</b>	<b>Conclusions and future work</b>	<b>186</b>
8.1	References	190
<b>9.</b>	<b>Experimental section</b>	<b>191</b>
9.1	General Procedures and Instrumentation	191
9.2	General experimental methods	193
9.3	References	254
	<b>Appendix I</b>	<b>256</b>
A)	The SYBR green assay	256
B)	Experimental procedure for <i>E. coli</i> RNA Polymerase Assay	257
C)	Selectivity and whole cell assays	258
D)	General experimental procedure for the selectivity assays	258
E)	Bacterial strains	259
F)	References	260
	<b>Appendix II</b>	<b>261</b>
A)	Commercially available esters	261
B)	Synthesized hydrazides	262
C)	Commercially available hydrazides	263
D)	Commercially available aldehydes	264
E)	Library of synthesised hydrazones	266
F)	References	273

## List of Tables

<b>Table 3.1</b> Results for the docking of myxopyronin A <b>1.25</b> inside its native co-crystal complex with PDB id: 3DXJ	73
<b>Table 3.2</b> Results for the docking of 7-desmethylmyxopyronin B <b>1.27</b> inside its native co-crystal complex with PDB id.: 3EQL	77
<b>Table 4.1</b> Overall score values and number of H-bonds of the best ranked compounds selected for synthesis	92
<b>Table 4.2</b> Sulphonamide-based library	98
<b>Table 4.3</b> Urea-based library	100
<b>Table 4.4</b> RNAP percentage inhibition of the synthesised small molecule inhibitors	102
<b>Table 5.1</b> Compounds selected for biological evaluation	113
<b>Table 5.2</b> <i>In vitro</i> RNAP percentage inhibition at 100 $\mu$ M of the purchased and synthesised molecules	124
<b>Table 6.1</b> Virtual hits selected for the <i>in vitro</i> percentage inhibition	134
<b>Table 6.2</b> <i>In vitro</i> RNAP percentage inhibition at 100 $\mu$ M of the virtual hits	140
<b>Table 6.3</b> Specificity assay on the selected hits on Malate Dehydrogenase and Chymotrypsin	141
<b>Table 6.4</b> Selected compounds from the similarity searching	146
<b>Table 6.5</b> Additional synthesized analogues	149
<b>Table 6.6</b> <i>In vitro</i> RNAP percentage inhibition of the synthesised analogues	150
<b>Table 6.7</b> Synthesized analogues with partial or total removal of benzyloxy groups	151
<b>Table 6.8</b> <i>In vitro</i> RNAP percentage inhibition of the synthesised compounds	152
<b>Table 6.9</b> Prediction of ALogP and LogSw values using ALOGPS 2.1 software <sup>9</sup> for some representative compounds	153
<b>Table 6.10</b> Synthesized hydrazone-based analogues with predicted improved solubility	155
<b>Table 6.11</b> Designed acylhydrazone analogues with structural variations on the phenyl ring	157
<b>Table 6.12</b> <i>In vitro</i> RNAP percentage inhibition of the designed acylhydrazone analogues	158

<b>Table 6.13</b> Prediction of ALogP and LogSw values using ALOGPS 2.1 software <sup>9</sup> for some representative compounds	159
<b>Table 6.14</b> MIC determination for selected compounds. <sup>a</sup> MIC values in µg/mL <sup>b</sup> <i>E. coli</i> TolC deficient strain	162
<b>Table 7.1</b> Virtual hits selected for biological evaluation	171
<b>Table 7.2</b> <i>In vitro</i> RNAP percentage inhibition of the vHTS hits	174
<b>Table 7.3</b> Close structural analogues selected for biological evaluation	176
<b>Table 7.4</b> <i>In vitro</i> RNAP percentage inhibition of <b>7.18</b> close structural analogues	178
<b>Table 7.5</b> MIC determination for <b>7.18</b> . <sup>a</sup> MIC values in µg/mL <sup>b</sup> <i>E. coli</i> TolC deficient strain	179
<b>Table 7.6</b> Chemscore value for both enantiomers of <b>7.18</b>	180
<b>Table 1</b> Selected bacterial strains description	259
<b>Table 6.15</b> Commercially available esters <b>6.94—6.102</b>	261
<b>Table 6.16</b> Synthesized hydrazides <b>6.103—6.111</b>	262
<b>Table 6.17</b> Commercially available hydrazides <b>6.112—6.116</b>	263
<b>Table 6.18</b> Commercially available aldehydes <b>6.33</b> and <b>6.117—6.134</b>	264
<b>Table 6.19</b> Library of synthesised hydrazones	266

## List of Figures

<b>Figure 1.1</b> Structures of new classes of antibiotics	6
<b>Figure 1.2</b> Structure of Gram-positive and Gram-negative bacteria	7
<b>Figure 1.3</b> Expression of genetic information	9
<b>Figure 1.4</b> Various subunits of the <i>T. aquaticus</i> RNA polymerase (PDB 1I6V) (see colour code) and the positions of the Cleft, Pore and RNA exit channel	10
<b>Figure 1.5</b> Structural elements of RNAP in the region of the DNA-RNA hybrid (PDB 1I6V)	11
<b>Figure 1.6</b> Intermediates in the transcription cycle. RNAP in grey, DNA in blue, nascent RNA in red, multidomain $\sigma$ factor in purple. Adapted from <i>Mooney et al.</i>	12
<b>Figure 1.7</b> 3-formylrifamycin SV <b>1.10</b> with numbering	16
<b>Figure 1.8</b> Structures of rifampicin <b>1.11</b> and rifabutin <b>1.12</b>	16
<b>Figure 1.9</b> Structure of rifapentine <b>1.13</b>	16
<b>Figure 1.10</b> Structure of rifalazil <b>1.14</b> and rifaximin <b>1.15</b>	17
<b>Figure 1.11</b> Schematic diagram of RNAP $\beta$ subunit interactions with rifampicin indicated in green (PDB 1I6V) <b>1.11</b>	19
<b>Figure 1.12</b> Two-pathway mechanism of rifamycins action. (A) The $\beta$ pathway, induced fit mechanism shown with the cyan lightning, RNA in yellow, DNA template in red and blue. (B) The $\sigma$ pathway, induced fit mechanism shown with the purple lightning, RNA in yellow, DNA template in red and blue	20
<b>Figure 1.13</b> Spiro-piperidyl-rifamicin <b>1.16</b> and spirorifamycin analogues <b>1.17</b>	22
<b>Figure 1.14</b> Lipiarmycin A3 <b>1.18</b> structure	25
<b>Figure 1.15</b> GE23077-A1 <b>1.19</b> and GE23077-B1 <b>1.20</b> structure	26
<b>Figure 1.16</b> Structure of streptolydigin <b>1.21</b>	27
<b>Figure 1.17</b> Structure of sorangicin <b>1.22</b>	28
<b>Figure 1.18</b> Structure and sequence of Microcin J25 <b>1.23</b> (PDB 1PP5)	29
<b>Figure 1.19</b> Capistruin <b>1.24</b> structure and its comparison with Microcin J25 <b>1.23</b> . Aminoacids reported in blue and red are highly mutable, in particular, the red ones are detrimental for the antibiotic activity. Adapted from <i>Kuznedelov et al</i>	30
<b>Figure 1.20</b> Structures of Myxopyronins	31

<b>Figure 1.21a</b> Crystal structure of <i>T. thermophilus</i> RNAP-dMyx B 1.27 complex (PDB 3EQL). dMyx B indicated in green	32
<b>Figure 1.21b</b> Schematic diagram of <i>T. thermophilus</i> RNAP-dMyx B 1.27 complex (PDB 3EQL). H-bonds in red dashed lines, van der Waals interactions in black	33
<b>Figure 1.22</b> Overlay of the co-crystal structure proposed by Mukhopadhyay <i>et al.</i> in cyan (PDB 3DXJ) with the co-crystal structure obtained by Belogurov <i>et al.</i> in green (PDB 3EQL)	34
<b>Figure 1.23</b> Structure of CorA 1.28	35
<b>Figure 1.24</b> Ripostatin A 1.29 and B 1.30 structure	36
<b>Figure 1.25</b> Structure of pyridyl-benzamide 1.31	37
<b>Figure 1.26</b> Structures of holomycin 1.32 and its derivative 1.33	38
<b>Figure 1.27</b> Structure of squaramide 1.34	38
<b>Figure 1.28</b> Tagetitoxin 1.35 structure	39
<b>Figure 1.29</b> CBR703 1.36, CBR9379 1.37 and CBR9393 1.38 structures	40
<b>Figure 1.30</b> Structure of 1.39	41
<b>Figure 1.31</b> Structures of 1.40, 1.41 and 1.42	43
<b>Figure 1.32</b> Structure of 1.43	43
<b>Figure 3.1</b> Overlay between crystallographic myxopyronin A 1.25 (C atoms in green) and docking solution obtained with Autodock (C atoms in yellow)	73
<b>Figure 3.2</b> Overlay between crystallographic myxopyronin A 1.25 (C atoms in green) and docking solution obtained with Dockblaster (C atoms in yellow)	74
<b>Figure 3.3</b> Overlay between crystallographic myxopyronin A 1.25 (C atoms in green) and docking solution obtained with eHiTS (C atoms in yellow)	74
<b>Figure 3.4</b> Overlay between crystallographic myxopyronin A 1.25 (C atoms in green) and docking solution obtained with F.R.E.D (C atoms in yellow)	75
<b>Figure 3.5</b> Overlay between crystallographic myxopyronin A 1.25 (C atoms in green) and docking solution obtained with GLIDE (C atoms in yellow)	75
<b>Figure 3.6</b> Overlay between crystallographic myxopyronin A 1.25 (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Goldscore as a scoring function	76

<b>Figure 3.7</b> Overlay between crystallographic myxopyronin A <b>1.25</b> (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Chemscore as a scoring function	76
<b>Figure 3.8</b> Overlay between crystallographic 7-desmethylmyxopyronin B <b>1.27</b> (C atoms in green) and docking solution obtained with Autodock (C atoms in yellow).	78
<b>Figure 3.9</b> Overlay between crystallographic 7-desmethylmyxopyronin B <b>1.27</b> (C atoms in green) and docking solution obtained with Dockblaster (C atoms in yellow).	78
<b>Figure 3.10</b> Overlay between crystallographic 7-desmethylmyxopyronin B <b>1.27</b> (C atoms in green) and docking solution obtained with eHiTS (C atoms in yellow)	79
<b>Figure 3.11</b> Overlay between crystallographic 7-desmethylmyxopyronin B <b>1.27</b> (C atoms in green) and docking solution obtained with F.R.E.D (C atoms in yellow)	79
<b>Figure 3.12</b> Overlay between crystallographic 7-desmethylmyxopyronin B <b>1.27</b> (C atoms in green) and docking solution obtained with GLIDE (C atoms in yellow)	80
<b>Figure 3.13</b> Overlay between crystallographic 7-desmethylmyxopyronin B <b>1.27</b> (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Goldscore as a scoring function	80
<b>Figure 3.14</b> Overlay between crystallographic 7-desmethylmyxopyronin B <b>1.27</b> (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Chemscore as a scoring function	81
<b>Figure 4.1</b> Structure of <b>4.1</b>	86
<b>Figure 4.2</b> Predicted binding mode of compound <b>4.1</b>	88
<b>Figure 4.3</b> Analogue design strategy for compound <b>4.1</b>	89
<b>Figure 4.4</b> Available starting materials for sulphonamide derivatives	91
<b>Figure 4.5</b> Starting materials for urea and thiourea derivatives	91
<b>Figure 4.6</b> Overlay of compound <b>4.1</b> (green sticks), <b>4.20</b> (purple sticks) and <b>4.30</b> (orange sticks) within myxopyronin binding site	95
<b>Figure 4.7</b> Predicted binding mode of <b>4.20</b>	96
<b>Figure 4.8</b> Predicted binding mode of <b>4.30</b>	97
<b>Figure 5.1</b> Structures of <b>1.25</b> and <b>1.27</b>	106
<b>Figure 5.2</b> Co-crystal structure of <b>1.27</b> (PDB 3EQL)	107
<b>Figure 5.3</b> Hydrogen bond network of <b>1.27</b>	108
<b>Figure 5.4</b> Design strategy based on <b>1.27</b>	109

<b>Figure 5.5</b> Structures of <b>5.1</b> and <b>5.2</b> , original $\alpha$ -pyrone core shown in blue, closure points and modified atoms in red	109
<b>Figure 5.6</b> Structures of <b>5.3</b> and <b>5.4</b> , original $\alpha$ -pyrone core shown in blue, closure points and modified atoms in red	109
<b>Figure 5.7</b> Shape query elaborated in vROCS represented as a grey cloud, structure of 7-desmethylmyxopyronin B <b>1.27</b> structure represented as sticks	110
<b>Figure 5.8</b> Pharmacophore query elaborated in vROCS represented as coloured spheres centred on 7-desmethylmyxopyronin B <b>1.27</b> structure represented as sticks	111
<b>Figure 5.9</b> Predicted binding mode of <b>5.8</b>	118
<b>Figure 5.10</b> Predicted binding mode of <b>5.15</b>	119
<b>Figure 5.11</b> Predicted binding mode of <b>5.18</b>	120
<b>Figure 5.12</b> Predicted binding mode of <b>5.29</b>	121
<b>Figure 6.1</b> Virtual screening protocol	129
<b>Figure 6.2</b> Scaffold hopping strategy	130
<b>Figure 6.3</b> Shape query elaborated in vROCS represented as a grey coloured volume. Structure of 7-desmethylmyxopyronin B <b>1.27</b> represented as sticks	132
<b>Figure 6.4</b> Pharmacophore query elaborated in vROCS represented as coloured spheres centred on 7-desmethylmyxopyronin B <b>1.27</b> structure represented as sticks	132
<b>Figure 6.5</b> Predicted binding mode of <b>6.3</b>	142
<b>Figure 6.6</b> Overlay of predicted binding pose for acylhydrazone <b>6.3</b> represented in green sticks and myxopyronin A <b>1.25</b> in yellow sticks.	143
<b>Figure 6.7</b> Myxopyronin binding site represented as a Connolly surface in yellow and acylhydrazide <b>6.3</b> in green sticks	144
<b>Figure 6.8</b> General analogue design strategy	144
<b>Figure 6.9</b> Predicted binding mode of <b>6.49</b>	154
<b>Figure 6.10</b> Putative binding mode of <b>6.75</b>	161
<b>Figure 7.1</b> Putative binding mode of the ( <i>S</i> ) enantiomer of <b>7.18</b>	181
<b>Figure 7.2</b> Putative binding mode of the ( <i>R</i> ) enantiomer of <b>7.18</b>	182
<b>Figure 7.3</b> Overlay between myxopyronin A <b>1.25</b> in yellow sticks and the ( <i>R</i> ) enantiomer of <b>7.18</b> in green	183
<b>Figure 1</b> The SYBR green assay	256

## Abbreviations

°C	Degrees Celsius
2D	Two dimensional
3D	Three dimensional
Å	Angstrom
AIDS	Acquired Immunodeficiency Syndrome
ALogP	Atomic prediction partition coefficient
AMRI	Albany Molecular Research Inc.
app.	Apparent
Ar	Aryl
ATP	Adenosine triphosphate
BIOS	Biology-Oriented synthesis
Bn	Benzyl
br	Broad
CBR	Cumbre pharmaceuticals
CCDC	Cambridge crystallographic data centre
CM	Cytoplasmic membrane
CorA	Corallopyronin A
COSY	Correlation spectroscopy
CTP	Cytidine triphosphate
δ	Chemical shift
d	Doublet
dd	Double doublet
ddd	Double double doublet
DCM	Dichloromethane
DEPT	Distortionless enhancement by polarization transfer
DMF	Dimethylformamide

DMSO	Dimethyl sulphoxyde
dMyxB	7- desmethylmyxopyronin B
DNA	Deoxyribonucleic acid
DOS	Diversity-Oriented synthesis
dt	Double triplet
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	<i>Exempli gratia</i> ; for example
eHITS	Electronic High Throughput Screening
EI	Electron impact
ELISA	Enzyme-Linked Immuno-Sorbent Assay
ES	Electron spray
Et	Ethyl
<i>et al.</i>	<i>Et alia</i> ; and others
EtOAc	Ethyl acetate
EtOH	Ethanol
eq.	Equivalent
F.R.E.D.	Fast Rigid Exhaustive Docking
FBDD	Fragment-Based Drug Discovery
FDA	Food and Drug Administration
GLIDE	Grid-Based Ligand Docking with Energetics
GOLD	Genetic Optimisation for Ligand Docking
GTP	Guanosine triphosphate
GUI	Graphical user interface
h	Hours
<i>H. influenzae</i>	<i>Haemophilus influenzae</i>
HGT	Horizontal gene transfer
HMBC	Heteronuclear multiple-bond correlation spectroscopy

HMQC	Heteronuclear multiple quantum coherence
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
Hz	Hertz
IC <sub>50</sub>	Half maximal inhibitory concentration
IR	Infra-red
<i>J</i>	Coupling constant
kDa	Kilo Dalton
LogP	Logarithm Partition coefficient
LogSw	Logarithm water solubility
M	molar
Me	methyl
MeOH	Methanol
mg	milligram
mL	millilitre
<i>M. leprae</i>	<i>Mycobacterium leprae</i>
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
M.W.	Molecular weight
MCCB	Medicinal chemistry and chemical biology technology group database
MIC	Minimum inhibitory concentration
m	Multiplet
mmol	Millimole
m.p.	Melting point
mRNA	Messenger RNA
Myx	Myxopyronins
MyxA	Myxopyronin A
MyxB	Myxopyronin B

<i>m/z</i>	Mass to charge ratio
NMR	Nuclear magnetic resonance
<i>Nocardia</i> sp.	<i>Nocardia</i> species
NOESY	Nuclear Overhauser effect spectroscopy
NTP	Nucleoside triphosphate
OM	Outer membrane
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PDB	Protein Data Bank
Ph	Phenyl
pp	Pyrophosphate
Ppm	Parts per million
Ppp	Triphosphate
Pppp	Tetraphosphate
PROPKA	Prediction of protein pKa values
pv.	Per volume
q	Quartet
qt	Quaternary
QSAR	Quantitative structure–activity relationship
quint	Quintet
$R_F$	Retention factor
RCSB	Research collaboratory for structural bioinformatics
r.t.	Room temperature
R.t.	Retention time
Rip	Ripostatin
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RNAP	RNA polymerase

ROCS	Rapid Overlay of Chemical Structures
r-RNA	Ribosomal ribonucleic acid
s	Singlet
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SAR	Structure-activity relationship
SBDD	Structure-based drug design
SBVS	Structure-based virtual screening
sept	Septet
sext	Sextet
t	Triplet
<i>T. aquaticus</i>	<i>Thermus aquaticus</i>
<i>T. thermophilus</i>	<i>Thermus thermophilus</i>
TLC	Thin layer chromatography
TMS	Trimethylsilane
t-RNA	Transfer ribonucleic acid
Um	Umbrelliferone
UTP	Uridine triphosphate
vHTS	Virtual high throughput screening
vROCS	Visual Rapid Overlay of Chemical Structures
$\lambda_{max}$	Maximum absorption wavelength
VS	Virtual screening
ZINC	Zinc Is Not Commercial

## Amino acids and their three and one letter codes

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

# 1. Introduction

## 1.1 Antibiotics

Antibiotics (literally 'against life') are substances produced by microorganisms and can be classified on the basis of their target and additionally on their capability of inducing cell death (bactericidal drugs) or inhibiting cell growth (bacteriostatic drugs).<sup>1</sup> Cell death caused by bactericidal drugs results from the interaction of the drug molecule with a specific target in the microorganism leading to alterations in the biochemistry of the bacterium both at the molecular and ultrastructural level.<sup>1</sup>

Common mechanisms of cell death induced by antibiotics include DNA damage *via* breaking double stranded DNA with inhibitors of topoisomerase II (DNA gyrase),<sup>2</sup> arrest of DNA-dependent RNA synthesis (rifamycins),<sup>3</sup> damage of the cell wall and alteration of its structural integrity (inhibitors of cell wall synthesis),<sup>4</sup> or impairment of cellular energetics, ribosome binding and protein translation using inhibitors of protein synthesis.<sup>1, 5</sup> Recently it has been discovered that the mode of action of all classes of bactericidal antibiotics features drug-induced stress including the production of free radicals which are responsible for oxidative damage and alteration of bacterial metabolism.<sup>1, 6, 7, 8</sup>

Since the discovery of penicillin other more effective antibiotics have been discovered and designed using chemical modification and often *via* an understanding of the interactions of the drugs with their target.<sup>1</sup> The successful clinical use of these new molecules lead to widespread optimism and confidence in the fight of modern medicine against infectious diseases during the 'era of antibiotics' between 1940 and 1970. However since that period, the increasing prevalence of drug-resistant bacteria<sup>9</sup> has underlined the importance of exploring and finding new antibacterials.<sup>1,10</sup>

## 1.2 Bacterial multi drug resistance

### 1.2.1 Current situation

Bacterial multi drug resistance is a global healthcare problem and despite the successful use of antibiotics for decades, bacterial infectious diseases are

the main cause of death worldwide, killing 13-17 million people per year and causing 25% of deaths according to the World Health Organization.<sup>11, 12, 13, 14</sup> The mortality and the socio-economical impact of infectious disease is enormous and the fight against multidrug resistant microorganisms is one of the most challenging global health problems.<sup>11, 12, 13, 14</sup> Paul Ehrlich said 'Drug resistance follows the drug like a faithful shadow'.<sup>15</sup> Resistance has developed within 1-4 years after the clinical introduction of the major classes of antibiotics and thanks to the consistent increase in international trade and travel, multidrug-resistant pathogens are representing a global problem.<sup>15</sup> In addition, about 50% of antibiotics are used for prophylaxis, chemotherapy and growth promotion in animals and this is worsening the problem.<sup>10</sup> Antibiotic resistance in Gram-positive and Gram-negative pathogens is growing rapidly and only six new antibiotics have been approved since 2003.<sup>13</sup> This reflects the challenge of identifying new drug classes and the low level of interest from pharmaceutical and biotechnology companies in antibacterial drug discovery for economic reasons.<sup>11, 12, 13, 14</sup> Among Gram-positive pathogens, *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Mycobacterium tuberculosis* and *Enterococci* constitute a significant public health concern while in Gram-negative pathogens, the opportunistic healthcare infections caused by *Pseudomonas aeruginosa*, *Acinetobacter baumannii*, *Stenotrophomonas maltophilia*, and *Burkholderia cepacia* are of particular concern.<sup>11, 12, 13, 14</sup> There is an urgent need to discover new therapeutic agents and the main strategies are the search for antibiotics with novel mechanisms of action, the identification of new molecular target sites, and the chemical modification of known effective molecules to override the mechanisms of resistance.<sup>11, 12, 13, 14</sup>

### **1.2.2 Molecular mechanisms of resistance**

Antibacterial resistance can be classified as intrinsic, where bacteria are naturally resistant to the antibiotic without any previous exposure, or acquired from environmental organisms by horizontal gene transmission (HGT).<sup>16</sup>

An example of intrinsic resistance is that utilised by *Pseudomonas aeruginosa*, whose low membrane permeability is the main reason for its innate resistance to many antibiotics. Acquired resistance is caused by acquisition of a

genetic element like a plasmid and/or transposon or by a chromosome mutation.<sup>10</sup> Mutations of DNA occurs randomly during its replication at a frequency of  $10^{-9}$ - $10^{-10}$  per base pair and copying errors may lead to the partial or complete deletion of individual genes.<sup>17</sup> Classical experiments have shown that production of pre-existing variants and not the emergence of new mutants occurs under the administration of antibiotics but the presence of bacterial strains which are hypermutable, as a consequence of the inactivation of the proofreading and DNA mismatch-repair system, make the mutation rate 200-fold higher than normal cells.<sup>17</sup> Specific resistance mechanisms include: I) chemical inactivation of the drug molecules by enzymes, II) decreased drug accumulation within the pathogen by decreasing uptake or increasing efflux, III) alteration of binding sites which reduces the affinity for antibiotics, IV) development of alternative metabolic pathways.<sup>15</sup>

### **1.2.3 Antibiotics: timeline, challenges and discovery strategy**

Most of the antibiotics used today were discovered before 1970 during the 'golden age' of discovery which started in 1945 and declined in 1965 where antibiotics were isolated from natural sources like soil *Streptomyces* and fungi.<sup>18</sup> Since 1970 only three new classes of antibiotics reached the market: oxazolidinones (discovered in 1978 and introduced in 2003), lipopeptides (discovered in 1986 and introduced in 2003) and carbapenems (discovered in 1975 and introduced in 1985) although this latter class has close chemical similarity with  $\beta$ -lactams and its novelty is debateable.<sup>18</sup> Most of the advances in antibacterial drug discovery since 1970 came *via* improvements of already known antibiotic classes with chemical analogues bearing increased potency and greater ability to overcome existing resistance, but during the last two decades this strategy has not delivered a sufficient number of antibiotics in particular against Gram-negative bacteria related infections.<sup>18</sup> The reasons for this discovery void are the partial withdrawal of 'Big Pharma' from this area of research which is mainly linked with regulatory and commercial challenges, the concomitant unsuccessful discovery strategy, and the simultaneous rise of multi-drug resistant bacteria.<sup>18</sup>

From an economic point of view, the development of new antibiotics is also facing some considerable challenges. These include difficulties in licensing and relatively low economical income following the launch of molecules used for short-course treatments which attack multiple target species developing resistance in short time scales.<sup>16, 18</sup> In addition, an antibiotic requires the ability to reach multiple body compartments and not to show toxicity for daily dosages which are usually higher than for other pharmaceuticals.<sup>16, 18</sup>

With regard to the discovery strategy, during the 'Golden Age' of discovery of antibiotics, the approach consisted of empirical screening of synthetic chemicals, natural compounds from fermentation broths, and extracts of microorganisms with no regard to mechanism of action, and selectivity was only tested later in assays or on animals.<sup>16</sup> Empirical screening from synthetic chemicals delivered important antibacterials like salvarsan, sulphonamides, nitrofurans, chloramphenicol, and quinolones, while many important antibiotic classes were discovered from natural products isolated from fermentation broths and extracts of microorganisms such as  $\beta$ -lactams, aminoglycosides, ansamycines and tetracyclines.<sup>16</sup> To avoid redundancy in the chemical diversity of the discovered scaffolds, in the early 1960s, a 'de-replication' strategy was attempted which focussed on targets or pathways (cell wall or protein synthesis inhibition) but the output of novel antibiotic classes started decreasing in 1977 and gradually moved into the 'genomic era' in the 1990s which mostly failed to deliver results.<sup>16</sup> This change was necessary because mining natural sources became less productive and did not offer good chemical diversity within the discovered scaffolds.<sup>16</sup>

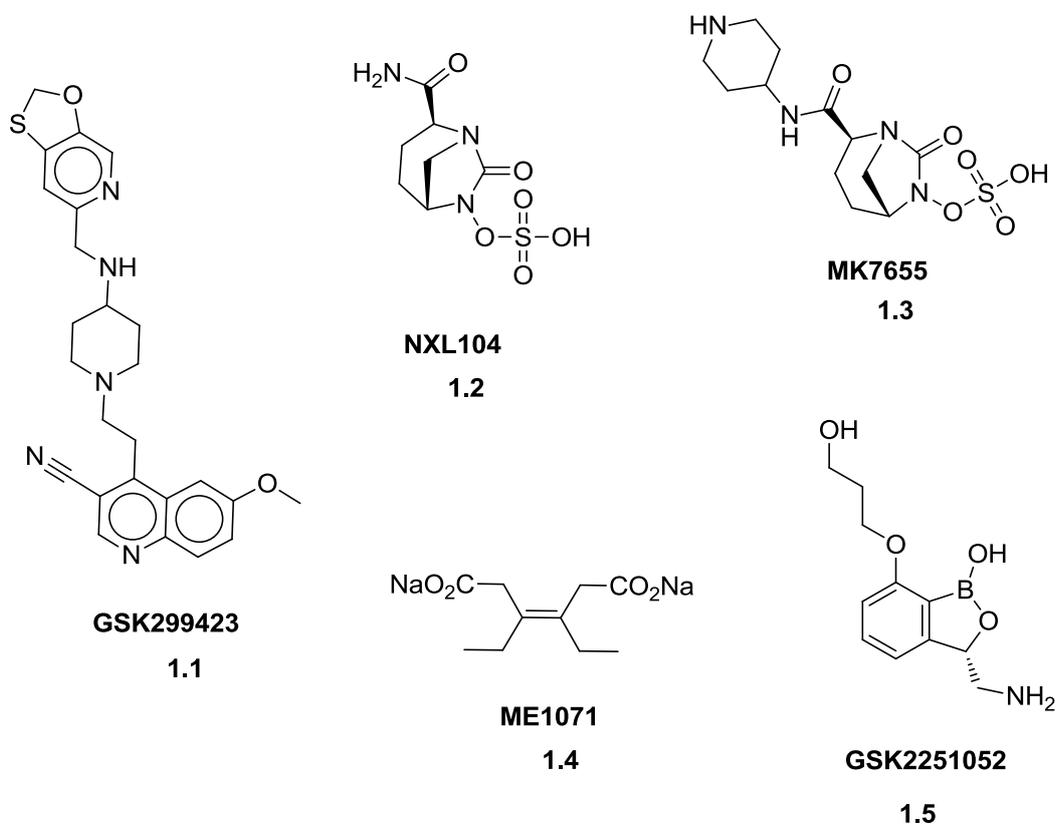
During the 'genomic era', the ability to clone genes and to produce purified proteins to be used for *in vitro* screening and assays allowed genomic-based and target-directed screening.<sup>16</sup> The systematic sequencing of pathogen genomes identified many targets not present in mammalian counterparts and an intense high-throughput screening of in-house synthetic chemical libraries was performed leaving apart natural compounds.<sup>16</sup> The compound libraries yielded very few hits when compared to other therapeutic areas and even fewer leads.<sup>16</sup> No antibiotic was developed using this approach during a 20 year time.<sup>16</sup>

One reason for this failure of discovery could be ascribed to the chemical composition of compound libraries. These libraries are generally biased towards molecules compliant with Lipinski's 'rule of five' with consequently good oral absorption but many existing antibiotics do not conform to this rule especially the ones which are administered *via* the parenteral route.<sup>16, 18</sup> According to some authors,<sup>19</sup> antibacterial agents reside within a unique physicochemical property space when compared to other therapeutics. Another problem is that a compound binding to a target does not necessarily show antibiotic activity because it may not be able to penetrate the bacterial membranes or to overcome removal from the cell by efflux processes mediated by membrane transporters.<sup>18</sup>

#### **1.2.4 Antibiotics in the pipeline and the future of antibacterial discovery**

Over the past 20 years (from 1983 to 2002) FDA approvals of new antibacterial agents decreased. Comparing the period from 1998 to 2008 with the period from 1983 to 1987, the approvals of antibacterial agents decreased by 56%.<sup>20</sup> During this period, forty compounds were under evaluation in clinical trials at various phases with an equal distribution between natural products and synthetically derived molecules.<sup>21</sup> Some new strategies were adopted by the companies for the compounds in the pipeline which included: (i) the search for new classes of molecules that bind well-established targets such as the non-quinolone topoisomerase inhibitors GSK299423 **1.1** (Figure 1.1), non  $\beta$ -lactam inhibitors of  $\beta$ -lactamases such as NXL104 **1.2**, MK7655 **1.3** and ME1071 **1.4** (Figure 1.1); (ii) inhibitors for multidrug efflux mediated mechanisms which could restore the activity of known antibiotics like quinolones towards bacterial species relying on this resistance mechanism such as *P. aeruginosa*; (iii) screening of unconventional chemical classes to overcome the limitations of classical compound libraries like the boron-based Leu-tRNA synthetase inhibitor GSK2251052 **1.5** (Figure 1.1).<sup>18</sup> A rediscovered interest in natural products has recently occurred which involves two main strategies: screening of various organism groups different from soil *Streptomyces* such as plants, deep sea bacteria and antifungal *Actinomycetes* which colonize the nests of ants or

manipulating the non-expressed regulatory gene clusters in *Streptomyces* since many existing antibiotics are only growth-phase dependent regulatory products.<sup>18</sup> Combinatorial synthesis through Diversity-Oriented Synthesis (DOS), and Biology-Oriented Synthesis (BIOS) will enhance the diversity of HTS libraries, while Fragment-Based Drug Discovery (FBDD), thanks to the growing number of solved crystallographic and nuclear magnetic resonance structures of bacterial targets with bound ligands, will deliver small molecules which are more likely to bind to selected targets when compared to larger molecules and will constitute a valuable starting point for optimization.<sup>16</sup>



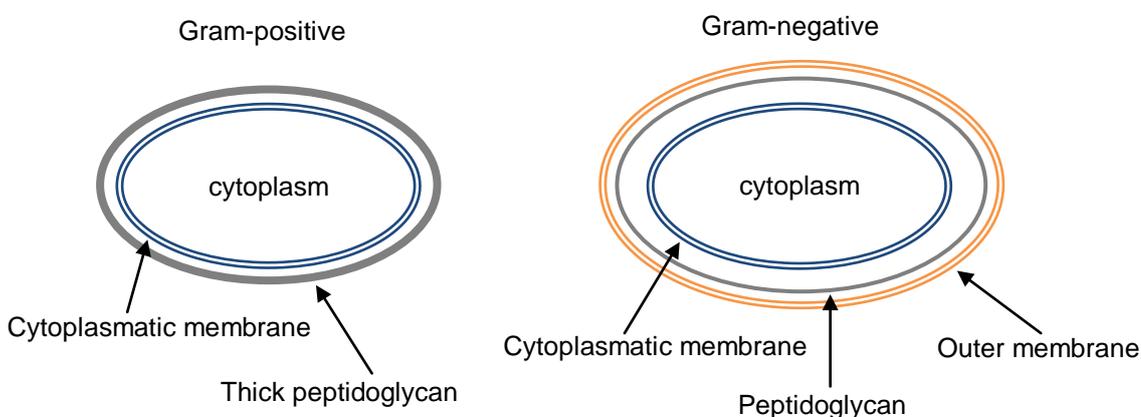
**Figure 1.1** Structures of new classes of antibiotics<sup>18</sup>

### 1.2.5 Requirements for an optimal antibacterial target and the problem of bacterial cell penetration

An antibacterial target should be considered amenable if it is (i) essential for the organism function to allow inhibition of bacterial growth or death; (ii) conserved among different bacterial species to provide a good spectrum of activity; (iii) lacking structural homology with the mammalian host in order to avoid toxicity, (iv) 'druggable' since it should possess a site or a structure where

small drug-like molecules can bind.<sup>16</sup> Additional desirable criteria are: (v) good solubility and stability of the enzyme in terms of delivering good quality and highly resolved X-ray crystallography data, (vi) an intrinsic low frequency of mutations to allow inhibitors to be effective for a reasonable time when used as monotherapeutic agents.<sup>16</sup> Unfortunately, in the field of antibacterials, the generally observed rapid development of resistance by single target enzymes constitutes a big challenge.<sup>16</sup> Developing inhibitors for multiple enzymes like fluoroquinolones (inhibition of DNA gyrase and topoisomerase IV) or targets encoded by multiple genes (such as rRNA inhibitors) could be advantageous with respect to single target-based antibacterial agents.<sup>16</sup>

The spectrum of activity of antibacterials is crucially dependent on cell permeability and the target distribution<sup>16</sup> within the bacterium. Targets can be external to the cell or in the cytoplasm and most antibiotics inhibit one among the several cellular targets.<sup>16</sup> Bacteria are prokaryotic cells with high intracellular pressure and are protected from osmolysis by a rigid peptidoglycan layer, constituted by strands of glycan and peptide covalently cross-linked, which contributes to cell wall rigidity.<sup>16, 22, 23</sup> Bacterial cells can be classified according to their ability to retain the violet Gram-staining and Gram-positive bacteria have a thicker peptidoglycan layer (30-100 nm thick) when compared to Gram-negative strains (20-30 nm thick).<sup>10, 16, 23</sup> An additional lipopolysaccharide outer membrane (OM) is present in Gram-negative bacteria and this represents an additional barrier to the cytoplasmatic membrane (CM) present in all bacteria<sup>16</sup> (Figure 1.2).



**Figure 1.2** Structure of Gram-positive and Gram-negative bacteria.

Uncharged and lipophilic molecules can readily diffuse through the CM while charged hydrophilic molecules can rely on active transport *via* solute-specific carriers and permeases.<sup>16</sup> The OM is globally impermeable but molecules can transit through porins, water-filled channels which are selective for hydrophilic and charged solutes with a size exclusion for high molecular weight molecules (in *E. coli* the upper limit is 600 Daltons).<sup>16</sup> Porins can be bypassed thanks to solute-specific facilitated diffusion channels and some natural antibiotics exploit these solute-specific routes.<sup>16</sup> The presence of specific efflux pumps in Gram-negative organisms contributes to the difficulties in the discovery of specific drugs for this class of bacteria.<sup>16</sup> Another important aspect of antibacterial agents is target-specific selectivity. When an inhibitor is discovered *via* empirical (phenotypic) screening for growth inhibition, it should not possess non specific antibacterial activity *via* cytotoxicity such as detergent action, alkylation or 'energy poisoning'.<sup>16</sup> This is equally important also for compounds identified *via in vitro* enzyme inhibition but a linkage between enzyme-based inhibitory potency and minimum inhibitory concentration (MIC) is not obvious for many compound classes because of the inability to cross the bacterial membranes or to concentrate inside the cell avoiding removal by efflux pumps.<sup>16</sup>

Possible approaches to overcome the problem of bacterial cell penetration include the alteration of the physicochemical properties of inhibitors to enhance uptake and limit efflux, but these modifications may reduce binding affinity to the molecular target.<sup>24</sup> The use of docking methodologies could help to predict whether the modifications are tolerated within the binding site.<sup>24</sup> Another approach is to exploit bacterial iron uptake through siderophores as a 'Trojan horse' strategy for drug delivery. Bacterial cells require iron as an essential element in a variety of metabolic pathways and high-affinity iron uptake is mediated by siderophore-dependent processes.<sup>25</sup> Some siderophore conjugates have already been made with existing drug classes, such as  $\beta$ -lactam antibiotics<sup>25</sup> and in some cases antibacterial activity was better when compared to the non-conjugated drugs. However, not all inhibitors will be chemically amenable to coupling with siderophores and drug release is then

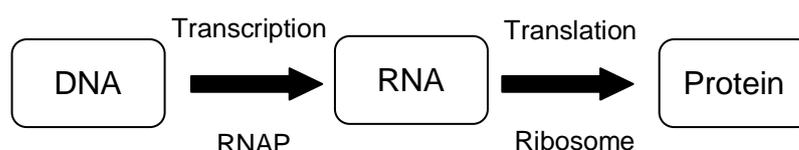
necessary within the bacterial cytoplasm. Moreover good pharmacokinetic properties are required for systemic administration.<sup>24</sup>

To conclude, the choice of a suitable target can be considered a rate-limiting step of any antibacterial drug discovery program.<sup>16</sup>

## 1.3 Bacterial DNA-dependent RNA polymerase

### 1.3.1 Structure

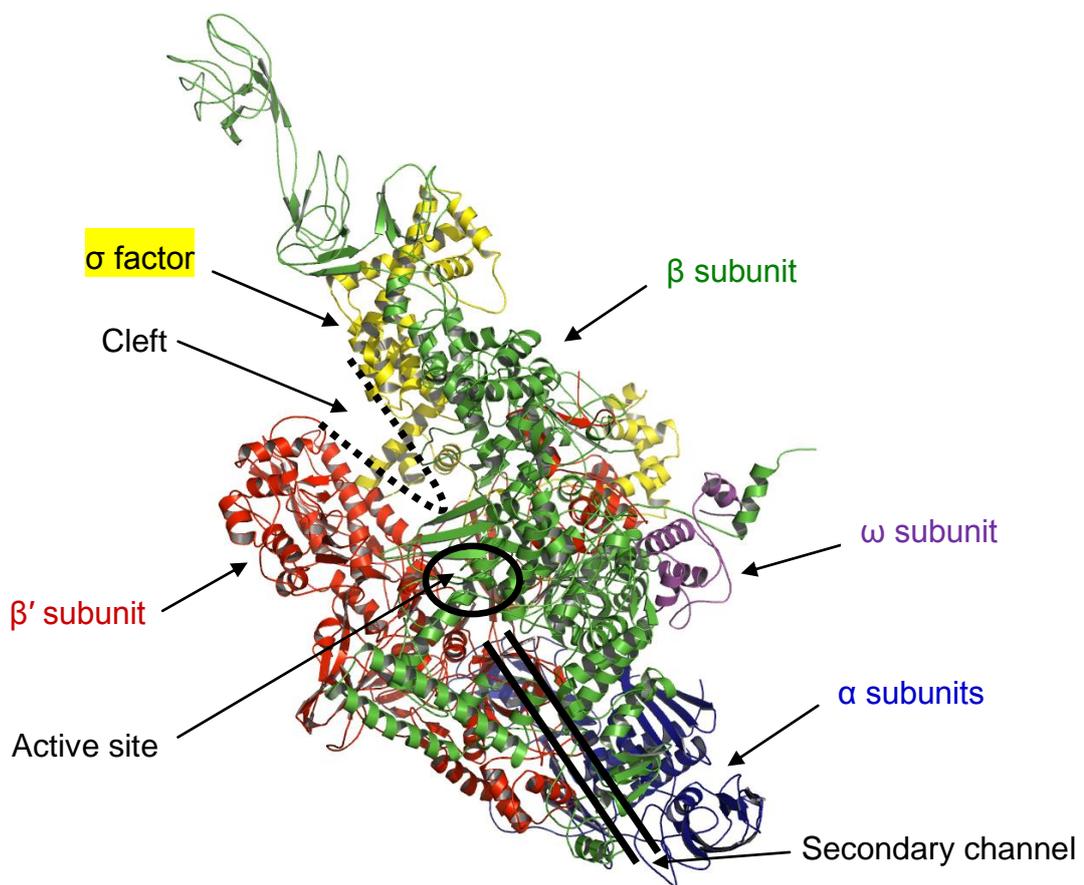
RNAP is a nucleotidyl transferase enzyme with a molecular mass of about 400kDa and is responsible for the transfer of genetic information from DNA to RNA in a process called transcription which is the first step of the expression of genetic information<sup>26</sup> (Figure 1.3).



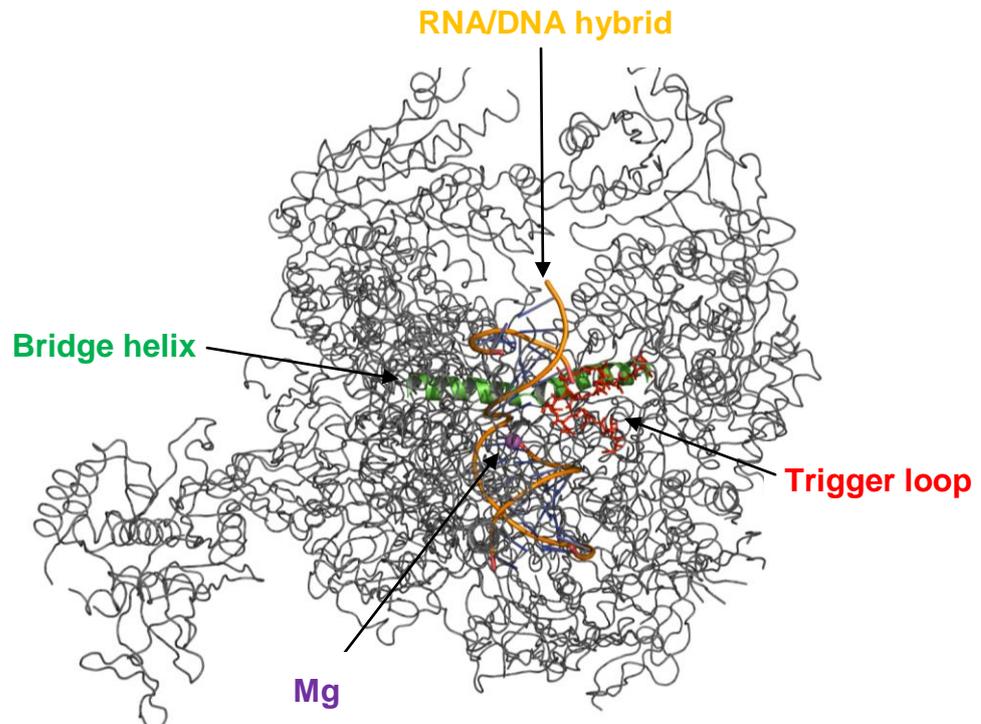
**Figure 1.3** Expression of genetic information

The DNA-dependent RNA polymerase (RNAP) in prokaryotic cells is responsible for the biosynthesis of all types of RNAs (r-RNA, mRNA, t-RNA) while in eukaryotic cells this enzyme is present in three different forms (RNAP-I, RNAP-II, RNAP-III) where each of these is respectively involved in the biosynthesis of r-RNAs, m-RNAs and t-RNAs.<sup>27</sup> Eukaryotic RNAPs contain 14 subunits in RNAP-I, 12 subunits in RNAP-II and 17 subunits in RNAP-III and all of these three enzymes are similar from a structural and functional point of view while prokaryotic core RNAPs consists of five subunits ( $\alpha_2 \beta \beta' \omega$ ) in association with the  $\sigma$  factor to form an holoenzyme which can bind a promoter<sup>27</sup> (Figure 1.4). In the eukaryotic cell the specific binding of a promoter is performed by transcription factors in association with the RNAPs.<sup>27</sup>

RNAP (Figure 1.4) is a 'crab claw'-shaped molecule and the large  $\beta$  and  $\beta'$  subunits form the 'pincers' with a large channel between them which locates the 3' hydroxyl of the RNA within the active site, the 8-9 base pair long RNA-DNA hybrid at the growing end of the transcript, a DNA duplex downstream of the hybrid and a 6-nucleotide long single-stranded RNA upstream of the hybrid.<sup>28</sup> The active site is directly connected to the surface of the enzyme by a secondary channel which is a passageway for the incoming nucleotide triphosphate substrates. This channel is required as the presence of the nucleic acids blocks access from the main channel.<sup>28</sup> The two 'pincers' of the 'crab claw' are bridged by a bridge helix (Figure 1.5), a long  $\alpha$  helix that spans the main channel near the active site.<sup>28</sup>



**Figure 1.4** Various subunits of the *T. aquaticus* RNA polymerase (PDB 116V) (see colour code) and the positions of the Cleft, Pore and RNA exit channel.

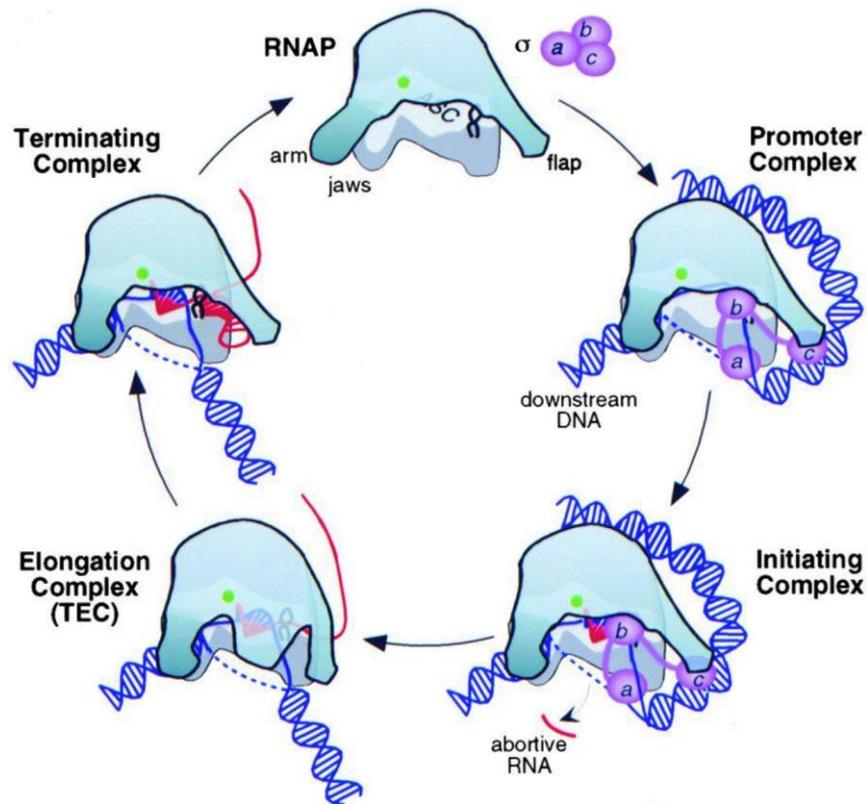


**Figure 1.5** Structural elements of RNAP in the region of the DNA-RNA hybrid (PDB 1I6V).

The active site of the RNAP is located on the floor of the cleft formed by the  $\beta$  and  $\beta'$  subunits.<sup>26</sup> The RNA-channel and the secondary channel for nucleotide triphosphate (NTP) substrates are the two access channels that connect the inner part of the enzyme with the external surface.<sup>26</sup> Three aspartate residues in the active centre coordinate a catalytic  $Mg^{2+}$  ion while another  $Mg^{2+}$  ion is bound to the incoming NTP.<sup>26</sup> The bridge helix (also called the F helix) and the trigger loop (also called G/G loop) (Figure 1.5), together with various other loops are important for the loading of NTPs and for catalysis and translocation and all these flexible structures are in the downstream face of the active centre.<sup>26</sup>

### 1.3.2 Catalytic reaction mechanism

The process of transcription consists of three main stages (Figure 1.6): initiation, elongation and termination.<sup>26</sup>

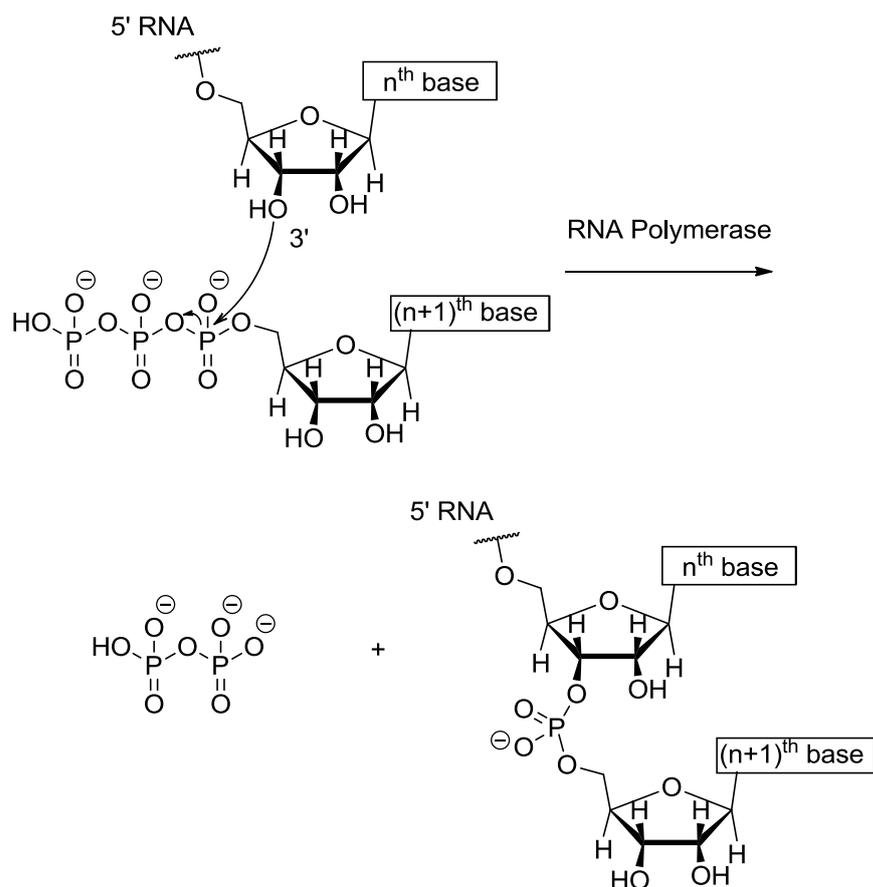


**Figure 1.6** Intermediates in the transcription cycle. RNAP in grey, DNA in blue, nascent RNA in red, multidomain  $\sigma$  factor in purple. Adapted from *Mooney et al.*<sup>29</sup>

The process begins when the  $\sigma$  initiation factor binds to the core enzyme and starts the recognition and melting of promoter sequences on DNA making possible the synthesis of the RNA transcript.<sup>26</sup> The initiation of transcription can be divided into several sub-steps and the first is the binding of a promoter which, after structural rearrangement, melts the double helix at the promoter site and forms the secondary complex.<sup>26</sup> Once the first nucleotide is associated with the nascent RNA chain, the first ternary complex is formed which evolves to an early ternary complex after the incorporation of the second nucleotide of the RNA chain with the subsequent formation of the first phosphodiester bond.<sup>26</sup> This complex can undergo abortive initiation *via* the release of its dinucleotide component or it can undergo elongation of the dinucleotide.<sup>26</sup> The abortive initiation stops after the incorporation of 9-11 nucleotides into the emerging RNA chain and at this stage the ternary complex is stabilised, the initiation has

ended and the elongation process can begin *via* the formation of a stable transcription elongation complex following a significant conformational change.<sup>26</sup> This complex is very stable and upon its formation the initiation factor  $\sigma$  is released allowing the enzyme to decouple from the promoter.<sup>26</sup> The highly efficient transcription cycle ends when a stop signal is recognized with the subsequent dissociation of the transcribing complex and release of the RNA polymerase which is readily available for a new round of transcription.<sup>26</sup>

As indicated earlier, two magnesium ions are involved in the mechanism of phosphodiester bond formation catalysed by RNA polymerase.<sup>27</sup> As indicated previously, the first  $Mg^{2+}$  ion is coordinated by three conserved aspartates included in the NADFDGD motif of the  $\beta'$  subunit and this ion catalyses the nucleophilic attack of the 3'-oxygen of the growing RNA chain on the 5'  $\alpha$ -phosphate of the incoming nucleotide (Scheme 1.1), whilst the second  $Mg^{2+}$  ion is chelated to the incoming NTP and is also coordinated by three aspartates (two from the  $\beta'$  and one from the  $\beta$  subunit located in a conserved ED motif).<sup>27</sup> After phosphodiester bond formation, the enzyme is translocated along the nucleic acid template for the addition of the next nucleotide.<sup>27</sup> The enzyme's active site is one of the regions which is highly conserved between prokaryotes and eukaryotes.<sup>27</sup> The nucleic acids are accommodated into the positively charged cleft formed by the  $\beta$  and  $\beta'$  subunits.<sup>27</sup> The 'wall' domain contained within the  $\beta$  subunit closes the upstream extremity of this cleft and the 'flap' element contained within this domain is a binding site for transcription factors and may have a role in obstructing the RNA exit channel *via* the  $\sigma$  factor.<sup>27</sup>



**Scheme 1.1** Mechanism of ribonucleotide addition to the RNA chain.

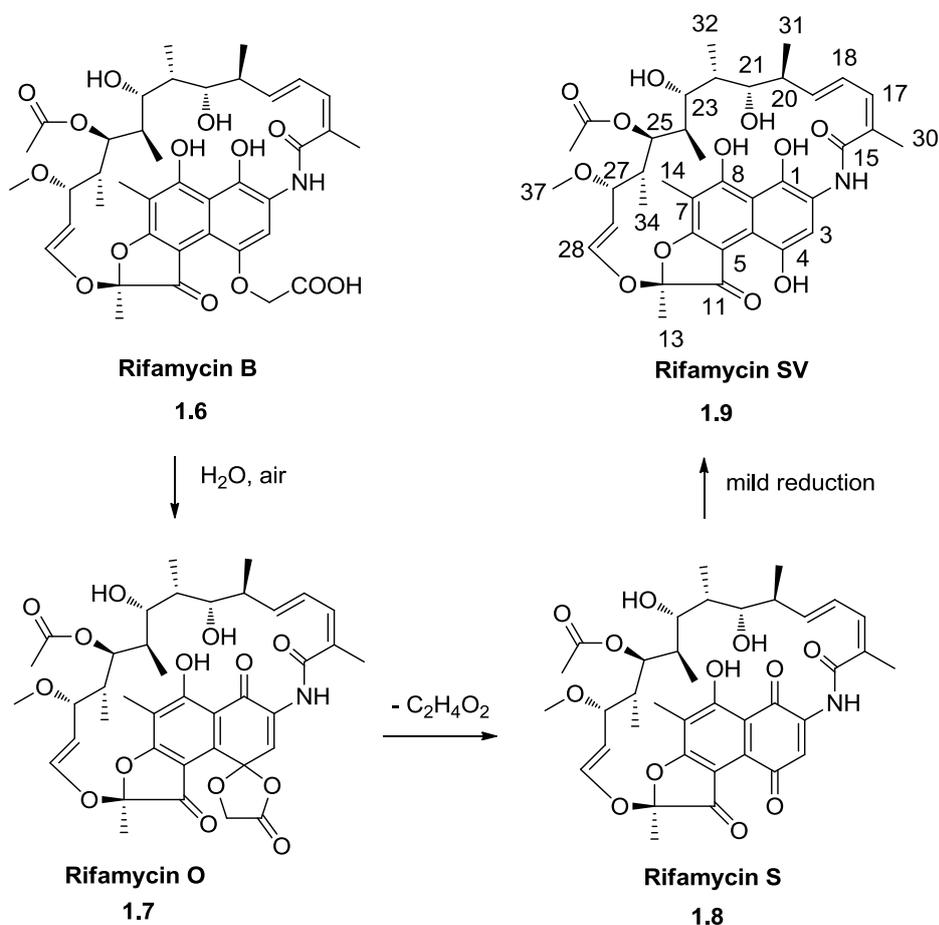
## 1.4 RNAP inhibitors and their mechanism of action

### 1.4.1 Rifamycins

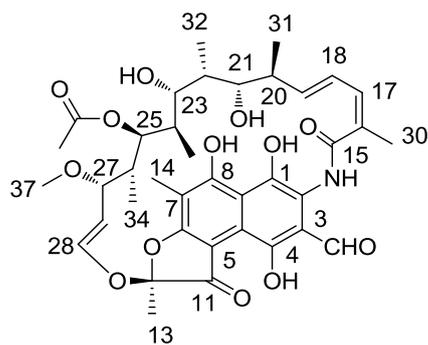
#### 1.4.1.1 Chemistry and biological activity

The ansamycin antibiotics are characterized by a basket-like molecular shape which comprises an aromatic moiety bridged by an aliphatic chain (ansa) in two nonadjacent positions.<sup>3</sup> The aromatic portion can be a naphthalene or a naphthoquinone (naphthalenic ansamycins) or a benzene or benzoquinone ring (benzenic ansamycins). The rifamycins are members of the ansamycins family and were isolated from *Amiclatopsis mediterranei* as a complex mixture of cognate compounds in 1959 by Sensi and co-workers at Lepetit SA in Italy.<sup>30</sup>

The rifamycins have a broad spectrum of antibiotic activity against Gram-positive and, to a lesser extent, against Gram-negative bacteria.<sup>31</sup> The lower activity with Gram-negative strains is due to the diminished penetration of the antibiotic through the outer cell membrane.<sup>31</sup> Interestingly, ansamycins are also active against the transcription machinery of the eukaryotic parasite *Plasmodium falciparum*.<sup>32, 33, 34</sup> Rifamycin SV **1.9** was the first rifamycin in clinical use and is a biosynthetic precursor of rifamycin B **1.6** which, in comparison has relatively modest activity.<sup>35, 36, 37, 38, 39</sup> However rifamycin B **1.6** can be easily converted to rifamycin SV **1.9** in various ways (chemically, microbiologically and by biotransformations).<sup>35, 36, 37</sup> With regards to the chemical transformation route (Scheme 1.2) rifamycin B **1.6** is converted spontaneously and reversibly to rifamycin O **1.7** in an aqueous oxygenated solution and the latter is hydrolysed with concomitant loss of glycolic acid to rifamycin S **1.8**. Rifamycin SV **1.9** is obtained using a mild reduction of rifamycin S **1.8**.

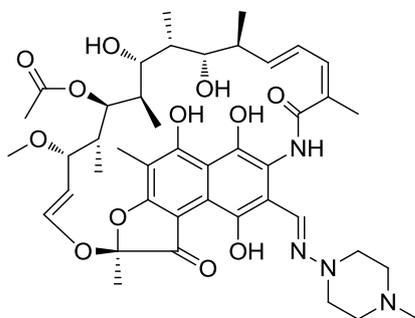


**Scheme 1.2** Conversion of rifamycin B **1.6** into rifamycin SV **1.9**.

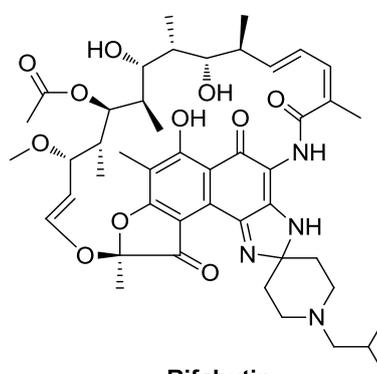


1.10

Figure 1.7 3-formylrifamycin SV 1.10 with numbering.

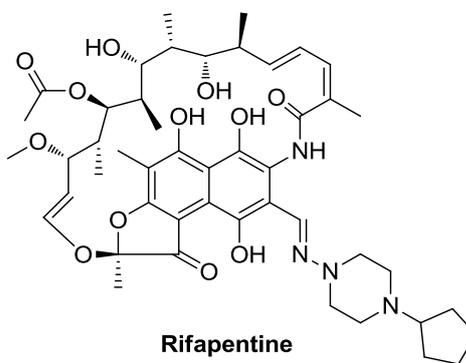


Rifampicin  
1.11



Rifabutin  
1.12

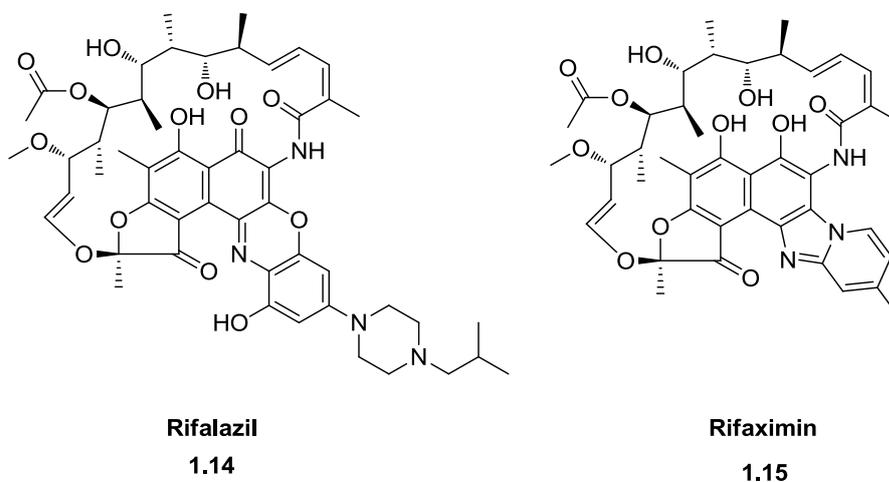
Figure 1.8 Structures of rifampicin 1.11 and rifabutin 1.12.



Rifapentine  
1.13

Figure 1.9 Structure of rifapentine 1.13.

Starting from the easily accessible intermediate 3-formylrifamycin SV **1.10** (Figure 1.7), many derivatives functionalised at the C-3 position have been prepared using semisynthesis and among these, rifampicin **1.11** (Figure 1.8) has a particularly pronounced activity towards Gram-positive bacteria (such as multidrug resistant *Staphylococcus aureus*) and in particular against mycobacteria. Moreover it has a weaker activity against Gram-negative bacteria.<sup>40, 41, 42, 43</sup> Rifampicin **1.11** exhibits good oral bioavailability and is one of the most important drugs for the treatment of tuberculosis, leprosy and mycobacterial infections associated with AIDS.<sup>3, 44</sup> Rifabutin **1.12** (Figure 1.8) and rifapentine **1.13** (Figure 1.9) are other semisynthetic derivatives introduced later to clinical use and rifabutin **1.12** has showed activity against a number of rifampicin-resistant clinical pathogens.<sup>34</sup> Nowadays the ansamycins currently in therapeutic use are rifampicin **1.11**, rifabutin **1.12**,<sup>34</sup> rifapentine **1.13**, rifalazil **1.14** and rifaximin **1.15** (Figure 1.10).



**Figure 1.10** Structure of rifalazil **1.14** and rifaximin **1.15**.<sup>45</sup>

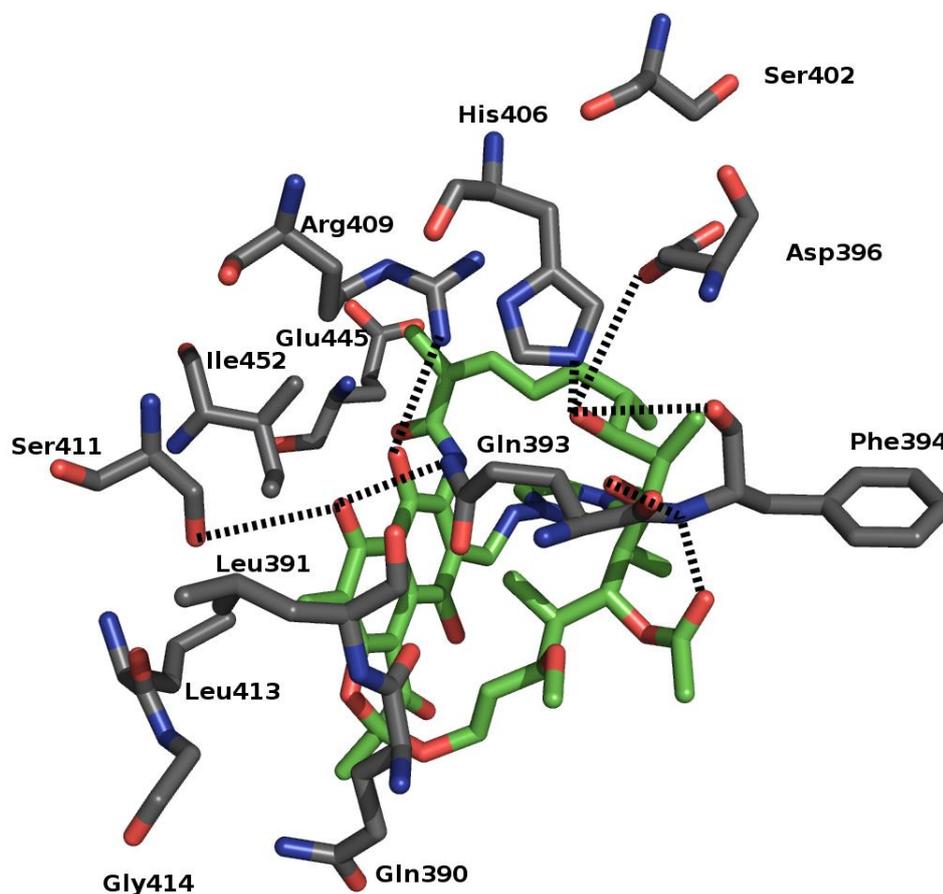
Rifampicin **1.11** exhibits its antibacterial activity by strongly binding the prokaryotic DNA-dependent RNA polymerase and inhibiting RNA synthesis.<sup>3, 46, 47</sup> This inhibition mechanism is shared by all the antibacterially active rifamycins despite the many structural modifications made in these molecules to alter the pharmacokinetics and the selectivity for the prokaryotic enzyme. The eukaryotic

DNA-dependent RNA polymerase has a weaker affinity to rifampicin **1.11** compared to prokaryotic RNAP allowing in this way selective toxicity towards bacteria.<sup>3</sup>

#### 1.4.1.2 Mechanism of action

Following genetic analysis of rifamycin resistant strains and the X-ray crystallographic based structure elucidation of the core enzyme in complex with various inhibitors, much information has been obtained concerning the mode of action of the rifamycins. In particular, it has been determined that rifampicin **1.11** binds to a site in the  $\beta$  subunit 12 Å away from the active site in the path of the nascent RNA. Twelve residues from the  $\beta$  subunit are involved in hydrogen bonds or van der Waals interactions with rifampicin **1.11**.<sup>26, 48</sup> Interestingly, the substituent in the C3 position of rifampicin **1.11** does not interact with the protein and this suggests that the rifampicin **1.11** is sterically blocking the synthesis of an RNA product longer than three nucleotides. Indeed, mutation of some residues confers resistance to rifampicin by altering the binding pocket geometry.<sup>48</sup> According to the crystal structure of rifampicin **1.11** in complex with the *Thermus aquaticus* core DNA dependent RNAP,<sup>48</sup> rifampicin **1.11** binding (Figure 1.11) involves hydrogen bonding interactions between the hydroxyl groups at C-1, C-8, C-21 and C-23 and the carbonyl oxygen of the C-25 acetoxy group with the amino acid residues Arg409, Ser411, Gln393, His406, Asp396 and Phe394. The binding of the antibiotic is reinforced by additional hydrophobic interactions with Glu445, Ile452, Gly414, Leu413, Leu391 and Gln390. The orientation of the bound antibiotic interferes sterically with the nascent oligonucleotide chain after the first or second chain elongation step and it is evident that there is no effect on the initiation or translocation step.<sup>49</sup> Unfortunately this steric model is not able to account for several amino acid substitutions that confer resistance to rifampicin **1.11** but not to rifabutin **1.12** and also in close analogues which have a fused ring in C3 and C4. Following analysis of crystal structures of rifabutin **1.12** and rifapentine **1.13** co-crystal structures (*T. Thermophilus* RNAP holoenzyme which includes the  $\sigma$ -subunit in addition to the  $\alpha_2\beta\beta'\omega$  assembly), an allosteric mechanism of action for

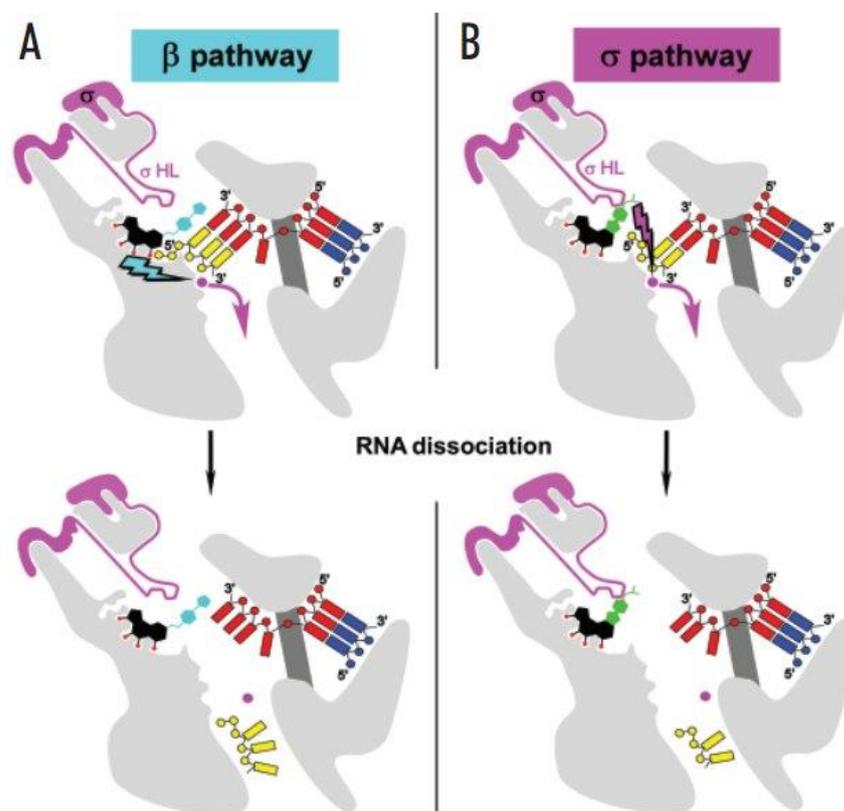
rifampicin **1.11**, rifabutin **1.12** and rifapentine **1.13** was recently proposed.<sup>50, 51</sup> While in rifapentine **1.13** there is no contact between the C3 substituent and the  $\sigma$ -subunit, in rifabutin **1.12** the C3/C4 substituent is able to interact with the  $\sigma$ -subunit.<sup>50, 51</sup>



**Figure 1.11** Schematic diagram of RNAP  $\beta$  subunit interactions with rifampicin indicated in green (PDB 1I6V) **1.11**.

Regarding rifampicin **1.11**, according to a recent study,<sup>50, 51</sup> the binding of this molecule to RNAP results in the propagation of an allosteric signal for 19 Å up to the active site inducing a decrease in affinity of the major catalytic  $Mg^{2+}$  ion with a consequent decrease in speed of the catalytic reaction and dissociation of the short DNA/RNA duplex. This mechanism of action has been called the  $\beta$ -pathway (Figure 1.12) and stops the transcription process *in vitro* at the second phosphodiester bond when the process starts from a dinucleotide.

This mechanism is activated by the interactions of the ansamycin core (aromatic ring and ansa) with the residues in the  $\beta$ -subunit and this mechanism is shared by all the rifamycins.<sup>50, 51</sup> The C-3/C-4 substituted rifamycins like rifabutin **1.12** can affect the formation of the first phosphodiester bond *via* the  $\sigma$ -pathway (Figure 1.12) thanks to the contacts between the aromatic substituent and the  $\sigma$ -subunit, with the same effect on the major catalytic  $Mg^{2+}$  and consequent decrease in speed of the catalytic reaction and dissociation of the short DNA/RNA duplex.<sup>50, 51</sup> This last mechanism is a unique characteristic of C3/C4 substituted rifamycins which are also able to act within the  $\beta$ -pathway at the same time. A recent study<sup>52</sup> has shown that this allosteric modulation for the binding affinity of the major catalytic  $Mg^{2+}$  is not possible and these researchers have cast doubt on the allosteric modulation mechanism.



**Figure 1.12** Two-pathway mechanism of rifamycins action. (A) The  $\beta$  pathway, induced fit mechanism shown with the cyan lightning, RNA in yellow, DNA template in red and blue. (B) The  $\sigma$  pathway, induced fit mechanism shown with the purple lightning, RNA in yellow, DNA template in red and blue.<sup>51</sup>

### 1.4.1.3 Structure activity relationships

Several rifamycin analogues have been prepared with the purpose of improving bioavailability and antimicrobial activity.

Rifamycin SV **1.9** (Scheme 1.2) was modified at various positions and these reveal that changes in the ansa chain generally negatively affect the activity and in particular, substitution and elimination of the two hydroxyl groups at C-21 and C-23 results in a large decrease of activity.<sup>26, 53, 54, 55</sup> All the modifications which leave hydroxyl groups unmodified but which alter the conformation of the ansa chain are also unfavourable.<sup>26, 53, 54, 55</sup> Alterations in the ansa rings are unfavourable with the exception of deacetylation at C-25 and the inversion of configuration at this stereocenter which have no negative influence upon activity. The hydroxyl group at C-8 is essential for antimicrobial activity while keto groups in the chromophoric moiety can be modified to hydroxyl groups in position C-1 and C-4 with no loss of activity.<sup>26</sup>

Many rifamycin derivatives with substitutions at position C-3 and/or position C4 have been prepared.<sup>56, 57, 58, 59, 60</sup> Among the C3/C4 position derivatives the most important are rifaximin **1.15**<sup>61, 62</sup> and rifabutin **1.12**.<sup>62</sup>

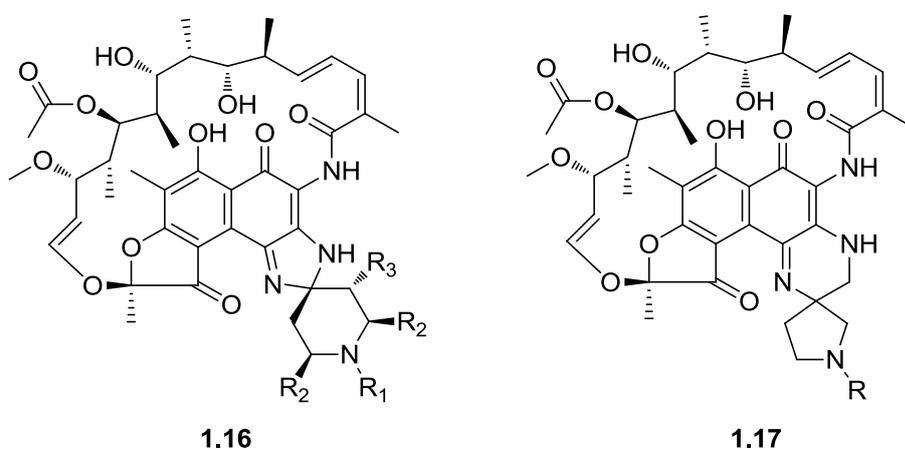
Rifapentine **1.13** is more active than rifampicin **1.11** against *Mycobacterium tuberculosis* and has a longer half-life which allows once weekly administration. It is approved for the treatment of tuberculosis but unfortunately some bacterial strains which are resistant to rifampicin **1.11** usually are also cross-resistant to rifapentine **1.13**.<sup>63</sup>

Rifaximin<sup>26</sup> **1.15** is active against bacterial enteropathogens and it is approved for this use while rifabutin **1.12** is active against certain rifampicin-resistant clinical pathogens and it is also in clinical use since 1994.

Rifalazil<sup>26</sup> **1.14**, a benzoxazino derivative, shows better pharmacokinetic properties and less side effects like interaction with liver microsomal enzymes compared to rifampicin **1.11** even if some rifampicin-resistant strains show cross-resistance with rifalazil **1.14**.

Some 25-hydroxyl analogues were synthesised and showed better efficacy than the 25-O-acetyl derivatives on some bacterial strains.<sup>26</sup>

Finally new rifabutin **1.12** analogues (Figure 1.13) such as spiro-piperidylrifamycins<sup>64</sup> **1.16**, spirorifamycins<sup>65</sup> **1.17**, C-11 oxime<sup>66</sup> derivatives and C-25 carbamates, were also prepared and generally showed activity on certain wild type strains of *S. aureus* and improved activity on some mutant strains. In particular, the C-25 carbamates diminish susceptibility to rifamycin-specific ADP-ribosyl transferases which are responsible for resistance by inactivation on some opportunistic pathogens including *Mycobacterium smegmatis* and *Pseudomonas aeruginosa*.<sup>67</sup>



**Figure 1.13** Spiro-piperidyl-rifamicin **1.16** and spirorifamicin analogues<sup>45</sup> **1.17**.

#### 1.4.1.4 Resistance to rifamycins

Bacteria and fungi exist in communities and they both need to communicate with each other and to compete as well. These organisms synthesize low molecular weight compounds to favour symbiosis or to inhibit the growth of competitors and bacteria have developed various strategies to survive in the presence of such antibiotics.<sup>34</sup>

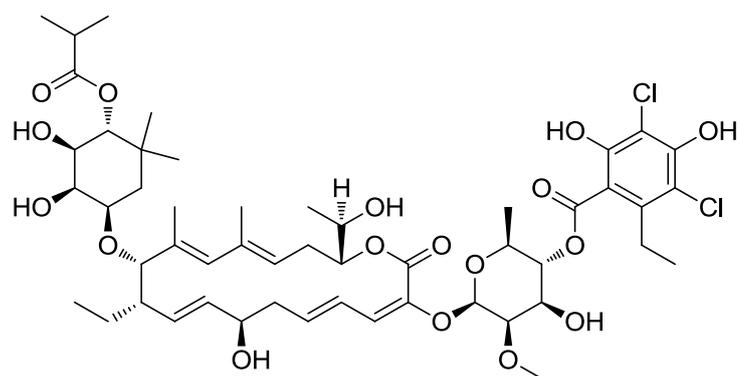
Rifampicin **1.11** blocks transcription initiation by binding to conserved amino acids near the active centre of RNAP and one of the main causes of bacterial resistance are the mutations of these amino acids.<sup>34,48</sup> The rate of resistance of pathogens to rifampicin **1.11** is  $10^{-8}$  to  $10^{-9}$  per bacterium per cell division. Many bacterial strains resistant to rifampicin **1.11** such as *Escherichia*

*coli*, *Staphylococcus aureus*, *Bacillus subtilis*, *Mycobacterium tuberculosis* etc. have mutations in the RNAP active centre located in the  $\beta$ -subunit which is encoded by the *rpoB* gene.<sup>34,48,3</sup> This binding site is highly conserved among bacteria.<sup>48</sup> Generally the binding of molecules to RNAP involves twelve residues and in most cases mutagenesis of each of these generates a spontaneous resistant phenotype. Single point mutations of amino acids are more common than insertions and deletions and 95% of these are located in four regions in the N-terminal half of the  $\beta$ -subunit.<sup>68</sup> In *E. coli* the mutations which confer resistance are located in the central region and in the N-terminus of the  $\beta$ -subunit while in *M. tuberculosis*, these are located in a 'hot-spot' within the central region where the mutation of Ser411, His406, Asp396 (according to the *Thermus aquaticus* numbering) account respectively for 41%, 36% and 9% of the clinically isolated resistant phenotypes.<sup>68,69,70</sup> High levels of resistance to rifampicin **1.11** have been reported in more than 1% of rifampicin resistant strains with the (V176F) mutation (*M. tuberculosis* numbering).<sup>71</sup> Following studies on the *rpoB* gene in *Thermus aquaticus*, it has been reported that twelve amino acids are involved in hydrogen bonding or van der Waals interactions with the bound rifampicin and all of them are susceptible to mutation with the mutation of Glu445 being lethal for the microorganism.<sup>3,34</sup> Interestingly, twelve of the twenty-three sites that can confer rifampicin **1.11** resistance when mutated do not make direct interactions with the bound antibiotic but are able to alter the overall conformation of the rifampicin binding site.<sup>3,34</sup> Other mutations which occur at high rate have been reported but most of them do not lead to clinical isolates due to reduced fitness of the microorganisms.<sup>34</sup> Duplication of the *rpoB* gene in *Nocardia sp.* seems to be a novel resistance mechanism confirmed using knockout studies.<sup>34</sup> In addition to the mutation of the target strategy other resistance mechanisms have been reported such as the expression of RNAP binding proteins induced by exposure to low concentration of rifampicin in *Streptomyces coelicolor*, *Corynebacterium diphtheriae*, *M. tuberculosis* and *M. leprae* confers resistance.<sup>3, 34, 67, 72</sup> Covalent modifications that neutralise rifampicin are also possible in a variety of ways such as glycosylation, ribosylation, phosphorylation. Finally and equally important, an additional strategy of resistance is the modification of membrane

permeability that prevents the entry of the antibiotic inside the cell or the over-expression of membrane-associated energy-driven efflux pumps.<sup>34</sup> Some examples of reduced efficacy resulting from the presence of efflux pumps have been reported in some species of *Mycobacteria* which are generally intrinsically resistant to many antibiotics thanks to the lipid-rich character of their cell wall but are susceptible to rifampicin **1.11**.<sup>34,73</sup>

#### 1.4.2 Lipiarmycin

Lipiarmycin **1.18** (Figure 1.14) is a natural compound mixture of four factors (B3, B4, A3, A4) from *Actinoplanes decanensis*<sup>74, 75</sup> and is active only against Gram-positive bacteria.<sup>34</sup> This antibiotic mixture interferes with the initiation step of RNA biosynthesis. As for rifampicin **1.11**, at much higher concentrations, lipiarmycin affects chain elongation during DNA synthesis.<sup>34</sup> Lipiarmycin is more effective in inhibiting the transcription of the core enzyme rather than the holoenzyme.<sup>34</sup> An important mutation within the RNAP, located in the DNA channel opposite to the rifampicin binding site and proximal to the  $\sigma$ -subunit, confers high lipiarmycin resistance.<sup>34</sup> Tiacumicin B, also known as fidaxomicin (OPT-80),<sup>76</sup> the major component of the tiacumicin complex from *Dactylosporangium auranticum*, is identical to lipiarmycin A3 **1.18** and is active against Gram-positive bacteria including *Clostridium difficile*.<sup>77, 78</sup> It has a low propensity to select for resistant mutants and showed no cross resistance with rifampicin **1.11** and other antibiotics.<sup>77, 78</sup> Fidaxomicin resistance-mutants have been mapped within the switch-2 region into the  $\beta'$  subunit and the  $\sigma^{70}$  subunit region 3.2.<sup>77</sup> fidaxomicin acts by blocking the initiation process only if added before the formation of the open promoter complex in which DNA strands are separated but RNA synthesis has not yet started.<sup>77</sup> These observations led to the conclusion that fidaxomicin operates with a mechanism distinct from the other known transcription initiation or elongation inhibitors.<sup>77, 78</sup> Fidaxomicin was recently approved by the US Food and Drug Administration for the treatment of *Clostridium difficile* infections.<sup>78</sup>



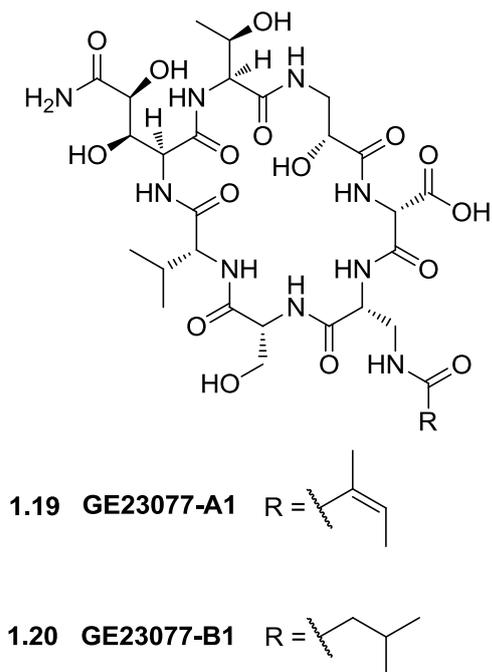
**Lipiarmycin A3**

**1.18**

**Figure 1.14** Lipiarmycin A3 **1.18** structure.<sup>34</sup>

### 1.4.3 GE23077

GE23077<sup>79, 80, 81</sup> (Figure 1.15), a cyclic heptapeptide containing four unnatural amino acids, is a potent and selective inhibitor of bacterial RNAP both from Gram-positive and Gram-negative derived enzymes. It is a mixture of four major components (A1 **1.19**, A2, B1 **1.20** and B2) and has a similar mechanism of action to rifampicin **1.11** but acts at a different binding site to that of rifampicin since it does not show any cross-resistance with the latter. Recent studies,<sup>82</sup> using a combination of genetic, biochemical, and structural approaches, showed that GE23077 binds directly to the RNAP active centre nucleotide binding site preventing transcription initiation. Notably, GE23077 resistance was unusually small, reflecting the fact that its binding site on RNAP includes residues of the RNAP active centre which are essential for RNAP activity.<sup>82</sup> Unfortunately, GE23077 shows only modest antimicrobial activity against some strains of *Moraxella catharralis*, *Neisseria gonorrhoeae* and *Mycobacterium smegmatis* in spite of its potent *in vitro* activity.<sup>81</sup> This is probably due to its hydrophilic nature that prevents crossing of bacterial membranes.

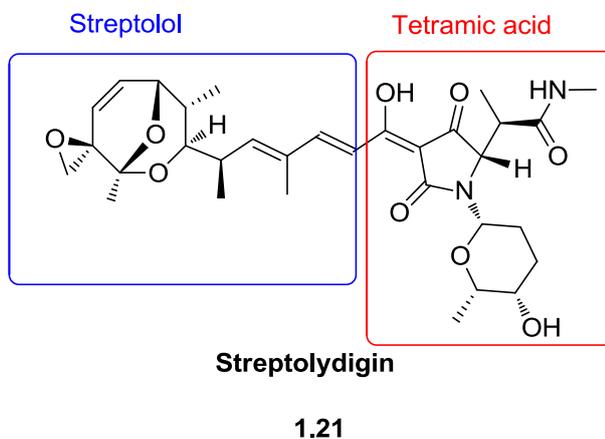


**Figure 1.15** GE23077-A1 **1.19** and GE23077-B1 **1.20** structures.<sup>79</sup>

#### 1.4.4 Streptolydigin

Streptolydigin **1.21**<sup>83, 84</sup> (Figure 1.16), isolated from *Streptomyces lydigus*, is a tetracyclic antibiotic which binds to a site adjacent (20 Å away) to, but not overlapping, the RNAP active site. Binding of streptolydigin stabilises the straight-bridge-helix conformational state as revealed *via* X-ray structure analysis of its complex with RNAP of *T. thermophilus*.<sup>85</sup> Biochemical, mutagenesis and modelling studies have revealed that the RNAP active site can exist in two alternative conformational states: straight-bridge-helix and bent-bridge-helix. This unique inhibition mechanism results in a lack of cross-resistance with the other RNAP inhibitors.<sup>86</sup> Streptolydigin **1.21** is not as medically important as rifampicin **1.11** because it is not able to cross bacterial membranes and its  $K_i$  is  $10^4$ - $10^5$  times higher when compared with that for rifampicin **1.11**. Streptolydigin **1.21** is a non-competitive allosteric inhibitor because it interacts directly with the bridging helix and nearby structures,

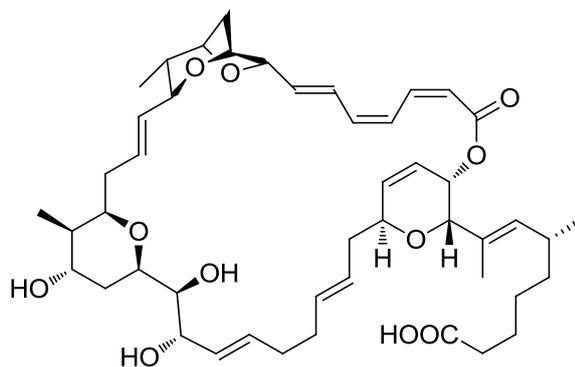
probably inducing conformational changes, during the nucleotide addition cycle.<sup>28</sup>



**Figure 1.16** Structure of streptolydigin **1.21**.

### 1.4.5 Sorangicin

Sorangicin **1.22** (Figure 1.17) is a selective bacterial RNAP inhibitor that acts mainly on Gram-positive bacteria. It is a macrolide polyether and was isolated from a *Sorangium cellulosum* strain.<sup>87</sup> Following the determination of the crystal structure of sorangicin **1.22** in complex with the *T. aquaticus* core RNAP enzyme<sup>88</sup> it was established that the antibiotic binds to a site located in the same RNAP  $\beta$ -subunit pocket as that for rifampicin **1.11**, partially overlapping its binding site and for this reason, sorangicin **1.22** shows partial cross-resistance with rifampicin **1.11**. Sorangicin **1.22** inhibits transcription by blocking the transcript at the length of 2-3 nucleotides and has an increased conformational flexibility when compared to rifampicin **1.11**, this can allow to sorangicin **1.22** to better adapt to conformational changes in the mutated targets.<sup>88</sup>



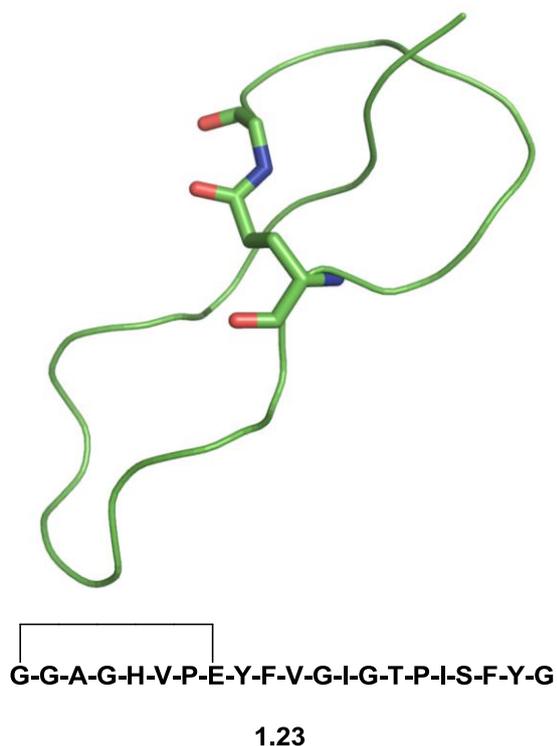
**Sorangicin**

**1.22**

**Figure 1.17** Structure of sorangicin **1.22**.

#### **1.4.6 Microcin J25**

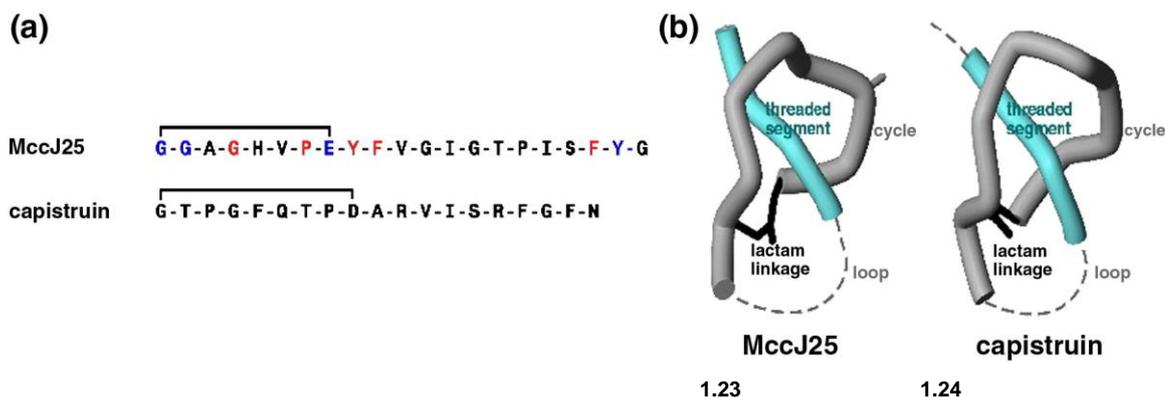
Microcin J25 **1.23** (Figure 1.18) is a cyclic peptide composed of twenty-one amino acid residues<sup>89, 90</sup> and is derived from the microcins, a miscellaneous group of antibiotics produced mostly by strains of *E. coli*. Microcin J25 binds inside the RNAP secondary channel blocking access to NTPs and halting transcription. It has a good range of bactericidal activity against Gram-negative bacteria although resistant strains with amino acidic substitutions both in the  $\beta$  and  $\beta'$  subunits have been isolated.<sup>28</sup>



**Figure 1.18** Structure and sequence of microcin J25 **1.23** (PDB 1PP5).

### 1.4.7 Capistruin

Capistruin<sup>91</sup> **1.24** (Figure 1.19) is a cyclic peptide (threaded lasso peptide) consisting of nineteen amino acid residues, is ribosomally synthesized and post-translationally modified by *Burkholderia thailandensis* E264.<sup>91</sup> Its structure is similar to that of Microcin J25 and it inhibits wild type *E. coli* RNAP but not mutant Microcin J25 resistant *E. coli* RNAP bearing a substitution in the RNAP secondary channel.<sup>91</sup> Despite the low sequence similarity between Capistruin and Microcin J25, in view of their similar three-dimensional structures and their inhibition of bacterial RNAP through possibly identical or overlapping binding sites within the RNAP secondary channel, it is possible to conclude that bacterial RNAP is the functional target of capistruin.<sup>91</sup>



**Figure 1.19** Capistruin **1.24** structure and its comparison with Microcin J25 **1.23**. Aminoacids reported in blue and red are highly mutable, in particular, the red ones are detrimental for the antibiotic activity. Adapted from *Kuznedelov et al.*<sup>91</sup>

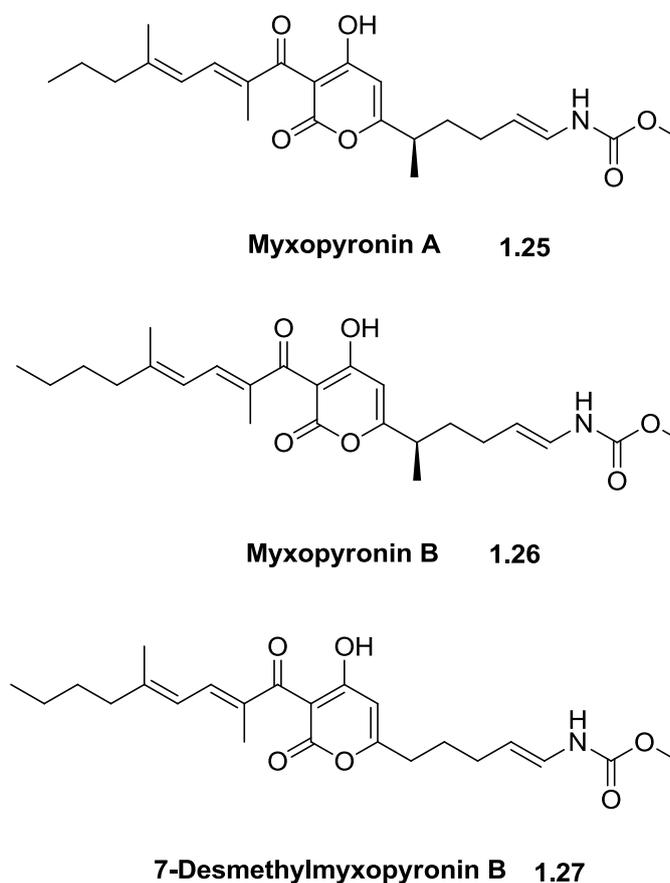
#### 1.4.8 ‘Switch region’ inhibitors

The recent discovery of myxopyronin B (MyxB) **1.26** (Figure 1.20), an  $\alpha$ -pyrone natural product antibiotic that inhibits bacterial RNAP and the determination of its X-ray co-crystal structure with *Thermus thermophilus* RNAP<sup>92, 93</sup> allowed the identification of the ‘switch region’, a new ligand binding site within the RNAP enzyme.<sup>94</sup> This highly mobile structural element, composed of five segments denominated ‘switch 1’ through ‘switch 5’, mediates the conformational changes and the interactions required for loading DNA into the RNAP during transcription initiation.<sup>92, 94</sup> The ‘switch region’ is located at the base of the RNAP  $\beta'$  subunit, known as the clamp, and operates as a ‘hinge’ mediating the opening and the closure of the clamp required respectively for loading the double-stranded DNA into the active centre cleft and to retain it during transcription.<sup>92, 94</sup> The ‘switch region’ residues are conserved in Gram-positive and Gram-negative RNAP allowing a broad spectrum antibacterial activity moreover they are not conserved in eukaryotic RNAPs.<sup>92, 94</sup> It has also been found that the structurally related natural antibiotic Corallopyronin A and the macrocyclic-lactone antibiotic Ripostatin are binding the ‘switch region’ and together with MyxB they do not exhibit cross-resistance with rifamycins since

this region is distinct and remote from the rifamycin binding site.<sup>92, 94</sup> All these three natural antibiotics will be discussed in detail below.

### 1.4.8.1 Myxopyronins

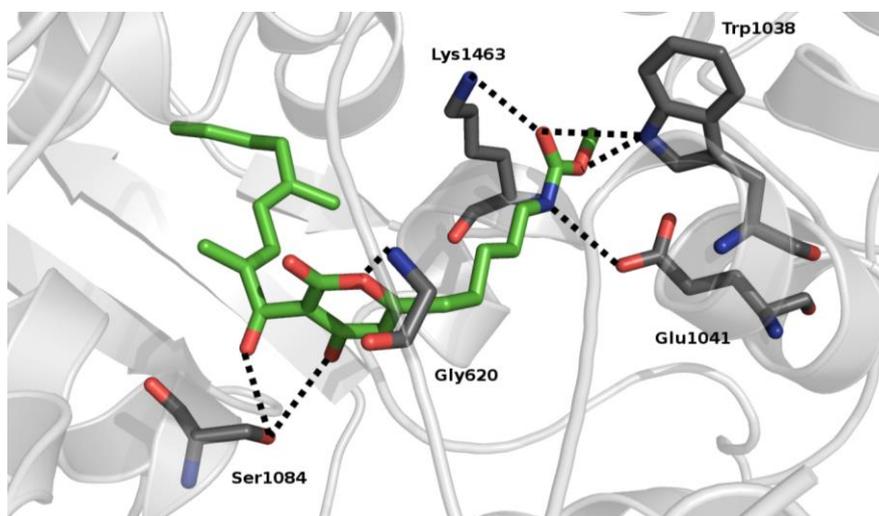
The myxopyronins **1.25-1.27**<sup>95</sup> (Figure 1.20), produced by the Myxobacterium *Myxococcus fulvus*, are polyketide-derived  $\alpha$ -pyrone antibiotics which inhibit transcription initiation by interacting with the RNAP 'switch region' and preventing the formation of a catalytically competent RNAP-promoter open complex.



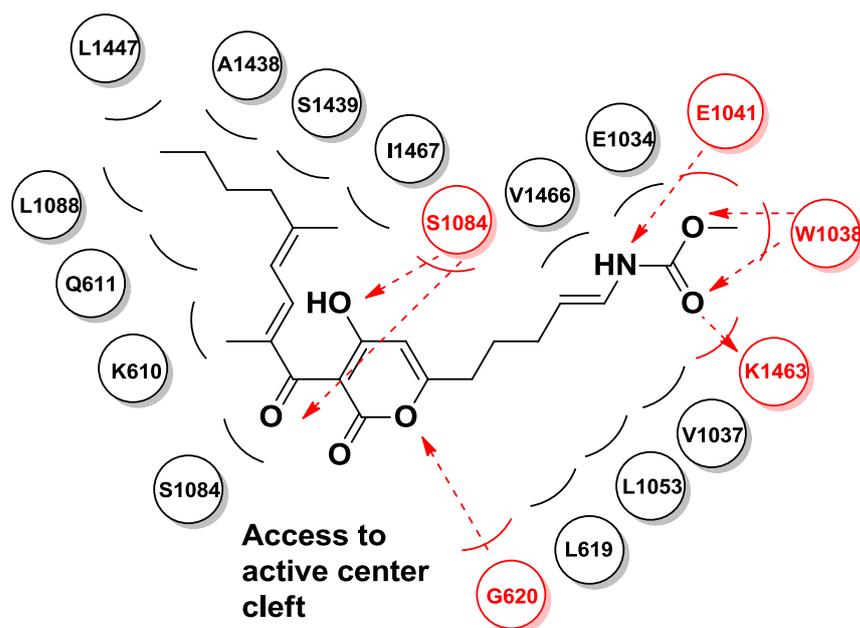
**Figure 1.20** Structures of Myxopyronins.<sup>95</sup>

These inhibitors are active against many Gram-positive and some Gram-negative bacteria<sup>26</sup> with MICs lower than 12.5  $\mu\text{g/mL}$  and inhibited the

transcription and growth of *S. aureus* with an  $IC_{50}$  of 14  $\mu$ M (7-desmethylmyxopyronin B **1.27**) (Figure 1.20).<sup>92,96</sup> Previous studies<sup>97</sup> have shown that these compounds lack *in vivo* efficacy in a mouse model as a consequence of the high affinity to serum albumin whilst no acute toxicity was exhibited at doses of up to at least 100 mg/kg. Two different mechanisms of action were proposed for MyxB **1.26**. Mukhopadhyay<sup>92</sup> *et al.* suggested that MyxB **1.26** prevents the interaction of RNAP with promoter DNA by locking the clamp in a closed conformation thus halting transcription initiation and does not efficiently inhibit the subsequent steps involving interaction of RNAP with promoter DNA. Belogurov<sup>93</sup> *et al.* proposed that MyxB **1.26** stabilizes the refolding of 'switch 2' impeding the accommodation of the melted template DNA near the transcription start site in the RNAP-promoter open complex during transcription initiation. Crystal structures,<sup>92, 93</sup> of myxopyronin A (MyxA) **1.25** and its 7-desmethyl analogue (dMyxB) **1.27** in complex with *T. thermophilus* (Figures 1.21a and 1.21b) have been determined and these reveal that binding occurs predominantly within a nearly completely enclosed hydrophobic crescent-shaped pocket.<sup>98</sup>



**Figure 1.21a** Crystal structure of *T. thermophilus* RNAP-dMyx B **1.27** complex (PDB 3EQL). dMyx B indicated in green.

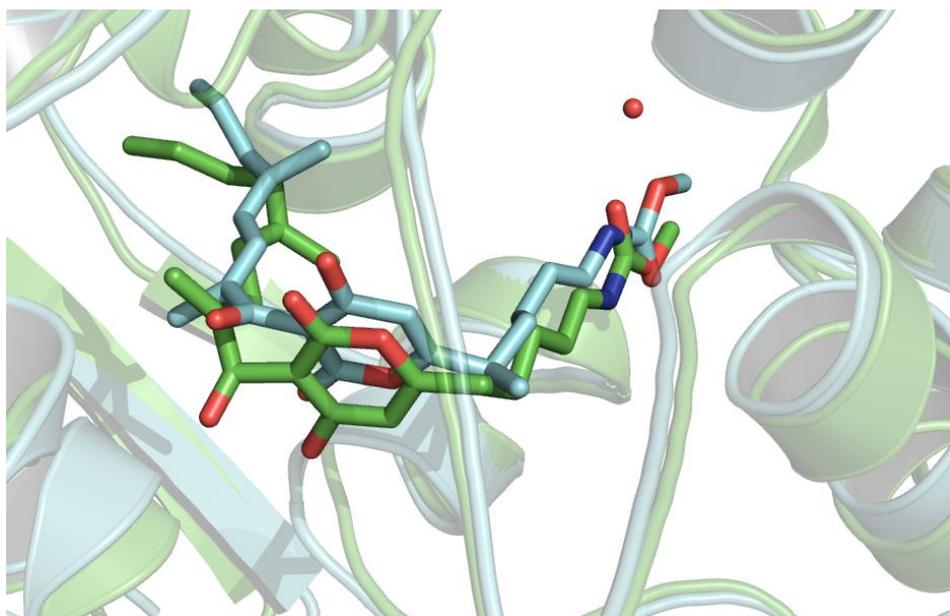


**Figure 1.21b** Schematic diagram of *T. thermophilus* RNAP-dMyx B 1.27 complex (PDB 3EQL). H-bonds in red dashed lines, van der Waals interactions in black.

The  $\alpha$ -pyrone ring is close to the active-centre cleft while the two lipophilic side chains are stretched into two different hydrophobic pockets. Whilst hydrophobic interactions are prevalent in binding, specific hydrogen bond interactions are performed by polar groups on the  $\alpha$ -pyrone ring and ene-carbamate group.

Interestingly, in the co-crystal structure proposed by Mukhopadhyay<sup>92</sup> *et al.* the ene-carbamate moiety is involved in a specific network of hydrogen bonds centred on an ordered water molecule while in the co-crystal structure obtained by Belogurov<sup>93</sup> *et al.* the water molecule is absent because the ene-carbamate moiety is rotated by 180 degrees along the axis perpendicular to the carbonyl group establishing direct hydrogen bond interactions with the nearby amino-acidic residues (Figure 1.22).

Those differences between the two co-crystal structures could be a consequence of the slightly dissimilar structures of MyxA 1.25 and dMyxB 1.27 or could reflect fitting uncertainties in the X-ray co-crystal structures.<sup>98</sup>



**Figure 1.22** Overlay of the co-crystal structure proposed by Mukhopadhyay<sup>92</sup> *et al.* in cyan (PDB 3DXJ) with the co-crystal structure obtained by Belogurov<sup>93</sup> *et al.* in green (PDB 3EQL).

The first total syntheses of MyxA **1.25** and MyxB **1.26** were reported<sup>95</sup> in 1998 and since then several MyxB analogues have been synthesized<sup>99,100</sup> without any knowledge of the structure within the binding site and the majority of them resulted less potent on *E. coli* RNAP with a decreased antibacterial activity when compared to the biological activity of the parent compound. MyxB **1.26** structure resulted very sensitive to structural modification and only a close analogue, dMyxB **1.27** possessed comparable biological activity.<sup>99</sup>

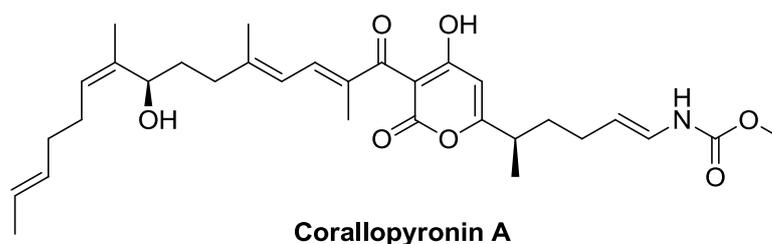
A recent report<sup>101</sup> has evaluated the resistance properties (frequency, spectrum, and fitness costs) of MyxB **1.26** in *Staphylococcus aureus*. The authors<sup>101</sup> located six substitutions in the RNAP  $\beta$  subunit and five in the RNAP  $\beta'$  subunit which conferred resistance to MyxB **1.26** and observed that the resistance rate for MyxB is comparable to the resistance rate for rifampicin **1.11**. Notably, all substitutions conferring MyxB resistance exhibit significant fitness costs (4 to 15% per generation) while at least three substitutions conferring rifampicin resistance exhibit no fitness costs.<sup>101</sup> These observations, together with the previously established inverse correlation between fitness cost and clinical prevalence, lead to the conclusion that MyxB resistance is likely to have

lower clinical prevalence than rifampicin resistance and suggest that the 'switch region' is a viable starting point for drug discovery.<sup>101</sup> The difference in fitness costs of MyxB resistance and rifampicin resistance presumably relates to the fact that the binding site on RNAP for MyxB (the RNAP 'switch region') is responsible for critical functions in opening and closing the RNAP active-centre cleft and in DNA binding whereas the binding site on RNAP for rifampicin performs no critical functions.<sup>101</sup>

### 1.4.8.2 Corallopyronins

The corallopyronins, isolated from *Corallocooccus coralloides*,<sup>97</sup> are  $\alpha$ -pyrone antibiotics structurally related to myxopyronin B (MyxB).<sup>102</sup> Corallopyronin A (CorA) **1.28** (Figure 1.23), has been found to inhibit bacterial RNAP *in vitro* (IC<sub>50</sub> 0.73  $\mu$ M) and is active against Gram-positive bacteria.<sup>102</sup> A recent total synthesis of CorA **1.28** has been reported.<sup>103</sup> Resistant mutants of *S. aureus* to CorA **1.28** were easily selected and the mutations were located in the  $\beta$  and  $\beta'$  subunits of RNAP.<sup>102</sup> Cross-resistance to the related  $\alpha$ -pyrone antibiotic MyxB **1.26** but not with the other known bacterial RNAP inhibitors was observed and therefore it is possible to conclude that CorA shares the same binding site of MyxB<sup>102</sup> within RNAP.

When compared to rifampicin **1.11**, CorA **1.28** has a limited spectrum of antibacterial activity, low activity both as an RNAP inhibitor and as an antibacterial agent and, like rifampicin, a high propensity for selection of resistance.<sup>102</sup> For these reasons, CorA is of limited interest for development as a future drug candidate.<sup>102</sup> However, the eventual synthesis of derivatives of CorA with improved physicochemical features could represent a viable starting point for future antibiotic candidates.<sup>102</sup>



**Corallopyronin A**

**1.28**

**Figure 1.23** Structure of CorA **1.28**.<sup>97</sup>

### 1.4.8.3 Ripostatins

Ripostatins<sup>104,105</sup> A **1.29** and B **1.30** (Figure 1.24) are two polyketide-derived macrocyclic lactones from a strain of *Sporangium cellulosum* with a narrow antibiotic spectrum on *S. aureus* and some *E. coli* strains due to their limited ability to cross membranes. Ripostatins (Rip) are structurally different to MyxB **1.26** and CorA **1.28** but share similar size and hydrophobic character.<sup>92</sup> The high levels of cross resistance with MyxB **1.26** and CorA **1.28** and the location of Rip resistant mutants in the switch region suggest that they probably share the same binding site to that of MyxB **1.26** and CorA **1.28** although no X-ray co-crystal structures have been reported.<sup>92</sup> Ripostatin A **1.29** selectively inhibits the initiation step of bacterial RNA synthesis ( $IC_{50}$  0.8  $\mu$ M), preventing interaction of RNAP with the promoter DNA.<sup>92</sup> Several total synthesis approaches have been reported for these molecules.<sup>106,106b,107,108,109</sup>

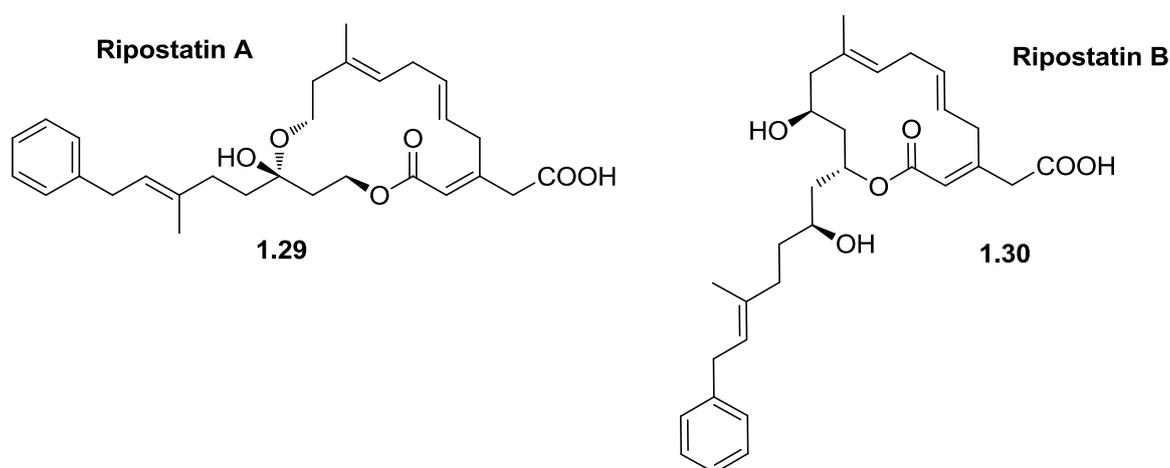


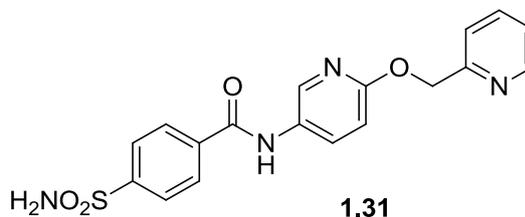
Figure 1.24 Ripostatin A **1.29** and B **1.30** structure.<sup>105</sup>

### 1.4.8.4 Synthetic inhibitors of the 'switch region'

#### 1.4.8.4.1 Pyridyl-Benzamides

A recent structure-based drug design approach (SBDD), and in particular a fragment based drug design (FBDD) strategy using the *de-novo* design SPROUT software, has been applied to the dMyxB **1.27** binding site.<sup>110</sup> Using the co-crystal structure of the *Thermus thermophilus* RNAP Myxopyronin

complex (PDB: 3EQL),<sup>93</sup> a weak *E. coli in vitro* inhibitor **1.31** (IC<sub>50</sub> 151 μM) (Figure 1.25) was identified and then chemically optimized.<sup>110</sup> Subsequently, SAR studies were performed to deliver some selective inhibitors of RNAP in the low micro-molar range but disappointingly none of these compounds exhibited antibacterial activity.<sup>24, 110</sup>

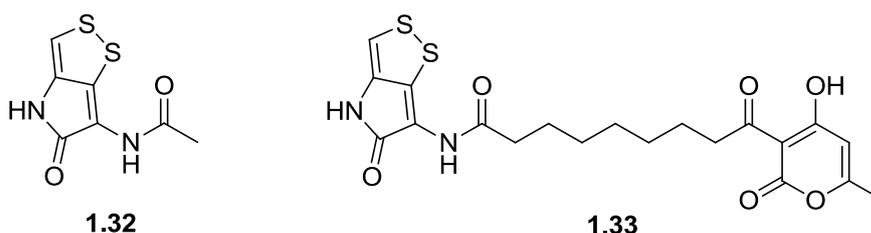


**Figure 1.25** Structure of pyridyl-benzamide **1.31**.<sup>110</sup>

#### 1.4.8.4.2 Myxopyronin based synthetic hybrids

A recent report<sup>111</sup> described the use of a hybrid strategy involving incorporating the antibiotic holomycin **1.32** (Figure 1.26) into a myxopyronin-type skeleton in order to improve antimicrobial activity, reduce lipophilicity and improve stability of the antibiotic myxopyronin A **1.25**. Holomycin **1.32** shows a moderate broad spectrum antibacterial activity against Gram-positive and Gram-negative bacteria, inhibits RNA synthesis and in addition, has a favourable CLogP value (2.0).<sup>111</sup> Two types of hybrid derivatives were designed based on the reported X-ray crystal structure of RNA polymerase bound with MyxA (PDB: 3DXJ)<sup>92</sup> and derivative **1.33** (Figure 1.26) showed moderate antibacterial activity on *B. subtilis* (MIC 16 μg/μL) and *in vitro* inhibition of *E. coli* RNAP (14 μM) whilst it was ineffective in terms of antibacterial activity against *E. coli* due to its low penetration ability in Gram-negative bacteria.<sup>111</sup> Subsequent docking studies on RNAP complex with MyxA (PDB: 3DXJ)<sup>92</sup> predicted the RNAP switch region as a putative binding site for **1.33**.<sup>111</sup>

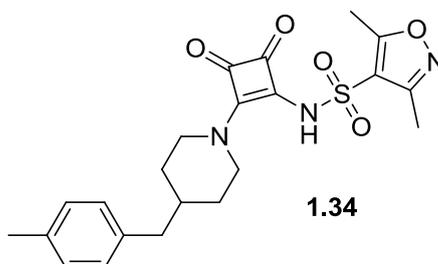
### Holomycin



**Figure 1.26** Structures of holomycin **1.32** and its derivative **1.33**.<sup>111</sup>

#### 1.4.8.4.3 Squaramides

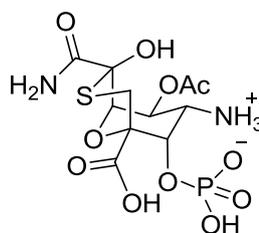
Following high-throughput screening for novel inhibitors of a transcription-coupled translation assay using *E. coli* S30 extracts, a series of *E. coli* RNAP inhibitors with a squaramide core **1.34** (Figure 1.27) were discovered by AstraZeneca.<sup>112</sup> The observation that this series had antibacterial activity against efflux-negative strains of *E. coli* and *H. influenzae* was exploited to show that squaramides exhibit their inhibitory activity within the ‘switch region’ of RNAP *via* the location of resistance mutations and the lack of cross resistance with the other known classes of RNAP inhibitors like rifampicin **1.11**.<sup>112</sup> Docking studies on homology models of the ‘switch region’ bearing the experimentally observed mutations delivered a binding mode consistent with the MIC profile on mutants.<sup>112</sup> Despite the low micro molar activity ( $IC_{50} = 0.3 \mu M$ ) observed with some analogues, no antimicrobial activity was observed against several wild type bacterial strains.<sup>112</sup>



**Figure 1.27** Structure of squaramide **1.34**.<sup>112</sup>

### 1.4.9 Tagetitoxin

The phytotoxin Tagetitoxin **1.35**<sup>113,114,115</sup> (Figure 1.28) produced by *Pseudomonas syringae* pv. *Tagetis* is a non-selective RNAP inhibitor due to its evolutionary conserved binding site. It was co-crystallized with the *T. thermophilus* RNAP<sup>116</sup> which revealed that tagetitoxin **1.35** binds to the base of the RNAP secondary channel, close to the active site and interacts with the incoming NTPs by altering the substrate loading or by stabilizing the enzymatic complex in an inactive state. Tagetitoxin **1.35** does not inhibit the growth of any bacteria despite having activity with a broad range of RNAP from different species.<sup>117</sup>



**Tagetitoxin**

**1.35**

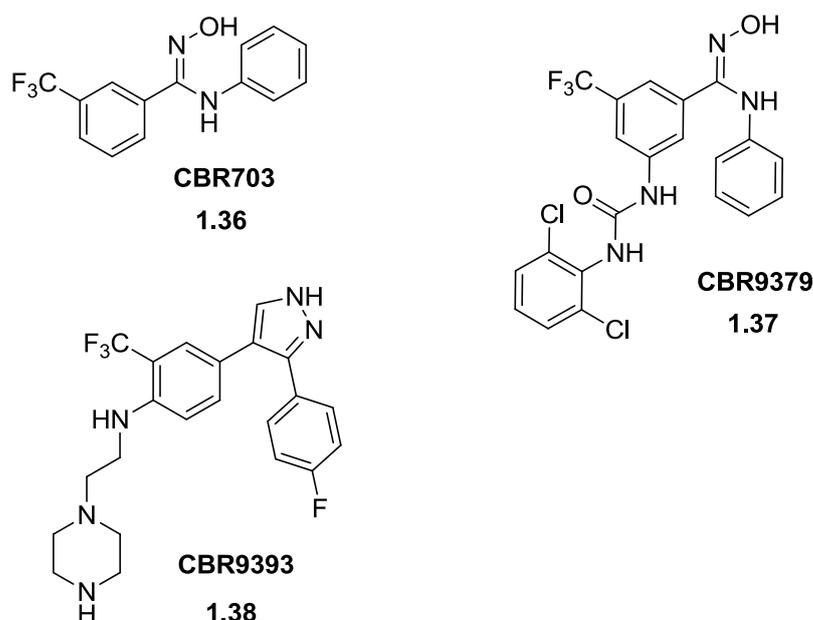
**Figure 1.28** Tagetitoxin **1.35** structure.<sup>117</sup>

### 1.4.10 Other synthetic RNAP inhibitors

#### 1.4.10.1 CBR703 and analogues

CBR703 **1.36**<sup>118</sup> (Figure 1.29) is the progenitor of a synthetic class of selective bacterial RNAP inhibitors identified by Cumbre Pharmaceuticals from a high-throughput screen of a library of ~300,000 compounds with the aim of inhibiting *E. coli* RNAP. The hit molecule, CBR703, showed selective inhibition of *E. coli* RNAP with an IC<sub>50</sub> value of 10 μM but did not show any antibacterial activity against Gram-positive or Gram-negative species. However, the

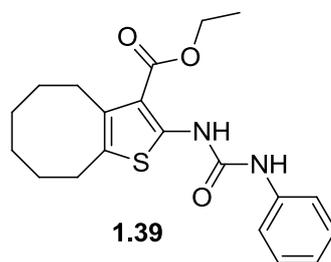
observation that CBR703 possessed antibacterial activity against *E. coli tolC* mutant strains led to the conclusion that the compound is specifically effluxed out from the cell. Synthetic variants of CBR703, like CBR9379 **1.37** and CBR9393 **1.38** (Figure 1.29), led to improved potency against the enzymatic target with no significantly improved antibiotic activity. Following studies on mutations which confer resistance, it has been proposed that these inhibitors bind to a surface-exposed groove at the junction of the  $\beta'$  bridge helix and the  $\beta$  subunit preventing the  $\beta'$  bridge helix to adopt the correct conformation for nucleotide addition by allosterically altering its conformation. A systematic SAR exploration has been performed by Cumbre Pharmaceuticals and reported in a patent<sup>119</sup> in 2002. Further development on this class was recently attempted by Zhu *et al.*<sup>120</sup> by synthesizing thirty analogues but the compounds which showed antibacterial activity correlated with a significant cytotoxicity toward HEK293 cells. Furthermore, the reported effects on biofilm formation by a previous study,<sup>121</sup> which were among the main reasons for synthesizing CBR703 analogues, were suspected to be artefacts due to compound precipitation. Finally, these authors<sup>120</sup> concluded that this class of compounds is unattractive for development as antibacterial agents.



**Figure 1.29** CBR703 **1.36**, CBR9379 **1.37** and CBR9393 **1.38** structures.<sup>118</sup>

### 1.4.10.2 Ureidothiophenes

Ureidothiophenes<sup>122</sup> are a class of bacterial RNAP inhibitors identified following high-throughput screening of a commercial library of compounds against *S. aureus* RNAP holoenzyme. The progenitor of this synthetic class is compound **1.39** (Figure 1.30), which displayed low micro-molar IC<sub>50</sub> value and potent selective antibacterial activity against some *S. aureus* strains but not on Gram-negative strains.<sup>122</sup>



**Figure 1.30** Structure of **1.39**.<sup>122</sup>

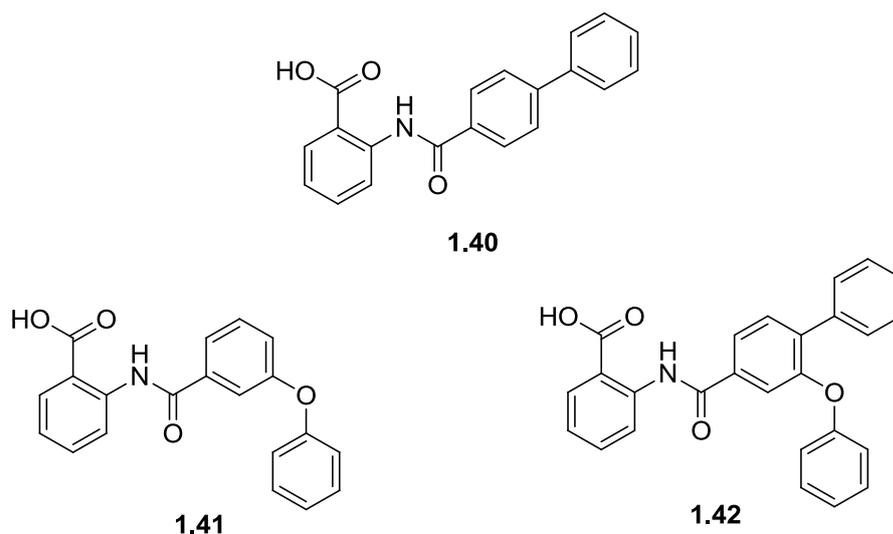
Antibacterial activity was abolished upon addition of human and mouse serum to the growth medium suggesting that the hydrophobic nature of these compounds results in poor pharmacokinetic properties.<sup>122</sup> The SAR exploration led to many analogues with *in vitro* nanomolar potency but very stringent structural requirements were needed to retain antibacterial activity which were not compatible with the need for derivatives with improved pharmacokinetic properties.<sup>122</sup> Ureidothiophenes showed good antibacterial activity against rifampicin-resistant strains of *S. aureus*, suggesting an alternate mechanism of action or binding site on the RNAP holoenzyme.<sup>122</sup> Macromolecular synthesis assays<sup>122</sup> concluded that the mode of action of these compounds is the inhibition of both RNA and protein biosynthesis. Studies conducted by our collaborators (unpublished data<sup>123</sup>) showed a preference for the inhibition of protein biosynthesis, suggesting off-target inhibition in the cell. We concluded that this scaffold is not a viable starting point for further development of potential RNAP inhibitors.

### 1.4.11 Peptides as protein-protein interaction inhibitors

A recent study,<sup>124</sup> considered the interface between the *E. coli* RNAP core enzyme and  $\sigma^{70}$  factor as a potential binding site for inhibiting the assembling of the RNAP holoenzyme which is essential for transcription inhibition. The lack of homology between the bacterial  $\sigma^{70}$  subunit and the mammalian homologues permits selective inhibition of bacterial RNAP with a broad spectrum of antibiotic activity due to its essential role in transcription.<sup>124</sup> A peptide approach was undertaken for generating new inhibitors of this crucial protein-protein interaction and sixteen peptides covering different regions of *E. coli* core enzyme and the  $\sigma^{70}$  interface were designed and some of them, derived from the  $\sigma^{70}$  2.2 region, showed strong RNAP inhibition with IC<sub>50</sub> values in the low micromolar range.<sup>124</sup> ELISA-based binding experiments and the observation that transcription initiation was inhibited strongly supported the  $\sigma^{70}$ -core enzyme interface as a target site.<sup>124</sup> These data could lead to the generation of novel RNAP inhibitors targeting RNAP with a new mode of action.<sup>124</sup>

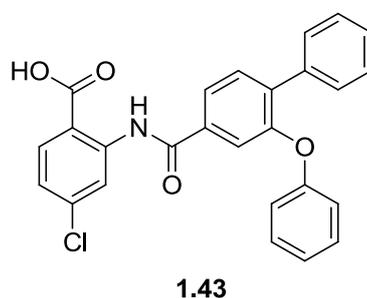
### 1.4.12 Small molecules as protein-protein interaction inhibitors<sup>125</sup>

Following the flexible alignment of structurally similar selected synthetic molecules that were known to inhibit bacterial RNAP, a pharmacophore model was developed by Hinsberger *et al.*<sup>125</sup> Despite the fact there were no empirical data confirming those molecules share the same binding region, this pharmacophore model was subsequently used to perform virtual screening. Three hit compounds bearing an anthranilic acid core, **1.40**, **1.41** and **1.42** (Figure 1.31), were identified.



**Figure 1.31** Structures of **1.40**, **1.41** and **1.42**.

Hit compounds were then chemically optimised yielding novel derivatives with improved in vitro potency. In particular, the most potent compound, **1.43** (Figure 1.32), showed an  $IC_{50}$  of 9  $\mu$ M on *E. coli* RNAP and fair antibacterial activity on Gram-positive strains (*S. aureus* MIC = 6  $\mu$ g / mL). The authors<sup>125</sup> demonstrated that the new inhibitors acted *via* a dual target effect *via* preventing the protein-protein interaction between  $\sigma^{70}$  and the RNAP core enzyme and inhibiting bacterial lipid biosynthesis. This dual target effect could account for the relatively low MIC values and the significantly lower resistance rate compared to the clinically used inhibitor Rifampicin **1.11**.



**Figure 1.32** Structure of **1.43**.

## 1.5 Thesis aims and objectives

The work presented in this thesis describes the application of *in silico* techniques to design rationally novel inhibitors of bacterial RNAP targeting the myxopyronin binding region of the enzyme. An iterative process has been devised which comprised *in silico* techniques to identify putative inhibitors, chemical synthesis and biological evaluation.

Within these categories, the aims and objectives of the project are:

### **Objective 1: To use *in silico* techniques to identify putative RNAP inhibitors**

The aims under this objective are:

- a) To learn and understand the principles of ligand and structure-based virtual screening including all the computational techniques required such as docking.
- b) To validate the software and the computational protocols.
- c) To formulate hypotheses for rational design on the basis of computational predictions and refine them after biological evaluation results.

### **Objective 2: To use synthetic chemistry to produce putative inhibitors**

The aims under this objective are:

- a) To apply modern organic synthesis techniques to prepare compound libraries corresponding the *in silico* designed molecular scaffolds.
- b) To conduct Structure Activity Relationship exploration and analysis using the synthesized analogues in an iterative process together with the *in silico* techniques.

- c) To ensure a high level of purity for the synthesized compounds to enable biological evaluation.
- d) To ensure full chemical characterization for all synthesized compounds to publication standard.

**Objective 3: To subject the synthesised molecules to preliminary biological evaluation**

The aims under this objective are:

- a) To measure *in vitro* percentage inhibition against *E. coli* RNAP at 100µM and then to select compounds with more than 50% inhibition for chemical optimization.
- b) To determine in parallel, for the most potent analogues, IC<sub>50</sub> values, MICs against a panel of representative Gram-positive and negative bacterial strains, and the potential for promiscuous inhibition on unrelated enzymes such as malate dehydrogenase and chymotrypsin.
- c) To progress promising compounds to further biological evaluation including membrane integrity assessment and specificity for inhibition of RNA biosynthesis.

## 1.6 References

1. Kohanski, M. A.; Dwyer, D. J.; Collins, J. J., *Nat. Rev. Micro.* **2010**, *8*, 423-435.
2. Drlica, K.; Malik, M.; Kerns, R. J.; Zhao, X., *Antimicrob. Agents Chemother.* **2008**, *52*, 385-392.
3. Floss, H. G.; Yu, T.-W., *Chem. Rev.* **2005**, *105*, 621-632.
4. Tomasz, A., *Annu. Rev. Microbiol.* **1979**, *33*, 113-137.
5. Vakulenko, S. B.; Mobashery, S., *Clin. Microbiol. Rev.* **2003**, *16*, 430-450.
6. Kohanski, M. A.; Dwyer, D. J.; Hayete, B.; Lawrence, C. A.; Collins, J. J., *Cell* **2007**, *130*, 797-810.
7. Dwyer, D. J.; Kohanski, M. A.; Hayete, B.; Collins, J. J., *Mol. Syst. Biol.* **2007**, *3*, 1-15.
8. Kohanski, M. A.; Dwyer, D. J.; Wierzbowski, J.; Cottarel, G.; Collins, J. J., *Cell* **2008**, *135*, 679-690.
9. Taubes, G., *Science* **2008**, *321*, 356-361.
10. Yoneyama, H.; Katsumata, R., *Biosci. Biotech. Bioch.* **2006**, *70*, 1060-1075.
11. Schmidt, F. R., *Appl. Microbiol. Biot.* **2004**, *63*, 335-343.
12. Woodford, N.; Livermore, D. M., *J. Infection* **2009**, *59*, S4-S16.
13. Donadio, S.; Maffioli, S.; Monciardini, P.; Sosio, M.; Jabes, D., *J. Antibiot.* **2010**, *63*, 423-430.
14. Trinh, V.; Langelier, M.-F.; Archambault, J.; Coulombe, B., *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 12-36.
15. Ebrahim, G. J., *J. Trop. Pediatrics* **2010**, *56*, 141-143.
16. Silver, L. L., *Clin. Microbiol. Rev.* **2011**, *24*, 71-109.
17. Livermore, D. M., *Clin. Infect. Dis.* **2003**, *36*, S11-S23.
18. Livermore, D. M.; British Soc, A., *J. Antimicrob. Chemoth.* **2011**, *66*, 1941-1944.
19. O'Shea, R.; Moser, H. E., *J. Med. Chem.* **2008**, *51*, 2871-2878.
20. Spellberg, B.; Powers, J. H.; Brass, E. P.; Miller, L. G.; Edwards, J. E., *Clin. Infect. Dis.* **2004**, *38*, 1279-1286.
21. Butler, M. S.; Cooper, M. A., *J. Antibiot.* **2011**, *64*, 413-425.

22. Walsh, C., *Nature* **2000**, 406, 775-781.
23. Singleton, P., *Bacteria in Biology, Biotechnology and Medicine*. 5th ed.; Wiley: 1999.
24. Chopra, I., *J. Antimicrob. Chemoth.* **2013**, 68, 496-505.
25. Miethke, M.; Marahiel, M. A., *Microbiol. Mol. Biol. Rev.* **2007**, 71, 413-451.
26. Mariani, R.; Maffioli, S. I., *Curr. Med. Chem.* **2009**, 16, 430-454.
27. Domecq, C.; Trinh, V.; Langelier, M. F.; Archambault, J.; Coulombe, B., *Curr. Chem. Biol.* **2008**, 2, 20-31.
28. Darst, S. A., *Trends Biochem. Sci.* **2004**, 29, 159-162.
29. Mooney, R. A.; Landick, R., *Cell* **1999**, 98, 687-690.
30. Sensi, P.; Margalith, P.; Timbal, M. T., *Farmaco, Ed. Sci.* **1959**, 14, 146-147.
31. Wehrli, W.; Staehelin, M., *Bacteriol. Rev.* **1971**, 35, 290-309.
32. Pukrittayakamee, S.; Viravan, C.; Charoenlarp, P.; Yeamput, C.; Wilson, R. J. M.; White, N. J., *Antimicrob. Agents Chemother.* **1994**, 38, 511-514.
33. Gardner, M. J.; Williamson, D. H.; Wilson, R. J. M., *Mol. Biochem. Parasitol.* **1991**, 44, 115-123.
34. Tupin, A.; Gualtieri, M.; Roquet-Banères, F.; Morichaud, Z.; Brodolin, K.; Leonetti, J.-P., *Int. J. Antimicrob. Ag.* **2010**, 35, 519-523.
35. Sensi, P.; Ballotta, R.; Greco, A. M.; Gallo, G. G., *Farmaco, Ed. Sci.* **1961**, 16, 165-180.
36. Maggi, N.; Pasqualucci, C. R.; Ballotta, R.; Sensi, P., *Chemotherapia* **1966**, 11, 285-292.
37. Sensi, P.; Thiemann, J. E., *Prog. Ind. Microbiol.* **1967**, 6, 21-59.
38. Bergamini, N.; Fowst, G., *Arzneim.-Forsch.* **1965**, 15, 951-1002.
39. Roos, R.; Ghisalba, O., *Experientia* **1980**, 36, 479-486.
40. Sensi, P., *Rev. Infect. Dis.* **1983**, 5, S402-S406.
41. Rinehart, K. L.; Jr, n., *Fortschr. Chem. Org. Naturst.* **1976**, 33, 231-307.
42. Traxler, P.; Vischer, V. A.; Zak, O., *Drugs Future* **1988**, 13, 845-856.
43. Lancini, G.; Zanichelli, W., *Structure activity relationships among the semisynthetic antibiotics*. Perlman, D. ed.; Academic Press, New York: 1977.

44. Binda, G.; Domenichini, E.; Gottardi, A.; Orlandi, B.; Ortelli, E.; Pacini, B.; Fowst, G., *Arzneim.-Forsch.* **1971**, *21*, 1907-1977.
45. Aristoff, P. A.; Garcia, G. A.; Kirchhoff, P. D.; Showalter, H. D. H., *Tuberculosis* **2010**, *90*, 94-118.
46. Calvori, C.; Frontali, L.; Leoni, L.; Tecce, G., *Nature* **1965**, *207*, 417-418.
47. Hartmann, G.; Honikel, K. O.; Knüsel, F.; Nüesch, J., *Biochim. Biophys. Acta* **1967**, *145*, 843-844.
48. Campbell, E. A.; Khorzheva, N.; Mustaev, A.; Murakami, K.; Nair, S.; Goldfarb, A.; Darst, S. A., *Cell* **2001**, *104*, 901-912.
49. Zhang, G.; Campbell, E. A.; Minakhin, L.; Richter, C.; Severinov, K.; Darst, S. A., *Cell* **1999**, *98*, 811-824.
50. Artsimovitch, I.; Vassilyeva, M. N.; Svetlov, D.; Svetlov, V.; Perederina, A.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Tahirov, T. H.; Vassilyev, D. G., *Cell* **2005**, *122*, 351-363.
51. Artsimovitch, I.; Vassilyev, D. G., *Cell Cycle* **2006**, *5*, 399-404.
52. Feklistov, A.; Mekler, V.; Jiang, Q. R.; Westblade, L. F.; Irschik, H.; Jansen, R.; Mustaev, A.; Darst, S. A.; Ebright, R. H., *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 14820-14825.
53. Brufani, M.; Cecchini, G.; Cellai, L.; Federici, M.; Guiso, M.; Segre, A., *J. Antibiot.* **1985**, *38*, 259-262.
54. Brufani, M.; Cellai, L.; Cozzella, L.; Federici, M.; Guiso, M.; Segre, A., *J. Antibiot.* **1985**, *38*, 1359-1362.
55. Brizzi, V.; Brufani, M.; Cellai, L.; Segre, A. L., *J. Antibiot.* **1983**, *36*, 516-521.
56. Cellai, L.; Cerrini, S.; Segre, A.; Battistoni, C.; Cossu, G.; Mattogno, G.; Brufani, M.; Marchi, E., *Mol. Pharmacol.* **1985**, *27*, 103-108.
57. Cavalleri, B.; Turconi, M.; Tamborini, G.; Ocelli, E.; Cietto, G.; Pallanza, R.; Scotti, R.; Berti, M.; Romano, G.; Parenti, F., *J. Med. Chem.* **1990**, *33*, 1470-1476.
58. Dellabruna, C.; Schioppacassi, G.; Ungheri, D.; Jabes, D.; Morvillo, E.; Sanfilippo, A., *J. Antibiot.* **1983**, *36*, 1502-1506.
59. Cellai, L.; Cerrini, S.; Segre, A. L.; Battistoni, C.; Cossu, G.; Mattogno, G.; Brufani, M.; Marchi, E.; Venturini, A. P., *Farmaco* **1989**, *44*, 97-107.

60. Marchi, E.; Montecchi, L.; Venturini, A. P.; Mascellani, G.; Brufani, M.; Cellai, L., *J. Med. Chem.* **1985**, *28*, 960-963.
61. Adachi, J. A.; DuPont, H. L., *Clin. Infect. Dis.* **2006**, *42*, 541-547.
62. Brogden, R. N.; Fitton, A., *Drugs* **1994**, *47*, 983-1009.
63. Jarvis, B.; Lamb, H. M., *Drugs* **1998**, *56*, 607-616.
64. Barluenga, J.; Aznar, F.; Garcia, A. B.; Cabal, M. P.; Palacios, J. J.; Menendez, M. A., *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5717-5722.
65. Kim, I. H.; Combrink, K. D.; Ma, Z.; Chapo, K.; Yan, D.; Renick, P.; Morris, T. W.; Pulse, M.; Simecka, J. W.; Ding, C. Z., *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1181-1184.
66. Li, J.; Ma, Z.; Chapo, K.; Yan, D.; Lynch, A. S.; Ding, C. Z., *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5510-5513.
67. Combrink, K. D.; Denton, D. A.; Harran, S.; Ma, Z.; Chapo, K.; Yan, D.; Bonventre, E.; Roche, E. D.; Doyle, T. B.; Robertson, G. T.; Lynch, A. S., *Bioorg. Med. Chem. Lett.* **2007**, *17*, 522-526.
68. Severinov, K.; Soushko, M.; Goldfarb, A.; Nikiforov, V., *Mol. Gen. Genet.* **1994**, *244*, 120-126.
69. Telenti, A.; Imboden, P.; Marchesi, F.; Lowrie, D.; Cole, S.; Colston, M. J.; Matter, L.; Schopfer, K.; Bodmer, T., *Abstr. Gen. Meet. Am. Soc. Microbiol.* **1993**, *93*, 171.
70. Ramaswamy, S.; Musser, J. M., *Tuberc. Lung Dis.* **1998**, *79*, 3-29.
71. Heep, M.; Brandstatter, B.; Rieger, U.; Lehn, N.; Richter, E.; Rusch-Gerdes, S.; Niemann, S., *J. Clin. Microbiol.* **2001**, *39*, 107-110.
72. Ishikawa, J.; Chiba, K.; Kurita, H.; Satoh, H., *Antimicrob. Agents Ch.* **2006**, *50*, 1342-1346.
73. Fernandes, P.; Ferreira, B. S.; Cabral, J. M. S., *Int. J. Antimicrob. Ag.* **2003**, *22*, 211-216.
74. Parenti, F.; Pagani, H.; Beretta, G., *J. Antibiot.* **1975**, *28*, 247-252.
75. Coronelli, C.; White, R. J.; Lancini, G. C.; Parenti, F., *J. Antibiot.* **1975**, *28*, 253-259.
76. Johnson, A. P., *Curr. Opin. Invest. Dr.* **2007**, *8*, 168-173.
77. Tupin, A.; Gualtieri, M.; Leonetti, J.-P.; Brodolin, K., *Embo J.* **2010**, *29*, 2527-2537.

78. Artsimovitch, I.; Seddon, J.; Sears, P., *Clin. Infect. Dis.* **2012**, *55*, S127-S131.
79. Sarubbi, E.; Monti, F.; Corti, E.; Miele, A.; Selva, E., *Eur. J. Biochem.* **2004**, *271*, 3146-3154.
80. Ciciliato, I.; Corti, E.; Sarubbia, E.; Stefanelli, S.; Gastaldo, L.; Montanini, N.; Kurz, M.; Losi, D.; Marinelli, F.; Selva, E., *J. Antibiot.* **2004**, *57*, 210-217.
81. Mariani, R.; Granata, G.; Maffioli, S. I.; Serina, S.; Brunati, C.; Sosio, M.; Marazzi, A.; Vannini, A.; Patel, D.; White, R.; Ciabatti, R., *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3748-3752.
82. Zhang, Y.; Degen, D.; Ho, M. X.; Sineva, E.; Ebright, K. Y.; Ebright, Y. W.; Mekler, V.; Vahedian-Movahed, H.; Feng, Y.; Yin, R.; Tuske, S.; Irschik, H.; Jansen, R.; Maffioli, S.; Donadio, S.; Arnold, E.; Ebright, R. H., *Elife* **2014**, *3*, 1-31.
83. Temiakov, D.; Zenkin, N.; Vassylyeva, M. N.; Perederina, A.; Tahirov, T. H.; Kashkina, E.; Savkina, M.; Zorov, S.; Nikiforov, V.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Severinov, K.; Vassylyev, D. G., *Mol. Cell* **2005**, *19*, 655-666.
84. Tuske, S.; Sarafianos, S. G.; Wang, X. Y.; Hudson, B.; Sineva, E.; Mukhopadhyay, J.; Birktoft, J. J.; Leroy, O.; Ismail, S.; Clark, A. D.; Dharia, C.; Napoli, A.; Laptenko, O.; Lee, J.; Borukhov, S.; Ebright, R. H.; Arnold, E., *Cell* **2005**, *122*, 541-552.
85. Vassylyev, D. G.; Sekine, S.; Laptenko, O.; Lee, J.; Vassylyeva, M. N.; Borukhov, S.; Yokoyama, S., *Nature* **2002**, *417*, 712-719.
86. Kyzer, S.; Zhang, J. W.; Landick, R., *Cell* **2005**, *122*, 494-496.
87. Jansen, R.; Wray, V.; Irschik, H.; Reichenbach, H.; Hofle, G., *Tetrahedron Lett.* **1985**, *26*, 6031-6034.
88. Campbell, E. A.; Pavlova, O.; Zenkin, N.; Leon, F.; Irschik, H.; Jansen, R.; Severinov, K.; Darst, S. A., *Embo J.* **2005**, *24*, 674-682.
89. Blond, A.; Cheminant, M.; Segalas-Milazzo, I.; Peduzzi, J.; Barthelemy, M.; Goulard, C.; Salomon, R.; Moreno, F.; Farias, R.; Rebuffat, S., *Eur. J. Biochem.* **2001**, *268*, 2124-2133.

90. Wilson, K. A.; Kalkum, M.; Ottesen, J.; Yuzenkova, J.; Chait, B. T.; Landick, R.; Muir, T.; Severinov, K.; Darst, S. A., *J. Am. Chem. Soc.* **2003**, *125*, 12475-12483.
91. Kuznedelov, K.; Semenova, E.; Knappe, T. A.; Mukhamedyarov, D.; Srivastava, A.; Chatterjee, S.; Ebright, R. H.; Marahiel, M. A.; Severinov, K., *J. Mol. Biol.* **2011**, *412*, 842-848.
92. Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M. Y.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H., *Cell* **2008**, *135*, 295-307.
93. Belogurov, G. A.; Vassilyeva, M. N.; Sevostyanova, A.; Appleman, J. R.; Xiang, A. X.; Lira, R.; Webber, S. E.; Klyuyev, S.; Nudler, E.; Artsimovitch, I.; Vassilyev, D. G., *Nature* **2009**, *457*, 332-335.
94. Srivastava, A.; Talaue, M.; Liu, S.; Degen, D.; Ebright, R. Y.; Sineva, E.; Chakraborty, A.; Druzhinin, S. Y.; Chatterjee, S.; Mukhopadhyay, J.; Ebright, Y. W.; Zozula, A.; Shen, J.; Sengupta, S.; Niedfeldt, R. R.; Xin, C.; Kaneko, T.; Irschik, H.; Jansen, R.; Donadio, S.; Connell, N.; Ebright, R. H., *Curr. Opin. Microbiol.* **2011**, *14*, 532-543.
95. Hu, T.; Schaus, J. V.; Lam, K.; Palfreyman, M. G.; Wuonola, M.; Gustafson, G.; Panek, J. S., *J. Org. Chem.* **1998**, *63*, 2401-2406.
96. Moy, T. I.; Daniel, A.; Hardy, C.; Jackson, A.; Rehrauer, O.; Hwang, Y. S.; Zou, D.; Kien, N.; Silverman, J. A.; Li, Q.; Murphy, C., *Fems Microbiol. Lett.* **2011**, *319*, 176-179.
97. Irschik, H.; Jansen, R.; Hofle, G.; Gerth, K.; Reichenbach, H., *J. Antibiot.* **1985**, *38*, 145-152.
98. Ho, M. X.; Hudson, B. P.; Das, K.; Arnold, E.; Ebright, R. H., *Curr. Opin. Struc. Biol.* **2009**, *19*, 715-723.
99. Doundoulakis, T.; Xiang, A. X.; Lira, R.; Agrios, K. A.; Webber, S. E.; Sisson, W.; Aust, R. M.; Shah, A. M.; Showalter, R. E.; Appleman, J. R.; Simonsen, K. B., *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5667-5672.
100. Lira, R.; Xiang, A. X.; Doundoulakis, T.; Biller, W. T.; Agrios, K. A.; Simonsen, K. B.; Webber, S. E.; Sisson, W.; Aust, R. M.; Shah, A. M.; Showalter, R. E.; Banh, V. N.; Steffy, K. R.; Appleman, J. R., *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6797-6800.

101. Srivastava, A.; Degen, D.; Ebricht, Y. W.; Ebricht, R. H., *Antimicrob. Agents Ch.* **2012**, *56*, 6250-6255.
102. Mariner, K.; McPhillie, M.; Trowbridge, R.; Smith, C.; O'Neill, A. J.; Fishwick, C. W. G.; Chopra, I., *Antimicrob. Agents Ch.* **2011**, *55*, 2413-2416.
103. Rentsch, A.; Kalesse, M., *Angew. Chem. Int. Edit.* **2012**, *51*, 11381-11384.
104. Irschik, H.; Augustiniak, H.; Gerth, K.; Hofle, G.; Reichenbach, H., *J. Antibiot.* **1995**, *48*, 787-792.
105. O'Neill, A.; Oliva, B.; Storey, C.; Hoyle, A.; Fishwick, C.; Chopra, I., *Antimicrob. Agents Ch.* **2000**, *44*, 3163-3166.
106. (a) Glaus, F.; Altmann, K.-H., *Angew. Chem. Int. Edit.* **2012**, *51*, 3405-3409; (b) Tang, W.; Prusov, E. V., *Org. Lett.* **2012**, *14*, 4690-4693.
107. Tang, W.; Prusov, E. V., *Angew. Chem. Int. Edit.* **2012**, *51*, 3401-3404.
108. Winter, P.; Hiller, W.; Christmann, M., *Angew. Chem. Int. Edit.* **2012**, *51*, 3396-3400.
109. Kujat, C.; Bock, M.; Kirschning, A., *Synlett* **2006**, *3*, 419-422.
110. McPhillie, M. J.; Trowbridge, R.; Mariner, K. R.; O'Neill, A. J.; Johnson, A. P.; Chopra, I.; Fishwick, C. W. G., *ACS Med. Chem. Lett.* **2011**, *2*, 729-734.
111. Yakushiji, F.; Miyamoto, Y.; Kunoh, Y.; Okamoto, R.; Nakaminami, H.; Yamazaki, Y.; Noguchi, N.; Hayashi, Y., *ACS Med. Chem. Lett.* **2013**, *4*, 220-224.
112. Buurman, E. T.; Foulk, M. A.; Gao, N.; Laganas, V. A.; McKinney, D. C.; Moustakas, D. T.; Rose, J. A.; Shapiro, A. B.; Fleming, P. R., *J. Bacteriol.* **2012**, *194*, 5504-5512.
113. Mitchell, R. E.; Durbin, R. D., *Physiol. Plant Pathol.* **1981**, *18*, 157-168.
114. Mathews, D. E.; Durbin, R. D., *J. Biol. Chem.* **1990**, *265*, 493-498.
115. Steinberg, T. H.; Mathews, D. E.; Durbin, R. D.; Burgess, R. R., *J. Biol. Chem.* **1990**, *265*, 499-505.
116. Vassylyev, D. G.; Svetlov, V.; Vassylyeva, M. N.; Perederina, A.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Artsimovitch, I., *Nat. Struct. Mol. Biol.* **2005**, *12*, 1086-1093.

117. Villain-Guillot, P.; Bastide, L.; Gualtieri, M.; Leonetti, J. P., *Drug Discov. Today* **2007**, *12*, 200-208.
118. Artsimovitch, I.; Chu, C.; Lynch, A. S.; Landick, R., *Science* **2003**, *302*, 650-654.
119. Li, L.; Mihalic, J. T., Antibacterial agents. US20020045749, 2002.
120. Zhu, W.; Haupenthal, J.; Groh, M.; Fountain, M.; Hartmann, R. W., *Antimicrob. Agents Chemother.* **2014**, *58*, 4242-4245.
121. Villain-Guillot, P.; Gualtieri, M.; Bastide, L.; Leonetti, J.-P., *Antimicrob. Agents Chemother.* **2007**, *51*, 3117-3121.
122. Arhin, F.; Belanger, O.; Ciblat, S.; Dehbi, M.; Delorme, D.; Dietrich, E.; Dixit, D.; Lafontaine, Y.; Lehoux, D.; Liu, J.; McKay, G. A.; Moeck, G.; Reddy, R.; Rose, Y.; Srikumar, R.; Tanaka, K. S. E.; Williams, D. M.; Gros, P.; Pelletier, J.; Parr, T. R., Jr.; Far, A. R., *Bioorg. Med. Chem.* **2006**, *14*, 5812-5832.
123. Chopra, I. *et. al.*, Unpublished data.
124. Hüsecken, K.; Negri, M.; Fruth, M.; Boettcher, S.; Hartmann, R. W.; Haupenthal, J., *ACS Chem. Biol.* **2013**, *8*, 758-766.
125. Hinsberger, S.; Huesecken, K.; Groh, M.; Negri, M.; Haupenthal, J.; Hartmann, R. W., *J. Med. Chem.* **2013**, *56*, 8332-8338.

## 2. Computational Methodologies

### 2.1 Virtual Screening approaches

High-throughput screening (HTS) of chemical libraries is a well-established method in drug discovery<sup>1</sup> but its main drawback is represented by the cost of randomly screening increasingly large databases.<sup>2</sup> Virtual screening (VS) is a computational technique capable of selecting the most promising compounds from an electronic chemical database for experimental screening.<sup>3</sup> VS approaches can be both ligand-based and structure-based.<sup>4,5</sup> Structure-based virtual screening (SBVS) involves the docking of molecules within the active site of target proteins for hit identification and is a complementary methodology to HTS allowing a higher hit rate and, at the same time, reducing the cost and duration of a screening campaign.<sup>6</sup> A SBVS protocol can be summarized in three main steps: 1) database preparation (filtering, protonation states, charges, most stable tautomers), 2) docking, 3) post-docking analysis (clustering of poses, consensus scoring,<sup>7</sup> visual inspection).

### 2.2 Docking

The availability of several techniques (X-ray crystallography, NMR, homology modeling) providing structural information about biological targets has opened up a vast field of drug discovery, Structure Based Drug Design (SBDD), which mainly relies in docking techniques.

All docking software are composed of two basic modules: a search algorithm and an evaluating algorithm (scoring function). The search algorithm identifies all possible interactions between the ligand and the biological target while the scoring function estimates the free energy of binding of the ligand poses generated within the binding site. The main uses of docking in drug discovery are:<sup>8,9</sup> 1) predicting the binding mode of a ligand with known biological activity, 2) identification of potential new ligands *via* SBVS, 3) predicting binding affinity of structurally related ligands with the same biological activity.

## 2.2.1 Challenges in docking

The success of docking techniques depends on the correct sampling of a flexible molecule within its biological target *via* the search algorithm and from the precise energetic evaluation of generated poses *via* the scoring function.<sup>8</sup> Any possible error deriving from those two algorithms will generate ‘hard’ and ‘soft’ errors in evaluating the energy of ligand-protein interactions when compared to the empirical data derived from the analysis of co-crystals structures of ligands with proteins.<sup>8</sup> A ‘soft’ error is generated when the predicted binding energy is underestimated while a ‘hard’ error comes from the opposite situation.<sup>8</sup> The consequence of ‘soft’ and ‘hard’ errors are the generation of false negatives and false positives in a virtual screening protocol respectively.

Docking can be visualized as a multi-step process where any consecutive step introduces additional degrees of complexity.<sup>10</sup>

## 2.2.2 Search algorithms

Generating ligand poses is challenging even for small molecules due to the presence of many conformational degrees of freedom. Conformational sampling must be performed rapidly for a large database of molecules and at the same time needs to be accurate.<sup>10</sup> Several strategies for solving this problem are adopted<sup>11</sup> and docking algorithms can be differentiated as stochastic or deterministic and are described in detail in the following Sections for each software evaluated in this thesis. Stochastic algorithms include a random factor and their results are not always reproducible while for deterministic algorithms, reproducible results are always given.<sup>11</sup> The main drawback of deterministic algorithms when compared with stochastic ones is the fact that they can generate a conformation which is trapped in a local minima of the potential energy surface.<sup>11</sup> With regard to protein flexibility, it is not feasible from a computational point of view and several approximations are needed (partial flexibility of selected amino acidic side chains, soft docking potentials,<sup>12</sup> ensemble docking<sup>11,13</sup>).

### 2.2.3 Scoring functions

Scoring functions are used for two purposes:<sup>11</sup> 1) to aid the generation of ligand poses *via* their optimization during conformational sampling, 2) to rank each solution of docked ligands included in a database.

The aim of a scoring function is to give a fast and possibly accurate estimation of the free energy of binding when protein-ligand complexes are generated during docking.<sup>11</sup>

The free energy of binding is expressed by the Gibbs-Helmholtz equation (2.1):

$$\Delta G = \Delta H - T\Delta S \quad (2.1)$$

where  $\Delta G$  represents the free energy of binding,  $\Delta H$  the enthalpy,  $T$  the temperature in Kelvin and  $\Delta S$  the entropy.

Scoring functions can be divided into three main classes:<sup>10</sup>

Force-field based: molecular mechanics force fields are used to quantify the receptor-ligand interaction which is generally the sum of two energies, the receptor-ligand interaction energy and the internal ligand energy. In most cases only a single conformation of the protein is considered in order to omit the calculation of internal protein energy.<sup>10</sup> Van der Waals and electrostatic energy terms are used to describe the interaction between ligand and receptor.<sup>10</sup>

Empirical scoring functions: are a sum of several parameterized functions for reproducing experimental data such as binding energies.<sup>10</sup> Coefficients of various terms are obtained from regression analysis of empirical data.<sup>10</sup>

Knowledge-based scoring functions: are designed to reproduce experimental structures rather than binding energies.<sup>10</sup> They are essentially based on data like defined atomic interaction pair potentials found in limited sets of protein-ligand complex structures.<sup>10</sup>

### 2.2.4 Main limitations of scoring functions

In order to give a fast evaluation of results, most of the scoring functions omit or simplify some of its energetic terms such as calculating enthalpic

contributions but neglecting entropic factors (often limited to loss of torsional entropy of ligands only) and solvation effects.<sup>10</sup> Other limitations are constituted by their dependence on molecular data sets used to gather coefficients of the terms used in functions for regression analysis and fitting.<sup>10</sup> As a consequence of those simplifications and limitations, a fast, robust and accurate scoring function for all biological targets does not exist.<sup>10</sup> In order to overcome the imperfections of current scoring functions, consensus scoring<sup>7</sup> is applied in docking and VS approaches. Consensus scoring consists of combining different and unrelated scoring functions (e.g. combining a force-field based with a knowledge-based scoring function) with the aim of reducing and balancing errors of the individual scoring functions and consequently, improving the chances of identifying 'true' ligands.<sup>10</sup>

### **2.3 Glide (Grid based ligand docking with energetic)<sup>14</sup>**

Glide is a protein-ligand docking program which approximates a complete systematic search over ligand positions, orientations and conformations in the receptor site.<sup>14</sup> A series of hierarchical filters are used to identify the possible locations of the ligand into the active site of the receptor which is represented as a grid derived by different sets of fields for the accurate scoring of any ligand pose.<sup>14</sup> To summarize, the search algorithm performs an initial rough positioning and scoring to reduce the search then a torsionally flexible energy optimization using the OPLS-AA<sup>15</sup> force field for a few hundred surviving candidate poses.<sup>14</sup> Only the best candidates are further refined *via* Monte Carlo sampling of pose conformation.<sup>14</sup> The best docked poses are selected by the scoring function (Glide Score) that combines empirical and force-field based terms.<sup>14</sup>

Glide XP<sup>16</sup> represents an extra precision mode in Glide which consists of an optimized search algorithm and scoring function. The goal of this methodology is to semi-quantitatively rank the ability of a candidate ligand to bind to a specified conformation of the protein receptor.<sup>16</sup>

The scoring function in Glide XP<sup>16</sup> is shown below (Equation 2.2):

$$\text{XP Glide Score} = E_{\text{coul}} + E_{\text{wdW}} + E_{\text{bind}} + E_{\text{penalty}} \quad (2.2)$$

$$E_{\text{bind}} = E_{\text{hyd\_enclosure}} + E_{\text{hb\_nn\_motif}} + E_{\text{hb\_cc\_motif}} + E_{\text{PI}} + E_{\text{hb\_pair}} + E_{\text{phobic\_pair}}$$

$$E_{\text{penalty}} = E_{\text{desolv}} + E_{\text{ligand\_Strain}}$$

Where  $E_{\text{hyd\_enclosure}}$  = hydrophobic enclosure interaction;  $E_{\text{hb\_cc\_motif}}$  = charged-charged hydrogen bond interaction motif;  $E_{\text{hb\_nn\_motif}}$  = neutral-neutral hydrogen bond interaction motif;  $E_{\text{PI}}$  = reward for pi-stacking and pi-cation interactions.

The key features of this scoring function<sup>16</sup> are:

- 1) the application of large desolvation penalties to both ligand and protein polar and charged groups.
- 2) the identification of specific structural motifs which contribute to binding affinity.

## 2.4 F.R.E.D. (Fast Rigid Exhaustive Docking)<sup>17,18</sup>

F.R.E.D. is a protein-ligand docking program which: a) takes a multi-conformer database and receptor file as input, b) performs a systematic, exhaustive, non-stochastic evaluation of all possible ligand poses within the protein active site, c) filters for shape complementarity and pharmacophoric features (if a known inhibitor is prompted) d) and finally select and optimize poses using the Chemgauss3 scoring function.<sup>17,18,19</sup> The protein is held rigid during the docking process as well as the conformers of the ligand. Ligand flexibility, however, is implicitly taken into account by docking the conformer ensemble of each molecule.<sup>17,18</sup>

The protein receptor is represented as the complementary image of the active site and it is created by contouring a shape potential field that complements the active site.<sup>17,18</sup>

The Chemgauss3 scoring function uses Gaussian smoothed potentials to measure the complementarity of ligand poses within the active site and it is shown below (Equation 2.3):<sup>18,19</sup>

$$\text{Chemgauss3 scoring function} = E_{\text{shape}} + E_{\text{hblp}} + E_{\text{hbs}} + E_{\text{metal}} \quad (2.3)$$

Where  $E_{\text{shape}}$  = shape interaction between the negative image of the protein and the ligand;  $E_{\text{hblp}}$  = Hydrogen bonding interaction between ligand and protein;  $E_{\text{hbs}}$  = Hydrogen bonding interactions with implicit solvent;  $E_{\text{metal}}$  = Metal-chelator interactions.

## 2.5 Autodock<sup>20,21</sup>

AutoDock is a suite of automated docking tools and consists of two main programs: AutoDock performs the docking of the ligand to a set of grids describing the protein while autogrid pre-calculates these grids.<sup>20,21</sup> AutoDock combines an empirical free energy force field with a Lamarckian Genetic Algorithm, giving a prediction of poses and free energies of binding.<sup>20,21,22</sup> The Lamarckian genetic algorithm provides the conformational search of the ligands creating a population of trial conformations, then genetic operators (mutations, crossovers etc.) are applied to exchange conformational parameters and to select conformations with the lowest binding energy.<sup>22</sup> The ‘Lamarckian’ feature allows individual conformations to search their local conformational space, finding local minima, and then to pass this information to later generations.<sup>22</sup>

The target protein is embedded in a grid representing the energy of interaction between a probe placed at any grid point and the target protein.<sup>22</sup>

The scoring function consists of a semi-empirical free energy force field designed to predict binding free energies of small molecules to macromolecular targets.<sup>22</sup> The free energy of binding is estimated to be equal to the difference between the energy of the ligand and the protein in a separated unbound state and the energy of the ligand–protein complex.<sup>22</sup> The force field includes six pair-wise evaluations ( $V$ ) and an estimate of the conformational entropy lost upon binding ( $\Delta S_{\text{conf}}$ ) as shown below (Equation 2.4):<sup>22</sup>

$$\Delta G = (V_{\text{L-L bound}} - V_{\text{L-L unbound}}) + (V_{\text{P-P bound}} - V_{\text{P-P unbound}}) + (V_{\text{P-L bound}} - V_{\text{P-L unbound}} + \Delta S_{\text{conf}}) \quad (2.4)$$

In the above equation L refers to the ligand and P refers to the protein in a protein–ligand complex and the pair-wise atomic terms ( $V$ ) include evaluations for dispersion/repulsion, hydrogen bonding, electrostatics, and desolvation.

## 2.6 eHiTS<sup>23,24,25,26</sup> (electronic high-throughput screening)

eHiTS is a docking program performing an exhaustive conformational and positional search of the space of the ligand, compatible with steric and chemical constraints, at a speed practical for virtual high-throughput screening.<sup>23,24</sup> A statistically derived empirical scoring function evaluates the score of each ligand pose during the docking sequence.<sup>23,24</sup>

A steric grid is built within the binding pocket, identifying pockets and possible interaction sites using thousands of geometric shapes (polyhedra) to describe cavities.<sup>23,24</sup>

The ligand is broken into rigid fragments while connecting flexible chains and all rigid fragments are docked independently.<sup>23,24</sup> A fast graph-matching algorithm identifies all matching solutions and reconstructs the original molecule by fitting flexible chains between the rigid fragments and satisfying all the sterical constraints.<sup>23,24</sup> The resulting poses are refined by a local energy minimization in the active site of the receptor and ranked by the scoring function which is constituted by the following terms:

- Hydrogen-bonding (distance and angle dependent energy function)
- Hydrophobicity (applied to surface contact points)
- Aromatic pi-stacking (applied to surface contact points)
- Electrostatic potential (based on Coulomb's law)
- Van der Waals contact energy (Lennard–Jones potential)
- Metal ion interactions (distance and angle dependent energy function)
- Penalty for incompatible contacts (e.g. polar–hydrophobic or same charge)
- Interaction energy of exposed surface atoms with solvent properties
- Intra-molecular interactions (both ligand and receptor)

## 2.7 GOLD<sup>27,28</sup> (Genetic Optimization for Ligand Docking)

GOLD<sup>27,28</sup> uses a genetic algorithm for identifying the global minimum of the energy of interaction between the ligand and the protein receptor.<sup>27</sup>

A genetic algorithm mimics the natural evolution process where the conformation of a ligand and the mapping of interaction points between ligand and receptor atoms are codified into a chromosome.<sup>27,29</sup> After the application of

genetic operators (mutations, crossovers and migrations), only the individuals with the best fitness are conserved.<sup>27,29</sup> The population of individuals is divided into smaller groups employing an island model and the fitness of each individual is assessed using the available scoring functions present in the software.<sup>27,29</sup>

The placing of the ligand in the binding site is based on fitting points which are generated on hydrogen bonding groups of protein and ligand where acceptor or donor points are mapped.<sup>27,29</sup> Additionally, hydrophobic fitting points in the protein cavity and ligand CH groups are mapped.<sup>27,29</sup>

For the purposes of this thesis, two out of the four available scoring functions were considered: GoldScore<sup>27</sup> and ChemScore.<sup>27, 30</sup>

The Goldscore<sup>27</sup> is a molecular mechanics–like scoring function with four terms (Equation 2.5):

$$\text{Goldscore Fitness} = S_{\text{hb\_ext}} + S_{\text{vdw\_ext}} + S_{\text{hb\_int}} + S_{\text{vdw\_int}} \quad (2.5)$$

where  $S_{\text{hb\_ext}}$  is the protein–ligand hydrogen bond score,  $S_{\text{vdw\_ext}}$  is the protein-ligand van der Waals score,  $S_{\text{hb\_int}}$  is the contribution to the Fitness due to intramolecular hydrogen bonds in the ligand and  $S_{\text{vdw\_int}}$  is the contribution to the Fitness due to intramolecular ligand van der Waals score.

The Chemscore<sup>14,17</sup> is an empirical based scoring function and estimates the free energy of binding of a ligand to a protein as follows (Equation 2.6):

$$\text{Chemscore Fitness} = \Delta G_{\text{binding}} + \Delta G_{\text{hbond}} S_{\text{hbond}} + \Delta G_{\text{metal}} S_{\text{metal}} + \Delta G_{\text{lipo}} S_{\text{lipo}} + \Delta G_{\text{rot}} H_{\text{rot}} \quad (2.6)$$

where  $S_{\text{hbond}}$ ,  $S_{\text{metal}}$ , and  $S_{\text{lipo}}$  are scores for hydrogen bonding, acceptor-metal, and lipophilic interactions, respectively.  $H_{\text{rot}}$  is a score representing the loss of conformational entropy of the ligand upon binding to the protein.

The Chemscore<sup>14,17</sup> scoring function is up to three times faster than Goldscore<sup>14</sup> to dock molecules, but the latter is more accurate in predicting binding modes with big ligands.<sup>27</sup> Combining both scoring functions has been shown to give better results than using a single function thanks to the complementarity of their parameters.<sup>27</sup> No case study has ever identified the

superiority of Goldscore<sup>14,17</sup> over Chemscore<sup>14,17</sup> and *vice versa* for all the biological targets.<sup>27</sup>

## 2.8 DOCK Blaster<sup>31</sup>

DOCK Blaster,<sup>31</sup> is a fully automated docking system which includes self-assessment. The software has been tested for reproducing experimental observed poses and for enrichment, defined as the ability to find active molecules included in a database of decoys. A decoy is a member of the database that does not bind to the target but shares similar physico-chemical properties with active molecules.<sup>31</sup> DOCK Blaster is accessible *via* a web-enabled user interface<sup>32</sup> and the docking program used is DOCK 3.6.<sup>33</sup> The DOCK Blaster pipeline is composed of six modules:<sup>31</sup> a) the parser, which identifies the receptor and ligand from a PDB file, b) the scrutinizer, which attempts to correct structural problems of supplied pdb structures, c) the preparer which assigns atomic parameters, calculates protonation states, 'hot spots' and grids, d) the calibrator, which assess docking performance and suggests optimal docking parameters, e) the docker, which performs the screening f) the assessor, which interprets and give results as a spreadsheet.

The docking is based on an incremental construction algorithm where the ligand is decomposed into single fragments and incrementally reconstructed inside the active site.<sup>34</sup> The shape characteristics of the protein active site are described as a series of overlapping spheres filling the available volume and an anchor fragment of the ligand is oriented independently from the other rigid fragments by matching its atoms with sphere centers.<sup>34</sup> The best scored anchor fragments are then used for subsequent growing of the ligand and the final conformation is scored.<sup>34</sup>

The DOCK blaster scoring function<sup>33</sup> is force field based and it is expressed as shown below (Equation 2.7) :

$$E_{\text{score}} = E_{\text{VdW}} + E_{\text{elec}} + \Delta G_{\text{desol}}^L \quad (2.7)$$

where  $E_{\text{VdW}}$  represents the van der Waals term based on the AMBER united-atom force field<sup>35</sup>, and  $E_{\text{elec}}$  is the electrostatics term that is based on Poisson

Boltzmann calculations performed by DELPHI software<sup>36</sup> and  $\Delta G_{\text{desol}}^L$  is the polar and apolar desolvation term for each ligand atom.

## 2.9 Ligand-based virtual screening

Ligand-based approaches are complementary to structure-based methods and are often applied when structural information on the protein target is missing or scarce and are focused on the biological and chemical properties of ligands.<sup>37,38</sup>

Ligand-based methods often use explicit parameters reflecting molecular properties and rely on the principle that ligands similar to an active molecule are more likely to be active than random ligands.<sup>37,38</sup>

Ligand-based approaches generally start from a series of active molecules targeted at a specific biological target which are then used to generate a query for ranking databases of molecules with a lower computational cost when compared with structure-based methods.<sup>37,38</sup> They include pharmacophores, quantitative structure activity relationship (QSAR) models as well as 3D or 2D similarity calculations based on physicochemical properties and molecular shapes.<sup>37</sup>

One of the advantages of shape-based approaches used for this thesis is that multiple active compounds and explicit parameters are not strictly required, resulting in limited computational time for screening large databases.<sup>37,38</sup> The most common drawbacks of ligand-based methods are derived from the fact that they 1) do not take into account the protein structure; 2) are biased towards existing ligands; 3) chemical and shape descriptors are dependent on input conformation; 4) results rely on the quality of training sets; 5) in shape-based approaches the problem of the false negatives is not uncommon since compounds showing unforeseen binding modes but structurally dissimilar to a known active molecule in terms of shape and chemistry could be easily missed.<sup>37</sup>

### 2.9.1 ROCS<sup>38</sup> (Rapid Overlay of Chemical Structures)

ROCS is a shape similarity ligand-based software which evaluates the molecular shape and the chemical features of one or more active molecules defined as a query.<sup>38</sup> Molecules are described as atom-centered Gaussian functions and the searching algorithm overlays molecular shapes by overlapping the centers of mass of the query molecule with a screening database of conformers.<sup>38</sup> After this preliminary overlay, the principal moments of inertia are aligned and the final overlapping is evaluated by the combined shape and chemistry (color) score defined as the combo score.<sup>38</sup> This score is expressed as the sum of the Tanimoto<sup>19</sup> coefficients for shape and chemical (color) similarity and it is a number between 0 and 2 of which 2 represents the ideal matching of shape and chemical features (self overlay of the query molecule).

## 2.10 References

1. Mander, T., *Drug Discov. Today* **2000**, *5*, 223-225.
2. Lahana, R., *Drug Discov. Today* **1999**, *4*, 447-448.
3. Walters, W. P.; Stahl, M. T.; Murcko, M. A., *Drug Discov. Today* **1998**, *3*, 160-178.
4. Gohlke, H.; Klebe, G., *Angew. Chem. Int. Ed.* **2002**, *41*, 2644-2676.
5. Bajorath, J., *Nat. Rev. Drug Discov.* **2002**, *1*, 882-894.
6. McInnes, C., *Curr. Opin. Chem. Biol.* **2007**, *11*, 494-502.
7. Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P., *J. Med. Chem.* **1999**, *42*, 5100-5109.
8. Verkhivker, G.; Bouzida, D.; Gehlhaar, D.; Rejto, P.; Arthurs, S.; Colson, A.; Freer, S.; Larson, V.; Luty, B.; Marrone, T.; Rose, P., *J. Comput. Aided Mol. Des.* **2000**, *14*, 731-751.
9. Moitessier, N.; Therrien, E.; Hanessian, S., *J. Med. Chem.* **2005**, *49*, 5885-5894.
10. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., *Nat. Rev. Drug Discov.* **2004**, *3*, 935-949.
11. H. Holtje, W. S., D. Rognan, G. Folkers, *Molecular Modeling: Basic Principles and Applications*. Wiley VCH: New York, 2008.
12. Kramer, B.; Rarey, M.; Lengauer, T., *Proteins* **1999**, *37*, 228-241.
13. Huang, S.-Y.; Zou, X., *Proteins* **2007**, *66*, 399-421.
14. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., *J. Med. Chem.* **2004**, *47*, 1750-1759.
15. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.
16. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., *J. Med. Chem.* **2006**, *49*, 6177-6196.
17. McGann, M., *J. Chem. Inf. Model.* **2011**, *51*, 578-596.
18. McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K., *Biopolymers* **2003**, *68*, 76-90.
19. Willett, P., *Drug Discov. Today* **2006**, *11*, 1046-1053.

20. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J., *J. Comput. Chem.* **1998**, *19*, 1639-1662.
21. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., *J. Comput. Chem.* **2009**, *30*, 2785-2791.
22. Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S., *J. Comput. Chem.* **2007**, *28*, 1145-1152.
23. Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, B. S.; Johnson, A. P., *Curr. Protein Pept. Sci.* **2006**, *7*, 421-435.
24. Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, S. B.; Johnson, A. P., *J. Mol. Graph.* **2007**, *26*, 198-212.
25. Zsoldos, Z., *Abstr. Pap. Am. Chem. Soc.* **2004**, *228*, U508-U508.
26. Zsoldos, Z.; Johnson, A. P.; Simon, A.; Szabo, I.; Szabo, Z., *Abstr. Pap. Am. Chem. Soc.* **2002**, *224*, U339-U339.
27. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., *Proteins* **2003**, *52*, 609-623.
28. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., *J. Mol. Biol.* **1997**, *267*, 727-748.
29. Jones, G.; Willett, P.; Glen, R. C., *J. Mol. Biol.* **1995**, *245*, 43-53.
30. Eldridge, M.; Murray, C.; Auton, T.; Paolini, G.; Mee, R., *J. Comput. Aided Mol. Des.* **1997**, *11*, 425-445.
31. Irwin, J. J.; Shoichet, B. K.; Mysinger, M. M.; Huang, N.; Colizzi, F.; Wassam, P.; Cao, Y., *J. Med. Chem.* **2009**, *52*, 5712-5720.
32. <http://blaster.docking.org/>.
33. Mysinger, M. M.; Shoichet, B. K., *J. Chem. Inf. Model.* **2010**, *50*, 1561-1573.
34. Ewing, T. A.; Makino, S.; Skillman, A. G.; Kuntz, I., *J. Comput. Aided Mol. Des.* **2001**, *15*, 411-428.
35. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P., *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
36. Gilson, M. K.; Sharp, K. A.; Honig, B. H., *J. Comput. Chem.* **1988**, *9*, 327-335.
37. Drwal, M. N.; Griffith, R., *Drug Discov. Today Technol.* **2013**, *10*, e395-e401.

38. Hawkins, P. C. D.; Skillman, A. G.; Nicholls, A., *J. Med. Chem.* **2006**, *50*, 74-82.

### 3. Critical assessment of the available docking software

The priority for developing a VS protocol and, more generally, for ligand-protein docking, is to critically assess, amongst the available docking software implementing different search and evaluating algorithms, which are the ones capable of reproducing co-crystal structures, when available, of a ligand with its protein. In particular, the success of a docking program is usually measured by the root-mean-square deviation (RMSD) between the predicted and the experimentally observed heavy atom positions.<sup>1</sup> Generally a successful result is obtained when docking solutions have RMSD < 2 Å but other authors<sup>2,3</sup> consider RMSDs which lie between 2 and 3 Å as a partial success.<sup>4</sup> Another important aspect is to consider the reproducibility of the protein-ligand interactions and in fact, an acceptable RMSD that lacks this requirement is not considered as a good result according to some authors.<sup>5</sup> Typically, no single docking algorithm works best for all targets and it is necessary to establish the optimal docking algorithm on the specific target of interest.<sup>1</sup>

For the purposes of this thesis, an RMSD < 3 Å was used as a 'rule-of-thumb' for evaluating docking performance and in addition, the reproducibility of the protein-ligand interactions limited to conserved hydrogen bond networks between protein-ligand H-Bond donors and acceptors was considered.

In this thesis, the following software were evaluated: GLIDE,<sup>6</sup> F.R.E.D.,<sup>7,8</sup> Autodock,<sup>9,10</sup> eHiTS,<sup>11,12,13,14</sup> GOLD<sup>15,16</sup> and Dockblaster.<sup>17</sup>

#### 3.1 Protein and ligand preparation protocol

Two crystal structures of the *T. thermophilus* RNA polymerase (RNAP) holoenzyme in complex with the antibiotic myxopyronin **1.25** (PDB id: 3DXJ<sup>18</sup> and PDB id.: 3EQL<sup>19</sup>) were available from the Protein Databank<sup>20</sup> and were both evaluated in the docking software assessment.

In particular, two different chemical forms of myxopyronin were present in the two crystal structures: myxopyronin A **1.25** for PDB id: 3DXJ<sup>18</sup> and 7-desmethylmyxopyronin B **1.27** for PDB id.: 3EQL.<sup>19</sup>

The binding site within the RNAP for both myxopyronin A and 7-desmethylmyxopyronin B was defined as a 20 Å spherical cut of the protein surrounding the co-crystallized ligand from the two previously mentioned crystal structures.

The two antibiotic molecules were manually built using Maestro (Schrödinger, LLC, New York, NY) and the resulting molecular structures were geometrically optimized and energy minimized using the software module MacroModel by applying the OPLS2005 force field in a simulated water dielectric. Ionization states were defined for a neutral pH using the Epik module and partial charges were assigned by applying the OPLS2005 force field.

Protein structures were prepared using the 'Protein preparation wizard' module included in Maestro. Preparation consisted of assigning bond orders, deleting water molecules, adding hydrogen atoms and optimizing hydrogen bond networks assuming neutral pH, using PROPKA. Finally, protein structures were subjected to a restrained energy minimization using the OPLS2005 force field with heavy atoms allowed to deviate up to 0.3 Å from their experimental positions.

### **3.2 Docking protocol used with Glide<sup>6</sup>**

From the previously prepared ligand and protein structures, a 10 Å<sup>3</sup> box centered on the mass-center of the crystallographic ligand was considered for calculating the docking grid. A fully flexible torsional geometry was allowed for the ligand, and 100 poses were generated using the Glide XP scoring function.

Cluster analysis with an average linkage method was performed using clustering scripts implemented within Maestro. The docking conformations were clustered on the basis of RMSD applying a cutting rule at 1.5 Å between the cartesian coordinates of the ligand atoms.

The best ranked pose of the most populated cluster was selected as a result and RMSD values between predicted and the experimentally observed heavy atom positions were calculated using Maestro.

### 3.3 Docking protocol used with eHiTS<sup>11,12,13,14</sup>

eHiTS v9.0 automatically evaluates all of the possible protonation states for ligand and enzyme and for this reason, previously prepared structures were not used. Active-site detection was carried out using the '-complex' parameter and docking was performed using the highest accuracy setting (set to 6). For allowing a more accurate scoring and conformational sampling of the known inhibitors myxopyronin A **1.25** and 7-desmethylmyxopyronin B **1.27**, the '-bindener' parameter was used for scoring with the eHiTS scoring function. The best ranked pose of the most populated cluster was selected as a result. Cluster analysis with an average linkage method was performed using clustering scripts implemented in Maestro. The docking conformations were clustered on the basis of RMSD applying a cutting rule at 1.5 Å between the cartesian coordinates of the ligand atoms. The best ranked pose of the most populated cluster was selected as a result and RMSD values between the predicted and the experimentally observed heavy atom positions were calculated using Maestro.

### 3.4 Docking protocol used with Autodock<sup>9,10</sup>

Autodock v.4.2.5.1 was used for docking calculations while ligand and protein structures were prepared with Autodock Tools v. 1.5.6 module. The docking area was defined using a box, centered on the cartesian coordinates of crystallographic myxopyronin A **1.25** and 7-desmethylmyxopyronin B **1.27**. A grid point box of 60 × 60 × 60 with 0.375 Å spacing was calculated around the docking area for all the ligand atom types using AutoGrid4. For each ligand, 100 separate docking calculations were performed using the Lamarckian Genetic Algorithm. All the other docking parameters were applied with the default settings. The docking conformations were clustered using Autodock Tools v. 1.5.6 module applying an average linkage rule on the basis of RMSD using a cutting rule at 1.5 Å between the cartesian coordinates of the ligand atoms and were ranked based on the AutoDock scoring function. RMSD values between

the predicted and the experimentally observed heavy atom positions were calculated using Maestro.

### **3.5 Docking protocol used with F.R.E.D<sup>7,8</sup>**

F.R.E.D v.2.2.5 was used for docking calculations. The receptor file was set up interactively using a GUI supplied with the software, specifying the location of the active site as a box of 20 Å<sup>3</sup> centered on the cartesian coordinates of crystallographic myxopyronin A **1.25** and 7-desmethylmyxopyronin B **1.27**. Ligand conformers were generated using OMEGA version 2.3.2 prior to running the docking and applying the default settings. All the default docking parameters were used with the only exception of the number of poses generated, which were set to 100. The Chemgauss scoring function was used for the exhaustive search, optimization, and final scoring. Cluster analysis with an average linkage method was performed using clustering scripts implemented in Maestro. The docking solutions were clustered on the basis of RMSD applying a cutting rule at 1.5 Å between the cartesian coordinates of the ligand atoms. The best ranked pose of the most populated cluster was selected as a result and RMSD values between the predicted and the experimentally observed heavy atom positions were calculated with Maestro.

### **3.6 Docking protocol used with GOLD<sup>15,16</sup>**

From the previously prepared ligand and protein structures as described in Section 3.1, the docking area was defined using a sphere of 20 Å diameter centered on the Cartesian coordinates of crystallographic myxopyronin A **1.25** and 7-desmethylmyxopyronin B **1.27**. Docking was performed using GOLD v4.0.1 applying 200% of accuracy for the genetic algorithm and 100 poses of the ligand were generated and ranked separately with both scoring functions, Goldscore and Chemscore.

Cluster analysis with an average linkage method was performed using a module implemented in GOLD. The docking solutions were clustered on the basis of RMSD applying a cutting rule at 1.5 Å between the cartesian coordinates of the ligand atoms. The best ranked pose of the most populated cluster was selected as a result and RMSD values between the predicted and the experimentally observed heavy atom positions were calculated with Maestro.

### **3.7 Docking protocol used with DOCK Blaster<sup>17</sup>**

The graphical interface of DOCK Blaster was accessed online and the PDB id of the enzyme (3DXJ<sup>18</sup> and 3EQL<sup>19</sup>) were submitted specifying the 3-letter code of the ligand (NE6 and MXP respectively). A job query number was assigned for the docking runs and results of the best docked poses with scoring values were generated automatically. No cluster analysis is available for this automated docking interface. RMSD values between the predicted and the experimentally observed heavy atom positions were automatically calculated.

### **3.8 Critical evaluation of results**

With regard to the docking of myxopyronin A **1.25** inside its native co-crystal complex with PDB id: 3DXJ,<sup>18</sup> the results in terms of RMSD between the predicted and the experimentally observed heavy atom positions and reproducibility of original protein-ligand interactions are summarized in Table 3.1 and docking poses are reported in Figures 3.1-3.7.

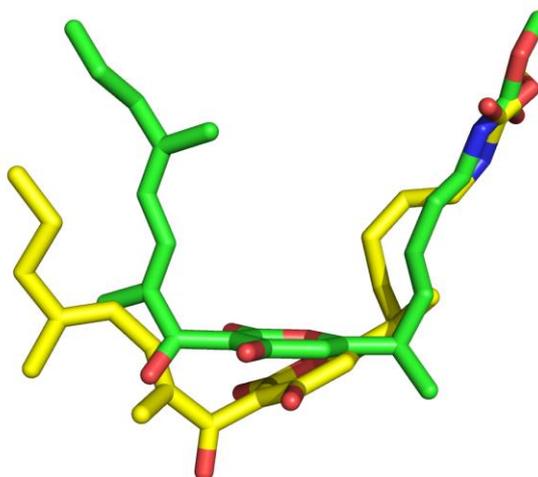
**Table 3.1** Results for the docking of myxopyronin A **1.25** inside its native co-crystal complex with PDB id: 3DXJ.<sup>18</sup>

Docking software	RMSD	Conserved interactions <sup>C</sup>
<b>Autodock</b>	6.44 Å	1 out of 4
<b>Dockblaster</b>	5.38 Å	2 out of 4
<b>eHiTS</b>	3.00 Å	1 out of 4
<b>F.R.E.D</b>	5.91 Å	2 out of 4
<b>GLIDE</b>	8.60 Å	0 out of 4
<b>GOLD<sup>A</sup></b>	1.28 Å	3 out of 4
<b>GOLD<sup>B</sup></b>	1.40 Å	3 out of 4

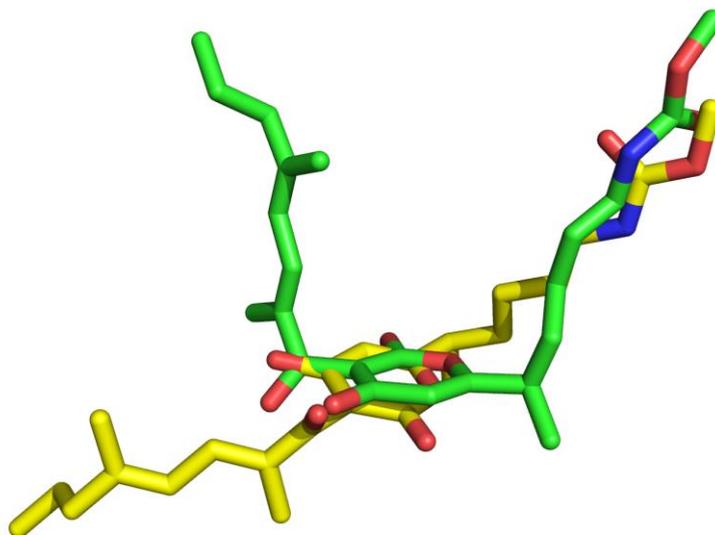
A Goldscore scoring function

B Chemscore scoring function

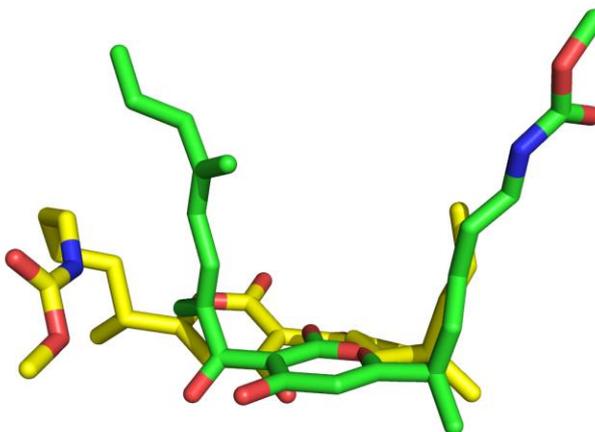
C Conserved interactions intended as polar contacts of docked ligand with same amino-acidic residues contacted in native co-crystal complex.



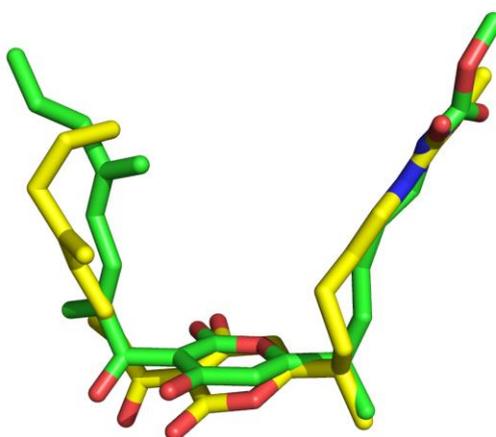
**Figure 3.1** Overlay between crystallographic myxopyronin A **1.25** (C atoms in green) and docking solution obtained with Autodock (C atoms in yellow).



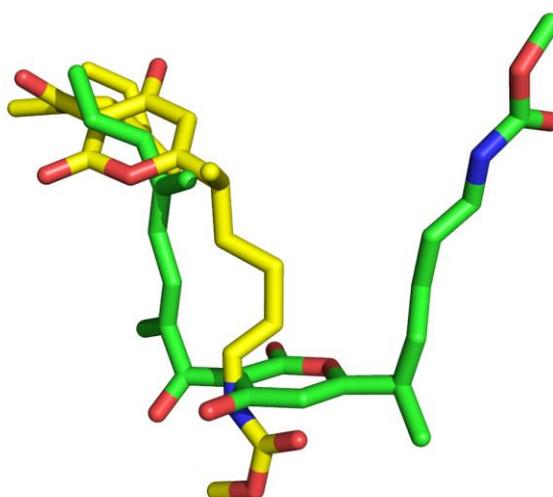
**Figure 3.2** Overlay between crystallographic myxopyronin A **1.25** (C atoms in green) and docking solution obtained with Dockblaster (C atoms in yellow).



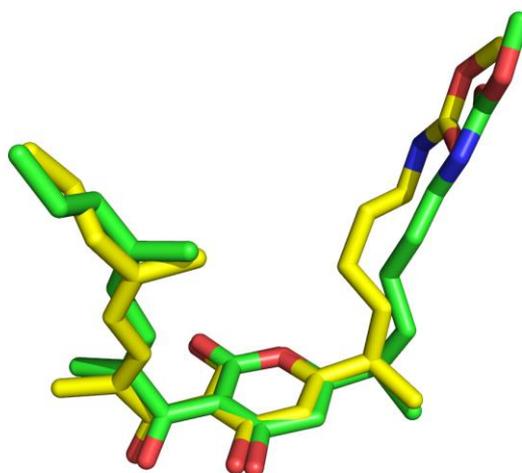
**Figure 3.3** Overlay between crystallographic myxopyronin A **1.25** (C atoms in green) and docking solution obtained with eHiTS (C atoms in yellow).



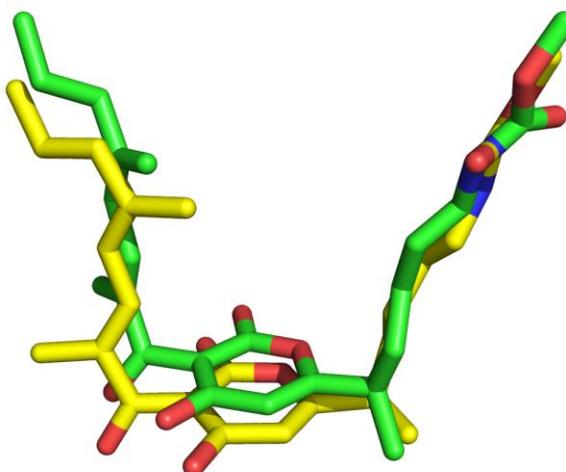
**Figure 3.4** Overlay between crystallographic myxopyronin A **1.25** (C atoms in green) and docking solution obtained with F.R.E.D (C atoms in yellow).



**Figure 3.5** Overlay between crystallographic myxopyronin A **1.25** (C atoms in green) and docking solution obtained with GLIDE (C atoms in yellow).



**Figure 3.6** Overlay between crystallographic myxopyronin A **1.25** (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Goldscore as a scoring function.



**Figure 3.7** Overlay between crystallographic myxopyronin A **1.25** (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Chemscore as a scoring function.

With regard to the docking of 7-desmethylmyxopyronin B **1.27** inside its native co-crystal complex with PDB id.: 3EQL,<sup>19</sup> the results in terms of RMSD between the predicted and the experimentally observed heavy atom positions and reproducibility of the original protein-ligand interactions are summarized in Table 3.2 and docking poses are reported in Figures 3.8-3.14.

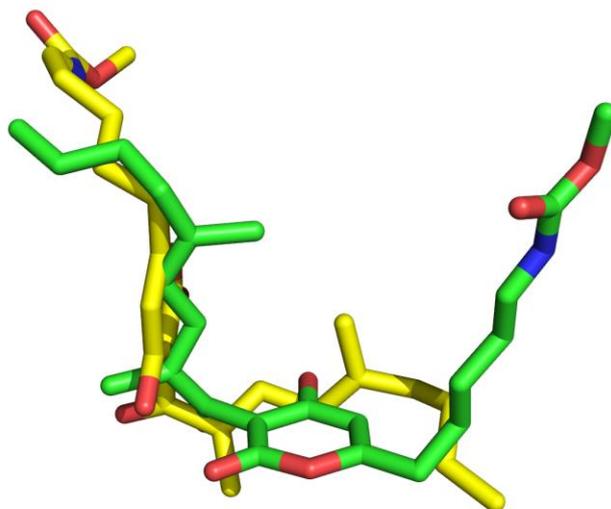
**Table 3.2** Results for the docking of 7-desmethylmyxopyronin B **1.27** inside its native co-crystal complex with PDB id.: 3EQL.<sup>19</sup>

Docking software	RMSD	Conserved interactions <sup>C</sup>
<b>Autodock</b>	6.30 Å	0 out of 5
<b>Dockblaster</b>	8.87 Å	1 out of 5
<b>eHiTS</b>	3.26 Å	1 out of 5
<b>F.R.E.D</b>	8.96 Å	2 out of 5
<b>GLIDE</b>	6.80 Å	2 out of 5
<b>GOLD<sup>A</sup></b>	1.76 Å	3 out of 5
<b>GOLD<sup>B</sup></b>	2.29 Å	4 out of 5

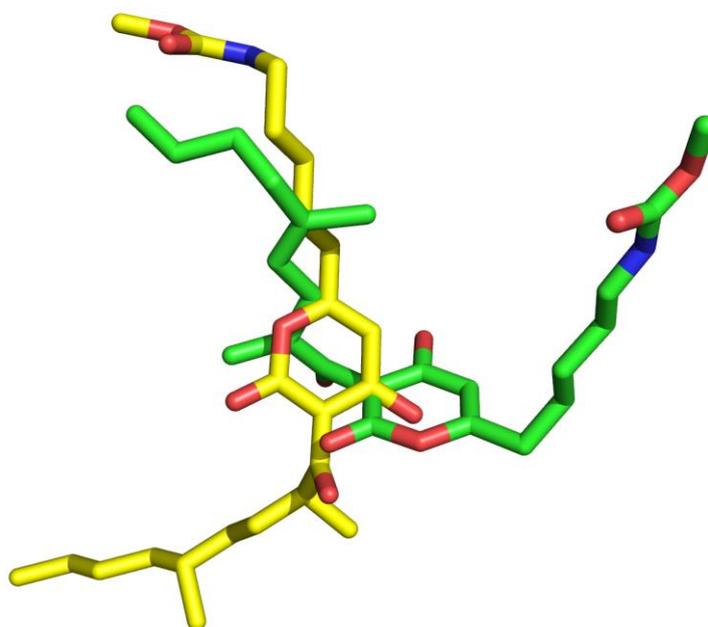
A Goldscore scoring function

B Chemscore scoring function

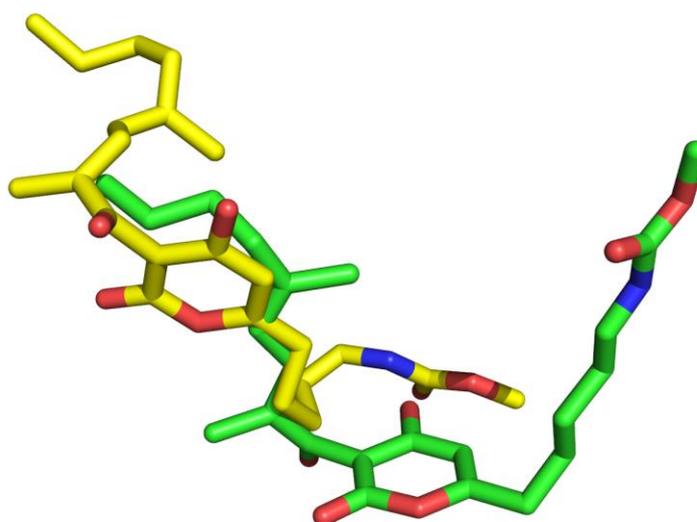
C Conserved interactions intended as polar contacts of docked ligand with same amino-acidic residues contacted in native co-crystal complex.



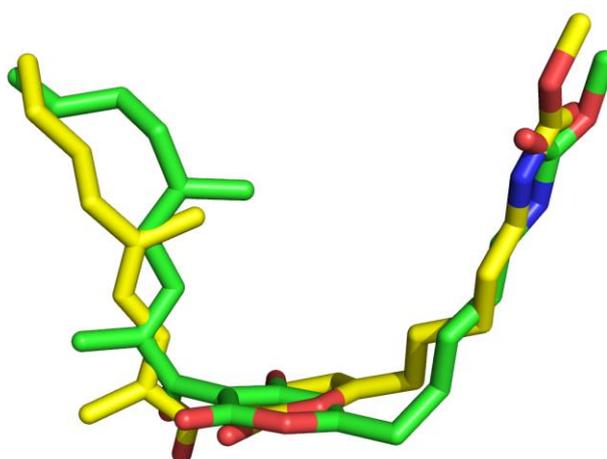
**Figure 3.8** Overlay between crystallographic 7-desmethylmyxopyronin B **1.27** (C atoms in green) and docking solution obtained with Autodock (C atoms in yellow).



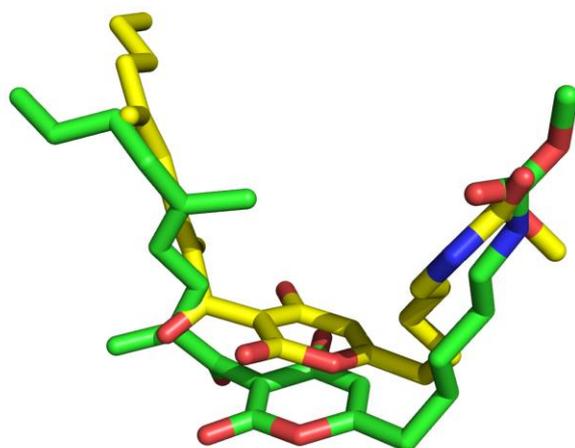
**Figure 3.9** Overlay between crystallographic 7-desmethylmyxopyronin B **1.27** (C atoms in green) and docking solution obtained with Dockblaster (C atoms in yellow).



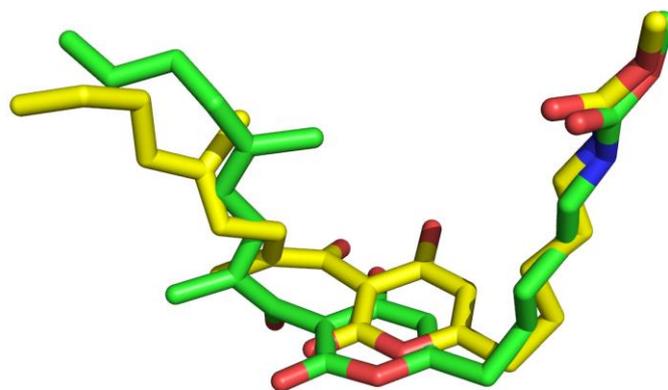
**Figure 3.10** Overlay between crystallographic 7-desmethylmyxopyronin B **1.27** (C atoms in green) and docking solution obtained with eHiTS (C atoms in yellow).



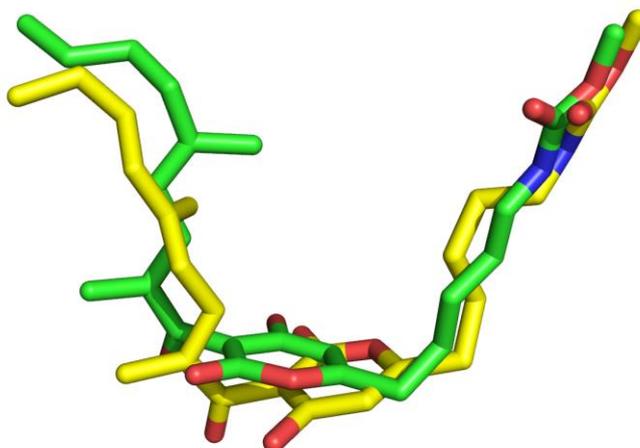
**Figure 3.11** Overlay between crystallographic 7-desmethylmyxopyronin B **1.27** (C atoms in green) and docking solution obtained with F.R.E.D (C atoms in yellow).



**Figure 3.12** Overlay between crystallographic 7-desmethylmyxopyronin B **1.27** (C atoms in green) and docking solution obtained with GLIDE (C atoms in yellow).



**Figure 3.13** Overlay between crystallographic 7-desmethylmyxopyronin B **1.27** (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Goldscore as a scoring function.



**Figure 3.14** Overlay between crystallographic 7-desmethylmyxopyronin B **1.27** (C atoms in green) and docking solution obtained with GOLD (C atoms in yellow) using Chemscore as a scoring function.

### 3.9 Conclusions

From the analysis of the docking results, it was possible to conclude that GOLD is the most suitable docking software for both crystal structures in terms of RMSD values and reproducibility of original hydrogen bond interactions, with a slightly better performance observed for the Goldscore scoring function over the Chemscore. Redocking of myxopyronin A **1.25** inside its native co-crystal complex (PDB id: 3DXJ<sup>18</sup>) has delivered slightly better results.

eHiTS can be considered as a second choice software in terms of RMSD values but as a consequence of the low reproducibility of the original hydrogen bond interactions, a different binding orientations are suggested for myxopyronin A **1.25** and 7-desmethylmyxopyronin B **1.27** to that observed in the co-crystal structures.

Interestingly, following comparison of RMSD values for both co-crystal structures, the majority of the tested software are performing better with myxopyronin A **1.25** co-crystal complex (PDB id: 3DXJ<sup>18</sup>) which has a slightly

lower resolution (3.0 Å) when compared to that for the 7-desmethylmyxopyronin B **1.27** co-crystal complex with PDB id.: 3EQL<sup>19</sup> (resolution 2.7 Å).

On the basis of these considerations, GOLD was used for performing docking studies and structure-based virtual screening protocols. Both scoring functions (Goldscore and Chemscore) were used alone or in combination for consensus scoring approaches and the crystal structure of the *T. thermophilus* RNA polymerase (RNAP) holoenzyme in complex with the antibiotic myxopyronin A<sup>18</sup> **1.27** (PDB id: 3DXJ<sup>18</sup>) was predominantly utilized.

### 3.10 References

1. Cole, J. C.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R., *Proteins* **2005**, *60*, 325-332.
2. Vieth, M.; Hirst, J. D.; Kolinski, A.; Brooks, C. L., *J. Comput. Chem.* **1998**, *19*, 1612-1622.
3. Bursulaya, B.; Totrov, M.; Abagyan, R.; Brooks, C., III, *J. Comput. Aided Mol. Des.* **2003**, *17*, 755-763.
4. Kontoyianni, M.; McClellan, L. M.; Sokol, G. S., *J. Med. Chem.* **2003**, *47*, 558-565.
5. Kroemer, R. T.; Vulpetti, A.; McDonald, J. J.; Rohrer, D. C.; Trosset, J.-Y.; Giordanetto, F.; Cotesta, S.; McMartin, C.; Kihlén, M.; Stouten, P. F. W., *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 871-881.
6. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., *J. Med. Chem.* **2004**, *47*, 1750-1759.
7. McGann, M., *J. Chem. Inf. Model.* **2011**, *51*, 578-596.
8. McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K., *Biopolymers* **2003**, *68*, 76-90.
9. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J., *J. Comput. Chem.* **1998**, *19*, 1639-1662.
10. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., *J. Comput. Chem.* **2009**, *30*, 2785-2791.
11. Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, B. S.; Johnson, A. P., *Curr. Protein Pept. Sci.* **2006**, *7*, 421-435.
12. Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, S. B.; Johnson, A. P., *J. Mol. Graph.* **2007**, *26*, 198-212.
13. Zsoldos, Z., *Abstr. Pap. Am. Chem. Soc.* **2004**, *228*, U508-U508.
14. Zsoldos, Z.; Johnson, A. P.; Simon, A.; Szabo, I.; Szabo, Z., *Abstr. Pap. Am. Chem. Soc.* **2002**, *224*, U339-U339.
15. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., *Proteins* **2003**, *52*, 609-623.
16. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., *J. Mol. Biol.* **1997**, *267*, 727-748.

17. Irwin, J. J.; Shoichet, B. K.; Mysinger, M. M.; Huang, N.; Colizzi, F.; Wassam, P.; Cao, Y., *J. Med. Chem.* **2009**, *52*, 5712-5720.
18. Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H., *Cell* **2008**, *135*, 295-307.
19. Belogurov, G. A.; Vassilyeva, M. N.; Sevostyanova, A.; Appleman, J. R.; Xiang, A. X.; Lira, R.; Webber, S. E.; Klyuyev, S.; Nudler, E.; Artsimovitch, I.; Vassilyev, D. G., *Nature* **2009**, *457*, 332-335.
20. Berman H.; Westbrook J.; Feng Z.; Gilliland G.; Bhat T.N.; Weissig H.; Shindyalov I. N.; P., B., *Nucl. Acids Res.* **2000**, *28*, 235-242.

## 4. Design of small molecule libraries as putative bacterial RNAP inhibitors

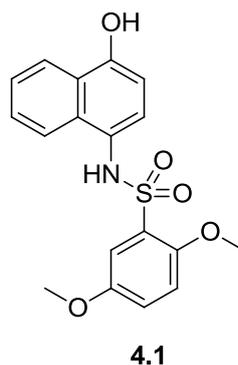
### 4.1 Reported inhibitors of bacterial RNAP

A high-throughput screen designed to discover novel inhibitors of *E. coli* RNA polymerase has been reported on Pubchem<sup>1</sup> by Penn Center for Molecular Discovery.<sup>2</sup>

The screen was based on an end-point assay monitoring the release of umbelliferone (Um), a fluorescent molecule. RNA polymerase catalyses the polymerization of RNA on a DNA template by incorporating nucleotide triphosphates, ATP, CTP, GTP, and UTP, with concomitant release of pyrophosphate (pp). For this assay, GTP was replaced with Um-pppp-G, which releases Um-ppp upon incorporation of guanosine into RNA. The subsequent addition of alkaline phosphatase cleaves Um-ppp to release the Um fluorophore which is fluorometrically detected.

A total of 62,232 compounds were screened and seventeen active molecules were reported. No putative binding site on RNAP was proposed for the active compounds and no synthetic chemistry has been developed to assess structure activity relationships.

It was envisaged that these inhibitors were attractive in terms of forming the basis of a study directed towards identifying the potential binding site within RNAP and therefore the development of more potent inhibitor scaffolds. Within this small set of reported inhibitors, the compound **4.1** (Figure 4.1) was selected on the basis of the RNAP *in vitro* inhibitory activity (18  $\mu$ M), synthetic accessibility, drug-likeness and suitability for designing small libraries to understand structure activity relationships and for analogue design.



**Figure 4.1** Structure of **4.1**.

In order to probe the potential binding region of the reported inhibitor **4.1** and to eventually allow structure-based rational design based upon its structure, docking studies were performed using the available X-Ray crystal structures of RNAP for the identification of its putative binding site. Two small libraries based on sulphonamide and urea scaffolds were designed, docked and synthesized as described below.

## **4.2 *In Silico* docking studies of reported inhibitors**

In order to identify the putative binding site of molecule **4.1**, an extensive docking study using GOLD was performed focussing on all the known inhibitor binding sites of RNAP and in particular, on the ansamycins,<sup>3</sup> streptolydigin,<sup>4</sup> myxopyronin,<sup>5</sup> tagetitoxin<sup>6</sup> and CBR<sup>7</sup> binding regions respectively. An X-Ray co-crystal structure is available for all these binding sites with the exception of the CBR703 class of inhibitors, where the putative binding site was hypothesized on the location of amino acid changes that yield altered sensitivity to CBR703 on *E. coli* tolC mutants.<sup>7</sup> The choice of docking parameters reflected the validation studies reported in Chapter 3.

## **4.3 General docking protocol**

The crystal structures of the *T. thermophilus* RNA polymerase holoenzyme in complex with the antibiotics streptolydigin (PDB id: 2A6H), myxopyronin (PDB

id: 3DXJ), and tagetitoxin (PDB id: 2BE5) were downloaded from the Protein Databank ([www.rcsb.org](http://www.rcsb.org)).

The binding sites within the RNAP for streptolidydigin, myxopyronin, tagetitoxin respectively were defined as a 20 Å cut of the protein surrounding the co-crystallised ligand from the structures with PDB code 2A6H, 3DXJ and 2BE5.

The binding site for the ansamycins is included within the 20 Å cut of the protein surrounding the streptolydigin molecule while the binding region for the CBR703 based inhibitors was specified as a 20 Å cut of the protein surrounding the amino acid residue 720 into the X-ray crystal structure with PDB code 2A6H.

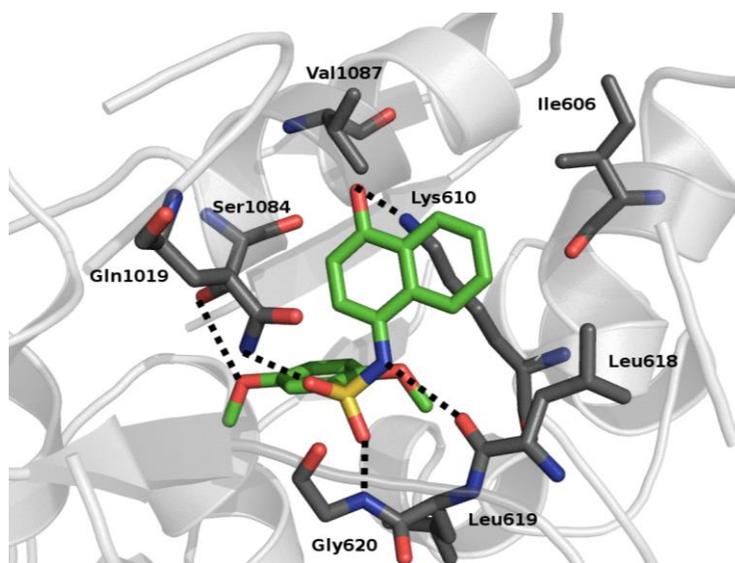
The hit molecule **4.1** was manually built using Maestro and the resulting molecular structure was geometrically optimised and energy minimised using the software module Macromodel by applying the OPLS2005 force field in a simulated water dielectric.

The docking runs were performed using GOLD<sup>8</sup> v4.0.1 (CCDC, Cambridge, UK) docking software by using the default settings for the genetic searching algorithm generating 100 poses for each compound and these were ranked with the Chemscore and rescored with Goldscore scoring function.

Due to the relatively large volumes of all the chosen binding regions as expected, the docking software gave plausible solutions for each binding site investigated and therefore in order to discriminate between the various possibilities a post-docking analysis was performed in each case.

Among the different docking solutions, a cluster analysis using an average linkage rule was performed using an internal command within the software. Only the best ranked poses of the most populated cluster were considered as the most likely and representative ones. In addition, the docked poses within each considered binding pocket were analysed in terms of overall score, shape complementarity to the cavity, possibility to establish specific H-bonds to the protein in addition to the less specific hydrophobic contacts, minimal exposure of hydrophobic groups to solvent and absence of an excessive torsional strain into the docked conformer. Following post-docking analysis of the results, it was

concluded that the predicted binding site for compound **4.1** is within the myxopyronin binding region of RNAP. The predicted binding mode of compound **4.1** within this region is shown in Figure 4.2 .



**Figure 4.2** Predicted binding mode of compound **4.1**.

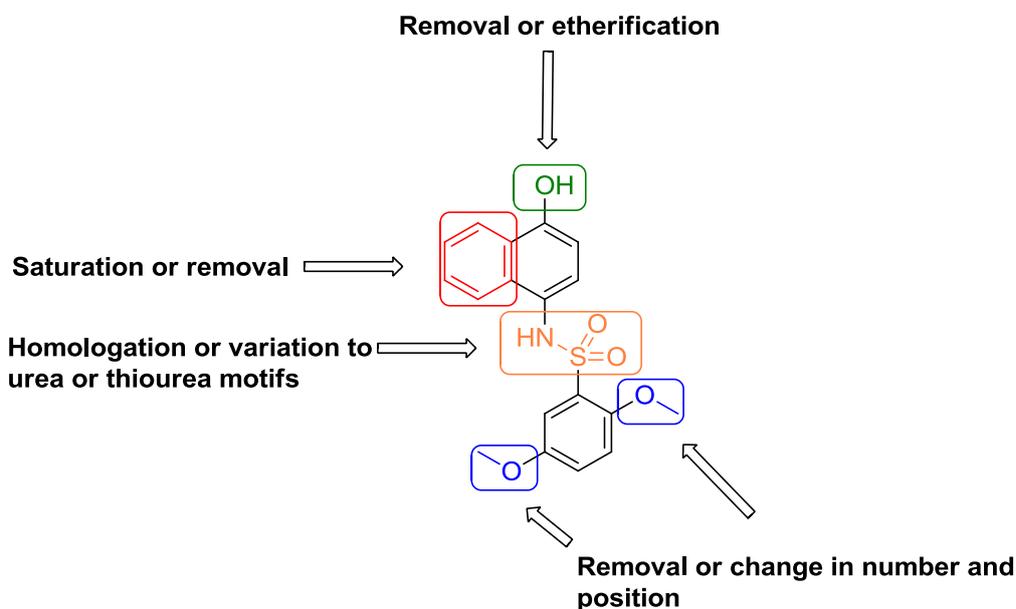
In particular, the following interactions between the ligand and the myxopyronin binding site were predicted: the NH group belonging to the sulphonamide moiety is predicted to H-bond to the Leu619 backbone carbonyl, the two sulphonamide oxygen atoms interact *via* H-bonds with the Gln1019 side chain and Gly620 backbone; the oxygen belonging to the methoxy group is predicted to make an H-bond with the hydroxyl of Ser1084, and the phenolic OH is predicted to form an H-bond with the side chain of Lys610. Additionally, the 4-amino-1-naphthol group is predicted to be involved in hydrophobic interactions with Val1087, Ile606 and Leu618 side chains.

#### **4.4 Design of analogues of compound 4.1**

In order to probe the validity of the predicted binding mode of compound **4.1** within the myxopyronin binding pocket of RNAP, a series of analogues of **4.1** was designed to probe these potential interactions.

In particular, following careful analysis of the predicted binding mode, it was reasoned that, although the sulphonamide moiety of the inhibitor is predicted to make a number of interactions with the protein, it may be possible to replace this moiety with other groups such as an urea and still maintain at least some of the predicted interactions with the protein. Additionally, inclusion of a thiourea spacer in place of the sulphonamide would test the importance of the oxygen atom predicted to act as an H-bond acceptor (Figure 4.3).

The hydroxyl group present in molecule **4.1** is not predicted to be essential and it can be removed or modified into an alkyl or phenyl ether which could establish additional interactions with the hydrophobic side chains of Val1087 and Ile1016. The two aromatic rings in **4.1** are predicted to be important for the binding although the naphthalene ring could be partially reduced to give a tetrahydro-derivative or simplified into a single aromatic ring. With regard to the two methoxy groups, modelling indicated that it is possible to vary their number or position within the ring or to substitute them with bioisosteres for obtaining useful structure activity relationships. This general analogues design strategy based upon compound **4.1** is summarised in Figure 4.3.

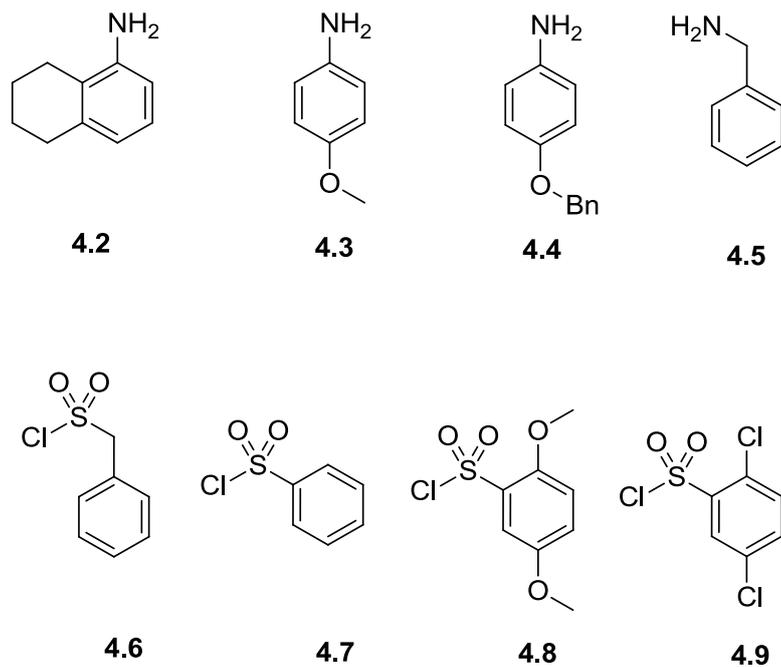


**Figure 4.3** Analogue design strategy for compound **4.1**.

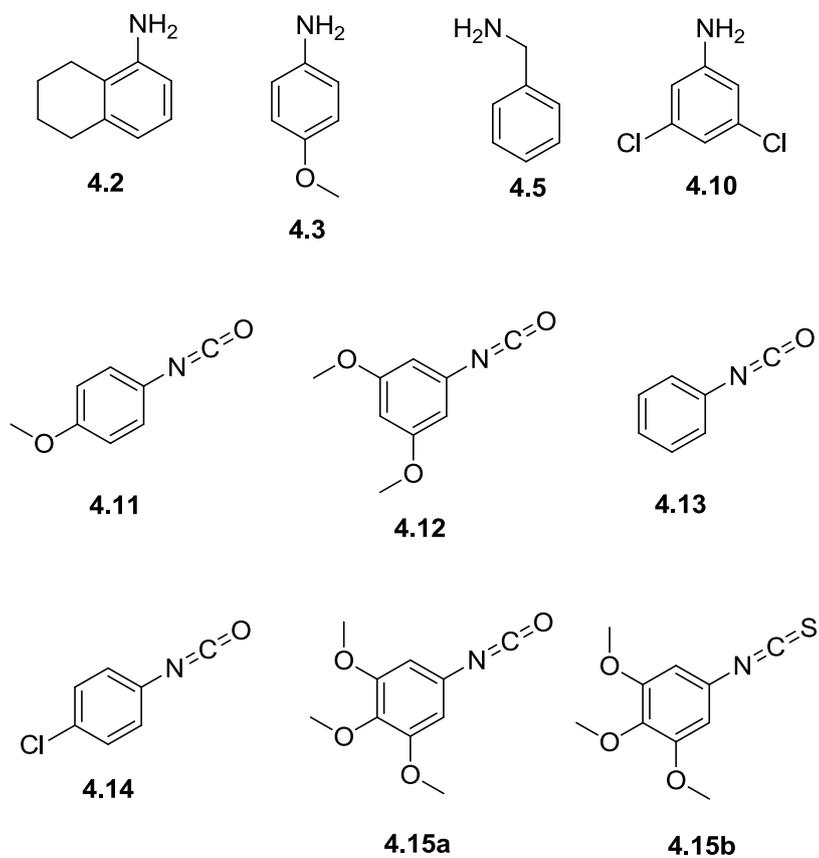
Following the above analysis and with reference to the ease of synthesis, two small libraries of sulphonamide and urea derivatives of **4.1** (Table 4.1) were designed. In order to obtain rapid SAR exploration, the composition of the planned libraries was fine-tuned following docking-based prioritization of a designed combinatorial library generated from selected commercially available starting materials (Figures 4.4 and 4.5).

#### **4.4.1 Design of the library and docking studies**

A library of 40 compounds was designed from starting materials reported in Figures 4.4 and 4.5 and each molecule in the library was docked into the myxopyronin binding site following a similar docking protocol and post-docking analysis as reported above (Section 4.3). In particular, the docking protocol in the present case was more accurate than that used earlier in order to prioritize the synthesis of the more promising candidates. The docking runs were performed using GOLD<sup>8</sup> v4.0.1 (CCDC, Cambridge, UK) docking software by using the best accuracy settings for the genetic searching algorithm which corresponded to the 200% of the default parameter values in order to get a more exhaustive search inside the binding cavity. 100 poses were generated for each compound and these were ranked using independently both the available scoring functions, Goldscore and Chemscore. Only compounds predicted to form at least two hydrogen bonds with binding site residues were progressed for visual inspection. A summary of the best ranked molecules selected for synthesis with relative score values and predicted H-bonds is reported in Table 4.1.



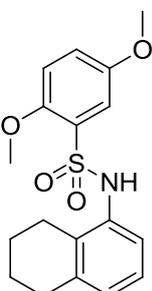
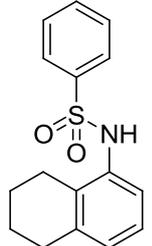
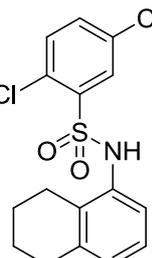
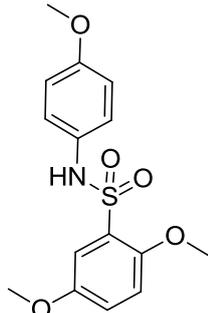
**Figure 4.4** Available starting materials for sulphonamide derivatives



**Figure 4.5** Starting materials for urea and thiourea derivatives

Among the designed structural analogues, compounds **4.20** and **4.30** appeared to be the most representative and promising ones in terms of overall score (Table 4.1) and binding mode analysis.

**Table 4.1** Overall score values and number of H-bonds of the best ranked compounds selected for synthesis.

Comp. number	Structure	Chemscore	Goldscore	Number of aminoacids contacted with H-bonds
4.16		26	47	3
4.17		28	46	2
4.18		29	43	2
4.19		24	45	3

**Table 4.1** continued

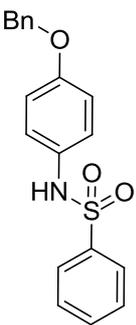
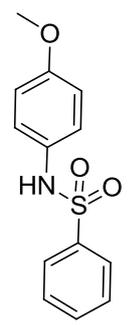
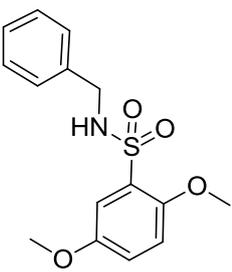
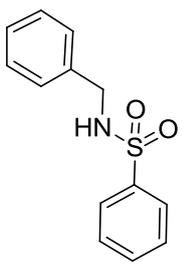
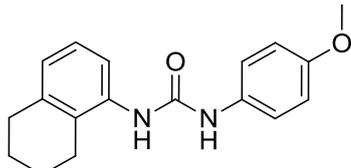
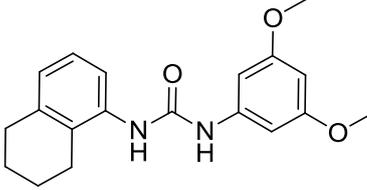
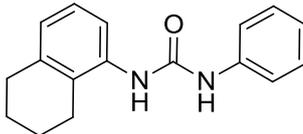
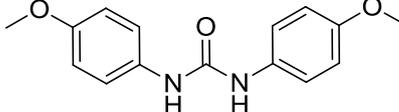
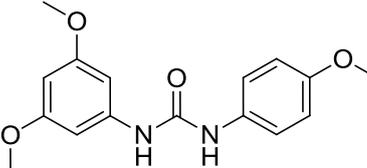
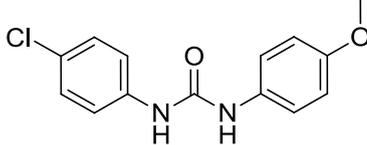
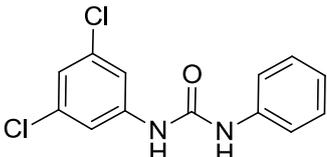
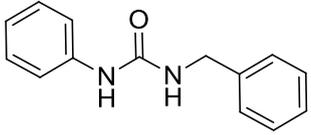
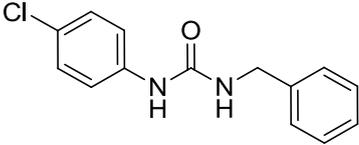
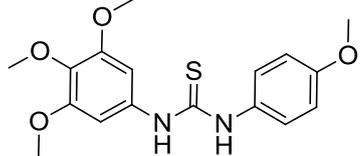
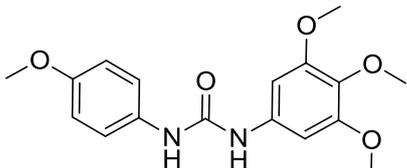
<p><b>4.20</b></p>		<p>29</p>	<p>58</p>	<p>3</p>
<p><b>4.21<sup>9</sup></b></p>		<p>25</p>	<p>42</p>	<p>3</p>
<p><b>4.22</b></p>		<p>25</p>	<p>44</p>	<p>3</p>
<p><b>4.23<sup>9</sup></b></p>		<p>23</p>	<p>45</p>	<p>3</p>
<p><b>4.24</b></p>		<p>26</p>	<p>41</p>	<p>2</p>

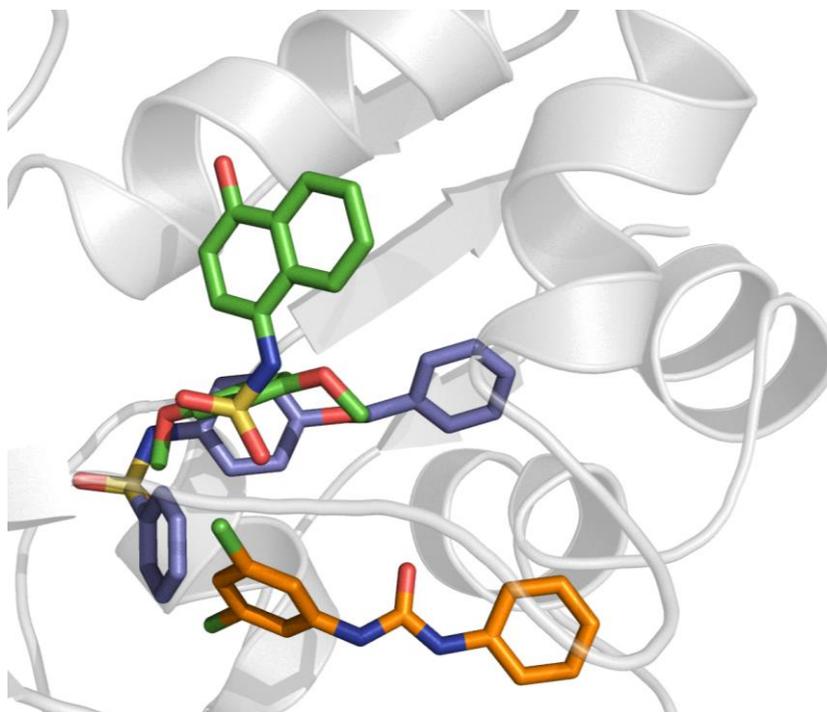
Table 4.1 continued

4.25		28	49	3
4.26		28	50	3
4.27 <sup>10</sup>		26	53	3
4.28		25	51	3
4.29 <sup>11</sup>		24	46	3
4.30		26	56	3
4.31 <sup>12</sup>		24	48	2
4.32 <sup>13</sup>		26	45	2

**Table 4.1** continued

<b>4.33</b>		25	43	3
<b>4.34</b>		25	49	3

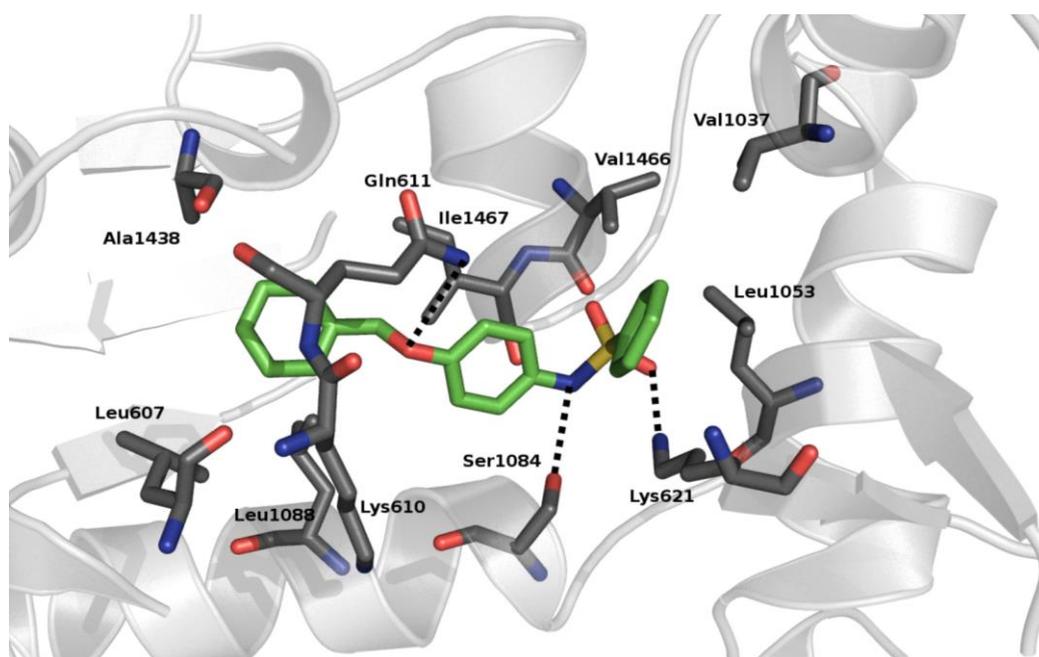
Interestingly, after visual inspection of the docked molecules it was possible to conclude that all the sulphonamide-based analogues are binding into a region close to that predicted for the binding of molecule **4.1** while the urea and thiourea derivatives are predicted to bind within a region which is distinct from that for compound **4.1** and the sulphonamide derivatives. An overlay within the myxopyronin binding site of the most promising sulphonamide and urea compounds, **4.20** and **4.30** respectively, along with the sulphonamide hit **4.1** is shown in Figure 4.6.



**Figure 4.6** Overlay of compound **4.1** (green sticks), **4.20** (purple sticks) and **4.30** (orange sticks) within myxopyronin binding site.

#### 4.4.2 Putative binding mode of compound 4.20

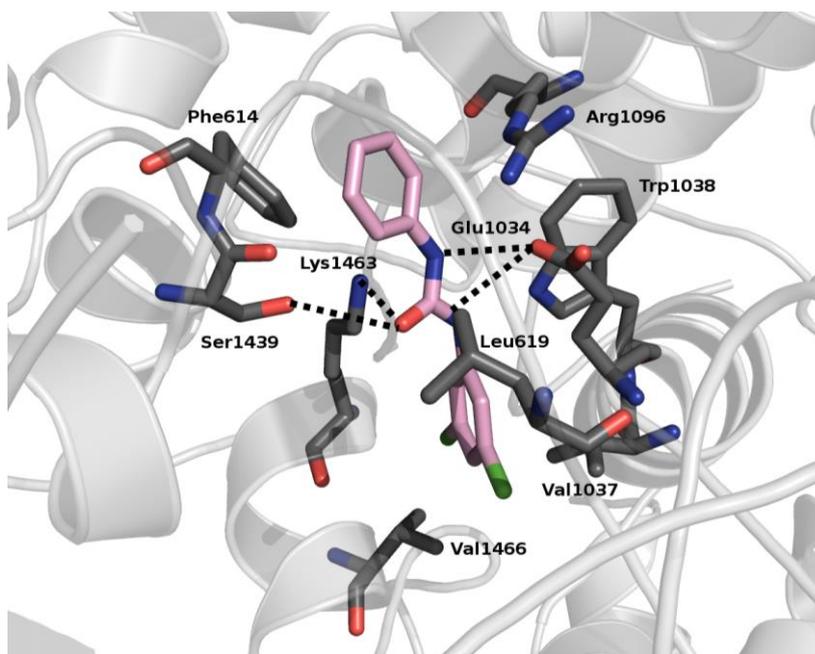
From the analysis of the proposed binding mode of compound **4.20** (Figure 4.7), the following interactions between the ligand and the myxopyronin binding site were predicted: the NH group belonging to the sulphonamide moiety H-bonded to Ser1084 side chain, the sulphonamide oxygen interacts *via* an H-bond with the Lys621 side chain; the aromatic ring belonging to the sulphonyl group is establishing hydrophobic interactions with the side chains of Val1466, Val1037 and Leu1053. The oxygen atom belonging to the benzyloxy moiety is H-bonding with the side chain of Gln611, while the aromatic portion of the benzyloxy group is establishing hydrophobic interactions with the side chains of Ile1467, Leu1088, Leu607 and Ala1438. The central aromatic ring bearing the benzyloxy substituent is establishing hydrophobic interactions with the side chain of Ile1467.



**Figure 4.7** Predicted binding mode of **4.20**.

#### 4.4.3 Putative binding mode of compound 4.30

From the analysis of the proposed binding mode of compound **4.30** (Figure 4.8), the following putative interactions between the ligand and the myxopyronin binding site of RNAP were predicted: the NH groups belonging to the urea moiety are H-bonded to Glu1034 side chain, the carbonyl oxygen interacted *via* H-bond with the Lys1463 and Ser1439 side chains while the aromatic ring belonging to the 3,5-dichlorophenyl moiety is establishing hydrophobic interactions with the side chains of Val1466, Val1037, Trp1038 and Leu619. The other phenyl ring is establishing cation- $\pi$  interactions with the positively charged side chains of Arg1096 and Lys1463 and additionally, it is predicted to be involved in a T-shaped  $\pi$ -stacking interaction with the side chain of Phe614.

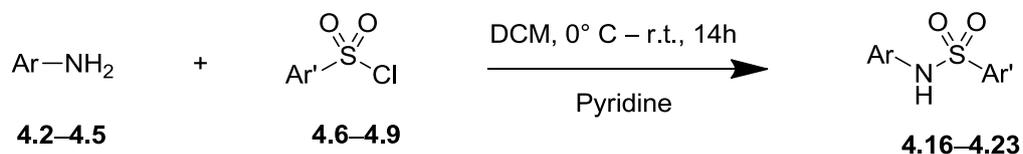


**Figure 4.8** Predicted binding mode of **4.30**.

#### 4.5 Synthesis of the sulphonamide-based library

Sulphonamides **4.16—4.23** were synthesized in good yield by adding dropwise the appropriate sulphonylchloride (**4.6—4.9**) to an excess of amine

(4.2—4.5) in presence of pyridine keeping the reaction mixture at 0 °C for 2h and then allowing it to warm at room temperature (Scheme 4.1 and Table 4.2).

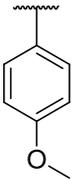
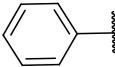
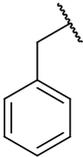
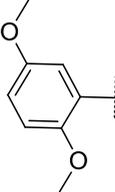
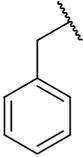
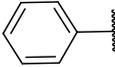


**Scheme 4.1** Synthesis of sulphonamides

**Table 4.2** Sulphonamide-based library

Compound	Ar	Ar'	Yield %
4.16			59
4.17			62
4.18			60
4.19			70
4.20			58

**Table 4.2 Continued**

4.21 <sup>9</sup>			72
4.22			51
4.23 <sup>9</sup>			56

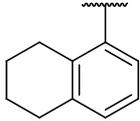
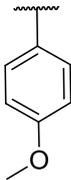
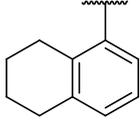
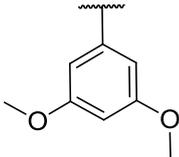
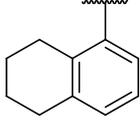
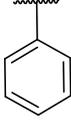
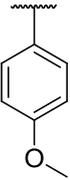
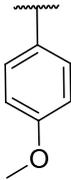
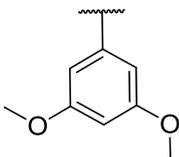
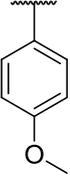
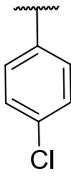
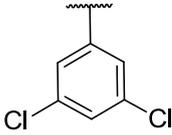
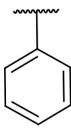
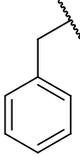
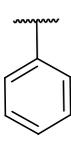
## 4.6 Synthesis of the urea-based library

The urea derivatives **4.24—4.32** and **4.34** were obtained in good yields by adding dropwise a dichloromethane solution of appropriate amine (**4.2**, **4.3**, **4.5**, **4.10**) into a solution of aryl isocyanate (**4.11—4.15a**) in dichloromethane using a slight excess of the latter at room temperature (Scheme 4.2 and Table 4.3).

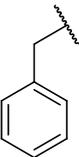
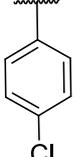
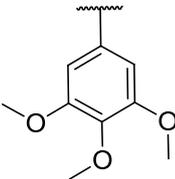


**Scheme 4.2** Synthesis of ureas

**Table 4.3** Urea-based library

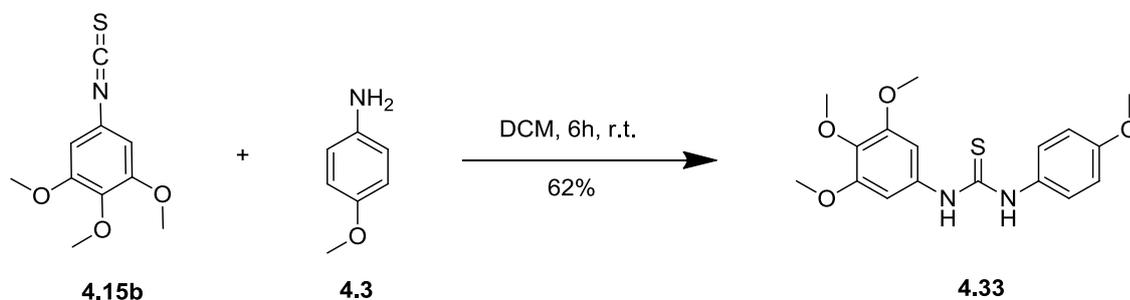
Compound	Ar	Ar'	Yield %
4.24			82
4.25			90
4.26			81
4.27 <sup>10</sup>			92
4.28			94
4.29 <sup>11</sup>			65
4.30			51
4.31 <sup>12</sup>			85

**Table 4.3** Continued

<b>4.32</b> <sup>13</sup>			69
<b>4.34</b>			86

## 4.7 Synthesis of the thiourea analogue

The thiourea **4.33** was obtained in moderate yield by adding dropwise a dichloromethane solution of amine **4.3** into a solution of aryl isothiocyanate **4.15b** in dichloromethane using a slight excess of the latter at room temperature (Scheme 4.3).



**Scheme 4.3** Synthesis of thiourea analogue **4.33**.

## 4.8 Biological evaluation of small molecule inhibitors

The synthesized analogues derived from **4.1** were evaluated for their *in vitro* RNAP activity against isolated *E. coli* RNAP using the SYBR Green assay (please see Appendix I for details) and the inhibitory activity is summarized below (Table 4.4). Disappointingly none of the molecules in this initial series of

derivatives were active against RNAP with the only exception being weak activity displayed compounds **4.19**, **4.22** and **4.32**.<sup>13</sup>

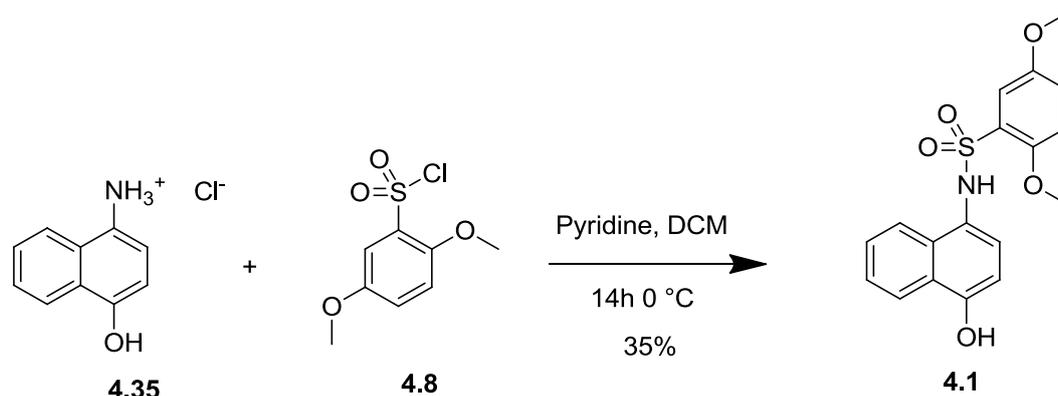
**Table 4.4** RNAP percentage inhibition of the synthesised small molecule inhibitors

Compound number	% of inhibition of RNAP at the concentration of 100 $\mu$ M	Compound number	% of inhibition of RNAP at the concentration of 100 $\mu$ M
<b>4.16</b>	0	<b>4.26</b>	0
<b>4.17</b>	0	<b>4.27</b> <sup>10</sup>	3.4
<b>4.18</b>	0	<b>4.28</b>	4.0
<b>4.19</b>	7.7	<b>4.29</b> <sup>11</sup>	2.9
<b>4.20</b>	0	<b>4.30</b>	0
<b>4.21</b> <sup>9</sup>	0	<b>4.31</b> <sup>12</sup>	3.2
<b>4.22</b>	5.1	<b>4.32</b> <sup>13</sup>	5.1
<b>4.23</b> <sup>9</sup>	0	<b>4.33</b>	0
<b>4.24</b>	0	<b>4.34</b>	0
<b>4.25</b>	3.2		

These negative preliminary results may reflect the intrinsic limitations of the docking software algorithms used and the limited resolution within the available X-ray crystal structure of *E. coli* RNAP, which may lead to predictions affected by a significant margin of error. Therefore it was decided to re-synthesize the original hit molecule **4.1** and to evaluate it in the assay. It should also be noted that the original hit molecule **4.1** was identified using a different assay<sup>14</sup> to that used at Leeds, which differs in several aspects when compared to the SYBR Green assay<sup>15</sup> and this may also account for differences in the biological activity.

## 4.9 Synthesis and biological evaluation of the sulphonamide hit

Compound **4.1** was synthesised in low yield by drop-wise addition of the aromatic sulphonylchloride **4.8** to an excess of amine **4.35** at 0 °C in the presence of pyridine (Scheme 4.4).



**Scheme 4.4** Synthesis of **4.1** from **4.35** and **4.8**.

Following the biological evaluation of compound **4.1** with the SYBR Green assay<sup>15</sup> (please see Appendix I for details) surprisingly, no inhibition activity of RNAP at the concentration of 100  $\mu\text{M}$  was observed. This is consistent with the lack of any inhibitory activity observed for the two synthesized analogue libraries and confirms the dependency of inhibitory activity for this compound series on the precise assay used. Specifically, the SYBR Green assay,<sup>15</sup> as opposed to the assay reported<sup>14</sup> for the RNAP inhibitory activity of compound **4.1**, utilizes core RNAP in the absence of promoter sequences, primers and sigma factors and for these reasons, this assay has a limited capacity to evaluate inhibitors acting on the initiation step of the transcription or inhibitors interfering with the protein-protein interactions between the core enzyme and the sigma factor to form a holoenzyme which can bind a promoter. This may lead to the conclusion that literature hit **4.1** is potentially an initiation or protein-protein interaction inhibitor.

Unfortunately, the previously reported assay was not available to us for evaluation of the library of analogues of compound **4.1**.

## 4.10 Conclusions

As a direct result of the ability to very rapidly generate analogues, a SAR study on a reported RNAP inhibitor was performed using a combination of *in silico* studies followed by synthesis. Although the *in silico* studies have predicted a putative binding site for the original hit molecule, compound **4.1**, the two synthesized libraries obtained by combining docking studies with classical medicinal chemistry approach unfortunately did not possess significant biological activity in the Leeds RNAP inhibition assay. Following synthesis of the original hit molecule **4.1** and its biological evaluation revealed the absence of inhibitory activity in our assay which may be explained by the fact that the original hit molecule **4.1** was identified using a different assay<sup>14</sup> which differs in several aspects when compared to the SYBR Green assay.<sup>15</sup> Whilst both assays were validated for HTS studies on the same biological target, surprisingly their results were not comparable.

## 4.11 References

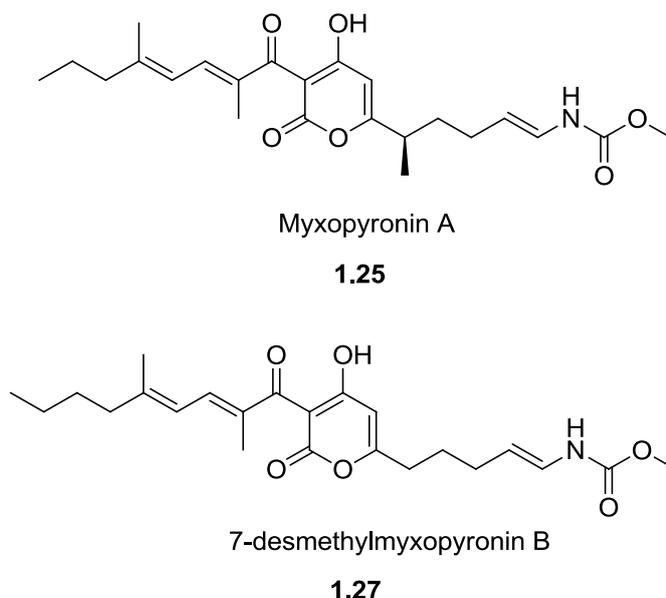
1. Wang, Y.; Suzek, T.; Zhang, J.; Wang, J.; He, S.; Cheng, T.; Shoemaker, B. A.; Gindulyte, A.; Bryant, S. H., *Nucl. Acids Res.* **2014**, *42*, 1075-1082.
2. <http://www.seas.upenn.edu/~pcmd/>.
3. Artsimovitch, I.; Vassilyeva, M. N.; Svetlov, D.; Svetlov, V.; Perederina, A.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Tahirov, T. H.; Vassilyev, D. G., *Cell* **2005**, *122*, 351-363.
4. Vassilyev, D. G.; Sekine, S.; Laptenko, O.; Lee, J.; Vassilyeva, M. N.; Borukhov, S.; Yokoyama, S., *Nature* **2002**, *417*, 712-719.
5. Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M. Y.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H., *Cell* **2008**, *135*, 295-307.
6. Vassilyev, D. G.; Svetlov, V.; Vassilyeva, M. N.; Perederina, A.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Artsimovitch, I., *Nat. Struct. Mol. Biol.* **2005**, *12*, 1086-1093.
7. Artsimovitch, I.; Chu, C.; Lynch, A. S.; Landick, R., *Science* **2003**, *302*, 650-654.
8. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., *Proteins* **2003**, *52*, 609-623.
9. Garcia Ruano, J. L.; Parra, A.; Marzo, L.; Yuste, F.; Mastranzo, V. M., *Tetrahedron* **2011**, *67*, 2905-2910.
10. Pasha, M. A.; Madhusudana Reddy, M. B., *Synth. Commun.* **2009**, *39*, 2928-2934.
11. Miyahara, M.; Nakadate, M.; Sueyoshi, S.; Tanno, M.; Kamiya, S., *Chem. Pharm. Bull.* **1982**, *30*, 4402-4406.
12. Yagodkin, A.; Lösckke, K.; Weisell, J.; Azhayev, A., *Tetrahedron* **2010**, *66*, 2210-2221.
13. Erb, B.; Rigo, B.; Pirotte, B.; Couturier, D., *J. Heterocycl. Chem.* **2002**, *39*, 15-28.
14. Kozlov, M.; Bergendahl, V.; Burgess, R.; Goldfarb, A.; Mustaev, A., *Anal. Biochem.* **2005**, *342*, 206-213.
15. Ohmichi, T.; Maki, A.; Kool, E. T., *Proc. Natl. Acad. Sci. U.S.A* **2002**, *99*, 54-59.



## 5. Scaffold-hopping and ligand-based virtual screening to assist the design and synthesis of small molecule inhibitors of the myxopyronin binding region

### 5.1 Physico-Chemical limitations of myxopyronin A

Myxopyronin A (Myx) **1.25**<sup>1</sup> (Figure 5.1) is a naturally occurring antibiotic which inhibits the initiation of transcription by interacting with the RNAP 'switch region', a hinge mediating the opening and the closure of the RNA 'clamp' which determines the open or closed state of the active centre cleft of the enzyme.



**Figure 5.1** Structures of **1.25** and **1.27**.

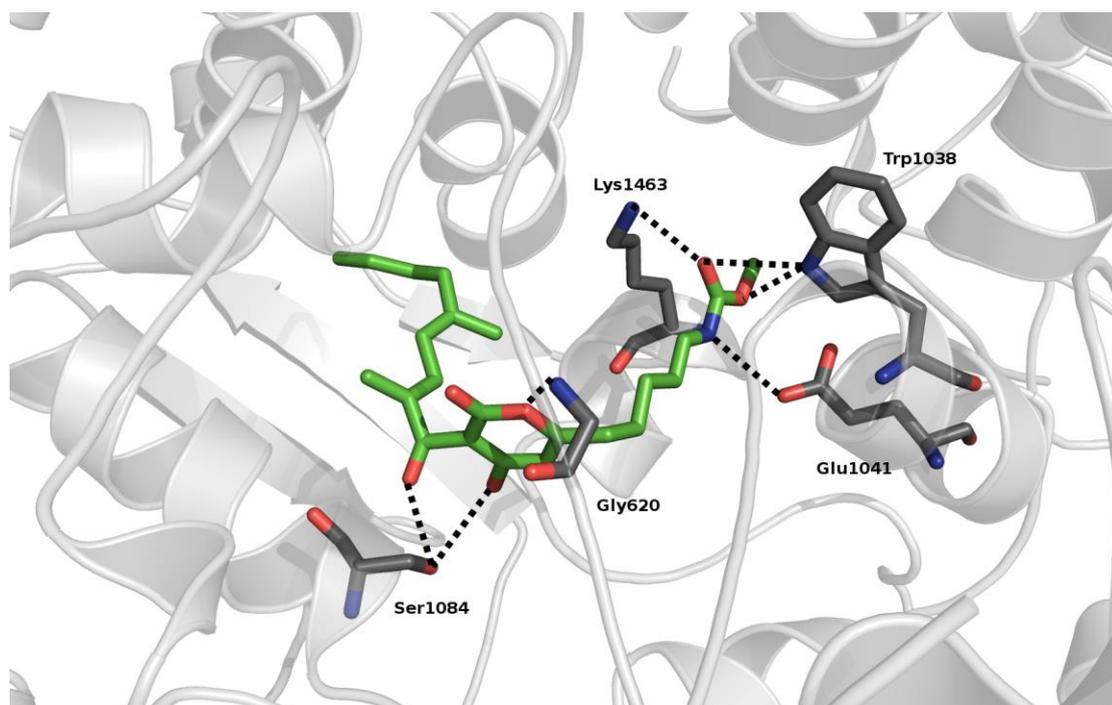
Myx **1.25** does not share cross-resistance with rifamycins. Unfortunately Myx **1.25** is not a viable drug lead because it does not possess suitable physicochemical properties. In particular, the dienone terminal side chain portion of the molecule which binds in a large hydrophobic pocket in the enzyme (Figure 5.2) results in poor pharmacokinetic properties due to the high LogP value and consequent strong binding to plasma proteins.<sup>2</sup> A previous study<sup>3</sup> revealed that binding to serum proteins significantly reduces the antibacterial activity of these compounds *in vivo*. The same authors<sup>3</sup> also underline issues with compound stability, in fact the Myx structure contains a

reactive Michael acceptor and a metabolically unstable carbamate group. The chemical synthesis of Myx analogues reported to date,<sup>1,4,5</sup> with the only exception of 7-desmethylmyxopyronin B derivative **1.27** (Figure 5.1), did not lead to more potent analogues and the design of more potent molecules has, until recently, been hampered by the lack of a high resolution structure of the RNAP-Myx complex.<sup>6</sup>

In the current project, in order to address the physico-chemical limitations of Myx, a scaffold hopping strategy was explored and the derived pharmacophoric hypothesis was used for a ligand-based virtual screening of commercial and in-house chemical libraries. The top ranking compounds were selected for biological evaluation on the basis of visual inspection and docking studies.

### 5.1.1 Scaffold-hopping strategy

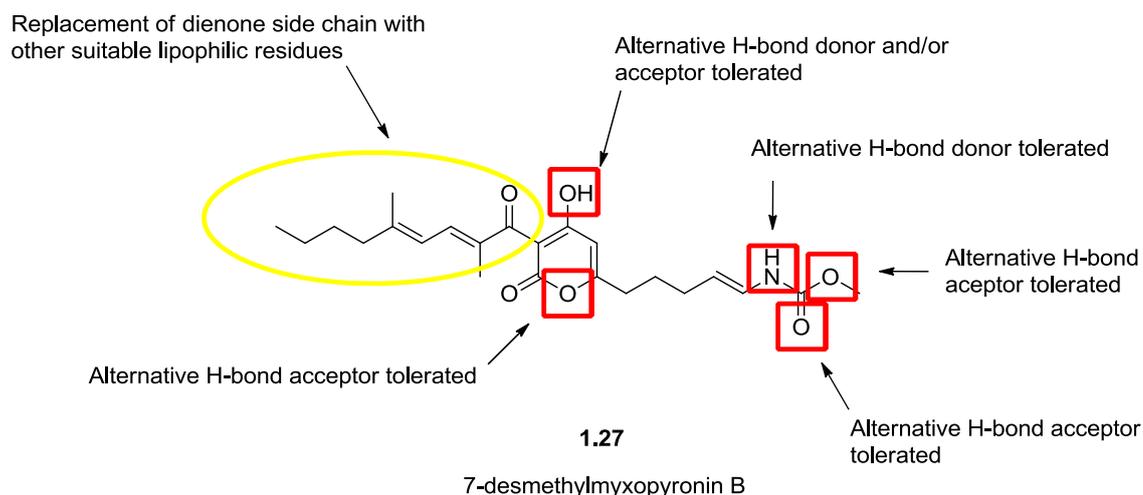
Analysis of the X-ray co-crystal structure of 7-desmethylmyxopyronin B **1.27** in complex with RNAP (PDB id: 3EQL<sup>7</sup>) reveals that 7-desmethylmyxopyronin B **1.27** is bound within an almost completely enclosed, predominantly hydrophobic crescent-shaped binding pocket (Figure 5.2).



**Figure 5.2** Co-crystal structure of **1.27**(PDB 3EQL).

Globally, there is a prevalence of hydrophobic interactions with the only exception being the  $\alpha$ -pyrone ring of 7-desmethylmyxopyronin B **1.27**, which is involved in key interactions near the cavity entrance *via* hydrogen bonds with Ser1084 and Gly620, while the carbonyl of the ene-carbamate side-chain is contacting Trp1038, Glu1041 and Lys1463 (Figure 5.2).

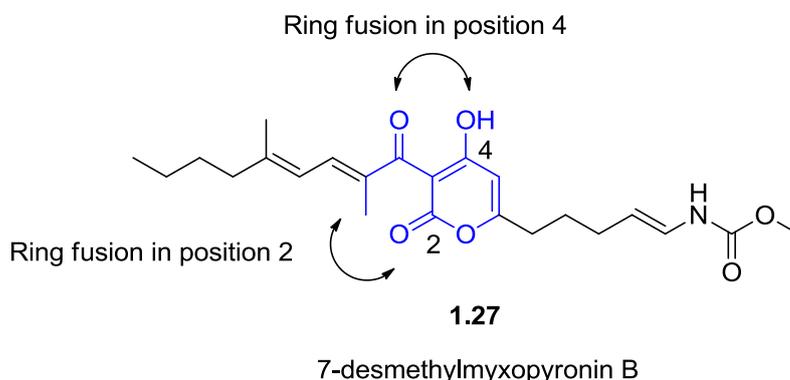
The scaffold-hopping strategy adopted in this project involved the preservation of the key specific hydrogen bonds at the level of the  $\alpha$ -pyrone ring and ene-carbamate side-chain and the replacement of the dienone side-chain with other suitable lipophilic residues. The key interaction points of the hydrogen bond network of 7-desmethylmyxopyronin B **1.27** (Figure 5.3) involve the hydroxyl group of the  $\alpha$ -pyrone which could be replaced by an hydrogen bond donor and/or acceptor and by the ring oxygen where an alternative hydrogen bond acceptor could be tolerated. With regard to the ene-carbamate moiety, it was reasoned that carbonyl oxygen could be replaced by any other suitable hydrogen bond acceptor as well as the methoxy group while the carbamate nitrogen could be substituted with any other suitable hydrogen bond donor capable of interacting with Glu1041 (Figure 5.2).



**Figure 5.3** Key interaction points of the hydrogen bond network of **1.27**.

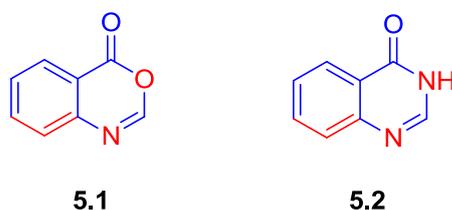
A plausible scaffold-hopping strategy could involve the replacement of the  $\alpha$ -pyrone core with an alternative aromatic or heteroaromatic moiety derived

from the ring fusion of the  $\alpha$ -pyrone at position 2 or alternatively, at position 4 with the concurrent substitution of the carbonyl belonging to the dienone with a carbon atom (Figure 5.4).



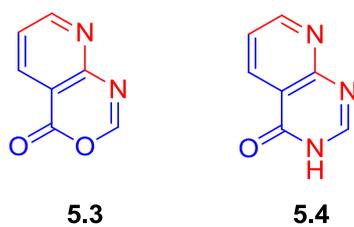
**Figure 5.4** Design strategy based on **1.27**.

As an example of this strategy, ring fusion at position 2 of the  $\alpha$ -pyrone could result in a 4H-benzo[d][1,3]oxazin-4-one or a quinazolin-4(3H)-one based heterocyclic core as shown in **5.1** and **5.2** (Figure 5.5). In particular, the ring nitrogen replaced the ring oxygen of 7-desmethylmyxopyronin B **1.27**.



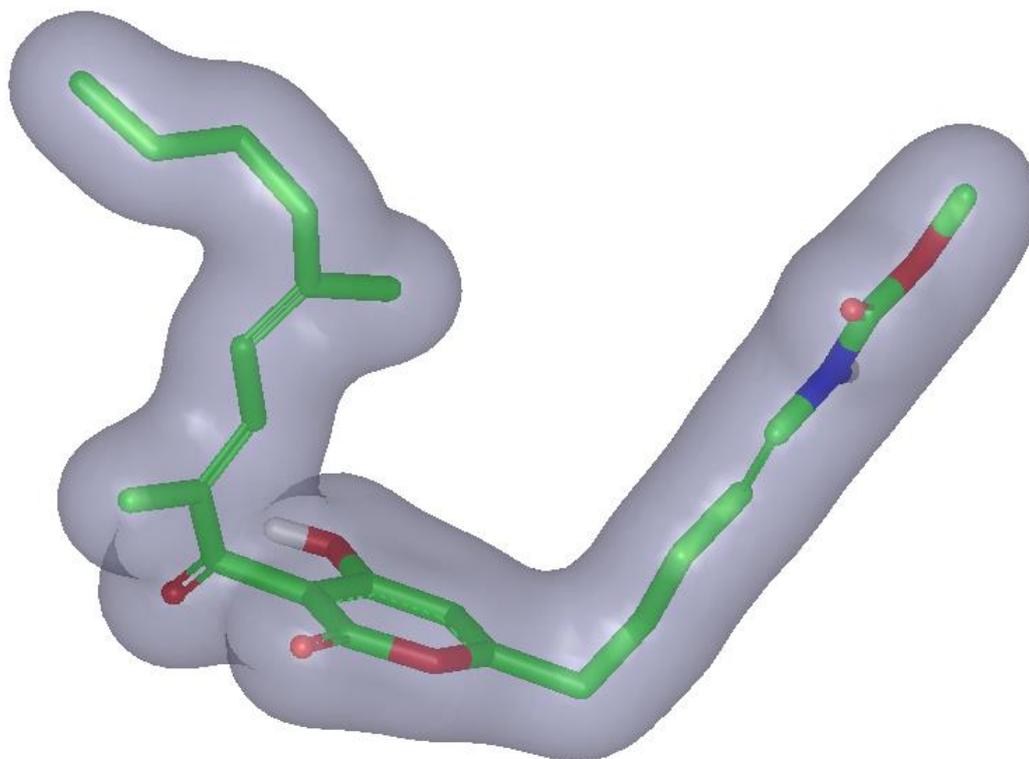
**Figure 5.5** Structures of **5.1** and **5.2**, original  $\alpha$ -pyrone core shown in blue, closure points and modified atoms in red.

Alternatively, following ring fusion at the position 4 of the  $\alpha$ -pyrone, a 4H-pyrido[2,3-d][1,3]oxazin-4-one or a pyrido[2,3-d]pyrimidin-4(3H)-one based heterocyclic core as shown in **5.3** and **5.4** could replace the  $\alpha$ -pyrone core (Figure 5.6) where a nitrogen atom replaced the hydroxyl of 7-desmethylmyxopyronin B **1.27**.

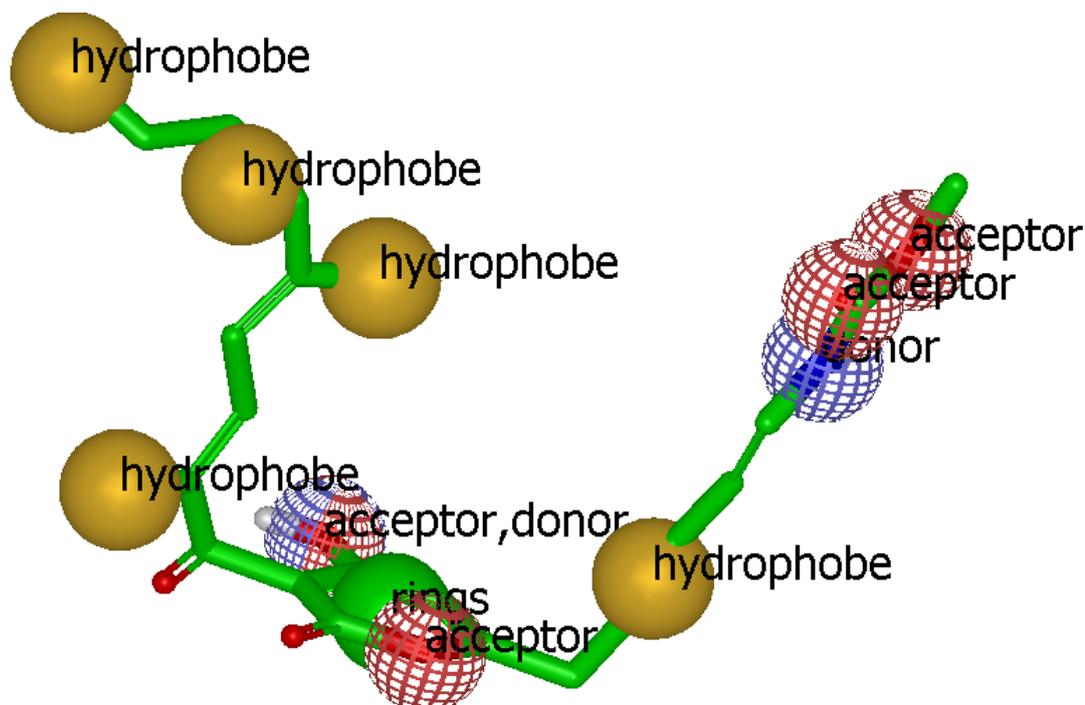


**Figure 5.6** Structures of **5.3** and **5.4**, original  $\alpha$ -pyrone core shown in blue, closure points and modified atoms in red.

In order to apply the previously outlined scaffold-hopping strategy, a pharmacophore and shape-based query was elaborated within the graphical interface of ROCS software, vROCS as shown below (Figure 5.7 and Figure 5.8).



**Figure 5.7** Shape query elaborated in vROCS represented as a grey cloud, structure of 7-desmethylmyxopyronin B **1.27** structure represented as sticks.



**Figure 5.8** Pharmacophore query elaborated in vROCS represented as coloured spheres centred on 7-desmethylmyxopyronin B **1.27** structure represented as sticks.

### 5.1.2 Ligand-based virtual screening protocol

Ligand-based virtual screening was performed for the specified chemical query taking in account the overall shape and volume of 7-desmethylmyxopyronin B **1.27** in addition to matching the pharmacophoric points.

The in-house medicinal chemistry and chemical biology technology group database (MCCB database) comprising 26,493 molecules from Albany Molecular Research Inc. (AMRI), Chembridge and Asinex commercial databases along with another two commercial databases, the Peakdale (15,339 molecules) and the Specs diversity set (17,520 structures with molecular similarity expressed as Tanimoto coefficient not greater than 0.7), were downloaded from the ZINC database.<sup>8</sup>

A library of conformers was generated from the previously downloaded databases using OMEGA software version 2.3.2, limiting the maximum number of conformers per molecule to 50 using the default settings.

The biologically active conformation of 7-desmethylmyxopyronin B was extracted from its crystallographic complex with RNAP (PDB id: 3EQL<sup>7</sup>) and atom types and bond orders were amended when necessary with Maestro software. The molecular structure was imported into vROCS version 3.1.2 where the pharmacophoric features were manually specified over the key atoms identified from the previous analysis (Section 5.1.1), and the overall shape and volume were automatically perceived by the software.

The previously prepared library of conformers was screened on the generated query using an Explicit Mills Dean<sup>9</sup> colour force field and leaving all the other settings to their default values. ROCS automatically ranked the top 500 molecules using the TanimotoCombo scoring function and the result was visually inspected using vROCS to assess the quality of overlapping to the shape and chemical features of the query molecule.

A short list of around 100 molecules was selected for further analysis and filtering *via* docking studies. The aim of these studies was to verify that compounds with good chemical and shape complementarity were also predicted to adopt a reasonable binding mode inside the targeted site, avoiding steric clashes and reproducing as much as possible the 7-desmethylmyxopyronin B interaction pattern.

### 5.1.3 General docking protocol

The crystal structure of the *T. thermophilus* RNA polymerase holoenzyme in complex with the antibiotic 7-desmethylmyxopyronin B (PDB id: 3EQL<sup>7</sup>) was downloaded from the Protein Databank ([www.rcsb.org](http://www.rcsb.org)).

The binding site within the RNAP for Myx was defined as a 15 Å cut of the protein surrounding the co-crystallised ligand.

The dockings runs were performed using GOLD<sup>10</sup> v4.0.1 (CCDC, Cambridge, UK) docking software by using the default settings for the genetic searching algorithm generating 100 poses for each ligand and the compounds were ranked with the Chemscore scoring function.

A post-docking analysis was performed on the different docking solutions for each ligand. Cluster analysis using an average linkage rule was performed

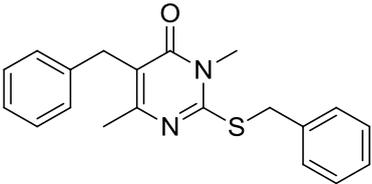
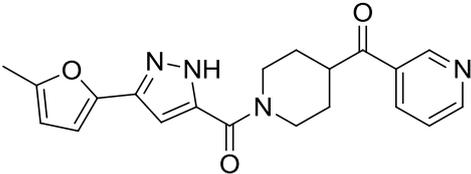
using an internal module within the software. Only the best ranked poses of the most populated cluster were considered as the most likely and representative ones. In addition, the docked poses were analysed in terms of overall score, good shape complementarity to the cavity, possibility to establish specific H-bonds to the protein in addition to the less specific hydrophobic contacts, minimal exposure of hydrophobic groups to solvent and absence of an excessive torsional strain into the docked conformer.

#### 5.1.4 Results

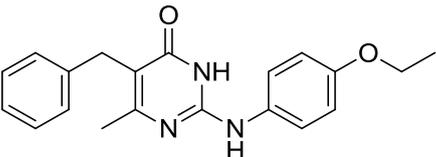
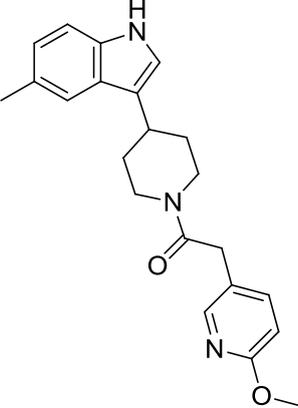
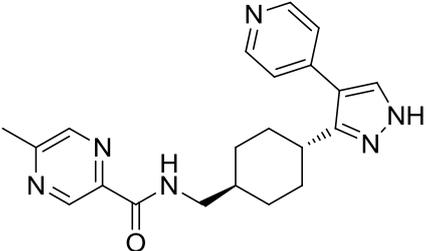
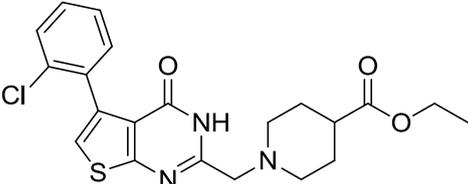
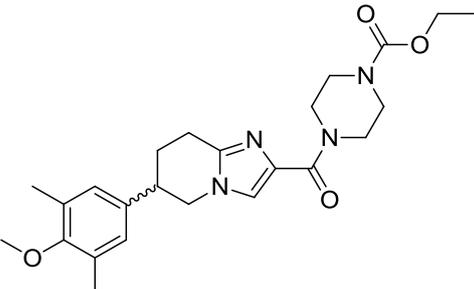
The list of selected compounds for biological evaluation along with predicted properties and docking score is reported in Table 5.1. Interestingly, the results are, in part, concordant with the previously mentioned scaffold hopping considerations (Section 5.1.1) and the number of 4H-benzo[d][1,3]oxazin-4-one and quinazolin-4(3H)-one-based scaffolds in the best-ranked virtual hits list could not be ignored when prioritizing the compounds for biological evaluation.

Compounds **5.15—5.21** were not available for purchase but because of their fair synthetic accessibility, good final ranking and relevant putative binding mode, were synthesized (Section 5.2).

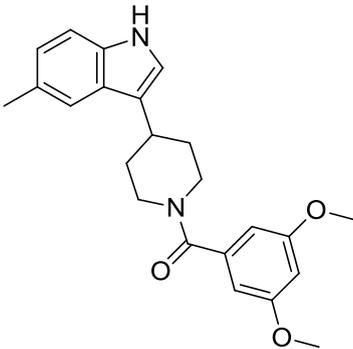
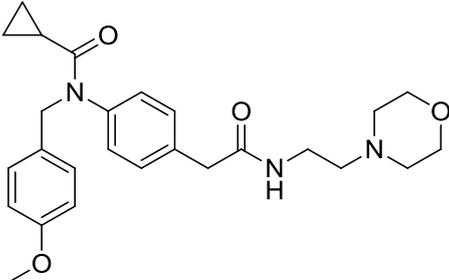
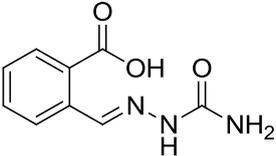
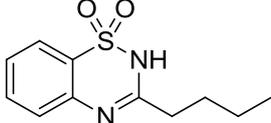
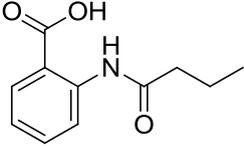
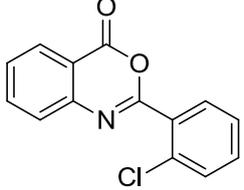
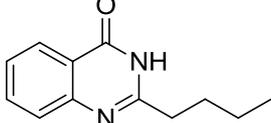
**Table 5.1** Compounds selected for biological evaluation

Comp. number	Structure	M.W.	ALogP (A)	Chem score	Rule of 5 violations (B)
5.5		336	3.23	36.4	0
5.6		364	2.36	30.9	0

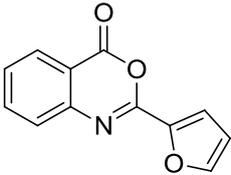
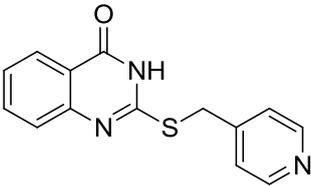
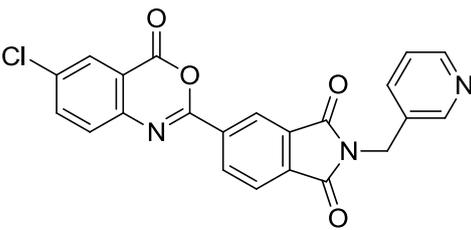
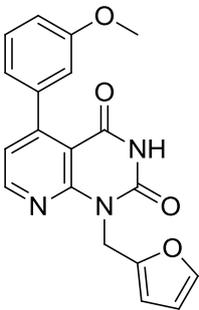
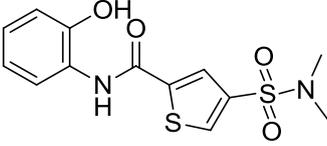
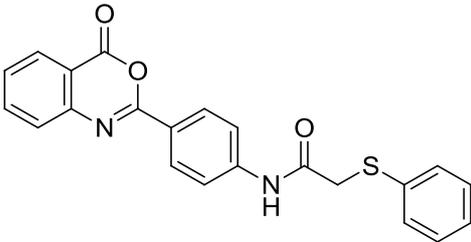
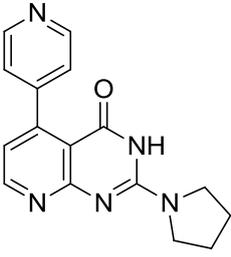
**Table 5.1 Continued**

5.7		335	2.97	33.1	0
5.8		363	3.90	32.9	0
5.9		376	2.49	31.8	0
5.10		420	4.04	32.1	0
5.11		431	4.09	30.4	0
5.12		440	2.70	30.2	0

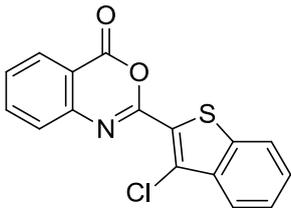
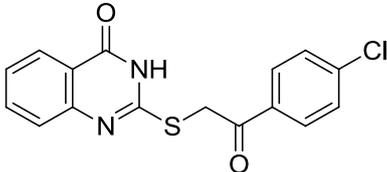
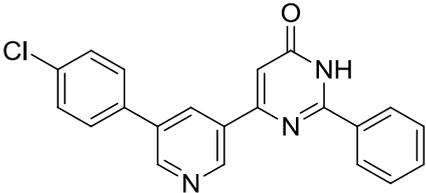
**Table 5.1 Continued**

<p><b>5.13</b></p>		<p>378</p>	<p>4.48</p>	<p>34.3</p>	<p>0</p>
<p><b>5.14</b></p>		<p>451</p>	<p>2.92</p>	<p>35.6</p>	<p>0</p>
<p><b>5.15<sup>11</sup></b></p>		<p>206</p>	<p>-0.42</p>	<p>29.5</p>	<p>0</p>
<p><b>5.16<sup>12</sup></b></p>		<p>238</p>	<p>1.72</p>	<p>23.7</p>	<p>0</p>
<p><b>5.17<sup>13</sup></b></p>		<p>207</p>	<p>2.09</p>	<p>25.5</p>	<p>0</p>
<p><b>5.18<sup>14</sup></b></p>		<p>257</p>	<p>3.49</p>	<p>29.7</p>	<p>0</p>
<p><b>5.19<sup>15</sup></b></p>		<p>202</p>	<p>2.22</p>	<p>24.3</p>	<p>0</p>

**Table 5.1 Continued**

<p><b>5.20<sup>14</sup></b></p>		<p>213</p>	<p>2.41</p>	<p>27.2</p>	<p>0</p>
<p><b>5.21</b></p>		<p>269</p>	<p>1.75</p>	<p>28.9</p>	<p>0</p>
<p><b>5.22</b></p>		<p>417</p>	<p>3.02</p>	<p>26.1</p>	<p>0</p>
<p><b>5.23</b></p>		<p>349</p>	<p>2.21</p>	<p>24.2</p>	<p>0</p>
<p><b>5.24</b></p>		<p>326</p>	<p>1.96</p>	<p>24.6</p>	<p>0</p>
<p><b>5.25</b></p>		<p>388</p>	<p>4.14</p>	<p>30.4</p>	<p>0</p>
<p><b>5.26</b></p>		<p>293</p>	<p>0.63</p>	<p>24.6</p>	<p>0</p>

**Table 5.1 Continued**

<b>5.27</b>		313	4.93	30.7	0
<b>5.28</b>		330	2.91	27.3	0
<b>5.29</b>		359	2.81	36.3	0

(A) ALogP calculated with ALOGPS 2.1 software<sup>16</sup>

(B) Rule of five defined as reported in Lipinski's paper,<sup>17</sup> parameters calculated with Molsoft Drug Likeness Prediction:<sup>18</sup>

150 ≤ M.W. ≤ 500

ALogP ≤ 5

Hydrogen bond donors ≤ 5

Hydrogen bond acceptors ≤ 10

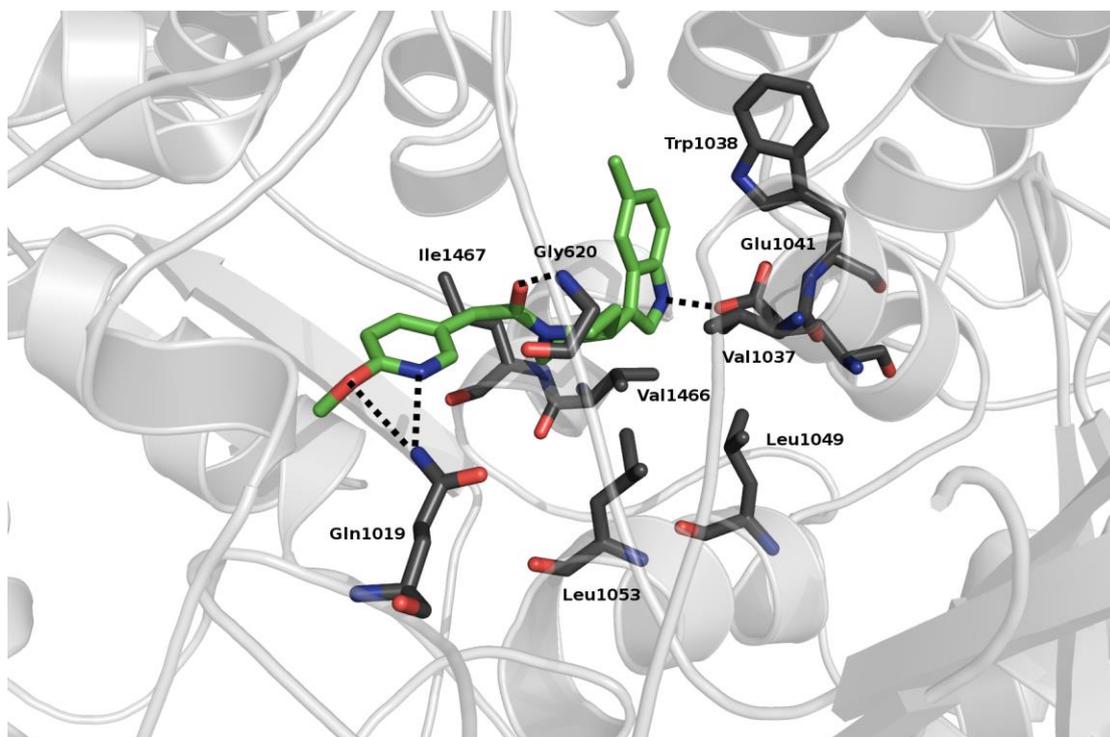
Rotatable bonds ≤ 7

Polar surface area ≤ 150

Examples of predicted binding mode for the most relevant and best ranked compounds are shown below.

The predicted binding mode of compound **5.8** is shown in Figure 5.9. The following interactions between the ligand **5.8** and the myxopyronin binding site were identified: the NH of the indole ring is predicted to form an H-bond with the side chain of Glu1041 while the carbonyl of the amide bond is interacting *via* an H-bond with the backbone NH of Gly620. The nitrogen of the pyridyl ring is predicted to form an H-bond with the side chain of Gln1019 while the oxygen of the methoxy group is interacting *via* an H-bond with the side chain of Gln1019.

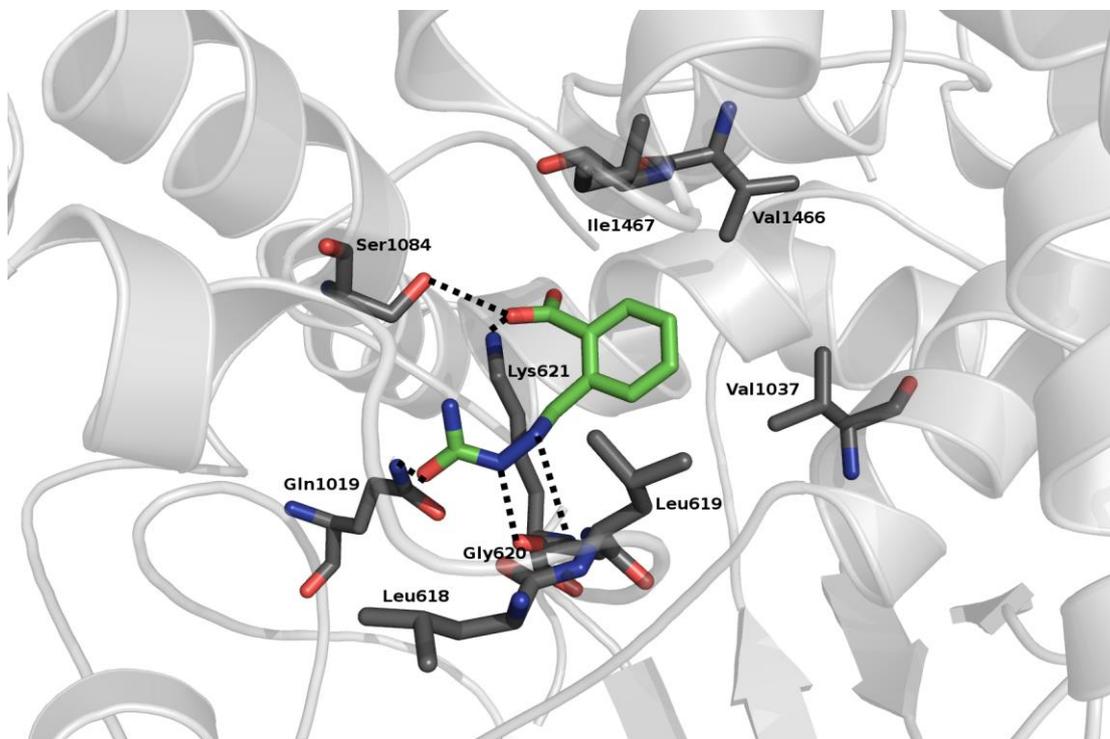
The side chains of Leu1049, Leu1053, Val1466, Ile1467, Val1037 and Trp1038 are predicted to be involved in hydrophobic interactions with the indole and the piperidine ring. Notably, H-bond interactions with Glu1041 and Gly620 are also present in the X-ray co-crystal structure of 7-desmethylmyxopyronin B **1.27** in complex with RNAP (PDB id: 3EQL<sup>7</sup>) (Figure 5.2) at the level of the  $\alpha$ -pyrone ring.



**Figure 5.9** Predicted binding mode of **5.8**.

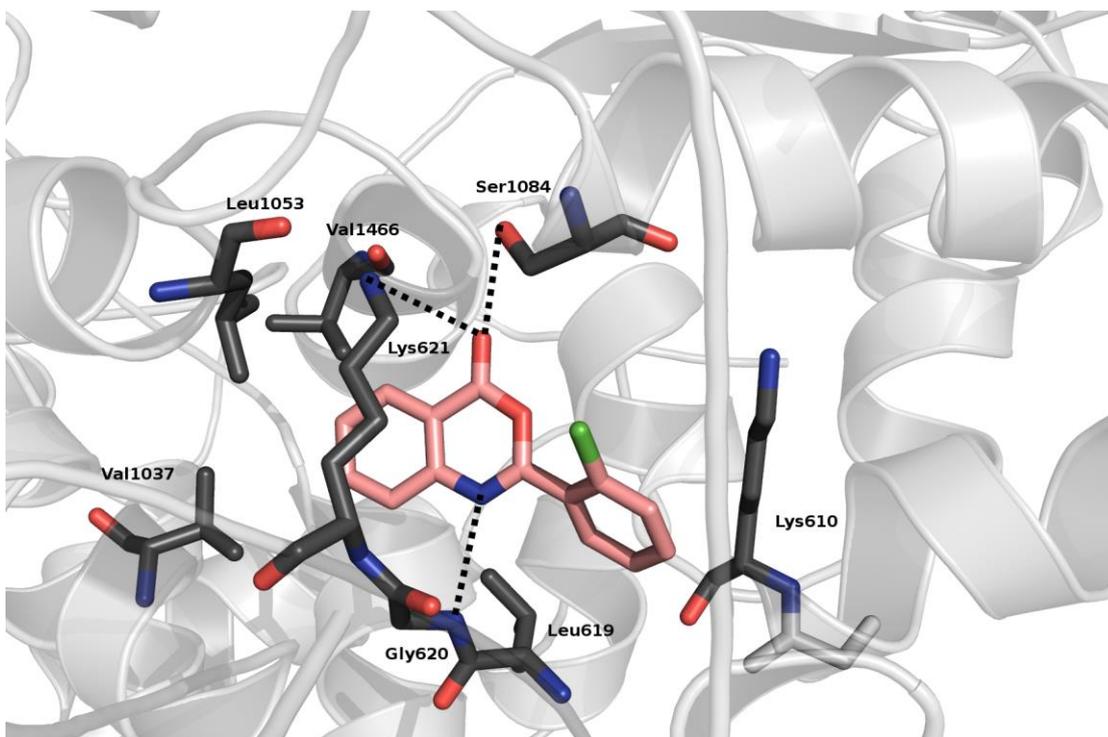
The binding mode of compound **5.15**<sup>11</sup> was predicted as shown in Figure 5.10. The following interactions between the ligand **5.15**<sup>11</sup> and the myxopyronin binding site were identified: the carboxylate is predicted to form an H-bond with the side chain of Ser1084, the carbonyl of the semicarbazide moiety is predicted to make an H-bond with the side chain of Gln1019 and Lys621 while the amidic nitrogen and the neighbouring azomethine nitrogen of the same moiety are interacting *via* H-bonds with the backbone carbonyl of Lys621 and the backbone NH of Gly620. The aromatic ring is predicted to be involved in hydrophobic interactions with the side chains of Val1466, Val1037 and

Leu1619. Interestingly the previously mentioned H-bond interactions with Ser1084 and Gly620 are also present in the X-ray co-crystal structure of 7-desmethylmyxopyronin B **1.27** in complex with RNAP (PDB id: 3EQL<sup>7</sup>) (Figure 5.2).



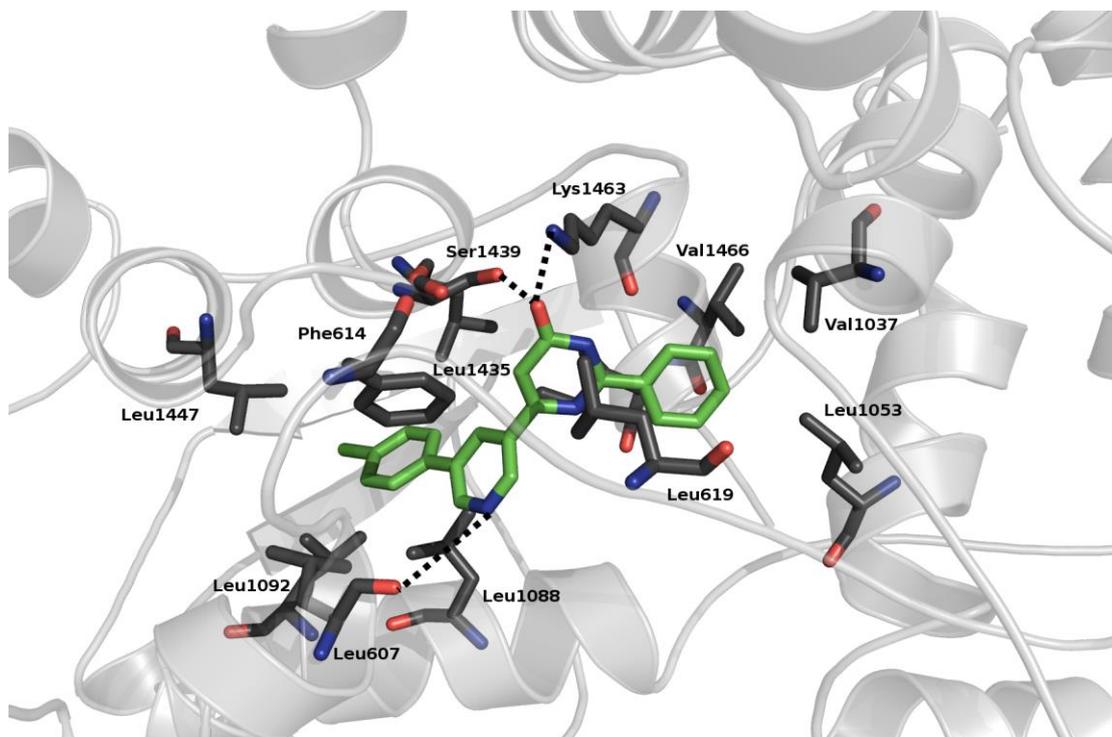
**Figure 5.10** Predicted binding mode of **5.15**.<sup>11</sup>

The binding mode of compound **5.18**<sup>14</sup> was predicted as shown in Figure 5.11. The following interactions between the ligand **5.18**<sup>14</sup> and the myxopyronin binding site were identified: the carbonyl of the pyrimidinone ring is predicted to establish an extended H-bond network with the side chains of Lys621, Ser1084 while the nitrogen of the same moiety is interacting *via* H-bond with the backbone NH of Gly620. The pyrimidinone and the 2-chlorophenyl moieties are involved in hydrophobic interactions with the side chains of Val1466, Leu1053, Val1037, Lys621 and Lys610. Interestingly, the two previously mentioned H-bond interactions are also present in the X-ray co-crystal structure of 7-desmethylmyxopyronin B **1.27** in complex with RNAP (PDB id: 3EQL<sup>7</sup>) (Figure 5.2) at the level of the  $\alpha$ -pyrone ring.



**Figure 5.11** Predicted binding mode of **5.18**.<sup>14</sup>

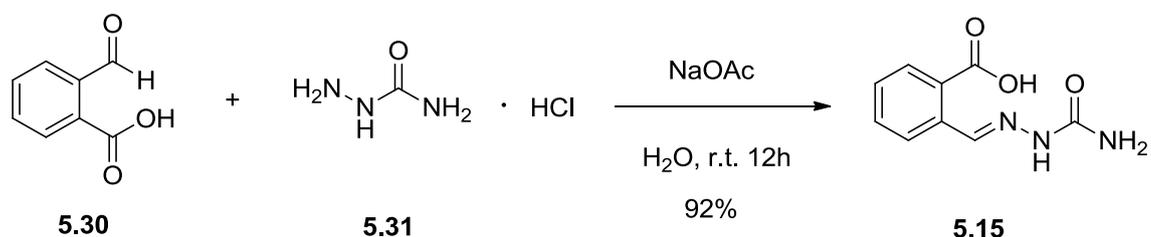
The binding mode of compound **5.29** was predicted as shown in figure 5.12. The following interactions between the ligand **5.29** and the myxopyronin binding site were identified: the carbonyl of the pyrimidinone ring is predicted to establish an H-bond network with the side chains of Ser1439 and Lys1463. Notably, the carbonyl of the pyrimidinone ring is mimicking the H-bond of the ene-carbamate carbonyl with Lys1463 present in the X-ray co-crystal structure of 7-desmethylmyxopyronin B **1.27** in complex with RNAP (PDB id: 3EQL<sup>7</sup>) (Figure 5.2). The nitrogen of the pyridyl ring is predicted to form an H-bond with the backbone carbonyl of Leu607; a T-shaped stacked interaction between the side chain aromatic ring of the Phe614 and the pyridyl group is observed while the other aromatic rings are predicted to be involved in very extended hydrophobic interactions with the side chains of Leu1088, Leu1092, Leu607, Leu1447, Leu1435, Val1466, Leu619, Leu1053 and Val1037.



**Figure 5.12** Predicted binding mode of **5.29**.

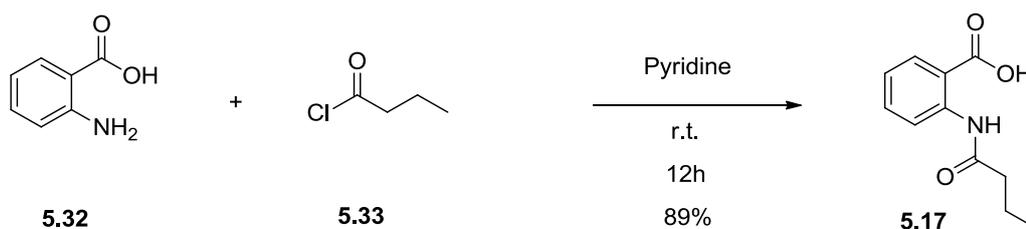
## 5.2 Synthesis

Compound **5.15**<sup>11</sup> was prepared in high yield following the procedure reported by *Kumar*,<sup>19</sup> reacting carboxybenzaldehyde **5.30** with semicarbazide hydrochloride **5.31** in an aqueous solution of sodium acetate (Scheme 5.1):



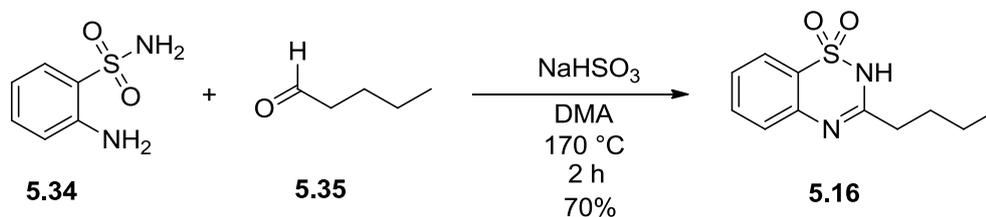
**Scheme 5.1** Synthesis of **5.15**.<sup>11</sup>

The reaction between anthranilic acid **5.32** with an excess of butyryl chloride **5.33** in pyridine at room temperature afforded the formation of **5.17**<sup>13</sup> in good yield (Scheme 5.2):



**Scheme 5.2** Synthesis of **5.17**.<sup>13</sup>

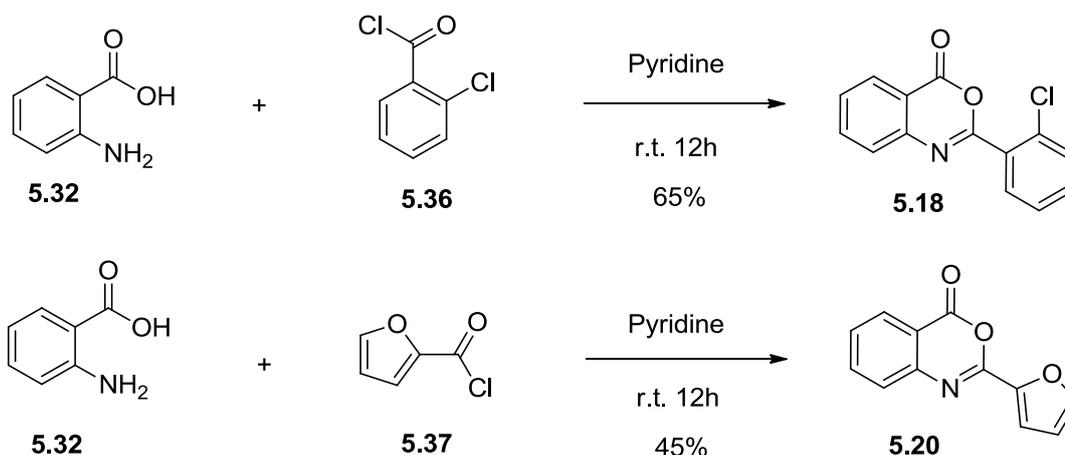
With regard to benzo[e][1,2,4]thiadiazine 1,1-dioxide derivative **5.16**,<sup>12</sup> this was prepared in good yield by reacting a solution of sulphonamide **5.34** in DMA with aldehyde **5.35** in presence of an excess of sodium bisulphite under reflux as reported by *Imai et al*<sup>20</sup> (Scheme 5.3).



**Scheme 5.3** Synthesis of **5.16**.<sup>12</sup>

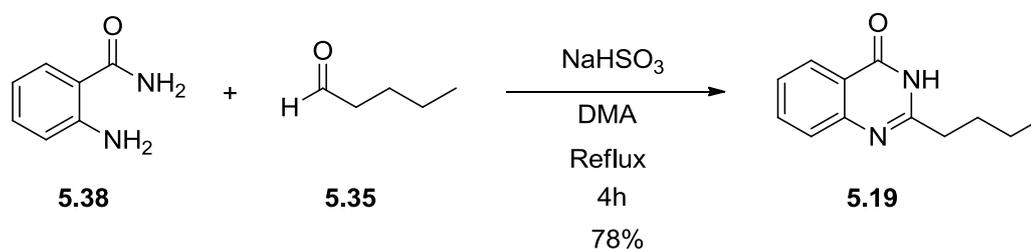
The synthesis of 2-aryl-4H-benzo[d][1,3]oxazin-4-one derivatives **5.18**<sup>14</sup> and **5.20**<sup>14</sup> was performed (Scheme 5.4) following the literature procedure as reported by Bain and Smalley.<sup>21</sup>

The one-pot reaction between equimolar quantities of anthranilic acid **5.32** with aroylchloride **5.36** or aroylchloride **5.37** in a solution of pyridine and toluene under reflux gave, respectively, the benzoxazinones **5.18**<sup>14</sup> and **5.20**<sup>14</sup> in moderate yields (Scheme 5.4):



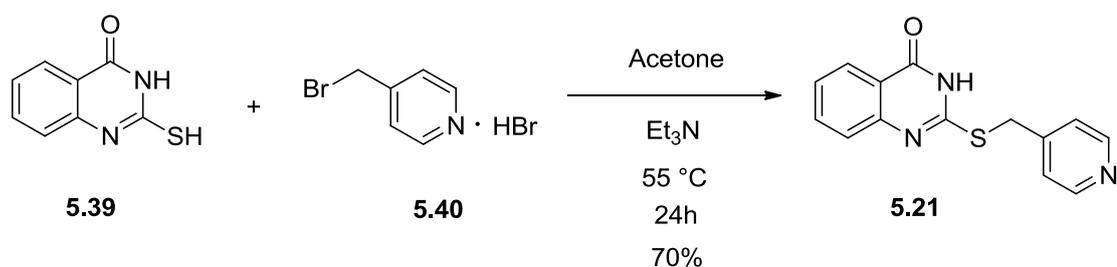
**Scheme 5.4** Synthesis of **5.18**<sup>14</sup> and **5.20**.<sup>14</sup>

Quinazolinone **5.19**<sup>15</sup> was prepared in good yield following the procedure reported by *Imai et al.*<sup>20</sup> by refluxing a solution of anthranilamide **5.38** in DMA and aldehyde **5.35** in the presence of an excess of sodium bisulphite (Scheme 5.5).



**Scheme 5.5** Synthesis of **5.19**.<sup>15</sup>

The nucleophilic substitution reaction by heating an acetone solution of equimolar quantities of 2-mercaptoquinazolin-4(3H)-one **5.39** and 4-(bromomethyl)pyridine hydrobromide **5.40** in presence of triethylamine afforded the quinazolinone **5.21** in good yield (Scheme 5.6).



**Scheme 5.6** Synthesis of **5.21**.

### 5.3 Biological results

Compounds **5.5—5.29** were evaluated in the RNAP SYBR green assay<sup>22</sup> (please see Appendix I for details) and weak or no inhibition activity was found (Table 5.2).

**Table 5.2** *In vitro* RNAP percentage inhibition at 100  $\mu$ M of the purchased and synthesised molecules.

Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M	Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M
<b>5.5</b>	7.4	<b>5.18</b> <sup>14</sup>	Insoluble
<b>5.6</b>	7.1	<b>5.19</b> <sup>15</sup>	0
<b>5.7</b>	10.2	<b>5.20</b> <sup>14</sup>	0
<b>5.8</b>	13.1	<b>5.21</b>	0
<b>5.9</b>	1.5	<b>5.22</b>	0
<b>5.10</b>	0.9	<b>5.23</b>	11.8
<b>5.11</b>	5.2	<b>5.24</b>	7.2
<b>5.12</b>	2.6	<b>5.25</b>	5.1
<b>5.13</b>	4.2	<b>5.26</b>	0
<b>5.14</b>	0	<b>5.27</b>	0
<b>5.15</b> <sup>11</sup>	0	<b>5.28</b>	0
<b>5.16</b> <sup>12</sup>	0	<b>5.29</b>	12.3
<b>5.17</b> <sup>13</sup>	0		

## 5.4 Conclusions

Following the analysis of the X-ray co-crystal structure of 7-desmethyl-myxopyronin B **1.27** in complex with RNAP (PDB id: 3EQL<sup>7</sup>) (Figure 5.2), a scaffold-hopping strategy has been elaborated and a pharmacophore- and shape-based virtual screening approach followed by docking studies has been performed in order to overcome the unfavourable pharmacokinetic and chemical properties of myxopyronin A (Myx) **1.25**.<sup>1</sup>

A series of molecules was selected from commercial databases or synthesized for biological evaluation. Whilst all the virtual hits were predicted to have improved physicochemical properties, good binding mode and better synthetic accessibility when compared to Myx, unfortunately no active molecules were identified. A possible reason for the lack of biological activity of the selected virtual hits could derive from the nature of the designed protocol. In fact, selection of compounds relied mainly on the ligand-based stage while docking studies confirmed shape, chemical complementarity and quality of binding mode within the targeted site on a narrow number of candidates. Whilst the shape screen and the matching of selected pharmacophoric points allowed the restriction of the number of candidates, the binding site sampling performed at this stage was limited to selected specific interactions of 7-desmethyl-myxopyronin B **1.27** within RNAP and, for this reason, underestimated the possibility of alternative binding modes for chemically dissimilar molecules within a large binding site.

## 5.5 References

1. Hu, T.; Schaus, J. V.; Lam, K.; Palfreyman, M. G.; Wuonola, M.; Gustafson, G.; Panek, J. S., *J. Org. Chem.* **1998**, *63*, 2401-2406.
2. Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M. Y.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H., *Cell* **2008**, *135*, 295-307.
3. Moy, T. I.; Daniel, A.; Hardy, C.; Jackson, A.; Rehrauer, O.; Hwang, Y. S.; Zou, D.; Nguyen, K.; Silverman, J. A.; Li, Q.; Murphy, C., *FEMS Microbiol. Lett.* **2011**, *319*, 176-179.
4. Doundoulakis, T.; Xiang, A. X.; Lira, R.; Agrios, K. A.; Webber, S. E.; Sisson, W.; Aust, R. M.; Shah, A. M.; Showalter, R. E.; Appleman, J. R.; Simonsen, K. B., *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5667-5672.
5. Lira, R.; Xiang, A. X.; Doundoulakis, T.; Biller, W. T.; Agrios, K. A.; Simonsen, K. B.; Webber, S. E.; Sisson, W.; Aust, R. M.; Shah, A. M.; Showalter, R. E.; Banh, V. N.; Steffy, K. R.; Appleman, J. R., *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6797-6800.
6. Srivastava, A.; Talaue, M.; Liu, S.; Degen, D.; Ebright, R. Y.; Sineva, E.; Chakraborty, A.; Druzhinin, S. Y.; Chatterjee, S.; Mukhopadhyay, J.; Ebright, Y. W.; Zozula, A.; Shen, J.; Sengupta, S.; Niedfeldt, R. R.; Xin, C.; Kaneko, T.; Irschik, H.; Jansen, R.; Donadio, S.; Connell, N.; Ebright, R. H., *Curr. Opin. Microbiol.* **2011**, *14*, 532-543.
7. Belogurov, G. A.; Vassilyeva, M. N.; Sevostyanova, A.; Appleman, J. R.; Xiang, A. X.; Lira, R.; Webber, S. E.; Klyuyev, S.; Nudler, E.; Artsimovitch, I.; Vassilyev, D. G., *Nature* **2009**, *457*, 332-335.
8. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., *J. Chem. Inf. Model.* **2012**, *52*, 1757-1768.
9. Mills, J. E. J.; Dean, P. M., *J Computer-Aided Mol Des* **1996**, *10*, 607-622.
10. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., *Proteins* **2003**, 609-623.
11. Andersen, R. A.; Coates, G. E., *J. Chem. Soc., Dalton Trans.* **1974**, 1171-1180.
12. Raffa, L.; Pecorari, P., *Farmaco-Edizione Scientifica* **1966**, *21*, 196-198.
13. Duffield, A. M.; Jefferies, P. R., *Aust. J. Chem.* **1963**, *16*, 292-294.

14. Bain, D. I.; Smalley, R. K., *J. Chem. Soc. C* **1968**, 1593-1594.
15. Dabiri, M.; Salehi, P.; Mohammadi, A. A.; Baghbanzadeh, M., *Synth. Commun.* **2005**, *35*, 279-287.
16. Tetko, I. V.; Tanchuk, V. Y., *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1136-1145.
17. Lipinski, C. A., *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235-249.
18. O'Neill, A.; Oliva, B.; Storey, C.; Hoyle, A.; Fishwick, C.; Chopra, I., *Antimicrob. Agents Ch.* **2000**, *44*, 3163-3166.
19. Kumar, S., *Turk. J. Chem.* **2011**, *35*, 99-108.
20. Imai, Y.; Sato, S.; Takasawa, R.; Ueda, M., *Synthesis* **1981**, *1981*, 35-36.
21. Bain, D. I.; Smalley, R. K., *J. Chem. Soc. C* **1968**, *13*, 1593-1594.
22. Ohmichi, T.; Maki, A.; Kool, E. T., *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 54-59.

## 6. Critical reassessment of a previously reported computational study<sup>1</sup> and development of a combined ligand and structure-based virtual screening protocol

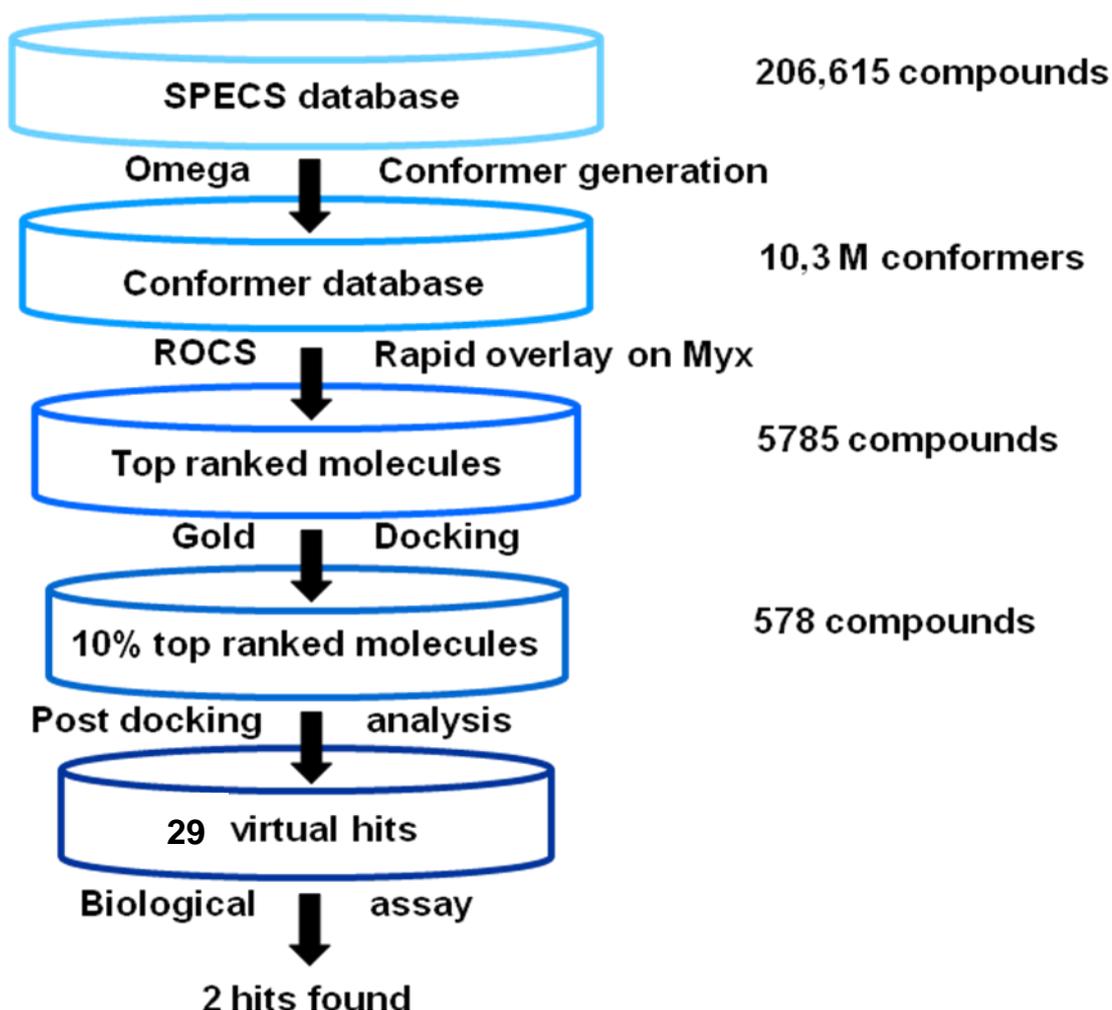
### 6.1 Limitations of a previously reported computational study<sup>1</sup>

A pharmacophore modelling and structure-based virtual screening study on the RNAP 'switch region' has been recently reported.<sup>1</sup> Although a list of 27 virtual hits out of the 321,374 compounds within the screened database were selected and underlined for further study by the authors<sup>1</sup> but no biological data were reported to support this suggestion or test the validity of the protocol. Following detailed inspection of this paper<sup>1</sup>, several questionable points and inaccuracies were found in addition to the lack of a biological validation, and these are summarized below:

- A) The screened database, according to the given protocol, was filtered prior to hit selection following the Lipinski Rule of five but several selected virtual hits have at least one or more violations of this rule.
- B) The developed pharmacophoric query did not include any hydrophobic features despite these interactions being important within the targeted binding site.
- C) The authors claimed a crystallographic water molecule close to the antibiotic molecule was conserved and this was included in the crystallographic protein structure used for docking calculations. Following analysis of the crystal structure of the *T. thermophilus* RNAP holoenzyme in complex with the antibiotic myxopyronin A **1.25** (PDB id: 3DXJ<sup>2</sup>), a water molecule interacting with the ene-carbamate group is found only in one biological assembly out of the two present in the asymmetric unit, while in the case of the crystal structure of the *T. thermophilus* RNAP holoenzyme in complex with the antibiotic 7-desmethylmyxopyronin **1.27** (PDB id.: 3EQL<sup>3</sup>) no water molecules are in proximity of the antibiotic molecule.

## 6.2 Critical reassessment of the protocol

A critical reassessment of this theoretical study<sup>1</sup> has been performed by elaborating a new and more rigorous virtual screening protocol which combined ligand and structure-based techniques and summarized in Figure 6.1.



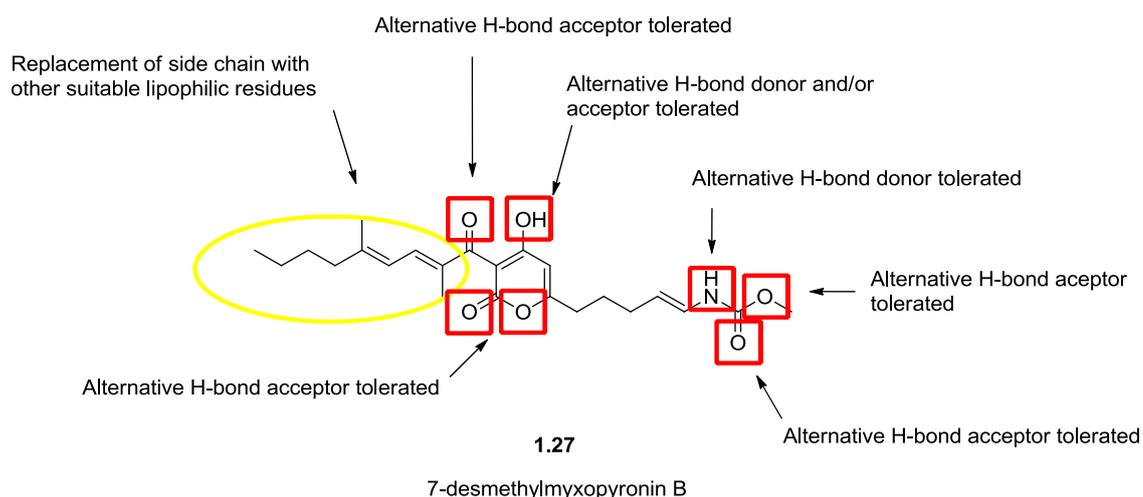
**Figure 6.1** Virtual screening protocol

A conformer database was generated using OMEGA software and an initial ligand-based virtual screening run with ROCS was applied with the aims of:

- A) reducing the computational time required for docking a relatively large database, and:

B) prioritizing molecules with similar shape and chemical features to myxopyronin, in order to maximize the chance of finding active molecules by mimicking the interactions of myxopyronin with RNAP.

The ligand-based virtual screening strategy was elaborated starting from the previously reported analysis (please see Chapter 5, Section 5.1.1) of the X-ray co-crystal structure of 7-desmethylmyxopyronin B **1.27** in complex with RNAP (PDB id: 3EQL<sup>3</sup>). In summary, pharmacophoric features in terms of donor/acceptor and lipophilic atoms were identified and a scaffold hopping strategy implying the preservation of key specific hydrogen bonds and hydrophobic interactions of 7-desmethylmyxopyronin B **1.27** was formulated, with the addition of two more features at the level of the carbonyl group (both at the level of the dienone and  $\alpha$ -pyrone ring) (Figure 6.2). These two additional features were included in order to bias any resulting virtual hits towards molecules that contain similar functionality to that within 7-desmethylmyxopyronin B **1.27**, and so to decrease dramatically the number of molecules to be docked in the subsequent step.



**Figure 6.2** Scaffold hopping strategy

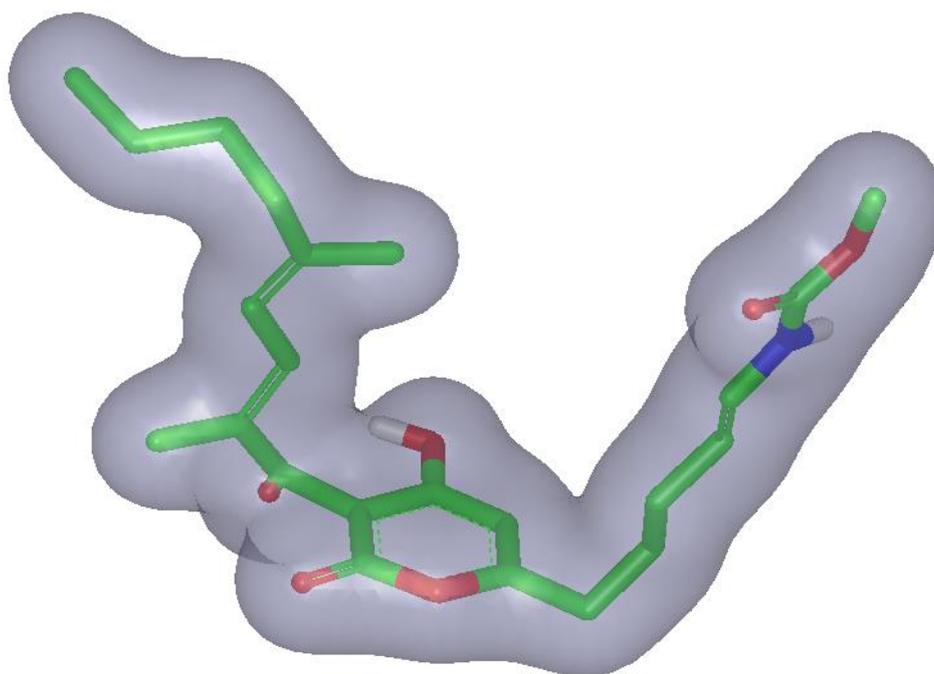
Docking was performed using GOLD docking software inside the myxopyronin binding region defined from the crystal structures of the *T. thermophilus* RNAP holoenzyme in complex with the antibiotic myxopyronin A (PDB id: 3DXJ<sup>2</sup>). The highest possible accuracy parameters with regard to the searching algorithm were used and consensus scoring was applied to rank the molecules. The Chemscore scoring function was used initially for ranking the compounds as it was found to be faster than Goldscore in terms of energetic evaluation during pose generation and it was even more appropriate because of its energetic terms for dealing with a predominantly hydrophobic binding site.<sup>6</sup> The Goldscore scoring function was then applied to re-rank the top scoring compounds in order to prioritize the best hydrogen bond network between the docked poses, given its superiority in identifying polar contacts when compared to Chemscore.<sup>6</sup> A post-docking analysis followed the consensus scoring strategy and consisted in cluster analysis and visual inspection.

### 6.2.1 General ligand-based virtual screening protocol

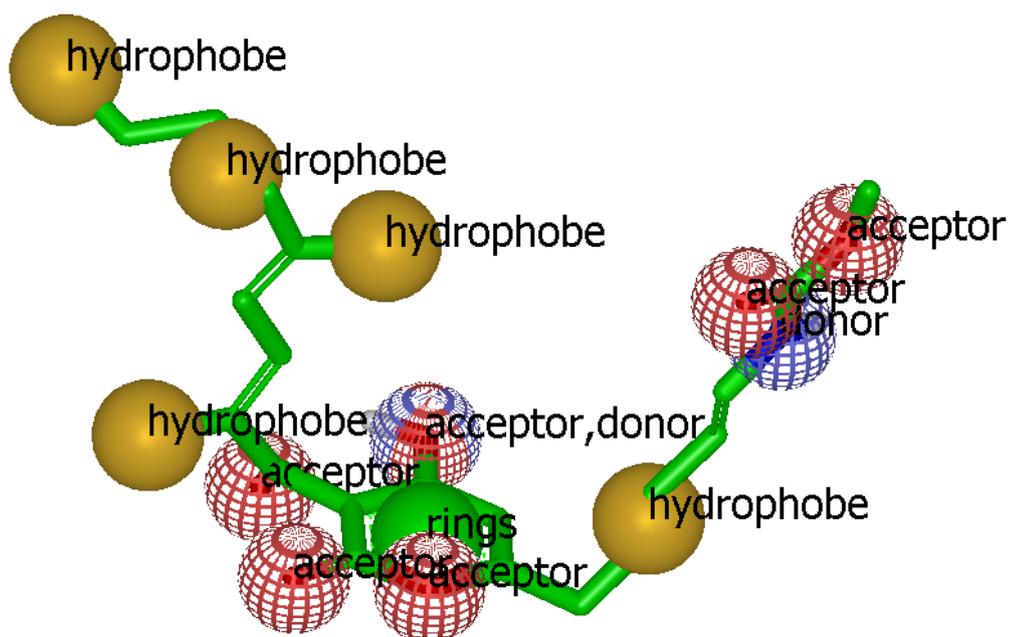
The Specs database (206,615 structures) was downloaded from ZINC database.<sup>4</sup>

A library of conformers was then generated from the previously downloaded databases using the OMEGA software version 2.3.2 limiting the maximum number of conformers per molecule to 50 using the default settings.

The biologically active conformation of 7-desmethylmyxopyronin B **1.27** was extracted from its crystallographic complex with RNAP (PDB id: 3EQL<sup>3</sup>) and atom types and bond orders were amended as necessary using Maestro software. The molecular structure was imported into vROCS version 3.1.2 where the pharmacophoric features were manually specified over the key atoms identified from the previous analysis (Chapter 5, Section 5.1.1) and the overall shape and volume of the molecule were automatically perceived by the software (Figures 6.3 and 6.4).



**Figure 6.3** Shape query elaborated in vROCS represented as a grey coloured volume. Structure of 7-desmethylmyxopyronin B **1.27** represented as sticks.



**Figure 6.4** Pharmacophore query elaborated in vROCS represented as coloured spheres centred on 7-desmethylmyxopyronin B **1.27** structure represented as sticks.

The previously prepared library of conformers was screened on the generated query using an Explicit Mills Dean<sup>5</sup> colour force field and leaving all the other settings to their default values. The top 10% of molecules ranked by ROCS using the TanimotoCombo scoring function were selected and after the elimination of duplicate conformers, a short list of 5,785 molecules was designated for structure-based virtual screening.

## 6.2.2 General docking protocol

Docking settings for the structure-based virtual screening protocol were chosen according to the docking validation studies reported in Chapter 3.

The crystal structure of the *T. thermophilus* RNAP holoenzyme in complex with the antibiotic myxopyronin A (PDB id: 3DXJ<sup>2</sup>), was downloaded from the Protein Databank ([www.rcsb.org](http://www.rcsb.org)).

All water molecules, cofactors and ions were manually removed using Maestro and the binding site within the RNAP was defined as the protein comprised in a sphere with a 15 Å radius surrounding the co-crystallised ligand.

The docking runs were performed using GOLD<sup>6</sup> v4.0.1 (CCDC, Cambridge, UK) docking software by using the best accuracy settings for the genetic searching algorithm which corresponded to the 200% of the default parameter values of the genetic algorithm in order to get a more exhaustive search inside the binding cavity. Compounds were ranked with the Chemscore scoring function generating 100 poses for each ligand. The final docked poses were re-ranked using the Goldscore scoring function after energetic minimization performed using the simplex algorithm implemented in the GOLD docking software.

A shortlist of 578 compounds corresponding to 10% of the top ranked molecules was selected for post-docking analysis which was performed for each ligand among its different docking solutions. Cluster analysis using an average linkage rule was performed using an internal module within the software. Only the best ranked poses of the most populated cluster were considered as the most likely and representative ones. In addition, the docked poses within the binding pocket were analysed in terms of overall score, shape complementarity to the cavity, possibility to establish specific H-bonds to the

protein in addition to the less specific hydrophobic contacts, minimal exposure of hydrophobic groups to solvent and absence of excessive torsional strain within the docked conformer.

### 6.2.3 vHTS results

After the post-docking analysis which included the selection of the potential virtual hits on the basis of their chemical diversity, 29 molecules were purchased for biological evaluation (please see Section 6.2.4 for the biological results). Interestingly, 10 out of the 27 virtual hits, **6.1—6.10**, reported in the previous study,<sup>1</sup> were present in the list of the purchased molecules (Table 6.1) and some compounds presented one or more violations of the Lipinski rule of five. The decision to include such non-Lipinski compliant molecules as part of the purchase selection reflected the observation that many existing clinically used antibiotics do not conform to this rule.<sup>7,8</sup> Moreover, subsequent exploration of SAR followed by hit optimization can be used to overcome the initial limitation of low oral bioavailability which may be associated with non-Lipinski compliance.

**Table 6.1** Virtual hits selected for the *in vitro* percentage inhibition

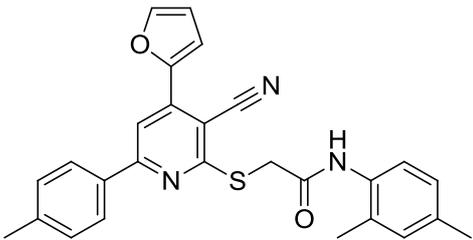
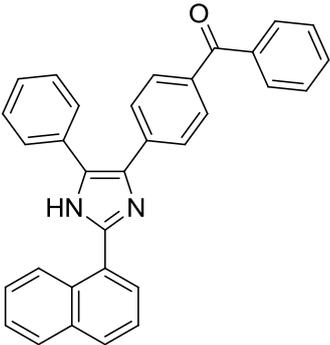
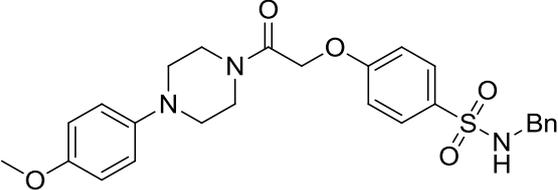
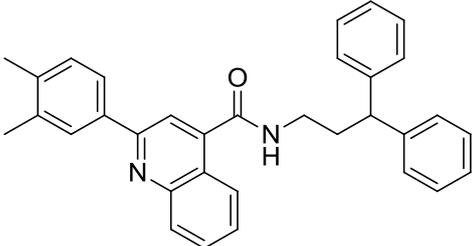
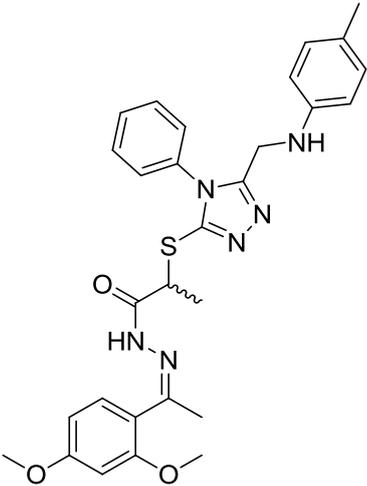
Comp. number	Structure	M.W.	ALogP <sub>(A)</sub>	Chem score	Rule of 5 violations <sub>(B)</sub>
6.1		453	5.50	38.7	1

Table 6.1 Continued

6.2		450	7.19	41.8	1
6.3		543	5.73	42.2	2
6.4		495	2.52	42.3	1
6.5		470	7.45	40.4	1
6.6		544	5.53	41.1	3



**Table 6.1 Continued**

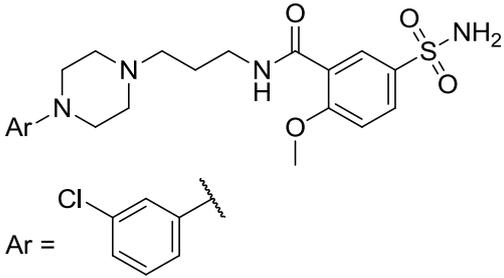
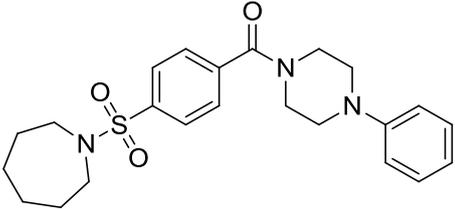
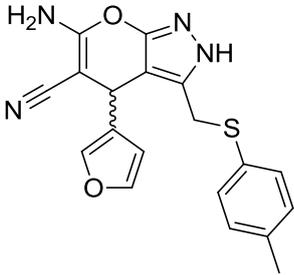
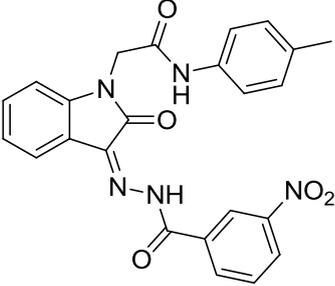
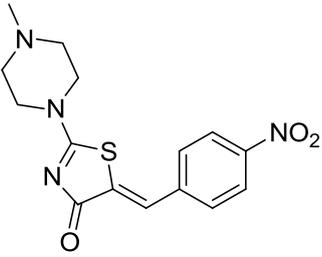
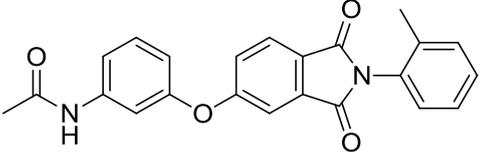
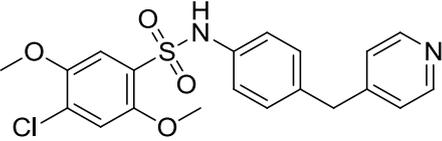
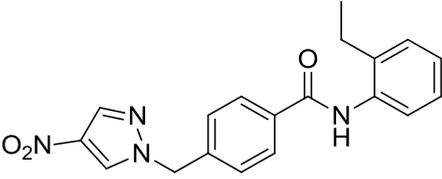
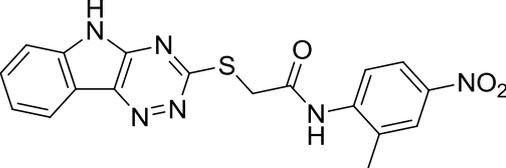
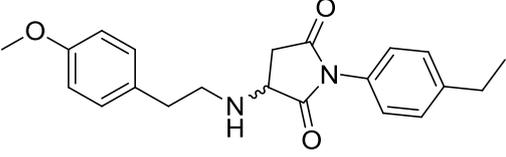
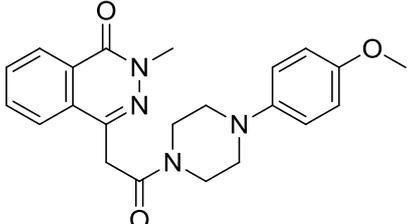
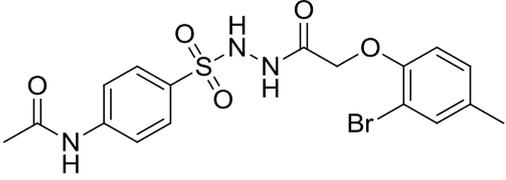
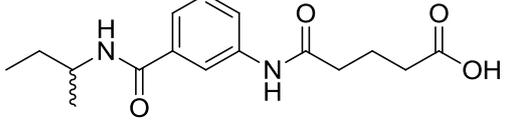
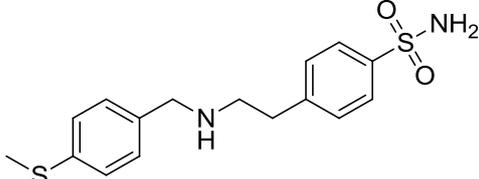
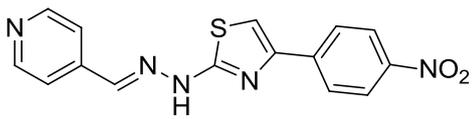
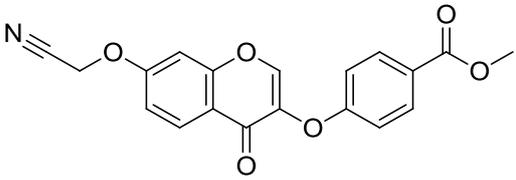
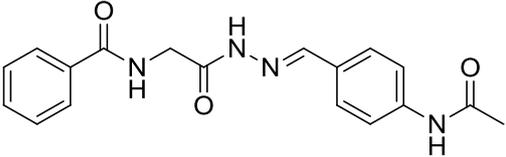
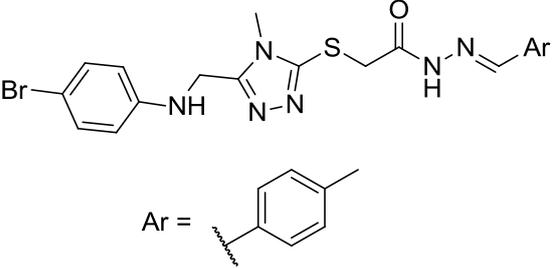
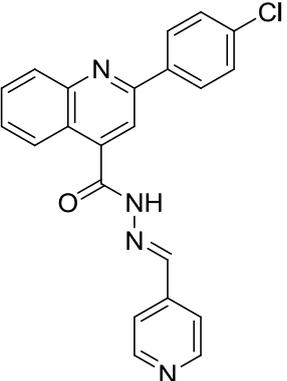
<p><b>6.11</b></p>		<p>466</p>	<p>2.80</p>	<p>33.6</p>	<p>1</p>
<p><b>6.12</b></p>		<p>427</p>	<p>3.07</p>	<p>35.2</p>	<p>0</p>
<p><b>6.13</b></p>		<p>364</p>	<p>3.89</p>	<p>34.5</p>	<p>0</p>
<p><b>6.14</b></p>		<p>458</p>	<p>4.07</p>	<p>35.0</p>	<p>0</p>
<p><b>6.15</b></p>		<p>333</p>	<p>2.04</p>	<p>35.4</p>	<p>0</p>
<p><b>6.16</b></p>		<p>386</p>	<p>3.52</p>	<p>29.0</p>	<p>0</p>

Table 6.1 Continued

6.17		418	3.71	30.9	0
6.18		351	3.95	34.9	0
6.19		395	3.25	38.1	0
6.20		352	2.86	31.0	0
6.21		392	1.76	37.3	0
6.22		455	2.62	37.5	1
6.23		306	1.47	33.7	1
6.24		336	2.62	28.3	0

**Table 6.1 Continued**

6.25		326	3.81	28.0	0
6.26		351	2.37	37.3	0
6.27		338	2.26	37.1	0
6.28		472	4.03	29.9	0
6.29		386	5.09	32.1	1

(A) ALogP calculated with ALOGPS 2.1 software<sup>9</sup>

(B) Rule of five defined as reported in Lipinski's paper,<sup>10</sup> parameters calculated with Molsoft Drug Likeness Prediction:<sup>11</sup>

150 ≤ M.W. ≤ 500

ALogP ≤ 5

Hydrogen bond donors ≤ 5

Hydrogen bond acceptors ≤ 10

Rotatable bonds ≤ 7

Polar surface area ≤ 150

## 6.2.4 Biological results

Compounds **6.1—6.29** were evaluated in the SYBR Green assay<sup>12</sup> (please see Appendix I for details) and the results are summarized in Table 6.2.

Compounds **6.3** and **6.10** were found to possess moderate inhibitory activity at the concentration of 100  $\mu$ M on *E.coli* RNAP core enzyme (compound **6.10** was not completely soluble under the assay conditions). To probe the selectivity of enzyme inhibition, compounds **6.3** and **6.10** were subjected to specificity assay analysis using malate dehydrogenase and chymotrypsin respectively, as unrelated enzymes, to identify promiscuous activity<sup>13</sup> (Table 6.3). Pleasingly, both compounds did not significantly inhibit these enzymes at the concentration of 100  $\mu$ M (please see Appendix I for details).

**Table 6.2** *In vitro* RNAP percentage inhibition at 100  $\mu$ M of the virtual hits.

Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M	Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M
<b>6.1</b>	21.7	<b>6.16</b>	0
<b>6.2</b>	0	<b>6.17</b>	0
<b>6.3</b>	56.4	<b>6.18</b>	0
<b>6.4</b>	5.9	<b>6.19</b>	0
<b>6.5</b>	0	<b>6.20</b>	0
<b>6.6</b>	0	<b>6.21</b>	0
<b>6.7</b>	16.4	<b>6.22</b>	0
<b>6.8</b>	0	<b>6.23</b>	0
<b>6.9</b>	2.8	<b>6.24</b>	0
<b>6.10</b>	54.3 <sup>†</sup>	<b>6.25</b>	0
<b>6.11</b>	0	<b>6.26</b>	0
<b>6.12</b>	0	<b>6.27</b>	0
<b>6.13</b>	0	<b>6.28</b>	0
<b>6.14</b>	0	<b>6.29</b>	17.1
<b>6.15</b>	0		

(†) Compounds showing solubility issues under the assay conditions.

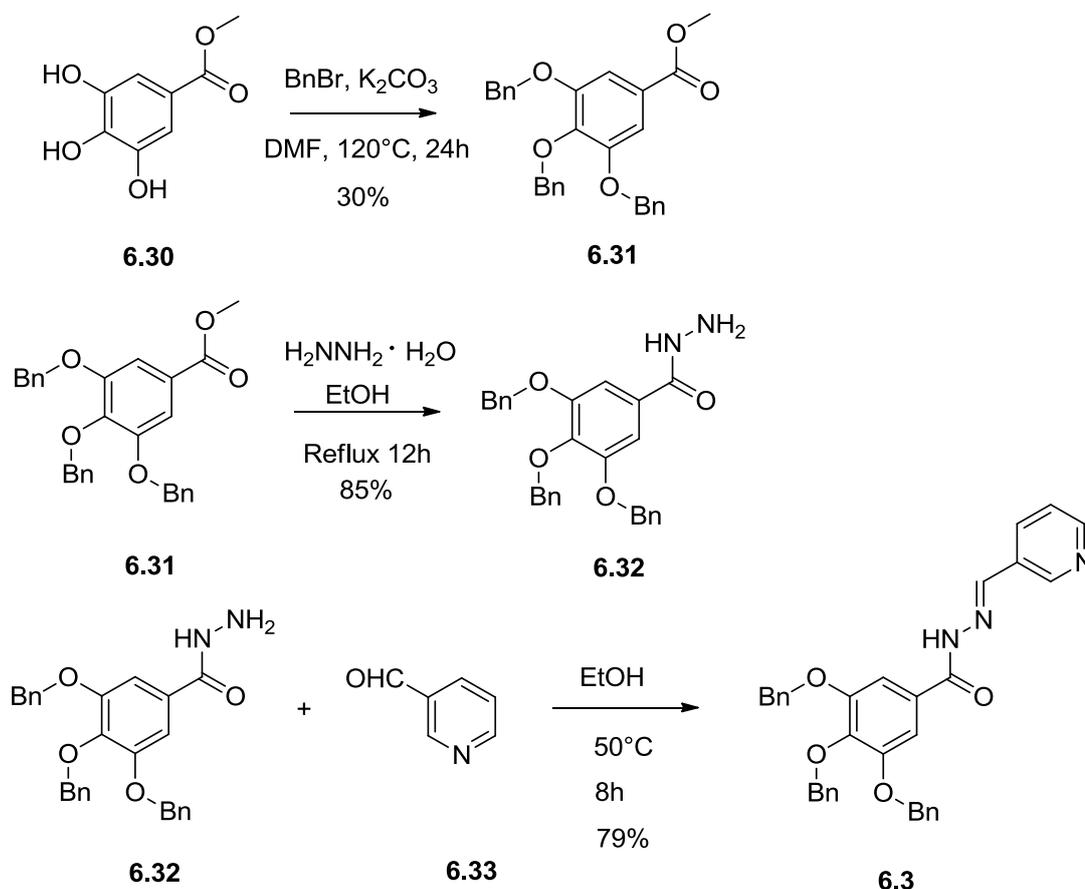
**Table 6.3** Specificity assay on the selected hits on Malate Dehydrogenase and Chymotrypsin.

Compound number	<i>In vitro</i> Malate Dehydrogenase % inhibition at 100 $\mu$ M	<i>In vitro</i> Chymotrypsin % inhibition at 100 $\mu$ M
<b>6.3</b>	2.9	4.3
<b>6.10</b>	0.3	21.4

From these two hits, acylhydrazone **6.3** was selected for SAR studies on the basis of its inhibitory activity, synthetic amenability, chemical stability, solubility properties and suitability for analogue design.

### 6.2.5 Synthesis of the acylhydrazone hit

In order to confirm the biological activity of the purchased compound **6.3**, this acylhydrazone was synthesized as shown below (Scheme 6.1):

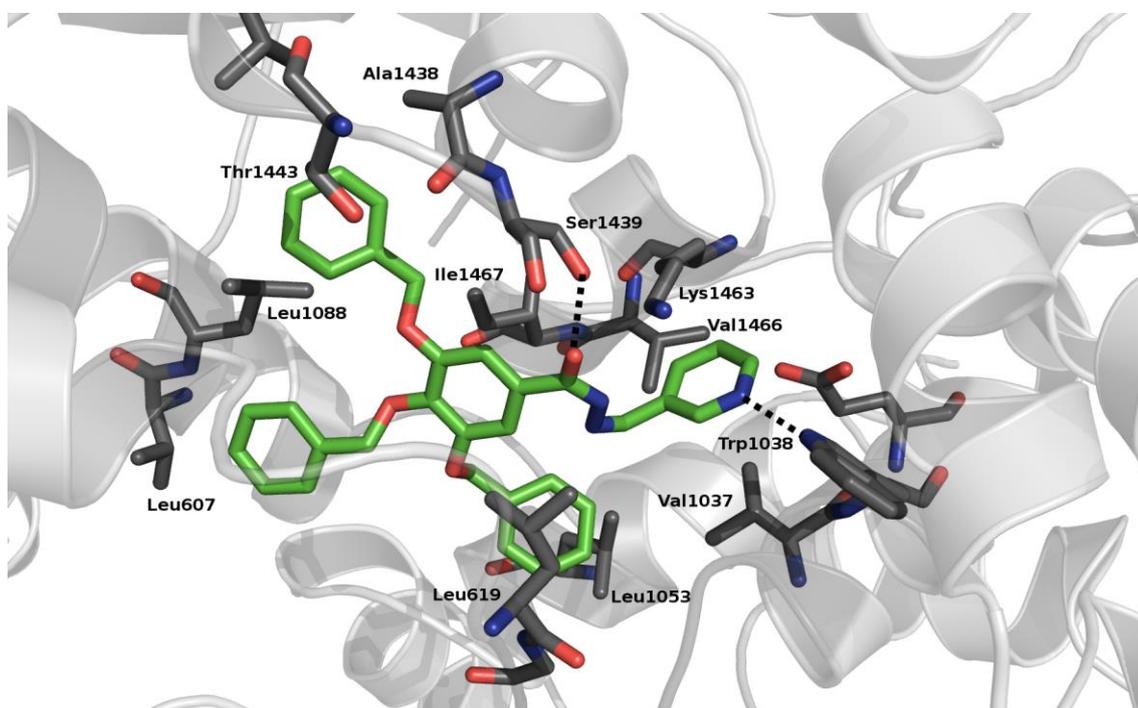


**Scheme 6.1** Synthesis of **6.3**.

A DMF solution of ester **6.30** was treated with an excess of benzylbromide in presence of an excess of potassium carbonate under reflux to give tri-ether **6.31**<sup>14</sup> in low yield. Intermediate hydrazide **6.32**<sup>15</sup> was obtained in excellent yield following the addition of excess hydrazine hydrate to ester **6.31**<sup>14</sup> and heating to reflux in ethanol. Finally, acylhydrazone **6.3** was obtained in good yield by reacting an ethanolic solution of hydrazide **6.32**<sup>15</sup> with aldehyde **6.33** under reflux.

### 6.2.6 Putative binding mode of the acylhydrazone hit

From the above mentioned docking studies (Section 6.2.2), the binding mode of acylhydrazone **6.3** was predicted as shown in Figure 6.5.

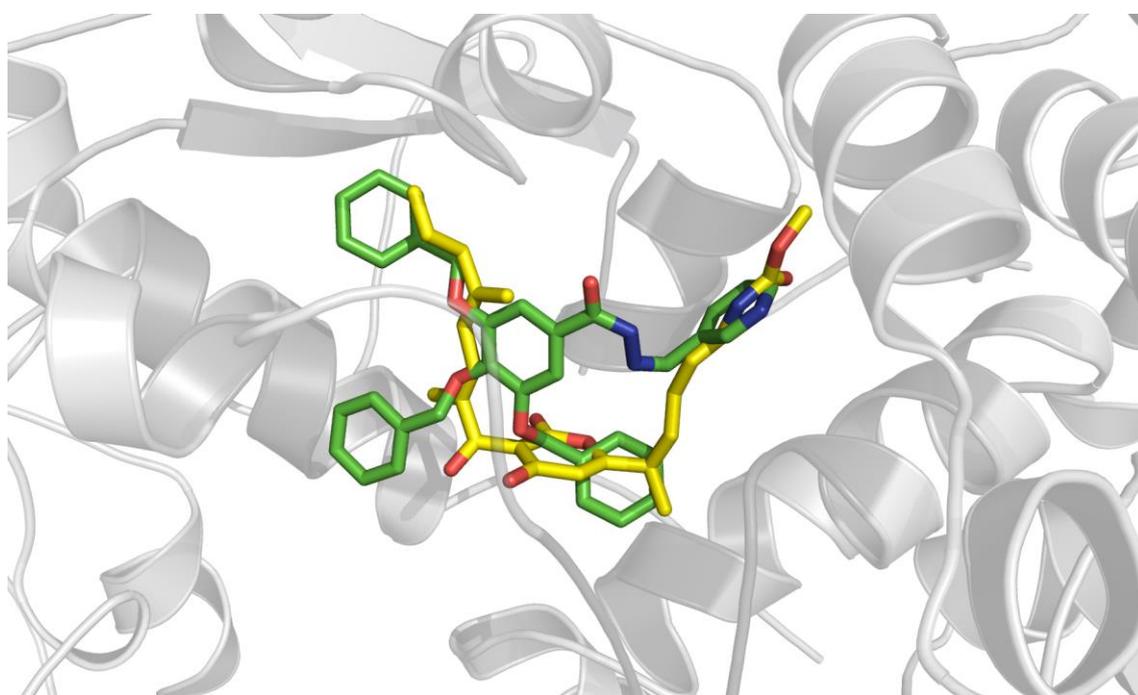


**Figure 6.5** Predicted binding mode of **6.3**.

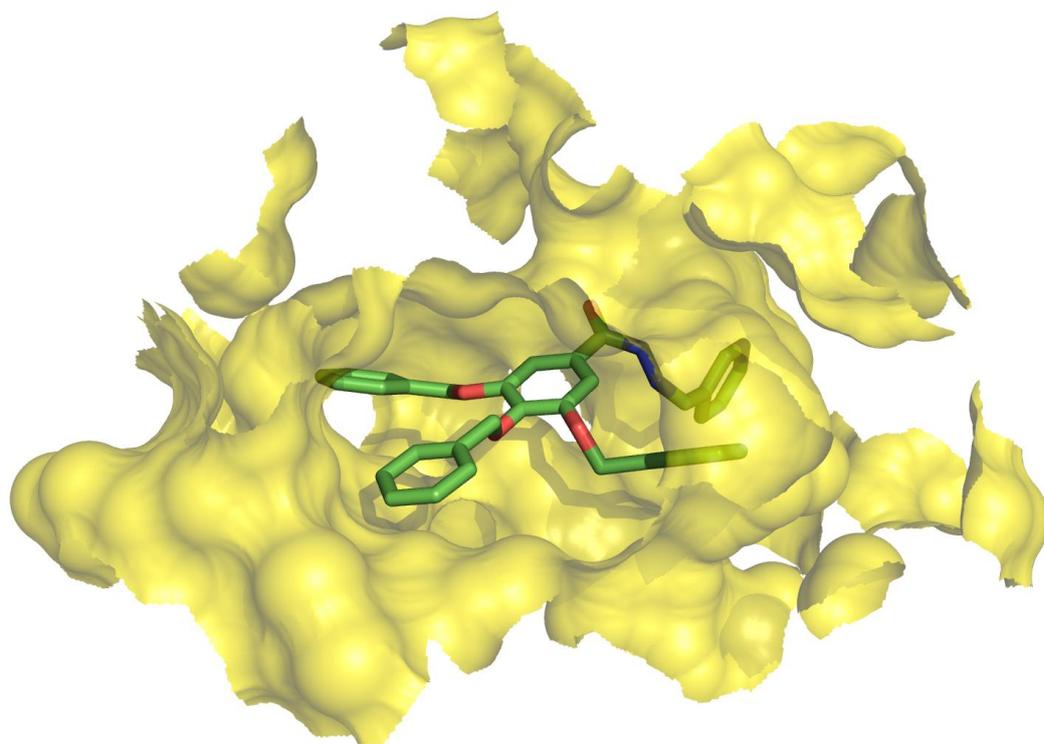
The following interactions between the ligand **6.3** and the myxopyronin binding site were identified: the carbonyl oxygen belonging to the acylhydrazone

moiety is predicted to H-bond to the side chain of Ser1439 while the nitrogen of the pyridyl unit interacts *via* H-bonding with the side chain NH of Trp1038. The protonated nitrogen belonging to the side chain of Lys1463 is performing a cation-pi interaction with the pyridyl ring which is also establishing hydrophobic interactions with the side chains of Val1466 and Val1037 respectively. The aromatic ring of one of the benzyloxy moieties is predicted to establish a OH-pi interaction with the hydroxyl moiety belonging to the side chain of Thr1443. All three benzyloxy groups are predicted to be involved in extensive hydrophobic interactions with the lipophilic side chains of Ala1438, Leu1088, Leu607, Leu619, and Leu1053 respectively.

As reported in Figure 6.6, acylhydrazone **6.3** is partially overlapping myxopyronin A **1.25** and the three benzyloxy groups are occupying three distinct lipophilic pockets inside the myxopyronin binding site (Figure 6.7).



**Figure 6.6** Overlay of predicted binding pose for acylhydrazone **6.3** represented in green sticks and myxopyronin A **1.25** in yellow sticks.

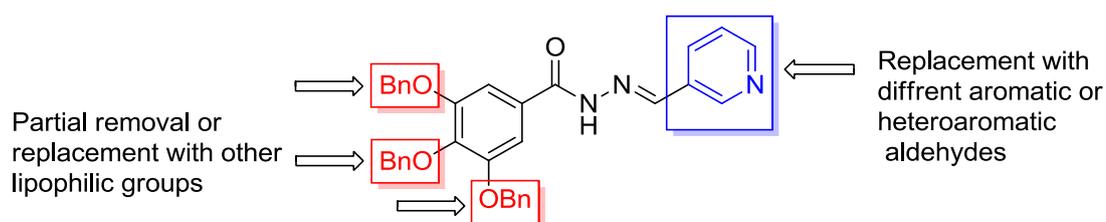


**Figure 6.7** Myxopyronin binding site represented as a Connolly surface in yellow and acylhydrazone **6.3** in green sticks.

### 6.2.7 General analogue design strategy

Based on the above cited putative binding mode a SAR study was performed in order to probe the predicted potential interactions and to increase potency and drug likeness properties of acylhydrazone **6.3**.

The general analogue design strategy is summarized in Figure 6.8:



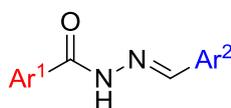
**Figure 6.8** General analogue design strategy.

In particular, it was reasoned that the partial replacement or complete removal of the three benzyloxy groups will help with improvement of the physicochemical properties and will also allow exploration of the importance of these extended hydrophobic interactions in term of potency. The substitution of the 3-pyridyl group with other aromatic or heteroaromatic groups will help the optimization and understanding of the hydrogen bonding network.

In order to optimize the SAR exploration and to maximize the identification of biologically active analogues, a similarity searching protocol, based on the structure of the virtual hit **6.3**, has been conducted using a focused library of acylhydrazones designed as shown in the next Section.

### 6.2.8 Similarity searching on a focused library of acylhydrazones

Two databases were considered for the preparation of the focused library, the ZINC<sup>4</sup> and Reaxys<sup>16</sup> databases. A chemical structure query was performed on both databases with the aim of selecting a chemical space composed of a central acylhydrazone scaffold substituted with aromatic or heteroaromatic moieties as shown in Scheme 6.2.



Ar<sup>1</sup> = any aromatic or heteroaromatic residue

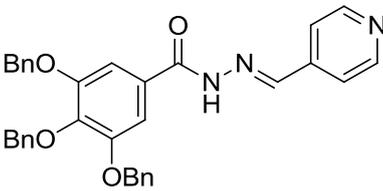
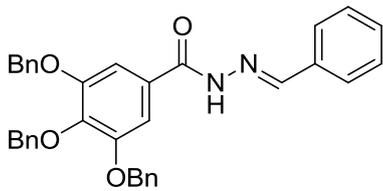
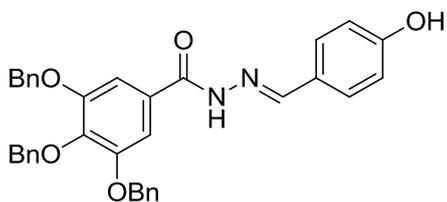
Ar<sup>2</sup> = any aromatic or heteroaromatic residue

**Scheme 6.2** General analogue design strategy.

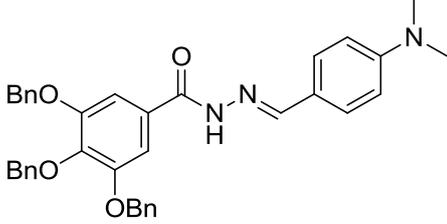
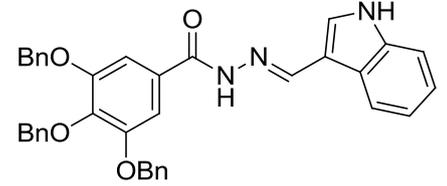
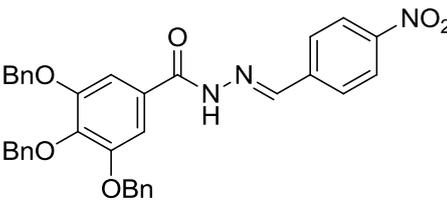
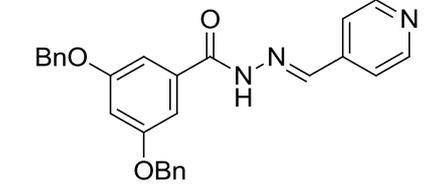
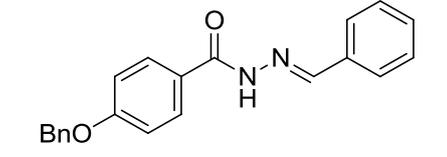
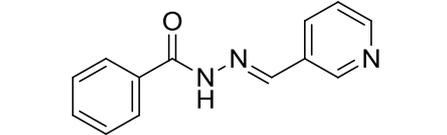
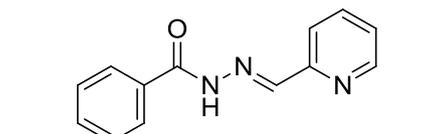
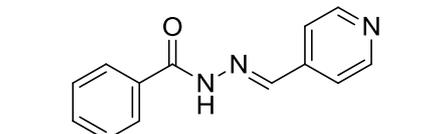
With regard to the Reaxys<sup>16</sup> database, additional filters were specified to this query and consisted of selecting only compounds for which a reported chemical preparation procedure was available. The above mentioned query retrieved 10,375 compounds derived from merging 4,932 compounds from the

ZINC database and 8,901 compounds from the Reaxys database after eliminating duplicates. The choice of these databases and the filtering criteria for creating the focused library satisfied the objective of short listing commercially available and synthetically accessible molecules. A similarity searching run using atom environment descriptors (MOLPRINT2D<sup>17</sup>) was then performed on this library using the structure of the acylhydrazone hit **6.3** as a reference structure and retrieving only molecules with a Tanimoto<sup>18</sup> similarity index  $\geq 0.3$  to the reference molecule. The resulting 321 compounds were docked following the same protocol reported in Section 6.2.2, using a more accurate post-docking analysis which selected only compounds showing at least the same number of H-bond interactions to that predicted for the parent compound **6.3**, and an overall Chemscore numerical value in the range of  $\pm 15$  related to the scoring value from the initial hit structure. After the post-docking analysis 17 molecules were selected and are reported in Table 6.4.

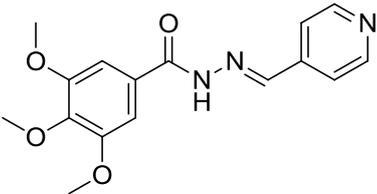
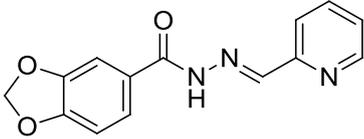
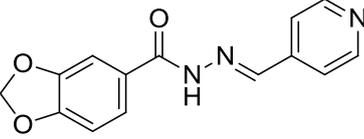
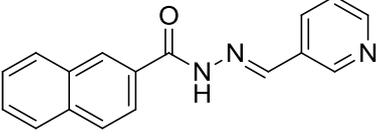
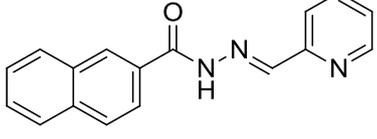
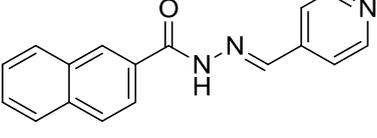
**Table 6.4** Selected compounds from the similarity searching.

Comp. number	Structure	Tanimoto similarity	Chem score	Number of H-bonds
<b>6.34</b> <sup>19</sup>		0.73	42.1	3
<b>6.35</b>		0.71	41.5	2
<b>6.36</b>		0.62	40.3	3

**Table 6.4 Continued**

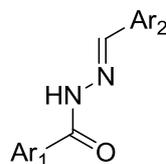
6.37		0.60	42.8	3
6.38		0.58	43.1	2
6.39		0.58	44.9	3
6.40 <sup>19</sup>		0.58	37.7	3
6.41 <sup>20</sup>		0.52	32.3	2
6.42 <sup>21</sup>		0.65	28.3	2
6.43 <sup>22</sup>		0.38	29.5	2
6.44 <sup>21</sup>		0.46	28.2	2

**Table 6.4** Continued

<b>6.45</b> <sup>23</sup>		0.46	28.1	2
<b>6.46</b> <sup>21</sup>		0.41	28.5	2
<b>6.47</b> <sup>21</sup>		0.44	28.2	2
<b>6.48</b> <sup>24</sup>		0.56	30.4	2
<b>6.49</b> <sup>24</sup>		0.33	30.0	2
<b>6.50</b> <sup>24</sup>		0.40	29.3	2

Interestingly, the selected virtual hits are predicted to share a similar binding mode to molecule **6.3** which suggests that the position of the pyridyl ring nitrogen is not essential and both the benzyloxy and pyridyl moieties can be substituted with other aromatic or heteroaromatic groups. In order to validate the similarity searching results, the previously formulated hypothesis concerning the position of the pyridyl ring nitrogen and the putative binding mode of **6.3**, the compounds in Table 6.4, along with other analogues (Table 6.5) bearing the generic formula reported in Scheme 6.3, were synthesized (please see Sections

6.4 and 6.5 for chemistry) also adopting the previously mentioned analogue design strategy (Section 6.2.7).



**6.51-6.60**

**Scheme 6.3** Generic formula of the additional synthesized analogues.

**Table 6.5** Additional synthesized analogues.

Compound number	Ar <sub>1</sub>	Ar <sub>2</sub>
<b>6.51</b>	3,4,5-tris(benzyloxy)phenyl	2-pyridyl
<b>6.52</b>	3,5-bis(benzyloxy)phenyl	2-pyridyl
<b>6.53</b>	3,5-bis(benzyloxy)phenyl	3-pyridyl
<b>6.54</b>	3,5-bis(benzyloxy)phenyl	phenyl
<b>6.55</b>	4-(benzyloxy)phenyl	2-pyridyl
<b>6.56</b>	4-(benzyloxy)phenyl	3-pyridyl
<b>6.57</b>	4-(benzyloxy)phenyl	4-pyridyl
<b>6.58</b>	3,4,5-trimethoxyphenyl	2-pyridyl
<b>6.59</b>	3,4,5-trimethoxyphenyl	3-pyridyl
<b>6.60</b>	3,4-methylenedioxyphenyl	3-pyridyl

### 6.2.9 Biological results

Compounds **6.34—6.60** were evaluated in the SYBR Green assay<sup>12</sup> (please see Appendix I for details) and results are summarized in Table 6.6.

**Table 6.6** *In vitro* RNAP percentage inhibition of the synthesised analogues.

Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M	Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M
<b>6.34</b> <sup>19</sup>	46.3	<b>6.48</b> <sup>24</sup>	38.4
<b>6.35</b>	14.8 <sup>†</sup>	<b>6.49</b> <sup>24</sup>	65.5
<b>6.36</b>	65.3 <sup>†</sup>	<b>6.50</b> <sup>24</sup>	50.2
<b>6.37</b>	4.8 <sup>†</sup>	<b>6.51</b>	53.8
<b>6.38</b>	0 <sup>†</sup>	<b>6.52</b>	15.3 <sup>‡</sup>
<b>6.39</b>	36.8 <sup>†</sup>	<b>6.53</b>	31.2 <sup>‡</sup>
<b>6.40</b> <sup>19</sup>	24 <sup>‡</sup>	<b>6.54</b>	14 <sup>†‡</sup>
<b>6.41</b> <sup>20</sup>	7.2	<b>6.55</b>	2.4
<b>6.42</b> <sup>21</sup>	0	<b>6.56</b>	24.3
<b>6.43</b> <sup>22</sup>	0	<b>6.57</b>	11.9
<b>6.44</b> <sup>21</sup>	0	<b>6.58</b>	2.7
<b>6.45</b> <sup>23</sup>	0	<b>6.59</b>	0
<b>6.46</b> <sup>21</sup>	3.1	<b>6.60</b>	7.6
<b>6.47</b> <sup>21</sup>	4.7		

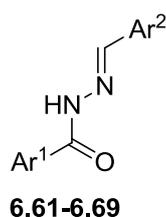
(†) Compounds showing solubility issues in the assay conditions

(‡) Compounds showing auto-fluorescence in the assay conditions

Unfortunately compounds **6.35**, **6.36**, **6.37**, **6.38** and **6.39** showed limited solubility in the assay conditions (please see Appendix I for details) moreover compounds **6.40**,<sup>19</sup> **6.52**, **6.53** and **6.54** showed auto-fluorescence. Solubility issues and auto-fluorescence exhibited by some compounds have a detrimental effect on the percentage of inhibition at the specified concentration due to the nature of the assay.<sup>12</sup>

### 6.2.10 Design and synthesis of analogues with partial or total removal of benzyloxy groups

Solubility problems reported in the previous series of compounds (Tables 6.4 and 6.5) were addressed by synthesizing structural analogues derived from partial or total removal benzyloxy groups (Table 6.7) bearing the generic formula reported in Scheme 6.4 (please see Sections 6.4 and 6.5 for chemistry).



**Scheme 6.4** Generic formula of the synthesized analogues with partial or total removal of benzyloxy groups.

**Table 6.7** Synthesized analogues with partial or total removal of benzyloxy groups.

Compound number	Ar <sup>1</sup>	Ar <sup>2</sup>
<b>6.61</b>	3,4,5-tris(benzyloxy)phenyl	5-nitrofuran-2-yl
<b>6.62</b>	3,5-bis(benzyloxy)phenyl	<i>p</i> -hydroxyphenyl
<b>6.63</b>	3,5-bis(benzyloxy)phenyl	1 <i>H</i> -indol-3-yl
<b>6.64</b> <sup>20</sup>	4-(benzyloxy)phenyl	<i>p</i> -nitrophenyl
<b>6.65</b>	4-(benzyloxy)phenyl	<i>p</i> -dimethylaminophenyl
<b>6.66</b>	4-(benzyloxy)phenyl	1 <i>H</i> -indol-3-yl
<b>6.67</b> <sup>25</sup>	2-naphthyl	<i>p</i> -nitrophenyl
<b>6.68</b>	2-naphthyl	<i>p</i> -dimethylaminophenyl
<b>6.69</b>	2-naphthyl	1 <i>H</i> -indol-3-yl

### 6.2.11 Biological results

Compounds **6.61—6.69** were evaluated in the SYBR Green assay<sup>12</sup> (please see Appendix I for details) and the results are summarized in Table 6.8.

**Table 6.8** *In vitro* RNAP percentage inhibition of the synthesised compounds.

Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M
<b>6.61</b>	0 <sup>†</sup>
<b>6.62</b>	59.5 <sup>†‡</sup>
<b>6.63</b>	2.2 <sup>†‡</sup>
<b>6.64</b> <sup>20</sup>	9.8 <sup>†</sup>
<b>6.65</b>	9.4 <sup>†</sup>
<b>6.66</b>	13.3 <sup>†</sup>
<b>6.67</b> <sup>25</sup>	30.4
<b>6.68</b>	13.2 <sup>†</sup>
<b>6.69</b>	33.1

(†) Compounds showing solubility issues in the assay conditions

(‡) Compounds showing auto-fluorescence in the assay conditions

Solubility issues and auto-fluorescence were observed for some compounds, which, for these cases, limited the accuracy of measurement of inhibition in the assay.

Disappointingly, the prediction of LogP and LogSw values using ALOGPS 2.1 software<sup>9</sup> (Table 6.9) for some representative compounds evaluated biologically (Tables 6.6, 6.7 and 6.8), was not able to explain why some compounds with better predicted solubility had limited solubility in the assay.

**Table 6.9** Prediction of ALogP and LogSw values using ALOGPS 2.1 software<sup>9</sup> for some representative compounds.

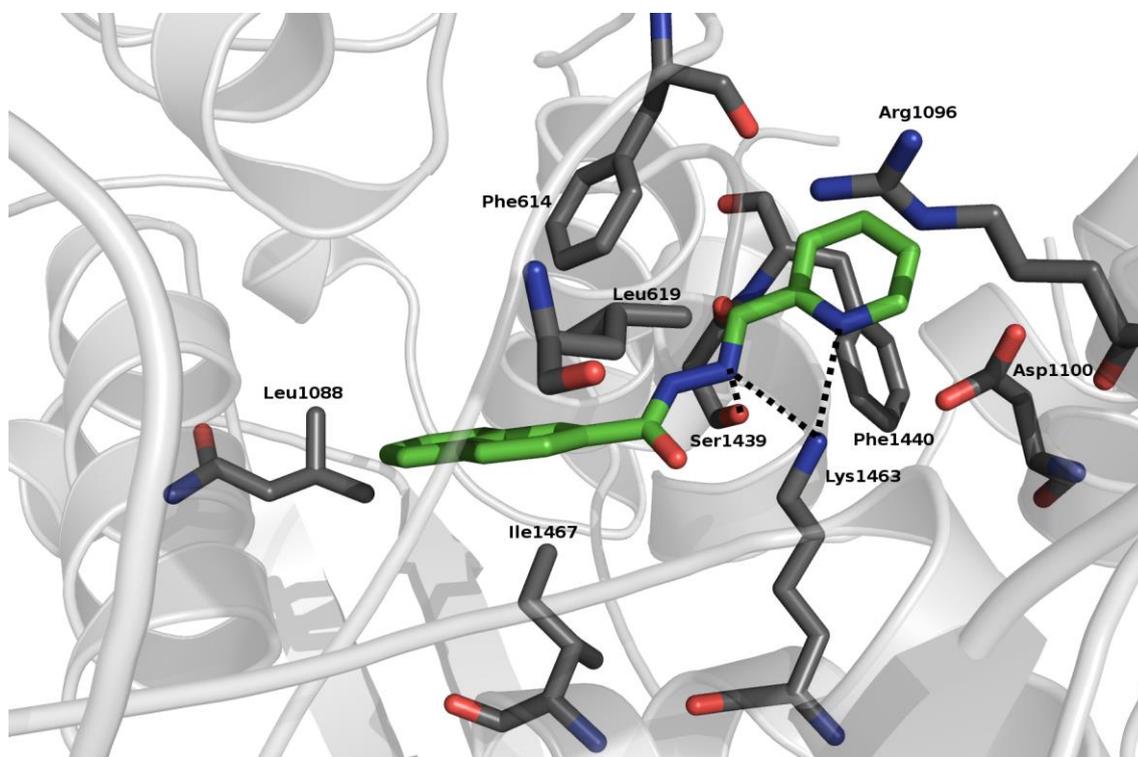
Compound number	ALogP	LogSw	Solubility Issues
<b>6.34</b> <sup>19</sup>	5.74	-6.63	No
<b>6.40</b> <sup>19</sup>	4.51	-6.14	Yes
<b>6.51</b>	5.95	-6.60	No
<b>6.52</b>	4.74	-6.12	Yes
<b>6.53</b>	4.50	-6.11	Yes
<b>6.62</b>	5.54	-6.12	Yes
<b>6.68</b>	4.00	-4.95	Yes
<b>6.69</b>	4.45	-5.68	No

### 6.2.12 SAR analysis

Considering the biological results reported in Sections 6.2.9 and 6.2.11, it was possible to conclude preliminary SAR considerations for this library (Schemes 6.3 and 6.4, Tables 6.4, 6.5 and 6.7). Replacement of the 3,4,5-(trisbenzyloxy)phenyl groups with a naphthalene maintains or slightly improves activity while complete or partial removal or replacement with other alkyl ethers is detrimental to the activity; the position of the nitrogen in the pyridyl ring is not essential for the activity. Interestingly, analogues bearing a 3,5-(bisbenzyloxy)phenyl moiety show auto-fluorescence and for this reason are unsuitable for further SAR studies. Whilst compounds **6.35**, **6.37**, **6.38** and **6.39** were predicted to be promising compounds on the basis of their putative binding mode and overall score, solubility issues in the assay encountered also in their less hydrophobic analogues prevented comparison of activity to the parent compound **6.3**, with the only exception of the soluble 2-naphthyl derivative **6.67**,<sup>25</sup> where the presence of the *p*-nitrophenyl moiety does not impair biological activity when compared to analogue **6.48**<sup>24</sup> bearing a 3-pyridyl moiety. Compounds **6.36** and **6.62**, despite their solubility problems, reveal that a

hydroxyl group in the para position bearing both electron donating properties and H-bond donor-acceptor features is tolerated.

Putative binding mode of the most active derivative, compound **6.49**<sup>24</sup> is reported in Figure 6.9.

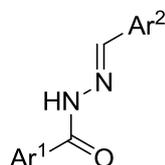


**Figure 6.9** Predicted binding mode of **6.49**.<sup>24</sup>

The following interactions between the ligand **6.49**<sup>24</sup> and the myxopyronin binding site were identified: the hydrazone nitrogen bound to the methinic carbon and the nitrogen belonging to the pyridyl ring are predicted to form an H-bond reinforced by the charge with the protonated nitrogen of the side chain of Lys1463. The guanidinium moiety belonging to the side chain of Arg1096 is predicted to form a cation- $\pi$  interaction with the pyridyl ring which is also performing a T-shaped displaced  $\pi$ -stacking interaction with the aromatic ring of the side chain of Phe1440. The naphthalenic ring is involved in hydrophobic interactions with the side chains of Phe614, Leu619, Leu1088 and Ile1467.

### 6.2.13 Design and synthesis of a new acylhydrazone library with improved solubility

A new acylhydrazone library (Table 6.10) bearing the generic formula reported in Scheme 6.5 was designed to further explore SAR in the light of the previous findings and to address some solubility issues encountered (please see Sections 6.4 and 6.5 for chemistry).



6.70-6.91a

**Scheme 6.5** Generic formula of the synthesized analogues with improved solubility.

**Table 6.10** Synthesized hydrazone-based analogues with predicted improved solubility

Compound number	Ar <sup>1</sup>	Ar <sup>2</sup>
6.70	3,4,5-tris(benzyloxy)phenyl	2-hydroxyphenyl
6.71	3,4,5-tris(benzyloxy)phenyl	3-hydroxyphenyl
6.72	3,4,5-tris(benzyloxy)phenyl	2,4-dihydroxyphenyl
6.73	3,4,5-tris(benzyloxy)phenyl	2-carboxyphenyl
6.74	3,4,5-tris(benzyloxy)phenyl	3-carboxyphenyl
6.75	3,4,5-tris(benzyloxy)phenyl	4-carboxyphenyl
6.76	2-naphthyl	2-carboxyphenyl
6.77	2-naphthyl	3-carboxyphenyl
6.78 <sup>25</sup>	2-naphthyl	4-carboxyphenyl

**Table 6.10** Continued

<b>6.79</b>	3,4,5-tris(benzyloxy)phenyl	2-cyanophenyl
<b>6.80</b>	3,4,5-tris(benzyloxy)phenyl	3-cyanophenyl
<b>6.81</b>	3,4,5-tris(benzyloxy)phenyl	4-cyanophenyl
<b>6.82</b>	3,4,5-tris(benzyloxy)phenyl	1 <i>H</i> -pyrrol-2-yl
<b>6.83</b>	2-naphthyl	1 <i>H</i> -pyrrol-2-yl
<b>6.84</b>	1-naphthyl	3-pyridyl
<b>6.85</b>	quinolin-6-yl	3-pyridyl
<b>6.86</b>	quinoxaline-6-yl	3-pyridyl
<b>6.87</b>	1 <i>H</i> -indol-3-yl	3-pyridyl
<b>6.88</b>	1 <i>H</i> -indol-5-yl	3-pyridyl
<b>6.89</b>	1 <i>H</i> -indol-6-yl	3-pyridyl
<b>6.90a</b>	3-morpholinophenyl	3-pyridyl
<b>6.91a</b>	4-(1 <i>H</i> -imidazol-1-yl)phenyl	3-pyridyl

The following structural variations were probed:

As compound **6.36** showed good inhibitory activity despite its solubility problems, other positions for the placement of an hydroxyl group on the phenyl ring were explored *via* compounds **6.70** and **6.71**, moreover a more soluble analogue, **6.72**, was prepared.

3,4,5-(Trisbenzyloxy)phenyl derivatives **6.73**, **6.74**, **6.75** and 2-naphthyl based analogues **6.76**, **6.77**, **6.78**<sup>25</sup> with the carboxylate group inserted in various positions of the phenyl ring were prepared with the aim of performing a bioisosteric replacement of the nitro group and to obtaining more soluble derivatives.

Compounds **6.79**, **6.80** and **6.81** were prepared to evaluate the inhibitory potency of derivatives bearing a substituent with electron withdrawing and H-bond acceptor properties.

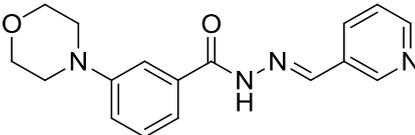
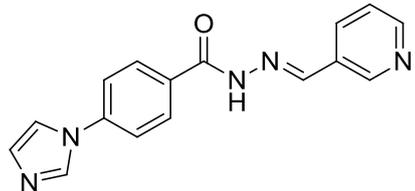
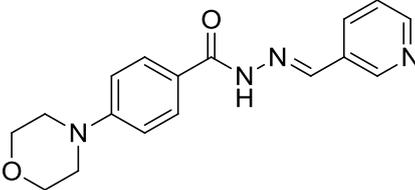
Substitution of the pyridyl group with a smaller ring was tested *via* synthesizing compounds **6.82** and **6.83**.

Compound **6.84** was prepared to assess whether a 2-naphthyl group is essential for the biological activity or if a 1-naphthyl substituent is tolerated.

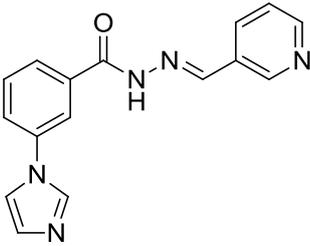
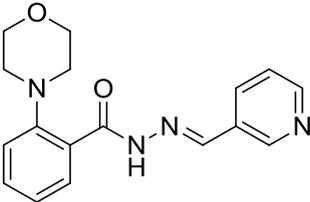
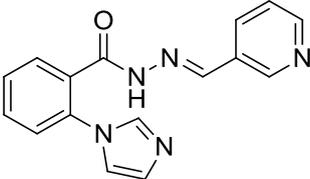
Heteroaromatic derivatives **6.85**, **6.86**, **6.87**, **6.88**, **6.89**, **6.90a**, **6.91a** were synthesized to improve solubility and to explore the effects of H-bond donors and acceptors on the putative binding region occupied by the 3,4,5-(trisbenzyloxy)phenyl group of parent compound **6.3**.

Compounds **6.90a** and **6.91a** were specifically shortlisted from a series of possible structural analogues (Table 6.11) following prediction from docking studies that they may undergo favourable binding to RNAP following the protocol reported in Section 6.2.2. The overall score value, putative binding mode and number of H-bond interactions were considered in prioritizing the synthesis of **6.90a** and **6.91a** as reported in Table 6.11.

**Table 6.11** Designed acylhydrazone analogues with structural variations on the phenyl ring

Comp. number	Structure	Chemscore	Number of H-bonds
<b>6.90a</b>		35.2	3
<b>6.91a</b>		34.0	2
<b>6.90b</b>		31.6	2

**Table 6.11** Continued

<b>6.91b</b>		28.8	1
<b>6.90c</b>		27.2	1
<b>6.91c</b>		27.6	1

### 6.2.14 Biological results

Compounds **6.70—6.91a** were evaluated in the SYBR Green assay<sup>12</sup> (please see Appendix I for details) and results are summarized in Table 6.12.

**Table 6.12** *In vitro* RNAP percentage inhibition of the designed acylhydrazone analogues.

Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M	Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M
<b>6.70</b>	0	<b>6.81</b>	17.1 <sup>†</sup>
<b>6.71</b>	15.1	<b>6.82</b>	9.7
<b>6.72</b>	5.9 <sup>†</sup>	<b>6.83</b>	0
<b>6.73</b>	53.8	<b>6.84</b>	0
<b>6.74</b>	18.7	<b>6.85</b>	7.9
<b>6.75</b>	72.7	<b>6.86</b>	0
<b>6.76</b>	11.8	<b>6.87</b>	13.3 <sup>†</sup>
<b>6.77</b>	16.4	<b>6.88</b>	0

**Table 6.12** Continued

<b>6.78<sup>25</sup></b>	30.1	<b>6.89</b>	7.5
<b>6.79</b>	0 <sup>†</sup>	<b>6.90a</b>	0
<b>6.80</b>	36.9	<b>6.91a</b>	0

(†) Compounds showing solubility issues in the assay conditions

Compounds **6.72**, **6.79**, **6.81** and **6.87** showed limited solubility in the buffer solution used for the assay, which would likely compromised their measured biological activity in the assay. As mentioned earlier, prediction of ALogP and LogSw values using ALOGPS 2.1 software<sup>9</sup> for some representative compounds (Table 6.13) was again not able to account for why some compounds with better predicted solubility showed solubility issues in the assay.

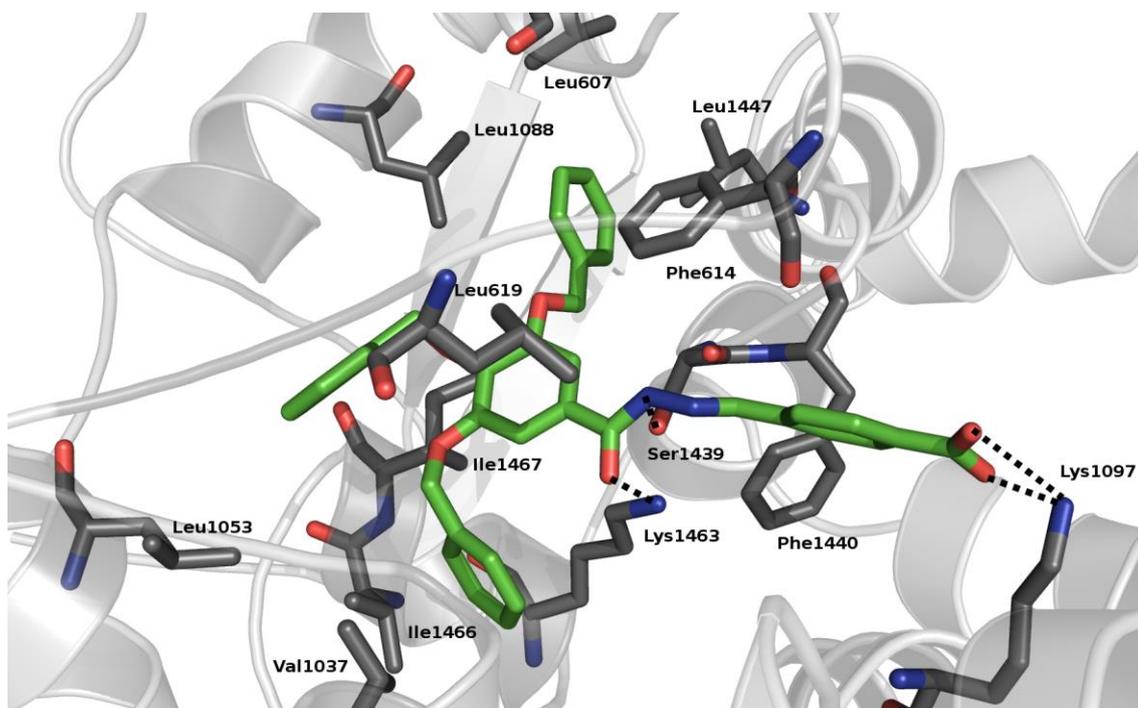
**Table 6.13** Prediction of ALogP and LogSw values using ALOGPS 2.1 software<sup>9</sup> for some representative compounds.

Compound number	ALogP	LogSw	Solubility Issues
<b>6.70</b>	6.45	-6.50	No
<b>6.71</b>	6.46	-6.56	No
<b>6.72</b>	6.36	-6.18	Yes
<b>6.79</b>	6.42	-6.08	Yes
<b>6.80</b>	6.43	-6.08	No
<b>6.81</b>	6.44	-6.08	Yes
<b>6.87</b>	2.37	-3.89	Yes
<b>6.88</b>	2.37	-3.98	No
<b>6.89</b>	2.37	-4.00	No

### 6.2.15 SAR analysis

In light of the biological results reported in Section 6.2.14, it was possible to make a number of deductions regarding SAR within this library (Scheme 6.5 and Table 6.10): with regard to 3,4,5-(trisbenzyloxy)phenyl-based derivatives, introduction of the hydroxyl group in the phenyl ring belonging to the benzylidene moiety in positions other than para, as in compounds **6.70** and **6.71** is detrimental for activity, while introduction of two hydroxyl groups somewhat surprisingly leads to solubility problems as in compound **6.72**. Introduction of a carboxyl group in either the ortho or para positions of the benzylidene unit increases the biological activity (compounds **6.73** and **6.75** respectively), whilst the placement of carboxyl at the meta position as in compound **6.74** reduces the biological activity relative to the parent compound **6.3**. Notably, introduction of a less strong H-bond acceptor and weaker electron withdrawing group relative to the carboxyl as in the cyano-substituted compounds **6.79**, **6.80** and **6.81**, is negative for the biological activity. With regard to the 2-naphthyl based derivatives, introduction of a carboxyl group in the same previously mentioned positions does not lead to more potent analogues in comparison with the parent compound **6.3** but, by analogy with compounds **6.37**, **6.38** and **6.39**, substitution at the para position of the benzylidene moiety is the most favorable in terms of biological activity. Interestingly, the 1-naphthyl based derivative **6.84** is inactive and this suggests that this scaffold is not amenable in contrast with 2-naphthyl based isomer. Substitution of the pyridyl group with a smaller ring system as in compounds **6.82** and **6.83** is detrimental for biological activity in both the 3,4,5-(trisbenzyloxy)phenyl and 2-naphthyl based derivatives. Substitution of the 3,4,5-(trisbenzyloxy)phenyl group of the parent compound **6.3** with other heteroaromatic derivatives conferring higher predicted solubility and the possibility to establish additional H-bond interactions led to inactive molecules as observed for compounds **6.85**, **6.86**, **6.87**, **6.88**, **6.89**, **6.90a**, **6.91a**.

The putative binding mode of the most active analogue, compound **6.75**, is shown in Figure 6.10.



**Figure 6.10** Putative binding mode of **6.75**.

The following interactions between the ligand **6.75** and the myxopyronin binding site were predicted: the protonated nitrogen of the Lys1097 side chain is predicted to form an H-bond/ionic interaction with the carboxylate moiety of **6.75**, while the hydrazone carbonyl is establishing an H-bond reinforced by charge with the protonated nitrogen belonging to the side chain of Lys1463. Hydrophobic interactions are predicted between the benzylidene aromatic ring and the side chain of Phe1440. A T-shaped pi-stacking interaction is established between the side chain of Phe614 and one benzyloxy group while the side chains of Leu1447, Leu607, Leu1088, Leu619, Leu1053, Ile1466, Ile1467 and Val1037 are performing hydrophobic interactions with the three benzyloxy groups.

### **6.2.16 Attempted evaluation of IC<sub>50</sub> values and antibacterial activity for selected compounds.**

As compounds **6.3**, **6.36**, **6.48**,<sup>24</sup> **6.49**,<sup>24</sup> **6.50**,<sup>24</sup> **6.51** and **6.75** exhibited the best percentage inhibition activity at the concentration of 100 μM, attempts were made to establish IC<sub>50</sub> values using *E. coli* RNA polymerase and also to

probe antibacterial activity on selected bacterial strains (please see Appendix I for details).

With regard to IC<sub>50</sub>, unfortunately a full dose response curve could not be determined for any of the compounds under study. This was due to the lack of 100% inhibition when compounds reached the saturation concentration in the assay conditions. This observation may reflect the relative poor solubility of these compounds at higher concentrations as coupled with the rather modest inhibitory activity.

Antibacterial activity for selected compounds is summarized in Table 6.14.

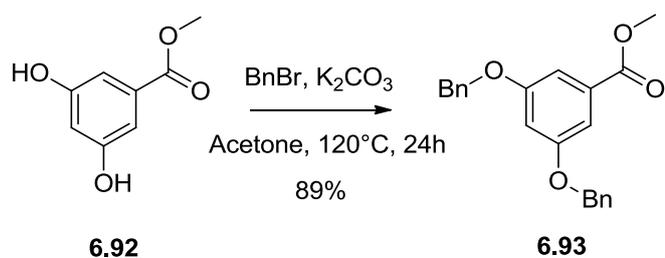
**Table 6.14** MIC determination for selected compounds. <sup>a</sup> MIC values in µg/mL <sup>b</sup> *E. coli* TolC deficient strain.

Compound number	<i>S. aureus</i> SH1000 <sup>a</sup>	<i>E. coli</i> SM1411 <sup>a,b</sup>
<b>6.3</b>	>128	>128
<b>6.36</b>	>128	>128
<b>6.48</b> <sup>24</sup>	>128	>128
<b>6.49</b> <sup>24</sup>	32	>128
<b>6.50</b> <sup>24</sup>	>128	>128
<b>6.51</b>	>128	>128
<b>6.75</b>	>128	>128

Disappointingly, none of the compounds showed significant antibacterial activity with the only exception being compound **6.49**,<sup>24</sup> which showed moderate inhibitory activity with the Gram positive bacterial strain of *S. aureus* SH1000.

### 6.3 Synthesis of intermediate 6.93

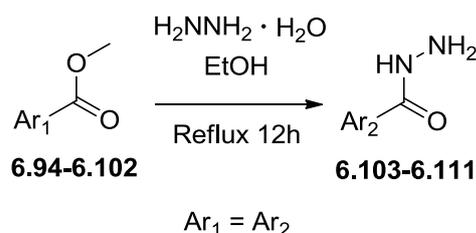
An acetone solution of ester **6.92** was treated with an excess of benzylbromide in presence of an excess of potassium carbonate under reflux to give di-ether **6.93**<sup>26</sup> in good yield (Scheme 6.6).



**Scheme 6.6** Synthesis of **6.93**.<sup>26</sup>

## 6.4 Synthesis of the hydrazides

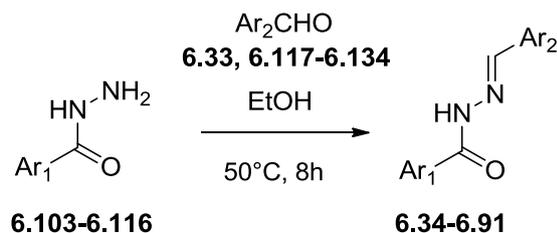
The intermediate hydrazides were readily prepared in excellent yields by adding an excess of hydrazine hydrate to an ethanolic solution of an appropriately prepared (**6.93**)<sup>26</sup> or commercially available ester (Table 6.15, Appendix II, Section A) and heating to reflux for 12 hours (Scheme 6.7). Recrystallization from a suitable solvent afforded the corresponding hydrazides (Table 6.16, Appendix II, Section B).



**Scheme 6.7** Synthesis of **6.103—6.111**.

## 6.5 Synthesis of the acylhydrazone library

A library of acylhydrazone analogues has been synthesized in good yield (Scheme 6.8) by heating to 50 °C for 8 hours an ethanolic solution of equimolar quantities of the prepared (**6.94-6.111**) or commercially available **6.112-6.116** hydrazide (Table 6.17, Appendix II, Section C) and commercially available aldehyde (Table 6.18, Appendix II, Section D). Recrystallization from a suitable solvent or purification *via* mass-directed preparative HPLC afforded the corresponding acylhydrazones (Table 6.19, Appendix II, Section E).



**Scheme 6.8** Synthesis of the hydrazones.

## 6.6 Conclusions

Following the application of a rigorous and accurate docking study, amongst the two discovered bacterial RNAP inhibitors, acylhydrazone **6.3** was selected on the basis of its RNAP inhibitor activity (56% of inhibition at the concentration of 100 $\mu$ M), synthetic amenability and suitability for analogue design. Based upon the predicted binding mode of this compound within RNAP, a SAR study was conducted in order to increase potency and drug likeness properties. A library of acylhydrazone analogues was synthesized after similarity searching and docking studies which allowed the exploration of the chemical space around the initial hit and to prioritize the synthesis of analogues on the basis of the predicted binding mode.

The biological results underlined the importance of hydrophobic interactions for the binding. In fact, any attempt to improve drug likeness by preparing compounds with predicted increased aqueous solubility, led systematically to inactive compounds or to derivatives with improved biological activity but which still suffered from poor solubility at higher concentrations which prevented the establishment of full IC<sub>50</sub> determination. Notably, the calculation of physicochemical descriptors like LogP and LogSw during the analogue design failed in some cases to predict empirical solubility in the specific assay conditions. It was possible to conclude from the SAR studies that effectively targeting a highly lipophilic binding site *via* designing compounds with better physicochemical properties when compared to myxopyronin antibiotics is very challenging.

Interestingly, compounds **6.36**, **6.49**,<sup>24</sup> **6.62** and **6.75** showed better inhibition activity when compared to the parent acylhydrazone **6.3** but disappointingly none of them showed antibacterial activity with the only exception being derivative **6.49**,<sup>24</sup> which possessed moderate activity towards a Gram-positive bacterial strain of *S. aureus* SH1000. It is clear that inhibiting a whole bacterial cell implies crossing biological barriers and in the case of Gram negative strains, additional barriers and active mechanisms of resistance are operating and this explains why it is more easier to obtain inhibitors active in the isolated enzyme as opposed to also exhibiting cellular activity.

With regard to the vHTS protocol, it was possible to conclude that this showed a reasonable hit identification ratio. The presence of false positives in the shortlist and the lack of correlation, in some cases, between the scoring values and the biological activity, underlined the actual limits of scoring functions in evaluating the free energy of binding. In addition, putative binding modes allowed to guide the design process and to explain the biological activity of some analogues, but the well known difficulties of search algorithms to sample a large binding site may lead to inaccuracies in the predictions. The lack of a potent inhibitor prevented the confirmation of binding modes with empirical data derived from X-ray diffraction studies of co-crystal structures of bound inhibitors.

In conclusion, none of the synthesized molecules possessed improved enzymatic inhibitory activity and antibacterial potency when compared to 7-desmethylmyxopyronin B **1.27** despite of their improved synthetical amenability, better binding energy on the basis of the scoring function and more suitable chemical-physical properties as predicted by their descriptors.

## 6.7 References

1. Li, Y. S.; Zhou, L.; Ma, X.; Song, H.; Tang, X. Y., *Med. Chem. Res.* **2012**, *21*, 642-652.
2. Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H., *Cell* **2008**, *135*, 295-307.
3. Belogurov, G. A.; Vassilyeva, M. N.; Sevostyanova, A.; Appleman, J. R.; Xiang, A. X.; Lira, R.; Webber, S. E.; Klyuyev, S.; Nudler, E.; Artsimovitch, I.; Vassilyev, D. G., *Nature* **2009**, *457*, 332-335.
4. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., *J. Chem. Inf. Model.* **2012**, *52*, 1757-1768.
5. Mills, J. E. J.; Dean, P. M., *J Computer-Aided Mol Des* **1996**, *10*, 607-622.
6. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., *Proteins* **2003**, *52*, 609-623.
7. Silver, L. L., *Clin. Microbiol. Rev.* **2011**, *24*, 71-109.
8. Livermore, D. M.; British Soc, A., *J. Antimicrob. Chemoth.* **2011**, *66*, 1941-1944.
9. Tetko, I. V.; Tanchuk, V. Y., *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1136-1145.
10. Lipinski, C. A., *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235-249.
11. Okawa, T.; Osakada, N.; Eguchi, S.; Kakehi, A., *Tetrahedron* **1997**, *53*, 16061-16082.
12. Ohmichi, T.; Maki, A.; Kool, E. T., *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 54-59.
13. Seidler, J.; McGovern, S. L.; Doman, T. N.; Shoichet, B. K., *J. Med. Chem.* **2003**, *46*, 4477-4486.
14. Su, X.; Surry, D. S.; Spandl, R. J.; Spring, D. R., *Org. Lett.* **2008**, *10*, 2593-2596.
15. Clinton, R. O.; Geissman, T. A., *J. Am. Chem. Soc.* **1943**, *65*, 85-87.
16. Reaxys, version 1.7.8. Elsevier, **2012**.
17. Bender, A.; Mussa, H. Y.; Glen, R. C.; Reiling, S., *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1708-1718.
18. Willett, P., *Drug Discovery Today* **2006**, *11*, 1046-1053.

19. Rajamalli, P.; Atta, S.; Maity, S.; Prasad, E., *Chem. Commun.* **2013**, *49*, 1744-1746.
20. Howlader, M. B. H.; Islam, M. S., *Indian J. Chem.* **2007**, *46*, 440-444.
21. Lima, P. C.; Lima, L. M.; da Silva, K. C. M.; Léda, P. H. O.; de Miranda, A. L. P.; Fraga, C. A. M.; Barreiro, E. J., *Eur. J. Med. Chem.* **2000**, *35*, 187-203.
22. Chaur, M. N.; Collado, D.; Lehn, J.-M., *Chem. – Eur. J.* **2011**, *17*, 248-258.
23. Mazzone, G.; Puglisi, G.; Marchetta, G.; Corsaro, A., *J. Het. Chem.* **1984**, *21*, 181-184.
24. Hu, Y.; Lu, X.; Chen, K.; Yan, R.; Li, Q.-S.; Zhu, H.-L., *Bioorg. Med. Chem.* **2012**, *20*, 903-909.
25. Avila, C. M.; Lopes, A. B.; Gonçalves, A. S.; da Silva, L. L.; Romeiro, N. C.; Miranda, A. L. P.; Sant'Anna, C. M. R.; Barreiro, E. J.; Fraga, C. A. M., *Eur. J. Med. Chem.* **2011**, *46*, 1245-1253.
26. Nguyen, T.-T.-T.; Simon, F.-X.; Schmutz, M.; Mesini, P. J., *Chem. Commun.* **2009**, 3457-3459.

## 7. Developing of a novel structure based virtual screening protocol

### 7.1 Virtual screening strategy

As described previously, in the light of the solubility problems and the general lack of potency required for future optimisation encountered in the previous series of synthesized acylhydrazones (Chapter 6), it was decided to develop a novel RNAP inhibitor using a structure-based virtual screening protocol.

In order to maximise the solubility of the newly predicted virtual hits, a preliminary library filtering based on physico-chemical descriptors was performed to restrict the number of screening candidates. After this preliminary step, a docking was then performed using GOLD docking software focussed inside the myxopyronin binding region defined from the crystal structure of the *T. thermophilus* RNA polymerase holoenzyme in complex with the antibiotic myxopyronin A **1.25** (PDB id: 3DXJ<sup>1</sup>).

In contrast to the protocol reported in the previous Chapter (Sections 6.2.1 and 6.2.2), no ligand-based stage was implemented in this virtual screening protocol. The definition of key pharmacophoric points based on the active conformation of 7-desmethyl-myxopyronin B **1.27** in the ligand-based virtual screening protocol reported in Section 6.2.1 permitted the prioritization of molecules showing similarity to the natural antibiotic in terms of pharmacophoric features and to decrease dramatically the number of molecules to be docked. The main disadvantage of this strategy was the identification of molecules sampling a limited number of interaction points present in a very large binding region. The full sampling of the binding site *via* docking allowed the selection of more structurally diverse molecules which also were predicted to interact with different residues when compared to the binding mode observed for the natural antibiotic 7-desmethylmyxopyronin B **1.27**.

In order to counterbalance the increased computational time required for screening a database of molecules without a preliminary ligand-based

screening, the protocol was limited to a small sized compound library, the Medicinal Chemistry and Chemical Biology technology group (MCCB) in-house chemical database, composed of 26,493 molecules. With regard composition, the MCCB database comprises selected collections of commercial compounds purchased from Albany Molecular Research Inc. (AMRI), Chembridge and Asinex databases.

### 7.1.1 Database filtering protocol

The initial database of 26,493 molecules was filtered using Canvas v.1.6, a chemoinformatic module included into Schrodinger Maestro software suite.

The specified parameters for filtering the database were calculated with Canvas v.1.6 with the exception of ALogP and LogSw (logarithm of the predicted solubility in water) values which were calculated with ALOGPS 2.1 software:<sup>2</sup>

$300 \leq \text{M.W.} \leq 500$

$\text{ALogP} \leq 4.0$

$\text{LogSw} \geq -5.0$

Hydrogen bond donors  $\leq 5$

Hydrogen bond acceptors  $\leq 10$

Rotatable bonds  $\leq 5$

Polar surface area  $\leq 150$

The choice of ALogP and LogSw parameters for filtering the database reflected the purpose of considering only molecules with good predicted solubility. Use of the specified M.W. range allowed the filtering of molecular fragments with a limited possibility to establish interactions within the binding site, while limiting the number of rotatable bonds allowed a significant reduction of the computational time spent by the searching algorithm for sampling the conformational space of the docked molecules.

Globally, the above specified physico-chemical parameters used for filtering the database were found to select a subgroup of 8,143 molecules which all satisfied the Lipinski's rule of five<sup>3</sup> and this database differed in more

restrictive range of values for the M.W., ALogP and the number of rotatable bonds.

### 7.1.2 Docking protocol

Docking settings for the structure-based virtual screening protocol were used as described in the docking validation studies reported in Chapter 3.

The crystal structure of the *T. thermophilus* RNA polymerase holoenzyme in complex with the antibiotic myxopyronin A **1.25** (PDB id: 3DXJ<sup>1</sup>), was downloaded from the Protein Databank ([www.rcsb.org](http://www.rcsb.org)).

All water molecules, cofactors and ions were manually removed using Maestro and the binding site within the RNAP was defined as the protein comprised in a sphere with a 15 Å radius surrounding the cocrystallised ligand.

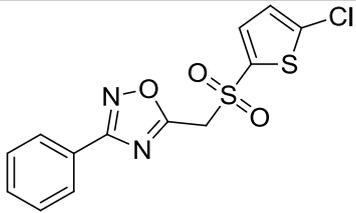
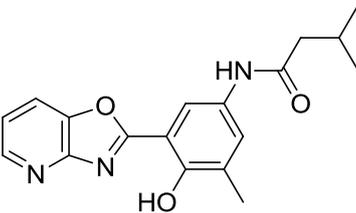
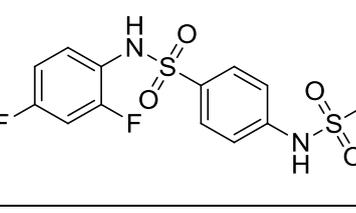
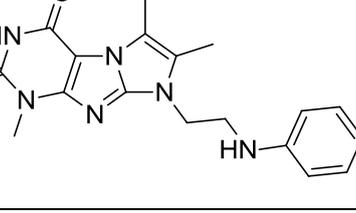
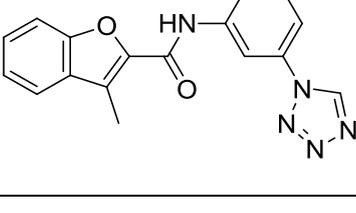
The docking runs were performed using GOLD<sup>4</sup> v4.0.1 (CCDC, Cambridge, UK) docking software by initially using accuracy settings for the genetic searching algorithm which corresponded to the 50% of the default parameters values of the genetic algorithm in order to get a quick and relatively accurate search inside the binding cavity. Compounds were ranked using the Chemscore scoring function generating 100 poses for each ligand. The top 10% of molecules, corresponding to 814 structures, were progressed to a further docking evaluation using the best accuracy settings for the genetic search algorithm which corresponded to the 200% of the default parameters values in order to get a more exhaustive search inside the binding cavity. The final docked poses were re-ranked using Goldscore scoring function after energetic minimization performed using the simplex algorithm implemented in GOLD.

A shortlist of 400 compounds which corresponded to around the 50% of the top-ranked molecules in this final step, was selected for post-docking analysis, which was performed for each ligand among its different docking solutions. Cluster analysis using an average linkage rule was performed using an internal module within the software. Only the best ranked poses of the most populated cluster were considered as the most likely and representative ones. In addition, the docked poses within the binding pocket were analysed applying the same post-docking criteria reported in Chapter 6, Section 6.2.2.

### 7.1.3 vHTS results

Following post-docking analysis, which included the selection of virtual hits on the basis of their chemical diversity, twenty molecules were selected for biological evaluation and are reported in Table 7.1 (please see Section 7.1.4 for biological results).

**Table 7.1** Virtual hits selected for biological evaluation

Comp. number	Structure	M.W.	A <sub>LogP</sub> <sup>(A)</sup>	LogS <sub>w</sub>	Chem score
7.1		340	3.67	-3.86	26.8
7.2		325	3.24	-3.42	30.2
7.3		362	2.26	-3.50	30.0
7.4		371	2.31	-2.53	31.0
7.5		319	2.49	-3.51	33.6

**Table 7.1 Continued**

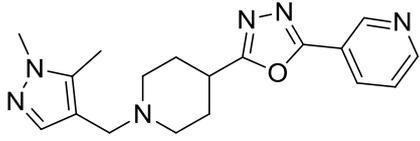
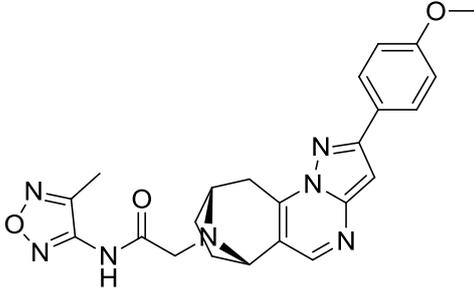
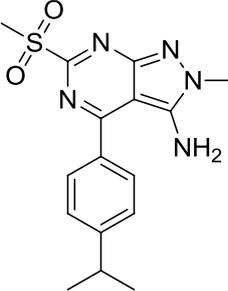
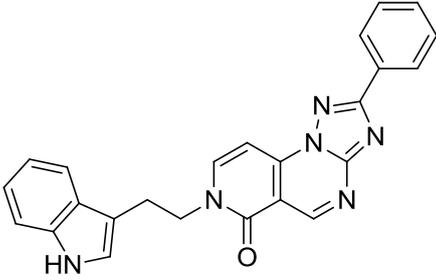
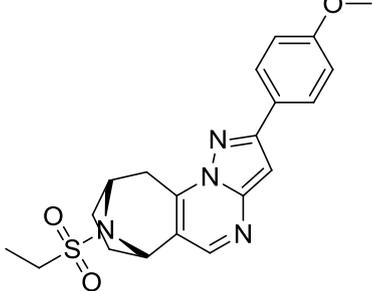
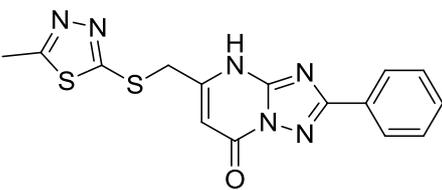
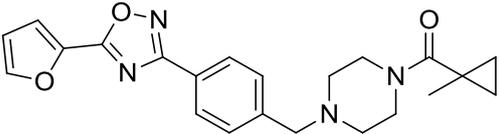
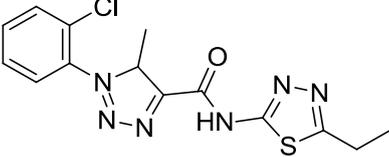
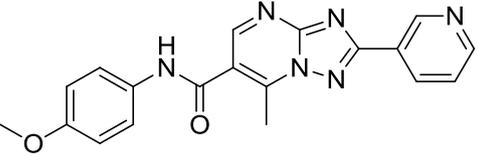
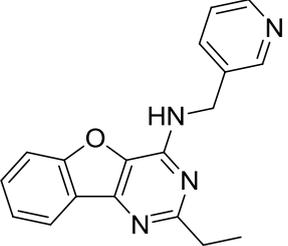
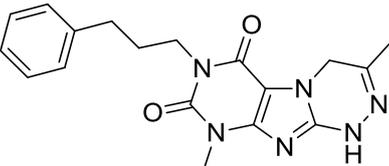
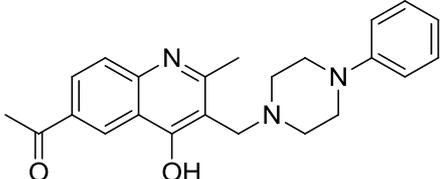
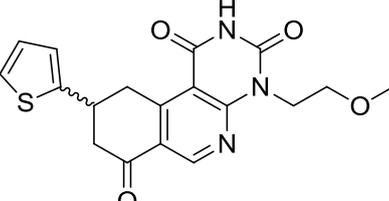
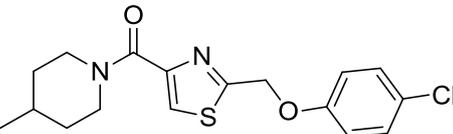
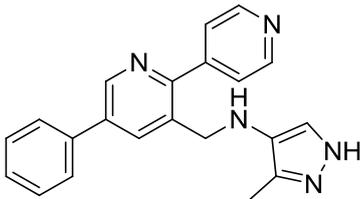
7.6		338	1.80	-3.30	33.8
7.7		445	2.69	-3.39	33.9
7.8		345	2.43	-3.32	28.3
7.9		406	3.62	-4.24	37.9
7.10		398	2.30	-3.55	35.5
7.11		356	2.88	-3.31	32.8

Table 7.1 Continued

7.12		392	2.95	-3.72	35.2
7.13		348	2.27	-3.75	29.0
7.14		360	2.05	-4.12	30.2
7.15		304	3.71	-3.71	33.5
7.16		352	2.53	-2.60	31.1
7.17		375	3.65	-3.48	40.7
7.18		371	1.98	-3.96	31.6 <sup>†</sup>
7.19		350	4.38	-4.74	38.9

**Table 7.1 Continued**

7.20		341	4.00	-4.56	38.6
------	---	-----	------	-------	------

(†) Compound **7.18** was purchased as a racemic mixture and the Chemscore value is referred to its mean value of the two single docked enantiomers.

### 7.1.4 Biological evaluation of the vHTS hits

Compounds **7.1—7.20** were evaluated in the SYBR Green assay<sup>5</sup> (please see Appendix I for details) and the results are summarized in Table 7.2.

**Table 7.2** *In vitro* RNAP percentage inhibition of the vHTS hits

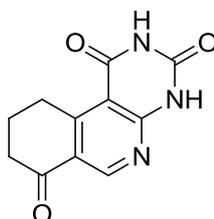
Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M	Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M
7.1	3.3	7.11	6.4
7.2	5.2	7.12	6.4
7.3	7.5	7.13	0
7.4	0.9	7.14	11.6
7.5	4.6	7.15	7.8
7.6	3.0	7.16	3.6
7.7	9	7.17	0
7.8	9.8	7.18	57.1 <sup>†</sup>
7.9	12.1	7.19	1.8
7.10	18.8	7.20	0

(†) % inhibition value for Compound **7.18** refers to the racemic mixture.

Disappointingly, most of the compounds showed only moderate inhibition activity with the only exception being **7.18** which was considered for further investigations and possible chemical optimization.

### 7.1.5 Substructure query-based search and docking of structural analogues of compound 7.18

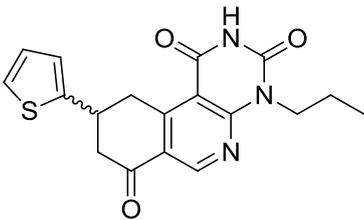
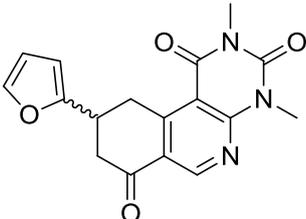
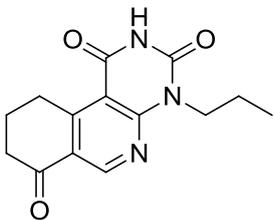
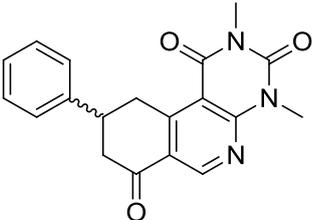
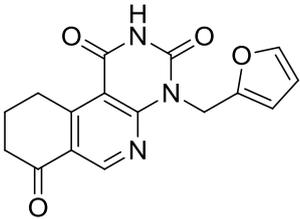
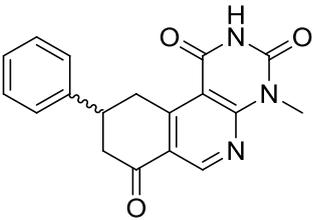
In order to explore the close chemical space of the identified hit from the previous virtual screening studies, a substructure query based search was performed both on the ZINC<sup>6</sup> and MCCB database using ZINC substructure query tool and Canvas module included in Schrodinger Maestro software respectively. Substructure query based search on all atoms was based on a reference structure represented by the central scaffold of the hit **7.18** (Scheme 7.1).



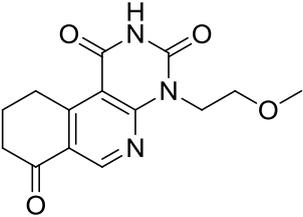
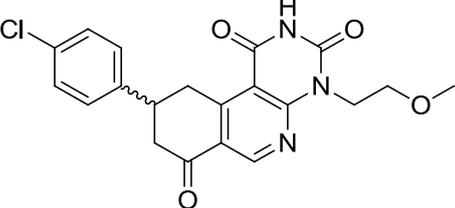
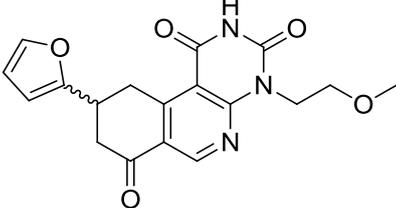
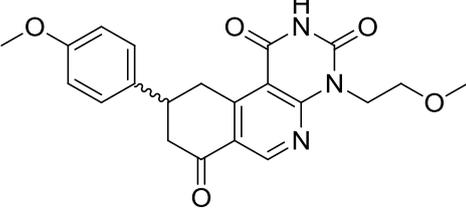
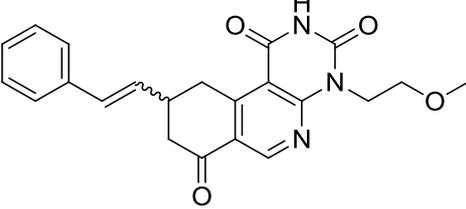
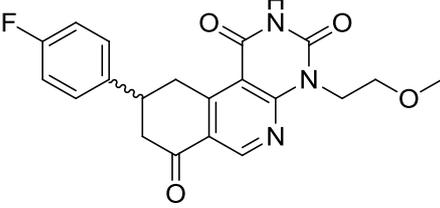
**Scheme 7.1** Central scaffold used for the substructure query based search

The substructure query based search of the MCCB library retrieved only fourteen compounds and they were all present within the Chembridge commercial library. Compounds were then docked following the same settings specified in Section 7.1.2 and it was observed that their predicted binding mode and overall score value was similar to those for compound **7.18** (Table 7.3). These compounds, were therefore all purchased for biological evaluation.

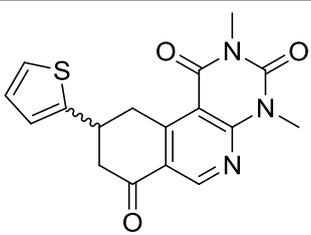
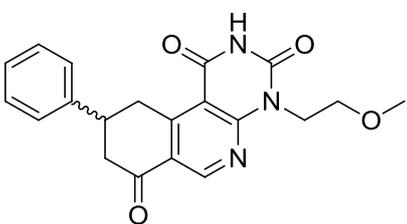
**Table 7.3** Close structural analogues selected for biological evaluation

Comp. number	Structure	M.W.	ALogP <sub>(A)</sub>	LogS <sub>w</sub>	Chem score <sup>†</sup>
7.21		355	2.43	-4.22	29.5
7.22		325	1.03	-2.78	28.6
7.23		273	0.51	-2.36	24.3
7.24		335	1.09	-3.32	28.1
7.25		311	0.82	-2.75	26.7
7.26		321	1.42	-3.42	28.5

**Table 7.3 Continued**

7.27		289	-0.05	-2.26	25.0
7.28		399	2.20	-3.43	28.4
7.29		355	0.71	-2.92	29.2
7.30		395	1.47	-3.75	29.3
7.31		391	1.72	-3.28	31.8
7.32		383	1.77	-3.51	28.5

**Table 7.3 Continued**

<b>7.33</b>		341	2.01	-4.06	30.6
<b>7.34</b>		365	1.40	-3.36	24.7

(†) All chiral compounds were purchased as a racemic mixture and the Chemscore value is referred as the mean value of the two single docked enantiomers.

### 7.1.6 Biological results

Compounds **7.21—7.34** were evaluated in the SYBR Green assay<sup>5</sup> (please see Appendix I for details) and results are summarized in Table 7.4.

**Table 7.4** *In vitro* RNAP percentage inhibition of **7.18** close structural analogues

Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M	Compound number	<i>In vitro</i> RNAP % inhibition at 100 $\mu$ M
<b>7.21</b>	27.8	<b>7.28</b>	30.4
<b>7.22</b>	0	<b>7.29</b>	11.2
<b>7.23</b>	0	<b>7.30</b>	11.5
<b>7.24</b>	6	<b>7.31</b>	4.5
<b>7.25</b>	8.2	<b>7.32</b>	28.7
<b>7.26</b>	10.3	<b>7.33</b>	3.6
<b>7.27</b>	23.5	<b>7.34</b>	7.1

Unfortunately, none of the purchased analogues showed better biological activity than the original hit **7.18**. In order to explore the potential for development of hit **7.18** and possible synthesis of novel analogues, this compound was further evaluated for antibacterial activity (Table 7.5) using selected bacterial strains (please see Appendix I for details).

**Table 7.5** MIC determination for **7.18**. <sup>a</sup> MIC values in  $\mu\text{g/mL}$  <sup>b</sup> *E. coli* TolC deficient strain.

Compound number	<i>S. aureus</i> SH1000 <sup>a</sup>	<i>E. coli</i> SM1411 <sup>a,b</sup>
<b>7.18</b>	>128	>128

Unfortunately, compound **7.18** did not possess antibacterial activity with both Gram positive and Gram negative representative strains which probably reflects the poor penetration of bacterial membranes despite the good inhibition activity displayed by these compounds in the presence of the isolated enzyme.

### 7.1.7 SAR analysis of the purchased analogues

On the basis of the previous biological results of the purchased analogues of hit **7.18** (Section 7.1.6), SAR were formulated. Globally, the thiophen-2-yl moiety along with the *N*-methoxyethyl chain are essential for the biological activity and no alternative moieties conferred better activity. In particular, the methoxy group at the level of the *N*-methoxyethyl chain is critical for the activity and *N*-alkyl derivatives are almost inactive, showing a negative trend of biological activity as the length of the chain is decreased as shown in compounds **7.21** and **7.33**. Interestingly, substitution of the thiophen-2-yl moiety with bioisosteres as shown in compounds **7.29** and **7.34**, is detrimental for the activity. Replacement of thiophen-2-yl moiety with groups showing increased steric hindrance in compounds **7.28**, **7.30**, **7.31** and **7.32** is negative for the activity, but the presence of a moderately electron withdrawing group in para position of the phenyl ring is better tolerated when comparing biological

activities of compounds **7.28** and **7.32** with **7.30** and **7.31**. When comparing the importance for the biological activity of the thiophen-2-yl and *N*-methoxyethyl moieties in compounds **7.27** and **7.33** respectively, the *N*-methoxyethyl chain is more critical for the activity. Methylation of the NH between the two keto groups is detrimental for the activity as shown by compounds **7.24** and **7.26**.

### 7.1.8 Putative binding mode of the identified hit

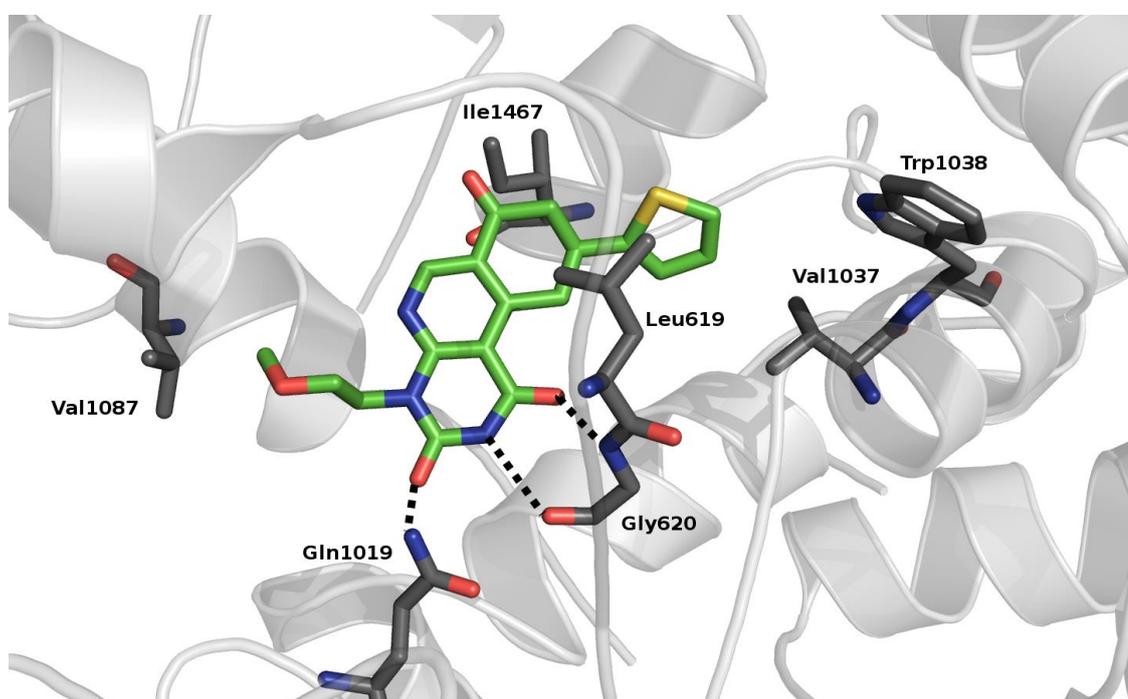
A putative binding mode of the most active derivative, compound **7.18**, is shown in Figures 7.1 and 7.2 for both the enantiomers, (*S*) and (*R*) respectively. Whilst compound **7.18** was available only as a racemic mixture from the vendor, a different binding mode was found for both enantiomers and a slightly better Chemscore value was found for the (*R*) enantiomer as shown in Table 7.6.

**Table 7.6** Chemscore value for both enantiomers of **7.18**

Compound number	Enantiomer	Chemscore
<b>7.18</b>	( <i>S</i> )	28.6
<b>7.18</b>	( <i>R</i> )	34.6

The following interactions between the (*S*) enantiomer of **7.18** and the myxopyronin binding site were identified (Figure 7.1): a H-bond was predicted between the side chain of Gln1019 and one carbonyl of the pyrimidine ring while the carbonyl and the NH of Gly620 are forming a H-bond with the NH and the carbonyl of the pyrimidine ring respectively. The *N*-methoxyethyl chain is establishing hydrophobic interactions with the side chain of Val1087 while the side chains of Ile1467, Leu619 and Val1037 are performing the same interactions with the dihydroisoquinolinone ring and the thiophen-2-yl moiety of **7.18**. Notably, the predicted binding mode for this enantiomer, is able to explain some aspects of the previous SAR analysis. The importance of the *N*-methoxyethyl chain and the thiophen-2-yl moiety in addition to the lower biological activity of shorter *N*-alkyl derivatives, may be explained by their predicted hydrophobic interactions. The detrimental effects upon biological

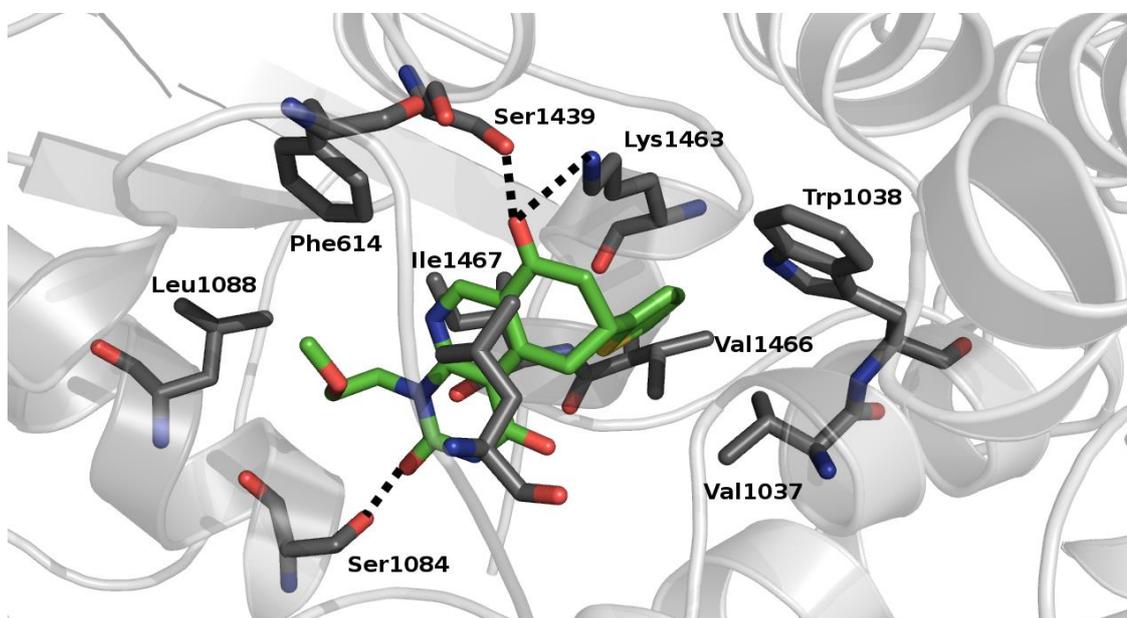
activity of moieties offering steric hindrance that is bigger than that predicted for the thiophen-2-yl group, is predicted by the possible steric clash with the side chain of Trp1038. Finally the loss of activity following the methylation of the pyrimidine NH is explained by the loss of its hydrogen bond with the backbone of Gly620.



**Figure 7.1** Putative binding mode of the (*S*) enantiomer of **7.18**

The following interactions between the (*R*) enantiomer of **7.18** and the myxopyronin binding site were identified (Figure 7.2): a H-bond was predicted between the side chain of Ser1084 and one carbonyl of the pyrimidine ring while the carbonyl and the dihydroisoquinolinone ring is forming H-bonds with the side chain of Ser1439 and Lys1463. The *N*-methoxyethyl chain is establishing hydrophobic interactions with the side chain of Val1087 while the side chains of Phe614, Val1466, Ile1467 and Val1466 are performing the same interactions with the dihydroisoquinolinone ring and the thiophen-2-yl moiety of **7.18**. A T-shaped pi-stacking interaction was predicted between the side chain of Trp1038 and the thiophen-2-yl moiety. Interestingly, the predicted binding mode for this enantiomer, is able to explain, by analogy with the previous

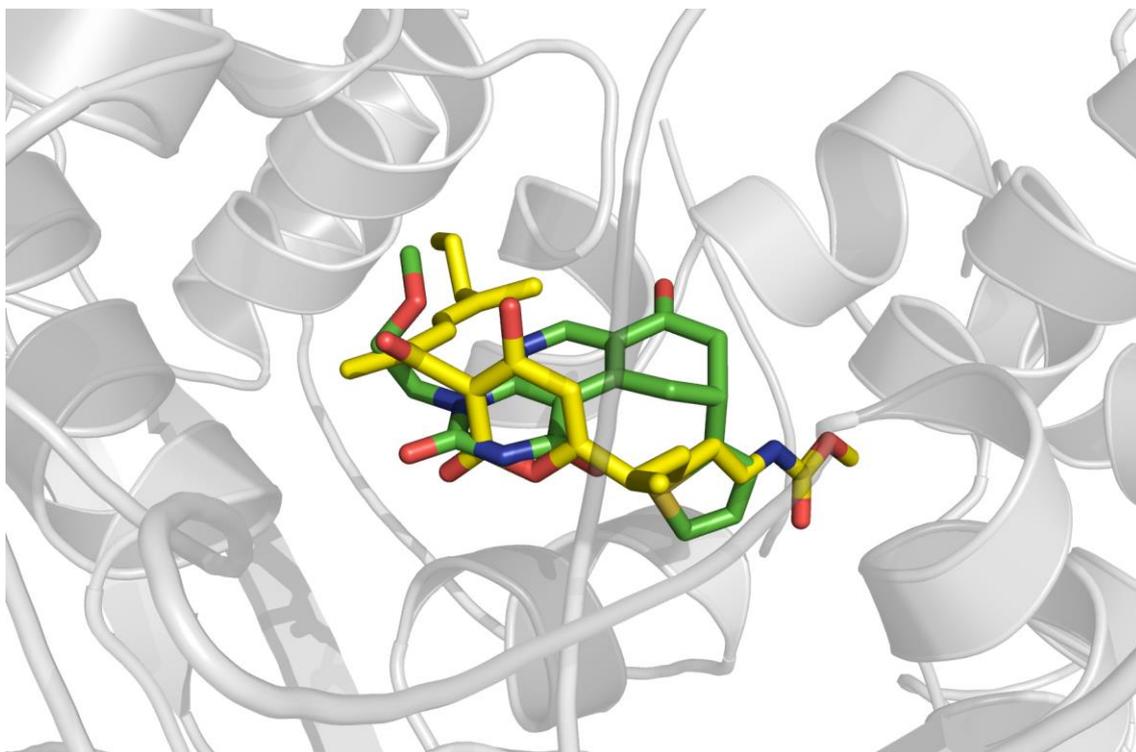
binding mode, some aspects of the previous SAR analysis. The importance of the *N*-methoxyethyl chain and thiophen-2-yl moiety in addition to the lower biological activity of shorter *N*-alkyl derivatives are explained by their predicted hydrophobic interactions. The detrimental effects on biological activity of moieties with bigger steric hindrance than the thiophen-2-yl group is predicted by the possible clash with the side chain of Val1466. The T-shaped pi-stacking interaction between the side chain of Trp1038 and the thiophen-2-yl moiety could explain why analogue **7.29** has lower activity than **7.18**, in fact, the electron rich furan ring could suffer from repulsion of the neighbouring electron rich indole ring while the thiophene moiety of **7.18**, having a lower electron density, could establish a more favourable interaction with the side chain of Trp1038.



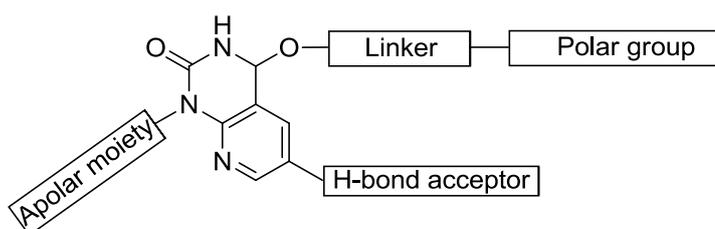
**Figure 7.2** Putative binding mode of the (*R*) enantiomer of **7.18**

Overlay of the (*R*) enantiomer of compound **7.18** with myxopyronin A **1.25** into its binding site (Figure 7.3) revealed that the pyrimidine ring is occupying the same region of the central  $\alpha$ -pyrone ring performing a similar network of H-bond interactions while the *N*-methoxyethyl chain and the thiophen-2-yl moiety is predicted to occupy the same hydrophobic regions of myxopyronin A. The

dihydroisoquinolinone ring system is involved in some important interactions within the binding site (Figure 7.3) and it is also providing the function of a rigid spacer for the thiophen-2-yl moiety. This putative binding mode could provide useful ideas for future analogue design based on a central *N*-substituted pyrido[2,3-*d*]pyrimidine-based scaffold linked to a suitable polar group mimicking the interactions of the ene-carbamate as shown in Scheme 7.2.



**Figure 7.3** Overlay between myxopyronin A **1.25** in yellow sticks and the (*R*) enantiomer of **7.18** in green.



**Scheme 7.2** General structure of potential analogues for future design

## 7.2 Conclusions

In order to avoid solubility problems in the assay, a structure-based virtual screening was conducted on the (MCCB) in-house chemical database applying a preliminary library filtering based on physico-chemical descriptors.

Application of this virtual screening protocol was successful and identified a hit between a short list of twenty molecules which did not show solubility issues.

In order to explore the close chemical space of the identified hit, a substructure query based search and docking studies were conducted on the on the ZINC<sup>6</sup> and MCCB database, identifying fourteen analogues and SAR analysis was performed on these molecules. Interestingly, the putative binding modes for both enantiomers of **7.18** were able to explain SAR and this appeared to validate the putative binding mode and to confirm the good performance of the searching algorithm. On the contrary, the lack of biological activity of the selected hit analogues showing a similar score underlines the limitations of the scoring function in correlating the biological activity with the estimated binding energy.

Considering the absence of antibacterial activity showed by the most active inhibitor, compound **7.18**, in addition to the inability to improve inhibitory activity *via* analysis of structural analogues, compound **7.18** is not a viable candidate for future optimisation by synthesis but its putative binding mode offers useful ideas for future design.

### 7.3 References

1. Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H., *Cell* **2008**, *135*, 295-307.
2. Tetko, I. V.; Tanchuk, V. Y., *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1136-1145.
3. Lipinski, C. A., *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235-249.
4. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., *Proteins* **2003**, *52*, 609-623.
5. Ohmichi, T.; Maki, A.; Kool, E. T., *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 54-59.
6. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., *J. Chem. Inf. Model.* **2012**, *52*, 1757-1768.

## 8. Conclusions and future work

The work described in this thesis was directed towards the identification of novel bacterial RNAP inhibitors *via* the application of rational computational design, synthesis and biological evaluation.

In particular, this research was focussed on the myxopyronin binding region, the 'switch region', a promising 'hot spot' which is essential for the conformational changes within RNAP required during DNA transcription.

The computational rational design phase of the work involved the combined use of ligand- and structure-based techniques. Given the availability of different docking algorithms as well as X-ray co-crystal structures, a validation study was conducted in Chapter 3 to identify the optimal software parameters and the most appropriate crystal structure for use in the subsequent inhibitor design studies.

Initially, the work was based around a small molecule RNAP inhibitor, **4.1**, reported on Pubchem.<sup>1</sup> The absence of information on the putative binding site of this molecule and the lack of any SAR data required attempts to be made to identify the binding region of this molecule within RNAP *via* the use of extended docking studies on all the known inhibitor binding sites followed by SAR exploration. Unfortunately, the apparent inhibitory activity of this compound could not be reproduced using the assay system at Leeds which may result from differences between the Leeds-based biological assay<sup>2</sup> and that reported in literature.<sup>3</sup>

A scaffold-hopping strategy was then conducted (Chapter 5) on the basis of a pharmacophore hypothesis based upon structural elements of myxopyronin. Ligand-based virtual screening followed by docking studies were performed in order to attempt to identify putative RNAP inhibitors which may overcome the unfavourable pharmacokinetic and chemical properties of myxopyronin A (Myx) **1.25**. Unfortunately, despite of the favourable physicochemical properties of the selected hits, no RNAP inhibition was observed. These results revealed the limitations of a virtual screening protocol relying mostly on ligand based techniques.

In Chapter 6, based upon a theoretical study described in the literature,<sup>4</sup> two bacterial RNAP inhibitors were identified *via* a rigorous and accurate combined ligand- and structure-based virtual screening protocol. Acylhydrazone **6.3** was selected and a SAR study was conducted in order to increase its potency and drug likeness on the basis of its putative binding mode. A library of acylhydrazone analogues was synthesized after similarity-based virtual-screening and docking studies, which allowed to explore the close chemical space of the initial hit and to prioritize the synthesis of analogues. Compounds **6.36**, **6.49**,<sup>5</sup> **6.62** and **6.75** showed better inhibition activity when compared to the parent acylhydrazone **6.3** and moreover, derivative **6.49**<sup>5</sup> possessed moderate activity towards a Gram-Positive bacterial strain of *S. aureus* SH1000. Unfortunately, during IC<sub>50</sub> determination studies, it was found that the most active compounds of this series had relatively poor aqueous solubility. Nonetheless, the results from the SAR study on this compound series underline the importance of hydrophobic interactions for the binding of these molecules to the myxopyronin binding site within RNAP and attempts to improve drug likeness by preparing compounds with predicted increased aqueous solubility (and corresponding decreased hydrophobicity) led systematically to inactive compounds or to derivatives with improved biological activity but which were poorly soluble. In addition, whilst a successful virtual screening protocol was conducted, this work showed that targeting effectively a highly lipophilic binding site *via* the design of compounds with better physicochemical properties than myxopyronin is very challenging. Furthermore, it was found that the use of physicochemical descriptors for predicting solubility is not always reliable when applied in circumstances different from model systems like the specific assay conditions.

In Chapter 7, in order to prevent solubility problems previously encountered in the assay, structure-based virtual screening was conducted using an in-house (MCCB) chemical database, applying preliminary library filtering based on physico-chemical descriptors. Given the small size of the screened library, a new structure-based virtual screening protocol was successfully applied and a hit, compound **7.18**, was identified. In order to explore the close chemical space of the identified hit, a substructure query

based search followed by docking studies were conducted identifying fourteen analogues and SAR analysis was performed on these molecules. Interestingly, the putative binding mode for **7.18** was generally able to explain the observed SAR in which appears to validate the modelling and design protocol. In contrast, the lack of biological activity of selected analogues showing a similar score to the original hit underlines the limitations of the scoring function in correlating the biological activity with the estimated binding energy. Compound **7.18** and its analogues did not show antibacterial activity and none of the analogues showed increased inhibition when compared with the parent molecule **7.18**. For the previously cited reasons, no additional synthetic exploration was undertaken even if the putative binding mode of **7.18** offers useful ideas for future design.

With the exception of compound **6.49**,<sup>5</sup> the lack of antibacterial activity among the synthesized enzyme inhibitors from this project underlines the challenge of crossing bacterial cell membranes and this explains why it is easier to obtain inhibitors active against the isolated enzyme as opposed to also exhibiting cellular activity.

Whilst there is no conclusive evidence that the synthesized inhibitors bind to the myxopyronin binding region, future work could involve the synthesis of new analogues with increased potency for determining a X-ray co-crystal complex with bacterial RNAP. These crystallography data would provide increased reliability of the predictions *via* computational studies and will represent an excellent starting point for future structure-based design.

Future analogue design could be based either on the ideas reported in Scheme 7.2 in Chapter 7 or applying the successful virtual screening protocols reported in Chapters 6 and 7 after preliminary library filtering based on physico-chemical descriptors.

The myxopyronin binding region of bacterial RNAP is a highly lipophilic binding site and the design of a ligand with a favourable aqueous solubility profile represents a challenge for drug design. A possible strategy to overcome these intrinsic limitations could be represented by focussing attention on alternative known inhibitor binding regions of bacterial RNAP, e.g. streptolydigin<sup>6</sup> and tagetitoxin<sup>7</sup> binding sites, characterised by a lower

hydrophobicity of their key residues which are still unexplored in terms of drug design and synthesis.

With regard to the lack of consensus between different biological assays encountered in Chapter 4, an interesting recent paper<sup>8</sup> offered useful insights into this issue. In fact, the authors<sup>8</sup> demonstrated that the template choice can influence the inhibitory potency of bacterial RNAP inhibitors. In particular, our assay, when compared with the other assays described in the paper,<sup>8</sup> underestimates the potency of myxopyronin while in the case of another inhibitor of the same binding region, corallopyronin, gave comparable results. The authors<sup>8</sup> recommended the use of a double-stranded, preferentially prokaryotic promoter-containing DNA template for the determination of inhibitory potencies of compounds targeting bacterial RNAP. In the light of these findings and considering that the template in our assay is a single strand of circular DNA lacking of promoter, it might be useful to develop an alternative assay in future for evaluating the synthesized molecules of this thesis which were designed and assumed as putative inhibitors of the myxopyronin binding region.

## 8.1 References

1. Wang, Y.; Suzek, T.; Zhang, J.; Wang, J.; He, S.; Cheng, T.; Shoemaker, B. A.; Gindulyte, A.; Bryant, S. H., *Nucl. Acids Res.* **2014**, *42*, D1075-D1082.
2. Ohmichi, T.; Maki, A.; Kool, E. T., *Proc. Natl. Acad. Sci. U.S.A* **2002**, *99*, 54-59.
3. Kozlov, M.; Bergendahl, V.; Burgess, R.; Goldfarb, A.; Mustaev, A., *Anal. Biochem.* **2005**, *342*, 206-213.
4. Li, Y. S.; Zhou, L.; Ma, X.; Song, H.; Tang, X. Y., *Med. Chem. Res.* **2012**, *21*, 642-652.
5. Hu, Y.; Lu, X.; Chen, K.; Yan, R.; Li, Q.-S.; Zhu, H.-L., *Bioorg. Med. Chem.* **2012**, *20*, 903-909.
6. Vassilyev, D. G.; Sekine, S.; Laptenko, O.; Lee, J.; Vassilyeva, M. N.; Borukhov, S.; Yokoyama, S., *Nature* **2002**, *417*, 712-719.
7. Vassilyev, D. G.; Svetlov, V.; Vassilyeva, M. N.; Perederina, A.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Artsimovitch, I., *Nat. Struct. Mol. Biol.* **2005**, *12*, 1086-1093.
8. Hauptenthal, J.; Huesecken, K.; Negri, M.; Maurer, C. K.; Hartmann, R. W., *Antimicrob. Agents Chemother.* **2012**, *56*, 4536-4539.

## 9. Experimental section

### 9.1 General Procedures and Instrumentation

All reagents were obtained from Sigma-Aldrich chemical company and/or Acros and/or Alfa-Aesar and/or TCI UK and were not further purified before use.

All reactions, unless otherwise specified, were performed under a positive pressure of dry, oxygen-free nitrogen.

Glassware for reactions carried out under dry conditions was washed with acetone, dried overnight at 125 °C and cooled under a stream of dry nitrogen prior to use.

Solvents were removed under reduced pressure using a Buchi rotary evaporator connected to a diaphragm pump. This was followed by drying under high vacuum using a rotary oil pump at 0.5 mmHg.

Analytical thin layer chromatography (TLC) was performed on Merck aluminium backed TLC silica gel 60 F<sub>254</sub> sheets and these were visualised using ultraviolet lamp ( $\lambda_{\text{max}} = 254 \text{ nm}$ ) or other developing agents (potassium permanganate, 2,4-dinitrophenylhydrazine or bromo-cresol green) where appropriate. Silica gel 60 (particle size 37-70  $\mu\text{m}$ ) supplied by E.M. Merck was employed for flash chromatography.

<sup>1</sup>H and <sup>13</sup>C NMR spectra were measured on a Bruker DPX300, Avance 500 or on a Bruker DRX500 Fourier transform spectrometer and chemical shifts are reported in parts per million (ppm) downfield from tetramethylsilane (TMS) in  $\delta$  units and coupling constants (*J*) are given in Hertz (Hz).

TMS was defined at 0 ppm for <sup>1</sup>H NMR spectra and the central peak of DMSO-*d*<sub>6</sub> septet was also defined as 39.5 ppm for <sup>13</sup>C NMR spectra.

The following abbreviations: s, singlet; br s, broad singlet; d, doublet; t, triplet; q, quartet; quint, quintet; sext, sextet; sept, septet; m, multiplet; dd, double doublet; dt, double triplet; ddd, double double doublet and apparent multiplicities (app., e.g. app. d denotes an apparent doublet); Ar, aromatic; qt, quaternary; are used when describing the <sup>1</sup>H NMR spectra. Where appropriate,

proton and carbon assignment has been based on COSY, DEPT 135, DEPT 90, HMQC, HMBC and NOESY spectra.

Infrared spectra (IR) were recorded on a Nicolet Avatar 300 Ft-IR or a Bruker Alpha Ft-IR Platinum ATR spectrophotometer. The vibrational frequencies are reported in wavenumbers ( $\text{cm}^{-1}$ ).

Mass spectra were recorded on a GTC Premier Micromass spectrometer for impact ionisation (EI) or fast atom bombardment (FAB) while for electron spray ionisation (ESI), a Bruker Daltonics microTOF or a Micromass LCT-KA11 spectrometer was used. Isotopic distributions in routine mass spectra were as expected.

HPLC analyses were carried out using:

**Method A;** Agilent 1290 infinity LC system equipped with a column Ascentis Express C18 (5 x 2.1mm, 2.7 $\mu\text{m}$ ) using a diode array detection system. Samples were eluted with a gradient of acetonitrile (5-95%) / water in the presence of 0.1% TFA at a flow rate of 0.5 ml/min over 5 minutes.

**Method B;** Agilent 1290 infinity LC system equipped with a column Ascentis Express C18 (5 x 2.1mm, 2.7 $\mu\text{m}$ ) using a diode array detection. Samples were eluted with a gradient of acetonitrile (5-95%) / water at a flow rate of 0.5ml/min over 5 minutes.

**Method C;** Agilent 1290 infinity LC system equipped with a column Ascentis Express C18 (5 x 2.1mm, 2.7 $\mu\text{m}$ ) using a diode array detection. Samples were eluted with a gradient of acetonitrile (50-95%) / water at a flow rate of 0.5 ml/min over 5 minutes.

Purifications *via* mass-directed preparative HPLC were carried out on Agilent 6100 series single quad mass spectrometer equipped with a XBridge Prep C18 (5 $\mu\text{m}$  OBD 19x100 mm) column at a flow rate of 20 ml/min using a mobile phase consisting of methanol / water at an appropriate gradient in the presence of 0.1% formic acid.

HPLC purity was reported as % values.

Melting points were determined on a Reichert Hot Stage apparatus and are uncorrected.

Combustion analyses were performed with a Carlo Erba elemental analyser MOD 1106 instrument.

## 9.2 General experimental methods

### Method A: Synthesis of sulphonamide derivatives

A solution of the sulphonyl chloride (1 eq) in dichloromethane (2.5 mL) was added dropwise to a stirred solution of the amine (2 eq) in dichloromethane (2.5 mL) and pyridine (5 mL) cooled in an ice bath. This solution was stirred at 0 °C for 2 hours and then allowed to warm to room temperature over 12 hours under constant stirring. The reaction mixture was poured into dichloromethane (50 mL) and the organic layer washed successively with 1M HCl solution (3 x 50 mL), water (50 mL), brine (50 mL), saturated aqueous NaHCO<sub>3</sub> (2 x 50 mL) and brine again (50 mL). The resulting solution was dried (MgSO<sub>4</sub>) and the solvent removed *in vacuo* to give the corresponding sulphonamides, which were recrystallised from MeOH / water.

### Method B: Synthesis of urea derivatives

A solution of amine (1 eq) in dichloromethane (4 mL) was added dropwise to a solution of isocyanate (1.1 eq) in dichloromethane (4 mL) at room temperature and stirred for 7 hours. The resulting white slurry was cooled at 0 °C, filtered and washed with cold dichloromethane. The resulting solid was recrystallised from dichloromethane.

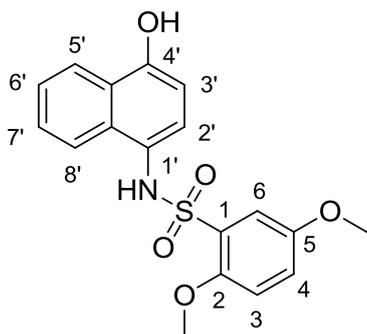
### Method C: Synthesis of hydrazide derivatives

To a solution of an appropriate ester (1 eq) in ethanol (10 mL) was added an excess of hydrazine hydrate. The mixture was stirred and heated at reflux for 12 hours unless otherwise specified. The reaction mixture was cooled to room temperature and the precipitate formed was removed *via* filtration and dried. The resulting residue was recrystallized from a suitable solvent to give the corresponding hydrazide.

### **Method D: Synthesis of acylhydrazone derivatives**

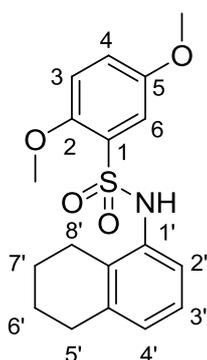
To a solution of hydrazide (1 eq) in ethanol (10 mL) was added the corresponding aldehyde derivative (1 eq). The mixture was heated to 50 °C and stirred for 8 hours unless otherwise specified. The reaction mixture was cooled to room temperature and poured into water and the precipitate formed was removed *via* filtration and dried. The resulting residue was recrystallized from a suitable solvent or purified *via* mass-directed preparative HPLC to give the corresponding acylhydrazone.

### ***N*-(4-hydroxynaphthalen-1-yl)-2,5-dimethoxybenzenesulfonamide (4.1)**



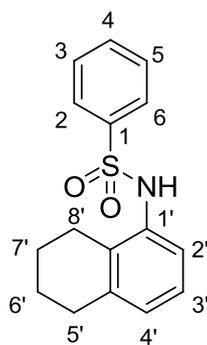
A solution of 2,5-dimethoxybenzenesulphonyl chloride **4.8** (236 mg, 1.0 mmol) in dichloromethane (5 mL) was added dropwise to a stirred solution of 4-aminonaphthalen-1-ol hydrochloride **4.35** (220 mg, 1.1 mmol) in dichloromethane (5 mL) and pyridine (10 mL) cooled in an ice bath. This solution was stirred for 2 hours and then allowed to warm to room temperature over 12 hours under constant stirring. The reaction mixture was poured into dichloromethane (50 mL) and the organic layer washed successively with 1M HCl solution (3 x 50 mL), water (50 mL), brine (50 mL), saturated aqueous NaHCO<sub>3</sub> (2 x 50 mL) and brine (50 mL). The resulting solution was dried (MgSO<sub>4</sub>) and the solvent removed *in vacuo* to give the corresponding dark red solid, which was purified using column chromatography on silica gel eluting with petroleum ether / EtOAc (80:20) and the resulting dark red solid was recrystallised from MeOH / water to give the title compound **4.1** (270 mg, 0.75 mmol, 75%) as light pink plates m.p. 119—121 °C. *R<sub>f</sub>* 0.60 (1:1 petroleum ether—EtOAc); HPLC (Method B), R.t. 3.17 min, (100%); δ<sub>H</sub> (300 MHz, DMSO-*d*<sub>6</sub>); 10.24 (1H, s, NH or OH), 9.60 (1H, s, NH or OH), 8.04-8.09 (2H, m, Ar-H<sub>5',8'</sub>), 7.38-7.48 (2H, m, Ar-H), 7.12-7.14 (2H, m, Ar-H), 6.94-6.99 (2H, m, Ar-H), 6.70 (1H, d, *J* = 8.1, Ar-H<sub>2'</sub>), 3.87 (3H, s, OCH<sub>3</sub>), 3.61 (3H, s, OCH<sub>3</sub>); δ<sub>C</sub> (75 MHz, DMSO-*d*<sub>6</sub>); 152.2 (4'-C), 151.9 (2-C or 5-C), 150.4 (2-C or 5-C), 131.7 (Ar-C<sub>qt</sub>), 128.1 (Ar-C<sub>qt</sub>), 126.1 (Ar-C), 125.2 (Ar-C), 124.7 (Ar-C), 123.3 (Ar-C<sub>qt</sub>), 123.2 (Ar-C), 122.0 (Ar-C), 119.4 (Ar-C), 114.6 (Ar-C), 113.8 (Ar-C), 107.2 (2'-C), 104.5 (Ar-C<sub>qt</sub>), 56.1 (OCH<sub>3</sub>), 55.5 (OCH<sub>3</sub>); ν<sub>max</sub>/cm<sup>-1</sup> (neat); 3404, 3287, 1610, 1510, 1310, 1280; *m/z* (ES) 382.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 382.0721. C<sub>18</sub>H<sub>17</sub>NO<sub>5</sub>S requires MNa, 382.0720).

### 2,5-dimethoxy-N-(5,6,7,8-tetrahydronaphthalen-1-yl)benzenesulfonamide (4.16)



Prepared *via* general method A using 5,6,7,8-tetrahydronaphthylamine **4.2** (0.280 mL, 2.0 mmol) and 2,5-dimethoxybenzenesulphonylchloride **4.8** (238 mg, 1.0 mmol). Recrystallization from methanol / water gave the title compound **4.16** (205 mg, 0.59 mmol, 59%) as light-pink needles m.p. 108—109 °C.  $R_f$  0.40 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 3.39 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 9.07 (1H, s, NH), 7.18-7.19 (2H, m, Ar-H benzene), 7.11-7.13 (1H, m, Ar-H benzene), 6.95 (1H, app. t,  $J = 7.6$ , Ar-H $_3$ ), 6.88 (1H, app. d,  $J = 7.6$ , Ar-H $_4$ ), 6.76 (1H, app. d,  $J = 7.6$ , Ar-H $_2$ ), 3.82 (3H, s, OCH $_3$ ), 3.70 (3H, s, OCH $_3$ ), 2.60-2.70 (4H, m, 2H $_{5'}$  and 2H $_{8'}$ ), 1.61-1.64 (4H, m, 2H $_{6'}$  and 2H $_{7'}$ ),  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 152.0 (2-C or 5-C), 150.3 (2-C or 5-C), 137.7 (1'-C), 134.8 (Ar-C $_{qt}$ ), 133.3 (Ar-C $_{qt}$ ), 128.9 (Ar-C $_{qt}$ ), 127.0 (Ar-C tetrahydronaphthalene), 125.2 (Ar-C tetrahydronaphthalene), 123.1 (Ar-C tetrahydronaphthalene), 119.4 (Ar-C benzene), 114.5 (Ar-C benzene), 114.0 (Ar-C benzene), 56.2 (OCH $_3$ ), 55.7 (OCH $_3$ ), 29.0 (5'-C), 24.2 (8'-C), 22.3 (6'-C or 7'-C), 22.2 (6'-C or 7'-C);  $\nu_{max}/cm^{-1}$  (neat); 3309, 3287, 2945, 1584, 1490, 1466, 1401, 1222, 1160, 1038;  $m/z$  (ES) 370.1 (100%, MNa $^+$ ); (Found MNa $^+$ , 370.1097. C $_{18}$ H $_{21}$ N $_4$ O $_4$ S requires MNa, 370.1083).

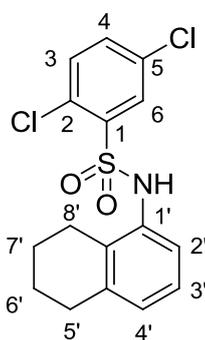
### N-(5,6,7,8-tetrahydronaphthalen-1-yl)benzenesulfonamide (4.17)



Prepared *via* general method A using 5,6,7,8-tetrahydronaphthylamine **4.2** (0.560 mL, 4.0 mmol) and benzenesulphonylchloride **4.7** (0.256 mL, 2.0 mmol). Recrystallization from methanol / water gave the title compound **4.17** (356 mg, 1.24 mmol, 62%) as orange platelets m.p. 151—153 °C.  $R_f$  0.52 (40:10 petroleum ether—EtOAc); (Found: C, 76.5; H, 6.75; N, 10.4; C $_{17}$ H $_{18}$ N $_2$ O requires C, 76.7; H, 6.81; N, 10.5%). HPLC (Method A), R.t. 3.36 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 9.42 (1H, s, NH), 7.63-7.68 (2H, m, Ar-H benzene), 7.54-7.58 (3H,

m, Ar-H *benzene*), 6.98 (1H, app. t,  $J = 7.6$ , Ar-H<sub>3</sub>), 6.92 (1H, dd,  $J = 7.6$ , 1.1, Ar-H<sub>4</sub>'), 6.75 (1H, dd,  $J = 7.6$ , 1.1, Ar-H<sub>2</sub>'), 2.66 (2H, t,  $J = 6.3$ , H<sub>5</sub>'), 2.39 (2H, t,  $J = 6.3$ , H<sub>8</sub>'), 1.49-1.60 (4H, m, 2H<sub>6</sub>' and 2H<sub>7</sub>'),  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 140.7 (1-C or 1'-C), 137.9 (1-C or 1'-C), 134.4 (Ar-C<sub>qt</sub>), 133.4 (Ar-C<sub>qt</sub>), 132.6 (Ar-C), 129.1 (Ar-C), 127.3 (Ar-C), 126.4 (Ar-C), 125.3 (Ar-C), 123.6 (Ar-C), 29.0 (5'-C), 24.2 (8'-C), 22.2 (6'-C or 7'-C), 22.1 (6'-C or 7'-C);  $\nu_{max}/cm^{-1}$  (neat); 3262, 2940, 1911, 1783, 1580, 1455, 1397, 1323, 1158, 1090;  $m/z$  (ES) 310.1 (90%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 310.0873. C<sub>16</sub>H<sub>17</sub>NO<sub>2</sub>S requires MNa, 310.0872).

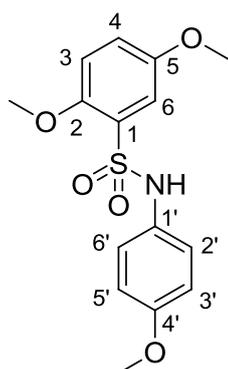
### 2,5-dichloro-N-(5,6,7,8-tetrahydronaphthalen-1-yl)benzenesulfonamide (4.18)



Prepared *via* general method A using 5,6,7,8-tetrahydronaphthylamine **4.2** (0.28 mL, 2.0 mmol) and 2,5-dichlorobenzenesulphonylchloride **4.9** (491 mg, 1.0 mmol). Recrystallization from methanol / water gave the title compound **4.18** (214 mg, 0.60 mmol, 60%) as orange-yellow platelets m.p. 190—191 °C dec.  $R_f$  0.48 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 3.88 min, (100%);  $\delta_H$

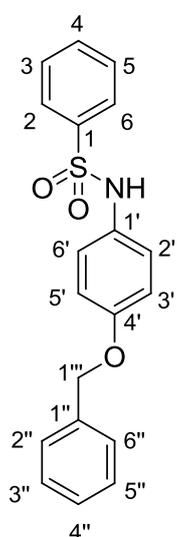
(300 MHz, DMSO-*d*<sub>6</sub>); 9.90 (1H, br s, NH), 7.78-7.77 (1H, m, Ar-H *benzene*), 7.75-7.70 (2H, m, Ar-H *benzene*), 6.95 (1H, app. t,  $J = 7.3$ , Ar-H<sub>3</sub>'), 6.88 (1H, app. d,  $J = 7.3$ , Ar-H<sub>4</sub>'), 6.76 (1H, app. d,  $J = 7.3$ , Ar-H<sub>2</sub>'), 2.61-2.70 (4H, m, 2H<sub>5</sub>' and 2H<sub>8</sub>'), 1.61-1.67 (4H, m, 2H<sub>6</sub>' and 2H<sub>7</sub>'),  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 142.5 (2-C or 5-C), 141.2 (2-C or 5-C), 137.8 (Ar-C<sub>qt</sub>), 133.5 (Ar-C *benzene*), 133.2 (Ar-C *benzene*), 131.7 (Ar-C<sub>qt</sub>), 129.8 (Ar-C *benzene*), 129.6 (Ar-C<sub>qt</sub>), 128.4 (Ar-C *tetrahydronaphthalene*), 126.2 (Ar-C<sub>qt</sub>), 125.3 (Ar-C *tetrahydronaphthalene*), 122.6 (Ar-C *tetrahydronaphthalene*), 29.1 (5'C or 8'-C), 24.5 (5'C or 8'-C), 22.4 (6'-C or 7'-C), 22.1(6'-C or 7'-C);  $\nu_{max}/cm^{-1}$  (neat); 3736, 3325, 1738, 1586, 1450, 1343, 1145, 1043;  $m/z$  (ES) 378.0 (75%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 378.0097. C<sub>16</sub>H<sub>15</sub>Cl<sub>2</sub>NO<sub>2</sub>S requires MNa, 378.0093).

### 2,5-dimethoxy-N-(4-methoxyphenyl)benzenesulfonamide (4.19)



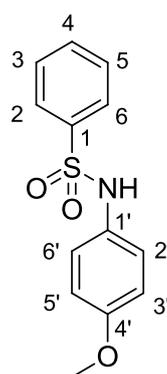
Prepared *via* general method A using *p*-anisidine **4.3** (246.3 mg, 2.0 mmol) and 2,5-dimethoxybenzenesulphonylchloride **4.8** (236.7 mg, 1.0 mmol). Recrystallization from methanol / water gave the title compound **4.19** (236.4 mg, 0.70 mmol, 70%) as light-pink needles m.p. 119—121 °C.  $R_f$  0.44 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 2.69 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 9.63 (1H, s, NH), 7.10-7.17 (3H, m, Ar-H<sub>3,4,6</sub>), 7.02 (2H, d,  $J = 9.0$ , Ar-H<sub>2',6'</sub>), 6.78 (2H, d,  $J = 9.0$ , Ar-H<sub>3',5'</sub>), 3.86 (3H, s, OCH<sub>3</sub> benzenesulfonamide), 3.69 (3H, s, OCH<sub>3</sub> benzenesulfonamide), 3.65 (3H, s, OCH<sub>3</sub> in C<sub>4'</sub>),  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 156.2 (4'-C), 152.0 (2-C or 5-C), 150.2 (2-C or 5-C), 130.2 (Ar-C<sub>qt</sub>), 126.8 (Ar-C<sub>qt</sub>), 122.7 (Ar-C), 119.6 (Ar-C), 115.0 (Ar-C), 114.1 (Ar-C), 113.9 (Ar-C), 56.3 (OCH<sub>3</sub>), 55.6 (OCH<sub>3</sub>), 55.0 (OCH<sub>3</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3249, 2939, 2840, 1888, 1612, 1582, 1434, 1273, 1153, 1016;  $m/z$  (ES) 324.1 (60%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 324.0914. C<sub>15</sub>H<sub>17</sub>NO<sub>5</sub>S requires *MH*, 324.0900).

### N-(4-phenoxyphenyl)benzenesulfonamide (4.20)



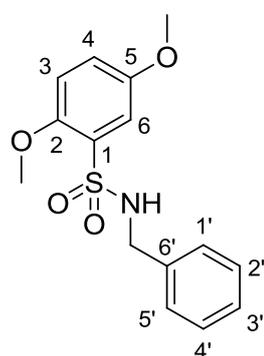
Prepared *via* general method A using 4-benzyloxyanilin hydrochloride **4.4** (472 mg, 2.0 mmol) and benzenesulphonylchloride **4.7** (0.128 mL, 1.0 mmol). Recrystallization from methanol / water gave the title compound **4.20** (273 mg, 0.58 mmol, 58%) as colourless platelets m.p. 98—100 °C.  $R_f$  0.46 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 3.45 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 9.95 (1H, s, NH), 7.68-7.71 (2H, m, Ar-H<sub>2,6</sub>), 7.59-7.63 (1H, m, Ar-H<sub>4</sub>), 7.52-7.56 (2H, m, Ar-H<sub>3,5</sub>), 7.30-7.42 (5H, m, Ar-H<sub>2'',3'',4'',5'',6''</sub>), 6.98 (2H, d,  $J = 9.0$ , Ar-H<sub>2',6'</sub>), 6.89 (2H, d,  $J = 9.0$ , Ar-H<sub>3',5'</sub>), 5.00 (2H, s, OCH<sub>2</sub>),  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 155.6 (4'-C), 139.4 (1-C), 136.9 (1'-C), 132.7 (4-C), 130.2 (1''-C), 129.1 (Ar-C), 128.4 (Ar-C), 127.8 (Ar-C), 126.6 (Ar-C), 123.3 (2'-C, 6'-C), 115.1 (3'-C, 5'-C), 69.3 (OCH<sub>2</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3273, 3063, 2886, 1894, 1610, 1510, 1326, 1157;  $m/z$  (ES) 362.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 362.0840. C<sub>19</sub>H<sub>17</sub>NO<sub>3</sub>S requires *MNa*, 362.0821).

### ***N*-(4-methoxyphenyl)benzenesulfonamide (4.21)<sup>1</sup>**



Prepared *via* general method A using *p*-anisidine **4.3** (246 mg, 2.0 mmol) and benzenesulphonylchloride **4.7** (0.256 mL, 1.0 mmol). Recrystallization from methanol / water gave the title compound **4.21** (189 mg, 0.72 mmol, 72%) as pink needles m.p. 91—92 °C (Lit.<sup>1</sup> 91—92 °C).  $R_f$  0.13 (40:10 petroleum ether—EtOAc); (Found: C, 59.2; H, 4.90; N, 5.2; S 12.2; C<sub>13</sub>H<sub>13</sub>NO<sub>3</sub>S requires C, 59.3; H, 4.98; N, 5.3; S, 12.2%). HPLC (Method A), R.t. 2.64 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 9.91 (1H, s, NH), 7.67-7.70 (2H, m, Ar-H benzenesulphonamide), 7.58-7.63 (1H, m, Ar-H benzenesulphonamide), 7.52-7.56 (2H, m, Ar-H benzenesulphonamide), 6.97 (2H, d,  $J = 9.0$ , Ar-H<sub>2',6'</sub>), 6.80 (2H, d,  $J = 9.0$ , Ar-H<sub>3',5'</sub>), 3.67 (3H, s, OCH<sub>3</sub> in C<sub>4'</sub>),  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 156.4 (4'-C), 139.4 (1-C), 132.7 (Ar-C benzenesulphonamide), 130.0 (1'-C), 129.1 (Ar-C benzenesulphonamide), 126.6 (Ar-C benzenesulphonamide), 123.4 (2'-C, 6'-C), 114.2 (3'-C, 5'-C), 55.1 (OCH<sub>3</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3259, 3004, 1613, 1506, 1466, 1445, 1333, 1289, 1147;  $m/z$  (ES) 286.0 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 286.0510. C<sub>13</sub>H<sub>13</sub>NO<sub>3</sub>S requires MNa, 286.0508).

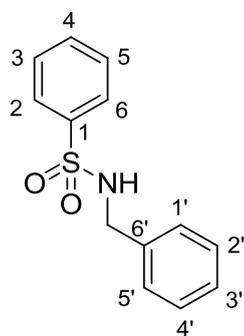
### ***N*-benzyl-2,5-dimethoxybenzenesulfonamide (4.22)**



Prepared *via* general method A using benzylamine **4.5** (0.22 mL, 2.0 mmol) and 2,5-dimethoxybenzenesulphonylchloride **4.8** (236 mg, 1.0 mmol). Recrystallization from methanol / water gave the title compound **4.22** (157 mg, 0.51 mmol, 51%) as light-yellow platelets m.p. 88—90 °C.  $R_f$  0.82 (40:10 petroleum ether—EtOAc); (Found: C, 58.7; H, 5.50; N, 4.4; C<sub>15</sub>H<sub>17</sub>NO<sub>4</sub>S requires C, 58.6; H, 5.57; N, 4.6%). HPLC (Method A), R.t. 2.82 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 8.19 (1H, t,  $J = 6.3$ , NH), 7.16-7.25 (6H, m, Ar-H), 7.12 (1H, dd,  $J = 9.0, 3.1$ , Ar-H<sub>4</sub>), 7.06 (1H, app. d,  $J = 9.0$ , Ar-H<sub>3</sub>), 4.05 (2H, d,  $J = 6.3$ , CH<sub>2</sub>), 3.78 (3H, s, OCH<sub>3</sub>), 3.75 (3H, s, OCH<sub>3</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 152.1 (2-C or 5-C), 150.1 (2-C or 5-C), 137.8 (6'-C), 128.9 (1-C), 127.9 (Ar-C benzyl), 127.4 (Ar-C benzyl), 126.9 (Ar-C benzyl), 119.3 (Ar-C benzenesulphonamide), 114.1 (Ar-C benzenesulphonamide), 114.0 (Ar-C

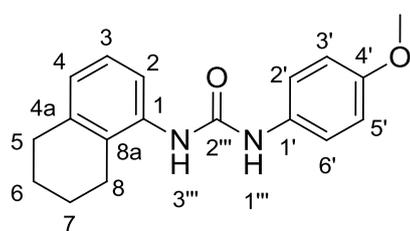
benzenesulphonamide), 56.2 (OCH<sub>3</sub>), 55.7 (OCH<sub>3</sub>), 46.2 (CH<sub>2</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3326, 1969, 1895, 1446, 1324, 1151, 1092, 687;  $m/z$  (ES) 308.1 (70%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 308.0964. C<sub>15</sub>H<sub>17</sub>NO<sub>4</sub>S requires MH, 308.0951).

### ***N*-benzylbenzenesulfonamide (4.23)<sup>1</sup>**



Prepared *via* general method A using benzylamine **4.5** (0.220 mL, 2.0 mmol) and benzenesulphonylchloride **4.7** (0.128 mL, 1.0 mmol). Recrystallization from methanol / water gave the title compound **4.23** (138 mg, 0.56 mmol, 56%) as colourless needles m.p. 87—88 °C (Lit.<sup>1</sup> 87—88 °C)  $R_f$  0.75 (40:10 petroleum ether—EtOAc); (Found: C, 63.2; H, 5.25; N, 5.6; S 13.0; C<sub>13</sub>H<sub>13</sub>NO<sub>2</sub>S requires C, 63.1; H, 5.30; N, 5.7, S 13.0%). HPLC (Method A), R.t. 2.81 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 8.17 (1H, t,  $J$  = 6.3, NH), 7.81-7.83 (2H, m, Ar-H benzenesulphonamide), 7.63-7.66 (3H, m, Ar-H benzenesulphonamide), 7.57-7.60 (5H, m, Ar-H benzyl), 4.00 (2H, d,  $J$  = 6.3, CH<sub>2</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 140.7 (1-C), 137.6 (6'-C), 132.3 (Ar-C benzenesulphonamide), 129.1 (Ar-C benzenesulphonamide), 128.2 (Ar-C benzenesulphonamide), 127.5 (Ar-C benzyl), 127.1 (Ar-C benzyl), 126.4 (Ar-C benzyl), 46.1(OCH<sub>2</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3338, 2930, 1969, 1896, 1447, 1320, 1150;  $m/z$  (ES) 270.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 270.0574. C<sub>13</sub>H<sub>13</sub>NO<sub>2</sub>S requires MNa, 270.0559).

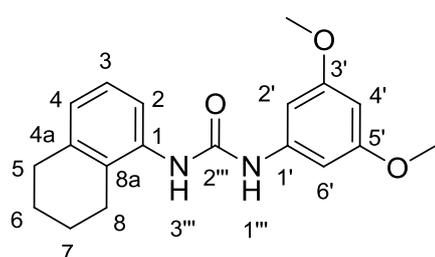
### **1-(4-methoxyphenyl)-3-(5,6,7,8-tetrahydronaphthalen-1-yl)urea (4.24)**



Prepared *via* general method B using 5,6,7,8-tetrahydronaphthylamine **4.2** (0.240 mL, 1.68 mmol) and 4-methoxyphenylisocyanate **4.11** (0.240 mL, 1.80 mmol). Recrystallization from dichloromethane gave the title compound **4.24** (408 mg, 1.38 mmol, 82%) as colourless needles m.p. 213—214 °C.  $R_f$  0.52 (40:10 petroleum ether—EtOAc); (Found: C, 73.0; H, 6.70; N, 9.4; C<sub>18</sub>H<sub>20</sub>N<sub>2</sub>O<sub>2</sub> requires C, 73.0; H, 6.80; N, 9.4%). HPLC (Method A), R.t. 3.21 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 8.83 (1H, s, NH<sub>1'''</sub>), 7.69 (1H, s, NH<sub>3'''</sub>), 7.66 (1H, app. d,  $J$  = 7.8, Ar-H<sub>2</sub>), 7.37 (2H, d,  $J$  = 9.0, Ar-H<sub>2',6'</sub>), 7.03 (1H, app. t,  $J$  = 7.8, Ar-H<sub>3</sub>),

6.88 (2H, d,  $J = 9.0$ , Ar- $H_{3',5'}$ ), 6.76 (1H, app. d,  $J = 7.8$ , Ar- $H_4$ ), 3.87 (3H, s, OCH<sub>3</sub>), 2.72 (2H, t,  $J = 6.2$ , H<sub>8</sub>), 2.58 (2H, t,  $J = 6.2$ , H<sub>5</sub>), 1.78-1.83 (2H, m, H<sub>6</sub>), 1.69-1.73 (2H, m, H<sub>7</sub>);  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 154.3 (4'-C), 152.8 (C=O), 137.1 (4a-C or 1-C), 137.0 (4a-C or 1-C), 133.0 (1'-C), 126.5 (8a-C), 125.1 (Ar-C tetrahydronaphthalene), 123.4 (Ar-C tetrahydronaphthalene), 119.6 (Ar-C), 118.4 (Ar-C), 114.0 (3'-C, 5'-C), 55.1 (OCH<sub>3</sub>), 29.4 (5-C), 24.2 (8-C), 22.4 (6-C or 7-C), 22.1 (6-C or 7-C);  $\nu_{max}/cm^{-1}$  (neat); 3854, 3326, 3272, 1873, 1644, 1560, 1509, 1246;  $m/z$  (ES) 319.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 319.1428. C<sub>18</sub>H<sub>20</sub>N<sub>2</sub>O<sub>2</sub> requires MNa, 319.1417).

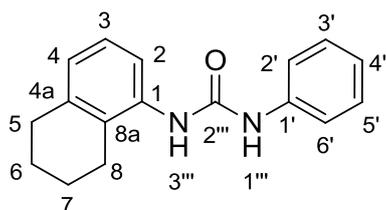
### 1-(3,5-dimethoxyphenyl)-3-(5,6,7,8-tetrahydronaphthalen-1-yl)urea (4.25)



Prepared *via* general method B using 5,6,7,8-tetrahydronaphthylamine **4.2** (0.240 mL, 1.68 mmol) and 3,5-dimethoxyphenylisocyanate **4.12** (331 mg, 1.85 mmol). Recrystallization from dichloromethane gave the title compound

**4.25** (483 mg, 1.51 mmol, 90%) as colourless needles m.p. 192—193 °C.  $R_f$  0.72 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 3.38 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 9.01 (1H, s, NH<sub>1'''</sub>), 7.75 (1H, s, NH<sub>3'''</sub>), 7.64 (1H, app. d,  $J = 7.8$ , Ar- $H_2$ ), 7.04 (1H, app. t,  $J = 7.8$ , Ar- $H_3$ ), 6.78 (1H, app. d,  $J = 7.8$ , Ar- $H_4$ ), 6.69 (2H, d,  $J = 2.2$ , Ar- $H_{2',6'}$ ), 6.14 (1H, t,  $J = 2.2$ , Ar- $H_{4'}$ ), 3.73 (6H, s, OCH<sub>3</sub> in 3' and 5'), 2.73 (2H, t,  $J = 6.2$ , H<sub>8</sub>), 2.57 (2H, t,  $J = 6.2$ , H<sub>5</sub>), 1.67-1.84 (4H, m, H<sub>6,7</sub>);  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 160.6 (3'-C, 5'-C), 152.5 (C=O), 141.6 (4a-C or 1-C), 137.1 (4a-C or 1-C), 136.8 (1'-C), 126.8 (8a-C), 125.1 (Ar-C tetrahydronaphthalene), 123.7 (Ar-C tetrahydronaphthalene), 118.7 (Ar-C tetrahydronaphthalene), 96.2 (2'-C, 6'-C), 93.8 (4'-C), 55.0 (OCH<sub>3</sub> in 3'-C and 5'-C), 29.4 (5-C), 24.2 (8-C), 22.4 (6-C or 7-C), 22.3 (6-C or 7-C);  $\nu_{max}/cm^{-1}$  (neat); 3852, 3736, 3350, 3294, 1744, 1650, 1605, 1564, 1208, 1151;  $m/z$  (ES) 349.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 349.1539. C<sub>19</sub>H<sub>22</sub>N<sub>2</sub>O<sub>3</sub> requires MNa, 349.1523).

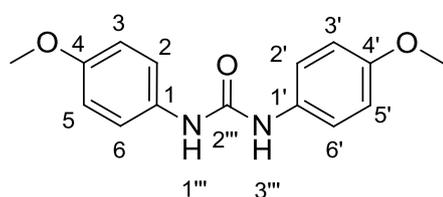
### 1-phenyl-3-(5,6,7,8-tetrahydronaphthalen-1-yl)urea (4.26)



Prepared *via* general method B using 5,6,7,8-tetrahydronaphthylamine **4.2** (0.240 mL, 1.68 mmol) and phenylisocyanate **4.13** (0.220 mL, 1.85 mmol). Recrystallization from dichloromethane gave the title compound **4.26** (362 mg, 1.36 mmol,

81%) as colourless needles m.p. 194—195 °C.  $R_f$  0.75 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 3.28 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 9.03 (1H, s, NH $_{1''''}$ ), 7.79 (1H, s, NH $_{3''''}$ ), 7.66 (1H, app. d,  $J = 7.4$ , Ar-H $_2$ ), 7.46 (2H, app. d,  $J = 7.9$ , Ar-H $_{2',6'}$ ), 7.28 (2H, app. t,  $J = 7.9$ , Ar-H $_{3',5'}$ ), 7.03 (1H, app. t,  $J = 7.4$ , Ar-H $_3$ ), 6.95 (1H, app. t,  $J = 7.9$ , Ar-H $_4$ ), 6.77 (1H, app. d,  $J = 7.4$ , Ar-H $_4$ ), 2.72 (2H, t,  $J = 6.0$ , H $_5$ ), 2.58 (2H, t,  $J = 6.0$ , H $_8$ ), 1.77-1.83 (2H, m, H $_6$ ), 1.71-1.75 (2H, m, H $_7$ );  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 152.6 (C=O), 139.9 (1'-C), 137.1 (4a-C or 1-C), 136.9 (4a-C or 1-C), 128.8 (3'-C, 5'-C), 126.7 (Ar-C $_{qt}$ ), 125.1 (2'-C, 6'-C), 123.6 (Ar-C *tetrahydronaphthalene*), 121.6 (Ar-C *tetrahydronaphthalene*), 118.6 (4'-C), 117.9 (2-C), 29.4 (5-C), 24.2 (8-C), 22.4 (6-C or 7-C), 22.1 (6-C or 7-C);  $\nu_{max}/cm^{-1}$  (neat); 3309, 2935, 1638, 1560, 1310, 1244;  $m/z$  (ES) 289.1 (100%, MNa $^+$ ); (Found MNa $^+$ , 289.1315. C $_{17}$ H $_{18}$ N $_2$ O requires MNa, 289.1311).

### 1,3-bis(4-methoxyphenyl)urea (4.27)<sup>2</sup>

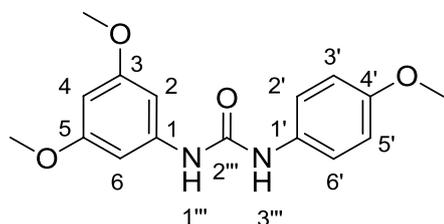


Prepared *via* general method B using *p*-anisidine **4.3** (246 mg, 2.0 mmol) and 4-methoxyphenylisocyanate **4.11** (0.290 mL, 2.2 mmol). Recrystallization from dichloromethane

gave the title compound **4.27** (551 mg, 1.84 mmol, 92%) as colourless needles m.p. 237—238 °C (Lit.<sup>2</sup> 238 °C).  $R_f$  0.38 (40:10 petroleum ether—EtOAc); (Found: C, 66.2; H, 5.92; N, 10.3; C $_{15}$ H $_{16}$ N $_2$ O $_3$  requires C, 66.0; H, 5.90; N, 10.4%). HPLC (Method A), R.t. 2.47 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 8.36 (2H, s, NH $_{1''''}$ , $_{3''''}$ ), 7.35 (4H, d,  $J = 9.0$ , Ar-H $_{2,2',6,6'}$ ), 6.86 (4H, d,  $J = 9.0$ , Ar-H $_{3,3',5,5'}$ ), 3.72 (6H, s, OCH $_3$ );  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 154.3 (4-C, 4'-C), 152.9 (C=O), 132.9 (1-C, 1'-C), 119.8 (2-C, 2'-C, 6-C, 6'-C), 113.9 (3-C, 3'-C, 5-C, 5'-C), 55.1 (OCH $_3$ );  $\nu_{max}/cm^{-1}$  (neat); 3325, 3289, 2959, 2837, 1876, 1643, 1609,

1247, 825, 652, 527;  $m/z$  (ES) 295.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 295.1055.  $C_{15}H_{16}N_2O_3$  requires  $MNa$ , 295.1053).

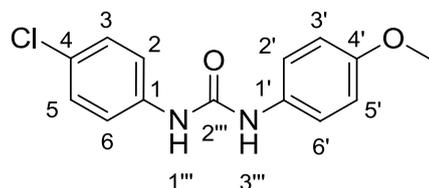
### 1-(3,5-dimethoxyphenyl)-3-(4-methoxyphenyl)urea (4.28)



Prepared *via* general method B using *p*-anisidine **4.3** (246 mg, 2.0 mmol) and 3,5-dimethoxyphenylisocyanate **4.12** (394 mg, 2.2 mmol). Recrystallization from dichloromethane gave the title compound **4.28** (568 mg, 1.9

mmol, 94%) as colourless needles m.p. 174—175 °C.  $R_f$  0.46 (40:10 petroleum ether—EtOAc); (Found: C, 63.2; H, 5.95; N, 9.0;  $C_{16}H_{18}N_2O_4$  requires C, 63.5; H, 6.00; N, 9.2%). HPLC (Method A), R.t. 2.69 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 8.57 (1H, s,  $NH_{3'''}$ ), 8.42 (1H, s,  $NH_{1'''}$ ), 7.35 (2H, d,  $J = 9.0$ , Ar- $H_{2',6'}$ ), 6.87 (2H, d,  $J = 9.0$ , Ar- $H_{3',5'}$ ), 6.67 (2H, d,  $J = 2.2$ , Ar- $H_{2,6}$ ), 6.13 (1H, t,  $J = 2.2$ , Ar- $H_4$ ), 3.73 (3H, s,  $OCH_3$  in 4'), 3.72 (6H, s,  $OCH_3$  in 3 and 5);  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 160.6 (3-C, 5-C), 154.5 (4'-C), 152.6 (C=O), 141.6 (1-C), 132.5 (1'-C), 120.1 (2'-C, 6'-C), 114.0 (3'-C, 5'-C), 96.3 (2-C, 6-C), 93.7 (4-C), 55.1 ( $OCH_3$ ), 55.0 ( $OCH_3$ );  $\nu_{max}/cm^{-1}$  (neat); 3306, 3255, 1634, 1601, 1567, 1208, 815;  $m/z$  (ES) 325.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 325.1169.  $C_{16}H_{18}N_2O_4$  requires  $MNa$ , 325.1159).

### 1-(4-chlorophenyl)-3-(4-methoxyphenyl)urea (4.29)<sup>3</sup>

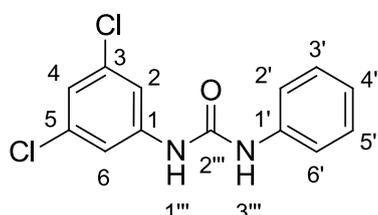


Prepared *via* general method B using *p*-anisidine **4.3** (123 mg, 1.0 mmol) and 4-chlorophenylisocyanate **4.14** (194 mg, 1.1 mmol). Recrystallization from dichloromethane

gave the title compound **4.29** (180 mg, 0.65 mmol, 65%) as colourless needles m.p. 254—255 °C (Lit.<sup>3</sup> 250—255 °C).  $R_f$  0.28 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 2.99 min, (100%);  $\delta_H$  (300 MHz, DMSO- $d_6$ ); 8.74 (1H, s,  $NH_{1'''}$ ), 8.52 (1H, s,  $NH_{3'''}$ ), 7.47 (2H, d,  $J = 8.9$ , Ar- $H_{2,6}$ ), 7.34 (2H, d,  $J = 9.0$ ,  $H_{2',6'}$ ), 7.31 (2H, d,  $J = 8.9$ , Ar- $H_{3,5}$ ), 6.87 (2H, d,  $J = 9.0$ , Ar- $H_{3',5'}$ ), 3.71 (3H, s,  $OCH_3$ );  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 154.5 (C=O), 152.6 (4'-C), 138.9 (1-C), 132.5 (4-C), 128.5 (2-C, 6-C), 125.0 (1'-C), 120.1 (3-C, 5-C), 119.5 (3'-C, 5'-C), 113.9

(2'-C, 6'-C), 55.1 (OCH<sub>3</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3302, 1637, 1560, 1509, 1244, 828;  $m/z$  (ES) 277.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 277.0725. C<sub>14</sub>H<sub>13</sub>ClN<sub>2</sub>O<sub>2</sub> requires MH, 277.0738).

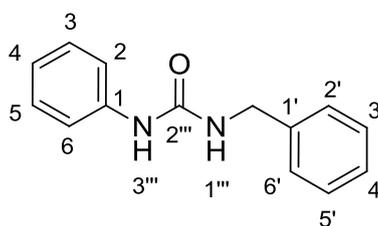
### 1-(3,5-dichlorophenyl)-3-phenylurea (4.30)



Prepared *via* general method B using 3,5-dichloroaniline **4.10** (162 mg, 1.0 mmol) and phenylisocyanate **4.13** (0.240 mL, 1.1 mmol). Recrystallization from dichloromethane gave the title compound **4.30** (143 mg, 0.51 mmol, 51%) as

colourless needles m.p. 200—201 °C.  $R_f$  0.56 (40:10 petroleum ether—EtOAc); (Found: C, 55.6; H, 3.50; N, 9.9; C<sub>13</sub>H<sub>10</sub>Cl<sub>2</sub>N<sub>2</sub>O requires C, 55.5; H, 3.59; N, 10.0%). HPLC (Method A), R.t. 3.63 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 9.04 (1H, s, NH<sub>1'''</sub>), 8.87 (1H, s, NH<sub>3'''</sub>), 7.54 (2H, d,  $J = 1.9$ , Ar-H<sub>2,6</sub>), 7.46 (2H, app. d,  $J = 8.0$ , Ar-H<sub>2',6'</sub>), 7.31 (2H, app. t,  $J = 8.0$ , Ar-H<sub>3',5'</sub>), 7.17 (1H, t,  $J = 1.9$ , Ar-H<sub>4</sub>), 7.02 (1H, app. t,  $J = 8.0$ , Ar-H<sub>4'</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 152.2 (C=O), 142.3 (1-C), 139.1 (1'-C), 134.1 (3-C, 5-C), 128.8 (3'-C, 5'-C), 122.3 (4-C), 120.8 (4'-C), 118.6 (2'-C, 6'-C), 116.3 (2-C, 6-C);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3325, 3183, 1650, 1606, 1588, 1544, 1448, 1310, 1216;  $m/z$  (ES) 303.0 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 303.0048. C<sub>13</sub>H<sub>10</sub>Cl<sub>2</sub>N<sub>2</sub>O requires MNa, 303.0062).

### 1-benzyl-3-phenylurea (4.31)<sup>4</sup>

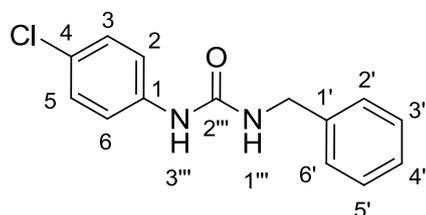


Prepared *via* general method B using benzylamine **4.5** (0.22 mL, 2.0 mmol) and phenylisocyanate **4.13** (0.24 mL, 2.2 mmol). Recrystallization from dichloromethane gave the title compound **4.31** (385 mg, 1.7 mmol, 85%) as colourless needles

m.p. 173—175 °C (Lit.<sup>4</sup> 173—175 °C).  $R_f$  0.20 (40:10 petroleum ether—EtOAc); (Found: C, 73.8; H, 6.25; N, 12.6; C<sub>14</sub>H<sub>14</sub>N<sub>2</sub>O requires C, 74.3; H, 6.24; N, 12.4%). HPLC (Method A), R.t. 2.52 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 8.55 (1H, s, NH<sub>3'''</sub>), 7.21-7.42 (9H, m, Ar-H), 6.89-6.92 (1H, m, Ar-H), 6.62 (1H, t,  $J = 5.9$ , NH<sub>1'''</sub>), 4.32 (2H, d,  $J = 5.9$ , CH<sub>2</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 155.2 (C=O), 140.4 (Ar-C<sub>qt</sub>), 140.3 (Ar-C<sub>qt</sub>), 128.6 (Ar-C), 128.3 (Ar-C), 127.1 (Ar-C),

126.7 (Ar-C), 121.0 (Ar-C), 117.6 (Ar-C), 42.7 (CH<sub>2</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3307, 3030, 1952, 1686, 1599, 1556, 1309, 696;  $m/z$  (ES) 249.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 249.0990. C<sub>14</sub>H<sub>14</sub>N<sub>2</sub>O requires MNa, 249.0998).

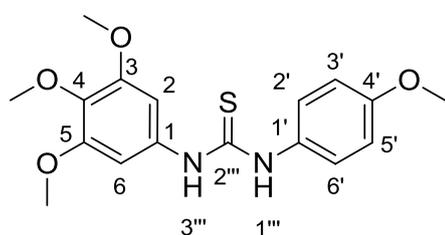
### 1-benzyl-3-(4-chlorophenyl)urea (**4.32**)<sup>5</sup>



Prepared *via* general method B using benzylamine **4.5** (0.22 mL, 2.0 mmol) and 4-chlorophenylisocyanate **4.14** (338 mg, 2.2 mmol). Recrystallization from dichloromethane gave the title compound **4.32** (360 mg, 1.7

mmol, 69%) as colourless needles m.p. 200—201 °C (Lit.<sup>5</sup> 201 °C).  $R_f$  0.31 (40:10 petroleum ether—EtOAc); (Found: C, 64.2; H, 4.95; N, 10.8; C<sub>14</sub>H<sub>13</sub>ClN<sub>2</sub>O requires C, 64.5; H, 5.03; N, 10.7%). HPLC (Method A), R.t. 2.97 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 8.71 (1H, s, NH<sub>3'''</sub>), 7.44-7.47 (2H, m, Ar-H), 7.24-7.36 (7H, m, Ar-H), 6.67 (1H, t,  $J = 5.9$ , NH<sub>1'''</sub>), 4.31 (2H, d,  $J = 5.9$ , CH<sub>2</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 155.0 (C=O), 140.2 (Ar-C<sub>qt</sub>), 139.4 (Ar-C<sub>qt</sub>), 128.4 (Ar-C), 128.3 (Ar-C), 127.1 (Ar-C), 126.7 (Ar-C), 124.5 (Ar-C<sub>qt</sub>), 119.1 (Ar-C), 42.7 (CH<sub>2</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3289, 1629, 1594, 1567, 1397, 1241, 810;  $m/z$  (ES) 261.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 261.0790. C<sub>14</sub>H<sub>13</sub>ClN<sub>2</sub>O requires MH, 261.0789).

### 1-(4-methoxyphenyl)-3-(3,4,5-trimethoxyphenyl)thiourea (**4.33**)

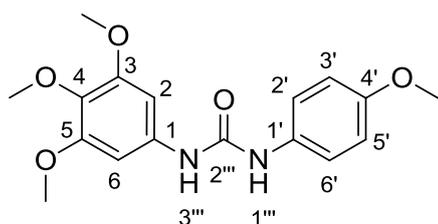


A solution of *p*-anisidine **4.3** (62 mg, 0.50 mmol) in dichloromethane (4 mL) was added dropwise to a solution of 3,4,5-trimethoxyphenylisothiocyanate **4.15b** (124 mg, 0.55 mmol) in dichloromethane (4 mL) at

room temperature and stirred for 7 hours. The resulting yellow-white slurry was concentrated *in vacuo* and recrystallized from ether to give the title compound **4.33** (119 mg, 0.34 mmol, 62%) as white-yellow platelets m.p. 100—102 °C.  $R_f$  0.40 (40:10 petroleum ether—EtOAc); HPLC (Method A), R.t. 2.41 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 9.53 (1H, s, NH<sub>1'''</sub>), 9.51 (1H, s, NH<sub>3'''</sub>), 7.29 (2H, d,  $J = 8.9$ , Ar-H<sub>2',6'</sub>), 6.90 (2H, d,  $J = 8.9$ , Ar-H<sub>3',5'</sub>), 6.83 (2H, s, Ar-H<sub>2,6</sub>), 3.76 (3H, s,

OCH<sub>3</sub> in 4), 3.75 (6H, s, OCH<sub>3</sub> in 3 and 5), 3.65 (3H, s, OCH<sub>3</sub> in 4');  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 179.6 (C=O), 156.6 (3-C, 5-C), 152.4 (4-C), 135.0 (1-C), 134.5 (1'-C), 132.1 (4'-C), 126.2 (2'-C, 6'-C), 113.6 (3'-C, 5'-C), 101.7 (2-C, 6-C), 60.0 (OCH<sub>3</sub>), 55.8 (OCH<sub>3</sub>), 55.2 (OCH<sub>3</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3276, 2995, 1598, 1505, 1234;  $m/z$  (ES) 371.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 371.1051. C<sub>17</sub>H<sub>20</sub>N<sub>2</sub>O<sub>4</sub>S requires MNa, 371.1036).

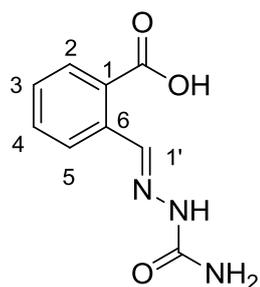
### 1-(4-methoxyphenyl)-3-(3,4,5-trimethoxyphenyl)urea (4.34)



Prepared *via* general method B using *p*-anisidine **4.3** (123 mg, 1.0 mmol) and 3,4,5-trimethoxyphenylisocyanate **4.15a** (230 mg, 1.1 mmol). Recrystallization from dichloromethane gave the title compound **4.34**

(286 mg, 0.86 mmol, 86%) as colourless needles m.p. 196—197 °C.  $R_f$  0.31 (40:10 petroleum ether—EtOAc); (Found: C, 61.0; H, 5.95; N, 8.2; C<sub>17</sub>H<sub>20</sub>N<sub>2</sub>O<sub>5</sub> requires C, 61.4; H, 6.07; N, 8.4%). HPLC (Method A), R.t. 2.33 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 8.53 (1H, s, NH<sub>1'''</sub>), 8.40 (1H, s, NH<sub>3'''</sub>), 7.34 (2H, d,  $J = 9.0$ , Ar-H<sub>2',6'</sub>), 6.86 (2H, d,  $J = 9.0$ , Ar-H<sub>3',5'</sub>), 6.78 (2H, s, Ar-H<sub>2,6</sub>), 3.74 (6H, s, OCH<sub>3</sub> in 3 and 5), 3.71 (3H, s, OCH<sub>3</sub> in 4), 3.60 (3H, s, OCH<sub>3</sub> in 4');  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 154.4 (Ar-C<sub>qt</sub>), 152.8 (Ar-C<sub>qt</sub>), 152.7 (Ar-C<sub>qt</sub>), 136.0 (Ar-C<sub>qt</sub>), 132.6 (Ar-C<sub>qt</sub>), 132.2 (Ar-C<sub>qt</sub>), 120.1 (2'-C, 6'-C), 113.9 (3'-C, 5'-C), 95.7 (2-C, 6-C), 60.1 (OCH<sub>3</sub> in 3 and 5), 55.6 (OCH<sub>3</sub> in 4), 55.1 (OCH<sub>3</sub> in 4');  $\nu_{max}/cm^{-1}$  (neat); 3271, 2952, 1626, 1614, 1573, 1511, 1411, 1230, 1132;  $m/z$  (ES) 355.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 355.1256. C<sub>17</sub>H<sub>20</sub>N<sub>2</sub>O<sub>5</sub> requires MNa, 355.1264).

### (E)-2-((2-carbamoylhydrazono)methyl)benzoic acid (5.15)<sup>6</sup>

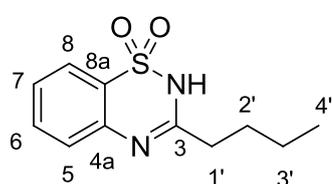


A solution of semicarbazide hydrochloride **5.31** (223 mg, 2.0 mmol) and sodium acetate (164 mg, 2.0 mmol) in water (20 mL) was stirred at room temperature for 30 minutes. 2-carboxybenzaldehyde **5.30** (300 mg, 2.0 mmol) was added to this solution and stirred for 12 hours. The resulting colourless slurry was filtered and the colourless solid

washed with water (50 mL) and petroleum ether (50 mL) and recrystallized from

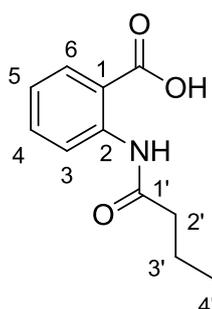
methanol / water to give the title compound 5.15 (381 mg, 1.8 mmol, 92%) as colourless plates m.p. 200 °C dec. (Lit.<sup>6</sup> 202 °C).  $R_f$  0.13 (1:1 petroleum ether—EtOAc); HPLC (Method B), R.t. 1.39 min, (100%);  $\delta_H$  (500 MHz, DMSO- $d_6$ ); 13.24 (1H, br s, COOH), 10.43 (1H, s, N-H), 8.57 (1H, s, C-H<sub>1'</sub>), 8.18 (1H, app. d,  $J = 7.9$ , Ar-H<sub>5</sub>), 7.81 (1H, dd,  $J = 7.9, 1.0$ , Ar-H<sub>2</sub>), 7.54 (1H, app. td,  $J = 7.9, 1.0$ , Ar-H<sub>4</sub>), 7.43 (1H, app. td,  $J = 7.9, 1.0$ , Ar-H<sub>3</sub>), 6.50 (2H, br s, NH<sub>2</sub>);  $\delta_C$  (125 MHz, DMSO- $d_6$ ); 168.3 (COOH), 156.7 (C=O), 138.1 (1'-C), 134.7 (6-C), 131.5 (4-C), 130.0 (1-C), 129.9 (2-C), 128.5 (3-C), 126.5 (5-C);  $\nu_{max}/cm^{-1}$  (neat); 3477, 2505, 1704, 1566, 1482, 1429, 1283;  $m/z$  (E-H<sup>-</sup>) 206.1 (100%, M-H<sup>-</sup>); (Found M-H<sup>-</sup>, 206.0563. C<sub>9</sub>H<sub>9</sub>N<sub>3</sub>O<sub>3</sub> requires M-H<sup>-</sup>, 206.0571).

### 3-butyl-2H-benzo[e][1,2,4]thiadiazine 1,1-dioxide (5.16)<sup>7</sup>



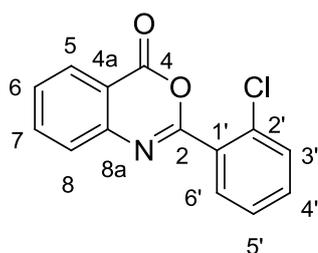
Sodium hydrogen sulphite (784 mg, 7.5 mmol) was added to a solution of 2-aminobenzenesulfonamide **5.34** (861 mg, 5.0 mmol) and pentanal **5.35** (0.520 mL, 5.0 mmol) in dimethylacetamide (10 mL). This mixture was heated with stirring at 170 °C for 2 hours and then poured into water (40 mL). The light brown precipitate formed was filtered off and washed with water (100 mL) and recrystallized from petroleum ether / EtOAc to give the title compound 5.16 (836 mg, 3.5 mmol, 70%) as light brown needles, m.p. 148—149 °C (Lit.<sup>7</sup> 150—151 °C);  $R_f$  0.31 (1:1 petroleum ether—EtOAc); HPLC (Method A), R.t. 1.74 min, (100%);  $\delta_H$  (500 MHz, DMSO- $d_6$ ); 11.93 (1H, s, N-H), 7.77 (1H, app. d,  $J = 8.0$ , Ar-H<sub>8</sub>), 7.65 (1H, app. t,  $J = 8.0$ , Ar-H<sub>7</sub>), 7.42 (1H, app. t,  $J = 8$ , Ar-H<sub>6</sub>), 7.31 (1H, app. d,  $J = 8.0$ , Ar-H<sub>5</sub>), 2.53 (2H, t,  $J = 7.5$ , H<sub>1'</sub>), 1.65 (2H, quint,  $J = 7.5$ , H<sub>2'</sub>), 1.35 (2H, sext,  $J = 7.5$ , H<sub>3'</sub>), 0.90 (3H, t,  $J = 7.5$ , H<sub>4'</sub>);  $\delta_C$  (75 MHz, DMSO- $d_6$ ); 160.3 (3-C), 135.1 (4a-C), 133.0 (7-C), 126.2 (6-C), 123.4 (8-C), 121.1 (8a-C), 117.3 (5-C), 34.9 (1'-C), 28.1 (2'-C), 21.4 (3'-C), 13.57 (4'-C);  $\nu_{max}/cm^{-1}$  (neat); 3185, 2958, 1941, 1824, 1616, 1479, 1274;  $m/z$  (ES) 239.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 239.0848. C<sub>11</sub>H<sub>14</sub>N<sub>2</sub>O<sub>2</sub>S requires MH<sup>+</sup>, 239.0849).

## 2-butyramidobenzoic acid (**5.17**)<sup>8</sup>



Butanoyl chloride **5.33** (0.42 mL, 4.0 mmol) was added dropwise to a stirred solution of anthranilic acid **5.32** (274 mg, 2.0 mmol) in pyridine (10 mL) at room temperature. This solution was stirred at room temperature for 12h. The reaction mixture was poured into EtOAc (10 mL) and the organic layer washed successively with 2M HCl (3 x 50 mL), water (50 mL) and brine (50 mL). The resulting solution was dried (MgSO<sub>4</sub>) and the solvent removed *in vacuo* and the colourless solid was recrystallised from MeOH / water to give the title compound **5.17** (373 mg, 1.8 mmol, 89%) as colourless plates m.p. 119—120 °C (Lit.<sup>8</sup> 117—118 °C); *R<sub>f</sub>* 0.08 (1:1 petroleum ether—EtOAc); (Found: C, 63.8; H, 6.25; N, 6.8; C<sub>11</sub>H<sub>13</sub>NO<sub>3</sub> requires C, 63.8; H, 6.32; N, 6.8%). HPLC (Method A), R.t. 2.24 min, (100%); δ<sub>H</sub> (300 MHz, DMSO-*d*<sub>6</sub>); 13.62 (1H, br s, COOH), 11.13 (1H, s, N-H), 8.50 (1H, app. d, *J* = 9.0, Ar-H<sub>6</sub>), 7.97 (1H, app. d, *J* = 9.0, Ar-H<sub>3</sub>), 7.58 (1H, app. t, *J* = 9.0, Ar-H<sub>4</sub>), 7.14 (1H, app. t, *J* = 9.0, Ar-H<sub>5</sub>), 2.36 (2H, t, *J* = 7.5, H<sub>2'</sub>), 1.64 (2H, sext, *J* = 7.5, H<sub>3'</sub>), 0.93 (3H, t, *J* = 7.5, H<sub>4'</sub>); δ<sub>C</sub> (75 MHz, DMSO-*d*<sub>6</sub>); 174.1 (1'-C), 169.5 (COOH), 140.9 (2-C), 134.0 (4-C), 131.0 (3-C), 122.5 (5-C), 119.8 (6-C), 116.2 (1-C), 39.4 (2'-C), 18.3 (3'-C), 13.5 (4'-C); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3171, 2961, 1947, 1842, 1682, 1606, 1537, 1414; *m/z* (EI) 207.1 (100%, M); (Found M, 207.0898. C<sub>11</sub>H<sub>13</sub>NO<sub>3</sub> requires M, 207.0895).

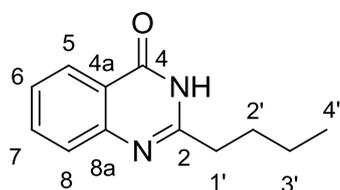
## 2-(2-chlorophenyl)-4H-benzo[d][1,3]oxazin-4-one (**5.18**)<sup>9</sup>



2-chlorobenzoyl chloride **5.36** (0.253 mL, 2.0 mmol) was added dropwise to a stirred solution of anthranilic acid **5.32** (274 mg, 2.0 mmol) in pyridine (10 mL) at room temperature. This solution was stirred at room temperature for 12h. The reaction mixture was poured into water (20 mL) and the colourless precipitate was filtered off. This colourless solid was washed with water (100 mL) and recrystallized from ethanol to give the title compound **5.18** (168 mg, 1.3 mmol, 65%) as colourless plates m.p. 137—138 °C (Lit.<sup>9</sup> 138 °C). *R<sub>f</sub>* 0.80 (1:1 petroleum Ether—EtOAc); (Found: C,

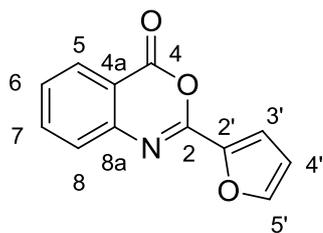
65.4; H, 3.10; N, 5.3; Cl, 13.6; C<sub>14</sub>H<sub>8</sub>ClNO<sub>2</sub> requires C, 65.3; H, 3.13; Cl, 13.8; N, 5.4%). HPLC (Method A), R.t. 3.41 min, (100%);  $\delta_{\text{H}}$  (500 MHz, CDCl<sub>3</sub>); 8.28 (1H, dd,  $J = 7.8, 1.6$ , Ar-H<sub>5</sub>), 7.91 (1H, dd,  $J = 7.8, 1.6$ , Ar-H<sub>8</sub>), 7.86 (1H, app. t,  $J = 7.8$ , Ar-H<sub>6</sub> or Ar-H<sub>7</sub>), 7.73 (1H, app. d,  $J = 8.1$ , Ar-H<sub>6'</sub>), 7.59 (1H, app. t,  $J = 7.8$ , Ar-H<sub>6</sub> or Ar-H<sub>7</sub>), 7.54 (1H, dd,  $J = 8.0, 1.3$ , Ar-H<sub>3'</sub>), 7.47 (1H, app. t,  $J = 8.0$ , Ar-H<sub>4'</sub>), 7.41 (1H, app. t,  $J = 8.0$ , Ar-H<sub>5'</sub>);  $\delta_{\text{C}}$  (125 MHz, CDCl<sub>3</sub>); 158.7 (4-C), 155.8 (2-C), 145.8 (Ar-C), 137.0 (Ar-C), 132.9 (Ar-C), 131.8 (2'-C), 131.7 (Ar-C), 130.6 (Ar-C), 130.2 (8a-C), 129.4 (Ar-C), 128.0 (Ar-C), 127.5 (Ar-C), 127.1 (Ar-C), 116.8 (4a-C);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 1767, 1625, 1605, 1475, 1316;  $m/z$  (ES) 280.0 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 280.0126. C<sub>14</sub>H<sub>8</sub>ClNO<sub>2</sub> requires MNa, 280.0136).

### 2-butylquinazolin-4(3H)-one (5.19)<sup>10</sup>



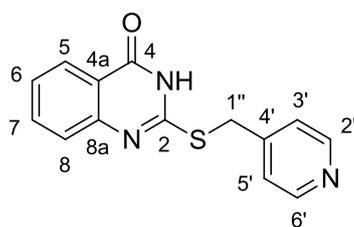
Sodium hydrogen sulphite (784 mg, 7.5 mmol) was added to a solution of anthranilamide **5.38** (681 mg, 5.0 mmol) and pentanal **5.35** (0.520 mL, 5.0 mmol) in dimethylacetamide (10 mL). This mixture was heated with stirring at 150 °C for 4 hours and then poured into water (40 mL). The light brown precipitate formed was filtered off and washed with water (100 mL) and recrystallized from petroleum ether / EtOAc to give the title compound **5.19** (158 mg, 3.9 mmol, 78%) as light brown needles m.p. 159—160 °C (Lit.<sup>10</sup> 159—160 °C).  $R_f$  0.40 (1:1 petroleum ether—EtOAc); HPLC (Method A), R.t. 1.62 min, (100%);  $\delta_{\text{H}}$  (300 MHz, DMSO-*d*<sub>6</sub>); 12.19 (1H, br s, NH), 8.09 (1H, dd,  $J = 7.8, 1.4$ , Ar-H<sub>5</sub>), 7.77 (1H, app. t,  $J = 7.8$ , Ar-H<sub>7</sub>), 7.59 (1H, app. d,  $J = 7.8$ , Ar-H<sub>8</sub>), 7.45 (1H, app. t,  $J = 7.8$ , Ar-H<sub>6</sub>), 2.61 (2H, t,  $J = 7.7$ , H<sub>1'</sub>), 1.72 (2H, quint,  $J = 7.7$ , H<sub>2'</sub>), 1.36 (2H, sext,  $J = 7.7$ , H<sub>3'</sub>), 0.96 (3H, t,  $J = 7.7$ , H<sub>4'</sub>);  $\delta_{\text{C}}$  (75 MHz, DMSO-*d*<sub>6</sub>); 161.8 (4-C), 157.5 (2-C), 148.9 (8a-C), 134.2 (7-C), 126.7 (5-C), 125.9 (6-C), 125.6 (8-C), 120.7 (4a-C), 34.1 (1'-C), 28.9 (2'-C), 21.7 (3'-C), 13.7 (4'-C);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 2925, 1845, 1679, 1620, 1564, 1503, 1468;  $m/z$  (ES) 225.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 225.0988. C<sub>12</sub>H<sub>14</sub>N<sub>2</sub>O requires MNa, 225.0998).

## 2-(furan-2-yl)-4H-benzo[d][1,3]oxazin-4-one (5.20)<sup>9</sup>



Furan-2-carbonyl chloride **5.37** (0.197 mL, 2.0 mmol) was added dropwise to a stirred solution of anthranilic acid **5.32** (274 mg, 2.0 mmol) in pyridine (10 mL) at room temperature. This solution was stirred at room temperature for 12h. The reaction mixture was poured into water (20 mL) and the colourless precipitate was filtered off. This colourless solid was washed with water (100 mL) and recrystallized from ethanol to give the title compound **5.20** (192 mg, 0.9 mmol, 45%) as colourless plates m.p. 101—102 °C (Lit.<sup>9</sup> 102 °C).  $R_f$  0.78 (1:1 petroleum Ether—EtOAc); (Found: C, 67.9; H, 3.40; N, 6.3;  $C_{12}H_7NO_3$  requires C, 67.6; H, 3.31; N, 6.6%). HPLC (Method A), R.t. 2.58 min, (100%);  $\delta_H$  (500 MHz,  $CDCl_3$ ); 8.14 (1H, dd,  $J = 7.8$ , 1.2, Ar- $H_5$ ), 7.74 (1H, app. t,  $J = 7.8$ , Ar- $H_6$  or Ar- $H_7$ ), 7.62-7.64 (2H, m, Ar-H), 7.50 (1H, app. t,  $J = 7.8$ , Ar- $H_6$  or Ar- $H_7$ ), 7.29 (1H, dd,  $J = 3.5$ , 0.7, Ar- $H_{3'}$  or Ar- $H_{4'}$ ), 6.55 (1H, dd,  $J = 3.5$ , 0.7, Ar- $H_{3'}$  or Ar- $H_{4'}$ );  $\delta_C$  (125 MHz,  $CDCl_3$ ); 158.2 (4-C), 149.2 (2-C), 147.9 (Ar-C), 146.3 (2'-C), 144.1 (8a-C), 136.9 (Ar-C), 128.3 (Ar-C), 128.2 (Ar-C), 126.6 (Ar-C), 117.0 (3'-C or 4'-C), 116.9 (4a-C), 112.9 (3'-C or 4'-C);  $\nu_{max}/cm^{-1}$  (neat); 1767, 1705, 1634, 1599, 1472, 1456, 1326, 1271;  $m/z$  (ES) 236.0 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 236.0311.  $C_{12}H_7NO_3$  requires  $MNa$ , 236.0318).

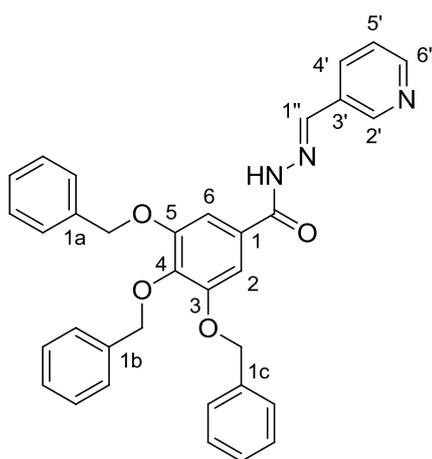
## 2-(pyridin-4-yl-methyl)quinazolin-4(3H)-one (5.21)



To a solution of 2-mercaptoquinazolin-4(3H)-one **5.39** (89 mg, 0.50 mmol) and triethylamine (5.0 mL, 0.50 mmol) in acetone (5 mL) at room temperature was added 4-(bromomethyl)pyridine hydrobromide **5.40** (126 mg, 0.50 mmol). The mixture was heated at 55 °C for 24 hours and then cooled to room temperature. The resulting brown slurry was dried under vacuo and the brown oily residue was recrystallized from water to give the title compound **5.21** (62 mg, 0.35 mmol, 70%) as colourless plates m.p. 204—205 °C.  $R_f$  0.40 (1:1 petroleum ether—EtOAc); HPLC (Method B), R.t. 1.30 min, (100%);  $\delta_H$  (300 MHz,  $DMSO-d_6$ ); 12.67 (1H, br s, N-

H), 8.50 (2H, dd,  $J = 4.5, 1.5$ , Ar-H<sub>2',6'</sub>), 8.02 (1H, dd,  $J = 7.9, 1.2$ , Ar-H<sub>5</sub>), 7.71-7.83 (1H, m, Ar-H<sub>7</sub>), 7.59 (1H, app. d,  $J = 7.9$ , Ar-H<sub>8</sub>), 7.51 (1H, dd,  $J = 4.5, 1.5$  Hz, Ar-H<sub>3',5'</sub>), 7.37-7.49 (1H, m, Ar-H<sub>6</sub>), 4.48 (2H, s, H<sub>1''</sub>);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 161.1 (4-C), 154.6 (2-C), 149.6 (2'-C, 6'-C), 148.2 (8a-C), 146.9 (4'-C), 134.7 (7-C), 126.0 (5-C, 6-C), 125.8 (8-C), 124.2 (3'-C, 5'-C), 120.0 (4a-C), 32.4 (1''-C);  $\nu_{max}/cm^{-1}$  (neat); 3061, 1682, 1579, 1562, 1455, 1241, 1117, 762;  $m/z$  (ES) 270.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 270.0689. C<sub>14</sub>H<sub>11</sub>N<sub>3</sub>OS requires *MH*, 270.0696).

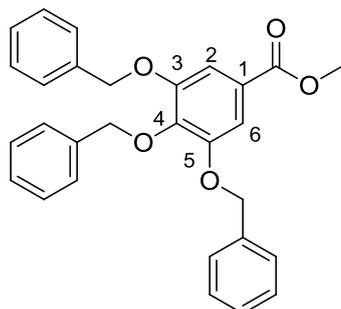
### (E)-3,4,5-tris(benzyloxy)-N'-(pyridin-3-yl-methylene)benzohydrazide (**6.3**)



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (227 mg, 0.5 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.06 mL, 0.5 mmol). Recrystallization from ethanol gave the title compound **6.3** (215 mg, 0.4 mmol, 79%) as colourless plates m.p. 200—201 °C.  $R_f$  0.36 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); (Found: C, 75.3; H, 5.10; N, 7.6; C<sub>34</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires C, 75.1; H, 5.38; N, 7.7%). HPLC

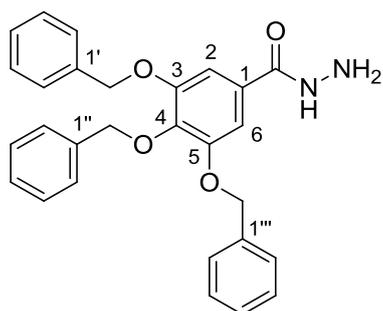
(Method B), R.t. 3.57 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.92 (1H, s, N-H), 8.87 (1H, app. s, Ar-H<sub>2'</sub>), 8.63 (1H, app. d,  $J = 4.5$ , Ar-H<sub>6'</sub>), 8.54 (1H, s, C-H<sub>1''</sub>), 8.17 (1H, app. d,  $J = 7.9$ , Ar-H<sub>4'</sub>), 7.26-7.53 (18H, m, Ar-H), 5.22 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.5 (C=O), 152.1 (3C, 5C), 150.7 (6'-C), 148.8 (2'-C), 145.0 (1''-C), 140.1 (4-C), 137.4 (1b-C), 136.7 (1a-C, 1c-C), 133.4 (4'-C), 130.2 (3'-C), 128.4 (Ar-C), 128.3 (1-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 124.0 (Ar-C), 106.9 (Ar-C), 74.2 (OCH<sub>2</sub> in 4), 70.4 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3231, 3033, 1646, 1580, 1540, 1426, 1371, 1334;  $m/z$  (ES) 544.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 544.2217. C<sub>34</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires *MH*, 544.2231).

### Methyl 3,4,5-tris(benzyloxy)benzoate (6.31)<sup>11</sup>



Potassium carbonate (7.38 g, 54.3 mmol) and benzyl bromide (1.35 mL, 67.9 mmol) were added to a solution of 3,4,5-trihydroxybenzoate methyl ester **6.30** (2.50 g, 13.6 mmol) in dry DMF (30 mL). This reaction mixture was stirred at 120 °C for seven hours, and then poured into water (90 mL). The mixture was extracted with EtOAc (4 x 60 mL). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The brown solid was purified using column chromatography on silica gel eluting with petroleum ether / ethylacetate (9:1) and recrystallized from hexane to give the title compound **6.31**<sup>11</sup> (1.85 g, 4.08 mmol, 30%) as colourless needles m.p. 97—98 °C (Lit.<sup>11</sup> 99—100 °C). *R<sub>f</sub>* 0.95 (1:1 petroleum ether—EtOAc); HPLC (Method A), R.t. 4.48 min, (100%); δ<sub>H</sub> (300 MHz, DMSO-*d*<sub>6</sub>); 7.45-7.47 (4H, m, Ar-H), 7.38-7.41 (4H, m, Ar-H), 7.32-7.36 (6H, m, Ar-H), 7.24-7.28 (3H, m, Ar-H), 5.18 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.04 (2H, s, OCH<sub>2</sub> in 4), 3.82 (3H, s, OCH<sub>3</sub>); δ<sub>C</sub> (75 MHz, DMSO-*d*<sub>6</sub>); 165.7 (C=O), 152.1 (3C, 5C), 141.3 (4-C), 137.3 (Ar-C<sub>qt</sub>), 136.7 (Ar-C<sub>qt</sub>), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 127.89 (Ar-C), 127.88 (Ar-C), 127.6 (Ar-C), 124.8 (1-C), 108.1 (2-C and 6-C), 74.2 (OCH<sub>2</sub>), 70.3 (OCH<sub>2</sub>), 52.2 (OCH<sub>3</sub>); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3411, 3066, 3031, 2949, 2879, 1967, 1715, 1587, 1495, 1428; *m/z* (ES) 477.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 477.1681. C<sub>29</sub>H<sub>26</sub>O<sub>5</sub> requires MNa, 477.1672).

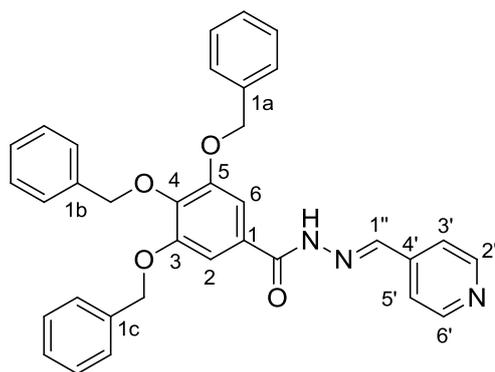
### 3,4,5-tris(benzyloxy)benzohydrazide (6.32)<sup>12</sup>



Prepared *via* general method C using methyl 3,4,5-tris(benzyloxy)benzoate **6.31** (2.0 g, 4.4 mmol) and hydrazine hydrate (2.3 ml, 44 mmol). Recrystallization from acetone / water to give the title compound **6.32**<sup>12</sup> (1.7 g, 3.7 mmol, 85%) as colourless fluffy solid m.p. 134—135 °C (Lit.<sup>12</sup> 137—137.5 °C). *R<sub>f</sub>* 0.72 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 0.40

min, (100%);  $\delta_{\text{H}}$  (500 MHz, DMSO- $d_6$ ); 9.73 (1H, s, N-H), 7.49-7.23 (17H, m, Ar-H), 5.17 (4H, s, OCH<sub>2</sub>), 4.99 (2H, s, OCH<sub>2</sub>), 4.49 (2H, s, NH<sub>2</sub>);  $\delta_{\text{C}}$  (125 MHz, DMSO- $d_6$ ); 165.2 (C=O), 151.9 (3-C, 5-C), 139.4 (4-C), 137.5 (1'-C, 1'''-C), 136.9 (1-C), 128.5 (1''-C), 128.4 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 127.6 (Ar-C), 106.1 (2-C, 6-C), 74.2 (OCH<sub>2</sub>), 70.2 (OCH<sub>2</sub>);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3278, 3065, 3032, 1951, 1630, 1583, 1499, 1345;  $m/z$  (ES) 455.2 (90%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 455.1968. C<sub>28</sub>H<sub>26</sub>N<sub>2</sub>O<sub>4</sub> requires MH, 455.1965).

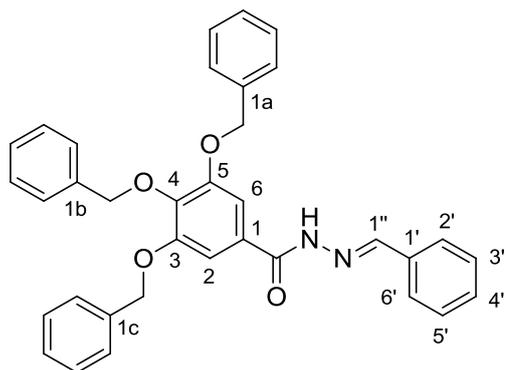
**(E)-3,4,5-tris(benzyloxy)-N'-(pyridine-4-yl-methylene)benzohydrazide (6.34)<sup>13</sup>**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (241 mg, 0.53 mmol) and 4-pyridinecarboxyaldehyde **6.118** (0.09 mL, 0.53 mmol). Recrystallization from ethanol gave the title compound **6.34** (271 mg, 0.50 mmol, 94%) as colourless needles

m.p. 206—207 °C (Lit.<sup>13</sup> not available).  $R_f$  0.28 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 3.47 min, (100%);  $\delta_{\text{H}}$  (500 MHz, DMSO- $d_6$ ); 12.00 (1H, s, N-H), 8.67 (2H, d,  $J$  = 4.5, Ar-H<sub>2',6'</sub>), 8.48 (1H, s, C-H<sub>1''</sub>), 7.68 (2H, d,  $J$  = 4.5, Ar-H<sub>3',5'</sub>), 7.48-7.51 (4H, m, Ar-H), 7.34-7.44 (10H, m, Ar-H), 7.26-7.32 (3H, m, Ar-H), 5.21 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_{\text{C}}$  (125 MHz, DMSO- $d_6$ ); 162.7 (C=O), 152.1 (3-C and 5-C), 150.3 (2'-C and 6'-C), 145.3 (1''-C), 141.5 (4'-C), 140.4 (4-C), 137.4 (1b-C), 136.7 (1a-C and 1c-C), 128.44 (Ar-C), 128.43 (1-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 121.0 (3'-C and 5'-C), 107.1 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3219, 3029, 1648, 1578, 1531, 1452, 1424, 1370, 1324, 1112, 680;  $m/z$  (ES) 544.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 544.2235. C<sub>34</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires MH, 544.2231).

### (E)-N'-benzylidene-3,4,5-tris(benzyloxy)benzohydrazide (6.35)

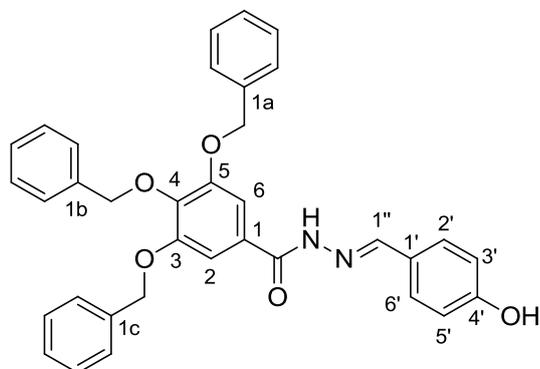


Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (227 mg, 0.50 mmol) and benzaldehyde **6.119** (0.05 mL, 0.50 mmol).

Recrystallization from ethanol gave the title compound **6.35** (239 mg, 0.44 mmol, 88%) as colourless needles m.p. 199—

200 °C.  $R_f$  0.62 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method C), R.t. 3.23 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.76 (1H, s, N-H), 8.49 (1H, s, C-H<sub>1''</sub>), 7.74–7.77 (2H, m, Ar-H), 7.09–7.57 (20H, m, Ar-H), 5.22 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.4 (C=O), 152.1 (3-C and 5-C), 147.7 (1''-C), 140.0 (4-C), 137.4 (1b-C), 136.7 (1a-C and 1c-C), 134.2 (1'-C), 130.1 (Ar-C), 128.9 (Ar-C), 128.5 (1-C), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 127.0 (Ar-C), 106.8 (2-C and 6-C), 74.2 (OCH<sub>2</sub> in 4), 70.4 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3209, 3033, 1640, 1577, 1549, 1422, 1363, 1099, 752;  $m/z$  (ES) 565.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 565.2091. C<sub>35</sub>H<sub>30</sub>N<sub>2</sub>O<sub>4</sub> requires MNa, 565.2098).

### (E)-3,4,5-tris(benzyloxy)-N'-(4-hydroxybenzylidene)benzohydrazide (6.36)

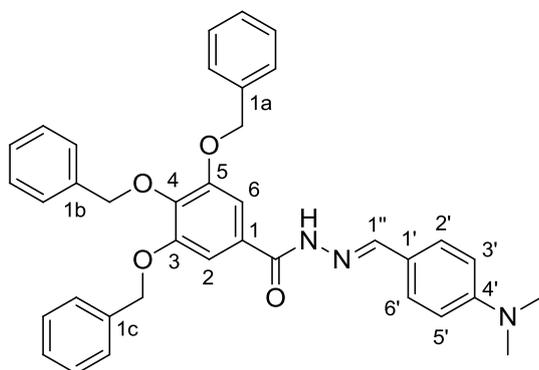


Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (300 mg, 0.66 mmol) and 4-hydroxybenzaldehyde **6.120** (80.6 mg, 0.66 mmol). Recrystallization from ethanol gave the title compound **6.36** (295 mg, 0.53 mmol, 80%) as colourless needles m.p. 204—205

°C.  $R_f$  0.24 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); (Found: C, 75.2; H, 5.40; N, 5.0; C<sub>35</sub>H<sub>30</sub>N<sub>2</sub>O<sub>5</sub> requires C, 75.2; H, 5.41; N, 5.0%). HPLC (Method B), R.t. 3.92 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.49 (1H, s, N-H), 9.91 (1H, s, O-H in 4'), 8.36 (1H, s, C-H<sub>1''</sub>), 7.56 (2H, d,  $J$  = 8.5, Ar-H<sub>2',6'</sub>), 7.47–7.50 (4H, m, Ar-H), 7.32–7.45 (10H, m,

Ar-H), 7.32-7.22 (3H, m, Ar-H), 6.84 (2H, d,  $J = 8.5$ , Ar-H<sub>3',5'</sub>), 5.20 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.01 (2H, s, OCH<sub>2</sub> in 4);  $\delta_{\text{C}}$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.1 (C=O), 159.4 (4'-C), 152.0 (3-C and 5-C), 148.1 (1''-C), 140.0 (4-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 128.82 (2'-C and 6'-C), 128.79 (1-C), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 125.3 (1'-C), 115.7 (3'-C and 5'-C), 106.9 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3486, 3419, 3030, 1648, 1596, 1510, 1424, 1233, 1073, 683;  $m/z$  (ES) 581.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 581.2034. C<sub>35</sub>H<sub>30</sub>N<sub>2</sub>O<sub>5</sub> requires MNa, 581.2047).

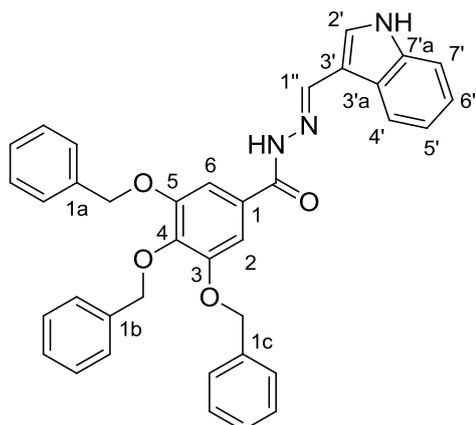
**(E)-3,4,5-tris(benzyloxy)-N'-(4-dimethylamino)benzylidene)benzohydrazide (6.37)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (119 mg, 0.26 mmol) and 4-(dimethylamino)benzaldehyde **6.130** (39 mg, 0.26 mmol). Recrystallization from ethanol gave the title compound **6.37** (116 mg, 0.20 mmol, 76%) as

light yellow plates m.p. 84—85 °C.  $R_f$  0.84 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); (Found: C, 75.7; H, 6.00; N, 7.2; C<sub>35</sub>H<sub>37</sub>N<sub>3</sub>O<sub>4</sub> requires C, 75.9; H, 6.02; N, 7.2%). HPLC (Method B), R.t. 4.40 min, (100%);  $\delta_{\text{H}}$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.45 (1H, s, N-H), 8.33 (1H, s, C-H<sub>1''</sub>), 7.56 (2H, d,  $J = 8.8$ , Ar-H<sub>2',6'</sub>), 7.46-7.50 (4H, m, Ar-H), 7.32-7.45 (10H, m, Ar-H), 7.23-7.31 (3H, m, Ar-H), 6.77 (2H, d,  $J = 8.8$ , Ar-H<sub>3',5'</sub>), 5.21 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.02 (2H, s, OCH<sub>2</sub> in 4), 2.98 (6H, s, CH<sub>3</sub>);  $\delta_{\text{C}}$  (125 MHz, DMSO-*d*<sub>6</sub>); 161.9 (C=O), 152.0 (3-C and 5-C), 151.5 (4'-C), 148.7 (1''-C), 139.9 (4-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 129.0 (2'-C and 6'-C), 128.42 (Ar-C), 128.38 (1-C), 128.0 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 127.7 (Ar-C), 121.5 (1'-C), 111.8 (3'-C and 5'-C), 106.7 (2-C and 6-C), 74.2 (OCH<sub>2</sub> in 4), 70.4 (OCH<sub>2</sub> in 3 and 5), 40.3 (NCH<sub>3</sub>);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3223, 3031, 2899, 1644, 1588, 1521, 1315, 1092;  $m/z$  (ES) 586.3 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 586.2680. C<sub>37</sub>H<sub>35</sub>N<sub>3</sub>O<sub>4</sub> requires MNa, 586.2676).

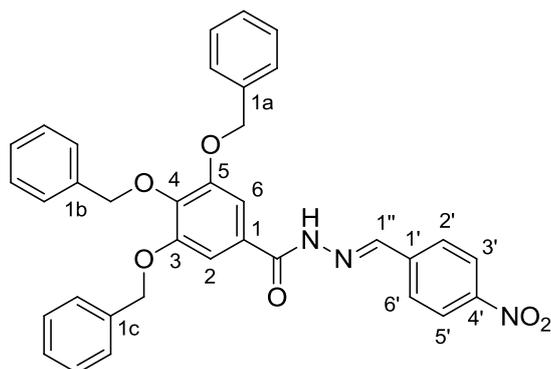
**(E)-N'-((1H-indol-3-yl)methylene)-3,4,5-tris(benzyloxy)benzohydrazide (6.38)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (156 mg, 0.50 mmol) and indole-3-carboxyaldehyde **6.132** (50 mg, 0.50 mmol). Recrystallization from ethanol / water gave the title compound **6.38** (238 mg, 0.41 mmol, 82%) as colourless needles m.p. 231 °C (dec).  $R_f$  0.48 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC

(Method B), R.t. 4.09 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.61 (1H, s, N-H *indole*), 11.42 (1H, s, N-H *hydrazide*), 8.66 (1H, s, 1''-C), 8.32 (1H, app. d,  $J = 7.5$ , Ar-H<sub>4'</sub>), 7.85 (1H, d,  $J = 2.7$ , Ar-H<sub>2'</sub>), 7.26-7.57 (18H, m, Ar-H), 7.23 (1H, app. t,  $J = 7.5$ , Ar-H<sub>5'</sub> or Ar-H<sub>6'</sub>), 7.18 (1H, app. t,  $J = 7.5$ , Ar-H<sub>5'</sub> or Ar-H<sub>6'</sub>), 5.24 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.05 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 161.8 (C=O), 152.0 (3-C and 5-C), 145.0 (1''-C), 139.7 (4-C), 137.4 (1a-C and 1c-C), 137.0 (Ar-C<sub>qt</sub>), 136.8 (Ar-C<sub>qt</sub>), 130.3 (2'-C), 129.2 (Ar-C<sub>qt</sub>), 128.4 (1b-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 124.3 (Ar-C<sub>qt</sub>), 122.6 (4'-C), 122.0 (Ar-C), 120.4 (Ar-C), 111.9 (Ar-C), 111.6 (Ar-C<sub>qt</sub>), 106.6 (Ar-C), 74.2 (OCH<sub>2</sub> in 4), 70.4 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3384, 3216, 3029, 1636, 1577, 1555, 1525, 1321, 1095, 749;  $m/z$  (ES) 604.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 604.2198. C<sub>37</sub>H<sub>31</sub>N<sub>3</sub>O<sub>4</sub> requires *MNa*, 604.2207).

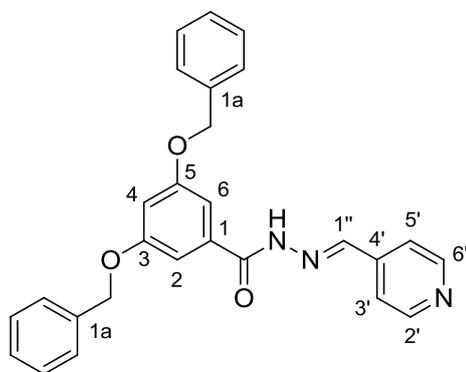
**(E)-3,4,5-tris(benzyloxy)-N'-(4-nitrobenzylidene)benzohydrazide (6.39)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (247 mg, 0.54 mmol) and 4-nitrobenzaldehyde **6.131** (82.1 mg, 0.54 mmol). Recrystallization from ethanol gave the title compound **6.39** (289 mg, 0.49 mmol, 91%) as colourless plates m.p. 234—235 °C.  $R_f$  0.34 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); (Found: C, 71.5; H, 5.00; N, 7.2; C<sub>35</sub>H<sub>29</sub>N<sub>3</sub>O<sub>6</sub> requires C, 71.5; H, 5.00; N, 7.2%). HPLC

(Method B), R.t. 4.38 min, (100%);  $\delta_{\text{H}}$  (500 MHz, DMSO- $d_6$ ); 12.00 (1H, s, N-H), 8.57 (1H, s, C-H $_{1''}$ ), 8.31 (2H, d,  $J = 8.3$ , Ar-H $_{3',5'}$ ), 8.00 (2H, d,  $J = 8.3$ , Ar-H $_{2',6'}$ ), 7.47-7.50 (4H, m, Ar-H), 7.32-7.44 (10H, m, Ar-H), 7.32-7.21 (3H, m, Ar-H), 5.21 (4H, s, OCH $_2$  in 3 and 5), 5.03 (2H, s, OCH $_2$  in 4);  $\delta_{\text{C}}$  (125 MHz, DMSO- $d_6$ ); 162.7 (C=O), 152.1 (3-C, 5-C), 147.9 (4'-C), 145.3 (1''-C), 140.6 (4-C or 1'-C), 140.4 (4-C or 1'-C), 137.4 (1b-C), 136.7 (1a-C and 1c-C), 128.4 (Ar-C), 128.23 (1-C), 128.17 (2'-C, 6'-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 127.7 (Ar-C), 124.1 (3'-C and 5'-C), 107.1 (2-C and 6-C), 74.3 (OCH $_2$  in 4), 70.5 (OCH $_2$  in 3 and 5);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3216, 3062, 3038, 1650, 1541, 1466, 1425, 1267, 1127, 694;  $m/z$  (ES) 610.2 (100%, MNa $^+$ ); (Found MNa $^+$ , 610.1921. C $_{35}$ H $_{29}$ N $_3$ O $_6$  requires MNa, 610.1949).

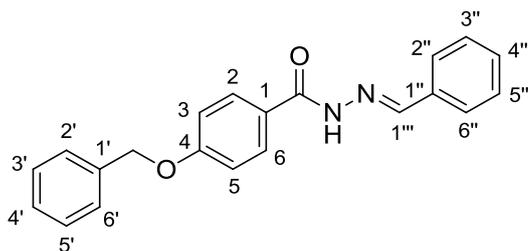
**(E)-3,5-bis(benzyloxy)-N'-(pyridin-4-yl-methylene)benzohydrazide (6.40)**<sup>13</sup>



Prepared *via* general method D using 3,5-bis(benzyloxy)benzohydrazide **6.109** (174 mg, 0.50 mmol) and 4-pyridinecarboxyaldehyde **6.118** (0.047 mL, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.40** (192 mg, 0.44 mmol, 88%) as colourless needles m.p. 183—184 °C (Lit.<sup>13</sup> not available).  $R_f$  0.35

(95:5 CH $_2$ Cl $_2$ —MeOH); HPLC (Method B), R.t. 4.12 min, (100%);  $\delta_{\text{H}}$  (500 MHz, DMSO- $d_6$ ); 12.01 (1H, s, N-H), 8.65 (2H, d,  $J = 4.5$ , Ar-H $_{2',6'}$ ), 8.44 (1H, s, C-H $_{1''}$ ), 7.65 (2H, d,  $J = 4.5$ , Ar-H $_{3',5'}$ ), 7.30-7.49 (10H, m, Ar-H), 7.18 (2H, d,  $J = 2.0$ , Ar-H $_{2,6}$ ), 6.93 (1H, t,  $J = 2.0$ , Ar-H $_4$ ), 5.16 (4H, s, OCH $_2$  in 3 and 5);  $\delta_{\text{C}}$  (125 MHz, DMSO- $d_6$ ); 162.8 (C=O), 159.5 (3-C and 5-C), 150.3 (2'-C and 6'-C), 145.5 (1''-C), 141.4 (4'-C), 136.7 (1aC), 135.0 (1-C), 128.5 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 121.0 (3'-C and 5'-C), 106.8 (2-C and 6-C), 105.2 (4-C), 69.6 (OCH $_2$ );  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3227, 3032, 1657, 1596, 1537, 1364, 1133, 729;  $m/z$  (ES) 438.2 (100%, MH $^+$ ); (Found MH $^+$ , 438.1824. C $_{27}$ H $_{23}$ N $_3$ O $_3$  requires MH, 438.1812).

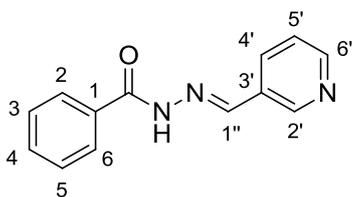
**(E)-N'-benzylidene-4-(benzyloxy)benzohydrazide (6.41)<sup>14</sup>**



Prepared *via* general method D using 4-(benzyloxy)benzohydrazide **6.106** (121 mg, 0.50 mmol) and benzaldehyde **6.119** (0.05 mL, 0.50 mmol). Recrystallization from ethanol gave the title

compound **6.41** (140 mg, 0.42 mmol, 85%) as colourless plates m.p. 186—187 °C (Lit.<sup>14</sup> m.p. 185).  $R_f$  0.50 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.74 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.70 (1H, s, N-H), 8.44 (1H, s, C-H<sub>1''</sub>), 7.90 (2H, d,  $J = 8.9$ , Ar-H<sub>2,6</sub>), 7.70-7.72 (2H, m, Ar-H), 7.30-7.50 (8H, m, Ar-H), 7.13 (2H, d,  $J = 8.9$ , Ar-H<sub>3,5</sub>), 5.19 (2H, s, OCH<sub>2</sub>);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.5 (C=O), 161.1 (4-C), 147.2 (1''-C), 136.6 (1'-C), 134.4 (1''-C), 129.9 (2-C and 6-C), 129.5 (Ar-C), 128.8 (Ar-C), 128.5 (Ar-C), 128.0 (Ar-C), 127.8 (Ar-C), 127.0 (Ar-C), 125.6 (1-C), 114.6 (3-C and 5-C), 69.4 (OCH<sub>2</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3166, 2944, 1634, 1608, 1559, 1509, 1487, 1287, 1249, 1183, 691;  $m/z$  (ES) 353.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 353.1274. C<sub>21</sub>H<sub>18</sub>N<sub>2</sub>O<sub>2</sub> requires MNa, 353.1260).

**(E)-N'-(pyridin-3-yl-methylene)benzohydrazide (6.42)<sup>15</sup>**

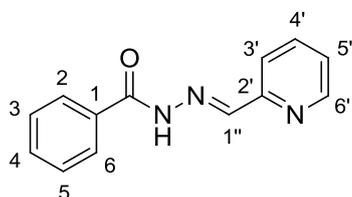


Prepared *via* general method D using benzoic hydrazide **6.112** (408 mg, 3.0 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.366 mL, 3.0 mmol).

Recrystallization from ethanol gave the title compound **6.42** (360 mg, 2.6 mmol, 88%) as colourless needles m.p. 185—186 °C (Lit.<sup>15</sup> m.p. 184—185 °C).  $R_f$  0.34 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 1.46 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 12.01 (1H, s, N-H), 8.86 (1H, app. s, Ar-H<sub>2</sub>), 8.61 (1H, app. d,  $J = 4.3$ , Ar-H<sub>6</sub>), 8.52 (1H, s, C-H<sub>1''</sub>), 8.14 (1H, app. d,  $J = 7.5$ , Ar-H<sub>4</sub>), 7.92 (2H, app. d,  $J = 7.4$ , Ar-H<sub>2,6</sub>), 7.59 (1H, app. t,  $J = 7.4$ , Ar-H<sub>4</sub>), 7.52 (2H, app. t,  $J = 7.4$ , Ar-H<sub>3,5</sub>), 7.48 (1H, dd,  $J = 7.5, 4.3$ , Ar-H<sub>5</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 163.2 (C=O), 150.7 (6'-C), 148.7 (2'-C), 145.0 (1''-C), 133.4 (4'-C), 133.2 (1-C), 131.9 (4-C), 130.2 (3'-C), 128.5 (3-C, 5-C), 127.6 (2-C, 6-C), 124.0 (5'-C);  $\nu_{max}/cm^{-1}$  (neat); 3180, 3015, 1671, 1602, 1591, 1274,

1137, 685;  $m/z$  (ES) 226.1 (100%,  $MH^+$ ); (Found  $MH^+$ , 226.0981.  $C_{13}H_{11}N_3O$  requires  $MH$ , 270.0975).

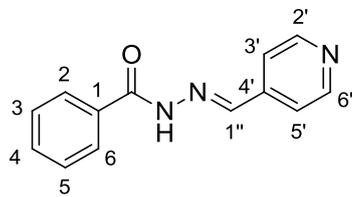
**(E)-N'-(pyridin-2-yl-methylene)benzohydrazide (6.43)<sup>16</sup>**



Prepared *via* general method D using benzoic hydrazide **6.112** (400 mg, 2.94 mmol) and 2-pyridinecarboxyaldehyde **6.117** (0.28 mL, 2.94 mmol). Recrystallization from ethanol gave the title

compound **6.43** (452 mg, 2.01 mmol, 76%) as colourless needles m.p. 107—108 °C (Lit.<sup>16</sup> m.p. 108—109 °C).  $R_f$  0.52 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 1.64 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 12.04 (1H, s, N-H), 8.61 (1H, app. d,  $J = 4.2$ , Ar-H<sub>6</sub>), 8.48 (1H, s, C-H<sub>1''</sub>), 7.84-8.00 (4H, m, Ar-H), 7.52-7.62 (3H, m, Ar-H), 7.31-7.46 (1H, m, Ar-H *pyridyl*);  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 163.4 (C=O), 153.3 (2'-C), 149.5 (6'-C), 148.0 (1''-C), 136.8 (Ar-C), 133.2 (1-C), 131.9 (Ar-C), 128.5 (Ar-C), 127.7 (Ar-C), 124.4 (Ar-C), 119.9 (Ar-C);  $\nu_{max}/cm^{-1}$  (neat); 3646, 3182, 3004, 1656, 1558, 1468, 1280, 1182, 698;  $m/z$  (ES) 248.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 248.0795.  $C_{13}H_{11}N_3O$  requires  $MNa$ , 248.0794).

**(E)-N'-(pyridin-4-yl-methylene)benzohydrazide (6.44)<sup>15</sup>**

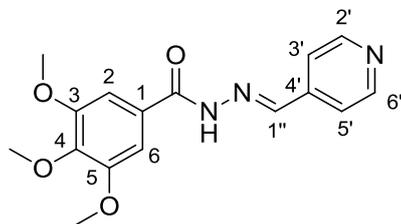


Prepared *via* general method D using benzoic hydrazide **6.112** (433 mg, 3.18 mmol) and 4-pyridinecarboxyaldehyde **6.118** (0.30 mL, 3.18 mmol). Recrystallization from ethanol gave the title

compound **6.44** (572 mg, 2.54 mmol, 80%) as colourless needles m.p. 162—163 °C (Lit.<sup>15</sup> m.p. 162—163 °C).  $R_f$  0.53 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 1.36 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 12.11 (1H, s, N-H), 8.65 (2H, d,  $J = 4.5$ , Ar-H<sub>2',6'</sub>), 8.45 (1H, s, C-H<sub>1''</sub>), 7.92 (2H, app. d,  $J = 7.3$ , Ar-H<sub>2,6</sub>), 7.69 (2H, d,  $J = 4.5$ , Ar-H<sub>3',5'</sub>), 7.60 (1H, app. t,  $J = 7.3$ , Ar-H<sub>4</sub>), 7.53 (2H, app. t,  $J = 7.3$ , Ar-H<sub>3,5</sub>);  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 163.4 (C=O), 150.3 (2'-C, 6'-C), 145.3 (1''-C), 141.5 (4'-C), 133.1 (1-C), 132.0 (4-C), 128.5 (3-C, 5-C), 127.7 (2-C, 6-C), 121.0 (3'-C, 5'-C);  $\nu_{max}/cm^{-1}$  (neat); 3257, 3038, 1646, 1546, 1489, 1306,

1277, 1140, 627;  $m/z$  (ES) 226.1 (100%,  $MH^+$ ); (Found  $MH^+$ , 226.0979.  $C_{13}H_{11}N_3O$  requires  $MH$ , 226.0975).

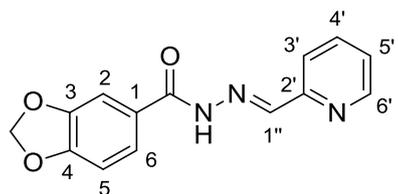
**(E)-3,4,5-trimethoxy-N'-(pyridin-4-yl-methylene)benzohydrazide (6.45)<sup>17</sup>**



Prepared *via* general method D using 3,4,5-trimethoxybenzohydrazide **6.113** (452 mg, 2.0 mmol) and 4-pyridinecarboxyaldehyde **6.118** (0.10 mL, 2.0 mmol). Recrystallization from ethanol gave the title compound **6.45** (385 mg,

1.7 mmol, 85%) as colourless plates m.p. 174—175 °C (Lit.<sup>17</sup> not available).  $R_f$  0.20 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 1.61 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 12.01 (1H, s, N-H), 8.67 (2H, d,  $J = 4.9$ , Ar- $H_{2',6'}$ ), 8.46 (1H, s, C- $H_{1''}$ ), 7.68 (2H, d,  $J = 4.9$ , Ar- $H_{3',5'}$ ), 7.25 (2H, s, Ar- $H_{2,6}$ ), 3.87 (6H, s,  $OCH_3$  in 3 and 5), 3.74 (3H, s,  $OCH_3$  in 4);  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 162.8 (C=O), 152.7 (3-C, 5-C), 150.3 (2'-C, 6'-C), 145.2 (1''-C), 141.5 (4'-C), 140.7 (4-C), 128.1 (1-C), 120.9 (3'-C, 5'-C), 105.4 (2-C, 6-C), 60.1 ( $OCH_3$  in 4), 56.1 ( $OCH_3$  in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3161, 3003, 1644, 1500, 1451, 1362, 1326, 1177, 1124;  $m/z$  (ES) 316.3 (100%,  $MH^+$ ); (Found  $MH^+$ , 316.1302.  $C_{16}H_{17}N_3O_4$  requires  $MH$ , 316.1292).

**(E)-N'-(pyridin-2-yl-methylene)benzo[d][1,3]dioxole-5-carbohydrazide (6.46)<sup>15</sup>**

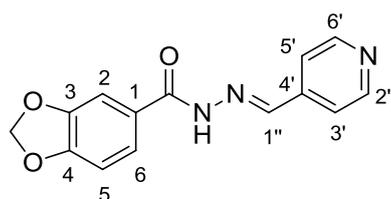


Prepared *via* general method D using 3,4-methylenedioxybenzhydrazide **6.114** (180 mg, 1.0 mmol) and 2-pyridinecarboxyaldehyde **6.117** (0.095 mL, 1.0 mmol). Recrystallization from

ethanol gave the title compound **6.46** (172 mg, 0.64 mmol, 64%) as colourless plates m.p. 181—182 °C (Lit.<sup>15</sup> m.p. 182—183 °C).  $R_f$  0.26 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 1.71 min, (100%);  $\delta_H$  (300 MHz,  $DMSO-d_6$ ); 11.92 (1H, s, N-H), 8.62 (1H, app. d,  $J = 4.8$ , Ar- $H_{6'}$ ), 8.45 (1H, s, C- $H_{1''}$ ), 7.97 (1H, app. d,  $J = 7.8$ , Ar- $H_3$ ), 7.89 (1H, app. t,  $J = 7.8$ , Ar- $H_4$ ), 7.54 (1H, dd,  $J = 8.1, 1.4$ , Ar- $H_6$ ), 7.47 (1H, d,  $J = 1.4$ , Ar- $H_2$ ), 7.36-7.45 (1H, m, Ar- $H_5$ ), 7.08 (1H,

d,  $J = 8.1$ , Ar-H<sub>5</sub>), 6.15 (2H, s, OCH<sub>2</sub>O);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.4 (C=O), 153.3 (4-C), 150.3 (2'-C), 149.5 (6'-C), 147.6 (1''-C), 147.4 (3-C), 136.8 (4'-C), 126.9 (1-C), 124.3 (2-C), 123.0 (6-C), 119.8 (3'-C), 108.1 (5-C), 107.7 (5'-C), 101.9 (OCH<sub>2</sub>O);  $\nu_{max}/cm^{-1}$  (neat); 3224, 3079, 1644, 1498, 1434, 1365, 1257, 1167, 1032;  $m/z$  (ES) 292.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 292.0693). C<sub>14</sub>H<sub>11</sub>N<sub>3</sub>O<sub>3</sub> requires MNa, 292.0693).

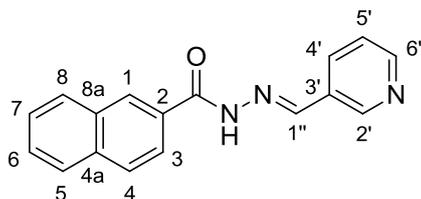
**(E)-N'-(pyridin-4-yl-methylene)benzo[d][1,3]dioxole-5-carbohydrazide (6.47)<sup>15</sup>**



Prepared *via* general method D using 3,4-methylenedioxybenzhydrazide **6.114** (180 mg, 1.0 mmol) and 4-pyridinecarboxyaldehyde **6.118** (0.094 mL, 1.0 mmol). Recrystallization from

ethanol gave the title compound **6.47** (215 mg, 0.8 mmol, 81%) as colourless plates m.p. 259—260 °C (Lit.<sup>15</sup> m.p. 257—258 °C).  $R_f$  0.36 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 1.52 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 12.00 (1H, s, N-H), 8.65 (2H, d,  $J = 4.9$ , Ar-H<sub>2',6'</sub>), 8.41 (1H, s, C-H<sub>1''</sub>), 7.66 (2H, d,  $J = 4.9$ , Ar-H<sub>3',5'</sub>), 7.54 (1H, app. d,  $J = 8.1$ , Ar-H<sub>6</sub>), 7.47 (1H, app. s, Ar-H<sub>2</sub>), 7.08 (1H, d,  $J = 8.1$ , Ar-H<sub>5</sub>), 6.15 (2H, s, OCH<sub>2</sub>O);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.4 (C=O), 150.4 (3-C or 4-C), 150.3 (2'-C, 6'-C), 147.4 (3-C or 4-C), 144.9 (1''-C), 141.5 (4'-C), 126.8 (1-C), 123.0 (6-C), 120.9 (3'-C, 5'-C), 108.1 (2-C), 107.7 (5-C), 101.9 (OCH<sub>2</sub>O);  $\nu_{max}/cm^{-1}$  (neat); 3236, 3028, 2904, 1645, 1538, 1497, 1481, 1404, 1288, 1211, 1167;  $m/z$  (ES) 270.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 270.0884). C<sub>14</sub>H<sub>11</sub>N<sub>3</sub>O<sub>3</sub> requires MH, 270.0873).

**(E)-N'-(pyridin-3-ylmethylene)-2-naphthohydrazide (6.48)<sup>18</sup>**

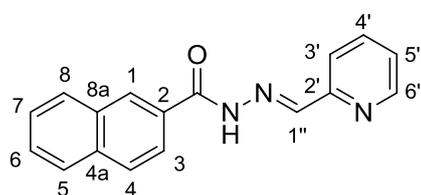


Prepared *via* general method D using 2-naphthohydrazide **6.105** (186 mg, 1.0 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.094 mL, 1.0 mmol). Recrystallization from methanol /

water gave the title compound **6.48** (201 mg, 0.73 mmol, 73%) as colourless needles m.p. 205—206 °C (Lit.<sup>18</sup> not available).  $R_f$  0.51 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH);

HPLC (Method B), R.t. 2.93 min, (100%);  $\delta_{\text{H}}$  (500 MHz, DMSO- $d_6$ ); 12.18 (1H, s, N-H), 8.88 (1H, app. s, Ar-H $_2$ '), 8.62 (1H, app. s,  $J = 4.5$ , Ar-H $_6$ '), 8.56 (1H, s, C-H $_1$ '), 8.55 (1H, app. s, Ar-H $_1$ ), 8.17 (1H, app. d,  $J = 7.5$ , Ar-H $_4$ '), 7.91-8.10 (4H, m, Ar-H $_{3,4,5,8}$ ), 7.57-7.68 (2H, m, Ar-H $_{6,7}$ ), 7.50 (1H, dd,  $J = 7.3, 4.5$ , Ar-H $_5$ ');  $\delta_{\text{C}}$  (125 MHz, DMSO- $d_6$ ); 163.3 (C=O), 150.7 (6'-C), 148.8 (2'-C), 145.1 (1''-C), 134.4 (Ar-C $_{\text{qt}}$ ), 133.5 (4'-C), 132.0 (2-C), 130.5 (Ar-C $_{\text{qt}}$ ), 130.3 (Ar-C $_{\text{qt}}$ ), 128.9 (1-C), 128.2 (Ar-C), 128.1 (6-C or 7-C), 128.0 (Ar-C), 127.7 (Ar-C), 126.9 (6-C or 7-C), 124.3 (Ar-C), 124.0 (5'-C);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3200, 3054, 1645, 1604, 1589, 1242, 1114, 740;  $m/z$  (ES) 276.1 (100%, MH $^+$ ); (Found MH $^+$ , 276.1124. C $_{17}$ H $_{13}$ N $_3$ O $_3$  requires *MH*, 276.1131).

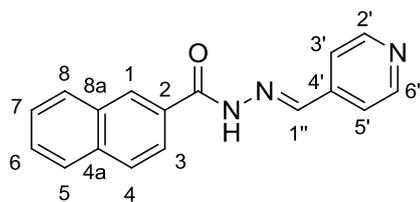
**(E)-N'-(pyridin-2-yl-methylene)-2-naphthohydrazide (6.49)<sup>18</sup>**



Prepared *via* general method D using 2-naphthohydrazide **6.105** (186 mg, 1.0 mmol) and 2-pyridinecarboxyaldehyde **6.117** (0.094 mL, 1.0 mmol). Recrystallization from methanol /

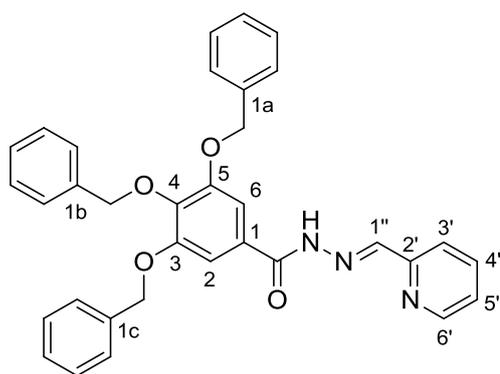
water gave the title compound **6.49** (245 mg, 0.89 mmol, 89%) as colourless needles m.p. 186—187 °C (Lit.<sup>18</sup> not available).  $R_f$  0.45 (95:5 CH $_2$ Cl $_2$ —MeOH); HPLC (Method B), R.t. 2.58 min, (100%);  $\delta_{\text{H}}$  (300 MHz, DMSO- $d_6$ ); 12.27 (1H, s, N-H), 8.64 (1H, app. d,  $J = 4.5$ , Ar-H $_6$ '), 8.58 (1H, app. s, Ar-H $_1$ ), 8.54 (1H, s, C-H $_1$ '), 7.99-8.12 (5H, m, Ar-H $_{4,5,8,3'}$ ), 7.91 (1H, app. t,  $J = 7.5$ , Ar-H $_4$ '), 7.52-7.74 (2H, m, Ar-H $_{6,7}$ ), 7.40-7.49 (1H, m, Ar-H $_5$ ');  $\delta_{\text{C}}$  (75 MHz, DMSO- $d_6$ ); 163.4 (C=O), 153.2 (2'-C), 149.5 (6'-C), 148.0 (1''-C), 136.9 (4'-C), 134.4 (Ar-C $_{\text{qt}}$ ), 132.0 (2-C), 130.4 (Ar-C $_{\text{qt}}$ ), 128.9 (1-C), 128.20 (Ar-C), 128.17 (Ar-C), 128.0 (Ar-C), 127.7 (6-C or 7-C), 127.0 (6-C or 7-C), 124.4 (5'-C), 124.3 (Ar-C), 119.9 (Ar-C);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3186, 3002, 1684, 1585, 1504, 1484, 1284, 1136;  $m/z$  (ES) 298.1 (100%, MNa $^+$ ); (Found MNa $^+$ , 298.0943. C $_{17}$ H $_{13}$ N $_3$ O requires *MNa*, 298.0951).

**(E)-N'-(pyridin-4-yl-methylene)-2-naphthohydrazide (6.50)**<sup>18</sup>



Prepared *via* general method D using 2-naphthohydrazide **6.105** (186 mg, 1.0 mmol) and 4-pyridinecarboxyaldehyde **6.118** (0.094 mL, 1.0 mmol). Recrystallization from methanol / water gave the title compound **6.50** (217 mg, 0.79 mmol, 79%) as colourless needles m.p. 224—225 °C (Lit.<sup>18</sup> not available).  $R_f$  0.46 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.00 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 12.34 (1H, s, N-H), 8.68 (1H, d,  $J = 4.4$ , Ar-H<sub>2',6'</sub>), 8.57 (1H, app. s, Ar-H<sub>1</sub>), 8.50 (1H, s, C-H<sub>1''</sub>), 7.96-8.14 (4H, m, Ar-H<sub>3,4,5,8</sub>), 7.71 (1H, d,  $J = 4.4$ , Ar-H<sub>3',5'</sub>), 7.59-7.68 (2H, m, Ar-H<sub>6,7</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 163.4 (C=O), 150.3 (2'-C and 6'-C), 145.3 (1''-C), 141.5 (4'-C), 134.4 (Ar-C<sub>qt</sub>), 132.0 (2-C), 130.3 (Ar-C<sub>qt</sub>), 128.9 (1-C), 128.23 (Ar-C), 128.21 (6-C or 7-C), 128.0 (Ar-C), 127.7 (Ar-C), 127.0 (6-C or 7-C), 124.3 (Ar-C), 121.0 (3'-C and 5'-C);  $\nu_{max}/cm^{-1}$  (neat); 3402, 3197, 3003, 1644, 1541, 1376, 965, 750;  $m/z$  (ES) 298.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 298.0946. C<sub>17</sub>H<sub>13</sub>N<sub>3</sub>O requires MNa, 298.0951).

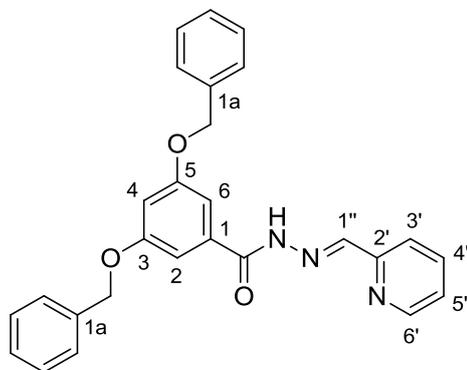
**(E)-3,4,5-tris(benzyloxy)-N'-(pyridin-2-yl methylene)benzohydrazide (6.51)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (241 mg, 0.53 mmol) and 2-pyridinecarboxyaldehyde **6.117** (0.05 mL, 0.53 mmol). Recrystallization from ethanol gave the title compound **6.51** (261 mg, 0.48 mmol, 91%) as colourless plates m.p. 166—167 °C.  $R_f$  0.60 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.82 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.95 (1H, s, N-H), 8.64 (1H, app. d,  $J = 4.6$ , Ar-H<sub>6'</sub>), 8.52 (1H, s, C-H<sub>1''</sub>), 8.00 (1H, app. d,  $J = 7.7$ , Ar-H<sub>3'</sub>), 7.90 (1H, app. t,  $J = 7.7$ , Ar-H<sub>4'</sub>), 7.21-7.56 (18H, m, Ar-H), 5.21 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 153.2 (2'-C), 152.1 (3-C and 5-C), 149.6 (6'-C), 147.9 (1''-C), 140.2 (4-C), 137.4 (1b-C), 136.9 (4'-C), 136.7 (1a-C and 1c-C), 128.4 (Ar-C), 128.3 (1-C), 128.2 (Ar-C), 128.1 (Ar-C),

128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 124.4 (5'-C), 119.9 (3'-C), 107.0 (2-C and 6-C), 74.2 (OCH<sub>2</sub> in 4), 70.4 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3209, 3061, 1644, 1580, 1547, 1501, 1433, 1326, 1108, 655;  $m/z$  (ES) 566.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 566.2055. C<sub>34</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires MNa, 544.2050).

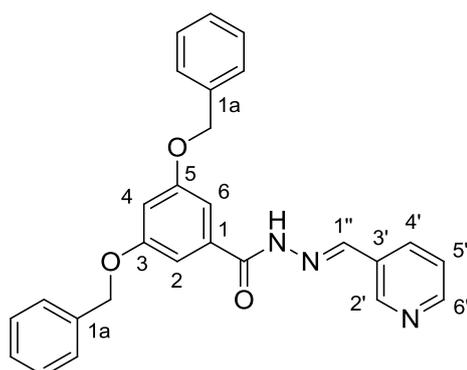
**(E)-3,5-bis(benzyloxy)-N'-(pyridin-2-yl-methylene)benzohydrazide (6.52)**



Prepared *via* general method D using 3,5-bis(benzyloxy)benzohydrazide **6.109** (174 mg, 0.50 mmol) and 2-pyridinecarboxyaldehyde **6.117** (0.047 mL, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.52** (197 mg, 0.45 mmol, 90%) as light yellow needles m.p.

73—74 °C.  $R_f$  0.34 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method C), R.t. 1.95 min, (96%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.94 (1H, s, N-H), 8.61 (1H, app. d,  $J$  = 4.6, Ar-H<sub>6</sub>), 8.48 (1H, s, C-H<sub>1''</sub>), 7.97 (1H, app. d,  $J$  = 7.7, Ar-H<sub>3</sub>), 7.88 (1H, app. t,  $J$  = 7.7, Ar-H<sub>4</sub>), 7.30-7.50 (11H, m, Ar-H), 7.18 (2H, d,  $J$  = 2.0, Ar-H<sub>2,6</sub>), 6.93 (1H, t,  $J$  = 2.0, Ar-H<sub>4</sub>), 5.17 (4H, s, OCH<sub>2</sub> in 3 and 5);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.7 (C=O), 159.5 (3-C and 5-C), 153.2 (2'-C), 149.5 (6'-C), 148.2 (1''-C), 136.9 (4'-C), 136.7 (1aC), 135.1 (1-C), 128.5 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 124.4 (Ar-C), 119.9 (Ar-C), 106.8 (2-C and 6-C), 105.1 (4-C), 69.7 (OCH<sub>2</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3452, 3200, 3031, 1674, 1591, 1546, 1350, 1165, 735;  $m/z$  (ES) 460.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 460.1647. C<sub>27</sub>H<sub>23</sub>N<sub>3</sub>O<sub>3</sub> requires MNa, 460.1632).

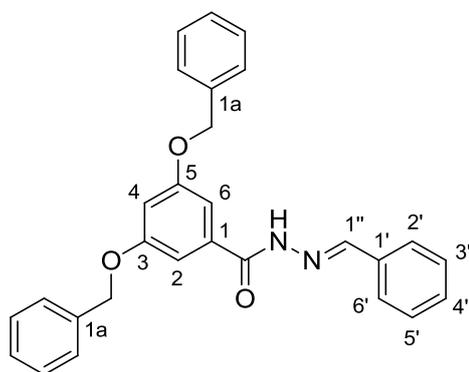
**(E)-3,5-bis(benzyloxy)-N'-(pyridin-3-yl-methylene)benzohydrazide (6.53)**



Prepared *via* general method D using 3,5-bis(benzyloxy)benzohydrazide **6.109** (174 mg, 0.50 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.047 mL, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.53** (184 mg, 0.42 mmol, 84%) as colourless needles m.p.

154—155 °C.  $R_f$  0.38 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method C), R.t. 1.86 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.91 (1H, s, N-H), 8.84 (1H, app. s, Ar-H<sub>2'</sub>), 8.61 (1H, dd,  $J = 4.6, 1.3$ , Ar-H<sub>6'</sub>), 8.50 (1H, s, C-H<sub>1''</sub>), 8.13 (1H, app. d,  $J = 7.9$ , Ar-H<sub>4'</sub>), 7.30-7.52 (11H, m, Ar-H), 7.17 (2H, d,  $J = 2.0$ , Ar-H<sub>2,6</sub>), 6.92 (1H, t,  $J = 2.0$ , Ar-H<sub>4</sub>), 5.16 (4H, s, OCH<sub>2</sub> in 3 and 5);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 159.5 (3-C and 5-C), 150.7 (6'-C), 148.7 (2'-C), 145.2 (1''-C), 136.7 (1aC), 135.2 (1-C), 133.4 (4'-C), 130.2 (3'-C), 128.5 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 124.0 (Ar-C), 106.8 (2-C and 6-C), 105.1 (4-C), 69.6 (OCH<sub>2</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3191, 3029, 1651, 1590, 1547, 1416, 1299, 1158, 694;  $m/z$  (ES) 460.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 460.1636. C<sub>27</sub>H<sub>23</sub>N<sub>3</sub>O<sub>3</sub> requires MNa, 460.1632).

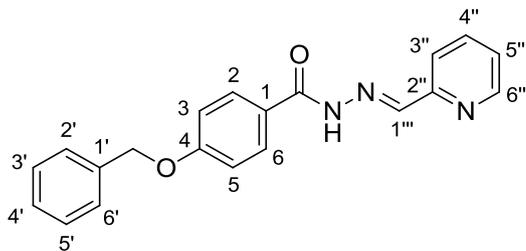
### (E)-N'-benzylidene-3,5-bis(benzyloxy)benzohydrazide (**6.54**)



Prepared *via* general method D using 3,5-bis(benzyloxy)benzohydrazide **6.109** (174 mg, 0.50 mmol) and benzaldehyde **6.119** (0.05 mL, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.54** (166 mg, 0.38 mmol, 76%) as colourless needles m.p. 100—101 °C.  $R_f$  0.52 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—

MeOH); HPLC (Method C), R.t. 2.54 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.74 (1H, s, N-H), 8.45 (1H, s, C-H<sub>1''</sub>), 7.67-7.76 (2H, m, Ar-H), 7.30-7.50 (13H, m, Ar-H), 7.17 (2H, d,  $J = 2.1$ , Ar-H<sub>2,6</sub>), 6.91 (1H, t,  $J = 2.1$ , Ar-H<sub>4</sub>), 5.16 (4H, s, OCH<sub>2</sub> in 3 and 5);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.5 (C=O), 159.5 (3-C and 5-C), 148.0 (1''-C), 136.7 (1a-C), 135.4 (1-C), 134.3 (1'-C), 130.1 (Ar-C), 128.8 (Ar-C), 128.4 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 127.1 (Ar-C), 106.7 (2-C and 6-C), 105.0 (4-C), 69.6 (OCH<sub>2</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3233, 3030, 1651, 1590, 1537, 1348, 1164, 691;  $m/z$  (ES) 459.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 459.1697. C<sub>28</sub>H<sub>24</sub>N<sub>2</sub>O<sub>3</sub> requires MNa, 459.1679).

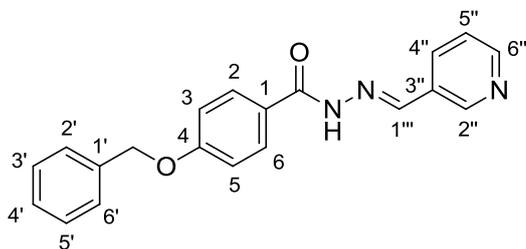
**(E)-4-(benzyloxy)-N'-(pyridin-2-yl-methylene)benzohydrazide (6.55)**



Prepared *via* general method D using 4-(benzyloxy)benzohydrazide **6.106** (121 mg, 0.5 mmol) and 2-pyridinecarboxyaldehyde **6.117** (0.047 mL, 0.5 mmol). Recrystallization from

methanol / water gave the title compound **6.55** (292 mg, 0.9 mmol, 88%) as colourless plates m.p. 161—162 °C.  $R_f$  0.58 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.36 min, (98%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.95 (1H, s, N-H), 8.62 (1H, app. d,  $J = 4.8$ , Ar-H<sub>6''</sub>), 7.81-8.00 (4H, m, Ar-H), 8.47 (1H, s, C-H<sub>1'''</sub>), 7.27-7.53 (6H, m, Ar-H), 7.16 (2H, d,  $J = 8.8$ , Ar-H<sub>3,5</sub>), 5.21 (2H, s, OCH<sub>2</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 161.2 (4-C), 153.3 (2''-C), 149.5 (6''-C), 147.4 (1'''-C), 136.8 (Ar-C), 136.6 (1'-C), 129.6 (Ar-C), 128.5 (Ar-C), 128.0 (Ar-C), 127.8 (Ar-C), 125.3 (1-C), 124.3 (Ar-C), 119.8 (Ar-C), 114.6 (3-C, 5-C), 69.4 (OCH<sub>2</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3454, 3180, 3064, 1653, 1602, 1471, 1281, 1247, 1021, 738;  $m/z$  (ES) 354.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 354.1219. C<sub>20</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub> requires MNa, 354.1213).

**(E)-4-(benzyloxy)-N'-(pyridin-3-yl-methylene)benzohydrazide (6.56)**

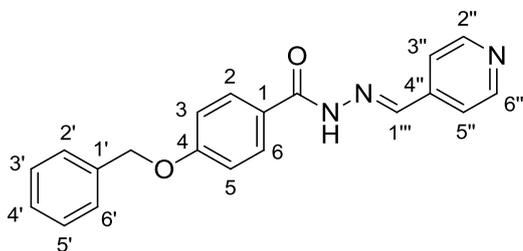


Prepared *via* general method D using 4-(benzyloxy)benzohydrazide **6.106** (121 mg, 0.5 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.047 mL, 0.5 mmol). Recrystallization from

ethanol gave the title compound **6.56** (131 mg, 0.4 mmol, 79%) as colourless plates m.p. 210—211 °C.  $R_f$  0.42 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.34 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.93 (1H, s, N-H), 8.85 (1H, app. s, Ar-H<sub>2''</sub>), 8.61 (1H, app. d,  $J = 4.2$ , Ar-H<sub>6''</sub>), 8.49 (1H, s, C-H<sub>1'''</sub>), 8.14 (1H, app. d,  $J = 7.5$ , Ar-H<sub>4''</sub>), 7.92 (2H, d,  $J = 8.8$ , Ar-H<sub>2,6</sub>), 7.27-7.55 (6H, m, Ar-H), 7.16 (2H, d,  $J = 8.8$ , Ar-H<sub>3,5</sub>), 5.21 (2H, s, OCH<sub>2</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 161.1 (4-C), 150.6 (6''-C), 148.6 (2''-C), 144.3 (1'''-C), 136.6 (1'-C), 133.3 (4''-C), 130.3 (3''-C), 129.6 (2-C, 6-C), 128.5 (Ar-C), 128.0 (Ar-C), 127.8 (Ar-C),

125.4 (1-C), 124.0 (Ar-C), 114.6 (3-C, 5-C), 69.4 (OCH<sub>2</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3410, 3190, 3010, 1660, 1605, 1508, 1216, 1177, 844;  $m/z$  (ES) 354.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 354.1215. C<sub>20</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub> requires MNa, 354.1213).

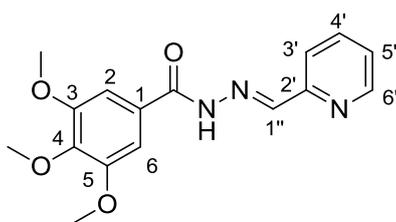
### (E)-4-(benzyloxy)-N'-(pyridin-4-ylmethylene)benzohydrazide (6.57)



Prepared *via* general method D using 4-(benzyloxy)benzohydrazide **6.106** (121 mg, 0.5 mmol) and 4-pyridinecarboxyaldehyde **6.118** (0.047 mL, 0.5 mmol). Recrystallization from

methanol / water gave the title compound **6.57** (272 mg, 0.8 mmol, 82%) as colourless plates m.p. 203—204 °C.  $R_f$  0.39 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.39 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.97 (1H, s, N-H), 8.65 (2H, d,  $J = 4.5$ , Ar-H<sub>2'',6''</sub>), 8.43 (1H, s, C-H<sub>1'''</sub>), 7.92 (2H, d,  $J = 8.9$ , Ar-H<sub>2,6</sub>), 7.66 (2H, d,  $J = 4.5$ , Ar-H<sub>3',5'</sub>), 7.31-7.53 (5H, m, Ar-H<sub>2',3',4',5',6'</sub>), 7.16 (2H, d,  $J = 8.9$ , Ar-H<sub>3,5</sub>), 5.20 (2H, s, OCH<sub>2</sub>);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.8 (C=O), 161.3 (4-C), 150.2 (2''-C and 6''-C), 144.6 (1''-C), 141.6 (4'-C), 136.6 (1'-C), 129.7 (2-C, 6-C), 128.5 (Ar-C), 128.0 (Ar-C), 127.8 (Ar-C), 125.3 (1-C), 120.9 (3''-C and 5''-C), 114.6 (3-C, 5-C), 69.4 (OCH<sub>2</sub>);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3253, 3035, 1645, 1602, 1487, 1247, 752;  $m/z$  (ES) 332.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 332.1386. C<sub>20</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub> requires MH, 332.1394).

### (E)-3,4,5-trimethoxy-N'-(pyridin-2-yl-methylene)benzohydrazide (6.58)

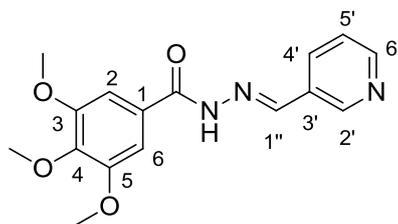


Prepared *via* general method D using 3,4,5-trimethoxybenzohydrazide **6.113** (679 mg, 3.0 mmol) and 2-pyridinecarboxyaldehyde **6.117** (0.360 mL, 3.0 mmol). Recrystallization from ethanol gave the title compound **6.58** (605 mg,

1.9 mmol, 64%) as colourless plates m.p. 193—194 °C.  $R_f$  0.68 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); (Found: C, 60.7; H, 5.45; N, 13.5; C<sub>16</sub>H<sub>17</sub>N<sub>3</sub>O<sub>4</sub> requires C, 60.9; H, 5.43; N, 13.3%). HPLC (Method B), R.t. 1.77 min, (96%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.89 (1H, br s, N-H), 8.62 (1H, app. d,  $J = 4.7$ , Ar-H<sub>6'</sub>), 8.49 (1H, s, C-H<sub>1''</sub>),

7.98 (1H, app. d,  $J = 7.8$ , Ar-H<sub>3</sub>), 7.88 (1H, app. t,  $J = 7.8$ , Ar-H<sub>4</sub>), 7.40-7.43 (1H, m, Ar-H<sub>5</sub>), 7.25 (2H, s, Ar-H<sub>2,6</sub>), 3.86 (6H, s, OCH<sub>3</sub> in 3 and 5), 3.73 (3H, s, OCH<sub>3</sub> in 4);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 153.3 (3-C, 5-C), 152.7 (2'-C), 149.6 (1''-C), 147.9 (6'-C), 140.6 (4-C), 136.9 (4'-C), 128.2 (1-C), 124.4 (5'-C), 119.9 (3'-C), 105.4 (2-C, 6-C), 60.1 (OCH<sub>3</sub> in 4), 56.1 (OCH<sub>3</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 2940, 1860, 1680, 1582, 1548, 1499, 1323, 1215;  $m/z$  (ES) 338.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 338.1104. C<sub>16</sub>H<sub>17</sub>N<sub>3</sub>O<sub>4</sub> requires MNa, 338.1111).

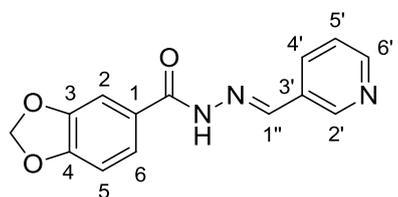
**(E)-3,4,5-trimethoxy-N'-(pyridin-3-ylmethylene)benzohydrazide (6.59)**



Prepared *via* general method D using 3,4,5-trimethoxybenzohydrazide **6.113** (679 mg, 3.0 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.360 mL, 3.0 mmol). Recrystallization from ethanol gave the title compound **6.59** (681 mg,

2.2 mmol, 72%) as colourless plates m.p. 195—196 °C.  $R_f$  0.52 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 1.50 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.90 (1H, br s, N-H), 8.87 (1H, app. s, Ar-H<sub>2'</sub>), 8.63 (1H, dd,  $J = 4.7, 1.3$ , Ar-H<sub>6'</sub>), 8.53 (1H, s, C-H<sub>1''</sub>), 8.16 (1H, app. d,  $J = 7.9$ , Ar-H<sub>4'</sub>), 7.51 (1H, dd,  $J = 7.9, 4.7$ , Ar-H<sub>5'</sub>), 7.25 (2H, s, Ar-H<sub>2,6</sub>), 3.87 (6H, s, OCH<sub>3</sub> in 3,5), 3.73 (3H, s, OCH<sub>3</sub> in 4);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 152.7 (3-C, 5-C), 150.7 (6'-C), 148.7 (2'-C), 144.9 (1''-C), 140.5 (4-C), 133.4 (4'-C), 130.2 (3'-C), 128.3 (1-C), 124.0 (5'-C), 105.2 (2-C, 6-C), 60.1 (OCH<sub>3</sub> in 4), 56.1 (OCH<sub>3</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3472, 3212, 2975, 1992, 1653, 1584, 1560, 1504, 1466, 1414;  $m/z$  (ES) 316.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 316.1289. C<sub>16</sub>H<sub>17</sub>N<sub>3</sub>O<sub>4</sub> requires MH, 316.1292).

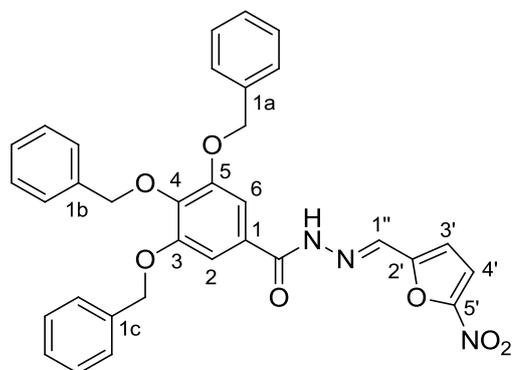
**(E)-N'-(pyridin-3-ylmethylene)benzo[d][1,3]dioxole-5-carbohydrazide (6.60)**



Prepared *via* general method D using 3,4-methylenedioxybenzhydrazide **6.114** (180 mg, 1.0 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.120 mL, 1.0 mmol). Recrystallization from ethanol

gave the title compound **6.60** (215 mg, 0.8 mmol, 80%) as colourless plates m.p. 196—197 °C.  $R_f$  0.43 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 1.42 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.84 (1H, s, N-H), 8.84 (1H, app. s, Ar-H<sub>2'</sub>), 8.60 (1H, dd,  $J = 4.7, 1.4$ , Ar-H<sub>6'</sub>), 8.47 (1H, s, C-H<sub>1''</sub>), 8.12 (1H, app. d,  $J = 7.9$ , Ar-H<sub>4'</sub>), 7.53 (1H, dd,  $J = 8.1, 1.4$ , Ar-H<sub>6</sub>), 7.48 (1H, dd,  $J = 7.9, 4.7$ , Ar-H<sub>5'</sub>), 7.45 (1H, app. s, Ar-H<sub>2</sub>), 7.05 (1H, d,  $J = 8.1$ , Ar-H<sub>5</sub>), 6.12 (2H, s, OCH<sub>2</sub>O);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.2 (C=O), 150.6 (6'-C), 150.3 (3-C or 4-C), 148.6 (2'-C), 147.4 (3-C or 4-C), 144.6 (1''-C), 133.4 (4'-C), 130.3 (3'-C), 126.9 (1-C), 124.0 (5'-C), 123.0 (6-C), 108.1 (5-C), 107.6 (2-C), 101.8 (OCH<sub>2</sub>O);  $\nu_{max}/cm^{-1}$  (neat); 3448, 3373, 3198, 3031, 1759, 1660, 1637, 1600, 1570, 1485, 1309, 1262;  $m/z$  (ES) 270.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 270.0865. C<sub>14</sub>H<sub>11</sub>N<sub>3</sub>O<sub>3</sub> requires *MH*, 270.0863).

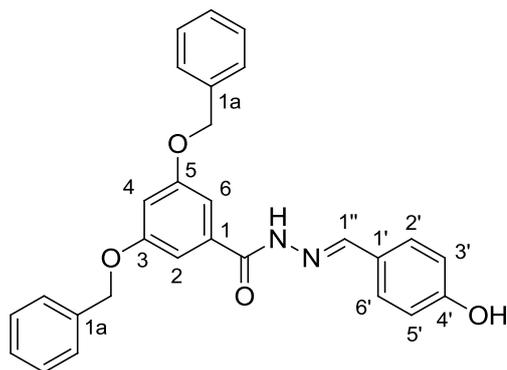
**(E)-3,4,5-tris(benzyloxy)-N'-((5-nitrofuran-2-yl)methylene)benzohydrazide (6.61)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (150 mg, 0.33 mmol) and 5-nitrofuran-2-carbaldehyde **6.134** (46.6 mg, 0.33 mmol). Recrystallization from ethanol / water gave the title compound **6.61** (166 mg, 0.29 mmol, 87%) as light yellow

plates m.p. 229—230 °C.  $R_f$  0.56 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); (Found: C, 68.9; H, 4.75; N, 7.2; C<sub>33</sub>H<sub>27</sub>N<sub>3</sub>O<sub>7</sub> requires C, 68.6; H, 4.71; N, 7.3%). HPLC (Method B), R.t. 4.13 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 12.13 (1H, s, N-H), 8.44 (1H, s, C-H<sub>1''</sub>), 7.83 (1H, d,  $J = 3.9$ , Ar-H<sub>4'</sub>), 7.46-7.53 (4H, m, Ar-H), 7.33-7.45 (10H, m, Ar-H), 7.25-7.31 (4H, m, Ar-H), 5.22 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.04 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.8 (C=O), 152.1 (3-C and 5-C), 151.9 (5'-C or 2'-C), 151.7 (5'-C or 2'-C), 140.5 (4-C), 137.3 (1b-C), 136.7 (1a-C and 1c-C), 135.5 (1''-C), 128.44 (Ar-C), 128.40 (1-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 115.4 (4'-C), 114.6 (3'-C), 107.2 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.6 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3140, 3029, 1647, 1584, 1558, 1478, 1331, 1124, 732;  $m/z$  (ES) 600.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 600.1744. C<sub>33</sub>H<sub>27</sub>N<sub>3</sub>O<sub>7</sub> requires *MNa*, 600.1741).

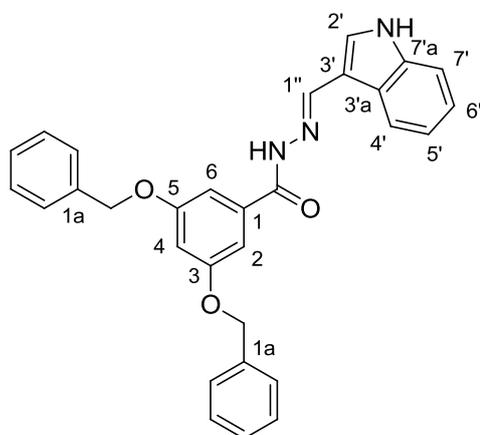
**(E)-3,5-bis(benzyloxy)-N'-(4-hydroxybenzylidene)benzohydrazide (6.62)**



Prepared *via* general method D using 3,5-bis(benzyloxy)benzohydrazide **6.109** (174 mg, 0.50 mmol) and 4-hydroxybenzaldehyde **6.120** (61 mg, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.62** (174 mg, 0.38 mmol, 77%) as light yellow plates m.p. 171—172 °C.  $R_f$  0.77 (100% EtOAc);

HPLC (Method B), R.t. 3.51 min, (100%);  $\delta_H$  (500 MHz, DMSO- $d_6$ ); 11.54 (1H, s, N-H), 9.91 (1H, br s, O-H in 4'), 8.34 (1H, s, C-H $_{1''}$ ), 7.54 (2H, d,  $J$  = 8.6, Ar-H $_{2',6'}$ ), 7.30-7.49 (10H, m, Ar-H), 7.15 (2H, d,  $J$  = 2.1, Ar-H $_{2,6}$ ), 6.89 (1H, t,  $J$  = 2.1, Ar-H $_4$ ), 6.83 (2H, d,  $J$  = 8.6, Ar-H $_{3',5'}$ ), 5.15 (4H, s, OCH $_2$  in 3 and 5);  $\delta_C$  (125 MHz, DMSO- $d_6$ ); 172.0 (C=O), 162.2 (4'-C), 159.4 (3-C and 5-C), 148.3 (1''-C), 136.7 (1a-C), 135.6 (1-C), 128.8 (2'-C, 6'-C), 128.4 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 125.2 (1'-C), 115.7 (3'-C, 5'-C), 106.6 (2-C and 6-C), 104.8 (4-C), 69.6 (OCH $_2$ );  $\nu_{max}/cm^{-1}$  (neat); 3649, 3231, 3063, 1649, 1589, 1495, 1438, 1275, 1149, 694;  $m/z$  (ES) 475.2 (100%, MNa $^+$ ); (Found MNa $^+$ , 475.1629. C $_{28}$ H $_{24}$ N $_2$ O $_4$  requires MNa, 475.1628).

**(E)-N'-((1H-indol-3-yl)methylene)-3,5-bis(benzyloxy)benzohydrazide (6.63)**

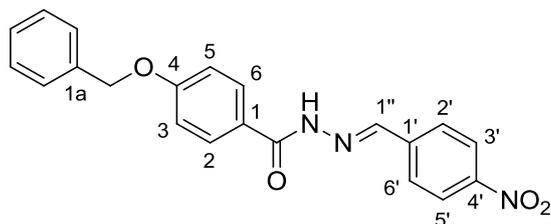


Prepared *via* general method D using 3,5-bis(benzyloxy)benzohydrazide **6.109** (174 mg, 0.50 mmol) and indole-3-carboxyaldehyde **6.132** (72 mg, 0.50 mmol). Recrystallization from ethanol / water gave the title compound **6.63** (178 mg, 0.38 mmol, 75%) as colourless plates m.p. 84 °C (dec).  $R_f$  0.41 (95:5 CH $_2$ Cl $_2$ —MeOH); HPLC (Method B), R.t. 3.72 min, (100%);  $\delta_H$  (500

MHz, DMSO- $d_6$ ); 11.57 (1H, s, N-H *indole*), 11.41 (1H, s, N-H *hydrazide*), 8.61 (1H, s, C-H $_{1''}$ ), 8.28 (1H, app. d,  $J$  = 7.9, Ar-H $_4$ ), 7.80 (1H, d,  $J$  = 2.8, Ar-H $_{2'}$ ),

7.27-7.50 (12H, m, Ar-H), 7.11-7.23 (3H, m, Ar-H), 6.88 (1H, t,  $J = 2.2$ , Ar-H<sub>4</sub>), 5.16 (4H, s, OCH<sub>2</sub> in 3 and 5);  $\delta_{\text{C}}$  (125 MHz, DMSO-*d*<sub>6</sub>); 161.8 (C=O), 159.4 (3-C and 5-C), 145.1 (1''-C), 137.0 (7'a-C), 136.8 (1a-C), 136.1 (3'a-C), 130.3 (2'-C), 128.5 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 124.3 (1-C), 122.6 (4'-C), 122.0 (Ar-C), 120.4 (Ar-C), 111.8 (Ar-C), 111.7 (3'-C), 106.6 (Ar-C), 104.6 (4-C), 69.6 (OCH<sub>2</sub> in 3 and 5);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3408, 3224, 3033, 1643, 1589, 1300, 1154, 741;  $m/z$  (ES) 498.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 498.1798. C<sub>30</sub>H<sub>25</sub>N<sub>3</sub>O<sub>3</sub> requires MNa, 498.1788).

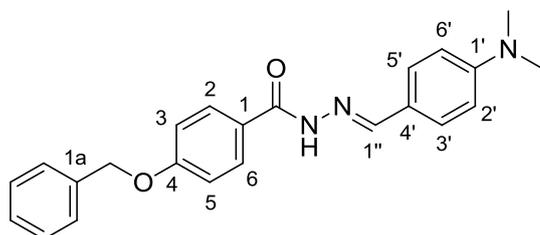
**(E)-4-(benzyloxy)-N'-(4-nitrobenzylidene)benzohydrazide (6.64)**<sup>14</sup>



Prepared *via* general method D using 4-(benzyloxy)benzohydrazide **6.106** (121 mg, 0.50 mmol) and 4-nitrobenzaldehyde **6.131** (75.5 mg, 0.50 mmol). Recrystallization from

ethanol gave the title compound **6.64** (69.5 mg, 0.46 mmol, 92%) as yellow needles m.p. 232—233 °C (Lit.<sup>14</sup> not available).  $R_f$  0.28 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.30 min, (100%);  $\delta_{\text{H}}$  (300 MHz, DMSO-*d*<sub>6</sub>); 12.06 (1H, s, N-H), 8.53 (1H, s, C-H<sub>1''</sub>), 8.31 (2H, d,  $J = 8.3$ , Ar-H<sub>3',5'</sub>), 7.99 (2H, d,  $J = 8.3$ , Ar-H<sub>2',6'</sub>), 7.93 (2H, d,  $J = 8.8$ , Ar-H<sub>2,6</sub>), 7.30-7.52 (5H, m, Ar-H), 7.17 (2H, d,  $J = 8.8$ , Ar-H<sub>3,5</sub>), 5.21 (2H, s, OCH<sub>2</sub>);  $\delta_{\text{C}}$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.7 (C=O), 161.2 (4-C), 147.8 (4'-C), 145.3 (1''-C), 140.8 (1'-C), 136.6 (1a-C), 129.7 (2-C, 6-C), 128.7 (2'-C, 6'-C), 128.5 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 125.3 (1-C), 124.1 (3'-C, 5'-C), 114.6 (3-C, 5-C), 69.4 (OCH<sub>2</sub>);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3255, 3065, 1650, 1606, 1587, 1538, 1463, 1316, 1144, 741;  $m/z$  (ES) 398.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 398.1110. C<sub>21</sub>H<sub>17</sub>N<sub>3</sub>O<sub>4</sub> requires MNa, 398.1111).

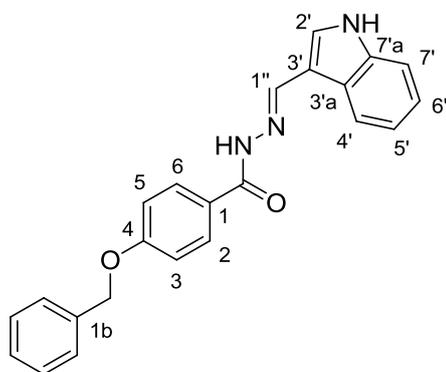
**(E)-4-(benzyloxy)-N'-(4-(dimethylamino)benzylidene)benzohydrazide (6.65)**



Prepared *via* general method D using 4-(benzyloxy)benzohydrazide **6.106** (121 mg, 0.50 mmol) and 4-(dimethylamino)benzaldehyde **6.130**

(74.6 mg, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.65** (149 mg, 0.40 mmol, 80%) as light yellow needles m.p. 247—248 °C.  $R_f$  0.57 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.22 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.45 (1H, s, N-H), 8.29 (1H, s, C-H<sub>1''</sub>), 7.88 (2H, d,  $J$  = 8.8, Ar-H<sub>2,6</sub>), 7.27-7.61 (7H, m, Ar-H), 7.13 (2H, d,  $J$  = 8.8, Ar-H<sub>3,5</sub>), 6.76 (2H, d,  $J$  = 8.8, Ar-H<sub>2',6'</sub>), 5.20 (2H, s, OCH<sub>2</sub>), 2.97 (6H, s, CH<sub>3</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 162.0 (C=O), 160.8 (4-C), 151.4 (1'-C), 148.0 (1''-C), 136.6 (1a-C), 129.3 (2-C and 6-C), 128.4 (Ar-C), 128.3 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 126.0 (1-C), 121.7 (4'-C), 114.4 (3-C, 5-C), 111.8 (2'-C and 6'-C), 69.3 (OCH<sub>2</sub>), 39.8 (NCH<sub>3</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3189, 3031, 1634, 1606, 1593, 1454, 1359, 1248, 1183, 658;  $m/z$  (ES) 396.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 396.1681. C<sub>23</sub>H<sub>23</sub>N<sub>3</sub>O<sub>2</sub> requires MNa, 396.1682).

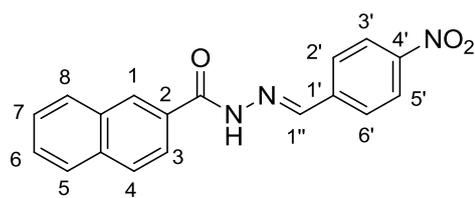
**(E)-N'-((1H-indol-3-yl)methylene)-4-(benzyloxy)benzohydrazide (6.66)**



Prepared *via* general method D using 4-(benzyloxy)benzohydrazide **6.106** (121 mg, 0.50 mmol) and indole-3-carboxyaldehyde **6.132** (72 mg, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.66** (151 mg, 0.41 mmol, 82%) as colourless plates m.p. 216—217 °C.  $R_f$  0.64 (92:8

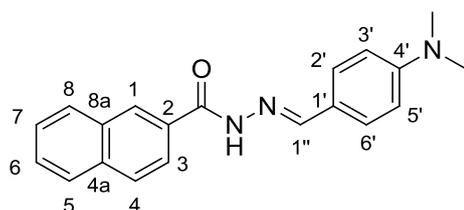
CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.07 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.59 (1H, s, N-H *indole*), 11.43 (1H, s, N-H *hydrazide*), 8.60 (1H, s, 1''-C), 8.30 (1H, app. d,  $J$  = 7.6, Ar-H<sub>4'</sub>), 7.91 (2H, d,  $J$  = 8.8, Ar-H<sub>2,6</sub>), 7.82 (1H, d,  $J$  = 2.6, Ar-H<sub>2</sub>), 7.29-7.55 (6H, m, Ar-H), 7.11-7.23 (4H, m, Ar-H), 5.20 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.0 (C=O), 160.7 (4-C), 144.4 (1''-C), 137.0 (7'a), 136.7 (1b-C), 130.0 (2'-C), 129.3 (2-C and 6-C), 128.5 (Ar-C), 127.9 (Ar-C), 127.7 (4'-C), 126.3 (3'a), 124.4 (1-C), 122.6 (Ar-C), 122.0 (Ar-C), 120.3 (Ar-C), 114.5 (Ar-C), 111.8 (3'-C), 111.7 (Ar-C), 69.4 (OCH<sub>2</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3302, 3032, 1627, 1601, 1566, 1246, 1045, 744;  $m/z$  (ES) 392.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 392.1372. C<sub>23</sub>H<sub>19</sub>N<sub>3</sub>O<sub>2</sub> requires MNa, 392.1369).

### (E)-N'-(4-nitrobenzylidene)-2-naphthohydrazide (**6.67**)<sup>19</sup>



Prepared *via* general method D using 2-naphthohydrazide **6.105** (100 mg, 0.54 mmol) and 4-nitrobenzaldehyde **6.131** (82 mg, 0.54 mmol). Recrystallization from ethanol gave the title compound **6.67** (140 mg, 0.44 mmol, 88%) as colourless plates m.p. 242—243 °C (Lit.<sup>19</sup> m.p. 244—246 °C).  $R_f$  0.38 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.03 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 12.24 (1H, s, N-H), 8.60 (1H, app. s, Ar-H<sub>1</sub>), 8.56 (1H, s, C-H<sub>1'</sub>), 8.31 (2H, d,  $J$  = 8.2, Ar-H<sub>3',5'</sub>), 7.97-8.10 (6H, m, Ar-H<sub>3,4,5,8,2',6'</sub>), 7.58-7.69 (2H, m, Ar-H<sub>6,7</sub>);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 163.5 (C=O), 147.9 (4'-C), 145.3 (1''-C), 140.6 (1'-C), 134.4 (Ar-C<sub>qt</sub>), 132.0 (2-C), 130.3 (Ar-C<sub>qt</sub>), 129.0 (Ar-C), 128.23 (Ar-C), 128.22 (Ar-C), 128.0 (6-C or 7-C), 127.7 (Ar-C), 127.0 (6-C or 7-C), 124.27 (2'-C, 6'-C), 124.26 (3'-C, 5'-C), 124.1 (Ar-C);  $\nu_{max}/cm^{-1}$  (neat); 3491, 3423, 3047, 1645, 1615, 1597, 1291, 1233, 696;  $m/z$  (ES) 342.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 342.0848. C<sub>18</sub>H<sub>13</sub>N<sub>3</sub>O<sub>3</sub> requires MNa, 342.0849).

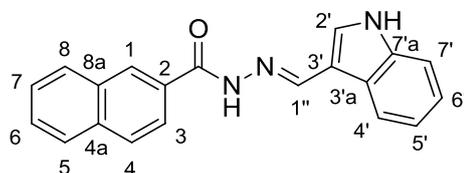
### (E)-N'-(4-(dimethylamino)benzylidene)-2-naphthohydrazide (**6.68**)



Prepared *via* general method D using 2-naphthohydrazide **6.105** (93 mg, 0.50 mmol) and 4-(dimethylamino)benzaldehyde **6.130** (75 mg, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.68** (127 mg, 0.40 mmol, 80%) as light yellow plates m.p. 212—213 °C.  $R_f$  0.34 (5:95 EtOAc—petroleum Ether); HPLC (Method B), R.t. 2.95 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.70 (1H, s, N-H), 8.50 (1H, app. s, Ar-H<sub>1</sub>), 8.36 (1H, s, C-H<sub>1'</sub>), 7.92-8.09 (4H, m, Ar-H<sub>3,4,5,8</sub>), 7.58-7.66 (2H, m, Ar-H<sub>6,7</sub>), 7.56 (2H, d,  $J$  = 8.8, Ar-H<sub>2',6'</sub>), 6.77 (2H, d,  $J$  = 8.8, Ar-H<sub>3',5'</sub>), 2.97 (6H, s, CH<sub>3</sub>);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 165.7 (C=O), 151.5 (4'-C), 148.7 (1''-C), 134.2 (Ar-C<sub>qt</sub>), 132.1 (2-C), 131.1 (Ar-C<sub>qt</sub>), 128.83 (Ar-C), 128.82 (Ar-C), 128.4 (2'-C, 6'-C), 128.0 (Ar-C), 127.7 (Ar-C), 127.6 (Ar-C), 126.8 (Ar-C), 124.3 (Ar-C), 121.6 (1'-C), 111.8 (3'-C, 5'-C), 40.3 (NCH<sub>3</sub>);  $\nu_{max}/cm^{-1}$  (neat); 3340, 3213, 3046, 2962, 1640, 1627, 1590,

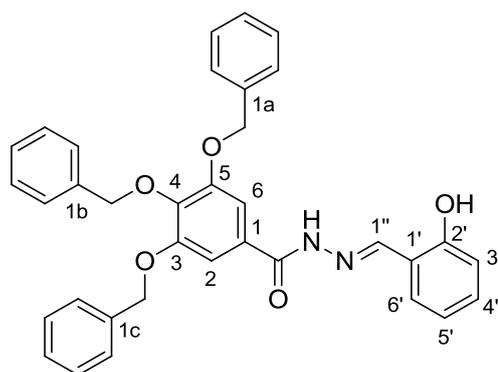
1550, 1323, 757;  $m/z$  (ES) 340.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 340.1423.  $C_{20}H_{19}N_3O$  requires  $MNa$ , 340.1420).

**(E)-N'-((1H-indol-3-yl)methylene)-2-naphthohydrazide (6.69)**



Prepared *via* general method D using 2-naphthohydrazide **6.105** (186 mg, 1.0 mmol) and indole-3-carboxyaldehyde **6.132** (145 mg, 1.0 mmol). Recrystallization from ethanol gave the title compound **6.69** (241 mg, 0.77 mmol, 77%) as colourless needles m.p. 238—239 °C.  $R_f$  0.47 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method B), R.t. 2.76 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 11.67 (1H, s, N-H *amide*), 11.58 (1H, s, N-H *indole*), 8.67 (1H, s, C-H $_{1''}$ ), 8.52 (1H, app. s, Ar-H $_1$ ), 8.33 (1H, app. d,  $J = 7.7$ , Ar-H $_{4'}$ ), 7.97-8.10 (4H, m, Ar-H), 7.83 (1H, d,  $J = 2.7$ , Ar-H $_2$ ), 7.57-7.68 (2H, m, Ar-H $_{6,7}$ ), 7.45 (1H, d,  $J = 7.9$ , Ar-H), 7.13-7.25 (2H, m, Ar-H *indole*);  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 162.5 (C=O), 145.0 (1''-C), 137.0 (7'a-C), 134.2 (2-C), 132.1 (Ar-C $_{qt}$ ), 131.9 (Ar-C $_{qt}$ ), 131.4 (3'a-C), 130.3 (2'-C), 128.8 (1-C), 128.0 (Ar-C), 127.7 (Ar-C), 127.6 (Ar-C), 127.5 (Ar-C), 126.8 (Ar-C), 124.4 (Ar-C), 122.6 (4'-C), 122.0 (Ar-C), 120.4 (Ar-C), 111.8 (Ar-C), 111.7 (3'-C);  $\nu_{max}/cm^{-1}$  (neat); 3312, 3175, 2930, 1633, 1618, 1570, 1547, 1415, 1361, 1247, 1146, 734;  $m/z$  (ES) 336.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 336.1119.  $C_{20}H_{15}N_3O$  requires  $MNa$ , 336.1107).

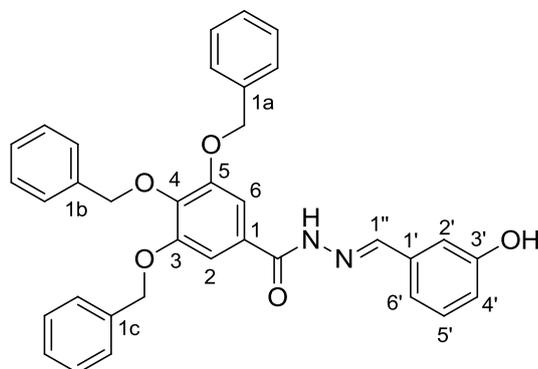
**(E)-3,4,5-tris(benzyloxy)-N'-(2-hydroxybenzylidene)benzohydrazide (6.70)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (180 mg, 0.40 mmol) and 2-hydroxybenzaldehyde **6.122** (0.03 mL, 0.40 mmol). Recrystallization from ethanol gave the title compound **6.70** (165 mg, 0.30 mmol, 74%) as creamy plates m.p. 170—171 °C.  $R_f$  0.39 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method B), R.t. 1.16 min, (98%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 11.94 (1H, s, N-H), 11.22 (1H, s, O-H), 8.66

(1H, s, C-H<sub>1''</sub>), 7.56 (1H, dd, *J* = 7.7, 1.5, Ar-H<sub>6'</sub>), 7.47-7.50 (4H, m, Ar-H), 7.23-7.44 (14H, m, Ar-H), 6.84-6.97 (2H, m, Ar-H *hydroxybenzyl*), 5.21 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4); δ<sub>C</sub> (125 MHz, DMSO-*d*<sub>6</sub>); 162.1 (C=O), 157.4 (2'-C), 152.1 (3-C and 5-C), 147.9 (1''-C), 140.3 (4-C), 137.4 (1b-C), 136.7 (1a-C and 1c-C), 131.4 (Ar-C), 129.2 (6'-C), 128.5 (1-C), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 119.3 (Ar-C *hydroxybenzyl*), 118.8 (1'-C), 116.4 (Ar-C *hydroxybenzyl*), 106.9 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3213, 3057, 1638, 1621, 1606, 1577, 1533, 1422, 1370, 1077; *m/z* (ES) 581.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 581.2049. C<sub>35</sub>H<sub>30</sub>N<sub>2</sub>O<sub>5</sub> requires *MNa*, 581.2047).

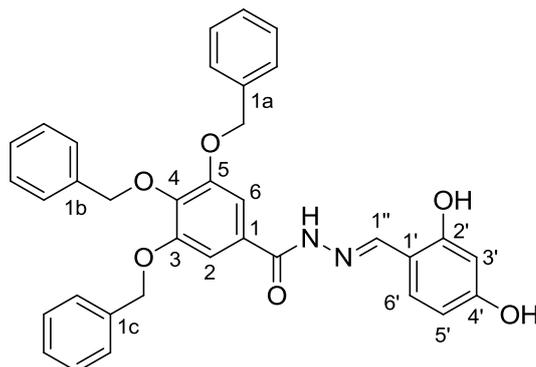
**(E)-3,4,5-tris(benzyloxy)-N'-(3-hydroxybenzylidene)benzohydrazide (6.71)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (127 mg, 0.28 mmol) and 3-hydroxybenzaldehyde **6.121** (34.1 mg, 0.28 mmol). Recrystallization from ethanol gave the title compound **6.71** (113 mg, 0.20 mmol, 72%) as

colourless needles m.p. 167—168 °C. *R*<sub>f</sub> 0.38 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 4.08 min, (100%); δ<sub>H</sub> (500 MHz, DMSO-*d*<sub>6</sub>); 11.64 (1H, s, N-H), 9.62 (1H, s, O-H), 8.38 (1H, s, C-H<sub>1''</sub>), 7.46-7.50 (4H, m, Ar-H), 7.32-7.44 (10H, m, Ar-H), 7.22-7.30 (4H, m, Ar-H), 7.20 (1H, app. s, Ar-H<sub>2'</sub>), 7.10 (1H, app. d, *J* = 7.4, Ar-H<sub>6'</sub>), 6.83 (1H, app. dd, *J* = 7.4, 1.8, Ar-H<sub>4'</sub>), 5.20 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.02 (2H, s, OCH<sub>2</sub> in 4); δ<sub>C</sub> (125 MHz, DMSO-*d*<sub>6</sub>); 162.3 (C=O), 157.7 (3'-C), 152.1 (3-C and 5-C), 147.8 (1''-C), 140.1 (4-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 135.6 (1'-C), 129.9 (5'-C), 128.6 (1-C), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 118.8 (6'-C), 117.4 (4'-C), 112.6 (2'-C), 106.9 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3669, 3627, 3211, 3032, 1626, 1605, 1554, 1497, 1451, 1336, 1221, 1118, 732; *m/z* (ES) 559.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 559.2237. C<sub>35</sub>H<sub>30</sub>N<sub>2</sub>O<sub>5</sub> requires *MH*, 559.2228).

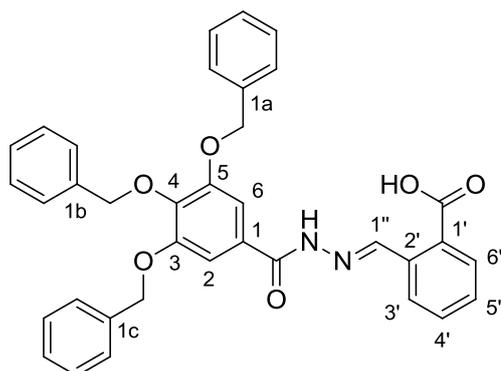
**(E)-3,4,5-tris(benzyloxy)-N'-(2,4-dihydroxybenzylidene)benzohydrazide (6.72)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (148 mg, 0.33 mmol) and 2,4-dihydroxybenzaldehyde **6.123** (45.1 mg, 0.33 mmol). A portion of the crude compound (60 mg out of 172mg) was purified *via* mass-directed preparative

HPLC eluting with a gradient of methanol in water (50—95%) in presence of formic acid (0.1%) to give a solid which was recrystallized from water to give the title compound **6.72** (45 mg, 0.078 mmol, 75%) as colourless plates m.p. 135—136 °C.  $R_f$  0.37 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 4.03 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.74 (1H, s, N-H), 11.39 (1H, s, O-H in 4'), 9.94 (1H, s, O-H in 2'), 8.51 (1H, s, C-H<sub>1''</sub>), 7.47-7.50 (4H, m, Ar-H), 7.24-7.44 (14H, m, Ar-H), 6.36 (1H, dd,  $J = 8.4, 2.3$ , Ar-H<sub>5'</sub>), 6.32 (1H, d,  $J = 2.3$ , Ar-H<sub>3'</sub>), 5.20 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.02 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 161.8 (C=O), 160.7 (4'-C), 159.4 (2'-C), 152.1 (3-C and 5-C), 148.9 (1''-C), 140.1 (4-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 131.1 (Ar-C), 128.7 (1-C), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 110.6 (1'-C), 107.7 (5'-C), 106.9 (2-C and 6-C), 102.6 (3'-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3578, 3422, 3210, 3231, 3031, 1623, 1602, 1581, 1280, 1184, 1099;  $m/z$  (ES) 575.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 575.2193. C<sub>35</sub>H<sub>30</sub>N<sub>2</sub>O<sub>6</sub> requires  $MH$ , 575.2178).

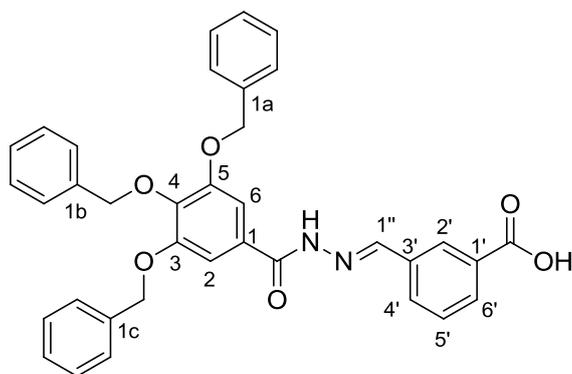
**(E)-2-((2-(3,4,5-tris(benzyloxy)benzoyl)hydrazono)methyl)benzoic acid (6.73)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (164 mg, 0.36 mmol) and 2-formylbenzoic acid **6.125** (54 mg, 0.36 mmol). Recrystallization from ethanol / water gave the title compound **6.73** (194 mg, 0.33 mmol, 92%) as colourless plates m.p.

208—209 °C.  $R_f$  0.29 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 4.00 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 13.39 (1H, br s, COOH), 11.97 (1H, s, N-H), 9.21 (1H, s, C-H<sub>1''</sub>), 8.08 (1H, app. d,  $J = 7.8$ , Ar-H<sub>6'</sub>), 7.92 (1H, app. d,  $J = 7.8$ , Ar-H<sub>3'</sub>), 7.67 (1H, app. t,  $J = 7.8$ , Ar-H<sub>4'</sub> or Ar-H<sub>5'</sub>), 7.23-7.60 (18H, m, Ar-H), 5.22 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 168.1 (C=O *carboxy*), 162.6 (C=O *hydrazono*), 152.1 (3-C, 5-C), 146.4 (1''-C), 140.2 (4-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 134.6 (2'-C), 132.0 (4'-C or 5'-C), 130.6 (1'-C), 130.3 (3'-C), 129.6 (Ar-C), 128.44 (Ar-C), 128.43 (1-C), 128.2 (Ar-C), 128.1 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 127.7 (Ar-C), 126.7 (6'-C), 107.1 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3133, 3026, 1681, 1628, 1582, 1496, 1315, 1226, 1124, 693;  $m/z$  (ES) 587.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 587.2169. C<sub>36</sub>H<sub>30</sub>N<sub>2</sub>O<sub>6</sub> requires *MH*, 587.2177).

**(E)-3-((2(3,4,5-tris(benzyloxy)benzoyl)hydrazono)methyl)benzoic acid (6.74)**

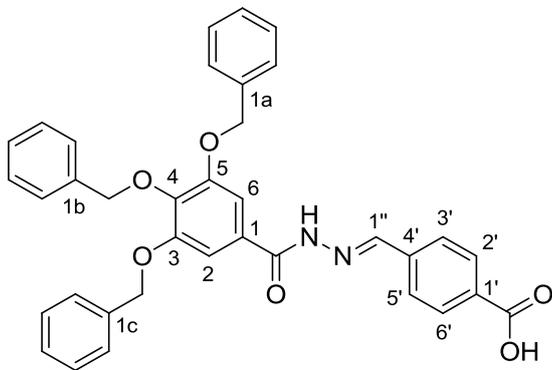


Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (227 mg, 0.50 mmol) and 3-formylbenzoic acid **6.124** (75 mg, 0.50 mmol). Recrystallization from ethanol / water gave the title compound **6.74** (249 mg, 0.42 mmol,

85%) as colourless fluffy solid m.p. 249—250 °C.  $R_f$  0.52 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.92 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 13.23 (1H, br s, COOH), 11.87 (1H, s, N-H), 8.53 (1H, s, C-H<sub>1''</sub>), 8.34 (1H, app. s, Ar-H<sub>2'</sub>), 8.00 (1H, app. d,  $J = 7.9$ , Ar-H<sub>4'</sub> or Ar-H<sub>6'</sub>), 7.96 (1H, app. d,  $J = 7.9$ , Ar-H<sub>4'</sub> or Ar-H<sub>6'</sub>), 7.61 (1H, app. t,  $J = 7.9$ , Ar-H<sub>5'</sub>), 7.20-7.54 (17H, m, Ar-H), 5.22 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 166.9 (C=O *carboxy*), 162.5 (C=O *hydrazono*), 152.1 (3-C, 5-C), 146.7 (1''-C), 140.2 (4-C), 137.4 (Ar-C<sub>qt</sub>), 136.8 (Ar-C<sub>qt</sub>), 134.8 (Ar-C<sub>qt</sub>), 131.5 (Ar-C), 130.6 (Ar-C), 129.2 (Ar-C), 128.7 (Ar-C<sub>qt</sub>), 128.5 (Ar-C<sub>qt</sub>), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (2'-C), 127.2 (Ar-C), 107.0 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3544,

3565, 3064, 1736, 1682, 1585, 1425, 1363, 1165;  $m/z$  (ES) 587.2 (100%,  $MH^+$ ); (Found  $MH^+$ , 587.2155.  $C_{36}H_{30}N_2O_6$  requires  $MH$ , 587.2177).

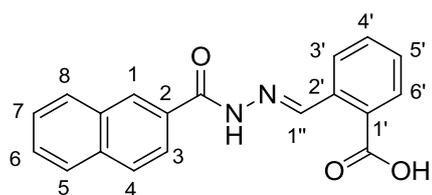
**(E)-4-((2-(3,4,5-tris(benzyloxy)benzoyl)hydrazono)methyl)benzoic acid (6.75)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (198 mg, 0.44 mmol) and 4-formylbenzoic acid **6.126** (65.5 mg, 0.44 mmol). Recrystallization from ethanol / water gave the title compound **6.75** (235 mg, 0.40 mmol, 91%) as colourless needles m.p. 226—227 °C.  $R_f$  0.38 (95:5  $CH_2Cl_2$ —

MeOH); HPLC (Method B), R.t. 3.93 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 13.06 (1H, br s, COOH), 11.86 (1H, s, N-H), 8.53 (1H, s, C-H $_{1''}$ ), 8.01 (2H, d,  $J$  = 7.9, Ar-H $_{2',6'}$ ), 7.85 (2H, d,  $J$  = 7.9, Ar-H $_{3',5'}$ ), 7.46-7.50 (4H, m, Ar-H), 7.31-7.44 (10H, m, Ar-H), 7.21-7.30 (3H, m, Ar-H), 5.21 (4H, s,  $OCH_2$  in 3 and 5), 5.03 (2H, s,  $OCH_2$  in 4);  $\delta_C$  (75 MHz,  $DMSO-d_6$ ); 166.9 (C=O *carboxy*), 162.6 (C=O *amide*), 152.1 (3-C and 5-C), 146.5 (1''-C), 140.3 (4-C), 138.3 (1'-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 131.7 (4'-C), 129.8 (2'-C and 6'-C), 128.44 (Ar-C), 128.37 (1-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 127.1 (3'-C and 5'-C), 107.0 (2-C and 6-C), 74.3 ( $OCH_2$  in 4), 70.5 ( $OCH_2$  in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3512, 3031, 1694, 1581, 1552, 1499, 1423, 1331, 1116, 639;  $m/z$  (ES) 587.2 (100%,  $MH^+$ ); (Found  $MH^+$ , 587.2162.  $C_{36}H_{30}N_2O_6$  requires  $MH$ , 587.2177).

**(E)-2-((2-(2-naphthoyl)hydrazono)methyl)benzoic acid (6.76)**

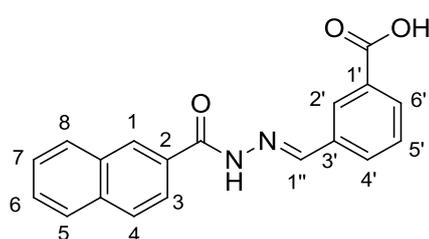


Prepared *via* general method D using 2-naphthohydrazide **6.105** (85 mg, 0.46 mmol) and 2-formylbenzoic acid **6.125** (69 mg, 0.46 mmol). Recrystallization from methanol / water

gave the title compound **6.76** (134 mg, 0.42 mmol, 92%) as colourless plates m.p. 213—214 °C.  $R_f$  0.22 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 2.53 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 13.38 (1H, br s, COOH), 12.23 (1H, s, N-

H), 9.24 (1H, s, C-H<sub>1''</sub>), 8.57 (1H, app. s, Ar-H<sub>1</sub>), 7.94-8.14 (5H, m, Ar-H), 7.91 (1H, app. d, *J* = 7.5, Ar-H<sub>6'</sub>), 7.58-7.70 (3H, m, Ar-H), 7.53 (1H, app. t, *J* = 7.5, Ar-H<sub>5'</sub>); δ<sub>C</sub> (125 MHz, DMSO-*d*<sub>6</sub>); 168.1 (C=O *carboxy*), 163.3 (C=O *hydrazono*), 146.6 (1''-C), 134.6 (Ar-C<sub>qt</sub>), 134.4 (Ar-C<sub>qt</sub>), 132.1 (Ar-C<sub>qt</sub>), 132.0 (Ar-C), 130.7 (Ar-C<sub>qt</sub>), 130.6 (Ar-C<sub>qt</sub>), 130.3 (6'-C), 129.6 (Ar-C), 128.9 (5'-C), 128.1 (1-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 126.9 (Ar-C), 126.7 (Ar-C), 124.4 (Ar-C); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3601, 3266, 3073, 1675, 1653, 1621, 1548, 1480, 1421, 1362, 1241, 1145, 751; *m/z* (ES) 341.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 341.0899. C<sub>19</sub>H<sub>14</sub>N<sub>2</sub>O<sub>3</sub> requires MNa, 341.0897).

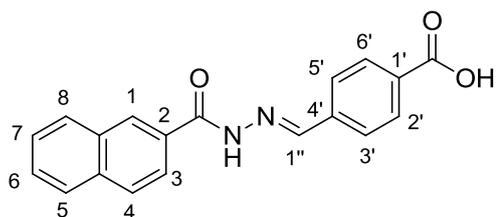
**(E)-3-((2-(2-naphthoyl)hydrazono)methyl)benzoic acid (6.77)**



Prepared *via* general method D using 2-naphthohydrazide **6.105** (186 mg, 1.0 mmol) and 3-formylbenzoic acid **6.124** (150 mg, 1.0 mmol). Recrystallization from ethanol / water gave the title compound **6.77** (274 mg, 0.86

mmol, 86%) as colourless plates m.p. 249—250 °C. *R*<sub>f</sub> 0.16 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 3.01 min, (100%); δ<sub>H</sub> (300 MHz, DMSO-*d*<sub>6</sub>); 13.24 (1H, br s, COOH), 12.19 (1H, s, N-H), 8.56 (1H, app. s, Ar-H<sub>2'</sub>), 8.55 (1H, s, C-H<sub>1''</sub>), 8.36 (1H, app. s, Ar-H<sub>1</sub>), 7.97-8.11 (6H, m, Ar-H), 7.49-7.79 (3H, m, Ar-H); δ<sub>C</sub> (125 MHz, DMSO-*d*<sub>6</sub>); 166.9 (C=O *carboxy*), 163.3 (C=O *hydrazono*), 146.8 (1''-C), 134.8 (Ar-C<sub>qt</sub>), 134.4 (Ar-C<sub>qt</sub>), 132.1 (Ar-C<sub>qt</sub>), 131.5 (Ar-C), 131.4 (Ar-C<sub>qt</sub>), 130.60 (Ar-C), 130.57 (Ar-C), 129.2 (Ar-C), 128.9 (2'-C), 128.14 (1-C), 128.08 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 127.4 (Ar-C), 126.9 (Ar-C), 124.3 (Ar-C); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3208, 3054, 2837, 1801, 1685, 1643, 1548, 1269, 1064, 689; *m/z* (ES) 341.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 341.0881. C<sub>19</sub>H<sub>14</sub>N<sub>2</sub>O<sub>3</sub> requires MNa, 341.0897).

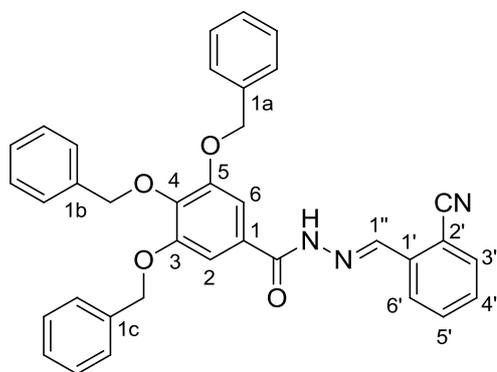
**(E)-4-((2-(2-naphthoyl)hydrazono)methyl)benzoic acid (6.78)<sup>19</sup>**



Prepared *via* general method D using 2-naphthohydrazide **6.105** (186 mg, 1.0 mmol) and 4-formylbenzoic acid **6.126** (150 mg, 1.0 mmol). Recrystallization from ethanol / water gave the title compound

**6.78** (290 mg, 0.91 mmol, 91%) as colourless plates m.p. >250 °C (Lit.<sup>19</sup> m.p. 150—151 °C).  $R_f$  0.36 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 2.49 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 13.15 (1H, br s, COOH), 12.22 (1H, s, N-H), 8.56 (2H, br s, Ar-H<sub>1</sub>, C-H<sub>1''</sub>), 7.94-8.15 (6H, m, Ar-H), 7.87 (2H, d,  $J = 7.8$ , Ar-H<sub>3',5'</sub>), 7.57-7.73 (2H, m, Ar-H);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 166.9 (C=O *carboxy*), 163.30 (C=O *hydrazono*), 146.6 (1''-C), 138.4 (4'-C), 134.4 (Ar-C<sub>qt</sub>), 132.0 (Ar-C<sub>qt</sub>), 131.7 (Ar-C<sub>qt</sub>), 130.5 (Ar-C<sub>qt</sub>), 129.8 (Ar-C), 128.9 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.7 (Ar-C), 127.1 (3'-C, 5'-C), 126.9 (Ar-C), 124.3 (Ar-C);  $\nu_{max}/cm^{-1}$  (neat); 3206, 3042, 1828, 1677, 1504, 1385, 1239, 691;  $m/z$  (ES) 319.1 (70%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 319.1065. C<sub>19</sub>H<sub>14</sub>N<sub>2</sub>O<sub>3</sub> requires *MH*, 319.1077).

**(E)-3,4,5-tris(benzyloxy)-N'-(2-cyanobenzylidene)benzohydrazide (6.79)**

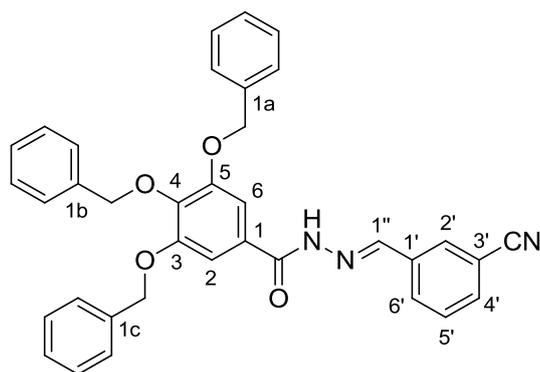


Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (106 mg, 0.23 mmol) and 2-formylbenzonitrile **6.128** (30.6 mg, 0.23 mmol). Recrystallization from ethanol gave the title compound **6.79** (115 mg, 0.20 mmol, 88%) as colourless needles m.p. 177—178 °C.  $R_f$  0.50 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—

MeOH); HPLC (Method A), R.t. 4.38 min, (100%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 12.15 (1H, s, N-H), 8.85 (1H, s, C-H<sub>1''</sub>), 8.16 (1H, app. d,  $J = 7.6$ , Ar-H<sub>6'</sub>), 7.94 (1H, app. d,  $J = 7.6$ , Ar-H<sub>3'</sub>), 7.82 (1H, app. t,  $J = 7.6$ , Ar-H<sub>4'</sub>), 7.64 (1H, app. t,  $J = 7.6$ , Ar-H<sub>5'</sub>), 7.47-7.53 (4H, m, Ar-H), 7.33-7.45 (10H, m, Ar-H), 7.24-7.31 (3H, m, Ar-H), 5.23 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.04 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (125 MHz,

DMSO-*d*<sub>6</sub>); 162.5 (C=O), 152.1 (3-C and 5-C), 142.8 (1''-C), 140.4 (4-C), 137.4 (1b-C), 137.1 (1'-C), 136.8 (1a-C and 1c-C), 133.6 (5'-C), 133.4 (3'-C), 130.4 (4'-C), 128.44 (Ar-C), 128.39 (1-C), 128.2 (Ar-C), 128.1 (6'-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 127.5 (Ar-C), 117.2 (CN), 110.8 (2'-C), 107.2 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3213, 3058, 3025, 2223, 1645, 1582, 1539, 1453, 1336, 1134, 730; *m/z* (ES) 568.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 568.2243. C<sub>36</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires *MH*, 568.2231).

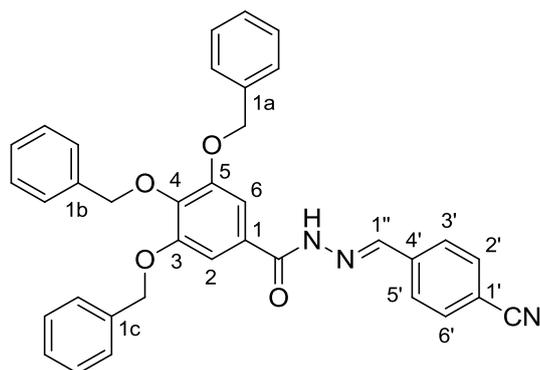
**(E)-3,4,5-tris(benzyloxy)-N'-(3-cyanobenzylidene)benzohydrazide (6.80)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (106 mg, 0.23 mmol) and 3-formylbenzonitrile **6.127** (30.6 mg, 0.23 mmol). Recrystallization from ethanol gave the title compound 6.80 (119 mg, 0.21 mmol, 90%) as

colourless needles m.p. 230—231 °C. *R<sub>f</sub>* 0.43 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 4.22 min, (100%);  $\delta_{\text{H}}$  (300 MHz, DMSO-*d*<sub>6</sub>); 11.97 (1H, s, N-H), 8.51 (1H, s, C-H<sub>1''</sub>), 8.15 (1H, app. s, Ar-H<sub>2'</sub>), 8.11 (1H, app. d, *J* = 7.9, Ar-H<sub>6'</sub>), 7.92 (1H, app. d, *J* = 7.9, Ar-H<sub>4'</sub>), 7.69 (1H, app. t, *J* = 7.9, Ar-H<sub>5'</sub>), 7.46-7.50 (4H, m, Ar-H), 7.32-7.46 (10H, m, Ar-H), 7.18-7.32 (3H, m, Ar-H), 5.22 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_{\text{C}}$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 152.1 (3-C and 5-C), 145.4 (1''-C), 140.3 (4-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 135.7 (1'-C), 133.3 (4'-C), 131.0 (6'-C), 130.7 (2'-C), 130.1 (5'-C), 128.4 (Ar-C), 128.3 (1-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 118.4 (CN), 112.0 (3'-C), 107.1 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/\text{cm}^{-1}$  (neat); 3171, 3062, 2230, 1684, 1644, 1580, 1424, 1336, 1121, 736; *m/z* (ES) 568.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 568.2245. C<sub>36</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires *MH*, 568.2231).

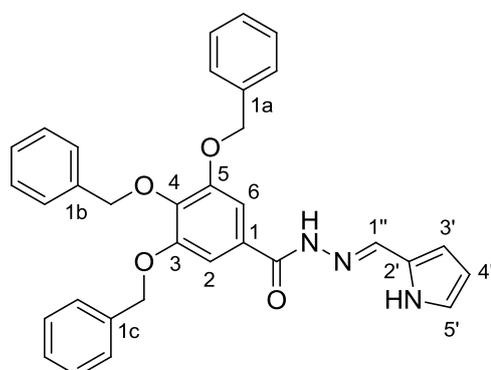
**(E)-3,4,5-tris(benzyloxy)-N'-(4-cyanobenzylidene)benzohydrazide (6.81)**



Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (167 mg, 0.37 mmol) and 4-formylbenzocyanide **6.129** (48 mg, 0.37 mmol). Recrystallization from ethanol gave the title compound **6.81** (183 mg, 0.32 mmol, 87%) as colourless

plates m.p. 223—224 °C.  $R_f$  0.46 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 4.22 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.94 (1H, s, N-H), 8.52 (1H, s, C-H<sub>1''</sub>), 7.92 (4H, app. br s, Ar-H<sub>2',3',5',6'</sub>), 7.46-7.50 (4H, m, Ar-H), 7.32-7.44 (10H, m, Ar-H), 7.19-7.30 (3H, m, Ar-H), 5.20 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.03 (2H, s, OCH<sub>2</sub> in 4);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 152.1 (3-C and 5-C), 145.7 (1''-C), 140.3 (4-C), 138.7 (4'-C), 137.3 (1b-C), 136.7 (1a-C and 1c-C), 132.8 (2'-C and 6'-C), 128.5 (3'-C and 5'-C), 128.45 (Ar-C), 128.36 (1-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 118.6 (CN), 111.9 (1'-C), 107.0 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.4 (OCH<sub>2</sub> in 3 and 5);  $\nu_{max}/cm^{-1}$  (neat); 3655, 3203, 3027, 2223, 1645, 1582, 1539, 1453, 1336, 730;  $m/z$  (ES) 568.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 568.2239. C<sub>36</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires *MH*, 568.2231).

**(E)-N'-((1H-pyrrol-2-yl)methylene)-3,4,5-tris(benzyloxy)benzohydrazide (6.82)**

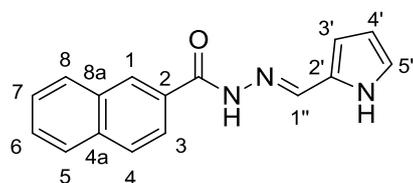


Prepared *via* general method D using 3,4,5-tris(benzyloxy)benzohydrazide **6.32** (205 mg, 0.45 mmol) and 1H-pyrrole-2-carbaldehyde **6.133** (42.9 mg, 0.5 mmol). Recrystallization from ethanol / water gave the title compound **6.82** (207 mg, 0.39 mmol, 87%) as colourless plates m.p. 217

°C (dec).  $R_f$  0.32 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); (Found: C, 74.8; H, 5.55; N, 7.8; C<sub>33</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires C, 74.6; H, 5.50; N, 7.9%). HPLC (Method B), R.t. 4.12 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.53 (1H, s, N-H *pyrrole*), 11.36 (1H, s, N-H *amide*), 8.29 (1H, s, C-H<sub>1''</sub>), 7.46-7.50 (4H, m, Ar-H), 7.33-7.43 (10H, m, Ar-

H), 7.24-7.30 (3H, m, Ar-H), 6.91 (1H, app. br s, Ar-H<sub>5'</sub>), 6.49 (1H, app. br s, Ar-H<sub>3'</sub>), 6.14 (1H, dd, *J* = 5.2, 2.4, Ar-H<sub>4'</sub>), 5.20 (4H, s, OCH<sub>2</sub> in 3 and 5), 5.01 (2H, s, OCH<sub>2</sub> in 4); δ<sub>C</sub> (125 MHz, DMSO-*d*<sub>6</sub>); 161.9 (C=O), 152.1 (3-C and 5-C), 141.0 (1''-C), 139.9 (4-C), 137.4 (1b-C), 136.8 (1a-C and 1c-C), 129.0 (1-C), 128.4 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 126.9 (2'-C), 122.6 (5'-C), 113.4 (3'-C), 109.2 (4'-C), 106.8 (2-C and 6-C), 74.3 (OCH<sub>2</sub> in 4), 70.5 (OCH<sub>2</sub> in 3 and 5); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3435, 3218, 1641, 1606, 1579, 1553, 1454, 1305, 1115, 728; *m/z* (ES) 554.2 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 554.2040. C<sub>33</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub> requires *MNa*, 554.2050).

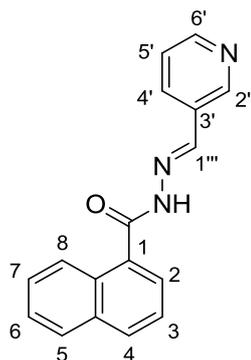
### (E)-N'-((1H-pyrrol-2-yl)methylene)-2-naphthohydrazide (**6.83**)



Prepared *via* general method D using 2-naphthohydrazide **6.105** (124 mg, 0.66 mmol) and 1H-pyrrole-2-carbaldehyde **6.133** (63.1 mg, 0.66 mmol). Recrystallization from ethanol gave

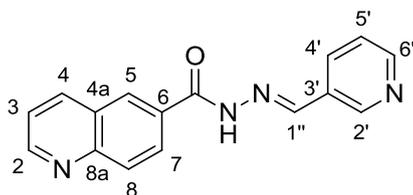
the title compound **6.83** (142 mg, 0.60 mmol, 91%) as light yellow plates m.p. 243 °C (dec). *R*<sub>f</sub> 0.33 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 2.59 min, (100%); δ<sub>H</sub> (500 MHz, DMSO-*d*<sub>6</sub>); 11.66 (1H, s, N-H *amide*), 11.54 (1H, br s, N-H *pyrrole*), 8.50 (1H, s, C-H<sub>1''</sub>), 8.33 (1H, app. s, Ar-H<sub>1</sub>), 7.87-8.13 (4H, m, Ar-H *naphthyl*), 7.72-7.28 (2H, m, Ar-H<sub>4</sub> *naphthyl*), 6.92 (1H, app. br s, Ar-H<sub>5'</sub>), 6.50 (1H, app. br s, Ar-H<sub>3'</sub>), 6.15 (1H, dd, *J* = 5.2, 2.4, Ar-H<sub>4'</sub>); δ<sub>C</sub> (75 MHz, DMSO-*d*<sub>6</sub>); 162.6 (C=O), 140.8 (1-C), 134.2 (2-C), 132.1 (4a-C or 8a-C), 131.0 (4a-C or 8a-C), 128.8 (Ar-C), 128.0 (Ar-C), 127.8 (Ar-C), 127.7 (1''-C), 127.6 (Ar-C), 127.00 (2'-C), 126.8 (Ar-C), 124.3 (Ar-C), 122.6 (5'-C), 113.4 (3'-C), 109.2 (4'-C); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 3445, 3359, 3224, 3059, 1619, 1562, 1435, 1286, 724; *m/z* (ES) 264.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 264.1134. C<sub>16</sub>H<sub>13</sub>N<sub>3</sub>O requires *MH*, 264.1131).

### (E)-N'-(pyridin-3-yl-methylene)-1-naphthohydrazide (6.84)



Prepared *via* general method D using 1-naphthohydrazide **6.108** (127 mg, 0.68 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.064 mL, 0.68 mmol). A portion of the crude compound (60 mg out of 182 mg) was purified *via* mass-directed preparative HPLC eluting with a gradient of methanol in water (5—95%) in presence of formic acid (0.1%) to give a solid which was recrystallized from ethanol to give the title compound **6.84** (52 mg, 0.19 mmol, 87%) as light yellow plates m.p. 221—222 °C.  $R_f$  0.33 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 2.01 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 12.16 (1H, s, N-H), 8.87 (1H, d,  $J = 1.6$ , Ar-H<sub>2'</sub>), 8.63 (1H, dd,  $J = 4.7$  Hz, 1.6, Ar-H<sub>6'</sub>), 8.40 (1H, s, C-H<sub>1''</sub>), 8.19-8.24 (1H, m, Ar-H), 8.17 (1H, app. dt,  $J = 7.9$ , 1.6, Ar-H<sub>4'</sub>), 8.10 (1H, d,  $J = 8.3$ , Ar-H), 8.02 (1H, dd,  $J = 6.5$ , 2.8, Ar-H), 7.76 (1H, d,  $J = 6.5$ , Ar-H), 7.56-7.64 (3H, m, Ar-H), 7.50 (1H, dd,  $J = 7.9$  Hz, 4.7, Ar-H<sub>5'</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 164.8 (C=O), 150.8 (6'-C), 148.8 (2'-C), 145.0 (1''-C), 133.5 (4'-C), 133.2 (Ar-C<sub>qt</sub>), 132.6 (Ar-C<sub>qt</sub>), 130.6 (Ar-C), 130.2 (3'-C), 130.0 (Ar-C<sub>qt</sub>), 128.4 (Ar-C), 127.1 (Ar-C), 126.4 (Ar-C), 126.0 (Ar-C), 125.1 (Ar-C), 125.0 (Ar-C), 124.0 (5'-C);  $\nu_{max}/cm^{-1}$  (neat); 3165, 3011, 1638, 1603, 1590, 1417, 1300, 698;  $m/z$  (ES) 276.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 276.1140. C<sub>17</sub>H<sub>13</sub>N<sub>3</sub>O requires *MH*, 276.1131).

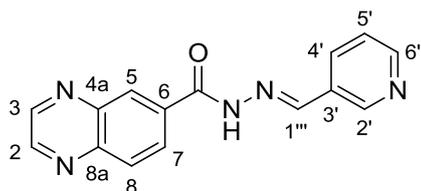
### (E)-N'-(pyridin-3-ylmethylene)quinoline-6-carbohydrazide (6.85)



Prepared *via* general method D using quinoline-6-carbohydrazide **6.107** (93.5 mg, 0.50 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.047 mL, 0.50 mmol). Recrystallization from ethanol gave the title compound **6.85** (105 mg, 0.38 mmol, 76%) as colourless needles m.p. 208—209 °C.  $R_f$  0.40 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 1.29 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 12.24 (1H, s, N-H), 9.01 (1H, dd,  $J = 4.2$ , 1.7, Ar-H<sub>2</sub>), 8.89 (1H, app. s, Ar-H<sub>2'</sub>), 8.63 (1H, app. d,  $J = 4.3$ , Ar-H<sub>6'</sub>), 8.60 (1H, app. s, Ar-H<sub>5</sub>), 8.55 (1H, s, C-H<sub>1''</sub>), 8.53 (1H, dd,  $J = 8.4$ , 1.7, Ar-H<sub>4</sub>), 8.08-8.29 (3H, m, Ar-H<sub>4'</sub>, Ar-H<sub>7</sub>, Ar-H<sub>8</sub>), 7.64 (1H, dd,  $J = 8.2$ , 4.2, Ar-H<sub>3</sub>), 7.51 (1H, dd,  $J =$

7.5, 4.3, Ar-H<sub>5</sub>);  $\delta_{\text{C}}$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.8 (C=O), 152.4 (2-C), 150.8 (6'-C), 148.8 (2'-C), 148.7 (8a-C), 145.4 (1''-C), 137.1 (4-C), 133.5 (4'-C), 131.0 (Ar-C<sub>qt</sub>), 130.2 (5-C), 129.3 (3'-C), 128.6 (7-C or 8-C), 127.9 (7-C or 8-C), 127.1 (Ar-C<sub>qt</sub>), 124.0 (5'-C), 122.4 (3-C);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3451, 3165, 3011, 1675, 1621, 1591, 1568, 1327, 701; *m/z* (ES) 277.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 277.1091. C<sub>16</sub>H<sub>12</sub>N<sub>4</sub>O requires *MH*, 277.1084).

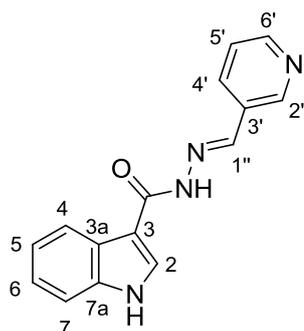
**(E)-N'-(pyridin-3-yl-methylene)quinoxaline-6-carbohydrazide (6.86)**



Prepared *via* general method D using quinoxaline-6-carbohydrazide **6.115** (103 mg, 0.55 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.05 mL, 0.55 mmol). A portion of the crude

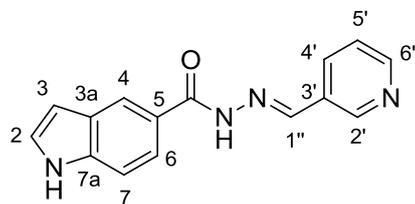
compound (50 mg out of 89 mg) was purified *via* mass-directed preparative HPLC eluting with a gradient of methanol in water (5—95%) in presence of formic acid (0.1%) to give a solid which was recrystallized from ethanol to give the title compound **6.86** (40 mg, 0.18 mmol, 80%) as colourless plates m.p. 106—107 °C. *R<sub>f</sub>* 0.18 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 1.35 min, (100%);  $\delta_{\text{H}}$  (500 MHz, DMSO-*d*<sub>6</sub>); 12.37 (1H, s, N-H), 9.02-9.10 (2H, m, Ar-H<sub>2,3</sub>), 8.90 (1H, app. s, Ar-H<sub>2'</sub>), 8.72 (1H, d, *J* = 1.6, Ar-H<sub>5</sub>), 8.63 (1H, app. d, *J* = 4.7, Ar-H<sub>6'</sub>), 8.58 (1H, s, C-H<sub>1'''</sub>), 8.33 (1H, dd, *J* = 8.7, 1.6, Ar-H<sub>7</sub>), 8.24 (1H, d, *J* = 8.7, Ar-H<sub>8</sub>), 8.18 (1H, app. d, *J* = 7.8, Ar-H<sub>4'</sub>), 7.51 (1H, dd, *J* = 7.8, 4.7, Ar-H<sub>5'</sub>);  $\delta_{\text{C}}$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.2 (C=O), 150.9 (6'-C), 148.9 (2'-C), 147.3 (2-C or 3-C), 146.9 (2-C or 3-C), 145.8 (1'''-C), 143.6 (4a-C), 141.5 (8a-C), 134.3 (6-C), 133.6 (4'-C), 130.1 (3'-C), 129.6 (8-C), 128.8 (7-C), 128.7 (5-C), 124.0 (5'-C);  $\nu_{\text{max}}/\text{cm}^{-1}$  (neat); 3506, 3208, 3010, 1731, 1446, 1369, 1209; *m/z* (EI) 277.1 (100%, M<sup>+</sup>); (Found M<sup>+</sup>, 277.0959. C<sub>15</sub>H<sub>11</sub>N<sub>5</sub>O requires *M*, 277.0964).

**(E)-N'-(pyridin-3-yl-methylene)-1H-indole-3-carbohydrazide (6.87)**



Prepared *via* general method D using 1H-indole-3-carbohydrazide **6.116** (175 mg, 1.0 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.094 mL, 1.0 mmol). Recrystallization from ethanol gave the title compound **6.87** (196 mg, 0.74 mmol, 74%) as colourless needles m.p. 248 °C (dec.).  $R_f$  0.37 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 2.04 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 12.05 (1H, s, N-H *amide*), 11.82 (1H, s, N-H *indole*), 8.89 (1H, app. s, Ar-H<sub>2'</sub>), 8.61 (1H, dd,  $J = 4.8, 1.6$ , Ar-H<sub>6'</sub>), 8.51 (1H, s, 1''-C), 8.16 (1H, app. d,  $J = 7.8$ , Ar-H<sub>4'</sub>), 7.68 (1H, app. d,  $J = 7.8$ , Ar-H<sub>5'</sub>), 7.42-7.54 (2H, m, Ar-H), 7.34 (1H, s, Ar-H<sub>2</sub>), 7.22 (1H, m, Ar-H), 7.07 (1H, m, Ar-H);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 157.8 (C=O), 150.6 (6'-C), 148.7 (2'-C), 144.3 (1''-C), 136.9 (Ar-C<sub>qt</sub>), 133.5 (4'-C), 130.3 (3'-C), 129.8 (Ar-C<sub>qt</sub>), 127.0 (3-C), 124.0 (Ar-C), 123.99 (Ar-C), 121.8 (5'-C), 120.00 (Ar-C), 112.4 (Ar-C), 103.9 (2-C);  $\nu_{max}/cm^{-1}$  (neat); 3279, 3048, 1650, 1606, 1577, 1354, 1227, 745;  $m/z$  (ES) 265.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 265.1074. C<sub>15</sub>H<sub>12</sub>N<sub>4</sub>O requires *MH*, 265.1084).

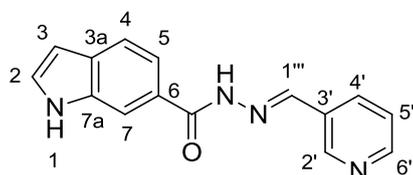
**(E)-N'-(pyridin-3-ylmethylene)-1H-indole-5-carbohydrazide (6.88)**



Prepared *via* general method D using 1H-indole-5-carbohydrazide **6.103** (175 mg, 1.0 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.094 mL, 1.0 mmol). Recrystallization from ethanol gave the title compound **6.88** (204 mg, 0.77 mmol, 77%) as light brown plates m.p. 151—152 °C.  $R_f$  0.32 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 1.55 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 11.90 (1H, s, N-H *amide*), 11.40 (1H, s, N-H *indole*), 8.84 (1H, d,  $J = 1.1$ , Ar-H<sub>2'</sub>), 8.59 (1H, dd,  $J = 4.7, 1.6$ , Ar-H<sub>6'</sub>), 8.51 (1H, s, 1''-C), 8.22 (1H, app. s, Ar-H<sub>4</sub>), 8.13 (1H, app. d,  $J = 7.3$ , Ar-H<sub>4'</sub>), 7.70 (1H, dd,  $J = 8.5, 1.6$ , Ar-H<sub>6</sub>), 7.45-7.51 (3H, m, Ar-H<sub>2,7,5'</sub>), 6.57-6.59 (1H, m, Ar-H<sub>3</sub>);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 164.3 (C=O), 150.4 (6'-C), 148.6 (2'-C), 143.7 (1''-C), 137.7 (Ar-C<sub>qt</sub>), 133.3 (4'-C), 130.5 (Ar-C<sub>qt</sub>), 127.01 (Ar-C<sub>qt</sub>), 126.97 (2-C or 7-C or 5'-C), 123.97 (2-C or 7-C or 5'-C), 123.95 (Ar-C<sub>qt</sub>), 120.9 (6-C), 120.5 (4-C),

111.2 (2-C or 7-C or 5'-C), 102.2 (3-C);  $\nu_{max}/cm^{-1}$  (neat); 3177, 3033, 1648, 1592, 1553, 1414, 1293, 1188, 700;  $m/z$  (ES) 265.1 (100%,  $MH^+$ ); (Found  $MH^+$ , 265.1073.  $C_{15}H_{12}N_4O$  requires  $MH$ , 265.1084).

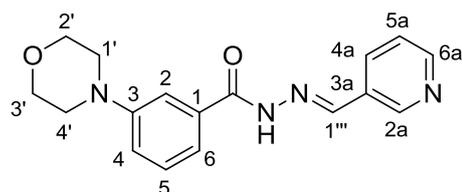
**(E)-N'-(pyridin-3-yl-methylene)-1H-indole-6-carbohydrazide (6.89)**



Prepared *via* general method D using 1H-indole-6-carbohydrazide **6.104** (88 mg, 0.50 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.047 mL, 0.50 mmol). Recrystallization from ethanol gave the

title compound **6.89** (108 mg, 0.41 mmol, 82%) as light pink plates m.p. 118 °C (dec).  $R_f$  0.17 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method B), R.t. 1.55 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 11.94 (1H, s, N-H *amide*), 11.46 (1H, s, N-H *indole*), 8.85 (1H, d,  $J = 1.4$ , Ar- $H_{2'}$ ), 8.60 (1H, dd,  $J = 4.7, 1.4$ , Ar- $H_{6'}$ ), 8.52 (1H, s, C- $H_{1''}$ ), 8.13 (1H, app. d,  $J = 8.3$ , Ar- $H_{4'}$ ), 8.04 (1H, app. s, Ar- $H_7$ ), 7.59-7.65 (2H, m, Ar- $H_{4,5'}$ ), 7.55 (1H, app. t,  $J = 2.8$ , Ar- $H_2$ ), 7.48 (1H, dd,  $J = 7.9, 4.8$ , Ar- $H_5$ ), 6.52 (1H, m, Ar- $H_3$ );  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 164.5 (C=O), 150.5 (6'-C), 148.6 (2'-C), 144.0 (1'''-C), 135.1 (7a-C), 133.3 (4'-C), 130.5 (6-C), 130.3 (3'-C), 128.52 (2-C), 125.7 (3a-C), 124.0 (5'-C), 119.6 (4-C), 118.3 (5-C), 111.8 (7-C), 101.4 (3-C);  $\nu_{max}/cm^{-1}$  (neat); 3400, 3160, 3056, 1642, 1610, 1556, 1306, 1267, 696;  $m/z$  (ES) 287.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 287.0904.  $C_{15}H_{12}N_4O$  requires  $MNa$ , 287.0903).

**(E)-3-morpholino-N'-(pyridin-3-yl-methylene)benzohydrazide (6.90a)**

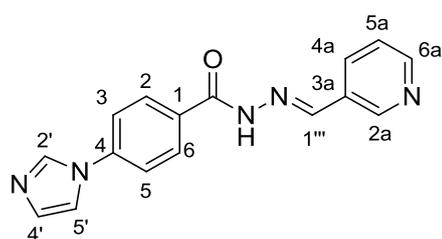


Prepared *via* general method D using 3-morpholinobenzohydrazide **6.110** (75 mg, 0.34 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.32 mL, 0.34 mmol). Recrystallization

from ether gave the title compound **6.90a** (90 mg, 0.29 mmol, 85%) as light yellow plates m.p. 183—184 °C.  $R_f$  0.13 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method B), R.t. 1.61 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 11.88 (1H, s, N-H), 8.84 (1H, app. s, Ar- $H_{2a}$ ), 8.61 (1H, app. d,  $J = 4.8$ , Ar- $H_{6a}$ ), 8.50 (1H, s, C- $H_{1''}$ ), 8.13 (1H, app. d,  $J = 7.6$ , Ar- $H_{4a}$ ), 7.48 (1H, dd,  $J = 7.6, 4.8$ , Ar- $H_{5a}$ ), 7.29-7.42 (3H,

m, Ar-H<sub>2,5,6</sub>), 7.17 (1H, app. d,  $J = 7.0$ , Ar-H<sub>4</sub>), 3.76 (4H, t,  $J = 5.0$ , CH<sub>2</sub> in 2' and 3'), 3.18 (4H, t,  $J = 5.0$ , CH<sub>2</sub> in 1' and 4');  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 163.6 (C=O), 151.0 (3-C), 150.7 (6a-C), 148.7 (2a-C), 144.9 (1'''-C), 134.0 (1-C), 133.4 (4a-C), 130.3 (3a-C), 129.14 (2-C or 5-C or 6-C), 124.0 (5a-C), 118.4 (4-C), 118.3 (2-C or 5-C or 6-C), 113.9 (2-C or 5-C or 6-C), 66.0 (2'-C and 3'-C), 48.2 (1'-C and 4'-C);  $\nu_{max}/cm^{-1}$  (neat); 3461, 3220, 3068, 1668, 1607, 1554, 1271, 1114, 691;  $m/z$  (ES) 311.2 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 311.1507. C<sub>17</sub>H<sub>18</sub>N<sub>4</sub>O<sub>2</sub> requires *MH*, 311.1502).

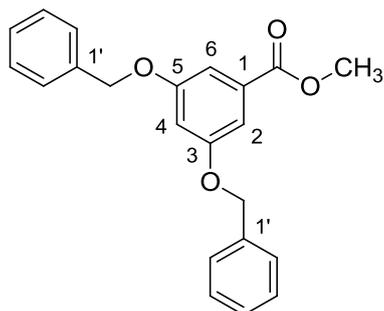
**(E)-4-(1H-imidazol-1-yl)-N'-(pyridin-3-yl-methylene)benzohydrazide (6.91a)**



Prepared *via* general method D using 4-(1H-imidazol-1-yl)benzohydrazide **6.111** (96 mg, 0.47 mmol) and 3-pyridinecarboxyaldehyde **6.33** (0.044 mL, 0.47 mmol). Recrystallization from ethanol gave the title compound **6.91a**

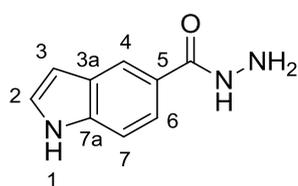
(116 mg, 0.40 mmol, 85%) as colourless needles m.p. 196—197 °C.  $R_f$  0.30 (95:5 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 1.18 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 12.07 (1H, s, N-H), 8.87 (1H, app. s, Ar-H<sub>2a</sub>), 8.62 (1H, app. d,  $J = 4.7$ , Ar-H<sub>6a</sub>), 8.53 (1H, s, C-H<sub>1'''</sub>), 8.41 (1H, app. s, Ar-H<sub>5'</sub>), 8.15 (1H, app. d,  $J = 7.6$ , Ar-H<sub>4a</sub>), 8.07 (2H, d,  $J = 8.3$ , Ar-H<sub>3,5</sub>), 7.88 (1H, t,  $J = 1.3$ , Ar-H<sub>2'</sub>), 7.86 (2H, d,  $J = 8.3$ , Ar-H<sub>2,6</sub>), 7.49 (1H, dd,  $J = 7.6, 4.7$ , Ar-H<sub>5a</sub>), 7.11-7.18 (1H, m, Ar-H<sub>4'</sub>);  $\delta_C$  (125 MHz, DMSO-*d*<sub>6</sub>); 162.2 (C=O), 150.8 (6a-C), 148.8 (2a-C), 145.2 (1'''-C), 139.4 (4-C), 135.6 (5'-C), 133.5 (4a-C), 131.1 (1-C), 130.3 (4'-C), 130.2 (3a-C), 129.5 (3-C and 5-C), 124.0 (5a-C), 119.7 (2-C and 6-C), 117.8 (2'-C);  $\nu_{max}/cm^{-1}$  (neat); 3451, 3240, 3090, 1652, 1513, 1426, 1367, 1105, 967;  $m/z$  (ES) 292.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 292.1203. C<sub>16</sub>H<sub>13</sub>N<sub>5</sub>O requires *MH*, 292.1192).

### Methyl 3,5-bis(benzyloxy)benzoate (6.93)<sup>20</sup>



Potassium carbonate (2073 mg, 15 mmol) and benzyl bromide (1.55 mL, 13 mmol) were added to a solution of 3,5-trihydroxybenzoate methyl ester **6.92** (1684 mg, 10 mmol) in acetone (30 mL). This reaction mixture was stirred at reflux for twenty-four hours and then poured into water (90 mL). The mixture was extracted with EtOAc (4 x 60 mL). The organic layer was dried over MgSO<sub>4</sub> and concentrated under reduced pressure. The light brown solid was purified using column chromatography on silica gel eluting with petroleum ether / ethylacetate (8:2) and recrystallized from hexane to give the title compound **6.93** (3100 mg, 8.9 mmol, 89%) as colourless plates m.p. 58—60 °C (Lit.<sup>20</sup> m.p. 69.2 °C). *R<sub>f</sub>* 0.56 (1:1 petroleum ether—EtOAc); HPLC (Method A), R.t. 1.28 min, (100%); δ<sub>H</sub> (300 MHz, DMSO-*d*<sub>6</sub>); 7.26-7.50 (10H, m, Ar-H), 7.17 (2H, d, *J* = 2.3, Ar-H<sub>2,6</sub>), 6.97 (1H, t, *J* = 2.3, Ar-H<sub>4</sub>), 5.15 (4H, s, OCH<sub>2</sub> in 3 and 5), 3.84 (2H, s, OCH<sub>3</sub>); δ<sub>C</sub> (75 MHz, DMSO-*d*<sub>6</sub>); 165.8 (C=O), 159.5 (3-C and 5-C), 136.6 (1'-C), 131.6 (1-C), 128.4 (Ar-C), 127.9 (Ar-C), 127.6 (Ar-C), 107.9 (2-C, 6-C), 106.9 (4-C), 69.5 (OCH<sub>2</sub>), 52.3 (OCH<sub>3</sub>); *v*<sub>max</sub>/cm<sup>-1</sup> (neat); 2947, 1711, 1498, 1377, 1351, 1235, 1105, 1044, 729; *m/z* (ES) 371.1 (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 371.1271. C<sub>22</sub>H<sub>20</sub>O<sub>4</sub> requires MNa, 371.1254).

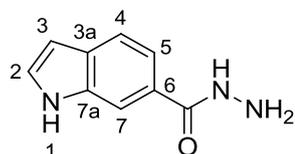
### 1H-indole-5-carbohydrazide (6.103)<sup>21</sup>



Prepared *via* general method C using methyl 1H-indole-5-carboxylate **6.94** (521 mg, 3.0 mmol) and hydrazine hydrate (1.44 mL, 30 mmol). Recrystallization from ethanol / water gave the title compound **6.103** (468 mg, 2.7 mmol, 89%) as colourless needles m.p. 173—175 °C (Lit.<sup>21</sup> m.p. 173—177 °C). *R<sub>f</sub>* 0.52 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 0.80 min, (96%); δ<sub>H</sub> (300 MHz, DMSO-*d*<sub>6</sub>); 11.33 (1H, s, N-H *indole*), 9.59 (1H, s, N-H *amide*), 8.09 (1H, app. s, Ar-H<sub>4</sub>), 7.60 (1H, dd, *J* = 8.5, 1.6, Ar-H<sub>6</sub>), 7.38-7.42 (2H, m, Ar-H<sub>2,7</sub>), 6.50-6.52 (1H, m, Ar-H<sub>3</sub>), 4.54 (2H, br s, NH<sub>2</sub>); δ<sub>C</sub> (75 MHz, DMSO-*d*<sub>6</sub>); 167.3 (C=O), 137.3 (5-C), 127.0 (7a-C), 126.6 (2-C or 7-C), 124.2 (3a-C), 120.2 (6-C),

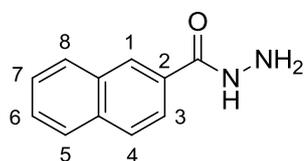
119.6 (4-C), 110.9 (2-C or 7-C), 102.0 (3-C);  $\nu_{max}/cm^{-1}$  (neat); 3296, 3213, 3028, 1635, 1614, 1583, 1466, 1327, 755;  $m/z$  (ES) 198.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 198.0638.  $C_9H_9N_3O$  requires  $MNa$ , 198.0638).

### 1H-indole-6-carbohydrazide (6.104)<sup>21</sup>



Prepared *via* general method C using methyl 1H-indole-6-carboxylate **6.95** (800 mg, 4.6 mmol) and hydrazine hydrate (4.50 mL, 91 mmol). Recrystallization from ethanol gave the title compound **6.104** (713 mg, 4.1 mmol, 89%) as light yellow needles m.p. 190—191 °C (Lit.<sup>21</sup> Not available).  $R_f$  0.46 (90:10  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 1.06 min, (100%);  $\delta_H$  (300 MHz,  $DMSO-d_6$ ); 11.40 (1H, s, N-H *indole*), 9.66 (1H, s, N-H *amide*), 7.92 (1H, app. s, Ar-H<sub>7</sub>), 7.55 (1H, app. d,  $J = 8.3$ , Ar-H<sub>4</sub>), 7.47-7.51 (2H, m, Ar-H<sub>2,5</sub>), 6.49-6.55 (1H, m, Ar-H<sub>3</sub>), 4.45 (2H, s,  $NH_2$ );  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 167.1 (C=O), 135.2 (7a-C), 129.7 (6-C), 127.8 (2-C), 126.1 (3a-C), 119.4 (4-C), 117.6 (5-C), 110.9 (7-C), 101.2 (3-C);  $\nu_{max}/cm^{-1}$  (neat); 3310, 3116, 3035, 1625, 1566, 1455, 1320, 1274, 1106, 731;  $m/z$  (ES) 198.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 198.0638.  $C_9H_9N_3O$  requires  $MNa$ , 198.0638).

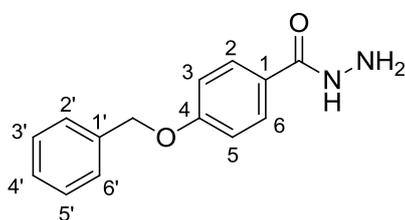
### 2-naphthohydrazide (6.105)<sup>22</sup>



Prepared *via* general method C using methyl-2-naphthoate **6.96** (1 g, 5.37 mmol) and hydrazine hydrate (6 mL, 123 mmol). Recrystallization from methanol gave the title compound **6.105** (0.9 g, 5.0 mmol, 93%) as colourless plates m.p. 146—148 °C (Lit.<sup>22</sup> m.p. 147—152 °C).  $R_f$  0.36 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 1.68 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 9.95 (1H, s, N-H), 8.44 (1H, app. s, Ar-H<sub>1</sub>), 7.96-8.04 (3H, m, Ar-H<sub>4,5,8</sub>), 7.92 (1H, dd,  $J = 8.6, 1.6$ , Ar-H<sub>3</sub>), 7.58-7.69 (2H, m, Ar-H<sub>6,7</sub>), 4.68 (2H, br s,  $NH_2$ );  $\delta_C$  (75 MHz,  $DMSO-d_6$ ); 165.8 (C=O), 134.0 (Ar-C<sub>qt</sub>), 132.1 (2-C), 130.6 (Ar-C<sub>qt</sub>), 128.8 (4-C or 5-C or 8-C), 127.8 (6-C or 7-C), 127.6 (4-C or 5-C or 8-C), 127.5 (1-C), 127.2 (6-C or 7-C), 126.7 (4-C or 5-C or 8-C), 123.8 (3-C);  $\nu_{max}/cm^{-1}$  (neat); 3311, 3177, 3036, 1650, 1571, 1506, 1267, 724;  $m/z$  (ES)

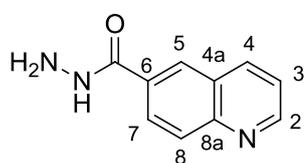
209.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 209.0690.  $C_{11}H_{10}N_2O$  requires  $MNa$ , 209.0685).

#### 4-(benzyloxy)benzohydrazide (6.106)<sup>23</sup>



Prepared *via* general method C using methyl-4-benzyloxybenzoate **6.97** (2000 mg, 8.25 mmol) and hydrazine hydrate (10 mL, 206 mmol). Recrystallization from ethanol gave the title compound **6.106** (1879 mg, 7.76 mmol, 94%) as colourless needles m.p. 139—140 °C (Lit.<sup>23</sup> m.p. 140).  $R_f$  0.42 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method B), R.t. 2.87 min, (100%);  $\delta_H$  (300 MHz,  $DMSO-d_6$ ); 9.63 (1H, s, N-H), 7.80 (2H, d,  $J = 8.9$ , Ar- $H_{2,6}$ ), 7.27-7.52 (5H, m, Ar- $H_{2',3',4',5',6'}$ ), 7.06 (2H, d,  $J = 8.9$ , Ar- $H_{3,5}$ ), 5.16 (2H, s,  $OCH_2$ ), 4.44 (2H, br s,  $NH_2$ );  $\delta_C$  (75 MHz,  $DMSO-d_6$ ); 165.5 (C=O), 160.4 (4-C), 136.7 (1'-C), 128.7 (2-C, 6-C), 128.4 (Ar-C), 127.9 (Ar-C), 127.8 (Ar-C), 125.6 (1-C), 114.3 (3-C, 5-C), 69.2 ( $OCH_2$ );  $\nu_{max}/cm^{-1}$  (neat); 3285, 3193, 1597, 1571, 1531, 1503, 1245, 1223, 834, 652;  $m/z$  (ES) 265.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 265.0943.  $C_{14}H_{14}N_2O_2$  requires  $MNa$ , 265.0947).

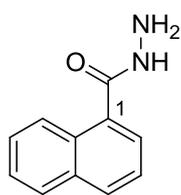
#### Quinoline-6-carbohydrazide (6.107)<sup>24</sup>



Prepared *via* general method C using methyl quinoline-6-carboxylate **6.98** (700 mg, 3.74 mmol) and hydrazine hydrate (3.70 mL, 74.8 mmol). Recrystallization from ethanol gave the title compound **6.107** (595 mg, 3.18 mmol, 85%) as colourless plates m.p. 192—194 °C (Lit.<sup>24</sup> m.p. 193—194 °C).  $R_f$  0.40 (90:10  $CH_2Cl_2$ —MeOH); HPLC (Method B), R.t. 0.72 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 10.04 (1H, s, N-H), 8.98 (1H, dd,  $J = 4.2, 1.7$ , Ar- $H_2$ ), 8.44-8.50 (2H, m, Ar- $H_4$  and Ar- $H_5$ ), 8.15 (1H, dd,  $J = 8.8, 1.9$ , Ar- $H_7$ ), 8.06 (1H, d,  $J = 8.8$ , Ar- $H_8$ ), 7.61 (1H, dd,  $J = 8.3, 4.2$ , Ar- $H_3$ ), 4.77 (2H, br s,  $NH_2$ );  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 165.3 (C=O), 151.9 (2-C), 148.6 (8a-C), 137.0 (4-C or 5-C), 131.1 (Ar- $C_{qt}$ ), 129.0 (8-C), 127.7 (4-C or 5-C), 127.4 (7-C), 127.1 (Ar- $C_{qt}$ ), 122.1 (3-C);  $\nu_{max}/cm^{-1}$  (neat); 3310, 3116, 3035, 1625, 1566, 1528, 1455, 1320,

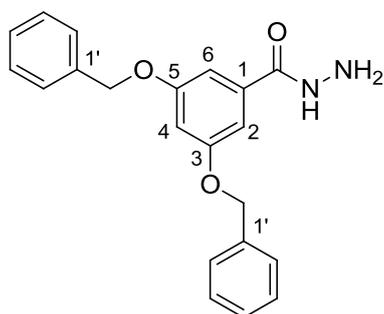
1334, 731;  $m/z$  (ES) 188.1 (100%,  $MH^+$ ); (Found  $MH^+$ , 188.0826.  $C_{10}H_9N_3O$  requires  $MH$ , 188.0818).

### 1-naphthohydrazide (6.108)<sup>25</sup>



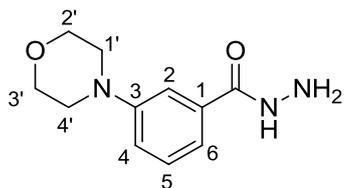
Prepared *via* general method C using ethyl 1-naphthoate **6.99** (0.20 mL, 1.10 mmol) and hydrazine hydrate (1.1 mL, 22.0 mmol). Recrystallization from ethanol gave the title compound **6.108** (197 mg, 1.06 mmol, 96%) as colourless plates m.p. 159—160 °C (Lit.<sup>25</sup> m.p. 160—163 °C).  $R_f$  0.78 (90:10  $CH_2Cl_2$ —MeOH); HPLC (Method B), R.t. 1.44 min, (100%);  $\delta_H$  (500 MHz,  $DMSO-d_6$ ); 9.71 (1H, s, N-H), 8.19-8.24 (1H, m, Ar-H), 8.02 (1H, d,  $J = 7.4$ , Ar-H), 7.96-8.00 (1H, m, Ar-H), 7.51-7.60 (4H, m, Ar-H), 4.61 (2H, br s,  $NH_2$ );  $\delta_C$  (125 MHz,  $DMSO-d_6$ ); 167.9 (C=O), 133.3 (Ar- $C_{qt}$ ), 133.1 (Ar- $C_{qt}$ ), 130.0 (Ar- $C_{qt}$ ), 129.8 (Ar-C), 128.2 (Ar-C), 126.6 (Ar-C), 126.2 (Ar-C), 125.4 (Ar-C), 125.3 (Ar-C), 125.0 (Ar-C);  $\nu_{max}/cm^{-1}$  (neat); 3274, 3046, 1644, 1608, 1588, 1515, 1208, 939, 734;  $m/z$  (ES) 209.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 209.0691.  $C_{11}H_{10}N_2O$  requires  $MNa$ , 209.0685).

### 3,5-bis(benzyloxy)benzohydrazide (6.109)<sup>13</sup>



Prepared *via* general method C using methyl 3,5-bis(benzyloxy)benzoate **6.100** (3.0 g, 8.6 mmol) and hydrazine hydrate (8.0 mL, 172 mmol). Recrystallization from ethanol gave the title compound **6.109** (2.8 g, 8.0 mmol, 93%) as colourless plates m.p. 118—119 °C (Lit.<sup>13</sup> not available).  $R_f$  0.40 (95:5  $CH_2Cl_2$ —MeOH); HPLC (Method A), R.t. 2.70 min, (100%);  $\delta_H$  (300 MHz,  $DMSO-d_6$ ); 9.75 (1H, s, N-H), 7.28-7.51 (10H, m, Ar-H), 7.09 (2H, d,  $J = 2.2$ , Ar- $H_{2,6}$ ), 6.81 (1H, t,  $J = 2.2$ , Ar- $H_4$ ), 5.13 (4H, s,  $OCH_2$  in 3 and 5), 4.51 (2H, br s,  $NH_2$ );  $\delta_C$  (75 MHz,  $DMSO-d_6$ ); 165.2 (C=O), 159.3 (3-C and 5-C), 136.8 (1'-C), 135.3 (1-C), 128.4 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 105.9 (2-C and 6-C), 104.6 (4-C), 69.4 ( $OCH_2$ );  $\nu_{max}/cm^{-1}$  (neat); 3276, 3031, 1625, 1453, 1353, 1162, 1058, 693;  $m/z$  (ES) 371.1 (100%,  $MNa^+$ ); (Found  $MNa^+$ , 371.1366.  $C_{21}H_{20}N_2O_3$  requires  $MNa$ , 371.1366).

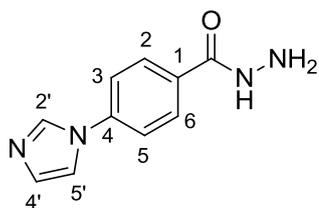
### 3-morpholinobenzohydrazide (6.110)<sup>26</sup>



Prepared *via* general method C using methyl 3-morpholinobenzoate **6.101** (256 mg, 1.16 mmol) and hydrazine hydrate (0.56 mL, 11.6 mmol).

Recrystallization from ethanol gave the title compound **6.110** (239 mg, 1.08 mmol, 93%) as light yellow needles m.p. 109—110 °C (Lit.<sup>26</sup> not available).  $R_f$  0.22 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method B), R.t. 1.08 min, (95%);  $\delta_H$  (300 MHz, DMSO-*d*<sub>6</sub>); 9.63 (1H, s, N-H), 7.13-7.32 (3H, m, Ar-H<sub>2,5,6</sub>), 6.96-7.03 (1H, m, Ar-H<sub>4</sub>), 4.42 (2H, br s, NH<sub>2</sub>), 3.66 (4H, t,  $J$  = 5.5, CH<sub>2</sub> in 2' and 3'), 3.05 (4H, t,  $J$  = 5.5, CH<sub>2</sub> in 1' and 4');  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 166.1 (C=O), 150.9 (3-C), 134.0 (1-C), 128.9 (Ar-C), 117.7 (Ar-C), 117.6 (Ar-C), 113.2 (Ar-C), 66.0 (2'-C and 3'-C), 48.2 (1'-C and 4'-C);  $\nu_{max}/cm^{-1}$  (neat); 3305, 3200, 2835, 1668, 1623 1597, 1574, 1516, 1487, 1240, 1121, 926;  $m/z$  (ES) 222.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 222.1235. C<sub>11</sub>H<sub>15</sub>N<sub>3</sub>O<sub>2</sub> requires *MH*, 222.1237).

### 4-(1H-imidazol-1-yl)benzohydrazide (6.111)<sup>27</sup>



Prepared *via* general method C using methyl 4-(1H-imidazol-1-yl)benzoate **6.102** (323 mg, 1.6 mmol) and hydrazine hydrate (0.82 mL, 16 mmol).

Recrystallization from ethanol gave the title compound **6.111** (303 mg, 1.5 mmol, 95%) as colourless needles m.p. 225—227 °C (Lit.<sup>27</sup> m.p. 223—225 °C).  $R_f$  0.16 (90:10 CH<sub>2</sub>Cl<sub>2</sub>—MeOH); HPLC (Method A), R.t. 0.65 min, (100%);  $\delta_H$  (500 MHz, DMSO-*d*<sub>6</sub>); 9.85 (1H, s, N-H), 8.36 (1H, app. s, Ar-H<sub>5</sub>), 7.95 (2H, d,  $J$  = 8.7, Ar-H<sub>3,5</sub>), 7.83 (1H, t,  $J$  = 1.1, Ar-H<sub>2</sub>), 7.75 (2H, d,  $J$  = 8.7, Ar-H<sub>2,6</sub>), 7.12 (1H, app. s, Ar-H<sub>4</sub>), 4.73 (2H, br s, NH<sub>2</sub>);  $\delta_C$  (75 MHz, DMSO-*d*<sub>6</sub>); 164.8 (C=O), 138.7 (4-C), 135.6 (5'-C), 131.2 (1-C), 130.2 (4'-C), 128.6 (3-C and 5-C), 119.6 (2-C and 6-C), 117.8 (2'-C);  $\nu_{max}/cm^{-1}$  (neat); 3318, 3276, 3128, 3111, 1651, 1605, 1506, 1485, 1320, 1252, 1061, 651;  $m/z$  (ES) 203.1 (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 203.0931. C<sub>10</sub>H<sub>10</sub>N<sub>4</sub>O requires *MH*, 203.0927).

### 9.3 References

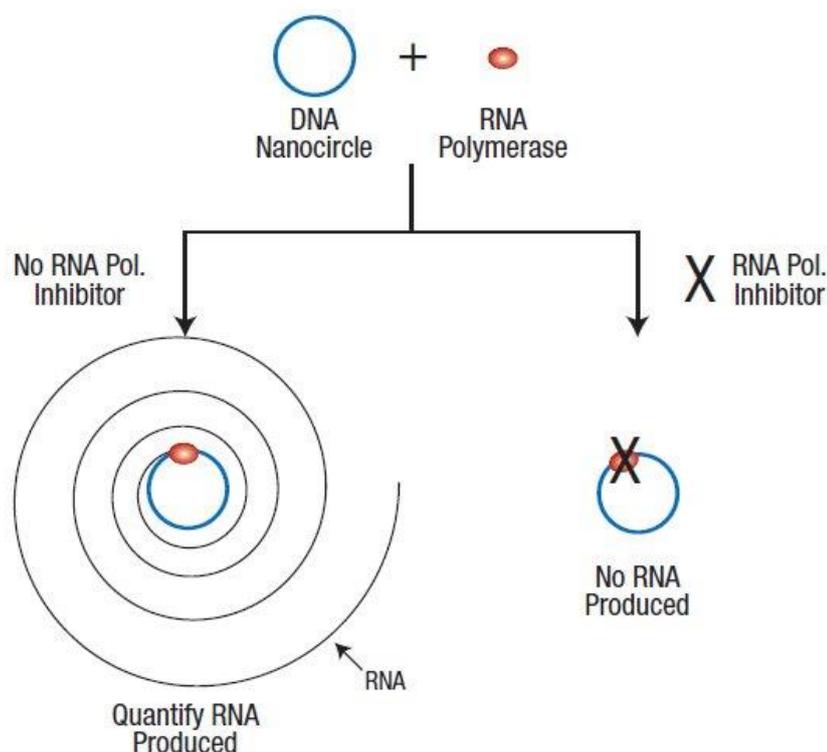
1. Garcia Ruano, J. L.; Parra, A.; Marzo, L.; Yuste, F.; Mastranzo, V. M., *Tetrahedron* **2011**, *67*, 2905-2910.
2. Pasha, M. A.; Madhusudana Reddy, M. B., *Synth. Commun.* **2009**, *39*, 2928-2934.
3. Miyahara, M.; Nakadate, M.; Sueyoshi, S.; Tanno, M.; Kamiya, S., *Chem. Pharm. Bull.* **1982**, *30*, 4402-4406.
4. Yagodkin, A.; Lösckke, K.; Weisell, J.; Azhayev, A., *Tetrahedron* **2010**, *66*, 2210-2221.
5. Erb, B.; Rigo, B.; Pirotte, B.; Couturier, D., *J. Heterocycl. Chem.* **2002**, *39*, 15-28.
6. Andersen, R. A.; Coates, G. E., *J. Chem. Soc., Dalton Trans.* **1974**, 1171-1180.
7. Raffa, L.; Pecorari, P., *Farmaco-Edizione Scientifica* **1966**, *21*, 196-198.
8. Duffield, A. M.; Jefferies, P. R., *Aust. J. Chem.* **1963**, *16*, 292-294.
9. Bain, D. I.; Smalley, R. K., *J. Chem. Soc. C* **1968**, 1593-1594.
10. Dabiri, M.; Salehi, P.; Mohammadi, A. A.; Baghbanzadeh, M., *Synth. Commun.* **2005**, *35*, 279-287.
11. Su, X.; Surry, D. S.; Spandl, R. J.; Spring, D. R., *Org. Lett.* **2008**, *10*, 2593-2596.
12. Clinton, R. O.; Geissman, T. A., *J. Am. Chem. Soc.* **1943**, *65*, 85-87.
13. Rajamalli, P.; Atta, S.; Maity, S.; Prasad, E., *Chem. Commun.* **2013**, *49*, 1744-1746.
14. Howlader, M. B. H.; Islam, M. S., *Indian J. Chem.* **2007**, *46*, 440-444.
15. Lima, P. C.; Lima, L. M.; da Silva, K. C. M.; Léda, P. H. O.; de Miranda, A. L. P.; Fraga, C. A. M.; Barreiro, E. J., *Eur. J. Med. Chem.* **2000**, *35*, 187-203.
16. Chaur, M. N.; Collado, D.; Lehn, J.-M., *Chem. – Eur. J.* **2011**, *17*, 248-258.
17. Mazzone, G.; Puglisi, G.; Marchetta, G.; Corsaro, A., *J. Het. Chem.* **1984**, *21*, 181-184.
18. Hu, Y.; Lu, X.; Chen, K.; Yan, R.; Li, Q.-S.; Zhu, H.-L., *Bioorg. Med. Chem.* **2012**, *20*, 903-909.

19. Avila, C. M.; Lopes, A. B.; Gonçalves, A. S.; da Silva, L. L.; Romeiro, N. C.; Miranda, A. L. P.; Sant'Anna, C. M. R.; Barreiro, E. J.; Fraga, C. A. M., *Eur. J. Med. Chem.* **2011**, *46*, 1245-1253.
20. Nguyen, T.-T.-T.; Simon, F.-X.; Schmutz, M.; Mesini, P. J., *Chem. Commun.* **2009**, 3457-3459.
21. Gillig, R.; Kinnick, D.; Morin, J.; Navarro, A. WO2005/40157 A2, 2005.
22. Caboni, L.; Egan, B.; Kelly, B.; Blanco, F.; Fayne, D.; Meegan, M. J.; Lloyd, D. G., *J. Chem. Inf. Model.* **2013**, *53*, 2116-2130.
23. Mazouz, F.; Gueddari, S.; Burstein, C.; Mansuy, D.; Milcent, R., *J. Med. Chem.* **1993**, *36*, 1157-1167.
24. Saitoh, M.; Kunitomo, J.; Kimura, E.; Hayase, Y.; Kobayashi, H.; Uchiyama, N.; Kawamoto, T.; Tanaka, T.; Mol, C. D.; Dougan, D. R.; Textor, G. S.; Snell, G. P.; Itoh, F., *Bioorg. Med. Chem.* **2009**, *17*, 2017-2029.
25. Li, X.; Zhao, Z.; Li, G.; Shi, P., *J. Chem. Res.* **2010**, *34*, 410-413.
26. Vankayalapati, H.; Sorna, V.; Warner, S. L.; Bearss, D. J.; Sharma, S.; Stephens, B. WO2013/25805 A1, 2013.
27. Artico, M.; Silvestri, R.; Stefancich, G.; Avigliano, L.; Di Giulio, A.; Maccarrone, M.; Agostinelli, E.; Mondovi, B.; Morpurgo, L., *Eur. J. Med. Chem.* **1992**, *27*, 219-228.

## Appendix I

### A) The SYBR green assay<sup>1</sup>

The **Kool NC-45 RNAP Activity and Inhibitor Screening Kit** was used for biological evaluation of the synthesized compounds. This assay is based on the Rolling Circle Transcription (RCT) Technology which takes advantage of the observations that certain small, circular, single-stranded DNAs are efficiently transcribed by RNAP *in vitro* through a rolling circle in the absence of promoter sequences, primers and sigma factors allowing to study the activity of inhibitors of the core RNAP. The kit is constituted of *E. coli* RNA Polymerase (core enzyme) and SYBR Green I dye (Molecular Probes) for real-time detection of RNA polymerase activity. The dye binds to RNA increasing the intensity of fluorescent emission and this signal is proportional to the quantity of RNA produced. Therefore in the presence of RNAP inhibitors this signal is decreased (Figure 1).



**Figure 1** The SYBR green assay<sup>1</sup>

The assay was run in the presence of a negative control constituted by the absence of ribonucleoside tri-phosphate (rNTPs) and a positive control represented by rNTPs. A known bacterial RNAP inhibitor represented by the antibiotic rifampicin, is also used to check the turnover of the assay. The full assay protocol is described in the next paragraph. A detergent, Triton X-100A, was present in the assay at the concentration of 0.01% to exclude aspecific inhibition activity due to aggregating compounds.<sup>2</sup>

The assay was developed by Prof. Ian Chopra and his collaborators<sup>3</sup> at the University of Leeds to screen compounds in a 384-well plate format for inhibition of bacterial RNAP.

Compounds showing good inhibition activity at 100  $\mu$ M (beyond 50% for this specified concentration), were shortlisted for a full dose-response curve determination in order to obtain the IC<sub>50</sub> value, the dose required for a 50% *in vitro* inhibition.

## **B) Experimental procedure for *E. coli* RNA Polymerase Assay<sup>4,5</sup>**

The ability of compounds to inhibit *E. coli* RNAP was determined in a 384-well plate (Greiner 781096) *in vitro* assay. Compounds, in a final concentration of 10% DMSO, were pre-incubated with buffer comprising 40mM TrisHCl, pH 7.5, 50mM KCl, 10mM MgCl<sub>2</sub>, 8mM DTT, 0.01% Triton X-100 with 20U/ml *E. coli* core RNA polymerase and 125ng/ml Kool<sup>TM</sup> NC-45<sup>TM</sup> Universal RNA polymerase template (Epicentre, Madison, WI, USA). The reaction was initiated with the addition of 0.5mM rNTPs (Roche Diagnostics Ltd., UK) and incubated for 2 h at 37°C and stopped with the addition of 20mM EDTA. RNA products were detected using SYBR Green I dye (Invitrogen Ltd., UK) in a PerkinElmer 2103 Multilabel reader with excitation and emission at 485nm and 531nm respectively. Assays were performed in duplicate and the % activity of compounds at 100 $\mu$ M was determined after deduction of background (no rNTPs) and comparison with no compound/DMSO control, designated having 100% activity. The IC<sub>50</sub> value against *E. coli* RNAP was determined using an 8-point 1:2 dilution series of compound and analysed using GraphPad Prism 6.

IC<sub>50</sub> values reported in this thesis are the mean value ± standard deviation of three independent measurements.

### **C) Selectivity and whole cell assays<sup>4,5</sup>**

Two structurally unrelated enzymes to bacterial RNAP, malate dehydrogenase and chymotrypsin, were used to assess selective inhibition of bacterial RNAP and to exclude promiscuous activity.<sup>6</sup> A detergent, Triton X-100A, was present in the assay at the concentration of 0.01% to exclude aspecific inhibition activity due to aggregating compounds.<sup>2</sup>

Antibacterial activity was then evaluated against a panel of Gram Positive and Gram Negative organisms. Minimum inhibitory concentration (MIC) was defined as the lowest concentration of a substance that inhibit visible growth of an organism after 18 hours of incubations at 37°C. Compounds were progressed only if they possessed a minimum suitable threshold, represented by MIC values against one or more test organisms comprised in a range between 32 and 64 µg/ml. The above cited minimum suitable threshold was established on the basis of possible improvement through analog refinement.<sup>7</sup>

### **D) General experimental procedure for the selectivity assays<sup>4,5</sup>**

Stocks of compounds were typically prepared at 10mM in 10% DMSO in malate dehydrogenase (MDH) and chymotrypsin assays and the results were controlled for the effect of DMSO.

Assays were performed in a 50 mM potassium phosphate solution containing 10% DMSO at pH 7.0. 100% activity control was constituted by absence of inhibitor into the above specified solution containing the enzyme while negative control consisted in no enzyme or inhibitor in the presence of 0.01% Triton-X100 at 25°C. All reactions were monitored on a Molecular Devices SPECTRAMax PLUS<sup>384</sup> spectrophotometer.

For MDH assays,<sup>6</sup> compound and 2nM enzyme were incubated in 50mM KPO<sub>4</sub> buffer (pH 7.35) containing 0.01% Triton-X100 for 5 minutes and the reaction initiated with 200nM oxaloacetic acid and 200nM NADH and progress

monitored at 340 nm. Oxaloacetic acid and NADH were each prepared as 2mM stocks in 50mM KPO<sub>4</sub> buffer (pH 7.35) containing 0.01% Triton-X100.

For chymotrypsin assays,<sup>8</sup> compound and 50nM enzyme were incubated for 5mins and the reaction initiated with 150nM succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide and the reaction progress was monitored at 410nm. Succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide was prepared as a 50mM stock in DMSO.

With regard to the MDH assay, oxaloacetic acid, NADH and MDH from porcine heart were purchased from Sigma-Aldrich.

For the Chymotrypsin assay,  $\alpha$ -chymotrypsin Type II from bovine pancreas and succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide were purchased from Sigma-Aldrich.

### E) Bacterial strains<sup>4,5</sup>

MIC determinations for compounds were performed by Liam Sharkey using the British Society for Antimicrobial Chemotherapy (BSAC) broth microdilution method<sup>9</sup> on selected bacterial strains (Tab. 1).

**Table 1** Selected bacterial strains description

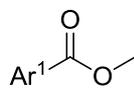
Strain	Description	Reference/Source
<i>S. aureus</i> SH1000	<i>rsbU</i> <sup>+</sup> derivative of 8325-4, common lab strain with genome sequenced	Horsburgh <i>et al</i> , <sup>10</sup>
<i>E. coli</i> SM1411	<i>lacI3, lacZ118, proB, trp, nalA, rpsL, ΔacrAB::Tn903kan<sup>r</sup></i> 1411 deficient in the AcrAB multidrug efflux pump component	O'Neill <i>et al</i> , <sup>11</sup>

## F) References

1. Ohmichi, T.; Maki, A.; Kool, E. T., *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 54-59.
2. McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K., *J. Med. Chem.* **2003**, *46*, 4265-4272.
3. Mariner, K. R.; Trowbridge, R.; Agarwal, A. K.; Miller, K.; O'Neill, A. J.; Fishwick, C. W. G.; Chopra, I., *Antimicrob. Agents Chemother.* **2010**, *54*, 4506-4509.
4. McPhillie, M. J.; Trowbridge, R.; Mariner, K. R.; O'Neill, A. J.; Johnson, A. P.; Chopra, I.; Fishwick, C. W. G., *ACS Med. Chem. Lett.* **2011**, *2*, 729-734.
5. Mariner, K.; McPhillie, M.; Trowbridge, R.; Smith, C.; O'Neill, A. J.; Fishwick, C. W. G.; Chopra, I., *Antimicrob. Agents and Chemother.* **2011**, *55*, 2413-2416.
6. Seidler, J.; McGovern, S. L.; Doman, T. N.; Shoichet, B. K., *J. Med. Chem.* **2003**, *46*, 4477-4486.
7. O'Neill, A. J.; Chopra, I., *Expert Opin. Investig. Drugs* **2004**, *13*, 1045-1063.
8. McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K., *J. Med. Chem.* **2002**, *45*, 1712-1722.
9. Phillips, I., *J. Antimicrob. Chemother.* **1991**, *27*, 1-50.
10. Horsburgh, M. J.; Aish, J. L.; White, I. J.; Shaw, L.; Lithgow, J. K.; Foster, S. J., *J. Bacteriol.* **2002**, *184*, 5457-5467.
11. O'Neill, A. J.; Bostock, J. M.; Moita, A. M.; Chopra, I., *J. Antimicrob. Chemother.* **2002**, *50*, 839-848.

## Appendix II

### A) Commercially available esters

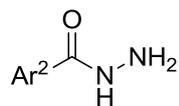


6.94-6.102

**Table 6.15** Commercially available esters **6.94—6.102**.

Compound number	Ar <sup>1</sup>	Compound number	Ar <sup>1</sup>
6.94		6.99	
6.95		6.100	
6.96		6.101	
6.97		6.102	
6.98			

## B) Synthesized hydrazides

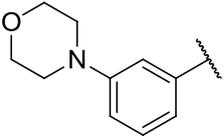
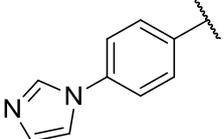


6.103-6.111

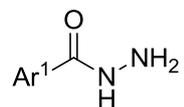
**Table 6.16** Synthesized hydrazides **6.103—6.111**.

Compound number	Ar <sup>2</sup>	Yield
6.103 <sup>1</sup>		89%
6.104 <sup>1</sup>		89%
6.105 <sup>2</sup>		93%
6.106 <sup>3</sup>		94%
6.107 <sup>4</sup>		85%
6.108 <sup>5</sup>		96%
6.109 <sup>6</sup>		93%

**Table 6.16** Continued

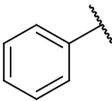
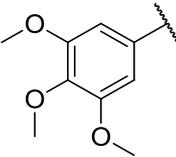
<b>6.110<sup>7</sup></b>		93%
<b>6.111<sup>8</sup></b>		95%

**C) Commercially available hydrazides**

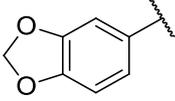
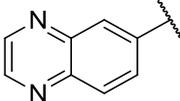
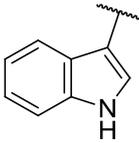


**6.112-6.116**

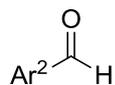
**Table 6.17** Commercially available hydrazides **6.112—6.116**.

<b>Compound number</b>	<b>Ar<sup>1</sup></b>
<b>6.112</b>	
<b>6.113</b>	

**Table 6.17** Continued

<p><b>6.114</b></p>	
<p><b>6.115</b></p>	
<p><b>6.116</b></p>	

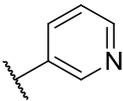
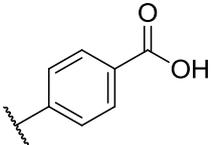
**D) Commercially available aldehydes**



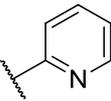
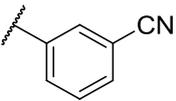
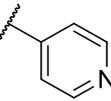
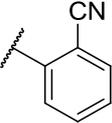
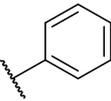
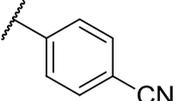
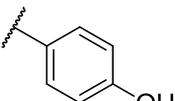
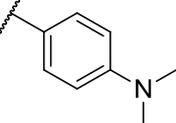
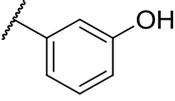
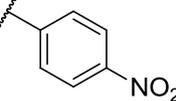
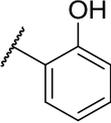
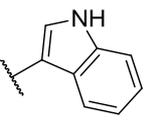
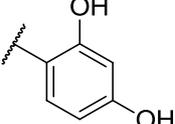
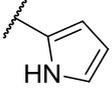
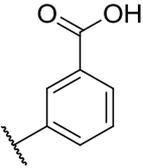
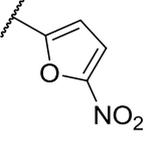
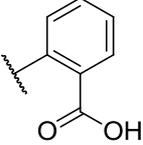
**6.33**

**6.117-6.134**

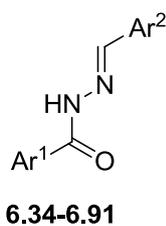
**Table 6.18** Commercially available aldehydes **6.33** and **6.117—6.134**.

Compound number	$\text{Ar}^2$	Compound number	$\text{Ar}^2$
<p><b>6.33</b></p>		<p><b>6.126</b></p>	

**Table 6.18 Continued**

6.117		6.127	
6.118		6.128	
6.119		6.129	
6.120		6.130	
6.121		6.131	
6.122		6.132	
6.123		6.133	
6.124		6.134	
6.125			

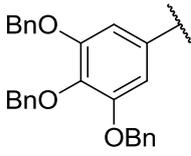
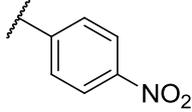
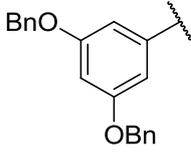
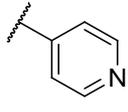
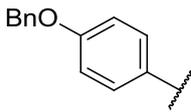
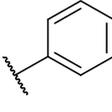
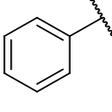
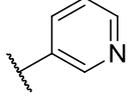
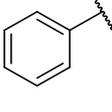
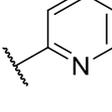
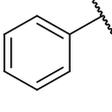
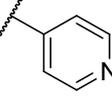
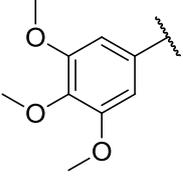
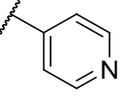
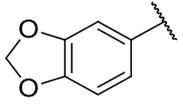
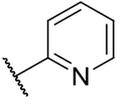
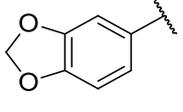
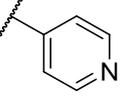
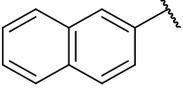
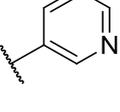
## E) Library of synthesised hydrazones



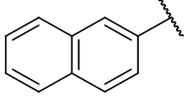
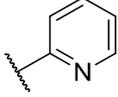
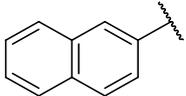
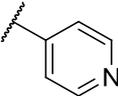
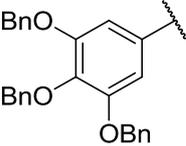
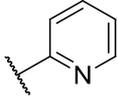
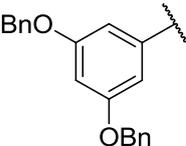
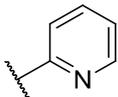
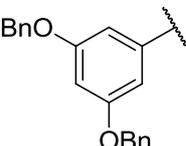
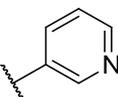
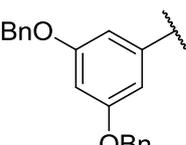
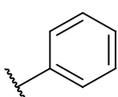
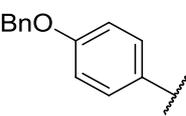
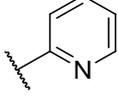
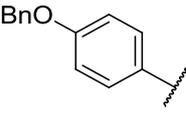
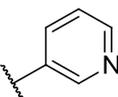
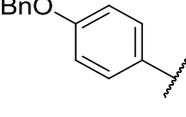
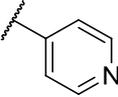
**Table 6.19** Library of synthesised hydrazones.

Compound number	Ar <sup>1</sup>	Ar <sup>2</sup>	Yield
<b>6.34<sup>6</sup></b>			94%
<b>6.35</b>			88%
<b>6.36</b>			80%
<b>6.37</b>			76%
<b>6.38</b>			82%

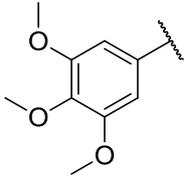
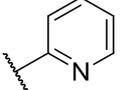
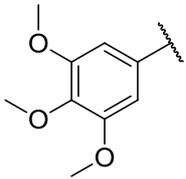
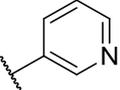
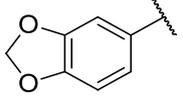
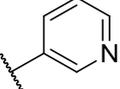
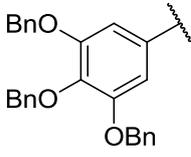
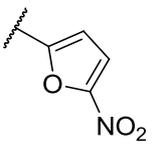
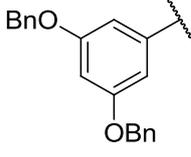
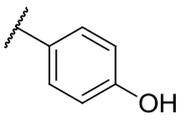
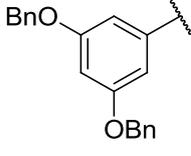
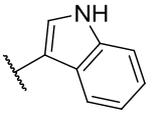
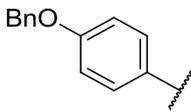
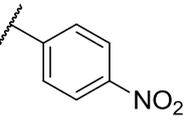
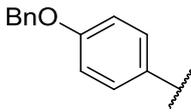
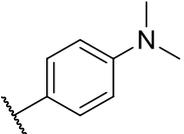
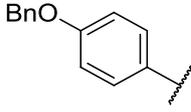
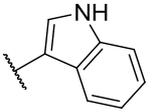
**Table 6.19** Continued

<b>6.39</b>			91%
<b>6.40<sup>6</sup></b>			88%
<b>6.41<sup>9</sup></b>			85%
<b>6.42<sup>10</sup></b>			88%
<b>6.43<sup>11</sup></b>			76%
<b>6.44<sup>10</sup></b>			80%
<b>6.45<sup>12</sup></b>			85%
<b>6.46<sup>10</sup></b>			64%
<b>6.47<sup>10</sup></b>			81%
<b>6.48<sup>13</sup></b>			73%

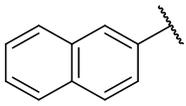
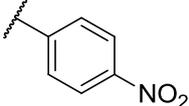
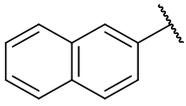
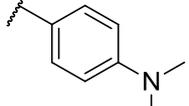
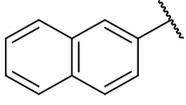
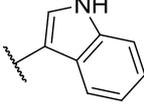
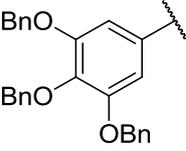
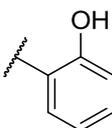
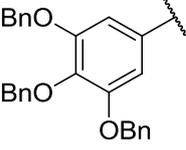
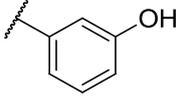
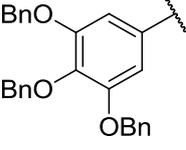
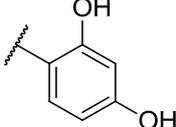
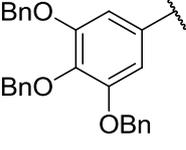
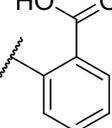
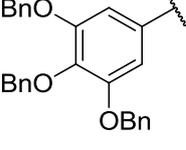
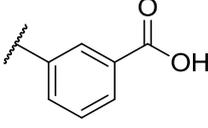
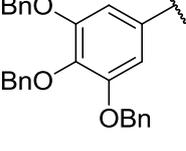
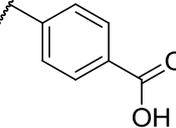
**Table 6.19** Continued

<b>6.49<sup>13</sup></b>			89%
<b>6.50<sup>13</sup></b>			79%
<b>6.51</b>			91%
<b>6.52</b>			90%
<b>6.53</b>			84%
<b>6.54</b>			76%
<b>6.55</b>			88%
<b>6.56</b>			79%
<b>6.57</b>			82%

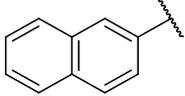
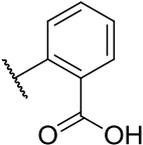
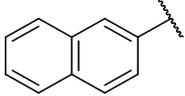
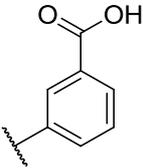
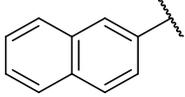
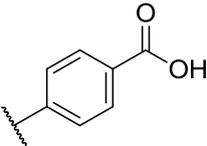
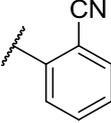
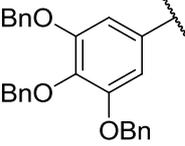
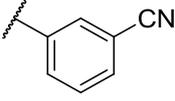
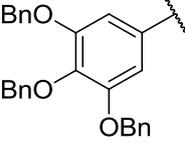
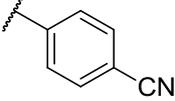
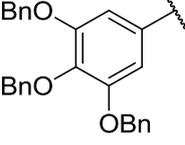
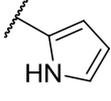
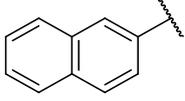
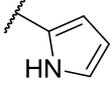
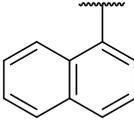
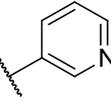
**Table 6.19** Continued

<b>6.58</b>			64%
<b>6.59</b>			72%
<b>6.60</b>			80%
<b>6.61</b>			87%
<b>6.62</b>			77%
<b>6.63</b>			75%
<b>6.64<sup>9</sup></b>			92%
<b>6.65</b>			80%
<b>6.66</b>			82%

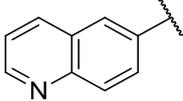
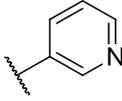
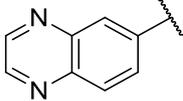
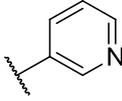
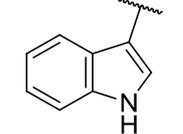
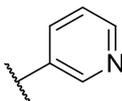
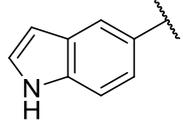
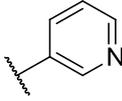
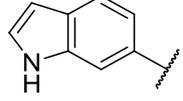
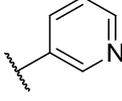
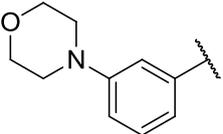
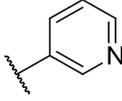
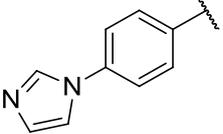
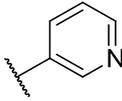
**Table 6.19** Continued

<b>6.67<sup>14</sup></b>			88%
<b>6.68</b>			80%
<b>6.69</b>			77%
<b>6.70</b>			74%
<b>6.71</b>			72%
<b>6.72</b>			75%
<b>6.73</b>			92%
<b>6.74</b>			85%
<b>6.75</b>			91%

**Table 6.19** Continued

<b>6.76</b>			92%
<b>6.77</b>			86%
<b>6.78<sup>14</sup></b>			91%
<b>6.79</b>			88%
<b>6.80</b>			90%
<b>6.81</b>			87%
<b>6.82</b>			87%
<b>6.83</b>			91%
<b>6.84</b>			87%

**Table 6.19** Continued

<b>6.85</b>			76%
<b>6.86</b>			80%
<b>6.87</b>			74%
<b>6.88</b>			77%
<b>6.89</b>			82%
<b>6.90a</b>			85%
<b>6.91a</b>			85%

## F) References

1. Gillig, R.; Kinnick, D.; Morin, J.; Navarro, A. WO2005/40157 A2, 2005.
2. Caboni, L.; Egan, B.; Kelly, B.; Blanco, F.; Fayne, D.; Meegan, M. J.; Lloyd, D. G., *J. Chem. Inf. Model.* **2013**, *53*, 2116-2130.
3. Mazouz, F.; Gueddari, S.; Burstein, C.; Mansuy, D.; Milcent, R., *J. Med. Chem.* **1993**, *36*, 1157-1167.
4. Saitoh, M.; Kunitomo, J.; Kimura, E.; Hayase, Y.; Kobayashi, H.; Uchiyama, N.; Kawamoto, T.; Tanaka, T.; Mol, C. D.; Dougan, D. R.; Textor, G. S.; Snell, G. P.; Itoh, F., *Bioorg. Med. Chem.* **2009**, *17*, 2017-2029.
5. Li, X.; Zhao, Z.; Li, G.; Shi, P., *J. Chem. Res.* **2010**, *34*, 410-413.
6. Rajamalli, P.; Atta, S.; Maity, S.; Prasad, E., *Chem. Commun.* **2013**, *49*, 1744-1746.
7. Vankayalapati, H.; Sorna, V.; Warner, S. L.; Bearss, D. J.; Sharma, S.; Stephens, B. WO2013/25805 A1, 2013.
8. Artico, M.; Silvestri, R.; Stefancich, G.; Avigliano, L.; Di Giulio, A.; Maccarrone, M.; Agostinelli, E.; Mondovi, B.; Morpurgo, L., *Eur. J. Med. Chem.* **1992**, *27*, 219-228.
9. Howlader, M. B. H.; Islam, M. S., *Indian J. Chem.* **2007**, *46*, 440-444.
10. Lima, P. C.; Lima, L. M.; da Silva, K. C. M.; Léda, P. H. O.; de Miranda, A. L. P.; Fraga, C. A. M.; Barreiro, E. J., *Eur. J. Med. Chem.* **2000**, *35*, 187-203.
11. Chaur, M. N.; Collado, D.; Lehn, J.-M., *Chem. – Eur. J.* **2011**, *17*, 248-258.
12. Mazzone, G.; Puglisi, G.; Marchetta, G.; Corsaro, A., *J. Het. Chem.* **1984**, *21*, 181-184.
13. Hu, Y.; Lu, X.; Chen, K.; Yan, R.; Li, Q.-S.; Zhu, H.-L., *Bioorg. Med. Chem.* **2012**, *20*, 903-909.
14. Avila, C. M.; Lopes, A. B.; Gonçalves, A. S.; da Silva, L. L.; Romeiro, N. C.; Miranda, A. L. P.; Sant'Anna, C. M. R.; Barreiro, E. J.; Fraga, C. A. M., *Eur. J. Med. Chem.* **2011**, *46*, 1245-1253.