

# **Using nature as a starting point for the development of novel antimicrobial agents**

**Christopher Ian Marriott**

School of Pharmacy

University of East Anglia



A thesis submitted for the degree of Doctor of Philosophy

July 2022

© "This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution."

## **DECLARATION**

This thesis is submitted to the University of East Anglia for the Degree of Doctor of Philosophy and has not been previously submitted at this or any University for assessment or for any other degree. Except where stated, and reference and acknowledgment is given, this work is original and has been carried out by the author alone.

Christopher Ian Marriott

## ABSTRACT

Antimicrobial resistance has caused an unprecedented need for the discovery of novel antimicrobial agents, especially those which are capable of acting against Gram-negative bacteria where strains have begun to emerge which have resistance pathways to all known antimicrobials. This can be achieved, broadly, by two means; either a new target, and hence a new class of antimicrobials, can be discovered and assessed, or by targeting an existing antimicrobial target but in such a way as to circumvent resistance. This project aimed to explore both means of overcoming resistance pathways, using two distinct targets.

Firstly, we investigated the inhibition of DNA gyrase, a topoisomerase enzyme found uniquely in bacteria which is essential for DNA replication. Using simocyclinone D8 (SD8), a natural product inhibitor of DNA gyrase that is not used therapeutically due to its poor *in vivo* activity, as a starting point, the work aimed to generate “pseudo-natural products” which were capable of emulating the activity of SD8 whilst being more readily synthetically prepared. Herein we illustrate an *in silico* fragment-based approach to the design of these mimics, with the selection of five angucyclinone-replacing alkyne fragments and five aminocoumarin-replacing amine fragments.

Secondly, a series of SD8-alkyne hybrids and “pseudo-natural products” were generated and examined in supercoiling assays against DNA gyrase, as well as other topoisomerase targets of SD8. Whilst these compounds did not possess any inhibitory activity, this work has furthered our understanding of DNA gyrase, and has laid the groundwork for future studies.

Finally, we investigated the inhibition of an important protein-protein interaction within the  $\beta$ -barrel assembly machinery (BAM) complex, an essential bacterial protein complex responsible for the insertion and correct folding of transmembrane  $\beta$ -barrel proteins in Gram-negative bacteria, using peptide fragments of important binding regions of BamA. The work aimed to use the peptide-directed binding methodology described in previous studies to generate small-molecule peptidomimetics capable of binding to BamD and inhibiting the complex. However, despite moderate success with the generation of these peptides, differential scanning fluorimetry studies revealed that the peptides selected were inefficient binders of BamD.

## **Access Condition and Agreement**

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

## TABLE OF CONTENTS

<b>DECLARATION</b> .....	<b>1</b>
<b>ABSTRACT</b> .....	<b>2</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>12</b>
<b>ABBREVIATIONS</b> .....	<b>13</b>
<b>CHAPTER 1: INTRODUCTION</b> .....	<b>17</b>
1.1. GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA .....	18
1.2. ANTIBIOTICS .....	19
1.3. ANTIMICROBIAL RESISTANCE .....	22
1.4. ESKAPE PATHOGENS .....	28
1.5. CURRENT CLINICAL TARGETS OF ANTIBIOTICS .....	29
1.6. PROTEIN-PROTEIN INTERACTIONS .....	36
1.7. DNA .....	38
1.8. GEOMETRY, TOPOLOGY AND HIGHER ORDER FEATURES OF DNA .....	40
1.9. DNA REPLICATION IN BACTERIA .....	43
1.10. TOPOISOMERASE ENZYMES .....	45
1.10.1. <i>Type IA</i> .....	46
1.10.2. <i>Type IB</i> .....	46
1.10.3. <i>Type IC</i> .....	46
1.10.4. <i>Type IIA</i> .....	47
1.10.5. <i>Type IIB</i> .....	47
1.11. DNA GYRASE .....	47
1.12. MECHANISM OF SUPERCOILING BY DNA GYRASE .....	48
1.13. CURRENT ANTIMICROBIALS TARGETING DNA GYRASE .....	49
1.13.1. <i>Quinolone antibiotics</i> .....	50
1.13.2. <i>Nybamycin antibiotics</i> .....	51
1.13.3. <i>Aminocoumarin antibiotics</i> .....	52
1.14. DRUG DISCOVERY APPROACHES .....	53
1.15. PEPTIDE-DIRECTED SYNTHESIS .....	56
1.16. AIMS OF THE PROJECT .....	57

## CHAPTER 2: FRAGMENT-BASED DESIGN OF SIMOCYCLINONE-INSPIRED NOVEL

<b>ANTIMICROBIAL AGENTS.....</b>	<b>59</b>
2.1. ISOLATION AND STRUCTURAL ELUCIDATION OF SIMOCYCLINONES SD4 AND SD8.....	60
2.2. MODE OF ACTION OF SD8 .....	62
2.3. BINDING OF SD8 TO DNA GYRASE .....	63
2.4. AIMS OF THE STUDY .....	64
2.5. COMPUTATIONAL DESIGN OF A REPLACEMENT LINKER.....	65
2.6. <i>IN SILICO</i> REPLACEMENT OF THE SD8 POLYKETIDE.....	70
2.7. <i>IN SILICO</i> REPLACEMENT OF THE SD8 AMINOCOUMARIN.....	75
2.8. CONCLUSIONS.....	79

## CHAPTER 3: AMINOCOUMARIN-LINKER-ALKYNE COMPOUNDS AS POTENTIAL MIMICS OF SD8..... 81

3.1. SYNTHESIS OF THE AMINOCOUMARIN.....	82
3.2. SYNTHESIS OF THE LINKER .....	87
3.3. GENERATION OF ANGUCYCLINONE-REPLACING ALKYNES .....	93
3.3.1. <i>Generation of alkynes 9 and 10</i> .....	94
3.3.2. <i>Generation of alkynes 11 and 12</i> .....	97
3.3.3. <i>Generation of alkyne 13</i> .....	99
3.4. ASSEMBLY OF AMINOCOUMARIN-LINKER-ALKYNE COMPOUNDS.....	103
3.5. BIOLOGICAL TESTING .....	107
3.6. CONCLUSIONS.....	110

## CHAPTER 4: SYNTHESIS OF “PSEUDO-NATURAL PRODUCT” COMPOUNDS BASED ON SIMOCYCLINONE D8..... 112

4.1. SYNTHESIS OF AMINOCOUMARIN-REPLACING AMINES .....	113
4.1.1. <i>Synthesis of compound 15</i> .....	113
4.1.2. <i>Synthesis of compound 16</i> .....	114
4.1.3. <i>Synthesis of compound 17</i> .....	118
4.1.4. <i>Synthesis of compound 18</i> .....	121
4.2. FINAL ASSEMBLY OF “PSEUDO-NATURAL PRODUCTS” .....	124
4.2.1. <i>Attachment of amines 14, 15, 17, 18 and 64 to linker 29</i> .....	124
4.2.2. <i>CuAAC attachment of alkyne 9 to amine-linker compounds 76-80</i> .....	129
4.3. BIOLOGICAL ASSESSMENT.....	134

4.3.1.	<i>Assessment of compounds 81, 82, and 84-86</i> .....	134
4.3.2.	<i>Further examination of all tested compounds</i> .....	137
4.4.	CONCLUSIONS.....	145
<b>CHAPTER 5: SYNTHESIS OF POTENTIAL BAMA/BAMD INTERACTION INHIBITORS.....</b>		<b>146</b>
5.1.	THE BAM COMPLEX.....	147
5.1.1.	<i>BamA</i> .....	148
5.1.2.	<i>BamB</i> .....	150
5.1.3.	<i>BamC</i> .....	152
5.1.4.	<i>BamD</i> .....	153
5.1.5.	<i>BamCD Complex</i> .....	155
5.1.6.	<i>BamE</i> .....	156
5.2.	TARGETING THE BAM COMPLEX.....	157
5.2.1.	<i>Monoclonal Antibodies (MAB1)</i> .....	158
5.2.2.	<i>Nitazoxanide</i> .....	159
5.2.3.	<i>Peptide binding of BamD</i> .....	160
5.3.	AIMS OF THE STUDY .....	161
5.4.	GENERATION OF BAMD SUBSTRATE SURFACE INHIBITORS.....	162
5.5.	INHIBITION OF THE BAMA/BAMD BINDING SURFACE .....	167
5.6.	<i>IN VITRO</i> ASSESSMENT .....	169
5.7.	CONCLUSIONS.....	172
<b>CHAPTER 6: CONCLUSIONS AND FUTURE WORK.....</b>		<b>174</b>
<b>CHAPTER 7: EXPERIMENTAL .....</b>		<b>179</b>
7.1.	GENERAL PROCEDURES.....	180
7.2.	CHROMATOGRAPHIC TECHNIQUES .....	180
7.3.	COMPUTATIONAL PROCEDURES.....	181
7.3.1.	<i>Preparation of the receptor and assessment of linker chain length</i> .....	181
7.3.2.	<i>Alkyne and amine fragment library generation and filtration</i> .....	181
7.3.3.	<i>Covalent docking</i> .....	182
7.3.4.	<i>Docking of fragment-SD8 libraries</i> .....	182
7.4.	SYNTHETIC PROCEDURES AND CHARACTERISATION.....	183
7.4.1.	<i>Solution phase synthesis</i> .....	183
7.4.2.	<i>Solid phase peptide synthesis</i> .....	209

7.5.	BIOLOGICAL PROCEDURES .....	211
7.5.1.	<i>Agarose gel electrophoresis</i> .....	211
7.5.2.	<i>BamD plasmid cloning</i> .....	211
7.5.3.	<i>BamD protein expression</i> .....	211
7.5.4.	<i>BamD buffer exchange</i> .....	213
7.5.5.	<i>Differential Scanning Fluorimetry buffer screening</i> .....	213
7.5.6.	<i>E. coli gyrase supercoiling assay</i> .....	213
7.5.7.	<i>E. coli Topo IV supercoiling relaxation assay</i> .....	214
7.5.8.	<i>S. aureus gyrase supercoiling assay</i> .....	215
7.5.9.	<i>Human Topo II<math>\alpha</math> and Topo II<math>\beta</math> decatenation assay</i> .....	215
7.5.10.	<i>Zone-of-inhibition assay</i> .....	216
<b>REFERENCES</b> .....		<b>217</b>

## FIGURES

<b>FIGURE 1.</b>	STRUCTURES OF GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIAL CELL WALLS.....	19
<b>FIGURE 2.</b>	THE STRUCTURES OF SALVARSAN AND NEOSALVARSAN .....	21
<b>FIGURE 3.</b>	THE STRUCTURE OF PENICILLIN G.....	21
<b>FIGURE 4.</b>	MOVEMENT OF ANTIBIOTICS INTO THE ENVIRONMENT .....	24
<b>FIGURE 5.</b>	SUMMARY OF EFFLUX PUMP CLASSES .....	27
<b>FIGURE 6.</b>	CLINICAL TARGETS OF ANTIBIOTICS AND EXAMPLES OF DRUGS. ....	30
<b>FIGURE 7.</b>	STRUCTURES OF THE CEPHALOSPORINS, CARBAPENEMS, AND AZTREONAM.....	31
<b>FIGURE 8.</b>	THE STRUCTURE OF VANCOMYCIN, A GLYCOPEPTIDE ANTIBIOTIC .....	31
<b>FIGURE 9.</b>	THE STRUCTURE OF GENTAMICIN, AN EXAMPLE OF AN AMINOGLYCOSIDE ANTIBIOTIC.....	32
<b>FIGURE 10.</b>	THE STRUCTURE OF TETRACYCLINE.....	33
<b>FIGURE 11.</b>	STRUCTURES OF ERYTHROMYCIN AND CLARITHROMYCIN .....	33
<b>FIGURE 12.</b>	GENERAL STRUCTURE OF THE LINCOSAMIDE ANTIBIOTICS.....	33
<b>FIGURE 13.</b>	THE MACROLIDE AND CYCLIC PEPTIDE COMPONENTS OF PRISTINAMYCIN .....	34
<b>FIGURE 14.</b>	THE STRUCTURE OF CHLORAMPHENICOL .....	34
<b>FIGURE 15.</b>	STRUCTURES OF LINEZOLID AND EPEREZOLID .....	35
<b>FIGURE 16.</b>	PRONTOSIL AND SULFAMETHOZAZOLE .....	35
<b>FIGURE 17.</b>	STRUCTURE OF TRIMETHOPRIM .....	35
<b>FIGURE 18.</b>	THE STRUCTURE OF RIFAMPICIN .....	36
<b>FIGURE 19.</b>	THE STRUCTURE OF DNA BASES AND THE BINDING OF NUCLEOTIDES TO FORM DNA .....	39
<b>FIGURE 20.</b>	BASE PAIR HYDROGEN BONDING.....	39

<b>FIGURE 21.</b> C2' ENDO AND C3' ENDO ARRANGEMENTS OF DNA SUGAR RING. ....	40
<b>FIGURE 22.</b> THE ANTI AND SYN CONFORMATIONS OF A PYRIMIDINE NUCLEOSIDE.....	40
<b>FIGURE 23.</b> STRUCTURES OF B-FORM, A-FORM AND Z-FORM DNA.....	42
<b>FIGURE 24.</b> DNA TWISTING AND WRITHING. ....	43
<b>FIGURE 25.</b> THE PROCESS OF DNA REPLICATION.....	44
<b>FIGURE 26.</b> MECHANISMS OF EACH CLASS OF TOPOISOMERASE ENZYME.....	45
<b>FIGURE 27.</b> SCHEMATIC OF THE DNA GYRASE A <sub>2</sub> B <sub>2</sub> COMPLEX .....	48
<b>FIGURE 28.</b> THE MECHANISM OF DNA GYRASE. ....	49
<b>FIGURE 29.</b> CLEAVAGE OF DNA BY DNA GYRASE. ....	49
<b>FIGURE 30.</b> EXAMPLES OF QUINOLONE ANTIBIOTICS.....	50
<b>FIGURE 31.</b> FRONT AND TOP VIEW OF DNA INTERCALATING WITH MOXIFLOXACIN WITHIN TOPO IV.....	51
<b>FIGURE 32.</b> THE STRUCTURES OF THE NYBOMYCIN SCAFFOLD, NYBOMYCIN AND DEOXYNYBOMYCIN .....	52
<b>FIGURE 33.</b> EXAMPLES OF AMINOCOUMARIN ANTIBIOTICS. ....	53
<b>FIGURE 34.</b> DRUG DESIGN AND DISCOVERY APPROACHES. ....	54
<b>FIGURE 35.</b> OUTLINE OF THE PEPTIDE-DIRECTED BINDING METHOD.....	57
<b>FIGURE 36.</b> STRUCTURE OF SIMOCYCLINONES SD8 AND SD4.....	60
<b>FIGURE 37.</b> STRUCTURES OF CLOROBIOCIN, FUMAGILLIN, AND AQUAYAMYCIN.....	61
<b>FIGURE 38.</b> GENERIC STRUCTURE OF SIMOCYCLINONES A-D. ....	61
<b>FIGURE 39.</b> VIEW OF THE GYRA-GYRA DIMER AS STAPLED BY SD8. ....	64
<b>FIGURE 40.</b> OUTLINE OF IN SILICO DESIGN OF SD8 MIMICS.....	65
<b>FIGURE 41.</b> BEST AMINOCOUMARIN-QUINOLONE HYBRID COMPOUND OBTAINED IN PREVIOUS WORK .....	66
<b>FIGURE 42.</b> AZIDO-CARBOXYLLIC ACID LINKERS EXPLORED DURING INITIAL DESIGN PROCESS. ....	66
<b>FIGURE 43.</b> OVERLAY OF RE-DOCKED SD8 ON SD8 BOUND TO GYRA IN CRYSTAL STRUCTURE .....	68
<b>FIGURE 44.</b> COMPOUNDS GENERATED IN SILICO AFTER "CORE-HOPPING" COMPOUNDS <b>1-4</b> INTO SD8.....	69
<b>FIGURE 45.</b> DOCKING POSES OF <b>5-8</b> COMPARED TO SD8 BOUND TO GYRA.....	69
<b>FIGURE 46.</b> SUMMARY OF ALKYNE LIBRARY PREPARATION WORK.....	71
<b>FIGURE 47.</b> STRUCTURE OF THE AMINOCOUMARIN-LINKER-AZIDE IN SILICO COMPOUND. ....	71
<b>FIGURE 48.</b> EXAMPLES OF ERRONEOUS STRUCTURES GENERATED BY COVALENT DOCKING .....	72
<b>FIGURE 49.</b> ALKYNE LIBRARY MEMBERS SELECTED FOR SYNTHESIS AND STUDY,.....	74
<b>FIGURE 50.</b> IN SILICO DOCKED POSES OF SD8-FRAGMENT HYBRIDS DERIVED FROM COMPOUNDS <b>9-13</b> .....	74
<b>FIGURE 51.</b> STRUCTURE OF THE LINKER-TRIAZOLE-ANGUCYCLINONE IN SILICO COMPOUND. ....	76
<b>FIGURE 52.</b> GENERAL STRUCTURE OF SD8-AMINE LIBRARY HYBRIDS .....	77
<b>FIGURE 53.</b> AMINE LIBRARY MEMBERS SELECTED FOR SYNTHESIS AND STUDY.....	78
<b>FIGURE 54.</b> IN SILICO DOCKED POSES OF SD8-FRAGMENT HYBRIDS DERIVED FROM COMPOUNDS <b>14-18</b> ...	78

<b>FIGURE 55. SELECTED LIBRARY MEMBERS 9-13</b> .....	94
<b>FIGURE 56. EFFECTS OF COMPOUNDS 52-55 ON DNA SUPERCOILING BY WT E. COLI GYRASE</b> .....	108
<b>FIGURE 57. ZONE-OF-INHIBITION ASSAY PLATES FOR COMPOUNDS 52 AND 53</b> .....	109
<b>FIGURE 58. ZONE-OF-INHIBITION ASSAY PLATES FOR COMPOUNDS 54 AND 55</b> .....	109
<b>FIGURE 59. AMINE FRAGMENTS SELECTED FOR FURTHER STUDY</b> .....	113
<b>FIGURE 60. EFFECTS OF COMPOUNDS 81, 82 AND 84 ON DNA SUPERCOILING BY WT E. COLI GYRASE</b> ...	134
<b>FIGURE 61. EFFECTS OF COMPOUNDS 85 AND 86 ON DNA SUPERCOILING BY WT E. COLI GYRASE</b> .....	134
<b>FIGURE 62. ZONE-OF-INHIBITION ASSAY PLATES FOR COMPOUNDS 81 AND 82</b> .....	135
<b>FIGURE 63. ZONE-OF-INHIBITION ASSAY PLATES FOR COMPOUNDS 84 AND 85</b> .....	136
<b>FIGURE 64. ZONE-OF-INHIBITION ASSAY PLATES FOR COMPOUND 86</b> .....	136
<b>FIGURE 65. EFFECTS OF 52-55 ON DNA SUPERCOILING RELAXATION BY WT E. COLI TOPO IV</b> .....	138
<b>FIGURE 66. EFFECTS OF 81, 82 AND 84 ON DNA SUPERCOILING RELAXATION BY WT E. COLI TOPO IV</b> ...	138
<b>FIGURE 67. EFFECTS OF 85 AND 86 ON DNA SUPERCOILING RELAXATION BY WT E. COLI TOPO IV</b> .....	138
<b>FIGURE 68. EFFECTS OF 52-55 ON DNA SUPERCOILING BY WT S. AUREUS GYRASE</b> .....	139
<b>FIGURE 69. EFFECTS OF 81, 82 AND 84 ON DNA SUPERCOILING BY WT S. AUREUS GYRASE</b> .....	139
<b>FIGURE 70. EFFECTS OF 85 AND 86 ON DNA SUPERCOILING BY WT S. AUREUS GYRASE</b> .....	140
<b>FIGURE 71. EFFECTS OF 52-55 ON kDNA DECATENATION BY HUMAN TOPO IIA</b> .....	141
<b>FIGURE 72. EFFECTS OF 81, 82 AND 84 ON kDNA DECATENATION BY HUMAN TOPO IIA</b> .....	142
<b>FIGURE 73. EFFECTS OF 85 AND 86 ON kDNA DECATENATION BY HUMAN TOPO IIA</b> .....	142
<b>FIGURE 74. EFFECTS OF 52-55 ON kDNA DECATENATION BY HUMAN TOPO IIB</b> .....	143
<b>FIGURE 75. EFFECTS OF 81, 82 AND 84 ON kDNA DECATENATION BY HUMAN TOPO IIB.</b> .....	144
<b>FIGURE 76. EFFECTS OF 85 AND 86 ON kDNA DECATENATION BY HUMAN TOPO IIB</b> .....	144
<b>FIGURE 77. STRUCTURE OF THE E. COLI BAM COMPLEX</b> .....	147
<b>FIGURE 78. A CARTOON SCHEMATIC OF THE BAM COMPLEX</b> .....	148
<b>FIGURE 79. STRUCTURE OF BAMA SHOWING THE B-BARREL AND POTRA DOMAINS 1-5</b> .....	149
<b>FIGURE 80. PROPOSED MECHANISMS OF OMP INSERTION</b> .....	150
<b>FIGURE 81. THE TOP-DOWN AND EDGE-ON STRUCTURES OF BAMB</b> .....	151
<b>FIGURE 82. THE C-TERMINAL AND N-TERMINAL REGIONS OF BAMC</b> .....	152
<b>FIGURE 83. CRYSTAL STRUCTURE OF BAMD</b> .....	153
<b>FIGURE 84. ELECTROSTATIC POTENTIAL SURFACE OF BAMD</b> .....	154
<b>FIGURE 85. THE BAMCD COMPLEX</b> .....	156
<b>FIGURE 86. A CRYSTAL STRUCTURE OF THE BAME LIPOPROTEIN</b> .....	157
<b>FIGURE 87. BAMA HIGHLIGHTING THE POSITION OF EXTRACELLULAR LOOP 4 WRT THE LATERAL GATE</b> .....	159
<b>FIGURE 88. THE STRUCTURE OF NITAZOXANIDE</b> .....	159

<b>FIGURE 89.</b> SECTIONS OF THE BAMA BARREL.....	161
<b>FIGURE 90.</b> STRUCTURE OF WANG RESIN AND DEPROTECTION TO FORM PEPTIDE ACIDS.....	163
<b>FIGURE 91.</b> STRUCTURE OF 2-CHLOROTRITYL RESIN.....	164
<b>FIGURE 92.</b> ANALYTICAL HPLC TRACE OF PEPTIDE <b>88</b> AFTER PURIFICATION.....	165
<b>FIGURE 93.</b> ANALYTICAL HPLC TRACE OF <b>87</b> BEFORE PURIFICATION.....	165
<b>FIGURE 94.</b> STRUCTURE OF THE POTRA5-BAMD INTERACTION SURFACE.....	168
<b>FIGURE 95.</b> ANALYTICAL HPLC TRACE OF <b>89</b> BEFORE PURIFICATION.....	169
<b>FIGURE 96.</b> STRUCTURE OF SYPRO ORANGE.....	170
<b>FIGURE 97.</b> REDUCTION OF MTT TO FORMAZAN.....	176
<b>FIGURE 98.</b> OVERVIEW OF SURFACE PLASMON RESONANCE.....	177

## **SCHEMES**

<b>SCHEME 1.</b> GENERATION OF AMINOCOUMARIN <b>27</b> .....	82
<b>SCHEME 2.</b> SYNTHESIS OF 1,3-DIETHYL-2-[[[(TERT-BUTOXY)CARBONYL]AMINO]PROPANEDIOATE ( <b>20</b> ).....	83
<b>SCHEME 3.</b> SYNTHESIS OF COMPOUND <b>21</b> .....	84
<b>SCHEME 4.</b> SYNTHESIS OF 2,4-BIS(ACTYLOXY) BENZOIC ACID ( <b>23</b> ).....	84
<b>SCHEME 5.</b> SYNTHESIS OF 2,4-BIS(ACETYLOXY) BENZOYL CHLORIDE ( <b>24</b> ).....	85
<b>SCHEME 6.</b> MECHANISM OF FORMATION OF ACYL CHLORIDE ( <b>28</b> ).....	85
<b>SCHEME 7.</b> SYNTHESIS OF COMPOUND <b>25</b> .....	86
<b>SCHEME 8.</b> CONJUGATE ADDITION OF MALONATE <b>25</b> TO ACYL CHLORIDE <b>24</b> .....	86
<b>SCHEME 9.</b> SYNTHESIS OF COMPOUND <b>26</b> .....	86
<b>SCHEME 10.</b> DEACETYLATION OF <b>25</b> AND FORMATION OF <b>26</b> .....	86
<b>SCHEME 11.</b> SYNTHESIS OF 4,7-DIHYDROXY-2-OXO-2H-CHROMEN-3-AMINIUM CHLORIDE ( <b>27</b> ).....	87
<b>SCHEME 12.</b> PROPOSED SYNTHESIS OF SD8-LIKE MOLECULES.....	88
<b>SCHEME 13.</b> SYNTHESIS OF 10-AZIDODECANOIC ACID ( <b>29</b> ).....	88
<b>SCHEME 14.</b> SYNTHESIS OF 10-AZIDODECANOYL CHLORIDE ( <b>30</b> ).....	89
<b>SCHEME 15.</b> ATTEMPTED SYNTHESIS OF COMPOUND <b>31</b> .....	89
<b>SCHEME 16.</b> PROPOSED PYRIDINE-CATALYSED FORMATION OF <b>31</b> .....	89
<b>SCHEME 17.</b> ATTEMPTED SYNTHETIC ROUTES TO <b>31</b> .....	90
<b>SCHEME 18.</b> COMPLETE SYNTHESIS PATHWAY OF 3-AMINOCOUMARIN ( <b>32</b> ).....	90
<b>SCHEME 19.</b> SYNTHESIS OF N-ACETYL-3-AMINOCOUMARIN ( <b>34</b> ).....	91
<b>SCHEME 20.</b> MECHANISM OF THE MODIFIED PERKIN REACTION TO FORM <b>34</b> .....	91
<b>SCHEME 21.</b> SYNTHESIS OF 3-AMINOCOUMARIN ( <b>32</b> ).....	92
<b>SCHEME 22.</b> SYNTHESIS OF 10-AZIDO-N-(2-OXO-2H-CHROMEN-3-YL)DECANAMIDE ( <b>35</b> ).....	93

<b>SCHEME 23.</b> GENERAL ROUTE TO ALKYNES <b>9</b> AND <b>10</b> .....	95
<b>SCHEME 24.</b> SYNTHESIS OF 5-OXO-5-(PROP-2-YN-1-YLOXY)PENTANOIC ACID ( <b>38</b> ) .....	95
<b>SCHEME 25.</b> ATTEMPTED SYNTHETIC ROUTES OF COMPOUND <b>9</b> .....	95
<b>SCHEME 26.</b> SYNTHESIS OF PROP-2-YN-1-YL-4-[(NAPHTHALEN-1-YL)CARBAMOYL]BUTANOATE ( <b>9</b> ).....	96
<b>SCHEME 27.</b> ATTEMPTED SYNTHESSES OF COMPOUND <b>10</b> .....	97
<b>SCHEME 28.</b> GENERAL ROUTE FOR THE GENERATION OF ALKYNES <b>11</b> AND <b>12</b> .....	97
<b>SCHEME 29.</b> SYNTHESIS OF 3-METHOXY-4-(PROP-2-YN-1-YLOXY)BENZALDEHYDE ( <b>44</b> ) .....	97
<b>SCHEME 30.</b> SYNTHESIS OF ALKYNE LIBRARY MEMBERS <b>11</b> AND <b>12</b> .....	98
<b>SCHEME 31.</b> GENERAL ROUTE FOR GENERATION OF ALKYNE <b>13</b> .....	100
<b>SCHEME 32.</b> ATTEMPTED SYNTHESSES OF INTERMEDIATE AMIDE <b>48</b> .....	100
<b>SCHEME 33.</b> SYNTHESIS OF 4-(PROP-2-YN-1-YLOXY)BENZOIC ACID ( <b>49</b> ) .....	100
<b>SCHEME 34.</b> ATTEMPTED SYNTHESSES OF AMIDE <b>13</b> .....	101
<b>SCHEME 35.</b> METHOD OF CDI AND IMIDAZOLIUM EFFECT ACTIVATION OF <b>49</b> TO HELP FORM AMIDE <b>13</b> .	102
<b>SCHEME 36.</b> SYNTHESIS OF N-(4-CYANOPHENYL)-4-(PROP-2-YN-1-YLOXY)BENZAMIDE ( <b>13</b> ).....	102
<b>SCHEME 37.</b> CONDITIONS USED FOR THE CUAAC REACTIONS OF <b>35</b> WITH ALKYNES <b>9, 11-13</b> .....	103
<b>SCHEME 38.</b> SYNTHESIS OF CLICK COMPOUND <b>52</b> .....	104
<b>SCHEME 39.</b> SYNTHESIS OF CLICK COMPOUND <b>53</b> .....	105
<b>SCHEME 40.</b> SYNTHESIS OF CLICK COMPOUND <b>54</b> .....	106
<b>SCHEME 41.</b> SYNTHESIS OF CLICK COMPOUND <b>55</b> .....	107
<b>SCHEME 42.</b> SYNTHESIS OF {[1,1'-BIPHENYL]-4-YL}METHYL(METHYL)AMINE ( <b>15</b> ).....	114
<b>SCHEME 43.</b> SYNTHESIS OF REGIOISOMERS <b>60</b> AND <b>61</b> .....	114
<b>SCHEME 44.</b> GENERATION OF THE DOUBLE BOC-PROTECTED INTERMEDIATE ( <b>63</b> ) .....	115
<b>SCHEME 45.</b> ATTEMPTED SELECTIVE BOC DEPROTECTION OF COMPOUND <b>63</b> .....	116
<b>SCHEME 46.</b> SYNTHESIS OF CBZ-PROTECTED L-NORVALINE ( <b>66</b> ).....	116
<b>SCHEME 47.</b> SYNTHESIS OF COMPOUND <b>67</b> .....	117
<b>SCHEME 48.</b> SYNTHESIS OF COMPOUND <b>64</b> .....	118
<b>SCHEME 49.</b> SYNTHETIC ROUTE TO COMPOUND <b>17</b> .....	119
<b>SCHEME 50.</b> SYNTHESIS OF COMPOUND <b>70</b> .....	119
<b>SCHEME 51.</b> SYNTHESIS OF COMPOUND <b>72</b> .....	120
<b>SCHEME 52.</b> SYNTHESIS OF COMPOUND <b>17</b> .....	120
<b>SCHEME 53.</b> SYNTHETIC ROUTE TO AMINE <b>18</b> .....	121
<b>SCHEME 54.</b> SYNTHESIS OF INTERMEDIATE COMPOUND <b>75</b> .....	121
<b>SCHEME 55.</b> POSSIBLE SITES OF ATTACK ON ETHYLENE CARBONATE ( <b>74</b> ) FOR PHENOL <b>73</b> .....	122
<b>SCHEME 56.</b> MECHANISM OF THE MITSUNOBU REACTION .....	123

<b>SCHEME 57. MECHANISM OF AMINE CLEAVAGE USING HYDRAZINE</b> .....	123
<b>SCHEME 58. GENERATION OF AMINE 18</b> .....	123
<b>SCHEME 59. GENERAL SYNTHESIS OF AMIDES 76-80</b> .....	124
<b>SCHEME 60. SYNTHESIS OF AMIDE 76</b> .....	125
<b>SCHEME 61. GENERATION OF COMPOUND 77</b> .....	125
<b>SCHEME 62. SYNTHESIS OF AMIDE 78</b> .....	126
<b>SCHEME 63. GENERATION OF COMPOUND 79</b> .....	127
<b>SCHEME 64. SYNTHESIS OF AMIDE 80</b> .....	128
<b>SCHEME 65. GENERAL ROUTE FOR THE GENERATION OF 81, 82, AND 84-86</b> .....	129
<b>SCHEME 66. SYNTHESIS OF COMPOUND 81 VIA COPPER-CATALYSED CLICK CHEMISTRY</b> .....	130
<b>SCHEME 67. SYNTHESIS OF COMPOUND 82 USING COPPER-CATALYSED CLICK CHEMISTRY</b> .....	130
<b>SCHEME 68. SYNTHESIS OF COMPOUND 83 USING COPPER-CATALYSED CLICK CHEMISTRY</b> .....	131
<b>SCHEME 69. DEPROTECTION OF COMPOUND 83 TO GENERATE COMPOUND 84</b> .....	132
<b>SCHEME 70. SYNTHESIS OF COMPOUND 85 BY COPPER-CATALYSED CLICK CHEMISTRY</b> .....	133
<b>SCHEME 71. SYNTHESIS OF COMPOUND 86 BY COPPER-CATALYSED CLICK CHEMISTRY</b> .....	133
<b>SCHEME 72. SOLID PHASE PEPTIDE SYNTHESIS COUPLING STEP</b> .....	162
<b>SCHEME 73. FMOC DEPROTECTION OF PEPTIDES USING PIPERIDINE</b> .....	163

## TABLES

<b>TABLE 1. MAIN CLASSES OF ANTIBIOTICS</b> .....	30
<b>TABLE 2. EXAMPLES OF SIMOCYCLINONES A-D</b> .....	61
<b>TABLE 3. GLIDESCORES OF COMPOUNDS 5-8</b> .....	70
<b>TABLE 4. SEQUENCES OF BAMD INHIBITORS BASED ON THE BAMA BARREL</b> .....	162
<b>TABLE 5. LIST OF CLEAVAGE COCKTAILS USED FOR THE ATTEMPTED SYNTHESIS OF 87 AND 88</b> .....	166
<b>TABLE 6. SUGGESTED SEQUENCES FOR THE BAMA-D INTERACTION SURFACE INHIBITORS</b> .....	168

## ACKNOWLEDGEMENTS

First of all, I'd like to thank Prof. Mark Searcey and Prof. Andrew Hemmings for allowing me to do this work, and for all of their help throughout the various projects that I've been involved in. Their passion for their subjects gave me the inspiration and motivation to help me through the more difficult moments of the PhD, and without their support there's no way that I would have finished.

I'd also like to thank Prof. Changjiang Dong and his PhD student, James Coleman, for allowing me access to the biology labs, as well as for all of their assistance in the expression and purification of the BamD protein samples. On the same vein, I'd like to thank Prof. Tony Maxwell and his research assistant, Lesley Mitchenall, for allowing me to work in their labs and for their assistance in performing the supercoiling assays. Thank you to you all for your patience and understanding whilst training me. Furthermore, I'd like to thank the UKRI-BBSRC Norwich Research Park Biosciences Doctoral Training Partnership for providing funding to support the project.

To my friends, family and loved ones, you will never know how grateful I am for your never ending support throughout the past four years. Thank you for putting up with me during my worst moments, and for sharing in the high moments. Most importantly though, thank you for believing in me, especially when I didn't believe in myself.

Finally, but most certainly not least, thank you to the members of the Searcey and Beekman groups, both past and present. I hope I haven't ruined the Youtube suggestions on the lab computer too much! Thank you for all of your guidance, banter and support. I won't forget all that you've done for me.

## ABBREVIATIONS

<b>Å</b>	Angstrom
<b>ABC</b>	ATP-binding cassette
<b>Ac</b>	Acetyl
<b>ADP</b>	Adenosine diphosphate
<b>ATP</b>	Adenosine triphosphate
<b>BAM</b>	B-barrel assembly machinery
<b>BLAST</b>	Basic local alignment search tool
<b>Boc</b>	<i>tert</i> -butoxycarbonyl
<b>br</b>	Broad
<b>Cbz</b>	Benzyl carbamate
<b>CDI</b>	Carbonyl diimidazole
<b>CFX</b>	Ciprofloxacin
<b>CuAAC</b>	Copper-catalysed azide-alkyne cycloaddition
<b>°C</b>	Degrees Celsius
<b>Da</b>	Dalton
<b>DCC</b>	N,N'-Dicyclohexylcarbodiimide
<b>DCM</b>	Dichloromethane
<b>DIAD</b>	Diisopropyl azodicarboxylate
<b>DIPEA</b>	N,N-Diisopropylethylamine
<b>DMAP</b>	4-dimethylaminopyridine
<b>DMF</b>	Dimethylformamide
<b>DMSO</b>	Dimethyl sulfoxide
<b>DNA</b>	Deoxyribonucleic acid
<b>DSF</b>	Differential scanning fluorimetry
<b>EDC</b>	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>ESBL</b>	Extended-spectrum $\beta$ -lactamases
<b>ESI</b>	Electrospray Ionisation
<b>Eq</b>	Equivalents
<b>ESKAPE</b>	<i>Enterococcus faecium</i> , <i>Staphylococcus aureus</i> , <i>Klebsiella pneumoniae</i> , <i>Acinetobacter baumannii</i> , <i>Pseudomonas aeruginosa</i> , <i>Enterobacter</i> species

## ABBREVIATIONS

<b>Et</b>	Ethyl
<b>FBDD</b>	Fragment based drug design
<b>FDA</b>	Food and Drug Administration
<b>Fmoc</b>	Fluorenylmethoxycarbonyl
<b>Glide</b>	Grid-based Ligand Docking with Energetics
<b>Gyr</b>	Gyrase
<b>h</b>	Hour
<b>HBTU</b>	Hexafluorophosphate benzotriazole tetramethyl uronium
<b>HOBT</b>	Hydroxybenzotriazole
<b>HPLC</b>	High-pressure liquid chromatography
<b>HTS</b>	High-throughput screening
<b>IgG</b>	Immunoglobulin G
<b>Imid</b>	Imidazole
<b>IPTG</b>	Isopropyl $\beta$ -d-1-thiogalactopyranoside
<b>IR</b>	Infrared spectroscopy
<b>ITC</b>	Isothermal titration calorimetry
<b>kDNA</b>	Kinetoplast DNA
<b>Lk</b>	Linking number
<b>LPS</b>	Lipopolysaccharide
<b>M</b>	Molar
<b>mAb</b>	Monoclonal antibody
<b>MALDI</b>	Matrix-assisted laser desorption/ionisation
<b>MATE</b>	Multidrug and toxic compound extrusion
<b>MDR</b>	Multi-drug resistant
<b>Me</b>	Methyl
<b>MFS</b>	Major facilitator superfamily
<b>mmol</b>	Millimoles
<b>Mp</b>	Melting point
<b>mRNA</b>	Messenger RNA
<b>MRSA</b>	Methicillin-resistant <i>Staphylococcus aureus</i>
<b>MTT</b>	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
<b><math>\mu</math>M</b>	Micromolar
<b>NADPH</b>	Nicotinamide Adenine Dinucleotide Phosphate
<b>NCBI</b>	National Center for Biotechnology Information

## ABBREVIATIONS

<b>NMR</b>	Nuclear Magnetic Resonance
<b>NOESY</b>	Nuclear Overhauser Effect spectroscopy
<b>Nu</b>	Nucleophile
<b>OM</b>	Outer membrane
<b>OMP</b>	Outer membrane protein
<b>PABA</b>	Para-aminobenzoic acid
<b>PAGE</b>	Polyacrylamide gel electrophoresis
<b>PAINS</b>	Pan-assay interference compounds
<b>PBP</b>	Penicillin-binding protein
<b>Ph</b>	Phenyl
<b>POTRA</b>	Polypeptide translocation-associated domain
<b>PPI</b>	Protein-protein interaction
<b>ppm</b>	Parts per million
<b>qPCR</b>	Quantitative polymerase chain reaction
<b>QSAR</b>	Quantitative structure-activity relationship
<b>REPLACE</b>	Replacement with partial ligand alternatives through computational enrichment
<b>RMSD</b>	Root-mean-square deviation
<b>RNA</b>	Ribonucleic acid
<b>RND</b>	Resistance-nodulation-cell division
<b>RP-HPLC</b>	Reverse phase HPLC
<b>rRNA</b>	Ribosomal RNA
<b>rt</b>	Room temperature
<b>SD4</b>	Simocyclinone D4
<b>SD8</b>	Simocyclinone D8
<b>SDS</b>	Sodium Dodecyl Sulfate
<b>SMARTS</b>	SMILES arbitrary target specification
<b>SMILES</b>	Simplified Molecular-Input Line-Entry System
<b>SMR</b>	Small multidrug resistance
<b>SPPS</b>	Solid-phase peptide synthesis
<b>SPR</b>	Surface-plasmon resonance
<b>STEB</b>	Sucrose, Tris hydrochloride, EDTA, bromophenol blue
<b><i>t</i>Bu</b>	<i>Tert</i> -butyl
<b>TFA</b>	Trifluoroacetic acid

## ABBREVIATIONS

<b>THF</b>	Tetrahydrofuran
<b>TIPS</b>	Triisopropylsilane
<b>TLC</b>	Thin-layer chromatography
<b>T<sub>m</sub></b>	Melting temperature
<b>TOF</b>	Time-of-flight
<b>Topo</b>	Topoisomerase
<b>TOPRIM</b>	Topo/primase
<b>T<sub>m</sub></b>	Melting temperature
<b>TPR</b>	Tetratricopeptide repeat
<b>tRNA</b>	Transfer RNA
<b>Trt</b>	Trityl
<b>Tw</b>	Twist
<b>UV</b>	Ultraviolet
<b>VRE</b>	Vancomycin-resistant Enterococci
<b>WHO</b>	World Health Organisation
<b>Wr</b>	Writhe

# **CHAPTER 1: INTRODUCTION**

## CHAPTER 1: INTRODUCTION

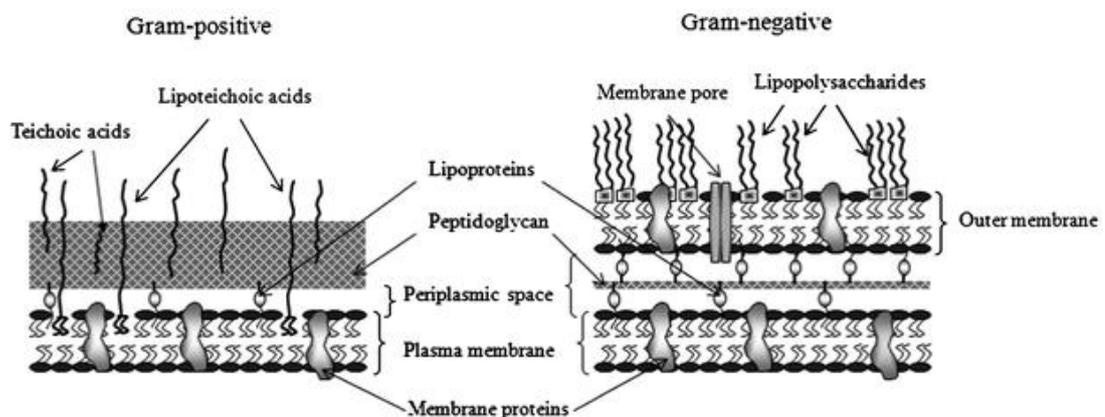
This chapter introduces bacteria, antibiotics and the resistance pathways of bacteria to these antibiotics. It also aims to orientate the reader to the multiple potential arrangements of DNA, showing the flexibility and variation that allows DNA to act as the fundamental building block to all life, as well as protein-protein interactions, highlighting to the reader that these interactions provide good targets for pharmaceuticals. Furthermore, this chapter covers DNA topology, including topoisomerase enzymes, and discusses inhibitors of these enzymes, before discussing potential approaches to creating new inhibitors.

### 1.1. Gram-positive and Gram-negative bacteria

All bacteria can be separated into two main categories; Gram-positive and Gram-negative. This classification system, developed by and named after Hans Christian Gram, separates bacteria by the colour observed when cells are dyed using crystal violet dye, and hence by the structural and physical characteristics of their cell walls. Gram-positive bacteria, as the name suggests, retain the crystal violet dye well, appearing purple following the procedure. This is due to the presence of a thick, rigid peptidoglycan layer, able to retain the dye, which surrounds the cell membrane of the bacteria. Peptidoglycan is made up of a series of linear glycan strands, made of alternating sugars N-acetylglucosamine and N-acetylmuramic acid which are linked by  $\beta$ -1,4 bonds. These glycan strands are cross-linked by short, 5-membered peptides, attaching to the D-lactoyl group of the muramic acid, and are most often composed of L-Ala,  $\gamma$ -D-Glu, *meso*-2,6-diaminopimelic acid ( $A_2pm$ ) or L-Lys, and two D-Ala residues, the last of which is lost in the matured polymer. These peptide crosslinks are added to the glycan strands by Mur ligases, with the two D-Ala residues being added as a dipeptide. The crosslinking occurs between the terminal D-Ala of one peptide chain and the  $A_2pm$  or L-Lys of another chain for Gram-positive and Gram-negative bacteria respectively.<sup>1</sup> Peptidoglycans not only provide the shape of the bacteria, but also allow for proteins and other cellular components to be anchored to the surface of the bacteria.

## CHAPTER 1: INTRODUCTION

Whilst Gram-positive bacteria possess a thick peptidoglycan layer, Gram-negative bacteria possess only a thin layer of peptidoglycan. This thin layer is unable to retain the crystal violet dye during Gram staining, instead appearing pink as a result of the counterstaining process (a negative test result). The thin peptidoglycan layer is encapsulated by a second, outer membrane (OM), with a periplasmic space separating the two cell walls. This OM, contrasting the cell membrane of Gram-positive bacteria, is decorated predominantly with lipopolysaccharides (LPS) in addition to proteins (termed outer membrane proteins, OMPs) and phospholipids.<sup>2</sup> LPS is made up of three components; a lipid which is used to anchor the molecule to the surface of the OM, a core oligosaccharide, and an O-antigen in the form of a repeated glycan monomer. The periplasmic space contains a mixture of enzymes, proteins and receptors forming the periplasm. The combination of the presence of LPS and the much more significant periplasm when compared to the Gram-positive cell wall is partially responsible for the increased difficulty in targeting Gram-negative bacteria with antibiotics.<sup>3</sup> A comparison of Gram-positive and Gram-negative bacterial cell walls is shown in **Figure 1**.



*Figure 1. Structures of Gram-positive and Gram-negative bacterial cell walls.<sup>4</sup>*

### 1.2. Antibiotics

The term “antibiotic” is used to describe compounds capable of inhibiting selectively the growth of bacterial strains in dilute solutions, in some case inhibiting to such an extent that the bacteria cannot survive. The term was first coined by Selman Waksman, one of the researchers responsible for the discovery of streptomycin, to describe compounds derived from bacteria and other microorganisms, however this definition has been altered over the years to include even totally synthesised compounds due to advances in synthetic technologies and knowledge.<sup>5</sup>

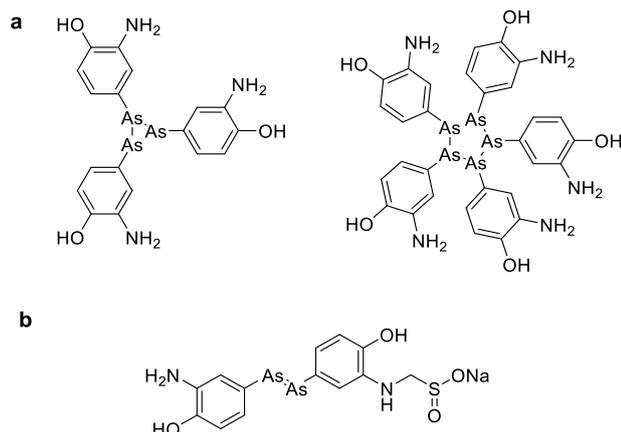
## CHAPTER 1: INTRODUCTION

It is commonly believed that the intake of antibiotic compounds has only relatively recently begun, with the discovery and marketing of penicillin in 1928 and 1941 respectively. However, antibiotic compounds have been inadvertently ingested for thousands of years, both from dietary sources and traditional medicinal formulations. For example, traces of tetracycline were found in the bones of Sudanese Nubians who were predicted to have lived between 350 and 550 BCE. The distribution of this antibiotic throughout these remains is indicative of dietary intake of tetracycline-containing foods and drinks, possibly in the form of contaminants from *Streptomyces*-containing soil or even from the alcoholic beverages consumed, the tetracycline being produced during the fermentation process.<sup>6-8</sup> This intake seemingly protected the population for the most part, as reported rates of diseases in the region were low, and bones found in nearby Egypt showed little-to-no signs of infection.<sup>7,9,10</sup>

Another example of pre-antibiotic era exposure to antibiotics is traditional Chinese medicine. Herb blends, such as *Fructus armeniaca mume* (wu mei) and *Radix isatidis* (ban lan gan), have been served as teas and ointments for thousands of years to serve as remedies for a wide variety of infections and ailments, with seemingly overall success. A study performed in 2010 examined the efficacy of 20 of these herb blends against bacteria found in the human mouth which cause tooth decay, and determined that 16 of these 20 displayed at least partial inhibitory effects against at least one of the four bacterial strains tested.<sup>11</sup>

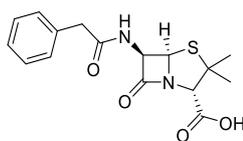
The dawn of the modern “antibiotic era”, however, began with the work of Paul Ehrlich in 1904. Ehrlich wanted to develop a so-called “magic bullet”, a compound which attacked only pathogens without affecting the host and inspired by the observation that certain synthetic dyes could stain specific bacterial cells but not others. To achieve this, he ran a large-scale systematic screening program to attempt to find compounds that would inhibit the syphilis-causing bacteria *Treponema pallidum*, screening over 600 compounds, all derivatives of Atoxyl, before finding one which successfully cured syphilis in rabbits. Human trials of the compound, although limited, showed the promise of a cure in humans,<sup>7,12</sup> and so the drug was released to market under the name Salvarsan (**Figure 2a**). Together with Neosalvarsan (**Figure 2b**), an analogue with improved solubility and lower toxicity, Salvarsan was the most frequently prescribed drug until penicillin entered the clinics in the 1940s.<sup>7,13</sup>

## CHAPTER 1: INTRODUCTION



**Figure 2.** The structures of Salvarsan, a mixture of two compounds (a), and Neosalvarsan (b).

Work on penicillin began in 1928, though its discovery was serendipitous. Alexander Fleming, whilst working with several species of *Staphylococcus* species within his laboratory, left a petri-dish containing one of his cultures open, causing it to become contaminated with a mold. When Fleming returned to the plate, he noticed that the bacterial growth had been inhibited selectively in the area surrounding this mold, leading Fleming to believe that the mold contained an antimicrobial compound. He therefore grew a pure culture of this mold, discovering it to be a *Penicillium* species, and isolated the antimicrobial compound by filtration from the culture broth. Fleming coined the term “penicillin” to describe this compound, which has now become known as penicillin G (**Figure 3**). Fleming’s work, unfortunately, did not receive much attention due to the relatively low yields of penicillin G produced, and it wasn’t until 1941, when the yield was sufficient to make the drug marketable, that penicillin was first used.<sup>14–17</sup>



**Figure 3.** The structure of penicillin G.

Following this, more and more research was devoted to identifying, discovering and generating new antimicrobial agents, with six new classes of antibiotics being discovered in the first 20 years following the discovery of penicillin.<sup>18</sup> As a result, modern society has become extremely reliant on antibiotic drugs to cure common yet fatal diseases such as pneumonia, tuberculosis and infective endocarditis caused by *Staphylococcus aureus* infections, as well as to safeguard vulnerable individuals such as those who have recently undergone surgery or childbirth.<sup>19,20</sup>

### 1.3. Antimicrobial resistance

Unfortunately, bacteria are beginning to become resistant to nearly all antimicrobial agents currently on the market, a phenomenon which is now widely considered to be one of the biggest threats to health.<sup>19</sup> Whilst some antimicrobial resistance is found to be caused simply by fortunate genetic mutations, changes in bacterial DNA during replication, or by DNA dissemination, antimicrobial use in society is the main cause of the increased rate of resistance. Societal use of antimicrobial agents allows resistant strains of a bacteria to flourish, whilst preventing non-resistant strains from reproducing effectively, thus creating selection pressure.<sup>19</sup>

Whilst the evolution of antimicrobial resistance is a natural process that would occur without the intervention of humanity, our use of antibiotics has expedited the process substantially. Human consumption of antibiotics has been higher than necessary worldwide, with more affluent countries over-prescribing antibiotics and developing countries giving easier access to antimicrobial agents, which would only be available by prescription in more developed nations. Self-medication is typically unnecessary and is often administered either too early or too late, leading to bacteria being exposed to the drug before necessary and hence allowing time for the bacteria to develop a resistance mechanism.<sup>21</sup>

The perception that antibiotics are “wonder drugs”, i.e. cure-alls, appears to be shared across both prescribed and self-medicated settings, leading to drugs being used even in the absence of bacterial or viral implications.<sup>21</sup> A 2009 survey of European citizens, 20% of those surveyed admitted self-administering antibiotics for flu-like symptoms, with 14% admitting that they had taken them to treat a common cold, despite knowing that antibiotics have no effect against viral infections.<sup>22</sup> Another survey, conducted in 2007 with respondents from the United Kingdom, showed that just under 40% of participants were unaware that antibiotics were ineffective against common colds or most coughs.<sup>23</sup> In addition to these factors, a common behaviour pattern appears to be that patients prescribed with antibiotics often fail to complete the full course of the drugs, or are prescribed shorter courses of antibiotics than required to completely overcome the infection. This can lead to incomplete inactivation or death of the bacterial infection, thus leaving a portion of the original strain which has now been exposed to antibiotic to reproduce, increasing the chance of resistance developing.<sup>21</sup>

## CHAPTER 1: INTRODUCTION

Hospitals tend to use antibiotics on intense and prolonged courses, which is possibly the main contributor to antimicrobial resistance in this setting. However, there are multiple other factors that also play important roles in the development of resistance. Patients such as those with cancer or AIDS tend to be highly susceptible to bacterial infections due to their immunocompromising ailments, whilst elderly patients, due to their fragility, also tend to be susceptible. There is also a relatively large problem with the control of infection spreading from patient to patient, also leading to more antibiotic use and therefore an increase in resistance.<sup>21,24</sup> This is further compounded by the limited number of antibiotics commonly used in hospital clinics. In a 2012 survey of European antibiotic use, it was discovered that 50% of all antimicrobials used belong to the  $\beta$ -lactams family, with amoxicillin being the antimicrobial agent most often used in all but two of the countries surveyed.

However, the development of antimicrobial resistance is not only assisted by humanity's direct consumption of antibiotics. A significant proportion of antibiotics are given to food-producing animals, both as a treatment for disease and as a prophylactic measure.<sup>25</sup> In addition, these antimicrobials are used to help promote the growth of these animals in some countries including the USA, despite the practice being banned in Europe in 2006.<sup>26</sup> In 2012 it was estimated that sales of antimicrobials for use in agriculture amounted to 8000 tonnes of active ingredients, with the most prevalent drugs being used found to be penicillins, sulphonamides and tetracyclines.<sup>27</sup>

All of the above factors aid in the dissemination of antimicrobial resistance into the environment. The soil microbiota is one of the most common sources of antibiotic compounds, and as such the bacteria present have developed resistance pathways to these antibiotics. However, when water that has potentially been contaminated with faeces or organic fertilizers, which have been used in crop production, antimicrobial-resistant bacterial strains may be added to the soil, allowing for horizontal gene transfer to occur and resistance pathways to be shared. Water that has been used for human sewage is also a major source of antimicrobial resistance, as antibiotics consumed by humans are seldom fully metabolised before being excreted.<sup>21</sup> **Figure 4** summarises the routes that antibiotics can take to enter the environment.

## CHAPTER 1: INTRODUCTION

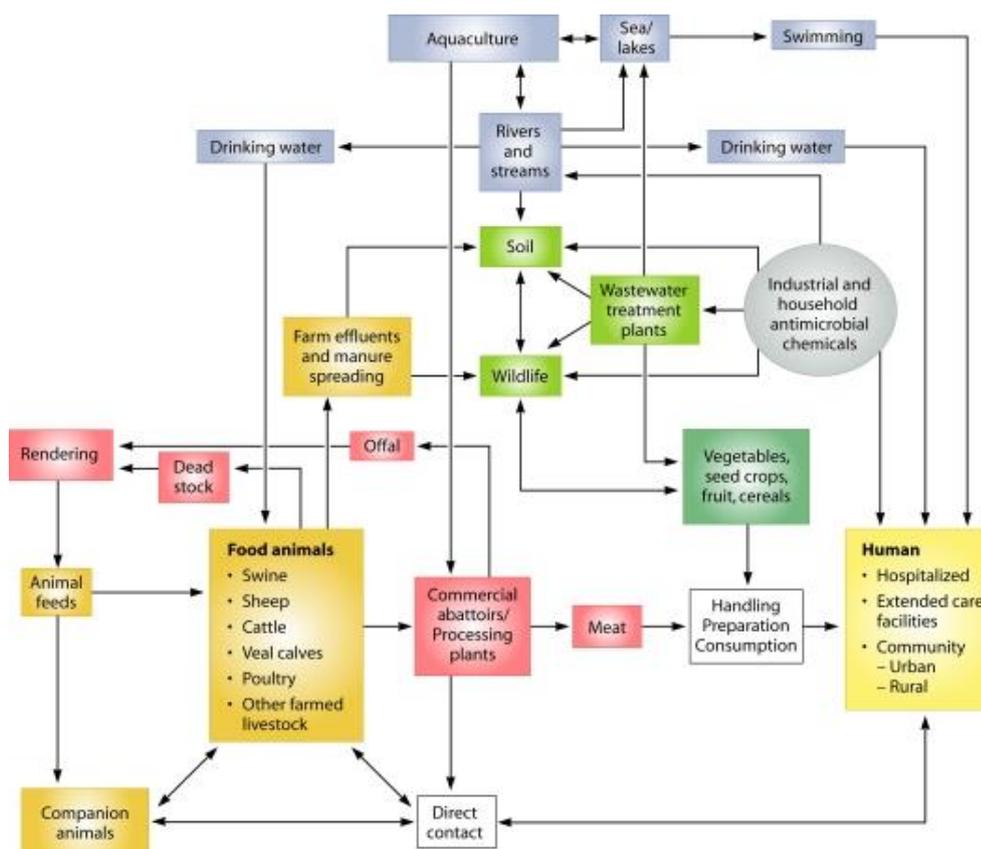


Figure 4. Movement of antibiotics into the environment.<sup>28</sup>

Intrinsic resistance, defined as resistance caused by bacterial features that are shared across a species independently of previous antibiotic exposure, and that is not related to horizontal gene transfer, is the process of sharing DNA between cells which is not due to reproduction.<sup>29,30</sup> The conventional example of intrinsic resistance is that of the outer cell membrane of Gram-negative bacteria. For example, daptomycin, a lipopeptide antibiotic compound that works by cell wall penetration and disruption, is only effective against Gram-positive bacteria. The outer cell membrane of Gram-negative bacteria possess a lower proportion of the required phosphatidylglycerol, and hence does not provide sufficient insertion sites for daptomycin to enter the cell.<sup>31</sup> In addition to outer membrane impermeability, Gram-negative bacteria also express a variety of multi-drug resistant (MDR) efflux pumps, which are capable of reducing the concentration of a given drug within the cell.<sup>32</sup> One such efflux pump exists for the tetracycline antibiotics in *E. coli*, as an example.<sup>33</sup>

## CHAPTER 1: INTRODUCTION

Acquired resistance, as the name suggests, is acquired through either the transfer of genetic material conferring resistance via horizontal transformation, transposition or conjugation, or through mutations to a cell's own bacterial DNA.<sup>34</sup> These mutated genes are typically the types responsible for encoding drug targets, drug transporters, regulators to control these transporters or antibiotic-modifying enzymes.<sup>29,35</sup> Possibly the most commonly referenced example of acquired resistance is that of the  $\beta$ -lactamase family of enzymes. These enzymes, first discovered in 1940 in the form of penicillinase, act by performing a ring-opening reaction of the  $\beta$ -lactam rings of certain classes of antibiotics, including penicillins, cephalosporins, and carbapenems, thus deactivating the antibiotic.<sup>36</sup> Whilst pharmaceutical research has found drugs that are capable of overcoming this initial  $\beta$ -lactamase emergence, these enzymes have also continued to develop and mutate, forming a sub-class of  $\beta$ -lactamases known as extended-spectrum  $\beta$ -lactamases (ESBL).<sup>37,38</sup>

The mechanisms of antimicrobial resistance can be split into four main categories: limiting drug uptake, target modification, drug inactivation and active drug efflux. Intrinsic resistance is capable of using limiting drug uptake, drug inactivation and drug efflux, whilst acquired resistance may be caused by target modification, drug inactivation and drug efflux.<sup>34</sup> There is variation not only between the type of resistance pathways available depending on whether the resistance is intrinsic or acquired, but also between Gram-positive and Gram-negative bacteria, with Gram-negative bacteria capable of using all four resistance categories. Gram-positive bacteria, on the other hand, typically are less capable of utilising drug uptake limitation due to their lack of LPS outer membrane, and do not possess the capability of using certain types of drug efflux mechanisms.<sup>34,39,40</sup>

Drug uptake limitation ability varies widely across the various species of bacteria, with structural features typically impacting drug uptake the most. As stated above, the LPS-containing outer membrane of Gram-negative bacteria helps this class of bacteria to resist the cell permeation of larger antimicrobial classes, whilst mycobacteria have a hydrophobic high-lipid outer membrane, allowing hydrophobic antimicrobial agents, such as fluoroquinolones, to enter the cell more easily whilst preventing hydrophilic agents, such as  $\beta$ -lactams, from entering.<sup>41-43</sup> *Mycoplasmas* and certain other bacterial species lack a cell wall, and hence are resistant to cell wall-targeting antibiotics such as  $\beta$ -lactam antibiotics and glycopeptides.<sup>44</sup>

## CHAPTER 1: INTRODUCTION

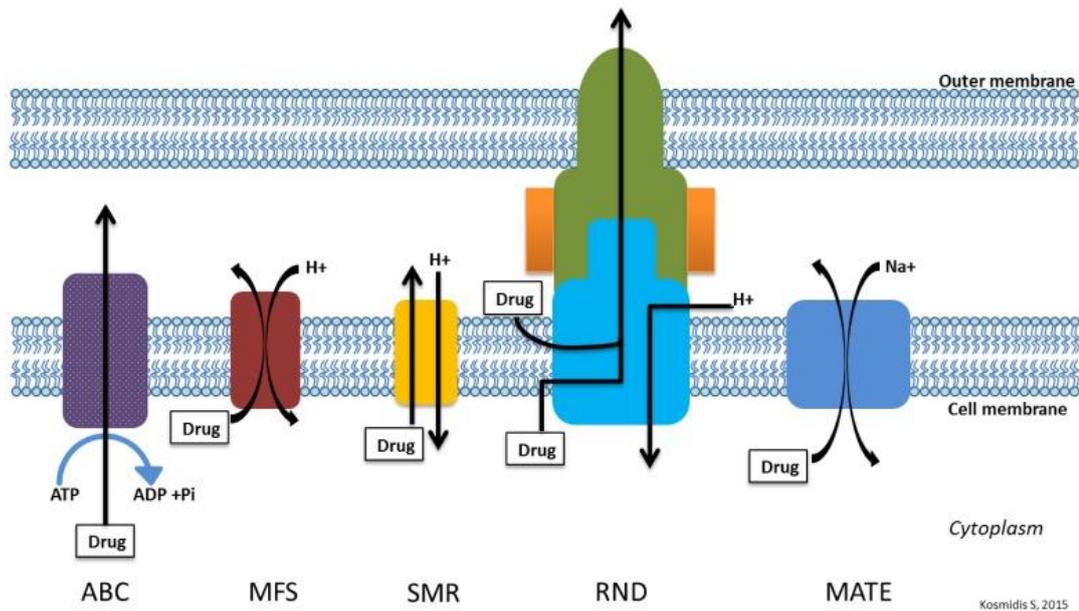
Whilst Gram-positive bacteria find limiting drug uptake more difficult than Gram-negative bacteria, some have found ways of developing resistance. For example, some species of *Staphylococcus aureus* have developed resistance to vancomycin by producing a thicker cell wall through a currently unknown mechanism. This makes the cell wall more difficult to penetrate, providing intermediate resistance to the drug.<sup>43,45</sup> Another potential route to resisting drug uptake is the modification of channels on the cell membrane usually associated with antibiotic agent transport. For example, in some strains of *Pseudomonas aeruginosa*, the imipenem transporter OprD2 is not present, preventing uptake of the drug, whilst in *E. aerogenes*, mutations to porin channels grant resistance to both imipenem and some cephalosporins.<sup>29,46</sup>

Drug target modification is a very varied pathway to resistance, as there are multiple mutations that can take place in each drug target and multiple drug targets to choose from. The most common pathway to resistance of  $\beta$ -lactam drugs in Gram-positive bacteria is the modification of the structure or concentration of penicillin-binding proteins (PBPs). Some PBPs are better at binding penicillin than others, and so increasing the concentration of those that decrease binding, or vice versa, will reduce the amount of drug able to bind. In addition, changes in structure, especially in the binding region of the protein, may decrease the binding affinity of the drug, or even prevent the drug from binding at all.<sup>35,47</sup> For drugs targeting ribosomal subunits, resistance is typically achieved by methylation of the subunit, or by protection of the ribosome,<sup>48,49</sup> whilst for drugs that target metabolic pathways, such as the synthesis of folate, resistance is achieved via mutations in the enzymes involved.<sup>34</sup> Another interesting example is that of the lipopeptide daptomycin, which requires the presence of calcium to initiate its binding. Some cells have developed a mutation in genes such as *mprF* which causes the cell membrane to become positively charged, thus preventing calcium from binding and hence bestows resistance to daptomycin.<sup>50-52</sup>

## CHAPTER 1: INTRODUCTION

Inactivation of antimicrobials can occur in two different ways; the drug can be degraded within the cell, or an additional functional group can be added to the drug changing its chemistry. As discussed above, a major route to the inactivation of  $\beta$ -lactam antibiotics is the  $\beta$ -lactamase enzymes, which essentially hydrolyse and hence degrade the drug. Another group of drugs inactivated by hydrolysis are the tetracyclines, which are hydrolysed by mutations in the tetX gene.<sup>48,53</sup> Transfer of a functional group onto a drug is typically achieved using transferase enzymes, which commonly transfer acetyl, phosphoryl and adenylyl groups. Acetylation is the most common process, and is known to affect a wide variety of antimicrobial drugs such as the aminoglycosides and the fluoroquinolones.<sup>53</sup>

Efflux pumps are chromosomally encoded into the bacterial genome, and function to remove toxic substances from cells. A high number of these efflux pumps transport a variety of drugs out of the cell, known as MDR efflux pumps, and the capability of these pumps to develop into resistance pathways is influenced predominantly on the carbon source available to the bacteria.<sup>41,54</sup> There are five classes of efflux pumps based on the structure and energy source of the pump, summarised in **Figure 5**.



*Figure 5. Summary of efflux pump classes.<sup>34</sup>*

The ATP-binding cassette (ABC) family contains both uptake and efflux transporters and are powered by the hydrolysis of ATP. These pumps, in addition to transporting necessary substances to bacteria such as amino acids and sugars, have been known to transport a variety of drugs such as fluoroquinolones and tetracyclines.<sup>55,56</sup> The multidrug and toxic compound extrusion (MATE) family utilise the intake of sodium ions as an energy source, and are capable of effluxing cationic dyes, fluoroquinolones, and in some cases aminoglycosides. This family contains very few members that are currently known, with most being found in Gram-negative bacteria.<sup>57,58</sup> The small multidrug resistance (SMR) family are a family of hydrophobic efflux pumps that are powered by the movement of protons ( $H^+$ ). These pumps have a narrow window of effectiveness, mainly removing lipophilic cationic species such as  $\beta$ -lactams and some aminoglycosides.<sup>41,42,59</sup> The major facilitator superfamily (MFS) are powered by solute/cation symport, that is, by the movement of both solute and cation in the same direction, or by solute/proton antiport, that is, by the movement of solute and protons in opposite directions to each other. These pumps are responsible for the movement of macrolides and tetracyclines, in addition to anions, sugars and metabolites such as bile salts.<sup>34</sup> The final class, the resistance-nodulation-cell division (RND) family, are powered by a substrate/proton antiport mechanism, and are found only in Gram-negative bacteria. These pumps consist of multiple components and are extremely complex, and are capable of transporting a wide variety of antimicrobials including  $\beta$ -lactams, tetracyclines, trimethoprim, and some fluoroquinolones.<sup>41,42,58</sup>

### 1.4. ESKAPE pathogens

The majority of infections found in hospital settings are caused by only 6 pathogenic species, known as the ESKAPE pathogens. These species, namely *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species, are of special relevance due to their growing resistance to a variety of antimicrobials, thus allowing them to “escape” from the bactericidal effects of most drugs.<sup>60</sup> The ESKAPE pathogens, therefore, have been found to be associated with the highest risk of mortality of all bacterial infections,<sup>61</sup> leading to the World Health Organisation (WHO) listing the pathogens as part of the 12 bacteria against which antimicrobial agents are desperately needed.<sup>62</sup>

The ESKAPE pathogens are a mixture of Gram-positive and Gram-negative bacteria, with *E. faecium* and *S. aureus* being Gram-positive bacteria and *K. pneumoniae*, *A. baumannii*, *P. aeruginosa* and the *Enterobacter* species being Gram-negative bacteria. Their mechanisms of resistance are therefore extremely varied, covering all four categories listed in **Section 1.3**.<sup>63</sup> In addition to these resistance mechanisms, these bacteria are also capable of generating biofilms, communities of bacteria which are densely packed and bound together with polymers secreted by the bacteria and which bind and grow on surfaces, including intestinal linings. These biofilms both physically prevent attack by antibiotics and immune responses, as well as protecting persister cells, bacterial cells which are dormant and hence do not take up most antibiotics, allowing culture re-growth and causing a relapse of the illness with a more resistant bacterial infection.<sup>64</sup>

Treatment of ESKAPE pathogen infections with antimicrobials either occurs singly or in combination, however the number of antibiotics available with activity against the pathogens is rapidly diminishing. Since 2010, many antibiotics recommended for use against the ESKAPE pathogens have been removed from the recommended list, with very few additions since that year. Even more concerning is the fact that resistance pathways have begun to reveal themselves against even these newly added compounds. There is therefore an urgency required in the search for new antimicrobial agents that can potentially work against these pathogens.<sup>65</sup>

### 1.5. Current clinical targets of antibiotics

Despite the breadth of antimicrobial classes available on the market (**Table 1**), there are currently only 5 drug targets which are widely exploited in the clinic (**Figure 6**). As suggested by **Figure 6**, one of the most common targets is cell wall biosynthesis. Penicillins specifically target transpeptidase enzymes, PBPs, responsible for the synthesis of peptidoglycan, a cell wall component, covalently binding to the serine active site of the enzyme. Penicillins are extremely effective at performing this role due to their stereochemical similarity to D-alanine and their highly reactive  $\beta$ -lactam rings, allowing them to be taken up as substrates for the enzyme. Due to the covalent binding, the enzyme then becomes inactive.<sup>66-68</sup> The result of this is that peptidoglycan can no longer cross-link to one another, causing the bacteria to burst under osmotic pressure.

## CHAPTER 1: INTRODUCTION

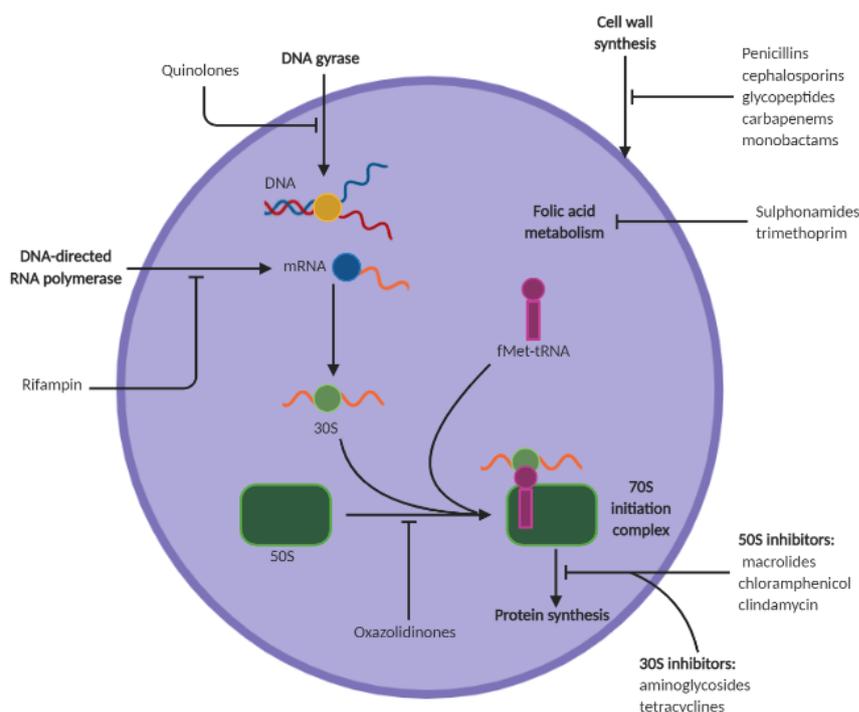


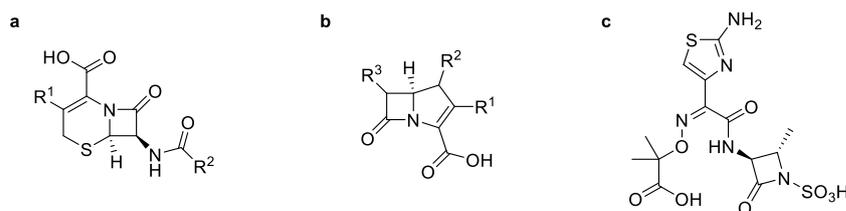
Figure 6. Clinical targets of antibiotics and examples of drugs that operate against each process.

Class	Examples
<b><i>β</i>-Lactams</b>	
Penicillins	Penicillin G, penicillin V, methicillin, oxacillin, cloxacillin, dicloxacillin, nafcillin, ampicillin, amoxicillin, carbenicillin, ticarcillin, mezlocillin, piperacillin, azlocillin, temocillin
Cephalosporins	Cephalothin, cephapirin, cephradine, cephaloridine, cefazolin
First generation	Cefamandole, cefuroxime, cephalexin, cefprozil, cefaclor, loracarbef, cefoxitin, cefmetazole
Second generation	Cefotaxime, ceftizoxime, ceftriaxone, cefoperazone, ceftazidime, cefixime, cefpodoxime, ceftibuten, cefdinir
Third generation	Cefpirome, cefepime
Fourth generation	
Carbapenems	Imipenem, meropenem
Monobactams	Astreconam
<b><i>β</i>-Lactamase inhibitors</b>	
	Clavulanate, sulbactam, tazobactam
<b>Aminoglycosides</b>	
	Streptomycin, neomycin, kanamycin, paromycin, gentamicin, tobramycin, amikacin, netilmicin, spectinomycin, sisomicin, dibekalin, isepamicin
<b>Tetracyclines</b>	
	Tetracycline, chlortetracycline, demeclocycline, minocycline, oxytetracycline, methacycline, doxycycline
<b>Rifamycins</b>	
	Rifampicin (also called rifampin), rifapentine, rifabutin, bezoxazinorifamycin, rifaximin
<b>Macrolides</b>	
	Erythromycin, azithromycin, clarithromycin
<b>Lincosamides</b>	
	Lincomycin, clindamycin
<b>Glycopeptides</b>	
	Vancomycin, teicoplanin
<b>Streptogramins</b>	
	Quinupristin, daflopristin
<b>Sulphonamides</b>	
	Sulphanilamide, <i>para</i> -aminobenzoic acid, sulfadiazine, sulfisoxazole, sulfamethoxazole, sulfathalidine
<b>Oxazolidinones</b>	
	Linezolid
<b>Quinolones</b>	
	Nalidixic acid, oxolinic acid, norfloxacin, pefloxacin, enoxacin, ofloxacin/levofloxacin, ciprofloxacin, temafloxacin, lomefloxacin, fleroxacin, grepafloxacin, sparfloxacin, trovafloxacin, clinafloxacin, gatifloxacin, moxifloxacin, sitafloxacin
<b>Others</b>	
	Metronidazole, polymyxin, trimethoprim

Table 1. Main classes of antibiotics.<sup>69</sup>

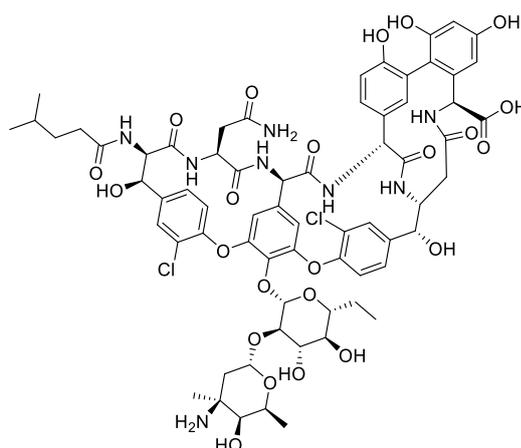
## CHAPTER 1: INTRODUCTION

This mode of action is also shared with the cephalosporin antibiotics, which feature a  $\beta$ -lactam fused with a dihydrothiazine ring (**Figure 7a**).<sup>66,70</sup> The carbapenems, another  $\beta$ -lactam-based antimicrobial class, also shares this mode of action, and, in contrast to penicillins and most cephalosporins, are potent broad-spectrum inhibitors (**Figure 7b**).<sup>71–73</sup> The final class of  $\beta$ -lactam antibiotics, the monobactams, also operate by this mechanism. These compounds, in contrast to the other  $\beta$ -lactam antibiotic classes presented above, do not possess a fused ring system (**Figure 7c**).<sup>74</sup>



**Figure 7.** Structures of the cephalosporins (a), carbapenems (b), and Aztreonam (c), a member of the monobactams.

Glycopeptides on the other hand, whilst also targeting cell wall biosynthesis, operate on a different mechanism of action. These molecules, consisting of a conserved heptapeptide domain in which only residues 1, 3, and the side chains of the other residues are allowed to change (**Figure 8**), possess a cleft which allow for binding of peptides of a specific L-D-D conformation, such as the C-terminus of the disaccharide pentapeptide repeating unit of peptidoglycan.<sup>75,76</sup> When peptidoglycan is bound in this way, the transpeptidase enzyme is unable to bind to the peptidoglycan, thus preventing cross-linking of the cell wall and inhibiting the formation of peptidoglycan.<sup>77,78</sup>

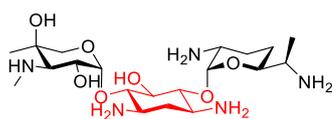


**Figure 8.** The structure of vancomycin, a glycopeptide antibiotic

## CHAPTER 1: INTRODUCTION

Another heavily utilised target is bacterial protein synthesis. As indicated in **Figure 6**, protein synthesis begins with mRNA binding to the 30S complex, before the 30S-mRNA complex itself binds to formylated methionyl-tRNA(fMet-tRNA) in such a way that the tRNA initiator is bound to the start codon. The 50S complex then joins to form the 70S complex, allowing the synthesis of proteins to begin.<sup>79-81</sup> These combinations are mediated by three initiation factors, which not only assist in the recruitment of the various components but also prevent premature binding of the 50S subunit.<sup>79,82,83</sup> Inhibition of the above processes, therefore, prevents the formation of proteins required for the correct functioning of the cell, or even allows the formation of polypeptides which are then released causing damage to the cell. Due to the variety of components available to target these processes, the mechanisms of action of drugs targeting protein synthesis is relatively broad.

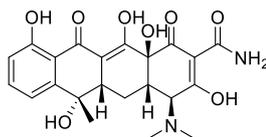
Aminoglycosides consist of aminated sugars connected by glycosidic linkages to a dibasic cyclitol, most commonly 2-deoxystreptamine (**Figure 9**).<sup>84,85</sup> These compounds bind, with high affinity, to a conserved sequence on the 16S ribosomal RNA (rRNA) of the 30S ribosome, altering the conformation of this section of the complex. This conformational change induces codon misreading of the aminoacyl tRNA, promoting mistranslation and causing errors in protein synthesis.<sup>84,86</sup> Other aminoglycosides, instead of causing the release of harmful proteins, are able to block the elongation processes or directly inhibit the initiation of the protein biosynthesis mechanism completely, hence preventing protein synthesis from occurring at all.<sup>86,87</sup>



*Figure 9. The structure of Gentamicin, an example of an aminoglycoside antibiotic. The 2-deoxystreptamine ring is highlighted in red*

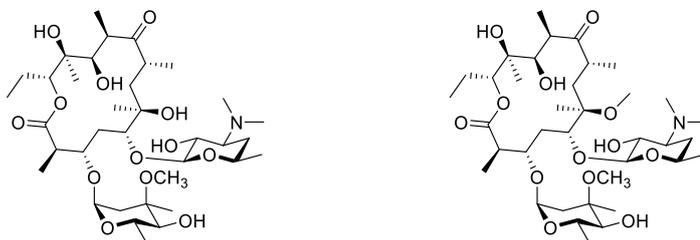
Tetracyclines also inhibit the 30S ribosome. These molecules, as the name suggests, consist of a linear fused ring system comprising of three cyclohexyl rings, named A, B, and C, and a phenolic ring, labelled as D by convention. Whilst these rings are widely functionalised, important features for maintaining antibiotic activity within these molecules are its linearity,  $\alpha$ -stereochemistry of the hydrogens at the A-B ring junction and the 4-dimethylamino positions, and the keto-enol system nearby to the D-ring.<sup>88</sup> The tetracyclines also interact with the 16S rRNA, binding in such a way as to create a steric block of the recognition of mRNA codons, stopping protein elongation processes.<sup>89</sup>

## CHAPTER 1: INTRODUCTION

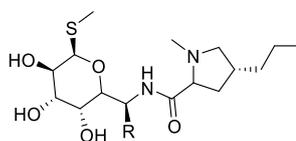


*Figure 10. The structure of tetracycline.*

The macrolide antibiotics, on the other hand, target the 50S ribosome, specifically binding close to the peptidyl transferase centre in the nascent peptide exit tunnel. This region allows for the passage of completed polypeptides out of the ribosome and into the bacteria.<sup>90–92</sup> The macrolides themselves are primarily formed of lactone rings, which are themselves adorned with a variety of sugars (**Figure 11**). The steric bulk of these molecules, therefore, completely blocks this channel in some cases, hence obstructing nascent peptide passage, and leads to a decline in overall protein concentration and an accumulation of peptidyl-tRNA.<sup>93,94</sup> However, the more general mode of action of these compounds is to change the structure of the peptidyl transferase centre, interfering with peptide bond formation.<sup>94–98</sup> The lincosamides, consisting of a pyrrolidine linked to a pyranose-type structure by an amide bond, operate in a similar manner, also inhibiting peptide growth in the peptidyl transferase centre. (**Figure 12**)<sup>90,99</sup> The pleuromutilins also bind in a similar manner.<sup>100</sup>



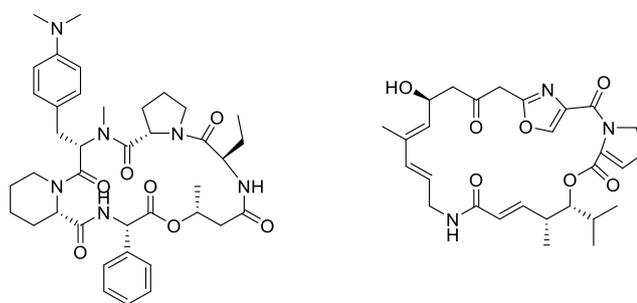
*Figure 11. Structures of Erythromycin (left) and Clarithromycin (right).*



*Figure 12. General structure of the lincosamide antibiotics*

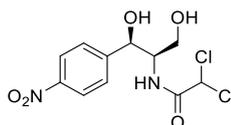
## CHAPTER 1: INTRODUCTION

The streptogramin antibiotics also bind to the 50S subunit. These drugs consist of two components, a cyclic peptide and a macrolactone (**Figure 13**), which are both required for the bactericidal effect, whilst performing bacteriostatic roles individually.<sup>101</sup> The macrolactone component binds to the 23S portion of the 50S subunit, in one of the free arms of the peptidyl transferase site, and works by blocking amino acid addition from the tRNA to the growing protein chain, thereby inhibiting early stage protein elongation. It also changes the conformation of the region, enhancing the binding of the cyclic peptide by a factor of roughly 100. This peptide then binds to a nearby site on the 50S subunit, working in a manner similar to the macrolides by further preventing peptide elongation and causing the release of incomplete peptides.<sup>102,103</sup>



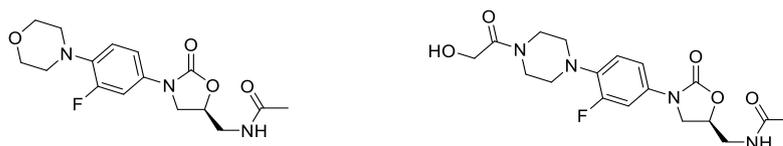
**Figure 13.** The macrolide (left) and cyclic peptide (right) components of Pristinamycin, a member of the streptogramin antibiotics

Chloramphenicol (**Figure 14**), whilst also binding to the 50S subunit, operates by a slightly different mechanism. This compound binds in a similar location to macrolides but, instead of preventing peptide bond formation by steric hindrance, as is the case for macrolides, interacts irreversibly with the receptor, thus preventing amino acid transfer from the peptidyl transferase centre.<sup>104–106</sup> Oxazolidinones, another antimicrobial class named after its characteristic oxazolidinone ring (**Figure 15**), also bind in the same binding site, competing with chloramphenicol, but have no effect on the peptidyl transferase centre.<sup>107</sup> Instead, these compounds prevent the formation of the 70S initiation complex, thus preventing the initiation of protein synthesis.<sup>107–110</sup>



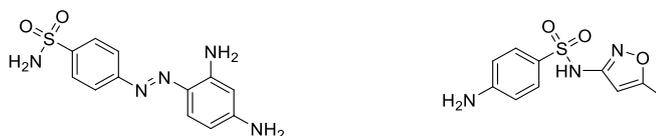
**Figure 14.** The structure of chloramphenicol

## CHAPTER 1: INTRODUCTION

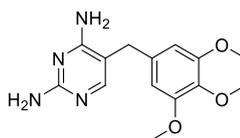


**Figure 15.** Structures of linezolid (left) and eperezolid (right), members of the oxazolidinone antibiotics.

The sulfonamide antibiotics (**Figure 16**) and trimethoprim (**Figure 17**) inhibit the bacterial formation of folic acid, required for the biosynthesis of both purines and pyrimidines, and therefore DNA. However, the mechanism of action of these two compounds exert themselves at different points of the biosynthesis. Sulfonamide antibiotics are structural analogues of para-aminobenzoic acid (PABA), a key precursor to folic acid, and act as competitive antagonists of PABA. Not only do they prevent PABA uptake, they are also themselves incorporated into folic acid precursors, generating antimicrobial pseudometabolites. Trimethoprim binds to dihydrofolate reductase, an enzyme involved in the conversion of dihydrofolic acid to tetrahydrofolic acid. Due to the synergistic activity between the sulfonamide antibiotics and trimethoprim, these drugs are often given together.<sup>111</sup>

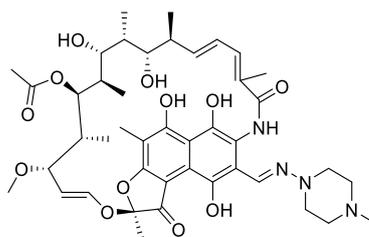


**Figure 16.** Prontosil (left) and sulfamethoxazole (right), examples of sulfonamide antibiotics



**Figure 17.** Structure of trimethoprim

Rifampicin, a semi-synthetic compound derived from the rifamycin family of antibiotics, selectively inhibits RNA polymerase, an enzyme involved in the synthesis of RNA in bacteria. The compound, a macrocyclic structure containing a naphthoquinone core and a long alkyl chain to close the ring (**Figure 18**), forms a stable complex with RNA polymerase, causing a loss in activity of the enzyme by sterically blocking the elongation of RNA.<sup>112–114</sup>



*Figure 18. The structure of Rifampicin*

The quinolone antibiotics are a class of synthetic compounds that inhibit the topoisomerase enzymes DNA gyrase and topo IV. Topoisomerases are enzymes required for correcting DNA topology, and as such are enzymes required for bacterial life. The quinolones and topoisomerases are discussed in more detail later in this chapter; however, it is first important to review some other fundamental aspects first.

### 1.6. Protein-protein interactions

Proteins are large, complex molecules consisting of a chain of amino acids, and are important biological molecules, performing a vast array of roles in prokaryotic and eukaryotic organisms. Due to their broad range of cellular functions, their structures are extremely varied, however most proteins fold into unique 3-dimensional structures. These structures can be described by four aspects, namely the primary, secondary, tertiary and quaternary structure. The primary structure describes simply the sequence of amino acids required to make up the protein, whilst the secondary structure describes locally repeating structures that are stabilised by hydrogen bonding interactions, for example  $\alpha$ -helices,  $\beta$ -sheets and turns. The tertiary structure gives the overall shape of the protein molecule, i.e. the relationship of the secondary structural features to one another, and so controls the basic function of the protein. Tertiary structures are typically stabilised by the formation of a hydrophobic core in addition to salt bridges and disulphide bonds.

## CHAPTER 1: INTRODUCTION

Quaternary structure describes the structures formed when multiple proteins, named sub-units, come together to form protein complexes. These sub-units interact with one another, forming one potential example of protein-protein interactions (PPIs). Protein complexes have a variety of biological functions that are essential for life and are as varied in structure as the sub-unit proteins that they consist of. As an example, proteasomes are the degradation complexes found inside all eukaryotic and archaeal organisms, as well as in some bacteria. These complexes destroy damaged or unnecessary proteins by breaking the amide bonds between the amino acids of the protein.<sup>115</sup> They do this by recognition of a polyubiquitin chain, attached to the protein to be degraded by ubiquitin ligases and constructed one ubiquitin molecule at a time, followed by binding of the polyubiquitin chain by the complex. The complex then undertakes nucleophilic attack using the hydroxyl functional group of one of the  $\beta$ -subunit's threonines, forming a series of short 7-9 residue polypeptides which can be further degraded and used in future protein syntheses.<sup>116</sup>

PPIs are also key features of a variety of other healthy biological processes. PPIs are responsible for the functioning of enzymes, allowing the incorporation of their substrates into their active sites by localised interactions between the receptor and the ligand. Ion channels, on the other hand, are examples of transmembrane multiprotein complexes that allow the transport of ions across membranes such as those found in the central nervous system.<sup>117</sup> When functioning incorrectly, these complexes can cause illnesses such as seizures.<sup>118</sup> PPIs are also utilised in a large number of receptors, including G protein - coupled receptors, a family of membrane proteins that are involved in various transductions of signals.<sup>119</sup> They are also directly involved in a wide array of disease-related processes in a similar way to ion channel malfunctioning described above. For example, cystic fibrosis is caused by malfunctioning of the cystic fibrosis transmembrane regulator, with symptoms worsening based on the level of dysfunction of the regulator.<sup>120,121</sup> In addition, the malfunction of the p53-MDM2 PPI leads to inhibition of the tumour-suppressing qualities of p53, leading to cancers which respond poorly to chemotherapy, and hence that have worse prognoses.<sup>122,123</sup>

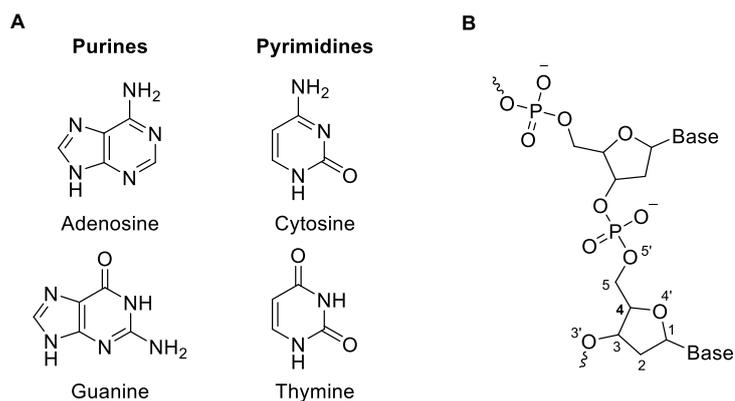
At first glance, therefore, PPIs would appear to be ideal targets for new pharmaceuticals, as interfering with these interactions could either assist in the correct function of healthy biological processes or inhibit disease-related processes. However, there are several problems associated with targeting PPIs. The interaction surface is typically large, relatively flat and hydrophobic, which makes binding with high affinity and selectivity difficult for standard “drug-like” small molecules.<sup>124</sup> Using peptides as drug leads also has its own challenges *in vivo*, such as loss of secondary structure which can lead to poor cellular uptake, and susceptibility to proteolysis.<sup>125</sup> In addition, computational screening and fragment screening, other common approaches to developing compounds affecting PPIs, require large libraries of compounds and the synthesis of a large number of compounds, most of which will be non-selective.<sup>124</sup>

That is not to say, however, that PPIs are un-druggable. Proteins have been shown to have so-called hot-spots, regions of the protein-protein interface where binding energies are more localised. These regions therefore contribute more to the binding of the proteins to one another than other regions where the binding energies are lower, hence allowing for drugs to be targeted more towards these hot spots.<sup>126</sup> In addition, many PPIs involve binding into a well-defined groove or a series of pockets, further allowing localised drug discovery studies to take place.<sup>127</sup> These observations have led to the development of a variety of proteomimetic and peptidomimetic approaches to drugging PPIs.<sup>128</sup> One approach, named peptide-directed synthesis, will be discussed in more detail in **Section 1.15**.

### 1.7. DNA

Deoxyribonucleic acid (DNA), discovered in 1953 by James Watson and Francis Crick,<sup>129</sup> is the component of life which allows the storage and passage of genetic information, and is composed of a series of repeated nucleic acid monomers. These acids are known as nucleotides, and are composed of a sugar, a phosphate group, and one of four nitrogen-containing heterocyclic bases; adenine (A), cytosine (C), guanine (G) and thymine (T). Adenine and guanine possess a purine base, whilst cytosine and thymine possess a pyrimidine base (**Figure 19A**). The monomers are linked together in DNA by utilisation of the phosphates, binding at the 3' and 5' positions of the sugars, forming a long nucleotide chain (**Figure 19B**).

## CHAPTER 1: INTRODUCTION



**Figure 19.** The structure of DNA bases (A) and the binding of nucleotides to form a single strand of DNA (B).

DNA exists as two chains of nucleotides held together by specific hydrogen bonding between base pairs; adenine and thymine bind to one another utilising two hydrogen bonding interactions whilst cytosine and guanine bind using three hydrogen bonds (**Figure 20**). This specificity is what allows DNA to store information that can be read by enzymes within an organism to direct the biosynthesis of various proteins, via an intermediate RNA chain; by changing the sequence slightly, the process of translating the sequence into a protein will result in a different protein being formed, or a protein not being formed at all. DNA is split into different areas, known as genes, and these genes are the sections of DNA that can be translated into a protein. This process begins with transcription, where RNA is synthesised from its DNA template. This RNA can then be read by enzymes, which will then build the protein. Genes are condensed together into a manageable sized formation, known as a chromosome.



**Figure 20.** Base pair hydrogen bonding between cytosine and guanine (left) and adenine and thymine (right).

The five-membered sugar ring of the nucleotides are neither planar, nor do they adopt the standard envelope structure associated with five-membered cyclic chemical structures. Instead, the ring adopts the most energetically favourable spatial arrangement of a selection of possible arrangements, all of which contain a puckering of some sort. This is due to steric interactions between the nearby nucleotides, requiring the ring twist and strain to accommodate these bulky features. Two of the most common arrangements are C2' *endo* puckering, in which the C2 atom of the ring shifts to a position above the plane of the ring, and C3' *endo* puckering, where the C3 atom shifts to a similar position. These spatial arrangements are depicted in **Figure 21**.<sup>130</sup>

## CHAPTER 1: INTRODUCTION

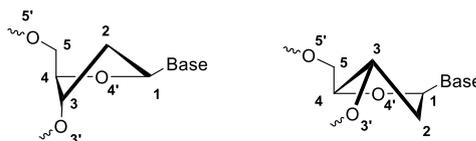


Figure 21. C2' endo (left) and C3' endo (right) arrangements of DNA sugar ring.

In a similar way, the bond linking the sugar to the base can be orientated axially ( $\alpha$ -nucleoside) or equatorially ( $\beta$ -nucleoside). However, *in vivo* only  $\beta$ -nucleosides occur. In addition, the bases have the ability to rotate around the glycosylated bond between the sugar and base, giving rise to *anti* and *syn* nucleoside conformations as shown in **Figure 22**.

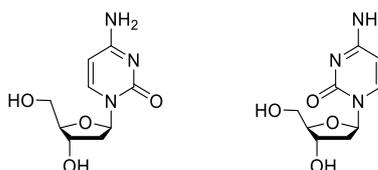


Figure 22. The *anti* (left) and *syn* (right) conformations of a pyrimidine nucleoside.

The purine bases are not co-planar, sitting at near right angles to the sugar, and are able to adopt either the *syn* or *anti* conformation without experiencing notable hindrances. However, the pyrimidine bases adopt only the *anti* conformation, owing to steric clashing between the carbonyl of the base and the hydrogen on the C2' position of the sugar if the base sits *syn* to the sugar ring. Finally, the  $sp^3$  hybridised sigma bonds making up the phosphodiester backbone are free to rotate on their axes, allowing any strain in the system to be spread across the entire backbone. This in turn means that the backbone is not straight, and in fact is curved. This contributes to DNA's biological function, as, for example, protein-DNA interactions, in a similar manner to PPIs, require binding into grooves. Overall, this flexibility influences the higher order structures that DNA can form both *in vitro* and *in vivo*.<sup>130</sup>

### 1.8. Geometry, topology and higher order features of DNA

There are three main forms of DNA that exist in nature. B-form DNA is believed to be the most common form observed in cells, and is the form described by Watson and Crick.<sup>131</sup> It consists of a right-handed anti-parallel  $\alpha$ -helix which proceeds through 10.5 base pairs per full turn. The radius of this helix is 10 Å, and the axial rise, the spacing between the base pairs, is 3.4 Å. The sugar backbone of B-form DNA also exhibits C2' endo puckering *anti* glycosidic bond configuration, and has intrastrand phosphate-phosphate distancing of 7 Å.<sup>132</sup> This form of DNA is extremely stable, as it possesses multiple stabilising interactions.

Firstly, the aromatic rings of the base pairs can interact with one another through  $\pi$ - $\pi$  stacking along the axis of the helix. The Watson-Crick base-pairing itself also contributes to stability. The phosphate groups electrostatically attract cations found in the DNA's surroundings. The DNA backbone also interacts in a hydrophilic manner, whilst the cylindrical core of the helix interacts in a hydrophobic manner. Finally, the presence of water and sodium ions is able to mitigate the electrostatic repulsions present between the phosphate groups, as well as further strengthening the  $\pi$ - $\pi$  stacking of the bases through the hydrophobic effect.<sup>133</sup>

A-form DNA is another alternative conformation that can potentially form *in vitro*, typically when lower levels of water are present.<sup>134</sup> This form of DNA once again consists of a right-handed  $\alpha$ -helix, however it is shorter and broader than the B-form, allowing for 11 base-pairs to fit into one full helix. It also possesses C3' *endo* puckering, in contrast to B-form's C2' *endo*, and has been shown to exhibit shorter phosphate-phosphate distances.<sup>132</sup> The existence of this form in more dehydrated conditions can be explained by the orientation of the base pairs. Due to the twist of the sugar due to the C3' *endo* puckering, the base pairs are tilted somewhat away from the central core of the helix, allowing for less hydrogen bonding interactions to take place between the DNA backbone and the surroundings. It makes sense, therefore, that in hydrated conditions B-form DNA would predominate, allowing for the maximum number of stabilising hydrogen bonds to form, whilst in dehydrated conditions the A-form would begin to form.

The third form, named Z-form DNA, was discovered in 1979 in the form of an alternating guanine-cytosine oligonucleotide possessing unusual structural features when compared to A-form and B-form DNA. The main difference between this form and the other two forms previously described is that the  $\alpha$ -helix is left-handed, with alternating *syn-anti* conformations of the bases, leading to a zig-zag-like backbone structure which gives Z-form DNA its name. This *anti-syn* conformational switch is due to the sugar puckering also alternating between C2' *endo* and C3' *endo*. Other structural features of the DNA are comparable, such as proceeding through 12 residues per turn, an axial rise of 3.7 Å, and the helix diameter being 18 Å.<sup>132,135</sup> The structures of A-, B-, and Z-form DNA are shown in **Figure 23**.

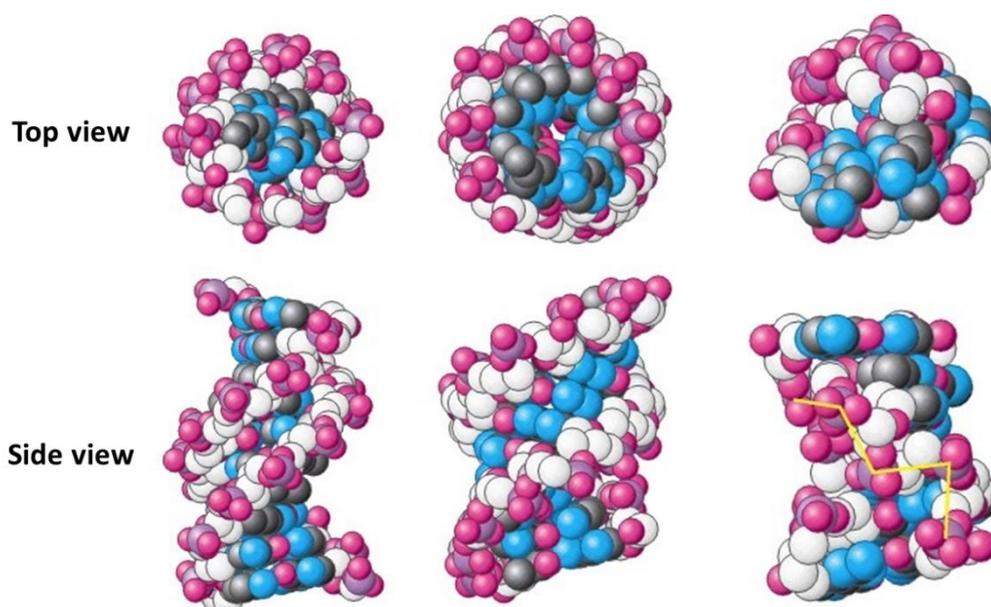


Figure 23. Structures of B-form, A-form and Z-form DNA respectively.<sup>136</sup>

When DNA strands in a double stranded structure are separated from one another, such as during transcription of DNA, rotation both ahead and behind the point of separation occurs. This results in strain, and ultimately yields over- or under-twisting of the DNA strand, causing structures known as supercoils to form. This supercoiling needs to be resolved before the transcription process can continue. Enzymes known as topoisomerases perform the role of removing this strain and will be covered in more detail in **Section 1.10**.

To fully understand this concept, however, it is first necessary to understand the topology and geometry of DNA, and how we can describe these features. DNA strands cross over one another in both linear and circular form when supercoiling occurs. The number of times this occurs is the same in both cases and is termed twist (**Tw**). The number of times double-stranded DNA crosses over itself to form supercoils is referred to as writhe (**Wr**). Twist and writhe are depicted below in **Figure 24**. The sum of these two values is known as the linking number (**Lk**), and therefore this value describes the topological changes that occur within the molecule (**Equation 1**). In addition, since this value cannot change without introduction of a break into the DNA strands, the linking number represents the number of times one strand would need to be passed through the other in order to fully separate the DNA into two linear strands.<sup>137</sup>

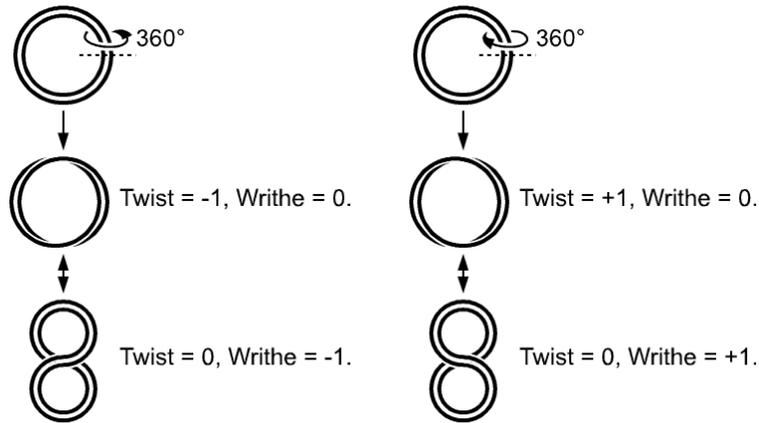


Figure 24. DNA twisting and writhing.

$$Lk = Tw + Wr$$

Equation 1. Relationship between linking number, twist, and writhe.

It is also common to refer to the change in linking number,  $\Delta Lk$ . This value refers to the difference in  $Lk$  when compared to a relaxed DNA molecule ( $Lk_0$ ) (Equation 2).

$$\Delta Lk = Lk - Lk_0$$

Equation 2. Change in linking number.

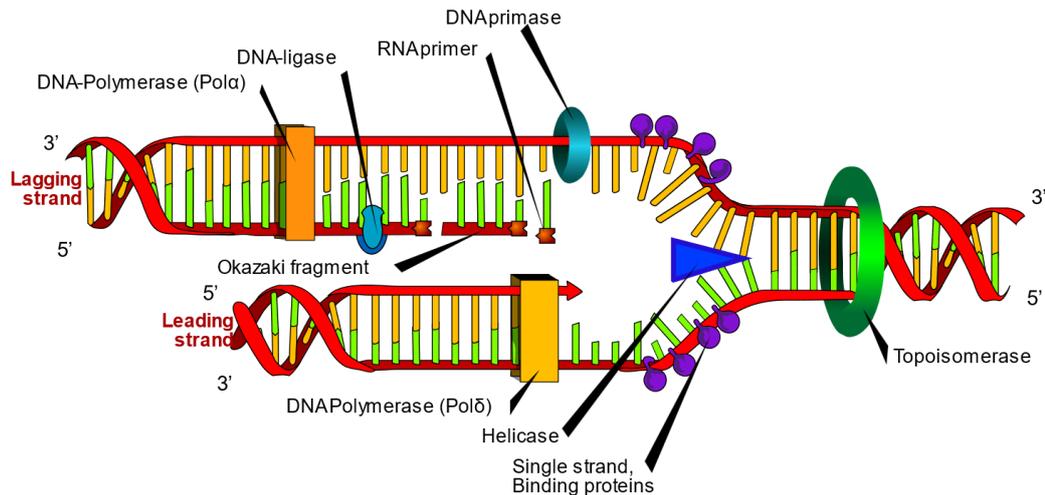
### 1.9. DNA replication in bacteria

DNA replication takes place over the course of three enzyme-catalysed and coordinated steps. The first step, initiation, begins with the binding of an initiator protein to a specific region of DNA, known as the origin of replication. This initiator protein then begins to unwind the origin DNA, and recruits other proteins to assist in the replication process such as helicase, an enzyme that breaks the base-pairing hydrogen bonds, and primase, an enzyme that adds primers to one strand of the DNA.<sup>138</sup>

The second step, elongation, begins once the DNA strands have begun to be separated from one another, leading to the generation of a replication fork. The two resulting strands, termed the leading strand and lagging strand, are then used to generate a complimentary strand for each of the leading and lagging strands by another enzyme called DNA polymerase. DNA polymerase operates by reading DNA from the 3' end to the 5' end, presenting an issue; the leading strand is generated from the 3' end, allowing correct reading, however the lagging strand is generated from the 5' end. The lagging strand is therefore synthesised in short segments, in contrast to the leading strand which is synthesised continuously. Primase enzymes read the lagging strand and initiate the synthesis of a complimentary RNA primer. DNA polymerase then extends these primed

## CHAPTER 1: INTRODUCTION

segments, generating short fragments known as Okazaki fragments. The RNA primers are then removed and replaced with DNA, which can then be joined to other fragments using DNA ligase.<sup>138</sup> The final step, termination, occurs when either there is no more DNA to replicate, or when a protein terminates the replication process. The process of DNA replication is shown below in **Figure 25**.

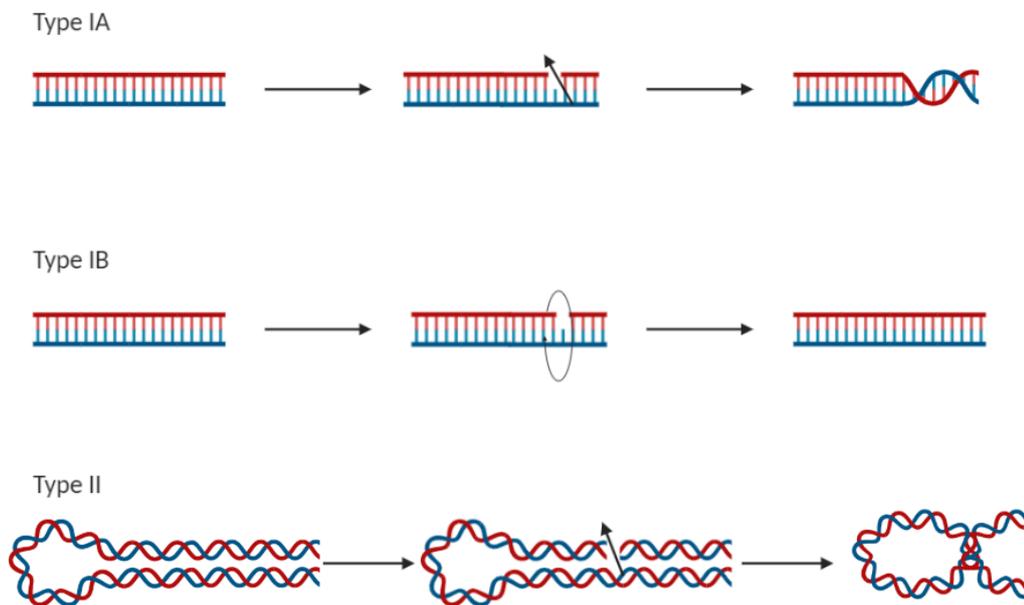


*Figure 25. The process of DNA replication.*

DNA replication therefore requires the separation of the two strands of DNA, unravelling the coiled structure at one end. However, this then causes overtwisting or undertwisting of DNA ahead of and behind the site of transcription respectively. This overtwisting and undertwisting causes a phenomenon known as supercoiling, with overtwisting being defined as positive supercoiling and undertwisting being defined as negative supercoiling.<sup>139</sup> The problem with supercoils arises as DNA strand separation progresses. As the helicase enzyme works its way along the DNA, more and more positive supercoiling occurs with less and less space available for this supercoiling to be distributed. This generates strain, and this strain can disseminate into the region in which DNA separation has already occurred, generating “precatenates”. These precede catenation of DNA, in which two chains of DNA link, leading to impeded cellular processes.<sup>3</sup> Nature therefore requires enzymes to regulate the topology of DNA.

### 1.10. Topoisomerase enzymes

Topoisomerase enzymes are nature's response to supercoiling. These enzymes are capable of modifying the linking number of DNA by cleaving and annealing DNA, reducing strain on the system by allowing these supercoils to unwind (in the case of positive supercoils) or wind (in the case of negative supercoils). This cleavage of DNA allows the two strands to pass through one another before being annealed back together.<sup>140</sup> The cleavage reaction is guaranteed to be reversible by the enzyme, as it covalently attaches itself to the cleaved DNA stand(s) through phosphodiester linkages utilising a tyrosine residue on the enzyme, thus keeping the two sides of the cleavage site sufficiently close together and enabling the catalysis of the annealing process.<sup>141</sup> DNA topoisomerases are split into two main categories based on whether they cleave one strand (Type I) or both strands (Type II). It is therefore implied that, as Type I topoisomerases split one strand and Type II topoisomerases split both strands,  $\Delta Lk$  for these enzymes is  $\pm 1$  and  $\pm 2$  respectively.<sup>142</sup> These types are then further split based on architecture and mechanism. The mechanisms for each type of topoisomerase enzyme are summarised in **Figure 26**.



*Figure 26. Mechanisms of each class of topoisomerase enzyme*

### 1.10.1. Type IA

Type IA topoisomerases possess a toroid-like structure with an internal lumen large enough to accommodate double-stranded DNA.<sup>141</sup> The single-stranded DNA substrate is held in place by the lower portion of the structure, with the strand then extending across the tip of a flexible arm in which the active-site tyrosine is located.<sup>143,144</sup> The enzyme cleaves a single strand of DNA, binding to the sugar phosphate backbone using the tyrosine residue, and then passes the second strand of DNA through this opening before then re-annealing the DNA resulting in a change of linking number of +1 ( $\Delta Lk = +1$ ). In the case of Type IA topoisomerases, this binding occurs at the 5' end of the sugar, releasing a 3' OH group.<sup>141,142,145,146</sup> When the two ends of the DNA strand are re-annealed, the strand which was passed remains in the lumen, meaning that the enzyme must open a second time to release the passed strand.<sup>146</sup> This explains why enzymes which fit into this category can only catalyse changes of  $\Delta Lk = +1$ ; the enzyme needs to reset after each cleavage-passage-annealing cycle. Type IA topoisomerases include enzymes such as topo I, primarily used to relax negatively supercoiled DNA, topo III, responsible for catenation and decatenation of replicating daughter DNA molecules, and reverse gyrase, an enzyme found uniquely in hyperthermophiles which introduces positive supercoils and renatures DNA.<sup>147-149</sup>

### 1.10.2. Type IB

Unlike Type IA topoisomerases, Type IB topoisomerases are capable of relaxing both positive and negative supercoils, and also form their phosphodiester linkages at the other side of the nucleotide sugar, instead binding to the 3' end releasing the 5' OH end. The mechanism of action of Type IB topoisomerases involves the clamping of DNA by four domains contained within the enzyme, followed by nicking a single strand of DNA by formation of the phosphodiester linkage.<sup>146,150</sup> The tension caused by DNA supercoiling is then relieved by rotation of the 5'-OH strand end with respect to the intact strand.<sup>146,151</sup>

### 1.10.3. Type IC

Type IC topoisomerases function in a similar way to topo IB enzymes; by a nicking-swivelling mechanism which is capable of relaxing both positively and negatively supercoiled DNA. There is only one member of this family; Topo V, found in archaea only.<sup>152,153</sup>

### 1.10.4. Type IIA

Type IIA topoisomerases exist as dimers, and bind and cleave duplex DNA by attachment of each subunit of the dimer to the 5' end of the DNA nucleotide, once again utilising a phosphodiester bond. As in Type IA topoisomerases, Type IIA topoisomerases use a strand passage mechanism, with a conformational change in the enzyme pulling the two cleaved ends of the DNA apart to create G-segment DNA. This G-segment DNA can then be passed through by a second duplex chain, either from the same molecule or a different molecule. This two-strand passage explains why  $\Delta Lk$  for these enzymes is  $\pm 2$ ; two strands pass through at a time, following which the enzyme needs to open up again to allow another segment to be separated.<sup>146,154</sup> Two examples of Type IIA topoisomerases are DNA gyrase, which will be discussed in **Section 1.11**, and topo IV, a structural homologue of DNA gyrase.

### 1.10.5. Type IIB

Topoisomerases that fit into Type IIB strongly resemble Type IIA topoisomerases, and possess two subunits that combine to form a heterotetramer. As is the case with Type IIA topoisomerases, Type IIB enzymes utilise the hydrolysis of ATP to fuel the supercoiling process.<sup>155,156</sup>

## 1.11. DNA Gyrase

As stated above, DNA gyrase is a Type IIA topoisomerase. It is found in bacteria but not in animals, and is also the only bacterial topoisomerase which is capable of inducing negative supercoiling, making this enzyme extremely important for bacterial DNA transcription processes. During bacterial transcription, strand separation introduces positive supercoiling, which can then be relieved by the negative supercoiling action of DNA gyrase.<sup>157</sup> It consists of two components, GyrA and GyrB, which are arranged in a heterotetramer structure ( $A_2B_2$ ) during the topoisomerism process. However, GyrA and GyrB are not associated with one another, and are actually randomly distributed throughout the cell.<sup>157-159</sup> GyrA has been found by X-ray scattering to exist in solution as a dimer, whilst GyrB remains as a monomer, further agreeing with the fact that the two components do not associate with one another besides in the presence of DNA.<sup>160,161</sup> A schematic of the DNA gyrase  $A_2B_2$  complex is shown below in **Figure 27**.

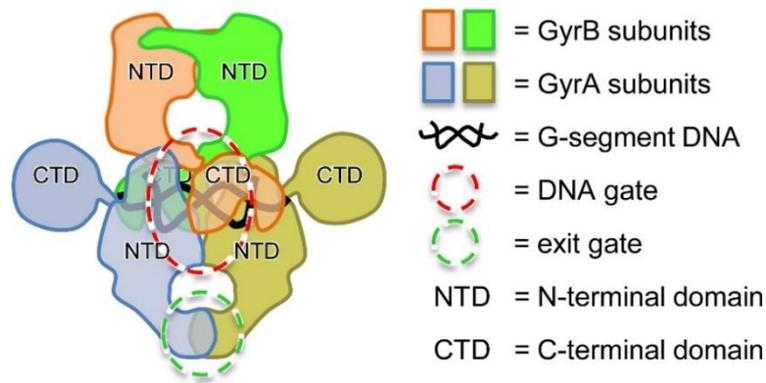
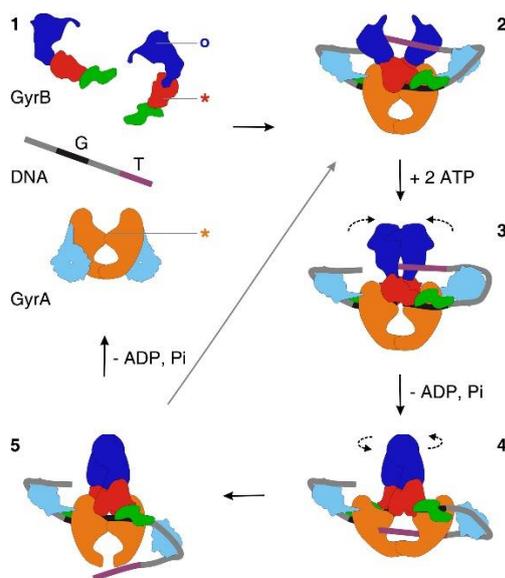


Figure 27. Schematic of the DNA gyrase  $A_2B_2$  complex.<sup>157</sup>

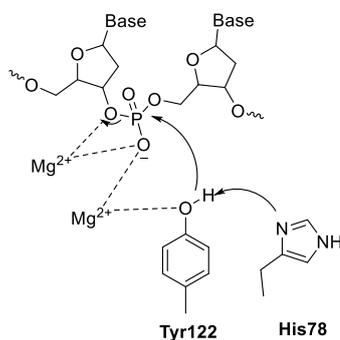
### 1.12. Mechanism of supercoiling by DNA gyrase

The full cycle of DNA supercoiling by DNA gyrase is shown below in **Figure 28**. DNA gyrase is thought to operate via a “two-gate” mechanism, i.e. the enzyme possesses two separate gates with one being used for entry and one being used for exiting the enzyme.<sup>162</sup> The first DNA strand, referred to as the G-segment, binds to the N-terminal region of GyrA, as well as the topo/primase (TOPRIM) region of GyrB.<sup>163,164</sup> The DNA then wraps around the C-terminal domain in a right-handed supercoil consisting of roughly 130 base pairs, allowing a secondary segment, known as the T-segment, to reach a second gate, known as the “N-gate”, above the G-segment of DNA.<sup>165,166</sup> Upon binding of ATP to the now fully assembled DNA gyrase, the N-gate closes and clamps the T-segment DNA in place before the enzyme cleaves the G-segment. This occurs using GyrA’s Tyr122 residue to hydrolyse the DNA backbone, utilising magnesium ions as co-factors, forming DNA-phosphotyrosyl bonds at a spacing of 4 base pairs (**Figure 29**).<sup>158,167,168</sup> The T-segment is then passed through the DNA gate, the broken G-segment, and finally out through the exit gate. The severed regions of the G-segment DNA are then re-annealed, using ATP to drive the energetically unfavourable reaction to completion. Once all ATP has been hydrolysed to form ADP the N-gate opens to release the DNA chain and to reset the enzyme for the next cycle.<sup>158</sup> One supercoiling cycle introduces two negative supercoils, and so requires two ATP molecules per cycle. In addition, when no ATP molecules are present DNA gyrase is capable of relaxing negatively supercoiled DNA.

## CHAPTER 1: INTRODUCTION



**Figure 28.** The mechanism of DNA gyrase: **1** proteins exist in their free state, **2** DNA associates around the C-terminal domain of GyrA in a right-handed supercoil presenting the T-segment above the G-segment, **3** GyrB dimerises in the presence of ATP cleaving the G-segment, **4** GyrB rotates widening the GyrA cleft and allowing transport of the T-segment through the G-segment driven by ATP, **5** DNA is reannealed with the introduction of two negative supercoils before the T-segment is released and the enzyme resets. Symbols: Circle; ATP binding pocket, Star; residues involved in DNA cleavage. Colour coding: Orange; GyrA N-terminal domain, Cyan; GyrA C-terminal domain, Navy Blue; GyrB N-terminal domain, Red: TOPRIM domain, Green; tail domain, Purple; T-segment, Black; G-segment.<sup>158</sup>



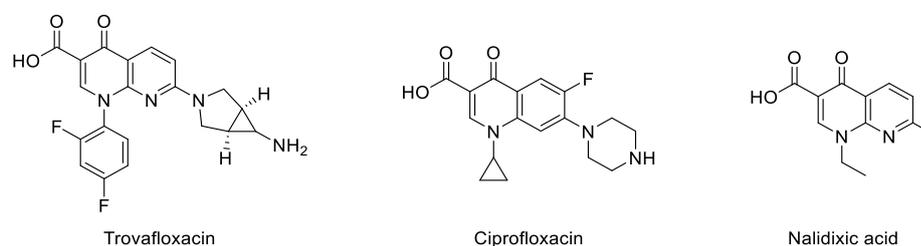
**Figure 29.** Cleavage of DNA by DNA gyrase.

### 1.13. Current antimicrobials targeting DNA gyrase

As a result of DNA gyrase being the only bacterial topoisomerase capable of inducing negative supercoiling, multiple attempts to target the enzyme have been studied. There are several points during the catalytic cycle at which a small molecule ligand can interfere with DNA binding, ATP binding, and/or DNA ligation. Three examples, quinolones, nybomycins and aminocoumarins will be discussed.

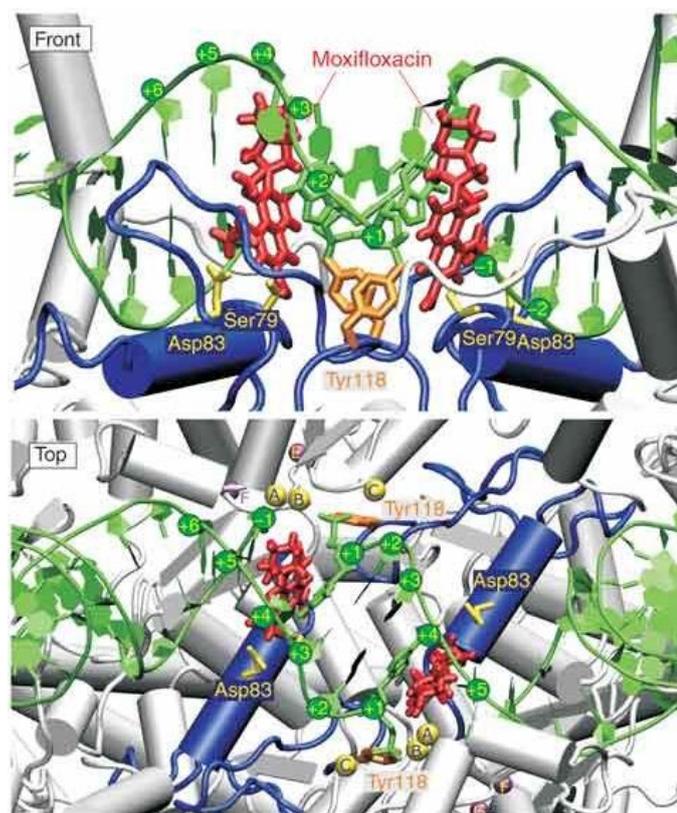
## 1.13.1. Quinolone antibiotics

The first quinolone antibiotic was discovered as a side product of the synthesis of chloroquine in 1962 in the form of nalidixic acid.<sup>169</sup> However, most quinolone antibiotics are derived from natural products, including examples such as trovafloxacin and ciprofloxacin (**Figure 30**). Whilst nalidixic acid is technically not a quinolone, containing instead of a quinoline core a 1,8-naphthyridine core, it is still considered a predecessor to current quinolone antibiotic compounds. In addition, as successive generations of quinolone antibiotics have been discovered and/or generated, the structures have deviated away from the quinoline core, with the classification of these compounds seemingly non-standardised.<sup>3</sup>



**Figure 30.** Examples of quinolone antibiotics.

Despite initial quinolones possessing multiple undesirable qualities such as limited potency and activity against Gram-negative bacteria, a short half-life and high binding affinity to proteins, more recent iterations of quinolone antibiotics have shown much improved pharmacokinetic properties, with ciprofloxacin now being used as a benchmark for efficacy. The key structural features of modern quinolones are the presence of a free carboxylic acid group and a fluorine bound to position 6 of the quinolone core.<sup>3,170,171</sup> The fluoroquinolone drugs interrupt the cleavage-ligation process by placing themselves between the DNA bases at both cleaved nucleotides, stacking themselves against the severed bases. Two molecules of the fluoroquinolone in question are required to do this, with each molecule stacking against one side of the break. This interaction is shown below in **Figure 31**.<sup>172</sup>



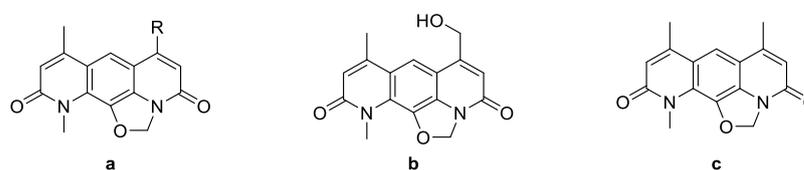
**Figure 31.** Front and top view of DNA intercalating with moxifloxacin within Topo IV. Colours: Red; moxifloxacin, Green; DNA strand, Orange; tyrosine residues. Reproduced with permission from reference 46.

However, certain strains of bacteria, including nearly all vancomycin-resistant enterococcus (VRE) and methicillin-resistant *Staphylococcus aureus* (MRSA) strains, have developed resistance pathways to fluoroquinolones.<sup>173</sup> This has been achieved by bacteria primarily via target site mutations, with almost 100% of MRSA substituting Ser84 of GyrA with a leucine residue, and nearly all VRE substituting Ser83 with Ile, Arg or Tyr.<sup>174–188</sup>

### 1.13.2. Nybomycin antibiotics

Nybomycins are natural products, first discovered in extracts from a *Streptomyces* species found in Missouri soil samples in 1955. They contain a characteristic ring system, consisting of two 5,6-dihydro-2-pyridones, one benzene and one oxazolidine ring, arranged as in **Figure 32a**. Nybomycins were discovered to possess antimicrobial activity against a variety of bacterial species, including several *Staphylococcus aureus* and *Escherichia coli* species, but showed little to no activity against other similar species for no obvious reason at the time.<sup>189</sup> However, in 2012 a study aiming to discover the mode of action of nybomycins discovered that this class of compounds were effective only against bacterial strains that had developed resistance to quinolone antibiotics. This was seemingly due to the mutations described in the previous section by which the bacteria is able to resist binding

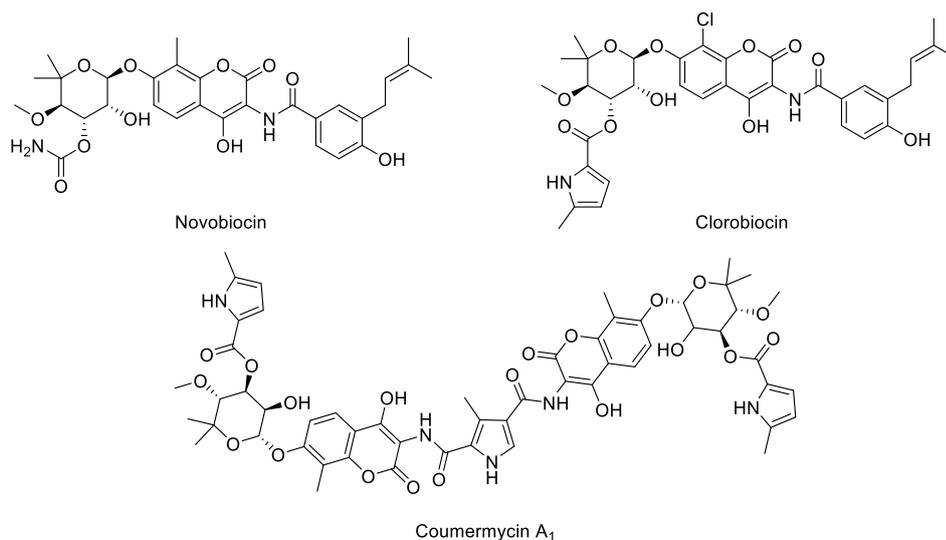
of quinolones to GyrA. Indeed, when the activity of nybomycins was examined with respect to the increasing mutation accumulation in both GyrA and another protein by which quinolone antibiotics are resisted, ParC, it was discovered that nybomycin activity was only affected by GyrA mutations. Interestingly, the mode of resistance of bacteria to nybomycins was found to be a back-mutation of GyrA back to the quinolone-sensitive variant. Thus, nybomycins have been designated “reverse antibiotics”, the first members of a novel family of antibiotics. This has led to the belief that combined therapies of quinolone and nybomycin antibiotics would be effective at treating bacterial infections, verified *in vitro* with ciprofloxacin and deoxynybomycin (**Figure 32c**) being used against 78 *Staphylococcus aureus* strains.<sup>181</sup> However, as of yet no nybomycin antibiotics have begun clinical trials.



**Figure 32.** The structures of the nybomycin scaffold (**a**), nybomycin (**b**) and deoxynybomycin (**c**).

### 1.13.3. Aminocoumarin antibiotics

Aminocoumarin antibiotics were discovered shortly before quinolone antibiotics, with the first member, Novobiocin, being discovered in 1957.<sup>190</sup> Two further aminocoumarins, discovered in 1965 and 1972 respectively, are coumermycin A<sub>1</sub> and clorobiocin. All three drugs are shown below in **Figure 33**. The aminocoumarins are natural products derived from *Streptomyces* species, and all possess a 4,7-dihydroxylated 3-aminocoumarin core. This is then attached to a sugar in all cases, typically by the 7-hydroxyl.



**Figure 33.** Examples of aminocoumarin antibiotics.

It was initially assumed that aminocoumarins had a mode of action of inhibiting cell wall formation, due to the accumulation of uridine nucleotides and reduction in the presence of protein and nucleic acids in cells, all of which are precursors to the construction of the cell wall. This would make the aminocoumarin's mechanism of action not dissimilar from penicillin.<sup>191,192</sup> However, in 1971 it was discovered that the gene responsible for aminocoumarin resistance was mapped to GyrB.<sup>193</sup> Upon discovery of DNA gyrase, this was further refined to indicate that aminocoumarin antibiotics inhibit DNA gyrase supercoiling by competition with ATP.<sup>194</sup> This was a surprising result due to a lack of similarity in terms of structure between the aminocoumarins and ATP. However, crystal structures of novobiocin bound to GyrB revealed that, whilst novobiocin was bound to a region of GyrB independent of ATP, the noviose sugar created a steric block for ATP by overlapping the ATP binding domain.<sup>195</sup> However, despite presenting good inhibitory properties *in vitro*, the aminocoumarins possess poor water solubility, poor stability, are toxic in eukaryotic cells and have low activity against Gram-negative bacteria. These properties make the aminocoumarins poor choices for use in clinical applications.<sup>196,197</sup>

#### 1.14. Drug discovery approaches

There are four main categories of conventional drug design and discovery approaches used by medicinal chemists. The decision of which of the categories to use is typically made based on the information available to the researcher regarding the structure of the target and the structure of a ligand for the target (**Figure 34**). In general, approaches utilising more structural information tend to generate more useful design features, and so

## CHAPTER 1: INTRODUCTION

approaches utilising both protein and ligand structures are more favourable. When neither the ligand or protein structure are known, then a good starting point for drug discovery is library screening such as high-throughput screening (HTS). This technique utilises large libraries of compounds at known concentrations, which can be added to known concentrations of the target protein in a suitable assay and the efficacy of the compound determined, to find active compounds which can then be further developed. Alternatively, a wide array of proteins or cells at known concentration could be screened against a given compound at a known concentration to determine what processes the compound in question could inhibit.

This process, whilst mostly automated in modern drug discovery experiments, is extremely time-consuming, requiring the synthesis or purchasing of multiple compounds (in the case of screening a library of compounds) or the expression, extraction and purification of a variety of proteins (in the case of screening a library of proteins), with a high possibility that the majority of library members will possess minimal to no observable activity. It is, however, a useful technique for finding hit compounds when little is known about the target protein, and allows for development of drugs against under-studied diseases. It also allows for the discovery of new modalities of compounds.

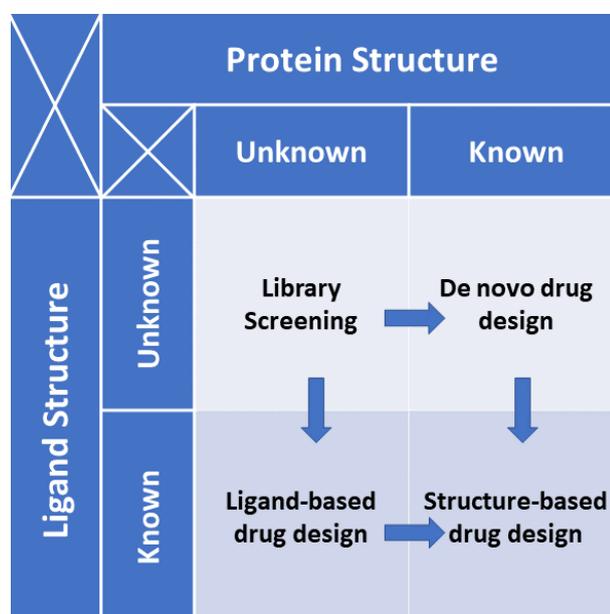


Figure 34. Drug design and discovery approaches.

## CHAPTER 1: INTRODUCTION

De novo drug design is a drug design technique used when the protein structure of the target is known but the ligand structure is unknown. In this design methodology, normally performed computationally, a new ligand molecule is constructed within the binding pocket of the target in a stepwise manner, with these steps introducing anything from single atoms to entire fragments. This allows for completely novel compounds to be suggested, allowing for the potential of an increased number of hits compared to libraries, due to libraries containing only known compounds in general.<sup>198</sup> Another potential drug design technique which could be used in the same case as de novo drug design is virtual library screening. This follows the same principles as the library screens described above, where a virtual library of compounds can be quickly docked into a target protein binding site to determine if any library members fit well into the binding site.

In a case where a ligand's structure is known but the target protein's structure is not, medicinal chemists can utilise ligand-based drug design as a starting point. This method, as the name suggests, uses the ligand as a starting point to determine how best to target the protein, and can be performed in a variety of ways. The ligand can be used to define the pharmacophore, "the ensemble of steric and electronic features that is required to ensure the optimal supramolecular interactions with a specific target structure, in this case the target protein, and to trigger or block its biological response".<sup>199</sup> It is important to note that the pharmacophore does not represent a new drug, rather it is the "common denominator" in previously reported binding interactions. Alternatively, if the target has multiple reported ligands with known structures and biological activity, then these compounds can be used to generate a quantitative structure-activity relationship (QSAR). These relationships, once again usually calculated computationally, compare computed values of various properties of the molecules with their experimentally derived values, giving an idea of the correlation between changing a component of the ligand and the biological activity. This in turn can then be used to assist in the design and prediction of activity of novel drugs.

The final methodology, used when both protein and ligand structures are known, is structure-based drug design. This can be achieved by a selective virtual screen, where compounds which are similar in shape and size to the ligand can be docked into the binding site and their docking score obtained, which can then be used to drive forward synthetic efforts. Alternatively, a de novo approach can be used in which a novel drug is built up in portions to mimic the structural features of the original ligand. A final approach could be the optimisation of a known ligand in a structure-activity relationship study.

Fragment based drug design (FBDD) is an alternative to the traditional structure-based drug design methods listed above. Instead of looking at a drug molecule as a whole, it considers the drug to be a series of binding fragments which can be altered or replaced to improve binding affinity and/or potency. FBDD allows for a more diverse chemical space to be covered than in HTS, and allows the efficiency of the ligand to be improved upon. In addition, whilst HTS tends to focus on high molecular weight compounds with potent activity, FBDD uses low molecular weight fragments (typically 140-300 Da), working on the theory that the energy contributed by each ligand atom to the binding energy of the ligand is inversely proportional to the molecular weight of the molecule.<sup>200-202</sup> Therefore, FBDD would appear to be a useful method for designing potential novel pharmaceuticals.

### 1.15. Peptide-directed synthesis

A similar approach to FBDD was developed in 2017 and was designed to exploit the strengths of both traditional drug discovery techniques and peptides, whilst also limiting their weaknesses. This approach, termed peptide-directed binding, utilises PPI-governing peptides or protein sections as a starting point. These peptides are then divided into half-length, semi-peptide fragments until binding affinity is lost, following which reactive terminals, in this case an azide and alkyne, are attached in place of the original amide bond. These reactive terminals undergo copper-catalysed azide-alkyne cycloaddition click reactions with libraries of alkynes and azides respectively to form a library of new small molecule/peptide hybrids. If these hybrids are then able to bind to the target once more, then the small molecule fragment could represent an appropriate mimic for the absent peptide fragment. Therefore, if one such alkyne fragment and azide fragment are bound together, the resulting molecule could mimic the entire peptide.<sup>203</sup> This process is shown below in **Figure 35**.

## CHAPTER 1: INTRODUCTION

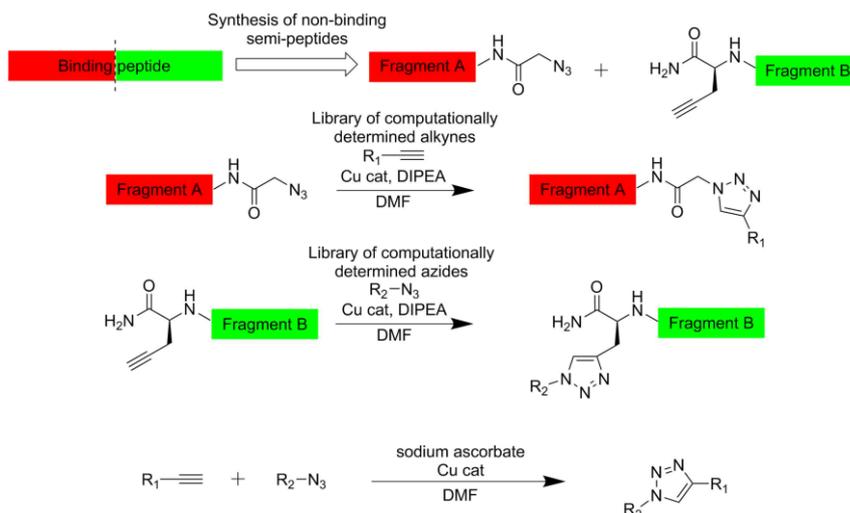


Figure 35. Outline of the peptide-directed binding method.

Peptide-directed binding was inspired by a similar strategy, known as REplacement with Partial Ligand Alternatives through Computational Enrichment (REPLACE) and developed in 2006. This strategy utilises *in silico* docking techniques to identify surrogates for specific segments of known peptide ligands by using a peptide-bound protein structure as a design tool. The work was piloted during an investigation into the inhibition of the CDK2-cyclin A PPI and was shown to provide peptidomimetic motifs for the inhibition of the target.<sup>204</sup> However, peptide-directed binding is an improvement, as a longer chain can be replaced with a single small molecule fragment and an entire peptide chain can be substituted for a small molecule rapidly. This method has been shown to work for a variety of PPIs, including the Mcl-1-Noxa PPI and the p53-*hDM2/hDMX* PPI.<sup>203,205</sup>

### 1.16. Aims of the project

As described above, antimicrobial resistance is a huge and growing problem in the modern world and as a result, new treatment options are required for the continuation of successful prophylactic and therapeutic treatment of bacterial infections. There are therefore two possibilities; either new bacterial targets must be discovered and successful binders found for these targets, or new methods of binding existing targets must be found to circumvent current resistance pathways. The aims of the project are therefore to explore bacterial targets and to find compounds that successfully inhibit these targets. Of the two proteins selected for this study, one currently has no pharmaceutical ligands whilst the other, DNA gyrase, has been discussed above and has several drugs targeting it, though resistance and/or poor properties are currently limiting their clinical use.

## CHAPTER 1: INTRODUCTION

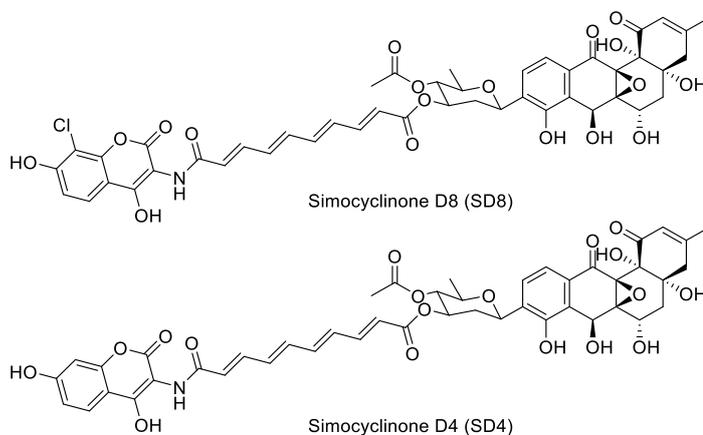
Another group of binders of DNA gyrase not discussed above is the simocyclinone class of natural products. These compounds have been shown to bind successfully *in vitro*, but have poor cell permeability and possess multiple challenging synthetic motifs. Using *in silico* design, and inspired by both peptide-directed binding and FBDD approaches, the aim was to generate a series of “pseudo natural-products” based on the structure of one of the simocyclinones SD8. The initial computational studies performed to design the molecules are discussed in **Chapter 2**, whilst the synthetic work and efficacy studies are discussed in **Chapter 3** and **Chapter 4**.

The BAM complex is an essential bacterial protein complex which is currently not targeted by any pharmaceuticals. Using peptide-directed binding, attempts were made to discover small-molecule inhibitors of an essential PPI within the complex, and to test the efficacy of the resulting compounds using differential scanning fluorimetry. This work is discussed in **Chapter 5**.

**CHAPTER 2: FRAGMENT-BASED DESIGN OF  
SIMOCYCLINONE-INSPIRED NOVEL  
ANTIMICROBIAL AGENTS**

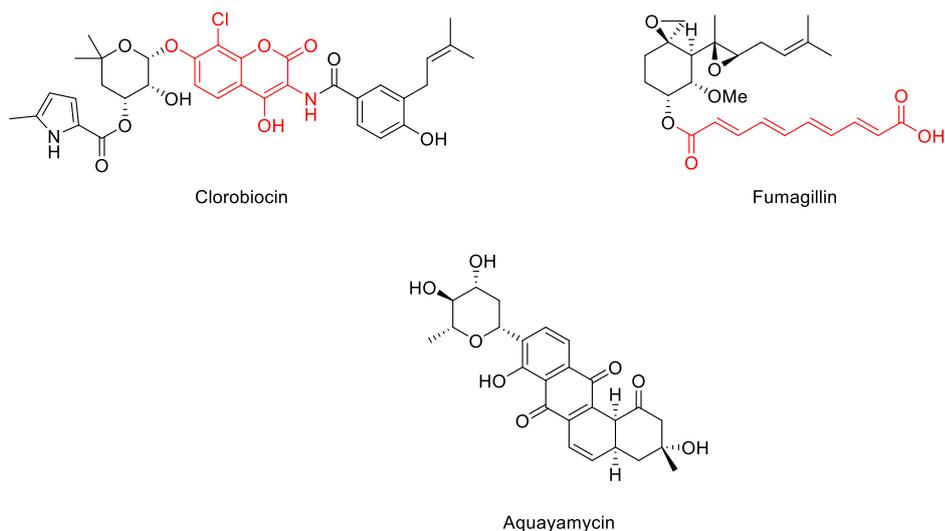
## 2.1. Isolation and structural elucidation of Simocyclinones SD4 and SD8

The simocyclinones were first discovered in a *Streptomyces antibioticus* strain found in an Argentinean soil sample using high-performance liquid chromatography (HPLC) and mass spectrometry.<sup>206</sup> UV signals detected during the HPLC step were compared to a database containing mostly antibiotics, with compounds of interest being further investigated using mass spectrometry. From these experiments, two novel compounds, simocyclinones D4 and D8 (SD4 and SD8 respectively), were revealed. The structures of these compounds are shown below in **Figure 36**.



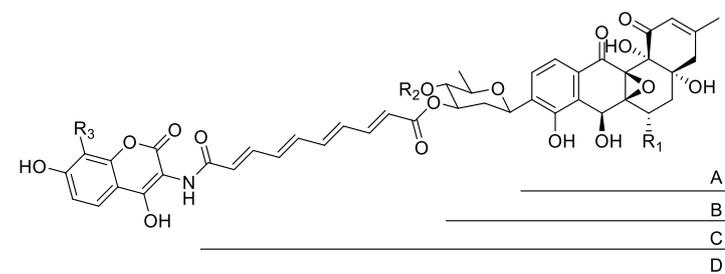
**Figure 36.** Structure of simocyclinones SD8 and SD4.

These structures are considered natural product hybrids, as they contain multiple structural motifs that are similar to previously identified antimicrobials. The left side of the molecule contains a dihydroxylated coumarin, with the coumarin found in SD8 being present in the aminocoumarin antibiotic clorobiocin. The unsaturated tetraene linker can be found in antimicrobials such as fumagillin, whilst the polyketide-sugar region is similar to aquayamycin, a naphthoquinolone antimicrobial (**Figure 37**).



**Figure 37.** Structures of clorobiocin, fumagillin, and aquayamycin. Regions also featuring in SD8 are highlighted in red.

Simocyclinones are split into four smaller groups, labelled A-D, depending on how much of the complete molecule is present. Simocyclinone A structures contain only the polyketide region, simocyclinone B structures contain the polyketide and the sugar, simocyclinone C structures contain all regions except the aminocoumarin, and simocyclinone D structures contain all four structural motifs. The classification is shown below in **Figure 38**. The members of these four families can be generated through biosynthesis by varying the carbon and nitrogen levels available to the bacteria. Examples of different members of the simocyclinone family are shown below in **Table 2**.<sup>207</sup>



**Figure 38.** Generic structure of simocyclinones A-D.

	A1	B2	C2	C4	D4	D6	D7	D8
R <sub>1</sub>	H	OH	OH	OH	OH	OH	H	OH
R <sub>2</sub>	-	H	H	COCH <sub>3</sub>	COCH <sub>3</sub>	H	COCH <sub>3</sub>	COCH <sub>3</sub>
R <sub>3</sub>	-	-	-	-	H	H	Cl	Cl

**Table 2.** Examples of simocyclinones A-D.

## 2.2. Mode of action of SD8

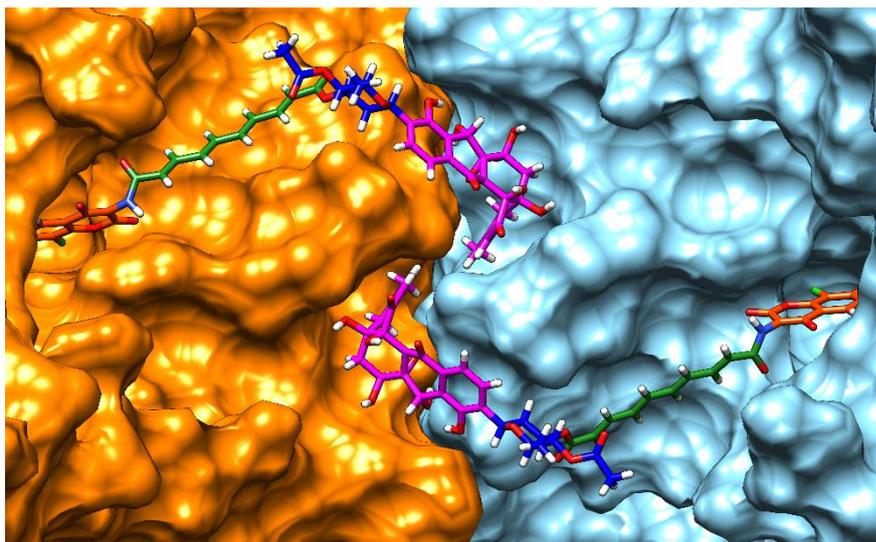
Due to the presence of an aminocoumarin group within the simocyclinones, it was expected that the mode of action of simocyclinones and the aminocoumarins would be similar, i.e. to competitively bind the ATPase active site within GyrB. Whilst SD8 and SD4 inhibited DNA supercoiling which would be catalysed by the ATP-dependent processes of DNA gyrase, it also inhibited DNA relaxation, a surprising result at the time as this process is ATP-independent. SD8 was also compared to the mode of action of fluoroquinolones, the stabilisation of the cleavage complex, however it was shown that SD8 once again did not act in this way. Instead, SD8 appeared to antagonise the activity of these agents to induce cleavage-complex formation.

It therefore appeared that SD8 inhibited the binding of DNA gyrase to the DNA strands rather than interfering with ATP binding. This was assessed using surface-plasmon resonance (SPR), during which DNA strands were tethered to a chip and the binding of DNA gyrase was monitored both in the presence and absence of SD8. This experiment showed that SD8 was able to inhibit binding at concentrations as low as 50 nM. The next step involved examining how DNA gyrase was bound to the SD8. Firstly, SPR was used to determine that SD8 prevents binding of GyrA to DNA immobilised on the chip surface. The observed inhibition was not observed when studying the binding of SD8 onto DNA, further confirming that SD8 binds only to GyrA. Isothermal titration calorimetry (ITC) studies involving the N-terminal domain of GyrA indicated that SD8 binds within the N-terminal domain, the region associated with G-segment DNA binding, with binding constants of between 50 and 100 nM. It was also determined that the number of SD8 molecules bound to complex was double the number of A<sub>2</sub>B<sub>2</sub> complex available, suggesting that two SD8 units bind to the heterotetramer. The results of these experiments therefore suggested that SD8 binds to the N-terminal region of GyrA, the region containing the binding site for G-segment DNA. This provided further evidence for the hypothesis that SD8 inhibited the binding of DNA gyrase to DNA by binding to the DNA gyrase region associated with DNA binding.<sup>157,208</sup>

### 2.3. Binding of SD8 to DNA gyrase

The above hypothesis was further confirmed by X-ray crystallographic structures of SD8 bound to N-terminal GyrA. The first structure, published in 2009 (PDB ID 2Y3P), appeared as a tetramer of GyrA, which consisted of two GyrA dimers cross-linked with four SD8 molecules. In each subunit, two binding sites were observed for SD8, with both sites located within the so-called DNA-binding “saddle” of the GyrA dimer, one site containing the aminocoumarin and the other the polyketide.<sup>163,209,210</sup> When resistant *E. coli* mutants were selected, the mutations occurred in both of these pockets, further confirming the validity of the binding site shown in the crystal structure. However, other site-directed mutants could not be fully rationalised, indicating that this tetrameric structure may not occur *in vivo*. It was therefore thought that the solid-state nature of a crystal structure stabilised the dimer-dimer interactions, and so the GyrA-SD8 complex was analysed in solution using nano-electro-spray ionisation mass spectrometry. This revealed that the tetramer could be reproduced in solution at high concentrations of SD8, but at lower concentrations a dimeric species appeared with two SD8 molecules bound to each dimer.<sup>209</sup>

This dimeric species was confirmed by a later crystal structure (PDB ID 4CKL) utilising a shorter N-terminal GyrA region which lacked the residues required to stabilise the tetrameric species. This structure was entirely consistent with the SD8 resistance mutations previously described, in addition to further mutants which were also selected. The conformation of the SD8 in the binding pockets differed from the previous structure, with the aminocoumarin oriented slightly differently and the binding pocket for the polyketide being shifted so that it spanned the interface of the two monomers. This new binding interaction indicates that SD8 staples the GyrA dimer closed, thus preventing the conformational changes required to bind DNA.<sup>211</sup> This stapled dimer is shown below in **Figure 39**. The authors also found evidence that SD8 bound to a binding site in GyrB, but roughly 1000-fold weaker than the binding to GyrA. It is therefore unlikely that this binding contributes to SD8's inhibitory effects on DNA gyrase.<sup>157,211</sup>

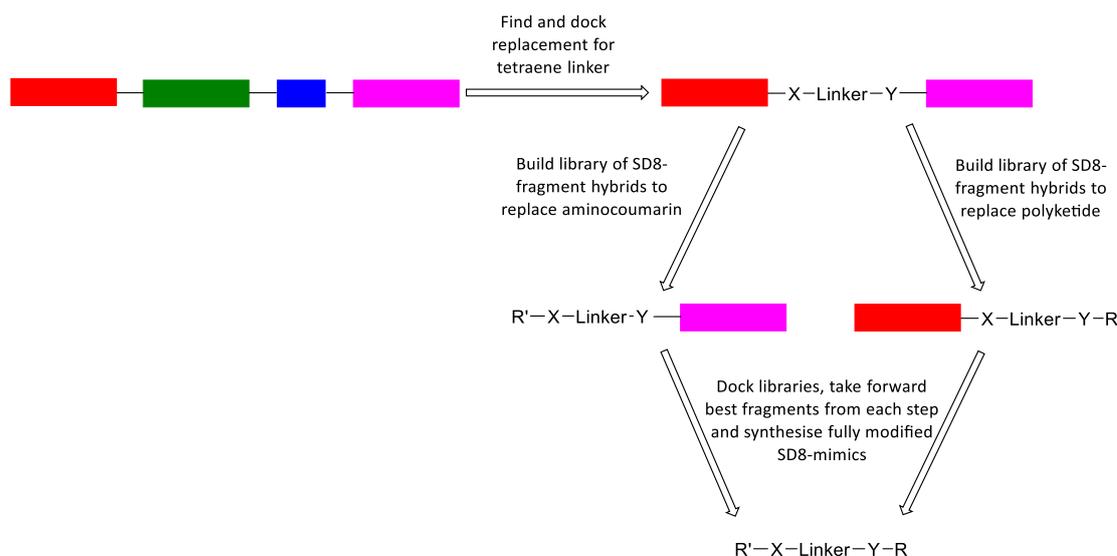


**Figure 39.** View of the GyrA-GyrA dimer as stapled by SD8. The aminocoumarin region is coloured orange, the tetraene green, the D-olivose blue, and the angucyclinone magenta. PDB ID: 4CKL.

#### 2.4. Aims of the study

Whilst SD8 is a potent antibiotic, its chemical synthesis can be quite problematic. Previous work within the group attempted to design a synthetic pathway for the aminocoumarin, polyketide and tetraene linker in an attempt to make a coumarin-quinolone hybrid. However, whilst this work was largely successful, the synthesis of the polyketide was lengthy and resulted in an extremely low yield, the synthetic pathway to aminocoumarins was unable to tolerate the addition of the chlorine atom in its position as found in SD8, and, whilst the tetraene linker was synthesised successfully, connection of the aminocoumarin to the tetraene linker was found to be challenging.<sup>3</sup> Furthermore, the synthesis and attachment of the sugar to both the linker and the polyketide regions were hypothesised to be complex, owing both to the challenges associated with generating the desired sugar and practically working with the sugar, in addition to the previously discussed difficulties with attachment of the tetraene. The aims of the project are therefore to design and generate a series of synthetically feasible mimics based on SD8, and to examine the biological activity of these compounds. The design work, conducted *in silico*, will be discussed in this chapter, with the synthetic work and biological testing conducted to be discussed in **Chapter 3** and **Chapter 4**.

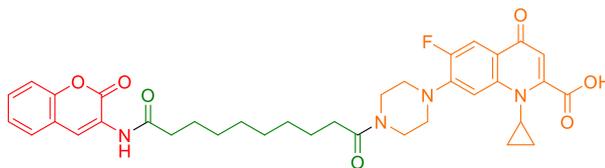
The design work can be split into three sections; the aminocoumarin region, the polyketide region, and a linker with simple attachments on both sides. The polyketide and aminocoumarin regions will be tackled separately from one another; however, the first region identified for modification was the tetraene-sugar region, the replacement of which will feature during the replacement of the polyketide and aminocoumarin regions. **Figure 40** outlines the planned method by which mimics of SD8 will be designed *in silico* during this project.



**Figure 40.** Outline of *in silico* design of SD8 mimics. X and Y represent linkages, whilst “Linker” represents a linker of unknown structure and length: Red; aminocoumarin, Green; tetraene, Blue; D-olivose sugar, Pink; polyketide.

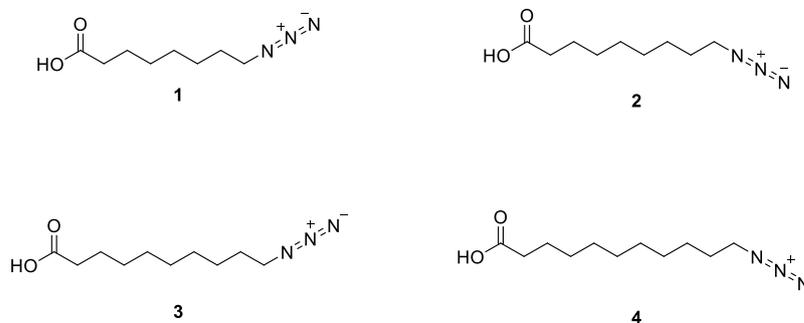
## 2.5. Computational design of a replacement linker

Previous work within the group focused on the design and synthesis of a small group of coumarin-quinolone hybrid molecules. These molecules, consisting of a selection of aminocoumarins, a sebacic acid linker, and a selection of quinolones including ciprofloxacin were found to possess inhibitory activity against GyrA *in vitro* (**Figure 41**),<sup>212</sup> and so this sebacic acid linker was used as a starting point for the linker design process. However, since the quinolone class of antibiotics and the polyketide region of SD8 do not share a binding pocket, it was decided that multiple lengths of aliphatic chains should be tested to determine the optimal length.



**Figure 41.** The structure of the best aminocoumarin-quinolone hybrid compound obtained in previous work. The molecule is made up of an aminocoumarin (red), a sebacic acid linker (green), and ciprofloxacin (orange).

In addition to this, the linker had to not only replace the tetraene region of SD8 but also the sugar ring. Examination of several functional groups based on their size and their ability to place the polyketide in a similar location when compared to the original SD8 molecule revealed that a triazole ring would be a suitable replacement for the sugar, mimicking the placement of the SD8 polyketide very closely. Therefore, one side of the linker would consist of an azide, allowing for a library of polyketide-replacing terminal alkynes to be attached using click chemistry *in silico*. The other side of the linker was to be used to form amide bonds similar to those found between the tetraene and aminocoumarin of SD8. As such, the other side of the linker would consist of a carboxylic acid group, which would then allow a library of primary and secondary amines to be attached *in silico*. These design choices, as well as the variety of linker lengths selected for examination, are summarised in **Figure 42**.



**Figure 42.** Azido-carboxylic acid linkers explored during initial design process.

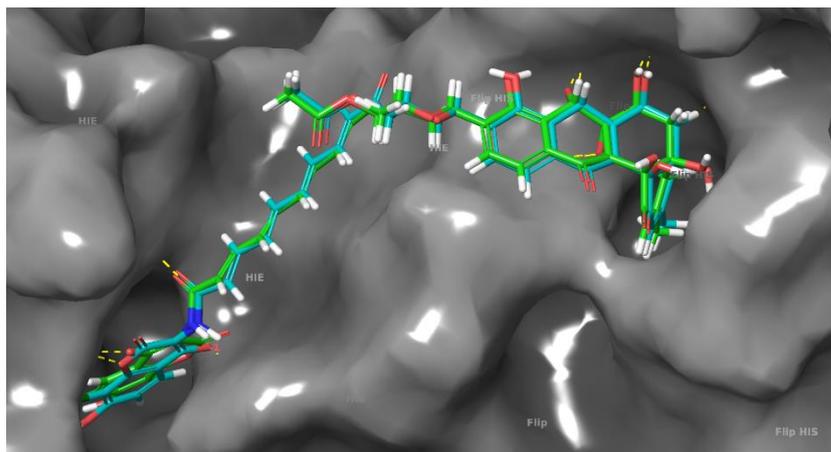
These linkers were then examined to determine how efficiently they are able to replace the tetraene region of SD8, as well as to obtain the linker which would be taken forward for future *in silico* work. For this, the crystal structure of the GyrA-SD8 complex obtained by Hearnshaw *et. al.* (PDB 4CKL) was used as a basis for docking experiments.<sup>211</sup> One molecule of SD8 was separated from the crystal structure, which was then minimized. This was performed using the LigPrep functionality of Schrodinger Maestro, and consisted of the addition of absent implicit hydrogens, removal of counterions and water molecules, neutralisation of charged groups, generation of possible ionization states at a given simulated pH (in this case, pH 7 ± 2), the generation of tautomers, finding low ring-energy conformations, and optimization of the geometries of the generated molecules to find the lowest energy conformation.<sup>213</sup>

Once the ligand was correctly minimized, the SD8 ligand was then docked into the protein using Glide (Grid-based Ligand Docking with Energetics) docking.<sup>214–216</sup> This step was performed to validate the Glide process, as well as the precise methodology which was expected to be used in later steps. Glide docking is a flexible ligand-fixed protein docking method, meaning that the atomic positions of the protein are kept constant whilst the atomic positions of the ligand are allowed to move. Each set of ligand positions, known as a pose, are passed through a series of filters to evaluate how well the pose interacts with the receptor, initially being assessed on how well the ligand fits into the protein structure in a spatial sense before moving on to examine how well the ligand-receptor interactions complement the pose. The structure is then minimized and scored, giving a value known as the Glidescore. This is a sum of a variety of terms, as shown in **Equation 3**, and provides an indication of how well the ligand is able to fit into the receptor.<sup>214</sup>

$$GScore = 0.05(vdW) + 0.15(Coul) + Lip + Hbon + Met + Rew + RotB + Site$$

**Equation 3.** Calculation of the GlideScore (GScore). Abbreviations: vdW (Van der Waals energy); Coul (Coulombic energy); Lip (Lipophilic interactions); HBon (Hydrogen-bonding interactions); Met (Metal-binding term); Rew (Rewards and penalties); RotB (Penalty for freezing rotatable bonds); Site (Polar interactions in active site).

The SD8 molecule was docked into the GyrA dimer using the extra-precise setting. This setting exposes the ligand to further refinement tools, further decreasing the number of false positive results yielded from the protocol. Following the docking of the SD8 molecule, the structures of the SD8 molecule before and after re-docking were overlaid and compared. As can be seen in **Figure 43**, the structures closely matched one another, and the root-mean-square deviation (RMSD) of the redocked molecule relative to the crystal structure was 1.372 Å, indicating that the methodology was suitable for this application.



**Figure 43.** Overlay of re-docked SD8 (cyan) on SD8 bound to GyrA in crystal structure (green). PDB ID 4CKL.

After ensuring that the docking platform used for this project was able to reproduce the binding interactions found in the original crystal structure, each of the linkers was then inserted into the SD8 molecule using the Core-hopping application.<sup>217</sup> This tool allows for part of a given molecule, known as the core, to be swapped out for one or more replacement cores provided by the user, whilst maintaining the rest of the molecule. The precise method used in this case was Glide-based core-hopping, in which the core molecules are aligned using a pre-prepared Glide grid, writing the resulting structures without the side-chains present. The side chains are then added in and the entire molecule docked into the receptor in a similar fashion to the SD8 molecule discussed previously. The resulting poses can then be visualised and examined, as well as the resulting Glidescores. In this case, the core was defined as containing both the tetraene and the D-olivose sugar region, allowing for the impact of replacing these regions whilst maintaining the aminocoumarin and polyketide regions. This required the adaption of linkers **1-4** to include the completed triazole ring in each case, resulting in the structures shown below in **Figure 44**.

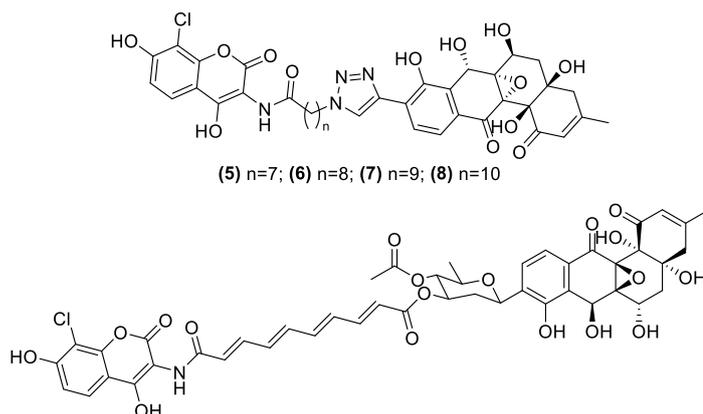


Figure 44. Compounds generated *in silico* after "core-hopping" compounds 1-4 into SD8 (bottom)

Following completion of the Glide-based core-hopping run, both the spatial arrangement and Glidescore of compounds 5-8 were examined. Whilst compounds 6-8 were found to present the aminocoumarin and polyketide regions to the GyrA binding pockets in a manner comparable to SD8, compound 5 showed that the shorter linker prevented the bulky binding groups from fully entering their respective pockets and so was theorised to be an ineffective linker. However, compounds 6-8 also demonstrated that the alkyl linkers were extremely flexible, possibly affecting the binding of the generated compounds. Nevertheless, this result demonstrated that an alkyl linker would be a viable replacement for the tetraene linker found in SD8. The poses of compounds 6-8 *in silico* are presented below in Figure 45, whilst the Glidescores of these molecules are given in Table 3.

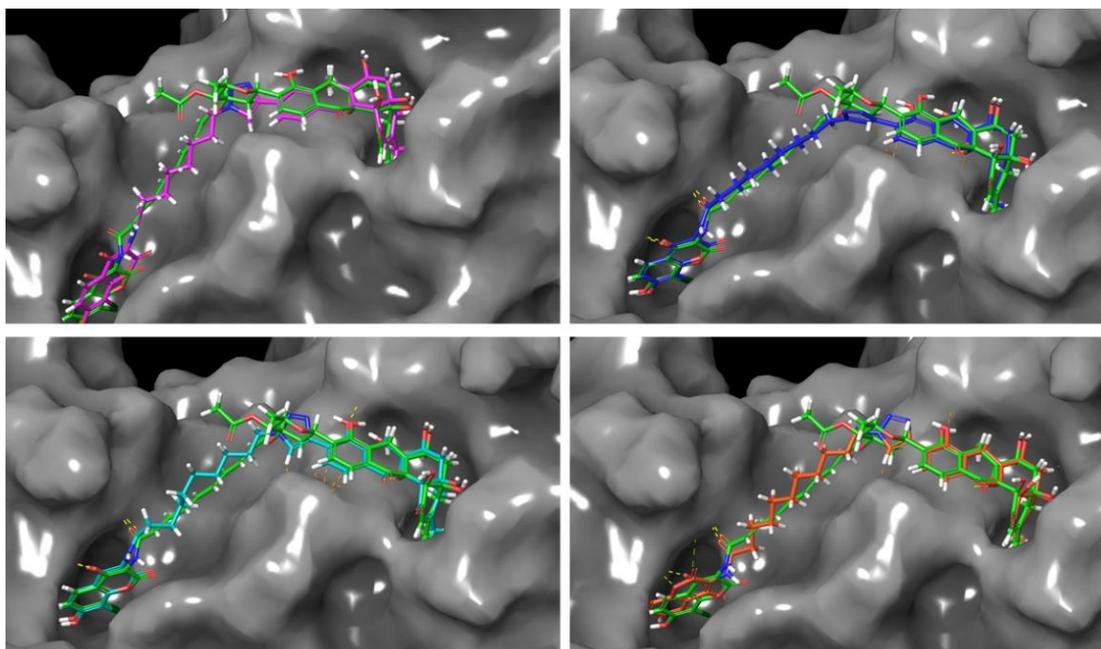


Figure 45. Docking poses of 5 (magenta), 6 (blue), 7 (cyan) and 8 (orange) compared to SD8 bound to GyrA (green). PDB ID: 4CKL

Compound	Glidescore
5	-5.977
6	-7.979
7	-8.590
8	-8.431

Table 3. Glidescores of compounds 5-8

As can be seen, the linker of **5** appeared to be too short, resulting in the coumarin being suspended over the binding site instead of being placed inside the site, with this being reflected in the Glidescore by a seemingly large penalty. The linkers present in compounds **6** and **7** placed both the polyketide and coumarin regions into their respective binding sites to an acceptable degree, with compound **7** more closely resembling the original SD8 conformation. Compound **7** was also found to have the best docking score. The linker of **8** appeared to be another good fit for replacement of the SD8 linker, but resulted in a different conformation of the coumarin to be adopted. As a result of these observations, the decision was made to proceed with the linker found in compound **7**, with the option to switch to the linkers found in compounds **6** or **8** if required.

## 2.6. *In silico* replacement of the SD8 polyketide

The polyketide region, which would need to be replaced with a readily synthesisable fragment possessing a terminal alkyne suitable for undergoing the copper-catalysed click reaction, was then examined. A library was constructed using Aldrich's market select service, an online database containing a large number of small molecules available for purchase. The library constructed contained alkynes of molecular weight between 100 and 1000 Da, and was filtered to remove unwanted non-terminal alkynes, leaving roughly 5000 compounds within the library. Upon preparing these ligands, this number rose to over 11000 unique poses, highlighting a new problem; libraries of this size are extremely computationally expensive to dock, and the library size had to be reduced. To do this, the range of molecular weights explored was reduced to 120-700 Da, somewhat more in line with the molecular weight of the polyketide. This also had the added benefit of removing lower weight small molecules which had little variation between each other, as well as removing heavier, larger molecules which would be more unsuited to the fragment-based methodology used in this work. This process yielded a library of just over 2000 compounds, and after ligand preparation this provided a library with ~4000 compounds. The process of obtaining the alkyne library is summarised below in **Figure 46**.

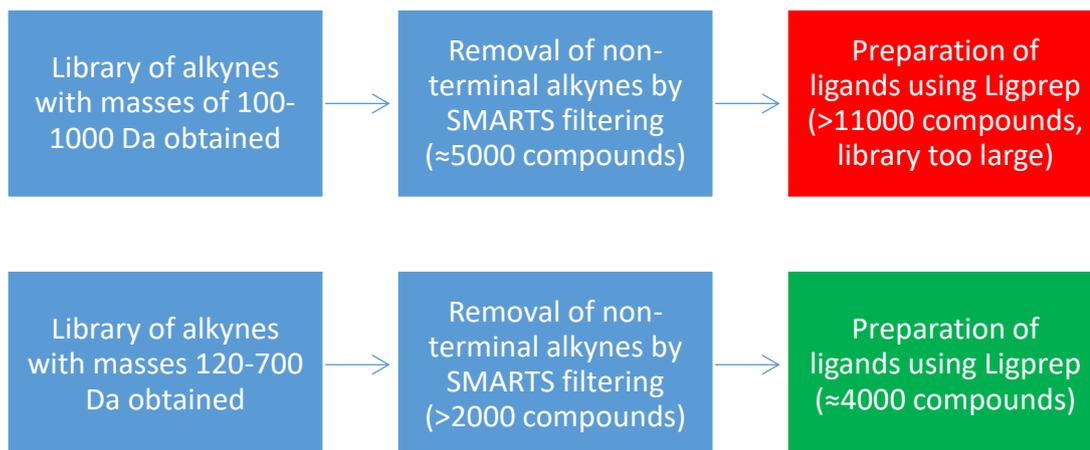


Figure 46. Summary of alkyne library preparation work.

Meanwhile, compound **7** was modified by removal of the triazole and polyketide groups and replaced with an azide group (**Figure 47**). This would then allow the addition of the generated library to the linker-aminocoumarin compound utilising another application known as Covalent Docking.<sup>218,219</sup> This software performs chemical reactions, provided in the form of SMARTS (SMILES arbitrary target specification) code by the user, between a ligand and a receptor. In the context of the work discussed here, the receptor was defined as the aminocoumarin-azide compound obtained by modifying the docked, core-hopped **7**, as shown below in **Figure 47**. The ligands were defined as the library of terminal alkynes. The reactive residue, in this case the azide, is mutated to an alanine to prevent any influencing of the ligand by the receptor, before the ligand is docked into the receptor using Glide, generating a series of poses. Once these poses have been identified, the azide residue is restored, and the poses are once again explored to determine the best fit. Finally, the covalent bonds are formed, generating a triazole, and the resulting structure is minimized to relieve strain. As the aminocoumarin-azide compound used was already docked into the GyrA crystal structure, it was anticipated that the resulting click compounds would also adopt poses in a similar fashion to if they had been docked into the crystal structure.

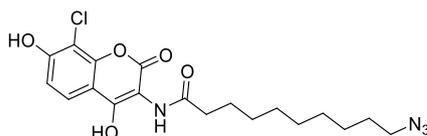
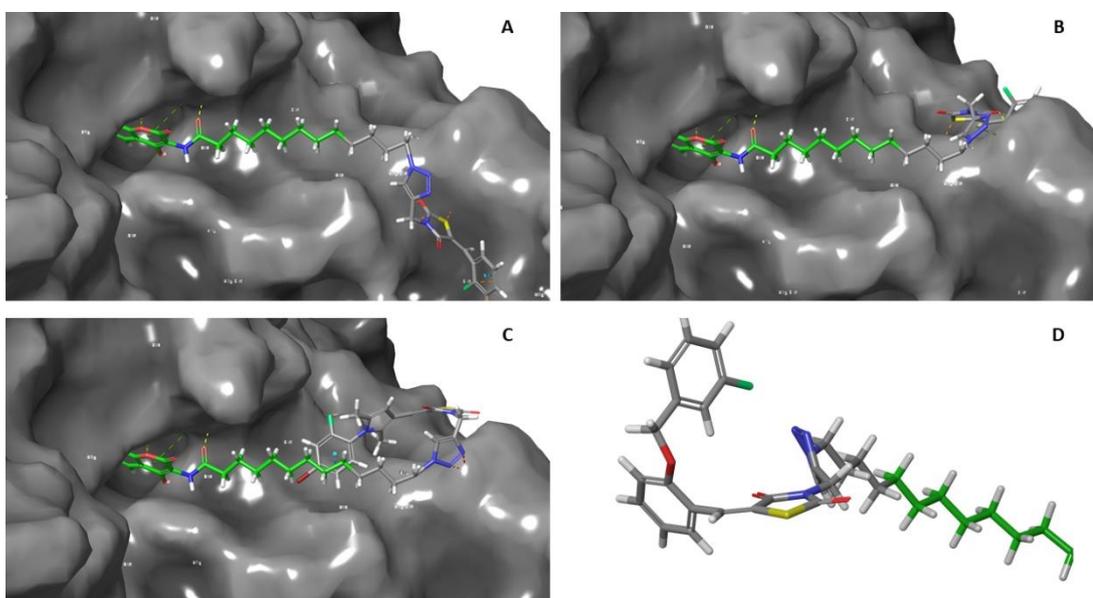


Figure 47. Structure of the aminocoumarin-linker-azide in silico compound.

Upon running the covalent docking job, however, it was discovered that the software was unable to reliably give reasonable structures for the new SD8 mimics, with structures with non-planar triazole rings, and stretched and compressed alkyl chains being suggested as potential binders by the program (**Figure 48**). This was determined to be due to the fact that the job was run using the “Virtual Screening” settings of the software, which skips certain steps including minimization to allow larger libraries to be run in a timely manner. The decision was made, therefore, to attempt the covalent docking again using the “Pose Prediction” settings, allowing the software to run all steps of the docking at the cost of a longer run time. This longer run time proved to be a problem, however, as the first 10 ligands took 2 days to run, albeit with good pose prediction.



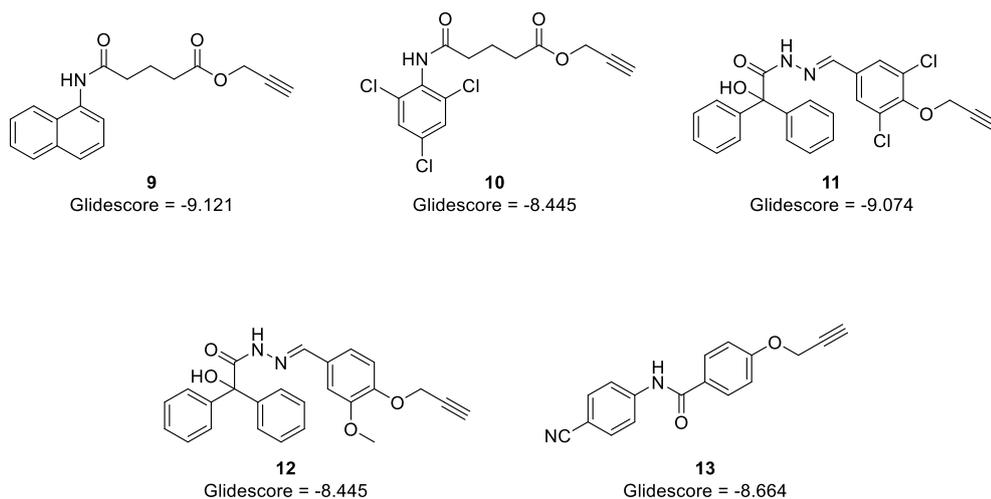
**Figure 48.** Examples of erroneous structures generated by Covalent docking. A-C possess inappropriate bond lengths and angles, whilst D possesses a puckered triazole ring. PDB ID: 4CKL

A new method was therefore required to both make the compounds *in silico* and to dock them into the binding site on GyrA. It was therefore decided to perform the *in silico* click reaction and the docking separately from each other. A Python script, written and provided by Marco Cominetti, was therefore used to generate the library of triazoles which could then be prepared and fed into the Glide docking procedure of Schrödinger Maestro as before. This script was not a perfect solution however; both the input and output of the script were two-dimensional, causing the loss of stereochemical information from the alkyne library. The preparation process converted the structures back into three-dimensional structures and so added stereochemistry back into the molecules, but there was no reasonable method to reintroduce only the stereochemistry found in the alkyne

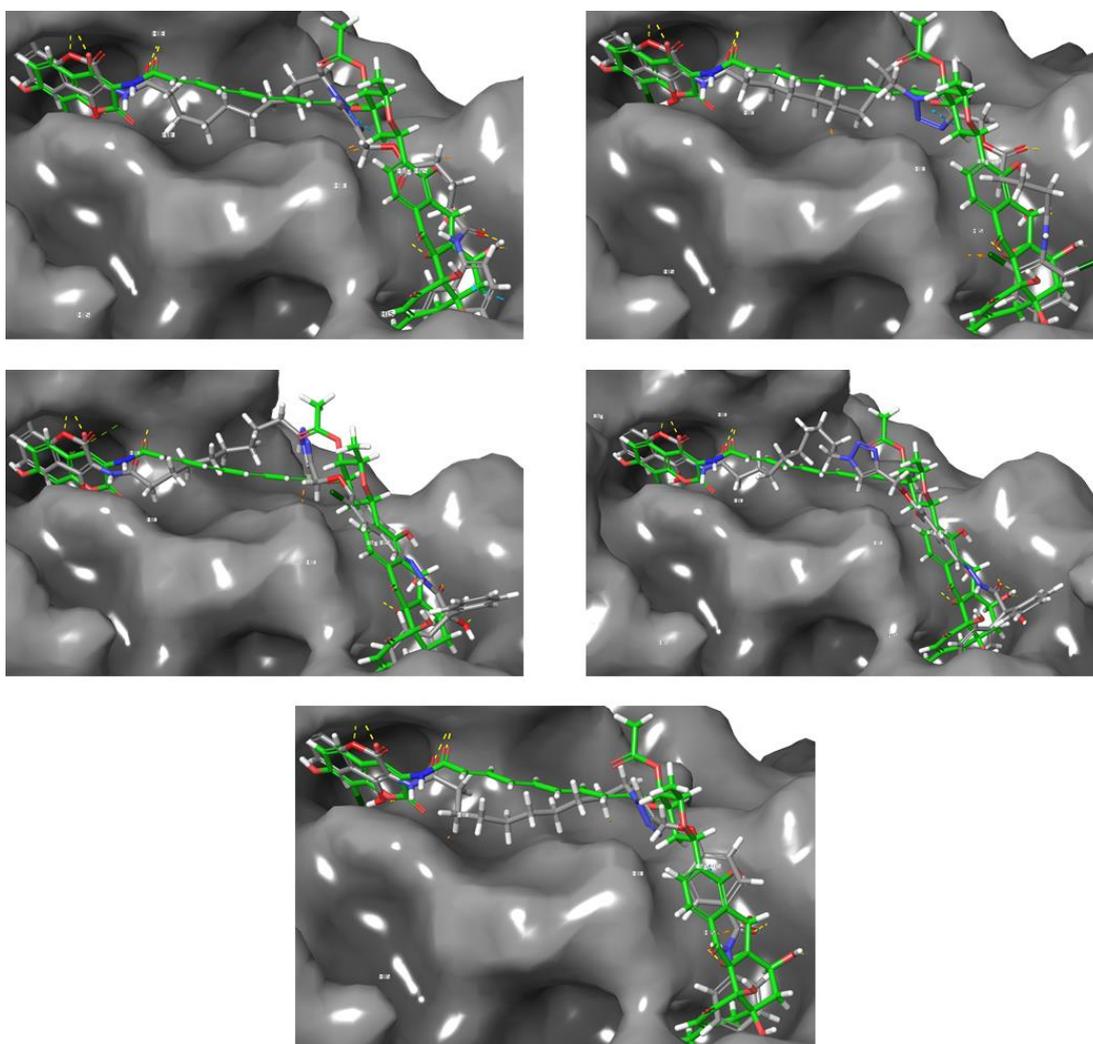
library. This issue, however, has now been identified and resolved within the group, so future work involving this script will no longer incur this problem.

Nevertheless, the 4039 compounds generated by the preparation process were then fed into a Glide docking run, using the original docked SD8 molecule as a reference ligand and constraining the aminocoumarin region into its binding region. This allowed the examination of the effect of the replacement of the angucyclinone region alone. From this run, 5607 poses were determined, all of which seemed to have maintained correct structural features (i.e. no ring puckering of the triazole, acceptable bond angles). The next task was then to identify the alkynes which had been predicted to fit best into the angucyclinone binding site. This was made more difficult by the fact that the alkyne library member had been taken from three-dimensional to two-dimensional, an *in silico* chemical reaction performed on it, and then prepared, causing sufficient alterations in the structure that the SMILES (Simplified Molecular-Input Line-Entry System) code no longer matched that of the original alkyne. In addition, some poses shared alkyne library members, further complicating the process.

It was therefore decided that the Python script responsible for performing the *in silico* click reaction could be modified to include an additional "Original SMILES" parameter, which could then be carried by the complete molecule through the ligand preparation and Glide docking processes. The steps listed above were therefore repeated and the results were then fed into another Python script, again provided by Marco Cominetti, which examined both the Glide score and Original SMILES parameters and displayed the top 30 scoring molecules derived from unique library members. However, when examining the cost of these library members it was discovered that none of the high scoring linkers could be purchased at a reasonable price. Work therefore began to design synthetic routes to the molecules which were both chemically and financially feasible. This work indicated five alkyne library members, **9-13**, shown in **Figure 49**, which would be suitable for initial synthesis and study. These compounds were selected due to their comparable docking scores to SD8 and compound **7**, their relatively simplistic synthetic pathways, as well as the low cost of the starting materials. The *in silico* docking poses of these compounds are shown below in **Figure 50**. As can be seen in **Figure 49**, three of the five molecules selected for further study possess higher Glidescores than compound **7**, suggesting that these compounds may actually bind to GyrA in a more efficient manner than the polyketide region of SD8.



**Figure 49.** Alkyne library members selected for synthesis and study, including the calculated Glidescores for the compounds.



**Figure 50.** In silico docked poses of SD8-fragment hybrids (grey) derived from compounds **9** (top left), **10** (top right), **11** (centre left), **12** (centre right) and **13** (bottom). The green structure in each image represents the pose of SD8 when bound to DNA gyrase. PDB ID: 4CKL.

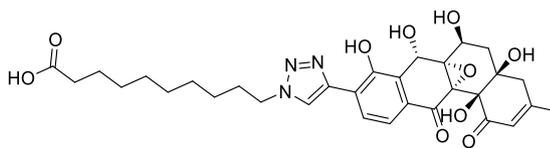
## 2.7. *In silico* replacement of the SD8 aminocoumarin

Separately, work to replace the aminocoumarin was also undertaken. These molecules would be required to be able to form amide bonds, and so would be primary or secondary amines. A library of compounds to replace the aminocoumarin region was also obtained using Aldrich Market Select, examining compounds of molecular weight between 120 and 700 Da, and was filtered to remove the unsuitable tertiary amine and quaternary ammonium salt compounds, yielding a library containing over 12,000 compounds. It was therefore necessary to impose further restrictions on the amine library to allow for computational study to proceed. Firstly, the library was filtered to remove compounds possessing motifs identified as pan-assay interference compounds (PAINS). Compounds containing PAINS motifs have been found to be frequent hitters of several assays, yielding false positives. Removal of these compounds, therefore, is important, as their presence in any further assays may generate false results.<sup>220</sup> However, filtering based on PAINS using Schrödinger Canvas<sup>221</sup> did not sufficiently lower the size of the library, and so further filtering was required.

The library was therefore filtered based on Lipinski's rule of five.<sup>222</sup> Compounds with molecular weight of over 500 Da were removed, as were structures containing more than 5 hydrogen bond donors or 10 hydrogen bond acceptors. The AlogP for the remaining members of the library was calculated using Schrödinger Canvas<sup>221</sup>, with only library members with AlogP values of between -0.4 and +5.6 being kept. AlogP is a method of predicting the logP of a molecule computationally by examining atomic contributions and by categorising atoms into one of several atom types depending on the environment that the atom is in within the molecule.<sup>223</sup> Whilst AlogP is less accurate than ClogP, another calculation method which uses the contribution of non-overlapping molecular fragments,<sup>224</sup> it has been shown to be only marginally so, and has also been shown to be more applicable in the presence of unexpected and unusual functional groups.<sup>225</sup> This led to a library of 3900 compounds which could then be used in covalent docking experiments.

Compound **7** was modified once again, this time by removing the aminocoumarin and converting the amide into a carboxylic acid, whilst also leaving the angucyclinone-triazole region attached to the linker (**Figure 51**). This carboxylic acid could then be used to generate amides using the *in silico* amine library. However, when covalent docking was attempted there were several problems discovered. Firstly, in a similar fashion to the alkyne library work, covalent docking resulted in poor structural prediction when run on

“Virtual Screening” mode, and was too computationally expensive using “Pose Prediction” mode. Secondly, it was discovered that filtering only to obtain primary and secondary amines still allowed other unsuitable functional groups, such as amides and hydrazines, to be carried forward. These would also need to be excluded using the ligand filtering tools available, or by further refinement of the reaction SMARTS code to ignore these functional groups.



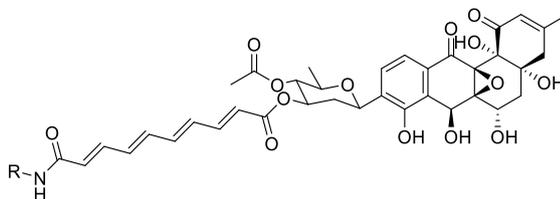
*Figure 51. Structure of the linker-triazole-angucyclinone in silico compound.*

To solve these problems, it was once again decided to perform the reaction separately from the Glide docking procedure, as was done for the alkyne library. The library was re-prepared from Aldrich Market Select, this time ensuring that as many compounds as possible with molecular weight between 120 and 500 Da were included, and was filtered for PAINS, Lipinski’s rule of five, and to remove an extended range of functional groups. This yielded a library of over 14000 compounds, and so was therefore also passed through the constraints of a modification of Lipinski’s rule of five designed to identify lead-like compounds. These constraints, together named the “Rule of Three”, include limiting the logP of the compounds to less than or equal to 3, the molecular weight to less than 300 Da, the number of hydrogen bond donors and acceptors to less than 3, and the number of rotatable bonds to less than 3.<sup>226</sup> This reduced the library size to 2990 members, suitably small to begin preparing for Glide docking.

Once again, the reaction between the library members and the modified compound **7** was performed using a Python script provided by Marco Cominetti, and the resulting compounds were then prepared using the ligand preparation tool, generating 9731 compounds. Glide docking was then attempted using these compounds, constraining the angucyclinone to its binding pocket, generating 1043 poses. Unfortunately, most of the compounds fed into the docking run failed to be included in the final pose list, with the software unable to correctly minimise the structures. This was initially thought to be an issue with the constraints applied, and so modification of the constraints was attempted. This came in the form of selecting angucyclinone atoms deeper into the binding pocket, as well as constraining the alkyl chain instead of the ring system. However, upon reattempting

the Glide docking protocol with these constraints, similar numbers of poses were obtained with an unacceptable failure rate.

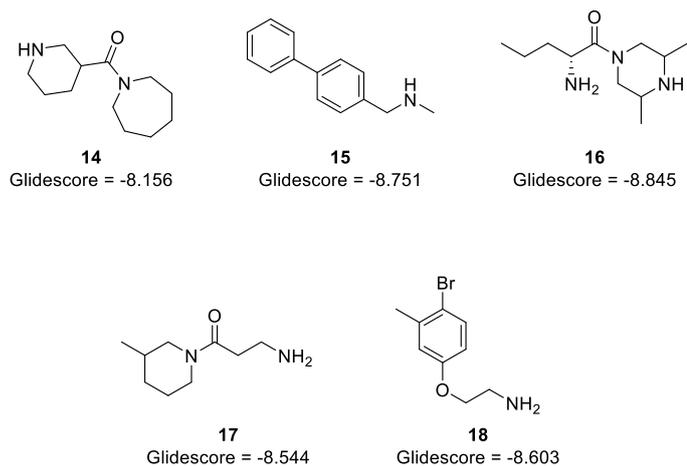
It was therefore decided to revert from the alkyl linker of compound **7** back to the tetraene originally found in SD8. This was thought to simplify the molecule substantially, due to the inherent geometric constraints of conjugated alkene systems. The tetraene would also hold the newly attached library members in place over their binding pocket, possibly furthering the reliance of the glide score on the suitability of the amine library member. The prepared library was attached in the same way as described before to a modified version of SD8, in which the aminocoumarin amide bond had been removed and replaced with the carboxylic acid required to attach the amine library, resulting in 2990 compounds following the general structure shown in **Figure 52**.



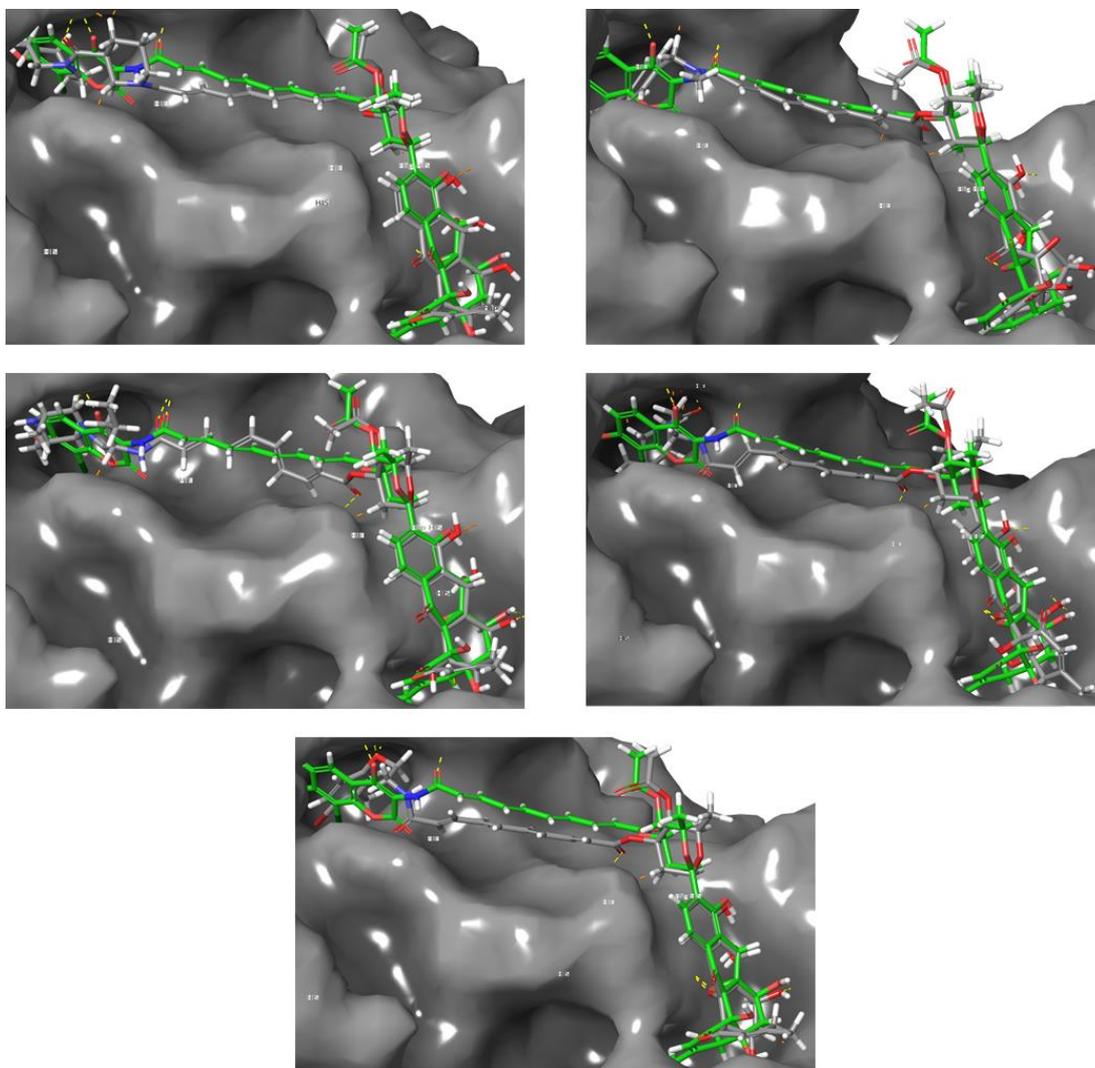
**Figure 52.** General structure of SD8-amine library hybrids, where *R* represents the library members

The Glide docking job was then run, constraining the angucyclinone into its binding pocket, resulting in the generation of 3196 poses with no observed structural errors. More importantly, however, no library members were excluded by the program due to minimisation errors, indicating that this change had satisfactorily corrected the errors observed prior. In a similar fashion to the poses predicted from the alkyne library, the resulting poses were fed into the result parse Python script to obtain the top 30 unique library members. From these, one of the members, **14**, was found to be purchasable, and so was taken forward. The synthetic feasibility and cost was then explored for the remaining 29 compounds, with four more compounds, **15-18**, identified by this process. The five selected compounds are shown below in **Figure 53**, with the docked poses of the structures related to these compounds shown in **Figure 54**. As can be seen by the docking scores given in **Figure 53**, the selected fragments were predicted to have similar docking affinity to the aminocoumarin, as the docking scores were close to that calculated for compound **7**. However, it is important to keep in mind that the docking experiments for compounds **14-18** were performed using a closer analogue to SD8 than compounds **9-13**, potentially leading to higher docking scores than would have been observed if the docking experiment was performed in a more similar fashion to the alkyne docking experiment.

## CHAPTER 2: FRAGMENT-BASED DESIGN OF SIMOCYCLINONE-INSPIRED NOVEL ANTIMICROBIAL AGENTS



*Figure 53. Amine library members selected for synthesis and study, including the calculated Glidescores for the selected compounds*



*Figure 54. In silico docked poses of SD8-fragment hybrids (grey) derived from compounds **14** (top left), **15** (top right), **16** (centre left), **17** (centre right) and **18** (bottom). The green structure in each image represents the pose of SD8 when bound to DNA gyrase.*

## 2.8. Conclusions

The work described in this chapter has provided an insight into the structures required to mimic all regions of the natural product SD8. It has shown that 10-azidodecanoic acid (**3**) acts as a suitable linker for both the aminocoumarin and polyketide regions of the molecule *in silico*, with 9-azidononanoic acid (**2**) and 11-azidodecanoic acid (**4**) also delivering these regions to their respective binding pockets adequately, albeit to a lesser extent than the decanoic acid. It has also been shown that 8-azidooctanoic acid (**1**) does not possess sufficient length to position both sides of the SD8 molecule in their respective binding pockets *in silico*, thus ruling it out as a potential substitute linker.

This work has also identified a selection of alkyne small molecules, **9-13**, that show the potential to mimic the binding activity of the polyketide region of SD8 *in vitro*. The molecules were found to possess good *in silico* binding affinity to the GyrA dimer, with the calculated Glidescores for the compounds closely resembling and in some cases surpassing the Glidescore observed for parent compound **7**. The molecules were also thought to be readily accessible by synthesis. All molecules possessed a propargyloxy motif, allowing the attachment of the alkyne region by reaction with either propargyl bromide or propargyl alcohol, and most of the molecules required only an amide formation in addition to the generation of the propargyloxy intermediates.

In addition, this work has identified a series of amine small molecules, **14-18**, found to suitably replace the aminocoumarin region of SD8 *in silico*, and therefore hoped to correctly mimic this region *in vitro*. Once again, these molecules were selected firstly by their binding *in silico*, with all selected compounds possessing similar Glidescores to the parent compound (**7**), as well as due to their synthetic feasibility and the relatively low cost of starting materials. In the case of the amines, however, the syntheses involved are more varied, due to the absence of a chemical motif common to all of the selected molecules. In spite of this, all molecules selected were thought to be accessible via low-cost synthesis, or were available for purchase. Furthermore, it is important to note that the docking experiment involving the amine library used a polyketide-linker-acid compound which more closely resembled SD8, which may have affected the Glidescores of the resulting amides.

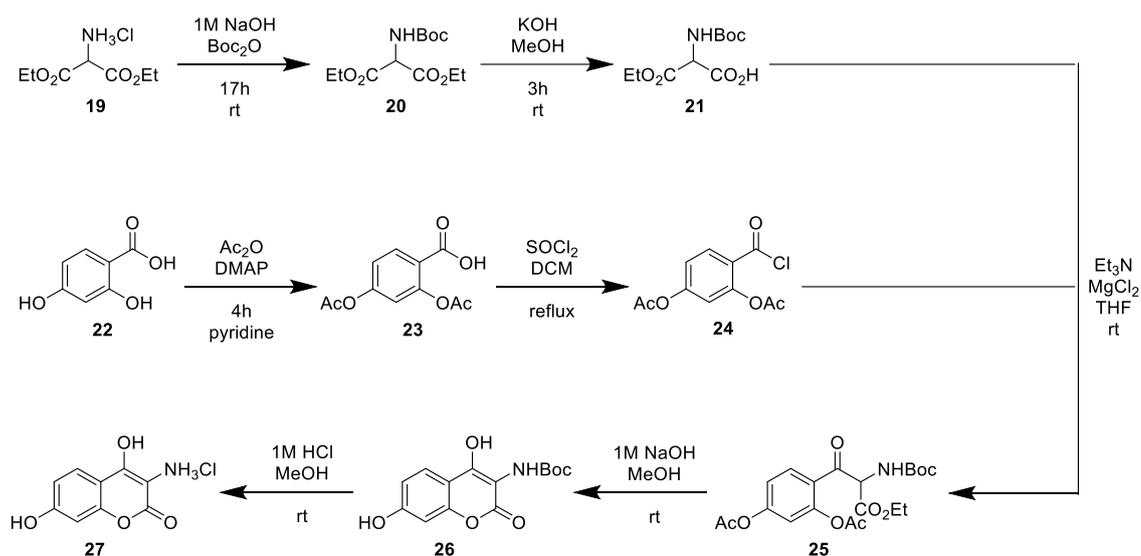
Ideally, the work following these *in silico* experiments would be the synthesis of SD8-small molecule hybrids, with the alkyne library being attached to the 10-azidodecanoic acid linker with the SD8 aminocoumarin attached to the acid, and the amine library being attached to the decanoic acid linker with the SD8 polyketide attached. However, due to the complexity of the synthesis of the SD8 polyketide, it was decided that the alkyne library members would first be synthesised and attached to the aminocoumarin-linker compound, and their activity assessed. The best alkyne library member(s) would then be taken forward and used to form the library of amine-linker-alkyne complexes required to assess the activity of the amine library members.

**CHAPTER 3: AMINOCOUMARIN-LINKER-  
ALKYNE COMPOUNDS AS POTENTIAL MIMICS  
OF SD8**

Assessment of the *in vitro* properties of the compounds found in **Chapter 2** began with the synthesis of aminocoumarin-linker-alkyne complexes. These compounds were similar in structure to compound **7** and SD8, with the aminocoumarin simplified and the polyketide replaced with fragments attached to terminal alkynes, which would be used to generate triazoles through the use of click chemistry. Previous work within the group indicated that the aminocoumarin found in SD8 was difficult to synthesise due to the presence of the chlorine atom, which deactivates an intermediate compound and hence prevents the synthesis of the aminocoumarin. However, a simplified aminocoumarin was successfully generated, and the work also presented a method for attaching aminocoumarins to sebacyl chloride under basic conditions, hence providing a potential route for attachment of the aminocoumarin onto the decanoic acid linker.<sup>3</sup>

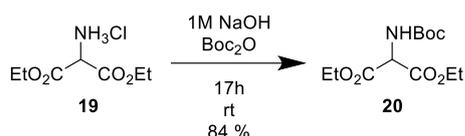
### 3.1. Synthesis of the aminocoumarin

The first step in generating these compounds was to construct the aminocoumarin-linker regions which would be common to the five final molecules. The aminocoumarin which was to be used for these studies was the ring system found in SD4, another member of the simocyclinone D family. Whilst SD4 possesses less binding affinity to DNA gyrase than SD8, the synthetic challenges associated with generating the chlorinated aminocoumarin have been explored in previous work within the group. The reaction scheme for the synthesis of the SD4 aminocoumarin (**27**) is outlined below in **Scheme 1**.



*Scheme 1. Generation of aminocoumarin 27*

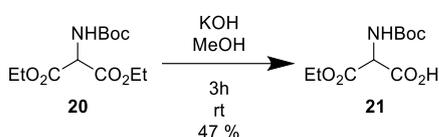
Compound **20** was generated as a colourless oil from commercially available **19** (**Scheme 2**).  $^1\text{H}$  NMR showed a signal at 7.63 ppm, which was assigned as the nitrogen proton. A doublet at 4.81 ppm was assigned to the single hydrogen found on the carbon linking the two ester groups to the protected amine group. A large singlet at 1.40 ppm corresponding to 9H was the confirmatory signal indicating that Boc had been introduced to the starting material. Finally, a triplet at 1.23 ppm was observed, corresponding to the  $\text{CH}_3$  groups present.  $^{13}\text{C}$  NMR showed seven resolved signals. A signal at 166.8 ppm was assigned to the ester carbonyl carbons, whilst another downfield signal at 155.0 ppm was assigned to the Boc carbonyl carbon. The tertiary carbon of the Boc group was observed at 79.2 ppm, whilst the ester  $\text{CH}_2$  and  $\text{CH}_3$  groups were observed at 61.8 ppm and 27.9 ppm respectively. Two further signals were observed at 57.7 ppm and 13.8 ppm, assigned to the tertiary carbon of the malonate region and the Boc methyl groups respectively. IR analysis showed a weak intensity absorption at  $3359\text{ cm}^{-1}$  which was attributed to the CON-H stretch. Absorption at  $2989\text{ cm}^{-1}$  was also observed, which was ascribed to C-H stretches. A further two strong absorptions were observed at  $1710\text{ cm}^{-1}$  and  $1679\text{ cm}^{-1}$  were assigned as two carbonyl groups. This signal was also broad, meaning that it could also mask other missing signals, including the missing carbonyl group. An absorption at  $1352\text{ cm}^{-1}$  was attributed to the C-N stretch, whilst two other prominent absorptions at  $1186\text{ cm}^{-1}$  and  $1155\text{ cm}^{-1}$  were attributed to the C-O stretches of the esters. A high-resolution accurate mass of 298.1292  $[\text{M}+\text{Na}]^+$  was found. These data closely matched that seen in previous work.<sup>3</sup>



**Scheme 2.** Synthesis of 1,3-diethyl-2-[[*tert*-butoxy]carbonyl]amino}propanedioate (**20**)

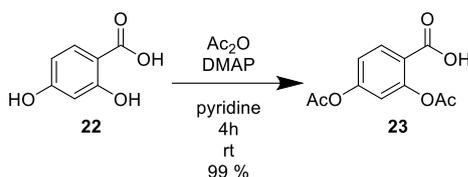
Compound **20** was then used to generate **21** by hydrolysis, which was isolated as a white solid (**Scheme 3**).  $^1\text{H}$  NMR showed a doublet at 4.74 ppm, ascribed as the single hydrogen found on the carbon linking the two carbonyl groups to the amine, whilst a multiplet between 4.19 and 4.13 ppm was once again ascribed to the two ester  $\text{CH}_2$  groups. However, the integration of this signal had changed from four to two, confirming that one ester had been hydrolysed. The Boc signal was present in the form of a singlet at 1.39 ppm, whilst the  $\text{CH}_3$  group found on the remaining ester appeared as a triplet at 1.21 ppm, once again with a reduction in integral value.  $^{13}\text{C}$  NMR further confirmed the hydrolysis of the ester, with an additional downfield signal observed at 167.8 ppm, ascribed to the acid

carbonyl carbon. Two further downfield signals, observed at 167.2 ppm and 155.1 ppm, were assigned to the ester carbonyl carbon and the Boc carbonyl carbon respectively. The tertiary carbon of the Boc group was represented by a signal at 79.1 ppm. The one remaining ester group was also observed, with a signal at 61.5 ppm ascribed to the CH<sub>2</sub> and another signal at 28.3 ppm ascribed to the CH<sub>3</sub>. The final two signals, observed at 57.7 ppm and 13.9 ppm, were assigned as the tertiary carbon of the malonate group and the methyl groups of the Boc group respectively. The hydrolysis of the ester was also confirmed by IR, with a new broad absorption at 3266 cm<sup>-1</sup> ascribed to the new O-H stretch. A second new signal, observed at 1651 cm<sup>-1</sup> was assigned as the acid C=O. These data closely matched that seen in previous work.<sup>3</sup>



*Scheme 3. Synthesis of 2-[[tert-butoxy]carbonyl]amino-3-ethoxy-3-oxopropanoic acid (**21**).*

Commercially available **22** was used to generate compound **23** (**Scheme 4**). Phenolic –OH groups as found in **22** are susceptible to deprotonation by organic bases such as pyridine, allowing nucleophilic attack from the phenoxy anion onto the acetic anhydride. Often DMAP is included in this reaction mixture for its catalytic properties. When exposed to moisture, heat and light, however, acetic anhydride will form acetic acid, and so the flask was covered to protect it from light exposure, the reaction was performed in a nitrogen atmosphere, and the acetic anhydride was added dropwise to limit the exothermic properties of the reaction.

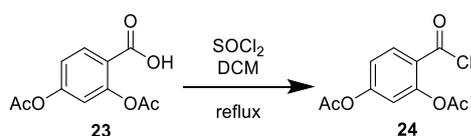


*Scheme 4. Synthesis of 2,4-bis(acetyloxy) benzoic acid (**23**)*

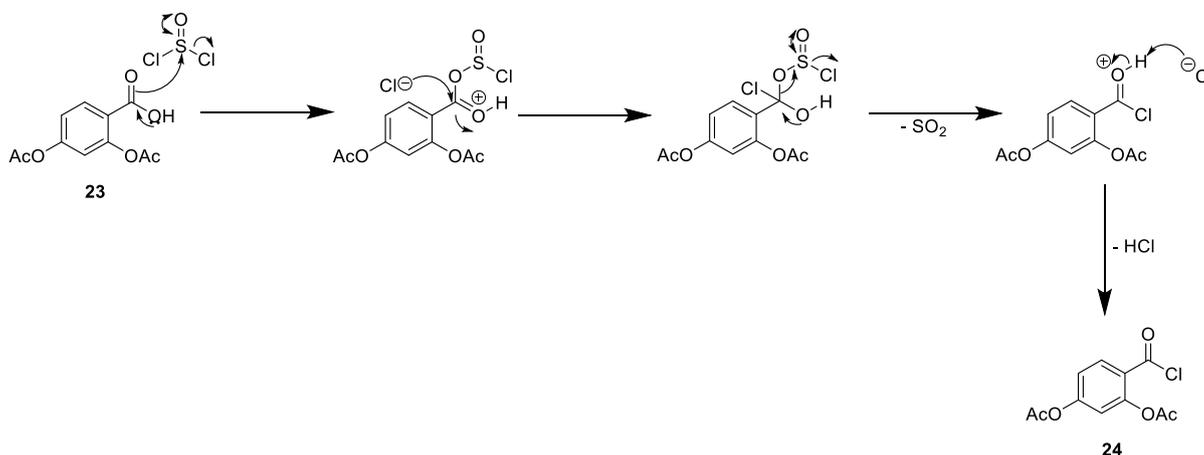
Compound **23** was isolated as a white powder. <sup>1</sup>H NMR showed a broad singlet at 13.15 ppm for the carboxylic acid hydrogen, whilst a doublet at 8.00 ppm was assigned as the aromatic hydrogen *ortho*- to the acid group. A doublet of doublets at 7.18 ppm and a further doublet at 7.10 ppm were assigned to the remaining aromatic hydrogens. Finally, two upfield signals at 2.29 ppm and 2.26 ppm were ascribed to the CH<sub>3</sub> groups found within the acetate groups. The <sup>13</sup>C NMR spectrum presented eleven resolved signals, with

three upfield signals at 169.0, 168.6 and 165.0 ppm corresponding to three carboxylic acids. The upfield signals at 20.9 and 20.8 ppm were ascribed to the two acetate groups. IR analysis showed a medium absorption at  $2593\text{ cm}^{-1}$ , which was ascribed to a C-H stretch. Two signals at  $1768\text{ cm}^{-1}$  and  $1679\text{ cm}^{-1}$  were thought to be C=O stretches, whilst absorptions at  $1607$  and  $1494\text{ cm}^{-1}$  were ascribed to aromatic C=C stretches. These data closely matched that observed in previous work.<sup>3</sup>

Subsequently, compound **23** was used to generate compound **24** by utilising thionyl chloride in anhydrous DCM (**Scheme 5**). This reaction occurs due to the electrophilic nature of the sulphur atom due to the combination of the mesomeric effect of the sulfur-oxygen double bond and the inductive electron withdrawing effect of the two chlorine atoms. It is therefore susceptible to nucleophilic attack from the carbonyl of **23**, generating an unstable tetrahedral intermediate which can then decompose to generate  $\text{SO}_2$  and hydrogen chloride (Error! Reference source not found.). Due to both side products being gases, this reaction proceeds irreversibly, making it thermodynamically favourable.



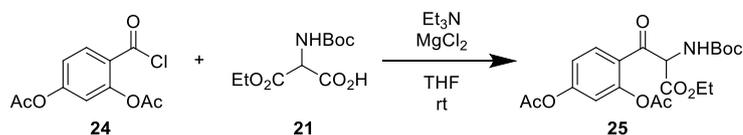
*Scheme 5. Synthesis of 2,4-bis(acetyloxy) benzoyl chloride (24).*



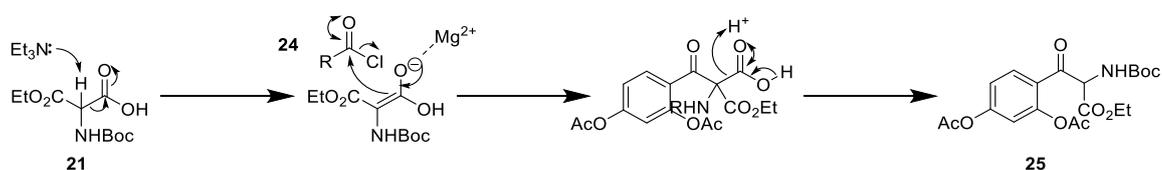
*Scheme 6. Mechanism of formation of acyl chloride (24).*

The acyl chloride (**24**) was used directly in the following reaction without any further purification, and was added to a solution of compound **21** in anhydrous THF to form compound **25** (**Scheme 7**). An acidic proton found at the 2-position of malonate **21** can be removed using a base such as triethylamine, allowing a negative charge to resonate

through the molecule. The magnesium cation was thought to localise this negative charge on the carbonyl oxygen of the carboxylic acid. Acyl chloride **24** can then be introduced onto the malonate in a 1,4-conjugate addition, with subsequent acid work-up removing the carboxylic acid as carbon dioxide and yielding **25** (Scheme 8).

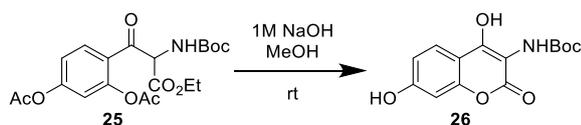


*Scheme 7. Synthesis of Ethyl-3-[2,4-bis(acetyloxy)phenyl]-2-[(tert-butoxy) carbonyl] amino}-3-oxopropanoate (25)*

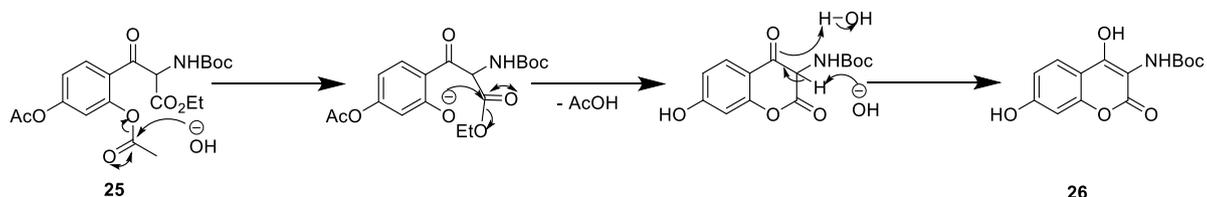


*Scheme 8. Conjugate addition of malonate 25 to acyl chloride 24*

Once again, **25** was used with no further purification for the following step to generate **26** by adding it to a solution of sodium hydroxide in methanol (Scheme 9). The strong base is able to cleave one of the acetate protecting groups of the molecule, allowing for the lactone ring to be formed by the resulting phenoxide anion intermediate. The lone pair of electrons is available to attack the nearby ester at the carbonyl group, removing the ethoxy group. In addition, the generation of the thermodynamically favourable  $\alpha,\beta$ -conjugation likely drives the protonation of the 4-position carbonyl (Scheme 10).



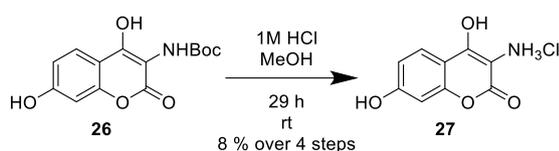
*Scheme 9. Synthesis of Tert-butyl N-(4,7-dihydroxy-2-oxo-2H-chromen-3-yl) carbamate (26)*



*Scheme 10. Deacetylation of 25 and formation of 26*

The *tert*-butoxycarbonyl protecting group was then removed from **26** using ethereal HCl, generating **27** (Scheme 11). The desired compound was yielded as a brown solid. Three

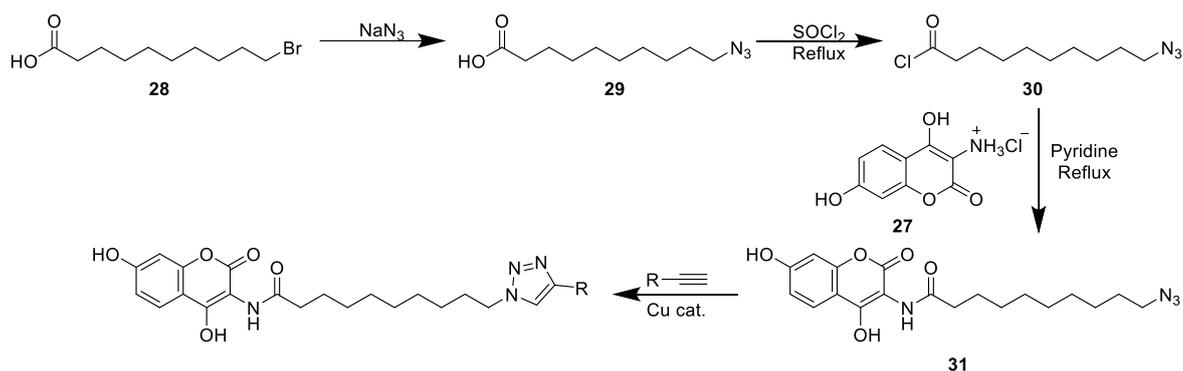
aromatic signals were observed in the  $^1\text{H}$  NMR spectrum at 7.89, 6.92 and 6.80 ppm, each corresponding to one hydrogen and appearing as a doublet, a doublet of doublets and a doublet respectively. These signals correspond to the aromatic hydrogens of the aminocoumarin ring, and the lack of any signals in the aliphatic region further indicates that the Boc group has been removed effectively. Furthermore, the  $^{13}\text{C}$  NMR of the product showed nine signals as would be expected. Signals at 126, 115 and 104 ppm were observed by HSQC experiments to be protonated, and so were assigned to be the three unsubstituted aromatic carbons. Analysis by IR exhibited two broad signals at  $3412\text{ cm}^{-1}$  and  $2847\text{ cm}^{-1}$ , indicating the presence of two O-H groups whilst also likely masking the N-H signal, as observed in previous work. A strong absorption at  $1707\text{ cm}^{-1}$  was assigned to the C=O bond, whilst a second strong signal at  $1639\text{ cm}^{-1}$  was assigned as an N-H bend. The aromatic C=C bonds were also observed in the form of two absorptions at  $1619\text{ cm}^{-1}$  and  $1530\text{ cm}^{-1}$ . These data closely matched that observed in previous work.<sup>3</sup>



*Scheme 11. Synthesis of 4,7-dihydroxy-2-oxo-2H-chromen-3-aminium chloride (27).*

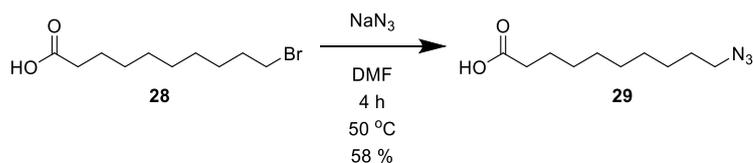
### 3.2. Synthesis of the linker

Following the synthesis of the aminocoumarin (**27**), preparation of the linker was begun. For the generation of these molecules, 10-bromodecanoic acid (**28**) would be converted to 10-azidodecanoic acid (**29**) using sodium azide. The carboxylic acid can then be reacted with thionyl chloride to form an acyl chloride (**30**) and then coupling under basic conditions with the aminocoumarin (**27**) would generate the amide (**31**). The azide will then be used to generate triazoles using click chemistry. An outline of this procedure is displayed in Error! Reference source not found..



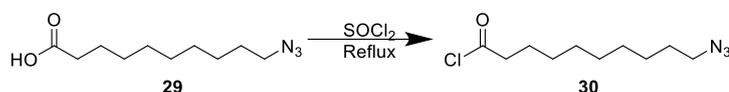
*Scheme 12. Proposed synthesis of SD8-like molecules.*

Compound **29** was prepared by treatment of commercially available **28** with sodium azide, and was isolated as a yellow oil (**Scheme 13**).  $^1\text{H}$  NMR presented a broad singlet at 9.33 ppm, identified as the carboxylic acid hydrogen. Two triplets at 3.17 and 2.24 ppm were assigned as the hydrogens on the chain closest to the azide and carboxylic acid respectively. The presence of a multiplet at 1.53 ppm corresponding to four hydrogens was ascribed to the next closest hydrogens on the chain to the two deshielding functional groups. Finally, a broad singlet-like signal was observed at 1.22 ppm. This signal corresponded to ten hydrogens, and so may correspond to the remaining hydrogens. Each of these hydrogen environments would be expected to form a triplet of triplet peak structure, which, when overlapped, may explain the presence of a broad singlet.  $^{13}\text{C}$  NMR showed ten signals. One deshielded signal at 180.3 ppm was assigned as the acid carbonyl carbon. The next most deshielded signal, found at 51.4 ppm, was ascribed to the carbon closest to the azide group. The remaining signals were assigned to the remainder of the carbon chain. IR analysis showed an absorption at  $3411\text{ cm}^{-1}$ , ascribed to the carboxylic acid O-H stretch. A strong absorption at  $2091\text{ cm}^{-1}$  was also observed, and was assigned as the characteristic azide  $\text{N}=\text{N}=\text{N}$  stretch. Finally, a stronger absorption corresponding to a  $\text{C}=\text{O}$  stretch was observed at  $1706\text{ cm}^{-1}$ . A high-resolution accurate mass of  $235.1302\text{ [M+Na]}^+$  was observed.



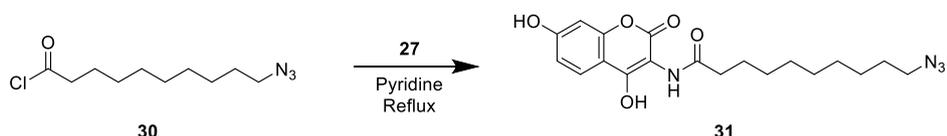
*Scheme 13. Synthesis of 10-azidodecanoic acid (29).*

Compound **29** was then used directly in the synthesis of **30** (**Scheme 14**). As for the synthesis of **24**, this was first attempted with thionyl chloride under anhydrous conditions, and was produced as an orange oil, which was used without further purification in the following steps.

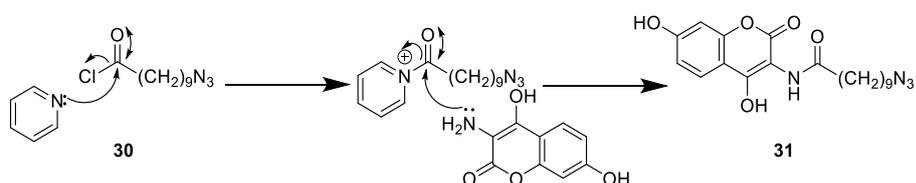


*Scheme 14. Synthesis of 10-azidodecanoyl chloride (30).*

Compound **30** was used to attempt to generate **31** (**Scheme 15**). Pyridine is known to be a good nucleophile for carbonyls and acts as a catalyst in this proposed reaction by nucleophilic attack of the carbonyl position, resulting in the loss of the chloride group. This nucleophilicity on the nitrogen of the pyridine is caused by the lone pairs on the nitrogen being unable to delocalise around the aromatic system of the ring. The amine of the aminocoumarin was then thought to be able to also react at the carbonyl, causing the loss of the charged pyridinium (**Scheme 16**). This methodology has been used in previous work within the group to attach aminocoumarins to sebacic acid linkers.<sup>212</sup>



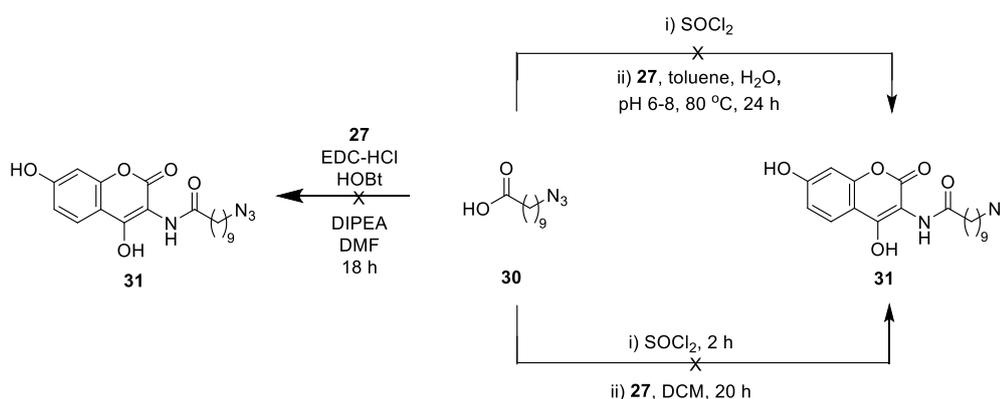
*Scheme 15. Attempted synthesis of 10-azido-N-(4,7-dihydroxy-2-oxo-2H-chromen-3-yl)decanamide (31)*



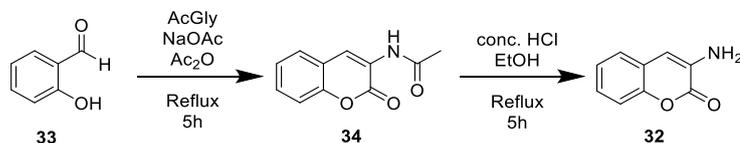
*Scheme 16. Proposed pyridine-catalysed formation of 31*

However, the reaction failed to proceed, instead yielding only starting materials **27** and **29**. Other potential routes were also explored, as highlighted in **Scheme 17**. Schotten-Baumann conditions were first attempted. These conditions involve the stirring of an organic solvent, in this case toluene, containing the starting materials, and basic water with pH correction to keep the aqueous phase between pH 6 and pH 8. These conditions have the advantage of ensuring that any acid formed does not slow or halt the reaction, and is instead

neutralised by the constant supply of base, in this case potassium carbonate. However, attempts to use these conditions returned only the starting materials, even upon heating to 80 °C. To rule out the possibility that the acid chloride (**30**) was either degrading too quickly to be used or was not being generated at all, coupling using EDC was also attempted. This coupling agent activates the acid group to nucleophilic attack by the amine nitrogen lone pair, resulting in the formation of the desired amide and a water-soluble urea by-product, allowing for easy separation from the crude product. However, this reaction once again returned the starting materials with no observed production of compound **31**. It is possible that the hydroxyl groups of the aminocoumarin deactivate the amine, reducing its nucleophilic ability sufficiently to prevent removal of the pyridinium intermediate. The decision was made therefore to simplify the aminocoumarin further. Previous work has found that 3-aminocoumarin (**32**) binds to DNA gyrase in a similar fashion to the SD8 aminocoumarin and has shown potency in work to generate aminocoumarin-fluoroquinolone hybrid inhibitors.<sup>212</sup> The two-step synthesis of **32** is shown below in **Scheme 18**.



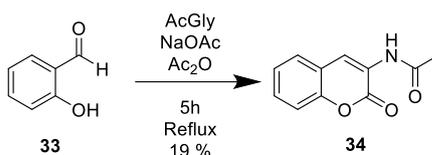
**Scheme 17.** Attempted synthetic routes to **31**



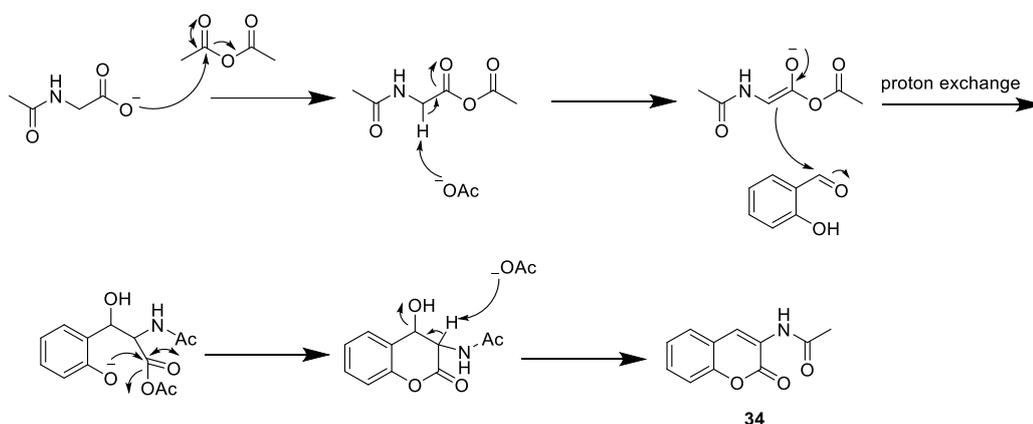
**Scheme 18.** Complete synthesis pathway of 3-aminocoumarin (**32**).

The first step of the reaction utilizes a modified Perkin reaction (**Scheme 19**). The Perkin reaction, discovered in 1868 by William Henry Perkin, allows for the generation of coumarins by condensation of an aromatic ortho-hydroxylaldehyde in the presence of a mild base with an anhydride.<sup>227,228</sup> The precise mechanism of this reaction remains unknown, but is thought to involve enolate formation on the anhydride by basic

deprotonation, followed by attack of this enolate onto the aldehyde. The resulting mixed anhydride can then be intramolecularly attacked by the hydroxyl group allowing for ring closure of the lactone ring and removal of the water and acetyl groups to form the coumarin (**Scheme 20**).<sup>229,230</sup> The modification in this case is the addition of N-acetylglycine, which is thought to react with the acetic anhydride before forming an enolate which is then able to react with the aldehyde in a similar fashion to the unmodified reaction.<sup>231</sup>



*Scheme 19. Synthesis of N-acetyl-3-aminocoumarin (34)*

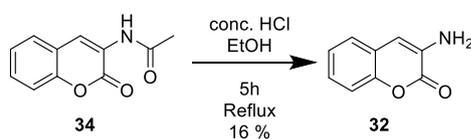


*Scheme 20. Mechanism of the modified Perkin reaction to form 34*

Compound **34** was isolated as a yellow solid.  $^1\text{H}$  NMR showed a singlet at 9.76 ppm, ascribed to the amide hydrogen. A further singlet was observed at 8.61 ppm, and was assigned to the hydrogen present in the lactone ring. Four aromatic signals were observed between 7.70 ppm and 7.36 ppm. A doublet of doublets at 7.70 ppm and a multiplet at 7.50 ppm were ascribed to the two closest aromatic hydrogens to the lactone ring, whilst a doublet at 7.40 ppm and a triplet at 7.36 ppm were ascribed to the furthest aromatic hydrogens from the lactone. A final singlet at 2.17 ppm was assigned to the methyl present on the acetyl group.  $^{13}\text{C}$  NMR showed eleven resolved signals. A deshielded signal at 170.7 ppm was assigned to the acetyl carbonyl, whilst another deshielded signal at 157.9 ppm was ascribed to the carbonyl carbon present on the lactone ring. Two signals, observed at 150.1 ppm and 120.0 ppm, were assigned to the two ring-joining carbons, whilst four signals at 130.0, 128.3, 125.4 and 116.2 ppm were ascribed to the remaining aromatic

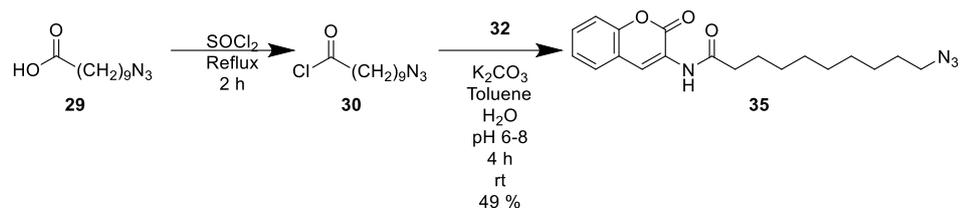
carbons. A further signal at 125.0 ppm was assigned as the lactone ring carbon attached to the amide, with another signal at 124.0 ppm ascribed to the remaining lactone carbon. A final signal at 24.4 ppm was assigned as the methyl group on the acetyl group. A sharp, strong signal at  $3329\text{ cm}^{-1}$  in the IR spectrum was ascribed to the amide N-H stretch, whilst two further strong signals at  $1708\text{ cm}^{-1}$  and  $1680\text{ cm}^{-1}$  were assigned to the two C=O environments. A final two strong signals at  $1604\text{ cm}^{-1}$  and  $1516\text{ cm}^{-1}$  were assigned to the aromatic C=C stretches. These data closely match that observed in previous work.<sup>212</sup>

**34** was then used to generate 3-aminocoumarin (**32**) by removal of the acetate group using hydrochloric acid (**Scheme 21**). The deprotected aminocoumarin was isolated as a pale yellow solid. No upfield signals were observed within the  $^1\text{H}$  NMR spectrum, suggesting that the acetate group had been completely removed. Three aromatic signals were observed between 7.46 ppm and 7.26 ppm. Multiplets between 7.46-7.40 ppm and 7.31-7.26 ppm were ascribed to the hydrogens closest to the lactone ring, whilst the remaining multiplet at 7.26-7.18 ppm were assigned to the two hydrogens furthest from the lactone ring. Two singlets at 6.72 ppm and 5.70 ppm were also observed, and were assigned to the lactone ring hydrogen and the amine hydrogens respectively.  $^{13}\text{C}$  NMR showed nine resolved signals. A signal at 159.1 ppm was assigned to the carbonyl carbon present in the lactone ring, whilst a signal at 148.4 ppm was ascribed to the amine carbon in the lactone ring. Two signals at 133.6 ppm and 115.9 ppm were assigned as the ring-joining carbons, and a further signal, observed at 122.2 ppm, was ascribed to the remaining unsaturated lactone carbon. The remaining signals, observed at 125.9, 125.3, 125.0 and 108.5 ppm, were assigned to the aromatic carbons. The IR spectrum showed a strong signal at  $1703\text{ cm}^{-1}$ , ascribed to the C=O stretch, and a further strong signal at  $1637\text{ cm}^{-1}$ , ascribed to the N-H stretch. A final signal at  $1588\text{ cm}^{-1}$  was assigned to the aromatic C=C stretches. These data closely match that observed in previous work.<sup>212</sup>



*Scheme 21. Synthesis of 3-aminocoumarin (32)*

Completed aminocoumarin **32** was then attached to linker **29**. As before, the acid chloride was generated (**30**), followed by coupling of aminocoumarin **32** using Schotten-Baumann conditions to generate aminocoumarin-linker compound **35** (**Scheme 22**).



*Scheme 22. Synthesis of 10-azido-N-(2-oxo-2H-chromen-3-yl)decanamide (35)*

Compound **35** was generated as an off-white solid.  $^1\text{H}$  NMR showed a deshielded singlet at 9.69 ppm, ascribed to the amide hydrogen. Two signals, observed at 3.31 ppm and 2.49 ppm, were masked by  $\text{H}_2\text{O}$  and DMSO respectively, but were observed by 2D NMR to correspond to the hydrogens closest to the azide and amide on the decane chain respectively. Two further multiplets were observed, and were assigned to the remaining hydrogens present in the decane chain.  $^{13}\text{C}$  NMR showed nineteen resolved signals. A deshielded signal at 173.7 ppm was ascribed to the amide carbonyl carbon, whilst another deshielded signal observed at 158.0 ppm was assigned as the carbonyl carbon within the lactone ring. A final downfield signal was observed at 150.1 ppm, and was ascribed to the lactone carbon attached to the amide nitrogen. Six aromatic signals were observed. The two most deshielded aliphatic signals, observed at 51.1 ppm and 36.4 ppm, were ascribed to the carbons closest to the azide and amide respectively, with the remaining signals being assigned to the remaining carbons in the decane chain. The IR spectrum showed a strong signal at  $3330\text{ cm}^{-1}$ , ascribed to the amide N-H stretch. A further strong signal at  $2091\text{ cm}^{-1}$  was assigned as the azide N=N=N stretch. Two signals at  $1709\text{ cm}^{-1}$  and  $1678\text{ cm}^{-1}$  were ascribed to the two C=O stretches present. Finally, two strong signals at  $1630\text{ cm}^{-1}$  and  $1525\text{ cm}^{-1}$  were assigned to be the aromatic C=C stretches. A high-resolution accurate mass of  $355.1755\text{ [M]}^+$  was observed.

### 3.3. Generation of angucyclinone-replacing alkynes

Upon completion of the linker, the purchase and preparation of the best-docking alkynes as implicated by the computational work described in **Chapter 2** was begun. As stated previously, a python script was used to extract the top 30 results of the Glide docking obtained from unique library members, and from these 5 members were taken forward. These members were both highly scored in the docking process, and could be constructed relatively quickly over the course of two or three steps, and at a low price, and are shown below in **Figure 55**.

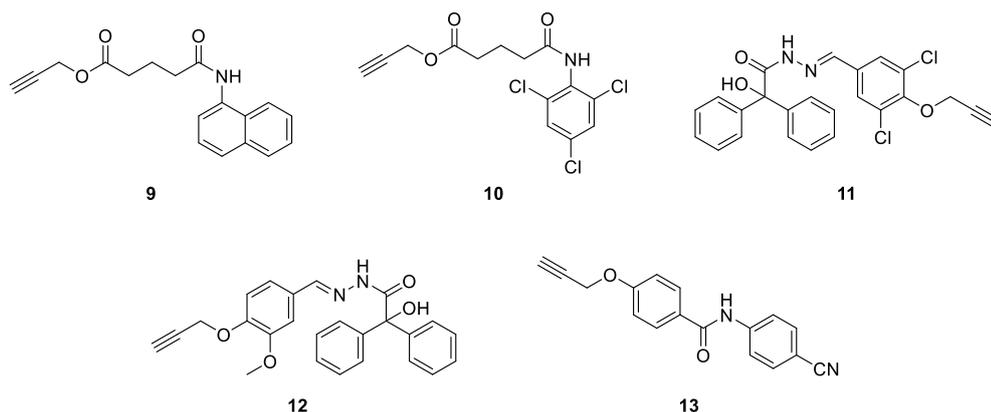
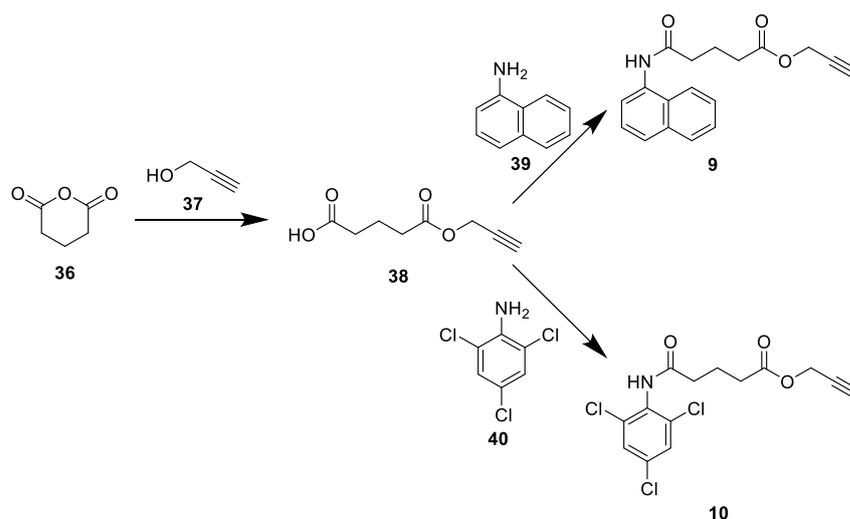


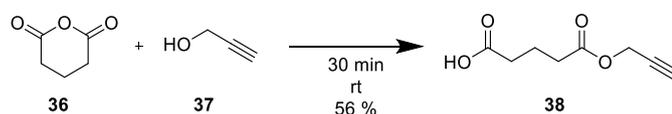
Figure 55. Selected library members 9-13

### 3.3.1. Generation of alkynes 9 and 10

Compounds **9** and **10** shared many structural features, and as such the syntheses of these two compounds were thought to be similar. The first step involved the ring opening of commercially available glutaric anhydride (**36**) using propargyl alcohol (**37**) to generate a carboxylic acid (**38**), before the relevant amine (**39** or **40**) was attached using amide coupling chemistry, as shown in **Scheme 23**. Compound **38** was therefore first generated as a pale yellow oil (**Scheme 24**).<sup>232</sup>  $^1\text{H}$  NMR showed a signal at 9.30 ppm, identified as the carboxylic acid hydrogen. A doublet at 4.65 ppm was ascribed as the propargyl  $\text{CH}_2$  group, whilst a triplet at 2.47 ppm with integral 1 was assigned to the alkyne CH hydrogen. A multiplet corresponding to four hydrogens was also observed, and assigned to the two  $\text{CH}_2$  groups of the main pentanoic acid region closest to the carbonyls. Finally, a quintet at 1.93 ppm was assigned as the central  $\text{CH}_2$  group of the pentanoic acid. The  $^{13}\text{C}$  NMR spectrum consisted of eight signals. Two upfield signals, at 178.9 and 172.2 ppm, were assigned as the carboxylic acid and the ester carbonyl carbons respectively. A signal at 76.8 ppm was assigned as the terminal alkyne carbon, whilst another signal at 75.0 ppm was determined to correspond to the alkyne carbon closest to the pentanoic acid motif. A further signal at 52.0 ppm was observed and was assigned to the remaining propargyl alcohol carbon. These data closely resembled that observed previously in literature.<sup>232</sup> Analysis by IR spectroscopy showed a broad signal at  $3290\text{ cm}^{-1}$ , assigned to the acidic O-H stretch. This signal potentially overlaps with the alkyne C-H stretch, with the presence of the alkyne indicated by a weak signal at  $2129\text{ cm}^{-1}$  in the spectrum. Finally, two signals at  $1733\text{ cm}^{-1}$  and  $1705\text{ cm}^{-1}$  were assigned as the ester and acid C=O stretches respectively. A high-resolution accurate mass of  $171.0641\text{ [M]}^+$  was observed.

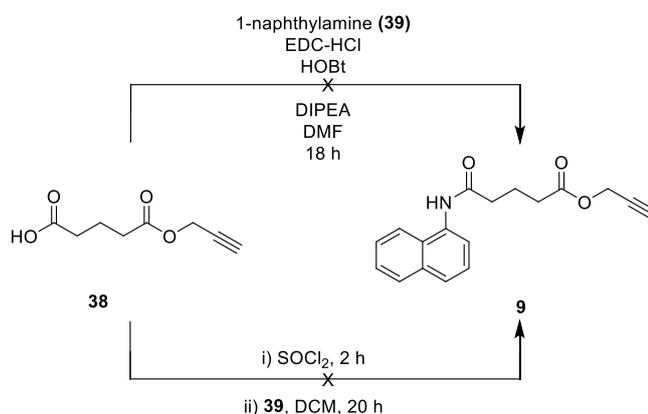


**Scheme 23.** General route to alkynes **9** and **10**

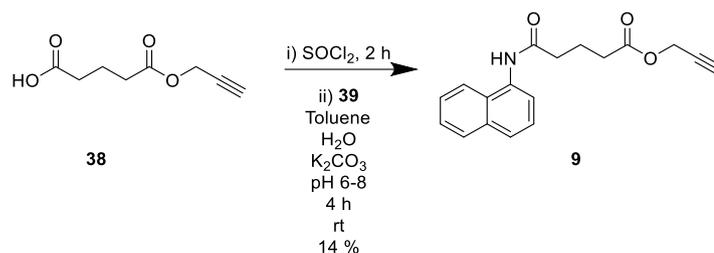


**Scheme 24.** Synthesis of 5-oxo-5-(prop-2-yn-1-yloxy)pentanoic acid (**38**)

The second step involved the formation of the amide bond between **38** and 1-naphthylamine (**39**). Initial attempts focused on utilising EDC chemistry, with subsequent attempts made to generate the acid chloride before stirring this with the naphthylamine (**Scheme 25**). However, both of these methods failed to generate **9**, instead returning the naphthylamine and acid upon workup. The bi-phasic Schotten-Baumann conditions described previously were then utilised, with compound **9** being successfully generated using this methodology (**Scheme 26**).



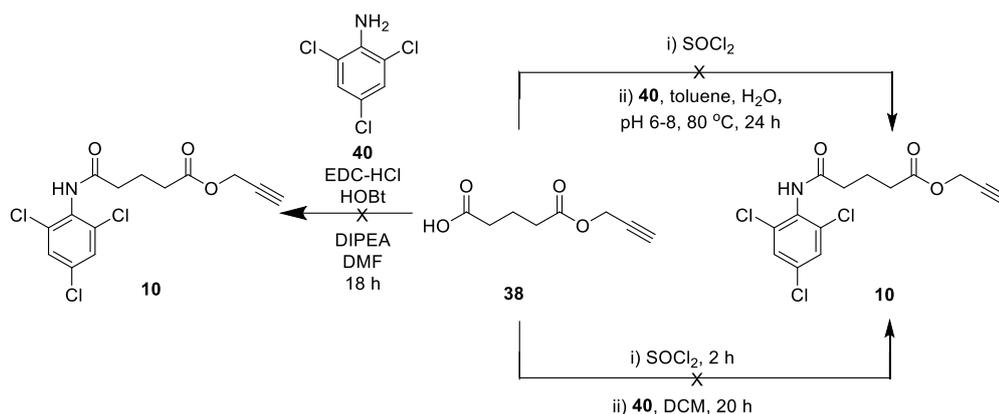
**Scheme 25.** Attempted synthetic routes of prop-2-yn-1-yl 4-[(naphthalen-1-yl)carbamoyl]butanoate (**9**)



*Scheme 26. Synthesis of prop-2-yn-1-yl-4-[(naphthalen-1-yl)carbamoyl]butanoate (9)*

Compound **9** was isolated as a pink solid.  $^1\text{H}$  NMR showed three doublets at 7.99, 7.91 and 7.79 ppm, as well as a multiplet corresponding to four hydrogens at 7.59-7.46 ppm. These signals were ascribed to the aromatic hydrogens present in the naphthalene system. A doublet at 4.73 ppm was assigned to the  $\text{CH}_2$  group present in the propargyl group, with a triplet at 2.93 ppm ascribed to the terminal alkyne hydrogen. Two triplets at 2.65 ppm and 2.56 ppm were assigned as the  $\text{CH}_2$  groups closest to the amide and ester groups respectively. A final quintet at 2.13 ppm was observed, and was ascribed to the central  $\text{CH}_2$  group.  $^{13}\text{C}$  NMR showed eighteen resolved signals. Two downfield signals at 174.8 ppm and 173.8 ppm were assigned as the carbonyl carbons of the amide and ester groups respectively. Ten aromatic carbons were observed, consistent with the naphthalene ring present in the structure. Two signals present at 78.9 ppm and 76.2 ppm were ascribed to the two carbons forming the alkyne group, whilst another signal at 52.8 ppm was assigned as the propargyl carbon closest to the oxygen. The IR spectrum showed a sharp signal at  $3247\text{ cm}^{-1}$ , assigned as the terminal alkyne C-H stretch. A further, weaker signal at  $2102\text{ cm}^{-1}$  was ascribed to the alkyne triple bond stretch. The ester C=O bond was observed at  $1729\text{ cm}^{-1}$ , with the acid C=O stretch shifting to  $1648\text{ cm}^{-1}$  following the functional group conversion to the amide. A high-resolution accurate mass of  $296.1340\text{ [M]}^+$  was observed.

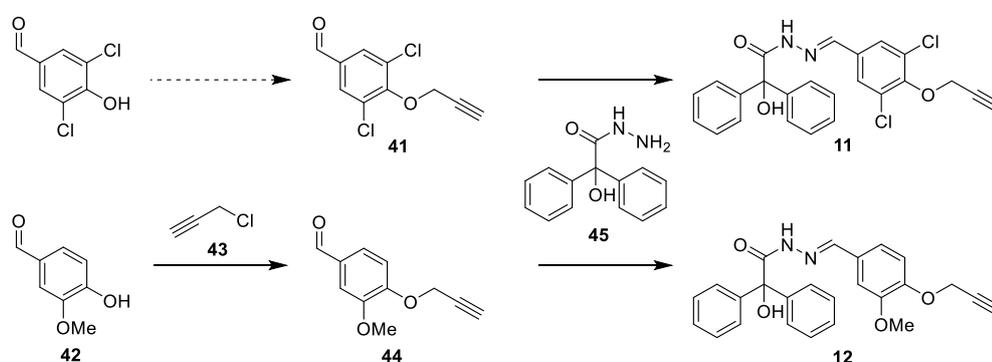
Attempts at generating alkyne **10**, however, were less successful. Multiple different methodologies for generating the amide bond between **38** and commercially available 2,4,6-trichloroaniline (**40**) resulted in the return of the starting materials only. These methodologies are summarised below in **Scheme 27**. It was therefore theorised that the three chloride groups decorating the aniline aromatic ring were deactivating the nucleophilic properties of the amine, as well as potentially sterically hindering the amine, resulting in a limited ability of the aniline to react with the acid/acid chloride. This alkyne was therefore abandoned.



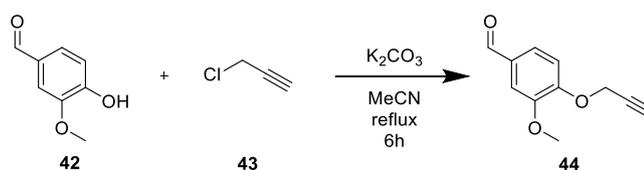
**Scheme 27.** Attempted syntheses of prop-2-yn-1-yl 4-[(2,4,6-trichlorophenyl)carbamoyl]butanoate (**10**)

### 3.3.2. Generation of alkynes **11** and **12**

Alkynes **11** and **12** were also generated using similar steps. The first step involved addition of a propargyl group to 4-hydroxybenzaldehydes, followed by the formation of the relevant hydrazones using benzylic acid hydrazide, as highlighted in **Scheme 28**. However, the propargylbenzaldehyde required for formation of **11** was found to be commercially available (**41**), and so the first step was skipped for the formation of this alkyne. Vanillin (**42**) was used to generate 3-methoxy-4-(prop-2-yn-1-yloxy)benzaldehyde (**44**) through reaction with propargyl chloride (**43**), as shown in **Scheme 29**.



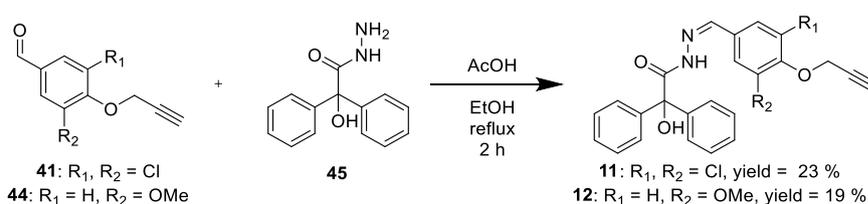
**Scheme 28.** General route for the generation of alkynes **11** and **12**



**Scheme 29.** Synthesis of 3-methoxy-4-(prop-2-yn-1-yloxy)benzaldehyde (**44**).

Compound **44** was isolated as a yellow solid.  $^1\text{H}$  NMR showed the presence of a singlet at 9.82 ppm, assigned as the aldehyde hydrogen. Three signals at 7.52 ppm, 7.46 ppm and 7.21 ppm were assigned as the three aromatic hydrogens. A doublet at 4.88 ppm suggested the presence of the propargyl  $\text{CH}_2$  group, whilst a triplet at 3.03 ppm was assigned as the alkyne  $\text{CH}$  group. Finally, a singlet at 3.89 ppm was assigned as the methoxy  $\text{CH}_3$  group.  $^{13}\text{C}$  NMR showed eleven signals. One upfield signal at 192.9 ppm was assigned as the aldehyde carbon, whilst six signals between 153.9 ppm and 111.2 ppm were determined to correspond to the aromatic carbons. Two further signals, found at 79.0 ppm and 77.7 ppm were assigned as the alkyne carbons. Finally, two downfield signals at 57.5 ppm and 56.5 ppm were assigned as the  $\text{CH}_2$  and  $\text{CH}_3$  groups respectively. Analysis by IR spectroscopy showed a signal at  $3248\text{ cm}^{-1}$ , ascribed to the terminal alkyne C-H stretch. The alkyne's presence was further confirmed by the presence of a weak signal at  $2110\text{ cm}^{-1}$ , corresponding to the alkyne carbon-carbon triple bond. A signal at  $1668\text{ cm}^{-1}$  was assigned to the aldehyde  $\text{C}=\text{O}$  stretch, whilst two strong signals at  $1587\text{ cm}^{-1}$  and  $1506\text{ cm}^{-1}$  were ascribed to the aromatic  $\text{C}=\text{C}$  stretches. Finally, a further two strong signals at  $1261\text{ cm}^{-1}$  and  $1243\text{ cm}^{-1}$  were assigned to the two ether C-O stretches. A high-resolution accurate mass of  $191.0633\text{ [M]}^+$  was observed.

Compounds **41** and **44** were then used to generate alkynes **11** and **12** respectively by reaction with benzylic acid hydrazide **45** in acidic conditions as shown in **Scheme 30**.



*Scheme 30. Synthesis of alkyne library members 11 and 12*

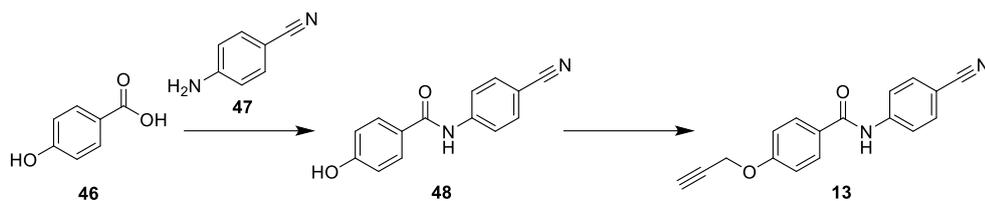
Compound **11** was isolated as a white solid.  $^1\text{H}$  NMR showed a singlet at 8.27 ppm which was ascribed to the hydrazone hydrogen, whilst a further singlet at 7.87 ppm corresponding to two hydrogens was assigned as the two aromatic hydrogens present on the dichlorinated ring. A doublet of doublets at 7.52 ppm and a multiplet at 7.38 ppm were ascribed to the ten hydrogens present in the benzylic acid ring system. A doublet at 4.86 ppm was assigned to the  $\text{CH}_2$  group of the propargyl region, whilst a triplet at 3.00 ppm was ascribed to the terminal alkyne hydrogen.  $^{13}\text{C}$  NMR showed fourteen signals. A downfield signal at 171.9 ppm was assigned to the carbonyl carbon. Eight aromatic signals were observed. A signal at 146.4 ppm was observed and assigned as the hydrazone carbon,

whilst another signal at 80.9 ppm was ascribed to the tertiary carbon within the benzylic acid group. Two signals at 77.2 ppm and 76.6 ppm were assigned to the alkyne carbons, whilst a final signal at 60.2 ppm was ascribed to the final carbon of the propargyl group. The IR spectrum showed a broad signal at  $3300\text{ cm}^{-1}$ , indicative of an acidic O-H stretch. A weak signal at  $2100\text{ cm}^{-1}$  was also observed, and was assigned as the characteristic  $\text{C}\equiv\text{C}$  stretch. Strong signals at  $1665\text{ cm}^{-1}$  and  $1656\text{ cm}^{-1}$  were ascribed to C=O and C=N stretches respectively, whilst a final strong signal at  $803\text{ cm}^{-1}$  was assigned as a C-Cl stretch. A high-resolution accurate mass of  $453.0773\text{ [M+H]}^+$  was obtained.

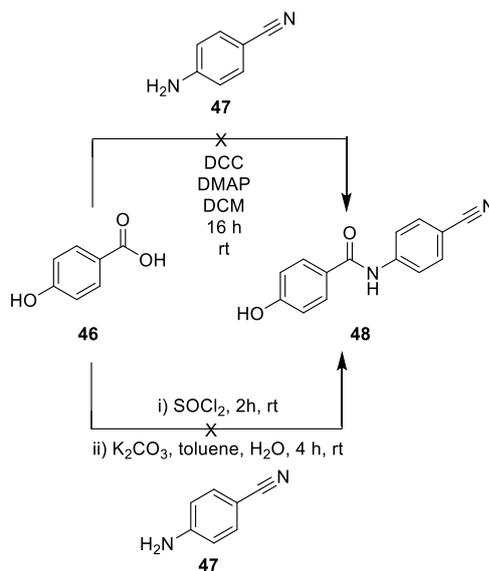
Compound **12** was isolated as an off-white solid.  $^1\text{H}$  NMR showed a singlet at 8.26 ppm which was ascribed to the hydrazone hydrogen. Six aromatic hydrogen signals were also observed. Two multiplets at 7.44 ppm and 7.36-7.26 ppm were assigned as the aromatic hydrogens present in the benzylic acid group. Finally, two signals were observed at 3.91 ppm and 2.53 ppm, corresponding to the  $\text{CH}_2$  group and the terminal alkyne hydrogen respectively.  $^{13}\text{C}$  NMR showed seventeen resolved signals. A downfield signal at 165.5 ppm was ascribed to the carbonyl carbon. Ten aromatic carbon signals were observed. A signal at 149.2 ppm was ascribed to the hydrazone carbon, whilst another signal at 80.5 ppm was assigned to the tertiary carbon within the benzylic acid group. Two signals at 78.0 ppm and 75.8 ppm were ascribed to the two carbons of the alkyne group, with two further signals at 56.1 ppm and 55.0 ppm corresponding to the  $\text{CH}_2$  of the propargyloxy group and the  $\text{CH}_3$  of the methoxy group respectively. A sharp, weak signal at  $3295\text{ cm}^{-1}$  on the IR spectrum was assigned to the terminal alkyne C-H stretch, whilst a broad signal at  $3100\text{ cm}^{-1}$  was ascribed to the acidic O-H stretch. Two strong signals at  $1658\text{ cm}^{-1}$  and  $1635\text{ cm}^{-1}$  were assigned as the C=O stretch and C=N stretches. A final two signals at  $1603\text{ cm}^{-1}$  and  $1509\text{ cm}^{-1}$  were ascribed to the aromatic C=C bonds. A high-resolution accurate mass of  $415.1620\text{ [M]}^+$  was obtained.

### 3.3.3. Generation of alkyne 13

Initial approaches to the synthesis of alkyne **13** began by attempting to couple 4-hydroxybenzoic acid (**46**) to 4-aminobenzonitrile (**47**), with the propargyl group to be added to the intermediate (**48**) following the formation of the amide as shown in **Scheme 31**. However, generation of amide **48** proved problematic, with all attempts to synthesise this molecule mostly returning the starting materials (**Scheme 32**). This was assumed to be due to the electron donating effect of the alcohol of **46** on the carboxylic acid, stabilising the acid group and preventing amide formation.

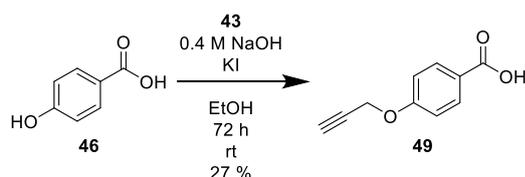


**Scheme 31.** General route for generation of alkyne **13**



**Scheme 32.** Attempted syntheses of intermediate amide **48**

To mitigate this, the propargyl group was added to the hydroxyl group of **46** first to form compound **49**, once again using compound **43** (Scheme 33).

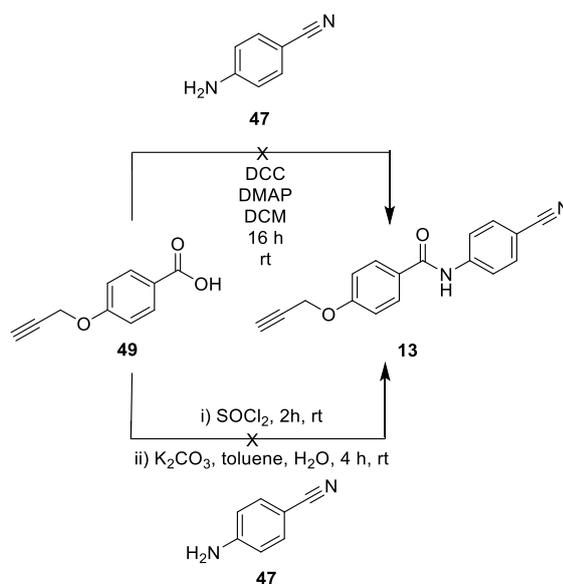


**Scheme 33.** Synthesis of 4-(prop-2-yn-1-yloxy)benzoic acid (**49**)

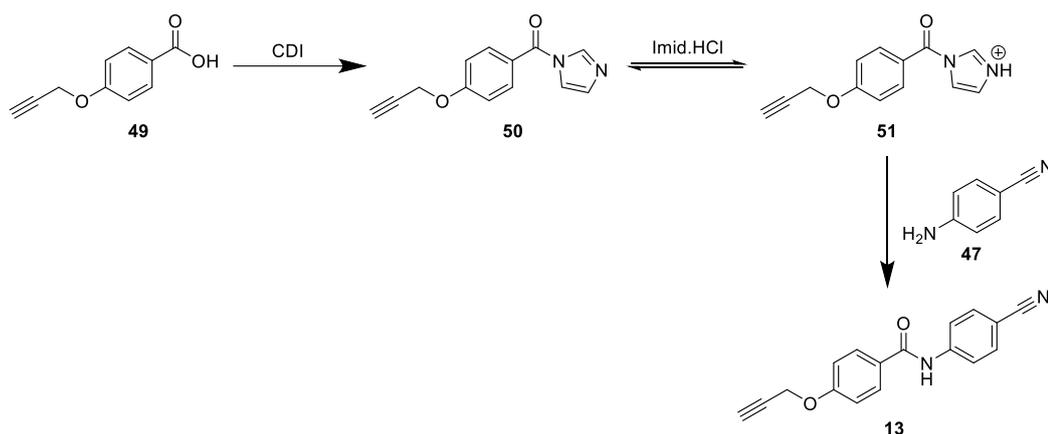
Compound **49** was isolated as a white solid.  $^1\text{H}$  NMR showed a doublet at 7.90 ppm corresponding to the two aromatic hydrogens adjacent to the acid group, whilst another doublet at 6.88 ppm corresponded to the two aromatic hydrogens adjacent to the propargyl group. A further signal at 4.90 masked by  $\text{H}_2\text{O}$  was ascribed to the  $\text{CH}_2$  of the propargyl group, and a triplet at 2.96 ppm was assigned to the terminal alkyne group.  $^{13}\text{C}$  NMR showed eight resolved signals. A downfield signal at 167.2 ppm was ascribed to the acid carbonyl group, whilst two further signals at 79.1 ppm and 76.1 ppm were ascribed to the two alkyne carbons. A final signal, observed at 52.9 ppm, were assigned to the  $\text{CH}_2$  carbon of the propargyl group. The IR spectrum showed a strong, sharp signal at  $3270\text{ cm}^{-1}$

corresponding to the terminal alkyne C-H stretch. A broad signal centred around  $3000\text{ cm}^{-1}$  was ascribed to the carboxylic acid OH stretch. A weak signal at  $2135\text{ cm}^{-1}$  was assigned as the alkyne carbon-carbon triple bond, whilst a strong signal at  $1674\text{ cm}^{-1}$  was assigned as the C=O stretch of the acid. Finally, two signals at  $1605\text{ cm}^{-1}$  and  $1485\text{ cm}^{-1}$  were assigned as the aromatic C=C stretches.

Following the generation of **49**, multiple attempts to synthesise alkyne **13** began (**Scheme 34**). However, this step once again proved to be problematic, with starting materials **47** and **49** again being returned as the only isolated products. The decision was made, therefore, to switch from base-catalysed amide formation to acid-catalysed amide formation. Carbonyl diimidazole (CDI) was used to generate an activated acyl-imidazole intermediate (**50**), with imidazole hydrochloride acting to improve the rate of the reaction by generating the protonated intermediate (**51**) (**Scheme 35**). This “imidazolium effect” has been shown to improve amide formation between aromatic acids and amines in previous work.<sup>233</sup>



**Scheme 34.** Attempted syntheses of amide **13**



*Scheme 35. Method of CDI and imidazolium effect activation of 49 to help form amide 13*

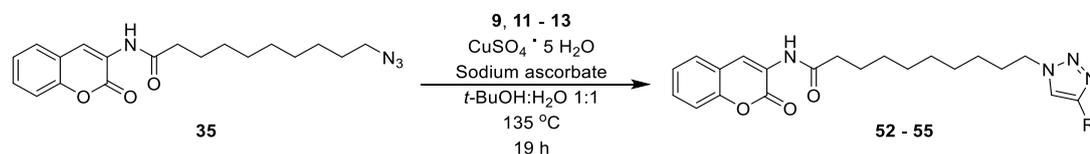
Compound **13** was isolated as a brown solid (**Scheme 36**).  $^1\text{H}$  NMR showed three aromatic signals corresponding to eight hydrogens, confirming the presence of both aromatic rings within the structure. A singlet at 4.82 ppm was ascribed to the  $\text{CH}_2$  group of the propargyl group, with a final signal at 3.31 masked by methanol assigned as the terminal alkyne hydrogen.  $^{13}\text{C}$  NMR showed thirteen resolved signals. A downfield signal at 167.8 ppm was ascribed to the carbonyl carbon of the amide, whilst another downfield signal at 162.0 ppm was assigned to the carbon adjacent to this. A signal at 143.3 ppm corresponded to the nitrile carbon. Two signals at 77.1 ppm and 75.3 ppm corresponded to the two carbons of the alkyne group, with a final signal at 55.4 ppm being assigned to the  $\text{CH}_2$  carbon of the propargyl group. The IR spectrum once again showed a strong, sharp signal at  $3245\text{ cm}^{-1}$ , ascribed to the terminal alkyne C-H stretch. A weak signal at  $2228\text{ cm}^{-1}$  was assigned as the nitrile  $\text{C}\equiv\text{N}$  stretch, whilst a strong, sharp signal at  $1646\text{ cm}^{-1}$  was ascribed to the amide  $\text{C}=\text{O}$  stretch. Finally, two signals at  $1592\text{ cm}^{-1}$  and  $1506\text{ cm}^{-1}$  were assigned to the aromatic  $\text{C}=\text{C}$  stretches. A high-resolution accurate mass of  $277.0935\text{ [M]}^+$  was obtained.



*Scheme 36. Synthesis of N-(4-cyanophenyl)-4-(prop-2-yn-1-yloxy)benzamide (13)*

### 3.4. Assembly of aminocoumarin-linker-alkyne compounds

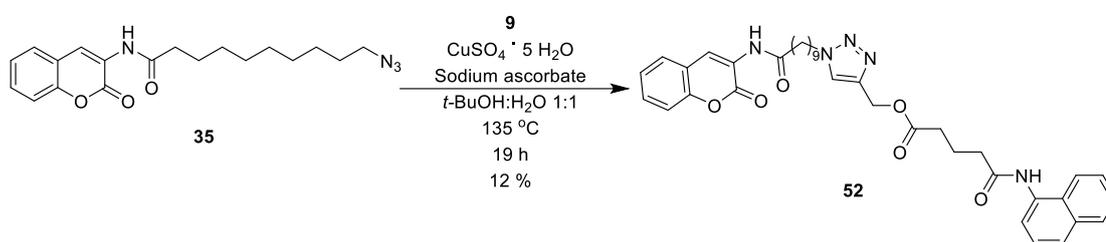
The final stages of synthesis involved connection of alkynes **9-13**, with the exception of compound **10** which was excluded from further study, to the aminocoumarin-coupled linker (**35**) to form compounds **52-55**. This was achieved using copper-catalysed azide-alkyne cycloadditions (CuAAC), possibly the most commonly used example of “Click chemistry”. Click chemistry was first described by Sharpless et al. In 2001, defined as “a set of powerful, highly reliable and selective reactions”, and further described as “spring-loaded” reactions.<sup>234</sup> Previous work within the group has utilised CuAACs to form amide-mimicking linkages between two small molecule fragments, and, in a similar way to this, the triazole was to be used to form a olivose-replacing linkage between the linker and the library members.<sup>203</sup> The conditions of this reaction (i.e. temperature, solvents, catalyst) vary across the reported literature, and so multiple different condition combinations were first attempted to find the optimal conditions for compounds **9-13**. From these tests, it was decided to use a 1:1 combination of *tert*-butanol and water as the solvent for the reactions, and copper(II) sulfate as the catalyst, with sodium ascorbate also being added to assist in the reduction of the copper(II) ions to the required reactive copper(I) species. The conditions are summarised in **Scheme 37**.



**Scheme 37.** Conditions used for the CuAAC reactions of **35** with alkynes **9, 11-13**

Compound **52** was generated as an off-white solid (**Scheme 38**). Two broad singlets were observed in the <sup>1</sup>H NMR spectrum at 9.89 ppm and 9.66 ppm, which were assigned to the aminocoumarin amide NH and the naphthylamide NH respectively. A singlet at 8.64 ppm was ascribed to the lone hydrogen of the lactone ring, whilst a further singlet at 8.15 ppm was assigned as the hydrogen of the triazole ring, indicating that the ring had been formed successfully. Aside from these signals, five further aromatic signals were observed, corresponding to the expected thirteen aromatic hydrogens. A singlet at 5.17 ppm was ascribed to the CH<sub>2</sub> group linking the triazole to the glutaric acid, with six further aliphatic hydrogen signals being assigned to the linker hydrogens and the glutaric acid chain hydrogens. <sup>13</sup>C NMR showed thirty five resolved signals. Three downfield signals at 173.6, 172.9 and 171.9 ppm were assigned as the carbonyls of the linker amide, glutaric ester and glutaric amide respectively. A further downfield signal at 150.1 ppm was ascribed to the

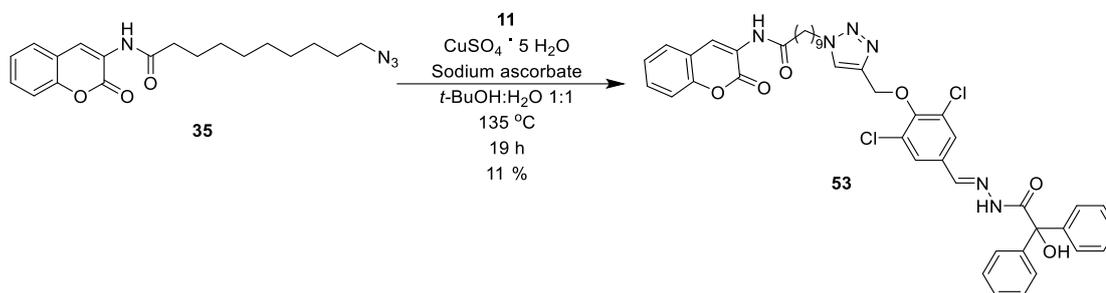
aminocoumarin carbonyl. Seventeen aromatic signals were observed, with the triazole carbons observed at 142.3 ppm and 125.0 ppm, the latter of these signals corresponding to the hydrogen-bearing carbon of this ring. A signal at 57.6 ppm was assigned to the CH<sub>2</sub> carbon linking the triazole and glutaric acid regions, with signals at 35.3, 33.7 and 21.6 ppm ascribed to the carbon closest to the ester, closest to the amide and the central carbon of the glutaric acid region respectively. The remaining nine upfield signals were assigned to the linker carbons. The IR spectrum of the compound did not show the characteristic signal of an azide at roughly 2100 cm<sup>-1</sup>, further indicating that the click reaction had taken place and that compound **35** was absent from the recovered material. A high-resolution accurate mass of 652.3163 [M]<sup>+</sup> was also obtained.



*Scheme 38. Synthesis of click compound 52*

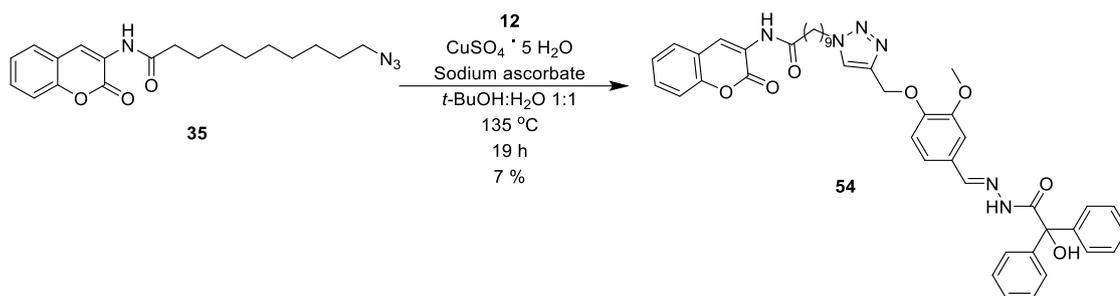
Compound **53** was generated as a light brown solid (**Scheme 39**). <sup>1</sup>H NMR showed the presence of three broad singlets, this time at 11.78, 9.70 and 7.08 ppm, ascribed to the OH group of the benzylic acid, the aminocoumarin-linker amide NH and the hydrazone NH respectively. Six aromatic signals were also observed, corresponding to the expected twenty hydrogens. Two singlets at 8.50 ppm and 8.30 ppm were ascribed to the carbon-bound hydrogen in the hydrazone and the triazole hydrogen respectively. A singlet at 5.22 ppm was assigned to the CH<sub>2</sub> group connecting the triazole to the dichlorinated ring, with the remaining five signals ascribed to the decanoic acid linker. <sup>13</sup>C NMR showed thirty three resolved signals. Two downfield signals at 173.7 ppm and 170.1 ppm were assigned as the linker and benzylic acid carbonyl carbons respectively, whilst a further downfield signal at 158.2 ppm was ascribed to the carbonyl carbon of the lactone ring. Eighteen aromatic signals were observed, with the triazole carbons observed at 142.1 ppm and 125.6 ppm, with the latter being ascribed to the hydrogen-bearing carbon. A further signal at 81.0 ppm corresponding to the carbon bound to the OH group. The OCH<sub>2</sub> carbon connecting the triazole and dichlorinated rings was ascribed to a signal at 66.7 ppm, with the remaining nine signals being assigned to the decanoic acid chain. Once again, the characteristic signal of an azide at 2100 cm<sup>-1</sup> in the IR spectrum was absent, indicating that the triazole had

been formed successfully. A high-resolution accurate mass of 809.2703  $[M]^+$  was also obtained.



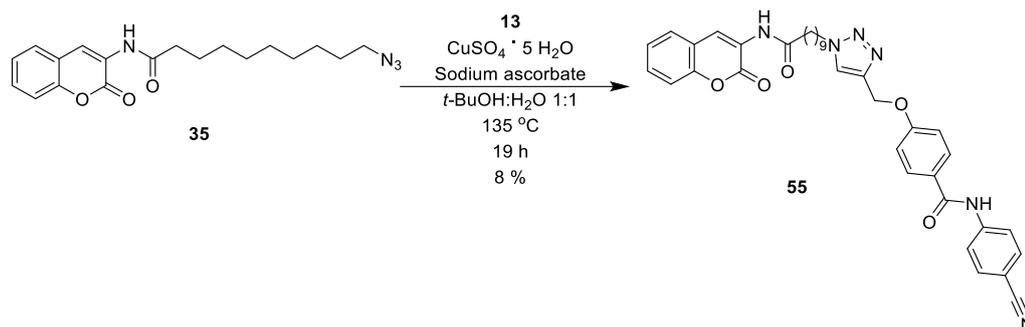
*Scheme 39. Synthesis of click compound 53*

Compound **54** was prepared as a green-brown solid (**Scheme 40**). <sup>1</sup>H NMR showed three broad singlets at 11.38, 9.66 and 6.98 ppm, ascribed to the OH group of the benzylic acid, the aminocoumarin-linker amide NH and the hydrazone NH respectively. Seven aromatic signals were observed, corresponding to the expected twenty hydrogens, with a singlet at 8.23 ppm being assigned as the triazole hydrogen, indicating that this ring had successfully been formed. A further singlet at 5.16 ppm was assigned as the CH<sub>2</sub> hydrogens linking the triazole to the vanillin ring. The remaining upfield signals were ascribed to the linker hydrogens. <sup>13</sup>C NMR showed thirty six resolved signals. Two downfield signals were observed at 173.6 ppm and 169.7 ppm, and were assigned as the linker-aminocoumarin amide carbonyl and hydrazone carbonyl carbons respectively. Twenty aromatic signals were observed, with the triazole carbons assigned to signals at 143.1 ppm and 124.7 ppm, the latter of which corresponding to the hydrogen-bearing carbon on the ring. A signal at 82.0 ppm was ascribed to the carbon of the benzylic acid bearing the OH group, whilst another signal at 62.3 ppm was assigned to the CH<sub>2</sub> group linking the triazole to the vanillin group. A signal at 55.6 ppm was ascribed to the methoxy carbon, with the remaining nine upfield signals being assigned to the linker carbons. The characteristic signal of an azide, a weak signal at 2100 cm<sup>-1</sup>, was absent from the IR spectrum of the compound, further suggesting that the intended product had been generated successfully. A high-resolution accurate mass of 771.3590  $[M]^+$  was also obtained.



*Scheme 40. Synthesis of click compound 54*

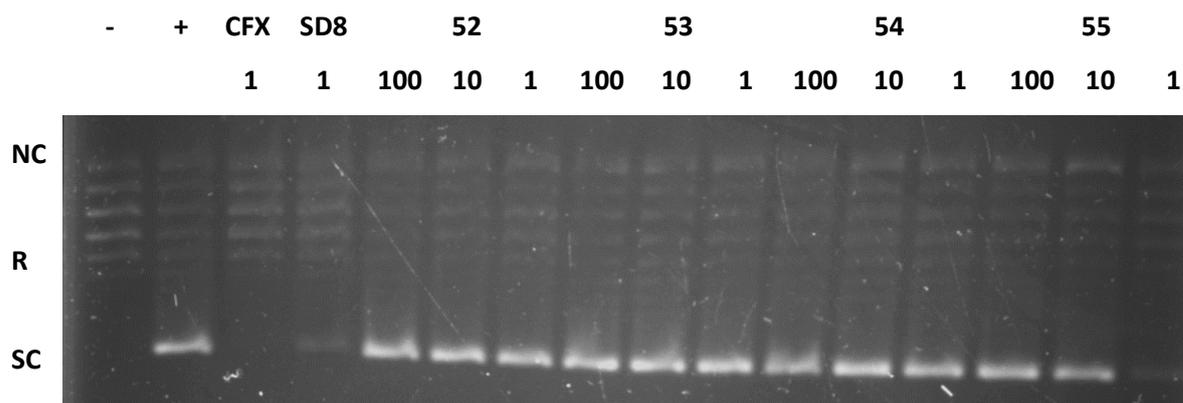
Compound **55** was generated as a white solid (**Scheme 41**). Two broad singlets were observed in the  $^1\text{H}$  NMR spectrum at 10.43 and 9.59 ppm, ascribed to the anilide and aminocoumarin amide hydrogens respectively. A further two singlets at 8.57 ppm and 8.21 ppm were assigned as the aminocoumarin lactone hydrogen and the triazole hydrogen respectively, suggesting that the click reaction had been completed successfully. The remaining aromatic hydrogens were observed in the form of six signals at 7.94, 7.76, 7.63, 7.46, 7.34-7.24 and 7.14 ppm. A singlet at 5.19 ppm was assigned as the  $\text{CH}_2$  group connecting the triazole and benzoic acid groups, whilst four further aliphatic signals were ascribed to the decanoic acid linker hydrogens.  $^{13}\text{C}$  NMR showed thirty two resolved signals. Two downfield signals at 173.7 ppm and 165.9 ppm were assigned as the linker carbonyl and anilide carbonyl carbons respectively. The triazole carbons were observed at 142.6 ppm and 125.02 ppm, with the latter signal corresponding to the hydrogen-bearing carbon of the triazole. Aside from these two carbons, seventeen further aromatic carbon signals were observed as expected. The nitrile carbon was observed at 144.2 ppm, with the carbon linking the triazole to the benzoic acid group corresponding to a signal at 61.8 ppm. Finally, nine upfield signals were ascribed to the linker carbons. The IR spectrum was once again lacking the characteristic signal of an azide, typically found at  $2100 \text{ cm}^{-1}$ , suggesting that the intended product had been formed. A high-resolution accurate mass of 633.2756  $[\text{M}]^+$  was also obtained.

Scheme 41. Synthesis of click compound **55**

### 3.5. Biological testing

Following the completion of synthetic work, the ability of the generated molecules to inhibit DNA gyrase was examined by supercoiling assay. The assay allows for qualitative determination of supercoiling activity, and involves addition of potential inhibitor at a known concentration to a mixture of relaxed circular DNA and *E. coli* DNA gyrase, followed by incubation of the mixture at  $37^\circ\text{C}$  for 30 minutes. The reaction is then arrested by the addition of iso-amyl alcohol in chloroform (1:24) and STEB buffer, a mixture of sucrose, Tris hydrogen chloride, EDTA and bromophenol blue. The various topoisomers of DNA present in the reaction mixture can then be separated using gel electrophoresis and visualised using ethidium bromide. Supercoiled DNA travels the furthest during this process due to the inherent decrease in surface area associated with supercoiling, with “nicked circle” DNA, where only one strand of the DNA is broken, travelling the least distance. Between these two signals are typically multiple bands of relaxed DNA. Therefore, in the presence of a DNA gyrase inhibitor, no supercoiling band should be observed, with only the nicked circle and relaxed DNA bands visible.

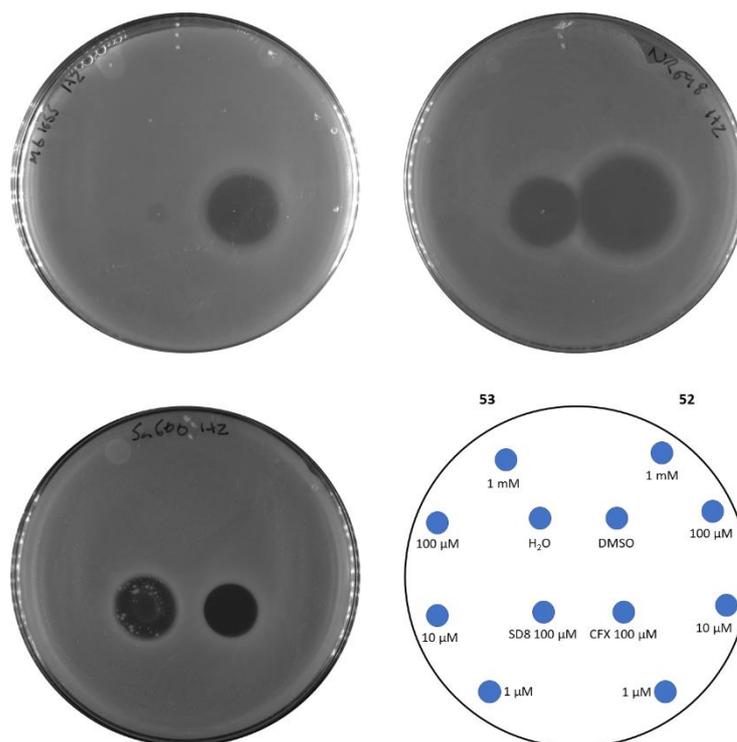
Supercoiling assays were carried out on compounds **52-55** at  $100\ \mu\text{M}$ ,  $10\ \mu\text{M}$  and  $1\ \mu\text{M}$  final concentrations, alongside  $1\ \mu\text{M}$  of SD8 and ciprofloxacin as comparators. The results are shown below in **Figure 56**. The controls showed a good response, with the negative control displaying only relaxed and nicked circle DNA and the positive control a strong supercoiled band. Both SD8 and ciprofloxacin showed good inhibition of supercoiling at  $1\ \mu\text{M}$  as expected. However, compounds **52-55** showed no appreciable inhibition of supercoiling activity at any of the tested concentrations.



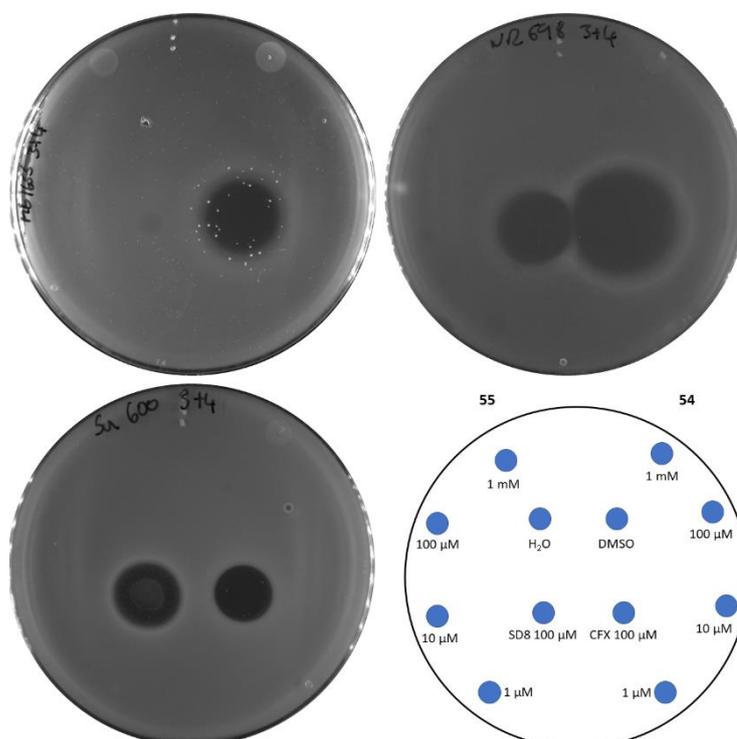
**Figure 56.** Effects of compounds **52-55** on DNA supercoiling by wild type *E. coli* gyrase. Relaxed pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and gyrase as a positive control. Concentrations are given in  $\mu\text{M}$ . CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.

To assess if compounds **52-55** possessed any potency against bacteria outside of inhibition of DNA gyrase, the compounds were used in a zone-of-inhibition assay. This phenotypic assay involves adding a small amount of compound to an LB-agar-plated bacterial culture, and allowing the culture to grow overnight before examination of the plate for areas in which the bacterial culture was unable to grow, indicating areas containing antimicrobials. Three strains of bacteria were used: an *E. coli* wild-type strain (gram-negative), a *S. aureus* wild-type strain (gram-positive), and a further *E. coli* strain which possessed a weakened membrane (NR698). This final strain would therefore allow compounds into the cell more readily, and hence a stronger response to any antimicrobials would be expected. This would also allow for the discounting of cell permeability as a factor for poor activity in the assay. In addition to compounds **52-55**, tested at 10-fold dilutions from 1 mM to 1  $\mu\text{M}$ , controls of water, DMSO, ciprofloxacin (30  $\mu\text{M}$ ) and SD8 (30  $\mu\text{M}$ ) were added to the plates for comparison.

Images of the plates are shown below in **Figure 57** and **Figure 58**. As can be seen, both ciprofloxacin and SD8 controls show good growth inhibition in both the *S. aureus* and *E. coli* NR698 strains, with only ciprofloxacin inhibiting the growth of the *E. coli* wild type strain. This is concordant with previous literature suggesting that SD8 is ineffective at inhibiting the growth of gram-negative bacteria, with the activity against the “leaky membrane” strain suggesting that this lack of effectiveness is likely to be due to the poor cell permeability of SD8.<sup>207,210</sup> In addition, water and DMSO had no observed growth inhibitory effects, as expected. However, no inhibition of growth of any of the strains was observed in the zones treated with compounds **52-55**, suggesting that these molecules possessed negligible activity against bacteria.



**Figure 57.** Zone-of-inhibition assay plates for compounds **52** and **53**. The top left plate contained an *E. coli* wild type strain; top right, *E. coli* NR698; bottom left, *S. aureus* wild type strain. Compound spotting pattern is shown in the bottom right.



**Figure 58.** Zone-of-inhibition assay plates for compounds **54** and **55**. The top left plate contained an *E. coli* wild type strain; top right, *E. coli* NR698; bottom left, *S. aureus* wild type strain. Compound spotting pattern is shown in the bottom right.

### 3.6. Conclusions

The work presented in this chapter aimed to assess the ability of the computational strategy used in **Chapter 2** to design suitable “pseudo-natural products” based on DNA gyrase by synthesis of the selected compounds and biological testing of the compounds by supercoiling assay. Four of the five fragments selected for study were successfully generated, with one, compound **10**, proving a challenge to synthesise, potentially due to the electron-withdrawing nature of the aniline-bound chlorides. In addition, whilst the dehydroxylated aminocoumarin found in SD4 was generated (**27**), attachment of the compound to linker **29** proved a challenge, resulting in further simplification of the aminocoumarin to 3-aminocoumarin (**32**), which was successfully attached to the linker. The four fragments were also successfully attached to the linker by CuAAC click chemistry. However, none of the molecules generated possessed any observed supercoiling inhibitory activity, nor did they appear to inhibit bacterial growth of wild type *E. coli* or *S. aureus* or an *E. coli* strain with a “leaky membrane”.

To improve upon these molecules, it is first important to understand which components of the molecule bind poorly to DNA gyrase. There are several methodologies that could be used to determine this. For example, running ITC of the small molecules selected would allow for the determination of whether these alkynes possess any binding affinity for DNA gyrase, whilst crystal structures of the compounds and DNA gyrase obtained by crystal soaking could provide an insight into where any observed binding interactions are taking place. Another potential factor is the flexibility of the linker, which may be therefore allowing too much movement of the aminocoumarin and fragment groups resulting in poor binding. Improving the rigidity may therefore also provide better inhibition, but may be synthetically challenging.

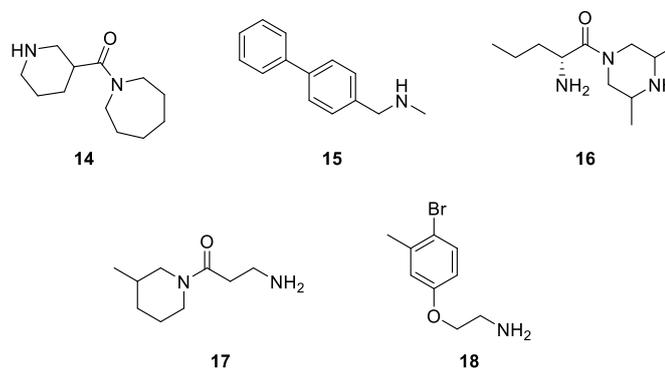
However, the failure of these molecules may not be completely attributed to a failure in the design process. Whilst the simplified 3-aminocoumarin (**32**) was used to generate DNA gyrase inhibitors in previous work, the molecules prepared also included quinolones, known inhibitors of DNA gyrase. It is therefore possible that the activity observed in this previous work could be attributed to the quinolone antibiotics present, as binding of the quinolone could drive the binding of the aminocoumarin. Alternatively, the activity observed could simply be due to the inhibitory effect of the quinolone, with the remainder of the molecule contributing nothing to the mode of action of the molecule. Future work to investigate this could involve obtaining the crystal structure of these aminocoumarin-linker-quinolone compounds and DNA gyrase by crystal soaking, allowing the position of the aminocoumarin within its expected binding pocket to be confirmed or refuted.

**CHAPTER 4: SYNTHESIS OF “PSEUDO-  
NATURAL PRODUCT” COMPOUNDS BASED ON  
SIMOCYCLINONE D8**

To allow for testing of the aminocoumarin-replacing compounds, it was decided that compound **9**, as the highest scoring compound in the *in silico* studies carried out in **Chapter 2**, would be taken forward for the next stage of synthesis and testing. This next stage involved the synthesis of amines **14-18**, identified by the *in silico* work carried out in **Chapter 2**, followed by the attachment of these amines to linker **29** by amide formation. The final synthetic step was to then attach alkyne **9** to these amides using click chemistry, as described previously in **Chapter 3**. These new “pseudo-natural products” would then be tested for their supercoiling inhibitory activity and bacterial growth inhibition.

#### 4.1. Synthesis of aminocoumarin-replacing amines

The work outlined in **Chapter 2** identified five small molecule amines that were predicted to be suitable to replace the aminocoumarin region of simocyclinone D8. These amines are shown below in **Figure 59**. From these compounds, one, compound **14**, was commercially available, and hence was purchased. Compounds **15-18** were also commercially available, but it was deemed cost ineffective to purchase these compounds as their syntheses were predicted to be relatively simplistic.

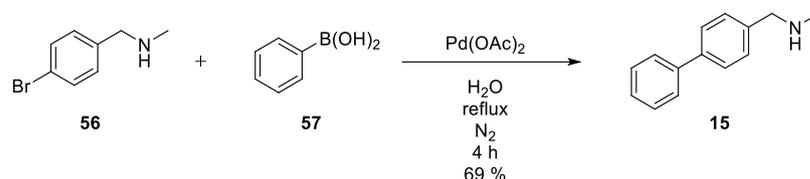


*Figure 59. Amine fragments selected for further study*

##### 4.1.1. Synthesis of compound **15**

Compound **15** was synthesised over the course of a single step using commercially available 4-bromo-N-methylbenzylamine (**56**) and phenylboronic acid (**57**) in a Suzuki reaction (**Scheme 42**). This palladium-catalysed cross-coupling reaction allows for carbon-carbon bond formation between the two aromatic rings, substituting the bromide and boronic acid groups from the starting materials and forming sodium bromide and a tetrahydroxyborate salt as side products. The compound was obtained as a yellow solid. <sup>1</sup>H NMR showed the presence of two multiplets at 7.68-7.58 ppm and 7.49-7.32 ppm, corresponding to the aromatic hydrogens of both rings. Two singlets at 3.68 ppm and 2.29

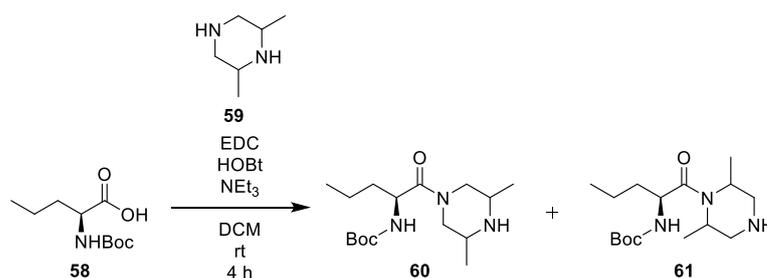
ppm were also observed, ascribed to the CH<sub>2</sub> group and CH<sub>3</sub> group of the N-methylbenzylamine group respectively. <sup>13</sup>C showed ten resolved signals. Eight downfield signals between 140.6 ppm and 126.8 ppm were assigned as the aromatic carbons, with the two remaining signals, observed at 55.2 ppm and 36.1 ppm, ascribed to the CH<sub>2</sub> and CH<sub>3</sub> carbons respectively.



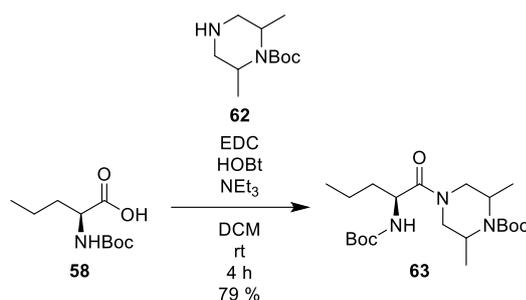
*Scheme 42. Synthesis of ([1,1'-biphenyl]-4-yl)methyl(methyl)amine (15)*

#### 4.1.2. Synthesis of compound 16

Initial attempts to generate compound **16** began with attachment of Boc-protected L-norvaline (**58**) to 2,6-dimethylpiperazine (**59**) to form the Boc-protected intermediate (**60**), as shown below in **Scheme 43**. However, the product of the reaction appeared to be a mixture of **60** and its regioisomer (**61**), with the mixture proving challenging to separate by column chromatography. As a result, the Boc-protected piperazine (**62**) was used in subsequent reactions (**Scheme 44**). It was also hypothesised at this stage that synthesis of **16** followed by attempted attachment of linker **29** would result in a mixture of regioisomers, which may then cause complications at a later stage of the synthesis. It was therefore decided that the Boc group of the piperazine would only be removed upon successful completion of both the attachment of the linker and the triazole formation.



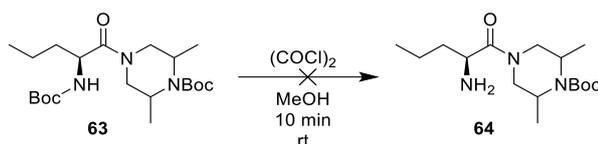
*Scheme 43. Synthesis of regioisomers 60 and 61*



**Scheme 44.** Generation of the double Boc-protected intermediate (**63**)

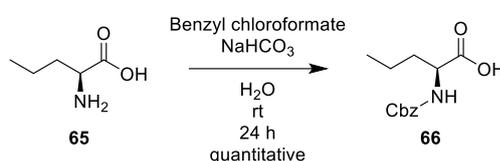
Compound **63** was isolated as a colourless oil. <sup>1</sup>H NMR showed a large upfield multiplet at 1.72-1.22 ppm corresponding to 22 hydrogens, ascribed to both sets of Boc group hydrogens and the two CH<sub>2</sub> hydrogen environments found within the norvaline chain. The two methyl groups present on the piperazine ring were ascribed to two doublets at 1.18 ppm and 1.08 ppm, whilst the methyl hydrogens of the norvaline chain were assigned to a triplet at 0.85 ppm. <sup>13</sup>C NMR showed seventeen resolved signals. A downfield signal at 172.5 ppm was ascribed to the amide carbonyl carbon, whilst two further downfield signals at 155.4 ppm and 154.3 ppm were assigned as the Boc group carbonyl carbons. The remaining carbons of the Boc groups were ascribed to signals at 80.0 ppm and 79.4 ppm for the tetra-substituted carbons and 28.2 ppm and 28.1 ppm for the methyl carbons. The methyl carbons of the dimethylpiperazine were assigned to signals at 19.9 ppm and 19.8 ppm, with the norvaline methyl being ascribed to a signal at 13.6 ppm. The IR spectrum showed a signal at 3313 cm<sup>-1</sup>, assigned as the N-H stretch. Two C=O stretches were observed at 1691 cm<sup>-1</sup> and 1639 cm<sup>-1</sup>. A high-resolution accurate mass of 436.2981 [M+Na]<sup>+</sup> was obtained.

The next stage of the synthesis required the removal of the norvaline Boc group, whilst retaining the piperazine protecting group. Previous work conducted by George *et al.* suggested a mild deprotection of N-Boc groups could be achieved using oxalyl chloride in methanol to generate hydrochloric acid *in situ* (**Scheme 45**).<sup>235</sup> It was hypothesised that the presence of the methyl groups around the piperazine Boc group may hinder the deprotection reaction in this position, resulting in selective deprotection of the norvaline protection. However, when this reaction was attempted both protecting groups were observed to be at least partially removed within 10 minutes by <sup>1</sup>H NMR.



*Scheme 45. Attempted selective Boc deprotection of compound 63*

It was therefore necessary to find a new protecting group for one of the protected positions, which would allow for selective removal of the norvaline protection whilst maintaining protection of the piperazine nitrogen. It was therefore decided that benzoxycarbonyl (Cbz) protection of the norvaline amine group would be attempted (**Scheme 46**). This protecting group is sensitive to hydrogenation, allowing for its selective removal in the presence of the piperazine Boc group.

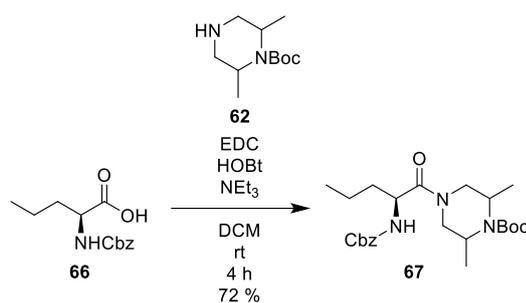


*Scheme 46. Synthesis of Cbz-protected L-norvaline (66)*

Compound **66** was obtained as a white solid.  $^1\text{H}$  NMR showed a singlet at 10.10 ppm, ascribed to the acidic hydrogen. A further signal at 7.28 ppm, partially masked by the solvent signal, was assigned to the Cbz aromatic hydrogens, suggesting that the protection had been successfully achieved. The four remaining signals, consisting of a multiplet at 4.32 ppm, a multiplet at 1.81-1.51 ppm, a sextet at 1.32 ppm, and a multiplet at 0.85 ppm, were ascribed to the hydrogens present on the L-norvaline chain.  $^{13}\text{C}$  NMR showed eleven fully resolved signals. A signal at 204.5 ppm was assigned to the Cbz carbonyl carbon, whilst a further signal at 187.0 ppm was ascribed to the norvaline acidic carbon. Four aromatic signals, observed at 166.6, 157.6, 157.1 and 156.9 ppm, were assigned to the carbons of the Cbz aromatic ring, with a further signal at 95.8 ppm being ascribed to the  $\text{CH}_2$  carbon of the Cbz group. The IR spectrum showed a signal at  $3359\text{ cm}^{-1}$ , assigned as the amide N-H stretch, whilst a further signal at  $3322\text{ cm}^{-1}$  was ascribed to the acid O-H stretch. Two C=O stretches were observed for the ester and acid carbonyls in the form of signals at  $1739\text{ cm}^{-1}$  and  $1688\text{ cm}^{-1}$  respectively.

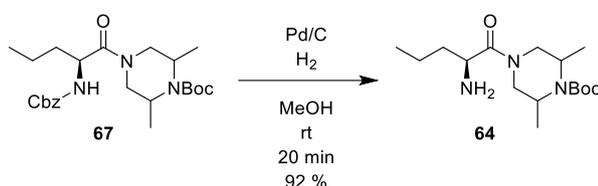
Compounds **62** and **66** were then used to generate the Cbz-protected analogue of compound **63** using a similar procedure, as shown below in **Scheme 47**. Compound **67** was isolated as a colourless oil.  $^1\text{H}$  NMR showed a multiplet at 7.23 ppm, assigned to the aromatic hydrogens of the Cbz group, with a further singlet at 4.98 ppm being ascribed to

the CH<sub>2</sub> hydrogens of the Cbz group. A large, upfield multiplet at 1.59-1.24 ppm was assigned to the hydrogens of the Boc group, as well as the two CH<sub>2</sub> groups of the norvaline chain, suggesting that the amide formation had been successfully achieved. The two methyl groups attached to the piperazine ring were observed in the form of two doublets at 1.14 ppm and 1.02 ppm, with the hydrogens of the norvaline methyl group being ascribed to a triplet at 0.82 ppm. <sup>13</sup>C NMR showed twenty two resolved signals. A downfield signal at 172.7 ppm was assigned to the norvaline-piperazine amide carbonyl carbons, whilst two further signals at 156.9 ppm and 154.7 ppm were ascribed to the Cbz and Boc carbonyl carbons respectively. 6 aromatic signals were observed, corresponding to the benzene ring of the Cbz group. The CH<sub>2</sub> group of the Cbz group was also observed at 66.4 ppm, whilst the quaternary carbon and the methyl carbons of the Boc group were observed at 66.4 ppm and 27.5 ppm respectively. The IR spectrum showed a signal at 3298 cm<sup>-1</sup>, assigned as an N-H stretch. Two C=O stretches were observed at 1688 cm<sup>-1</sup> and 1638 cm<sup>-1</sup>, whilst two further signals at 1342 cm<sup>-1</sup> and 1089 cm<sup>-1</sup> were ascribed to C-N and C-O stretches respectively. A high-resolution accurate mass [M]<sup>+</sup> of 448.2878 was also obtained.



**Scheme 47.** Synthesis of tert-butyl 4-[(2S)-2-[[[(benzyloxy)carbonyl]amino]pentanoyl]-2,6-dimethylpiperazine-1-carboxylate (**67**)

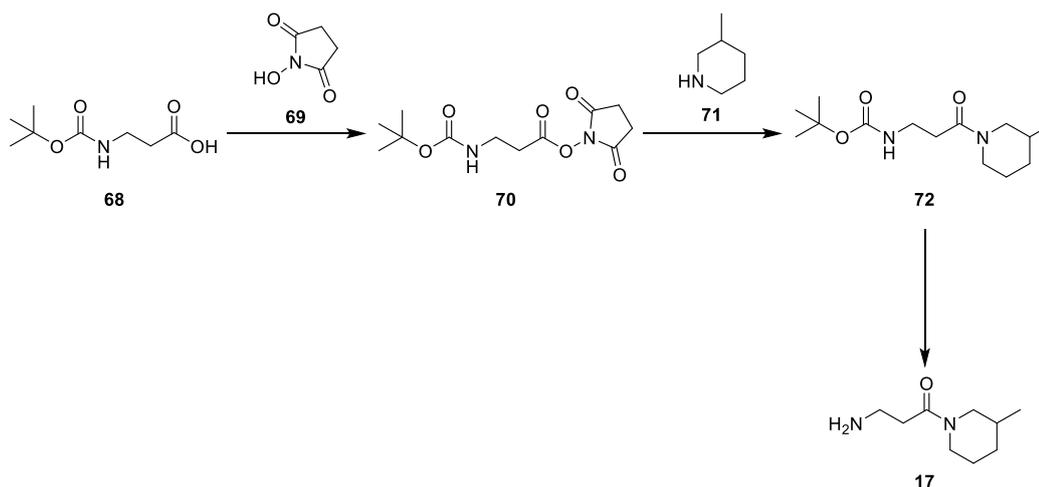
The final step was to selectively remove the Cbz group whilst maintaining the Boc protection. As suggested above, this was achieved by hydrogenation. Compound **64** was generated as a colourless oil (**Scheme 48**).  $^1\text{H}$  NMR showed a lack of aromatic signals, suggesting that the aromatic system of the Cbz group was no longer present. However, the presence of a multiplet at 1.84-1.38 ppm, ascribed to the Boc hydrogens and one of the  $\text{CH}_2$  groups of the norvaline chain, indicated that the Boc protection had been maintained.  $^{13}\text{C}$  NMR showed fourteen resolved signals. The presence of two downfield signals at 175.5 ppm and 154.4 ppm, assigned as the norvaline carbonyl carbon and the Boc carbonyl carbon respectively, further suggested that the Cbz group had been removed successfully. The quarternary and methyl carbons of the Boc group were also observed at 80.0 ppm and 28.4 ppm respectively. A high-resolution accurate mass  $[\text{M}+\text{H}]^+$  of 314.2397 was also obtained.



*Scheme 48. Synthesis of compound 64*

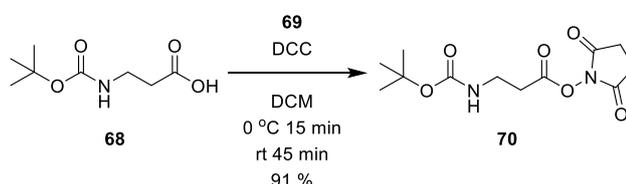
#### 4.1.3. Synthesis of compound 17

Compound **17** was to be generated over the course of three steps. Firstly, the acid group of Boc-protected  $\beta$ -alanine (**68**) was activated by addition of N-hydroxysuccinimide (**69**), followed by attachment of 3-methylpiperidine (**70**) to the  $\beta$ -alanine chain. The final step was to then remove the Boc protection to yield **17**. The overall route is highlighted below in **Scheme 49**.



*Scheme 49. Synthetic route to compound 17*

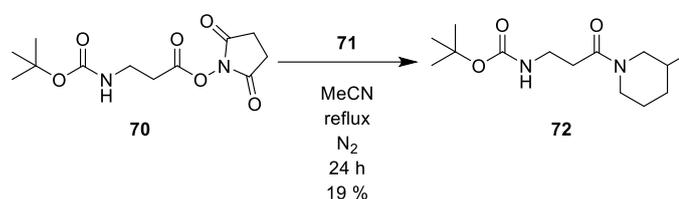
The first step was therefore to attach **69** to **68**. This was achieved using DCC chemistry, a coupling agent similar in both chemistry and structure to EDC. The compound was generated as a white solid (**Scheme 50**). A broad singlet present in the  $^1\text{H}$  NMR spectrum was ascribed to the amine hydrogen, whilst a further singlet at 1.43 ppm was assigned as the Boc hydrogens. Two multiplets at 3.51 ppm and 2.90-2.80 ppm were ascribed to the remaining hydrogens of the  $\beta$ -alanine chain and the succinimide ring.  $^{13}\text{C}$  NMR showed eight resolved signals. Three downfield signals at 169.0, 167.6, and 155.7 ppm were assigned at the three carbonyl carbon environments on the succinimide ring, alanine chain and Boc group respectively. Two further signals at 79.7 ppm and 28.3 ppm were ascribed to the remaining carbons of the Boc group, whilst a signal at 25.6 ppm was assigned as the remaining carbon environment of the succinimide ring. The IR spectrum showed a signal at  $3354\text{ cm}^{-1}$ , corresponding to an N-H stretch. Three C=O stretches were observed in the form of three signals at 1822, 1781 and  $1729\text{ cm}^{-1}$ . The N-O stretch of the N-hydroxysuccinimide group was also observed in the form of a signal at  $1687\text{ cm}^{-1}$ .



*Scheme 50. Synthesis of 2,5-dioxopyrrolidin-1-yl 3-[[tert-butoxy]carbonyl]amino]propanoate (**70**)*

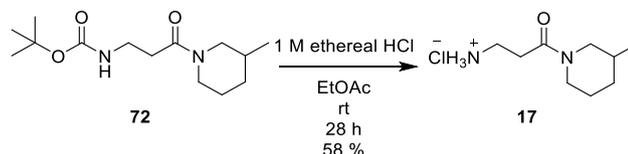
3-Methylpiperidine (**71**) was then bound to the activated  $\beta$ -alanine chain, replacing the N-hydroxysuccinimide. Compound **72** was generated as a yellow oil (**Scheme 51**). A triplet at 0.84 ppm in the  $^1\text{H}$  NMR was ascribed to the methyl group of the 3-methylpiperidine, with the remaining hydrogens of the ring being observed in the form of a doublet of doublets at

3.64 ppm, and multiplets at 4.31, 2.63-2.52, 1.75, 1.68-1.55, 1.54-1.42, and 1.13-0.99 ppm, and a part of a multiplet at 1.35 ppm.  $^{13}\text{C}$  NMR showed twelve resolved signals. The presence of only two downfield signals, observed at 169.4 ppm and 158.8 ppm and assigned to the carbonyl carbons of the alanine chain and Boc group respectively, further indicates that the succinimide ring was removed completely. Signals at 52.6, 48.7, 32.7, 31.4, and 25.6 ppm were ascribed to the five carbons of the piperidine ring, with a further signal at 18.8 ppm being assigned to the methyl attached to the piperidine ring. The IR spectrum showed a signal at  $3330\text{ cm}^{-1}$ , ascribed to an N-H stretch. Two C=O stretches were also observed at  $1708\text{ cm}^{-1}$  and  $1627\text{ cm}^{-1}$ . A high-resolution accurate mass  $[\text{M}+\text{Na}]^+$  of 293.1793 was also obtained.



**Scheme 51.** Synthesis of tert-butyl N-[3-(3-methylpiperidin-1-yl)-3-oxopropyl]carbamate (**72**)

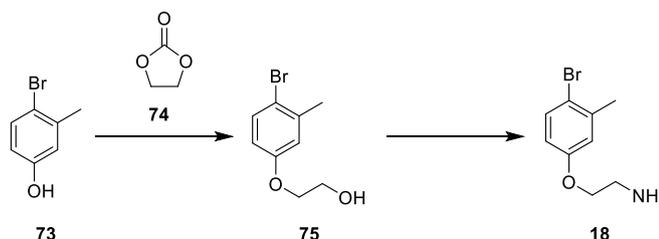
The final step of the synthesis was to remove the Boc protection to yield **17**, achieved using ethereal hydrochloric acid (**Scheme 52**). The compound was yielded as a white solid.  $^1\text{H}$  NMR showed a broad singlet at 8.20 ppm, ascribed to the free amine group and indicating that the Boc group had been successfully removed. This was further confirmed by the loss of the characteristic upfield singlet corresponding to the Boc methyl groups, previously observed for compounds **70** and **72** at 1.43 ppm and 1.35 ppm respectively. The loss of the Boc group was also confirmed by  $^{13}\text{C}$  NMR, with only one downfield signal, corresponding to the piperidinamide carbonyl carbon, observed. Only a single C=O stretch was observed in the IR spectrum, further suggesting that the Boc group had been removed completely. Finally, a high-resolution accurate mass of 171.1419  $[\text{M}]^+$  was obtained.



**Scheme 52.** Synthesis of 3-(3-methylpiperidin-1-yl)-3-oxopropan-1-aminium chloride (**17**)

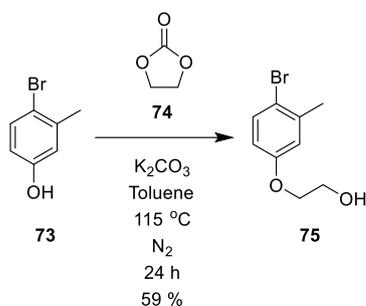
## 4.1.4. Synthesis of compound 18

Compound **18** was generated over the course of two steps (**Scheme 53**). Firstly, 4-bromo-3-methylphenol (**73**) was reacted with ethylene carbonate (**74**) to form the phenoxyethanol intermediate (**75**), following which the alcohol group underwent functional group interconversion to yield the final product **18**.

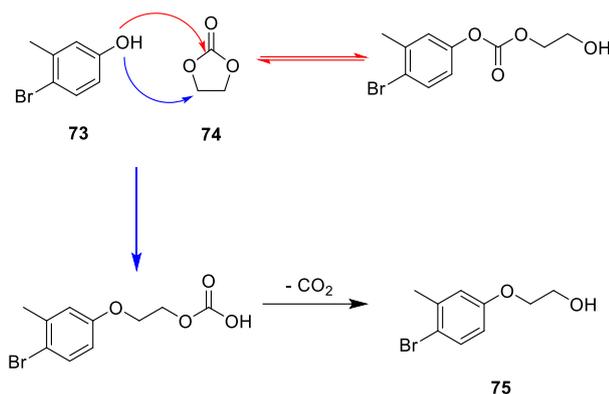


*Scheme 53. Synthetic route to amine 18*

The first step was therefore to convert the phenol group into an ether using ethylene carbonate under basic conditions (**Scheme 54**).<sup>236</sup> The reaction between the alcohol and the carbonate can occur at both the carbonyl carbon and the ethylene carbon, with both sites electrophilic enough to accept attack due to the adjacent electron withdrawing oxygens (**Scheme 55**). However, when the reaction proceeds via attack of the ethylene group, the side product is carbon dioxide gas, making this route the thermodynamically favoured pathway.



*Scheme 54. Synthesis of intermediate compound 75*

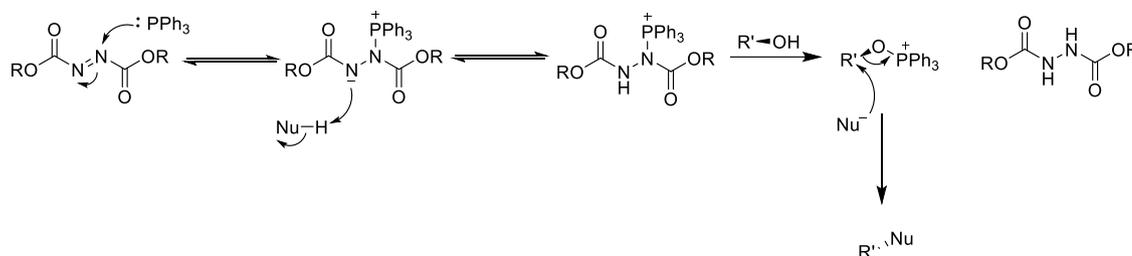


**Scheme 55.** Possible sites of attack on ethylene carbonate (**74**) for phenol **73**, and the resulting intermediates

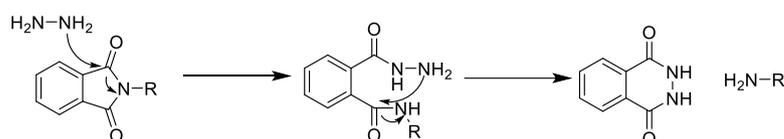
Compound **75** was isolated as a yellow oil. <sup>1</sup>H NMR showed two aromatic doublets and an aromatic doublet of doublets, observed at 7.44 ppm, 6.84 ppm and 6.67 ppm respectively. The doublet of doublets was assigned to the aromatic hydrogen in the *ortho* position to the oxygen and in the *para* position to the methyl group, with the remaining doublets being ascribed to the remaining aromatic hydrogens. Two multiplets at 4.07 ppm and 3.98 ppm were assigned to the ethyl chain hydrogens, with the signal at 4.07 ppm corresponding to the hydrogens closest to the aromatic ring. A final singlet at 2.39 ppm was ascribed to the methyl group. <sup>13</sup>C NMR showed nine clearly resolved signals. A signal at 157.8 ppm was assigned to the aromatic carbon attached to the oxygen, whilst a signal at 139.0 ppm was ascribed to the carbon attached to the methyl group. Signals at 133.0, 117.2, and 113.6 ppm were assigned to the hydrogen-bearing aromatic carbons, with a final aromatic signal observed at 115.9 ppm assigned to the carbon attached to the bromide. Three aliphatic signals were also observed at 69.4, 61.4 and 23.1 ppm, ascribed to the two ethyl group carbons and the methyl carbon respectively. These data are similar to those observed in previous work, but are subtly different due to the use of differing solvents for the NMR analyses.<sup>237</sup> The IR spectrum showed a signal at 3293 cm<sup>-1</sup>, corresponding to an O-H stretch. Two signals at 1589 cm<sup>-1</sup> and 1477 cm<sup>-1</sup> were assigned as aromatic C=C stretches, whilst a further signal at 1450 cm<sup>-1</sup> was ascribed to the methyl C-H bend. A high-resolution accurate mass of 230.9703 [M+H]<sup>+</sup> was also obtained.

The final step of the synthesis of **18** was to convert the alcohol group of intermediate **75** to an amine. This was achieved using a Mitsunobu reaction, first reported by Oyo Mitsunobu and Masaaki Yamada in 1967.<sup>238</sup> The methodology uses an azodicarboxylate, in this instance diisopropyl azodicarboxylate (DIAD), and triphenylphosphine to allow for displacement of the alcohol group by a given nucleophile in an S<sub>N</sub>2-type reaction. The precise mechanism has not yet been confirmed, however one possibility is that shown

below in **Scheme 56**. To achieve conversion of an alcohol to an amine using the Mitsunobu reaction, phthalimide is used as the nucleophile, following which the amine functional group is introduced using hydrazine. This opens the phthalimide ring in a similar fashion to the second stage of the Gabriel synthesis, generating a stable bicyclic byproduct (**Scheme 57**).

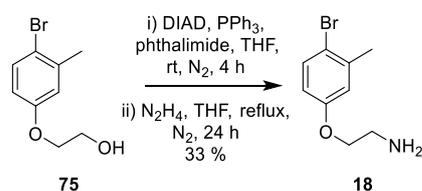


*Scheme 56. Mechanism of the Mitsunobu reaction*



*Scheme 57. Mechanism of amine cleavage using hydrazine*

Compound **18** was generated as an off-white solid (**Scheme 58**). Both the  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra closely matched previously reported spectroscopic data, with the key signal, a multiplet at 1.36 ppm on the  $^1\text{H}$  NMR spectra, being ascribed to the primary amine.<sup>237</sup> The IR spectrum of the compound showed a signal at  $1592\text{ cm}^{-1}$ , assigned as an N-H bending wavenumber, with a further signal at  $1135\text{ cm}^{-1}$  ascribed to a C-N stretch, further suggesting the presence of the amine.

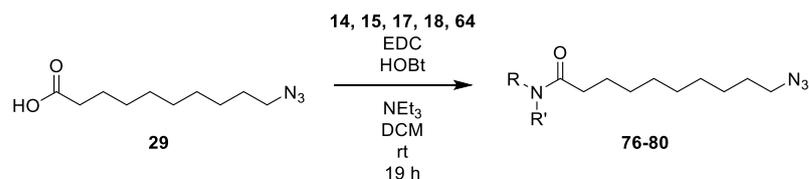


*Scheme 58. Generation of amine 18*

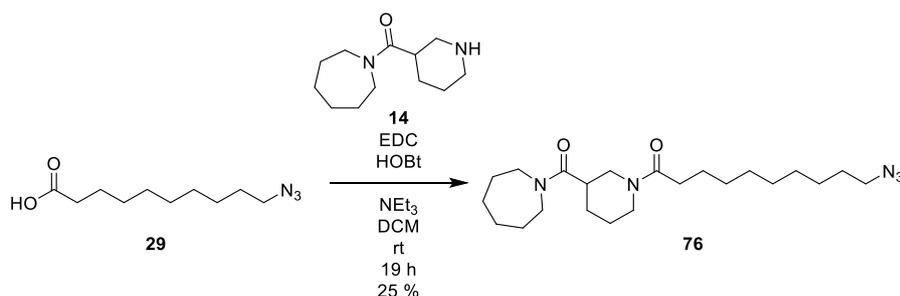
## 4.2. Final assembly of “pseudo-natural products”

### 4.2.1. Attachment of amines 14, 15, 17, 18 and 64 to linker 29

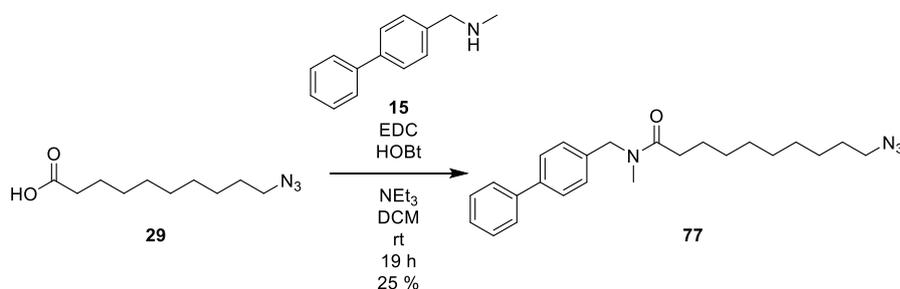
Upon successful completion of the syntheses of compounds **14**, **15**, **17**, **18** and **64**, the amines were attached to the linker using EDC coupling chemistry (**Scheme 59**). Compound **76**, generated from amine **14**, was isolated as a brown solid (**Scheme 60**).  $^1\text{H}$  NMR showed a multiplet at 4.59 ppm corresponding to a single hydrogen, ascribed to the substituted position of the piperidine ring. A triplet at 3.77 ppm, a doublet of triplets at 2.98 ppm, a multiplet at 2.63 ppm and part of a multiplet at 1.89 ppm were assigned to the remaining hydrogens of the piperidine ring. The hydrogens of the azepane ring were observed in the form of a multiplet at 3.53 ppm and part of a multiplet at 1.89 ppm. The hydrogens of the decane chain were observed in the form of a triplet at 3.18 ppm, ascribed to the closest hydrogens to the azide group, a multiplet at 2.30 ppm, assigned to the next closest hydrogens to the azide, and the remaining hydrogens contained within the multiplets at 1.89 ppm and 1.31 ppm.  $^{13}\text{C}$  NMR showed twenty two resolved signals, as expected. Two downfield signals at 172.6 ppm and 171.7 ppm, ascribed to the carbonyl carbons of the linker and amine respectively, suggested that the linker had been successfully incorporated into the molecule. A signal at 51.4 ppm was assigned to the carbon closest to the azide, with the adjacent carbon to this being observed as a signal at 33.3 ppm. The remainder of the decane chain could not be specifically assigned, but was observed as part of a series of signals between 29.3 ppm and 25.2 ppm. Signals at 47.4 ppm and 46.1 ppm were ascribed to the carbons of the azepane ring closest to the nitrogen, with the remaining carbons of the ring once again being observed within the series of upfield signals. Finally, signals at 46.2, 44.8 and 41.9 ppm, as well as some signals located between 29.3 ppm and 25.2 ppm, were assigned to the carbons of the piperidine ring. IR analysis showed a signal at  $2091\text{ cm}^{-1}$ , ascribed to the characteristic  $\text{N}=\text{N}=\text{N}$  azide stretch, further confirming the presence of the linker chain within the molecule. Two  $\text{C}=\text{O}$  stretches were also observed at  $1647\text{ cm}^{-1}$  and  $1626\text{ cm}^{-1}$ . Finally, a high-resolution accurate mass of  $406.3119\text{ [M]}^+$  was obtained.



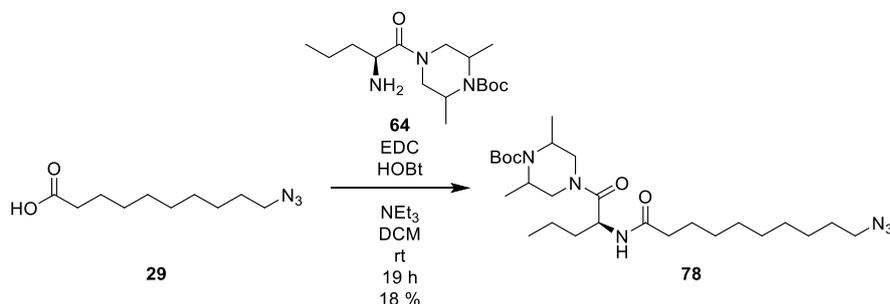
*Scheme 59. General synthesis of amides 76-80*

Scheme 60. Synthesis of amide **76**

Compound **77**, generated from amine **15**, was generated as a yellow oil (Scheme 61). <sup>1</sup>H NMR showed three multiplets at 7.78, 7.49 and 7.31 ppm, ascribed to the aromatic hydrogens of the biphenyl region. A further multiplet at 4.54 ppm was assigned to the CH<sub>2</sub> group between the biphenyl system and the amide, whilst a signal partially masked by the water signal at 3.32 ppm was ascribed to the nitrogen-bound methyl group. A multiplet at 2.94 ppm was assigned as the hydrogens of the decane chain closest to the azide group, with the remaining hydrogens of the chain observed as two large multiplets at 1.58 ppm and 1.35 ppm. <sup>13</sup>C NMR showed twenty resolved signals. One downfield shift was observed at 173.5 ppm, assigned as the carbonyl carbon. Eight aromatic signals were observed, with five of these, observed at 127.7, 127.2, 126.0, 125.9 and 125.7 ppm, confirmed by HSQC analysis to correspond to the hydrogen-bearing carbons of the biphenyl system. Signals at 52.0 ppm and 32.5 ppm were ascribed to the CH<sub>2</sub> and CH<sub>3</sub> carbons connected to the amide nitrogen respectively. The two closest carbons to the azide group were also identified and assigned to signals at 31.7 ppm and 23.9 ppm. The IR spectrum once again showed a strong signal at 2093 cm<sup>-1</sup>, suggesting the presence of an azide. A signal observed at 1736 cm<sup>-1</sup> was ascribed to a C=O stretch, whilst a further signal at 1644 cm<sup>-1</sup> was assigned as an aromatic C=C stretch. Finally, a high-resolution accurate mass of 393.2621 [M]<sup>+</sup> was obtained.

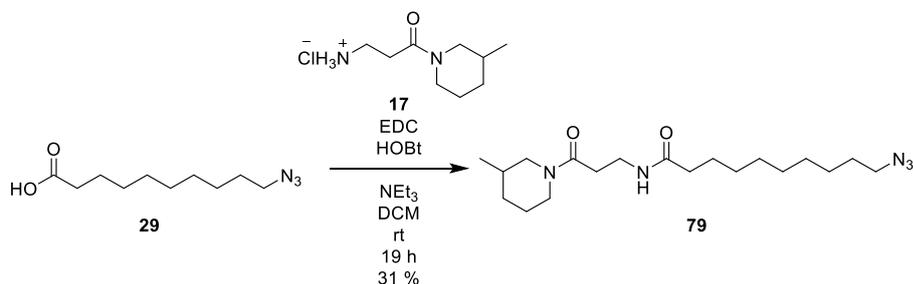
Scheme 61. Generation of compound **77**

Compound **78**, generated from amine **64**, was isolated as a pale yellow oil (**Scheme 62**). A multiplet at 5.05 ppm in the  $^1\text{H}$  NMR spectrum was ascribed to the hydrogen adjacent to the amide nitrogen on the norvaline chain. Multiplets at 4.45, 4.28, 2.89 and 2.22 ppm were assigned to the piperazine ring hydrogens, with a further multiplet at 1.17 ppm and part of a multiplet at 1.43 ppm ascribed to the methyl groups attached to this ring. The Boc hydrogens were observed in the form of a singlet at 1.49 ppm, whilst the methyl of the norvaline chain was observed as a multiplet at 0.94 ppm. A triplet at 3.28 ppm was assigned to the hydrogens closest to the azide on the decane chain, with the remaining hydrogens of the chain being observed as part of multiplets at 1.87 ppm and 1.43 ppm.  $^{13}\text{C}$  NMR showed twenty four resolved signals. Two downfield signals observed at 172.9 ppm and 162.1 ppm were assigned to the norvaline carbonyl carbon and the decanamide carbonyl carbon respectively. A further downfield signal at 154.8 ppm was ascribed to the Boc carbonyl carbon, with the remaining carbons of the Boc group being assigned to signals at 80.1 ppm and 28.4 ppm. Signals at 46.6, 46.4, 46.1 and 36.7 ppm were ascribed to the carbons of the piperazine ring, with two further signals at 20.1 ppm and 19.7 ppm assigned to the methyl carbons attached to the ring. The norvaline chain carbons were observed at 48.1, 35.3, 18.5 and 13.9 ppm, whilst the carbon closest to the azide on the decane chain was observed at 51.5 ppm. The remaining carbons of the decane chain were ascribed to signals between 29.3 ppm and 20.2 ppm. The IR spectrum once again showed a strong  $\text{N}=\text{N}=\text{N}$  stretch at  $2092\text{ cm}^{-1}$ , indicative of the presence of an azide group and hence the azidodecanamide region. Three  $\text{C}=\text{O}$  stretches were observed at  $1711$ ,  $1684$  and  $1643\text{ cm}^{-1}$ , with an ester  $\text{C}-\text{O}$  stretch also observed at  $1344\text{ cm}^{-1}$ . Finally, a high-resolution accurate mass of  $531.3845\text{ [M+Na]}^+$  was obtained.



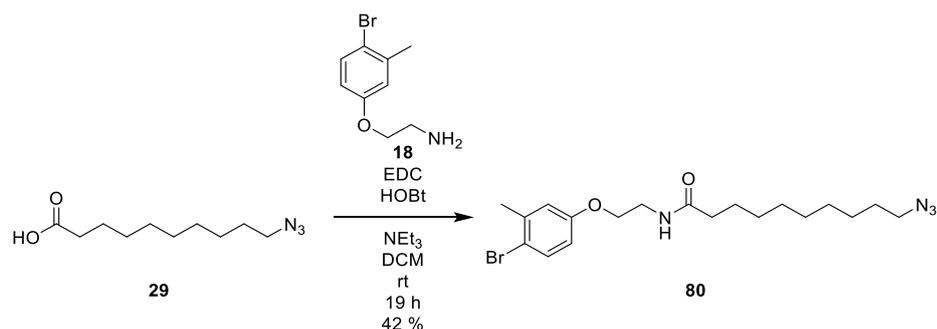
*Scheme 62. Synthesis of amide 78*

Compound **79**, generated from amine **17**, was isolated as a white solid (**Scheme 63**). The  $^1\text{H}$  NMR spectrum showed the hydrogens of the piperidine ring in the form of multiplets at 4.38, 3.73, 3.00 and 2.66 ppm and forming part of multiplets at 1.86 ppm and 1.48 ppm. A doublet at 0.91 ppm was assigned as the methyl hydrogens of the methylpiperidine group, whilst two triplets at 3.56 ppm and 2.51 ppm were ascribed to the two  $\text{CH}_2$  groups of the beta-alanine. A further triplet at 3.27 ppm was assigned as the closest hydrogens of the decane chain to the azide group, whilst a further triplet at 2.14 ppm was ascribed to the  $\text{CH}_2$  closest to the amide group on this chain. The remaining hydrogens of the chain formed part of multiplets observed at 1.86 ppm and 1.48 ppm.  $^{13}\text{C}$  NMR showed nineteen resolved signals. Two downfield signals observed at 173.2 ppm and 170.0 ppm were assigned as the carbonyl carbons of the decanamide and alanine regions respectively. Signals at 45.9, 42.2, 32.9, 29.3 and 25.7 ppm were ascribed to the carbons of the piperidine ring, with a further signal at 18.8 ppm assigned to the methyl carbon attached to this ring. The  $\text{CH}_2$  carbons of the beta alanine region were ascribed to signals at 35.0 ppm and 32.9 ppm, whilst the  $\text{CH}_2$  carbon closest to the azide group on the decane chain was assigned to a signal at 51.5 ppm. The  $\text{CH}_2$  group adjacent to this carbon was also identified at 31.0 ppm, whilst the two carbons closest to the amide group on the decane chain were assigned to signals at 36.8 ppm and 24.7 ppm. The remaining carbons of this chain were ascribed to signals between 29.2 ppm and 25.8 ppm. The IR spectrum showed a broad signal at  $3308\text{ cm}^{-1}$ , assigned to an N-H stretch. A strong signal at  $2087\text{ cm}^{-1}$  was ascribed to the characteristic  $\text{N}=\text{N}=\text{N}$  stretch of an azide, further suggesting that the azidodecane region was successfully incorporated into the product. Two  $\text{C}=\text{O}$  stretches were observed at  $1638\text{ cm}^{-1}$  and  $1549\text{ cm}^{-1}$ . Finally, a high-resolution accurate mass of  $366.2795\text{ [M]}^+$  was obtained.



*Scheme 63. Generation of compound 79*

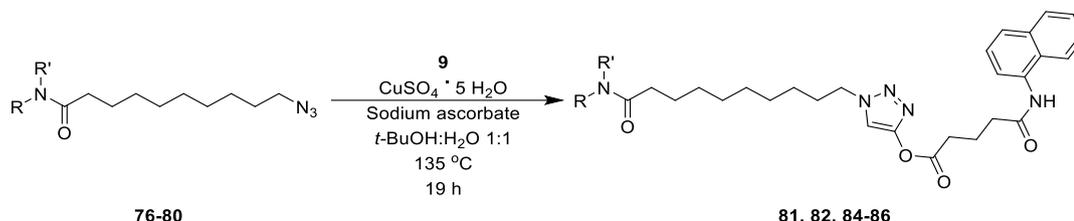
Compound **80**, generated from amine **18**, was isolated as a brown oil (**Scheme 64**). A broad singlet at 8.02 ppm in the  $^1\text{H}$  NMR spectrum was ascribed to the amide hydrogen, the shift downfield from the spectrum for compound **18** suggesting that the amine had been successfully bound to the acid group of **29**. Two doublets at 7.44 ppm and 6.95 ppm, as well as a doublet of doublets at 6.74 ppm, were assigned to the aromatic hydrogens. A triplet at 3.97 ppm was ascribed to the  $\text{CH}_2$  group closest to the aromatic ring, whilst a signal masked by the HDO signal was assigned to the adjacent  $\text{CH}_2$  hydrogens, as well as the hydrogens closest to the azide on the decane chain. A singlet at 2.30 ppm was ascribed to the methyl group attached to the aromatic ring, with the remaining signals being assigned to the hydrogens of the decane chain.  $^{13}\text{C}$  NMR showed nineteen resolved signals. A downfield signal at 177.7 ppm was ascribed to the amide carbonyl carbon, whilst a further downfield signal at 163.1 ppm was assigned to the aromatic carbon closest to the ether group. Five further aromatic signals were observed at 143.5, 137.8, 122.5, 119.8 and 119.3 ppm. The two carbons between the aryl ether and amide groups were assigned to signals at 71.8 ppm and 43.9 ppm, whilst the methyl carbon bound to the aromatic ring was ascribed to a signal at 27.8 ppm. The remaining signals were assigned to the decane chain. The IR spectrum showed a signal at  $3315\text{ cm}^{-1}$ , ascribed to an N-H stretch, whilst a signal at  $2095\text{ cm}^{-1}$  was once again assigned as an azide  $\text{N}=\text{N}=\text{N}$  stretch. A carbonyl  $\text{C}=\text{O}$  stretch was observed at  $1639\text{ cm}^{-1}$ , as well as an aromatic  $\text{C}=\text{C}$  stretch at  $1609\text{ cm}^{-1}$ . A C-Br stretch was also observed at  $638\text{ cm}^{-1}$ . Finally, a high-resolution accurate mass of  $424.1197\text{ [M]}^+$  was obtained, with the bromine isotope signal also observed at 426.1534.



*Scheme 64. Synthesis of amide 80*

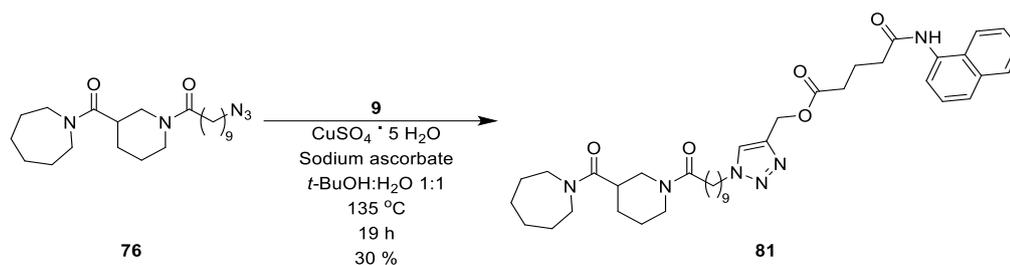
4.2.2. CuAAC attachment of alkyne **9** to amine-linker compounds **76-80**

The final step for compounds **76**, **77**, **79** and **80** was to attach alkyne **9** via copper-catalysed click chemistry, as discussed previously in **Chapter 3** and shown below in **Scheme 65**. For compound **78**, an additional deprotection step would be required following the click reaction to remove the Boc group.



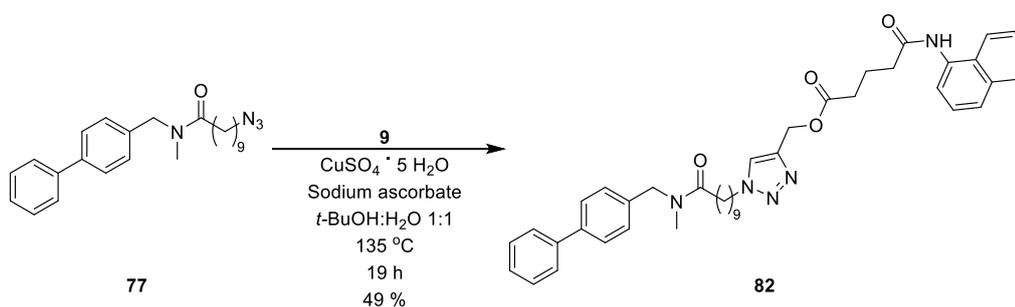
*Scheme 65. General route for the generation of "pseudo-natural products" **81**, **82**, and **84-86***

Compound **76** was used to generate “pseudo-natural product” **81** as a red-brown oil (**Scheme 66**).  $^1\text{H}$  NMR showed a multiplet at 8.69 ppm, assigned to the triazole hydrogen. Two further multiplets at 7.83 ppm and 7.38 ppm were ascribed to the aromatic hydrogens of the naphthalene ring system. A singlet at 5.13 ppm was assigned to the  $\text{CH}_2$  hydrogens between the triazole and the glutaric acid linker, whilst a triplet at 4.15 ppm was ascribed to the hydrogens closest to the triazole on the decane linker. The hydrogens of the glutaric acid linker were observed as part of a multiplet at 2.62 ppm, a multiplet at 2.18 ppm and a multiplet at 1.96 ppm, whilst the hydrogens of the azepane ring were ascribed to a multiplet at 3.50 ppm and part of a multiplet at 1.84 ppm. The decane linker hydrogens were assigned to part of the 2.62 ppm multiplet, part of the 1.84 ppm multiplet, and a multiplet at 1.23 ppm.  $^{13}\text{C}$  NMR showed forty resolved signals. Four downfield signals at 173.1, 172.7, 172.2 and 172.0 ppm were observed and ascribed to the four carbonyl carbons. 12 aromatic signals were observed, with two signals at 145.1 ppm and 125.49 ppm being assigned to the triazole carbons, further suggesting that the click reaction had been completed successfully. The IR spectrum showed a signal at  $3281 \text{ cm}^{-1}$ , ascribed to an N-H stretch. Two C=O stretches were observed at  $1734 \text{ cm}^{-1}$  and  $1634 \text{ cm}^{-1}$ , with two further signals at  $1599 \text{ cm}^{-1}$  and  $1502 \text{ cm}^{-1}$  assigned as aromatic C=C stretches. Finally, a high-resolution accurate mass of  $701.4362 [\text{M}]^+$  was obtained.



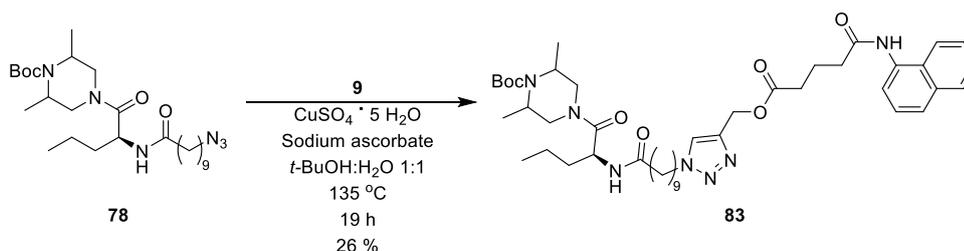
*Scheme 66. Synthesis of compound 81 via copper-catalysed click chemistry*

Compound **77** was used to generate compound **82** as an off-white solid (**Scheme 67**).  $^1\text{H}$  NMR showed a multiplet at 8.34 ppm, ascribed to the triazole hydrogen. Multiplets at 7.88, 7.76 and 7.40 ppm, as well as a doublet at 7.59 ppm, were assigned to the aromatic hydrogens of the biphenyl ring system, whilst two further multiplets at 7.52 ppm and 7.41 ppm were ascribed to the aromatic hydrogens of the naphthalene rings. The  $\text{CH}_2$  hydrogens between the triazole and the glutaric acid linker were once again observed as a singlet at 5.19 ppm, whilst the N-methyl hydrogens were assigned to a multiplet at 4.52 ppm.  $^{13}\text{C}$  NMR showed thirty eight resolved signals. Three downfield signals at 173.6, 173.2 and 172.6 ppm were ascribed to the three carbonyl carbons present in the structure. Twenty aromatic signals were observed, with two signals at 140.8 ppm and 128.4 ppm being assigned to the two triazole carbons. The N-methyl carbon was also observed as a signal at 29.9 ppm, with the  $\text{CH}_2$  carbon linking the biphenyl system to the amide being ascribed to a signal at 50.7 ppm. The IR spectrum showed a signal at  $3289 \text{ cm}^{-1}$ , corresponding to an N-H stretch. Three C=O stretches were observed at  $1720 \text{ cm}^{-1}$ ,  $1658 \text{ cm}^{-1}$  and  $1640 \text{ cm}^{-1}$ , with a further signal at  $1599 \text{ cm}^{-1}$  corresponding to an aromatic C=C stretch. Finally, a high-resolution accurate mass of  $688.3936 [\text{M}]^+$  was obtained.



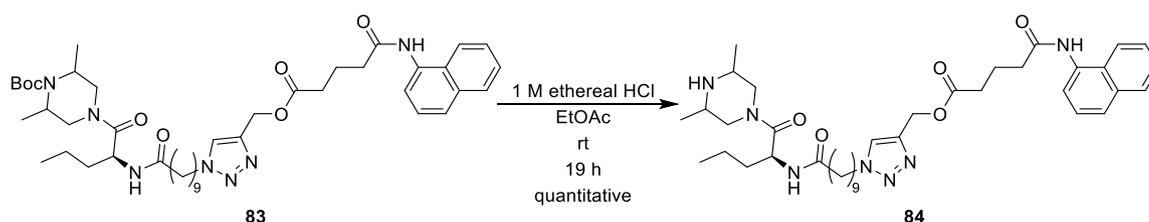
*Scheme 67. Synthesis of compound 82 using copper-catalysed click chemistry*

Compound **78** was used to generate compound **83** as a yellow oil (**Scheme 68**).  $^1\text{H}$  NMR showed a multiplet at 8.65 ppm, ascribed to the triazole hydrogen. Multiplets at 7.92, 7.73, 7.53 and 7.43 ppm, as well as a doublet at 7.61 ppm, were assigned to the aromatic hydrogens of the naphthalene rings, whilst a singlet at 5.18 ppm was ascribed to the hydrogens of the  $\text{CH}_2$  group linking the triazole to the glutaric acid linker. A multiplet at 4.86 ppm was assigned as the norvaline hydrogen geminal to the amide nitrogen, with the remaining hydrogens of the norvaline side chain being found as part of multiplets at 2.09 ppm and 1.63 ppm, as well as a multiplet at 0.83 ppm which was ascribed to the methyl hydrogens of the chain. The hydrogens of the piperazine ring were also observed as part of the multiplets at 2.09 ppm and 1.63 ppm, as well as forming part of a multiplet at 1.08 ppm. The large singlet expected of a Boc group was also observed as part of the multiplet at 1.63 ppm.  $^{13}\text{C}$  NMR showed forty two resolved signals. Four downfield signals were observed at 173.2, 172.6, 171.6 and 170.2 ppm, and were ascribed to the carbonyl carbon of the decanoic acid linker, the glutaric acid linker, and the norvaline chain. A final downfield signal at 153.6 ppm was assigned as the Boc carbonyl carbon, with the remaining carbons of the protecting group being observed at 80.4 ppm and 19.4 ppm. Twelve aromatic signals were observed, with signals at 142.7 ppm and 125.9 ppm being ascribed to the triazole carbons. The IR spectrum showed a strong signal at  $3281\text{ cm}^{-1}$ , assigned to an N-H stretching signal. Two C=O stretches were observed at  $1734\text{ cm}^{-1}$  and  $1634\text{ cm}^{-1}$ , whilst a further two signals at  $1598\text{ cm}^{-1}$  and  $1502\text{ cm}^{-1}$  were ascribed to aromatic C=C stretches. Finally, a high-resolution accurate mass of  $804.5148\text{ [M+H]}^+$  was obtained.



*Scheme 68. Synthesis of compound **83** using copper-catalysed click chemistry*

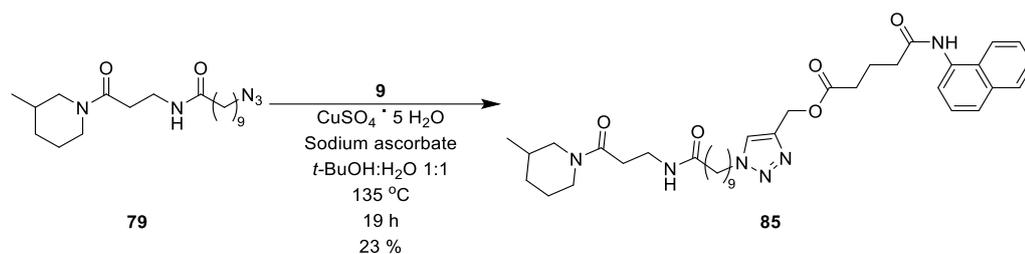
Compound **83** was then deprotected using ethereal hydrochloric acid, as described in previous Boc deprotection procedures (**Scheme 69**). Compound **84** was yielded as a yellow syrup. As expected, both the  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra closely matched that of compound **83**. However, the large upfield signal corresponding to the Boc group hydrogens was no longer present in the  $^1\text{H}$  NMR, suggesting that the group had been successfully removed. This was further confirmed by the loss of signals corresponding to the Boc carbonyl and the quaternary carbon on the  $^{13}\text{C}$  NMR spectrum. A high-resolution accurate mass of 726.4332  $[\text{M}+\text{Na}]^+$  was also obtained.



*Scheme 69. Deprotection of compound **83** to generate compound **84***

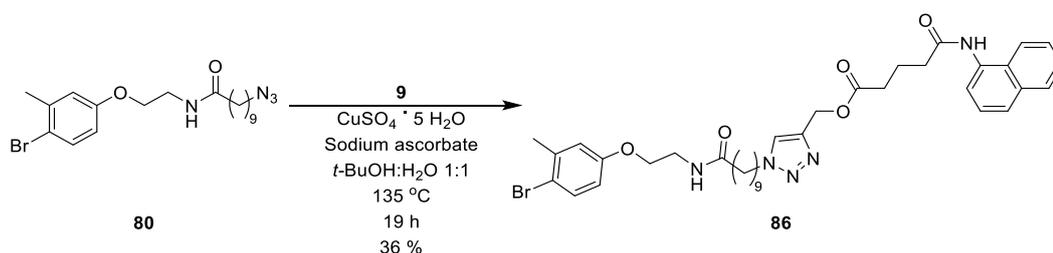
Compound **79** was used to generate compound **85** as a brown solid (**Scheme 70**).  $^1\text{H}$  NMR showed a multiplet at 8.47 ppm, ascribed to the triazole hydrogen. Multiplets at 7.88, 7.75, 7.60, 7.49 and a signal masked by the  $\text{CDCl}_3$  solvent signal were assigned to the naphthalene hydrogens, whilst a singlet at 5.16 ppm was ascribed to the  $\text{CH}_2$  group between the triazole and glutaric acid linker. The hydrogens of the beta alanine chain were observed as part of multiplets at 3.51, 2.77 and 2.22 ppm, whilst the hydrogens of the piperidine ring were observed as part of multiplets at 3.51, 2.56, 2.22, 1.58 and 1.32 ppm. The methyl group hydrogens were also observed as a multiplet at 0.79 ppm.  $^{13}\text{C}$  NMR showed thirty seven resolved signals. Four downfield signals at 173.3, 173.1, 171.5 and 170.0 ppm were assigned as the four carbonyl carbons. Twelve aromatic signals were observed, with signals at 142.9 ppm and 123.4 ppm ascribed to the triazole carbons. The carbon between the triazole and the glutaric acid linker was assigned to a signal at 57.6 ppm, whilst the  $\text{CH}_2$  carbons of the beta-alanine chain were ascribed to signals at 36.7 ppm and 35.0 ppm. The methyl group attached to the piperidine ring was also assigned to a signal at 18.9 ppm. The IR spectrum showed a signal at  $3272 \text{ cm}^{-1}$ , ascribed to an N-H stretch. Three C=O stretches were observed at  $1727 \text{ cm}^{-1}$ ,  $1638 \text{ cm}^{-1}$  and  $1627 \text{ cm}^{-1}$ , with a further signal at  $1503 \text{ cm}^{-1}$  assigned to an aromatic C=C stretch. Finally, a high-resolution accurate mass of 661.4144  $[\text{M}]^+$  was obtained.

CHAPTER 4: SYNTHESIS OF “PSEUDO-NATURAL PRODUCT” COMPOUNDS BASED ON SIMOCYCLINONE D8



*Scheme 70. Synthesis of compound 85 by copper-catalysed click chemistry*

Compound **80** was used to generate compound **86** as an off-white solid (**Scheme 71**).  $^1\text{H}$  NMR showed a multiplet at 8.44 ppm, ascribed to the triazole hydrogen. Part of a multiplet at 7.45 ppm, a multiplet at 6.66 ppm and a doublet of doublets at 6.50 ppm were all assigned to the aromatic hydrogens of the 4-bromo-3-methylphenol ring, with the remainder of the multiplet at 7.45 ppm, multiplets at 7.86 ppm and 7.75 ppm and a doublet at 7.59 ppm being ascribed to the aromatic hydrogens of the naphthalene ring. The hydrogens of the  $\text{CH}_2$  group linking the triazole and glutaric acid groups together were observed as a singlet at 5.16 ppm, whilst the hydrogens of the two  $\text{CH}_2$  groups connecting the phenol and amide groups were observed as two multiplets at 3.87 ppm and 3.50 ppm.  $^{13}\text{C}$  NMR showed thirty seven resolved signals. Three downfield signals at 173.5, 173.2 and 171.7 ppm were ascribed to the three carbonyl carbons, whilst a further downfield aromatic signal at 157.7 ppm was assigned to the aromatic carbon attached to the ether oxygen. Seventeen further aromatic signals were observed, with signals at 149.7 ppm and 125.7 ppm ascribed to the carbons of the triazole ring. The IR spectrum showed a signal at  $3290 \text{ cm}^{-1}$ , assigned as an N-H stretch. Three C=O stretches were observed at  $1720 \text{ cm}^{-1}$ ,  $1657 \text{ cm}^{-1}$  and  $1645 \text{ cm}^{-1}$ , with an aromatic C=C stretch observed at  $1500 \text{ cm}^{-1}$ . Finally, a high-resolution accurate mass of  $722.2825 [\text{M}]^+$  was obtained.

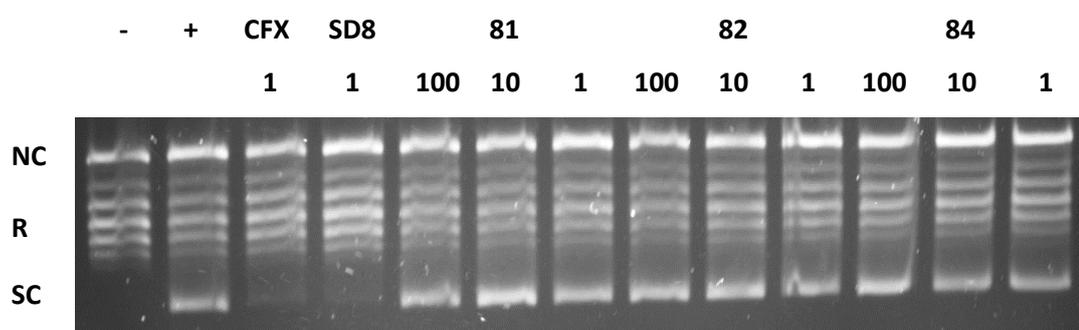


*Scheme 71. Synthesis of compound 86 by copper-catalysed click chemistry*

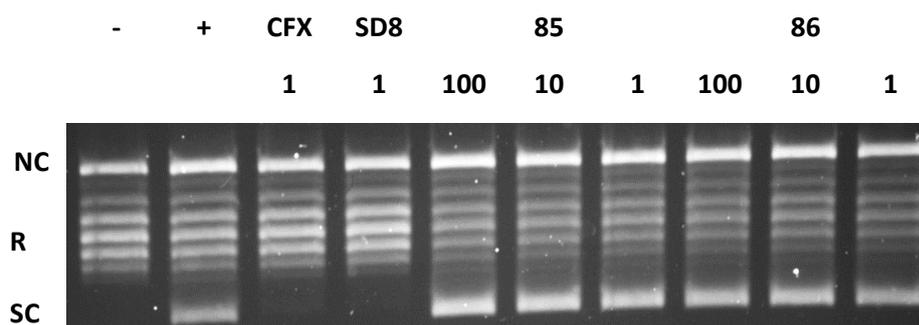
### 4.3. Biological assessment

#### 4.3.1. Assessment of compounds 81, 82, and 84-86

Following the successful synthesis of the final five “pseudo-natural products”, the inhibitory activity of the compounds was once again examined by *E. coli* gyrase supercoiling assay. As for compounds 52-55, these assays were carried out at 100  $\mu$ M, 10  $\mu$ M and 1  $\mu$ M final concentrations, with 1  $\mu$ M SD8 and ciprofloxacin once again used as comparators. The results are shown below in **Figure 60** and **Figure 61**. The controls once again showed a good response, with the negative control showing only relaxed and nicked circle DNA and the positive control a strong supercoiled band. Both SD8 and ciprofloxacin once again showed good inhibitory activity at 1  $\mu$ M as expected. However, compounds **81**, **82** and **84-86** showed no appreciable inhibition of supercoiling activity at any of the tested concentrations.

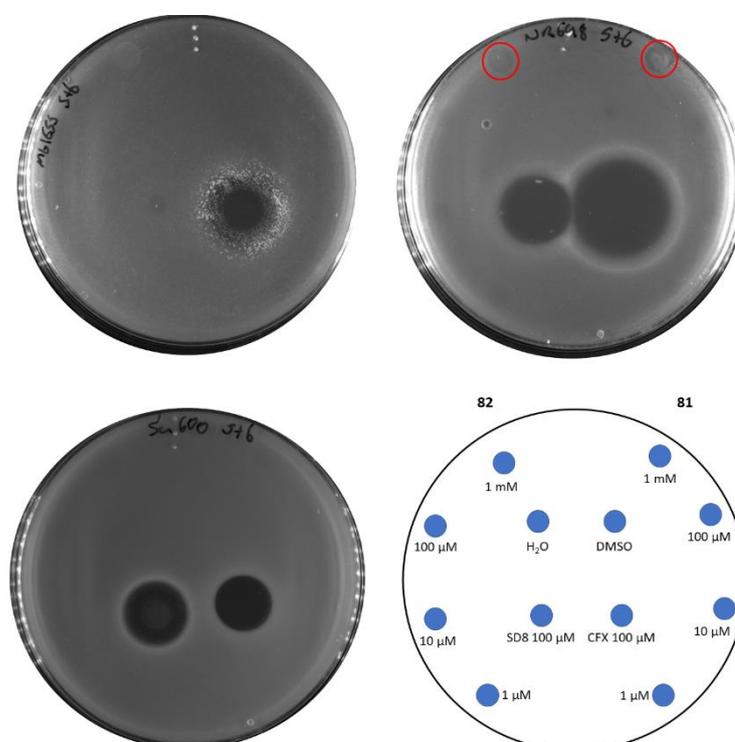


**Figure 60.** Effects of compounds **81**, **82** and **84** on DNA supercoiling by wild type *E. coli* gyrase. Relaxed pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and gyrase as a positive control. Concentrations are given in  $\mu$ M. CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.

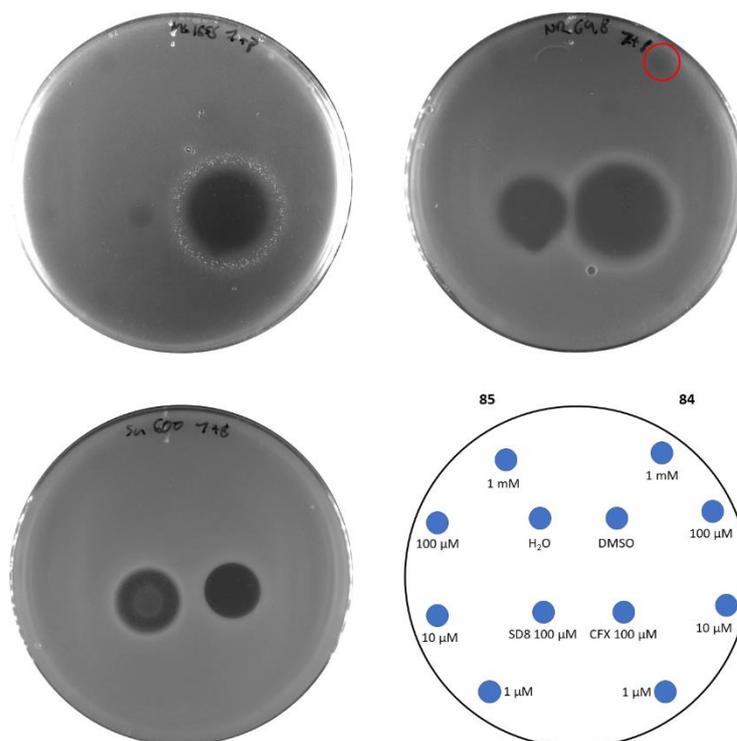


**Figure 61.** Effects of compounds **85** and **86** on DNA supercoiling by wild type *E. coli* gyrase. Relaxed pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and gyrase as a positive control. Concentrations are given in  $\mu$ M. CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.

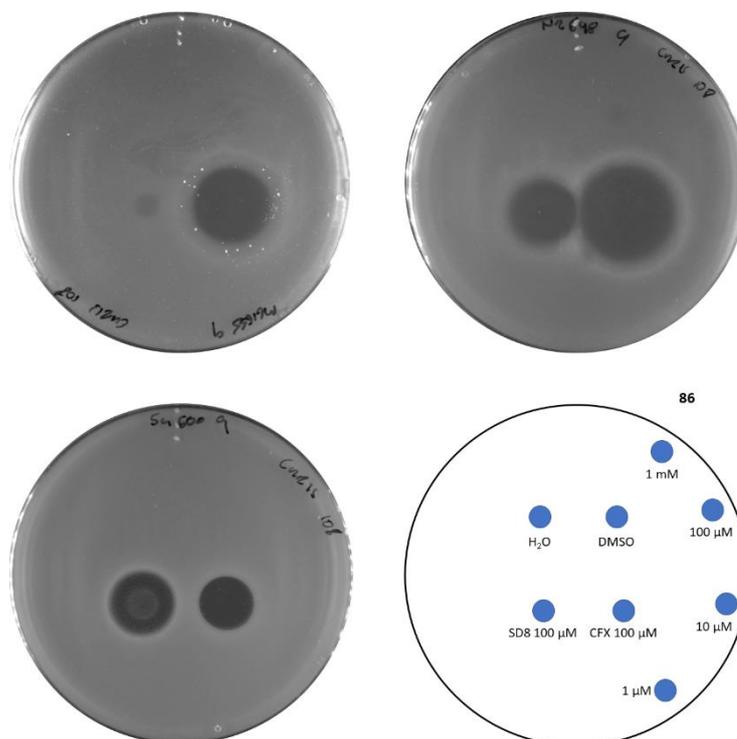
To assess if the compounds possessed any potency against bacteria outside of inhibition of DNA gyrase, the compounds were once again used in a zone-of-inhibition assay. As before, three strains were used: two *E.coli* strains, one with a weakened membrane, and one *S. aureus* strain. Compounds **81**, **82** and **84-86** were tested in 10-fold dilutions from 1 mM to 1  $\mu$ M, with controls of water, DMSO, ciprofloxacin (30  $\mu$ M) and SD8 (30  $\mu$ M) used as comparators. Images of the plates are shown below in **Figure 62**, **Figure 63** and **Figure 64**. As can be seen, both ciprofloxacin and SD8 controls show good growth inhibition in both the *S. aureus* and *E. coli* “leaky membrane” NR698 strains, with only ciprofloxacin inhibiting the growth of the *E. coli* wild type strain, as discussed before. In addition, water and DMSO had no observed growth inhibitory effects, as expected. Some potential minor inhibition of growth was observed in the areas of the *E. coli* NR698 plates treated with 1 mM **83**, **84** and **86**, suggesting that these compounds may possess some antibacterial activity, albeit at a concentration that would be clinically irrelevant. The lack of inhibition in the corresponding *E. coli* wild type plates potentially indicates that the compounds are unable to penetrate the cell membrane, and hence are unable to act against their target. Furthermore, based on the *E. coli* gyrase assays discussed above, any observed growth inhibition is not due to inhibition of DNA gyrase, acting through an alternative mode of action.



**Figure 62.** Zone-of-inhibition assay plates for compounds **81** and **82**. The top left plate contained an *E. coli* wild type strain; top right, *E. coli* NR698; bottom left, *S. aureus* wild type strain. Compound spotting pattern is shown in the bottom right. Suspected inhibition is highlighted in red circles.



**Figure 63.** Zone-of-inhibition assay plates for compounds **84** and **85**. The top left plate contained an *E. coli* wild type strain; top right, *E. coli* NR698; bottom left, *S. aureus* wild type strain. Compound spotting pattern is shown in the bottom right. Suspected inhibition is highlighted in red circles.

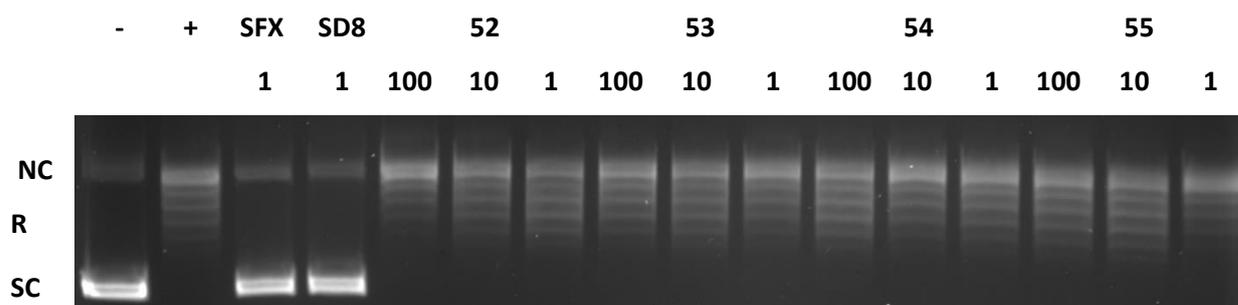


**Figure 64.** Zone-of-inhibition assay plates for compound **86**. The top left plate contained an *E. coli* wild type strain; top right, *E. coli* NR698; bottom left, *S. aureus* wild type strain. Compound spotting pattern is shown in the bottom right.

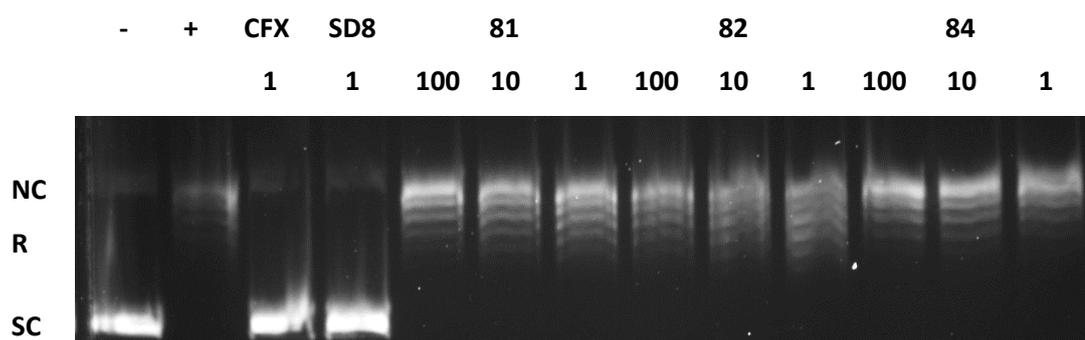
#### 4.3.2. Further examination of all tested compounds

It was therefore decided that compounds **52-55**, as well as compounds **81, 82** and **84-86** would undergo further biological examination to determine if the compounds had any activities missed in prior study. Since compounds **81, 82** and **84** had been observed to have a small growth inhibition effect against *E. coli* in the zone-of-inhibition assay, all compounds were first subjected to an *E. coli* topoisomerase IV (Topo IV) activity assay. Like DNA gyrase, Topo IV is a type II topoisomerase found in bacteria. However, unlike DNA gyrase, Topo IV is unable to introduce supercoiling, and serves to relax previously introduced supercoils and to decatenate DNA.<sup>239</sup> As a result, the assay used for DNA gyrase, where relaxed DNA is supercoiled, cannot be used for Topo IV. Instead, a DNA relaxation assay was used to determine the activity of the compounds. In a similar fashion to the supercoiling assay, a potential inhibitor at a known concentration was added to a mixture of *E. coli* Topo IV and supercoiled DNA, followed by incubation of the mixture at 37 °C for 30 minutes. The reaction is then arrested by addition of iso-amyl alcohol in chloroform (1:24) and STEB buffer, before separation of the topoisomers by gel electrophoresis and visualisation using ethidium bromide.

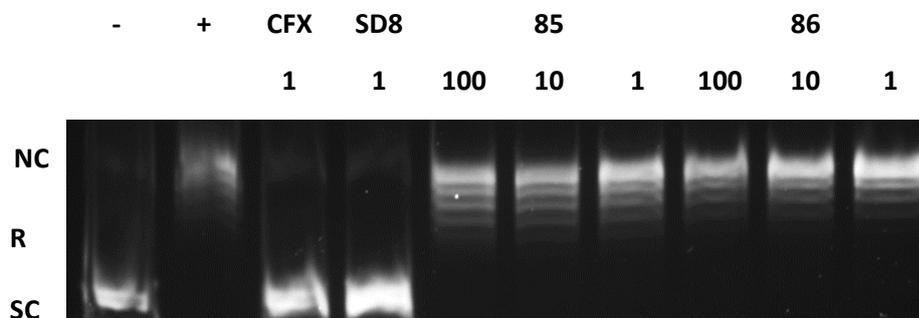
The results for compounds **52-55** are shown below in **Figure 65**, whilst results for compounds **81, 82** and **84** are shown in **Figure 66**. Finally, the results for compounds **85** and **86** are shown in **Figure 67**. In each of these assays, SD8 and sparfloxacin, another quinolone antibiotic, were used as comparitors. As can be seen, the controls worked well for each assay, with the negative controls showing only supercoiled DNA and the positive controls showing an absence of supercoiling. Both of the SD8 and sparfloxacin comparitors also worked well, with only nicked circle and supercoiled DNA present. However, none of the compounds possessed any inhibitory activity at the highest concentration tested, suggesting that any antibacterial activity observed against *E. coli* was not due to Topo IV inhibition.



**Figure 65.** Effects of compounds 52-55 on DNA supercoiling relaxation by wild type *E. coli* Topo IV. Supercoiled pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu\text{M}$ . SFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.



**Figure 66.** Effects of compounds 81, 82 and 84 on DNA supercoiling relaxation by wild type *E. coli* Topo IV. Supercoiled pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu\text{M}$ . CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.



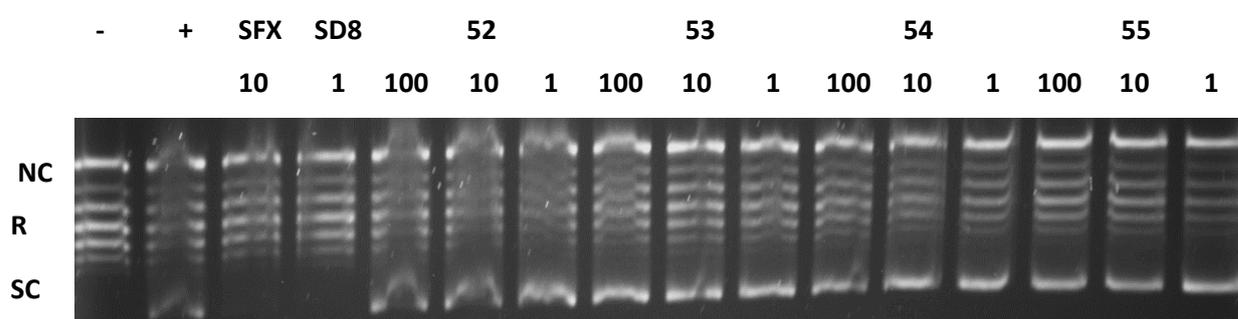
**Figure 67.** Effects of compounds 85 and 86 on DNA supercoiling relaxation by wild type *E. coli* Topo IV. Supercoiled pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu\text{M}$ . CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.

The zone-of-inhibition assays showed that, in some cases, the *E. coli* with a weakened membrane was more susceptible to growth inhibition than its wild-type counterpart due to its increased permeability. Therefore, one potential reason for a lack of inhibition of growth of the *S. aureus* wild-type strain in all zone-of-inhibition assays performed is that the compounds of interest were unable to enter the cell. To examine the effect of all nine compounds on *S. aureus*, a further DNA gyrase supercoiling assay was performed using enzyme from an *S. aureus* strain. These assays were performed in an almost identical

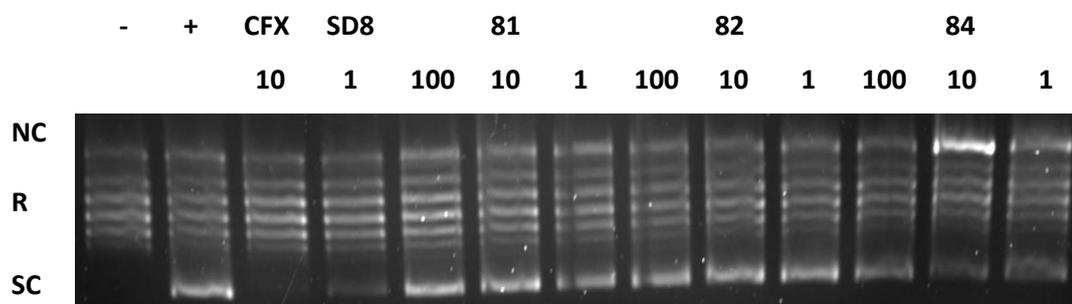
**CHAPTER 4: SYNTHESIS OF “PSEUDO-NATURAL PRODUCT” COMPOUNDS BASED ON SIMOCYCLINONE D8**

manner to the assays performed using *E. coli* gyrase, with the key difference taking place at the gel electrophoresis stage. Instead of loading the gel and running it immediately, as was the case for the *E. coli* gyrase assays, the loaded gels were allowed to sit in the running buffer for 30 minutes. This was done to allow the potassium glutamate present in the assay buffer to diffuse throughout the running buffer, a step which, if not carried out, causes streaking of the electrophoresis bands, making the assay results uninterpretable.

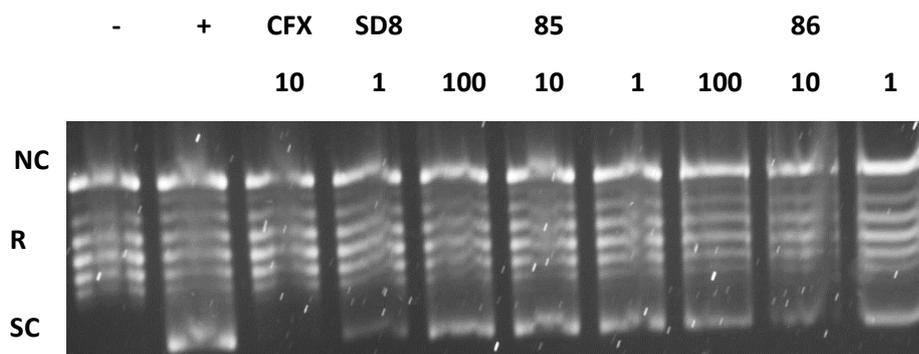
The results for compounds **52-55** are shown below in **Figure 68**, whilst results for compounds **81, 82** and **84** are shown in **Figure 69**. Finally, the results for compounds **85** and **86** are shown in **Figure 70**. In each of these assays, SD8 and ciprofloxacin were used as comparators. Due to *S. aureus* gyrase being less sensitive to ciprofloxacin when compared to its *E. coli* counterpart, a concentration of 10  $\mu\text{M}$  was used instead of the 1  $\mu\text{M}$  used in previous assays. As can be seen, the controls worked well for each assay, with the negative controls showing only relaxed and nicked circle DNA and the positive controls showing the presence of supercoiled DNA. Both of the SD8 and ciprofloxacin comparators also worked well, with little-to-no supercoiled DNA observed. However, none of the compounds possessed any inhibitory activity at the highest concentration tested.



**Figure 68.** Effects of compounds **52-55** on DNA supercoiling by wild type *S. aureus* gyrase. Relaxed pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and gyrase as a positive control. Concentrations are given in  $\mu\text{M}$ . CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.



**Figure 69.** Effects of compounds **81, 82** and **84** on DNA supercoiling by wild type *S. aureus* gyrase. Relaxed pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and gyrase as a positive control. Concentrations are given in  $\mu\text{M}$ . CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.



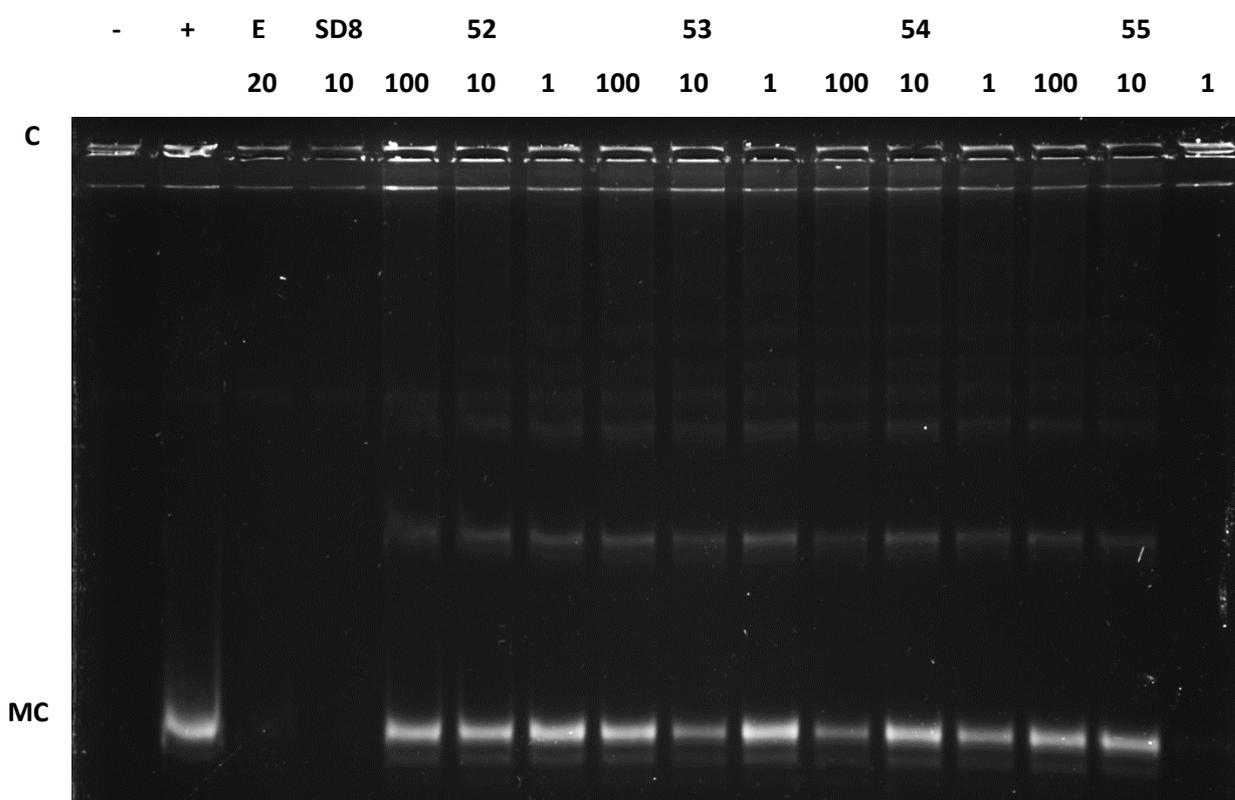
**Figure 70.** Effects of compounds **85** and **86** on DNA supercoiling by wild type *S. aureus* gyrase. Relaxed pBR322 plasmid DNA was used as a negative control, and an incubated sample of DNA and gyrase as a positive control. Concentrations are given in  $\mu\text{M}$ . CFX represents ciprofloxacin; NC, nicked circle DNA; R, relaxed DNA; SC, supercoiled DNA.

Finally, the activity of the nine compounds against human topoisomerase II (Topo II) was determined. Topo II is also a type II topoisomerase, and exists in two homologous isoforms, Topo II $\alpha$  and Topo II $\beta$ .<sup>240</sup> Topo II $\alpha$  is typically overexpressed in cells which are undergoing proliferation, serving as a marker for cell proliferation, whilst Topo II $\beta$  is distributed equally across all cells. Together, Topo II has been shown to be an effective target for the broad-spectrum treatment of cancer and is targeted by a variety of drugs including doxorubicin, daunorubicin and etoposide. Since Topo II $\alpha$  is overexpressed during cell proliferation, it is considered to be the isoform which is targeted by the anticancer therapies shown above.<sup>241</sup> SD8 has also been shown to inhibit Topo II *in vitro*, with SD8 thought to bind to the DNA-binding region of Topo II.<sup>242</sup> It was therefore hoped that compounds **52-55**, **81**, **82** and **84-86**, designed to mimic SD8, would similarly inhibit Topo II.

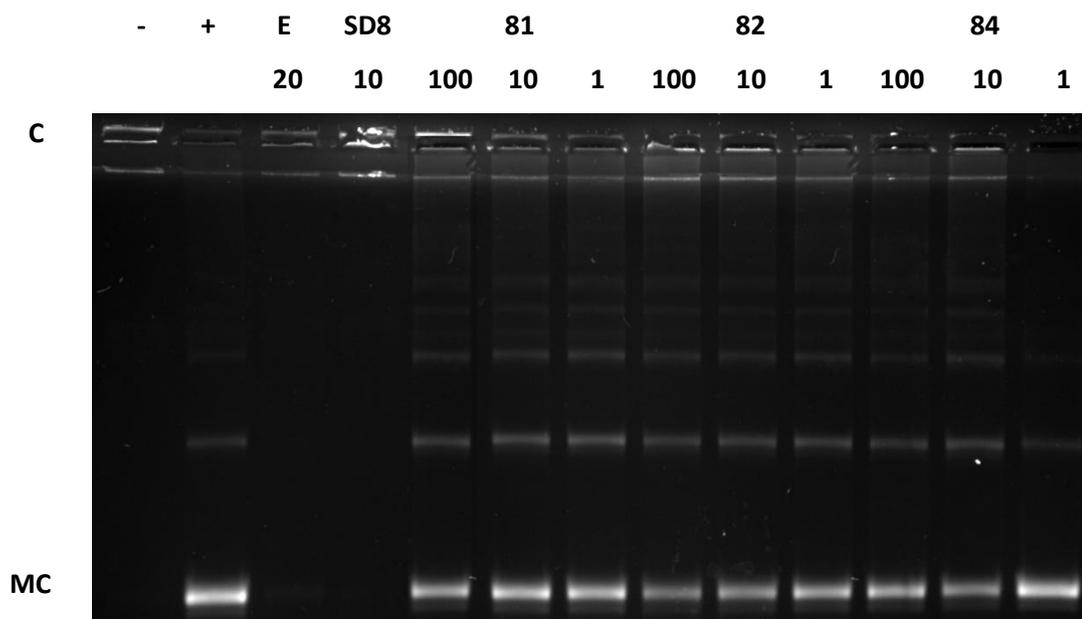
To assess the level of inhibition of Topo II, both Topo II $\alpha$  and Topo II $\beta$  were examined separately using decatenation assays. These assays used kinetoplast DNA (kDNA), consisting of two types of circular DNA, maxicircles and minicircles, which are catenated together. Whilst the decatenated DNA circles are susceptible to movement during gel electrophoresis, catenated kDNA is unable to move, remaining in the wells. Since both Topo II isoforms are capable of decatenating kDNA, inhibition of Topo II would result in a lack of decatenation and hence a reduced intensity of bands outside of the wells. In a similar fashion to the topoisomerase assays described previously, compounds of a known concentration were added to a mixture of kDNA and enzyme, and the mixture incubated at 37 °C for 30 minutes. The reaction was then arrested with iso-amyl alcohol in chloroform (1:24) and STEB buffer before the decatenated DNA was separated using gel electrophoresis and visualised using ethidium bromide.

**CHAPTER 4: SYNTHESIS OF “PSEUDO-NATURAL PRODUCT” COMPOUNDS BASED ON SIMOCYCLINONE D8**

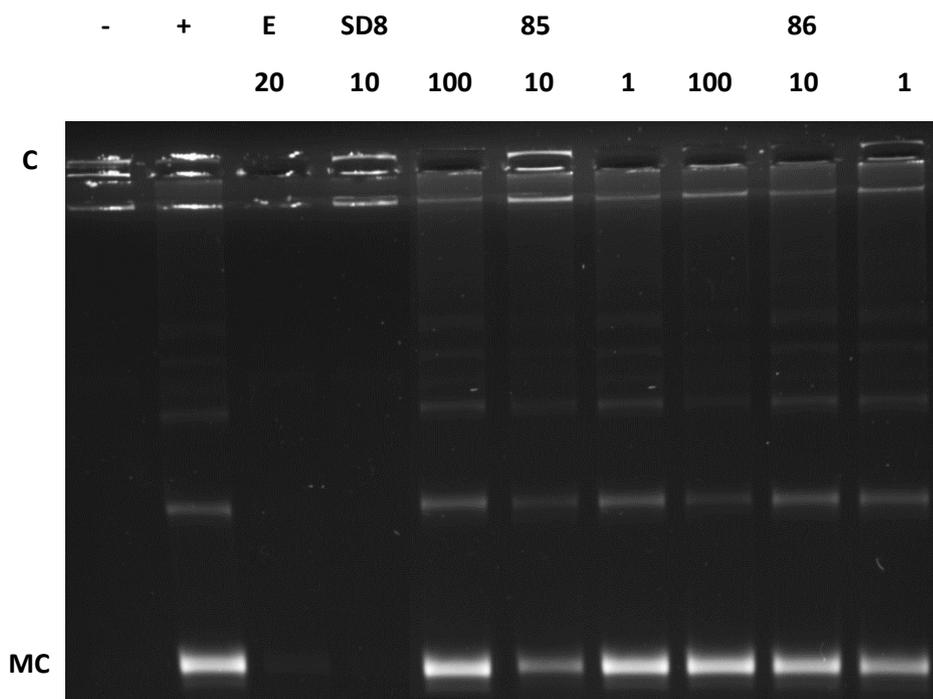
The results for compounds **52-55** against Topo II $\alpha$  are shown below in **Figure 71**, whilst results for compounds **81, 82** and **84** are shown in **Figure 72**. Finally, the results for compounds **85** and **86** are shown in **Figure 73**. Assays were carried out alongside 20  $\mu$ M etoposide, a known inhibitor of Topo II $\alpha$ , as well as SD8, to act as comparators. As expected, the controls gave a good response, with the negative control displaying no decatenated circles and the positive control showing a distinct decatenated band. Similarly, the known inhibitors worked well in each assay, with little-to-no decatenation observed. However, in the case of all tested compounds, no appreciable inhibition of decatenation activity was observed at any of the tested concentrations.



**Figure 71.** Effects of compounds **52-55** on kDNA decatenation by human Topo II $\alpha$ . kDNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu$ M. E represents etoposide; C, catenated kDNA; MC, decatenated minicircles.



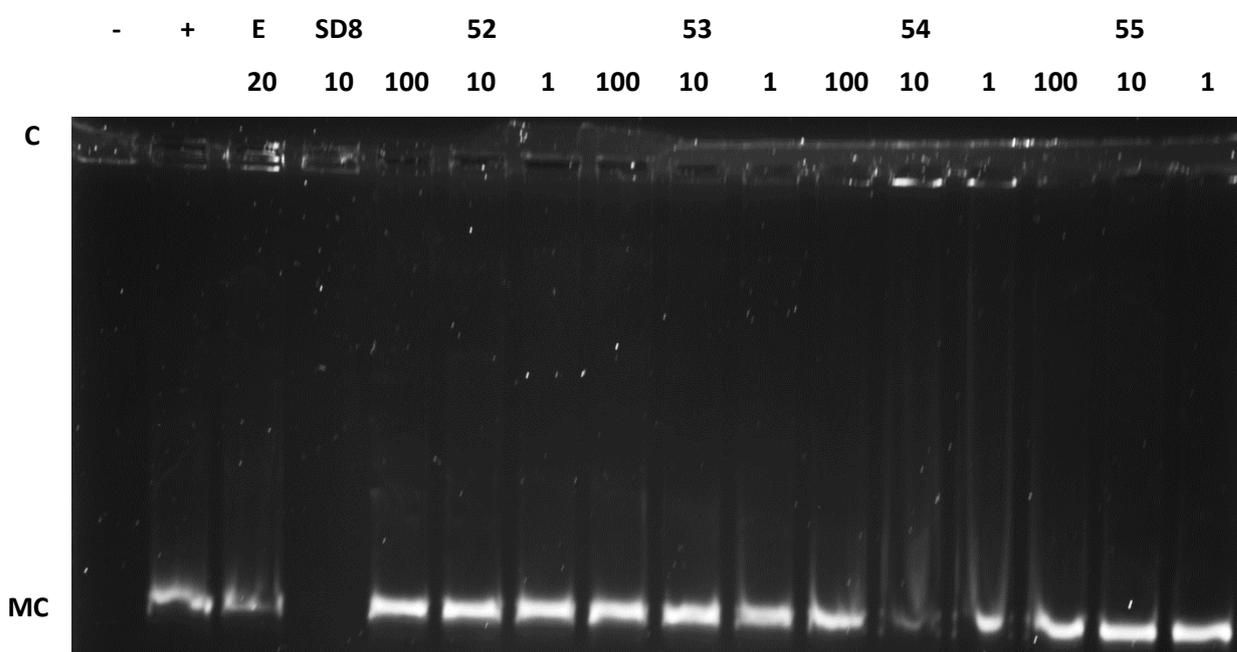
**Figure 72.** Effects of compounds **81**, **82** and **84** on kDNA decatenation by human Topo II $\alpha$ . kDNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu$ M. E represents etoposide; C, catenated kDNA; MC, decatenated minicircles.



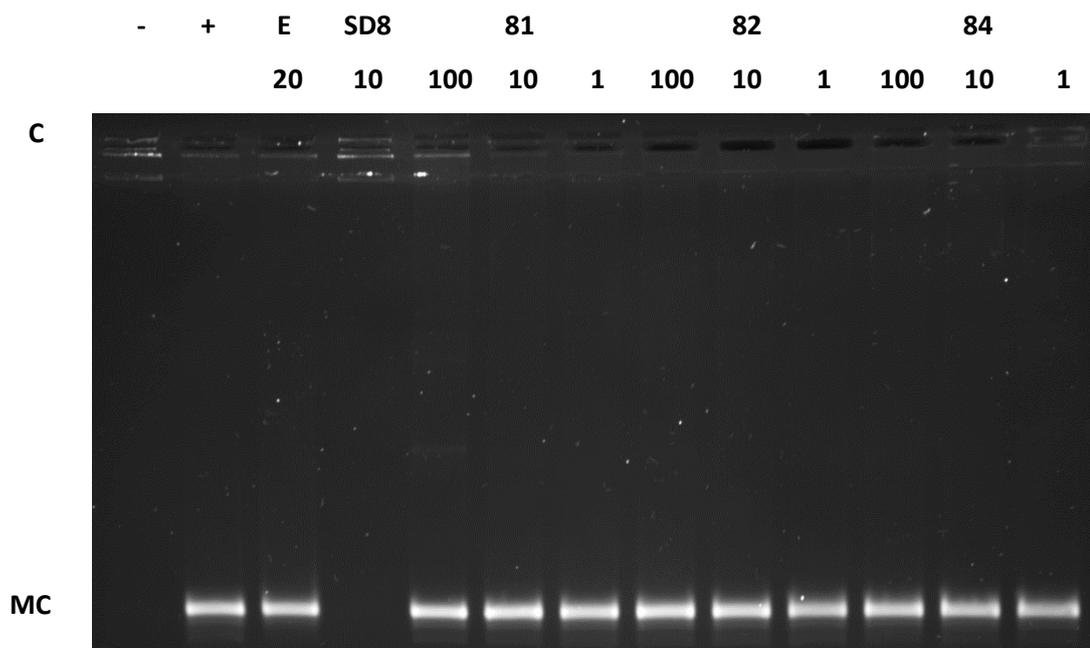
**Figure 73.** Effects of compounds **85** and **86** on kDNA decatenation by human Topo II $\alpha$ . kDNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu$ M. E represents etoposide; C, catenated kDNA; MC, decatenated minicircles.

**CHAPTER 4: SYNTHESIS OF “PSEUDO-NATURAL PRODUCT” COMPOUNDS BASED ON SIMOCYCLINONE D8**

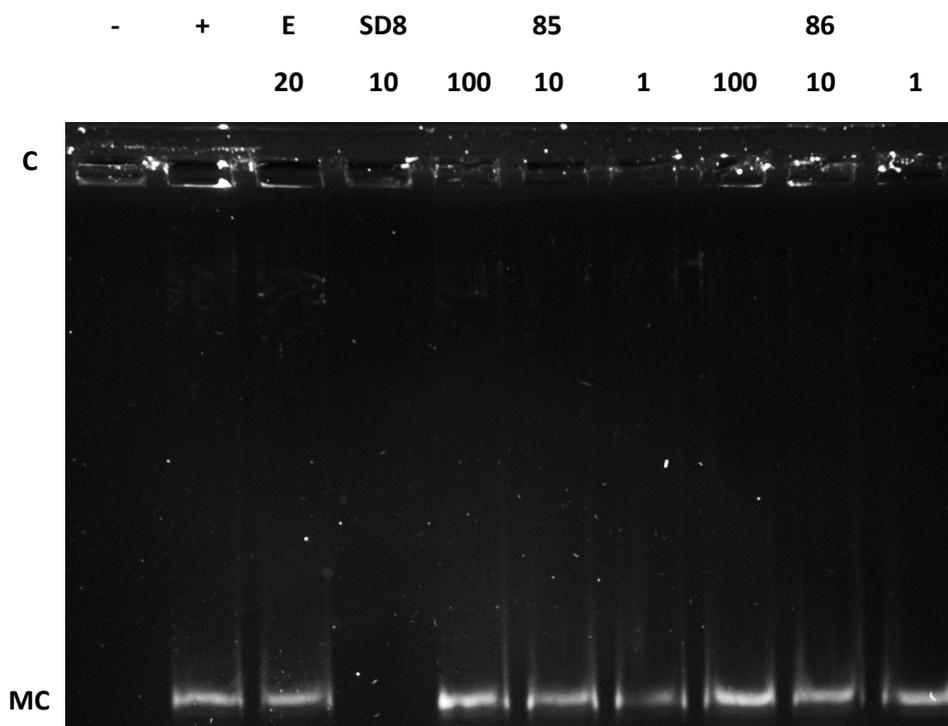
The results for compounds **52-55** against Topo II $\beta$  are shown below in **Figure 74**, whilst results for compounds **81, 82** and **84** are shown in **Figure 75**. Finally, the results for compounds **85** and **86** are shown in **Figure 76**. Assays were once again carried out alongside 20  $\mu$ M etoposide as well as SD8 to act as comparators. As expected, the controls gave a good response, with the negative control displaying no decatenated circles and the positive control showing a distinct decatenated band. Similarly, SD8 worked well in each assay, with little-to-no decatenation observed. However, the etoposide comparator failed to inhibit Topo II $\beta$  in any assay, suggesting that etoposide’s cytostatic effect is solely due to inhibition of Topo II $\alpha$ . Unfortunately, the tested compounds were also ineffective at inhibiting decatenation activity of Topo II $\beta$  at any concentration tested.



**Figure 74.** Effects of compounds **52-55** on kDNA decatenation by human Topo II $\beta$ . kDNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu$ M. E represents etoposide; C, catenated kDNA; MC, decatenated minicircles.



**Figure 75.** Effects of compounds **81**, **82** and **84** on kDNA decatenation by human Topo II $\beta$ . kDNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu$ M. E represents etoposide; C, catenated kDNA; MC, decatenated minicircles.



**Figure 76.** Effects of compounds **85** and **86** on kDNA decatenation by human Topo II $\beta$ . kDNA was used as a negative control, and an incubated sample of DNA and enzyme as a positive control. Concentrations are given in  $\mu$ M. E represents etoposide; C, catenated kDNA; MC, decatenated minicircles.

#### 4.4. Conclusions

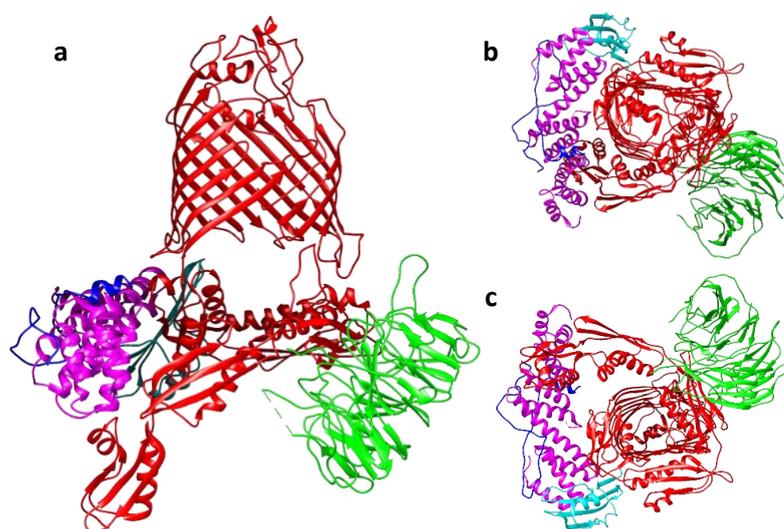
The work presented in this chapter aimed to further assess the ability of the computational strategy used in **Chapter 2** to design suitable “pseudo-natural products” based on DNA gyrase by synthesis of the selected compounds and biological testing of the compounds by supercoiling, relaxation and decatenation assays. Continuing from the work performed in **Chapter 3**, the five amine fragments selected for study were successfully generated and attached to the decanoic acid linker. Furthermore, these five amine-linker compounds were then successfully attached to the highest scoring alkyne in the computational work performed in **Chapter 1**, compound **9**, by CuAAC click chemistry. However, despite a potentially promising inhibition of growth of a “leaky membrane” *E. coli* strain observed in zone-of-inhibition assays for compounds **81**, **82** and **84**, both the five compounds generated in this chapter and the four compounds generated in **Chapter 3** failed to inhibit the activities of *E. coli* or *S. aureus* DNA gyrase, *E. coli* Topo IV, or human Topo II.

Improving upon the molecules generated would require more detailed information regarding how the molecules fit and bind with DNA gyrase *in vitro*. A potential methodology for examining this would be obtaining crystal structures of DNA gyrase with the compounds using crystal soaking techniques. Further examinations could also look at the fragments directly, such as using ITC to determine if the alkynes or amines possess binding affinity for DNA gyrase. Once again, the flexibility of the linker could cause issues, allowing the two bulky groups to move more freely than in the original natural product. Improving rigidity may help with this problem, but may prove more synthetically challenging.

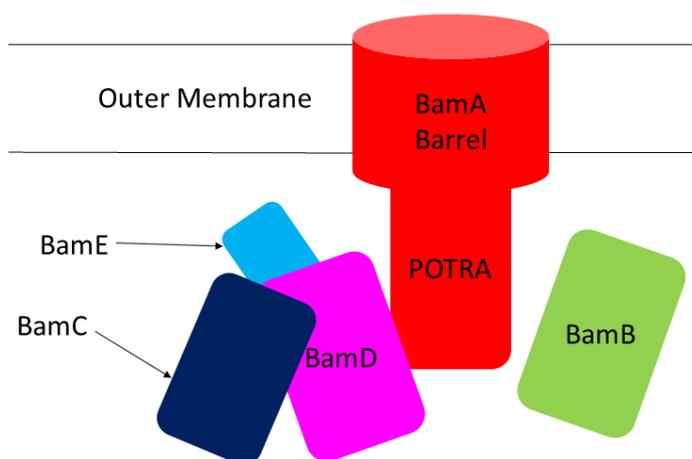
**CHAPTER 5: SYNTHESIS OF POTENTIAL  
BAMA/BAMD INTERACTION INHIBITORS**

### 5.1. The BAM complex

The BAM complex allows for the formation of beta-barrel-folded proteins, that is, protein beta-sheets which are folded into a tube-like shape where the two terminal beta-strands are non-covalently linked. It is also responsible for insertion into the outer membrane, and forms a top hat-shaped structure. This process must be regulated by the complex, as, if the proteins were inserted into the membrane in an unfolded state, a large number of hydrogen bond donors and acceptors would be unsatisfied. Conversely, if the proteins were folded before being inserted, the resulting beta-barrels would be “inside-out”, as the barrel would need to be pushed through itself to be inserted.<sup>243</sup> The structure of the complete BAM complex, as found in *E. coli*, is shown in **Figure 77**. A cartoon schematic of the complex is also given in **Figure 78**.



**Figure 77.** Structure of the *E. coli* BAM complex as viewed from the membrane plane (a), the extracellular side (b) and the periplasmic side (c). BamA (red), BamB (green), BamC (blue), BamD (magenta), BamE (cyan). PDB ID: 5D00



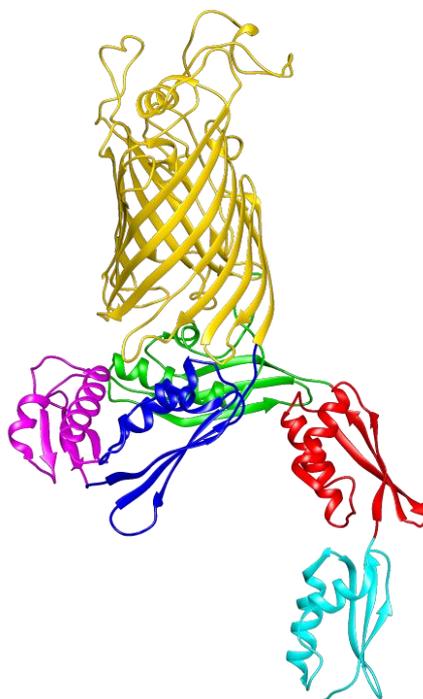
**Figure 78.** A cartoon schematic of the BAM complex. Abbreviations: POTRA (polypeptide translocation-associated domains).

The BAM complex is essential for numerous biological functions within bacterial cells.  $\beta$ -barrels inserted within the outer membrane allow the transport of nutrients into the cell, as well as allowing the secretion of waste products which would otherwise increase in concentration within the cell and become toxic.<sup>244,245</sup> In addition,  $\beta$ -barrels act as an entry point for proteins that act as biosynthetic precursors within the cell.<sup>244</sup> Hence, inhibition of this complex could potentially lead to cell death and give rise to new classes of antibiotic compounds. The BAM complex consists of a protein, BamA (also known as YaeT), whose C-terminal barrel is situated within the outer membrane and whose N-terminal is located within the periplasm, which is surrounded by up to four lipoprotein subunits, BamB-E (also known as YfgL, NlpB, YfiO and SmpA respectively), all of which are found in the periplasm and are anchored to the outer membrane by a lipid anchor.<sup>246,247</sup> For maximum efficiency of the complex, all five subunits must be present,<sup>243,248–250</sup> however BamA and BamD are the only essential subunits, with one copy of each being required to form the BAM complex.<sup>246,251,252</sup> In addition to this, the removal of any of BamBCE appears to cause the cell to become hypersensitive to certain antibiotics.<sup>253</sup>

### 5.1.1. BamA

The 16  $\beta$ -strand barrel domain of BamA is nearly empty, forming a volume of roughly  $13,000 \text{ \AA}^3$ , with extracellular loops creating a dome-shaped barrier between the inside of the barrel and extra-cellular space.<sup>247</sup> The N-terminal domain of BamA forms a series of five polypeptide translocation-associated (POTRA) domains arranged in a fishhook-like shape, with each successive domain situated next to its predecessor in a right-handed rotation. Each POTRA domain consists of a three-stranded  $\beta$ -sheet with two antiparallel

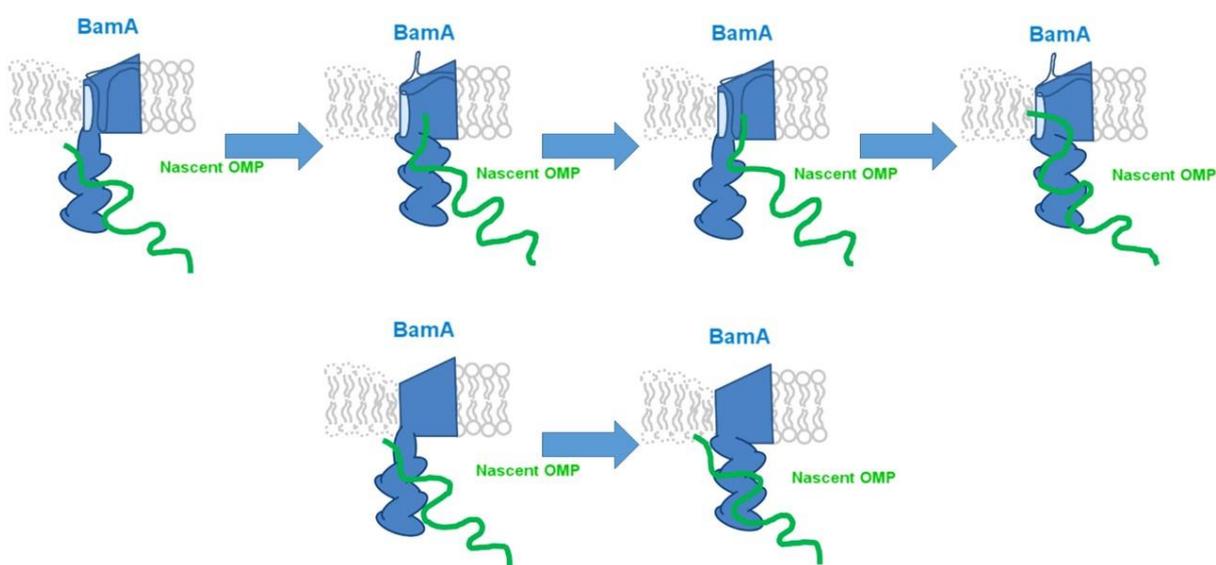
helices arranged in the order  $\beta$ - $\alpha$ - $\alpha$ - $\beta$ - $\beta$ , i.e. the first and second  $\beta$ -strands form the edges of the sheet, with the third strand and the helices sandwiched between them. However, the domains possess low sequence similarity, with the only conserved residues being found in the hydrophobic core or the loops, indicating that these regions are structurally important. In addition, by utilising deletion constructs each lacking one of the POTRA domains, it has been discovered that POTRA 1 and POTRA 2 can be deleted whilst still allowing partial functionality of BamA. POTRA domains 3-5, however, appear to be necessary for the function of BamA, as mutants with these domains deleted resulted in cell death.<sup>254</sup> The complete structure of BamA is shown below in **Figure 79**.



**Figure 79.** Structure of BamA showing the  $\beta$ -barrel (gold) and POTRA domains 1-5 (cyan, red, green, purple, blue respectively). PDB ID: 5D00.

The conformations of the POTRA domains appear to alter the structure of the conformation of the C-terminal barrel, seemingly forming a gating mechanism for the interior of the barrel. For example, in studies conducted on BamA crystals containing all five POTRA domains, the N-terminal domain interacts with several periplasmic loops within the barrel stabilising the closed conformation, whilst if the first three POTRA domains are removed there are no observed interactions. This closed conformation, however, destabilises the C-terminal  $\beta$ -strand due to the strand becoming twisted and “tucked into” the barrel itself, as well as modifying the conformation of one of the extracellular loops, allowing access from the interior of BamA to the lipid phase of the outer membrane between strands 1 and 16, which form a lateral opening, potentially highlighting a

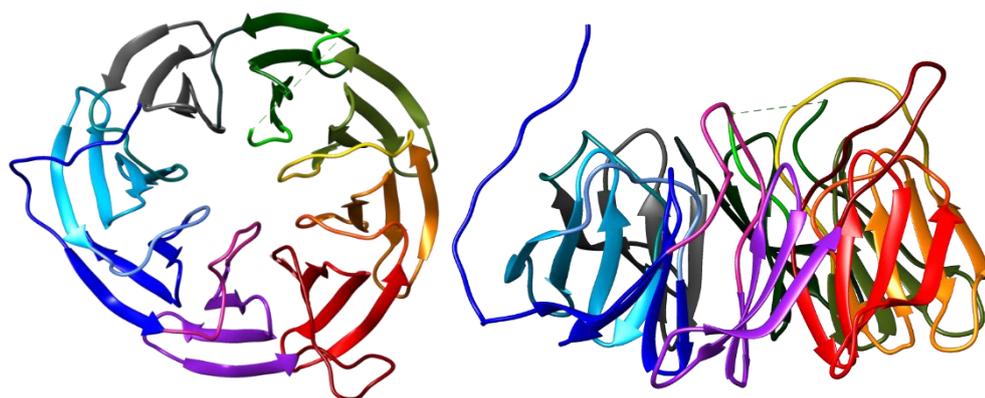
mechanism for outer membrane protein insertion. Moreover, a notable decrease in the thickness of the outer membrane close to the C-terminal strand and a decrease in lipid order in this region also suggests that the  $\beta$ -barrel also primes the membrane for the insertion. An alternative mechanism has also been proposed for simpler OMPs, in which the membrane proteins are directly inserted into the destabilised membrane by interactions with the POTRA domains of the BamA complex.<sup>247</sup> Both mechanisms are shown below in **Figure 80**. In both these contexts, it appears that the POTRA 3 domain is the binding site for proteins, whilst the POTRA 5 domain is required for interactions with BamB and BamD.<sup>254–256</sup>



**Figure 80.** Proposed mechanisms of OMP insertion. Top: Lateral opening of barrel and folding using POTRA domain and extracellular loop. Bottom: Direct insertion into lipid membrane using POTRA domain.<sup>247</sup>

### 5.1.2. BamB

BamB, whilst not being necessary for the BAM complex to function, appears to play an important role in the assembly of OMPs, with activity of the complex BamACDE being much reduced when compared to the five-protein complex.<sup>243</sup> The protein takes on the shape of a  $\beta$ -propeller with eight blades, that is, eight blade-shaped  $\beta$ -sheets which are arranged in a toroid shape, and resembles a flattened cylinder with a shallow groove on the top side and a deep opening on the bottom side. Each blade consists of four anti-parallel  $\beta$ -strands, and the eighth blade is formed by  $\beta$ -strands of both the N- and C-terminus, a feature observed in other eight-bladed  $\beta$ -propellers.<sup>257–259</sup> The N-terminus of Bam B, located to the top of the  $\beta$ -propeller, is attached to the cell membrane by a palmitate lipid site.<sup>252</sup> The chain then joins the  $\beta$ -propeller in the last strand of blade eight before continuing onto the first strand of blade one.<sup>257,260</sup> The structure of BamB is shown below in **Figure 81**.



**Figure 81.** The top-down (left) and edge-on (right) structures of BamB. Blade 1 (purple), blade 2 (red), blade 3 (orange), blade 4 (olive), blade 5 (dark green), blade 6 (grey), blade 7 (cyan), blade 8 (blue). PDB ID: 3PRW.

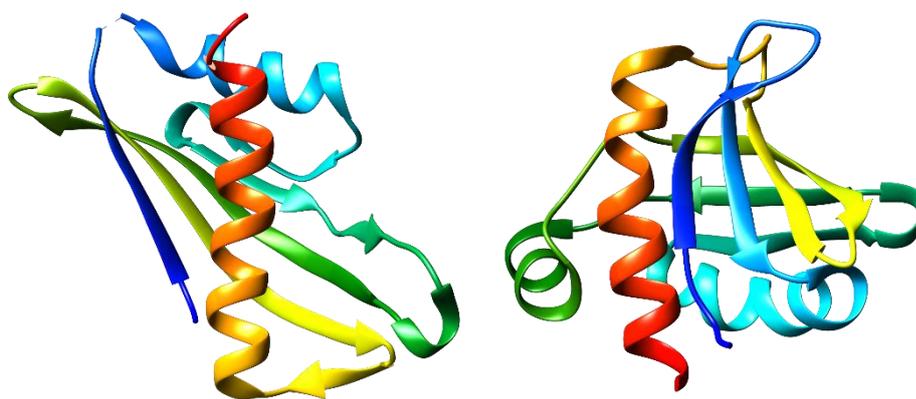
The protein binds to BamA utilising two separate peptide segments, both of which border the central groove on the top face of the propeller.<sup>257,260</sup> The first section, containing residues L192, L194 and R195, is part of a loop linking blades 3 and 4 which appears to be fairly flexible. The residues that interact with BamA are spatially close to the second binding region, defined by residues D246 and D248, located on another loop linking blades 4 and 5. These two segments make up the complete binding site for BamA, which appears to bind using an induced-fit mechanism due to the flexibility of the loops on BamB, both those directly involved in binding and those in the surrounding area.<sup>257</sup> The binding sites for BamB appear to be any of the POTRA domains of BamA aside from POTRA 1, as the deletion of any of POTRA 2-5 resulted in the loss of BamB.<sup>254</sup> Also notable is the region including residues V319 to H328, which, whilst not participating in any known interactions with other components or OMPs, is important for the structure of BamB.<sup>260</sup>

An insight into the function of BamB can be found in the form of an F104 residue, found situated in a deep hydrophobic pocket of its neighbouring peptide when a BamB lattice is crystallised. This residue forms part of another loop that connects blades 1 and 2, and extends a fair distance from the core of the propeller. The hydrophobic pocket is lined with residues A203, A206, V208, V219, I227 and L272, with additional hydrophobic and van der Waals interactions between N225 and I227, and W103 and L220. These interactions appear to indicate that the BamB lipoprotein is capable of recognising aromatic peptide residues, and the orientation of the hydrophobic pockets seemingly presents an ideal site for the binding of peptide  $\beta$ -sheets, allowing for BamB to assist in the guidance of unfolded OMPs to BamA for folding and insertion.<sup>257</sup>

### 5.1.3. BamC

The BamC complex, contrary to BamB's interactions, does not directly interact with the POTRA domains of BamA, with the deletion of any of the POTRA1-4 domains having no effect on the presence of BamC.<sup>254</sup> Instead, the protein binds to BamD, which is itself dependent on the presence of POTRA5, explaining BamC's dependence on this domain.<sup>254,256</sup> Further discussion both of this interaction and of BamD will be covered later in **Section 5.1.5**.

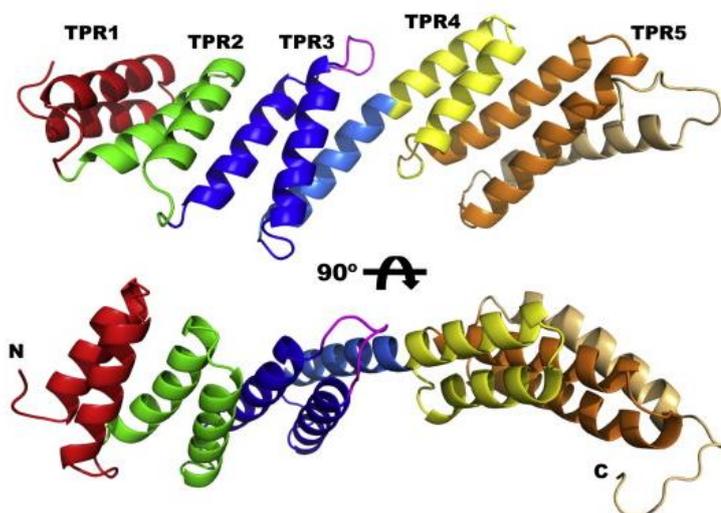
NOESY NMR studies of BamC reveal the presence of two defined domains, consisting of the N-terminal domain (residues 101-210) and the C-terminal domain (residues 229-346) with a flexible linker domain (residues 211-228). Further to this, there does not appear to be any evidence that these domains interacted with one another, and so these domains are treated as isolated regions (**Figure 82**). Residues 1-100 do not appear to be structured in any discernible way, with the first 24 residues forming the periplasmic localisation palmitate lipid anchor and the 25<sup>th</sup> residue, typically a cysteine residue which is modified to include this lipid anchor. The remaining 75 residues appear to form a structure which serves only to connect the ordered regions to the lipoprotein.<sup>261,262</sup> The N-terminal domain of BamC contains two  $\alpha$ -helices which are packed against a  $\beta$ -sheet consisting of five  $\beta$ -strands, a structure similar to the helix-grip fold found in major latex proteins, a protein family found only in plants associated with pathogen defence.<sup>261,263</sup> The C-terminal domain is also made up of a similar motif, with the addition of a short additional  $\beta$ -strand before  $\beta$ 1 and a further seven-residue helix located between the third and fourth  $\beta$ -strands.<sup>261</sup>



*Figure 82. The C-terminal (left) and N-terminal (right) regions of BamC. Colours proceed from red to blue as distance to the N-terminus decreases. PDB IDs: 2LAE and 2LAF respectively*

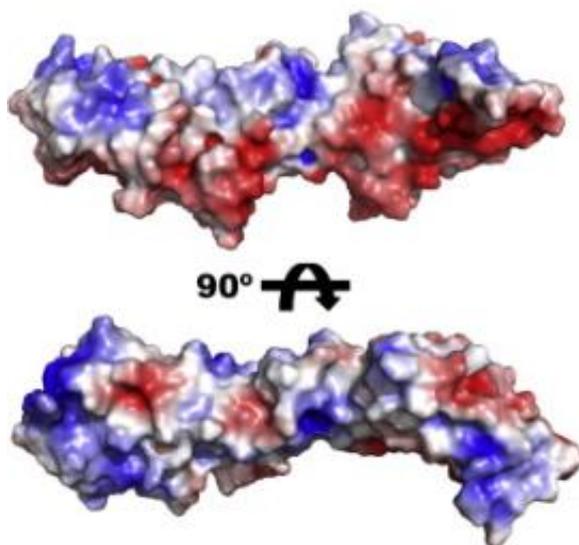
#### 5.1.4. BamD

As stated above, BamD is the only lipoprotein required for BAM functionality,<sup>246,251,252</sup> and is also responsible for incorporating both BamC and BamE into the BAM complex.<sup>254,256</sup> The lipid-free structure (residues 23-280) is made up of five tetratricopeptide repeats (TPR), a type of sequence formed from 34 residues folded into a helix-turn-helix formation which are typically arranged so that the two anti-parallel helices from one repeat pack alongside the first helix of the next repeat.<sup>264-266</sup> The TPR array is often terminated with a capping helix, resulting in a right-handed superhelix, that is, a clockwise helix structure formed from the individual helical protein structures.<sup>267</sup> However, in the case of BamD, this does not occur. Instead, three N-terminal TPRs are capped with a helix which then leads directly into the final two, C-terminal TPRs, resulting in the C-terminal region being offset from the N-terminal region and the elongation of the molecule. In addition to this, the loop connecting the two helices in TPR3 is extended compared to the other connecting loops, which in general are short and typical.<sup>264</sup> The overall arrangement of BamD is relatively similar to that of the POTRA domains of BamA, suggesting that BamD may function as a protein-protein interaction scaffold, both for the components of the BAM complex and the target OMPs.<sup>268</sup> The structure of BamD is shown in **Figure 83**.



**Figure 83.** Crystal structure of BamD. The N-terminal TPR1-3 (red, green, blue respectively) region is capped with a helix (light blue) which then leads directly into TPR4-5 (yellow, orange respectively). The extended TPR3 loop is also shown in magenta.<sup>264</sup>

Mapping the electrostatic surface potential of BamD (**Figure 84**) shows that the region containing the links between the helices in the TPRs is positively charged, whilst the opposite side, linking one TPR to another, is negatively charged. In addition, the offset between the N- and C-terminal regions results in the formation of a negative semi-circular crease close to the underside of BamD.<sup>264</sup> The C-terminal region also appears to be largely negatively charged, however the final helix of TPR5 has an amphoteric charge distribution, i.e. the positively charged residues are located on one side whilst the other side is composed of negatively charged residues from both helices of TPR5. It has therefore been suggested that the C-terminal region, especially  $\alpha 10$ , interacts with the positive region of the BamA POTRA5 domain to link itself to the complex. However, even when  $\alpha 10$  is deleted the BAM complex is still able to function, suggesting that the active component of BamD is found in the N-terminal region.<sup>268</sup> This hypothesis is supported further by the high level of conservation between *E. coli* BamD and its homologues in several other strains of Gram-negative bacteria, with most of these conserved residues found on the surface of the protein.<sup>264,268</sup>



**Figure 84.** Electrostatic potential surface of BamD. The orientation is the same as that found in Figure 5. The negative (red) semi-circular groove can be seen, as well as the negative regions caused by the linking loops. Blue indicates positive regions, whilst white regions indicate neutrality.<sup>264</sup>

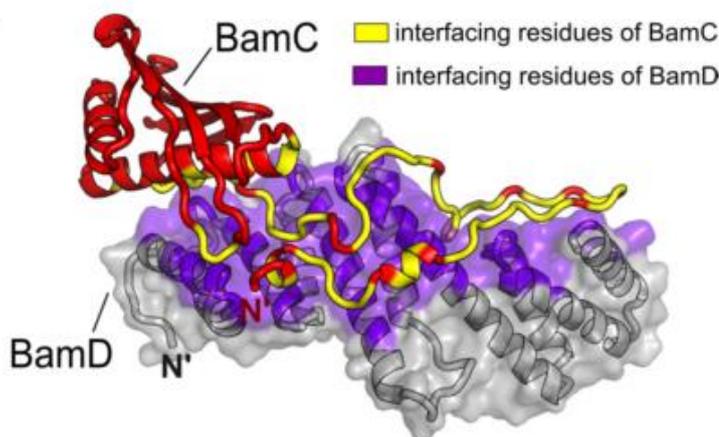
The structure of BamD is very similar to several proteins which are involved in recognising and binding extended polypeptides, potentially suggesting a role for BamD in the folding and insertion of OMPs. The right-handed twist of BamD forms a concave surface lined by the first helix of each TPR, similar to that found in proteins such as the heat shock protein (Hsp)-organising protein, Tom70, a component of a complex designed to allow proteins to move through the outer membrane of mitochondria, and FKBP52, a member of a family of

proteins which plays a role in immunoregulation. All of these complexes use their concave surface to recruit various molecular chaperones by binding their C-terminal tails into an extended conformation, i.e. binding not using  $\alpha$ -helices but using a  $\beta$ -strand structure.<sup>264,269–271</sup> The role of BamD, therefore, may be to recognise the C-terminal sequence which is conserved in OMPs across a wide range of species, and it is hypothesised that this interaction could take place in a pocket formed by TPR1 and TPR2.<sup>264,272</sup> Due to the conservation, it is likely that this C-terminal region, and therefore BamD, is important for at least one of the steps involved in the folding and insertion of the OMPs into the outer membrane.<sup>264</sup>

Comparison of the N-terminal region of BamD with those found in chaperones of the type three secretion system, an appendage which, in pathogenic Gram-negative bacteria, allows detection of eukaryotic organisms and the secretion of proteins to aid in infection, shows that these two structures are very structurally similar.<sup>264</sup> These type three secretion system chaperones bind to short chain length peptides utilising the concave face of their TPR tandem structures.<sup>273–275</sup> This suggests that BamD may bind unfolded OMPs by forming an extended conformation prior to folding and insertion of the OMPs into the outer membrane.<sup>264</sup>

### 5.1.5. BamCD Complex

The BamCD complex is formed, somewhat surprisingly, by the utilisation of the unstructured region on BamC in addition to the N-terminal domain (i.e. residues 26-217). This unstructured region forms a U-shaped structure which appears to fit into a trail of creases that run longitudinally around BamD, allowing extensive interactions to take place. The N-terminal region of BamC that immediately follows the unstructured region also sits adjacent to the N-terminal region of BamD, suggesting that both the C- and N-terminal regions of BamD are important for BamC binding. Furthermore, when the BamCD complex is formed BamD undergoes a significant conformational change localised primarily in the C-terminal region (the final helix of TPR3 and both helices in TPR4 and TPR5) when compared to its monomeric form.<sup>276</sup> The BamCD interface is illustrated in **Figure 85**.



*Figure 85. The BamCD complex, including an indication of interfacing surfaces of both components. The long, unstructured region of BamC is used extensively in the binding to the negative surface of BamD.<sup>276</sup>*

At first glance, BamC appears to merely stabilise the multi-component BAM complex, possessing no other known interactions than those linking it to BamD.<sup>261</sup> However, upon closer inspection and comparison to the suggested interaction domains of BamD, it may be possible that BamC takes part in some form of binding regulation, blocking or exposing the binding site depending on requirements. For instance, the pocket in which BamD is thought to bind the C-terminal region of unfolded OMPs described above is occupied in the BamCD complex by part of the U-shaped N-terminal region of BamC.<sup>30,38,42</sup>

#### 5.1.6. BamE

BamE appears to be the simplest of the lipoproteins in terms of structure, whose core consists solely of a single anti-parallel  $\beta$ -sheet consisting of three  $\beta$ -strands packed against a pair of  $\alpha$ -helices in the sequence  $\alpha 2$ - $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\beta 3$ . The linking region between  $\alpha 2$  and  $\beta 1$  also appears to adopt a helical-like turn, whilst the N-terminal region, situated at the end of  $\alpha 1$ , and the C-terminal region, located at  $\beta 3$ , both appear to be highly disordered.<sup>277,278</sup> The structure of BamE is shown in **Figure 86**. Unfortunately, due to the current lack of data regarding the functionality of BamE, its function can only currently be speculated upon by comparison with other BAM complex components. BamB utilises its  $\beta$ -strands to recognise OMPs via  $\beta$ -augmentation, a phenomenon which involves the extension of the binding  $\beta$ -sheet by addition of a further  $\beta$ -sheet provided by the binding protein.<sup>279</sup> The  $\beta$ -strands in BamE also appear to be well suited to experience  $\beta$ -augmentation, suggesting that BamE may have a similar function.<sup>251</sup>



**Figure 86.** A crystal structure of the BamE lipoprotein. Colours change progressively from the N-terminal (blue) to the C-terminal (red) region. PDB ID: 5WAM.

Coordination of BamE with BamD was seen to induce line broadening within an NMR experiment, as well as perturbation of the observed chemical shift at specific residues indicating that the BamDE complex is strongly bound together. By utilising alanine scanning, it was discovered that the residues important for structural viability of BamE were also the residues responsible for binding to BamD. These residues are found in the disordered N-terminal region (R29, I32, Y37, L38), the loop linking  $\alpha 2$  and  $\beta 1$  (T61, L63, F68, N71) and the three  $\beta$  strands (T72, W73, R78, T92, F95).<sup>268,277,278</sup> Electrostatic studies of BamE suggest that there is a large region of negative potential which overlaps with these important residues, which, combined with the amphoteric charge distribution around BamD's  $\alpha 10$  region, suggests that BamE may bind to BamD in the above regions, and therefore that the  $\alpha 10$  helix of BamD potentially becomes clamped between the POTRA 5 domain of BamA and the large, negatively charged region of BamE.<sup>268,277</sup>

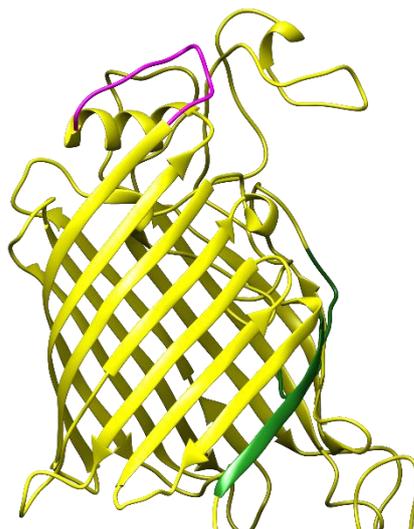
## 5.2. Targeting the BAM Complex

The BAM complex appears to be a very important component of Gram-negative bacterial cells, inserting and folding the  $\beta$ -barrels necessary for cell life to continue. Whilst the precise mechanism by which this insertion and folding occurs is currently unknown, it appears that the lipoprotein components of the complex direct the unfolded OMPs towards BamA with the assistance of the POTRA domains. This appears to happen by recognition of aromatic regions and by binding to unfolded  $\beta$ -sheets utilising hydrophobic regions. In addition, some lipoprotein components may have a regulatory effect on BamA, ensuring that the folding and insertion machinery is only operating when needed.

Due to the evidence suggesting that BamA and BamD are the only required components of the BAM complex, it is not surprising that previous efforts to target the complex have predominantly focused on these proteins. However, whilst the inhibition of the BAM complex has been explored, as discussed below in the following examples, no pharmaceuticals have been licensed by the FDA for use against the protein complex.

### 5.2.1. Monoclonal Antibodies (MAB1)

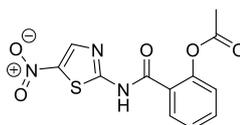
Monoclonal antibodies (mAbs) have become useful biological and chemical biology tools owing to their versatility and broad range of applications, such as blocking of functions of targeted molecules inducing apoptosis of cells containing the target, and the modulation of signalling pathways.<sup>280,281</sup> As such, they have also found usage in treatments for a variety of diseases such as cancer and autoimmune diseases including rheumatoid arthritis and allergic asthma.<sup>280,282</sup> They also possess high target affinity and selectivity, making them prime candidates to examine inhibition of BamA by targeting the extracellular region of the protein. In a screen of over 1600 BamA immunoglobulin G (IgG) mAbs against a strain of *E. coli* performed in 2018, 7 were found to completely inhibit growth of the bacteria, with the representative antibody, named MAB1, found to be able to reduce the number of colony-forming units of a culture by roughly 50-fold within 4 hours of addition. In contrast, cultures treated with a non-inhibitory BamA mAb were found to increase the number of colony-forming units by 50-fold in the same space of time, demonstrating the efficacy of MAB1. Furthermore, MAB1 was found to be a selective BamA-binding antagonist, binding to an ion pair found on extracellular loop 4 of the  $\beta$ -barrel (**Figure 87**), causing increased OM permeability to ethidium bromide and rifampicin at sub-lethal concentrations.<sup>283</sup> However, a limitation of this study is that the *E. coli* strain used displays the lowest possible concentration of LPS, due to the known prevention of mAb-OMP binding by LPS.<sup>283,284</sup> MAB1 therefore, whilst showing high selectivity and potency against BamA, currently has limited use as a therapeutic. It does, however, prove that BamA is targetable by extracellular agents.<sup>285</sup>



*Figure 87. BamA barrel highlighting the position of extracellular loop 4 (magenta) with respect to the lateral gate (green). PDB ID: 5D0O.*

### 5.2.2. Nitazoxanide

Nitazoxanide (**Figure 88**) is a thiazolide which was first approved by the FDA in 2004 for the treatment of giardiasis, but which has since been approved for use as a broad-spectrum antiparasitic and broad-spectrum antiviral drug.<sup>286,287</sup> However, work conducted in 2009 presented evidence that Nitazoxanide may also find use as an anti-virulence antibiotic, an alternative approach to traditional antibiotics which attempts to prevent or slow the bacterial infection by reducing bacterial virulence instead of outright killing the bacteria. In this study, biofilm production and hemagglutination of enteroaggregative *E. coli* strains was inhibited by Nitazoxanide, and the mechanism of action was found to be the inhibition of an assembly mechanism responsible for the construction of fimbriae, the structures responsible for attachment of the bacteria to their host as well as to each other.<sup>288</sup> Further work then showed that Nitazoxanide was able to inhibit the biogenesis of type 1 and type P pili expressed in uropathogenic *E. coli*, again by inhibition of their assembly machinery. These pili are involved in host kidney and bladder adhesion mediation respectively.<sup>289</sup>



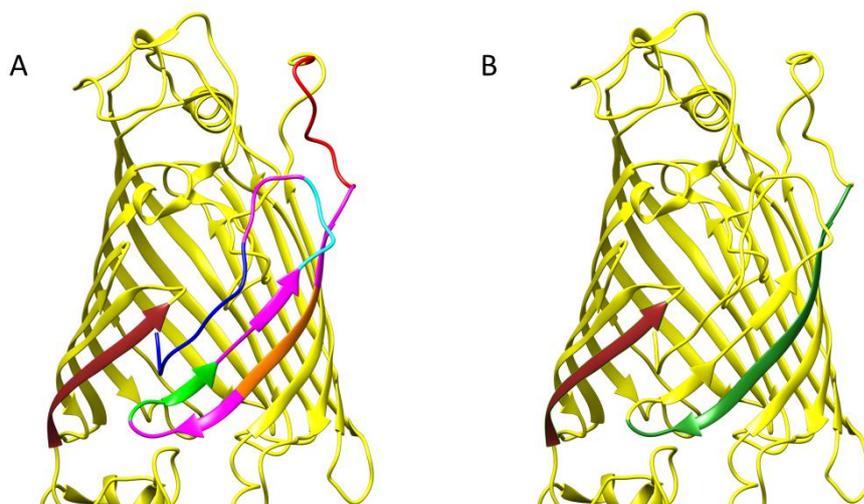
*Figure 88. The structure of Nitazoxanide.*

The two bacterial assembly processes given above both involve the inhibition of OMPs, and hence further work to explore the effect of Nitazoxanide on the BAM complex was undertaken, using an *E. coli* strain as a model. From this study, it was discovered that BAM was the target for Nitazoxanide, with bacterial sensitivity to the drug increasing in mutants with reduced BAM complex levels and decreasing in those with increased levels. This sensitivity is observed in the form of a pili-assembling OMP becoming less or more resistant to the effects of Nitazoxanide, an effect which was observed not to be due to more or less OMP being produced respectively, but instead suggesting that the drug affects a pili assembly OMP-specific aspect of the function of the BAM complex. In particular, Nitazoxanide's target seems to be not only the essential BamD protein, as determined by a P100S mutation which caused decreased sensitivity for the drug, but also BamE. In a mutant strain lacking BamB, the drug had no effect, however this was attributed not to a lack of binding but to an apparent lack of folded protein, whilst a mutant strain lacking BamE showed even greater resistance to Nitazoxanide than the wild type bacteria. Together, these results suggest that Nitazoxanide may operate as an anti-virulence antibiotic, targeting the BAM complex in a way which allows other OMPs to form but preventing the formation of pili.<sup>290</sup>

### 5.2.3. Peptide binding of BamD

One potential method of targeting a protein-protein interaction such as those found in the BAM complex is by using fragments of one of the constituent proteins to target another member. This, to some extent, simplifies the process, as once an interaction surface is identified from a crystal structure the corresponding peptide sequences can then be found, allowing synthesis and testing to occur without computational prediction. Work performed in 2015 was successful in finding a BamD-binding peptide derived from BamA. This peptide, a segment of the  $\beta$ -barrel, was found by expressing the barrel and POTRA domains separately, and upon discovering that the barrel contained the binding region, expression of four segments of the barrel's sequence constituting the entire barrel. Only the C-terminal 96 amino acids of the barrel were found to bind to BamD, resulting in inhibition of BamA folding. This region was therefore further divided into 5 smaller peptides consisting of overlapping sequences of 15-16 residues (**Figure 89A**), and these smaller peptides tested. Only one of these inhibited the folding of BamA, as well as another OMP, indicating that this peptide could be used as a general inhibitor of the BAM complex (**Figure 89B**). The work also highlighted that this inhibitory activity is more sensitive to changes in the

sequence of the central region of the peptide, potentially indicating that the true interaction surface of BamA required for an inhibitory effect may be even smaller.<sup>291</sup>



**Figure 89.** Sections of the BamA barrel. (A) The five peptides generated from the C-terminal 96 amino acid sequence. Red, orange, green, cyan and blue sections are regions which do not overlap with any other peptide, whilst the magenta regions are overlapping segments of the sequences. (B) The complete inhibitory peptide. The 61 side of the lateral gate (brown) is shown for reference. PDB ID: 5D00.

### 5.3. Aims of the study

Whilst the peptide inhibitor of BamD shows good activity, it has already been discussed that peptides are typically not taken forward as good candidates for pharmaceuticals due to their poor *in vivo* properties (see **Section 1.6**). However, it was hypothesised that the utilisation of these and other results in a peptide-directed binding study (see **Section 1.15**) may yield small molecule peptidomimetics able to inhibit the BAM complex. The aims of this study are therefore:

- To identify sequences involved in important processes within the BAM complex.
- To perform peptide-directed binding studies on these sequences and to generate small molecule peptidomimetics based on these peptides.
- To examine the activity of these small molecules using *in vitro* and *in vivo* studies.

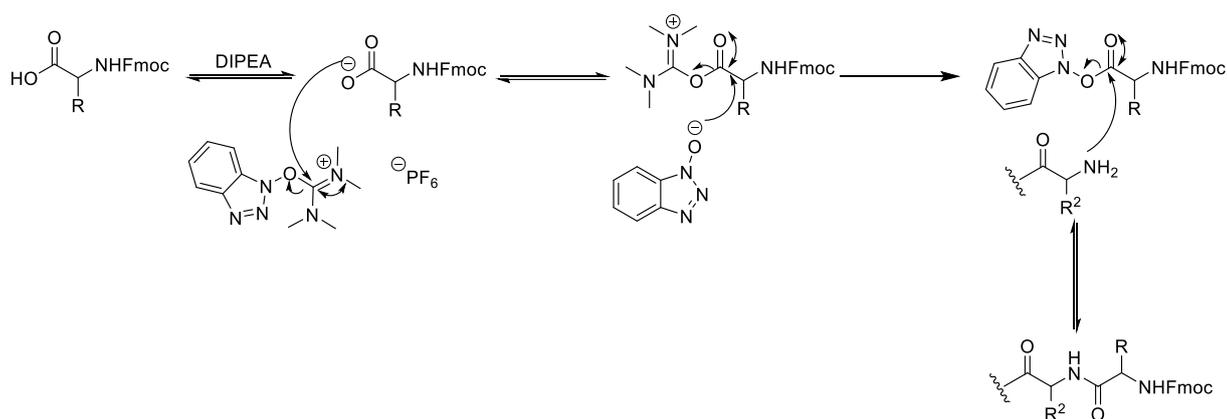
#### 5.4. Generation of BamD substrate surface inhibitors

As discussed in **Section 5.2.3.**, a 15-residue peptide inhibitor of the BamD substrate surface was found in the form of a fragment of the BamA barrel. This substrate surface is required for the correct function of the BAM complex, and so seems an ideal starting point for peptide-directed binding. In addition to this peptide (residues 765-779 of the BamA barrel), it was observed that a mutation within a small section of this peptide (residues 769-776) dramatically reduced the binding properties of the peptide. It was therefore decided that both the full length peptide and the shorter segment of this peptide should be synthesised and examined for binding. The sequences of these peptides are provided in **Table 4.**

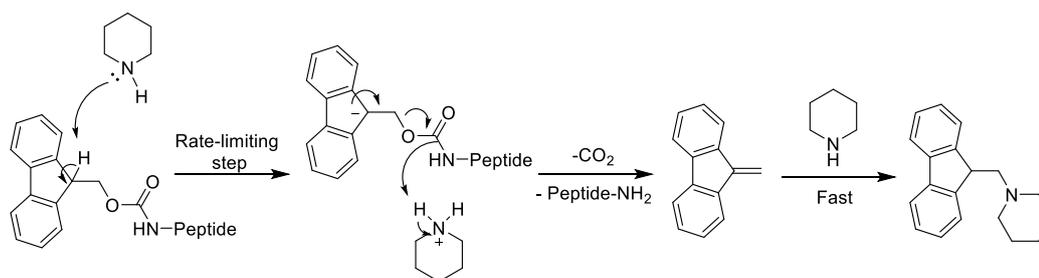
*Table 4. Sequences of BamD inhibitors based on the BamA barrel.*

Compound ID	Sequence
<b>87</b>	NIRMSAGIALQWMSP
<b>88</b>	SAGIALQW

Attempts to synthesise these peptides utilised solid-phase peptide synthesis. This method uses a polystyrene-based resin to bind the growing peptide in place, allowing for the sequence to be constructed quickly and with a high level of automation. After each residue is attached to the resin-bound peptide using coupling reagents, such as hydroxybenzotriazole (HOBt) and hexafluorophosphate benzotriazole tetramethyl uronium (HBTU), and a base such as diisopropylethylamine (DIPEA) (**Scheme 72**), it is first washed and then the Fmoc group protecting the amine group is removed to allow for the attachment of the next residue (**Scheme 73**), allowing for a minimum of side products.

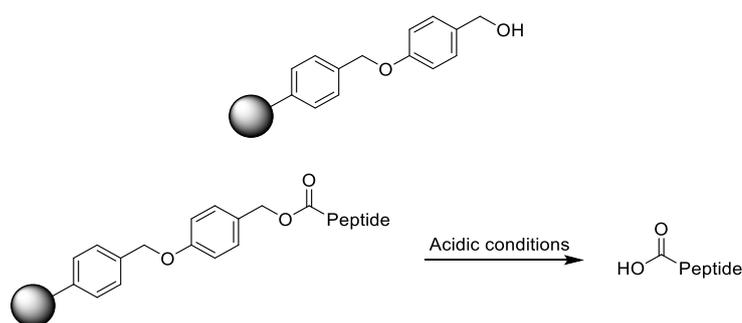


*Scheme 72. Solid phase peptide synthesis coupling step*



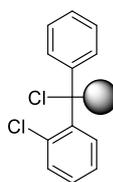
*Scheme 73. Fmoc deprotection of peptides using piperidine*

Initial synthetic attempts to obtain **87** and **88** used a Wang resin. This resin, upon cleavage of the peptide from the functionalised polystyrene, yields peptide acids such as those found in nature (**Figure 90**). The synthesis was carried out under automation, following which the peptides were cleaved. This was achieved using a cleavage cocktail consisting predominantly of trifluoroacetic acid (TFA) in addition to water and triisopropylsilane (TIPS). Under these conditions, the acid-labile protecting groups attached to the amino acid side chains are removed, and the bonds connecting the resin to the peptide are cleaved, allowing the peptide to be isolated by filtration of the resin and removal of solvents. This process results in the formation of cationic species resulting from the cleaved protecting groups which, if the cleavage reaction was performed in only TFA, could result in unwanted modifications to the peptide chain. The water and TIPS in the cleavage cocktail act as scavengers, deactivating these species. In addition to these two scavengers, ethane-1,2-dithiol must be used as an additional scavenger for peptides containing cysteine or methionine, both to protect tryptophan residues from sulfonation and to suppress oxidation of the thiols on these residues.<sup>292</sup>



*Figure 90. Structure of Wang resin (top) and the deprotection of the resin to form peptide acids (bottom).*

Whilst **88** was successfully synthesised on this resin, as determined by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry, **87** was not, instead yielding several shorter peptide fragments. Upon closer analysis, it was found that none of these fragments contained the proline residue, indicating that this amino acid was not successfully attached to the resin. This was further suggested by previous publications, indicating that proline-containing peptides were difficult to synthesise when the C-terminal residue was proline.<sup>292</sup> Therefore, a 2-chlorotrityl (2-Cl-Trt) resin (**Figure 91**) with pre-loaded proline was used for future attempts to synthesise **87**. These attempts appeared to be successful, with the correct mass observed using MALDI-TOF.



*Figure 91. Structure of 2-chlorotrityl resin.*

Purification was achieved using reverse-phase preparative HPLC, with first attempts at purification utilising a 5-95% water:methanol gradient. Peptide **88** was able to be mostly purified using these conditions, showing as a major product of the reaction and appearing with only small shoulders on analytical HPLC, which could not be separated from the major peak by preparative or semi-preparative HPLC (**Figure 92**), but peptide **87** proved to be very difficult to dissolve in these solvents. This difficulty was further increased by the fact that the signal of interest is localised in a region containing multiple signals. In addition, based on the HPLC trace, it appeared that the desired compound was also a minor product of the reaction (**Figure 93**). An attempt to replace methanol with acetonitrile proved to decrease solubility further, however in both cases a small amount of peptide was isolated, further indicating that this is not the major product of the synthesis. It was hypothesised that this was due to the harsh cleavage conditions that the peptide was exposed to, and so the synthesis of **87** was repeated using a less acidic cleavage mixture, as well as adding a  $\beta$ -alanine residue to the C-terminus to attempt to remove the problems associated with a C-terminal proline. This once again yielded an insoluble product, and the desired peptide was also once again a minor product of the reaction.

CHAPTER 5: SYNTHESIS OF POTENTIAL BAMA/BAMD INTERACTION INHIBITORS

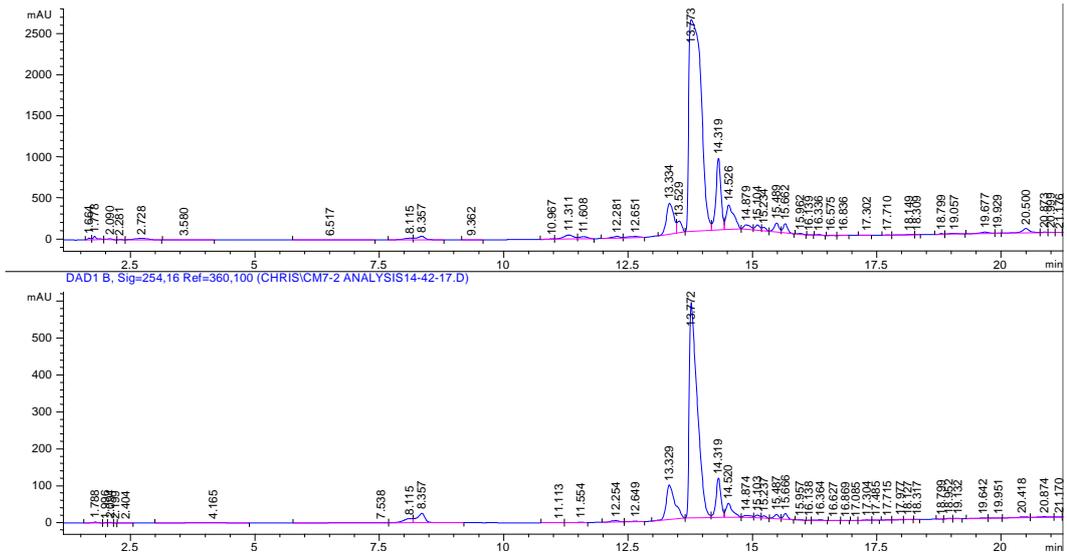


Figure 92. Analytical HPLC trace of peptide 88 after purification

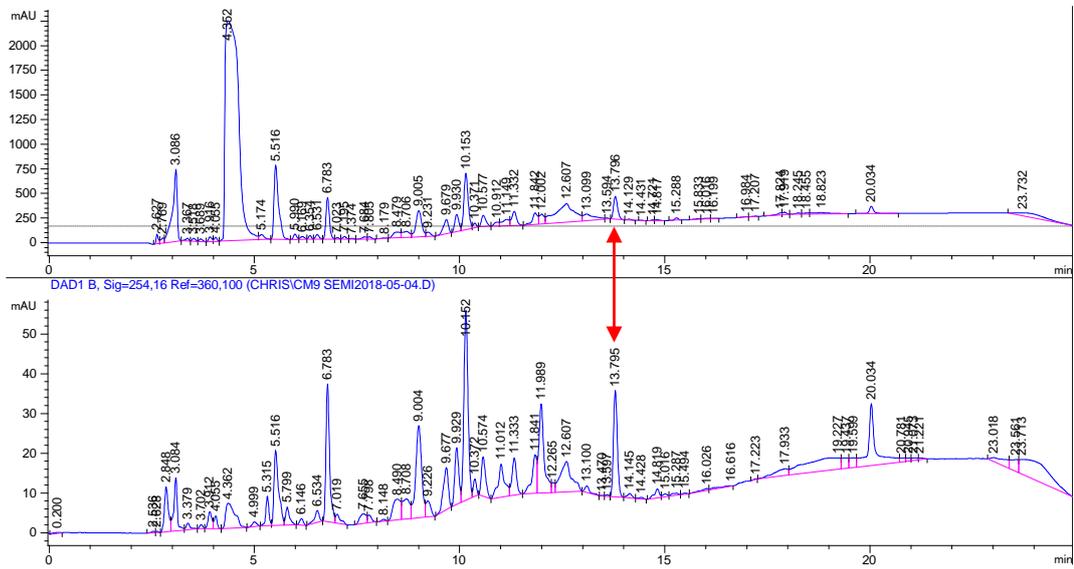


Figure 93. Analytical HPLC trace of 87 before purification. The red arrow indicates the signal corresponding to the peptide

## CHAPTER 5: SYNTHESIS OF POTENTIAL BAMA/BAMD INTERACTION INHIBITORS

Due to impurities remaining after preparative HPLC of **88**, the synthesis was repeated using a 2-Cl-Trt resin preloaded with Fmoc-Trp(Boc), keeping the same deprotection and cleavage conditions. This synthesis was once again performed under automation. Analytical HPLC of the product showed a similar spread of impurities, as well as the product signal also appearing as a multiplet. It was therefore hypothesised that the deprotection mixture or the cleavage mixture was the cause of this high array of side products. An alteration of the cleavage cocktail scavengers, for example, may result in more efficient scavenging of unwanted side products generated during the side chain deprotections. Alternatively, if the cleavage mixture is not acidic enough, it could cause incomplete side chain deprotection. Finally, if the deprotection mixture is too concentrated peptide aggregation could occur, whilst if it is not concentrated enough it could cause incomplete Fmoc deprotection.

To address this problem, the reaction was repeated once more on 2-Cl-Trt resin with pre-loaded tryptophan. Two batches of this reaction were run in parallel, using 40% piperidine and 20% piperidine in DMF as the deprotection mixtures. Once these reactions had been run, the products were once again purified using preparative HPLC. The HPLC traces of these products indicated that using 40% piperidine provided a product with fewer signals than 20% piperidine, suggesting that 40% piperidine deprotection is the most efficient deprotection conditions. However, additional signals corresponding to impurities were still observed, and so the cleavage conditions were next examined. Cleavage conditions were modified several times, as summarised in **Table 5**. However, in all cases impurity signals were observed in HPLC, prompting an investigation into problematic residues.

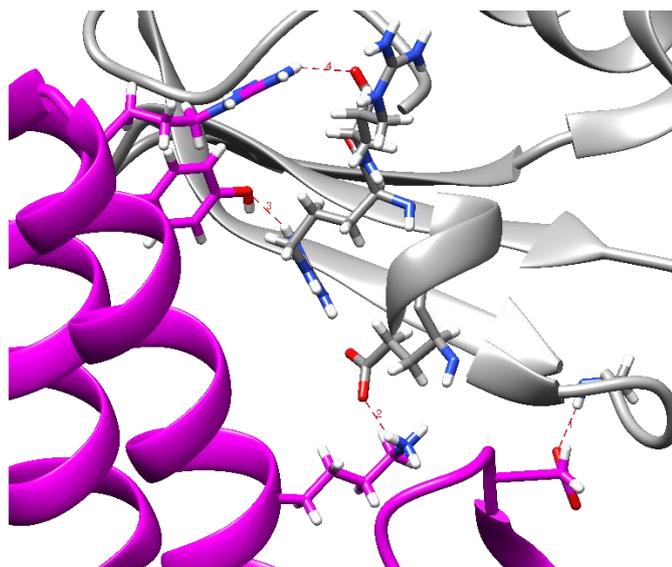
*Table 5. List of cleavage cocktails used for the attempted synthesis of **87** and **88***

	<b>Cleavage cocktail</b>
<b>(a)</b>	95 % TFA, 2.5 % TIPS, 2.5 % H <sub>2</sub> O
<b>(b)</b>	95 % DCM, 1.5 % TFA
<b>(c)</b>	70 % DCM, 20 % TFA, 10 % TIPS
<b>(d)</b>	85 % DCM, 10 % TIPS, 5 % TFA

The shorter peptide (**88**) was prepared manually using Wang resin and a 40% piperidine deprotection cocktail, with a small aliquot of resin removed after every 2 couplings. The process was also monitored using the Kaiser test. This chemical test uses potassium cyanide, ninhydrin and phenol, which, when added to a sample containing a free amine, undergoes a colour change to become intense blue due to the formation of Ruhemann's purple dye. The test was used following each amino acid coupling, with a successful coupling being indicated by the lack of this blue coloration owing to the Fmoc protection of the amino acids. Following this stepwise synthesis, all aliquots were deprotected using cleavage cocktail (**a**) listed in **Table 5**. HPLC analyses of these aliquots showed that in all cases impurities were present. At this point, attempts to improve the purity of the crude peptide mixture were abandoned, with purification of **88** by preparative HPLC removing the vast majority of impurities. However, due to the low yields associated with purification of **87**, this peptide was abandoned.

### 5.5. Inhibition of the BamAD binding surface

Meanwhile, another potential starting point was found by literature review. The structure of the BamAD interface in *Rhodothermus marinus*, elucidated in 2016 by Bergal and co-workers,<sup>293</sup> identified a number of key residues for the correct interaction between BamA and BamD (**Figure 94**). The side chain of D128 of BamD interacts with the amide bond of G398 of the BamA POTRA domain, whilst the side chains of R180 and Y177 of BamD interact with the chain carbonyl of R409 and the side chain guanidinium of R408 of BamA respectively. Furthermore, the side chain amine of K169 of BamD and the side chain carbonyl of E404 of BamA possess a salt bridge, stabilising the interaction surface further.<sup>293</sup>



**Figure 94.** Structure of the POTRA5 (grey)-BamD (magenta) interaction surface. Dashed red lines indicate protein-protein interactions, with interaction 1 corresponding to G398-D128; label 2 to E404-K169; label 3 to R408-Y177; label 4 to R409-R180. PDB ID: 5EFR.

From these interactions, two peptides were designed in an attempt to include as many of the interacting residues as possible, whilst also ensuring that the resulting peptide was not excessively long. However, due to the complexity of performing assays on membrane proteins, only the peptide derived from the POTRA5 domain of BamA that was thought to bind to BamD was taken forward. This peptide, including all four of the important binding residues, is shown below in **Table 6**. Whilst the peptide would therefore be an ideal starting point for peptide-directed binding based solely on the retention of the key residues, it is also excessively long, which may result in complications during peptide synthesis. In addition, whilst E404, R408 and R409 are all situated within the same  $\alpha$ -helical region, G398 is located on the preceding  $\beta$ -bend, a structural feature which may not be replicated within the peptide sequence. It was therefore decided that G398 would not be included in the final used peptide, thus shortening the peptide to eight residues. The sequence of this peptide (**89**), as well as the unused peptide, is given below in **Table 6**.

**Table 6.** Suggested sequences for the BamA-D interaction surface inhibitors

Compound ID	Sequence
Unused full-length BamD-inhibiting peptide	VGNQKTKEHVIRRE
<b>89</b>	KEHVIRRE
Unused BamA-inhibiting peptide	RKQYEAARLYERRE

Peptide **89** was synthesised under automation on a Wang resin using 40% piperidine and HOBt/HBTU. The peptide was then cleaved from the resin using 95:2.5:2.5 TFA:TIPS:H<sub>2</sub>O. Purification by HPLC appeared to indicate that synthesis of this peptide had been successful, owing to the small number of signals observed in the trace. However, upon analysis of these signals by MALDI-TOF mass spectrometry, it appeared that the peptide had not been successfully generated, instead indicating that a compound of mass 858 Da was the major product of the reaction. It was therefore assumed that one or more of the residues were incorrectly coupling, and so the preparation was repeated manually, with monitoring by the Kaiser test and with aliquots of resin removed after every two residues. However, upon performing the reaction manually, the expected product was observed with good purity (**Figure 95**), most likely due to the arginine residues requiring longer periods of time to couple correctly as determined by Kaiser testing during the synthesis.

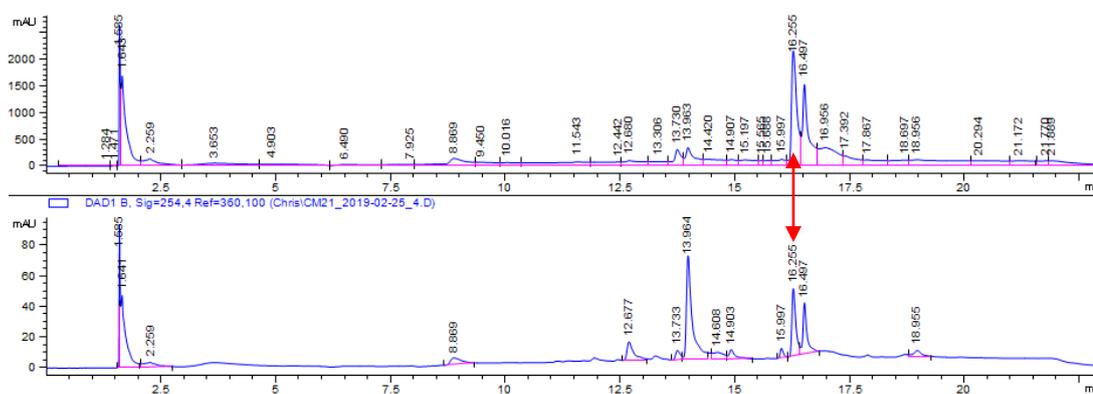


Figure 95. Analytical HPLC trace of **89** before purification

## 5.6. *In vitro* assessment

To assess the activity of peptides **88** and **89**, it was first necessary to express and purify samples of BamD. This was achieved by first cloning the BamD plasmid. Both forward and reverse primers were designed based on DNA sequence data obtained from the NCBI GenBank.<sup>294</sup> Following this, the primers and BAM genome were inserted into two separate plasmids, pHisTEV and pLeu3, which were then cloned using Tpp10 cells. These plasmids both encoded an ampicillin resistant gene, allowing for the cells to be grown selectively by addition of ampicillin to the growth media, killing most other cells. The plasmids also encoded a His tag, a chain of 6 histidines, to the BamD generated by the cells. This allows for separation of the protein of interest by nickel column chromatography. The Tpp cells were selected due to their fast culture growth rate, allowing for large quantities of competent plasmid to be produced in a shorter space of time.



The first step in this process is to perform a buffer screening experiment to determine which buffer is most suitable for the protein, and so will therefore be used for any further ligand binding studies using DSF. Buffers have been shown to be important in stabilising the dissolved protein during DSF, with the ideal buffer helping to increase the change in  $T_m$  observed between samples with the inhibitor present and samples with the inhibitor absent.<sup>297</sup> Sixteen buffers were prepared and were added to a mixture of protein, peptide, DMSO and SYPRO Orange dye, as well as a series of blanks containing only buffer, protein, DMSO and SYPRO Orange. These samples were then loaded into a quantitative polymerase chain reaction (qPCR) machine and the DSF analysis run on temperatures between 20 °C and 95 °C. The buffer screen for peptide **88** was the first to be performed. However, upon examining the results it was clear that binding of the peptide to the protein had not occurred in any of the buffers, as there was no significant difference between the  $T_m$  of the protein in the presence of and absence of peptide. This was originally assumed to be due to the presence of trisaminomethane(tris) from the lysis buffer remaining following the protein expression, and so the samples of BamD were concentrated further and the buffer replaced with another buffer with a lower tris concentration. Even after this buffer exchange, the peptide failed to bind within any of the buffer mixtures prepared. Since peptide **88** is only a smaller segment of peptide **87** which is thought to contain the binding region, it is possible that the entire sequence of peptide **87** is required for binding. However, due to the difficulties associated in its synthesis and purification, this hypothesis was not proven.

In a similar fashion, peptide **89** was used in DSF buffer screens against BamD. Once again, no binding was observed via DSF for the peptide. Peptide **89** was derived from the sequence of an *R. marinus* species, and was shown to be important for the BamAD interaction within that species only. Due to the somewhat sequence-specific nature of protein-protein interactions, it was thought that the BamAD interaction surface was not conserved across different strains and species of bacteria. Indeed, when the sequences of BamA from *E. coli* and *R. marinus* were compared by NCBI BLAST (Basic Local Alignment Search Tool), they were found to only match by 23 % throughout the entire sequence, and 24 % within the POTRA5 domain.<sup>298-300</sup> Furthermore, in *E. coli* the corresponding sequence to peptide **89** replaces the histidine with an alanine, suggesting an additional difference to the interaction surface in *E. coli*. An alternative cause of this failure to bind could be a mismatch in the secondary structure between the synthesised peptide and the section of the POTRA5 domain that this sequence forms a part of. In the crystal structure, the sequence is part of an  $\alpha$ -helical motif, and so there is no guarantee that this peptide would also adopt this secondary structure.

### 5.7. Conclusions

The BAM complex is an essential bacterial protein complex which currently possesses relatively few confirmed targetable features. This work aimed to use peptide-directed binding to explore the possibility of using small molecule peptidomimetic compounds to inhibit the action of the BAM complex. The first peptide explored (**87**), a segment of the BamA transmembrane barrel, was found to be difficult to synthesise using SPPS, and was found to be the minor product of the reaction. Future work could attempt this reaction using microwave-assisted SPPS, as microwave irradiation has been shown to improve the performance of a variety of reactions, including the formation of peptide amide bonds.<sup>301</sup> There is also scope for further resins and cleavage cocktails to be examined.

A shorter chain within this barrel fragment (**88**), suggested in previous work to be the residues responsible for the binding affinity of the fragment, was synthesised using SPPS, allowing examination of the binding by DSF analysis using *E. coli*-expressed BamD. However, upon attempting the buffer screening step to identify the most efficient buffer for the protein (i.e. the buffer which provides the greatest change in  $T_m$ ), the peptide exhibited no observed binding effects. This could potentially indicate that the complete peptide segment of the BamA barrel is required for correct binding to BamD, but due to the challenging synthesis of the longer peptide this was not confirmed. Further studies, such as

the slow expansion of this peptide to include more residues found in the original peptide, could potentially yield a peptide which is both readily synthesised and shows good activity against BamD, thus allowing the peptide-directed synthesis of a small molecule to occur.

A second peptide (**89**), discovered in previous literature to constitute part of the *R. marinus* BamA POTRA5 domain which binds to BamD, was also synthesised using SPPS. However, as for peptide **88**, this peptide showed no binding affinity in a DSF buffer screen. This could be due to the peptide not matching the secondary structure found in the POTRA5 region, hence preventing binding due to the required residues not being able to contact their BamD counterparts correctly, or could be due to sequence differences between *R. marinus* and *E. coli* thus providing an unsuitable binding surface. It may therefore be more beneficial if future work focused on BamD proteins from species with higher conservation of the binding region in an attempt to make a more universally applicable inhibitor.

# **CHAPTER 6: CONCLUSIONS AND FUTURE WORK**

## CHAPTER 6: CONCLUSIONS AND FUTURE WORK

Antimicrobial resistance has caused an unprecedented need for the discovery of novel antimicrobial agents, especially those which are capable of acting against Gram-negative bacteria where strains have begun to emerge which have resistance pathways to all known antimicrobials. This can be achieved, broadly, by two means; either a new target, and hence a new class of antimicrobials, can be discovered and assessed, or by targeting an existing antimicrobial target but in such a way as to circumvent resistance. The work described in the previous chapters sought to explore both methods, using nature as a starting point for the generation of novel antibiotic agents.

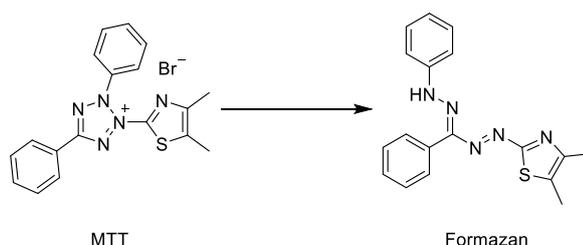
**Chapters 2-4** focused on DNA gyrase, a bacterial Type II topoisomerase found in all bacterial wild-type strains that is currently therapeutically targeted by the quinolone antibiotics. The work performed aimed to use SD8, a member of the simocyclinone family of natural products, as a starting point for the design and synthesis of “pseudo-natural product” inhibitors of DNA gyrase. An *in silico* library screening project found five angucyclinone-replacing alkynes and five aminocoumarin-replacing amines, selected for both their docking scores and synthetic feasibility, before four of five of the alkyne fragments and all five amine fragments were synthetically prepared. Beginning with aminocoumarin-linker-alkyne hybrid molecules, the prepared alkynes were attached, through CuAAC click chemistry, to a molecule consisting of an aminocoumarin connected to a 10-azidodecanoic acid linker. These hybrids were then examined in *E. coli* DNA gyrase supercoiling assays, determining that none of the four compounds possessed any observable supercoiling inhibitory activity at <100  $\mu$ M. However, since a simplified version of the aminocoumarin was used, the alkyne with the highest docking score, compound **9**, was taken forward.

The final stage involved attachment of the five amine fragments to the 10-azidodecanoic acid linker, before compound **9** was attached, as before using CuAAC click chemistry. These compounds, **81**, **82**, and **84-86**, were then examined in *E. coli* DNA gyrase supercoiling assays as before. Unfortunately, these compounds once again possessed no observable supercoiling inhibitory activity at <100  $\mu$ M.

Since SD8 acts on Type II topoisomerases other than DNA gyrase, all nine compounds were also examined for other topoisomerism inhibitory activity. Firstly, the compounds were used in *E. coli* Topo IV supercoiling relaxation assays, followed by *S. aureus* DNA gyrase supercoiling assays. Finally, the compounds were used in human Topo II decatenation assays. In all cases, whilst SD8 showed potent inhibitory activity, compounds **52-55** and **81**,

**82** and **84-86** showed no inhibitory activity at concentrations  $<100 \mu\text{M}$ . However, compounds **81**, **82** and **84** may possess some antimicrobial activity, as observed by zone-of-inhibition assays conducted with all nine compounds. However, as the results of these assays were unclear, further work would be required to ensure that these compounds were active, and then to elucidate the mode of action of these molecules.

To confirm that compounds **81**, **82** and **84** are active, a cell viability and proliferation assay, such as an MTT assay, could be used. This assay utilises 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), a yellow tetrazolium salt which is reduced by the NAD(P)H-dependent oxidoreductase enzymes of viable cells to formazan, a purple solid (**Figure 97**).<sup>302,303</sup> This solid can be solubilised and the solution quantified by measuring its absorbance at 500-600 nanometres, with the higher the measured absorbance indicating a greater number of viable cells. Not only would this assay confirm whether the compounds are in fact active, it would also quantify the level of activity.



*Figure 97. Reduction of MTT to formazan*

The lack of DNA gyrase inhibitory activity, whilst disheartening, is not necessarily surprising. Comparing the structures of SD8 and compounds **52-55**, or compounds **81**, **82** and **84-86**, it is clear that a large number of modifications have taken place to simplify the synthesis of the molecule, with each modification adding a further degree of ambiguity as to what the effect of each change has been. In an ideal situation, one section of the molecule would have been modified at any given time, with the remaining sections being held constant. With the advancement of synthetic techniques, it is plausible that the synthesis of each component, and the total synthesis, of SD8 may be achieved in the near future. However, prior to this development it may be possible, through the use of ITC or SPR, to determine the binding affinity of each individual component to DNA gyrase.

## CHAPTER 6: CONCLUSIONS AND FUTURE WORK

Isothermal titration calorimetry allows for the quantitative examination of binding through extremely small temperature changes that take place during these events. A reference cell, containing water, and a sample cell, containing the protein of interest, are placed in a calorimeter which is set to the desired experimental temperature, before a potential ligand is titrated accurately into the sample cell through the addition of a series of small aliquots. Any temperature changes are detected, with these changes being directly proportional to the amount of binding.

Surface plasmon resonance is another quantitative technique which examines binding through changes in mass. Polarised light is focused on a gold-plated glass receptor, with enzyme immobilised on the gold plating. The ligand of interest is then flowed across the gold-enzyme surface, with binding affecting the mass of the enzyme and hence the refractive index of the surface, enabling real-time detection and accurate determination of the amount of bound ligand, the ligand's binding affinity and kinetic details regarding association and dissociation of the ligands. Importantly, the technique requires very little enzyme, and does not require tagging or radiolabelling.<sup>304</sup> An overview of SPR is shown below in **Figure 98**.

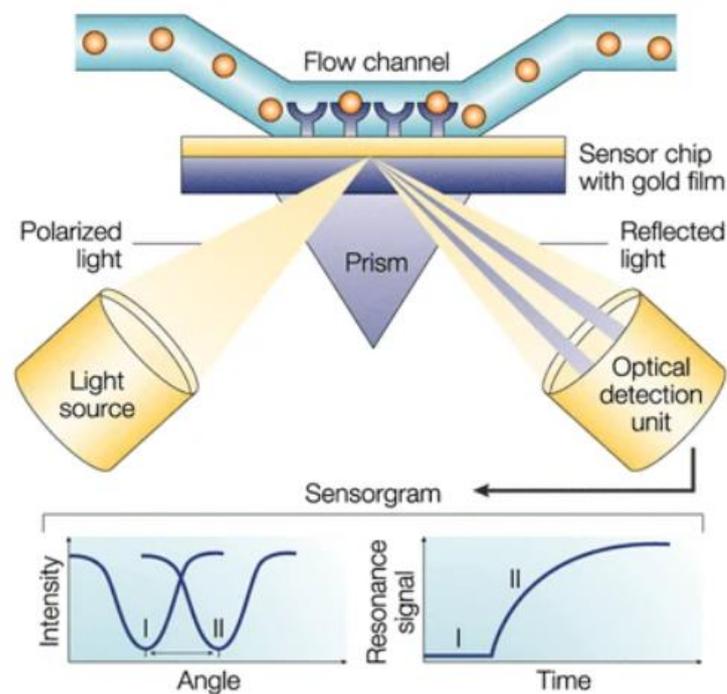


Figure 98. Overview of surface plasmon resonance (SPR) <sup>304</sup>

Finally, if any of these fragments show binding affinity, a crystal structure of the compound bound to the protein would provide further information regarding the specific orientation and mode of docking of the compounds, allowing informed modifications to be made to improve the efficacy of the molecules.

**Chapter 5** focused on the BAM complex, a bacterial protein complex found in Gram-negative bacteria that is currently not therapeutically targeted. The work aimed to use peptides with sequences derived from important binding regions of one component of the complex, BamA, as a starting point to generate peptidomimetics that targeted the BamAD protein-protein interaction. Two peptides, **88** and **89**, were successfully synthesised, and were initially used in DSF buffer screens to determine which buffer would be used in further experiments. However, neither peptide showed appreciable binding to BamD, as measured by the  $\Delta T_m$ , and hence would not serve to inhibit the interaction. Compounds **87**, a sequence that proved too challenging to generate, and **88** were closely related, with peptide **88** being the binding sequence of peptide **87** which had been shown in previous work to successfully bind to BamD and inhibit the BamAD interaction. Further work could therefore attempt to expand the number of residues from peptide **87** included in peptide **88**, aiming to improve binding affinity without sacrificing synthetic feasibility. Peptide **87** could also be generated using microwave-assisted SPPS, a technique which would now be possible within the group.

## **CHAPTER 7: EXPERIMENTAL**

### 7.1. General Procedures

All chemicals were purchased from Fluorochem or Sigma Aldrich. Anhydrous solvents were commercially purchased and assumed to conform to manufacturers specifications.  $^1\text{H}$  and  $^{13}\text{C}$  NMR were obtained in Fourier Transform mode on a Bruker™ Ultrashield PLUS 400 spectrometer operating at a normal frequency of 400 MHz using the specified deuterated solvent. All spectra have been calibrated to the residual deuterated solvent peak and chemical shifts reported in ppm. Spectra were analysed and processed using Topspin 3.2 software. Multiplicities in the NMR are described as: br s = broad singlet, d = doublet, dd = doublet of doublets, t = triplet, q=quartet, qu=quintet, se=sextet m = multiplet; coupling constants are reported in Hz. Infrared spectra were recorded from neat samples using a Perkin-Elmer Spectrum BX FT-IR spectrometer and analysed using Spectrum v5.3.1 software. MALDI mass spectra were recorded using a Kratos Analytical Maxima MALDI-TOF. High-resolution mass spectra were recorded using a Waters Synapt XS mass spectrometer fitted with a Water Acquity I-Class UPLC.

### 7.2. Chromatographic techniques

Flash chromatography was performed using a Biotage Isolera and silica cartridges (particle size 60  $\mu\text{m}$ ). Analytical RP-HPLC was performed on an Agilent 1200 using an Agilent eclipse XDB-C18 column, 4.6 x 150 mm, 5  $\mu\text{M}$  and a flow rate of 1 mL/min. The gradient for these separations used water with 0.05 % TFA (solvent A) and MeOH with 0.05 % TFA (solvent B), and involved 5 % solvent B increasing to 95 % over 20 minutes. Detection wavelengths were set to 214 nm and 254 nm. Preparative RP-HPLC was performed on an Agilent 1200 using an Agilent eclipse XDB-C18 column, 21.2 x 150 mm, 5  $\mu\text{M}$  and a flow rate of 20 mL/min. The gradient for these separations used 95 % water, 5 % MeOH, 0.05 % TFA (solvent A) and 5 % water, 95 % MeOH, 0.05 % TFA (solvent B), and involved 5 % solvent B increasing to 95 % B over 20 minutes.

### 7.3. Computational procedures

Computational studies were performed using Schrödinger Maestro 2016-4 software, with core-hopping, glide and covalent docking functionalities.<sup>214-219</sup> Amine library PAINS filtration, calculation of AlogP and filtration based on these AlogP values were performed using Schrödinger Canvas.<sup>221</sup>

#### 7.3.1. Preparation of the receptor and assessment of linker chain length

The X-ray crystal structure of DNA GyrA with SD8 (PDB ID 4CKL) was utilized for docking calculations. The ligand was removed and the protein prepared using the Protein Preparation application. The SD8 was modified where necessary using the Maestro drawing tools, and the resulting ligand was prepared using the LigPrep application. The Glide grid was generated for the prepared protein structure, and the SD8 was docked into the protein using the Glide application to verify that the Glide grid functioned correctly. The new cores were drawn using the Maestro drawing tools and were inserted into the SD8 structure using the core-hopping application. The suitability of the new cores was judged based on the resulting Glide score function of the docked molecules.

#### 7.3.2. Alkyne and amine fragment library generation and filtration

Ligands were taken from the Aldrich Market Select virtual library, filtered using the filtration application of Schrödinger Maestro 2016-4 (for filtration based on functional groups and hydrogen-bond donors/acceptors in the case of the amine library) or Schrödinger Canvas (for filtration of PAINS, molecular weight, calculation of AlogP and subsequent filtration of compounds with undesirable AlogP values), and prepared using the LigPrep application. Filtering was performed using the following SMILES codes, using matching molecules:

*Alkynes:*

[C#CH]

*Amines:*

[C,c][NX3,NX4;H3,H2,H1;!\$(NC=O);!\$(NC=S);!\$(NC=N);!\$(NCS(=O)(=O));!\$(N=\*);!\$(NO);!\$(N(C)(C)(C));!\$(N(c)(C)(C));!\$(N-N);!\$(N(C)(C)(S));!\$(NS(=O)(=O))]

**7.3.3. Covalent docking**

Covalent docking was performed using the CovDock tool from the Schrödinger Suite 2016-4, using the following custom code in attempts to generate the desired connections:

*Azide receptor, alkyne ligand:*

```
RECEPTOR_SMARTS_PATTERN 1,N=[N+]=[N-]
LIGAND_SMARTS_PATTERN 3,[C,c]-C#C
CUSTOM_CHEMISTRY (“<1>=[N+]=[N-]”,(“charge”, 0, 3))
CUSTOM_CHEMISTRY (“<1>=[N+]”,(“charge”, 0, 2))
CUSTOM_CHEMISTRY (“<1>|<2>”,(“bond”, 1, (1,2)))
CUSTOM_CHEMISTRY (“<2>#C”,(“bond”, 2, (1,2)))
CUSTOM_CHEMISTRY (“<1>=[N]”,(“bond”, 1, (1,2)))
CUSTOM_CHEMISTRY (“<1>([C]=[C])-[N]=[N]”,(“bond”, 1, (3,5)))
```

*Carboxylic acid receptor, amine ligand:*

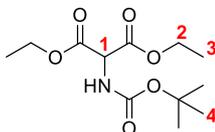
```
RECEPTOR_SMARTS_PATTERN 1,[C](=O)-[OH]
LIGAND_SMARTS_PATTERN
2,[C,c][NX3,NX4;H3,H2,H1;$(NC=O);$(NC=N);$(NC=S);$(N=*);$!(NCS);$!(N(C)(C)
(C));$(N(c)(C)(C));$(N-N);!$(N(C)(C)(S));!$(NS(=O)(=O))]
CUSTOM_CHEMISTRY (“<1>-[O;H1]”,(“delete”,2))
CUSTOM_CHEMISTRY (“<2>”,(“charge,0,1))
CUSTOM_CHEMISTRY (“<1>|<2>”,(“bond”, 1, (1,2)))
```

**7.3.4. Docking of fragment-SD8 libraries**

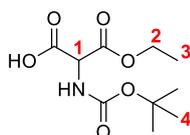
Fragment-SD8 libraries were prepared by generation of the required connections between the fragment library and SD8, modified through the use of the drawing tools of Maestro, by using a previously prepared Python script. The resulting library was then prepared using the Ligprep functionality before the molecules were docked into the protein crystal structure using Glide. In the case of the alkyne-SD8 library, the aminocoumarin was held stationary, whilst in the case of the amine-SD8 library the angucyclinone was held stationary. The resulting poses were then filtered using a second previously prepared Python script to include only one pose per original library member, and to rank the structures in order of their Glide docking score. Molecules were then assessed in order of decreasing Glide score on their synthetic feasibility.

## 7.4. Synthetic procedures and characterisation

## 7.4.1. Solution phase synthesis

1,3-Diethyl-2-[[*tert*-butoxy]carbonyl]amino}propanedioate (**20**)

Diethyl aminomalonate hydrochloride (10.2 g, 48.1 mmol) was dissolved in a mixture of 1M sodium hydroxide (48 cm<sup>3</sup>) and 1,4-dioxane (40 cm<sup>3</sup>). Following this, a solution of Boc<sub>2</sub>O (11.5 g, 52.7 mmol) dissolved in 1,4-dioxane (20 cm<sup>3</sup>) was added dropwise to the stirred solution of starting material, causing the mixture to become cloudy. The reaction was stirred for 17 h at room temperature, before the solvent was removed under reduced pressure. The residue was then taken up in ethyl acetate, and the insoluble white solid was filtered off before the organic layer was washed with 5% KHSO<sub>4</sub>, saturated NaHCO<sub>3</sub> and brine. The organic layer was dried over sodium sulfate, after which the solvent was removed under reduced pressure to yield the desired product as a colourless oil (11.15 g, 84%). <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>) δ<sub>H</sub> ppm: 7.63 (s, 1 H, N-H), 4.83 (d, *J* = 8 Hz, 1 H, 1-H), 4.23-4.11 (m, 4 H, 2-H), 1.39 (br s, 9 H, Boc), 1.23 (t, *J* = 6 Hz, 6 H, 3-H). <sup>13</sup>C NMR (100 MHz, DMSO-d<sub>6</sub>) δ<sub>C</sub> ppm: 166.8 (ester C=O), 155.0 (Boc C=O), 79.2 (Boc C(CH<sub>3</sub>)<sub>3</sub>), 61.8 (2), 57.7 (1), 27.9 (3), 13.8 (4). IR ν<sub>max</sub> (neat)/ cm<sup>-1</sup> 3383 (N-H), 2981 (C-H), 1715 (C=O), 1393 (C-N), 1158 (C-O). HRMS (ESI+) calculated for C<sub>12</sub>H<sub>20</sub>NO<sub>6</sub>Na [M+Na]<sup>+</sup> 298.1266, found 298.1292.

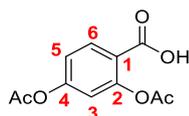
2-[[*Tert*-butoxy]carbonyl]amino}3-ethoxy-3-oxopropanoic acid (**21**)

Compound **20** (10.13 g, 36.7 mmol) was added dropwise to a stirred solution of potassium hydroxide (2.05 g, 36.5 mmol) in ethanol (20 cm<sup>3</sup>), following which the mixture was stirred for 3 h at room temperature. 90% of the solvent was then removed under pressure before ethyl acetate and 1M sodium bicarbonate added, with the organic layer removed to eliminate any unreacted starting material. The aqueous layer was acidified with 5M potassium bisulfate, and the product extracted with ethyl acetate before the organic layer was dried over sodium sulfate and the solvent removed under pressure. The flask was then stored at -20 °C to furnish a white solid (4.32 g, 47%). <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>) δ<sub>H</sub> ppm:

## CHAPTER 7: EXPERIMENTAL

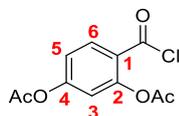
7.46 (d,  $J = 7.9$  Hz, 1 H, NH), 4.74 (d,  $J = 7.9$  Hz, 1 H, 1-H), 4.19 – 4.13 (m, 2 H, 2-H), 1.39 (s, 9 H, Boc), 1.20 (t,  $J = 7.0$  Hz, 3 H, 3-H).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_c$  ppm: 167.8 (acid  $\text{C}=\text{O}$ ), 167.2 (ester  $\text{C}=\text{O}$ ), 155.1 (Boc  $\text{C}=\text{O}$ ), 79.1 (Boc  $\text{C}(\text{CH}_3)_3$ ), 61.5 (2), 57.7 (1), 28.3 (3), 13.9 (4). IR  $\nu_{\text{max}}$  (neat)/  $\text{cm}^{-1}$  3266 (N-H), 2982 (C-H), 1753 (C=O), 1652 (C=O), 1398 (C-N), 1243 (C-O), 1183 (C-O). Mp 78-80 °C (lit 75-79 °C).<sup>3</sup>

### 2,4-Bis(acetyloxy) benzoic acid (**23**)

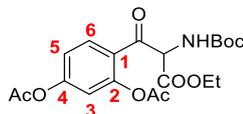


2,4-dihydroxybenzoic acid (2.77 g, 14.7 mmol) was dissolved in a solution of DMAP (16 mg, 0.13 mmol) in pyridine (12  $\text{cm}^3$ ). Following this, acetic anhydride (6  $\text{cm}^3$ , 63.4 mmol) was added drop-wise to the reaction mixture. The reaction vessel was then flushed with nitrogen, covered in foil and the reaction mixture stirred for 4 h at room temperature. Ice was added to the reaction mixture and the mixture was acidified with 3M HCl before being washed with EtOAc. The organic layer was dried over sodium sulfate, then the solvent was removed under reduced pressure immediately to furnish a white solid (3.46 g, 99%).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta_H$  ppm: 13.14 (s, 1 H, COOH), 8.00 (d,  $J = 8.7$  Hz, 1 H, 6-H), 7.18 (dd,  $J_1 = 8.6$  Hz,  $J_2 = 2.3$  Hz, 1 H, 5-H), 7.10 (d,  $J = 2.2$  Hz, 1 H, 3-H), 2.30 (s, 3 H,  $\text{CH}_3$ ), 2.26 (s, 3 H,  $\text{CH}_3$ ).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_c$  ppm: 169.0 (Acetyl  $\text{C}=\text{O}$ ), 168.6 (Acetyl  $\text{C}=\text{O}$ ), 165.0 (acid  $\text{C}=\text{O}$ ), 154.1 (2), 151.1 (4), 132.6 (6), 121.6 (1), 119.7 (5), 117.7 (3), 20.9 (Acetyl  $\text{CH}_3$ ), 20.8 (Acetyl  $\text{CH}_3$ ). IR  $\nu_{\text{max}}$  (neat)/  $\text{cm}^{-1}$  2593 (C-H), 1768 (C=O), 1679 (C=O), 1607 (Ar C=C), 1494 (Ar C=C), 1177 (C-O), 1138 (C-O). Mp 126-130 °C (lit 149-150 °C).<sup>3</sup>

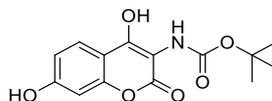
### 2,4-Bis(acetyloxy) benzoyl chloride (**24**)



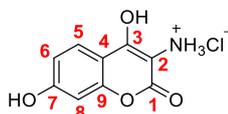
Compound **23** (3.35 g, 14.1 mmol) was taken up in anhydrous DCM (35  $\text{cm}^3$ ), following which thionyl chloride (21  $\text{cm}^3$ ) was added dropwise to the stirred mixture. The mixture was then refluxed for 4 h before the reagent and solvent were removed under reduced pressure to yield a sticky oil. This oil was then taken up in anhydrous DCM and used directly in the next procedure.

**Ethyl-3-[2,4-bis(acetyloxy)phenyl]-2-[[(*tert*-butoxy) carbonyl] amino]-3-oxopropanoate****(25)**

Compound **21** (3.89 g, 15.7 mmol) was added to a mixture of anhydrous THF (40 cm<sup>3</sup>), triethylamine (12.8 cm<sup>3</sup>, 91.2 mmol) and MgCl<sub>2</sub> (4.76 g, 50.0 mmol) and the slurry stirred for 2.5 h. Crude **24** was then added dropwise to this mixture, resulting in the grey suspension changing colour to an orange-brown suspension. The mixture was stirred for 15.5 h, following which the reaction was quenched with saturated NH<sub>4</sub>Cl causing the mixture to clarify. The product was extracted with ethyl acetate, dried over sodium sulfate, and the solvent removed under reduced pressure to yield a crude brown oil (3.11 g). This was used directly in the following reaction without any further purification.

***Tert*-butyl N-(4,7-dihydroxy-2-oxo-2H-chromen-3-yl) carbamate (26)**

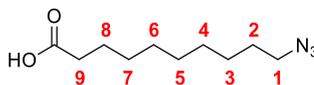
To crude **25** (3.11 g) was added a mixture of MeOH (3 cm<sup>3</sup>) and 1.5 M NaOH (4 cm<sup>3</sup>), following which the mixture was stirred for 3.5 h. The reaction was then acidified with 1 M HCl causing a precipitate to form. The product was extracted with ethyl acetate and the organic fraction dried over sodium sulfate. The solvent was then removed under pressure to yield a crude orange-brown solid (1.73 g) which was used directly in the next step without further purification.

**4,7-Dihydroxy-2-oxo-2H-chromen-3-aminium chloride (27)**

Crude **26** (0.295 g) was added to a mixture of ethereal hydrogen chloride (1 M, 2.9 cm<sup>3</sup>) and MeOH (2.1 cm<sup>3</sup>) and stirred at room temperature for 29 h. The resulting precipitate was then filtered and washed with a small amount of ethyl acetate to give the desired compound as a brown solid (214 mg, 8% over 3 steps). <sup>1</sup>H NMR (400 MHz, MeOD-d<sub>4</sub>) δ<sub>H</sub> ppm: 7.89 (d, *J* = 9 Hz, 1H, 8-H), 6.92 (dd, *J*<sub>1</sub> = 9 Hz, *J*<sub>2</sub> = 2 Hz, 1H, 6-H), 6.80 (d, *J* = 3 Hz, 1H, 5-H). <sup>13</sup>C NMR (100 MHz, MeOD-d<sub>4</sub>) δ<sub>C</sub> ppm: 164.6 (1), 162.3 (7), 161.6 (3), 155.9(4), 126.1

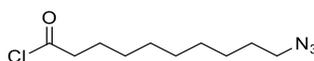
(5), 115.0 (6), 108.2 (2), 104.0 (8), 96.7(9). IR  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$ : 3412 (N-H), 3105 (O-H), 2847 (C-H), 1707 (C=O), 1639 (N-H bend), 1619 (C=C), 1530 (C=C). Mp 210-214 °C (lit 237-238 °C).<sup>3</sup>

### 10-Azidodecanoic acid (29)



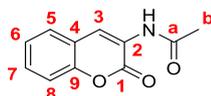
10-bromodecanoic acid (2.00 g, 1.99 mmol) and sodium azide (0.672 g, 10.4 mmol) were dissolved in DMF (16  $\text{cm}^3$ ). The resulting orange solution was then stirred at 50 °C for 24 h, after which the reaction was allowed to cool before water was added. The product was extracted using diethyl ether, with the organic fraction being washed with water exhaustively, dried with sodium sulfate and the solvent removed under pressure to yield a yellow oil (992 mg, 58 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 9.35 (br s, 1H, O-H), 3.17 (t,  $J$  = 7 Hz, 2H, 1-H), 2.24 (t,  $J$  = 7 Hz, 2H, 9-H), 1.52-1.40 (m, 4H, 2-H, 8-H), 1.26-1.13 (m, 10H, 3,4,5,6,7-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 180.3 (C=O), 51.4 (1), 34.3 (9), 29.2 (4), 29.10 (7), 29.05 (6), 29.0 (5), 28.8 (2), 26.7 (3), 24.8 (8). IR  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$ : 3411 (O-H), 2849 (C-H), 2091 (N=N=N), 1706 (C=O). HRMS (ESI+) calculated for  $\text{C}_{10}\text{H}_{18}\text{N}_3\text{NaO}_2$  [ $\text{M}+\text{Na}$ ]<sup>+</sup> 235.1301, found 235.1302.

### 10-Azidodecanoyl chloride (30)



To (19) (1.05 g, 4.92 mmol) was added dropwise thionyl chloride (15  $\text{cm}^3$ ). The mixture was then refluxed for 4 h before the reagent and solvent were removed under reduced pressure to yield a sticky orange-brown oil. This oil was used with no further purification.

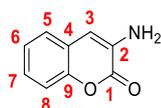
### N-(2-oxo-2H-chromen-3-yl) acetamide (34)



Salicylaldehyde (20.5 g, 168 mmol) was added to a mixture of N-acetylglycine (19.68 g, 168 mmol), anhydrous sodium acetate (53.83 g, 656 mmol) and acetic anhydride (100  $\text{cm}^3$ ), following which the mixture was allowed to stir at reflux for 5 h. The mixture was then allowed to cool to room temperature before being quenched and triturated with ice water. Small portions of ethyl acetate were added, and the resulting precipitate collected and

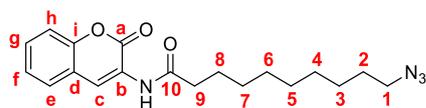
washed with small amounts of ethyl acetate to yield the product as a yellow solid (6.35 g, 19%).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta_{\text{H}}$  ppm: 9.76 (s, 1 H, NH), 8.61 (s, 1 H, 3-H), 7.70 (dd,  $J_1 = 8.0$  Hz,  $J_2 = 1.5$  Hz, 1 H, 5-H), 7.53-7.47 (m, 1 H, 8-H), 7.40 (d,  $J = 8$  Hz, 1 H, 6-H), 7.36 (t,  $J = 8$  Hz, 1 H, 7-H), 2.17 (s, 3 H, b-H).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_{\text{C}}$  ppm: 170.7 (a), 157.9 (1), 150.1 (4), 130.0 (8), 128.3 (5), 125.4 (7), 125.0 (2), 124.0 (3), 120.0 (9), 116.2 (6), 24.4 (b). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3329 (N-H), 1708 (C=O), 1680 (C=O), 1604 (C=C), 1516 (C=C). Mp. 203-205 °C (lit 203-204 °C).<sup>305</sup>

### 3-Amino-2H-chromen-2-one (32)



(38) (6.35 g, 31.2 mmol) was added to a 2:1 mixture of conc. HCl and EtOH (100  $\text{cm}^3$ ) before being stirred at reflux for 5 h. The solution was then allowed to cool to room temperature, before the pH of the solution was adjusted to pH 7 using NaOH. The resulting precipitate was filtered and desiccated to yield the product as a brown solid (0.82 g, 16%).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta_{\text{H}}$  ppm: 7.46-7.40 (m, 1 H, 5-H), 7.31-7.26 (m, 1 H, 8-H), 7.26-7.18 (m, 2 H, 6-H, 7-H), 6.72 (s, 1 H, 3-H), 5.70 (s, 2 H,  $\text{NH}_2$ ).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_{\text{C}}$  ppm: 159.1 (1), 148.4 (2), 133.6 (4), 125.9 (8), 125.3 (5), 125.0 (7), 122.2 (3), 115.9 (9), 108.5 (6). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3431 (N-H), 3322 (N-H), 1703 (C=O), 1637 (N-H), 1588 (C=C). Mp. 139-141 °C (lit 138-140 °C).<sup>306</sup>

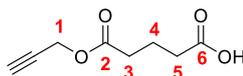
### 10-Azido-N-(4,7-dihydroxy-2-oxo-2H-chromen-3-yl)decanamide (35)



(36) (0.62 g, 3.85 mmol) was taken up in water (5  $\text{cm}^3$ ) and the pH adjusted to pH 6-8, before (34) (1.05 g, 4.92 mmol) in toluene (5  $\text{cm}^3$ ) was added to the stirring mixture. The resulting solution was stirred for 4 h with regular monitoring of pH, making adjustments if necessary using aliquots of potassium carbonate to keep the pH within the range of pH 6-8. Following this, 1M HCl was added to quench the reaction, and the product was extracted using toluene, washed with 1M NaOH and brine and dried. Removal of the solvent under vacuum yielded the product as an off-white solid (0.67 g, 49%).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta_{\text{H}}$  ppm: 9.69 (s, 1 H, NH), 8.65 (s, 1 H, c-H), 7.71 (dd,  $J_1 = 7.6$  Hz,  $J_2 = 1.4$  Hz, 1 H, e-H), 7.52-7.48 (m, 1 H, h-H), 7.42-7.32 (m, 2 H, f-H, g-H), 3.31 (masked by  $\text{H}_2\text{O}$ , 1-H), 2.49

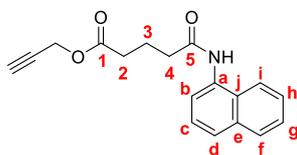
(masked by DMSO, 9-H), 1.62-1.48 (m, 4 H, 7-H, 2-H), 1.35-1.24 (m, 10 H, 8-H, 6-H, 5-H, 4-H, 3-H)  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_c$  ppm: 173.7 (10), 158.0 (b), 150.1 (a), 130.0 (h), 128.3 (e), 125.4 (g), 125.0 (d), 124.1 (c), 120.1 (i), 116.3 (f), 51.1 (1), 36.4 (9), 29.2, 29.1, 29.01, 29.00, 28.7 (2), 26.6 (3), 25.5 (7). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3330 (N-H), 2091 (N=N=N), 1709 (C=O), 1678 (C=O), 1630 (C=C), 1525 (C=C). Mp. 87-90 °C. HRMS (ESI+) calculated for  $\text{C}_{19}\text{H}_{24}\text{N}_4\text{O}_3$   $[\text{M}]^+$  355.1764, found 355.1755.

### 5-Oxo-5-(prop-2-yn-1-yloxy)pentanoic acid (38)



Glutaric anhydride (1.00 g, 8.78 mmol) was dissolved in propargyl alcohol (1  $\text{cm}^3$ , 17.2 mmol) and the mixture stirred at 105 °C for 30 minutes. The reaction was then allowed to cool, following which saturated sodium bicarbonate was added and excess propargyl alcohol removed by washing with diethyl ether. The reaction was brought to pH 3 by addition of dilute HCl, and the product extracted using diethyl ether. The organic layer was dried over sodium sulfate before the solvent was removed under reduced pressure to yield the product as a pale yellow oil (0.83 g, 56%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 9.30 (br s, 1H, O-H), 4.65 (d,  $J = 2$  Hz, 2H, O-CH $_2$ ), 2.47 (t,  $J = 3$  Hz, 1H, C $\equiv$ CH), 2.44-2.38 (m, 4H, 2-H, 4-H), 1.93 (qu,  $J = 7$  Hz, 2H, 3-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_c$  ppm: 178.9 (C=O), 172.2 (C=O), 76.8 (C $\equiv$ CH), 75.0 (C $\equiv$ CH), 52.0 (O-CH $_2$ ), 32.8 (2), 32.7 (4), 19.6 (3). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3290 (O-H), 2129 (C $\equiv$ C), 1733 (ester C=O), 1705 (acid C=O). HRMS (ESI+) calculated for  $\text{C}_8\text{H}_{10}\text{O}_4$   $[\text{M}]^+$  171.0657, found 171.0641.

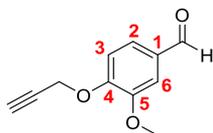
### Prop-2-yn-1-yl 4-[(naphthalen-1-yl)carbamoyl]butanoate (9)



(42) (0.70 g, 4.87 mmol), was taken up in thionyl chloride (1  $\text{cm}^3$ ), and the resulting solution was stirred under nitrogen for 2 h at room temperature. The solvent was subsequently removed under vacuum, before the residue was taken up in toluene (2  $\text{cm}^3$ ) and added to a stirred mixture of 1-naphthylamine (0.49 g, 3.42 mmol) in  $\text{H}_2\text{O}$  (2  $\text{cm}^3$ ) at pH 6-8, causing the pink solution to become yellow in colour. The reaction was stirred at room temperature for 4 h, with continual pH correction to pH 6-8, then quenched with 1 M HCl. The product was extracted using toluene, washed with 1 M NaOH, and the organic layer dried over

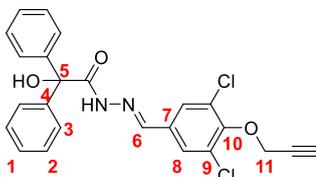
sodium sulfate before the solvent was removed under reduced pressure to yield the crude product. The product was purified using flash chromatography (0-100% EtOAc in DCM) to yield the pure product as a pink solid (0.15 g, 14 %).  $^1\text{H}$  NMR (400 MHz, MeOD- $d_4$ )  $\delta_{\text{H}}$  ppm: 7.99 (d,  $J = 9$  Hz, 1 H, b-H), 7.91 (d,  $J = 8$  Hz, 1 H, f-H), 7.79 (d,  $J = 8$  Hz, 1 H, d-H), 7.59-7.46 (m, 4 H, c-H, g-H, h-H, i-H), 4.73 (d,  $J = 3$  Hz, 2 H, OCH<sub>2</sub>), 2.93 (t,  $J = 3$  Hz, 1 H, C $\equiv$ CH), 2.65 (t,  $J = 7$  Hz, 2 H, 4-H), 2.56 (t,  $J = 7$  Hz, 2 H, 2-H), 2.13 (qu,  $J = 8$  Hz, 2 H, 3-H).  $^{13}\text{C}$  NMR (100 MHz, MeOD- $d_4$ )  $\delta_{\text{C}}$  ppm: 174.8 (5), 173.8 (1), 135.7 (a), 134.1 (j), 130.2 (e), 129.4 (f), 127.6 (d), 127.3 (h), 127.1 (g), 126.5 (i), 124.2 (c), 123.4 (b), 78.9 (C $\equiv$ CH), 76.2 (C $\equiv$ CH), 52.8 (OCH<sub>2</sub>), 36.2 (4), 34.0 (2), 22.1 (3). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3247 (alkyne C-H), 2102 (C $\equiv$ C), 1729 (ester C=O), 1648 (amide C=O), 1505 (C=C). Mp 106-110 °C. HRMS (ESI+) calculated for  $\text{C}_{18}\text{H}_{17}\text{NO}_3$   $[\text{M}]^+$  296.1287, found 296.1340.

### 3-Methoxy-4-(prop-2-yn-1-yloxy)benzaldehyde (44)



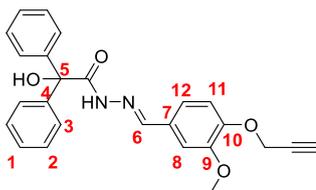
Vanillin (0.50 g, 3.31 mmol) and potassium carbonate (1.36 g, 9.86 mmol) were suspended in acetonitrile (8  $\text{cm}^3$ ) and propargyl chloride (0.29  $\text{cm}^3$ , 4.01 mmol) was added dropwise to this suspension. The reaction was then refluxed for 6 h, following which the solvent was removed under vacuum. The resulting solid was suspended in water and extracted with EtOAc, with the organic extract being washed with water and brine before being dried over sodium sulfate and the solvent removed under vacuum to yield the product as a yellow solid (0.29 g, 46%).  $^1\text{H}$  NMR (400 MHz, MeOD- $d_4$ )  $\delta_{\text{H}}$  ppm: 9.82 (s, 1 H, COH), 7.52 (dd,  $J_1 = 9$  Hz,  $J_2 = 2$  Hz, 1 H, 2-H), 7.46 (d,  $J = 2$  Hz, 1 H, 6-H), 7.21 (d,  $J = 8$  Hz, 1 H, 3-H), 4.88 (d,  $J = 3$  Hz, 2 H, OCH<sub>2</sub>), 3.89 (s, 3 H, OCH<sub>3</sub>), 3.03 (t,  $J = 2$  Hz, 1 H, C $\equiv$ CH).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 192.9 (COH), 153.9 (1), 151.5 (4), 132.3 (5), 127.0 (2), 114.4 (3), 111.2 (6), 79.0 (C $\equiv$ CH), 77.7 (C $\equiv$ CH), 57.5 (OCH<sub>2</sub>), 56.5 (OCH<sub>3</sub>). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3248 (alkyne C-H), 2110 (C $\equiv$ C), 1668 (C=O), 1587 (C=C), 1506 (C=C), 1261 (C-O), 1243 (C-O). Mp 76-80 °C (lit 81-82 °C).<sup>307</sup> HRMS (ESI+) calculated for  $\text{C}_{11}\text{H}_{10}\text{O}_3$   $[\text{M}]^+$  191.0708, found 191.0633.

**N'-[3,5-dichloro-4-(prop-2-yn-1-yloxy)phenyl]methylidene]-2-hydroxy-2,2-diphenylacetohydrazide (11)**



Benzilic acid hydrazide (0.20 g, 0.84 mmol) was dissolved in ethanol (5 cm<sup>3</sup>) and stirred at 70 °C for 20 minutes, following which 3,5-dichloro-4-(prop-2-yn-1-yloxy)benzaldehyde (0.19 g, 0.84 mmol) in ethanol (3 cm<sup>3</sup>) and three drops of acetic acid were added to the stirring mixture. The reaction was refluxed for 3 h, then cooled to allow a white precipitate to form. This precipitate was filtered off, washed with cold ethanol and hexanes and dried yielding the product as a white solid (89.4 mg, 23 %). <sup>1</sup>H NMR (400 MHz, MeOD-d<sub>4</sub>) δ<sub>H</sub> ppm: 8.27 (s, 1 H, 6-H), 7.87 (s, 2 H, 8-H), 7.52 (dd, *J*<sub>1</sub> = 8 Hz, *J*<sub>2</sub> = 1 Hz, 4 H, 3-H), 7.39-7.32 (m, 6 H, 1-H, 2-H), 4.86 (d, *J* = 2 Hz, 2 H, 11-H), 3.00 (t, *J* = 5 Hz, 1 H, C≡CH). <sup>13</sup>C NMR (100 MHz, MeOD-d<sub>4</sub>) δ<sub>C</sub> ppm: 171.9 (C=O), 151.2 (10), 146.4 (6), 143.2 (4), 132.5 (7), 130.1 (9), 127.6 (1), 127.53 (2), 127.50 (3), 127.48 (masked by 3, 8) 80.9 (5), 77.2 (C≡CH), 76.6 (C≡CH), 60.2 (11). IR ν<sub>max</sub> (neat)/cm<sup>-1</sup>: 3300 (O-H), 2100 (C≡C), 1665 (C=O), 1656 (C=N), 803 (C-Cl). Mp 203-205 °C. HRMS (ESI+) calculated for C<sub>24</sub>H<sub>19</sub>Cl<sub>2</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup> 453.0765, found 453.0773.

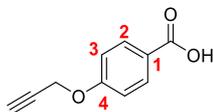
**N'-[3-methoxy-4-(prop-2-yn-1-yloxy)phenyl]methylidene]-2-hydroxy-2,2-diphenylacetohydrazide (12)**



Prepared as for (15), using (36) (0.17 g, 0.84 mmol) and benzylic acid hydrazide (0.21 g, 0.85 mmol) and yielding the product as an off-white solid (71.7 mg, 19 %). <sup>1</sup>H NMR (400 MHz, MeOD-d<sub>4</sub>) δ<sub>H</sub> ppm: 8.26 (s, 1 H, 6-H), 7.68 (d, *J* = 2 Hz, 1 H, 11-H), 7.44-7.38 (m, 4 H, 3-H), 7.36-7.28 (m, 6 H, 1-H, 2-H), 7.18 (dd, *J*<sub>1</sub> = 8 Hz, *J*<sub>2</sub> = 2 Hz, 1 H, 12-H), 7.08 (d, *J* = 8 Hz, 1 H, 8-H), 7.01 (d, *J* = 8.2 Hz, 11-H), 3.91 (s, 2 H, OCH<sub>2</sub>), 2.53 (s, 1 H, C≡CH). <sup>13</sup>C NMR (100 MHz, MeOD-d<sub>4</sub>) δ<sub>C</sub> ppm: 165.6 (C=O), 149.7 (9), 149.2 (6), 143.3 (10), 128.2 (4), 127.53 (3), 127.50 (2), 127.4 (1), 124.7 (7), 122.3 (12), 113.6 (8), 109.0 (11), 80.5 (5), 78.0 (C≡CH), 75.8 (C≡CH), 56.1 (OCH<sub>2</sub>), 55.0 (OCH<sub>3</sub>). IR ν<sub>max</sub> (neat)/cm<sup>-1</sup>: 3295 (alkyne C-H), 3100 (O-H), 2219

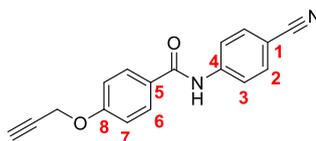
(C≡C), 1658 (C=O), 1635 (C=N), 1603 (C=C), 1509 (C=C). Mp 181-185 °C. HRMS (ESI+) calculated for C<sub>25</sub>H<sub>22</sub>N<sub>2</sub>O<sub>4</sub> [M]<sup>+</sup> 415.1658, found 415.1620.

#### 4-(Prop-2-yn-1-yloxy)benzoic acid (**49**)



4-hydroxybenzoic acid (2.06 g, 14.9 mmol) was dissolved in ethanol (80 cm<sup>3</sup>) and heated to 50 °C. Potassium iodide (2.49 g, 16.0 mmol) in 0.4 M NaOH (70 cm<sup>3</sup>) was then added to the solution before propargyl chloride (1.30 cm<sup>3</sup>) was added dropwise to the stirred mixture. The reaction was allowed to stir at room temperature for 72 h, following which 1 M HCl was added causing a white precipitate to form. This precipitate was collected by filtration, washed with water and dried to give (**53**) as a white solid (0.79 g, 27 %) <sup>1</sup>H NMR (400 MHz, MeOD-d<sub>4</sub>) δ<sub>H</sub> ppm: 7.90 (d, *J* = 9 Hz, 2 H, 2-H), 6.88 (d, *J* = 9 Hz, 2 H, 3-H), 4.90 (masked by H<sub>2</sub>O, OCH<sub>2</sub>), 2.96 (t, *J* = 3 Hz, 1 H, C≡CH). <sup>13</sup>C NMR (100 MHz, MeOD-d<sub>4</sub>) δ<sub>C</sub> ppm: 167.2 (C=O), 163.8 (1), 133.0 (2), 121.6 (4), 116.3 (3), 79.1 (C≡CH), 76.1 (C≡CH), 52.9 (OCH<sub>2</sub>). IR ν<sub>max</sub> (neat)/cm<sup>-1</sup>: 3270 (alkyne C-H), 3000 (O-H), 2135 (C≡C), 1674 (C=O), 1605 (C=C), 1485 (C=C). Mp 209-211 °C (lit 212-214 °C).<sup>308</sup>

#### N-(4-cyanophenyl)-4-(prop-2-yn-1-yloxy)benzamide (**13**)



(**53**) (0.52 g, 2.70 mmol), carbonyldiimidazole (CDI) (0.57 g, 3.51 mmol) and imidazole hydrochloride (0.42 g, 4.04 mmol) were stirred in anhydrous DMF (7 cm<sup>3</sup>) under nitrogen at room temperature for 5.5 h. The reaction was monitored for production of the imidazolidine intermediate by reaction of an aliquot of the reaction with butylamine for 5 mins and TLC. Upon adequate production of the imidazolidine intermediate, the reaction mixture was heated to 100 °C before 4-cyanoaniline (0.51 g, 4.31 mmol) was added and the reaction stirred under nitrogen at 100 °C for 18 h. The mixture was then cooled to room temperature and the reaction drowned in water causing a precipitate to form. The precipitate was filtered and dried under suction yielding (**17**) as a brown solid (0.24 g, 21 %). <sup>1</sup>H NMR (400 MHz, MeOD-d<sub>4</sub>) δ<sub>H</sub> ppm: 7.95-7.85 (m, 4 H, 2-H, 7-H), 7.70 (d, *J* = 10 Hz, 2 H, 3-H), 7.11 (d, *J* = 10 Hz, 2 H, 6-H), 4.82 (s, 2 H, OCH<sub>2</sub>), 3.31 (masked by MeOD, C≡CH). <sup>13</sup>C

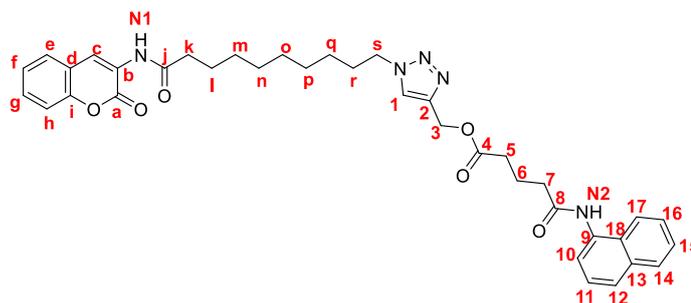
NMR (100 MHz, MeOD-d<sub>4</sub>)  $\delta_c$  ppm: 167.8 (C=O), 162.0 (5), 143.3 (C=N), 132.7 (3), 129.3 (6), 127.1 (8), 120.4 (2), 118.5 (4), 114.5 (7), 106.3 (1), 77.1 (C $\equiv$ CH), 75.3 (C $\equiv$ CH), 55.4 (OCH<sub>2</sub>). IR  $\nu_{\max}$  (neat)/cm<sup>-1</sup>: 3245 (alkyne C-H), 2228 (C $\equiv$ N), 2128 (C $\equiv$ C), 1646 (C=O), 1592 (C=C), 1506 (C=C). Mp 198-201 °C. HRMS (ESI+) calculated for C<sub>17</sub>H<sub>12</sub>N<sub>2</sub>O<sub>2</sub> [M]<sup>+</sup> 277.0977, found 277.0935.

### General procedure A for formation of triazoles using click chemistry

Alkyne (1 eq), azide (1.1 eq), sodium ascorbate (0.5 eq) and copper(II) sulfate pentahydrate (1 eq) were taken up in a 1:1 mixture of *tert*-butanol and water (5 cm<sup>3</sup>) and the mixture stirred at 135 °C for 19 h. Upon completion, the reaction was cooled to 45 °C and water added before the product was extracted in dichloromethane. The organics were dried and the solvent removed under vacuum to yield the crude product. Compounds were then purified using flash chromatography.

#### (1-{9-[(2-Oxo-2H-chromen-3-yl)carbamoyl]nonyl}-1H-1,2,3-triazol-4-yl)methyl

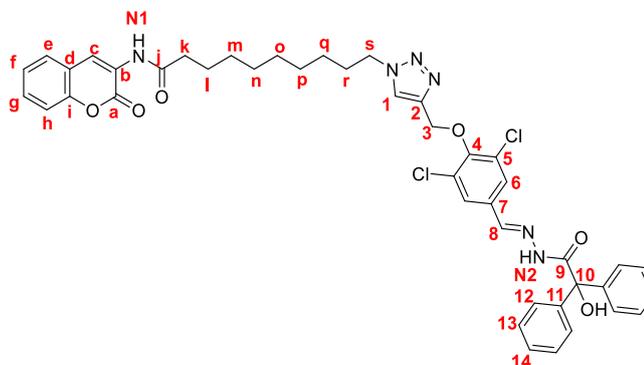
#### 4-[(naphthalen-1-yl)carbamoyl]butanoate (52)



Prepared using General Procedure A and purified using a gradient of 0-100 % EtOAc in DCM. Isolated as an off-white solid (33 mg, 12 %). <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>)  $\delta_H$  ppm: 9.89 (br s, 1 H, N2-H), 9.66 (br s, 1 H, N1-H), 8.64 (s, 1 H, c-H), 8.15 (s, 1 H, 1-H), 8.13-8.08 (m, 1 H, 10-H), 8.00-7.95 (m, 1 H, 14-H), 7.78-7.65 (m, 2 H, e-H, 12-H), 7.56-7.45 (m, 5 H, h-H, 11-H, 15-H, 16-H, 17-H), 7.42-7.30 (m, 2 H, f-H, g-H), 5.17 (s, 2 H, 3-H), 4.33 (t, *J* = 7.1 Hz, 2 H, s-H), 2.49 (masked by DMSO, k-H, 5-H, 7-H), 2.00-1.88 (m, 2 H, 6-H), 1.82-1.75 (m, 2 H, r-H), 1.60-1.52 (m, 2 H, l-H), 1.32-1.15 (m, 10 H, m-H, n-H, o-H, p-H, q-H). <sup>13</sup>C NMR (100 MHz, DMSO-d<sub>6</sub>)  $\delta_c$  ppm: 173.6 (j), 172.9 (4), 171.9 (8), 158.0 (a), 150.1 (b), 142.3 (2), 137.2 (18), 134.14 (9), 134.05 (16), 130.0 (17), 128.6 (e), 128.2 (14), 126.2 (11), 126.1 (13), 126.0 (h), 125.6 (12), 125.4 (f), 125.0 (1), 124.1 (c), 123.2 (10), 122.2 (d), 120.1 (i), 116.3 (g), 57.6 (3), 49.9 (s), 36.4 (k), 35.3 (5), 33.3 (7), 30.1 (r), 29.2, 29.0, 28.80, 28.75, 26.3, 25.5 (l), 21.1 (6). IR  $\nu_{\max}$  (neat)/cm<sup>-1</sup>: 3301 (N-H), 2917 (C-H), 1734 (C=O), 1720 (C=O), 1675 (C=O), 1661

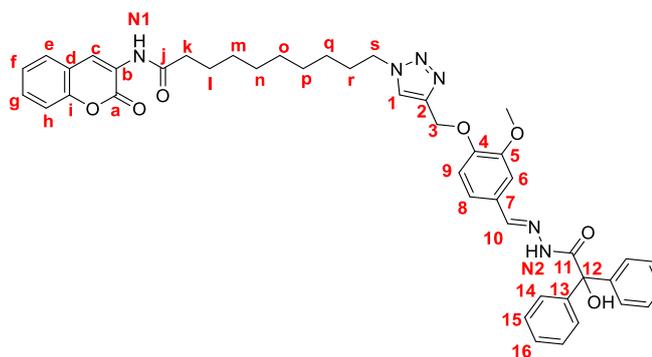
(C=C), 1600 (C=C), 1519 (C=C), 1160 (C-N). Mp 80-82 °C. HRMS (ESI+) calculated for  $C_{37}H_{41}N_5O_6$   $[M]^+$  652.3135, found 652.3163.

**10-[4-({2,6-Dichloro-4-[[2-hydroxy-2,2-diphenylacetamido]imino]methyl}phenoxy)methyl)-1H-1,2,3-triazol-1-yl]-N-(2-oxo-2H-chromen-3-yl)decanamide (53)**



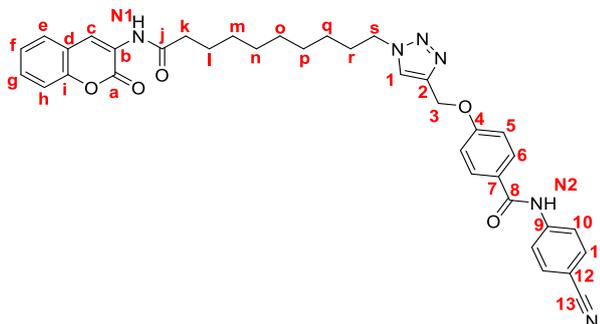
Prepared using General Procedure A and purified using a gradient of 0-100 % EtOAc in DCM. Isolated as a light brown solid (20 mg, 11 %).  $^1H$  NMR (400 MHz, DMSO- $d_6$ )  $\delta_H$  ppm: 11.78 (br s, 1 H, OH), 9.70 (br s, 1 H, N1-H), 8.67 (s, 1 H, c-H), 8.50 (s, 1 H, 8-H), 8.30 (s, 1 H, 1-H), 7.74-7.70 (m, 3 H, 6-H, e-H), 7.55-7.45 (m, 5 H, 11-H, f-H), 7.44-7.32 (m, 8 H, 13-H, 14-H, g-H, h-H), 7.08 (br s, 1 H, N2-H), 5.22 (s, 2 H, 8-H), 4.41 (t,  $J = 6.7$  Hz, 2 H, s-H), 2.48 (masked by DMSO, k-H), 1.90-1.78 (m, 2 H, r-H), 1.68-1.53 (m, 2 H, l-H), 1.39-1.17 (m, 10 H, m-H, n-H, o-H, p-H, q-H).  $^{13}C$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_C$  ppm: 173.7 (j), 170.1 (9), 158.2 (a), 151.4 (4), 150.2 (b), 145.8 (8), 144.1 (11), 142.1 (2), 133.1 (7), 130.0 (d), 129.9 (5), 128.2 (g), 128.1 (e), 127.92 (i), 127.87 (12), 127.5 (6), 125.6 (1), 125.4 (h), 125.0 (f), 124.1 (c), 120.1 (14), 116.2 (13), 81.0 (10), 66.7 (3), 49.8 (s), 36.4 (k), 30.2 (r), 29.2, 29.1, 29.0, 28.7, 26.1, 25.5 (l). IR  $\nu_{max}$  (neat)/ $cm^{-1}$ : 3315 (N-H), 3231 (N-H), 3061 (O-H), 2927 (C-H), 1713 (C=O), 1682 (C=N), 1628 (C=O), 1606 (C=C), 1516 (C=C), 1448 (O-H bend), 1179 (C-O), 753 (C-Cl), 698 (C-Cl). Mp 106-108 °C. HRMS (ESI+) calculated for  $C_{43}H_{42}Cl_2N_6O_6$   $[M]^+$  809.2621, found 809.2703.

**10-[4-({4-[[[(2-Hydroxy-2,2-diphenylacetamido)imino]methyl]-2-methoxyphenoxy]methyl}-1H-1,2,3-triazol-1-yl)-N-(2-oxo-2H-chromen-3-yl)decanamide (54)**



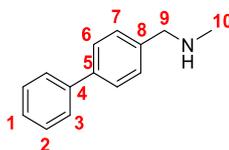
Prepared using General Procedure A and purified using a gradient of 0-10 % MeOH in DCM. Isolated as a green-brown solid (13 mg, 7 %).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta_{\text{H}}$  ppm: 11.38 (br s, 1 H, OH), 9.66 (br s, 1 H, N1-H), 8.65 (s, 1 H, c-H), 8.46 (s, 1 H, 10-H), 8.23 (s, 1 H, 1-H), 7.70 (d,  $J = 7.7$  Hz, 1 H, e-H), 7.53-7.41 (m, 6 H, g-H, h-H, 14-H), 7.41-7.27 (m, 9 H, f-H, g-H, 6-H, 15-H, 16-H), 7.21 (d,  $J = 7.2$  Hz, 1 H, 9-H), 7.12 (d,  $J = 7.1$  Hz, 1 H, 8-H), 6.98 (br s, 1 H, N2-H), 5.16 (s, 2 H, 3-H), 4.36 (t,  $J = 4.4$  Hz, 2 H, s-H), 3.77 (s, 3 H, OCH<sub>3</sub>), 2.49 (masked by DMSO, k-H), 1.85-1.76 (m, 2 H, r-H), 1.68-1.51 (m, 2 H, l-H), 1.33-1.14 (m, 10 H, m-H, n-H, o-H, p-H, q-H).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_{\text{C}}$  ppm: 173.6 (j), 169.7 (11), 158.4 (a), 150.5 (b), 149.7 (5), 149.0 (10), 144.5 (4), 143.1 (2), 131.4 (13), 130.1 (i), 128.5 (f), 128.3 (e), 128.2 (g), 128.1 (6), 128.0 (14), 125.4 (h), 125.1 (d), 124.7 (1), 124.3 (c), 122.2 (8), 120.1 (16), 116.4 (15), 113.5 (9), 108.5 (7), 81.0 (12), 62.3 (3), 55.6 (OCH<sub>3</sub>), 50.0 (s), 36.0 (k), 30.1 (r), 29.6, 28.7, 28.1, 26.7, 25.8, 25.5 (l). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3315 (N-H), 3232 (N-H), 3056 (O-H), 2928 (C-H), 1717 (C=O), 1668 (C=N), 1603 (C=C), 1505 (C=C), 1448 (O-H bend), 1265 (C-O), 1182 (C-O). Mp 102-104 °C. HRMS (ESI+) calculated for C<sub>44</sub>H<sub>46</sub>N<sub>6</sub>O<sub>7</sub> [M]<sup>+</sup> 771.3506, found 771.3590.

**N-(4-cyanophenyl)-4-[(1-{9-[(2-oxo-2H-chromen-3-yl)carbamoyl]nonyl}-1H-1,2,3-triazol-4-yl)methoxy]benzamide (55)**



Prepared using General Procedure A and purified using a gradient of 0-100 % EtOAc in DCM. Isolated as a white solid (10 mg, 8 %).  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-d}_6$ )  $\delta_{\text{H}}$  ppm: 10.43 (br s, 1 H, N2-H), 9.59 (br s, 1 H, N1-H), 8.57 (s, 1 H, c-H), 8.21 (s, 1 H, 1-H), 8.01-7.94 (m, 4 H, 6-H, 10-H), 7.76 (d,  $J = 8.0$  Hz, 2 H, 11-H), 7.63 (d,  $J = 8.1$  Hz, 1 H, e-H), 7.53-7.46 (m, 1 H, f-H), 7.34-7.24 (m, 2 H, g-H, h-H), 7.14 (d,  $J = 9.0$  Hz, 2 H, 5-H), 5.19 (s, 2 H, 3-H), 4.30 (t,  $J = 7.1$  Hz, 2 H, s-H), 1.86-1.76 (m, 2 H, r-H), 1.60-1.50 (m, 2-H, k-H), 1.31-1.17 (m, 12 H, l-H, m-H, n-H, o-H, p-H, q-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{DMSO-d}_6$ )  $\delta_{\text{C}}$  ppm: 173.7 (j), 165.9 (8), 161.5 (4), 158.0 (a), 150.1 (b), 144.2 (13), 142.6 (2), 133.5 (11), 130.3 (10), 130.0 (f), 128.2 (e), 127.1 (7), 125.4 (g), 125.02 (1), 124.96 (d), 124.1 (c), 120.5 (6), 120.1 (i), 119.6 (12), 116.3 (h), 114.9 (5), 105.5 (9), 61.8 (3), 49.9 (s), 36.4, 30.1 (r), 29.2, 29.1, 29.0, 28.8, 26.2 (k), 25.4. IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3374 (N-H), 2938 (C-H), 2221 ( $\text{C}\equiv\text{N}$ ), 1698 (C=O), 1682 (C=O), 1604 (C=C), 1505 (C=C). Mp 190-193 °C. HRMS (ESI+) calculated for  $\text{C}_{36}\text{H}_{36}\text{N}_6\text{O}_5$   $[\text{M}]^+$  633.2825, found 633.2756.

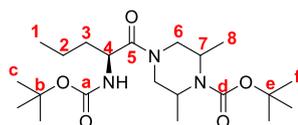
**([1,1'-Biphenyl]-4-yl)methyl(methyl)amine (15)**



4-bromo-N-methylbenzylamine (0.52 g, 2.60 mmol), phenylboronic acid (0.48 g, 3.93 mmol) and palladium(II) acetate (4 mol %) were taken up in water and the mixture stirred at reflux for 4 h under nitrogen. The reaction mixture was then filtered through celite to remove the palladium, and the filtrate was treated with 30 % NaOH. The product was thoroughly extracted in chloroform, the organic phases dried and the solvents removed under vacuum to yield the product as a yellow solid (0.36 g, 69 %).  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-d}_6$ )  $\delta_{\text{H}}$  ppm: 7.68-7.58 (m, 4 H, 6-H, 7-H), 7.49-7.32 (m, 5 H, 1-H, 2-H, 3-H), 3.68 (s, 2

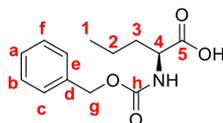
H, 9-H), 2.29 (s, 3 H, 10-H).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_{\text{C}}$  ppm: 140.6 (8), 140.5 (5), 138.9 (4), 129.4 (2 or 3), 129.0 (2 or 3), 127.7 (1), 127.0 (6 or 7), 126.8 (6 or 7), 55.2 (9), 36.1 (10). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3026 (N-H), 2795 (C-H), 1487 (C=C), 759 (C-H bend). Mp. 87-89 °C.

***Tert*-butyl-4-[(2S)-2-[[*tert*-butoxy]carbonyl]amino]pentanoyl]-2,6-dimethylpiperazine-1-carboxylate (63)**



Boc-protected L-norvaline (0.29 g, 1.33 mmol) was dissolved in a mixture of 1-Boc-2,6-dimethylpiperazine (0.25 g, 1.17 mmol), HOBt (0.25 g, 1.63 mmol), EDC (0.25 g, 1.63 mmol) and triethylamine (0.6  $\text{cm}^3$ ) in DCM (5  $\text{cm}^3$ ). The resulting solution was stirred at room temperature for 4 h before being quenched with water. The product was extracted with DCM, the organic phases dried over  $\text{Na}_2\text{SO}_4$  and the solvent removed under vacuum to yield the crude product. Purification was achieved using flash chromatography (0-99% MeOH in DCM, 1% TEA) to yield the pure product as a colourless oil (381 mg, 79 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 4.55-4.40 (m, 1 H, 6-H), 4.36-4.21 (m, 1 H, 6-H), 4.14-4.00 (m, 2 H, 7-H, 9-H), 3.79-3.70 (m, 1 H, 6-H), 3.21-3.04 (m, 1 H, 6-H), 2.74-2.60 (m, 1 H, 4-H), 1.72-1.22 (m, 22 H, c-H, f-H, 3-H, 2-H), 1.18 (d,  $J = 5.5$  Hz, 3 H, 8-H), 1.08 (d,  $J = 6.7$  Hz, 3 H, 8-H), 0.85 (t,  $J = 7.0$  Hz, 3 H, 1-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 172.5 (5), 155.4 (a/d), 154.3 (a/d), 80.0 (b/e), 79.4 (b/e), 49.5 (6), 46.8 (7), 46.5 (7), 46.4 (6), 45.9 (4), 35.3 (3), 28.2 (c/f), 28.1 (c/f), 19.9 (8), 19.8 (8), 18.4 (2), 13.6 (1). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3313 (N-H), 2973 (C-H), 1691 (C=O), 1639 (C=O), 1343 (C-N), 1168 (C-O). HRMS (ESI+) calculated for  $\text{C}_{21}\text{H}_{38}\text{N}_3\text{O}_5\text{Na}$  [ $\text{M}+\text{Na}$ ] $^+$  436.2787, found 436.2981.

**(2S)-2-[[*tert*-butoxy]carbonyl]amino]pentanoic acid (66)**



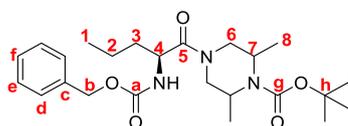
Sodium hydrogen carbonate (8.20 g, 97.6 mmol) was added to a mixture of L-norvaline (5.00 g, 32.5 mmol) in water (65  $\text{cm}^3$ ) and the reaction cooled to 0 °C. Benzyl chloroformate (7  $\text{cm}^3$ , 48.8 mmol) was added dropwise to the stirred mixture, and the reaction stirred at 0 °C for 45 minutes, followed by stirring at room temperature for 24 h. The aqueous layer was then washed with ethyl acetate before being taken to pH 1 with 1M HCl. The product

was extracted using ethyl acetate, the organic phase dried over sodium sulfate, and the solvent removed under vacuum to yield the product as a white solid (8.17 g, quantitative).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 10.10 (s, 1 H, OH), 7.28 (masked by  $\text{CDCl}_3$ , a-H, b-H, c-H, e-H, f-H), 5.09-4.98 (m, 2 H, g-H), 4.36-4.27 (m, 1 H, 4-H), 1.81-1.51 (m, 2 H, 3-H), 1.32 (se,  $J = 7$  Hz, 2 H, 2-H), 0.88-0.77 (m, 3 H, 1-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 204.5 (h), 187.0 (5), 166.6 (d), 157.6 (Ar-H), 157.1 (Ar-H), 156.9 (Ar-H), 95.8 (g), 83.4 (4), 63.0 (3), 48.4 (2), 42.0 (1). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3359 (N-H), 3322 (O-H), 2956 (C-H), 1739 (ester C=O), 1688 (acid C=O), 1648 (C=C), 1543 (C=C), 1265 (ester C-O). Mp 87-89 °C (lit 85-86 °C).<sup>309</sup>

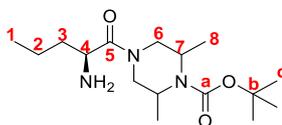
### *Tert*-butyl

#### 4-[(2S)-2-[[[(benzyloxy)carbonyl]amino]pentanoyl]-2,6-dimethylpiperazine-1-carboxylate

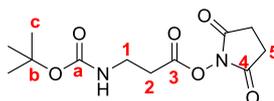
(67)



EDC (2.15 g, 11.2 mmol) was added to a mixture of **(68)** (2.00 g, 7.96 mmol), 1-Boc-2,6-dimethylpiperazine (1.91 g, 8.91 mmol), HOBT (1.71 g, 11.2 mmol) and TEA (2.25  $\text{cm}^3$ ) in DCM (10  $\text{cm}^3$ ), and the resulting mixture was stirred at room temperature for 4 h. The reaction was then quenched with water and the product extracted using DCM. The organic phases were dried over sodium sulfate and concentrated under vacuum to yield the crude product. The product was purified using flash chromatography (0-99% EtOAc in DCM, 1% TEA) to yield the product as a colourless oil (2.56 g, 72 %).  $^1\text{H}$  NMR (400 MHz,  $\text{MeOD-d}_4$ )  $\delta_{\text{H}}$  ppm: 7.26-7.14 (m, 5 H, Ar-H), 4.98 (s, 2 H, b-H), 4.55-4.47 (m, 1 H, 4-H), 4.26-4.12 (m, 1 H, 6-H), 4.09-4.01 (m, 2 H, 7-H), 3.77-3.67 (m, 1 H, 6-H), 3.27-3.12 (m, 1 H, 6-H), 2.73-2.64 (m, 1 H, 6-H), 1.59-1.24 (m, 13 H, i-H, 2-H, 3-H), 1.14 (d,  $J = 5.6$  Hz, 3 H, 8-H), 1.02 (d,  $J = 5.0$  Hz, 3 H, 8-H), 0.82 (t,  $J = 6.6$  Hz, 3 H, 1-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{MeOD-d}_4$ )  $\delta_{\text{C}}$  ppm: 172.7 (5), 156.9 (a), 154.7 (g), 136.8 (c), 128.8 (d), 128.1 (d), 127.7 (e), 127.6 (e), 127.4 (f), 80.0 (h), 66.4 (b), 50.1 (4), 49.1 (7), 48.9 (7), 46.8 (7), 45.8 (6), 33.8 (3), 27.5 (i), 18.9 (8), 18.8 (8), 18.6 (2), 12.9 (1). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3298 (N-H), 2973 (C-H), 1688 (C=O), 1638 (C=O), 1342 (C-N), 1089 (C-O). HRMS (ESI+) calculated for  $\text{C}_{24}\text{H}_{37}\text{N}_3\text{O}_5$   $[\text{M}]^+$  448.2812, found 448.2878.

***Tert*-butyl 4-[(2*S*)-2-aminopentanoyl]-2,6-dimethylpiperazine-1-carboxylate (64)**

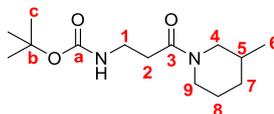
(69) (2.56 g, 5.72 mmol) was dissolved in methanol (10 cm<sup>3</sup>) and palladium on activated charcoal (20 wt %) was added. The reaction was stirred at room temperature under H<sub>2</sub> (1 atm) for 20 minutes before the palladium was removed by filtration through celite. The mixture was then concentrated to yield the product as a colourless oil (1.65 g, 92 %). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ<sub>H</sub> ppm: 4.43-4.27 (m, 1 H, 4-H), 4.26-4.15 (m, 2 H, 7-H), 3.89-3.60 (m, 1 H, 6-H), 3.42-3.38 (m, 1 H, 6-H), 3.36-3.31 (m, 1 H, 6-H), 2.94-2.81 (m, 1 H, 6-H), 2.15-2.10 (m, 2 H, 3-H), 1.84-1.38 (m, 11 H, c-H, 2-H), 1.25 (d, *J* = 6.0 Hz, 3 H, 8-H), 1.18 (d, *J* = 4.6 Hz, 3 H, 8-H), 0.93 (t, *J* = 5.9 Hz, 3 H, 1-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ<sub>C</sub> ppm: 175.5 (5), 154.4 (a), 80.0 (b), 50.6 (6), 49.3 (6), 47.0 (7), 46.6 (7), 46.4 (4), 37.4 (3) 28.4 (c), 20.1 (8), 19.8 (8), 18.9 (2), 13.9 (1). IR ν<sub>max</sub> (neat)/cm<sup>-1</sup>: 3364 (N-H), 2970 (C-H), 1685 (C=O), 1642 (C=O), 1453 (Me C-H). HRMS (ESI+) calculated for C<sub>16</sub>H<sub>32</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup> 314.2444, found 314.2397.

**2,5-Dioxopyrrolidin-1-yl 3-[[*tert*-butoxy]carbonyl]amino}propanoate (70)**

N-boc-β-alanine (2.00 g, 10.6 mmol) was dissolved in DCM (20 cm<sup>3</sup>) and cooled in an ice bath. N-hydroxysuccinimide (1.50 g, 13.0 mmol) followed by DCC (3.02 g, 14.6 mmol) was added, and the reaction was then stirred in an ice bath for 15 minutes and at room temperature for a further 45 minutes. Following this, the reaction was cooled once again in an ice bath, resulting in the formation of a precipitate. The precipitate was removed by vacuum filtration, washed with cold DCM and then discarded. The filtrate was washed with water and saturated NaHCO<sub>3</sub> before being dried over sodium sulfate. The organic phases were then concentrated to a syrup, and cold diethyl ether was added causing the formation of a precipitate. The precipitate was collected by filtration, washed with cold hexanes and diethyl ether, and dried, yielding the product as a white solid (2.61 g, 91 %). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ<sub>H</sub> ppm: 5.14 (br s, 1 H, NH), 3.55-3.46 (m, 2 H, 1-H), 2.90-2.80 (m, 6 H, 2-H, 5-H), 1.43 (s, 9 H, c-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ<sub>C</sub> ppm: 169.0 (4), 167.6 (3), 155.7 (a), 79.7

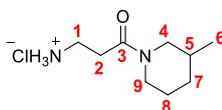
(b), 36.1 (1), 32.1 (2), 28.3 (c), 25.6 (5). IR  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$ : 3354 (N-H), 2982 (C-H), 1822 (C=O), 1781 (C=O), 1729 (C=O), 1687 (N-O). Mp 100-102 °C.

***Tert*-butyl N-[3-(3-methylpiperidin-1-yl)-3-oxopropyl]carbamate (72)**



(72) (2.12 g, 7.83 mmol) was dissolved in anhydrous acetonitrile (300  $\text{cm}^3$ ), and 3-methylpiperidine (1.8  $\text{cm}^3$ ) was added to this solution. The reaction was stirred at reflux under a  $\text{N}_2$  atmosphere for 24 h before being concentrated and the residue dissolved in DCM. The organic phase was then washed with  $\text{H}_2\text{O}$  and 1.5 N HCl before being dried over sodium sulfate and concentrated under vacuum to yield the crude product. The product was purified using flash chromatography (0-99% DCM in hexanes, 1 % TEA) to yield the pure product as a yellow oil (0.41 g, 19 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 5.55 (br s, 1 H, NH), 4.35-4.23 (m, 1 H, 7-H), 3.64 (dd,  $J_1 = 12.6$  Hz,  $J_2 = 3.2$  Hz, 1 H, 4-H), 3.31 (q,  $J = 5.8$  Hz, 2 H, 1-H), 2.63-2.52 (m, 1 H, 4-H), 2.44 (t,  $J = 6.0$  Hz, 2 H, 2-H), 1.78-1.70 (m, 1 H, 9-H), 1.68-1.55 (m, 1 H, 8-H), 1.54-1.42 (m, 1 H, 5-H), 1.39-1.29 (m, 10 H, c-H, 8-H), 1.13-0.99 (m, 1 H, 9-H), 0.84 (t,  $J = 6.8$  Hz, 3 H, 6-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 169.4 (3), 158.8 (a), 78.4 (b), 52.6 (4), 48.7 (7), 36.3 (1), 33.1 (2), 32.7 (9), 31.4 (5), 28.2 (c), 25.6 (8), 18.8 (6). IR  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$ : 3330 (N-H), 2976 (C-H), 1708 (C=O), 1627 (C=O), 1441 (methyl C-H bend), 1248 (C-O), 1168 (C-N). HRMS (ESI+) calculated for  $\text{C}_{14}\text{H}_{25}\text{N}_2\text{O}_3\text{Na}$   $[\text{M}+\text{Na}]^+$  293.1841, found 293.1793.

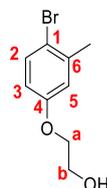
**3-(3-Methylpiperidin-1-yl)-3-oxopropan-1-aminium chloride (17)**



(74) (0.87 g, 3.22 mmol) was dissolved in a mixture of ethereal HCl (1M, 10  $\text{cm}^3$ ) and ethyl acetate (6  $\text{cm}^3$ ), and the solution was stirred for 28 h at room temperature resulting in the formation of a white precipitate. This precipitate was collected by filtration, washed with ethyl acetate, and dried under suction to yield the product as a white solid (0.33 g, 58 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 8.20 (br s, 3 H, NH<sub>3</sub>), 4.80-4.76 (m, 1 H, 9-H), 4.32-4.18 (m, 1 H, 8-H), 3.73-3.55 (m, 1 H, 9-H), 3.40-3.27 (m, 2 H, 2-H), 3.01-2.84 (m, 2 H, 1-H), 2.77-2.58 (m, 1 H, 8-H), 1.81-1.77 (m, 1 H, 7-H), 1.72-1.36 (m, 3 H, 4-H, 5-H), 1.17-1.03 (m, 1 H, 7-H), 0.92-0.81 (m, 3 H, 6-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 168.8 (3), 49.2 (9), 42.5 (8),

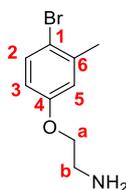
36.7 (1), 32.8 (7), 29.7 (2), 25.6 (4), 24.7 (5), 19.0 (6). IR  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$ : 2947 (N-H) 1625 (C=O), 1437 (Me C-H bend), 1251 (C-N). Mp 141-143 °C. HRMS (ESI+) calculated for  $\text{C}_9\text{H}_{18}\text{N}_2\text{O}$   $[\text{M}]^+$  171.1497, found 171.1419.

### 2-(4-Bromo-3-methylphenoxy)ethan-1-ol (**75**)



4-bromo-3-methylphenol (1.96 g, 10.5 mmol) and ethylene carbonate (2.03 g, 41.7 mmol) were dissolved in dry toluene (75  $\text{cm}^3$ ), following which  $\text{K}_2\text{CO}_3$  (1.85 g, 21.0 mmol) was added to the solution. The mixture was stirred at 115 °C for 24 h before being cooled to room temperature. Water was then added, and the product was extracted using ethyl acetate. The organic phase was dried over sodium sulfate and the solvents were removed under vacuum to yield the crude product. Purification was achieved by flash chromatography (2:1 hexanes:EtOAc), yielding the pure product as a yellow oil (1.44 g, 59 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 7.44 (d,  $J = 8.8$  Hz, 1 H, 2-H), 6.84 (d,  $J = 3.0$  Hz, 1 H, 5-H), 6.67 (dd,  $J_1 = 8.7$  Hz,  $J_2 = 3.0$  Hz, 1 H, 3-H), 4.08-4.04 (m, 2 H, a-H), 3.98-3.95 (m, 2 H, b-H), 2.39 (s, 3 H, CH<sub>3</sub>).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 157.8 (4), 139.0 (6), 133.0 (2), 117.2 (5), 115.9 (1), 113.6 (3), 69.4 (a), 61.4 (b), 23.1 (CH<sub>3</sub>). IR  $\nu_{\max}$  (neat)/ $\text{cm}^{-1}$ : 3293 (O-H), 1589 (C=C), 1477 (C=C), 1450 (Me C-H), 1236 (C-O), 1075 (C-O), 638 (C-Br). HRMS (ESI+) calculated for  $\text{C}_9\text{H}_{12}\text{BrO}_2$   $[\text{M}+\text{H}]^+$  231.0021, found 230.9703.

### 2-(4-Bromo-3-methylphenoxy)ethan-1-amine (**18**)



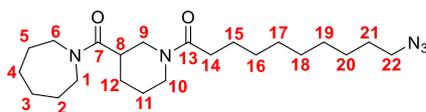
DIAD (1.2  $\text{cm}^3$ ) in dry THF (16  $\text{cm}^3$ ) was added to a stirred solution of (**77**) (1.40 g, 6.06 mmol), phthalimide (0.90 g, 6.12 mmol) and  $\text{PPh}_3$  (1.59 g, 6.06 mmol) in dry THF (16  $\text{cm}^3$ ). The reaction was stirred under a  $\text{N}_2$  atmosphere at room temperature for 4 h and monitored by TLC (DCM mobile phase). Upon completion, the reaction was concentrated under vacuum and the residue taken up in methanol, causing a precipitate to form. This precipitate was collected by vacuum filtration, washed with methanol and then dried. This

precipitate was then taken up in dry THF (32 cm<sup>3</sup>) under a N<sub>2</sub> atmosphere and hydrazine (2.4 cm<sup>3</sup>) was added to this solution. The reaction was stirred at reflux for 24 h before being quenched using a saturated NaHCO<sub>3</sub> solution. The product was extracted with diethyl ether, followed by acidification and extraction using 10 % HCl. The aqueous phase was then neutralised with excess sodium hydroxide before extracting once again using diethyl ether. The organic phases were dried over K<sub>2</sub>CO<sub>3</sub> before concentration under vacuum to yield the product as an off-white solid (0.46 g, 33 %). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ<sub>H</sub> ppm: 7.40 (d, *J* = 8.5 Hz, 1 H, 2-H), 6.80 (d, *J* = 2.8 Hz, 1 H, 5-H), 6.63 (dd, *J*<sub>1</sub> = 8.7 Hz, *J*<sub>2</sub> = 3.0 Hz, 1 H, 3-H), 3.94 (t, *J* = 5.2 Hz, 2 H, a-H), 3.06 (t, *J* = 5.0 Hz, 2 H, b-H), 2.36 (s, 3 H, CH<sub>3</sub>), 1.43-1.29 (m, 2 H, NH<sub>2</sub>). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ<sub>C</sub> ppm: 158.1 (4), 138.9 (1), 132.8 (2), 117.1 (5), 115.6 (3), 113.5 (6), 70.3 (a), 41.5 (b), 23.1 (CH<sub>3</sub>). IR ν<sub>max</sub> (neat)/cm<sup>-1</sup>: 1592 (N-H), 1578 (C=C), 1478 (C=C), 1463 (Me C-H), 1238 (C-O). Mp 84-86 °C.

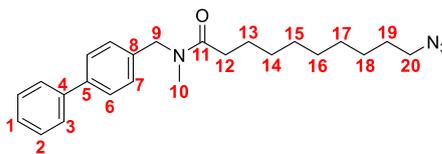
#### General Procedure B for the formation of amine library-linker compounds

To a stirred solution of amine (1 eq), acid (1.1 eq), HOBt (1.4 eq) and triethylamine (1.5 eq) in DCM was added EDC (1.4 eq). The resulting mixture was stirred for 19 h at room temperature before the reaction was quenched with water. The product was then extracted with DCM, dried over sodium sulfate, and concentrated in vacuo to yield the crude product. This crude product was then purified using flash chromatography to yield the pure product.

#### 1-[3-(Azepane-1-carbonyl)piperidin-1-yl]-10-azidodecan-1-one (76)

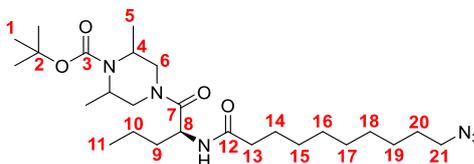


Prepared using General Procedure B and purified using a gradient of 0-9% MeOH in DCM, 1% TEA. Isolated as a brown solid (0.13 g, 25 %). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ<sub>H</sub> ppm: 4.49-4.38 (m, 1 H, 8-H), 3.77 (t, *J* = 16 Hz, 1 H, 10-H), 3.53-3.31 (m, 4 H, 1-H, 6-H), 3.18 (t, *J* = 8.0 Hz, 2 H, 22-H), 2.98 (dt, *J*<sub>1</sub> = 13.5 Hz, *J*<sub>2</sub> = 2.2 Hz, 1 H, 10-H), 2.63-2.38 (m, 2 H, 9-H), 2.30-2.15 (m, 2 H, 21-H), 1.89-1.34 (m, 16 H, 2-H, 3-H, 4-H, 5-H, 11-H, 12-H, 14-H, 20-H), 1.31-1.17 (m, 10 H, 15-H, 16-H, 17-H, 18-H, 19-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ<sub>C</sub> ppm: 172.6 (13), 171.7 (7), 51.4 (22), 47.4 (1 or 6), 46.2 (10), 46.1 (1 or 6), 44.8 (9), 41.9 (8), 33.3 (21), 29.3, 29.2, 29.0, 28.7, 28.5, 27.7, 27.4, 27.3, 26.9, 26.6, 25.6, 25.4, 25.2. IR ν<sub>max</sub> (neat)/cm<sup>-1</sup>: 2919 (C-H), 2091 (N=N=N), 1647 (C=O), 1626 (C=O). Mp 53-56 °C. HRMS (ESI+) calculated for C<sub>22</sub>H<sub>39</sub>N<sub>5</sub>O<sub>2</sub> [M]<sup>+</sup> 406.3182, found 406.3119.

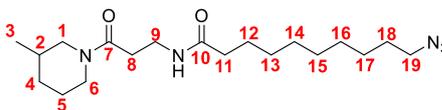
**10-Azido-N-([1,1'-biphenyl]-4-yl)methyl)-N-methyldecanamide (77)**

Prepared using General Procedure B and purified using a gradient of 0-100 % EtOAc in DCM. Isolated as a yellow oil (0.26 g, 25 %).  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta_{\text{H}}$  ppm: 7.78-7.60 (m, 4 H, 2-H, 6-H), 7.49-7.43 (m, 2 H, 3-H), 7.31-7.25 (m, 3 H, 1-H, 7-H), 4.55-4.51 (m, 2 H, 9-H), 3.32 (masked by HDO, 10-H), 2.96-2.92 (m, 2 H, 20-H), 1.58-1.45 (m, 4 H, 12-H, 19-H), 1.35-1.18 (m, 12 H, 13-H, 14-H, 15-H, 16-H, 17-H, 18-H).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta_{\text{C}}$  ppm: 173.5 (11), 139.6 (4), 139.5 (5), 139.3 (8), 127.7 (3), 127.2 (7), 126.0 (1), 125.9 (6), 125.7 (2), 52.0 (9), 32.5 (10), 32.1, 31.7 (20), 28.3, 28.2, 28.1, 28.0, 27.7 (12), 25.6, 23.9 (19). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 2926 (C-H), 2093 (N=N=N), 1736 (C=O), 1644 (C=C), 1488 (Me C-H). HRMS (ESI+) calculated for  $\text{C}_{24}\text{H}_{32}\text{N}_4\text{O}$   $[\text{M}]^+$  393.2654, found 393.2621.

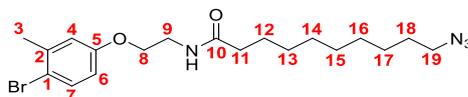
***Tert*-butyl-4-[(2S)-2-(10-azidodecanamido)pentanoyl]-2,6-dimethylpiperazine-1-carboxylate (78)**



Prepared using General Procedure B and purified using a gradient of 0-10 % MeOH in DCM. Isolated as a pale yellow oil (0.47 g, 18 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 5.10-4.94 (m, 1 H, 8-H), 4.45-4.30 (m, 1 H, 6-H), 4.28-4.20 (m, 2 H, 4-H), 3.28 (t,  $J = 7.0$  Hz, 2 H, 21-H), 2.89-2.80 (m, 1 H, 6-H), 2.28-2.16 (m, 2 H, 6-H), 1.87-1.53 (m, 7 H, 9-H, 13-H, 14-H, 20-H), 1.49 (s, 9 H, 1-H), 1.43-1.24 (m, 16 H, 5-H, 10-H, 14-H, 15-H, 16-H, 17-H, 18-H, 19-H), 1.21-1.15 (m, 3 H, 5-H), 1.00-0.87 (m, 3 H, 11-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 172.9 (7), 162.1 (12), 154.8 (3), 80.1 (2), 51.5 (21), 48.1 (8), 46.6 (4), 46.4 (4), 46.1 (6), 36.7 (6), 35.3 (9), 29.3, 29.18, 29.16, 29.06, 28.8, 28.4 (1), 26.7, 25.6, 20.2, 20.1 (5), 19.7 (5), 18.5 (10), 13.9 (11). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 2973 (C-H), 2092.1 (N=N=N), 1711 (ester C=O) 1684 (C=O), 1643 (C=O), 1344 (C-O), 1090 (C-N). HRMS (ESI+) calculated for  $\text{C}_{26}\text{H}_{47}\text{N}_6\text{O}_4\text{Na}$   $[\text{M}+\text{Na}]^+$  531.3635, found 531.3845.

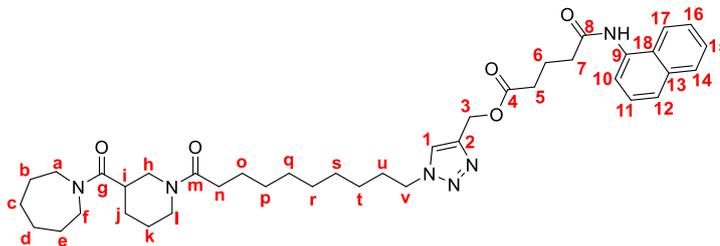
**10-Azido-N-[3-(3-methylpiperidin-1-yl)-3-oxopropyl]decanamide (79)**

Prepared using General Procedure B and purified using a gradient of 0-9 % MeOH in DCM, 1 % TEA. Isolated as a white solid (0.30 g, 31 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 4.38-4.33 (m, 1 H, 5-H), 3.73-3.57 (m, 1 H, 6-H), 3.56 (t,  $J = 5.7$  Hz, 2 H, 8-H), 3.27 (t,  $J = 7.0$  Hz, 2 H, 19-H), 2.97-2.86 (m, 1 H, 6-H), 2.66-2.59 (m, 1 H, 5-H), 2.51 (t,  $J = 5.6$  Hz, 2 H, 9-H), 2.14 (t,  $J = 7.5$  Hz, 2 H, 11-H), 1.86-1.49 (m, 7 H, 1-H, 2-H, 12-H, 18-H), 1.48-1.06 (m, 12 H, 4-H, 13-H, 14-H, 15-H, 16-H, 17-H), 0.91 (d,  $J = 6.7$  Hz, 3 H, 3-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 173.2 (10), 170.0 (7), 51.5 (19), 45.9 (6), 42.2 (5), 36.8 (11), 35.0 (8), 32.9 (1), 31.7 (9), 31.0 (18), 29.3 (4), 29.2, 29.1, 28.8, 26.7, 25.8, 25.7 (2), 24.7 (12), 18.8 (3). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3308 (N-H), 2920 (C-H), 2087 (N=N=N), 1638 (C=O), 1549 (C=O), 1442 (C-H bend). HRMS (ESI+) calculated for  $\text{C}_{19}\text{H}_{35}\text{N}_5\text{O}_2$   $[\text{M}]^+$  366.2869, found 366.2795. Mp 47-49 °C.

**10-Azido-N-[2-(4-bromo-3-methylphenoxy)ethyl]decanamide (80)**

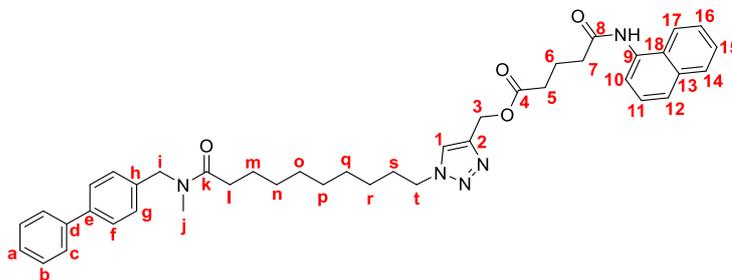
Prepared using General Procedure B and purified using a gradient of 0-100 % EtOAc in DCM. Isolated as a brown oil (0.47 g, 42 %).  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-d}_6$ )  $\delta_{\text{H}}$  ppm: 8.02 (br s, 1 H, NH), 7.44 (d,  $J = 8.6$  Hz, 1 H, 7-H), 6.95 (d,  $J = 2.8$  Hz, 1 H, 4-H), 6.74 (dd,  $J_1 = 8.8$  Hz,  $J_2 = 2.8$  Hz, 1 H, 6-H), 3.97 (t,  $J = 5.7$  Hz, 2 H, 8-H), 3.33 (masked by HDO, 9-H, 19-H), 2.30 (s, 3 H, 3-H), 2.08 (t,  $J = 7.4$  Hz, 2 H, 11-H), 1.55-1.42 (m, 4 H, 12-H, 18-H), 1.33-1.15 (m, 10 H, 13-H, 14-H, 15-H, 16-H, 17-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{DMSO-d}_6$ )  $\delta_{\text{C}}$  ppm: 177.7 (10), 163.1 (5), 143.5 (2), 137.8 (7), 122.5 (4), 119.8 (1), 119.3 (6), 71.8 (8), 55.8 (19), 43.3 (9), 40.5 (11), 34.0, 33.9, 33.8, 33.7, 33.4 (18), 31.3, 30.5 (12), 27.8 (3). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3315 (N-H), 2923 (C-H), 2095 (N=N=N), 1639 (C=O), 1609 (C=C), 1468 (C-H bend), 1288 (C-O), 1241 (C-N), 638 (C-Br). HRMS (ESI+) calculated for  $\text{C}_{19}\text{H}_{28}\text{BrN}_4\text{O}_2$   $[\text{M}]^+$  424.1248, found 424.1197.

**(1-{10-[3-(Azepane-1-carbonyl)piperidin-1-yl]-10-oxodecyl}-1H-1,2,3-triazol-4-yl)methyl  
4-[(naphthalen-1-yl)carbamoyl]butanoate (81)**



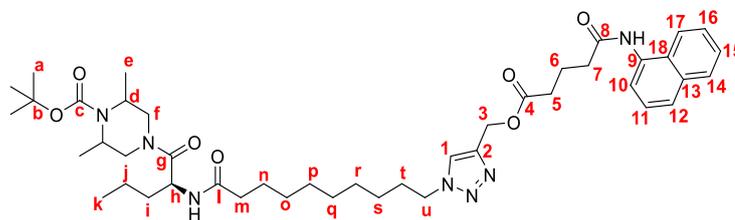
Prepared using General Procedure A and purified using a gradient of 0-9 % MeOH in DCM, 1% TEA. Isolated as a red-brown oil (64 mg, 30 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 8.84-8.49 (m, 1 H, 1-H), 7.83-7.50 (m, 4 H, Ar-H), 7.38-7.23 (masked by  $\text{CDCl}_3$ , Ar-H), 5.13 (s, 2 H, 3-H), 4.57-4.42 (m, 1 H, i-H), 4.15 (t,  $J = 6.9$  Hz, 2 H, v-H), 3.74-3.64 (m, 1 H, k-H), 3.50-3.25 (m, 4 H, a-H, f-H), 2.62-2.27 (m, 7 H, h-H, k-H, n-H, 5-H), 2.24-2.10 (m, 2 H, 7-H), 2.04-1.89 (m, 2 H, 6-H), 1.84-1.36 (b-H, c-H, d-H, e-H, j-H, l-H, t-H, u-H), 1.23-1.07 (m, 10 H, o-H, p-H, q-H, r-H, s-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 173.1 (m), 172.7 (4), 172.2 (8), 172.0 (g), 145.1 (2), 134.0, 132.70, 132.66, 128.4, 127.9, 126.0, 125.9, 125.51, 125.49 (1), 121.83, 121.77, 57.6 (3), 50.6 (v), 48.5 (a or f), 47.8 (k), 47.7 (a or f), 44.8 (h), 42.0 (i), 35.9 (5), 33.4 (7), 33.3 (u), 30.0 (j), 29.5 (l), 29.3, 28.8, 28.4, 27.7 (b-e), 27.3 (b-e), 27.0, 26.9 (b-e), 26.6, 25.5 (o), 25.4 (b-e), 24.5(b-e), 21.1 (6). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3281 (N-H), 2972 (C-H), 1734 (ester C=O), 1634 (C=O), 1599 (C=C), 1502 (C=C), 1532 (N-H bend), 1343 (C-N). HRMS (ESI+) calculated for  $\text{C}_{40}\text{H}_{56}\text{N}_6\text{O}_5$   $[\text{M}]^+$  701.4390, found 701.4362.

**(1-{9-[[{1,1'-Biphenyl]-4-yl)methyl](methyl)carbamoyl]nonyl}-1H-1,2,3-triazol-4-yl)methyl  
4-[(naphthalen-1-yl)carbamoyl]butanoate (82)**



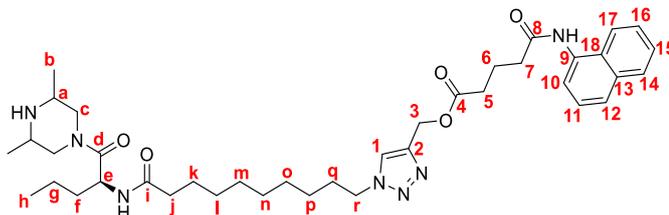
Prepared using General Procedure A and purified using a gradient of 0-10 % MeOH in DCM. Isolated as an off-white solid (92 mg, 49 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 8.41-8.27 (m, 1 H, 1-H), 7.89-7.83 (m, 2 H, b-H), 7.78-7.71 (m, 2 H, c-H), 7.62-7.56 (d, 2 H, f-H), 7.52-7.48 (m, 3 H, Ar-H), 7.41-7.31 (m, 4 H, Ar-H), 7.30-7.08 (m, 3 H, a-H, g-H), 5.19 (s, 2 H, 3-H), 4.58-4.43 (m, 3 H, j-H), 4.25-4.13 (m, 2 H, i-H), 3.19-3.12 (m, 1 H, t-H), 2.91-2.81 (m, 1 H, t-H), 2.49-2.41 (m, 2 H, l-H), 2.32-2.21 (m, 2 H, 5-H), 2.10-1.98 (m, 1 H, 7-H), 1.82-1.68 (m, 1 H, 7-H), 1.62-1.45 (m, 4 H, 5-H, 6-H), 1.29-1.13 (m, 12 H, m-H, n-H, o-H, p-H, q-H, r-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 173.6 (k), 173.2 (4), 172.6 (8), 140.8 (2), 139.6 (d), 139.2 (e), 139.1 (h), 137.9 (18), 136.6 (9), 133.0 (16), 132.5 (14), 131.7 (17), 129.0 (11), 128.4 (1), 127.7 (c), 127.2 (a), 126.7 (13), 125.9 (12), 125.2 (15), 124.3 (b), 123.6 (g), 123.6 (f), 121.3 (10), 53.1 (3), 51.6 (t), 50.7 (i), 34.2 (5), 33.2 (7), 33.1 (s), 29.9 (j), 29.1 (l), 28.9, 28.8, 27.9, 27.5, 26.9, 26.8, 24.8 (6). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3289 (N-H), 2930 (C-H), 1720 (ester C=O), 1658 (C=O), 1640 (C=O), 1599 (C=C), 1163 (C-O). HRMS (ESI+) calculated for  $\text{C}_{42}\text{H}_{49}\text{N}_5\text{O}_4$   $[\text{M}]^+$  688.3863, found 688.3936. Mp 91-93 °C.

***Tert*-butyl 2,6-dimethyl-4-[(2*S*)-2-(10-{4-[(naphthalen-1-yl)carbamoyl]butanoyl}oxy)methyl]-1*H*-1,2,3-triazol-1-yl]decanamido]pentanoyl]piperazine-1-carboxylate (**83**)**



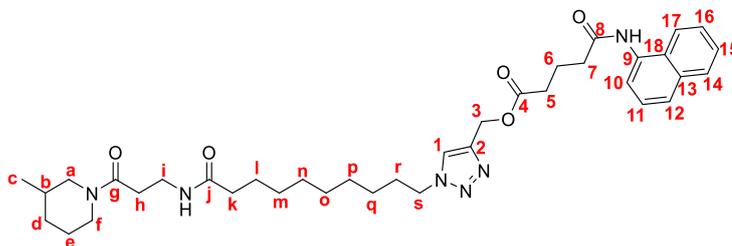
Prepared using General Procedure A and purified using a gradient of 0-9 % MeOH in DCM, 1 % TEA. Isolated as a yellow oil (160 mg, 26 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 8.67-8.60 (m, 1 H, 1-H), 7.94-7.88 (m, 1 H, Ar-H), 7.73-7.73 (m, 2 H, Ar-H), 7.61 (d,  $J = 8.2$  Hz, 1 H, Ar-H), 7.53-7.50 (m, 1 H, Ar-H), 7.43-7.35 (m, 2 H, Ar-H), 5.18 (s, 2 H, 3-H), 4.90-4.80 (m, 1 H, h-H), 4.22 (t,  $J = 7.2$  Hz, 2 H, u-H), 2.85-2.40 (m, 12 H, f-H, m-H, 5-H, 6-H, 7-H), 2.10-2.00 (m, 4 H, d-H, i-H), 1.81-1.72 (m, 2 H, t-H), 1.63-1.17 (m, 18 H, a-H, e-H, j-H, r-H, s-H), 1.08-0.93 (m, 13 H, e-H, n-H, o-H, p-H, q-H), 0.87-0.79 (m, 3 H, k-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 173.2 (4), 172.6 (l), 171.6 (8), 170.2 (g), 153.6 (c), 142.7 (2), 134.1 (18), 132.8 (9), 128.4, 127.7 (13), 126.1, 125.9 (1), 125.8, 125.6, 123.3, 121.7, 121.6, 80.4 (b), 57.6 (3), 52.0 (u), 51.3 (f), 50.7 (f), 50.4 (h), 48.5 (d), 48.4 (d), 36.7 (5), 35.8 (i), 35.0 (7), 33.2 (t), 30.1 (m), 29.1, 28.8, 26.3, 25.6 (s), 21.2 (6), 19.4 (a), 19.3 (e), 19.2 (e), 18.3, 13.9, 11.3 (j), 7.8 (k). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3281 (N-H), 2972 (C-H), 1734 (ester C=O), 1634 (C=O), 1598 (C=C), 1502 (C=C), 1343 (C-N), 1142 (C-O). HRMS (ESI+) calculated for  $\text{C}_{44}\text{H}_{65}\text{N}_7\text{O}_7$   $[\text{M}+\text{H}]^+$  804.5024, found 804.5148.

**[1-(9-(((2S)-1-(3,5-dimethylpiperazin-1-yl)-1-oxopentan-2-yl)carbamoyl)nonyl)-1H-1,2,3-triazol-4-yl]methyl 4-[(naphthalen-1-yl)carbamoyl]butanoate (84)**



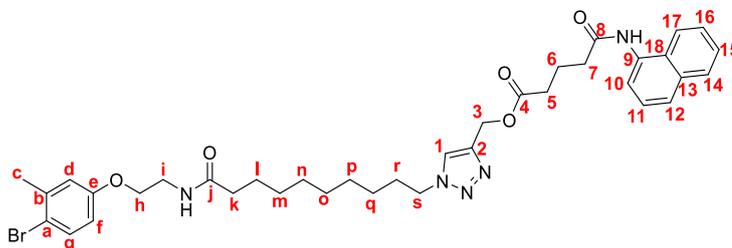
**(85)** (160 mg, 0.20 mmol) was taken up in ethyl acetate (2 cm<sup>3</sup>), following which ethereal HCl (1 M, 3.4 cm<sup>3</sup>) was added. The mixture was stirred overnight at room temperature, resulting in the formation of a sticky residue. Diethyl ether was then added, and the liquid phase was removed. The residue washed with ethyl acetate to isolate the product as a yellow syrup (141 mg, quantitative). <sup>1</sup>H NMR (400 MHz, MeOD-d<sub>4</sub>) δ<sub>H</sub> ppm: 8.46 (s, 1 H, 1-H), 8.02 (dt, *J*<sub>1</sub> = 7.5 Hz, *J*<sub>2</sub> = 2.0 Hz, 1 H, Ar-H), 7.93-7.89 (m, 1 H, Ar-H), 7.81 (d, *J* = 8.1 Hz, 1 H, Ar-H), 7.63-7.46 (m, 4 H, Ar-H), 4.83 (s, 2 H, 3-H), 4.68-4.56 (m, 2 H, r-H), 4.41-4.33 (m, 1 H, e-H), 3.65-3.50 (m, 2 H, 7-H), 3.29-3.16 (m, 2 H, 5-H), 2.81-2.61 (m, 2 H, 6-H), 2.59-2.46 (m, 2 H, j-H), 2.30-2.20 (m, 2 H, q-H), 2.14-1.96 (m, 4 H, c-H), 1.87-1.56 (m, 4 H, a-H, f-H), 1.51-1.25 (m, 20 H, b-H, g-H, k-H, l-H, m-H, n-H, o-H, p-H), 0.96 (t, *J* = 7.3 Hz, 3 H, h-H). <sup>13</sup>C NMR (100 MHz, MeOD-d<sub>4</sub>) δ<sub>C</sub> ppm: 174.3 (4), 174.0 (i), 173.6 (8), 173.5 (d), 145.0 (2), 134.3 (18), 132.8 (9), 128.8 (13), 128.0 (Ar), 126.3 (Ar), 126.0 (Ar), 125.8 (Ar), 125.3 (Ar), 125.1 (1), 122.8 (Ar), 122.1 (Ar), 53.4 (3), 52.6 (r), 52.3 (c), 52.0 (c), 49.1 (e), 46.5 (5), 44.7 (7), 35.2 (q), 34.9 (6), 33.1 (f), 32.7 (j), 29.2, 28.7, 28.5, 25.8 (a), 25.5 (a), 21.0, 20.8, 18.8, 18.6 (g), 12.6 (h), 7.9 (b), 6.4 (b). IR ν<sub>max</sub> (neat)/cm<sup>-1</sup>: 3269 (N-H), 2931 (C-H), 1731 (ester C=O), 1648 (C=O), 1598 (C=C), 1503 (C=C), 1462 (C-H bend), 1343 (C-N), 1156 (C-N). HRMS (ESI+) calculated for C<sub>39</sub>H<sub>56</sub>N<sub>7</sub>O<sub>5</sub>Na [M+Na]<sup>+</sup> 726.4319, found 726.4332.

**[1-(9-{{[3-(3-Methylpiperidin-1-yl)-3-oxopropyl]carbamoyl}nonyl)-1H-1,2,3-triazol-4-yl]methyl 4-{{(naphthalen-1-yl)carbamoyl}butanoate (85)}**



Prepared using General Procedure A and purified using a gradient of 0-10 % MeOH in DCM. Isolated as a brown solid (51 mg, 23 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 8.47-8.29 (m, 1 H, 1-H), 7.91-7.82 (m, 1 H, Ar-H), 7.73-7.66 (m, 2 H, Ar-H), 7.62-7.54 (m, 2 H, Ar-H), 7.53-7.42 (m, 1 H, Ar-H), 7.27 (masked by  $\text{CDCl}_3$ , Ar-H), 5.16 (s, 2 H, 3-H), 4.29-4.10 (m, 2 H, s-H), 3.51-3.36 (m, 3 H, f-H, i-H), 2.76-2.58 (m, 1 H, i-H), 2.56-2.24 (m, 8 H, e-H, k-H, 6-H, 7-H), 2.22-1.90 (m, 6 H, d-H, h-H, 5-H), 1.74-1.63 (m, 2 H, r-H), 1.61-1.49 (m, 1 H, b-H), 1.32-0.90 (m, 14 H, a-H, l-H, m-H, n-H, o-H, p-H, q-H), 0.83-0.72 (m, 3 H, c-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 173.3 (4), 173.1 (j), 171.5 (8), 170.0 (g), 142.9 (2), 134.1 (18), 132.1 (9), 128.5 (Ar), 127.7 (13), 126.1 (Ar), 125.9 (Ar), 125.8 (Ar), 125.6 (Ar), 123.4 (1), 121.6 (Ar), 121.4 (Ar), 57.6 (3), 52.8 (k), 50.4 (s), 49.1 (d), 45.9 (f), 42.2 (e), 36.7 (h), 35.8 (7), 35.0 (i), 33.2, 32.9 (6), 31.7, 31.0, 30.1 (r), 29.1, 28.8, 26.3 (q), 25.6 (b), 24.7, 21.2 (5), 18.9 (c). IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3272 (N-H), 2930 (C-H), 1727 (ester C=O), 1638 (C=O), 1627 (C=O), 1503 (C=C), 1455 (C-H bend), 1224 (C-N). HRMS (ESI+) calculated for  $\text{C}_{37}\text{H}_{52}\text{N}_6\text{O}_5$   $[\text{M}]^+$  661.4077, found 661.4144. Mp 134-136 °C.

**[1-(9-{{[2-(4-Bromo-3-methylphenoxy)ethyl]carbamoyl}nonyl)-1H-1,2,3-triazol-4-yl]methyl 4-{{(naphthalen-1-yl)carbamoyl}butanoate (86)}**



Prepared using General Procedure A and purified using a gradient of 0-10 % MeOH in DCM. Isolated as an off-white solid (67 mg, 36 %).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  ppm: 8.44 (s, 1 H, 1-H), 7.88-7.79 (m, 1 H, Ar-H), 7.77-7.67 (m, 2 H, Ar-H), 7.59 (d,  $J = 8.1$  Hz, 1 H, Ar-H), 7.45-7.24 (m, 4 H, d-H, Ar-H), 6.68-6.61 (m, 1 H, g-H), 6.50 (dd,  $J_1 = 8.7$  Hz,  $J_2 = 2.8$  Hz, 1 H, f-H),

5.16 (s, 2 H, 3-H), 4.19 (t,  $J = 7.1$  Hz, 2 H, s-H), 3.88-3.78 (m, 2 H, h-H), 3.50 (d,  $J = 5.1$  Hz, 2 H, i-H), 2.47-2.38 (m, 4 H, 6-H, 7-H), 2.25 (s, 3 H, c-H), 2.11-1.94 (m, 4 H, k-H, 5-H), 1.78-1.67 (m, 2 H, r-H), 1.54-1.43 (m, 2 H, l-H), 1.23-1.05 (m, 10 H, m-H, n-H, o-H, p-H, q-H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  ppm: 173.5 (4), 173.2 (j), 171.7 (8), 157.7 (e), 149.7 (2), 138.9 (b), 134.1 (18), 132.9 (9), 128.5 (Ar), 126.3 (13), 126.2 (Ar), 126.1 (Ar), 125.9 (Ar), 125.7 (1), 125.6 (d), 121.6 (Ar), 117.0 (g), 115.8 (Ar), 113.6 (Ar), 113.5 (f), 113.3 (a), 66.9 (h), 57.6 (3), 50.5 (s), 38.9 (i), 36.6 (5), 35.8 (7), 33.2 (6), 30.1 (r), 29.2, 29.1 (k), 28.9 (l), 28.7, 26.3 (q), 25.6, 23.1 (c), 21.2. IR  $\nu_{\text{max}}$  (neat)/ $\text{cm}^{-1}$ : 3290 (N-H), 2919 (C-H), 1720 (ester C=O), 1657 (C=O), 1645 (C=O), 1500 (C=C), 1242 (C-H bend), 1165 (C-N), 637 (C-Br). HRMS (ESI+) calculated for  $\text{C}_{37}\text{H}_{46}\text{BrN}_5\text{O}_5$   $[\text{M}]^+$  722.2748, found 722.2825. Mp 117-120 °C.

#### 7.4.2. Solid phase peptide synthesis

SPPS was performed, unless stated otherwise, under automation using a Biotage Syro I peptide synthesiser. Amino acid solutions were prepared at 0.5 M in NMP, with the exception of Arg which was dissolved in DMF. A mixture of HOBt and HBTU was prepared at 0.45 M HOBt and 0.44 M HBTU in DMF. A solution of 2 M DIPEA in NMP was also prepared, as well as a 20 % or 40 % piperidine solution in DMF. Resin was swelled using 2  $\text{cm}^3$  DMF with shaking for 20 minutes, before the DMF was removed. If required, the Fmoc protection of the resin was then removed using the previously prepared piperidine in DMF mixture (1.5  $\text{cm}^3$ ). The reaction was vortexed for 15 minutes before the solution was removed under suction. The resin was then washed three times with DMF. Four equivalents of amino acid solution were added, followed by four equivalents of HOBt/HBTU mixture and eight equivalents of DIPEA solution, before the reaction was allowed to continue with vortexing for 45 minutes. The solution was then drained before the process was repeated. Following this second coupling step, piperidine in DMF (1.5  $\text{cm}^3$ ) was added and the reaction allowed to continue with vortexing for 10 minutes. The solution was removed and another batch of piperidine in DMF (1.5  $\text{cm}^3$ ) added, with the reaction once again being allowed to continue with agitation for 10 minutes. Finally, the resin was washed three times with DMF (2  $\text{cm}^3$ ) before the coupling-deprotection cycle was started with the next amino acid. Manual SPPS also followed this procedure.



## **7.5. Biological procedures**

### **7.5.1. Agarose gel electrophoresis**

1 % (w/v) agarose was added to a pre-mixed solution of 40 mM Tris Base, 20 mM acetic acid and 1 mM EDTA (TAE buffer) and the mixture brought to the boil until the agarose had completely dissolved. The solution was allowed to cool briefly before being cast in a gel rack with comb. The gel was allowed to solidify for 1 hour before the comb was removed and the gel placed into an electrophoresis tank. The gel was covered with further TAE buffer and experimental samples were loaded into the wells.

### **7.5.2. BamD plasmid cloning**

Plasmids were cloned using Tpp10 cells. 1  $\mu\text{L}$  of a 100 ng/ $\mu\text{L}$  solution of pLeu3 plasmid was added to the cells before incubating the culture on ice for 30 minutes. The culture was then heat shocked at 42 °C for 70 seconds before being cooled on ice once again for 5 minutes. LB (700  $\mu\text{L}$ ) was added to the culture before the culture was allowed to grow at 37 °C with agitation (200 rpm) for 1 hour. Following incubation, the culture was centrifuged at 4000 rpm for 5 minutes, and all but roughly 200  $\mu\text{L}$  of the supernatant discarded and the pellet resuspended. A 20  $\text{cm}^3$  plate of 1.5 % Agar-LB with 0.1 % ampicillin was prepared by mixing 300 mg Agar with 20 mL of a pre-mixed LB-ampicillin mixture and heating in a microwave until all solid had dissolved. The mixture was then poured into a petri-dish and allowed to cool for 45 minutes, following which the resuspended pellet was plated. The culture was allowed to grow overnight at 37 °C with no agitation before 3 colonies were collected and each placed in 20  $\mu\text{L}$  of LB with 0.1 % ampicillin. These liquid cultures were then once again incubated at 37 °C with shaking before being centrifuged at 4000 rpm and the supernatant removed. A plasmid extraction kit was then used as per the manufacturer's instructions to obtain the plasmids.

### **7.5.3. BamD protein expression**

BamD was expressed using SB21 cells. 1  $\mu\text{L}$  of a 100 ng/ $\mu\text{L}$  solution of the previously prepared BamD plasmid was added to the cells before incubating the culture on ice for 30 minutes. The culture was then heat shocked at 42 °C for 70 seconds before being cooled on ice once again for 5 minutes. LB (700  $\mu\text{L}$ ) was added to the culture before the culture was allowed to grow at 37 °C with agitation (200 rpm) for 1 hour. Following incubation, the culture was centrifuged at 4000 rpm for 5 minutes, and all but roughly 200  $\mu\text{L}$  of the supernatant discarded and the pellet resuspended. A 20  $\text{cm}^3$  plate of 1.5 % Agar-LB with 0.1

## CHAPTER 7: EXPERIMENTAL

% ampicillin was prepared by mixing 300 mg Agar with 20 mL of a pre-mixed LB-ampicillin mixture and heating in a microwave until all solid had dissolved. The mixture was then poured into a petri-dish and allowed to cool for 45 minutes, following which the resuspended pellet was plated. The culture was allowed to grow overnight at 37 °C with no agitation before 4 colonies from each plate were collected and each placed in roughly 10 mL LB with 0.1 % ampicillin. These liquid cultures were once again incubated at 37 °C with agitation (200 rpm) overnight before a further 10 mL of 50 % glycerol was added, with these new glycerol stocks stored at -20 °C.

An aliquot of one glycerol stock was added to 300 mL LB with 0.1 % ampicillin and the culture grown overnight at 37 °C with agitation (200 rpm) before adding 25 mL of the culture to 1 L of LB with 0.1 % ampicillin. This larger culture was once again grown at 37 °C with agitation whilst monitoring the optical density. Once this reached 0.6-0.8, 200 µL of 1 M IPTG was added, and the culture incubated for a further 18 h. The culture was then centrifuged at 4500 rpm for 15 minutes and the supernatant discarded. The pellet was resuspended in a lysis buffer consisting of 20 mM Tris Base (pH 7.9), 10 mM imidazole and 300 mM sodium chloride and a small amount of DNAase1 and lysozyme, as well as a tablet of protease inhibitor, were added. Cells were mechanically lysed by passage through a Constant Systems TS Series cell disruptor at 30 kPa and 20 kPa, and the resulting solution was centrifuged at 20000 rpm for 30 minutes. The protein was then purified by nickel column chromatography, using a wash buffer consisting of 20 mM Tris Base (pH 7.9), 300 mM sodium chloride and 50 mM imidazole, and an elution buffer consisting of 20 mM Tris Base (pH 7.9), 290 mM sodium chloride and 300 mM imidazole. An aliquot of the eluted protein was assessed using SDS-PAGE to ensure that the protein had been expressed and was pure, before the protein solution was concentrated to less than 10 mL. The solution was then desalted by passage through a desalting column, following which protease was added to the solution and the mixture agitated overnight in a cold room. The mixture was once again passed through a nickel column using the lysis buffer described above before the solution was concentrated to less than 8.5 mL and passed through a gel filtration column. The relevant signal was identified by SDS-PAGE and the concentration of the solution determined through the use of a Thermo Scientific Nanodrop 2000. The protein was then snap frozen for storage at -80 °C.

#### 7.5.4. BamD buffer exchange

A gel filtration column was equilibrated with a buffer consisting of 20 mM Tris Base (pH 7.9) and 150 mM NaCl. The protein solution was concentrated to less than 3 mL before being loaded into the column. The buffer was then exchanged by elution with the new buffer, and the fractions containing the protein combined and concentrated to less than 8.5 mL. The concentration of the solution was determined using a Thermo Scientific Nanodrop 2000 before the solution was snap frozen for storage at -80 °C.

#### 7.5.5. Differential Scanning Fluorimetry buffer screening

A mixture of 5 µM BamD, 5X SYPRO orange, 25 µM peptide in DMSO and the relevant buffer (final volume 20 µL) were added to rotor gene-Q RT-PCR sample tubes. Samples were incubated at 25 °C for 30 minutes before being scanned at a ramp of 0.5 °C/min from 25 °C to 95 °C. The intensity of fluorescence at an excitation wavelength of 465 nm and an emission wavelength of 580 nm was recorded. All measurements were run in duplicate, alongside a duplicate of blank runs with no peptide added. The first derivative of the resulting melting curve was analysed to identify the point of inflexion, correlated to the melting temperature. The screened buffers included sodium acetate pH 5.0, pyridine pH 5.5, sodium cacodylate pH 6.0, MES pH 6.5, sodium citrate pH 6.5, imidazole pH 7.0, bis-tris pH 7.2, MOPS pH 7.2, HEPES pH 7.5, sodium phosphate pH 7.5, tricine pH 8.0, tris pH 8.0, glycine pH 9.0, borax pH 9.0 and CHES pH 9.5. Buffers contained 50 mM of buffer compound, 1 mM TCEP and 150 mM NaCl.

#### 7.5.6. *E. coli* gyrase supercoiling assay

*E. coli* gyrase supercoiling assay kits were purchased from Inspiralis. Dilution buffer consisted of 50 mM Tris.HCl (pH 7.5), 100 nM KCl, 2mM DTT, 1 mM EDTA and 50 % (w/v) glycerol. Supercoiling assay buffer consisted of 35 mM Tris.HCl (pH 7.5), 24 mM MgCl<sub>2</sub>, 2 mM DTT, 1.8 mM spermidine, and 1 mM ATP. All enzymes were stored at -80 °C. Relaxed pBR322 DNA was supplied as part of the assay kits. Assay kits were used following the manufacturer's instructions. Compounds were weighed into an Eppendorf tube and dissolved in DMSO. Serial dilutions were performed using DMSO. For each molecule to be tested a master mix was prepared using 0.5 µL of relaxed pBR322 DNA, 2 µL of dilution buffer, 2 µL of assay buffer and 16.5 µL of water. Diluted samples and the master mix were prepared in ice prior to incubation to reduce any intrinsic supercoiling activity. A 25 µL aliquot of master mix was used for the negative control and the volume made up to 30 µL

with 1  $\mu\text{L}$  of DMSO and 4  $\mu\text{L}$  of assay buffer. Following this, 4  $\mu\text{L}$  of enzyme was added for every compound to be tested including the positive control and known inhibitors (ciprofloxacin, simocyclinone D8). The master mix was homogenised, and 29  $\mu\text{L}$  of master mix was dispensed into an Eppendorf along with 1  $\mu\text{L}$  of compound to be tested. Samples were centrifuged at 13,000 rpm for 5 seconds before incubation at 37  $^{\circ}\text{C}$  for 30 minutes. The reaction was then stopped using 30  $\mu\text{L}$  of iso-amyl alcohol in chloroform (1:24) and 20  $\mu\text{L}$  of 40 % (w/v) sucrose, 0.1 M Tris.HCl (pH 8.0), 0.1 M EDTA and 0.5 mg/mL bromophenol blue (STEB buffer). The samples were once again centrifuged at 13,000 rpm for 2 minutes before 15  $\mu\text{L}$  of the aqueous layer was loaded onto a 1 % agarose gel. The topoisomers and supercoiled DNA were then separated by electrophoresis before the gel was stained in ethidium bromide in TAE buffer (1  $\mu\text{g}/\text{mL}$ ) for 30 minutes. Following this, the gel was de-stained using TAE buffer for 15 minutes before visualisation of the gel under UV light.

### 7.5.7. *E. coli* Topo IV supercoiling relaxation assay

*E. coli* Topo IV relaxation assay kits were purchased from Inspiralis. Dilution buffer consisted of 40 mM HEPES.KOH (pH 7.6), 100 mM potassium glutamate, 1 mM DTT, 1 mM EDTA and 40 % (v/v) glycerol. Assay buffer consisted of 50 mM HEPES.KOH (pH 7.6), 100 mM potassium glutamate, 10 mM magnesium acetate, 10 mM DTT, 1 mM ATP and 50  $\mu\text{g}/\text{mL}$  albumin. All enzymes were stored at -80  $^{\circ}\text{C}$ . Supercoiled pBR322 DNA was supplied as part of the assay kits. Assay kits were used following the manufacturer's instructions. Compounds were weighed into an Eppendorf tube and dissolved in DMSO. Serial dilutions were performed using DMSO. For each molecule to be tested a master mix was prepared using 0.5  $\mu\text{L}$  of supercoiled pBR322 DNA, 2  $\mu\text{L}$  of dilution buffer, 2  $\mu\text{L}$  of assay buffer and 16.5  $\mu\text{L}$  of water. Diluted samples and the master mix were prepared in ice prior to incubation to reduce any intrinsic supercoiling activity. A 25  $\mu\text{L}$  aliquot of master mix was used for the negative control and the volume made up to 30  $\mu\text{L}$  with 1  $\mu\text{L}$  of DMSO and 4  $\mu\text{L}$  of assay buffer. Following this, 4  $\mu\text{L}$  of enzyme was added for every compound to be tested including the positive control and known inhibitors (sparfloxacin, simocyclinone D8). The master mix was homogenised, and 29  $\mu\text{L}$  of master mix was dispensed into an Eppendorf along with 1  $\mu\text{L}$  of compound to be tested. Samples were centrifuged at 13,000 rpm for 5 seconds before incubation at 37  $^{\circ}\text{C}$  for 30 minutes. The reaction was then stopped using 30  $\mu\text{L}$  of iso-amyl alcohol in chloroform (1:24) and 20  $\mu\text{L}$  of 40 % (w/v) sucrose, 0.1 M Tris.HCl (pH 8.0), 0.1 M EDTA and 0.5 mg/mL bromophenol blue (STEB buffer). The samples were once again centrifuged at 13,000 rpm for 2 minutes before 15  $\mu\text{L}$  of the aqueous layer was loaded onto a 1 % agarose gel. The topoisomers and supercoiled DNA were then separated

by electrophoresis before the gel was stained in ethidium bromide in TAE buffer (1 µg/mL) for 30 minutes. Following this, the gel was de-stained using TAE buffer for 15 minutes before visualisation of the gel under UV light.

### 7.5.8. *S. aureus* gyrase supercoiling assay

*S. aureus* gyrase supercoiling assay kits were purchased from Inspiralis. Dilution buffer consisted of 50 mM Tris.HCl (pH 7.5), 1 mM DTT, 1 mM EDTA and 40 % (w/v) glycerol. Supercoiling assay buffer consisted of 40 mM HEPES.KOH (pH 7.6), 10 mM magnesium acetate, 10 mM DTT, 2 mM ATP, 500 mM potassium glutamate and 0.05 mg/mL albumin. All enzymes were stored at -80 °C. Relaxed pBR322 DNA was supplied as part of the assay kits. Assay kits were used following the manufacturer's instructions. Compounds were weighed into an Eppendorf tube and dissolved in DMSO. Serial dilutions were performed using DMSO. For each molecule to be tested a master mix was prepared using 0.5 µL of relaxed pBR322 DNA, 2 µL of dilution buffer, 2 µL of assay buffer and 16.5 µL of water. Diluted samples and the master mix were prepared in ice prior to incubation to reduce any intrinsic supercoiling activity. A 25 µL aliquot of master mix was used for the negative control and the volume made up to 30 µL with 1 µL of DMSO and 4 µL of assay buffer. Following this, 4 µL of enzyme was added for every compound to be tested including the positive control and known inhibitors (ciprofloxacin, simocyclinone D8). The master mix was homogenised, and 29 µL of master mix was dispensed into an Eppendorf along with 1 µL of compound to be tested. Samples were centrifuged at 13,000 rpm for 5 seconds before incubation at 37 °C for 30 minutes. The reaction was then stopped using 30 µL of iso-amyl alcohol in chloroform (1:24) and 20 µL of 40 % (w/v) sucrose, 0.1 M Tris.HCl (pH 8.0), 0.1 M EDTA and 0.5 mg/mL bromophenol blue (STEB buffer). The samples were once again centrifuged at 13,000 rpm for 2 minutes before 15 µL of the aqueous layer was loaded onto a 1 % agarose gel. The topoisomers and supercoiled DNA were then separated by electrophoresis before the gel was stained in ethidium bromide in TAE buffer (1 µg/mL) for 30 minutes. Following this, the gel was de-stained using TAE buffer for 15 minutes before visualisation of the gel under UV light.

### 7.5.9. Human Topo II $\alpha$ and Topo II $\beta$ decatenation assay

Human Topo II $\alpha$  and Topo II $\beta$  decatenation assay kits were purchased from Inspiralis. Dilution buffers consisted of 50 mM Tris.HCl (pH 7.5), 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 50 % (v/v) glycerol and 50 µg/mL albumin. Assay buffer consisted of 50 mM Tris.HCl (pH 7.5), 125 mM NaCl, 10 mM MgCl<sub>2</sub>, 5 mM DTT and 100 µg/mL albumin. ATP stock

contained 30 mM ATP. All enzymes were stored at -80 °C. Kinetoplast DNA (kDNA) was supplied as part of the assay kits. Assay kits were used following the manufacturer's instructions. Compounds were weighed into an Eppendorf tube and dissolved in DMSO. Serial dilutions were performed using DMSO. For each molecule to be tested a master mix was prepared using 2 µL of kDNA, 2 µL of dilution buffer, 3 µL of assay buffer, 1 µL of ATP stock and 17 µL of water. Diluted samples and the master mix were prepared in ice prior to incubation to reduce any intrinsic supercoiling activity. A 25 µL aliquot of master mix was used for the negative control and the volume made up to 30 µL with 1 µL of DMSO and 4 µL of assay buffer. Following this, 4 µL of enzyme was added for every compound to be tested including the positive control and known inhibitors (etoposide, simocyclinone D8). The master mix was homogenised, and 29 µL of master mix was dispensed into an Eppendorf along with 1 µL of compound to be tested. Samples were centrifuged at 13,000 rpm for 5 seconds before incubation at 37 °C for 30 minutes. The reaction was then stopped using 30 µL of iso-amyl alcohol in chloroform (1:24) and 20 µL of 40 % (w/v) sucrose, 0.1 M Tris.HCl (pH 8.0), 0.1 M EDTA and 0.5 mg/mL bromophenol blue (STEB buffer). The samples were once again centrifuged at 13,000 rpm for 2 minutes before 15 µL of the aqueous layer was loaded onto a 1 % agarose gel. The topoisomers and supercoiled DNA were then separated by electrophoresis before the gel was stained in ethidium bromide in TAE buffer (1 µg/mL) for 30 minutes. Following this, the gel was de-stained using TAE buffer for 15 minutes before visualisation of the gel under UV light.

### 7.5.10. Zone-of-inhibition assay

Cultures of *E. coli* wild-type (MG 1655), *E. coli* with a weakened membrane (NR698) and *S. aureus* wild type (ATCC292/13) were grown from glycerol stocks overnight in LB media and incubated at 37 °C with agitation. A 5 mL portion of these liquid cultures was then added to 45 mL LB before again being incubated at 37 °C for 4 h with agitation. Meanwhile, a mixture of 0.7 % agar (w/v) in LB was heated until homogenous before being aliquoted into 3 mL portions and cooled to 52 °C. A 50 µL extract of the liquid culture was then added to the cooled LB-agar mix, quickly shaken and transferred to a pre-poured LB-agar plate, distributing the liquid across the plate evenly. Once dry, 5 µL aliquots of the relevant compounds were deposited onto the surface of the plate, along with 5 µL aliquots of H<sub>2</sub>O, DMSO, ciprofloxacin (100 µM) and SD8 (100 µM). The plates were then incubated at 37 °C overnight. Inhibition was determined by visual inspection of the plates.

## REFERENCES

- 1 W. Vollmer, D. Blanot and M. A. De Pedro, *FEMS Microbiol. Rev.*, 2008, **32**, 149–167.
- 2 T. J. Beveridge, *J. Bacteriol.*, 1999, **181**, 4725–4733.
- 3 M. J. Austin, PhD thesis, University of East Anglia, 2015.
- 4 V. Vadillo-Rodríguez and J. R. Dutcher, *Soft Matter*, 2011, **7**, 4101–4110.
- 5 M. K. Bhattacharjee, *Chemistry of Antibiotics and Related Drugs*, Springer International Publishing, Cham, 2016.
- 6 M. L. Nelson, A. Dinardo, J. Hochberg and G. J. Armelagos, *Am. J. Phys. Anthropol.*, 2010, **143**, 151–154.
- 7 R. I. Aminov, *Front. Microbiol.*, 2010, **1**, 1–7.
- 8 E. J. Bassett, M. S. Keith, G. J. Armelagos, D. L. Martin and A. R. Villanueva, *Science (80-. )*, 1980, **209**, 1532–1534.
- 9 G. J. Armelagos, *Science (80-. )*, 1969, **163**, 255–259.
- 10 M. Cook, E. Molto and C. Anderson, *Am. J. Phys. Anthropol.*, 1989, **80**, 137–143.
- 11 R. W. K. Wong, U. Hägg, L. Samaranayake, M. K. Z. Yuen, C. J. Seneviratne and R. Kao, *Int. J. Oral Maxillofac. Surg.*, 2010, **39**, 599–605.
- 12 S. Hata and P. Ehrlich, *Experimentelle Grundlage der Chemotherapie der Spirillosen*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1910.
- 13 J. F. Mahoney, R. C. Arnold and A. Harris, *Am. J. Public Heal. Nations Heal.*, 1943, **33**, 1387–1391.
- 14 R. Bud, *Penicillin : triumph and tragedy*, Oxford University Press, New York, 1st edn., 2007.
- 15 E. Lax, *The mold in Dr. Florey's coat : the story of the penicillin miracle*, Henry Holt, 2005.
- 16 A. Fleming, *Br. J. Exp. Pathol.*, 1929, **10**, 226.
- 17 Making Penicillin Possible: Norman Heatley Remembers,  
<https://web.archive.org/web/20070221041204/http://www.sciencewatch.com/inte>

## CHAPTER 7: EXPERIMENTAL

- rviews/norman\_heatly.htm, (accessed 24 March 2017).
- 18 K. Lewis, *Nat. Rev. Drug Discov.*, 2013, 12, 371–387.
  - 19 WHO, Antibacterial agents in clinical development: An analysis of the antibacterial clinical development pipeline, including tuberculosis, <http://apps.who.int/iris/bitstream/handle/10665/258965/WHO-EMP-IAU-2017.11-eng.pdf?sequence=1>, (accessed 6 April 2018).
  - 20 WHO, Antimicrobial Resistance: Global Report on Surveillance, <https://apps.who.int/iris/handle/10665/112642>, (accessed 19 February 2017).
  - 21 F. Prestinaci, P. Pezzotti and A. Pantosti, *Pathog. Glob. Health*, 2015, 109, 309–318.
  - 22 T. O. & Social, *Special Eurobarometer 338: Antimicrobial Resistance, November-December 2009*, Brussels, 2010.
  - 23 C. A. M. McNulty, P. Boyle, T. Nichols, P. Clappison and P. Davey, *J. Antimicrob. Chemother.*, 2007, **59**, 727–738.
  - 24 J. L. Vincent, *Lancet*, 2003, 361, 2068–2077.
  - 25 B. M. Marshall and S. B. Levy, *Clin. Microbiol. Rev.*, 2011, 24, 718–733.
  - 26 European Commission, 2005.
  - 27 E. M. Agency, *Sales of veterinary antimicrobial agents in 26 EU/EEA countries in 2012: Fourth ESVAC report*, 2014.
  - 28 J. Davies and D. Davies, *Microbiol. Mol. Biol. Rev.*, 2010, **74**, 417–433.
  - 29 J. L. Martinez, *Drug Discov. Today Technol.*, 2014, 11, 33–39.
  - 30 G. Cox and G. D. Wright, *Int. J. Med. Microbiol.*, 2013, 303, 287–292.
  - 31 J. Pogliano, N. Pogliano and J. A. Silverman, *J. Bacteriol.*, 2012, **194**, 4494–4504.
  - 32 H. Nikaido, *Science (80-. )*, 1994, **264**, 382–388.
  - 33 L. McMurry, R. E. Petrucci and S. B. Levy, *Proc. Natl. Acad. Sci. U. S. A.*, 1980, **77**, 3974–3977.
  - 34 W. C. Reygaert, *AIMS Microbiol.*, 2018, **4**, 482–501.
  - 35 W. Reygaert, *Clin. Lab. Sci.*, 2009, **22**, 115–119.

## CHAPTER 7: EXPERIMENTAL

- 36 E. P. Abraham and E. Chain, *Nature*, 1940, **146**, 837.
- 37 C. C. Sanders and W. E. Sanders, *Antimicrob. Agents Chemother.*, 1979, **15**, 792–797.
- 38 C. L. Emery and L. A. Weymouth, *J. Clin. Microbiol.*, 1997, **35**, 2061–2067.
- 39 S. T. Chancey, D. Zähler and D. S. Stephens, *Future Microbiol.*, 2012, **7**, 959–978.
- 40 C. R. Mahon and D. C. Lenman, in *Textbook of Diagnostic Microbiology*, Elsevier, St Louis, 6th edn., 2019, pp. 254–273.
- 41 J. M. A. Blair, G. E. Richmond and L. J. V. Piddock, *Future Microbiol.*, 2014, **9**, 1165–1177.
- 42 A. Kumar and H. P. Schweizer, *Adv. Drug Deliv. Rev.*, 2005, **57**, 1486–1513.
- 43 P. A. Lambert, *J. Appl. Microbiol.*, 2002, **92**, 46S–54S.
- 44 C. M. Bébéar and S. Pereyre, *Curr. Drug Targets - Infect. Disord.*, 2005, **5**, 263–271.
- 45 W. R. Miller, J. M. Munita and C. A. Arias, *Expert Rev. Anti. Infect. Ther.*, 2014, **12**, 1221–1236.
- 46 A. Thiolas, C. Bornet, A. Davin-Régli, J. M. Pagès and C. Bollet, *Biochem. Biophys. Res. Commun.*, 2004, **317**, 851–856.
- 47 A. Beceiro, M. Tomás and G. Bou, *Clin. Microbiol. Rev.*, 2013, **26**, 185–230.
- 48 S. Kumar, M. M. Mukherjee and M. F. Varela, *Int. J. Bacteriol.*, , DOI:10.1155/2013/204141.
- 49 M. C. Roberts, *Appl. Biochem. Biotechnol. - Part B Mol. Biotechnol.*, 2004, **28**, 47–62.
- 50 S. Yang, B. N. Kreiswirth, G. Sakoulas, M. R. Yeaman, Y. Q. Xiong, A. Sawa and A. S. Bayer, *J. Infect. Dis.*, 2009, **200**, 1916–1920.
- 51 S. Stefani, F. Campanile, M. Santagati, M. L. Mezzatesta, V. Cafiso and G. Pacini, *Int. J. Antimicrob. Agents*, 2015, **46**, 278–289.
- 52 N. N. Mishra, A. S. Bayer, C. Weidenmaier, T. Grau, S. Wanner, A. Stefani, V. Cafiso, T. Bertuccio, M. R. Yeaman, C. C. Nast and S. J. Yang, *PLoS One*, , DOI:10.1371/journal.pone.0107426.
- 53 J. M. A. Blair, M. A. Webber, A. J. Baylay, D. O. Ogbolu and L. J. V. Piddock, *Nat. Rev. Microbiol.*, 2015, **13**, 42–51.

## CHAPTER 7: EXPERIMENTAL

- 54 N. A. Villagra, J. A. Fuentes, M. R. Jofré, A. A. Hidalgo, P. García and G. C. Mora, *J. Antimicrob. Chemother.*, 2012, **67**, 921–7.
- 55 J. Lubelski, W. N. Konings and A. J. M. Driessen, *Microbiol. Mol. Biol. Rev.*, 2007, **71**, 463–476.
- 56 M. Putman, H. W. van Veen and W. N. Konings, *Microbiol. Mol. Biol. Rev.*, 2000, **64**, 672–693.
- 57 T. Kuroda and T. Tsuchiya, *Biochim. Biophys. Acta - Proteins Proteomics*, 2009, 1794, 763–768.
- 58 L. J. V. Piddock, *Clin. Microbiol. Rev.*, 2006, **19**, 382–402.
- 59 D. C. Bay, K. L. Rommens and R. J. Turner, *Biochim. Biophys. Acta - Biomembr.*, 2008, 1778, 1814–1838.
- 60 L. B. Rice, *J. Infect. Dis.*, 2008, **197**, 1079–1081.
- 61 R. C. Founou, L. L. Founou and S. Y. Essack, *PLoS One*, 2017, **12**.
- 62 E. Tacconelli, E. Carrara, A. Savoldi, S. Harbarth, M. Mendelson, D. L. Monnet, C. Pulcini, G. Kahlmeter, J. Kluytmans, Y. Carmeli, M. Ouellette, K. Outterson, J. Patel, M. Cavalieri, E. M. Cox, C. R. Houchens, M. L. Grayson, P. Hansen, N. Singh, U. Theuretzbacher, N. Magrini, A. O. Aboderin, S. S. Al-Abri, N. Awang Jalil, N. Benzonana, S. Bhattacharya, A. J. Brink, F. R. Burkert, O. Cars, G. Cornaglia, O. J. Dyar, A. W. Friedrich, A. C. Gales, S. Gandra, C. G. Giske, D. A. Goff, H. Goossens, T. Gottlieb, M. Guzman Blanco, W. Hryniewicz, D. Kattula, T. Jinks, S. S. Kanj, L. Kerr, M. P. Kieny, Y. S. Kim, R. S. Kozlov, J. Labarca, R. Laxminarayan, K. Leder, L. Leibovici, G. Levy-Hara, J. Littman, S. Malhotra-Kumar, V. Manchanda, L. Moja, B. Ndoye, A. Pan, D. L. Paterson, M. Paul, H. Qiu, P. Ramon-Pardo, J. Rodríguez-Baño, M. Sanguinetti, S. Sengupta, M. Sharland, M. Si-Mehand, L. L. Silver, W. Song, M. Steinbakk, J. Thomsen, G. E. Thwaites, J. W. van der Meer, N. Van Kinh, S. Vega, M. V. Villegas, A. Wechsler-Fördös, H. F. L. Wertheim, E. Wesangula, N. Woodford, F. O. Yilmaz and A. Zorzet, *Lancet Infect. Dis.*, 2018, **18**, 318–327.
- 63 S. Santajit and N. Indrawattana, *Biomed Res. Int.*, , DOI:10.1155/2016/2475067.
- 64 K. Lewis, *Nat. Rev. Microbiol.*, 2007, **5**, 48–56.
- 65 M. S. Mulani, E. E. Kamble, S. N. Kumkar, M. S. Tawre and K. R. Pardesi, *Front.*

## CHAPTER 7: EXPERIMENTAL

- Microbiol.*, DOI:10.3389/fmicb.2019.00539.
- 66 D. J. Waxman and J. L. Strominger, *Annu. Rev. Biochem.*, 1983, **52**, 825–869.
- 67 D. J. Tipper and J. L. Strominger, *Proc. Natl. Acad. Sci.*, 1965, **54**, 1133–1141.
- 68 R. R. Yocum, J. R. Rasmussen and J. L. Strominger, *J. Biol. Chem.*, 1980, **255**, 3977–3986.
- 69 A. Coates, Y. Hu, R. Bax and C. Page, *Nat. Rev. Drug Discov.*, 2002, **1**, 895–910.
- 70 W. O. Foye, T. L. Lemke and D. A. Williams, *Foye's principles of medicinal chemistry*, Lippincott Williams & Wilkins, Philadelphia, 6th edn., 2008.
- 71 K. M. Papp-Wallace, A. Endimiani, M. A. Taracila and R. A. Bonomo, *Antimicrob. Agents Chemother.*, 2011, **55**, 4943–4960.
- 72 F. P. Tally, N. V. Jacobus and S. L. Gorbach, *Antimicrob. Agents Chemother.*, 1978, **14**, 436–438.
- 73 H. Kropp, J. G. Sundelof, J. S. Kahan, F. M. Kahan and J. Birnbaum, *Antimicrob. Agents Chemother.*, 1980, **17**, 993–1000.
- 74 N. S. Brewer and W. C. Hellinger, *Mayo Clin. Proc.*, 1991, **66**, 1152–1157.
- 75 J. C. J. Barna and D. H. Williams, *Annu. Rev. Microbiol.*, 1984, **38**, 339–357.
- 76 P. E. Reynolds, *Structure, Biochemistry and Mechanism of Action of Glycopeptide Antibiotics*, 1989, vol. 8.
- 77 D. C. Jordan and P. E. Reynolds, in *Mechanism of Action of Antimicrobial and Antitumor Agents*, eds. J. W. Corcoran, F. E. Hahn, J. F. Snell and K. L. Arora, Springer Berlin Heidelberg, Berlin, Heidelberg, 1975, pp. 704–718.
- 78 A. N. Chatterjee and H. R. Perkins, *Biochem. Biophys. Res. Commun.*, 1966, **24**, 489–494.
- 79 S. Arenz and D. N. Wilson, *Cold Spring Harb. Perspect. Med.*, , DOI:10.1101/cshperspect.a025361.
- 80 A. Simonetti, S. Marzi, L. Jenner, A. Myasnikov, P. Romby, G. Yusupova, B. P. Klaholz and M. Yusupov, *Cell. Mol. Life Sci.*, 2009, **66**, 423–436.
- 81 B. S. Laursen, H. P. Sorensen, K. K. Mortensen and H. U. Sperling-Petersen,

## CHAPTER 7: EXPERIMENTAL

- Microbiol. Mol. Biol. Rev.*, 2005, **69**, 101–123.
- 82 R. Karimi, M. Y. Pavlov, R. H. Buckingham and M. Ehrenberg, *Mol. Cell*, 1999, **3**, 601–609.
- 83 A. Dallas and H. F. Noller, *Mol. Cell*, 2001, **8**, 855–864.
- 84 K. M. Krause, A. W. Serio, T. R. Kane and L. E. Connolly, *Cold Spring Harb. Perspect. Med.*, , DOI:10.1101/cshperspect.a027029.
- 85 M. P. Mingeot-Leclercq, Y. Glupczynski and P. M. Tulkens, *Antimicrob. Agents Chemother.*, 1999, **43**, 727–737.
- 86 L. P. Kotra, J. Haddad and S. Mobashery, *Antimicrob. Agents Chemother.*, 2000, **44**, 3249–3256.
- 87 D. N. Wilson, *Nat. Rev. Microbiol.*, 2014, **12**, 35–48.
- 88 I. Chopra and M. Roberts, *Microbiol. Mol. Biol. Rev.*, 2001, **65**, 232–260.
- 89 F. Nguyen, A. L. Starosta, S. Arenz, D. Sohmen, A. Dönhöfer and D. N. Wilson, *Biol. Chem.*, 2014, **395**, 559–575.
- 90 F. Schlünzen, R. Zarivach, J. Harms, A. Bashan, A. Tocilj, R. Albrecht, A. Yonath and F. Franceschi, *Nature*, 2001, **413**, 814–821.
- 91 C. M. T. Spahn and C. D. Prescott, *J. Mol. Med.*, 1996, **74**, 423–439.
- 92 D. Tu, G. Blaha, P. B. Moore and T. A. Steitz, *Cell*, 2005, **121**, 257–270.
- 93 J. R. Menninger, *J. Antimicrob. Chemother.*, 1985, **16**, 23–34.
- 94 K. Kannan, P. Kanabar, D. Schryer, T. Florin, E. Oh, N. Bahroos, T. Tenson, J. S. Weissman and A. S. Mankin, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15958–15963.
- 95 S. Sothiselvam, B. Liu, W. Han, H. Ramu, D. Klepacki, G. C. Atkinson, A. Brauer, M. Remm, T. Tenson, K. Schulten, N. Vázquez-Laslop and A. S. Mankin, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 9804–9809.
- 96 S. Arenz, H. Ramu, P. Gupta, O. Berninghausen, R. Beckmann, N. Vázquez-Laslop, A. S. Mankin and D. N. Wilson, *Nat. Commun.*, 2014, **5**, 3501.
- 97 H. Ramu, N. Vázquez-Laslop, D. Klepacki, Q. Dai, J. Piccirilli, R. Micura and A. S. Mankin, *Mol. Cell*, 2011, **41**, 321–330.

## CHAPTER 7: EXPERIMENTAL

- 98 N. Vazquez-Laslop, C. Thum and A. S. Mankin, *Mol. Cell*, 2008, **30**, 190–202.
- 99 T. Tenson, M. Lovmar and M. Ehrenberg, *J. Mol. Biol.*, 2003, **330**, 1005–1014.
- 100 K. S. Long, L. H. Hansen, L. Jakobsen and B. Vester, *Antimicrob. Agents Chemother.*, 2006, **50**, 1458–1462.
- 101 M. Gumustas, C. T. Sengel-Turk, A. Gumustas, S. A. Ozkan and B. Uslu, in *Multifunctional Systems for Combined Delivery, Biosensing and Diagnostics*, Elsevier, 2017, pp. 67–108.
- 102 D. G. Allison and P. A. Lambert, in *Molecular Medical Microbiology*, Elsevier, 2nd edn., 2014, vol. 1, pp. 583–598.
- 103 B. E. Murray, C. A. Arias and E. C. Nannini, in *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases*, Elsevier, 8th edn., 2014, vol. 1, pp. 377–400.
- 104 O. Jardetzky, *J. Biol. Chem.*, 1963, **238**, 2498–2508.
- 105 F. E. Hahn, C. L. Wisseman and H. E. Hopps, *J. Bacteriol.*, 1955, **69**, 215–223.
- 106 A. D. Wolfe and F. E. Hahn, *BBA Sect. Nucleic Acids Protein Synth.*, 1965, **95**, 146–155.
- 107 A. H. Lin, R. W. Murray, T. J. Vidmar and K. R. Marotti, *Antimicrob. Agents Chemother.*, 1997, **41**, 2127.
- 108 J. S. Daly, G. M. Eliopoulos, E. Reiszner and R. C. Moellering Jr, *J. Antimicrob. Chemother.*, 1988, **21**, 721–730.
- 109 D. C. Eustice, P. A. Feldman, L. Zajac and A. M. Slee, *Antimicrob. Agents Chemother.*, 1988, **32**, 1218–1222.
- 110 A. M. Slee, M. A. Wuonola, R. J. McRipley, I. Zajac, M. J. Zawada, P. T. Bartholomew, W. A. Gregory and M. Forbes, *Antimicrob. Agents Chemother.*, 1987, **31**, 1791–1797.
- 111 J. E. Maddison, A. D. J. Watson and J. Elliott, in *Small Animal Clinical Pharmacology*, Elsevier, 2nd edn., 2008, pp. 148–185.
- 112 W. Wehrli, F. Knüsel, K. Schmid and M. Staehelin, *Proc. Natl. Acad. Sci. U. S. A.*, 1968, **61**, 667–673.
- 113 E. A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb and S. A.

## CHAPTER 7: EXPERIMENTAL

- Darst, *Cell*, 2001, **104**, 901–912.
- 114 W. R. McClure and C. L. Cech, *J. Biol. Chem.*, 1978, **253**, 8949–8956.
- 115 H. Lodish, A. Berk, C. A. Kaiser, M. Krieger, M. P. Scott, S. L. Zipursky and J. Darnell, *Molecular Cell Biology*, W. H. Freeman and Company, New York, 5th edn., 2004.
- 116 D. Voges, P. Zwickl and W. Baumeister, *Annu. Rev. Biochem.*, 1999, **68**, 1015–1068.
- 117 M. G. Shapiro, S. J. Frazier and H. A. Lester, *ACS Chem. Neurosci.*, 2012, **3**, 619–629.
- 118 Q. Lu, J. Peevey, F. Jow, M. M. Monaghan, G. Mendoza, H. Zhang, J. Wu, C. Y. Kim, J. Bicksler, L. Greenblatt, S. S. Lin, W. Childers and M. R. Bowlby, *Bioorganic Med. Chem.*, 2008, **16**, 3067–3075.
- 119 L. Shen, Y. Yuan, Y. Guo, M. Li, C. Li and X. Pu, *Front. Pharmacol.*, 2019, **10**, 1310.
- 120 J. R. Riordan, J. M. Rommens, B. S. Kerem, N. O. A. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. I. Lok, N. Plavsic, J. L. Chou, M. L. Drumm, M. C. Iannuzzi, F. S. Collins and L. C. Tsui, *Science (80-. )*, 1989, **245**, 1066–1073.
- 121 J. L. Kreindler, *Pharmacol. Ther.*, 2010, **125**, 219–229.
- 122 J. Momand, G. P. Zambetti, D. C. Olson, D. George and A. J. Levine, *Cell*, 1992, **69**, 1237–1245.
- 123 C. J. Brown, S. Lain, C. S. Verma, A. R. Fersht and D. P. Lane, *Nat. Rev. Cancer*, 2009, **9**, 862–873.
- 124 A. J. Wilson, *Chem. Soc. Rev.*, 2009, **38**, 3289.
- 125 A. Muppidi, K. Doi, C. P. Ramil, H.-G. Wang and Q. Lin, *Tetrahedron*, 2014, **70**, 7740–7745.
- 126 A. A. Bogan and K. S. Thorn, *J. Mol. Biol.*, 1998, **280**, 1–9.
- 127 T. L. Blundell, B. L. Sibanda, R. W. Montalvão, S. Brewerton, V. Chelliah, C. L. Worth, N. J. Harmer, O. Davies and D. Burke, *Philos. Trans. R. Soc. B Biol. Sci.*, 2006, **361**, 413–423.
- 128 S. Fletcher and A. D. Hamilton, *Curr. Opin. Chem. Biol.*, 2005, **9**, 632–638.
- 129 J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 737–738.
- 130 S. Neidle, *DNA structure and recognition*, IRL Press at Oxford University Press, 1st

## CHAPTER 7: EXPERIMENTAL

- edn., 1994.
- 131 T. J. Richmond and C. A. Davey, *Nature*, 2003, **423**, 145–150.
- 132 R. R. Sinden, *DNA Structure and Function*, Academic Press Limited, London, 1st edn., 1994.
- 133 G. Barone, C. Fonseca Guerra and F. M. Bickelhaupt, *ChemistryOpen*, 2013, **2**, 186–193.
- 134 W. Fuller, W. H. F. Wilkins, H. R. Wilson and L. D. Hamilton, *J. Mol. Biol.*, 1965, **12**, 60–76.
- 135 A. H-J Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boomt, G. van der Marelt and A. Rich, *Simpson. H. 1. in Global Chemical Cycles and their Alteration* edited by Stumm, W. } 235 (Dahlem Konferenzen, Wiley, 1967, vol. 36.
- 136 J. M. Berg, J. L. Tymoczko and L. Stryer, in *Biochemistry*, W H Freeman, New York, 5th edn., 2002.
- 137 A. D. Bates and A. Maxwell, *DNA Topology*, Oxford University Press, New York, 2nd edn., 2005.
- 138 G. M. Cooper, in *The Cell: A molecular Approach*, Sinauer Associates, Sunderland, 2nd edn., 2000.
- 139 J. Ma and M. D. Wang, *Biophys. Rev.*, 2016, **8**, 75–87.
- 140 J. C. Wang, *Q. Rev. Biophys.*, 1998, **31**, 107–144.
- 141 J. J. Champoux, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11998–12000.
- 142 N. H. Dekker, V. V Rybenkov, M. Duguet, N. J. Crisona, N. R. Cozzarelli, D. Bensimon and V. Croquette, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12126–12131.
- 143 C. D. Lima, J. C. Wang and A. Mondragón, *Nature*, 1994, **367**, 138–146.
- 144 A. Changela, R. J. DiGate and A. Mondragón, *Nature*, 2001, **411**, 1077–1081.
- 145 Y. C. Tse, K. Kirkegaard and J. C. Wang, *J. Biol. Chem.*, 1980, **255**, 5560–5565.
- 146 J. J. Champoux, *Annu. Rev. Biochem.*, 2001, **70**, 369–413.
- 147 T. Hsieh and J. L. Plank, *J. Biol. Chem.*, 2006, **281**, 5640–7.

## CHAPTER 7: EXPERIMENTAL

- 148 A. Kikuchi and K. Asai, *Nature*, 1984, **309**, 677–681.
- 149 K. Terekhova, K. H. Gunn, J. F. Marko and A. Mondragón, *Nucleic Acids Res.*, 2012, **40**, 10432–10440.
- 150 J. J. Champoux, *Prog. Nucleic Acid Res. Mol. Biol.*, 1998, **60**, 111–132.
- 151 L. Stewart, M. R. Redinbo, X. Qiu, W. G. Hol and J. J. Champoux, *Science*, 1998, **279**, 1534–1541.
- 152 A. I. Slesarev, K. O. Stetter, J. A. Lake, M. Gellert, R. Krah and S. A. Kozyavkin, *Nature*, 1993, **364**, 735–737.
- 153 P. Forterre, *Trends Biotechnol.*, 2006, **24**, 245–7.
- 154 J. C. Wang, *Annu. Rev. Biochem.*, 1996, **65**, 635–692.
- 155 P. Forterre, S. Gribaldo, D. Gadelle and M.-C. Serre, *Biochimie*, 2007, **89**, 427–446.
- 156 A. Bergerat, B. de Massy, D. Gadelle, P.-C. Varoutas, A. Nicolas and P. Forterre, *Nature*, 1997, **386**, 414–417.
- 157 M. J. Buttner, M. Schäfer, D. M. Lawson and A. Maxwell, *FEMS Microbiol. Rev.*, 2018, **42**, fux055.
- 158 F. Collin, S. Karkare and A. Maxwell, *Appl. Microbiol. Biotechnol.*, 2011, **92**, 479–497.
- 159 M. Thornton, M. Armitage, A. Maxwell, B. Dosanjh, A. J. Howells, V. Norris and D. C. Sigee, *Microbiology*, 1994, **140**, 2371–2382.
- 160 L. Costenaro, J. G. Grossmann, C. Ebel and A. Maxwell, *Structure*, 2007, **15**, 329–339.
- 161 L. Costenaro and J. G. Grossmann, *Structure*, 2005, **13**, 287–296.
- 162 J. Roca and J. C. Wang, *Cell*, 1994, **77**, 609–616.
- 163 J. H. M. Cabral, A. P. Jackson, C. V. Smith, N. Shikotra, A. Maxwell and R. C. Liddington, *Nature*, 1997, **388**, 903–906.
- 164 B. D. Bax, P. F. Chan, D. S. Eggleston, A. Fosberry, D. R. Gentry, F. Gorrec, I. Giordano, M. M. Hann, A. Hennessy, M. Hibbs, J. Huang, E. Jones, J. Jones, K. K. Brown, C. J. Lewis, E. W. May, M. R. Saunders, O. Singh, C. E. Spitzfaden, C. Shen, A. Shillings, A. J. Theobald, A. Wohlkonig, N. D. Pearson and M. N. Gwynn, *Nature*, 2010, **466**, 935–940.

## CHAPTER 7: EXPERIMENTAL

- 165 G. Orphanides and A. Maxwell, *Nucleic Acids Res.*, 1994, **22**, 1567–1575.
- 166 J. G. Heddle, S. Mittelheiser, A. Maxwell and N. H. Thomson, *J. Mol. Biol.*, 2004, **337**, 597–610.
- 167 L. Brino, A. Urzhumtsev, M. Mousli, C. Bronner, A. Mitschler, P. Oudet and D. Moras, *J. Biol. Chem.*, 2000, **275**, 9468–75.
- 168 D. B. Wigley, G. J. Davies, E. J. Dodson, A. Maxwell and G. Dodson, *Nature*, 1991, **351**, 624–629.
- 169 G. Y. Leshner, E. J. Froelich, M. D. Gruett, J. H. Bailey and R. P. Brundage, *J. Med. Pharm. Chem.*, 1962, **5**, 1063–1065.
- 170 J. M. Domagala, L. D. Hanna, C. L. Heifetz, M. P. Hutt, T. F. Mich, J. P. Sanchez and M. Solomon, *J. Med. Chem.*, 1986, **29**, 394–404.
- 171 J. S. Wolfson and D. C. Hooper, *Antimicrob. Agents Chemother.*, 1985, **28**, 581–6.
- 172 I. Laponogov, M. K. Sohi, D. A. Veselkov, X.-S. Pan, R. Sawhney, A. W. Thompson, K. E. McAuley, L. M. Fisher and M. R. Sanderson, *Nat. Struct. Mol. Biol.*, 2009, **16**, 667–669.
- 173 G. A. Jacoby, *Clin. Infect. Dis.*, 2005, **41**, S120–S126.
- 174 Y. G. Kwak, Q. C. Truong-Bolduc, H. Bin Kim, K.-H. Song, E. S. Kim and D. C. Hooper, *J. Antimicrob. Chemother.*, 2013, **68**, 2766–2772.
- 175 S. Wang, Y. Wang, J. Shen, Y. Wu and C. Wu, *FEMS Microbiol. Lett.*, 2013, **341**, 13–17.
- 176 G. Werner, C. Fleige, B. Ewert, J. A. Laverde-Gomez, I. Klare and W. Witte, *Int. J. Antimicrob. Agents*, 2010, **35**, 119–125.
- 177 P. Grohs, S. Houssaye, A. Aubert, L. Gutmann and E. Varon, *Antimicrob. Agents Chemother.*, 2003, **47**, 3542–7.
- 178 C.-L. Tremblay, A. Charlebois, L. Masson and M. Archambault, *Front. Microbiol.*, 2013, **4**, 245.
- 179 I. U. Rathnayake, M. Hargreaves and F. Huygens, *Syst. Appl. Microbiol.*, 2012, **35**, 326–333.
- 180 E. I. Parkinson, J. S. Bair, B. A. Nakamura, H. Y. Lee, H. I. Kuttub, E. H. Southgate, S.

## CHAPTER 7: EXPERIMENTAL

- Lezmi, G. W. Lau and P. J. Hergenrother, *Nat. Commun.*, 2015, **6**, 6947.
- 181 K. Hiramatsu, M. Igarashi, Y. Morimoto, T. Baba, M. Umekita and Y. Akamatsu, *Int. J. Antimicrob. Agents*, 2012, **39**, 478–485.
- 182 K. Baba, K. Ishihara, M. Ozawa, M. Usui, M. Hiki, Y. Tamura and T. Asai, *J. Vet. Med. Sci.*, 2012, **74**, 561–565.
- 183 S. Costa, C. Falcão, M. Viveiros, D. Machado, M. Martins, J. Melo-Cristino, L. Amaral and I. Couto, *BMC Microbiol.*, 2011, **11**, 241.
- 184 C. M. Sanfilippo, C. K. Hesje, W. Haas and T. W. Morris, *Chemotherapy*, 2011, **57**, 363–371.
- 185 M. Aligholi, A. Mirsalehian, S. Halimi, H. Imaneini, M. Taherikalani, F. Jabalameli, P. Asadollahi, B. Mohajer, A. Abdollahi and M. Emaneini, *Med. Sci. Monit.*, 2011, **17**, PH71–PH74.
- 186 E.-J. Yoon, C.-Y. Lee, M.-J. Shim, Y.-H. Min, A.-R. Kwon, J. Lee and E.-C. Choi, *Chemotherapy*, 2010, **56**, 153–157.
- 187 F. F. Coskun-Ari and G. Bosgelmez-Tinaz, *Eur. J. Med. Res.*, 2008, **13**, 366–70.
- 188 E. Sadowy, A. Sieńko, I. Gawryszewska, A. Bojarska, K. Malinowska and W. Hryniewicz, *Eur. J. Clin. Microbiol. Infect. Dis.*, 2013, **32**, 1193–1203.
- 189 F. Strelitz, H. Flon and I. N. Asheshov, *Proc. Natl. Acad. Sci. U. S. A.*, 1955, **41**, 620–4.
- 190 T. W. Sery, S. D. Paul, I. H. Leopold and M. Ronca, *AMA. Arch. Ophthalmol.*, 1957, **57**, 100–109.
- 191 R. M. Wishnow, J. L. Strominger, C. H. Birge, R. H. Threnn and J. Bacteriol, *J. Bacteriol.*, 1965, **89**, 1117–1123.
- 192 J. L. Strominger and R. H. Threnn, *Biochim. Biophys. Acta*, 1959, **36**, 83–92.
- 193 M. J. Ryan, *Biochemistry*, 1976, **15**, 3769–3777.
- 194 R. J. Lewis, F. T. F. Tsai and D. B. Wigley, *BioEssays*, 1996, **18**, 661–671.
- 195 R. J. Lewis, O. M. Singh, C. V Smith, T. Skarzynski, A. Maxwell, A. J. Wonacott and D. B. Wigley, *EMBO J.*, 1996, **15**, 1412–20.
- 196 A. Maxwell, *Mol. Microbiol.*, 1993, **9**, 681–686.

## CHAPTER 7: EXPERIMENTAL

- 197 J. D. Mullins and T. J. Macek, *J. Am. Pharm. Assoc.*, 1960, **49**, 245–248.
- 198 M. Hartenfeller and G. Schneider, in *Chemoinformatics and Computational Chemical Biology*, ed. J. Bajorath, Humana Press, Totowa, 1st edn., 2011, vol. 672, pp. 299–323.
- 199 C. G. Wermuth, C. R. Gabellin, P. Lindberg and L. A. Mitscher, *Pure Appl. Chem.*, 1998, **70**, 1129–1143.
- 200 W. P. Jencks, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 4046–50.
- 201 P. R. Andrews, D. J. Craik and J. L. Martin, *J. Med. Chem.*, 1984, **27**, 1648–1657.
- 202 I. D. Kuntz, K. Chen, K. A. Sharp and P. A. Kollman, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 9997–10002.
- 203 A. M. Beekman, M. A. O’Connell and L. A. Howell, *Angew. Chemie Int. Ed.*, 2017, **56**, 10446–10450.
- 204 M. J. I. Andrews, G. Kontopidis, C. McInnes, A. Plater, L. Innes, A. Cowan, P. Jewsbury and P. M. Fischer, *ChemBioChem*, 2006, **7**, 1909–1915.
- 205 A. M. Beekman, M. M. D. Cominetti, S. J. Walpole, S. Prabhu, M. A. O’Connell, J. Angulo and M. Searcey, *Chem. Sci.*, 2019, **10**, 4502–4508.
- 206 J. Schimana, H.-P. Fiedler, I. Groth, R. Submuth, W. Beil, M. Walker and A. Zeeck, *J. Antibiot. (Tokyo)*, 2000, **53**, 779–787.
- 207 U. Theobald, J. Schimana and H. P. Fiedler, *Antonie Van Leeuwenhoek*, 2000, **78**, 307–13.
- 208 R. H. Flatman, A. J. Howells, L. Heide, H.-P. Fiedler and A. Maxwell, *Antimicrob. Agents Chemother.*, 2005, **49**, 1093–1100.
- 209 M. J. Edwards, M. A. Williams, A. Maxwell and A. R. McKay, *Biochemistry*, 2011, **50**, 3432–3440.
- 210 M. J. Edwards, R. H. Flatman, L. A. Mitchenall, C. E. M. Stevenson, T. B. K. Le, T. A. Clarke, A. R. McKay, H.-P. Fiedler, M. J. Buttner, D. M. Lawson and A. Maxwell, *Science*, 2009, **326**, 1415–8.
- 211 S. J. Hearnshaw, M. J. Edwards, C. E. Stevenson, D. M. Lawson and A. Maxwell, *J. Mol. Biol.*, 2014, **426**, 2023–2033.

## CHAPTER 7: EXPERIMENTAL

- 212 M. J. Austin, S. J. Hearnshaw, L. A. Mitchenall, P. J. McDermott, L. A. Howell, A. Maxwell and M. Searcey, *Medchemcomm*, 2016, **7**, 1387–1391.
- 213 Schrodinger Release 2016-4: Ligprep, Schrodinger LLC, New York, NY, 2015.
- 214 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrín and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177-6196.
- 215 T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750–1759.
- 216 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrín and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177–6196.
- 217 Schrodinger Release 2016-4: Core Hopping, Schrodinger LLC, New York, NY, 2015
- 218 K. Zhu, K. W. Borrelli, J. R. Greenwood, T. Day, R. Abel, R. S. Farid and E. Harder, *J. Chem. Inf. Model.*, 2014, **54**, 1932–1940.
- 219 Schrodinger Release 2016-4: Maestro, Schrodinger LLC, New York, NY, 2015.
- 220 J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719–2740.
- 221 J. Duan, S. L. Dixon, J. F. Lowrie and W. Sherman, *J. Molec. Graph. Model*, 2010, **29**, 157-170
- 222 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Deliv. Rev.*, 2001, **46**, 3–26.
- 223 A. K. Ghose and G. M. Crippen, *J. Comput. Chem.*, 1986, **7**, 565–577.
- 224 C. Hansch and A. Leo, in *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley & Sons Ltd., New York, 1979.
- 225 A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, *J. Phys. Chem. A*, 1998, **102**, 3762–3772.
- 226 M. Congreve, R. Carr, C. Murray and H. Jhoti, *Drug Discov. Today*, 2003, **8**, 876–877.
- 227 W. H. Perkin, *J. Chem. Soc.*, 1868, **21**, 53–63.
- 228 W. H. Perkin, *J. Chem. Soc.*, 1877, **32**, 660–674.
- 229 D. S. Breslow and C. R. Hauser, *J. Am. Chem. Soc.*, 1939, **61**, 786–792.

## CHAPTER 7: EXPERIMENTAL

- 230 C. R. Hauser and D. S. Breslow, *J. Am. Chem. Soc.*, 1939, **61**, 793–798.
- 231 A. A. Kudale, J. Kendall, C. C. Warford, N. D. Wilkins and G. J. Bodwell, *Tetrahedron Lett.*, 2007, **48**, 5077–5080.
- 232 S. V. Vasilyeva, A. S. Levina, N. S. Li-Zhulanov, N. V. Shatskaya, S. I. Baiborodin, M. N. Repkova, V. F. Zarytova, N. A. Mazurkova and V. N. Silnikov, *Bioorganic Med. Chem.*, 2015, **23**, 2168–2175.
- 233 E. K. Woodman, J. G. K. Chaffey, P. A. Hopes, D. R. J. Hose and J. P. Gilday, *Org. Process Res. Dev.*, 2009, **13**, 106–113.
- 234 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chemie Int. Ed.*, 2001, **40**, 2004–2021.
- 235 N. George, S. Ofori, S. Parkin and S. G. Awuah, *RSC Adv.*, 2020, **10**, 24017–24026.
- 236 R. L. M. Teeuwen, S. S. van Berkel, T. H. H. van Dulmen, S. Schoffelen, S. A. Meeuwissen, H. Zuilhof, F. A. de Wolf and J. C. M. van Hest, *Chem. Commun.*, 2009, 4022–4024.
- 237 A. Smeets, P. Willot, J. De Winter, P. Gerbaux, T. Verbiest and G. Koeckelberghs, *Macromolecules*, 2011, **44**, 6017–6025.
- 238 O. Mitsunobu and M. Yamada, *Bull. Chem. Soc. Jpn.*, 1967, **40**, 2380–2382.
- 239 J.-I. Kato, H. Suzuki and H. Ikeda, *J. Biol. Chem.*, 1992, **267**, 25676–25684.
- 240 U. Kellner, P. Rudolph and R. Parwaresch, *Onkologie*, 2000, **23**, 424–430.
- 241 Ž. Skok, N. Zidar, D. Kikelj and J. Ilaš, *J. Med. Chem.*, 2019, **63**, 884–904.
- 242 A. A. Sadiq, M. R. Patel, B. A. Jacobson, M. Escobedo, K. Ellis, L. M. Oppegard, H. Hiasa and R. A. Kratzke, *Invest. New Drugs*, 2010, **28**, 20–25.
- 243 C. L. Hagan, S. Kim and D. Kahne, *Science*, 2010, **328**, 890–2.
- 244 D. M. Walther, D. Rapaport and J. Tommassen, *Cell. Mol. Life Sci.*, 2009, **66**, 2789–2804.
- 245 J. Tommassen, *Microbiology*, 2010, **156**, 2587–2596.
- 246 Y. Gu, H. Li, H. Dong, Y. Zeng, Z. Zhang, N. G. Paterson, P. J. Stansfeld, Z. Wang, Y. Zhang, W. Wang and C. Dong, *Nature*, 2016, **531**, 64–69.

## CHAPTER 7: EXPERIMENTAL

- 247 N. Noinaj, A. J. Kuszak, J. C. Gumbart, P. Lukacik, H. Chang, N. C. Easley, T. Lithgow and S. K. Buchanan, *Nature*, 2013, **501**, 385–390.
- 248 L. M. McMorran, D. J. Brockwell and S. E. Radford, *Arch. Biochem. Biophys.*, 2014, **564**, 265–280.
- 249 C. L. Hagan, D. B. Westwood and D. Kahne, *Biochemistry*, 2013, **52**, 6108–6113.
- 250 G. Roman-Hernandez, J. H. Peterson and H. D. Bernstein, *Elife*, 2014, **3**, e04234.
- 251 T. J. Knowles, A. Scott-Tucker, M. Overduin and I. R. Henderson, *Nat. Rev. Microbiol.*, 2009, **7**, 206–214.
- 252 T. Wu, J. Malinverni, N. Ruiz, S. Kim, T. J. Silhavy and D. Kahne, *Cell*, 2005, **121**, 235–245.
- 253 E. S. Charlson, J. N. Werner and R. Misra, *J. Bacteriol.*, 2006, **188**, 7186–94.
- 254 S. Kim, J. C. Malinverni, P. Sliz, T. J. Silhavy, S. C. Harrison and D. Kahne, *Science* (80-. ),, 2007, **317**, 961–964.
- 255 P. Z. Gatzeva-Topalova, T. A. Walton and M. C. Sousa, *Structure*, 2008, **16**, 1873–1881.
- 256 J. C. Malinverni, J. Werner, S. Kim, J. G. Sklar, D. Kahne, R. Misra and T. J. Silhavy, *Mol. Microbiol.*, 2006, **61**, 151–164.
- 257 A. Heuck, A. Schleiffer and T. Clausen, *J. Mol. Biol.*, 2011, **406**, 659–666.
- 258 B. Hao, S. Oehlmann, M. E. Sowa, J. W. Harper and N. P. Pavletich, *Mol. Cell*, 2007, **26**, 131–143.
- 259 Z. Jawad and M. Paoli, *Structure*, 2002, **10**, 447–454.
- 260 P. Vuong, D. Bennion, J. Mantei, D. Frost and R. Misra, *J. Bacteriol.*, 2008, **190**, 1507–1517.
- 261 L. R. Warner, K. Varga, O. F. Lange, S. L. Baker, D. Baker, M. C. Sousa and A. Pardi, *J. Mol. Biol.*, 2011, **411**, 83–95.
- 262 T. J. Knowles, D. M. McClelland, S. Rajesh, I. R. Henderson and M. Overduin, *Biomol. NMR Assign.*, 2009, **3**, 203–206.
- 263 B. L. Lytle, J. Song, N. B. de la Cruz, F. C. Peterson, K. A. Johnson, C. A. Bingman, G. N.

## CHAPTER 7: EXPERIMENTAL

- Phillips and B. F. Volkman, *Proteins Struct. Funct. Bioinforma.*, 2009, **76**, 237–243.
- 264 C. M. Sandoval, S. L. Baker, K. Jansen, S. I. Metzner and M. C. Sousa, *J. Mol. Biol.*, 2011, **409**, 348–357.
- 265 L. D. D’Andrea and L. Regan, *Trends Biochem. Sci.*, 2003, **28**, 655–662.
- 266 E. R. G. Main, Y. Xiong, M. J. Cocco, L. D’Andrea and L. Regan, *Structure*, 2003, **11**, 497–508.
- 267 R. K. Allan and T. Ratajczak, *Cell Stress Chaperones*, 2011, **16**, 353–67.
- 268 C. Dong, H.-F. Hou, X. Yang, Y.-Q. Shen and Y.-H. Dong, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2012, **68**, 95–101.
- 269 B. Wu, P. Li, Y. Liu, Z. Lou, Y. Ding, C. Shu, S. Ye, M. Bartlam, B. Shen and Z. Rao, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 8348–53.
- 270 Y. Wu and B. Sha, *Nat. Struct. Mol. Biol.*, 2006, **13**, 589–593.
- 271 C. Scheufler, A. Brinker, G. Bourenkov, S. Pegoraro, L. Moroder, H. Bartunik, F. U. Hartl and I. Moarefi, *Cell*, 2000, **101**, 199–210.
- 272 R. Albrecht and K. Zeth, *J. Biol. Chem.*, 2011, **286**, 27792–803.
- 273 J. S. Zarins-Tutt, T. T. Barberi, H. Gao, A. Mearns-Spragg, L. Zhang, D. J. Newman and R. J. M. Goss, *Nat. Prod. Rep.*, 2016, **33**, 54–72.
- 274 M. Lunelli, R. K. Lokareddy, A. Zychlinsky and M. Kolbe, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 9661–6.
- 275 V. Job, P.-J. Matteï, D. Lemaire, I. Attree and A. Dessen, *J. Biol. Chem.*, 2010, **285**, 23224–32.
- 276 K. H. Kim, S. Aulakh and M. Paetzel, *J. Biol. Chem.*, 2011, **286**, 39116–21.
- 277 K. H. Kim, H.-S. Kang, M. Okon, E. Escobar-Cabrera, L. P. McIntosh and M. Paetzel, *Biochemistry*, 2011, **50**, 1081–1090.
- 278 T. J. Knowles, D. F. Browning, M. Jeeves, R. Maderbocus, S. Rajesh, P. Sridhar, E. Manoli, D. Emery, U. Sommer, A. Spencer, D. L. Leyton, D. Squire, R. R. Chaudhuri, M. R. Viant, A. F. Cunningham, I. R. Henderson and M. Overduin, *EMBO Rep.*, 2011, **12**, 123–8.

## CHAPTER 7: EXPERIMENTAL

- 279 D. Cowburn, *Curr. Opin. Struct. Biol.*, 1997, **7**, 835–838.
- 280 F. C. Breedveld, *Lancet*, 2000, **355**, 735–740.
- 281 S. Swaminathan and S. Riminton, *Aust. Prescr.*, 2006, **29**, 130–133.
- 282 H. P. Rang, *Pharmacology*, Churchill Livingstone, Edinburgh, 5th edn., 2003.
- 283 K. M. Storek, M. R. Auerbach, H. Shi, N. K. Garcia, D. Sun, N. N. Nickerson, R. Vij, Z. Lin, N. Chiang, K. Schneider, A. T. Weckler, E. Skippington, G. Nakamura, D. Seshasayee, J. T. Koerber, J. Payandeh, P. A. Smith and S. T. Rutherford, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 3692–3697.
- 284 A. T. Bentley and P. E. Klebba, *J. Bacteriol.*, 1988, **170**, 1063–1068.
- 285 K. M. Lehman and M. Grabowicz, *Antibiotics*, , DOI:10.3390/antibiotics8040163.
- 286 J. F. Rossignol, *Antiviral Res.*, 2014, **110**, 94–103.
- 287 A. C. White Jr, *Expert Rev. Anti. Infect. Ther.*, 2004, **2**, 43–49.
- 288 E. R. Shamir, M. Warthan, S. P. Brown, J. P. Nataro, R. L. Guerrant and P. S. Hoffman, *Antimicrob. Agents Chemother.*, 2010, **54**, 1526–1533.
- 289 P. Chahales, P. S. Hoffman and D. G. Thanassi, *Antimicrob. Agents Chemother.*, 2016, **60**, 2028–2038.
- 290 J. J. Psonis, P. Chahales, N. S. Henderson, N. W. Rigel, P. S. Hoffman and D. G. Thanassi, , DOI:10.1074/jbc.RA119.009616.
- 291 C. L. Hagan, J. S. Wzorek and D. Kahne, *Proc. Natl. Acad. Sci.*, 2015, **112**, 2011–2016.
- 292 W. C. Chan and P. D. White, *Fmoc Solid Phase Peptide Synthesis*, Oxford University Press, Oxford, 2004.
- 293 H. T. Bergal, A. H. Hopkins, S. I. Metzner and M. C. Sousa, *Structure*, 2016, **24**, 243–251.
- 294 D. A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell and E. W. Sayers, *Nucleic Acids Res.*, 2013, **41**, D42.
- 295 F. H. Niesen, H. Berglund and M. Vedadi, *Nat. Protoc.*, 2007, **2**, 2212–2221.
- 296 S. M. Yarmoluk, V. B. Kovalska and K. D. Volkova, in *Advanced Fluorescence Reporters in Chemistry and Biology III: Applications in Sensing and Imaging*, ed. A. P.

## CHAPTER 7: EXPERIMENTAL

- Demchenko, Springer Heidelberg Dordrecht, New York, 2011, pp. 161–200.
- 297 M. Ristic, N. Rosa, S. A. Seabrook and J. Newman, *Acta Crystallogr. Sect. Struct. Biol. Commun.*, 2015, **71**, 1359–1364.
- 298 S. F. Altschul, J. C. Wootton, E. M. Gertz, R. Agarwala, A. Morgulis, A. A. Schäffer and Y. K. Yu, *FEBS J.*, 2005, **272**, 5101–5109.
- 299 S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J. Mol. Biol.*, 1990, **215**, 403–410.
- 300 S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic Acids Res.*, 1997, **25**, 3389–3402.
- 301 G. S. Vanier, in *Peptide Synthesis and Applications*, Humana Press Inc., 2013, vol. 1047, pp. 235–249.
- 302 D. T. Vistica, P. Skehan, D. Scudiero, A. Monks, A. Pittman and M. R. Boyd, *Cancer Res.*
- 303 M. V. Berridge and A. S. Tan, *Arch. Biochem. Biophys.*, 1993, **303**, 474–482.
- 304 M. A. Cooper, *Nat. Rev. Drug Discov.* 2002 **17**, 2002, **1**, 515–528.
- 305 H. D. Dakin, *J. Biol. Chem.*, 1929, **82**, 439–446.
- 306 D. K. Das, S. Sarkar, M. Khan, M. Belal and A. T. Khan, *Tetrahedron Lett.*, 2014, **55**, 4869–4874.
- 307 T. S. Rodrigues, A. G. M. da Silva, L. C. de Oliveira, A. M. da Silva, R. R. Teixeira and P. H. C. Camargo, *Tetrahedron Lett.*, 2017, **58**, 590–595.
- 308 *Justus Liebigs Ann. Chem.*, 1955, **596**, 38–79.
- 309 A. M. King, C. Salomé, J. Dinsmore, E. Salomé-Grosjean, M. De Ryck, R. Kaminski, A. Valade and H. Kohn, *J. Med. Chem.*, 2011, **54**, 4815–4830.