NMR structural studies of the binding of peptidyl transferase antibiotics to conserved secondary structural motifs of 23S ribosomal RNA

A thesis submitted to The University of Manchester for the Degree of PhD in the Faculty of Physical Sciences and Engineering

2011

Misbah Nareen

School of Chemistry
Contents

NMR structural studies of the binding of peptidyl transferase antibiotics to conserved secondary structural motifs of 23S ribosomal RNA .................................................. 1

List of Figures .................................................................................................................. 6
List of Tables .................................................................................................................... 13
Symbols and Abbreviations ............................................................................................. 15

Abstract ............................................................................................................................ 19
Declaration ......................................................................................................................... 20
Copyright statement .......................................................................................................... 20
Acknowledgements ........................................................................................................... 21
Dedication ......................................................................................................................... 22

Chapter 1  Introduction ..................................................................................................... 23

1.1 New Biology ................................................................................................................. 23
  1.1.1 Advances made in biological science and tackling antibiotic resistance ............... 23
  1.1.2 Nobel Prize in chemistry 2009 .............................................................................. 26
  1.1.3 Ribosome crystal structures ................................................................................... 27

1.2 The importance of ribosomal RNA in protein synthesis ............................................. 28
  1.2.1 Biophysical properties of RNA ............................................................................. 28
  1.2.2 The Ribosome and protein synthesis .................................................................... 34

1.3 PT inhibitor antibiotics ............................................................................................... 40
  1.3.1 Antibiotic resistance ............................................................................................ 40
  1.3.2 Amicetin .............................................................................................................. 42
  1.3.3 Bamicetin ............................................................................................................ 44
  1.3.4 Sparsomycin ....................................................................................................... 45
  1.3.5 Anisomycin ......................................................................................................... 48

1.4 Structures and binding sites of the PT inhibitor antibiotics ........................................ 49
  1.4.1 Amicetin and Bamicetin ...................................................................................... 49
  1.4.2 Sparsomycin ....................................................................................................... 50
  1.4.3 Anisomycin ......................................................................................................... 53

1.5 Structures of the RNAs ............................................................................................ 55
  1.5.1 B.subtilis 27mer RNA motif ................................................................................. 55
  1.5.2 GTPase centre of H.hal 25mer RNA .................................................................... 57
  1.5.3 NMR structures of RNA in databases ................................................................ 59

1.6 Cross-hypersensitivity ............................................................................................. 60
  1.6.1 Cross-hypersensitivity phenomenon ................................................................... 60
  1.6.2 The effect of thiostrepton and amicetin ................................................................. 61

1.7 Introduction to NMR ............................................................................................... 64
  1.7.1 Chemical shift and J-coupling .............................................................................. 66
  1.7.2 T1 and T2 relaxation ........................................................................................... 70
  1.7.3 Nuclear Overhauser Effect (NOE) ....................................................................... 71
Chapter 2 Materials and Methods

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Antibiotic and RNA sample preparation for the NMR studies</td>
<td>83</td>
</tr>
<tr>
<td>2.1.1 Antibiotics</td>
<td>83</td>
</tr>
<tr>
<td>2.1.2 <em>B. subtilis</em> 27mer RNA</td>
<td>84</td>
</tr>
<tr>
<td>2.1.3 <em>H. hal</em> 25mer RNA</td>
<td>85</td>
</tr>
<tr>
<td>2.2 NMR instrumentation and experiments</td>
<td>85</td>
</tr>
<tr>
<td>2.2.1 NMR spectrometers</td>
<td>85</td>
</tr>
<tr>
<td>2.2.2 NMR Acquisition Parameters</td>
<td>87</td>
</tr>
<tr>
<td>2.2.3 NMR experiments</td>
<td>88</td>
</tr>
<tr>
<td>2.2.4 Processing and Displaying the NMR spectra</td>
<td>89</td>
</tr>
<tr>
<td>2.2.5 Analysis</td>
<td>90</td>
</tr>
<tr>
<td>2.3 NMR techniques</td>
<td>91</td>
</tr>
<tr>
<td>2.3.1 1D $^1$H experiments</td>
<td>91</td>
</tr>
<tr>
<td>2.3.2 1D $^{13}$C { $^1$H} and $^{31}$P { $^1$H} decoupled experiments</td>
<td>92</td>
</tr>
<tr>
<td>2.3.3 2D homonuclear and heteronuclear experiments</td>
<td>93</td>
</tr>
<tr>
<td>2.3.4 $^1$H-$^1$H Correlated Spectroscopy (COSY)</td>
<td>95</td>
</tr>
<tr>
<td>2.3.5 $^1$H-$^1$H Double-Quantum Filtered Correlation Spectroscopy (DQF-COSY)</td>
<td>98</td>
</tr>
<tr>
<td>2.3.6 $^1$H-$^1$H Total Correlation Spectroscopy (TOCSY)</td>
<td>99</td>
</tr>
<tr>
<td>2.3.7 $^1$H-$^{13}$C Heteronuclear Single Quantum Coherence (HSQC)</td>
<td>100</td>
</tr>
<tr>
<td>2.3.8 $^1$H-$^{13}$C HSQC-TOCSY</td>
<td>101</td>
</tr>
<tr>
<td>2.3.9 $^1$H-$^1$H Nuclear Overhauser Effect Spectroscopy (NOESY)</td>
<td>103</td>
</tr>
<tr>
<td>2.3.10 $^1$H-$^1$H Rotating-frame Overhauser Effect Spectroscopy (ROESY)</td>
<td>105</td>
</tr>
<tr>
<td>2.3.11 $^1$H-$^{31}$P CPMG-HSQC-NOESY</td>
<td>106</td>
</tr>
<tr>
<td>2.3.12 3D NOE/2QC</td>
<td>107</td>
</tr>
<tr>
<td>2.4 Assignment strategy and constraint generation</td>
<td>111</td>
</tr>
<tr>
<td>2.4.1 NMR Resonance assignment</td>
<td>111</td>
</tr>
<tr>
<td>2.4.2 NOE based distance constraints</td>
<td>114</td>
</tr>
<tr>
<td>2.4.3 Dihedral angle restraints</td>
<td>116</td>
</tr>
<tr>
<td>2.4.4 Hydrogen bond restraints</td>
<td>119</td>
</tr>
<tr>
<td>2.4.5 Database potential restraints</td>
<td>119</td>
</tr>
<tr>
<td>2.5 Structure calculation protocol-Antibiotics</td>
<td>120</td>
</tr>
<tr>
<td>2.5.1 Overview</td>
<td>120</td>
</tr>
<tr>
<td>2.5.2 Starting structure generation</td>
<td>122</td>
</tr>
<tr>
<td>2.5.3 Randomisation</td>
<td>122</td>
</tr>
<tr>
<td>2.5.4 Simulated Annealing</td>
<td>123</td>
</tr>
<tr>
<td>2.5.5 Refinement</td>
<td>123</td>
</tr>
</tbody>
</table>
Chapter 3  NMR structure determination of Antibiotics  

3.1  Bamicetin  

3.1.1  NMR assignment of Bamicetin  

3.1.2  Variable temperature effects observed  

3.1.3  Experimental constraints  

3.1.4  NMR solution structure of Bamicetin  

3.2  Sparsomycin  

3.2.1  NMR assignment of Sparsomycin  

3.2.2  Variable temperature effects on the amide and methylene protons  

3.2.3  Experimental constraints  

3.2.4  NMR solution structure of Sparsomycin  

3.3  Anisomycin  

3.3.1  NMR assignment of Anisomycin  

3.3.2  Experimental constraints  

3.3.3  NMR solution structure of Anisomycin  

3.4  Discussion and comparison of Bamicetin, Sparsomycin and Anisomycin structures  

Chapter 4  NMR Structure determination of the B.subtilis 27mer RNA and its interactions with bamicetin antibiotic  

4.1  NMR assignment of the 27mer RNA motif of Bacillus subtilis rRNA  

4.1.1  Secondary structure  

4.1.2  NMR assignment of the exchangeable proton resonances  

4.1.3  Assignment of the non exchangeable proton and carbon resonances  

4.1.4  Assignment of the phosphorus resonances  

4.1.5  Assignment table  

4.1.6  Experimental constraints  

4.1.7  NMR structure of the 27mer RNA motif of B.subtilis rRNA  

4.1.8  Conformational analysis  

4.2  27mer RNA motif of B.subtilis rRNA-bamicetin complex  

4.2.1  1H NMR titration  

-4-
Chapter 5  NMR Structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic 221

5.1  NMR assignment of the *H. hal* 25mer RNA motif of 23S rRNA 221

   5.1.1  Secondary structure 221
   5.1.2  Assignment of the exchangeable proton resonances 222
   5.1.3  Assignment of the non-exchangeable proton resonances 227
   5.1.4  Assignment table 235
   5.1.5  Geometrical constraints 236
   5.1.6  NMR solution structure of the *H. hal* 25mer RNA 238

5.2  25mer motif of *H. hal* rRNA-amicetin complex 244

   5.2.1  $^1$H NMR titration 244

Chapter 6  Conclusion and Future work 253

References 258

Appendices 265

Appendix I-Experimental scripts 265

(1) NMR data processing script 265

(2) Structure calculation scripts 267

Appendix II-Dissemination 284

Word count = 50,336
List of Figures

Figure 1.1.1: A diagram illustrating the combination of two major factors (blue and green circles) which are causing antibiotic resistance (orange circle). .......................................................... 25

Figure 1.2.1: A growing RNA chain showing the attachment of the phosphate groups to the ribose at position 5’ of one nucleotide and 3’ of the next nucleotide ...................... 29

Figure 1.2.2: Scheme showing the various secondary structures that RNA molecules can adopt. ................................................................................................................................. 30

Figure 1.2.3: Conformation of the ribose ring of RNA and the deoxyribose ring of DNA. ................................................................................................................................. 30

Figure 1.2.4: Watson-Crick base pairs A-U and G-C are responsible for the duplexes in RNA. ................................................................................................................................. 31

Figure 1.2.5: Some examples of common base pairs which can be formed containing at least two hydrogen bonds in RNA. .............................................................................................. 32

Figure 1.2.6: A double stranded RNA secondary structure comprising of the WC base pairs (A-U and G-C) and the phosphate groups running along the outer side of the structure ................................................................................................................................. 33

Figure 1.2.7: Diagrammatic representation of the ribosome comprised of the large ribosomal subunit (50S) and the small ribosomal subunit (30S) showing the important sites of protein synthesis ................................................................................................................................. 35

Figure 1.2.8: Scheme for protein synthesis in the ribosome ........................................................................................................................................................................................................... 36

Figure 1.2.9: Schematic secondary structure of 23S rRNA endowed with the six domains (I to VI) and the 101 helices are numbered sequentially (1-101). .............................................. 39

Figure 1.3.1: The chemical structure of the chloramphenicol antibiotic .............................................................. 41

Figure 1.3.2: Electrophoretic pattern of E.coli 70S ribosome cells without and with exposure to chloramphenicol ................................................................................................................................. 42

Figure 1.3.3: The chemical structure of amicetin antibiotic with the common structural features observed with the amicetin family of antibiotics ......................................................................................... 42

Figure 1.3.4: Graph displaying the amicetin-resistant mutant cells (Ami') being less susceptible to the antibiotic in comparison to the wild-type (WT) cells ................................................................................................................................. 43

Figure 1.3.5: The effects of plicacetin amicetin and bamicetin on the PT activity of B.subtilis 70S ribosome ........................................................................................................................................................................................................... 44

Figure 1.3.6: The chemical structure of sparsomycin containing a methylated uracil ring and a sulphur ‘rich’ tail. ........................................................................................................................................................................................................... 45

Figure 1.3.7: The chemical structure of sparsomycin containing a methylated uracil ring and a sulphur ‘rich’ tail. ........................................................................................................................................................................................................... 48

Figure 1.4.1: The NMR structure of amicetin and the intramolecular hydrogen bonds which help to maintain the antibiotic in a folded conformation ................................................................................................................................. 50

Figure 1.4.2: Secondary structure of the peptidyl transferase centre of H. Marismortui 23S rRNA ........................................................................................................................................................................................................... 51
Figure 1.4.3: The X-ray crystal structure of the *H. marismortui* 50S ribosome bound sparsomycin and the X-ray crystal structure of *D. radiodurans* 50S bound to sparsomycin. .......................................................... 52

Figure 1.4.4: The molecular structure of N-acetylbromoanisomycin ........................................ 53

Figure 1.4.5: X-ray crystal structure of anisomycin bound to *H. marismortui* 23S rRNA... 54

Figure 1.4.6: The location of the exit tunnel of the *H. marismortui* 50S ribosomal subunit, where anisomycin binds and interacts with the surrounding nucleotides ............... 54

Figure 1.5.1: The secondary structure of the peptidyl transferase centre located in the domain V region of *B. subtilis* 23S rRNA with the 27mer RNA strand ..................... 55

Figure 1.5.2: Homologous sequences of the 23S rRNAs ......................................................... 56

Figure 1.5.3: The secondary structure of the 25mer RNA motif residing in domain II of *H. hal* 23S rRNA and displaying the thiostrepton resistant mutant ......................... 57

Figure 1.5.4: The chemical structure of the cyclic peptide thiostrepton antibiotic ............ 58

Figure 1.5.5: Graph displaying the rapid growth of structures deposited into the RCSB database since its formation in 1976 ........................................................................ 59

Figure 1.6.1: The cross-hypersensitivity effects observed between the amicetin antibiotic and the thiostrepton mutated site A1159G (G11) and conversely with the thiostrepton antibiotic and the amicetin mutated site U2457C (U26) ........................................ 61

Figure 1.6.2: Illustration displaying the location of the thiostrepton resistant mutants in domain II and the amicetin resistant mutant in domain V of the *H. hal* 23S rRNA.... 62

Figure 1.7.1: The precessing nuclei has a longitudinal component of the magnetisation along the z-direction and an orthogonal component in the x-y plane ......................... 65

Figure 1.7.2: Nuclei with spin I=½ produce two quantised energy levels (+½, α and -½, β) when placed in the applied magnetic field (B₀) ........................................ 65

Figure 1.7.3: The EM radiation is applied in the form of a short RF pulse (B₁) along the x-axis ...................................................................................................................... 66

Figure 1.7.4: The electrons (e⁻) in the molecule shield the nucleus from the applied field B₀ by opposing it ........................................................................................................ 67

Figure 1.7.5: A doublet peak arising for nuclear spin A due to the neighbouring nuclear spin B. ...................................................................................................................... 69

Figure 1.7.6: A doublet peak arising for nuclear spin A due to the neighbouring nuclear spin B. ...................................................................................................................... 70

Figure 1.7.7: A short T₂ results in a greater spread in the frequency dimension producing a broader line A long T₂ will produce a sharper peak due to less spread of the frequencies. ........................................................................................................ 71

Figure 1.7.8: Schematic energy level diagrams and population differences for two spins I and S which share a dipolar coupling. ........................................................................ 72

Figure 1.7.9: Diagram illustrating the lineshapes in an NMR spectrum for slow, intermediate and fast exchange processes ......................................................... 74
**Figure 1.8.1:** The potential energy of the system is calculated by a series of molecular mechanic force fields which not only include the energy of the bonded terms but the long range non bonded interactions. .................................................................75

**Figure 1.8.2:** An ‘improper’ dihedral angle which can arise from carbonyl and aromatic groups..................................................................................................................77

**Figure 1.8.3:** Simulated annealing is a combination of molecular dynamics calculations and the change in temperature.............................................................................79

**Figure 2.2.1:** The Bruker Avance 1000MHz (1GHz) spectrometer at the Centre de RMN à Très Hauts Champs in Lyon (France). .........................................................................86

**Figure 2.3.1:** The 1D $^1$H NMR pulse sequence.................................................................................................................................92

**Figure 2.3.2:** The $^1$H-decoupled 1D NMR pulse sequence.........................................................93

**Figure 2.3.3:** The general scheme for any 2D NMR experiment..................................................94

**Figure 2.3.4:** The $^1$H-$^1$H COSY pulse sequence.................................................................95

**Figure 2.3.5:** A schematic 2D NMR spectrum illustrating the absorptive and dispersive nature of the diagonal and cross peaks.................................................................97

**Figure 2.3.6:** The pulse sequence for $^1$H-$^1$H double-quantum filtered COSY (DQF-COSY).................................................................................................................................99

**Figure 2.3.7:** $^1$H-$^1$H TOCSY pulse sequence...............................................................................100

**Figure 2.3.8:** The NMR pulse sequence for the $^1$H-$^13$C experiment ................................101

**Figure 2.3.9:** The $^1$H-$^13$C HSQC-TOCSY pulse experiment................................................102

**Figure 2.3.10:** An outline of the scalar coupling pathway mapped in a 2D HSQC-TOCSY experiment .................................................................................................................................102

**Figure 2.3.11:** A $^1$H-$^1$H NOESY pulse sequence........................................................................103

**Figure 2.3.12:** The magnetisation is inverted along the $-z$ axis in a NOESY experiment and in the magnetisation is spin-locked along the $y$-axis and parallel to the $B_1$ field in a ROESY experiment.................................................................................................105

**Figure 2.3.13:** $^1$H-$^1$H ROESY pulse sequence.................................................................106

**Figure 2.3.14:** The $^1$H-$^31$P CPMG-HSQC-NOESY pulse sequence........................................107

**Figure 2.3.15:** The pulse sequence for 3D NOE/2QC experiment ..............................................109

**Figure 2.3.16:** Schematic illustrating the connectivities observed in the 3D NOE/2QC experiment ........................................................................................................................................110

**Figure 2.4.1:** A schematic chart illustrating the various steps of the protocol followed to generate the NMR structures of the RNA motifs .............................................................................113

**Figure 2.4.2:** The Karplus curve.........................................................................................................................117

**Figure 2.5.1:** The general scheme outlining the various steps required to initiate the structure determination calculation of the antibiotics ..................................................................121

**Figure 2.6.1:** Base pair parameters describing the geometry of base pairs and sequential base pairs.........................................................................................................................128

**Figure 2.7.1:** A square hyperbola representing the relationship between the concentration of the bound ligand [RL] and free ligand [L] to reveal the $K_D$ ......................................................................131

**Figure 2.7.2:** Binding isotherms produced by monitoring the change in chemical shift of the bound ligand or receptor with respect to the free ligand concentration .........................132
Figure 3.1.1: The chemical structure of the bamicetin antibiotic..........................................................137
Figure 3.1.2: 600MHz 1D $^1$H NMR spectrum of bamicetin. The labile amide proton (V) appears as a broad resonance at 10.7ppm due to rapid exchange with the water solvent protons. .............................................................................................................................138
Figure 3.1.3: 600MHz $^1$H,$^13$C HSQC-TOCSY ($\tau_m$=75ms) spectrum of bamicetin The spectrum shows the assignment of the anomeric carbons C7 and C13 with their directly attached protons H7 and H13.................................................................139
Figure 3.1.4: 600MHz $^1$H-$^1$H TOCSY ($\tau_m$=75ms) spectrum of bamicetin. The methyl groups 3CH$_3$ and 9CH$_3$ show multiple bond $^1$J$_{H-H}$ scalar couplings to all of the protons.................................................................140
Figure 3.1.5: 600MHz $^1$H-$^1$H ROESY ($\tau_m$ = 400ms) spectrum of bamicetin. The plot shows the dipolar couplings observed between the methyl 18CH$_3$ and the phenyl H16 proton.................................................................140
Figure 3.1.6: 600MHz $^1$H-$^1$H ROESY ($\tau_m$ = 400ms) spectrum of bamicetin. The plot exhibits the dipolar couplings observed for the cytosine proton (H14) and the protons on the second saccharide ring (II). .........................................................................................141
Figure 3.1.7: (a) 400MHz 2D $^1$H-$^13$C HMBC spectrum and (b) 1D $^13$C {$^1$H} NMR spectrum of bamicetin The three bond scalar coupling between the H16 proton and the adjacent carbonyl carbon is observed and confirmed........................................................................142
Figure 3.1.8: 600MHz $^1$H-$^13$C HSQC-TOCSY ($\tau_m$=75ms) spectrum of bamicetin. The plot displays the assignment of the pyranose carbons C2, C9, C11 and C12 and directly attached protons. .........................................................................................143
Figure 3.1.9: 600MHz $^1$H-$^1$H ROESY ($\tau_m$=400ms) spectrum of bamicetin. Important distance based constraints were generated from the dipolar correlated cross peaks observed between protons on both saccharide rings ........................................................................144
Figure 3.1.10: 600MHz 1D $^1$H NMR spectra illustrating the effect of increase in temperature of bamicetin..............................................................................................................................................146
Figure 3.1.11: The energy minimised NMR structure an overlay of the ten best structures of bamicetin antibiotic..............................................................................................................................................................150
Figure 3.1.12: The energy minimised NMR structure an overlay of the ten best structures of bamicetin antibiotic..............................................................................................................................................................151
Figure 3.2.1: The chemical structure of sparsomycin consists of a methylated uracil ring and a sulphur rich ‘tail’ linked by an amide group. .............................................................................................................152
Figure 3.2.2: The 600MHz 1D $^1$H NMR spectrum of sparsomycin antibiotic .................153
Figure 3.2.3: 600MHz $^1$H-$^1$H ROESY ($\tau_m$=400ms) spectrum of sparsomycin. The plot shows the through space dipolar connectivities observed between the different proton groups of the antibiotic........................................................................................................................................154
Figure 3.2.4: 600MHz $^1$H-$^13$C HSQC spectrum of sparsomycin. The spectrum displays the resolved $^1$J$_{C-H}$ couplings which distinguished the methylene group protons (H11/H12, H13/H14 and H16/H17).............................................................................................................155
Figure 3.2.5: The three staggered conformations of the enantiomeric protons of sparsomycin. ........................................................................................................................................................................156
Figure 3.2.6: 600MHz ^1H-^1H ROESY (\(\tau_m=400\)ms) spectrum of sparsomycin. The spectrum illustrates a number of ROEs identified amongst the various protons of the antibiotic.................................................................157

Figure 3.2.7: 1D ^1H NMR variable temperature series of sparsomycin. .........................159

Figure 3.2.8: 600MHz ^1H-^1H ROESY (\(\tau_m = 400\)ms) spectrum of sparsomycin. The plot shows the various ROE connectivities observed between the methylene protons (13/14H) to the nearby protons. .................................................................160

Figure 3.2.9: The average NMR structure and overlay of ten structures of sparsomycin. .................................................................163

Figure 3.3.1: The chemical structure of anisomycin consists of a range of functional groups.................................................................165

Figure 3.3.2: 1D ^1H NMR spectra of anisomycin recorded on first day of sample preparation and after 3 months.................................................................166

Figure 3.3.3: 400MHz 1D ^1H NMR spectrum of anisomycin. ........................................167

Figure 3.3.4: 500MHz 1D ^1H NMR spectrum of sodium acetate..............................168

Figure 3.3.5: 400MHz ^1H-^1H TOCSY spectrum (\(\tau_m = 75\)ms) of anisomycin.............170

Figure 3.3.6: A 400MHz ^1H-^1H ROESY spectrum (\(\tau_m=400\)ms) of anisomycin. A number of ROEs were obtained from the spectrum which provided valuable distance constraints to link the pyrrolidine and aromatic residues.........................171

Figure 3.3.7: A 2D ^1H-^13C HSQC spectrum of anisomycin. ........................................172

Figure 3.3.8: The energy minimised NMR structure of anisomycin and the 10 best accepted structures. .................................................................175

Figure 3.4.1: The NMR structures of the bamicetin, sparsomycin and anisomycin antibiotics. .................................................................................................177

Figure 4.1.1: The secondary structure of the double stranded B.subtilis 27mer RNA motif used for study by NMR. ........................................................................180

Figure 4.1.2: 600MHz 1D ^1H NMR spectrum of B.subtilis 27mer RNA displaying the resonances of different proton groups.................................................................181

Figure 4.1.3: 600MHz ^1H-^1H NOESY (\(\tau_m =100\)ms) spectrum of B.subtilis 27mer RNA. illustrates the sequential assignment of the imino proton resonances and the observed sequential imino-imino connectivities. .................................................................183

Figure 4.1.4: 600MHz ^1H-^1H NOESY (\(\tau_m =100\)ms) spectrum of B.subtilis 27mer RNA, the imino proton resonance assignment were confirmed by the identification of connectivities to the NH2, aromatic and ribose protons........................................184

Figure 4.1.5: 800MHz 1D ^1H NMR spectrum of B.subtilis 27mer RNA. Adenine H2 protons do not share scalar coupling with other protons and so produce sharp singlet peaks in the aromatic region. .................................................................185

Figure 4.1.6: 600MHz ^1H-^13C HSQC spectrum of the B.subtilis 27mer RNA. The four adenine H2 protons were unambiguously assigned based on their correlation to the C2 resonances........................................................................................................187
Figure 4.1.7: 600MHz $^1$H-$^1$H TOCSY ($\tau_m = 75$ms) spectrum of the *B.subtilis* 27mer RNA. The figure displays the strong H5-H6 couplings due to the pyrimidine bases cytosine and uracil..................................................................................................................188

Figure 4.1.8: 600MHz $^1$H-$^{13}$C HSQC spectrum of the *B.subtilis* 27mer RNA displays the $^1J$ H5-C5 correlations of the cytosine bases and the uracil bases..........................189

Figure 4.1.9: 600MHz $^1$H-$^{13}$C HSQC spectrum of the *B.subtilis* 27mer RNA. The glycosidic H1’ protons were identified by their C1’ correlation to the respective carbons. ..............................................................................................................190

Figure 4.1.10: 600MHz $^1$H-$^1$H DQF-COSY spectrum of *B.subtilis* 27mer RNA. The scalar correlated H1’-H2’ cross peaks were constrained with C$_2$-endo sugar puckers. ..................................................................................................................191

Figure 4.1.11: An illustration of the NOE contacts produced between the H6 base proton to the intraresidual glycosidic H1’ proton and a second NOE from H1’ to the H8 aromatic proton on the subsequent nucleotide in the NOESY spectrum........192

Figure 4.1.12: 600MHz $^1$H-$^1$H NOESY ($\tau_m = 400$ms) spectrum of *B.subtilis* 27mer RNA. Sequential connectivities observed between the aromatic H6/H8 and the sugar H1’ protons ......................................................................................................................194

Figure 4.1.13: 600MHz $^1$H-$^1$H NOESY spectrum of *B.subtilis* 27mer RNA. Sequential connectivities observed between the aromatic H6/H8 and the sugar H1’ protons. ..195

Figure 4.1.14: 81MHz 1D {[H-decoupled]} $^{31}$P –NMR spectrum of *B.subtilis* 27mer RNA...............................................................................................................................196

Figure 4.1.15: 2D $^1$H-$^{31}$P CPMG-HSQC-NOESY ($\tau_m = 500$ms) spectrum of *B.subtilis* 27mer RNA ..................................................................................................................................................198

Figure 4.1.16: The NMR structure of *B.subtilis* 27mer RNA and an overlay of the ten lowest energy NMR structures ................................................................................................................202

Figure 4.1.17: Illustration of the U6 internal loop in the NMR structure of *B.subtilis* 27mer RNA ..................................................................................................................................................204

Figure 4.1.18: Non canonical Watson-Crick base pairing between A9 and A27 ........205

Figure 4.2.1: 600MHz 1D $^1$H NMR spectrum of *B.subtilis* 27mer RNA and containing varying amounts of bamicetin (in molar equivalent)............................................209

Figure 4.2.2: 600MHz $^1$H-$^1$H NOESY spectra of *B.subtilis* 27mer RNA, *B.subtilis* 27mer RNA-bamicetin (1:1) complex and *B.subtilis* 27mer RNA-bamicetin (1:10) complex. .................................................................212

Figure 4.2.3: The changes in the line widths of the G10 and G7 imino proton resonances of *B.subtilis* 27mer RNA with incremental addition of bamicetin. .........................213

Figure 4.2.4: 400MHz 1D $^1$H NMR spectrum of *B.subtilis* 27mer RNA and containing varying amounts of bamicetin (in molar equivalents).............................215

Figure 4.2.5: 600MHz $^1$H-$^1$H TOCSY spectra of *B.subtilis* 27mer RNA and *B.subtilis* 27mer RNA-bamicetin (1:10) complex. .................................................................216

Figure 4.2.6: NOE growth curves for the U6 and U26 H5-H6 resonances of *B.subtilis* 27mer RNA ..................................................................................................................217
Figure 4.2.7: 81MHz 1D $^{31}$P spectra of B.subtilis 27mer RNA (0.59mM) and B.subtilis 27mer RNA-bamicetin (1:10) complex. .................................................................218

Figure 4.2.8: 400MHz 1D $^1$H NMR spectra of bamicetin, B.subtilis 27mer RNA, (c-g) 27mer B.subtilis RNA containing varying amounts of bamicetin (in molar equivalents). .................................................................219

Figure 5.1.1: The secondary structure of the 25mer stem-loop motifs of the GTPase domain II H.hal 23S rRNA. .................................................................................................................222

Figure 5.1.2: 1D $^1$H NMR spectrum of the H.hal 25mer RNA. .................................................................223

Figure 5.1.3: 800MHz $^1$H-$^1$H NOESY ($\tau_m = 250$ms) spectrum of the H.hal 25mer RNA. .................................................................................................................225

Figure 5.1.4: 800MHz $^1$H-$^1$H NOESY ($\tau_m = 250$ms) spectrum of H.hal 25mer RNA. The assignment of the imino protons was confirmed in the amino region with correlations to the exchangeable amino and non exchangeable ribose and base protons........226

Figure 5.1.5: 800MHz 1D $^1$H NMR spectrum of the H.hal 25mer RNA. .................................................................228

Figure 5.1.6: 600MHz $^1$H-$^1$H TOCSY ($\tau_m$=75ms) spectrum displaying the pyrimidine H5-H6 correlated resonances of the H.hal 25mer RNA. .................................................................229

Figure 5.1.7: 600MHz $^1$H-$^{13}$C HSQC spectrum of the H.hal 25mer RNA, the plot displays the assigned H1'-'C1' correlated resonances.................................................................230

Figure 5.1.8: 600MHz 3D NOE/2QC spectrum of H.hal 25mer RNA.................................................................232

Figure 5.1.9: 600MHz $^1$H-$^1$H NOESY spectrum ($\tau_m = 250$ms) of H.hal 25mer RNA illustrating the H6/H8-H1' sequential assignment.................................................................233

Figure 5.1.10: The NMR structure of the H.hal 25mer RNA and an overlay of the ten best converged NMR structures. .................................................................................................................238

Figure 5.1.11: Conformation of the nanoloop motif of H.hal 25mer RNA. .................................................................239

Figure 5.1.12: The NMR structure of the 25mer RNA.................................................................................................................240

Figure 5.2.1: 400MHz 1D $^1$H NMR spectra of (a-c) H.hal 25mer RNA (0.6mM) in its complex with amicetin. .................................................................................................................244

Figure 5.2.2: 400MHz $^1$H-$^1$H NOESY ($\tau_m = 250$ms) spectra of H.hal 25mer RNA (0.6mM) and varying amounts of amicetin.................................................................246

Figure 5.2.3: 400MHz 1D $^1$H NMR line width plots of the imino proton resonances U4, G7 and U9 measured.................................................................247

Figure 5.2.4: 600MHz $^1$H-$^1$H NOESY spectrum ($\tau_m = 250$ms) of H.hal 25mer RNA with no amicetin illustrating the H6/H8-H1' sequential assignment and displaying the amicetin induced broadening of the NOE cross peaks.................................................................248

Figure 5.2.5: 162MHz 1D $^{31}$P NMR spectra of H.hal 25mer RNA and containing varying amounts of amicetin.................................................................249

Figure 5.2.6: 1GHz $^1$H NMR spectrum of H.hal 25mer RNA-amicetin complex (1:5)..250

Figure 5.2.7: 1GHz $^1$H-$^1$H NOESY ($\tau_m = 250$ms) spectra of 25mer H.hal RNA-amicetin (1:5) complex and B.subtilis 27mer RNA-bamicetin.................................................................251
List of Tables

Table 1.1.1: A summary of the X-ray crystal and cryo-EM structures of different ribosomes solved at increasingly higher resolution. ...............................................................27
Table 1.5.1: A summary of the released entries in the RCSB PDB database as of 16th August 2011. .................................................................................................................................60
Table 2.4.1: Proton distance types in RNA. ......................................................................................112
Table 3.1.1: $^1$H, $^{13}$C chemical shifts ($\delta$, ppm) and coupling constants ($^3$J, Hz) of bamicetin..........................................................................................................................147
Table 3.1.2: The NMR experimental ROE distance and dihedral angle constraints input into the XPLOR-NIH program for the structure determination of bamicetin.............................149
Table 3.2.1: $^1$H, $^{13}$C chemical shifts ($\delta$, ppm) and coupling constants ($^3$J, Hz) obtained from the 1D $^1$H NMR spectrum of sparsomycin........................................................................158
Table 3.2.2: A grid displaying the couplings obtained from the scalar based $^1$H-$^1$H TOCSY and the dipolar coupled $^1$H-$^1$H ROESY experiments................................................161
Table 3.2.3: The total number of NMR experimental ROEs and dihedral angle constraints input into the XPLOR-NIH structure determination program to generate the solution structure of sparsomycin. ..............................................................................................162
Table 3.3.2: The total number of NMR experimental ROEs and dihedral angle constraints input into the XPLOR-NIH structure determination program to determine the structure of anisomycin. ....................................................................................................................174
Table 4.1.1: $^1$H, $^{13}$C, $^{31}$P chemical shifts ($\delta$, ppm) of the B.subtilis 27mer RNA..............199
Table 4.1.2: A summary of the experimental constraints used as input into the structure calculation of the B.subtilis 27mer RNA. ..................................................................................201
Table 4.1.3: Average RMSD values of the whole motif and the nucleotides of the two bulges. ........................................................................................................................................203
Table 4.1.4: Hydrogen bonding between the canonical and non canonical base pairs of the NMR structure of B.subtilis 27mer RNA..............................................................................205
Table 4.1.5: Local helical parameters of the NMR structure of B.subtilis 27mer RNA. .....................206
Table 4.1.6: Local helical parameters of the NMR structure of B.subtilis 27mer RNA. .....................208
Table 5.1.1: $^1$H, $^{13}$C and $^{31}$P chemical shifts ($\delta$, ppm) of the H.hal 25mer RNA..............235
Table 5.1.2: A summary of the experimental constraints used as input into the structure determination calculation of the \textit{H.hal} 25mer RNA. .......................................................237

Table 5.1.3: Summary of the hydrogen bonds observed within the base pairs. ...............239

Table 5.1.4: Local base-pair and base-pair step parameters for the 25mer RNA. ..........242

Table 5.1.5: Local base-pair helical parameters for the 25mer RNA..........................243
Symbols and Abbreviations

1D  One dimensional
2D  Two dimensional
3D  Three dimensional
A^{260}  Absorbance at a wavelength of 260nm
Å  Angstrom
°  Degree
B  Applied magnetic field
Da  Daltons
Dx  Shift
Dy  Slide
Dz  Rise
E  Energy
e^-  Electron
F_1  First frequency dimension
F_2  Second frequency dimension
F_3  Third frequency dimension
Hz  Hertz
h-rise  Helical rise
I  Nuclear spin quantum number
J_{A,B}  Scalar coupling between spins A and B
k  Rate of inter conversion
K  Kelvin
K^+  Potassium ions
K_D  Dissociation constant
m  Quantum magnetic moment
M  Molar
mM  millimolar
M_y  Transverse magnetisation
M_Z  Bulk magnetisation along z-axis
M_w  Molecular weight
NH  Imino
NH_2  Amino
p  Para
P  Angular momentum
ppm  Frequency parts per million
ps  Picoseconds
P_0  Zero-order phase correction
P_1  First-Order phase correction
R  Relaxation matrix
Sx  Shear
Sy  Stretch
Sz  Stagger
SLy  Spin-lock mixing time
S/N  Signal to noise ratio
t₁, t₂, t₃  Evolution and detection periods
T  Temperature
Tₘ  UV melting temperature
T  Tilt
T₁  Longitudinal relaxation
T₂  Transverse relaxation
W₀  Zero quantum transition
W₁  Single quantum transition
W₂  Double quantum transition
X-disp  x-displacement
Y-disp  y-displacement
{^{1}H}  Decoupled proton nuclei
α, β, δ, ε, ζ  Sugar phosphate backbone dihedral angles
ν₀, ν₁, ν₂, ν₃, ν₄, χ  Ribose sugar ring dihedral angles
δ  Chemical shift
η  Inclination
γ  Gyromagnetic ratio
κ  Buckle
μ  Magnetic moment
μM  Micromolar
μL  Micro litres
Ω  Twist
π  Propeller
ρ  Roll
σ  Opening
σ  Shielding constant
τ  Delay
τₑ  Rotational correlation time
τₘ  Mixing Time
θ  Tip
ν  Frequency
Δ₁,₂  Delay times in ^1H-^13C HSQC experiment
ΔG  Gibbs free energy
α  Lower energy state of nuclei
β  Higher energy state of nuclei
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>U</td>
<td>Uracil</td>
</tr>
<tr>
<td>A-site</td>
<td>Aminoacyl site</td>
</tr>
<tr>
<td>Ami&lt;sup&gt;i&lt;/sup&gt;</td>
<td>Amicetin resistant mutant</td>
</tr>
<tr>
<td>B3LYP</td>
<td>Becke three-parameter Lee-Yang-Parr</td>
</tr>
<tr>
<td>B.&lt;i&gt;subtilis&lt;/i&gt;</td>
<td>&lt;i&gt;Bacillus subtilis&lt;/i&gt;</td>
</tr>
<tr>
<td>BMRB</td>
<td>Biological magnetic resonance bank</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>COSY</td>
<td>Correlation spectroscopy</td>
</tr>
<tr>
<td>CPMG</td>
<td>Carr-Purcell-Meiboom-Gill</td>
</tr>
<tr>
<td>DANTE</td>
<td>Delays alternating with nutation for tailored excitation</td>
</tr>
<tr>
<td>DFT</td>
<td>Density functional theory</td>
</tr>
<tr>
<td>DQ</td>
<td>Double quantum</td>
</tr>
<tr>
<td>DQF-COSY</td>
<td>Double quantum filtered correlation spectroscopy</td>
</tr>
<tr>
<td>DMS</td>
<td>Dimethyl sulphate</td>
</tr>
<tr>
<td>&lt;i&gt;D.radiodurans&lt;/i&gt;</td>
<td>&lt;i&gt;Deinococcus radiodurans&lt;/i&gt;</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribose nucleic acid</td>
</tr>
<tr>
<td>&lt;i&gt;E.coli&lt;/i&gt;</td>
<td>&lt;i&gt;Escherichia coli&lt;/i&gt;</td>
</tr>
<tr>
<td>&lt;i&gt;E.faecium&lt;/i&gt;</td>
<td>&lt;i&gt;Enterococcus faecium&lt;/i&gt;</td>
</tr>
<tr>
<td>EF</td>
<td>Elongation factors</td>
</tr>
<tr>
<td>EM</td>
<td>Electromagnetic</td>
</tr>
<tr>
<td>E-site</td>
<td>Exit-site</td>
</tr>
<tr>
<td>fMet</td>
<td>Formylmethionine</td>
</tr>
<tr>
<td>GAFF</td>
<td>General amber force field</td>
</tr>
<tr>
<td>GARP</td>
<td>Globally optimised alternating phase rectangular pulses</td>
</tr>
<tr>
<td>GHz</td>
<td>Gigahertz</td>
</tr>
<tr>
<td>&lt;i&gt;H.hal&lt;/i&gt;</td>
<td>&lt;i&gt;Halobacterium halobium&lt;/i&gt;</td>
</tr>
<tr>
<td>&lt;i&gt;H.marismortui&lt;/i&gt;</td>
<td>&lt;i&gt;Halobacterium marismortui&lt;/i&gt;</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear single quantum coherence</td>
</tr>
<tr>
<td>IDSA</td>
<td>Infectious disease society of America</td>
</tr>
<tr>
<td>IF</td>
<td>Initiation factors</td>
</tr>
<tr>
<td>ISPA</td>
<td>Isolated spin pair approximation</td>
</tr>
<tr>
<td>LP</td>
<td>Linear prediction</td>
</tr>
<tr>
<td>mer</td>
<td>Oligomer</td>
</tr>
<tr>
<td>MHz</td>
<td>Megahertz</td>
</tr>
<tr>
<td>MLEV</td>
<td>Malcolm Levitts CPD sequence</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MK</td>
<td>Merzsingh Kollman</td>
</tr>
<tr>
<td>MQ</td>
<td>Multiple quanta</td>
</tr>
<tr>
<td>MRSA</td>
<td><em>Methicillin-resistant staphylococcus aureus</em></td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MP</td>
<td>Møller Plesset</td>
</tr>
<tr>
<td>NDB</td>
<td>Nucleic acid database</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>Nuclear overhauser effect</td>
</tr>
<tr>
<td>NOESY</td>
<td>Nuclear overhauser effect spectroscopy</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotides</td>
</tr>
<tr>
<td>P-site</td>
<td>Peptidyl-site</td>
</tr>
<tr>
<td>PFGs</td>
<td>Pulsed field gradients</td>
</tr>
<tr>
<td>PT</td>
<td><em>Peptidyl transferase</em></td>
</tr>
<tr>
<td>PTC</td>
<td><em>Peptidyl transferase centre</em></td>
</tr>
<tr>
<td>RCSB PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>ref</td>
<td>Reference</td>
</tr>
<tr>
<td>RF</td>
<td>Radiofrequency</td>
</tr>
<tr>
<td>rMD</td>
<td>Restrained molecular dynamics</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>ROE</td>
<td>Rotating-frame overhauser effect</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>Thi'</td>
<td>Thioistrepton-resistant mutant</td>
</tr>
<tr>
<td>TOCSY</td>
<td>Total correlation spectroscopy</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
<tr>
<td><em>T. thermophilus</em></td>
<td><em>Thermus thermophilus</em></td>
</tr>
<tr>
<td>UV</td>
<td>Ultra violet</td>
</tr>
<tr>
<td>WATERGATE</td>
<td>Water suppression by gradient tailored excitation</td>
</tr>
<tr>
<td>WC</td>
<td>Watson Crick</td>
</tr>
<tr>
<td>W.T.</td>
<td>Wild type</td>
</tr>
<tr>
<td>ZF</td>
<td>Zero filling</td>
</tr>
</tbody>
</table>
Abstract
The University of Manchester
Misbah Nareen
Doctor of Philosophy PhD
NMR structural studies of the binding of peptidyl transferase antibiotics to conserved secondary structural motifs of 23S ribosomal RNA
24/08/2011

The peptidyl transferase centre (PTC) of 23S ribosomal RNA is the target for a number of antibiotics which inhibit protein synthesis. The precise mode of binding of these antibiotics is largely unknown and hence is an active area of research in structural biology.

The NMR solution structures of three PT antibiotics, bamicetin, sparsomycin and anisomycin have been successfully characterised using a range of two-dimensional NMR techniques and restrained molecular dynamics. The NMR structures of these antibiotics provided valuable first hand insight into their conformations, since no X-ray crystal structures of the antibiotics in their free states have been determined so far. Bamicetin adopts a folded conformation possibly held by intramolecular hydrogen bonds and similar to the published NMR structure of amicetin.

These antibiotics generate spontaneous single nucleotide mutants upon prolonged exposure and bamicetin and sparsomycin are universal PT inhibitors, interacting with all three evolutionary domains of 23S rRNAs. The amicetin antibiotic produces a spontaneous single mutation U2457C in the Halobacterium halobium (H.hal) 23S rRNA and the binding site is predicted to be very close to this nucleotide. The similarity in chemical structure with amicetin, suggests bamicetin to target the same binding site on the 23S rRNA. Both bamicetin and sparsomycin show exchange retarded amide proton resonances in the NMR spectrum, akin to other amicetin family antibiotics, indicating the retarded exchange to be a characteristic feature in the native solution state.

The Bacillus subtilis (B.subtilis) 70S ribosomes have strong affinity for bamicetin and so a highly conserved 27mer RNA motif containing the possible binding site was selected for NMR structure determination and bamicetin binding studies. The greater number of imino proton resonances observed together with the high quality of the determined structure of the motif proved that B.subtilis rRNA is more stable than E.coli and H.hal rRNAs.

The B.subtilis 27mer rRNA-bamicetin interaction studies revealed a fast exchange, weak binding system and careful analysis of line width and chemical shifts indicated changes at the local conformation of the RNA after binding. To probe the cross-hypersensitivity phenomenon, a 25mer RNA corresponding to the thiostrepton-resistant mutant (G1159) residing in the domain II of H.hal 23S rRNA was chosen for NMR structure determination and amicetin binding. Discrete chemical shift changes and NOESY experiments using ultrahigh field 1GHz NMR revealed weak interactions.

The structures of the antibiotics and analysis of their dynamics as well as interactions with the RNA motifs of different organisms have yielded important information in understanding their binding and inhibitory activities at the atomic level. The results can be used for generating new or hybrid antibiotics to tackle the escalating problem of antibiotic resistance.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/about us/regulations) and in The University’s policy on presentation of Theses.
Acknowledgements

First and foremost I would like to thank Allah (swt) for giving me the opportunity to undertake this research and to complete it successfully. I sincerely thank my supervisor Dr V. Ramesh for his immense dedication, enthusiasm and for completely pushing me outside the comfort zone throughout this research project. I express thanks to my senior colleagues John King and Tony Cheung for their unselfish assistance and guidance throughout my research. I would like to thank my ‘slumdog PhD’ group members; Misbah Ghafoor, Sadia Mohammad, Usman Rasul and Nick Chan for being great friends to work with on a daily basis and made the demanding research less hardwork and more fun to tackle. I also appreciate the many hours they spent tackling the abundance of section breaks, which they will forever remember as my thesis.

I am grateful to Tom Frankiel, Geoff Kelly and Alain Origiani at the National Institute Medical Research Centre (NIMR) in Mill Hill, London; for providing both invaluable support and NMR spectrometer time. I appreciate the generous NMR spectrometer time and support provided by Roger Speak at the NMR department in University of Manchester. For access to the 1GHz NMR spectrometer, I would like to thank Moreno Lelli at the Centre de RMN à Très Hauts Champs in Lyon, France.

I would like to sincerely thank Ayeesha Mujeeb for her unwavering support throughout the last four years, while undergoing many testing times of her own and a family friend, Rahil Jameel for providing invaluable assistance with some difficult formatting issues.

Finally, I would like to wholeheartedly acknowledge my family for their love, encouragement, guidance, constructive criticism and irreplaceable support both emotionally and financially.
Dedication

To my mum, Sughran Bibi
I want to thank you for being the most beautiful and important influence in my life. You have always supported me, pushed me to achieve better things and to never be afraid of hard work. For one to give up their own happiness for the sake of others is a very rare and difficult thing to do, but you have done this on countless occasions and I hope my PhD is some compensation to you for your many sacrifices. I love you and will always try to make you proud.
Chapter 1  Introduction

1.1  New Biology

1.1.1  Advances made in biological science and tackling antibiotic resistance

The use of the penicillin antibiotic in 1942 opened the doors for successfully treating infections, undergoing cancer chemotherapies, routine operations and transplants were able to be carried out without the threat of deadly staphylococcus aureus (S. aureus). For the remainder of the 20th century, stronger and more effective antibiotics were discovered, synthesised and used and the future looked promising in tackling infections. However, at the turn of the 21st century, infections were becoming more diverse and resistant to a range of antibiotics. In 2004, a 70 year old man died from the vancomycin-resistant Enterococcus faecium (E. faecium) infection, despite being treated with the best possible antibiotics.1 This example was amongst many others indicating that a full circle had been reached with antibiotics, from successfully treating infections to the present where it has become difficult to tackle multi-drug resistant superbugs.

In recent years, the pace of development of new drugs has slowed down in comparison to the fairly rapid growth in demand. Prominent infectious disease specialists are warning the general public about the increasing threat of antibiotic resistance and believe this is the number one problem faced by the medical community. It may reach a situation where health care professionals are unable to treat all patients who develop infections with the appropriate medication.2,3

A new superbug originating in India and Pakistan was found in UK hospitals in 2010 which was resistant to all known antibiotics.4 The NDM-1 is an enzyme produced by the bacteria Enterobacteriaceae and can exist in various bacteria such as Escherichia coli (E. coli). The ability to transfer to other bacteria makes it an effective resistance mechanism against the most powerful group of antibiotics, the Carbapenems. Carbapenems are reserved for emergencies and as a last resort to combat multi-resistant bacterial infections when other antibiotics have failed to yield positive results.
At the present time, most of the bacteria carrying the NDM-1 enzyme have been successfully treated through a combination of antibiotics, quarantining the patient and increasing the levels of hygiene in hospitals. Similar NDM-1 infections were observed in the USA, Canada, Australia and the Netherlands and with unprecedented human air travel and migration; it is a grave worry that NDM-1 may become endemic worldwide.

The gram positive bacteria methicillin-resistant *staphylococcus aureus* (MRSA) and *E. faecium* are difficult problems for the pharmaceutical community to tackle. More than 50% of *S. aureus* isolates recovered in US hospitals (2003) belong to MRSA.\(^1\) Unfortunately, none of the new antibiotics including Linezolid and Daptomycin are better than the current drug vancomycin at treating MRSA.

*E. faecium* which was previously treated by vancomycin and ampicillin has developed increasing drug resistance. The situation is more disheartening for gram negative infections, with no new antibiotics in the advanced stages of clinical development to tackle multi-drug resistant organisms. Even some strains have adapted to be effective against the carbapenem antibiotics.

The problem of eradicating infections caused by the antibiotic resistant superbugs has intensified by the lack of new drugs and through their overuse causing a loss of effectiveness (Figure 1.1.1). Unnecessarily prescribing antibiotics to patients who do not need them (i.e. to treat a common cold which is a virus and unaffected by antibiotics) will make their immune system weaker and as a result more prone to infections.\(^2\) The Infectious Diseases Society of America (IDSA) has made fighting antibiotic resistance one of its top priorities. It launched the ‘bad bugs, no drugs’ campaign in 2004 to educate health care professionals and the general public about the dangers of antibiotic resistance and to discourage their excessive usage.\(^5\)
Figure 1.1.1: A diagram illustrating the combination of two major factors (blue and green circles) which are causing antibiotic resistance (orange circle). A reduction in the excessive use of antibiotics and increasing the investments in their research is required in order to fight the antibiotic resistance.

To tackle the problem of multi-drug resistant superbugs, it is most likely that old compounds originally abandoned due to toxicity are revisited or hybrid antibiotics are developed. This must be done in conjunction with better diagnostics, infection controls and immunisation to effectively suppress the problem. Unquestionably, the problem of multi-drug resistant superbugs has global consequences and so it must be a united effort on the part of academic researchers and institutions, pharmaceutical industries and governments to gain control in the battle against bacterial infections.
1.1.2 Nobel Prize in chemistry 2009

The 2009 Nobel Prize in chemistry was awarded to three scientists, Ada E. Yonath, Thomas A. Steitz and Venkatraman Ramakrishnan for their work in mapping the ribosome at the atomic level. Ribosomes are the macromolecular machines which produce proteins to control the chemistry of all living creatures from bacteria to humans. Since no living organisms can survive without ribosomes, they are the perfect target for antibacterial drugs.

In 1980, A. Yonath obtained three dimensional crystals of *Geobacillus stearothermophilus* and was the instrumental person joined by T. Steitz and V. Ramakrishnan in initiating the discovery of high resolution crystalline structures of ribosomal subunits.

The 50S subunit of *Haloarcula marismortui* (*H.marismortui)* at 2.4 Å was reported by Steitz and his collaborators which proposed a mechanistic model for peptidyl transfer during protein synthesis. Ramakrishnan and Yonath both separately published 30S structures of the *Thermus thermophilus* (*T.thermophilus*) at 3.0 Å and 3.2 Å, respectively. The 30S subunit of *T.thermophilus* provided knowledge of the poorly understood mechanism of accurate codon reading of mRNA during peptide bond formation. Yonath also determined the crystalline structure of the 50S subunit of *Deinococcus radiodurans* (*D.radiodurans*), while Ramakrishnan reported the high resolution (2.8 Å) X-ray crystal structure of the 70S ribosome of *T.thermophilus* in the pre-translocation state. The structure demonstrated how the tRNA and mRNA interact with the ribosome.

Obtaining the structures of the two ribosomal subunits at high resolution revealed new mechanistic insights of the ribosome, the catalysis of peptide bond formation and binding sites of the antibiotics as well as providing support for previous predictions. One finding was that the peptidyl transferase centre (site of peptide bond formation) lacked ribosomal protein components and this was proven, with no protein within ~18 Å of the peptidyl transferase centre.
Chapter 1 - Introduction

The work carried out by the Nobel Laureates has provided new insight into the ribosomes innermost workings, allowing us to gain a greater understanding of their mechanism and for developing new antibiotics.

1.1.3 Ribosome crystal structures

The last thirty years has yielded increased activity in the development of crystalline structures of the ribosomal subunits. The table below is a short summary of the determined X-ray crystal structures of various ribosomes.

Table 1.1.1: A summary of the X-ray crystal and cryo-EM structures of different ribosomes solved at increasingly higher resolution.

<table>
<thead>
<tr>
<th>Year</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984-1985</td>
<td>3D crystals of 50S subunit of <em>T.thermophilus</em> and <em>H.marismortui</em> ribosomes.</td>
</tr>
<tr>
<td>1991</td>
<td>50S subunit of <em>H.marismortui</em> at 3.0Å.</td>
</tr>
<tr>
<td>2000</td>
<td>50S subunit of <em>H.marismortui</em> at 2.4Å, 30S subunit of <em>T.thermophilus</em> at 3.0Å.</td>
</tr>
<tr>
<td>2001</td>
<td>50S subunit of <em>D.radiodurans</em> at 3.1Å.</td>
</tr>
<tr>
<td>2003</td>
<td>Cryo-EM <em>E.coli</em> ribosome with bound peptidyl-tRNA in the A-site.</td>
</tr>
<tr>
<td>2005</td>
<td>70S of <em>E.coli</em> at 3.5Å.</td>
</tr>
<tr>
<td>2006</td>
<td>70S of <em>T.thermophilus</em> at 2.8Å and 50S docking to 30S of <em>E.coli</em> subunit alongside mRNA and IFs.</td>
</tr>
<tr>
<td>2009</td>
<td>Cryo-EM of yeast 80S ribosome.</td>
</tr>
</tbody>
</table>
1.2 The importance of ribosomal RNA in protein synthesis

1.2.1 Biophysical properties of RNA

RNA is a polymeric molecule with alternating monosaccharide and phosphate units and each monosaccharide carries one of four heterocyclic nitrogenous bases. Two bases are derivatives of purine (adenine and guanine) and two are derivatives of pyrimidine (cytosine and uracil). The monosaccharide is a five carbon sugar (pentose), it exists in a β-D-ribofuranose orientation and is simply known as ribose, so the nucleic acid is referred to as ribonucleic acid (RNA). The ribose is closely related to the monosaccharide of DNA, deoxyribose, in which the hydroxyl group on the C2 position is replaced by a hydrogen atom.

The ribose-heterocyclic base units (nucleosides) are covalently linked together by phosphodiester bonds between the 5’-hydroxyl group of one sugar and 3’-hydroxyl group of the next sugar, thus producing a linear polymer of nucleotides (Figure 1.2.1). However, in RNA the nucleosides can also bind to the 2’-OH position of the ribose ring. The chain is directional with its two ends being non identical and the structure of the RNA is written in the 5’→3’ direction. The individual nucleotides are covalently bonded via the phosphate groups to form the RNA strands.
The overall structures of RNA and DNA are very similar; however, they differ in three important structural details.

- RNA is usually single-stranded, while DNA is double-stranded. But, RNA can form complex secondary and tertiary structures (Figure 1.2.2), increasing the complexity of roles it can undertake.
RNA nucleotides contain the ribose sugar while DNA has the deoxyribose sugar as in Figure 1.2.1. The hydroxyl group at the C2’ position of RNA is replaced by a hydrogen in DNA. The presence of the hydroxyl group in RNA pushes the ribose ring to exist in a C3’-endo conformation (Figure 1.2.3). The ribose is a five carbon sugar (pentose), in which the aldehyde and ketone functional groups react with hydroxyl functional groups to produce ring structures.

Figure 1.2.3: Conformation of the ribose ring of RNA (left) and the deoxyribose ring of DNA (right). The ribose ring contains the 3’-carbon (C3’-endo conformers) on the same side of the C4’-O4’-C1’ plane as the base and the C5’, whereas the sugar ring of DNA contains the 2’-carbon (C2’-endo conformers).

RNA has the pyrimidine uracil base rather than thymine, which is present in DNA. The chemical structures of the two bases differ only by the presence of a methyl group on the thymine.
There are several properties which the bases of nucleic acids all adopt. The most important features are that the bases are planar or almost planar. They are conjugated systems, converting between keto and enol forms (keto-enol tautomerism) in high pH. The bases absorb strongly in the UV region, as a result, are highly susceptible to damage and breakdown by background radiation. RNA can take part in alkali hydrolysis due to the 2’-OH group on the ribose rendering the RNA susceptible to strand cleavage in alkali solutions.

RNA does not produce DNA type helical structures with the base pairs perpendicular to the helical axis due to the hydroxyl group positioned on the C2 of the ribose ring and instead forms duplexes with base pairs tilted at ~20° from the axis. As a result, this action forces the ribose into a C3’-endo sugar conformation unlike the C2’-endo conformation (Figure 1.2.3) adopted by DNA.

The regular double-stranded RNA structures consist of canonical Watson-Crick (WC) base pairs, as shown in Figure 1.2.4. Their geometry allows any sequence of base pairs to fit into the nucleic acid helix without distortion; the distances between C1’ atoms of sugars on opposite strands is essentially the same for A-U and G-C base pairs.

**Figure 1.2.4:** Watson-Crick base pairs A-U (left) and G-C (right) are responsible for the duplexes in RNA. WC base pairs are stabilised due to hydrogen bonding, however, a GC base pair contains three hydrogen bonds compared to the AU base pair with two hydrogen bonds and so is the more stable base pair out of the two.

Many other hydrogen bonded base pairs are possible if the double-stranded helix is not required. There are many possibilities of base pairs formed by two hydrogen bonds (Figure 1.2.5). Amongst these are reverse WC, Hoogsteen, reverse Hoogsteen and Wobble mispairs.
Figure 1.2.5: Some examples of common base pairs which can be formed in RNA that contain at least two hydrogen bonds.

The complementary (WC) base pairs form duplex regions of the strand, while causing the non-complementary sections of the RNA to form loops, bulges or mismatches and thus possibly form one of the structures as shown in Figure 1.2.2. Figure 1.2.6 displays the hydrogen bonds between the complementary bases to form the double strand.
Figure 1.2.6: A double stranded RNA secondary structure comprising of the WC base pairs (A-U and G-C) and the phosphate groups running along the outer side of the structure. The strands run in the antiparallel (5’→3’ direction).

The ribose-phosphate groups form a backbone, running on the outer side of the secondary structure. Each phosphodiester group has one hydroxyl group which easily deprotonates at pH 7 resulting in the phosphate backbone to be negatively charged. The almost planar bases stack on top of one another inside the double helix.
The conformational variability of RNA has important implications. It allows RNA to adopt many diverse roles in the cell as that compared to DNA. One important aspect of RNA is the regulatory role it undertakes. It inhibits translation and consequently prevents peptide bond formation.

1.2.2 The Ribosome and protein synthesis

There are tens of thousands of proteins in the body and they have different forms and functions ranging from oxygen carrying haemoglobin, hormones such as insulin, collagen forming the skin and enzymes used to break down sugars. RNA plays the central role in transcribing genetic information from the DNA so that it can be translated into the structure of proteins, in a process known as the central dogma of molecular biology. Hence, ribosomes are crucial to life and accordingly they are a major target for antibiotics.

The ribosome is a macromolecular machine composed of RNA and ribosomal proteins and is the site of protein synthesis. In 2000, the tertiary structure of the *H. marismortui* 50S ribosomal subunit was solved which revealed that the key catalytic sites of the ribosome are composed of RNA, with proteins holding no major structural role.7

The prokaryotic ribosome has a molecular weight of approximately 2.5MDa and consists of two subunits; the large ribosomal subunit (50S) and the small ribosomal subunit (30S) as shown in Figure 1.2.7.
Figure 1.2.7: Diagrammatic representation of the ribosome comprised of the large ribosomal subunit (50S) and the small ribosomal subunit (30S) showing the important sites of protein synthesis: A (aminoacyl/acceptor), P (peptidyl) and E (exit). The initiator tRNA (formylmethionine-tRNA) occupies the peptidyl site and the messenger RNA (mRNA) containing the transcribed genetic information from a DNA template.

The 50S large ribosomal subunit contains two RNA components; 23S rRNA (2904 nt) and 5S rRNA (120 nt) and 34 ribosomal proteins. The peptidyl transferase centre (PTC) which is the site of protein biosynthesis resides in the 23S rRNA. The PTC contains both the aminoacyl (A-site) and peptidyl (P-site) sites for the peptide bond formation.

There are three types of RNA which are important for protein synthesis, messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA). Messenger RNA is the template for protein synthesis and adopts a linear structure which can vary in size. The template contains codons which represent a particular amino acid. A start codon indicates where to start the protein synthesis and a stop codon to indicate the termination stage.

Transfer RNA is a globular tertiary structure resembling a cloverleaf; it contains an amino acid attached to one end and an anticodon. It reads the mRNA template and acts as an ‘adaptor’ by attaching the anticodon to the complementary codon for the peptide bond formation.
Ribosomal RNA is the largest in size amongst the three types of RNA and the most abundant and conserved. It has structural roles within the ribosome but importantly it catalyses the protein synthesis reaction with the aid of protein factors. These are initiation factors (IF-1, IF-2, and IF-3), elongation factors (EF-Tu, EF-Ts and EF-G) and release factors (RF-1/2 and RF-3). Protein synthesis has four stages: Initiation, Elongation, Translocation and Termination and the mechanism is shown in Figure 1.2.8.

**Figure 1.2.8:** Scheme for protein synthesis in the ribosome. Peptide bond formation is initiated by the initiator tRNA (fMet-tRNA) occupying the P-site. The next step is an aminoacyl-tRNA which contains the correct anticodon to occupy the A-site and this step requires EF-Tu and GTP. A peptide bond is formed during the translocation step, with the fMet transferring from its tRNA to the aminoacyl-tRNA in the A-site. The mRNA shifts one codon’s length towards P-site and a new aminoacyl-tRNA can enter the A-site. The process continues until the stop codon of the mRNA is reached and the peptide is then released by the aid of RF-1/2 and RF-3 through the nascent peptide exit tunnel.
Protein synthesis is initiated with the initiator tRNA (fMet-tRNA, formylmethionine-tRNA) occupying the P-site of the ribosome and this process is facilitated by IF-1 and IF-2. The elongation step begins with the binding of the aminoacyl-tRNA to the A-site and requires GTP (for energy) and EF-Tu. To form the peptide bond, the fMet is transferred from its tRNA in the P-site to the aminoacyl-tRNA in the A-site, resulting in the deacylated tRNA transferring to the E-site and exiting the ribosome. This step forms a two-amino acid unit called a dipeptide (dipeptidyl-tRNA) in the A-site.

Next is the translocation step, which requires EF-G and GTP energy. The mRNA moves one codon’s length towards the P-site, resulting in the new dipeptidyl-tRNA moving from the A-site to the P-site allowing a new aminoacyl-tRNA to enter the A-site and continue the elongation step. The process is repeated to add another amino acid and continues until the stop codon on the mRNA is reached. The polypeptide is released with the aid of RF-1/2 and RF-3, from the ribosome in the termination step.17,19

Some Peptidyl transferase inhibitor antibiotics are able to interrupt peptide bond formation as their structures resemble that of the aminoacyl-tRNAs and so can be accepted in the A-site of the ribosome. For example, the Puromycin antibiotic’s chemical structure resembles that of tyrosyl-tRNA, thus a peptide bond between the peptide in the P-site and puromycin in the A-site is formed, yielding a peptidyl-puromycin. Puromycin has no association with the mRNA and cannot be translocated to the P-site. Therefore, the peptidyl-puromycin dissociates from the ribosome, aborting translation prematurely.19
1.2.3 23S rRNA and its conserved motifs

RNA is responsible for most critical functions of the ribosome; including the decoding of genetic information and catalysis of peptide bond formation. Hence, it is not surprising that ribosome-targeted antibiotics interact primarily with rRNA. Although the molecular size of the ribosome is approximately $10^4$ times larger than an antibiotic, thus containing numerous possible binding sites, only a few sites in the large ribosomal subunit (50S) are targeted by known antibiotics.

The 23S rRNA which resides in the 50S subunit contains six large domains that interlock with 34 proteins and the 5S rRNA to form an enormous structure which is approximately 250Å across (Figure 1.2.8). There are three main sites which have been identified in the structural domains. The domain II region is where the thiostrepton antibiotic inhibits the GTP-hydrolysis events associated with the binding of elongation factors (discussed in section 1.5.2). Domain V highlighted in red in the Figure below contains the peptidyl transferase centre which is the site of protein synthesis. The antibiotic inhibitors act at the central loop region to prevent peptide bond formation. The stem loop structure in domain VI is the site where α-sarcin, ricin and other inhibitors covalently modify the rRNA and consequently inhibit elongation factor-dependent reactions. These three sites (domains II, V and VI) participate in aminoacyl-tRNA binding, peptide elongation and translocation and therefore are functionally interrelated.
Figure 1.2.9: Schematic secondary structure of 23S rRNA endowed with the six domains (I to VI) and the 101 helices are numbered sequentially (1-101).\textsuperscript{20} The catalytically important \textit{peptidyl transferase} centre is situated in domain V (Dom V, red).
1.3 PT inhibitor antibiotics

1.3.1 Antibiotic resistance

Antibiotics can cure various infectious diseases such as pneumonia, whooping cough and meningitis by blocking the function of bacterial ribosomes. These antibiotics are produced by soil microorganisms which in order to survive, evolve antibacterial compounds to fend off predators trying to attack them. They selectively interfere with other microorganisms by terminating protein synthesis, the formation of the cell walls and/or directly interfering with bacterial DNA or RNA. Humans have discovered these natural products and converted them into antibiotic drugs.

Bacteria have lived on Earth for billions of years and during this lengthy time, they have encountered a wide range of naturally occurring antibiotics. To survive and protect their niche in the environment, they have developed antibiotic resistance mechanisms. As a result, it is no surprise that they have become resistant to most of the natural antibacterial agents discovered and developed over the past 50 years. Resistance occurs when the antibiotic loses its efficiency or cannot bind due to random mutation. Resistance can occur by an alteration in the antibiotic, an alteration in the target molecule or the cell boundary appearing less permeable to the antibiotic. Any alteration to the structure of the antibiotic can cause reduction in binding affinity between the antibiotic and the target molecule. Alteration in the target molecule can also cause resistance consequently reducing the antibiotic’s affinity. Small molecules can form an impermeable layer on the bacterial cell’s outer membrane. This antibiotic resistance requires the formation of diffusion channels to allow nutrients to pass into the cell.\textsuperscript{21} Hybrid antibiotics provide the advantage of introducing further species-specific anchors thus increasing selectivity and providing alternative interactions and in doing so reducing the rate of resistance.

*Peptidyl transferase* antibiotics target rRNA since it is less liable to develop antibiotic resistance by mutation. The rRNA is encoded in multiple gene copies and mutation usually occurs only in one gene copy, therefore, only a small area of the ribosome carries the mutant rRNA while the majority of the ribosome will contain wild-type rRNA.
transcribed from non mutated genes. Since most antibiotic resistance mutations in rRNA are recessive, the cell remains sensitive to the drug. Consequently, in order to become resistant to an antibiotic, the cell must simultaneously acquire the same mutation in several gene copies. Therefore, ribosomal antibiotics are highly efficient weapons since it is generally very hard to find defence against them. Chloramphenicol was the first antibiotic to be identified as inhibiting protein synthesis and its chemical structure is displayed in Figure 1.3.1.

![Chemical structure of chloramphenicol](image)

*Figure 1.3.1*: The chemical structure of the chloramphenicol antibiotic consists of an aromatic ring (blue) and an amide group (red).

*E.coli* cells were exposed to chloramphenicol and examined by electrophoresis and ultracentrifugation analysis against control *E.coli* cell extracts (Figure 1.3.2). Immediately after introducing chloramphenicol to the bacteria, protein synthesis terminated. In Figure 1.3.2a, the addition of the antibiotic yielded chloramphenicol inhibited ribosome (3) and can be clearly observed when measured with the control cells. The 50S large ribosomal subunit of the ribosome is completely inhibited by the antibiotic and barely visible in Figure 1.3.2b.
Figure 1.3.2: (a) Electrophoretic pattern of *E. coli* 70S ribosome cells without and with exposure to chloramphenicol: (1) 5S, (2) 70S, (3) chloramphenicol-inhibited ribosome and (4) boundary. (b) Ultracentrifuge pattern of control and antibiotic inhibited cells: (1) 5S, (2) 30S, (3) 50S, (4) 15S.\textsuperscript{22}

### 1.3.2 Amicetin

Amicetin is part of a family of five nucleoside antibiotics (Plicacetin, Bamicetin, Gougerotin and Blasticidin S). They are isolated from the microorganism *Streptomyces vinaceus drappus* cultures from soil samples\textsuperscript{23} and are universal drugs acting on all three evolutionary domains of life. Several common structural features arise in the five antibiotics: a cytosine moiety, at least one saccharide ring or a pseudo-saccharide ring and a methylated amino acid residue (Figure 1.3.3).

Figure 1.3.3: The chemical structure of amicetin antibiotic with the common structural features observed with the amicetin family of antibiotics, saccharide rings (orange), cytosine moiety (blue) and the amino acid residue, α-methyl-serine moiety (green).
Due to the structural similarities, these five antibiotics share related inhibitory activity, this is particularly true for amicetin and bamicetin as their chemical structures are comparable. The antibiotics act on the ribosomal *peptidyl transferase* centre (PTC) by inhibiting *peptidyl transferase* activity and therefore are capable of interfering with the binding of substrates to the PT centre.\(^{24}\)

*Halobacterium halobium* (*H. hal*) cells were exposed for 1.5 months to sub-inhibiting concentrations of amicetin antibiotic to determine the binding site of the drug. RNA sequencing of the individual colonies revealed drug resistance occurred by a U2457→C mutation (Figure 1.3.4). The U2457 is a universally conserved nucleotide and is labelled U26 on the RNA motifs displayed in Figure 1.5.2, (see page 56).

![Figure 1.3.4:](image)

**Figure 1.3.4:** (A) Graph displaying the amicetin-resistant mutant cells (Ami\(^ {r} \)) being less susceptible to the antibiotic in comparison to the wild-type (WT) cells. (B) RNA sequencing revealed drug resistance occurred by a single nucleotide change (U2457→C) for amicetin resistant cells.\(^ {25}\)

Further investigation was carried out to see whether the single mutant U2457C is solely responsible for drug resistance. Three possible mutations were engineered by site-directed mutagenesis at the nucleotide position 2457. The results revealed only the U2457C mutant was stable and its resistance was similar to the natural spontaneous mutant. The U2457A mutation grew slower than U2457C while the U2457G mutation was very sick and did not maintain stability.\(^ {25}\)
The U2457C mutation is predicted to be close to the binding site of the amicetin drug. Although the antibiotic does not produce chemical footprints, in *E. coli* ribosomes, amicetin-dependent protection of G2096 (G8, Figure 1.5.2) is observed and the A2458 (A27, Figure 1.5.2) is prevented from dimethylsulphate (DMS) attack in the 50S large ribosomal subunit of *T. thermophilus*.

### 1.3.3 Bamicetin

Bamicetin inhibits the growth of gram negative and gram positive bacteria including *Bacillus subtilis* (*B. subtilis*), *E. coli*, *Mycobacterium tuberculosis*, and *S. aureus*.

Bamicetin and amicetin possess very similar chemical structures (the difference being amicetin replaces the hydrogen with a methyl group on the amine moiety on the first saccharide ring), but the former possesses greater microbiological activity compared to amicetin. The most important difference is that bamicetin is the stronger PT inhibitor of the two antibiotics. Figure 1.3.5 below shows the *peptidyl transferase* inhibitor activity of amicetin, bamicetin and plicacetin. Approximately 20µM of bamicetin can be observed to reduce the PT activity of the *B.subtilis* 70S ribosome to 10%, triple the concentration of amicetin is required to achieve the same effect.

![Figure 1.3.5: The effects of plicacetin (blue), amicetin (red) and bamicetin (green) on the PT activity of *B.subtilis* 70S ribosome (0.55mM). Approximately 20µM of bamicetin is required to reduce the PT activity to 10% (dashed green lines), in contrast to ~ 60µM of amicetin (dashed red lines) illustrating that the bamicetin antibiotic to be a stronger PT inhibitor.](image)
1.3.4 Sparsomycin

The antibiotic sparsomycin is an inhibitor of protein synthesis, preventing peptide transfer by interfering with the PTC in the ribosome. The antibiotic was isolated in 1962 from the fermentation broths of the microorganism *Streptomyces sparsogene* \(^{27}\) and its structure was first reported in 1970.\(^{28}\) Although the PTC is a target for a considerable number of antibiotics which inhibit peptide bond formation (puromycin, chloramphenicol, virginiamycin M and blasticidin S), sparsomycin does not share a common binding site with any of them.

![Figure 1.3.6: The chemical structure of sparsomycin containing a methylated uracil ring (blue) and a sulphur ‘rich’ tail (red). It contains two chiral centres with the ScRs configurations.](image)

Sparsomycin is one of very few antibiotics (amicetin, bamicetin, blasticidin S, and gougerotin) which have universal PT inhibitor activity indicating the existence of a highly conserved target for the antibiotic action in the PTC of all ribosomes. Assays were carried out to observe the peptide bond formation using N-acetyl-aminoacylpentanucleotide and puromycin antibiotic as donor and acceptor substrates, respectively, in the *E.coli* 50S large ribosomal subunit.\(^{29}\) The test proved sensitivity to sparsomycin and a proof of the antibiotic interaction site at the PTC of a bacterial ribosome. Later, inhibitory results were observed from ribosomes of yeast, wheat germ and mammalian cells.

Sparsomycin has attracted considerable interest because of its structural characteristics and biological activity. The antibiotic is composed of a modified uracil ring, a \textit{trans}-olefin bond, a chiral carbon atom (S configuration) and a chiral sulphur atom (R configuration)
and is displayed in Figure 1.3.6 The absolute configuration was determined by Circular Dichroism (CD) spectroscopic studies and X-ray crystallographic studies.\(^{30}\) It has a thioacetal moiety \(\text{RS(O)}\text{CH}_2\text{SMe}\) which is rarely encountered in natural products and is one of the antibiotic’s diverse functions correlated with its PT inhibitory power.

The drug was found to be active against KB human epidermoid carcinoma cells in tissue culture but lacked sufficient species-specificity for use in treatment of human infections. High eye toxicity (retinopathy) in man discontinued Phase I clinical studies, whereas in animals (mouse, rat, dog, monkey) organ damage was observed. Sensitivity towards tumours, antiviral activity, inhibitory activity to growing gram positive and gram negative bacteria and a number of fungi\(^{27}\) prompted additional studies to elucidate its mode of action.

Structure-activity relationship studies have shown that the activity of sparsomycin is dependent upon the configuration of the chiral carbon atom and the sulphoxide function.\(^{30}\) The increase in activity of the derivatives has stirred up new interest in this drug as an antitumour agent. Two chiral centres produce four potential stereoisomers, but only the \(\text{ScRs}\) isomer confers activity with the other three isomers practically inactive.

The antibiotic contains two important interaction sites with the ribosome, the uracil ring and the hydrophobic \(\text{S-CH}_3\) group which is linked by an aliphatic chain with a strict steric configuration that allows the alignment of these two interacting sites in the correct position.

Sparsomycin binds only to the intact ribosome and only in the presence of an N-blocked aminoacyl-tRNA bound to the P-site.\(^{31}\) It binds at the central circle of the PTC, making it competitive with the binding of tRNAs to either the A- or P-site.\(^{32}\) Therefore, it can affect any of the three stages of peptide bond formation; binding of donor substrate, binding of the acceptor substrate and the transfer reaction with the exact one as yet unknown.
The uracil ring probably forms hydrogen bonds with the rRNA and the S-CH₃ group may interact with a hydrophobic ribosomal domain. Increasing the lipophilicity in the sulphur tail region dramatically increases the inhibitory activity of the drug. At the A-site, the drug inhibits the binding of small size substrates (unacetylated fragment) and distorts the interaction of the large aminoacyl-tRNA with the ribosome.³³

At the P-site, the binding of both small and large substrates is stimulated by the antibiotic and is possibly an allosteric effect which induces conformational change.³³ This predicted allosteric effect illustrates the close interrelation between the A- and P-sites in the peptidyl transferase centre and reveals a mutual interaction between the two sites during antibiotic binding.

A large number of drug derivatives has been synthesised and tested to understand the interactions with the ribosome and characterise the drug binding site.³¹ The results reveal the antibiotic interacts with the ribosome by the uracil ring and sulphur tail with orientation of the lateral chain to be important. The trans geometry of the double bond is necessary with the cis isomer inactive. Removal of the uracil C(6) methyl group decreased activity by 2-3 fold and derivatives of the uracil ring and sulphur tail caused drug inactivation.

Increasing the hydrophobicity at the S-CH₃ end by substituting with a phenyl ring (S-Ph) increased the anti-tumour activity of the sparsomycin in vivo and in vitro³⁴ and binds to the hydrophobic pocket (recognition site for hydrophobic amino acids) of the PTC in the ribosome.

The site of action was investigated by inducing cross-link and primer extension with reverse transcriptase between sparsomycin and archaeal, bacterial and eukaryotic ribosomes complexed with P-site bound tRNA on irradiating with low energy UV₃⁶⁵ light in the E.coli 70S ribosome.³⁵ The interaction site was localised to the universally conserved nucleotide A2602 of the central circle. Cross linking of several sparsomycin derivatives with modifications near the sulfoxy group implicated the uracil ring to be interacting with the rRNA.
Chapter 1- Introduction

The A2602 base is universally conserved and explains the strong inhibitory power of sparsomycin in all three evolutionary domains. It is situated in the central circle of the PTC and proximal to the 3’ ends of both donor and acceptor tRNA substrates on the ribosome. It is predicted that sparsomycin inhibits the protein synthesis by blocking the 3’ terminus of the P-site bound tRNA at the A2602 position.\(^3\)

1.3.5 Anisomycin

Anisomycin (2-\(p\)-methoxyphenylmethyl-3-acetoxy-4-hydroxypyrrolidine) is isolated from fermentation broths of Streptomyces griseolus and Streptomyces roseochromogenes and inhibits peptide bond formation on eukaryotic ribosomes.\(^3\) Anisomycin is relatively small in size compared to most of the PT inhibitor antibiotics, however, the chemical structure is composed of various functional groups; a chiral pyrrolidine ring, an aromatic ring, an acetoxy group and a methoxy group (Figure 1.3.6).\(^3\)

![Figure 1.3.7: The chemical structure of anisomycin antibiotic hosting a pyrrolidine ring (red) and an aromatic group (green).](image)

Titrations carried out on the antibiotic\(^3\) determined a pK\(_a\) value of 7.9 on protonation of pyrrolidine, revealing the antibiotic has basic properties and thus confirming presence of
the amine group. Relative stereochemistry determined by chemical studies and the absolute stereochemistry (2S, 3R, 4R) was determined by comparison with L-tyrosine.\textsuperscript{37}

Anisomycin has been successful in clinical trials against pathogenic protozoa and fungi for treatment of amoebic dysentery and fungicides in bean mildew. Research on the antibiotic and several derivatives has progressed over the years as they have shown positive results in treating human tumourous cells. Several of the derivatives have been patented due to their improved stability and activity \textit{in vivo}.\textsuperscript{38}

\section*{1.4 Structures and binding sites of the PT inhibitor antibiotics}

\subsection*{1.4.1 Amicetin and Bamicetin}

A number of structural studies have been carried out on the Amicetin antibiotic; X-ray crystallography, NMR spectroscopy and purely computational techniques have been employed.

The first structural study of amicetin was a Dreiding stereo-model.\textsuperscript{39} An X-ray crystallography study of amicetin in its free state has been produced. The structure revealed a highly extended conformation with a number of intramolecular hydrogen bonds.\textsuperscript{40}

A partial aqueous NMR solution structure revealed an elongated, rigid structure, similar to the X-ray crystal structure.\textsuperscript{41} A molecular modelling study of amicetin (binding to a possible binding site) has also been carried out.\textsuperscript{42} The result was a folded conformation, with the molecule literally bent double about the amide group IV, held together by a series of intramolecular hydrogen bonds between the two ends of the structure. The most recent study was a complete NMR solution structure study.\textsuperscript{43} The antibiotic produced a folded conformation, mediated by a network of hydrogen bonds (Figure 1.4.1), different to the previous study.\textsuperscript{43}
Figure 1.4.1: The NMR structure of amicetin and the intramolecular hydrogen bonds which help to maintain the antibiotic in a folded conformation are shown by the green dashed lines. Unlike amicetin, there has not been a range of structural studies accomplished for bamicetin. The similarity in the chemical structure of the two antibiotics is likely to have influenced the opinion that the X-ray or NMR structures of the two antibiotics would also be very similar.

1.4.2 Sparsomycin

Although there is a single X-ray crystallography study on a precursor of sparsomycin to determine the configuration of the antibiotic, there are no X-ray or NMR structures of sparsomycin in the free state.

There exist two bound X-ray crystal structures which are displayed in Figure 1.4.3 of sparsomycin and the large ribosomal subunits (50S) of *H.marismortui* (at 2.8Å) and *D.radiodurans* (at 3.7Å). The structure of the sparsomycin bound to the 50S subunit of *H.marismortui* also revealed the importance of the universally conserved A2602 base. Two hydrophobic crevices were revealed to play a part in the binding of the antibiotic and
lie midway between the two crevices; A2602 (A2103 *H. marismortui*) varies in its conformation and thereby contacts the antibiotic bound at either crevice.

The 23S *H. marismortui* rRNA contains two crevices which are formed by the imperfect stacking of pairs of adjacent bases, thus producing a wedge-shaped hydrophobic gap. The first crevice is at the PTC and is shown in Figure 1.4.2, between bases A2486 (A2451, *E. coli* numbering) and C2487 (C2452). These nucleotides interact with the side chains of the aminoacyl-tRNAs bound in the A-site.

![Diagram indicating elements of secondary structure in the peptidyl transferase centre of *H. marismortui* 23S rRNA. Nucleotides highlighted in red are the sparsomycin resistant mutants which are located in the central circle of domain V and the black triangles indicate the position of the two crevices (active site and exit tunnel).](image)

**Figure 1.4.2:** Diagram indicating elements of secondary structure in the *peptidyl transferase* centre of *H. marismortui* 23S rRNA. Nucleotides highlighted in red are the sparsomycin resistant mutants which are located in the central circle of domain V and the black triangles indicate the position of the two crevices (active site and exit tunnel).\(^\text{32}\)

The second crevice is at the entrance of the peptidyl exit tunnel, between nucleotides G2099 (A2058) and A2100 (A2059). Both crevices contain hydrophobic and aromatic interiors making them ideal for binding to the hydrophobic tail of sparsomycin. They are spatially very close (~13Å) and this may explain why the antibiotic is able to interact at both the A- and P-sites. The sparsomycin resistant mutants occur solely in the central circle of the PTC.
Chapter 1- Introduction

The antibiotic binds to the *H.marismortui* ribosome by forming hydrogen bonds with the phosphate group of P-site tRNA and the the N3 of the pseudouracil ring. It also hydrogen bonds with the phosphate group of U2620 (U2585) of 23S rRNA, producing a bifurcated system. The oxygen atom (O4) of the uracil ring is also involved in hydrogen bonding with the ribose 2’-OH of U2619 (U2584) as well as interacting with a magnesium ion which is bound to the phosphate oxygens of the P-site substrate. The methyl group of the uracil base contributes to binding by forming van der Waals contacts with the bases of the P-site substrate.

As discussed in section 1.3.4, modifications to the sparsomycin structure reduced its inhibitory prowess. In both of the bound X-ray crystal complexes, methylating the N3 and O4 atoms of uracil ring of sparsomycin which are involved in hydrogen bonding with rRNA residues resulted in a decrease in the binding affinity and removing its methyl group, subsequently reduced the affinity by three fold.

![Figure 1.4.3](image1.png)  
**Figure 1.4.3:** (a) The X-ray crystal structure of sparsomycin bound to the *H.marismortui* 50S ribosome. The antibiotic only binds in the presence of a P-site substrate. (b) The X-ray crystal structure of sparsomycin bound to *D.radiodurans* 50S ribosome.
1.4.3 Anisomycin

A three-dimensional X-ray diffraction analysis has been carried out on N-acetylbromoanisomycin (Figure 1.4.4) to determine the structure of anisomycin.\(^{37}\)

![Molecular structure of N-acetylbromoanisomycin](image)

**Figure 1.4.4:** The molecular structure of N-acetylbromoanisomycin. The X-ray crystallography studies on this molecule, aided in determining the correct chemical structure of anisomycin.\(^{37}\)

An X-ray crystal structure of anisomycin bound to the 50S large ribosomal subunit *H.marismortui* at 3.0 Å has been determined\(^{32}\) (Figure 1.4.5). The antibiotic inserts its hydrophobic \(p\)-methoxy-phenyl group into the A-site crevice and fully occupies the hydrophobic crevice which is the abode for the incoming amino acid side chains of the A-site bound aminoacyl-tRNAs and as a result anisomycin is able to interfere with the binding of A-site substrates. The evidence for anisomycin interacting at the A-site of the PTC is the resistance to the antibiotic seen on mutation of the nucleotides which stabilise the A-site hydrophobic crevice. The strongest resistance occurred for the mutation of base C2487U (C2452 *E.coli* numbering) which eliminates the hydrogen bonding to the nitrogen group of the pyrrolidine ring.
Figure 1.4.5: X-ray crystal structure of anisomycin bound to *H. marismortui* 23S rRNA.\(^{32}\)

Figure 1.4.6 shows the anisomycin antibiotic binding to residues in the 50S large ribosomal subunit of *H. marismortui*. Evidence of binding interactions was demonstrated by conformational changes to the A2486 (A2451) and A2488 (A2453) nucleotides.\(^{46}\)

Figure 1.4.6: The location of the exit tunnel of the *H.marismortui* 50S ribosomal subunit, where anisomycin binds and interacts with the surrounding nucleotides. The antibiotic inserts the *p*-methoxy-phenyl group in between A2451 and A2452 residues, altering their conformations. The conformations of the nucleotides in the absence of anisomycin (grey) and in the presence of anisomycin are shown (yellow).\(^{32}\)
1.5 Structures of the RNAs

1.5.1 B. subtilis 27mer RNA motif

*B. subtilis* is a Gram-positive bacterium commonly found in soil and has the ability to form a tough, protective endospore, allowing the organism to tolerate extreme environmental conditions. It has proven highly amenable to genetic manipulation, and has therefore become widely adopted as a model organism for laboratory studies. The inhibitory power of bamicetin is very high when interacted with the *B. subtilis* 70S ribosomes\(^2\) and this is an important reason to study this RNA.

Figure 1.5.1 displays the PT centre which is located in domain V of the *B. subtilis* 23S rRNA. The 27mer RNA motif which is to be studied in this research project is situated on the stem region of the central circle of domain V. This 27mer RNA is highly stabilised due to an abundance of stable canonical G-C base pairs.

![Secondary structure of the peptidyl transferase centre](image)

**Figure 1.5.1**: The secondary structure of the *peptidyl transferase* centre located in the domain V region of *B. subtilis* 23S rRNA with the 27mer RNA strand indicated by the curly bracket.\(^4\) Universally conserved nucleotides (> 98%) in archaea, bacteria and eukarya ribosomes are indicated on the diagram by the black filled circles.\(^4\)
The secondary structures of *B.subtilis*, *E.coli*, *H.hal* and *H.marismortui* RNA motifs are displayed in Figure 1.5.2 to observe the similarities between all five of the motifs. The RNA motifs are part of the same helical section of the PTC centre in domain V of their respective 23S rRNAs. The NMR structures of the 29mer motifs of *E.coli* and *H.hal* have been determined in the research group; hence, it is important that the NMR structure of the *B.subtilis* 27mer RNA is established for further understanding of the ribosome-antibiotic affinity. As can be observed in the Figure below, the 27mer RNA motif of *B.subtilis* contains the greater number of canonical WC G-C base pairs signifying that it is probably the most stable RNA motif amongst the four RNAs.

![Figure 1.5.2: Homologous sequences of the 23S rRNAs: (A) 35mer and (B) H.hal 29mer RNA; (C) E.coli 29mer RNA; and (D) H. marismortui 27mer RNA compared to the (E) 27mer motif of B. subtilis RNA illustrated with the internationally recognised number of bonds for canonical WC GC and AU base pairs. The numbering is kept consistent with the H.hal 35mer RNA in conformity with previous work and to allow easier identification of bases on any of the four motifs.](image-url)
1.5.2 GTPase centre of *H. hal* 25mer RNA

The GTP hydrolysis centre (GTPase centre) is an important functional site on the 50S large ribosomal subunit of *H. hal* 23S rRNA, participating in hydrolysis reactions involving ribosomal elongation factors, EF-Tu and EF-G. It is located on a double hairpin structure within domain II of 23S rRNA and contains two thiostrepton mutation sites A1159G and A1187G. The 25mer motif of the GTPase centre that is to be studied contains the thiostrepton-resistant mutant at the nucleotide position A1159G (Figure 1.5.3).

![Figure 1.5.3: The secondary structure of the 25mer RNA motif residing in domain II of *H. hal* 23S rRNA. The thiostrepton resistant mutant A1159G is highlighted by the red box.](image)

Thiostrepton is a cyclic peptide antibiotic (Figure 1.5.4) and has shown multiple inhibitory mechanisms *in vivo* and *in vitro*, suggesting that the GTPase centre undergoes functionally important conformational transitions, one or more of which is blocked by thiostrepton.
Several separate studies indicate that thiostrepton binds strongly and primarily to rRNA in this region. The important observations made are:

- The thiostrepton producing microorganism *Streptomyces azureus* (*S. azureus*), methylates the A1159 (A1067 *S. azureus*) residue to prevent the drug binding to its own ribosome.⁵²
- rRNA footprinting results revealed drug protection effects around bases A1159 (A1067) and A1187 (A1095) of *S. azureus*.
- Mutation of these nucleotides *in vivo* *H. hal*, and *in vitro* in *E. coli*, produced high drug resistance levels to the antibiotic.
- Thiostrepton exhibits a high binding affinity for 23S rRNA of archea, bacteria and eukarya ribosomes.

Based on the above evidence for drug-binding, it can be established that the functional transitions in the ribosomal GTP hydrolysis centre occur at an rRNA level.⁵⁰,⁵³
1.5.3 NMR structures of RNA in databases

The determined NMR structures can be deposited into three web accessible databases and the majority of journals make this a requirement. These are the Protein Data Bank (RCSB PDB),\textsuperscript{54} Nucleic Acid Database (NDB)\textsuperscript{55} and Biological Magnetic Resonance Bank (BMRB).\textsuperscript{56} The RCSB PDB is the largest of the three databases. Founded in 1976, its archive contains thousands of entries of X-ray and NMR structures (Table 1.5.1) of nucleic acids and protein structures, which have exponentially grown over the years (Figure 1.5.5). NDB contains only 3D structural information of nucleic acids and nucleic acids complexed with proteins or small molecules.

For submission of an NMR structure, the Cartesian coordinates and constraint data files are processed and assigned a unique Identification (ID) number by RCSB PDB. The database contains a validation server (ADIT-NMR) which pre-checks, validates and deposits the structure. It also includes details of the experiments, sample conditions (pH, temperature and ionic strength) and software used. NMR spectral data including chemical shifts, scalar $J$-coupling constants and relaxation parameters are processed and archived by the BMRB database.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1_5_5.png}
\caption{Figure 1.5.5: Graph displaying the rapid growth of structures deposited into the PDB database since its formation in 1976.\textsuperscript{54}}
\end{figure}
Table 1.5.1: A summary of the released entries in the RCSB PDB database as of 16th August 2011.

<table>
<thead>
<tr>
<th>Methods</th>
<th>X-ray</th>
<th>NMR</th>
<th>Electron Microscopy</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecule type</td>
<td>Proteins</td>
<td>Nucleic Acids (NA)</td>
<td>Protein/NA complexes</td>
<td>Other</td>
<td>Total</td>
</tr>
<tr>
<td>Proteins</td>
<td>61365</td>
<td>1304</td>
<td>2981</td>
<td>3</td>
<td>65653</td>
</tr>
<tr>
<td>NMR</td>
<td>7882</td>
<td>955</td>
<td>178</td>
<td>7</td>
<td>9022</td>
</tr>
<tr>
<td>Electon Microscopy</td>
<td>254</td>
<td>22</td>
<td>95</td>
<td>0</td>
<td>371</td>
</tr>
<tr>
<td>Other</td>
<td>173</td>
<td>7</td>
<td>6</td>
<td>14</td>
<td>200</td>
</tr>
<tr>
<td>Total</td>
<td>69674</td>
<td>2288</td>
<td>3260</td>
<td>24</td>
<td>75246</td>
</tr>
</tbody>
</table>

1.6 Cross-hypersensitivity

1.6.1 Cross-hypersensitivity phenomenon

Two thiostrepton resistant mutations (A1159G and A1187G) which inhibit the ribosomal GTPase centre in domain II were hypersensitive to the amicetin antibiotic which acts at the PTC in domain V of the H.hal 23S rRNA (Figure 1.6.1). This cross-hypersensitivity phenomenon was a fascinating result as both mutated sites are widely separated in the primary and secondary structures of the 23S rRNA (Figure 1.6.2).
The cross-hypersensitivity effects observed between the amicetin antibiotic and the thiostrepton mutated site A1159G (G11) and conversely with the thiostrepton antibiotic and the amicetin mutated site U2457C (U26). The mutated bases are labelled according to the *H. hal* 23S rRNA numbering and the simplified numbering adopted with the *H. hal* 25mer RNA and *B. subtilis* 27mer RNA motifs are labelled in brackets.

**Figure 1.6.1**: The cross-hypersensitivity effects observed between the amicetin antibiotic and the thiostrepton mutated site A1159G (G11) and conversely with the thiostrepton antibiotic and the amicetin mutated site U2457C (U26). The mutated bases are labelled according to the *H. hal* 23S rRNA numbering and the simplified numbering adopted with the *H. hal* 25mer RNA and *B. subtilis* 27mer RNA motifs are labelled in brackets.

### 1.6.2 The effect of thiostrepton and amicetin

The thiostrepton resistant mutants (A1159G and A1187G) displayed hypersensitivity to amicetin, and were more sensitive to the antibiotic than the wild type RNA. They also possessed affinity to two other PT antibiotics, anisomycin and chloramphenicol which act in the central circle of domain V. Similar to amicetin and thiostrepton, both antibiotics impair or inhibit the binding of aminoacyl-tRNAs at the A-site.

The sole amicetin resistant mutant residing in domain V also displayed sensitivity to the thiostrepton drug providing support for a functional and structural linkage between the thiostrepton site in domain II and the amicetin site in domain V and hence must contribute to a mutual site for the binding of aminoacyl-tRNA substrates.
In order to establish whether the effects observed between the two mutated sites were attributed to long range effects, the thiostrepton binding region from the amicetin mutant and the amicetin binding region on the thiostrepton resistant ribosome were probed with nucleotide specific chemicals. Minor local changes at the sites were observed but no long range effect was identifiable.

So far the halobacteria *H. hal* is the only organism for which it has been possible to identify spontaneous rRNA mutations conferring resistance to thiostrepton antibiotics. The A1159 base has been strongly implicated in thiostrepton binding, since the microorganism *S. azureus* which produces the antibiotic protects its own ribosomes by methylating this
nucleotide. The hypersensitivity effects observed between the thioestrepton resistant mutant and the PT antibiotics, amicetin, anisomycin and chloramphenicol is fascinating. The mutated sites lie in separate domains (II and V) and are widely separated in the primary structure (>1400 nucleotides). However, importantly their degree of separation in the tertiary structures of the ribosome is uncertain. Also, the antibiotics have different modes of action, but they all interfere with binding of aminoacyl-tRNAs and therefore have common interference with the A-site.
1.7 Introduction to NMR\textsuperscript{59, 60}

The magnetic properties of a nucleus form the basis of nuclear magnetic resonance spectroscopy (NMR). The nuclei of all atoms can be characterised by their nuclear spin quantum number (I) which must have values > 0 to exhibit nuclear magnetic resonance. The spinning nucleus possesses angular momentum (P) and a magnetic moment (μ) which are linked by the gyromagnetic ratio (γ), and is dependent on the type of nuclei (Equation 1.7.1).

\[ \mu = \gamma P \quad \text{Equation 1.7.1} \]

When the nucleus is placed in an external magnetic field (B\textsubscript{0}), the field imposes a torque on the magnetic moment which traces a circular path about B\textsubscript{0} as well as spinning on its own axis (Figure 1.7.1). This precession is known as Larmor precession. The rate of precession is dependent on the magnetic field strength (B\textsubscript{0}) and the gyromagnetic ratio (γ). The individual microscopic magnetic moments of the molecule align relative to the field with quantised orientations according to the 2I + 1 rule. For the \textsuperscript{1}H and \textsuperscript{13}C nuclei, which have I = \( \frac{1}{2} \), this results in 2 energy states; α, β (Figure 1.7.2). The nuclei can exist in two possible orientations, one being aligned with the applied field (B\textsubscript{0}) and of lower energy (α) and one against the field and of higher energy (β). According to the Boltzmann thermodynamic distribution, there is a slight excess of spins in the lower state, producing bulk magnetisation (M\textsubscript{z}) in the z-direction.
Figure 1.7.1: The precessing nuclei will have a longitudinal component of the magnetisation along the $z$-direction and an orthogonal component in the $x$-$y$ plane. The applied static magnetic field is shown by $B_0$.

Figure 1.7.2: Nuclei with spin $I=\frac{1}{2}$ produce two quantised energy levels ($\pm \frac{1}{2}$, $\alpha$ and $-\frac{1}{2}$, $\beta$) when placed in the applied magnetic field ($B_0$). The nuclei in the $\alpha$ state contain lower energy and align with the $B_0$, whereas those in the $\beta$ state are higher in energy and align perpendicular to $B_0$.

NMR occurs by applying energy to the nucleus in the form of electromagnetic (EM) radiation. The frequency of the EM radiation must match the resonant frequency (Larmor frequency) of the nucleus, at which it absorbs energy and the nucleus is able to move from a lower state ($\alpha$) to a higher state ($\beta$) (Figure 1.7.3).
Figure 1.7.3: The EM radiation is applied in the form of a short RF pulse ($B_1$) along the x-axis. The energy is absorbed by the nuclei resulting in excitation of the spins from the lower to the upper state. In order for absorption to take place, the frequency of $B_1$ must be equivalent to the energy gap ($\Delta E$) between the two states.

To induce transitions between the energy levels, the $B_1$ field is transmitted via a coil along the x-axis. This rotates the bulk magnetisation away from the z-axis and onto the x-y plane. The RF pulse is then switched off and the magnetisation vector will start to freely precess back to the equilibrium state, +z-axis; as it does so, it induces a current in the coil which is recorded as the NMR signal.

1.7.1 Chemical shift and $J$-coupling
Chemical shift values and scalar $J$-coupling constants are two important experimental parameters which can be extracted from the NMR spectrum to achieve the assignment of the molecule in question. When the molecule is placed in a magnetic field ($B_0$), the electrons within the molecule shield the nuclei from the external field (Figure 1.7.4). The electrons are able to oppose the magnetic field as their circulation around the nucleus produces an electric current with its own secondary magnetic field (B).
Chapter 1 - Introduction

Figure 1.7.4: The electrons (e⁻) in the molecule shield the nucleus from the applied field B₀ by opposing it.

As a result, the magnetic field felt by the nucleus (B_{eff}) is less than the applied field (B₀) and takes into consideration the shielding effect produced by the electrons (shielding constant) (2.7.1).

\[ B_{\text{eff}} = B_0 - \sigma B \]

\[ \sigma = \text{shielding constant} \]

The resonance frequency is proportional to the magnetic field at the nucleus and a shift in the resonance frequency is thus known as the Chemical shift (\( \delta \)). Depending on the electron density experienced by a chemically non-equivalent nucleus, it will be shielded to varying extents and therefore produce individual resonance signals in the NMR spectrum.

The chemical shift (2.7.2) is defined as the resonant frequency (\( v_{\text{ref}} \)) of the nuclei divided by the applied field and is a molecular quantity, dependent on the sample conditions and not on the spectrometer frequency (\( v_0 \)).

\[ \delta \text{ (ppm)} = \frac{V - V_{\text{ref}}}{V_0} \times 10^6 \]
The internationally recognised reference sample for $^1$H and $^{13}$C NMR experiments is tetramethylsilane (TMS) with a chemical shift value of 0 ppm. The chemical shift value of a nucleus is affected by the circulating electrons of neighbouring substituents which influence the magnetic field at the nucleus. The $^1$H proton nucleus contains only an ‘s’ electron which has a spherically symmetrical charge distribution and results in an upfield shift of the nucleus. Every molecule contains s electrons and so this is a universal effect. The $^{13}$C, $^{15}$N and $^{31}$P atoms contain p-orbital electron distributions which are non spherical resulting in a greater spread of electron densities and thus producing a larger range of chemical shifts which are shifted to low field. This is a deshielding effect known as the paramagnetic shift. $^1$H NMR spectra contain only a diamagnetic effect and so have a smaller chemical shift range ~15ppm. The $^{13}$C NMR spectra have a much larger chemical shift range ~250ppm owing to a combination of diamagnetic and paramagnetic shift.

Scalar $J$-coupling is a phenomenon indicating which nuclei are close to each other along the bonding network. The magnetisation dipoles of neighbouring nuclei interact with other through the chemical bonds splitting the NMR signal into a multiplet. Taking an example of two spins $H_A$ and $H_B$ which are covalently attached, the resonance signal of $H_A$ will be split into two producing a doublet (Figure 1.7.5). The magnitude of the splitting is represented as the scalar coupling constant, $J_{A,B}$. 
Figure 1.7.5: A doublet peak arising for nuclear spin A due to the neighbouring nuclear spin B. Vertical arrow (up) represents the spin B in the $\alpha$ state (deshielded) and appears lower field and the vertical arrow (down) is due to spin B in the $\beta$ state and therefore shielded and appears higher field.

The spins are in a bulk sample, with the number of A spins in $\alpha$ ($\uparrow$) and $\beta$ ($\downarrow$) state nearly equal according to the Boltzmann distribution. So the two lines of the doublet appear with identical intensity in the NMR spectrum (Figure 1.7.5). The splitting of the signal is due to the nuclear magnetic moments and so $J_{A,B}$ does not depend on the magnetic field density ($B_0$) and is therefore expressed in the frequency unit Hertz (Hz).
1.7.2 $T_1$ and $T_2$ relaxation

Immediately after pulse excitation of the nuclear spins, the bulk magnetisation vector moves away from the equilibrium $+z$-axis corresponding to a change in the spin populations. In order to re-establish the equilibrium populations, the spins must lose energy and they do this by two ways. The recovery of the $z$-component of the magnetisation along the $+z$-axis is known as longitudinal relaxation ($T_1$) and the $x$-$y$ (transverse) magnetisation component along the $x$-$y$ plane is the transverse relaxation ($T_2$). During longitudinal relaxation the energy lost by the spins is transferred to the surroundings in the form of heat (change in temperature is very small). The magnetisation gradually shrinks in the $x$-$y$ plane and fully recovers along the $+z$-axis at a rate dependent on the nuclei (Figure 1.7.6a).

![Diagram of longitudinal and transverse relaxation](image)

**Figure 1.7.6**: a. Longitudinal relaxation ($T_1$). b. Transverse relaxation ($T_2$) following the application of a 90° rotating RF pulse to the bulk magnetisation $M_z$.

After the 90° RF pulse is applied, the transverse magnetisation is on resonance along the $y$-axis. The bulk magnetisation vector results from the addition of many microscopic vectors from the individual nuclei that possess phase coherence. The magnetic field experienced by each spin in the sample is not the same due to the magnetic field inhomogeneity and results in a fanning out of the individual magnetisation vectors in the $x$-$y$ plane (Figure 1.7.6b). Some spins will experience a slightly greater local field causing them to have higher frequency than the larmor frequency and move in front of spins with
smaller frequencies. The spread of the individual magnetisation vectors leads to no net magnetisation causing decay of the NMR signal by transverse relaxation $T_2$.

The width of the NMR resonance is inversely proportional to $T_2$ and can be used to calculate the magnitude of the transverse relaxation (Figure 1.7.7). $T_2$ can never be longer than $T_1$, as for longitudinal magnetisation to be restored, it must all return to the $+z$-axis leaving none in the $x$-$y$ plane.

![Diagram of NMR resonance with $T_2$ values](image)

**Figure 1.7.7:** (a). A short $T_2$ results in a greater spread in the frequency dimension producing a broader line while in (b). A long $T_2$ will produce a sharper peak due to less spread of the frequencies.

### 1.7.3 Nuclear Overhauser Effect (NOE)

The NOE phenomenon occurs from cross-relaxation between spins that are spatially close. By irradiating one proton (S) with a strong RF pulse, the intensity of the neighbouring proton (I) is affected. To explain this effect, Figure 1.7.8 displays a homonuclear two spin system that are not scalar coupled with each other. The $^1$H spectrum of two inequivalent protons consists of two singlets at their respective chemical shifts (Figure 1.7.8a). While recording the spectrum, the S spins are saturated by exciting the nuclei with a strong RF field oscillating at the S resonance frequency. This destroys the NMR signal of S (Figure 1.7.8b) and affects the intensity of the I resonance. After excitation, the two spins can
return to equilibrium via two relaxation transitions, \( W_0 \) and \( W_2 \). Both involve the inversion of two spin populations I and S and therefore are known as cross-relaxation pathways. Cross-relaxation from the double quantum transition \( W_2 \), yields positive NOEs (Figure 1.7.8c) and negative NOEs arise from cross-relaxation by zero quantum transitions, \( W_0 \) (Figure 1.7.8d). Single quantum transitions \( W_1 \), also occur but they do not produce NOEs as they only involve inversion of one spin.

**Figure 1.7.8:** Schematic energy level diagrams and population differences for two spins I and S which share a dipolar coupling: (a) I and S spins at equilibrium and have excess population (\( \Delta \)), (b) after instantaneous saturation of the S spins, (c) after relaxation via \( W_2 \) processes and (d) after
relaxation via $W_0$ processes. Below each energy level diagram is the corresponding NMR spectrum.

Two factors affect the size of the NOE coupling commonly known as dipolar coupling. The first factor is the coupling is observed through space (distance dependent $\leq 5\text{Å}$) and so provides information about the three dimensional structure of the molecule. The second factor is the rotational correlation time ($\tau_c$) of the molecule, which is the time it takes the molecule to rotate one radian. The rate of rotation is affected by the size and shape of the molecule and so affects the NOE signal.

### 1.7.4 Rate Processes

NMR spectra of many molecules are temperature dependent making it possible to study their kinetics using NMR Spectroscopy. This is known as Dynamic NMR spectroscopy with the resonance line shape (line width and peak intensity) sensitive to the exchange processes. In an intramolecular equilibrium process, a molecule can interconvert between two chemically distinguishable states A and B (2.7.3), with the change in Gibbs energy ($\Delta G$) and rate of interconversion ($k$) defining the equilibrium.$^{61}$

![Diagram of A ↔ B](image)

2.7.3

The molecule in state A will have chemical shift $\nu_A$ and in state B it will have chemical shift $\nu_B$. Three exchange rate processes can occur: slow, intermediate and fast (Figure 1.7.9).$^{62}$ In the slow exchange process, the rate of interconversion of A and B is slow on the NMR timescale and the interconversion rates $k_A$ and $k_B$ are much smaller than the difference in chemical shifts ($\delta \nu$) of states A and B and so both peaks appear at separate chemical shift values. The intensity of each peak depends on the molar concentrations of the nuclei in states A and B. The intermediate exchange occurs when the rates $k_A$ and $k_B$ are approximately equal. This produces broadened NMR signals of both peaks due to partial averaging of the chemical shifts.
When the rate of interconversion is fast (fast exchange process), the rates $k_A$ and $k_B$ are much greater than $\delta\nu$ producing a single peak representing both states. The chemical shifts are the molar weighted averages of states A and B, thus producing a NMR signal at $\nu_{AB}$.

Intermolecular exchange reactions can also be studied by dynamic NMR. For example, the interactions between a ligand and nucleic acid or protein. The exchange rate of the ligand proton exchanging with the macromolecule can be monitored if the interconversion is slow on the NMR timescale.

**Figure 1.7.9:** Diagram illustrating the lineshapes in an NMR spectrum for slow, intermediate and fast exchange processes. The nucleus interconverts between states A↔B as a function of the average lifetimes ($\tau$).
1.8 Molecular modelling

1.8.1 Forcefields

Force fields implement Newtonian mechanics to model molecular systems. They are parameters for calculating the potential energy \( E \) of a system in a given conformation as a sum of individual energy terms and are applicable to both small and large biological systems.

Force fields ignore electronic motions and calculate the energy of the system only as a function of the nuclear positions. It is therefore used to perform calculations on systems containing significant number of atoms. The force fields can be categorised as intramolecular and intermolecular forces within the system and hence the energy may consist of covalent (bonded) and non covalent (non bonded) contributions (Figure 1.8.1 and Equation 1.8.1).\(^6\)

\[ E = E_{\text{bonded}} + E_{\text{non bonded}} \]

\( E_{\text{bonded}} \) includes the energy of covalent bonds, bond stretching, angle bending, and bond rotation. 

\( E_{\text{non bonded}} \) includes non covalent interactions such as electrostatic forces and van der Waals forces.

\[ E_{\text{bonded}} = \sum E_{\text{bond}} \]

\[ E_{\text{non bonded}} = \sum E_{\text{vdW}} + \sum E_{\text{elec}} \]

**Figure 1.8.1**: The potential energy of the system is calculated by a series of molecular mechanic force fields which not only include the energy of the bonded terms but the long range non bonded interactions.
Equation 1.8.1: Typical potential energy of a system is calculated by a set of force fields, including covalent and non covalent terms. $E_{\text{miscellaneous}}$ incorporates charges and NOE distance values.

Energetic penalties indicate deviation of bonds and angles away from their reference or equilibrium values and energy changes as the bonds are rotated. It is important that the force fields are transferrable to make modelling of related molecules possible. They are empirical entities and great consideration is taken when selecting the set of force fields for a calculation. A balance between accuracy and computational efficiency is required. As more powerful computers are developed, it will become possible to incorporate more sophisticated models.

Another feature involved in force fields is that of an atom type. The atom type not only includes the atomic number of an atom, but it contains information about its hybridisation state and sometimes the local environment. The force fields differentiate between sp$^3$ hybridisation carbons (tetrahedral, 109.5°), sp$^2$ hybridisation carbons (trigonal planar, 120°) and sp hybridisation carbons (linear 180°).

The atom types were manually placed for the structure calculations of the three antibiotics (bamicetin, sparsomycin and anisomcyin), while the structure determination program XPLOR-NIH$^{64}$ determined the atom types of the two RNA motifs ($B.\text{subtilis}$ 27mer RNA and $H.\text{hal}$ 25mer RNA).

Planar groups such as carbonyl or aromatic groups produce ‘improper’ dihedral angles with one of the atoms above or below the molecular plane and so an additional force field is required to ensure planarity. Three of the four atoms in the dihedral angle are in the plane; however, the branch atom (i.e. oxygen of the carbonyl group) becomes distorted (Figure 1.8.2). The distortion is measured as the height of the central atom above the plane formed by the other three atoms and the correct geometry is invoked.
**Figure 1.8.2**: An ‘improper’ dihedral angle which can arise from carbonyl and aromatic groups. A, B, C and D are the four atoms in the dihedral angle and the branch atom D (i.e. oxygen of a carbonyl group) can appear above or below the plane and therefore an additional force field is input to maintain planarity.

1.8.2 *Ab initio* calculations

Total energy simulations (*ab initio*) involve both the electronic and nuclear motions of the molecule. The electronic energy of the ion is calculated with a given fixed internuclear distance and the total energy is obtained by adding the nuclear-nuclear repulsion term. Three examples of *ab initio* calculations are Hartree-Fock\(^\text{65}\), Møller-Plesset\(^\text{66}\) and B3LYP.\(^\text{67}\) The Hartree-Fock method consists of Hamiltonian operators and the Schrödinger equation to calculate the kinetic and potential energy of the system. Exchange terms in addition to Coulombic interelectron interactions are included.

The Møller-Plesset (MP) method advances on the Hartree-Fock scheme by tackling the problem of electron correlation using the Rayleigh-Schrödinger perturbation theory. If the correlation is curtailed at the second order, the calculation is known as an MP2, if at the third it is MP3 and so on.

The B3LYP method is a hybrid functional technique; it is a combination of the Hartree-Fock method and Density functional method (DFT). DFT calculations simplify the Hartree-Fock scheme by considering electron density as opposed to numerous electron wave functions. The hybrid combination increases the accuracy in comparison to pure
DFT methods and reduces the calculation time of a Hartree-Fock calculation. It is a good method for generating correct conformations of energy minimised structures.

1.8.3 restrained Molecular Dynamics

Restrained molecular dynamics (rMD) is an important application in determining the X-ray and NMR structures of large biological molecules. The aim is to determine the conformation that best illustrates the experimental data.

In rMD, additional terms known as penalty functions are added to the potential energy function to penalise conformations that do not agree with the experimental data. It explores the conformational space using distance geometry in order to find conformations that not only have a low intrinsic energy but are consistent with the experimental data.

Harmonic restraint terms such as bond distance \( k(d-d_o)^2 \), where \( d \) is the distance between the atoms and \( d_o \) is the desired distance derived from the NMR spectrum, \( k \) is the force constant and determines how tightly the restraint should be applied. Torsional restraints are also incorporated in the calculation alongside the bond distance information. The atoms are prevented from coming too close to each other by the van der Waals terms included in the force fields. Distances are obtained from the NOESY peak intensity and placed with lower and upper error bounds (explained in section 2.4). Distances appearing outside the ranges are subjected to restraint using harmonic potentials with possibly different force constants and steepness.

1.8.4 Simulated annealing

Simulated annealing method is applied to effectively search the conformational space of the molecule to find the most stable structure. The technique simulates heating of the molecular system and then the temperature is slowly decreased to encourage the molecule to adopt a low energy stable conformation. At high temperatures (~1500K) the system is able to occupy high energy regions of conformational space and is able to pass over high
energy barriers (Figure 1.8.3). As the temperature is reduced, the lower energy states become more populated according to the Boltzmann distribution.

![Simulated annealing diagram](image)

**Figure 1.8.3:** Simulated annealing is a combination of molecular dynamics calculations and the change in temperature. At high temperatures, the molecule is permitted to a large search of conformational space in order to minimise the probability of becoming ‘trapped’ in a local minima. Following this, the temperature of the system is slowly reduced; the cycle continues until the molecule is annealed to a lower energy conformation.

Careful control is required when the energy of the system is similar in height to the barriers that separate one region of conformational space from another. It is very difficult to obtain the global minimum structure as it is not possible to distinguish between the local minima and the global minimum. However, good overlay of simulated structures strongly indicate that the global minimum of the molecule has been acquired.
1.8.5 Energy minimisation

The energy obtained from the molecular dynamics calculation is likely to be high and not representative of the actual structure. To obtain more reliable geometries and lower energies, energy minimisation of the system is carried out. The energy minimisation step works in the Cartesian coordinate space and optimises the atomic positions subject to the restraining forces of the force fields. The nearest energy minimised structure is generated and is used as the initial starting conformation. A molecule possesses $3N - 6$ degrees of freedom ($N$ is the number of atoms and 6 corresponds to the translational and rotational degrees), during the calculation. $N$ is varied in order to find the energy minima.

The energy varies with coordinates producing a potential energy surface and the minimum points on the energy surface correspond to the stable states of the system, with deviation from a minimum giving a conformation of a higher energy. A minimisation algorithm is used to identify the geometries of the system that correspond to the minimum points on the energy surface.
1.9 Background and Previous work on this project

The research into the mode of action of the amicetin family of antibiotics has been a decade long project. James Donarski synthesised a *H. hal* 35mer RNA motif containing the predicted binding site of the amicetin antibiotic and carried out unconstrained molecular modelling of the two species. He published a model where the drug interacts with the unpaired U6 in the bulge of the 35mer RNA.\(^{43}\)

Christos Shammas determined the complete NMR assignment of the amicetin antibiotic\(^ {44}\) and a *H. hal* 29mer RNA motif analogous to the *H. hal* 35mer RNA. He successfully produced the small molecule and RNA structure determination scripts using the standard XPLOR-NIH scripts and determined the NMR structures of the free antibiotic\(^ {44}\) and RNA motif. The *H. hal*-amicetin complex was studied and the calculated binding constant (0.43mM\(^{-1}\)) indicated a weak low affinity system.\(^ {68}\)

The interactions between the amicetin family antibiotics and the 23S rRNA motifs were further elaborated by John King. He studied blasticidin S and gougerotin antibiotics and *E. coli* 29mer RNA and the same *H. hal* 29mer RNA motifs.\(^ {69}\) The structure determination scripts were updated with newer information and subsequently used to produce better converged structures of amicetin and *H. hal* RNA motif structures than previous studies.
Chapter 1- Introduction

1.10 Aim of the Project

This project involves elucidating the mode of interaction between a highly conserved secondary structural RNA motif of *B. subtilis* 23S ribosomal RNA and the peptidyl transferase inhibitor bamicetin antibiotic using NMR spectroscopy as a three dimensional structural tool. The project is summarised into three main aims:

1) Characterise the NMR structures of the PT antibiotics: bamicetin, sparsomycin and anisomycin.

2) Elucidate the NMR structures of the 25mer and 27mer motifs of 23S *H. hal* and *B. subtilis* rRNA, respectively.

3) Determine the binding interactions of the RNA-bound antibiotic complexes.

These objectives will be implemented by the use of a range of multidimensional NMR techniques, including measurements on the only highest field 1GHz spectrometer and restrained Molecular Dynamics.

Structure determination of the two RNA motifs will be completed by using a similar structure determination protocol to that of the antibiotics, with modifications where necessary. The mode of action between the *B. subtilis* 23S rRNA and bamicetin is not yet fully understood and so NMR titrations will be carried out to monitor the binding affinity of the bamicetin-*B. subtilis* 27mer RNA complex. The *H. hal* 25mer RNA contains a thiostrepton resistant mutant which is hypersensitive to the amicetin antibiotic and so interaction studies will take place between the 25mer RNA motif and the antibiotic. Similar NMR titration procedures to that carried out on the bamicetin-*B. subtilis* 27mer RNA will be implemented.

Determining the three dimensional structures and dynamics of the individual antibiotics and their interactions with the RNA motifs of different organisms will be an essential step in understanding their binding and inhibitory activities at the atomic level and therefore paving the way for producing more effective drugs.
Chapter 2  Materials and Methods

The materials and methods implemented to obtain the NMR assignment and structure determination of the antibiotics and the RNA motifs are described in this chapter. NMR spectrometers at different field strengths and several multidimensional (1D, 2D and 3D) experiments were utilised. Computational modelling techniques and protocols to obtain the spatial molecular structures are described in detail and conclude the Chapter.

2.1 Antibiotic and RNA sample preparation for the NMR studies

2.1.1 Antibiotics

Bamicetin was a gift from Professor P. Lovett at the University of Pennsylvania, USA and Sparsomycin was supplied by N.I.H, USA and Harry Ottenheijm (Technical University, Eindhoven and University of Maastricht, The Netherlands). The Amicetin and Anisomycin antibiotics were commercially bought from Pharmacia and Sigma-Aldrich, respectively. The antibiotics were supplied as powdered forms and used without further purification.

The NMR samples of the antibiotics were prepared in filtered $Q^{-1}\text{H}_2\text{O} + \text{H}_2\text{O}$ (90%:10%) to observe exchangeable protons and in 100% $\text{H}_2\text{O}$ (99.9% purity) solvent to observe the non exchangeable protons. To replace the solvent, the sample was repeatedly lyophilised in a freeze dryer (E-C Modulyo) at room temperature and pressure applied of 2-4torr. Phosphate buffer ($\text{PO}_4^{III} = 20\text{mM}$, pH 6.2) was added to the samples and the pH of bamicetin (4.8), amicetin (6.2) and sparsomycin (6.2) was measured on a pH/Ion Analyser (Mettler Toledo MA 235). Sample concentrations of bamicetin (5.5mM), amicetin (2.5mM), sparsomycin (4.6mM) and anisomycin (13.3mM) were calculated.

NMR titration studies were carried out with the bamicetin and amicetin antibiotics to investigate the binding interactions with $B.\text{subtilis}$ 27mer RNA and 25mer motif of $H.\text{hal}$ RNA, respectively. A stock solution (500μl) of bamicetin was prepared and microlitre aliquots were systematically added to an NMR tube containing the RNA (0.59mM). 1D
$^1$H NMR experiments were carried out at each titration point to monitor changes to the RNA proton resonances and to detect the saturation point.

Bamicetin (1.65mg) in powdered form was weighed out and placed in a microcentrifuge vial along with filtered Q-$^1$H$_2$O (500μL). The solution was vortexed before being stored in the fridge. The initial titration point was the addition of bamicetin (20μl) producing a bamicetin-\textit{B.subtilis} rRNA complex in 1:0.3 ratio. Further aliquots of bamicetin were added yielding complexes with ratios of 1:0.3, 1:0.6, 1:1, 1:3 and 1:10.

Amicetin (2mg) was carefully added to a microcentrifuge tube along with 1.3ml of filtered Q-$^1$H$_2$O and the tube underwent centrifugation to remove any undissolved particles. The concentration of the stock solution was 2.5mM. To make the first titration point, 100μl of amicetin was added to the NMR tube of the \textit{H.hal} 25mer RNA using a glass pipette. This produced a 1:0.5 ratio of the RNA-antibiotic. Further titrations were carried out with the antibiotic concentration producing a 1:5 ratio.

2.1.2 \textit{B.subtilis} 27mer RNA

The 27mer motif of the \textit{B.subtilis} RNA was commercially bought from Metabion International AG. It was chemically synthesised by solid phase synthesis as two individual strands (13mer and 14mer) and supplied salt-free in separate vials. The samples were purified by high performance liquid chromatography (HPLC) by the suppliers. The RNA was dissolved in microcentrifuge tubes containing filtered Q-water and PO$_4^{3-}$ buffer (500μl, 20mM, pH 6.0).

UltraViolet (UV) absorbance assays using quartz cuvettes$^{70}$ were carried out to determine the concentration of the two strands and combined in equimolar ratios. The concentration of the RNA (0.59mM) was calculated from the Beer-Lambert relationship. The sample was centrifuged on a desktop microcentrifuge (Sigma) to remove any undissolved fragments and subsequently annealed in a water bath at 75ºC for four minutes. It was then cooled to room temperature to ensure the RNA motif adopted a homogeneous, correctly folded conformation free of aggregation. The sample was lyophilised in a freeze dryer at
room temperature and pressure applied (6-8torr) before the volume was made up to 600µl of filtered Q.H₂O + ²H₂O (90%:10%). Mfold server\(^7\) is an algorithm based program which combines experimental measurements of structural RNA elements with a thermodynamic programming tool and was used to predict the Empirical Gibbs free energy (ΔG) and melting temperature (T\(_m\)) of the 27mer RNA motif. These prediction calculations were carried out to assess the stability of the RNA motif (ΔG = -12.8kcal mol\(^-\) and T\(_m\) = 82.5°C). The samples in H₂O were repeatedly lyophilised to remove any traces of the solvent before making solutions in high purity (99.9%) ²H₂O.

2.1.3 H. hal 25mer RNA
The 25mer single stranded motif was purchased from Metabion International AG. The RNA was chemically synthesised by solid phase synthesis and purified by HPLC. It is a single stranded motif and was supplied in 2 vials. To make the NMR sample, the same method was implemented as that applied to the B.subtilis 27mer RNA and concentration of the motif (0.6mM) was determined by UV absorption measurements. All the samples were stored in the fridge at 2-4°C.

2.2 NMR instrumentation and experiments
2.2.1 NMR spectrometers
The majority of the NMR experiments were conducted at the National Institute of Medical Research centre (NIMR) in Mill Hill, London. The usage of high-field NMR spectrometers is indispensible for acquiring high quality data. Increasing field strength allows greater dispersion of the NMR resonances. Access to high field spectrometers included the Bruker Avance 600MHz (14.1T) and 700MHz (16.4T) and the Varian Inova 600MHz (14.1T) and 800MHz (18.8T), all equipped with four RF channels. All spectrometers are accessible for \(^1\)H, \(^13\)C, \(^15\)N triple resonance experiments with pulse field z-gradients (PFGs), deuterium decoupling and electronic variable temperature units. Three of the spectrometers (Bruker Avance 600MHz and 700MHz and the Varian 800MHz) are equipped with cryogenically cooled 5mm \(^1\)H, \(^13\)C and \(^15\)N triple-resonance PFG probes.
Chapter 2 - Materials and methods

The RNA-antibiotic interaction studies were carried out on the latest Bruker Avance 1000MHz (1GHz) spectrometer housed at the Centre de RMN à Très Hauts Champs in Lyon (France). It is the only ultra high-field NMR spectrometer in the world (Figure 2.1.1) incorporating a 23.5T (54mm) superconducting NMR magnet.

![Figure 2.1.1: The Bruker Avance 1000MHz (1GHz) spectrometer at the Centre de RMN à Très Hauts Champs in Lyon (France).](image)

Some of the 1D and 2D experiments were carried out on the Bruker Avance 200MHz, 400MHz and 500MHz spectrometers at the University of Manchester. They are all equipped with broadband probes capable of detecting $^1$H, $^{13}$C, $^{15}$N and $^{31}$P nuclei.

Optimal sample stability for the experiments is paramount, ensuring the molecule is stable in $^1$H$_2$O and $^2$H$_2$O solvent and can withstand a range of temperatures and pH. Experiments were carried out in 90% $^1$H$_2$O: 10% $^2$H$_2$O at 2°C to observe the labile protons as the rate of proton exchange with the solvent is reduced at low temperatures. However, one can execute experiments in $^1$H$_2$O at elevated temperatures to probe the dynamic processes of protons such as the imino protons; but the increase in temperature leads to line broadening. To study the non-labile protons, experiments in 100% $^2$H$_2$O solvent at 25°C
were carried out. This has two advantages, firstly resonances belonging to the exchangeable protons are absent, therefore clearing up some regions of the spectrum and the second reason is at higher temperatures, the NMR peaks are narrower, hence producing a sharper spectrum. The spectra remained invariant over prolonged periods thus attesting to the stability of the samples.

2.2.2 NMR Acquisition Parameters

The acquisition parameters for the 1D and 2D experiments were kept at standard values. 1D $^1$H experiments were typically carried out with 512 or 1024 number of scans to produce adequate signal to noise (S/N) ratio and relaxation delay time of 1.5 seconds was used to ensure sufficient relaxation of all the spins. Samples dissolved in 90% $^1$H$_2$O: 10% $^2$H$_2$O had spectral widths of 24ppm to prevent aliasing of the low field shifted imino protons and those samples present in 100% $^2$H$_2$O had spectral widths of 12ppm. The $^2$H$_2$O resonance served as a field frequency lock signal to maintain the field homogeneity. The pulse calibration, relaxation delay times and probe tuning and matching (essential for good sensitivity) were optimised for each sample. Shimming of the NMR coils was performed for each experiment to obtain the maximum absorption line shape. The 90° pulse widths had optimum durations of ~ 6.8-8.2 seconds.

The transmitter offset was placed on the water resonance to minimise any signal artefacts and provide a suitable spectral width. All the experiments contained water suppression pulse sequences (Presaturation$^{72}$ and WATERGATE$^{73}$) to suppress any residual $^1$H$_2$HO resonances arising due to the large concentration of protons in the solvent.

1D experiments of low natural abundance $^{13}$C {$^1$H} and $^{31}$P {$^1$H} nuclei were carried out with large number of scans (4096 and 32000 scans, respectively) in 100% $^2$H$_2$O at 25°C to obtain the maximum signal sensitivity. The measurement of $^{31}$P nuclei was carried out on the Bruker 200MHz spectrometer. A lower field strength spectrometer is preferred as $^{31}$P nuclei display broadened resonances proportional to increasing field strength in a relaxation mechanism known as Chemical Shift Anisotropy (CSA).$^{74}$ A $^1$H decoupling
pulse was a component of the 1D $^{13}$C and $^{31}$P measurements, with the purpose of eliminating the large number of scalar couplings arising from $^1$H-$^{13}$C and $^1$H-$^{31}$P connectivities, thus producing a clearer spectrum.

The 2D data were acquired using 4096 complex points in $t_2$ and 512 complex points in $t_1$ dimension to obtain optimal spectral resolution. Sensitivity was achieved by acquiring spectra with as large a number of scans as possible (typically 32 or 48).

### 2.2.3 NMR experiments

A range of multidimensional (2D and 3D) homonuclear and heteronuclear NMR experiments was carried out to achieve the sequence specific assignment of the antibiotic and RNA motifs. Through bond (scalar) experiments; $^1$H-$^1$H DQF-COSY, $^1$H-$^1$H TOCSY, natural abundance $^1$H-$^{13}$C HSQC, $^1$H-$^{13}$C HSQC-TOCSY in 100% $^2$H$_2$O at 25°C were particularly important for identifying the individual spin systems of the small antibiotics. The conformation of the antibiotic was determined by the presence of sequential connectivities of non-labile protons in the through-space (dipolar) 2D $^1$H-$^1$H ROESY ($\tau_m$=200 and 400ms) spectra in 100% $^2$H$_2$O at 25°C. To observe the exchange retarded imino (NH) protons, 1D $^1$H NMR spectra in 90% $^1$H$_2$O:10% $^2$H$_2$O at 2°C were measured.

The labile imino and amino protons of the RNA were assigned from the $^1$H-$^1$H NOESY ($\tau_m$=150 and 400ms) in 90% $^1$H$_2$O + 10% $^2$H$_2$O. To identify the non-exchangeable protons of the RNA motifs, through-space (dipolar) $^1$H-$^1$H NOESY ($\tau_m$ 150 and 400ms) experiments in 100% $^2$H$_2$O at 25°C were acquired along with the scalar correlated experiments. The backbone phosphorus assignment of the RNA motif was made using the 2D $^1$H-$^{31}$P CPMG-HSQC-NOESY spectrum. Along with identifying the phosphorus resonances, the HSQC and NOESY components of the sequence identified the ribose and base non-exchangeable protons and their assignment is correlated with those obtained in the NOESY spectra.
Chapter 2 - Materials and methods

A 3D homonuclear NOE-2QC\textsuperscript{82} experiment was obtained to further aid the \textit{H.hal} 25mer RNA ribose proton assignment. This experiment has a particular advantage in that it can observe both the scalar and dipolar connectivities and so previous resonance characterisation achieved from the TOCSY/NOESY spectra are corroborated with this spectrum.

The large range of multidimensional experiments provided the complete characterisation of the individual spin systems of the ribose and bases and the geometry of the base pairs.

\subsection{2.2.4 Processing and Displaying the NMR spectra}

The 2D and 3D NMR spectra were processed on the UNIX based NMRPipe\textsuperscript{83} and NMRDraw\textsuperscript{83} software. NMRPipe is a software system for processing, analysing and illustrating NMR data. It consists of a shell based script and operates \textit{via} a pipeline of commands to provide Fourier processing of the raw time domain data from Bruker and Varian spectrometers. The dedicated Bruker and Varian conversion interfaces interpret and extract the acquisition parameters and the raw data is converted into a common format with a uniform organisation of real and imaginary sections. The package includes comprehensive facilities to process, rephase and display the multidimensional data.

NMRDraw is the interactive interface for NMRPipe allowing visualisation of the multidimensional spectra. The 1D spectral graphics are overlaid on the 2D contour display, allowing real time manipulation of the 1D vectors within the viewed data.

The processing scripts comprised of a number of commands. Zero filling (ZF) was placed on each of the dimensions to enhance the apparent digital resolution by increasing the number of data points. Some spectra were treated with Linear Prediction (LP) which is a more thorough technique than ZF as it predicts additional points from information extracted from the previous data points. The large water resonance present in some spectra distorted the base line and so a polynomial processing function (Poly auto) was applied to automatically correct the base line by discarding points below an estimated threshold as noise. The SOL function was applied to suppress the intense water peak and also to
remove distortion artefacts caused by the residual water resonance in some spectra which may have not been completely removed by the water suppression pulse. Phase correction parameters, zero order phase ($P_0$) and first order phase ($P_1$), were applied to phase the 1D vectors and obtain the pure absorption mode lineshapes of the resonances.

All the processing schemes are constructed as simple shell scripts. The time domain data is then converted into frequency domain by a mathematical complex operation (FT) and saved as an .ft2 file.

### 2.2.5 Analysis

The 1D spectra were all analysed on the Spinworks program, it was fully utilised for deducing the chemical shift assignment of individual resonances, integration of line intensities and measurement of vicinal scalar ($^3J$) couplings from well-resolved multiplets. A range of tools was available to manipulate the digitised data, such as zooming into peaks, phasing different regions of the spectrum and stack plotting of multiple 1D spectra.

The multidimensional data processed as a ft2 file was converted into the ucsf format to be analysed and visualised in SPARKY. SPARKY is a graphical NMR assignment and integration program and can output the analysed data in text peak lists illustrating the assignment, chemical shift values and integrated peak volumes of the resonances. The spectrum can be observed with different contour levels and an optimum threshold level was chosen to reduce errors from peak overlap. Peak volumes are interpreted from the integrated 1D peak line shapes and these are carefully monitored to minimise errors from wrongly fitted peaks.

CcpNMR Analysis program was employed to assign the RNA peaks from multidimensional spectra and convert them into distance constraints for structure calculations. It is a Linux-based graphical user interface and contains additional features not available within SPARKY which were greatly utilised to enhance the quality of the spectra. For example, several spectra may be superimposed in the same window and one can simultaneously and readily toggle between their contours and plots. Chemical shifts
can be assigned and matched on several peaks at once; preventing mistakes and improving the user’s efficiency. The RNA motif sequences were entered so all of their NMR assignable atoms were available for peak annotation and listing. The assignment was converted into distance constraints using the intensity-distance conversion tool for the structure generation.

The ucsf file can be viewed and printed in the GSview program. It can then be transported to other drawing programs for producing figures and diagrams.

### 2.3 NMR techniques

#### 2.3.1 1D $^1$H experiments

A 1D $^1$H NMR experiment consists of two sections, a preparation and a detection stage (Figure 2.3.1). During the preparation stage, a relaxation delay was used to allow the spin system (bulk magnetisation) to fully return to the equilibrium state along the longitudinal z-axis ($M_z$) in a static magnetic field ($B_0$). Following this, a 90° pulse consisting of a very short burst ($<20\mu$s) of high power RF ($B_1$ field) along the x axis, rotates the $M_z$ onto the x-$y$ transverse plane. The transverse magnetisation ($M_y$) in the x-$y$ plane precesses under the influence of $B_0$ at the individual larmor frequencies ($\nu_0$) of the nuclei and induces an electric current in the detection coil placed along the y-axis.

The detected signal decays exponentially with time and is called the Free Induction Delay (FID). The NMR spectrum is obtained by the Fourier transformation (FT) of the detected time domain data (FID).
Chapter 2 - Materials and methods

Figure 2.3.1: The 1D $^1$H NMR pulse sequence is composed of two sections: preparation and detection. The preparation stage is the relaxation delay (typically lasting a few seconds), the spins are polarised resulting in net macroscopic magnetisation ($M_z$) in the direction of $B_0$. Following this, a 90° pulse is applied to rotate $M_z$ on to the transverse plane. The detection period which lasts between 50ms and a second measures the freely precessing signal (FID) relaxing back to equilibrium.

A single FID produces a weak NMR spectrum and so the experiment is repeated multiple times in a procedure known as signal averaging, the resulting FIDs are co-added to improve the (S/N) ratio and hence increase the sensitivity of the peaks.

2.3.2 1D $^{13}$C ($^1$H) and $^{31}$P ($^1$H) decoupled experiments
1D $^{13}$C and $^{31}$P experiments were routinely carried out with broadband proton decoupling and the pulse sequence is shown in Figure 2.3.2. The nucleus which experiences the decoupling RF pulse is placed in curly brackets and so $^{13}$C ($^1$H) denotes carbon observation in the presence of proton decoupling. The decoupling removes the $^1$H-X (X= $^{13}$C, $^{31}$P) couplings thus concentrating all of the heteronucleus resonance intensity into a single line. As a result, a significant enhancement in signal sensitivity is observed and the spectrum is simplified. Another advantage arising from $^1$H decoupling is the
continuous saturation of the proton spins provides further NOE enhancement of the heteronuclei signal.

![Decoupling pulse](image)

**Figure 2.3.2:** The $^1\text{H}$-decoupled 1D NMR pulse sequence applied for the detection of heteronuclei $X$. The decoupling $^1\text{H}$ pulse is employed after the $90^\circ$ pulse and prior to data acquisition.

The decoupling pulse works on the basis of the mutual scalar coupling arising between $^1\text{H}$ and $X$ with a magnitude of $J$ Hz. Thus, if the $^1\text{H}$ spins are irradiated during data acquisition, the $X$ resonance displays no coupling to the proton and the $^1\text{H}$ spins are therefore decoupled.

### 2.3.3 2D homonuclear and heteronuclear experiments

Two dimensional (2D) NMR experiments have been extensively utilised in deducing structure determination of biomolecules such as proteins and nucleic acids (DNA and RNA). In 1D spectra, the chemical shift and spin-spin coupling fine structure is obtained whereas 2D experiments relay the magnetisation between spins in different environments providing scalar (through-bond), dipolar (through-space) and chemical exchange information. As a result, the experiments provide complete mapping of the molecular stereochemistries and the local and global conformation of the macromolecule. Alongside
the standard $^1$H nuclei, 2D experiments are versatile in detecting heteronuclei such as $^{13}$C, $^{15}$N, $^{19}$F and $^{31}$P. Regardless of the nature of the interaction to be mapped, all the 2D sequences can be divided into four well defined time periods as illustrated in Figure 2.3.3.

**Figure 2.3.3**: The general scheme for any 2D NMR experiment. The nuclei are aligned with the applied magnetic field ($B_o$) during the preparation (P) stage, immediately after the NMR (90°) pulse is applied, the spins evolve during the incremental $t_1$ values in the evolution period (E). The mixing period comprising of a 90° pulse achieves coherence transfer producing an observable signal. Each incremented FID is acquired as a time variable $t_2$ during the detection period (D).

The preparation period is identical to the 1D, it prepares the equilibrium magnetisation ($M_z$) and the initial 90° pulse generates the transverse magnetisation ($M_y$) which evolves during the evolution period ($t_1$). No observations are made during the evolution period and therefore this coherence can be the unobservable multiple quantum (MQ) of which the evolution is an important feature of two-dimensional NMR. The mixing period may be as simple as a single pulse or contain an array of pulses, transferring the magnetisation between spins. The mixing period is also the location where MQ coherence is transformed into an observable signal. The signal is then recorded during the detection period, $t_2$.  

-94-
A 1D experiment consists of a plot of the signal intensity against one frequency whereas a 2D experiment is recorded as two time domains and FT to produce the intensity along two frequency axes. The influence of the first pulse on the FID depends on the length of the evolution period (t₁). A second time dimension is created by repeating the pulse sequence with incrementations of t₁. For each value of t₁, FID vs (t₂) is stored producing a time domain 2D data matrix s(t₁,t₂). Fourier transformation of the data matrix produces the 2D frequency spectrum s(F₁,F₂).

2.3.4 ¹H-¹H Correlated Spectroscopy (COSY)

The ¹H-¹H CORrelated SpectroscopY⁸⁷,⁸⁸ (COSY) was the first multidimensional sequence developed and the most widely used 2D method (Figure 2.3.4). It identifies spins which have a mutual scalar coupling by correlating their chemical shifts.

![Diagram of ¹H-¹H COSY pulse sequence](image)

**Figure 2.3.4:** The ¹H-¹H COSY pulse sequence. The equilibrium magnetisation (Mₐ) is rotated onto the y axis by the initial 90° pulse where it evolves during the t₁ interval. The second 90° pulse (mixing period) distributes the magnetisation amongst the spin states of the coupled nuclei before signals are acquired in the t₂ detection period.

Following the initial excitation, the first 90° pulse creates magnetisation in the transverse plane which evolves during the t₁ according to its chemical shift. The second 90° pulse (mixing sequence) transfers some of the magnetisation (coherence transfer) to the neighbouring spins to which it is coupled. Cross-peaks appear at the intersection of the
two resonance frequencies of the directly coupled spins. This is explained by an example consisting of a two $J$-coupled spin system: A and X at their chemical shifts $\nu_A$ and $\nu_B$. The magnetisation which is associated with spin A (after the first 90° pulse) will precess during $t_1$ according to its $\nu_A$. The second 90° pulse transfers some of the magnetisation of A to its coupled spin X. The magnetisation remaining on A precesses in the detection period at its chemical shift appearing in the spectrum at $\nu_A$ in both dimensions and thus producing the diagonal.

The transferred magnetisation precesses in $t_2$ at the frequency of spin X and as a result produces a peak which corresponds to two different chemical shift values ($\delta_A, \delta_B$) in the two dimensions and produces a cross peak. This process is repeated in reverse to produce cross peaks located at the intersections of the diagonal peaks ($\nu_A, \nu_B$). The cross peaks are a direct evidence of scalar coupling between the two spins A and X.

This coherence transfer process creates antiphase cross peak multiplets, which contrast with the in-phase structure of the diagonal peak. As a result, the lines of the cross peak tend to cancel one another out, leading to a reduction in intensity whereas, in the case of the diagonals, the lines are reinforced. Therefore, there may be a considerable difference in the overall intensity of the cross and diagonal peaks. This problem is illustrated in Figure 2.3.5.
Figure 2.3.5: A schematic 2D NMR spectrum illustrating the absorptive (bottom left) and dispersive (bottom right) nature of the diagonal and cross peaks, respectively. Cross peaks which have coupling constants and linewidths of similar value, their positive and negative regions of the peak can cancel each other out resulting in weakened cross peak intensities.

There are two major drawbacks of a COSY experiment:

1) The fine structure components that produce each cross peak appear in anti-phase resulting in poor sensitivity unless the coupling is at least as large as the line width.

2) The diagonal and cross peaks are 90° out of phase and strong weighting functions must be applied to balance the line shapes, again this will result in further reduction of sensitivity.
2.3.5 $^1$H-$^1$H Double-Quantum Filtered Correlation Spectroscopy (DQF-COSY)

The basic COSY sequence suffers from a major drawback, the diagonal peaks exhibit pure absorption mode line shapes while the cross peaks are in dispersion mode. As a result, the broad tails associated with the diagonal peaks can end up masking nearby cross peaks.

This is solved by the application of a third pulse and the use of a modified phase-cycle or gradient sequence, forming the Double-Quantum Filtered Correlation Spectroscopy ($^1$H-$^1$H DQF-COSY) homonuclear experiment. This experiment produces both the diagonal and cross peak multiplets in absorption mode and possessing the same lineshape. The DQF-COSY spectrum is often preferred to the conventional $^1$H-$^1$H COSY as the pure absorption line shapes provide the highest possible resolution and allow one to extract coupling information from the fine structure within the cross peaks multiplets.

The pulse sequence is shown in Figure 2.3.6. Following the $t_1$ frequency labelling, the second 90° pulse generates MQ coherence which is not observed in the COSY experiment. The phase cycle detects only these signals and the third pulse reconverts them to single quantum coherence which is detected.

Singlets (i.e. methyl protons and H$_2$O protons) are unable to create double quantum coherence and thus are not observed in the spectrum. This has an advantage as sharp intense singlets which usually produce intense $t_1$ noise bands are omitted improving the clarity of the spectrum. An additional benefit is that the pulse sequence also acts as a very good water suppression technique by suppressing the large singlet solvent proton resonance. On the negative side, the signal to noise ratio is decreased by two fold due to the generation of the double quantum coherence, which detects only half of the available signal. However, the clarity generated by the suppressed diagonal outweighs this hindrance.
Figure 2.3.6: The pulse sequence for $^1\text{H}-^1\text{H}$ double-quantum filtered COSY (DQF-COSY). The observed signal arises from the double-quantum coherence present during the short time interval between the second and third 90° pulses. This delay lasts a few microseconds, (20µs).

2.3.6 $^1\text{H}-^1\text{H}$ Total Correlation Spectroscopy (TOCSY)

$^1\text{H}-^1\text{H}$ Total Correlation Spectroscopy (TOCSY) experiment is similar to COSY in that it detects scalar couplings between nearby protons. However unlike COSY, additional cross peaks between spins which are connected by an unbroken chain of $^1\text{H}-^1\text{H}$ couplings can also be observed.

In principle, a single cross-section taken at the chemical shift of one spin should show through bond correlations to all the other resonances which are part of its coupling network. The key part of this experiment is the spin-lock mixing period (SLy) (Figure 2.3.7). During this period, the evolving $z$-magnetisation is transferred from one spin to another at a rate which depends on the magnitude of the coupling constant and the duration of the mixing time. It is this magnetisation transfer which produces the cross peaks in the spectrum.
Chapter 2 - Materials and methods

Figure 2.3.7: $^1$H-$^1$H TOCSY pulse sequence. The spin-lock mixing time (SLy) utilises an MLEV-17 sequence$^{89}$ consisting of a long, low-powered pulse train of 180° pulses.

2.3.7 $^1$H-$^{13}$C Heteronuclear Single Quantum Coherence (HSQC)

The $^1$H-$^{13}$C HSQC experiment is widely used to detect the one-bond correlation between a proton and its directly scalar coupled heteronucleus. There is no diagonal and all the signals represent correlations from $^1$H to $^{13}$C, $^{15}$N or $^{31}$P. The transverse heteronuclear magnetisation is generated by a polarisation transfer from the attached protons via the Insensitive Nuclei Enhancement Polarisation Transfer (INEPT) sequence.$^{90}$ The pulse sequence shown in Figure 2.3.8, creates the proton antiphase magnetisation which is transferred to the directly attached heteronucleus. The heteronucleus (X) magnetisation evolves at its chemical shift during $t_1$ with a 180° RF pulse at its midpoint to remove the effect of $^1$H-X coupling. The removal of the $^1$H-X interaction ensures only the heteronuclear chemical shifts remain in the $F_1$ dimension and accordingly improves the resolution of the X-nuclei resonances.
Figure 2.3.8: The NMR pulse sequence for the $^1$H-$^{13}$C experiment. The black boxes represent 90° pulses which are applied along the x-axis (except the ‘y’ pulse which is directed along the y-axis and is indicated in the figure) and the blue boxes represent 180° pulses. Delay times ($\Delta_{1,2}$) are dependent upon $J_{HX}$ and values of $\frac{1}{4} J$ are generally used. The sequence applies an INEPT step to generate transverse X magnetisation which evolves during $t_1$ and is then transferred back to the proton magnetisation. The composite decoupling pulse (shaded grey) is the GARP sequence.

Following $t_1$, the heteronuclear magnetisation is transferred back onto the $^1$H by an INEPT step in reverse to produce the in-phase $^1$H magnetisation. The $^1$H-X couplings were removed by the Gated Alternating Rectangular Pulse (GARP) decoupling pulse.

### 2.3.8 $^1$H-$^{13}$C HSQC-TOCSY

Proton spectra obtained from 2D homonuclear correlation experiments can sometimes be overcrowded, hindering the complete interpretation of the resonances and even with a HSQC experiment, the complete overlap of the proton resonances can fail to provide unambiguous identification of the parent heteroatom. In such instances, greater dispersion of the heteroatom chemical shifts can be utilised to provide additional separation of the proton-proton correlations.
Chapter 2 - Materials and methods

**Figure 2.3.9:** The $^1$H-$^{13}$C HSQC-TOCSY pulse experiment consists of a TOCSY spin-lock mixing sequence appended to the HSQC spectrum, characterising the scalar coupling.

In the $^1$H-$^{13}$C HSQC-TOCSY pulse sequence (Figure 2.3.9), a TOCSY spin-lock mixing period is added after the HSQC sequence. This allows the magnetisation which has returned to the $^1$H to be transferred onto neighbouring scalar-coupled protons, thus producing a map of the entire proton coupling network as illustrated in Figure 2.3.10.

**Figure 2.3.10:** An outline of the scalar coupling pathway mapped in a 2D HSQC-TOCSY experiment. Direct correlations are observed for the proton bound to the spin $\frac{1}{2}$ heteronucleus ($C_a$) and further correlations are produced for those protons through the TOCSY magnetisation transfer ($H_b, H_c$).

Extended mixing periods will relay magnetisation along the proton network and can therefore provide a complete proton subspectrum of the molecular fragment to which the heteroatom belongs. The proton correlations appear along rows at the chemical shift.
values (F₁ dimension) of the hetero spins. Overlap encountered in the 2D TOCSY spectrum is removed due to sufficient chemical shift dispersion in the X-spin dimension.

The decoupling pulse is applied to remove multiplets arising from the ¹³C nuclei scalar coupling with nearby protons and favourably increases the sensitivity of the resonances. However, the experiment is not as sensitive as the 2D ¹H-¹H TOCSY since the cross peaks originate from the ¹³C nuclei which has a low natural abundance and low gyromagnetic value. Apart from the TOCSY step, COSY, NOESY and ROESY mixing schemes can also be utilised.

### 2.3.9 ¹H-¹H Nuclear Overhauser Effect Spectroscopy (NOESY)

A two-dimensional ¹H-¹H Nuclear Overhauser Effect Spectroscopy (NOESY) experiment generates dipolar couplings arising from cross relaxation via the dipolar mechanism. Therefore, a NOESY cross peak between two spins indicates the two nuclei are in spatial proximity (typically ≤ 5.0Å).

![Figure 2.3.11: A ¹H-¹H NOESY pulse sequence. The important part of the sequence is the mixing period τₘ, during which the NOE develops generating the cross peaks.](image)

The NOESY pulse sequence (Figure 2.3.11) is similar to that of the COSY experiment with the exception being a mixing period containing a fixed time interval and an additional 90° pulse (Figure 2.3.11). Following the initial excitation and t₁ evolution, the
magnetisation vector exists in the transverse plane. The second 90° pulse rotates one component of the magnetisation onto the \( -z \) axis (longitudinal magnetisation), thereby, generating the required population inversion that enables the NOEs to develop during the subsequent mixing period. The magnetisation transfer takes place during the mixing time, generating the cross peaks at the chemical shifts of the nuclei that are relaxing each other. Observable transverse magnetisation is created from the remaining longitudinal magnetisation by the application of the third 90° pulse and the signal is recorded during the detection \( (t_2) \) period.

The intensity of the cross peak and internuclear distance share a linear relationship up to the mixing time of \( \tau_m < 150 \text{ms} \). At this point the maximum intensity is reached and will start to decrease at longer \( \tau_m \) values due to the onset of spin diffusion. Thus, the length of the mixing time delay is chosen according to the transfer process that is to be observed. The experiment identifies the proximity of the protons and is important for macromolecules where the secondary or tertiary structures are dictated by non-bonded interactions. Furthermore, the diagonal and cross peaks are obtained with absorption mode line shapes and hence the fine structure multiplets are in phase and possess good resolution, ideal for studying large molecules.
2.3.10 $^1$H-$^1$H Rotating-frame Overhauser Effect Spectroscopy (ROESY)

The $^1$H-$^1$H ROESY pulse sequence is analogous to the NOESY experiment, except that instead of generating cross peaks by cross relaxation between the $z$-magnetisation of different spins, the cross peaks in ROESY arise from cross relaxation between spin-locked transverse magnetisation (Figure 2.3.12).

![Diagram](image)

**Figure 2.3.12:** (a) the magnetisation is inverted along the $-z$ axis in a NOESY experiment, (b) the magnetisation is spin-locked along the $y$-axis and parallel to the $B_1$ field in a ROESY experiment.

The 90° pulse tips the magnetisation into the transverse plane and a continuous low-power spin lock pulse allows the ROE to develop whilst the magnetisation is held static in the transverse plane (Figure 2.3.13). The mixing time is chosen according to the cross-relaxation times of the nuclei.

The experiment is a useful alternative to NOESY as the ROE is always positive, regardless of the molecular size or rotational correlation time of the spins. This has the benefit of looking for ROE enhancements in small to medium sized molecules whose correlation times reduce the conventional NOEs close to zero or unobservable.
2.3.11 $^1$H-$^{31}$P CPMG-HSQC-NOESY

The $^1$H-$^{31}$P CPMG – HSQC – NOESY experiment is designed to carry out the sequence specific assignment of phosphorus resonances in nucleic acids. The experiment transfers magnetisation from the ribose protons, $H_3'$ and $H_5'/5''$ to scalar coupled phosphorus nuclei. This is executed by the $^1$H-$^{31}$P HSQC component of the pulse sequence which is a series of $180^\circ$ pulses known as the Carr – Purcell – Meiboom – Gill (CPMG) pulse trains as shown in Figure 2.3.14. These pulse trains prevent dephasing of the spins and keep them closer to each other to allow more efficient magnetisation transfer. In particular, the CPMG trains provide optimal refocusing between nuclei correlated along the phosphodiester backbone, which may experience significant conformational exchange. The CPMG train prevents the build up of $^1$H-$^1$H scalar coupling which has a similar magnitude to $^1$H-$^{31}$P coupling (5-7Hz) and can therefore lead to distortions in the spectrum.

The NOESY segment of the experiment is performed to relay magnetisation to protons which are close in space to the phosphorus nuclei and the GARP sequence removes the $^1$H-$^1$H scalar couplings.

---

**Figure 2.3.13**: $^1$H-$^1$H ROESY pulse sequence. The ROE develops during the long spin-lock pulse (a low level RF field, typically 1-2 kHz) which constitutes the mixing period ($\tau_m$).
Figure 2.3.14: The $^1$H-$^{31}$P CPMG-HSQC-NOESY pulse sequence. $\tau$ is a delay time consisting of 100ms and mixing time ($\tau_m$) of 500ms. The black filled bars represent the 90° pulses and the 180° pulses are shown by the blue filled bars. The NOESY sequence is applied prior to detecting the $^1$H nuclei.

2.3.12 3D NOE/2QC

A common problem encountered when attempting the NMR assignment of RNA is the extremely overcrowded small chemical shift window (~ 0.8ppm) in which the majority of the ribose protons (H2', H3', H4', H5' and H5'') reside. This can result in severe overlap of cross peaks in a 2D $^1$H-$^1$H NOESY spectrum which can be resolved by spreading the signals in the third dimension, leading to an increase in signal dispersion.

The 3D NOE/2QC experiment comprises a NOESY and double quantum (2QC) component and the pulse sequence is illustrated in Figure 2.3.15. A DANTE presaturation pulse is applied at the beginning of the experiment to suppress the water signal. A short delay follows the presaturation step to allow saturated signals near to the water resonance to partially recover. The chemical shift labelling occurs during the evolution period ($t_1$) after which the magnetisation is transferred through the NOE contacts during the NOESY mixing sequence ($\tau_m$). For the duration of the $\tau_{DQ}$, double quantum coherence is generated for the scalar coupled spins. Single and multiple quanta (MQ) coherence is removed by phase cycling while labelling in $t_2$ takes place by the sum frequency ($\delta_A + \delta_X$) of two coupled spins A and X. Finally coherence is transferred back to the corresponding single quantum frequency and is detected during $t_3$. 

-107-
Double quantum selection is accomplished in four phase-cycling steps. A further four steps are required to remove the single and MQ coherences during the NOESY mixing period ($\tau_m$). Aliasing is applied along the single-quantum dimension (F1) and the double-quantum dimension (F2) to maximise the spectral resolution with the least possible overlap.

The 2QC sequence identifies the though-bond transfers which are restricted to the intraregional connectivities while the NOESY section facilitates through-space sequential assignments determining the secondary structure of the RNA motif. It has several advantages over the conventional $^1$H-$^1$H COSY/TOCSY sequences; working well for low coupling constant/linewidth ratio resonances which are particularly useful for H1’-H2’ connectivities in RNA. It does not produce a diagonal and so nearby cross peaks are not obscured and the magnetisation transfer is limited to one or two steps unlike the mixing step of the TOCSY experiment which diffuses magnetisation throughout the entire ring, potentially reducing the peak intensities.

It is a proficient experiment allowing the assignment of almost all of the non-exchangeable protons, in particular the ribose protons and thereby eliminating the need for expensive isotope-labelling of the RNA which can lead to broadened lines due to heteronuclear $J$-coupling.
Figure 2.3.15: The pulse sequence for 3D NOE/2QC experiment. The DANTE sequence (orange bars) consists of small flip angle pulses. The narrow black bars are 90° RF pulses while the 180° RF pulses are represented by the wider green blocks. The 180° RF pulse in the middle of the NOE mixing time ($\tau_m=200\text{ms}$) and DQ excitation delay ($\tau_{DQ}$) are used as refocusing pulses. The carrier frequency was set between the chemical shift region of the H6 and H5/H1' resonances.
Chapter 2 - Materials and methods

Figure 2.3.16 is a diagram illustrating how the information is displayed in the 3D NOE/2QC spectra. The 3D spectrum consists of NOESY and 2QC planes which are combined to provide a comprehensive NMR assignment of the molecule.

Figure 2.3.16: Schematic drawing illustrating the connectivities observed in the 3D NOE/2QC experiment. (a) dipolar connectivities in the 2D NOESY planes can be linked to the scalar connectivities in the 2QC planes. Scalar $^3J_{H5-H6}$ correlations in green circles can display NOE correlations to the ribose protons (red circles). (b) The sequential connectivities in the entire nucleotide can be mapped.
2.4 Assignment strategy and constraint generation

2.4.1 NMR resonance assignment

Antibiotics

The NMR assignment of the antibiotics was initiated with the full analysis and interpretation of the 1D $^1$H NMR spectra. The chemical shift values and scalar coupling constants were measured, laying the foundation for the assignment of the 2D spectra. 2D homonuclear $^1$H-$^1$H DQF-COSY and $^1$H-$^1$H TOCSY ($\tau_m=75$ms) and heteronuclear $^1$H-$^{13}$C HSQC and $^1$H-$^{13}$C HSQC-TOCSY ($\tau_m=75$ms) experiments were acquired to assign the scalar coupled protons. Confirmation of the assignment and the spatial orientation of the molecules were determined by the dipolar based $^1$H-$^1$H ROESY experiment at various spin-lock mixing times ($\tau_m=100, 200$ and $400$ms). The behaviour of the labile protons was monitored by 1D variable temperature (2-50°C) $^1$H experiments.

RNA motifs

The information extracted from the assignment of the RNA protons falls into two categories:

1) Labile protons (imino and amino) residing on the aromatic bases and the adenine H2 protons characterise the base pairing scheme of the duplex structure.

2) The non-labile protons (base and ribose) except the H2 of adenine provide $^1$H-$^1$H NOEs between neighbouring or closely spaced nucleotides in the sequence.

The NMR assignments of the 25mer and 27mer motifs of 23S H.hal and B.subtilis 23S rRNAs, respectively, were initiated by recording 2D $^1$H-$^1$H NOESY spectra (90% $^1$H2O + 10% $^2$H2O) at various mixing times ($\tau_m=100, 200$ and $400$ms) with the prime purpose of carrying out the assignment of the labile imino and amino protons. The imino-imino and the imino-amino connectivities gave evidence for the base pairing arrangement of the RNA motifs. Strong NOE correlations were also observed between the imino protons to pyrimidine H5 protons and the adenine H2 base protons. To determine the secondary structure of the RNA motif, sequential NOE connectivities were observed between the non
labile aromatic base protons H6/H8 and the ribose glycosidic proton H1’ and this was observed in the $^1$H-$^1$H NOESY (τm= 150 and 400ms) spectra measured in 100% $^2$H2O. The assignment of the exchangeable and non exchangeable protons simultaneously established the secondary structure of the RNA motif.

The NOESY spectra ($^1$H2O and $^2$H2O) provided information corresponding to intranucleotide, sequential and long-range or inter-strand short distances between the same proton types and these can be summarised into four distance types.

Table 2.4.1: Proton distance types in RNA. $i, j$ represent two adjacent nucleotides.

<table>
<thead>
<tr>
<th>Distance type ($^1$H-$^1$H)</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intranucleotide (non exchangeable)</td>
<td>ribose H1’-H2’, purine base-ribose (H8-H1’)</td>
</tr>
<tr>
<td>Sequential between non-exchangeable and exchangeable protons</td>
<td>ribose $i$ (H1’)-base $j$ (H8), base $i$ (H5)- base $j$ (H8), imino-imino</td>
</tr>
<tr>
<td>Interstrand between protons within a base pair (exchangeable)</td>
<td>pyrimidine (NH3)- purine (H2), (imino – amino connectivity)</td>
</tr>
<tr>
<td>Exchangeable protons or adenine H2 in adjoining stacked base pairs</td>
<td>U (H3)-U(H1) same strand, U(H3)-C(H4) opposite strand</td>
</tr>
</tbody>
</table>

The four types of $^1$H-$^1$H distances outlined in Table 2.4.1 provide the complete assignment of the RNA resonances, thus allowing the structure determination of the molecule. The full assignment was achieved with the aid of the 2D $^1$H-$^1$H DQF-COSY, $^1$H-$^1$H TOCSY, $^1$H-$^{13}$C HSQC and $^1$H-$^1$H NOESY experiments, while the 3D NOE-2QC experiment was measured for the ribose proton assignment of the H.hal 25mer RNA. This experiment works effectively for unpaired nucleotides which are characteristically known to produce weaker cross peaks in 2D experiments.
Figure 2.4.1 displays the protocol which was employed to carry out the structure determination of the RNA motifs, starting with measurement of the NMR spectra and finishing with a full conformational analysis of the determined structure.

**Figure 2.4.1**: A scheme illustrating the various steps of the protocol followed to generate the NMR structures of the RNA motifs of the above, the rate determining step is the NMR assignment; incorrect assignment will produce inaccurate constraints resulting in imprecise structures with increased number of violations.

A full range of NMR experiments is carried out to ensure the correct assignment of the molecule and to yield accurate experimental constraints for the production of high quality final structures.
2.4.2 NOE based distance constraints

The distance-based NOE constraints were acquired from the dipolar coupled experiments: NOESY and ROESY. The antibiotics are small molecules and so the ROESY experiment was implemented whereas the $^1$H-$^1$H NOESY spectra ($\tau_m = 150$ and $400\text{ms}$) were utilised for the RNA motifs.

Antibiotics

The $^1$H-$^1$H distance dependent ROE constraints were obtained from 2D $^1$H-$^1$H ROESY spectra with different spin-lock mixing times ($\tau_m = 100$, $200$ and $400\text{ms}$). The intensities of the ROE cross peaks were integrated to obtain interproton distances. The cross peak intensity ($I$) in a NOESY or ROESY spectrum is directly proportional to the rate of magnetisation transfer between each pair of dipolar coupled nuclear spins (i,j). The cross peak intensities ($I_{ij}$) are related to distances between spins i and j via the mixing time ($\tau_m$) and the relaxation matrix ($R$) and shown below in 2.4.1.

$$I_{ij} \propto \exp (-R\tau_{\text{mix}})_{ij} \quad 2.4.1$$

The relaxation matrix is a function of the transition rates and is determined by spectral densities and dipolar coupling magnitude. The initial rate approximation states that the intensity of a particular ROE/NOE cross peak grows linearly with increasing mixing time and this can be directly related to one-sixth of the interproton distance ($r_{ij}^{-6}$) (2.4.2).

$$\text{NOE} \propto r_{ij}^{-6} \quad 2.4.2$$

Known covalent distances obtained from fixed non-labile protons were used as a reference to calibrate the experimental intensities to be consistent with their respective interproton distances. The molecular ruler used for determining the ROE intensities of bamicetin was the interproton distance of H13-H14 of the modified cytosine base (H5-H6 numbering in pyrimidines). The olefinic protons H6-H7 in sparsomycin were applied as a molecular
ruler and likewise the inter proton distance of the aromatic protons (H10-H11) was used for determining the distance based ROE constraints of anisomycin. Inter-proton distances ($r_{ij}$) were computed from the 2D ROE/NOE cross peak volumes using the Isolated Spin Pair Approximation (ISPA) technique. The NMR analysis programs, SPARKY and CcpNMR Analysis incorporate the ISPA model to estimate the inter-proton distances for the production of distance-based constraints.

The inter-proton distance of two spins (i and j) are calculated from their peak volumes ($a_{ij}$) using a reference nuclei distance ($r_{ref}$) and intensity ($a_{ref}$) and displayed in 2.4.3.

$$r_{ij} = r_{ref} \left(\frac{a_{ref}}{a_{ij}}\right)^{1/6}$$

Fixed interproton distances (i.e. methylene and methyl protons) were measured from the calculated starting structures (B3LYP calculations in section 2.5.2). These empirical distances were plotted against their ROE cross peak volume; producing an ROE growth curve. The growth curves were constructed for each $^1$H-$^1$H ROESY spectrum with individual mixing times to produce a consistent set.

At long mixing times ($\tau_m > 150$ms) ROE intensities become a complicated function of many internuclear distances due to the onset of spin diffusion. At this point, the relationship between the cross peak intensity and the mixing time is not linear. A second limitation to the ROESY experiment is that at higher mixing times some cross peaks can arise due to TOCSY correlations as artifacts. Therefore, the distance constraints obtained from the 400ms spectrum were given large error bounds. Other ROE constraints were placed with increased error bounds if they possessed low intensities, making it difficult to distinguish them from noise.
RNA motifs

NOE distance dependent constraints were extracted from the assigned exchangeable and non-exchangeable 2D $^1$H-$^1$H NOESY spectra. Non-exchangeable proton cross peak volumes were obtained from spectra at short ($\tau_m=150$ms) and long ($\tau_m=400$ms) mixing times and these were converted into inter-proton distances using the repulsive algorithm in CcpNMR Analysis. The distances were classified according to their cross-peak intensities: strong 1.8-2.5 Å (strong at 150ms and 400ms), medium 1.8-3.0 Å (medium at 150ms and medium/strong at 400ms), weak 1.8-5.0 Å (weak at 150ms and medium/weak at 400ms) and very weak 1.8-7.0 Å (weak at 400ms). The intensities are placed in bands with upper and lower bounds to account for any uncertainty in the estimated distance. The molecule will possess dynamics in a solution and so it is not possible for most protons to have fixed distances.

NOE constraints obtained from the cross-peaks arising from the labile protons were placed with larger error bounds as these protons are likely to exchange with the solvent protons. They were further classified as hard or soft constraints, where the former had a greater weight placed on them in the structure generation calculation with respect to the softer constraints. The softer constraints were categorised for protons arising from non-base paired nucleotides or those present on the terminal bases.

2.4.3 Dihedral angle restraints

Antibiotics

Additional constraints were derived from vicinal scalar ($^3J_{H,H}$) coupling constants, characterising the scalar interactions between the nuclei linked by three covalent bonds (H-C-C-H). The three-bond vicinal coupling constant represents the magnitude of the spin-spin coupling ($^3J_{H,H}$) and the intervening torsional angle ($\varphi$).

1D $^1$H spectra recorded with high digital resolution were acquired to accurately determine the dihedral angle constraints. The MestRe-C$^{92}$ software implements the Karplus
Chapter 2 - Materials and methods

principle\textsuperscript{93} and was used to arbitrate the dihedral angles from the experimental data. The Karplus relation (2.4.4) describes the correlation between vicinal scalar coupled ($^{3}J_{\text{H,H}}$) protons and dihedral torsion angles and is represented diagrammatically by the Karplus curve in Figure 2.4.2.

$$^{3}J (\varphi) = A\cos^{2} \varphi + B\cos \varphi + C \tag{2.4.4}$$

$^{3}J$ = vicinal scalar coupling constant, it is field independent and quoted in hertz (Hz).

$\varphi$ = dihedral angle

A, B and C = empirically derived parameters depending on the two atoms involved and the neighbouring substituents neighbouring the bond.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{karplus_curve.png}
\caption{The Karplus curve illustrates the correlation between the vicinal scalar coupling constant ($^{3}J$) and the dihedral angle ($\Phi$). The magnitude of the coupling is smallest ($\sim 2$Hz) when the dihedral angle is at 90° while the $^{3}J$ shows the maximum value (13Hz) at 0 and 180°.}
\end{figure}
The coupling constants are dependent on other factors, such as the electronegativity of the attached substituents and bond lengths and therefore the information was used conservatively. The dihedral angles contained error limits (±10° to ±30°) depending on resolved multiplets and the reliability of the value.

**RNA motifs**

The torsional angle restraints consisted of determining the sugar puckers of the ribose rings. The conformations were determined from the scalar coupling constants measured from the ribose H1‘-H2’ protons in the 2D ¹H-¹H DQF-COSY spectrum. The magnitude of the ³JH1‘-H2’ coupling constants determined the sugar puckers and the five backbone torsional angles. Residues with ³JH1‘-H2’ > 7 Hz were restrained to C2‘-endo sugar pucker (δ = 140° ± 20°, ν₁ = 35° ± 5° and ν₂ = -35° ± 5° ) and those with ³JH1‘-H2’ < 2 Hz were restrained to the C3‘-endo sugar pucker (δ =80° ± 20°, ν₁ = -25° ± 30° and ν₂ = 37.3° ± 30°). A-form geometry was invoked for all C3‘-endo nucleotides and the backbone torsional angles were set with the presence of the ³¹P chemical shifts of each nucleotide in the -1 to -5ppm region accordingly for C3‘-endo conformation (α = -68° ± 17°, β = 178° ± 17°, γ = 54° ± 17°, ε = -153° ± 17°, ζ = -71° ± 17°) and C2‘-endo riboses (α =-80° ± 17°, β = -150° ± 17°, γ = 54° ± 17°, ε = 170° ± 17°, ζ = -90° ± 17°). The glycosidic angles of all the stem residues were constrained to anti conformation (-150 ± 17°) based on the observation of canonical Watson-Crick base-pairs. The unpaired nucleotides in the internal loops of the *B.subtilis* 27mer RNA motif, U6, A27 and C28 did not have their backbone constrained allowing greater flexibility during the structure determination calculation. Likewise, the backbone constraints of the non base paired residues in the hairpin and internal loops of the *H.hal* 25mer RNA were left absent.
2.4.4 Hydrogen bond restraints

Hydrogen bonding restraints were placed on the bases which possessed imino-imino and/or imino-amino connectivities in the 2D $^1$H-$^1$H NOESY spectrum and therefore adopted the canonical Watson-Crick base pairing. The constraints were placed with large error bounds as these protons are prone to exchange with the solvent thus potentially leading to distorted peak intensities. Depending on the magnitude of the NOE intensity, the constraints were placed in two batches: ‘soft’ and ‘hard’ and those with weak peak intensities were placed in the soft category and a lower weight was placed on them during the structural calculation. The ‘hard’ constraints had a larger weight and thus greater emphasis placed on them in the structural calculation.

2.4.5 Database potential restraints

Database potential restraints\textsuperscript{34} were included in the refinement step to compensate for the lack of exchangeable proton based NOE constraints and to encourage the bases to adopt the canonical Watson-Crick base pairing. The non-bonded contact terms applied in the simulated annealing and refinement steps of the structure calculation can have a major impact on the generated structures, particularly since the density of short inter-proton distances in RNA is quite limited. These restraints improve the coordinate accuracy of the NMR structures. The restraints consist of two sets of statistical potentials derived from high resolution (2-3Å) nucleic acid X-ray crystal structures. The first set is the nucleic acid torsion angle database potential consisting of various multidimensional surfaces describing various torsion angle correlations. It consists of sugar-phosphate backbone, sugar puckers and glycosidic bond torsional angles. The second set of restraints is the RNA specific base-base positioning potentials providing simple geometric, statistically based descriptions of sequential and non-sequential base-base interactions. They provide complete three dimensional positions of the two bases of a base pair defined by the Cartesian coordinates of three atoms on each base (purines: N7, N6/O6 and N3 and pyrimidines: C6, N4/O4 and O2).
The potentials seek to bias sampling during the refinement stage to physically resemble regions of the conformational space within the range of possibilities that are consistent with the NMR experimental constraints. Significant improvements can be made in the accuracy of the NMR structures of the nucleic acids by incorporating the two potentials with the non-bonded interactions.

### 2.5 Structure calculation protocol-Antibiotics

#### 2.5.1 Overview

The XPLOR-NIH software package was chosen to perform the structure determination calculations for the antibiotics. There are five files which are required to generate the NMR structure (Figure 2.5.1). All the files were converted to the XPLOR-NIH format prior to structure calculation.

1. The standard topology file (topallhdg.dna) was obtained from XPLOR-NIH and modified for each antibiotic. This file incorporates specific information regarding the molecule; such as atom masses, General Amber Force Field (GAFF) atom types, charges, bonding, hydrogen bond donors and acceptors and improper angles.

2. The GAFF parameter file. This file includes the ideal bond lengths, angles, dihedral and improper angles and defines the energy penalty of deviation for each of the variables.

3. The PDB coordinate file. This file is generated from the pdb file of the starting structure and defines the antibiotic’s atoms in the 3D space. It is employed as a reference structure to promote correct stereochemistry of the generated structures.

4. The Psf structure file. This file is similar in information content to that of the topology file, but the information is molecule specific.
(5) The experimental ROE and dihedral angle constraint files.

Figure 2.5.1: The general scheme outlining the various files required to initiate the structure determination calculation of the antibiotics.

After creating the above files, the structure calculation can begin. Initially, a large number of random structures is produced, which are subjected to a short period of restrained molecular dynamics (rMD). This is followed by simulated annealing, before several refinement steps are carried out. The structures are then accepted or rejected based on whether they violate the ROE and dihedral angle constraints, the ideal bond lengths, angles, dihedral and improper angles and how well they overlay upon each other. If the structures are rejected, the experimental constraints and possibly the structure determination script are reviewed and any modifications applied.
2.5.2 Starting structure generation

The starting coordinates of the antibiotics were generated by running a B3LYP $^{97}$ 6-31G* energy minimisation on Gaussian 03$^{98}$ and Gabedit 2.0.7.$^{99}$ They were converted before incorporating into the XPLOR-NIH program. Charges were obtained from an MP2 $^{100}$ 6-31G* single point energy calculation to be used in the topology file, along with the Merz-Singh-Kollman scheme (MK or ESP) of population analysis.$^{101}$

2.5.3 Randomisation

The structure calculation began with the randomisation step producing 100 random coordinate structures of the antibiotic using the Distance Embedding $^{102}$ method. Along with interproton distances calculated from NOE intensities, it incorporates covalent bond distances, bond angles and coupling constants to produce approximate structures of the molecule.

The above structures were then put through a short period (4.5ps) of restrained molecular dynamics simulation at a temperature of 1500K. This was divided into three parts; during the first 1.5ps, the weights of the covalent bond and bond angle terms were increased from 0 to 0.02, followed by a 1.5ps simulation and increasing the weights to 0.05 and the final 1.5ps simulation increased the weights of the same two terms to 0.1. The NOE constraints also had their scale increased from 1 to 5, while improper angles and electrostatic terms were set to zero. The weights were gradually raised in order to induce convergence within the structures and van der Waals energy, electrostatic and improper angles were omitted to allow the atoms to move through each other to freely facilitate the sampling of conformational space and to adopt stable arrangements.

Although the structures are globally correct, the local structural elements will contain violations as the input information is incomplete, since it lacks torsional angles and hydrogen bonding information and therefore undergoes the simulated annealing step.
2.5.4 Simulated Annealing

Each of the random structures was subjected to a simulated annealing period with the temperature set to 1500K. The high temperature stage was set to a total of 100,000 steps and the slow cooling stage consisted of 3,000,000 steps with the time interval at 0.001ps. First, an energy minimisation step was run consisting of 100 steps, minimising the energies of the covalent bonds, ROE and dihedral angle constraints. This was commenced by a longer energy minimisation section consisting of 1000 steps and now including the bond angles. The rMD step followed with a total of 10,000 steps (100ps) minimising all the previous terms and now including the improper angles. The weight placed on the NMR experimental constraints (ROE and dihedral angle) and improper angles were slowly incremented (0.2→0.4→1.0) and the weight of the van der Waals interactions was set very low (0.003).

The simulated annealing step was implemented by steadily decreasing the temperature from 1500K to 300K in steps of 50K. The scale of the dihedral angle constraints was increased from 5 to 20 during this period. Once the temperature reached 300K, the weights of the non-bonded repulsive terms were modified; electrostatic potentials slightly decreased from 0.9 to 0.75, in contrast to the van der Waals weight which was increased from 0.003 to 4.0. Finally, a long energy minimisation consisting of 100K steps was carried out.

2.5.5 Refinement

The resulting annealed structures were subjected to a second set of rMD consisting of three refinement steps. The ROE and dihedral angle constraints were established with increasing weight after each step. The first refinement step commenced with an energy minimisation of 500 steps followed by rMD for a short period (0.01ps) at 1000K. Next, another short period of rMD (7.5ps) was executed with the dihedral angle scale steadily increasing from 5 to 45, in steps of 5.
A 500 step energy minimisation and a 7.5ps rMD at 1000K initiated the second refinement step. Followed by a 2.5ps simulated annealing period (5000 steps of 0.0005ps time duration) whereby the temperature was reduced from 1000.1K to 300K in steps of 25K. A long energy minimisation (10K steps) concluded the second refinement step.

The final refinement step was a lengthy energy minimisation (100K steps) which included the charges obtained from the MP2 calculation and the non-bonded energy terms (electrostatic and van der Waals). The simulations were followed by a cooling step to refine the structures and the total number of steps was equal to 50,000 with the step size equal to 0.5fs.

2.5.6 Acceptance and Validation
The refinement structures were used for analysis and validation. The Acceptance script contained violation cut-offs and therefore was able to discard structures which contained large violations on ROE and dihedral angles constraints, improper angles, bond lengths and bond angles. The final accepted structures were subjected to a final energy minimisation in order to produce an average structure.

2.5.7 Analysis
The accepted structures were assessed manually and those displaying incorrect stereochemistries or bonding were discarded. The final set of structures was evaluated by several criteria. The calculated structures must be consistent with the experimental data, possess low energies and exhibit a good overlay. A balanced judgement of the quality of the final structures was reached by taking into consideration the constraint violations, geometric accuracy and precision. The energy minimised average structure along with the overlaid structures represented the NMR structure of the antibiotic.

The overall precision among the set of NMR structures was measured by the Root Mean Square Deviation\(^{103}\) (RMSD) method. The RMSD demonstrates how similar the
individual fits are to one another, indicating the degree of variation for each arrangement.

2.6 Structure calculation protocol-RNA motifs

2.6.1 Overview

The aim of NMR structure determination is to achieve the minimum energy conformation of a RNA molecule using experimental NMR constraints, covalent geometry and non-bonded contacts. The structure calculation was run on the Python interface of the XPLOR-NIH software and standard topology and parameter files for nucleic acids were included in the program and have very similar information to that of the antibiotics (section 2.5.1). The topology file contains information on the atom types, charges, mass of atoms and the location of bonds and parameters. Similarly, the parameter file contains the ideal bond lengths, angles, dihedral and improper angles. Specific files for the B.subtilis 27mer RNA were constructed consisting of the NMR geometrical constraints and a sequence file. The software used distance geometry (DG) and rMD to calculate the three-dimensional structure.

2.6.2 Randomisation and Simulated Annealing

The initial temperature was set at 3500K and run with an energy minimisation step of 12.5ps for a period of 800ps. During this step, only the non-bonded phosphorus – phosphorus terms were incorporated to allow better sampling of the conformational space. The non-bonded repulsion energy constants were ramped up from 0.004 to 4. This allowed the phosphorus atoms to come close to each other and the other atoms to move freely in the three-dimensional space. The non-exchangeable proton NOE constraints were placed with a greater weight and increased from 2 to 30, while those from the exchangeable protons and the hydrogen bond restraints were ramped from 0.2 to 30. The scale of the dihedral and improper angles was inclined from 0.4 to 1.0 and 0.1 to 1.0, respectively. The parameters were ramped to prevent the RNA from being trapped in local minima. Following this, a torsional angle simulated annealing step was applied. The initial temperature at 3500K was dropped 12.5K every 0.2ps until the final temperature of 25K
was reached. The scale of the dihedral angle restraints was increased to 200. This was succeeded by a torsional angle energy minimisation step. Subsequently, a Cartesian-space simulated annealing step (same set up as the torsional angle simulated annealing) and energy minimisation was applied.

The simulated annealing section produced 100 structures. The top 20 structures were selected on the basis of lowest overall all atom average RMSD, lower energies and minimal constraint violations. The violations stemming from these structures were examined and error bounds adjusted. Any corrections applied to the input files resulted in the simulated annealing part of the structure calculation to be repeated. Once the simulated annealed structures satisfied the majority of the violations and ideal RNA parameters, the structure with the lowest all atom average RMSD was selected for the refinement stage.

2.6.3 Refinement
This section of the calculation yielded 100 refined structures. The same constraints input in the randomisation and simulated annealing stage were used with the addition of database orientation restraints to aid the alignment of the canonical Watson-Crick base pairs. Initial energy minimisation (20 steps) was followed by initial high temperature steps for the duration of 10ps. The dihedral angle constraints scale was set to 200. A torsional angle simulated annealing calculation was applied with the initial temperature set to 2000K and decreased in steps of 25K every 0.2ps until the final temperature of 25K was reached. A torsional angle energy minimisation and a Cartesian space energy minimisation were performed to produce the final structures.

2.6.4 Acceptance and Final structure
An ensemble of the top ten converged structures satisfying zero constraint violations, correct stereochemistry and low energies were selected from a pool of 100 refined structures. The same selection procedure as applied in the simulated annealing section was carried out to extract the best group of structures. RMSD values were calculated for the
ten best structures and the final structure was chosen which possessed the lowest value. As well as the average RMSD values calculated for the entire RNA motif, individual sections such as the flexible loops were also measured for comparison with the stem regions of the molecule.

2.6.5 Conformational analysis

The average structure was subjected to a full geometrical analysis using the web-based 3DNA\textsuperscript{104} program and the internal loop parameters monitored by the CURVES\textsuperscript{105} program. The 3DNA package uses the coordinate files of structures entered in the RCSB PDB database and describes the local and global conformations of the nucleotides. It represents the nucleic acid base pair geometry and the sugar-phosphate backbone \textit{via} a reference frame and hence defines the complete position and orientation of each base pair in the structure (Figure 2.6.1). Each base pair is uniquely characterised by a set of six rigid body parameters, Shear, Stretch, Stagger, Buckle, Propeller and Opening which describe the geometry of the complementary base pairs. The relative position and orientation of successive base pair planes can be described by the six dimer reference frames (Shift, Tilt, Slide, Roll, Rise and Twist) or it can be represented by the four helical frames (x-displacement, Inclination, y-displacement and Tip).\textsuperscript{106}
Chapter 2- Materials and methods

**Figure 2.6.1**: Base pair parameters describing the geometry of base pairs and sequential base pairs. Base pair reference frame (black) determines the parameters in the conformational analysis of the nucleic acid structure. The base pair is represented by the two cuboids. Base pair axis parameters are highlighted in green, local base pair parameters in blue and helical base pair parameters in orange.

The final structure is uploaded as a pdb file on the web based 3DNA program and the output analysis is presented in a tabulated format and as a diagrammatical representation illustrating the orientations of the bases and backbone nucleotides.

The geometrical parameters of the non-base paired hairpin loop of the *H. hal* 25mer rRNA were obtained from CURVES. It is an algorithm-based program which calculates the helical parameters of an irregular nucleic acid segment with respect to an optimised global helical axis.
Chapter 2 - Materials and methods

2.7 Quantitative measurement and binding constant determination

2.7.1 Introduction

NMR Spectroscopy is a powerful technique for understanding the binding affinity between a ligand and a macromolecule in the solution state. Changes in the target receptor binding to the ligand can be monitored by the differences in chemical shifts, line widths of the resonances and NOE intensities. The binding site can be deduced and the affinity between a ligand and receptor can be measured and represented as a binding constant. The binding constants are a special case of the equilibrium constant $K$ and represent the balance between the ligand and the receptor in their free and bound states as a complex (2.7.1).

$$
[L] + [R] \xleftrightarrow{k_1}{k_{-1}} [LR]
$$

$L= \text{Ligand}$

$R=\text{Receptor}$

The ratio of the rate constants for the association and dissociation is equal to the equilibrium constant ($K$) of the reaction. The two molecules must collide in order to react and therefore the rate of the reaction is proportional to the product of their individual concentrations ($[L]$ and $[R]$) (2.7.2).

$$
K_D = \frac{k_{-1}}{k_1} = \frac{[L][R]}{[LR]}
$$

The dissociation constant ($K_D$) is a measure of the propensity of a complex $[LR]$, dissociating into the separate free $[L]$ and $[R]$ components. As represented in the equation 2.7.2, the smaller the magnitude of the dissociation constant, the stronger the binding between the ligand and the receptor. It is the representation of the affinity between the
ligand and the receptor. These affinities are influenced by non-covalent intermolecular interactions, such as hydrogen bonding, electrostatic interactions and van der Waal contacts. The $K_D$ can be defined in two ways; thermodynamically by relating the concentration of the free and bound states of the two molecules and kinetically by assessing the association and dissociation rates.

### 2.7.2 Two-state processes

The types of a two site reaction are demonstrated in Schemes 1 and 2.

**Scheme 1**

\[ [L] + [R] \rightarrow [LR] \]

Scheme 1 is a simple model, where the reactants and the complex are present in 1:1 ratio. Unfortunately not all systems reflect this simple biomolecular binding reaction, instead scheme 2 is a realistic model for a two state process.

**Scheme 2**

\[ [L] + [R] \xrightleftharpoons[k_1]{k^{-1}} [LR]_{nonspecific} \xrightarrow[k_2]{k^{-2}} [LR] \]

Scheme 2 illustrates a non-specific binding of the ligand to the biomolecule followed by a slow conformational change leading to the complex. There may be a slow exchange of one resonance with compared to faster exchange of the other resonances. For example, an aromatic group on the ligand may induce slow conformational unstacking of the base pairs following rapid non-specific binding of the ligand and receptor.$^{107}$

However, ligands may have multiple binding sites with different affinities, where the binding to one site may change the likelihood (affinity) that the ligand binds to the other related sites. The binding of one site may enhance the affinity of a second binding site to the same ligand or it may produce a barrier for binding.
2.7.3 Binding constant calculation

The dissociation constant $K_D$, characterises the ligand binding in terms of the concentration of ligand at which half of the receptor binding sites are occupied. The solution conditions (temperature, pH and buffer concentration) can significantly alter the strength of intermolecular interactions holding a complex together.

The 2.7.2 equation is expanded to calculate the concentration of the bound species (2.7.3):

$$[RL] = \frac{[L][R]}{KD} \quad 2.7.3$$

When $K_D = [L]$, 50% of R is bound and so the binding constant is interpreted as the ligand concentration that leads to 50% occupancy of the receptor’s binding site. The lower the value of $K_D$, the less the ligand concentration is required to achieve 50% occupancy, indicating a higher affinity (stronger binding) between the receptor and ligand. The $[RL]$ is observed and it is the concentration of the free ligand [L] which is calculated, producing a square hyperbola binding isotherm (Figure 2.7.1).108

![Figure 2.7.1: A square hyperbola representing the relationship between the concentration of the bound ligand [RL] and free ligand [L] to reveal the $K_D$. The dissociation constant is measured in molar quantity (M) corresponding to the concentration of the ligand at which 50% of the binding sites on the receptor is occupied. The binding strength ranges from $10^{-3}$M (weak) to $10^{-6}$M (medium) to strong binding at $10^{-9}$M.](image-url)
The equation can be applied instead of measuring the concentration of the bound ligand which is not known, one can measure the change in chemical shift ($\Delta\delta$) as the population of the bound ligand increases. So, instead of plotting $[RL]$ vs $[L]$, a graph is produced where the $\Delta\delta$ vs $[L]$ is plotted.

**Figure 2.7.2:** Binding isotherms produced by monitoring the change in chemical shift of the bound ligand or receptor with respect to the free ligand concentration. As a greater volume of ligand is titrated into the receptor, the binding is likely to increase, causing a change in the chemical shift value.

The type of binding can be deduced from the binding isotherm in Figure 2.7.2. High affinity complexes will have strong binding, hence creating the greatest chemical shift change and the plot will resemble a square hyperbola. If only certain protons are involved in the binding then it is a specific binding and a change in chemical shift is not as great but the plot still resembles a square hyperbola. Non-specific binding arises when the majority of the protons are involved in the binding and so there is small but similar chemical shift changes producing a linear plot. However, with a weak binding (low affinity) system, the ligand will exchange rapidly between the free and bound states; existing in the two states for a very small period of time and consequently a large excess of ligand is required to drive the reaction to completion. There is likely to be no measurable change in the chemical shift.
The change in line width of a resonance as a function of ligand concentration at constant receptor concentration can be used to determine the binding constant.

### 2.7.4 Line width analysis

The 1D $^1$H NMR resonances can provide information on the type of exchange system of the complex studied. The appearance of the NMR spectrum of the complex will depend on whether chemical exchange of the ligand between the free and bound state is slow or fast.

A strongly bound complex is a slow exchanging system with respect to the NMR timescale (lifetimes of the individual species are tens of seconds to minutes). As a result, the NMR spectrum will reveal resonances belonging to the free and bound species. If the concentration of the receptor is kept constant and the ligand increased, then the peaks corresponding to the bound receptor will increase in intensity and simultaneously the free receptor peak will diminish in intensity. The bound receptor peak will increase until a 1:1 ratio (receptor: ligand) is reached and the system is saturated. All of the receptor sites are occupied and any additional ligand will simply result in an increase in the intensity of the free ligand peaks.

In the case of a weak binding fast exchange (lifetimes in $\leq 10^{-3}$s) system, the chemical environment experienced by a receptor proton will be averaged over the environments of the free and bound states, producing an NMR singlet at the frequency that is a weighted average of the time spent in the two states.

$$\delta_{\text{observed}} = \alpha_{\text{free}}\delta_{\text{free}} + \sum \alpha_{\text{bound}}\delta_{\text{bound}} \quad \text{(2.7.4)}$$

The relationship outlined in 2.7.4 is the chemical shift value ($\delta$) observed for a resonance of a fast exchanging system and is devised from the components of the ligand ($\alpha$) in its free and bound state and the corresponding chemical shifts of the free and bound resonances of the ligand.
Three possible situations are observed with fast exchange systems. The first situation is both the free and bound resonances can be measured independently; this is very rarely the case as the nucleus is rapidly interconverting between the two states. The second situation is where only the resonances belonging to the free molecules are clearly observed. The receptor peak is shifted very slightly (<0.1ppm) and the difference in line intensity is very small. Additional evidence for binding can be proved if the free ligand peaks are absent and the bound resonances and the $K_D$ can be determined. The final situation is that neither the free nor bound resonance may be observed and therefore a computational program must be used.

The line width plots are measured from high resolution 1D $^1$H NMR spectra obtained with a large number of scans. Phasing was optimised to achieve the maximum absorption mode line shapes of well-resolved, non-overlapping resonances. The line widths were evaluated manually for the isolated peaks and the values obtained can be used to calculate the transverse ($T_2$) relaxation rates for the individual protons. The plots were constructed for both the antibiotic and the RNA at increasing ligand concentration. Analysis of the plots will provide an understanding of how the binding affects the two molecules; similar trends in both the ligand and RNA line widths will suggest a non-specific binding in which a range of protons is involved.

### 2.7.5 NOE growth

NOE growth curves are an important method which can be implemented to analyse the binding between a ligand and a receptor. The NOE arises from the cross-relaxation rate between two spins, demonstrating that the rate of the NOE growth is a determination of how quickly the enhancement is transferred between the two spins, directly relating this with the internuclear distance.

The NOE intensity increases linearly with the mixing time ($\leq \tau_m= 250$ms) reaching a maximum, before decreasing due to spin diffusion. To compute the unknown interproton distances of two nuclei, it is fundamental that the linear region of the curve is obtained.\textsuperscript{60}
A series of 2D $^1$H-$^1$H NOESY spectra was acquired with different mixing times ($\tau_m = 25, 50, 100, 200$ and $400\text{ms}$) of the *B.subtilis* 27mer RNA-bamicetin complex (1:10) (described in section 4.2). The small mixing time values are important as here the NOE intensity will grow linearly producing a slope where the cross-relaxation is correlated to the interproton distance. NOE build-up curves were generated for protons of the antibiotic and the RNA.
Chapter 3  NMR structure determination of the antibiotics

The PT inhibitor antibiotics act at the site of protein synthesis in the ribosome of the microorganism and are thus very effective in preventing the bacteria from developing or multiplying. Therefore, three antibiotics; Bamicetin (C$_{28}$H$_{40}$N$_{6}$O$_{9}$, Mw= 604.66Da), Sparsomycin (C$_{12}$H$_{19}$N$_{3}$O$_{5}$S$_{2}$, Mw=349.44Da) and Anisomycin (C$_{14}$H$_{19}$N$_{1}$O$_{4}$, Mw = 277.31Da), from the Streptomyces microorganism family, with strong PT inhibitor potential were studied in this research project. A description of the NMR assignment, derived from a range of multidimensional experiments and structures determined using molecular modelling calculations for each antibiotic is presented in this Chapter. There are no X-ray crystal structures of these antibiotics in their uncomplexed states, making it all the more necessary to determine their solution structures via NMR spectroscopy.

The NMR structure of amicetin has been solved before and studies with bamicetin, which possesses a near identical chemical structure were undertaken to see whether there are similar structural features shared between the two antibiotics and to gain an understanding as to why bamicetin is a stronger PT inhibitor antibiotic than amicetin. The bamicetin antibiotic was chosen for binding studies with a conserved 27mer motif of B.subtilis 23S rRNA which are described in Chapter 4.

Sparsomycin is one of the most potent PT inhibitor antibiotics, active against all three evolutionary domains; therefore it must recognise a universally conserved functional site in the PT centre. A number of analogues of the antibiotic have been used for binding studies and positive results suggest that it may be a promising antitumour agent for clinical use. Determining the NMR structure will provide information concerning which regions of the antibiotic are involved in interactions with the ribosomal RNA of the peptidyl transferase centre.

The dynamic features of the individual antibiotics were considered and an analogous thermodynamic effect shown by anisomycin was also investigated. The chapter concludes with a collective discussion of the NMR structures of the three antibiotics.
3.1 Bamicetin

The chemical structure of bamicetin contains two pyranose rings, a nucleic acid cytosine base, a phenyl ring and an α-methyl substituted serine amino acid moiety and is shown in Figure 3.1.1.

\[\text{Figure 3.1.1: The chemical structure of bamicetin. Hydrogen atoms are labelled by Arabic numerals, amide groups and the two saccharide rings (I and II) are numbered with Roman numerals. The inequivalent methylene protons present on the pyranose rings are labelled with lower case } a \text{ (axial) and } e \text{ (equatorial) respectively, and the oxygen atoms are labelled alphabetically. This labelling scheme is used throughout in the assignment of the NMR spectra.}\]

Figure 3.1.2 displays the assigned 1D $^1$H NMR spectrum of bamicetin in $^1$H$_2$O and is also summarised in Table 3.1.1. Protons in different chemical environments resonate in the characteristic chemical shift regions, the methyl protons producing intense singlet peaks are the most shielded and appear high field (1-3ppm) in comparison to the deshielded, low field aromatic protons (7-8.5ppm) and the retarded amide (V) proton at 10.70ppm.
Figure 3.1.2: 600MHz 1D $^1$H NMR spectrum of bamicetin (5.5mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 4.8 at 2°C. The labile amide proton (V) appears as a broad resonance at 10.7ppm due to rapid exchange with the water solvent protons.

3.1.1 NMR assignment of Bamicetin

The 2D scalar correlated experiments $^1$H-$^1$H DQF-COSY, $^1$H-$^1$H TOCSY and $^1$H-$^{13}$C HSQC were used to unambiguously assign the individual spin systems, followed by a dipolar $^1$H-$^1$H ROESY experiment to confirm the assignment and identify connectivities between protons in different parts of the antibiotic. The NMR assignment began by identifying the low field shifted resonances H7 and H13 appearing at 5.18ppm and 5.74ppm, respectively (Figure 3.1.2). The anomeric protons are adjacent to the electronegative oxygen atom of the saccharide ring thus shifting the respective resonances lower field. Relayed coherence transfer between coupled protons resulted in scalar connectivities to be observed in the $^1$H-$^1$H TOCSY spectrum between the H7 and protons on the first saccharide ring H4 and H6 (not shown). Further confirmation of the anomeric proton assignment was obtained in the analogous $^1$H-$^{13}$C HSQC-TOCSY spectrum which displayed through bond correlations between C13 and the methylene protons; H11$a$, H11$e$, H12$a$ and H12$e$ (Figure 3.1.3), together with the one bond C-H correlation to H13.
Figure 3.1.3: 600MHz $^1$H-$^{13}$C HSQC-TOCSY ($\tau_m=75$ms) spectrum of bamicetin (5.5mM) in 100% $^2$H$_2$O, pH 4.8 at 25°C. The spectrum shows the assignment of the anomeric carbons C7 and C13 with their directly attached protons H7 and H13 respectively, highlighted in red, in addition to remote protons of their spin systems.

The next assignment consisted of the four intense and well resolved resonances in the high field aliphatic region (1-3ppm) (Figure 3.1.2). On the basis of their chemical shift values and line intensities they were reliably assigned to the four methyl group protons (1, 3, 9 and 18) of bamicetin as described below. The methyl 1CH$_3$ is attached to an amine functional group and experiences a similar deshielding effect to the anomeric protons and thus it appears lower field at 2.70ppm. The two signals at 1.35ppm and 1.28ppm were identified as 3CH$_3$ and 9CH$_3$ situated on the saccharide rings (I and II). This assignment was made possible by the observation of scalar connectivities between the methyl group to remote protons in their individual spin systems in the $^1$H-$^1$H TOCSY spectrum (Figure.3.1.4). This experiment yielded long range couplings, the magnetisation was relayed so effectively that up to 6 bond scalar connectivities ($^6J_{H3-H6}$, $^6J_{9CH3-H12a}$, $^6J_{H7-H12e}$) were observed.
Figure 3.1.4: 600MHz $^1$H-$^1$H TOCSY ($\tau_m$=75ms) spectrum of bamicetin (5.5mM) in 100% $^2$H$_2$O, pH 4.8 at 25°C. The methyl groups 3CH$_3$ and 9CH$_3$ show multiple bond $^1J_{H-H}$ scalar couplings to all of the protons except the anomeric protons (H7 and H13).

The $\alpha$-methylserine region of the antibiotic consists of a methylene group with enantiotopic protons H19a (3.83ppm) and H19b (4.14ppm). The electronegative hydroxyl group (OH) and the adjacent bulky -NH$_2$ and -CH$_3$ groups render H19a and H19b non-equivalent protons. The methyl 18CH$_3$ possessed strong ROE couplings to both H19a and H19b as well as to H16 of the phenyl group (Figure 3.1.5).

Figure 3.1.5: 600MHz $^1$H-$^1$H ROESY ($\tau_m$ = 400ms) spectrum of bamicetin (5.5mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 4.8 at 2°C. The plot shows the dipolar couplings observed between the methyl 18CH$_3$ and the phenyl H16 proton.
A number of dipolar connectivities between the cytosine proton H14 and protons of the saccharide ring (II) were observed in the $^1$H-$^1$H ROESY spectrum, displayed in Figure 3.1.6. The correlations included long distance ROEs to H8 and 9CH$_3$ which are indicative of the spatial proximity between the pyranose ring and the cytosine base. The correlations provided important structural information which was input as distance based constraints in the structure determination calculation. Of interest is the intensity of the H12$a$-H14 connectivity which is greater than that belonging to H12$e$-H14, demonstrating clearly that the correlations are indicated through space and, in this instance, H12$a$ is closer to H14.

Figure 3.1.6: 600MHz $^1$H-$^1$H ROESY ($\tau_m = 400$ms) spectrum of bamicetin (5.5mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 4.8 at 2°C. The plot exhibits the dipolar couplings observed for the cytosine proton (H14) and the protons on the second saccharide ring (II).
The aromatic protons of the cytosine (H14, H15) and phenyl ring (H16, H17) exhibit the characteristic resonance patterns in the low-field (7-8.5ppm) region of the 1D $^1$H NMR spectrum. Inspection of the line integrals enabled the identification of these protons (Figure 3.1.2). The aromatic group contains two sets of identical H16 (7.64ppm) and H17 (7.92ppm) protons and so, instead of four resonances, they produce two, one for each pair of protons. The resonances represent two protons each; and this is depicted in the area of the peaks which are twice the area of the peaks of H14 (8.23ppm) and H15 (7.50ppm), which denote one proton each. The cytosine ring protons H14 and H15 were assigned using the $^1$H-$^1$H ROESY spectrum, in which only proton H14 produced a ROE correlation to H13, therefore confirming its assignment (Figure 3.1.6). The H16 (7.64ppm) and H17 (7.92ppm) protons demonstrate strong scalar correlations to each other in the $^1$H-$^1$H TOCSY spectrum and, in addition to the ROE connectivity between H16 and 18CH3, in the $^1$H-$^1$H ROESY ($\tau_m$=200 and 400ms) spectra. To further confirm the assignment of the H16 and H17 protons, a 2D $^1$H-$^13$C HMBC experiment was carried out (Figure 3.1.7a). This experiment reveals only scalar couplings between $^{13}$C and $^1$H nuclei separated by 2-3 bonds and one-bond HSQC couplings are absent. The carbonyl adjacent to the H16 should reveal connectivity in the spectrum due to being three bonds away but no coupling between the carbonyl carbon and H17 would be observed as they are removed.

**Figure 3.1.7:** (a) 400MHz 2D $^1$H-$^13$C HMBC spectrum of bamicetin (5.5mM) in 100% $^2$H2O, pH 4.8 at 21°C. The three bond scalar coupling between the H16 proton and the adjacent carbonyl carbon is observed and confirmed in (b) 1D $^{13}$C {$^1$H} NMR spectrum of (5.5mM) in 100% $^2$H2O, pH 4.8 at 25°C.
The presence of the three bond coupling between the H16 proton of the aromatic ring and the adjacent carbonyl carbon in Figure 3.1.7 unambiguously verifies the assignment of the H16 and H17 protons.

The saccharide protons were assigned predominantly with the aid of the $^1$H-$^{13}$C HSQC-TOCSY experiment, which, in addition to the direct $^{1}J_{H,C}$ coupling (HSQC), also relays coherence transfer to the remaining protons of the coupled spin system via the TOCSY segment of the pulse sequence. It thereby allowed scalar connectivities between protons of each pyranose ring spin system to be identified. For example, Figure 3.1.8, C11 (28.9ppm) displays correlation to its directly attached geminal protons H11a (1.60ppm) and H11e (2.43ppm), as well as to remote protons H8, H10 and H12a, H12e and H13 within the same spin system.

![Figure 3.1.8: 600MHz $^1$H-$^{13}$C HSQC-TOCSY ($\tau_{m}=75\text{ms}$) spectrum of bamicetin (5.5mM) in 100% $^2$H$_2$O, pH 4.8 at 25°C. The plot displays the assignment of the pyranose carbons C2, C9, C11 and C12 and directly attached protons; H2, H10, H11a/H11e and H12a/H12e, respectively, highlighted in red in addition to scalar coupling to remote protons in the saccharide ring.](image)
To verify the assignment obtained from the $^1$H-$^13$C HSQC-TOCSY spectrum, and to identify through space correlations of the saccharide protons, a $^1$H-$^1$H ROESY spectrum (Figure 3.1.9) was measured and interpreted. A number of intersaccharide dipolar correlations were observed, including 9CH$_3$-H4, H8-H4 and H11e-H6. The ROESY spectrum also revealed connectivities between the methyl 1CH$_3$ and protons on the first saccharide ring (I). The two bond coupling constants were measured for the methylene protons; H11a, H11e and H12a, H12e from the 1D $^1$H NMR spectrum. The magnitude of the coupling constants $^2J_{11a-11e} = 12.6$Hz, $^2J_{11e-11a} = 12.2$Hz, $^2J_{12a-12e} = 12.9$Hz and $^2J_{12e-12a} = 12.9$Hz are consistent with expected values and confirmed that the protons are geminal pairs.

Figure 3.1.9: 600MHz $^1$H-$^1$H ROESY (τm=400ms) spectrum of bamicetin (5.5mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 4.8 at 2°C. Important distance based constraints were generated from the dipolar correlated cross peaks observed between protons on both saccharide rings.
It is acknowledged that ideally it would have been best to carry out the NMR assignment of bamicetin in the protonated states of the amines and for this charge densities for each nitrogen atom would have to be calculated. However, protonated amines exert local effects and so may not affect the conformation of the molecule in the solution state. The methodology applied in this project was based on previous published work carried out on amicetin\textsuperscript{43,68,69} in which the pKas of the primary and tertiary amino groups were 7.27 and 7.57, respectively and based on the chemical structural similarities it was predicted that the secondary amine groups of bamicetin will be of similar value. Also, NMR experiments were carried out on bamicetin at low pH (~ 5) since amicetin is known to degrade at alkaline pH.

The labile amide proton (NHV) at 10.70ppm (Figure 3.1.2) was assigned based on the broad low field shifted value and its absence in the 100\% \textsuperscript{2}H\textsubscript{2}O spectrum. No connectivities were observed to other protons of bamicetin, however, the assignment labelling was based on the NMR assignment of amicetin.\textsuperscript{44} The resonance was previously assigned by a natural abundance \textsuperscript{1}H-\textsuperscript{15}N HSQC experiment on amicetin and this assignment was transferred to bamicetin.
3.1.2 Variable temperature effects observed
To probe the dynamic effects of bamicetin in solution a series of 1D $^1$H NMR spectra were measured in the temperature range of 2-25°C. The resonances of the NHV proton, and the N-methyl (1CH$_3$) group protons attached to the first pyranose ring, were carefully monitored. Despite an increase in temperature (Figure 3.1.10), the NHV proton (panel A) demonstrates retardation to exchange with the solvent protons and the residual proton resonance is clearly visible even at 20°C. The retardation to solvent exchange suggests that the amide proton maybe involved in intramolecular hydrogen bonding interactions. The line width of the 1CH$_3$ group peak displayed in panel B decreases and the resonance sharpens due to the faster rotation of the single C-N bond resulting in the three protons of the methyl group being indistinguishable. The resonances of the aromatic or other aliphatic protons of bamicetin did not show any significant changes with respect to the temperature increase, thereby indicating that the bamicetin antibiotic is a thermodynamically stable structure.

![Figure 3.1.10: 600MHz 1D $^1$H NMR spectra illustrating the effect of increase in temperature; (a) 2°C, (b) 5°C, (c) 10°C, (d) 15°C, (e) 20°C and (f) 25°C, on the resonances of bamicetin (5.5mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 4.8. Panel A shows the exchange-retarded amide (V) proton and panel B displays the aliphatic amino methyl proton, both revealed changes to the chemical shift values and line widths of the peaks induced by the rise in temperature. C is the line width plot of the amide (V) proton and D illustrates the line width changes of the methyl proton.](image-url)
The change in linewidths is represented as graph plots in panels C and D. The line width of the amide proton (V) has increased three fold, from 28.1Hz to 93.1Hz. This is in contrast, to the methyl group (1CH₃) resonance which decreased from 6.8Hz to 3.2Hz, a reduction of nearly 50%.
Table 3.1.1: $^1$H, $^{13}$C chemical shifts ($\delta$, ppm) and coupling constants ($^3J$, Hz) of bamicetin. The values are based on the 600MHz 1D $^1$H NMR spectrum of bamicetin (5.5mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 4.8 at 2°C (Figure 3.1.2).

<table>
<thead>
<tr>
<th>Assignment</th>
<th>$^1$H Chemical shift $\delta$ (ppm)</th>
<th>$^{13}$C Chemical shift $\delta$ (ppm)</th>
<th>$^1$H Coupling constant $^3J$ (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.70</td>
<td>32.7</td>
<td>s</td>
</tr>
<tr>
<td>2</td>
<td>3.11</td>
<td>65.9</td>
<td>t, 10.5</td>
</tr>
<tr>
<td>3</td>
<td>1.35</td>
<td>19.6</td>
<td>d, 9.4</td>
</tr>
<tr>
<td>4</td>
<td>4.17</td>
<td>66.8</td>
<td>d, 6.1</td>
</tr>
<tr>
<td>5</td>
<td>3.99</td>
<td>69.7</td>
<td>t, 9.9</td>
</tr>
<tr>
<td>6</td>
<td>3.69</td>
<td>74.3</td>
<td>d, 9.9</td>
</tr>
<tr>
<td>7</td>
<td>5.18</td>
<td>97.2</td>
<td>d, 9.4</td>
</tr>
<tr>
<td>8</td>
<td>3.51</td>
<td>77.0</td>
<td>m, 8.5</td>
</tr>
<tr>
<td>9</td>
<td>1.28</td>
<td>20.5</td>
<td>d, 8.3</td>
</tr>
<tr>
<td>10</td>
<td>3.91</td>
<td>79.5</td>
<td>d, 11.8</td>
</tr>
<tr>
<td>11a</td>
<td>1.60</td>
<td>28.9</td>
<td>d, 12.6</td>
</tr>
<tr>
<td>11e</td>
<td>2.43</td>
<td>28.9</td>
<td>d, 12.2</td>
</tr>
<tr>
<td>12a</td>
<td>1.85</td>
<td>31.1</td>
<td>dd, 12.9</td>
</tr>
<tr>
<td>12e</td>
<td>2.12</td>
<td>31.1</td>
<td>d, 12.9</td>
</tr>
<tr>
<td>13</td>
<td>5.74</td>
<td>86.2</td>
<td>d, 11.9</td>
</tr>
<tr>
<td>14</td>
<td>8.23</td>
<td>148.7</td>
<td>d, 8.1</td>
</tr>
<tr>
<td>15</td>
<td>7.50</td>
<td>101.6</td>
<td>d, 8.1</td>
</tr>
<tr>
<td>16</td>
<td>7.64</td>
<td>124.3</td>
<td>d, 8.5</td>
</tr>
<tr>
<td>17</td>
<td>7.92</td>
<td>132.1</td>
<td>d, 8.4</td>
</tr>
<tr>
<td>18</td>
<td>1.61</td>
<td>22.2</td>
<td>-</td>
</tr>
<tr>
<td>19a</td>
<td>3.83</td>
<td>20.6</td>
<td>t, 6.7</td>
</tr>
<tr>
<td>19b</td>
<td>4.14</td>
<td>20.6</td>
<td>-</td>
</tr>
<tr>
<td>20*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>V</td>
<td>10.70</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(-) not observable. Singlet peaks are indicated by s and hence do not possess coupling values. Doublet, triplet and doublet of doublets are indicated by the symbols d, t and dd, respectively and the multiplet (m) value is an estimate. Proton 20 was not identified.
3.1.3 Experimental constraints

The distance-based ROE and scalar coupling constant-based dihedral angle constraints were obtained from the 2D $^1$H-$^1$H ROESY spectra and 1D $^1$H NMR spectrum, respectively, based on the methodology described in section 2.4. A total of 64 experimental constraints (Table 3.1.2) were used with the majority (47) being the ROE based distance constraints measured from 2D $^1$H-$^1$H ROESY ($\tau_m$=150 and 400ms) spectra. The dihedral angle constraints (12) derived from the coupling constants were important in securing the correct conformation of the generated NMR structures.

ROE connectivities observed between the saccharide-saccharide and/or saccharide-aromatic protons aided the orientations of different regions of the antibiotic. Some of these constraints were H4-H8, H6-H9 and H14-9CH$_3$.

The structure generation protocol outlined in section 2.5 was implemented to calculate the NMR structure of bamicetin. A set of 100 random structures was subjected to simulated annealing, followed by three sets of refinement calculations with increasing energy minimisation steps, to produce the NMR structure which was viewed in the Unix-based VMD software.$^{110}$

**Table 3.1.2:** The NMR experimental ROE distance and dihedral angle constraints input into the XPLOR-NIH program for the structure determination of bamicetin.

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ROE distances</strong></td>
<td>47</td>
</tr>
<tr>
<td><em>Strong:</em> 1.8-3.0Å</td>
<td>18</td>
</tr>
<tr>
<td><em>Medium:</em> 1.8-4.0Å</td>
<td>16</td>
</tr>
<tr>
<td><em>Weak:</em> 1.8-5.0Å</td>
<td>11</td>
</tr>
<tr>
<td><em>Very weak:</em> 1.8-7.0Å</td>
<td>2</td>
</tr>
<tr>
<td><strong>Dihedral angles</strong></td>
<td>17</td>
</tr>
<tr>
<td><strong>Total constraints</strong></td>
<td>64</td>
</tr>
</tbody>
</table>
3.1.4 **NMR solution structure of Bamicetin**

The energy minimised and ten best NMR structures of the bamicetin antibiotic are displayed in Figure 3.1.11. The structures were assessed and selected based on three criteria; low energy, low average all-atom RMSD and no constraint violations.

![Figure 3.1.11](image)

**Figure 3.1.11:** (a) The energy minimised NMR structure and (b) An overlay of the ten best structures of bamicetin antibiotic. The antibiotic adopts a folded stable conformation which is possibly held by three hydrogen bonds. The top ten structures of bamicetin possess an average all-atom RMSD value of 0.82Å.

The RMSD value, along with the comparable overlay of structures, manifested the flexibility of the saccharide rings and the serine moiety in comparison to the cytosine and phenyl rings, which are more rigid. The NMR structure is a stable, folded conformation with the serine amino acid residue tilted towards the saccharide groups.

The structure satisfied the experimental constraints as well as providing an explanation for the retardation of exchange of amide proton NHV which may be involved in hydrogen bonding with amine VI group.

The amino (NHI) situated on the first saccharide ring has potential to bind to the ribosome *via* hydrogen bond interactions with the electronegative nitrogens or oxygens of the ribosomal RNA bases and phosphate backbone. This may perhaps explain why bamicetin is a stronger PT inhibitor antibiotic than amicetin which does not contain such a proton at this position in the chemical structure.
The antibiotic contains a number of exposed heavy atoms situated on the peptide moiety, cytosine ring and the amide groups (NIV-Oe, NV-Of) which are potential hydrogen bond donors and acceptors to the peptidyl transferase rRNA.

The turn position of the cytosine and phenyl ring reveal functionally significant conformations for predicted interactions with the RNA motifs of 23S rRNA. The cytosine nitrogen (III) and oxygen (d) atoms are exposed to make potential hydrogen bond interactions with the bulged nucleotides of the RNA motif (i.e U6, A27, C28) (Figure 4.1.1).

<table>
<thead>
<tr>
<th>Proton group</th>
<th>Dihedral angle</th>
<th>Value (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st saccharide ring (I)</td>
<td>C5-C2-C3-Oa</td>
<td>58.3</td>
</tr>
<tr>
<td></td>
<td>Oa-C7-C8-C5</td>
<td>57.7</td>
</tr>
<tr>
<td>Glycosidic bridge C-O</td>
<td>C7-Ob-C8-C9</td>
<td>165.5</td>
</tr>
<tr>
<td>2nd saccharide ring (II)</td>
<td>C11-C8-C9-Oc</td>
<td>57.7</td>
</tr>
<tr>
<td></td>
<td>Oc-C13-C12-C11</td>
<td>58.9</td>
</tr>
<tr>
<td>Glycosyl C-N</td>
<td>Oc-C13-N2-C14</td>
<td>46.4</td>
</tr>
<tr>
<td>Cytosine and phenyl link</td>
<td>N4-Coe-C15-C16a</td>
<td>91.6</td>
</tr>
<tr>
<td></td>
<td>N4-Coe-C15-C16b</td>
<td>87.8</td>
</tr>
<tr>
<td>Methyl-serine moiety</td>
<td>N5-C0F-C18-C19</td>
<td>62.7</td>
</tr>
</tbody>
</table>

Figure 3.1.12: Table of dihedral angles which assist the antibiotic to adopt a bent stable conformation. The dashed lines on the NMR structure indicate the dihedral angles presented in the table.

The dihedral angles of the saccharide rings (-58.3°, -57.7°, -57.7° and 58.9°) displayed in Figure 3.1.12 show that both exist in stable chair conformations. The glycosyl bridge C7-Ob-C8-C9 dihedral angle (-165.5) reveals the H7 atom to be equatorial and the H8 axial. The H7 may have altered its position so that it can interact with the peptide moiety and maintain the folded conformation. The torsional angle between the cytosine and phenyl rings is -91.6°, indicating that the two rings are almost perpendicular to each other and this can be observed in the figure above. The H14 and H15 protons resonate lower field in the 1D 1H NMR spectrum than the cytosine H5-H6 proton resonances usually appear, and this is most likely due to the edge to face stacking between the cytosine and phenyl ring. This
increases the stability of these protons, shifting the resonances to lower field. The methyl
serine moiety (N5-COF-C18-C19) adopts a stable gauche torsional angle (-62.7°).

3.2 **Sparsomycin**

Sparsomycin is one of the most potent PT inhibitor antibiotics and so it is informative to
determine its solution structure and understand the reasons behind its potency.

![Chemical structure of sparsomycin](image)

**Figure 3.2.1:** The chemical structure of sparsomycin consists of a methylated uracil ring and a
sulphur rich ‘tail’ linked by an amide group. The hydrogen, carbon, nitrogen and oxygen atoms are
labelled with Arabic numerals and this scheme is used throughout the NMR assignment. R and R’
define the two separate regions of the antibiotic for the staggered conformations in Figure 3.2.5.

The chemical structure shown in Figure 3.2.1 consists of a methyl substituted uracil ring,
*trans*-olefenic group, an amide group, several methylene groups and two sulphur groups.
These groups are all potential interaction sites with RNA residues, as they contain
hydrogen bonding or van der Waals interaction capabilities.
3.2.1 NMR assignment of Sparsomycin

1D $^1$H NMR and several 2D NMR experiments were carried out to determine the assignment of the sparsomycin antibiotic. Sparsomycin contains three labile amide protons (2, 5, 9), and so these were assigned to the peaks appearing lower field in the 1D $^1$H NMR spectrum (Figure 3.2.2). As is characteristic of resonances of exchangeable protons, these peaks showed a change in chemical shifts and line intensities in the variable temperature (VT) spectra (discussed in 3.2.2).

![Figure 3.2.2](image)

**Figure 3.2.2**: The 600MHz 1D $^1$H NMR spectrum of sparsomycin antibiotic (4.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C.

The first resonances assigned were the two singlet peaks appearing highest field at 2.25ppm and 2.40ppm. The methyl 18CH$_3$ (2.25ppm) was assigned using the $^1$H-$^1$H ROESY spectrum with the observation of dipolar couplings to the methylene H16 and H17 protons (Figure 3.2.3). The lower field singlet at 2.40ppm was assigned to the methyl 4CH$_3$ group, supported by connectivities to trans-olefinic protons, H6 and H7 and is displayed in the ROESY spectrum, below.
Chapter 3 - NMR structure determination of the antibiotics

Figure 3.2.3: 600MHz $^1$H-$^1$H ROESY ($\tau_{m}$=400ms) spectrum of sparsomycin (4.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. The plot shows the through-space dipolar connectivities observed between the different protons of the antibiotic. The methyl 4CH$_3$ group was assigned due to the presence of correlations to nearby H6 and H7 protons, while 18CH$_3$ displayed couplings to H13/H14 in addition to H16 and H17 protons.

Olefinic protons appear low field, typically around 6-8ppm, depending on the chemical environments of the individual protons. Hence, H6 was assigned to the resonance at 7.08ppm and H7 at 7.42ppm, both peaks are scalar correlated and so split into intense doublets ($^3J_{H6-H7}$=15.6Hz) (Figure 3.2.2). The H6 and H7 protons show dipolar connectivities to 4CH$_3$ in the ROESY spectrum, with the more intense ROE from H7 (Figure 3.2.3). This demonstrates that the H4 proton is closer in space to H7 than it is to H6. The amide proton NH9 also exhibited couplings to H6 and H7 in the ROESY spectrum, with a much stronger connectivity to H6 (Figure 3.2.6). Figure 3.2.4 is the $^1$H-$^{13}$C HSQC spectrum showing the $^1J_{H-C}$ coupling used to assign the various methylene protons.
Figure 3.2.4: 600MHz $^1$H-$^{13}$C HSQC spectrum of sparsomycin (4.6mM) in 100% $^2$H$_2$O, pH 6 at 25°C. The spectrum displays the resolved $^1J_{C,H}$ couplings which distinguished the methylene group protons (H11/H12, H13/H14 and H16/H17).

Protons H10, H11 and H12 behave as an ABX system due to the neighbouring chiral atom (C10), where the geminal protons H11 (A) and H12 (B) appear non equivalent and are known as diastereotopic protons (Figure 3.2.5).
Figure 3.2.5: The three staggered conformations of the enantiomeric protons, R and R’ defined in Figure 3.2.1.

Two protons attached to a carbon atom which is adjacent to an asymmetric centre often exhibit non-equivalent environments. The reason for the steric disposition of groups in the various possible conformers, is demonstrated by the Newman projections in Figure 3.2.5. By keeping the position of the rear carbon stationary and rotating the carbon atom in front, one obtains the three staggered conformations. Two protons attached to a carbon adjacent to a chiral centre therefore often appear at different, but close, chemical shifts and exhibit scalar coupling with one another. The H11 and H12 protons possess close chemical shifts values but H10 (X) has a distinctly different chemical shift. Both protons H11 and H12 exhibit connectivities to H13/H14 and H10. Proton H11 also displays connectivities to the amide proton NH9 in the ROESY spectrum (Figure 3.2.6, below). The hydroxyl proton H10 has the potential to hydrogen bond to the carbonyl atom producing a closed ring-like structure thus stabilising this proton. H10 is involved in dipolar couplings with several neighbouring protons, NH9, H11, H12 and H13 in the ROESY spectrum and the resonance appears as a multiplet peak in the 1D NMR spectrum (Figure 3.2.6).

Protons H13 and H14 were assigned to the overlapped resonance peaks at 3.16-3.17ppm. A long range ROE was observed between protons H13/H14 and the exchangeable amide proton NH9, additional dipolar couplings were observed to H12 and H17 in the ROESY spectrum (Figure 3.2.6).
Protons H16 and H17 demonstrate strong coupling to each other. Proton H16 was assigned to the resonance at 3.93ppm and H17 at 4.11ppm and both protons exhibited strong scalar couplings in the TOCSY spectrum (not shown). Figure 3.2.6 illustrates the range of connectivities observed between the various protons in the ROESY spectrum.

Figure 3.2.6: 600MHz $^1$H-$^1$H ROESY ($\tau_m=400$ms) spectrum of sparsomycin (4.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. The spectrum illustrates a number of ROEs identified amongst the various protons of the antibiotic.

Two resonances appearing low field at 8.43ppm and 8.58ppm (Figure 3.2.2) were missing from the 1D $^1$H NMR spectra measured in $^2$H$_2$O solvent. These were assigned to the amide groups NH2/NH5 (8.43ppm) and NH9 (8.58ppm) of sparsomycin. The assignment was confirmed by the ROEs observed in Figure 3.2.6 between NH9 to H6, H10, H11 and H13/14 protons.
Table 3.2.1: $^1$H, $^{13}$C chemical shifts (δ, ppm) and coupling constants ($^3J$, Hz) obtained from the 1D $^1$H NMR spectrum of sparsomycin (4.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 40°C.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>$^1$H Chemical shift δ (ppm)</th>
<th>$^{13}$C Chemical shift δ (ppm)</th>
<th>$^1$H Coupling constant $^3J$ (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH2/5</td>
<td>8.43</td>
<td>-</td>
<td>s</td>
</tr>
<tr>
<td>4</td>
<td>2.40</td>
<td>19.7</td>
<td>s</td>
</tr>
<tr>
<td>6</td>
<td>7.08</td>
<td>123.6</td>
<td>d, 15.6</td>
</tr>
<tr>
<td>7</td>
<td>7.42</td>
<td>135.4</td>
<td>d, 15.7</td>
</tr>
<tr>
<td>NH9</td>
<td>8.58</td>
<td>-</td>
<td>d, 9.1</td>
</tr>
<tr>
<td>10</td>
<td>4.50</td>
<td>49.5</td>
<td>m, 5.1*</td>
</tr>
<tr>
<td>11</td>
<td>3.62</td>
<td>65.8</td>
<td>dd, 5.9</td>
</tr>
<tr>
<td>12</td>
<td>3.70</td>
<td>65.8</td>
<td>dd, 5.0</td>
</tr>
<tr>
<td>13</td>
<td>3.16</td>
<td>55.7</td>
<td>t, 7.1</td>
</tr>
<tr>
<td>14</td>
<td>3.16</td>
<td>55.7</td>
<td>t, 7.1</td>
</tr>
<tr>
<td>16</td>
<td>3.93</td>
<td>57.5</td>
<td>d, 14.0</td>
</tr>
<tr>
<td>17</td>
<td>4.11</td>
<td>57.5</td>
<td>d, 14.2</td>
</tr>
<tr>
<td>18</td>
<td>2.25</td>
<td>18.8</td>
<td>s</td>
</tr>
</tbody>
</table>

* obtained from the 1D $^1$H NMR spectrum in 100% $^2$H$_2$O, pH 6.2 at 5°C. Singlet (s) peaks do not possess coupling values. Doublet (d) triplet (t) and doublet of doublets (dd) peaks are indicated in the table. The values for the multiplet (m) is an estimate.
3.2.2 Variable temperature effects on the amide and methylene protons

A series of 1D $^1$H NMR spectra at different temperatures (Figure 3.2.7) were carried out to probe the dynamics of the sparsomycin antibiotic. The labile protons generally show change and so the resonances of the amide protons (NH2, NH5 and NH9) were monitored. Figure 3.2.7(a) reveals that two amide peaks, at 8.43ppm (NH2/NH5) and 8.58ppm (NH9) broaden and shift to high field as the temperature increased. The resonances were still visible at 50°C establishing that the NH protons are involved in strong intramolecular hydrogen bonding. Switching the temperature back to 2°C resulted in the original peaks returning, verifying that the induced changes are not permanent. The NH2/5 resonance shifted slightly to lower field but the peak intensity was very similar throughout the rise in temperature. In contrast the resonance of NH9 not only broadened significantly, but also shifted to the high field by a considerable amount ($\Delta\delta \approx 0.4$ppm). NH9 produces a doublet peak due to scalar coupling to the H10 protons.

![Figure 3.2.7: 1D $^1$H NMR variable temperature series of sparsomycin (4.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2-50°C. (a) The resonances of the labile amide protons NH2/5 and NH9 show pattern change in their chemical shifts and line intensities induced by the rise in temperature and (b) Changes to the multiplet resulting for the geminal protons 13/14H as the temperature is increased.](image)
Non labile protons usually show little dynamic change but the variable temperature series revealed a fascinating result with regard to the methylene H13/14 protons. The peak assigned to these geminal protons splits into a well defined pair of double doublets as the temperature is increased (Figure 3.2.7b). The geminal protons are attached to a carbon atom which is in a saturated system, the rotation barrier decreases as the temperature is increased, allowing the methylene protons to take part in $^3J$ coupling with nearby proton H10.

![Figure 3.2.8: 600MHz $^1$H-$^1$H ROESY (τ_m = 400ms) spectrum of sparsomycin (4.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 40°C. The plot shows the various ROE connectivities observed between the methylene protons (H13/H14) to the nearby protons.](image-url)
3.2.3 Experimental constraints
Following the above assignment, distance-based ROE constraints were extracted from the two $^1$H$^{-}$H ROESY spectra at individual mixing times ($\tau_m$= 250 and 400ms). The advantage of using two ROESY spectra is that one is able to identify, and if necessary discard anomalies which may have only appeared in one spectrum. Table 3.2.2 summarises the range of scalar and dipolar connectivities (27) observed between the various protons with the majority (18) of those converted into distance-based ROE constraints.

**Table 3.2.2**: A grid displaying the couplings obtained from the scalar based $^1$H-$^1$H TOCSY and the dipolar coupled $^1$H-$^1$H ROESY experiments.

<table>
<thead>
<tr>
<th>Proton</th>
<th>4CH₃</th>
<th>6</th>
<th>7</th>
<th>NH9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>4CH₃</td>
<td>x, W</td>
<td>x</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>x, M</td>
<td>x</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH9</td>
<td>x, M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>x, M</td>
<td>x</td>
<td>S</td>
<td>x, M</td>
<td>x, S</td>
<td>x, M</td>
<td>x, S</td>
<td>x, M</td>
<td>x, S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>x, W</td>
<td>x</td>
<td></td>
<td>x, S</td>
<td>x, W</td>
<td>x, S</td>
<td>x, W</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>x</td>
<td>x</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>x</td>
<td>x</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x, M</td>
<td>x, W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x, W</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x, S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18CH₃</td>
<td>x, W</td>
<td>x</td>
<td>S</td>
<td>x, S</td>
<td>x, S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

x= scalar or dipolar connectivity observed  
S, M, W= internuclear distance bands according to the ROE intensities  
(S= strong, 1.8-2.4 Å, M= medium, 2.4-4.0 Å; W= weak, 3.4-5.0 Å)
Chapter 3-NMR structure determination of the antibiotics

Table 3.2.3: The total number of NMR experimental ROEs and dihedral angle constraints input into the XPLOR-NIH molecular dynamics program to generate the solution structure of sparsomycin.

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ROE distances</strong></td>
<td>18</td>
</tr>
<tr>
<td>Strong: 1.8-3.0Å</td>
<td>5</td>
</tr>
<tr>
<td>Medium: 1.8-4.0Å</td>
<td>8</td>
</tr>
<tr>
<td>Weak: 1.8-5.0Å</td>
<td>5</td>
</tr>
<tr>
<td>Very weak: 1.8-7.0Å</td>
<td>0</td>
</tr>
<tr>
<td><strong>Dihedral angles</strong></td>
<td>7</td>
</tr>
<tr>
<td><strong>Total restraints</strong></td>
<td>25</td>
</tr>
</tbody>
</table>
3.2.4 NMR solution structure of Sparsomycin

The structure determination protocol described in section 2.4 was implemented to generate the NMR structure of sparsomycin (Figure 3.2.9).

Figure 3.2.9: A. The average NMR structure and B. Overlay of ten structures of sparsomycin with the protons labelled according to the chemical structure in Figure 3.2.1. The solution structure displays an extended conformation similar to the bound X-ray crystal structure.

The generated structure satisfies all of the experimental constraints and yields a very reasonable all-atom average RMSD of 0.52Å. The input of good quality constraints, was essential as the majority of the structure is composed of flexible single C-C bonds.

The NMR structure of sparsomycin is similar to the published sparsomycin-ribosomal RNA bound X-ray crystal structure (Figure 1.4.3), with respect to the orientation of the sulphur tail. The tail adopts a linear arrangement in the crystalline state and in the solution state. The antibiotic interacts mainly via the uracil ring and the sulphur tail and therefore the position of the tail will have direct implications on the binding affinity of the antibiotic. The tail is composed of single bonds and with no hydrogen bonds to constrain
the fold it has the ability to flex out and interact fully with the rRNA residues. The methylated uracil ring can be involved in aromatic base stacking interactions with the RNA, with the folded tail acting as an anchor to stabilise the contacts.

In Figure 3.2.3, the methyl 4CH\(_3\) selectively produces a strong ROE to H7, and an intense ROE between NH9 and H6 was displayed (Fig. 3.2.6). These interactions are satisfied in the NMR structure with the olefinic bond orientated so that the H6 is closer to NH9 and likewise H7 is nearer to 4CH\(_3\).

Sparsomycin possesses several possible hydrogen atom donors and acceptors and could be involved in binding by forming hydrogen bonds with the RNA nucleotides and phosphate backbone. The methyl groups (4CH\(_3\) and 18CH\(_3\)) have been shown to play an important role in producing hydrophobic interactions\(^{112}\) and both groups are well exposed in the NMR structure so as to potentially fulfil this hypothesis. Additional interactions may be possible between the rRNA nucleotides and the methylene (H11/H12, H13/H14 and H16/H17) protons.
3.3 Anisomycin

Figure 3.3.1: The chemical structure of anisomycin consists of a range of functional groups: a substituted pyrrolidine ring, an aromatic ring, a methoxy and an acetoxy group. The hydrogen atoms are distinguished by Arabic numerals and the oxygen atoms are labelled by alphabets. The labelling format is used for the assignment of the NMR spectra. The stereochemistry of the three chiral carbons is indicated by dashed and wedged lines and the R and S configurations in blue lettering.

3.3.1 NMR assignment of Anisomycin

1D and various 2D NMR spectra of anisomycin (Figure 3.3.1) were measured using Bruker 400MHz and 500MHz spectrometers. To begin the assignment, a 1D $^1$H NMR experiment was carried out on the antibiotic in $^1$H$_2$O solvent to observe the secondary amine (NH1) proton. The labile proton is not visible in the $^1$H$_2$O spectrum and so all of the subsequent experiments were carried out in $^2$H$_2$O solvent, thus eliminating the possibility of the water peak affecting the intensities of nearby resonances.

The 1D $^1$H NMR spectrum of anisomycin showed time dependent changes to its resonances as shown in Figure 3.3.2. The spectrum measured on day 1 underwent a progressive change over a period of 3 months which was monitored at regular intervals. As observed in the spectrum after 3 months, the changes began to slow down and remained permanent with no further changes being detected after 6 months. Thus, the spectrum after 3 months was taken up for NMR structure determination. Initially it was thought that the progressive changes may perhaps be due to anisomycin existing in a metastable state upon dissolution and gradually reaching a thermodynamically stable
conformation over a slow period of 3 months. This was also probed by varying temperature, solvents and varying field strengths but no differences were detected.

Figure 3.3.2: 1D $^1$H NMR spectra recorded on first day of sample preparation (at 500MHz, top) and after 3 months (at 400MHz, bottom). The conformation changed after 3 months with the original resonance peaks at 2.2ppm, 7.2ppm almost disappeared and new peaks have emerged at slightly lower field values.
Chapter 3 - NMR structure determination of the antibiotics

Figure 3.3.3: 400MHz 1D $^1$H NMR spectrum of anisomycin (13.3mM) in 100% $^2$H$_2$O at 20°C. The sharp singlet peaks at 1.90ppm and 3.82ppm corresponded to the methyl groups; 7CH$_3$ and 12CH$_3$, respectively.

However, as the 7CH$_3$ methyl group is attached to an ester this functional group can hydrolyse into an alcohol in the presence of water in a reaction known as ester hydrolysis. In the case of anisomycin this was a likelihood as in Figure 3.3.2, there is the original methyl peak at ~2.2ppm disappearing and a new peak appearing at 1.90ppm after 3 months. All of the other proton resonances are unaffected or have a change in their chemical shift values indicating the ester group has cleaved from the antibiotic producing a carboxylic group and leaving a hydroxyl group attached to the pyrrolidine ring. A 1D $^1$H NMR spectrum of sodium acetate was carried out to observe the chemical shift value of the methyl group and a similar value would indicate the ester group of anisomycin has indeed been hydrolysed (Figure 3.3.4).
Figure 3.3.4: 500MHz 1D $^1$H NMR spectrum of sodium acetate (0.6M) in 100% $^2$H$_2$O at 21°C. The methyl of the molecule is represented by the sharp singlet peak at 1.90ppm and the $^1$H$_2$O resonance appears at 4.0ppm, respectively.

The presence of the methyl group of the sodium acetate appearing at the same chemical shift as that of the ester group of anisomycin in Figure 3.3.4 confirms hydrolysis has taken place. The anisomycin sample was purchased from Sigma-Aldrich, which was the same source used to carry out the bound X-ray crystal structure with $H. marismortui$ 50S by the Nobel prize winner T.Steitz and co-workers. The antibiotic hydrolyses in solution, whereas in the crystalline state it remained intact. There is a possibility that the sample was not completely purified and some alkaline substituents remained causing the hydrolysis to take place. However, for the purpose of the structure determination of the antibiotic, the compound was treated like an acetate. In view of the above hydrolysis, the structure determination of the hydroxyl form of anisomycin should be repeated in the future.

The 1D $^1$H NMR spectrum of anisomycin revealed well dispersed resonances (Figure 3.3.3); with different proton groups were appearing within their standard chemical shift ranges: the methyl protons, 7CH$_3$ and 12 CH$_3$ (1-4ppm), the methylene H2, H3 and H8, H9 and methine protons, H4 and H5, (2-5ppm) and the aromatic protons, H10 and H11 (7-
9ppm). The protons were initially assigned using the 1D $^1$H NMR spectrum displayed above in Figure 3.3.3, along with the peak integral values.

The $^1$H-$^1$H TOCSY experiment (Figure 3.3.5) was carried out to initiate the assignment of the pyrrolidine group protons and the NMR assignment of the methyl $12\text{CH}_3$ (3.82ppm) was confirmed by dipolar connectivities to the aromatic H10 and H11 protons in the ROESY spectrum (Figure 3.3.6).

The methylene protons H2, H3 were assigned to resonances appearing at 3.14ppm and 3.64ppm, respectively. Both peaks show connectivities to the resonance appearing at 4.37ppm, which was assigned as proton H4. There is an overlap of peaks at 2.97ppm (H9) and 3.14ppm (H8) and they both displayed connectivities to the resonance of proton H6 at 3.93ppm.

In addition to couplings between H2, H3, and H4, additional connectivities to the peak at 4.14ppm, were assigned to H5. Protons H8, H9 also showed strong couplings to H5 and to H4, the latter of weaker intensity.
Figure 3.3.5: 400MHz $^1$H-$^1$H TOCSY spectrum ($\tau_m = 75$ms) of anisomycin (13.3mM) in 100% $^2$H$_2$O at 20ºC. The spectrum displayed resolved peaks, but it was not possible to individually assign the methylene protons H2 and H8, due to peak overlap.

The intensities of the H2-H3 and H3-H4 connectivities are stronger than those between H2-H5 and H3-H5 as the former are separated by three bonds, while the latter are separated by four bonds.

The 2D $^1$H-$^1$H ROESY spectrum displayed in Figure 3.3.6 confirmed the assignment of the aromatic protons H10 (7.31ppm) and H11 (7.0ppm). The H10 resonance displayed dipolar couplings to the resonances of H5, H6, H8 and H9 protons and a strong coupling was observed between H11 proton and methyl 12CH$_3$. 

-170-
**Figure 3.3.6:** A 400MHz $^1$H-$^1$H ROESY spectrum ($\tau_m=400$ms) of anisomycin (13.3mM) in 100% $^2$H$_2$O at 20ºC. A number of ROEs were obtained from the spectrum which provided valuable distance constraints to link the pyrrolidine and aromatic residues.

Figure 3.3.7 shows the $^1$H-$^{13}$C HSQC spectrum, which confirmed the assignment of the methylene protons. Both protons H8 and H9 resonances showed coupling to a carbon with a chemical shift value of 33.4ppm, likewise, both H2 and H3 coupled to the carbon at 53.4ppm.
Figure 3.3.7: A 2D $^1$H-$^{13}$C HSQC spectrum of anisomycin (13.3mM) in 100% $^2$H$_2$O at 20°C.
Table 3.3.1: $^1$H, $^{13}$C chemical shifts (δ, ppm) and coupling constants ($^3$J, Hz) of anisomycin antibiotic. The values are based on the 400MHz 1D $^1$H NMR spectrum of anisomycin in 100% $^2$H$_2$O at 21°C.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>$^1$H chemical shift δ (ppm)</th>
<th>$^{13}$C chemical shift δ (ppm)</th>
<th>$^1$H coupling constant $^3$J (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>3.14</td>
<td>53.4</td>
<td>m</td>
</tr>
<tr>
<td>3</td>
<td>3.64</td>
<td>53.4</td>
<td>dd, 4.7</td>
</tr>
<tr>
<td>4</td>
<td>4.37</td>
<td>77.1</td>
<td>d, 4.9</td>
</tr>
<tr>
<td>5</td>
<td>4.14</td>
<td>77.7</td>
<td>d, 2.5</td>
</tr>
<tr>
<td>6</td>
<td>3.93</td>
<td>66.3</td>
<td>m, 6.4</td>
</tr>
<tr>
<td>7*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>3.14</td>
<td>33.4</td>
<td>m</td>
</tr>
<tr>
<td>9</td>
<td>2.97</td>
<td>33.4</td>
<td>dd, 9.1</td>
</tr>
<tr>
<td>10</td>
<td>7.31</td>
<td>132.9</td>
<td>d, 8.3</td>
</tr>
<tr>
<td>11</td>
<td>7.00</td>
<td>117.2</td>
<td>d, 8.3</td>
</tr>
<tr>
<td>12</td>
<td>3.82</td>
<td>58.0</td>
<td>s</td>
</tr>
</tbody>
</table>

(-) not measurable. Singlet (s) peaks do not possess coupling value.
Doublet (d) triplet (t) and doublet of doublets (dd) peaks are indicated in the table.
Multiplet (m) peaks are estimates.*The ester group hydrolysed and so no value was obtained.
3.3.2 Experimental constraints

A total of 18 experimental constraints were input into the structure determination calculation to generate the NMR structures of anisomycin. The ROESY experiment provided the ROE based distance constraints and the correct geometry was invoked by the torsional angle constraints, calculated from the $^3J$ scalar coupling constants measured from the 1D $^1$H NMR spectrum. The same methodology (section 2.4), used to produce the NMR structures of bamicetin and sparsomycin was adopted to generate the solution structure of anisomycin.

Smaller number of constraints (18) was used to produce the anisomycin structure than those that generated the bamicetin (64) and sparsomycin (25) structures. This was due to the limited number of protons present in the chemical structure. Nevertheless, the structure of anisomycin revealed a low energy stable conformation.

**Table 3.3.2:** The total number of NMR experimental ROEs and dihedral angle constraints input into the XPLOR-NIH structure determination program to determine the structure of anisomycin.

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROE distances</td>
<td>14</td>
</tr>
<tr>
<td>Strong: 1.8-3.0Å</td>
<td>5</td>
</tr>
<tr>
<td>Medium: 1.8-4.0Å</td>
<td>5</td>
</tr>
<tr>
<td>Weak: 1.8-5.0Å</td>
<td>3</td>
</tr>
<tr>
<td>Very weak: 1.8-7.0Å</td>
<td>1</td>
</tr>
<tr>
<td>Dihedral angles</td>
<td>4</td>
</tr>
<tr>
<td>Total constraints</td>
<td>18</td>
</tr>
</tbody>
</table>
3.3.3 NMR solution structure of Anisomycin

The energy-minimised NMR structure of anisomycin was produced from the ten best converged structures (Figure 3.3.8).

![Figure 3.3.8](image)

Figure 3.3.8: (A) The energy minimised NMR structure of anisomycin. (B) The structure was generated from an average of 10 best accepted structures. The bifurcated hydrogen bonds (2.3Å and 2.7Å) are indicated by the black dashed lines.

Despite the small number of experimental constraints, anisomycin displays a low average global RMSD value of 0.35Å. The NMR structure possesses the correct stereochemistries of the chiral centres (R, R and S configuration) displayed in Figure 3.3.8.

The antibiotic consists of a range of functional groups which can be involved in intermolecular interactions with the RNA residues in several ways. It can be hypothesised that the heavy atoms (N1, Oa, Ob, Oc, Od) of anisomycin act as hydrogen bond donors and acceptors and maybe involved in electrostatic interactions with the PT rRNA residues. Additionally, the oxygen atoms may be involved in interactions with RNA nucleotides. The methyl group (12CH₃) and methylene protons (H6, H7) have the potential to bind via van der Waal interactions.

The X-ray crystal structure of the large ribosomal subunit of *H.marismortui* (H50S) bound to anisomycin at 3.0Å has been described in section 1.4.3. The NMR structure shares similarities with the X-ray crystal structure of anisomycin. The methoxy group is orientated in the same way in both structures, strongly implying that this is the most stable
orientation. As the two structures are similar, it can be predicted that they bind to the RNA residues of the PT centre in an analogous manner.

The benzyl ring in the NMR structure has the capability to stabilise interactions with the ribosome through base stacking with the RNA nucleotides.
3.4 Discussion and comparison of Bamicetin, Sparsomycin and Anisomycin structures

Figure 3.4.1: The NMR structures of the bamicetin, sparsomycin and anisomycin antibiotics.

The NMR experiments of bamicetin, sparsomycin and anisomycin produced stable, low energy structures exhibiting small all average RMSD values, good overlay of structures and thermodynamic stability. The NMR structures of bamicetin and sparsomycin contain possible intramolecular hydrogen bonds established by the exchange retarded amide protons NH\textsubscript{V} and NH\textsubscript{9} respectively which serve to enhance the conformational stability of the molecules.

The resonances of both amide group protons display significant temperature dependent chemical shift and line intensity changes. The bamicetin NH\textsubscript{V} proton disappears at 25°C, in contrast the NH\textsubscript{9} proton of sparsomycin is observed at much higher temperature (50°C).

This states that the amine proton of sparsomycin is most likely involved in stronger intramolecular hydrogen bonding whereas the amine of bamicetin is displays weak hydrogen bonding. Thus suggesting that the NH\textsubscript{V} proton is free to be involved in interactions with the peptidyl transferase ribosomal RNA residues.

The variable temperature series of the non labile resonances of bamicetin (1CH\textsubscript{3}) and sparsomycin (H13/H14) provided additional evidence for the stability of the two antibiotics. The NMR peak of the methyl group proton of bamicetin decreased in
linewidth and the methylene (H13/H14) protons of sparsomycin displayed scalar couplings to nearby proton at elevated temperature.

The chemical structure of bamicetin is very similar to that of the amicetin antibiotic and naturally, their NMR structures are similar. The two structures differ in the α-methylserine moiety section, with amicetin displaying greater range of flexibility. The N-methyl (I) of bamicetin is involved in intramolecular hydrogen bonding and this is one plausible reason as to why bamicetin is a stronger PT inhibitor than amicetin, as amicetin does not contain a proton at this position in the chemical structure. However, both structures are fairly well conserved in the saccharide and aromatic regions of the molecules.

The sparsomycin NMR structure displays resemblance to its X-ray crystal structure with respect to the sulphur tail. However, the antibiotic is likely to alter conformation upon binding to the ribosomal subunits.

All three antibiotics produced good degree of convergence between the energy-minimised structures. The low RMSD values and zero constraint violations highlight the importance of the type and quantity of NMR experimental constraints input and the structure generation protocol used.
Chapter 4  NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

The effect of bamicetin on the peptidyl transferase (PT) activity of *B. subtilis* 70S ribosome was monitored in a PT assay, the results revealed high affinity between the two species. Based on previous studies with amicetin, bamicetin is predicted to bind near to or at the same site, close to the peptidyl transferase centre. A 27mer RNA motif located in the stem region of the PTC of *B. subtilis* 23S rRNA was chosen to study its interactions with bamicetin. A description of the NMR assignment and structure determination of the 27mer RNA, followed by a detailed analysis of the binding studies is presented in this Chapter. Conformational analysis of the NMR structure was carried out using the w3DNA and CURVES software.

4.1  NMR assignment of the 27mer RNA motif of *B. subtilis* rRNA

4.1.1  Secondary structure

Figure 4.1.1 shows the stable *B. subtilis* 27mer RNA motif and is obtained from the predicted secondary structure of the 23S ribosomal RNA. The motif is double stranded, comprising of 13 nt and 14 nt in each strand with two bulges in the middle. It consists of 12 stable canonical Watson-Crick base pairs and 3 unpaired residues. The universally conserved nucleotides (C1, C2, U6, U12, U13, U26, A27, C28, G34 and G35) as described in section 1.5 are indicated on the secondary structure.
Figure 4.1.1: The secondary structure of the double stranded *B.subtilis* 27mer RNA motif used for study by NMR. For the purpose of NMR assignment, the nucleotides are labelled numerically as shown in Figure 1.5.2, instead of the absolute numbers (four digits) found in the *B.subtilis* 23S ribosomal RNA. The universally conserved nucleotides are indicated by the red dots.

4.1.2 NMR assignment of the exchangeable proton resonances

The NMR assignment of the 27mer RNA was initiated by acquiring a 1D $^1$H NMR spectrum in $^1$H$_2$O solvent (Figure 4.1.2). The sample was dissolved in $^1$H$_2$O to observe all the RNA proton resonances.
Chapter 4- NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

Figure 4.1.2: 600MHz 1D $^1$H NMR spectrum of *B. subtilis* 27mer RNA (0.59mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. The resonances of different proton groups; imino NH protons (green box), amino NH$_2$ and aromatic protons (blue box), pyrimidine H5 and glycosidic proton H1' (red box) and ribose protons (brown box) appear in distinct regions of the spectrum.

Figure 4.1.2 displays the separate chemical shift regions in which the different proton groups of the RNA appear within the NMR spectrum. The ribose proton peaks (H2’, H3’, H4’, H5’, H5’’) appear at highest field (3.8-5.0ppm) whilst the glycosidic proton H1’ and pyrimidine H5 resonate in the 5-6.5ppm range. The aromatic and labile amino NH$_2$ protons resonate in the 6.0-9.0ppm region, with the imino NH protons appearing at lowest field (9.5-15.0ppm) due to the stabilising ring current effect induced by base stacking.

A series of NOESY spectra at different mixing times ($\tau_m = 100$ms, 150ms and 250ms) in $^1$H$_2$O was acquired. The shortest mixing time (100ms) spectrum produced the optimum results with the maximum number of NOE connectivities observed in the labile region, so this spectrum was selected for the assignment of the exchangeable protons.
Imino proton resonances

The chemical shift of the imino NH proton resonance reveals whether the proton is base paired or in an unpaired state and thus reveals the secondary structure of the motif. Base paired guanine NH protons resonate at 12-13.5ppm and uracil NH protons appear in the 13-15ppm region. Unpaired or non canonical base paired NH protons appear with weaker intensity at higher field; 9.5-12ppm. Imino protons are located only on the aromatic guanine and uracil bases and so each canonical Watson-Crick base pair will produce one peak in the imino region of the NOESY spectrum.

The NH proton resonances are assigned by the presence of dipolar connectivities to the imino protons of the adjacent base pair. Figure 4.1.3 displays the assigned NH proton resonances and the imino-imino connectivities are highlighted on the secondary structure of the 27mer RNA motif. The terminal uracil (U13) imino proton was assigned to the resonance at 14.12ppm, overlapping with the resonance of U12. A strong cross peak between the U12/U13-NH and the imino resonance at 13.55ppm was used to assign G24, which is adjacent to the U12-A23 base pair. G24 displayed a strong imino-imino connectivity at 13.65ppm, which was assigned to the adjacent G10-NH proton. Although no connectivities were observed to the U26 imino proton (13.35ppm), the assignment was supported by a strong NOE to the base paired A9-H2, which is described in the next section (amino proton resonances).
Chapter 4- NMR structure determination of the *B.subtilis* 27mer RNA and its interactions with bamicetin

**Figure 4.1.3**: 600MHz $^1$H-$^1$H NOESY ($\tau_m = 100$ms) spectrum of *B.subtilis* 27mer RNA (0.59mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. The figure illustrates the sequential assignment of the imino proton resonances and the observed sequential imino-imino connectivities are illustrated by the blue arrows on the adjoining secondary structure of the 27mer RNA.

The terminal G35 NH proton produced a weak resonance at 13.26ppm due to rapid solvent exchange with water protons. The G34 NH proton resonance (12.92ppm) was assigned by the presence of a cross peak with the G35 resonance. An imino-imino cross peak between G34 and the cross peak at 12.68ppm was used as the basis for assignment of G33 NH proton, which in turn possesses a strong NOE to the neighbouring G32 imino proton resonance (12.38ppm). A sequential NOE connectivity was observed between G32 to G5 (12.80ppm). The NH proton resonances of U6 (10.72ppm), G7 (12.50ppm) and G8 (12.11ppm) did not produce imino-imino correlations however; the assignments were verified by connectivities in the amino region (described below). Ten out of thirteen imino proton resonances were identified either from sequential NOE imino connectivities and/or correlations to the amino and aromatic protons (Figure 4.1.4).
Amino proton resonances

Amino proton resonances appear with weaker intensity, due to broadening from exchange with the solvent protons and rotation about the C-N bond. The NH$_2$ of cytosine residues are more commonly observed since they are often in a slow exchange regime. Guanine and adenine NH$_2$ protons are in intermediate exchange at room temperature and therefore the corresponding resonances are broadened.

**Figure 4.1.4**: 600MHz $^1$H-$^1$H NOESY ($\tau_m=100$ms) spectrum of *B.subtilis* 27mer RNA (0.59mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. The imino proton resonance assignment were confirmed by the identification of connectivities to the NH$_2$, aromatic and ribose protons. The assigned imino proton resonances are labelled along the top of the spectrum.
The assignment of the imino NH proton resonances was confirmed by the sequential and intraresidual dipolar correlations to the NH₂, aromatic and sugar protons (Figure 4.1.4). Strong correlations are observed in Figure 4.1.4 between the U12 imino proton and the adenine A23-H2 and U26(NH)-A9(H2) since the interproton distance between the two groups of protons is < 4Å.

### 4.1.3 Assignment of the non exchangeable proton and carbon resonances

A 1D ¹H NMR spectrum (Figure 4.1.5) in conjunction with 2D homonuclear and heteronuclear experiments; ¹H-¹H DQF-COSY, ¹H-¹H TOCSY, ¹H-¹³C HSQC and ¹H-¹H NOESY were acquired in ²H₂O at 25°C to assign the non exchangeable protons. A set of NOESY spectra with two different mixing times (τₘ = 150 and 400ms) was obtained. The 400ms spectrum contained a large number of NOE correlations, in addition to weaker peaks in the 150ms spectrum which were enhanced and used to confirm the NMR assignment.

![Figure 4.1.5: 800MHz 1D ¹H NMR spectrum of B.subtilis 27mer RNA (0.59mM, pH 6.2) in 100% ²H₂O at 25°C. Adenine H2 protons do not share scalar coupling with other protons and so produce sharp singlet peaks in the aromatic region (see below).](image)
Figure 4.1.5 displays the 1D $^1$H NMR spectrum of the 27mer RNA in $^2$H$_2$O, the deuterated solvent removes the labile NH and NH$_2$ proton resonances, thus minimising the spectral width and simplifying the aromatic region of the spectrum. The pyrimidine H6 protons resonate in the 7-7.7 ppm region and the H8 proton resonances appear in the 7-8.5 ppm range, but those present on unpaired or non canonical base pairs may be shifted to lower field.

**Adenine H$_2$ proton resonances**

The adenine H2 protons produce sharp singlet peaks due to no neighbouring scalar coupled protons and are generally easy to distinguish in the 1D $^1$H NMR spectrum (7-8.5 ppm). The base paired H2 protons produced NOE correlations to the uracil NH protons in the NOESY ($^1$H$_2$O) spectrum in Figure 4.1.4 and so were the first non exchangeable protons to be assigned.

The heteronuclear one bond scalar correlated $^1$H-$^{13}$C HSQC spectrum displays the directly bonded carbons (C2) of the adenine H2 protons in the 154-156 ppm region (Figure 4.1.6) where no signals arising from other resonances are observed.
Chapter 4- NMR structure determination of the *B.subtilis* 27mer RNA and its interactions with bamicetin

**Figure 4.1.6:** 600MHz $^1$H-$^{13}$C HSQC spectrum of the *B.subtilis* 27mer RNA (0.59mM) in 100% $^2$H₂O, pH 6.2 at 25°C. The four adenine H2 protons were unambiguously assigned based on their correlation to the C2 resonances appearing in a non overlapped region of the spectrum.

All four adenine (A9, A22, A23 and A27) H2 protons were identified by the scalar correlations to the attached C2 carbons in the HSQC spectrum. Additional weaker peaks observed are due to a small amount of RNA adopting a minor thermodynamic conformation. Adenine H2 protons are at least 4.5Å away from all sugar protons from the same nucleotide and display connectivity only to the H1’ proton. Therefore, the H2 proton assignment of A9, A22, A23 and A27 was confirmed by the presence of sharp resonances in the 1D $^1$H NMR spectrum (Figure 4.1.5), the appearance of low field C2 resonances (Figure 4.1.6) and no detectable NOE connectivities to the ribose proton (H2’, H3’, H4’, H5’, H5”’) resonances.
Pyrimidine $H_5$,$H_6$ proton resonances

Pyrimidine $H_5$-$H_6$ connectivities are readily identified through strong mutual NOE cross peaks corresponding to a covalently fixed 2.4Å distance and scalar coupling (~7Hz) in $^1H$-$^1H$ DQF-COSY and $^1H$-$^1H$ TOCSY spectra. The resonances are slightly broader than the $H_2$ and $H_8$ singlets due to the neighbouring scalar coupled proton. The assignment was carried out in the TOCSY spectrum displayed in Figure 4.1.7 as singlet ($H_2$, $H_8$) resonances are not observed.

![Figure 4.1.7](image)

**Figure 4.1.7**: 600MHz $^1H$-$^1H$ TOCSY ($\tau_m = 75$ms) spectrum of the *B.subtilis* 27mer RNA (0.59mM) in 100% $^2H_2$O, pH 6.2 at 25°C. The figure displays the strong H5-H6 couplings due to the pyrimidine bases cytosine (green circles) and uracil (blue circles).

The H5-H6 resonances were further assigned to cytosine and uracil bases using the HSQC spectrum. The H5 resonances are identified from the $^{13}C$ chemical shifts of the attached C5 resonances. Uracil C5 resonances resonate downfield (100-105ppm) from cytosines (95-100ppm) and the assignment is displayed in Figure 4.1.8.
Chapter 4 - NMR structure determination of the *B.subtilis* 27mer RNA and its interactions with bamicetin

The assignment of all the 14 pyrimidine H5-H6 resonances was achieved successfully and this enabled the purine H8 protons to be determined by elimination.

**Ribose proton resonances**

The ribose protons (H1’, H2’, H3’, H4’, H5’, H5’”) were assigned with the aid of the scalar correlated HSQC, DQF-COSY and TOCSY experiments. The dispersion of $^{13}$C chemical shifts in the HSQC spectrum results in the different ribose carbons (C1’, C2’, C3’, C4, C5’) appearing in separate regions. Figure 4.1.9 displays the H1’ proton resonance assignment obtained from the HSQC spectrum.

**Figure 4.1.8**: 600MHz $^1$H-$^{13}$C HSQC spectrum of the *B.subtilis* 27mer RNA (0.59mM) in 100% $^2$H$_2$O, pH 6.2 at 25ºC. (A) displays the $^1$J H5-C5 correlations of the cytosine bases and (B) shows the $^1$J H5-C5 correlations of the uracil bases.
Figure 4.1.9: 600MHz $^1$H-$^{13}$C HSQC spectrum of the *B. subtilis* 27mer RNA (0.59mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C. The glycosidic H1’ protons were identified by their C1’ correlation to the respective carbons.

The ribose ring is a coupled spin system and further assignment of the protons was achieved through scalar couplings identified in the DQF-COSY and TOCSY spectra. However, due to overlap in the $^1$H chemical shift region (3.8-5ppm), unambiguous assignment of all the ribose protons was not possible. The magnitude of the H1’-H2’ coupling identified in the $^1$H-$^1$H DQF-COSY (Figure 4.1.10) is dependent on the dihedral angle formed by the H1’-C1’-C2’-H2’ atom plane and therefore determined the pucker of the ribose. Therefore, H1’-H2’ vicinal couplings will reveal structural information about the sugar puckers of the ribose rings.
Chapter 4- NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

![Diagram of NMR spectrum](image)

**Figure 4.1.10**: 600MHz $^1$H-$^1$H DQF-COSY spectrum of *B. subtilis* 27mer RNA (0.59mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C. The scalar correlated H1'-'H2' cross peaks of the assigned nucleotides indicated along the top of the figure were constrained with C$_2$-endo sugar puckers.

Vicinal $^3$J$_{H1'-H2'}$ coupling values for C$_3$-endo sugar puckers, typically found in A-form helices are <3Hz and almost invisible in the DQF-COSY spectrum, while C$_2$-endo contain larger $^3$J$_{H1-H2}$ coupling values ~7-8Hz and thus show more intense resonance cross peaks. Therefore, as observed in Figure 4.1.10, the ribose rings of nucleotides such as A27, C30 and G35 were constrained to C$_2$-endo sugar puckers.

**Sequence specific assignment**

The sequence specific assignment of the RNA was achieved through the identification of aromatic and sugar protons and confirmed the secondary structure of the motif. The H1’ protons display NOE contacts to the H6 and H8 protons of their own base (H6/H8$_n$-H1’$_n$) and to the succeeding nucleotide (H6/H8$_{n+1}$-H1’$_{n+1}$) in the 5’→3’ direction. The intraresidual base to ribose pattern is observed in helical regions of the nucleic acid structure with stronger NOE connectivities for the intranucleotide protons (3.4-3.9Å) due to a shorter intraproton distance than the internucleotide protons (4.4-5.2Å) (Figure 4.1.11).
Figure 4.1.11: An illustration of the NOE contacts produced between the H6 base proton to the intraresidual glycosidic H1’ proton (blue arrow) and a second NOE from H1’ to the H8 aromatic proton on the subsequent nucleotide (green arrow) in the NOESY spectrum. Shorter distance between the intranucleotide protons resulted in stronger connectivities as displayed in the NOESY diagram in the figure. The chain of sequence specific NOE connectivities maps out the secondary structure of the motif.

The aromatic (H6, H8) protons were observed in the $^1$H-$^1$H NOESY ($\tau_m = 150$ and 400ms) spectra in 100% $^2$H$_2$O solvent (Figure 4.1.12 and Figure 4.1.13). The intranucleotide connectivity of the terminal C1 nucleotide was designated to the resonance at 8.2-5.62ppm since it only contained two connectivities in the aromatic H6 region; one corresponding to the mentioned intranucleotide and one to the H5-H6 connectivity. C1 (H1’) proton displayed internucleotide cross-peak to the H6 resonance at 8.07ppm which was assigned to the proceeding base C2. The C2 aromatic proton in turn possesses intranucleotide H1’connectivity and internucleotide correlation to the C3 H1’ proton. An uninterrupted scheme was obtained for the entire strand (C1-U13) with overlap resonances for U6-G7 and G8-A9. The adenine H2-H1’ connectivities are observed in the NOESY spectrum and indicated on both Figures 4.1.12 and 4.1.13.
Figure 4.1.12 displays the intraresidual sequential assignment of one strand of the 27mer motif and the determined resonances were confirmed by the assignments made in the heteronuclear $^1$H-$^{13}$C HSQC and $^1$H-$^{31}$P CPMG-HSQC-NOESY spectra. The sequence specific assignment was carried out for both strands of the 27mer motif and the sequential assignment of the second strand is displayed in Figure 4.1.13.

In helical regions, aromatic-aromatic connectivities may be observed and can be used to validate the sequence specific assignment. In Figure 4.1.12 C11(H5)-G10(H8) and in Figure 4.1.13 C25(H6)-G24(H8) connectivities were observed. The sequential H6/H8-H1’ assignment was implemented to achieve a similar sequential assignment with the aromatic H6/H8 and ribose H2’ protons (not shown).
Figure 4.1.12: (a) 600MHz $^1$H-$^1$H NOESY ($\tau_m = 400$ms) spectrum of *B.subtilis* 27mer RNA (0.59mM) dissolved in 100% $^2$H$_2$O, pH 6.2 at 25°C. Sequential connectivities observed between the aromatic H6/H8 and the sugar H1’ protons are shown for the top strand of the 27mer sequence. Blue filled circles represent intranucleotide and green filled circles represent internucleotide connectivities. The adenine (A9) H2 resonance is indicated by the brown filled circle. (b) 600MHz $^1$H-$^{13}$C HSQC spectrum shows the H6/H8-C6/C8 correlated spectrum. (c) and (e) 600MHz $^{31}$P CPMG-HSQC-NOESY ($\tau_m = 500$ms) spectra displays the NOE connectivites observed between the glycosidic H1’ and base H6/H8 protons and the phosphorus atom. (d) 600MHz $^1$H-$^{13}$C correlated HSQC spectrum showing the H1’-C1’ region.
Figure 4.1.13: (a) 600MHz $^1$H-$^1$H NOESY spectrum of *B.subtilis* 27mer RNA (0.59mM) dissolved in 100% $^2$H$_2$O, pH 6.2 at 25 °C. Sequential connectivities observed between the aromatic H6/H8 and the sugar H1' protons are shown for the bottom strand of the 27mer sequence. Blue filled circles represent intranucleotide and green filled circles represent internucleotide connectivities. Brown filled circles highlight the resonances of A22, A23 and A27 H2 protons. (b) 600MHz $^1$H-$^{13}$C HSQC spectrum showing the H6/H8-C6/C8 correlated spectrum. (c) and (e) 600MHz $^{31}$P CPMG-HSQC-NOESY ($\tau_m = 500$ms) spectra displaying the NOE connectivities observed between the ribose H1’ and base H6/H8 protons and the phosphorus atom. (d) 600MHz $^1$H-$^{13}$C correlated HSQC spectrum shows the H1’-C1’ region.
4.1.4 Assignment of the phosphorus resonances

Chemical shifts of $^{31}$P nuclei are sensitive to temperature, pH and conformation and appear in the -1 to -5ppm region. Figure 4.1.14 is the 1D $^{31}$P {¹H} spectrum which displays the assigned phosphorus resonances all appearing in the A-form characteristic chemical shift region.

![Figure 4.1.14: 81MHz 1D {¹H-decoupled} $^{31}$P –NMR spectrum of B.subtilis 27mer RNA (0.59mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C. All the phosphorus atoms of the 27 residues were obtained with the aid of the 2D CPMG-HSQC-NOESY ($\tau_m = 500$ms) spectrum (Figure 4.1.15).](image)

A 2D $^1$H-$^{31}$P CPMG-HSQC-NOESY ($\tau_m = 500$ms) spectrum was obtained to assign the backbone phosphorus resonances (Figure 4.1.15). The spectrum displays scalar correlations between the $^{31}$P nuclei and ribose H3’, H5’ and H5” protons in addition to dipolar couplings to the base (H2, H6, H8) protons. Unfortunately, it was not possible to obtain a sequential assignment as observed in the NOESY ($\tau_m = 250$ms, $^2$H$_2$O) spectrum in Figures 4.1.12/13 but nevertheless most of the aromatic protons were identified.
An interesting trend was observed with the adenine A27 protons, the H2, H8, H1’ and H2’ proton resonances appeared low field in the NMR spectra. The H8 proton (8.43ppm) was the most low field shifted resonance amongst the aromatic protons (Figure 4.1.13) and the H2 proton (8.10ppm) was also the lowest field shifted resonance out of the adenine H2 protons (Figure 4.1.6). The H1’ proton resonated lowest field (6.14ppm) amongst the glycosidic protons (Figure 4.1.9) and the H2’ proton resonance appearing third lowest at 4.58ppm after A9 (4.69ppm) and G8 (4.66ppm) in Figure 4.1.10.

This effect was also observed with the nearby A9 base with the H2’ proton resonated lowest field (4.69ppm) out of the H2’ protons. The aromatic H2 proton appeared at 7.75ppm with only A27-H2 resonating lower field. The observations imply that the two nucleotides are in a different local environment from the other residues of the 27mer RNA motif. The NMR structure revealed a non Watson-Crick base pairing between A9 and A27 and this is explained in section 4.1.8.
Figure 4.1.15: 600MHz 2D $^1$H-$^{31}$P CPMG-HSQC-NOESY ($\tau_n=500\text{ms}$) spectrum of B.subtulis 27mer RNA (0.59mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C. (A) $^{31}$P edited NOE connectivities to the H1’ and H5 protons and (B) $^{31}$P edited NOE correlations to aromatic (H6/H8) protons of the RNA.
4.1.5 Assignment table

Table 4.1.1: \(^1\)H, \(^{13}\)C, \(^{31}\)P chemical shifts (\(\delta\), ppm) of the \textit{B.subtilis} 27mer RNA (0.59mM). Labile imino and amino proton resonances were measured in 90\% \(^1\)H\(_2\)O + 10\% \(^2\)H\(_2\)O, pH 6.2 at 2\(^\circ\)C and non labile aromatic and sugar resonances in 100\% \(^2\)H\(_2\)O at 25\(^\circ\)C. All \(^1\)H assignments referenced to the residual water peak according to temperature.\(^{113}\) \(^{13}\)C and \(^{31}\)P assignment at 25\(^\circ\)C were referenced indirectly to \(^1\)H.

![Table Image]

-199-
4.1.6 Experimental constraints

236 NOE constraints, 234 dihedral angle constraints and 24 hydrogen bond restraints outlined in Table 4.1.2, derived from the above NMR experiments were used to produce the NMR structure of the \textit{B. subtilis} 27mer RNA. The main source of structural data was obtained from the NOE correlations, with 55 intranucleotide and 49 internucleotide NOE constraints providing secondary structural information. Cross-peak volumes were extracted from short ($\tau_m = 100$ms) and long ($\tau_m = 400$ms) mixing time NOESY spectra and converted into interproton distances by the CCPNMR Analysis software. Due to spectral overlap, errors in peak intensity are sometimes possible and so the NOEs were assigned into distance ranges instead of absolute distances for each NOE, as outlined in section 2.4.2.

The torsion angle constraints consisted of ribose puckers ($\upsilon_1, \upsilon_2, \chi$) and backbone ($\alpha, \beta, \gamma, \varepsilon, \zeta$) dihedral angles. A-form geometry was invoked for all C$_3$-\textit{endo} nucleotides and the backbone angles ($\alpha = -68^\circ \pm 17^\circ$, $\beta = 178^\circ \pm 17^\circ$, $\gamma = 54^\circ \pm 17^\circ$, $\varepsilon = -153^\circ \pm 17^\circ$, $\zeta = -71^\circ \pm 17^\circ$) were set when the $^{31}$P chemical shifts of each nucleotide fell in the -1 to -5ppm region. Glycosidic angles of the nucleotides were constrained to \textit{anti} conformation ($150^\circ \pm 90^\circ$) with the presence of intraresidue H6/H8-H1’ correlations.

Labile NH and NH$_2$ protons display changes in chemical shift value according to hydrogen bonding interactions. If the resonances appeared in the characteristic Watson-Crick base pairing regions (12-15ppm and 7-9ppm, respectively), their identification was used to constrain the base pairs by means of hydrogen bonding restraints. Hydrogen bond constraints were placed on the residues which contained imino-imino (Figure 4.1.3) and/or imino-amino connectivities (Figure 4.1.4) in the NOESY ($^1$H$_2$O) spectrum. The constraints contained loose error bounds as the labile protons can be prone to exchange with the solvent $^1$H$_2$O protons and susceptible to distorted peak intensities. Database orientation restraints were included in the refinement step to compensate for the lack of exchangeable constraints. Both sets of constraints were input in the structure
determination calculation to encourage the bases to adopt the canonical Watson-Crick base pairing.

**Table 4.1.2**: A summary of the experimental constraints used as input into the structure calculation of the *B. subtilis* 27mer RNA.

<table>
<thead>
<tr>
<th>Constraint type</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE distances</td>
<td>236</td>
</tr>
<tr>
<td>Strong</td>
<td>47</td>
</tr>
<tr>
<td>Medium</td>
<td>116</td>
</tr>
<tr>
<td>Weak</td>
<td>70</td>
</tr>
<tr>
<td>Very weak</td>
<td>3</td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>24</td>
</tr>
<tr>
<td>Glycosidic (χ)</td>
<td>27</td>
</tr>
<tr>
<td>Ribose pucker (ν1, ν2, δ)</td>
<td>81</td>
</tr>
<tr>
<td>Helix (α, β, γ, ε, ζ)</td>
<td>126</td>
</tr>
<tr>
<td>NOEs per residue</td>
<td></td>
</tr>
<tr>
<td>Constraints per residue</td>
<td>18</td>
</tr>
<tr>
<td>Total constraints</td>
<td><strong>494</strong></td>
</tr>
</tbody>
</table>
4.1.7 NMR structure of the 27mer RNA motif of *B.subtilis* rRNA

The structure calculation protocol detailed in section 2.6 was implemented to produce 200 simulated annealed and 200 refined NMR structures of the *B.subtilis* 27mer RNA. The NMR structure (A) and an overlay of the ten best structures of the 27mer RNA (B) chosen from the refined structures are displayed in Figure 4.1.16.

![Figure 4.1.16](image)

**Figure 4.1.16:** (A) The NMR structure of *B.subtilis* 27mer RNA with low RMSD of 0.30Å. (B) Overlay of the ten lowest energy NMR structures with a RMSD of 0.32Å. The nucleotides are colour coded to match the bases in the secondary structure on the left hand side of the figure.

In assessing the quality of the final NMR structures and prior to acceptance, a number of statistics were considered: average global RMSD, potential energies, empirical bond violations (angles, lengths or impropers) and NOE and torsion angle constraint violations. An overlay of the ten, best violation-free structures displayed an average all atom RMSD value of 0.32Å. The low RMSD value signifies the stability of the motif and that this is the major conformation of the molecule (Figure 4.1.16). Local RMSD values (Table...
Chapter 4- NMR structure determination of the *B.subtilis* 27mer RNA and its interactions with bamicetin

4.1.3) were also calculated to observe the dynamic nature of the flexible internal bulge regions and the values indicate good convergence and resemble the overall global RMSD.

**Table 4.1.3**: Average RMSD values of the whole motif and the nucleotides of the two bulges.

<table>
<thead>
<tr>
<th>Residues</th>
<th>RMSD value (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>27mer motif</td>
<td>0.32</td>
</tr>
<tr>
<td>G5-G8</td>
<td>0.27</td>
</tr>
<tr>
<td>U26-C30</td>
<td>0.28</td>
</tr>
</tbody>
</table>

The highly conserved U6 base is flipped out of the helix in the NMR structure of the 29mer *E.coli* rRNA\(^69\) but in the *B.subtilis* 27mer RNA it is flipped in and the position is stabilised by the G-C base pairs on its either side (Figure 4.1.17). The low value of the local RMSD for residues G5-G8 which compose the U6 bulge also implies that the unpaired U6 is held in a stable arrangement.
Chapter 4- NMR structure determination of the *B.subtilis* 27mer RNA and its interactions with bamicetin

**Figure 4.1.17**: (A) Illustration of the U6 internal loop in the NMR structure of *B.subtilis* 27mer RNA. (B) Displays the conformation of the U6 base amongst the G-C rich Watson-Crick base pairs. Blue circle sugar moiety; red circle phosphate group; blue box nucleic acid base; red filled box demonstrates base stacking, blue solid lines; glycosidic and phosphodiester bonds and blue dotted lines show hydrogen bonds.

### 4.1.8 Conformational analysis

W3DNA and CURVES conformational analysis programmes were used to determine the helical and torsional angle parameters of the 27mer RNA motif. The results confirmed that the C1 and G35 bases are not base paired, this is due to the terminal base pair being more prone to exchange with the solvent protons and thus weakening its hydrogen bonding interactions. Further, the NMR structure reveals a non-canonical base pair (A-A) between A9 in the upper part of the helix and A27 in the bulge. The unpaired adenine (A27) in the bulge base pairs with the A9 (Figure 4.1.18) and in doing so disrupts the canonical base pairing of A9-U26. The interaction also explains the low RMSD value of the A27/C28 bulge and the low field shifted proton resonances of the two nucleotides are due to the presence of A-A interactions which ensure structural conservation of the bulge.
Chapter 4- NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

**Figure 4.1.18**: (a) Non canonical Watson-Crick base pairing between A9 and A27. (b) The sole hydrogen bond between A9 (N1) and A27 (NH₂) produces an A•A N1-N6 (Amino) symmetric base pair.

Table 4.1.4 includes two non Watson-Crick base pairs: G8-C28 and A9-A27. Canonical base pairings between G-C bases are as a result of three hydrogen bonds; however, G8-C28 produces non canonical base pairing with three weak hydrogen bonds, one between the G8 amino (N2) and C28 ribose oxygen (O2’).
Table 4.1.4: Hydrogen bonding between the canonical and non canonical base pairs of the NMR structure of *B.subtilis* 27mer RNA. BS indicates the hydrogen bond is between the base of one nucleotide and sugar of the second nucleotide.

| Basepair | Hydrogen bonds | | | |
|----------|----------------|----------------|----------------|
| C2-G34   | O2-N2          | 2.85           | N3-N1          | 2.69           | N4-O6          | 2.43           |
| C3-G33   | O2-N2          | 2.52           | N3-N1          | 2.66           | N4-O6          | 2.68           |
| C4-G32   | O2-N2          | 3.28           | N3-N1          | 2.90           | N4-O6          | 2.45           |
| G5-C31   | N2-O2          | 3.06           | N1-N3          | 2.78           | O6-N4          | 2.44           |
| G7-C30   | N2-O2          | 2.70           | N1-N3          | 2.76           | O6-N4          | 2.96           |
| G8-C28   | N2-O2' (BS)    | 5.15           | N1-O2          | 3.64           | O6-N3          | 4.06           |
| A9-A27   | N1-N6          | 2.84           |                |                |                |                |
| G10-C25  | N2-O2          | 2.62           | N1-N3          | 2.71           | O6-N4          | 2.87           |
| C11-G24  | O2-N2          | 2.84           | N3-N1          | 2.79           | N4-O6          | 2.71           |
| U12-A23  | N3-N1          | 2.88           | O4-N6          | 2.69           | O4-N6          | 2.43           |
| U13-A22  | N3-N1          | 2.92           |                |                |                |                |

The sugar puckers of the ribose rings range from C₃-endo to C₄-exo for the majority of the base paired nucleotides, which is typical for A-form helices. The terminal residues (G35, A22, A23) have greater flexibility than those in the duplex and so the sugar puckers can exist in the B-form C₂-endo conformation.
Chapter 4- NMR structure determination of the *B.subtilis* 27mer RNA and its interactions with bamicetin

### Table 4.1.5: Local helical parameters of the NMR structure of *B.subtilis* 27mer RNA.

<table>
<thead>
<tr>
<th>Base pair</th>
<th>Propeller ($\pi$)</th>
<th>Opening ($\sigma$)</th>
<th>Buckle ($\kappa$)</th>
<th>Stagger ($S_z$)</th>
<th>Stretch ($S_y$)</th>
<th>Shear ($S_x$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2-G34</td>
<td>-2.25</td>
<td>-5.75</td>
<td>-12.17</td>
<td>-0.97</td>
<td>-0.60</td>
<td>-0.30</td>
</tr>
<tr>
<td>C3-G33</td>
<td>-4.84</td>
<td>0.31</td>
<td>12.77</td>
<td>-0.42</td>
<td>-0.35</td>
<td>0.10</td>
</tr>
<tr>
<td>C4-G32</td>
<td>7.09</td>
<td>-20.88</td>
<td>-2.57</td>
<td>-1.50</td>
<td>-0.67</td>
<td>-0.89</td>
</tr>
<tr>
<td>G5-C31</td>
<td>6.50</td>
<td>-12.78</td>
<td>2.29</td>
<td>-0.97</td>
<td>-0.50</td>
<td>0.16</td>
</tr>
<tr>
<td>G7-C30</td>
<td>27.02</td>
<td>-4.13</td>
<td>32.35</td>
<td>-0.40</td>
<td>-0.24</td>
<td>0.68</td>
</tr>
<tr>
<td>G8-C29</td>
<td>3.71</td>
<td>11.02</td>
<td>8.04</td>
<td>-1.02</td>
<td>0.37</td>
<td>-3.00</td>
</tr>
<tr>
<td>A9-A27</td>
<td>1.03</td>
<td>15.47</td>
<td>-1.50</td>
<td>-1.05</td>
<td>1.01</td>
<td>2.76</td>
</tr>
<tr>
<td>G10-C25</td>
<td>-15.72</td>
<td>2.39</td>
<td>8.65</td>
<td>0.11</td>
<td>-0.27</td>
<td>-0.10</td>
</tr>
<tr>
<td>C11-G24</td>
<td>-14.65</td>
<td>-0.25</td>
<td>13.26</td>
<td>-0.82</td>
<td>-0.29</td>
<td>-0.08</td>
</tr>
<tr>
<td>U12-A23</td>
<td>7.97</td>
<td>-8.96</td>
<td>-12.55</td>
<td>-0.46</td>
<td>-0.15</td>
<td>-0.48</td>
</tr>
<tr>
<td>U13-A22</td>
<td>11.35</td>
<td>-15.27</td>
<td>-3.39</td>
<td>0.04</td>
<td>-0.20</td>
<td>0.15</td>
</tr>
</tbody>
</table>

There are three local base pair helical parameters which are critical in characterising the base pairings of the motif; Shear, Stretch and Opening (Table 4.1.5). For an A-RNA conformation, Shear values have a range of $0.01 \pm 0.23$; it can be observed the values for the bases surrounding the bulges and terminal bases diverge from the standard range due to the increased flexibility in these regions of the motif. In particular, G8-C29 base pair has a value of $-3.00$, although the G8 is base paired to C29 and satisfies the secondary structure, it also interacts with C28 in the bulge by means of three hydrogen bonds (Table 4.1.4).

The values for the Opening parameter are much larger than the standard value of $0.6 \pm 2.8$, this is due to the two bulges widening the helix and thus increasing the distance between the Watson-Crick base pairs. The Stagger parameter describes the non-planarity of a base pair ($0.02 \pm 0.25$) and the NMR structure contains values which satisfy this range with only base pairs C11-G24 and U12-A23 which digress at $-0.82$ and $-0.46$, respectively.
Table 4.1.6: Local helical parameters of the NMR structure of *B.subtilis* 27mer RNA.

<table>
<thead>
<tr>
<th>Step</th>
<th>Shift (Dx)</th>
<th>Slide (Dy)</th>
<th>Rise (Dz)</th>
<th>Tilt (T)</th>
<th>Roll (ρ)</th>
<th>Twist (Ω)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1-C2/G34-G35</td>
<td>1.37</td>
<td>-1.26</td>
<td>6.17</td>
<td>-18.72</td>
<td>-0.59</td>
<td>42.07</td>
</tr>
<tr>
<td>C2-C3/G33-G34</td>
<td>0.89</td>
<td>-1.01</td>
<td>2.91</td>
<td>8.96</td>
<td>11.13</td>
<td>32.37</td>
</tr>
<tr>
<td>C3-C4/G32-G33</td>
<td>0.89</td>
<td>-1.01</td>
<td>2.91</td>
<td>8.96</td>
<td>11.13</td>
<td>32.37</td>
</tr>
<tr>
<td>C4-G5/C31-G32</td>
<td>-1.13</td>
<td>-1.72</td>
<td>3.18</td>
<td>3.80</td>
<td>17.49</td>
<td>14.97</td>
</tr>
<tr>
<td>G5-U6/C30-C31</td>
<td>1.08</td>
<td>-0.97</td>
<td>4.56</td>
<td>-7.92</td>
<td>15.50</td>
<td>31.96</td>
</tr>
<tr>
<td>U6-G7/C29-C30</td>
<td>-3.08</td>
<td>-1.97</td>
<td>2.83</td>
<td>-1.96</td>
<td>18.27</td>
<td>9.54</td>
</tr>
<tr>
<td>G7-G8/C28-C29</td>
<td>2.29</td>
<td>0.38</td>
<td>2.16</td>
<td>22.87</td>
<td>3.44</td>
<td>44.41</td>
</tr>
<tr>
<td>G8-A9/A27-C29</td>
<td>0.56</td>
<td>-1.53</td>
<td>6.85</td>
<td>-32.55</td>
<td>23.49</td>
<td>27.02</td>
</tr>
<tr>
<td>A9-G10/C25-A27</td>
<td>2.53</td>
<td>-0.46</td>
<td>3.89</td>
<td>1.56</td>
<td>4.56</td>
<td>48.16</td>
</tr>
<tr>
<td>G10-C11/G24-C25</td>
<td>-0.05</td>
<td>-2.03</td>
<td>4.23</td>
<td>-11.41</td>
<td>23.19</td>
<td>24.01</td>
</tr>
<tr>
<td>C11-U12/A23-G24</td>
<td>-0.16</td>
<td>-1.47</td>
<td>2.60</td>
<td>9.89</td>
<td>17.62</td>
<td>22.67</td>
</tr>
<tr>
<td>U12-U13/A22-A23</td>
<td>-0.33</td>
<td>-1.35</td>
<td>3.96</td>
<td>-12.84</td>
<td>15.20</td>
<td>30.47</td>
</tr>
</tbody>
</table>

Table 4.1.6 reports the local base pair helical parameters which describe the position and orientation of one base pair relative to another. The values deviated from the standard A-RNA range are indicated in red.

The parameters which show whether the motif adopts an A-RNA or B-DNA conformation are the base pair helical parameters: **x-displacement** and **helical twist** (not shown). The majority of the base pairs satisfied the standard range (standard value $-4.17 \pm 1.22$) of the x-displacement parameter. The helical twist is the angle of rotation about the helical axis that brings successive base pairs into coincidence and the standard value is $32.5 \pm 3.8$ for an A-form RNA. Apart from the base pairs adjacent to the internal loops, the remaining Watson-Crick base pairs possessed values which satisfy the standard range.
Chapter 4- NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

4.2 27mer RNA motif of *B. subtilis* rRNA-bamicetin complex

4.2.1 $^1$H NMR titration

NMR titrations of the *B. subtilis* 27mer RNA with bamicetin were carried out as described in section 2.1.1. Changes to the imino region were the first to be monitored as the imino proton resonances appear in a non overlapped region of the NMR spectrum. Mapping these changes is important as the protons are involved in base pairing and changes caused to the peaks indicate the secondary structure of the RNA is affected by the antibiotic.

Progressive titration was carried out with the antibiotic and 1D $^1$H NMR spectra were measured of RNA:bamicetin ratios of 1:0, 1:0.3, 1:0.6, 1:1, 1:3 and 1:10.

![Figure 4.2.1](image)

**Figure 4.2.1:** (a) 600MHz 1D $^1$H NMR spectrum of *B. subtilis* 27mer RNA (0.59mM), (b-f) 400MHz $^1$H NMR spectra of *B. subtilis* 27mer RNA containing varying amounts of bamicetin (in molar equivalent). All spectra obtained in 90% $^1$H$_2$O +10% $^2$H$_2$O, pH 6.2 at 2°C. The G7 (12.50ppm), G10 (13.65ppm) and U6 (10.72ppm) imino proton resonances (indicated by the black vertical lines) were monitored.
Figure 4.2.1 displays the imino proton region of the $^1$H NMR spectra with the previously assigned *B. subtilis* 27mer RNA imino proton resonances (section 4.1.3). All the imino proton resonance peaks undergo broadening with successive addition of bamicetin but with maximum effect on G7, G10 and U6. The broadening of resonances indicates weakened hydrogen bonds and hence disruption to the canonical Watson-Crick base pairing.

The unpaired U6 imino proton is broad and only visible in the 600MHz spectrum of the free RNA (Figure 4.2.1a); it disappears after the first 0.3 equivalents of the antibiotic. Two imino proton resonances were selected for analysis, G7 and G10; both are situated in the middle of the RNA motif and predicted to be the antibiotic binding pocket. Both peaks broaden at 1:1 equivalents, on further addition of the antibiotic (3 molar equivalents) the peaks are very broad and they are barely visible at 10 equivalents of bamicetin. The broadening of the two peaks is a consequence of weak hydrogen bonding and this was exemplified by measuring NOESY spectra of the 1:1 and 1:10 complex in $^1$H$_2$O (Figure 4.2.2). The amino NH$_2$ protons of G10 disappear in the 1:1 complex (Figure 4.2.2b) and its base paired C25 residue protons are also affected. At 1 molar equivalent of bamicetin, the NH$_2$ protons of C25 are very weak and in the presence of excess of antibiotic (10 molar equivalents) (Figure 4.2.2c) only one NH$_2$ proton resonance is almost visible.

This trend was also observed with G7 and its base paired C30 amino proton resonances (Figure 4.2.2). The NH$_2$ proton resonances of both nucleotides are weakened but still possess good intensity in the 1:1 complex, however, upon addition of 10 molar equivalents of bamicetin, only the G7 NH$_2$ proton resonance is observed with very weak intensity while the C30 amino proton resonances have disappeared. The NMR assignment in the green lettering in Figure 4.2.2(b-c) are additional NOE correlations observed in the complex spectra and absent in the free *B. subtilis* 27mer RNA spectrum (Figure 4.2.2.a).
Chapter 4- NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

The uniform broadening resonances and no emergence of new bound peaks indicate that the binding kinetics belong to a fast-to-intermediate rate of exchange. The large concentration of bamicetin required to diminish the intensities of the imino proton resonances suggests a weak binding system.
Figure 4.2.2: 600MHz $^1$H-$^1$H NOESY spectra (a) *B.subtilis* 27mer RNA (0.59mM), (b) *B.subtilis* 27mer RNA-bamicetin (1:1) complex, (c) *B.subtilis* 27mer RNA-bamicetin (1:10) complex. All spectra obtained in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. The NOE correlations between the imino protons of G10 and G7 to their respective base paired amino protons of C25 and C30 are highlighted in the green shaded boxes.
Chapter 4- NMR structure determination of the *B.subtilis* 27mer RNA and its interactions with bamicetin

The G7 and G10 imino proton peaks were selected as representative examples for line width measurement and analysis. The line width of the resonances was taken at each titration point and plotted against the ligand concentration at 2°C.

![Figure 4.2.3](image)

**Figure 4.2.3:** The changes in the line widths of the G10 and G7 imino proton resonances of *B.subtilis* 27mer RNA with incremental addition of bamicetin. G7 produced a characteristic binding curve, an important result since this base is adjacent to the unpaired U6 nucleotide.

Both plots illustrate the larger changes to the line width occur by 3 molar equivalents of antibiotic with further addition (10 equivalents) not having a significant change to the line width. The G7 graph indicates the system has reached the saturation point at 3 equivalents of bamicetin and G10 at 4 equivalents. The 3-4 molar equivalents of bamicetin required signifies the binding system is weak-intermediate and justifies the observation made in the 1D $^1$H NMR spectra of the imino and amino protons. The binding curves produced in the line width plots can be correlated with the binding plots described in section 2.7.3 (Figure 2.7.2) in support of the affinity of bamicetin for the 27mer RNA motif.

The amino/aromatic region of the 1D $^1$H NMR spectra of the titrations (Figure 4.2.4) illustrated changes to the proton resonances and complements the observations made about the G7 and G10 nucleotides. The aromatic proton resonances of C25 (H6) and C30 (H6) overlap in the spectra at 7.60-7.61ppm, but interestingly in the presence of the antibiotic, the intensity of the resonance reduces and at 1:3 is not visible. This is consistent with the bamicetin induced missing amino protons peaks of C25 and C30 (Figure 4.2.2)
and strongly indicates that these nucleotides may be involved in binding with the bamicetin antibiotic.

Additional nucleotides which displayed changes to the aromatic and amino proton resonances included G8-H8, G24-H2(2)/A27-H1’, A27-H8 and C31-H4(2) and are displayed in Figure 4.2.4. The G8, G24, C31 residues neighbour G7, C25 and C30, respectively, and as a result this changes the local environment and so changes in the chemical shift values are observed.
Figure 4.2.4: (a) 400MHz 1D $^1$H NMR spectrum of B.subtilis 27mer RNA (0.59mM), (b-f) 400MHz 1D $^1$H NMR spectra of B.subtilis 27mer RNA containing varying amounts of bamicetin (in molar equivalents). All spectra carried out in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. Changes to the aromatic protons of specific nucleotides are indicated by the black vertical lines.
Chapter 4- NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

The A27 nucleotide resides in the bulge neighbouring G8, C25 and U26 and hence in the binding site area. Therefore, it is important to map the changes to the A27 proton resonance line shapes. A27-H1' proton overlaps with the G24-H2(2), nevertheless the addition of bamicetin reduces the peak and this is also observed with the well resolved A27-H8 proton resonance which at 1:10 ratio has reduced considerably in intensity.

A TOCSY spectrum was obtained of the 27mer RNA motif in the presence of 10 molar equivalents of bamicetin (Figure 4.2.5).

![TOCSY spectra](image)

**Figure 4.2.5:** 600MHz $^1$H-$^1$H TOCSY spectra of (a) *B. subtilis* 27mer RNA (0.59mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C and (b) *B. subtilis* 27mer RNA-bamicetin (1:10) complex in 100% $^2$H$_2$O, pH 6.2 at 25°C.

The U6, C25 and U26 nucleotides have small chemical shift changes and supports the observations made in the 1D $^1$H NMR spectra of the aromatic region of the RNA (Figure 4.2.2). The C25 nucleotide is perturbed by the antibiotic and the effect is relayed by the disappearance of the amino proton resonances (Figure 4.2.2) and the shift in the H5-H6 resonance (Figure 4.2.5). The analysis of the TOCSY spectrum and the aromatic proton resonances strongly indicate that the antibiotic is interacting predominantly in the middle region of the 27mer RNA motif. The U6 and U26 H5-H6 resonances were further
analysed by NOE growth curves (Figure 4.2.6). The U26 resonance increases steadily as the mixing time is increased, however in the case of the U6 the NOE builds up rapidly at 100ms before decreasing very quickly at longer mixing times ($\tau_m = 200\text{ms}$ and $400\text{ms}$). The U6 nucleotide relaxes very quickly due to neighbouring protons and thus the NOE fails to develop like the U26.

**Figure 4.2.6:** NOE growth curves for the U6 and U26 H5-H6 resonances of *B. subtilis* 27mer RNA

The $^{31}\text{P}$ resonances were probed to observe changes with the backbone of the RNA and the results are consistent with the imino and aromatic protons with broadening and disappearing of several peaks (Figure 4.2.7).
Chapter 4 - NMR structure determination of the *B. subtilis* 27mer RNA and its interactions with bamicetin

**Figure 4.2.7:** 81MHz 1D $^{31}$P spectra of (a) *B. subtilis* 27mer RNA (0.59mM) and (b) *B. subtilis* 27mer RNA-bamicetin (1:10) complex, both carried out in 100% $^2$H$_2$O, pH 6.2 at 27°C. The considerable broadening of the resonances suggests a change in conformation of the phosphorus backbone of the RNA upon bamicetin binding.

Figure 4.2.7 shows the 1D $^{31}$P NMR spectra of the *B. subtilis* 27mer RNA and in the presence of excess (10 molar equivalents) bamicetin. There is a considerable amount of change in the phosphorus peaks, in particular to G5, U6, G7, G8 and can be correlated with the broadened imino peaks. The changes to the phosphate backbone of the 27mer RNA indicate interactions between the nucleic acid motif and the antibiotic.
Figure 4.2.8: 400MHz 1D $^1$H NMR spectra of (a) bamicetin (5.5mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 4.8 at 2°C, (b) *B.subtilis* 27mer RNA (0.59mM), (c-g) 400MHz 1D $^1$H NMR spectra of 27mer *B.subtilis* RNA containing varying amounts of bamicetin (in molar equivalents). All spectra (b-g) carried out in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C.
The ligand proton resonances were analysed for the observation of intermolecular effects with the RNA. Figure 4.2.8(a) displays the assigned aliphatic region of the 1D $^1$H NMR spectra of bamicetin. The protons of $B.subtilis$ 27mer RNA do not resonate in the 1.0-3.8ppm region and so this region is a clear window for the observation of the bamicetin aliphatic protons during the RNA titration. At 1:1 complex, some bamicetin peaks are appearing and as the bamicetin concentration is increased to 3 molar equivalents, more bamicetin proton resonances are visible and this is consistent with the previous observation that 3 equivalents of the ligand are required to saturate the system. All of the bamicetin proton resonances are observed in the 1D $^1$H NMR spectrum and the 2D NOESY spectrum of the 1:10 complex (not shown).

The data presented unambiguously provides evidence for binding of the bamicetin antibiotic to the 27mer $B.subtilis$ RNA motif. The broadening of the G7 and G10 imino protons and disappearance of their base paired amino protons designates the binding area. This was substantiated by the changes to the pyrimidine U6 and U26 H5-H6 resonances. The large amount of antibiotic required (3-4 molar equivalents) strongly suggests that the binding complex is a weak affinity system.
Chapter 5  NMR structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic

The 25mer RNA motif resides in the GTPase hydrolysis domain II region of the *H. hal* 23S rRNA and generates a thiostrepton-resistant mutant (G1159, *H. hal* numbering). The thiostrepton resistant mutant (A1159G) displays cross-hypersensitivity effects with the PT antibiotic amicetin; likewise the amicetin resistant mutant (U2457) situated in a stem region of the central circle of domain V has affinity for the thiostrepton antibiotic. The two mutation sites are in different domains, and separated by hundreds of nucleotides in their primary structures. Interaction studies were conducted with the 25mer RNA motif containing the thiostrepton-resistant mutation and the amicetin antibiotic. Their mode of interaction at the tertiary structural level is important for understanding the cross-hypersensitivity phenomenon.

5.1  NMR assignment of the *H. hal* 25mer RNA motif of 23S rRNA

5.1.1  Secondary structure

The *H. hal* 25mer RNA motif is a hairpin and its secondary structure is endowed with a symmetric four membered internal loop and a nine membered hairpin loop held together in a stable helical stem by six base pairs (Figure 5.1.1A). Also, the stable conformation of the 25mer RNA motif was predicted by the web based MC-Sym computational program. The predicted stem loop structure is displayed in Figure 5.1.1B and reveals a closure base pairing between U9 and A17 to produce a seven membered hairpin loop. The formation of the U9-A17 base pair is possible; being adjacent to the stem they are likely to possess both base pairing and non base pairing capabilities.
Chapter 5- NMR structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic

### Figure 5.1.1: (A) The secondary structure of the 25mer stem-loop motifs of the GTPase domain II *H. hal* 23S rRNA. The thiostrepton resistant mutant base G11 (G1159 *H. hal* numbering) is highlighted by the red box. (B) The predicted secondary structure and the UV\textsubscript{260} absorbance melting temperature ($T_m = 67.6^\circ$C) from the Mc-Sym program. The motif is labelled numerically instead of the four digit numbering found in the *H. hal* ribosome.

### 5.1.2 Assignment of the exchangeable proton resonances

A 1D $^1$H NMR spectrum in $^1$H$_2$O solvent (Figure 5.1.2) displayed all the proton resonances of the *H. hal* 25mer RNA to initiate the NMR assignment. The imino protons appear lowest field in a non overlapped region of the spectrum (10-15ppm) which is not shared by any other proton group signals.
Figure 5.1.2: (A) 1D $^1$H NMR spectrum of the $H$.hal 25mer RNA (0.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. (B) Expanded imino region with the canonical, Watson-Crick base paired imino protons appearing in the characteristic chemical shift range (12-15ppm) and broader resonances of the unpaired and non canonical imino protons resonating in the highfield (10-12ppm) region.
Imino proton resonances

The chemical shift values observed for the imino NH proton resonances in the 1D $^1$H NMR spectrum (Figure 5.1.2B) verify the base pairing of the nucleotides. Those present in the 12-15ppm region are canonical Watson-Crick base pairs and appear with greater intensity owing to the stability conferred by the hydrogen bonds while weaker and broader resonances appear higher field (10-12ppm) and these are due to non canonical imino protons on unpaired nucleotides.

The 25mer RNA motif consists of 6 base paired imino protons (2 uracils and 4 guanines) and 7 unpaired imino protons (3 uracils and 4 guanines) and inspection of the 1D $^1$H imino region reveals seven resonances in the base paired region and five peaks in the non base paired region suggesting that there is the possibility of an additional base pair.

A $^1$H-$^1$H NOESY spectrum in $^1$H$_2$O solvent was carried out to carry out the sequence specific assignment of the labile imino NH protons. In regions of A-form helical RNA, neighbouring NH protons in the spectrum are ~4Å apart and thus give rise to NOE cross peaks. Figure 5.1.3 shows the imino region of the NOESY spectrum which displays imino-imino connectivities between some of the base paired nucleotides. The first NH resonance assignment was that of the U20 which appears at 13.32ppm and displayed a strong correlation to the cross strand G7 (13.60ppm). The imino resonance of G2 appearing at 12.77ppm exhibited NOE to the sequential G3-NH resonance (13.10ppm).

The imino resonance of the terminal U25 was assigned to the resonance at 13.96ppm. G18 base (13.42ppm) is adjacent to the G7 (13.60ppm) and displayed an imino-imino connectivity. The U9 residue is represented as a non base paired nucleotide in the secondary structure; however, its NH proton resonance (13.24ppm) appears in the base paired region of the NOESY spectrum. This observation indicates the likelihood that U9 being involved in base pairing in the tertiary structure.
Figure 5.1.3: 800MHz $^1$H-$^1$H NOESY ($\tau_m = 250$ms) spectrum of the *H.hal* 25mer RNA (0.6mM) in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. The plot displays the assigned resonances of the base paired labile imino protons.

Amino proton resonances

It was not possible to assign the unpaired NH protons of the loops to specific nucleotides, since they did not produce any NOEs to the amino NH$_2$ or aromatic protons. Identification of the NH protons enabled the assignment of the labile NH$_2$ resonances (Figure 5.1.4). The NH protons of base pairs produced strong NOE correlations to the intrabase NH$_2$ protons. The assignment of the NH protons was also confirmed by correlating the nonlabile aromatic proton resonances in the NOESY $^1$H$_2$O spectrum to the same aromatic protons assigned separately in the NOESY $^2$H$_2$O spectrum. The NH$_2$ protons resonate in the 6-9ppm region and the chemical shift values distinguish the hydrogen bonded geminal amino protons. Further, the labile NH and NH$_2$ proton NOEs provided important distance based constraints for the generation of the NMR structure.
Chapter 5- NMR structure determination of the \textit{H.hal} 25mer RNA and its interactions with amicetin antibiotic

The NH$_2$ protons in RNA nucleobases are observed with less intensity than the NH protons, due to broadening from exchange with the solvent protons and rotation about the C-N bond. However, as can be observed in Figure 5.1.4 below, the cytosine NH$_2$ protons are usually observed with sharp intensity due to slower exchange in comparison to the guanine and adenine NH$_2$ protons which resonate in the intermediate/fast exchange regime and are therefore generally severely broadened.

\textbf{Figure 5.1.4:} 800MHz $^1$H-$^1$H NOESY ($\tau_m=250\text{ms}$) spectrum of \textit{H.hal} 25mer RNA (0.6mM) in 90\% $^1$H$_2$O + 10\% $^2$H$_2$O, pH 6.2 at 2°C. The assignment of the imino protons was confirmed in the amino region with correlations to the exchangeable amino and non exchangeable ribose and base protons.
Chapter 5- NMR structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic

5.1.3 Assignment of the non-exchangeable proton resonances

The assignment of the labile NH and NH$_2$ protons is important in establishing the base pairing and base stacking of the RNA motif, while the sequence specific assignments of non-labile aromatic and sugar protons help to confirm the secondary structure of the motif.

The identification of the non exchangeable proton resonances was achieved by performing a combination of 2D and 3D experiments. The scalar $^1$H-$^1$H TOCSY spectrum was acquired for the assignment of pyrimidine H5-H6 connectivities. The $^{13}$C chemical shift dispersion achieved in the $^1$H-$^{13}$C HSQC experiment provided the unambiguous identification of the ribose proton resonances and also in distinguishing the uracil from cytosine H5 resonances. The sugar puckering of the ribose rings was determined by the 2QC connectivities in the homonuclear 3D NOE/2QC experiment. In addition to the sugar puckering, the 3D experiment exhibited both scalar and dipolar correlations and provided further assignment of the ribose protons (*ribose proton resonance* section) and independently confirmed the assignment obtained from the 2D experiments.

The $^1$H NMR spectrum in $^2$H$_2$O in Figure 5.1.5 shows the different RNA proton groups in the expected chemical shift ranges for the 25mer RNA, analogous to the identification of the various proton groups shown in Figure 4.1.2.
Chapter 5- NMR structure determination of the *H.hal* 25mer RNA and its interactions with amicetin antibiotic

**Figure 5.1.5:** 800MHz 1D $^1$H NMR spectrum of the *H.hal* 25mer RNA (0.6mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C.

*Pyrimidine H5, H6 resonances*

The $^1$H NMR spectrum in $^2$H$_2$O (Figure 5.1.5) is a better quality spectrum than in $^1$H$_2$O solvent as the small residual $^1$H$^2$HO solvent signal is effectively suppressed leading to higher sensitivity.
The H5-H6 connectivities appear with strong intensities due to the three bond scalar coupling and the short covalent distance (2.4Å) between the two protons (Figure 5.1.6A). The assignment was achieved in conjunction with the HSQC spectrum (Figure 5.1.6B/C) where the dispersion in $^{13}$C chemical shifts of cytosine and uracil C5 caused the resonances to appear in separate regions of the NMR spectrum. The assignments were verified in the 3D NOE/2QC spectrum and they were subsequently transferred to the NOESY spectrum to calibrate the intensities of the NOE resonances for input into the structure determination calculation.

Figure 5.1.6: (A) 600MHz $^1$H-$^1$H TOCSY ($\tau_{m}$=75ms) spectrum displaying the pyrimidine H5-H6 correlated resonances of the H.hal 25mer RNA (0.6mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C. (B) and (C) 600MHz $^1$H-$^{13}$C HSQC in 100% $^2$H$_2$O, pH 6.2 at 25°C panels displaying the cytosine and uracil $^1 J$ H5-C5 correlations of the cytosine and uracil residues.
Chapter 5- NMR structure determination of the \textit{H.hal} 25mer RNA and its interactions with amicetin antibiotic

\textit{Ribose proton resonances}

The $^1\text{H}$-$^{13}\text{C}$ HSQC spectrum was the initial spectrum that was used to identify the sugar proton resonances and the observations were transposed and verified in the TOCSY and NOESY spectra. Figure 5.1.7 displays one bond C1’-H1’ cross peaks in the C1’ region ($^{13}\text{C} = 85-95$ ppm), the assignment of the ribose anomeric protons was transferred to the $^2\text{H}_2\text{O}$ NOESY spectrum to resolve any ambiguities and to carry out the intraresidual H6/H8-H1’ assignment.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5_1_7.png}
\caption{600MHz $^1\text{H}$-$^{13}\text{C}$ HSQC spectrum of the \textit{H.hal} 25mer RNA (0.6mM) in 100% $^2\text{H}_2\text{O}$, pH 6.2 at 25°C. The plot displays the assigned H1’-C1’ correlated resonances.}
\end{figure}
3D experiment

The 3D NOE/2QC experiment was carried out to obtain the assignment of the ribose proton resonances, since the majority of the resonances occur in a narrow chemical shift window (~1ppm) and are difficult to identify in the 2D spectra. The 3D planes remove the overlap of the resonances that is common in 2D spectra by dispersing the peaks into the third dimension. This results in resolution of ambiguities in the 2D spectra as well as gaining additional proton resonance assignments.

The combination of the scalar (2QC) and dipolar (NOE) sequences of the experiment resulted in correlations observed across the nucleotide. For example, in Figure 5.1.8 the U25 H6 proton produces correlations to the U25 H5 proton as well as to the intranucleotide ribose protons (U25-H2’/H3’/H4’). Similar connectivities were observed with the other nucleotides including C19, C8 and C16 (Figure 5.1.8). The 3D experiment was valuable in securing the assignment of the residues in the hairpin loop of the 25mer RNA motif, the NOEs of which usually appear weak in the 2D spectra.
Chapter 5 - NMR structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic

**Figure 5.1.8:** 600MHz 3D NOE/2QC spectrum of *H. hal* 25mer RNA (0.6mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C. (a) U26-H6 to H5, H2, H3’ and H4’. (b) C19-H6 to H5, H1’ and H3’. (c) C8-H6 to H5, H2’ and H3’. (d) C16-H2’ to H3’, H4’ and H5’.

*Sequence specific assignment*

The strategy applied to the *B. subtilis* 27mer duplex RNA (chapter 4) was used to carry out the sequential assignment of the 25mer single stranded RNA and Figure 5.1.9 displays the sequential H6/H8-H1’ assignments.
Chapter 5- NMR structure determination of the H.hal 25mer RNA and its interactions with amicetin antibiotic

Figure 5.1.9: (A) 600MHz $^1$H-H NOESY spectrum ($\tau_m$=250ms) of H.hal 25mer RNA (0.6mM) in 100% $^2$H$_2$O, pH 6.2 at 25°C illustrating the H6/H8-H1’ sequential assignment. The intraresidue (H6/H8$_n$-H1$_n$”) connectivities are represented by blue circles and the internucleotide (H6/H8$_{n+1}$-H1$_{n’}$) correlations are indicated by the green circles. (B) 600MHz $^1$H-$^{13}$C correlated HSQC spectrum shows the H6-C6/H8-C8 correlated spectrum. (C) 600MHz $^1$H-$^{13}$C correlated HSQC spectrum displays the H1’-C1’ region.
The intranucleotide (H8\textsubscript{n}-H1\textsubscript{n}’) connectivity of the terminal nucleotide (A1) was assigned to the resonance at 8.37-6.03 ppm as it contained only one peak in this region. The A1 (H1’) proton displayed a sequential NOE cross peak to a resonance at 7.48 ppm which was assigned to the succeeding base G2. Similarly this procedure was repeated for the entire 25 nucleotide RNA motif as illustrated in Figure 5.1.9. There were some overlapping H1’ and H5 peaks which were resolved and confirmed in the scalar TOCSY (Figure 5.1.6) and HSQC (Figure 5.1.7) spectra.

The sequential H6/H8-H1’ assignment was verified by carrying out a similar assignment in the NOESY spectrum between the aromatic H6/H8 and ribose H2’ protons. The H2’ resonances in the NOESY spectrum were confirmed by the assignment carried out in the ribose regions of the 3D NOE/2QC and 2D TOCSY spectra (not shown).

A 2D \textsuperscript{1}H-\textsuperscript{31}P CPMG-HSQC-NOESY (\(\tau_m = 500\) ms) spectrum was acquired to assign the backbone phosphorus resonances. The appearance of the \textsuperscript{31}P nuclei in the chemical shift region -1 to -5 ppm identified the 25mer RNA to have A-form geometry. The presence of scalar coupled \textsuperscript{31}P-H3’/H5’/H5’’ resonances allowed some phosphorus nuclei to be resolved (data not shown). Unfortunately, dipolar connectivities to the H1’ protons or aromatic were not observed and so it was not possible to conduct a separate sequential H6/H8-H1’ assignment by this backbone route.
5.1.4 Assignment table

Table 5.1.1: $^1$H, $^{13}$C and $^{31}$P chemical shifts ($\delta$, ppm) of the *H.hal* 25mer RNA (0.6mM). Labile imino and amino proton resonances were measured in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C and non labile aromatic and sugar proton resonances measured in 100% $^2$H$_2$O, pH 6.2 at 25°C. All $^1$H assignments referenced to the residual water peak according to temperature. $^{113}$ $^{13}$C and $^{31}$P assignment at 25°C were referenced directly to $^1$H.

| H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 | H9 | H10 | H11 | H12 | H13 | H14 | H15 | H16 | H17 | H18 | H19 | H20 | H21 | H22 | H23 | H24 | H25 |
|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 | 7.48 |
| 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 |
| 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 | 6.08 |
| 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 | 7.18 |
| 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 | 6.06 |

-235-
5.1.5 Geometrical constraints

The CcpNMR analysis software was utilised to obtain the cross peak volumes from the NOESY (τ_m = 250ms) spectrum and converted them into interproton distance constraints. The NOE distance dependent constraints were generated from the exchangeable and non-exchangeable proton resonances and empirically classified into distance bands with upper bounds, based on their cross peak intensities (described in section 2.4.2) and shown in Table 5.1.2. NOEs derived from imino protons were treated with caution as their ability to exchange with solvent water protons can distort peak intensities and to account for any spin diffusion, larger error bounds were placed. However, the presence of strong imino-imino cross peaks of sequential (G7-U20 and G2-G3) base pairs provided distance based NOE constraints which helped align the base pairs in the structure determination (Figure 5.1.3).

Torsion angle constraints were applied to constrain the ribose puckers (ν_1, ν_2, χ) and the backbone phosphate angles (α, β, γ, ε, ζ) to establish the backbone orientation. The sugar puckers were inferred from the scalar H_1″-H_2″ assignment in the 3D NOE/2QC spectrum. Nucleotides possessing $^3J_{H1″-H2″} > 7\text{Hz}$ were constrained C_2-endo ($145° \pm 20°$) and those with $^3J_{H1″-H2″} < 3\text{Hz}$ were as C_3′-endo ($80° \pm 20°$). The glycosidic torsion angle (χ) defines the orientation of the base relative to the ribose and is either in the anti or syn conformation. The 25mer RNA nucleotides were constrained anti conformation (-150° ± 90°) if intranucleotide NOEs between the aromatic H6/H8 and sugar H1′ protons were observed. The backbone torsion angles were restrained to A-form values when the ribose adopted the C_3-endo conformation ($α = -68° \pm 17°$, $β = 178° \pm 17°$, $γ = 54° \pm 17°$, $ε = -153° \pm 17°$, $ζ = -71° \pm 17°$). The final structure contained all five backbone angles constrained for the base paired stem nucleotides and the backbone angle ζ was excluded for the nucleotides in the hairpin loop as its inclusion resulted in an increase in constraint violations.

Hydrogen bond restraints were placed on nucleotides displaying imino-imino and/or imino-amino NOE connectivities and database orientation restraints were also added to
define the planarity of the base pairs, and hence improve the accuracy of the generated structures.

Table 5.1.2 summarises the experimental constraints used as input into the structure determination calculation to generate the NMR structure of the *H. hal* 25mer RNA. A total of 113 NOE based distance constraints, 197 dihedral angle constraints and 9 hydrogen bond constraints were used.

**Table 5.1.2**: A summary of the experimental constraints used as input into the structure determination calculation of the *H. hal* 25mer RNA.

<table>
<thead>
<tr>
<th>Constraint type</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE distances</td>
<td>113</td>
</tr>
<tr>
<td>Strong</td>
<td>32</td>
</tr>
<tr>
<td>1.8 – 2.5Å</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>46</td>
</tr>
<tr>
<td>1.8 – 3.3Å</td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td>29</td>
</tr>
<tr>
<td>1.8 – 5Å</td>
<td></td>
</tr>
<tr>
<td>Very weak</td>
<td>6</td>
</tr>
<tr>
<td>1.8 – 7Å</td>
<td></td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>9</td>
</tr>
<tr>
<td>Dihedral angles</td>
<td>197</td>
</tr>
<tr>
<td>Glycosidic (χ)</td>
<td>25</td>
</tr>
<tr>
<td>Ribose pucker (ν₁, ν₂, δ)</td>
<td>75</td>
</tr>
<tr>
<td>Helix (α, β, γ, ε, ζ)</td>
<td>97</td>
</tr>
<tr>
<td>NOEs per residue</td>
<td>4.5</td>
</tr>
<tr>
<td>Constraints per residue</td>
<td>13</td>
</tr>
<tr>
<td><strong>Total constraints</strong></td>
<td>319</td>
</tr>
</tbody>
</table>
Chapter 5 - NMR structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic

### 5.1.6 NMR solution structure of the *H. hal* 25mer RNA

The structure calculation protocol, as described in section 2.6, was followed to generate 200 simulated annealed and 200 refined structures. A rigorous acceptance protocol was applied to identify the converged NMR structures amongst the refined structures. Figure 5.1.10 shows the NMR structure of the *H. hal* 25mer RNA, characterised by an average global RMSD value of 0.67Å amongst the ten best low energy structures.

![Figure 5.1.10](image)

**Figure 5.1.10**: A. The NMR structure of the *H. hal* 25mer RNA with zero constraint violations and low RMSD 0.50Å. B. An overlay of the ten best converged NMR structures with an average all atom RMSD 0.67Å.

As can be seen in Figure 5.1.10, the residues of the hairpin loop display less conservation of geometries between the structures demonstrating the dynamic behaviour of the loop and Figure 5.1.11 displays the nucleotides A14 to A17 involved in base stacking, thus increasing the stability of the hairpin loop.
Figure 5.1.11: Conformation of the nanoloop motif of *H. hal* 25mer RNA as shown in (A) the NMR structure. (B) a schematic illustration; black circles, sugar moiety; blue circles, phosphate group; blue box, base; red filled box, base stacking; blue dotted line, hydrogen bond; black solid lines, glycosidic and phosphodiester bonds.

Seven Watson-Crick and three non-canonical base pairs (U9-A17, G11-C16 and A13-G15) are observed in the NMR structure and are reported in Table 5.1.3.

Table 5.1.3: Summary of the hydrogen bonds observed within the base pairs.

<table>
<thead>
<tr>
<th>Base pair</th>
<th>Hydrogen bonds</th>
<th>N1-N3</th>
<th>N6-O4</th>
<th>N1-N3</th>
<th>O6-N4</th>
<th>2.82</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1-U25</td>
<td>2</td>
<td>N1-N3</td>
<td>3.06</td>
<td>N6-O4</td>
<td>3.07</td>
<td></td>
</tr>
<tr>
<td>G2-C24</td>
<td>3</td>
<td>N2-O2</td>
<td>2.79</td>
<td>N1-N3</td>
<td>2.85</td>
<td>O6-N4</td>
</tr>
<tr>
<td>G3-C23</td>
<td>3</td>
<td>N2-O2</td>
<td>3.22</td>
<td>N1-N3</td>
<td>3.07</td>
<td>O6-N4</td>
</tr>
<tr>
<td>G5-C22</td>
<td>3</td>
<td>N2-O2</td>
<td>2.75</td>
<td>N1-N3</td>
<td>2.92</td>
<td>O6-N4</td>
</tr>
<tr>
<td>A6-U20</td>
<td>2</td>
<td>N1-N3</td>
<td>2.67</td>
<td>N6-O4</td>
<td>2.53</td>
<td></td>
</tr>
<tr>
<td>G7-C19</td>
<td>3</td>
<td>N2-O2</td>
<td>2.55</td>
<td>N1-N3</td>
<td>2.80</td>
<td>O6-N4</td>
</tr>
<tr>
<td>C8-G18</td>
<td>4</td>
<td>O2-N2</td>
<td>3.51</td>
<td>N3-N2</td>
<td>3.45</td>
<td>N3-N1</td>
</tr>
<tr>
<td>U9-A17</td>
<td>2</td>
<td>O2*N1</td>
<td>5.12</td>
<td>N3*N6</td>
<td>4.66</td>
<td></td>
</tr>
<tr>
<td>G11-C16</td>
<td>1</td>
<td>O6-N4</td>
<td>5.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A13-G15</td>
<td>2</td>
<td>N3*N7</td>
<td>3.91</td>
<td>N1*O6</td>
<td>3.15</td>
<td></td>
</tr>
</tbody>
</table>
The thiostrepton resistant mutant nucleotide G11 (green shading in Table 5.1.3) is involved in one hydrogen bond with the unpaired loop nucleotide C16. The O6-N4 non-canonical base pairing is also known as a GC carbonyl-amino base pair and is present in ten structure logs in the database of non-canonical RNA structures. The Figure 5.1.12 below displays the non W-C base pairs observed in the NMR structure and indicated in Table 5.1.3.

**Figure 5.1.12**: A. The NMR structure with the non-canonical base pairs labelled and residues colour coded to match the secondary structure displayed in B. B. The secondary structure of the 25mer RNA indicates the number of hydrogen bonds (red dashed lines) which are present in each of the non canonical base pair.

Non canonical Watson-Crick base pairing is observed between the U9 and A17 nucleotides. The base pairing is composed of weak base-base hydrogen bonds arising from U9(O2)-A17(N1) and U9(N3)-A17(N6). Although the two bases are fairly planar, the weak hydrogen bonds are unable to prevent the two residues from kinking upwards. A G-A sheared base paring between G15-A13 is produced by two hydrogen bonds; G15(N7)-
A13(N3) and G15(carbonyl)-A13(N1). Both types of base pairing (U9-A17 and G15-A13) are unusual and were not found in the structures deposited in the database. Full geometrical analyses of the NMR structure were obtained from the w3DNA and CURVES programs.

To characterise the hydrogen bonding, the influence of three base pair parameters; Shear, Stretch and Opening is critical (Table 5.1.4). The Shear values obtained for the canonical Watson-Crick base pairs are close to standard values, however the largest deviations were observed for the non canonical and terminal base pairs. The values for the Stretch and Opening parameters on the whole are quite high due to the presence of the unpaired residues in the loops which increased the distance between the nucleotides involved in base pairing.
Table 5.1.4: Local base-pair and base-pair step parameters for the 25mer RNA. Values deviating from the standard A-form RNA values range are highlighted in red.

<table>
<thead>
<tr>
<th>Base pair</th>
<th>Propeller (ω)</th>
<th>Opening (σ)</th>
<th>Buckle (κ)</th>
<th>Stagger (Sz)</th>
<th>Stretch (Sy)</th>
<th>Shear (Sx)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1-U25</td>
<td>12.68</td>
<td>-6.17</td>
<td>2.35</td>
<td>-0.83</td>
<td>-0.08</td>
<td>-0.18</td>
</tr>
<tr>
<td>G2-C24</td>
<td>-2.63</td>
<td>-0.58</td>
<td>0.08</td>
<td>-0.98</td>
<td>-0.42</td>
<td>-0.38</td>
</tr>
<tr>
<td>G3-C23</td>
<td>26.80</td>
<td>-9.08</td>
<td>-18.05</td>
<td>-1.26</td>
<td>-0.13</td>
<td>-0.15</td>
</tr>
<tr>
<td>G5-C22</td>
<td>18.77</td>
<td>25.12</td>
<td>7.74</td>
<td>1.61</td>
<td>-0.56</td>
<td>-1.06</td>
</tr>
<tr>
<td>A6-U20</td>
<td>4.15</td>
<td>-5.89</td>
<td>1.38</td>
<td>-0.05</td>
<td>-0.32</td>
<td>0.13</td>
</tr>
<tr>
<td>G7-C19</td>
<td>6.43</td>
<td>1.07</td>
<td>-2.96</td>
<td>-1.08</td>
<td>-0.38</td>
<td>-0.15</td>
</tr>
<tr>
<td>C8-G18</td>
<td>28.68</td>
<td>-19.76</td>
<td>-6.18</td>
<td>-1.24</td>
<td>-0.25</td>
<td>-0.80</td>
</tr>
<tr>
<td>U9-A17</td>
<td>42.91</td>
<td>-35.31</td>
<td>-26.37</td>
<td>-1.53</td>
<td>0.84</td>
<td>2.51</td>
</tr>
<tr>
<td>G11-C16</td>
<td>60.70</td>
<td>63.86</td>
<td>14.82</td>
<td>-1.36</td>
<td>2.74</td>
<td>-0.79</td>
</tr>
<tr>
<td>A13-G15</td>
<td>-34.56</td>
<td>-31.70</td>
<td>-54.45</td>
<td>1.33</td>
<td>2.32</td>
<td>4.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Base pair step</th>
<th>Shift (Dx)</th>
<th>Slide (Dy)</th>
<th>Rise (Dz)</th>
<th>Tilt (τ)</th>
<th>Roll (ρ)</th>
<th>Twist (Ω)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG/CU</td>
<td>0.50</td>
<td>-1.69</td>
<td>3.82</td>
<td>-1.24</td>
<td>13.76</td>
<td>25.61</td>
</tr>
<tr>
<td>GG/CC</td>
<td>0.00</td>
<td>-1.75</td>
<td>4.38</td>
<td>5.90</td>
<td>0.25</td>
<td>29.74</td>
</tr>
<tr>
<td>GG/CC</td>
<td>1.59</td>
<td>-2.05</td>
<td>4.86</td>
<td>3.11</td>
<td>19.16</td>
<td>33.58</td>
</tr>
<tr>
<td>GA/UC</td>
<td>-0.11</td>
<td>-2.40</td>
<td>5.46</td>
<td>-9.45</td>
<td>12.42</td>
<td>44.53</td>
</tr>
<tr>
<td>AG/CU</td>
<td>1.07</td>
<td>-1.65</td>
<td>3.69</td>
<td>7.49</td>
<td>4.38</td>
<td>29.99</td>
</tr>
<tr>
<td>GC/GC</td>
<td>-0.88</td>
<td>-1.65</td>
<td>3.87</td>
<td>8.67</td>
<td>12.00</td>
<td>23.22</td>
</tr>
<tr>
<td>CU/AG</td>
<td>-1.07</td>
<td>-1.45</td>
<td>5.94</td>
<td>-2.73</td>
<td>6.37</td>
<td>37.83</td>
</tr>
<tr>
<td>UG/CA</td>
<td>-3.23</td>
<td>-1.03</td>
<td>5.83</td>
<td>30.87</td>
<td>38.43</td>
<td>18.67</td>
</tr>
<tr>
<td>GA/GC</td>
<td>-3.34</td>
<td>0.44</td>
<td>0.89</td>
<td>77.18</td>
<td>81.82</td>
<td>33.73</td>
</tr>
</tbody>
</table>
The Buckle and Propeller parameters (Table 5.1.4) represent the planar geometry of a helical structure and the fluctuated values from the standard are due to the input of a smaller number of hydrogen bond constraints which try to keep the base pairs planar.

The non canonical base pairs present in the hairpin loop resulted in structural changes observed for that region and thus produced deviated values. The large flexible hairpin loop contains nine bases and the dynamic behaviour is indicated in the large Opening, Propeller, Tilt and Stagger values. The analysis of the base pair step (Table 5.1.4) and helical parameters (Table 5.1.5) reveals similar behaviour, with deviations for the loop and terminal residues observed.

**Table 5.1.5**: Local base-pair helical parameters for the 25mer RNA. Numbers in red show deviation from standard A-form RNA values; h, helical; disp, displacement; Inclination (η), Tip (θ).

<table>
<thead>
<tr>
<th>Base pair helical</th>
<th>X-disp</th>
<th>Y-disp</th>
<th>h-Rise</th>
<th>η</th>
<th>θ</th>
<th>h-Twist</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG/CU</td>
<td>-6.52</td>
<td>-1.28</td>
<td>2.56</td>
<td>28.56</td>
<td>2.56</td>
<td>29.04</td>
</tr>
<tr>
<td>GG/CC</td>
<td>-3.41</td>
<td>1.59</td>
<td>4.29</td>
<td>0.47</td>
<td>-11.36</td>
<td>30.31</td>
</tr>
<tr>
<td>GG/CC</td>
<td>-6.29</td>
<td>-1.99</td>
<td>3.38</td>
<td>30.23</td>
<td>-4.90</td>
<td>38.65</td>
</tr>
<tr>
<td>GA/UC</td>
<td>-4.62</td>
<td>-1.10</td>
<td>4.60</td>
<td>15.82</td>
<td>12.04</td>
<td>47.05</td>
</tr>
<tr>
<td>AG/CU</td>
<td>-4.02</td>
<td>-0.38</td>
<td>3.57</td>
<td>8.25</td>
<td>-14.11</td>
<td>31.19</td>
</tr>
<tr>
<td>GC/GC</td>
<td>-6.50</td>
<td>4.04</td>
<td>2.30</td>
<td>26.62</td>
<td>-19.22</td>
<td>27.49</td>
</tr>
<tr>
<td>CU/AG</td>
<td>-3.63</td>
<td>1.01</td>
<td>5.69</td>
<td>9.72</td>
<td>4.17</td>
<td>38.44</td>
</tr>
<tr>
<td>UG/CA</td>
<td>-4.37</td>
<td>6.22</td>
<td>-0.49</td>
<td>54.99</td>
<td>-44.17</td>
<td>52.50</td>
</tr>
<tr>
<td>GA/GC</td>
<td>0.46</td>
<td>1.78</td>
<td>-1.55</td>
<td>57.49</td>
<td>-54.22</td>
<td>115.74</td>
</tr>
</tbody>
</table>

Slide (Table 5.1.4) and X-displacement parameters (Table 5.1.5) characterise the helical conformation. For an A-form geometry, the standard value of Slide is -1.53 (± 0.34) and X-displacement is 4.17 (± 1.22). The base pair residues of the *H.hal* 25mer RNA structure yield values which fall within the standard range.
Chapter 5- NMR structure determination of the *H*. *hal* 25mer RNA and its interactions with amicetin antibiotic

5.2 25mer motif of *H*. *hal* rRNA-amicetin complex

5.2.1 $^1$H NMR titration

The NMR titration of *H*. *hal* 25mer RNA with amicetin was carried out as described in section 2.1.1. The amicetin was progressively titrated into the RNA in 0.5, 1, 2 and 5 molar equivalents. The $^1$H NMR titration spectra revealed the RNA was saturated at 5 equivalents of amicetin and so no further additions were carried out.

![NMR spectra](image)

**Figure 5.2.1:** 400MHz 1D $^1$H NMR spectra of (a-c) *H*. *hal* 25mer RNA (0.6mM) in its complex with amicetin in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C. (a) no ligand, (b) 1:0.5 and (c) 1:1.

The first region of the 1D $^1$H NMR spectra inspected for changes was the imino NH proton resonance region. Figure 5.2.1 shows the imino proton resonances of 25mer *H*. *hal* RNA and the small changes induced to the NH protons in the presence of the amicetin antibiotic. This observation gave the first insight that the binding is weak. No additional peaks in the imino region corresponding to bound nucleotides were present.
Chapter 5- NMR structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic

Changes to the U4, G7, U9 imino proton resonances were observed in Figure 5.2.1. The resonances broaden quite significantly with the addition of the antibiotic. The G7 and U9 nucleotides are involved in base pairing as shown in the NMR structure determined in section 5.1.7 and the broadening of their peaks represent weakening of the base pairs in the presence of amicetin.

The labile amino NH$_2$ proton region was also examined to follow the changes as observed with the imino protons (Figure 5.2.2). The addition of a small amount of amicetin (0.5 equivalents, Figure 5.2.2b) severely affects the amino proton peaks with the majority of them disappearing. This indicates that the antibiotic has disrupted the hydrogen bonds between base pairs and hence the secondary structure of the RNA. The G7 is base paired to C19 nucleotide and upon addition of 0.5 equivalents of amicetin, the amino protons of C19 disappear and this is also comparable with the weakening of A17 NH$_2$ proton resonances which are hydrogen bonded to U9. However, U9 is less affected by amicetin with the presence of the very weak A17-NH$_2$ proton resonance at 5 molar equivalents of the antibiotic. At 1:5 ratio, only the imino proton resonances of U9 and G18 remain thereby indicating that the region encompassing the internal loop of the RNA is involved in interactions with amicetin.

The increased exchange between the imino protons and the water solvent protons became apparent after the first addition of bamicetin (0.5 equivalents) when strong cross peaks at 5ppm were observed (Figure 5.2.2b). This implies that the introduction of the antibiotic disrupted the base pairing of the 25mer RNA motif and allowed the imino protons to exchange rapidly with the water solvent protons.
Chapter 5- NMR structure determination of the *H.hal* 25mer RNA and its interactions with amicetin antibiotic

Figure 5.2: 400MHz $^1$H-$^1$H NOESY ($\tau_m = 250$ms) spectra (a) *H.hal* 25mer RNA (0.6mM), (b) *H.hal* 25mer-amicetin (1:0.5) complex, (c) 1:1 complex, (d) 1:2 complex and (e) 1:5 complex. All spectra carried out in 90% $^1$H$_2$O + 10% $^2$H$_2$O, pH 6.2 at 2°C.
Linewidth measurements were produced for the U4, G7 and U9 imino proton resonances for further analysis of these three nucleotides (Figure 5.2.3). The steady increase in line widths with increasing amicetin reconfirms the changes noted to the imino and amino regions of the NMR spectra and provides evidence for antibiotic binding at or near these residues.

**Figure 5.2.3:** 400MHz 1D $^1$H NMR line width plots of the imino proton resonances U4, G7 and U9 measured at 2°C.

The relationship observed in the line width plots can be correlated with the binding isotherms in section 2.7.2 (Figure 2.7.2). U4 and U9 indicate non-specific binding while G7 displays a total binding curve. The graphs illustrate weak non-specific binding.

Following the changes to the labile proton resonances, the non-exchangeable proton resonances were monitored for amicetin binding. Figure 5.2.4 is the sequential region of
the free *H.hal* 25mer RNA and in the presence of excess amicetin. As observed in Figure 5.2.4.B, the sequential region was very difficult to analyse after the addition of amicetin. It was not possible to elucidate a sequential scheme as was the case with the free 25mer RNA motif. This effect is similar to the previously observed broadening of the amino proton resonances which signify that the antibiotic after binding has affected the tertiary structure of the motif.

![Figure 5.2.4: A. 600MHz ¹H-¹H NOESY spectrum (τ_m =250ms) of *H.hal* 25mer RNA (0.6mM) in 100% ²H₂O, pH 6.2 at 25°C with no amicetin illustrating the H6/H8-H1” sequential assignment. B. 1GHz ¹H-¹H NOESY spectrum (τ_m =250ms) of *H.hal* 25mer RNA-amicetin (1:5) complex in 100% ²H₂O at 25°C displaying the amicetin induced broadening of the NOE cross peaks.](image)

However, the intranucleotide (H6/H8-H1”) NOE peaks were observed for the A14, G15, C16, A17 nucleotides indicating that the hairpin loop is sufficiently stable even after the addition of amicetin. No internucleotide NOE peaks between the RNA motif and antibiotic were observed in this region.

To observe the effect on the RNA’s backbone, the ³¹P resonances were also monitored at each titration point (Figure 5.2.5). The broad resonances make it difficult to assign the
peaks to individual nucleotides. However, there are some distinct changes with each addition of amicetin and this complements the changes observed before to the NH and NH₂ proton resonances.

**Figure 5.2.5**: 162MHz 1D $^{31}$P NMR spectra (a) *H.hal* 25mer RNA (0.6mM), (b,c) *H.hal* 25mer RNA containing varying amounts of amicetin (in molar equivalents) in 100% $^2$H₂O at 25°C.

Figure 5.2.6 is the 1D $^1$H NMR spectrum of the *H.hal* 25mer RNA in the presence of 5 molar equivalents of amicetin at 2°C. Due to this excess the proton resonances of the antibiotic appear at their normal unbound chemical shifts with much greater intensity than the RNA peaks, thus confirming that the system has passed the saturation point.
Chapter 5- NMR structure determination of the *H. hal* 25mer RNA and its interactions with amicetin antibiotic

![Figure 5.2.6](image)

**Figure 5.2.6**: 1GHz $^1$H NMR spectrum of *H. hal* 25mer RNA-amicetin complex (1:5) in 100% $^2$H$_2$O, pH 6.2 at 2°C.

The amicetin proton resonances were analysed to detect the presence of any intermolecular NOEs to the RNA. There were more intramolecular NOE peaks observed between amicetin protons in contrast to the bamicetin protons. In Figure 5.2.7, dipolar connectivities are observed between the antibiotic aromatic protons (H14, H15, H16 and H17) and saccharide (II) protons. Bamicetin displays NOEs only between H13 and H14, but amicetin possesses NOE cross peaks to all four aromatic protons. The majority of the NOE peaks seen in the Figure 5.2.7 below for amicetin are not observed in the antibiotic’s free state. This signifies that there is rapid exchange of the antibiotic between its bound and free chemical states.
Chapter 5 - NMR structure determination of the \textit{H.hal} 25mer RNA and its interactions with amicetin antibiotic

Figure 5.2.7: 1GHz $^1$H-$^1$H NOESY (\(\tau_m = 250\)ms) spectra of A. 25mer \textit{H.hal} RNA-amicetin (1:5) complex in 100\% $^2$H$_2$O, pH 6.2 at 25°C and B. \textit{B.subtilis} 27mer RNA-bamicetin (1:10) complex in 100\% $^3$H$_2$O, pH 6.2 at 25°C. evidence between free and bound ligand?
The absence of intermolecular NOE cross peaks between the RNA and the amicetin protons, the disappearance of RNA amino proton resonances and the disruption to the sequential assignment (H6/H8_{n+1}-H1'_{n}) after the addition of amicetin provide evidence that there are interactions between the two molecular components.
Chapter 6  Conclusion and Future work

Conclusion
The main aim of the project was to elucidate the mode of interaction between a structurally important 27mer RNA motif of \textit{B.subtilis} 23S ribosomal RNA and a \textit{peptidyl transferase} inhibitor antibiotic, bamicetin. Further, the project gave an opportunity to investigate a thiostrepton-resistant \textit{H.hal} 25mer RNA mutant involved in a biologically significant phenomenon \textit{viz.} cross-hypersensitivity, in three dimensional structural terms. This project was investigated using state of the art high field NMR spectrometers (600MHz, 700MHz and 800MHz) equipped with latest accessories. Access was gained to the ultra highfield 1GHz NMR instrument at Lyon (France) to probe the RNA-antibiotic interactions.

Antibiotics
The NMR assignment of the three antibiotics bamicetin, sparsomycin and anisomycin was successfully carried out and their three dimensional structures determined. There are no published X-ray crystal structures of the antibiotics in their uncomplexed free states and so this study is the first insight revealing the conformations of the molecules in the native solution state.

The structures of all three antibiotics possessed low RMSD and low potential energy values indicating stable conformations. Bamicetin adopted a well-folded conformation possibly held by intramolecular hydrogen bonds. The exchange retarded amide protons most likely to be involved in intramolecular hydrogen bonding in sparsomycin were stable at high temperatures of 50°C and in the case of bamicetin they were stable up to 25°C. The retention of hydrogen bonds even at high temperatures signifies the conformational stability of the antibiotics which may be essential for binding to the conserved RNA motifs.

The modified cytosine and uracil bases of bamicetin and sparsomycin, respectively, are both well exposed in the NMR structures and signify the important role of these bases in the antibiotics as they are able to interact favourably with the ribosome. The sulphur tail of
sparsomycin is extended in both the NMR and X-ray crystal structures of 50S ribosomal RNA-sparsomycin complex. The determined NMR structures of all three antibiotics should help provide information for generating modified antibiotics, since clinical trials were aborted due to high toxicity of the original antibiotics.

**RNA motifs**

The complete NMR assignment and structure determination were carried out on the *B.subtilis* 27mer RNA and the *H.hal* 25mer RNA motifs. The project provided an option of studying a double stranded RNA and a hairpin RNA and the structure determination of both motifs revealed stable A-form helices with low structural RMSD values and good overlay of structures.

The 27mer RNA structure contained a non canonical A-A base pairing between the A9 and the A27 residues. The A9 base was previously thought to be involved in base pairing with U26 but the NMR structure revealed this not to be the case, instead it forms an A·A N1-N6 symmetric base pair with A27. The 27mer RNA structure revealed the unpaired U6 to be flipped inside the duplex and its position stabilised by the surrounding canonical G-C base pairs. The NMR structure of the 25mer RNA motif displayed non canonical base pairing and base stacking between several nucleotides (A14 to A17) in the hairpin loop providing evidence for the stability of the loop region of the RNA motif.

The 27mer and 25mer RNA converged structures demonstrate the stability of the small RNA motifs although they are a very small part of the intact ribosome.

**RNA-antibiotic complexes**

The 27mer RNA-bamicetin and 25mer-amicetin complex data were acquired on the only 1GHz NMR spectrometer. The clear, well resolved and dispersed resonances indicate the experiments were carried out under the best possible conditions with excellent water suppression and at the highest spectrometer frequency.
Two different RNA-antibiotic models were studied with both RNA motifs residing in two separate domains of the 23S rRNA. The \textit{B.subtilis} 27mer-bamicetin complex contained the predicted binding site of the antibiotic and the \textit{H.hal} 25mer -amicetin system with the antibiotic displaying hypersensitivity to the G11 base of the 25mer RNA motif.

Changes to the NMR spectra of 27mer RNA indicated the binding region and analysis revealed it to be near to the A27 bulge, very close to the site of mutation that confers amicetin resistance (U26C). The \textit{B.subtilis}-bamicetin complex displayed greater exchange with water as the concentration of the ligand increased. The absence of some nucleotide peaks- i.e. U26 labile proton resonances indicated this base was affected more strongly than the other nucleotides.

A most important result of this research project was the cross-hypersensitivity effect observed between the thiostrepton-resistant \textit{H.hal} 25mer RNA and amicetin antibiotic. It is the first time that this phenomenon has been structurally probed and the results in chapter 5 clearly represent affinity between the two molecular components. Due to the binding being in the fast-intermediate rate it was not possible to observe intermolecular NOEs which would have distinguished the specific RNA residues and antibiotic protons involved in binding. Despite using a small but well conserved RNA motif, the amicetin binding results were positive and justify further studies with the intact ribosome.

**Future Work**

The project can be advanced by adopting a number of routes. The stability of the small RNA motifs and their affinities for the antibiotics signifies additional experiments can be carried out to further probe the complexes and determine their binding site. A \textsuperscript{19}F-substituted \textit{H.hal} 29mer RNA motif containing isotopically labelled \textsuperscript{19}F on U6, U26 and U30 residues could be titrated with the amicetin antibiotic. All three nucleotides are situated around the internal loops and the predicted binding pocket of the antibiotic. The imino proton resonances of the three \textsuperscript{19}F modified residues would be monitored by


$^{19}$F-NMR and changes to the peaks would confirm the region of binding with the antibiotic.

To observe the bound bamicetin proton resonances in the $B$.subtilis 27mer RNA complex, a Filtered NOE experiment would be carried out. Changes to the bound proton resonances would be monitored and these would be used as distance constraints to generate the conformation of the antibiotic in the bound state with the RNA. An additional advantage is that any possible intermolecular NOEs between the bound antibiotic proton and RNA would also be observed and not overlapped or obscured by unbound proton resonances.

The bamicetin and amicetin antibiotics can be isotopically labelled with $^2$H, $^{13}$C and $^{15}$N and NMR binding studies carried out with the 70S $B$.subtilis and $H$.hal ribosomes. This would be very exciting research to carry out since no antibiotic-ribosome complexes have been studied via NMR spectroscopy. Changes to the isotopically labelled protons will be observed in $^1$H-$^{13}$C HSQC and $^1$H-$^{15}$N HSQC spectra. Along with the changes to the $^2$H nuclei, the enhancement in the $^{13}$C and $^{15}$N resonances of the antibiotic will locate the binding site in the ribosome.
References


2 C. Shekhar, Bacteria: Drug resistance spreads, but few new drugs emerge, *Chemistry and Biology.*, 2010, 17, 413-414.


23 T.H. Haskell, Amicetin, bamicetin and plicacetin, chemical studies, contribution from the research division, Parke, Davis and Co, 1958, 747-751.


32 J.L. Hansen, P.B. Moore, T.A. Steitz, Structures of five antibiotics bound at the peptidyl transferase centre of the large ribosomal subunit, *J.Mol.Biol.*, 2003, **330**, 1061-1075.


J. Keeler, in *Understanding NMR spectroscopy*, John Wiley and Sons Ltd. 2006, ch.1, pp. 5-15.


76 A. Bax and D.G. Davis, MLEV based two dimensional homonuclear magnetisation transfer spectroscopy, *J. Magn. Reson.*, 1985, 65, 355-360.


T.D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco.


http://www.ghostscript.com


-264-


http://www.gaussian.com

http://gabedit.sourceforge.net


D.R. Hare, B.R. Reid, Three-dimensional structure of a DNA hairpin in solution: Two-dimensional NMR studies and distance geometry calculations on d(CGCGTTTTCGCG), *Biochemistry.*, 1986, **25**, 5341-5350.


Appendices

Appendix I-Experimental Scripts

(i) NMR data processing script

**NMRpipe script**

*Typical bruker conversion*

```
bruk2pipe -in /home/misbah/Desktop/MNdata_1_02_08/22/ser \
-bad 0.0 -noaswap -DMX -decim 12 -dspfvs 12 -grpdly -1 \
-xN 6144 -yN 800 \
xT 2999 -yT 400 \
xMODE DQD -yMODE States-TPPI \
xSW 13227.513 -ySW 13227.513 \
xOBS 599.927 -yOBS 599.927 \
xCAR 4.971 -yCAR 4.971 \
xLAB 1Hx -yLAB 1Hy \
-ndim 2 -aq2D States \ 
-out ./test.fid -verb -ov
```

Typical varian conversion

```
var2pipe -in /fid -noaswap \ 
-xN 4096 -yN 600 \ 
xT 2048 -yT 300 \ 
xMODE Complex -yMODE Complex \ 
xSW 6600.116 -ySW 6600.116 \
```
-xOBS 599.893 -yOBS 599.893 \
-xCAR 4.773 -yCAR 4.773 \
-xLAB H1x -yLAB H1y \
-ndim 2 -aq2D States \
-out ./test.fid -verb -ov 

*Typical water suppression 2D processing script*

```
nmrPipe -verb -in /home/misbah/Desktop/MNdata_060608/tocsys_75ms.fid/test.fid \
| nmrPipe -fn SOL -fl 94 \
| nmrPipe -fn POLY -time \
| nmrPipe -fn GMB -lb -1.8 -gb 0.08 -c 0.8 \
| nmrPipe -fn ZF -auto \
| nmrPipe -fn FT -auto \
| nmrPipe -fn PS -p0 65.0 -p1 21.0 -di \
| nmrPipe -fn TP \
| nmrPipe -fn SP -off 0.4 -end 1.0 -c 0.6 \
| nmrPipe -fn ZF -auto \
| nmrPipe -fn FT -auto \
| nmrPipe -fn PS -p0 25.0 -p1 -23.6 -di \
| nmrPipe -fn POLY -auto -ord 0 \
| nmrPipe -fn TP \
| nmrPipe -ov -verb -out \
/home/misbah/Desktop/MNdata_060608/tocsys_75ms.fid/bamicetin_toc75.ft2
```
**NMR draw script**

csh
source /disk1/NMRPipe/com/nmrInit.linux9.com
source /disk1/NMRPipe/com/font.com
if (-e /disk1/NMRPipe/com/nmrInit.linux9.com) then
source /disk1/NMRPipe/com/nmrInit.linux9.com
endif
if (-e /disk1/NMRPipe/dynamo/com/dynInit.com) then
source /disk1/NMRPipe/dynamo/com/dynInit.com
endif
nmrDraw

**Sparky conversion**

2D NMR spectra processed on NMRpipe and saved as test.ft*. To assign the spectra, the scripts were converted to the ucsf format. The conversion was carried out at the terminal using the command /usr/local/sparky/bin/pipe2ucsf followed by the file name (test.ft*) and its new name (uscf).

(ii) Structure calculation scripts

**Antibiotics-simulated annealing**

xplor19

evaluate ($init_t = 1500)
evaluate ($high_steps = 100000)
evaluate ($cool_steps = 3000000)

parameter @TOPPAR:parallhdgimprbodge.dna end
structure @bam.psf end

set echo=false end
set message=off end
set echo=off end

restraints plane
@plane.tbl
?
end

flags include plan end
set message=on end

set echo=on end

noe
nres=600
class=all
@NOE_bam.tbl
end

restraints dihedral

nassign=50

@Bamtor.tbl
end

vector do (fbeta=10) (all)
vector do (mass=100) (all)

noe
ceiling=100
averaging * cent
potential * square
sqconstant * 1
 sqexponent * 2
 scale * 60.
end

parameter
nbonds
repel=0.5
rexpl=2 irexpl=2 rcon=1.
nbxmod=-2
wmin=0.01
cutnb=4.5 ctonnb=2.99 ctofnb=3.
tolerance=0.5
end
end

restraints dihedral
scale=5
end

evaluate ($end_count=100)
evaluate ($count=0)
while ($count < $end_count ) loop main
evaluate ($count=$count+1)
evaluate ($filename="random_"+encode($count)+".pdb")
for $image in ( 1 - 1 ) loop imag
  coor initialize end
  coor @$filename
  vector do (x=x * $image) ( known )
  vector identity (store1) (not known)  {*Set store1 to unknowns.*}

  {* ============================= Minimization of bonds, VDWs, and NOEs.*}
  restraints dihedral   scale=5. end
  parameter nbonds nbxmod=-2 repel=0.5 end end
  flags exclude * include bond vdw noe cdih end
  constraints interaction (all) (all) weights * 1. vdw 20. end end
  minimize powell nstep=100 nprint=10 end

  {* ================================================== Include angles. *}
  flags include angl end
  minimize powell nstep=1000 nprint=10 end
  flags include impr plan end
  evaluate ($nstep1 = int($high_steps/8))
  evaluate ($nstep2 = int($high_steps/2))
  constraints inter (all) (all) weights * 0.1 impr 0.05 vdw 20. end end
  dynamics  verlet
    nstep=$nstep1 time=0.001 iasvel=maxwell firstt=$init_t
tcoup=true tbath=$init_t nprint=100 iprfrq=0
  end
  constraints inter (all) (all) weights * 0.2 impr 0.1 vdw 20. end end
  dynamics  verlet
    nstep=$nstep1 time=0.001 iasvel=current firstt=$init_t
tcoup=true tbath=$init_t nprint=100 iprfrq=0
  end
  parameter nbonds repel=0.9 end end
  constraints inter (all) (all) weights * 0.2 impr 0.2 vdw 0.01 end end
  dynamics  verlet
    nstep=$nstep1 time=0.001 iasvel=maxwell firstt=$init_t
tcoup=true tbath=$init_t nprint=100 iprfrq=0
  end
  parameter nbonds nbxmod=-3 end end
  constraints inter (all) (all) weights * 0.4 impr 0.4 vdw 0.003 end end
  dynamics  verlet
    nstep=$nstep2 time=0.001 iasvel=current firstt=$init_t
tcoup=true tbath=$init_t nprint=100 iprfrq=0
  end
  constraints inter (all) (all) weights * 1.0 impr 1.0 vdw 0.003 end end
  dynamics  verlet
    nstep=$nstep1 time=0.001 iasvel=current firstt=$init_t
tcoup=true tbath=$init_t nprint=100 iprfrq=0
dend

if ($image = 1) then
   vector do (store7=x) ( all )        {*Store first image in stores.*}
   vector do (store8=y) ( all )
   vector do (store9=z) ( all )
   vector do (store4=vx) ( all )
   vector do (store5=vy) ( all )
   vector do (store6=vz) ( all )
end if

dend loop imag

energy end
evaluate ($e_minus=$ener)
coor copy end
vector do (x=store7) ( all )
vector do (y=store8) ( all )
vector do (z=store9) ( all )
energy end

{* Uncomment the following lines if a test for the correct*}
{* enantiomer is desired based on an rms difference from a*}
{* reference structure. *}
coor disp=comp @bam.pdb                  {*Read reference structure.*}
coor fit sele=( known ) end
coor rms sele=( known ) end
evaluate ($old_rms=$result)
vector do (x=-x) ( known )
coor fit sele=( known ) end
coor rms sele=( known ) end
if ($result > $old_rms) then
   vector do (x=-x) ( known )
end if

restraints dihedral scale=200. end

evaluate ($final_t = 300)  { K }
evaluate ($tempstep = 50)  { K }
evaluate ($ncycle = ($init_t - $final_t)/$tempstep)
evaluate ($nsstep = int($cool_steps/$ncycle))
evaluate ($ini_rad = 0.9)   evaluate ($fin_rad = 0.75)
evaluate ($ini_con=  0.003)       evaluate ($fin_con=  4.0)
evaluate ($bath  = $init_t)
evaluate ($k_vdw = $ini_con)
evaluate ($k_vdwfact = ($fin_con/$ini_con)^(1/$ncycle))
evaluate ($radius=    $ini_rad)
evaluate ($radfact = ($fin_rad/$ini_rad)^(1/$ncycle))
evaluate ($i_cool = 0)
while ($i_cool < $ncycle) loop cool
  evaluate ($i_cool=$i_cool+1)
  evaluate ($bath = $bath -$tempstep)
  evaluate ($k_vdw=min($fin_con,$k_vdw*$k_vdwfact))
  evaluate ($radius=max($fin_rad,$radius*$radfact))

parameter nbonds repel=$radius end end
constraints interaction (all) (all) weights * 1. vdw $k_vdw end end

dynamics verlet
  nstep=$nstep timestep=0.0005 iasvel=current firstt=$bath
tcoup=true tbath=$bath nprint=$nstep iprfreq=0
end

evaluate ($critical=$temp/$bath)
  if ($critical > 10. ) then
    display ****&**& rerun job with smaller timestep (i.e., 0.003)
    stop
  end if
end loop cool

minimize powell nstep= 100000 nprint=25 end

print threshold=0.1 noe
  evaluate ($rms_noe=$result)
  evaluate ($violations_noe=$violations)
print threshold=5. cdih
  evaluate ($rms_cdih=$result)
  evaluate ($violations_cdih=$violations)
print thres=0.05 bonds
  evaluate ($rms_bonds=$result)
print thres=5. angles
  evaluate ($rms_angles=$result)
print thres=5. impropers
  evaluate ($rms_impropers=$result)

remarks ===============================================================
remarks            overall,bonds,angles,improper,vdw,noe,cdih
remarks energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih
remarks ==================
=============================================
remarks            bonds,angles,impropers,noe,cdih
remarks rms d: $rms_bonds,$rms_angles,$rms_impropers,$rms_noe,$rms_cdih
remarks ===============================================================
remarks               noe, cdih
remarks violations.: $violations_noe, $violations_cdih
remarks ==============================================================

evaluate ($filename="bamidgsa_"+encode($count)+".pdb")
write coordinates output =$filename end
end loop main

stop

**Antibiotics-refinement script**

xplor19

structure @bam.psf end

parameter @TOPPAR:parallhdgimprbodge.dna end

set echo=false end
set message=off end
set echo=off end

restraints plane
@plane.tbl
?
end

flags include plan end
set message=on end
set echo=on end

evaluate ($knoe = 75.0)
evaluate ($kedi = 50.0)

noe

nrestraints=1000
class=all

@NOE_bam.tbl

scale * $knoe
celling 1000

potential * squarewell
sqconstant * 1.0
sqexponent * 2
average * center
end
flags
include cdih noe
end

restraints dihe

nassign=50

@Bamtor.tbl

scale $kcdi
end

evaluate ($rcon=100)

parameters

nbonds

wmin=1.4
cutnb=8.5
toler 0.3
repel=0.0
rexp=2
irex=2
rcon=$rcon
end

end

constraints interaction (all) (all)

weights

angles 2.0
impropers 2.0
bonds 2.0
end

end

parameter nbonds nbxmod=5 end end

evaluate ($end_count=100)
evaluate ($count=0)

while ($count < $end_count ) loop main

evaluate ($count=$count+1)

evaluate ($filename="bamirefa_"+encode($count)+".pdb")

coordinates @@$filename

minimize powell nstep=100000 drop=10 nprint=500 end

print threshold=0.1 noe

evaluate ($rms_noe=$result)

evaluate ($violations_noe=$violations)

print threshold=5. cdih

evaluate ($rms_cdih=$result)

evaluate ($violations_cdih=$violations)

print thres=0.05 bonds

evaluate ($rms_bonds=$result)

print thres=5. angles

evaluate ($rms_angles=$result)

print thres=5. impropers

evaluate ($rms_impropers=$result)

remarks ===============================================================

remarks energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih

remarks ===============================================================

remarks           noe,  cdih

remarks violations.: $violations_noe, $violations_cdih

remarks ===============================================================

evaluate ($filename="bamiref_"+encode($count)+".pdb")

write coordinates output =$filename end

end loop main

stop
RNA-simulated annealing script

from pdbTool import PDBTool
from xplorPot import XplorPot
from rdcPotTools import create_RDCPot
from varTensorTools import create_VarTensor
import varTensorTools
from ivm import IVM
from potList import PotList
import protocol
from protocol import initMinimize
from ivm import IVM
from xplor import command
import random
from atomAction import SetProperty
from simulationTools import StructureLoop
from vec3 import Vec3
from psfGen import seqToPSF
from xplorPot import XplorPot
from varTensorTools import create_VarTensor
import varTensorTools
from ivm import IVM
from potList import PotList
import protocol
from avePot import AvePot
from simulationTools import MultRamp, StaticRamp, InitialParams, StructureLoop, AnnealIVM, FinalParams
from simulationTools import AnnealIVM
from monteCarlo import randomizeTorsions
from noePotTools import create_NOEPot
xplor.parseArguments()
# this checks for typos on the command-line. User-customized arguments can
# also be specified

command = xplor.command

protocol.initParams("nucleic")
protocol.initTopology("nucleic")

# parameters to ramp up during the simulated annealing protocol
#
rampedParams=[]
highTempParams=[]

init_t = 3500.  # Need high temp and slow annealing to converge
final_t=25
bathTemp=2000

seqToPSF(open('13mer.seq').read(), seqType='rna')
seqToPSF(open('14mer.seq').read(), seqType='rna', startResid=14)
#seqToPSF(open('27mer.seq').read(), seqType='rna', startResid=1)

command("write psf output=Bs27mer.psf end")

for atom in AtomSel("all"):  
    atom.setPos( Vec3(float(atom.index())/10, 
                   random.uniform(-0.5,0.5), 
                   random.uniform(-0.5,0.5)) )  
    pass

protocol.fixupCovalentGeom(useVDW=1,maxIters=100)

pots = PotList()

noex = create_NOEPot("noex", "Bs27ex.tbl")
noex.setPotType("soft")  #if incorrect noes suspected set soft if not set hard
rampedParams.append( MultRamp(2,30.,"noex.setScale( VALUE )") )

noen = create_NOEPot("noen", "Bs27nonex.tbl")
noen.setPotType("soft")  #if incorrect noes suspected set soft if not set hard
rampedParams.append( MultRamp(2,30.,"noen.setScale( VALUE )") )

hbon = create_NOEPot("hbon", "hbon27.tbl")
hbon.setPotType("hard")
rampedParams.append( MultRamp(2,30.,"hbon.setScale( VALUE )") )

hbs = create_NOEPot("hbs", "hsoft27.tbl")
hbs.setPotType("soft")
rampedParams.append( MultRamp(2,30.,"hbs.setScale( VALUE )") )

protocol.initDihedrals("tor27.tbl", scale=5, useDefaults=0)

highTempParams.append( StaticRamp("pots['CDIH'].setScale(10)") )
rampedParams.append( StaticRamp("pots['CDIH'].setScale(200)") )

# set custom values of threshold values for violation calculation
#
pots.add( XplorPot('CDIH') )
pots['CDIH'].setThreshold( 5 )

#xplor.command("@plane16.tbl")

## radius of gyration term
## protocol.initCollapse(Rtarget=10.16)
#pots.append( XplorPot('COLL') )

pots.add( XplorPot("BOND") )
pots.add( XplorPot("DIHE") )
pots.add( XplorPot("ANGL") )
pots.add( XplorPot("IMPR") )

rampedParams.append( MultRamp(0.4,1.0,"pots['ANGL'].setScale(VALUE)"))
rampedParams.append( MultRamp(0.1,1.0,"pots['IMPR'].setScale(VALUE)"))

pots.add( XplorPot("VDW") )
rampedParams.append( MultRamp(0.9,0.78,
    "xplor.command('param nbonds repel VALUE end end')") )
rampedParams.append( MultRamp(.004,4,
    "xplor.command('param nbonds rcon VALUE end end')") )

pots.add(noex)
pots.add(noen)
pots.add(hbon)
pots.add(hbs)
pots.append(AvePot(XplorPot("plan",xplor.simulation)) )

# IVM setup
# the IVM is used for performing dynamics and minimization in
# torsion-angle
# space, and in Cartesian space.
# intergrator = IVM() #create an IVM object
cooling = IVM() #create an IVM object
dyn = IVM()

from selectTools import IVM_groupRigidSidechain
minc = IVM()
protocol.initMinimize(minc,potList=pots)
IVM_groupRigidSidechain(minc)
protocol.cartesianTopology(minc,"not resname ANI")

# object which performs simulated annealing
# from simulationTools import AnnealIVM
cool = AnnealIVM(initTemp =init_t,
    finalTemp=final_t,
    tempStep =12.5,
    ivm=dyn,
    rampedParams = rampedParams)

cart_cool = AnnealIVM(initTemp =init_t,
    finalTemp=25,
    tempStep =12.5,
    ivm=minc,
    rampedParams = rampedParams)

def calcOneStructure( structData ):
randomizeTorsions(dyn)
# initialize parameters for high temp dynamics.
InitialParams( rampedParams )
# high-temp dynamics setup - only need to specify parameters which
# differ from initial values in rampedParams
InitialParams( highTempParams )

# high temperature bit - using only P-P nonbonded terms
protocol.initNBond(repel=1.2,
cutnb=100,
tolerance=45,
selStr="name P")

protocol.initDynamics(dyn,
potList=pots, # potential terms to use
bathTemp=init_t,
initVelocities=1,
finalTime=800,    # stops at 800ps or 8000 steps
numSteps=8000,    # whichever comes first
printInterval=100)

dyn.setETolerance( init_t/100 )  # used to det. stepsize. default: t/1000
   dyn.run()

protocol.initNBond() # reset to include all atoms
# initialize parameters for cooling loop
InitialParams( rampedParams )

# initialize integrator for simulated annealing
# protocol.initDynamics(dyn,
potList=pots,       # at each temp: 100 steps
numSteps=100,     # at each temp: 100 steps
or
finalTime=.2,     # .2ps, whichever is
less
printInterval=100)

# perform simulated annealing
# cool.run()

# final torsion angle minimization
# protocol.initMinimize(dyn,
# printInterval=50)
dyn.run()
protocol.initDynamics(minc,
potList=pots,
numSteps=100,    # at each temp: 100 steps
or
finalTime=.4,       # .2ps, whichever is less
printInterval=100)
cart_cool.run()
# final all- atom minimization
#
protocol.initMinimize(minc,
potList=pots,
dEPred=10)
minc.run()
structData.writeStructure(pots)
simWorld.setRandomSeed( 785 )
outPDBFilename = 'SCRIPT_STRUCTURE.pdb'
StructureLoop(numStructures=100,
pdbTemplate=outPDBFilename,
structLoopAction=calcOneStructure,
genViolationStats=1,
averageTopFraction=0.3, #report stats on best 30% of structs
averageContext=FinalParams(rampedParams),
averageSortPots=[pots['BOND'],pots['ANGL'],pots['IMPR'],
one, noex, pots['CDIH'], hbon, hbs],
averagePotList=pots).run()

RNA- refinement script

from pdbTool import PDBTool
from xplorPot import XplorPot
from rdcPotTools import create_RDCPot
from varTensorTools import create_VarTensor
import varTensorTools
from ivm import IVM
from potList import PotList
from protocol import initMinimize
from ivm import IVM
from xplor import comm
from random import random
from atomAction import SetProperty
from simulationTools import StructureLoop
from vec3 import Vec3
from psfGen import seqToPSF
from xplorPot import XplorPot
from ivm import IVM
from potList import PotList
import protocols
from avePot import AvePot
from simulationTools import MultRamp, StaticRamp, InitialParams,
StructureLoop, AnnealIVM
from simulationTools import AnnealIVM
xplor.parseArguments()
# this checks for typos on the command-line. User-customized arguments can also be specified.

```python
command = xplor.command
from noePotTools import create_NOEPot

protocol.initParams("nucleic")
protocol.initTopology("nucleic")

seed=10
numberOfStructures=200
startStructure=0
outFilename = "SCRIPT_STRUCTURE.pdb"

rampedParams=[]

init_t=2000
final_t=25
bathTemp=2000

startFile="annealBs27mer_45.pdb"

simWorld.setRandomSeed(seed)

seqToPSF(open('13mer.seq').read(), seqType='rna')
seqToPSF(open('14mer.seq').read(), seqType='rna', startResid=14)

command("write psf output=27mer.psf end")

# starting coords
# protocol.initCoords(startFile)
protocol.covalentMinimize()

# list of potential terms used in refinement
pots = PotList()
crossTerms=PotList('cross terms') # can add some pot terms which are not refined against- but included in analysis

noex = create_NOEPot("noex",
"Bs27mer_ex.tbl")
noex.setPotType("soft") #if incorrect noes suspected
rampedParams.append( MultRamp(0.2,30.,"noex.setScale( VALUE )") )

noen = create_NOEPot("noen",
"Bs27nonex.tbl")
noen.setPotType("soft") #if incorrect noes suspected
rampedParams.append( MultRamp(2,30.,"noen.setScale( VALUE )") )
```
hbon = create_NOEPot("hbon",
    "hbon27.tbl")
hbon.setPotType("hard")
hbon.setScale(1000)
rampedParams.append( MultRamp(0.2,30.,"hbon.setScale( VALUE )") )

hbs = create_NOEPot("hbs",
    "hsoft27.tbl")
hbs.setPotType("soft")
hbon.setScale(1000)
rampedParams.append( MultRamp(0.2,30.,"hbs.setScale( VALUE )") )

protocol.initDihedrals("tor27_a.tbl",
    scale=5)        #initial force constant
pots.append(AvePot(XplorPot,"cdih") )
rampedParams.append( StaticRamp("pots['CDIH'].setScale(200)"") )

protocol.initRamaDatabase('nucleic')
pots.append(AvePot(XplorPot,"rama") )
rampedParams.append( MultRamp(1,0.9,"xplor.command('rama scale VALUE end')") )

xplor.command("@rna_orient1.setup")
pots.append(AvePot(XplorPot,"orie") )
rampedParams.append( StaticRamp("pots['ORIE'].setScale(0.2)"") )
rampedParams.append( MultRamp(0.002,0.3,"xplor.command('orie scale VALUE end')") )

#xplor.command("@plane16.tbl")

pots.add( XplorPot("BOND") )
pots.add( XplorPot("DIHE") )
pots.add( XplorPot("ANGL") )
pots.add( XplorPot("IMPR") )
rampedParams.append( MultRamp(0.4,1.0,"pots['ANGL'].setScale(VALUE)") )
rampedParams.append( MultRamp(0.1,1.0,"pots['IMPR'].setScale(VALUE)") )

protocol.initNBond(cutnb=4.5)
pots.add( XplorPot("VDW") )
rampedParams.append( MultRamp(0.9,0.78,
    "xplor.command('param nbonds repel VALUE end end')") )
rampedParams.append( MultRamp(.004,4,
    "xplor.command('param nbonds rcon VALUE end end')") )

pots.add(noen)
pots.add(noex)
pots.add(hbon)
pots.add(hbs)
pots.append(AvePot(XplorPot("plan",xplor.simulation)) )

mini = IVM()             #initial alignment of orientation tensor axes
from selectTools import IVM_groupRigidSidechain
from selectTools import IVM_breakRiboses

IVM_groupRigidSidechain(mini)
#IVM_breakRiboses(mini, sel=0, breakSelStr="name O4' or name C1'")

protocol.cartesianTopology(mini,"not resname ANI")

protocol.initMinimize(mini,
    numSteps=20)
mini.fix("not resname ANI")
mini.run()               #this initial minimization is not strictly necessary

#uncomment to allow Da, Rh to vary
#for medium in ('bic1','phg1'): media[medium].setFreedom("varyDa, varyRh")
#for medium in ('bic2',):
#    media[medium].setFreedom("varyDa, varyRh, fixAxisTo bic1")
#for medium in ('phg2','phg3',):
#    media[medium].setFreedom("varyDa, fixAxisTo phg1, fixRhTo phg1")

dyn = IVM()
protocol.initDynamics(dyn,potList=pots)
IVM_groupRigidSidechain(dyn)
#IVM_breakRiboses(dyn, sel=0, breakSelStr="name O4' or name C1'")
#protocol.cartesianTopology(dyn,"not resname ANI")
protocol.torsionTopology(dyn)

from selectTools import IVM_groupRigidSidechain
minc = IVM()
protocol.initMinimize(minc,potList=pots)
IVM_groupRigidSidechain(minc)
#IVM_breakRiboses(minc, sel=0, breakSelStr="name O4' or name C1'")
protocol.cartesianTopology(minc,"not resname ANI")

anneal= AnnealIVM(initTemp =init_t,
    finalTemp=25,
    tempStep =25,
    ivm=dyn,
    rampedParams = rampedParams)

# initialize parameters for initial minimization.
InitialParams( rampedParams )

# initial minimization
protocol.initMinimize(dyn,
    numSteps=1000)
dyn.run()

from simulationTools import testGradient
#testGradient(potList,eachTerm=1)

def calcOneStructure( structData ):
# initialize parameters for high temp dynamics.
InitialParams( rampedParams )

# high temperature bit - using only P-P nonbonded terms
protocol.initNBond(repel=1.2,
cutnb=100,
tolerance=45,
selStr="name P")

protocol.initDynamics(dyn,
    initVelocities=1,
bathTemp=init_t,
potList=pots,
    finalTime=10)

dyn.run()
protocol.initNBond() #reset to include all atoms

# perform simulated annealing
#
protocol.initDynamics(dyn,
    finalTime=0.2,  #time to integrate at a given
temp.
    numSteps=0,     # take as many steps as
necessary
    eTol_minimum=0.001 # cutoff for auto-TS det.
)

anneal.run()

#
# torsion angle minimization
#
protocol.initMinimize(dyn)
dyn.run()
#
# all atom minimization
#
minc.run()

structData.writeStructure(pots,crossTerms)

from simulationTools import StructureLoop, FinalParams
StructureLoop(numStructures=numberOfStructures,
    startStructure=startStructure,
    structLoopAction=calcOneStructure,
pdbTemplate=outFilename,
genViolationStats=1,
    averageTopFraction=0.3,
    averagePotList=pots,
    averageSortPots=[pots['BOND'],pots['ANGL'],pots['IMPR'],
        noex, noen,pots['CDIH'],hbon,hbs],
    #averageAccept=accept,  #only use structures which pass
accept()
    averageContext=FinalParams(rampedParams),

-285-
Appendix II - Dissemination

(i) Seminars presented and the courses and meetings attended

I have successfully presented the results of this project on several occasions through postgraduate seminars and poster presentations with high commendation. The most significant of these being the poster presented at the Royal Society of Chemistry (U.K.) NMR Discussion Group meeting held in the University of Cambridge in April 2008. Following the review of my poster abstract, I was awarded a competitive, full bursary by the RSC. Additionally an image of the antibiotic NMR structures was selected and used on the front cover of the meetings programme.

(ii) Publications

Two research papers covering the results of this project described in Chapters 3-5 will be published in peer reviewed international journals and they are listed below.

Misbah Nareen, John King and Vasudevan Ramesh, (2011) NMR structure determination of peptidyl transferase inhibitor antibiotics (to be submitted).

Misbah Nareen and Vasudevan Ramesh (2011) NMR studies of the structure and interaction of the conserved secondary structural motifs of B. subtilis and H. hal 23S rRNAs (to be submitted).