Lawrence Berkeley National Laboratory
Recent Work

Title
Self-Cleaving RNA: Characterization of a Family of Hammerhead Self-Cleaving Ribonucleic Acids by $^1$H NMR Spectroscopy

Permalink
https://escholarship.org/uc/item/2kj0c5tf

Author
Caviani, A.M.

Publication Date
1990-08-01
Self-Cleaving RNA: Characterization of a Family of Hammerhead Self-Cleaving Ribonucleic Acids by \( ^1H \) NMR Spectroscopy

A.M. Caviani  
(Ph.D. Thesis)  

August 1990
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
Self-Cleaving RNA:
Characterization of a Family
of Hammerhead Self-Cleaving Ribonucleic Acids by $^1$H NMR Spectroscopy

by

Ann Maria Caviani

PH.D. Thesis

August, 1990

Chemical Biodynamics Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94708

This work was supported by the Director, Office of Energy Research, Office of Biological and Environmental Research, Structural Biology Division of the U. S. Department of Energy Under Contract No. DE-AC03-76SF00098.
For my family
# Table of Contents

Chapter 1 Introduction to NMR

1.1 What is Spectroscopy? 2
1.2 Nuclear Spin 2
1.3 Larmor Precession 5
1.4 NMR is an Inherently Insensitive Spectroscopy 9
1.5 The Simple Fourier Transform NMR Experiment 10
1.6 Parameters in Biological NMR 13
   1.6.1 Chemical Shift 13
   1.6.2 Ring Currents 16
   1.6.3 Spin-Spin Coupling Constants 16
   1.6.4 Intensity of an NMR Resonance 18
   1.6.5 Relaxation 18
      1.6.5.1 Dipolar Relaxation 18
      1.6.5.2 Quadrupole Relaxation 19
      1.6.5.3 Paramagnetic Relaxation 20
1.7 The Nuclear Overhauser Effect 20
1.8 NMR Time Scales 25
1.9 Two Dimensional Spectroscopy 27

Chapter 2 Nucleic Acid Structure and NMR

2.1 Introduction 34
2.2 The Anatomy of Nucleic Acids 34
2.3 Nucleic Acids and Nuclear Spin 36
2.4 NMR Parameters and DNA/RNA Spectroscopy 41
2.5 Spectroscopy of Exchangeable Protons 43
2.6 Imino Protons are a Probe of Secondary Structure 44
2.7 Assignment of DNA and RNA Imino Proton Resonances 48
2.8 Assignment of the Nonexchangeable Protons of DNA 51
2.9 Spectroscopy of RNA 56

Chapter 3 Experimental Methods 60

3.1 Introduction 61
3.2 Isolation of RV/RI/plBl20 Vector DNA 61
3.3 Construction of Insert DNA 68
3.4 Construction and Propagation of Plasmid DNA 71
   Ligation of Vector and Insert DNA 71
3.5 Large Scale Purification of Plasmid DNA 71
3.6 Linearization of Plasmid DNA with EcoRI 73
   Analytical EcoRI Digests 73
   Large Scale EcoRI Digests 75
3.7 Purification of T7 RNA Polymerase 75
3.8 Large Scale Transcription Reactions 78
   Analytical Transcription Reactions 78
   Large Scale Transcription Reactions, RNA1, 3/4-RNA2, RNA3, and RNA4 78
   Large Scale Transcription Reactions, RNA6 79
   Large Scale Purification, RNA1, RNA3, RNA4 79
   Large Scale Purification, 3/4-RNA2, RNA6 81
3.9 Demonstration of Cleavage at the Appropriate Site 81
   Purification of $^{32}$P-Body-Labelled RNA 82
   Demonstration of Cleavage 83
Conclusion and Discussion 154

Chapter 6 NMR Investigation of RNA3 157

6.1 Introduction 158

6.2 Spectroscopy of RNA3 158

Results, NOE Difference Spectroscopy 161

Temperature Data 169

Conclusion 172

Chapter 7 NMR Investigation of RNA4 174

7.1 Introduction 175

7.2 Spectroscopy of RNA4 176

Results, NOE Difference Spectroscopy, No MgCl₂ 180

Results, Temperature Studies, No MgCl₂ 191

Results, Temperature Studies, 10 mM MgCl₂ 193

Chapter 8 NMR Investigation of RNA6 201

8.1 Introduction 202

8.2 Determination of the Cleavage Half-Life 203

Analytical Transcription Reactions 203

Results 205

8.3 Spectroscopy of RNA6 205

Sample Preparation 205

Results, NOESY Spectroscopy, No MgCl₂ 208

Results, Melting Data, No MgCl₂ 219

Results, NOE Difference Spectroscopy, 10 mM MgCl₂ 223

Conclusions 226

8.4 Determination of NMR Cleavage Conditions 228
Preparation of the RNA
Buffer and pH Dependence of RNA6 Cleavage
Results
MgCl₂ Concentration Dependence of Cleavage in Phosphate Buffer
Results
Spermidine Concentration Dependence of RNA6 Cleavage
Results
Conclusions and Discussion
8.5 NMR Spectroscopy of Cleaving RNA6
8.6 Spectroscopy of Cleaved RNA6
8.7 Final Conclusions and Discussion
Appendix A Biochemical Protocols
Purification of Plasmid DNA
Linearization of Plasmid DNA
Purification of T7 RNA Polymerase
Transcription Buffers and Conditions
Large Scale RNA Transcription
Large Scale Purification of RNA
Appendix B GN-500 Pulse Sequences
1331 Experiment
1331NO Experiment
11NOE1 Experiment
13NSY Experiment
References
List of Figures

Figure 1.1 Quantized Nuclear Spin 6
Figure 1.2 Behavior of $\mu$ in a Magnetic Field 8
Figure 1.3 The Simple NMR Experiment 12
Figure 1.4 Chemical Shift 15
Figure 1.5 Ring Current 17
Figure 1.6 The Nuclear Overhauser Effect 22
Figure 1.7 Spin Diffusion 24
Figure 1.8 Schemes for FT Pulse Experiments 28
Figure 1.9 Contour Plot Representation of a NOESY Experiment 31
Figure 1.10 Schemes for 2D Pulse Experiments 32
Figure 2.1 The Basic Components of DNA and RNA 35
Figure 2.2 Diagrammatic Representation of an RNA Single Strand 37
Figure 2.3 Watson-Crick Base Pairs 39
Figure 2.4 Space Filling Representation of A and B Double Helices 40
Figure 2.5 Proton Chemical Shift Ranges of DNA and RNA 42
Figure 2.6 Some non-Watson-Crick Base Pairs of yeast tRNA$^{Phe}$ 46
Figure 2.7 The 1331 Experiment 47
Figure 2.8 Assignment Strategy for a 1D NOE Experiment 50
Figure 2.9 COSY Spectrum of a Duplex DNA Oligomer 53
Figure 2.10 NOESY Spectrum of a Duplex DNA Oligomer 54
Figure 3.1 plB120, Insert DNA, and RNA Sequences 62
Figure 3.2 Cleavage of RNA1, RNA3, RNA4, and RNA6 84

Figure 4.1 Sequences of minSTobRV RNA and RNA1 92
Figure 4.2 GN500 Scheme for Processing 1D NOE Difference Spectra
Figure 4.3 Sequence and 500 MHz Imino Protons Spectrum of RNA1
Figure 4.4 25°C RNA1 Stem II NOE Difference Spectra
Figure 4.5 30°C RNA1 Stem II NOE Difference Spectra
Figure 4.6 25°C RNA1 Stem III NOE Difference Spectra
Figure 4.7 30°C RNA1 Stem III NOE Difference Spectra
Figure 4.8 RNA1 Difference Spectra of Resonance F at 25°C and 30°C
Figure 4.9 The Rest of the 25°C RNA1 NOE Difference Spectra
Figure 4.10 The Rest of the 30°C RNA1 NOE Difference Spectra
Figure 4.11 Imino-Aromatic Region of RNA1 2D NOESY Spectrum
Figure 4.12 Imino-Imino Region of RNA1 2D NOESY Spectrum
Figure 4.13 Temperature Dependence of RNA1, No MgCl₂
Figure 4.14 MHz Imino Proton Spectra of RNA1 in 0 mM MgCl₂
and 10 mM MgCl₂ at 30°C
Figure 4.15 RNA1 NOE Difference Spectra, 10 mM MgCl₂
Figure 4.16 RNA1 NOE Difference Spectra, 10 mM MgCl₂
Figure 4.17 Temperature Dependence of RNA1, 10 mM MgCl₂
Figure 4.18 Temperature Dependence of RNA1, 20 mM MgCl₂
Figure 4.19 Imino Proton Assignments of RNA1
Figure 5.1 Sequence of RNA2
Figure 5.2 18 Hour Cleavage Reactions of rC- and dC-RNA2
Figure 5.3 1 Hour Cleavage Reactions of rC-RNA2
Figure 5.4 500 MHz Imino Proton Spectrum of 3/4-RNA2
Figure 5.5 500 MHz Spectra of 3/4-RNA2 and dC-RNA2
Figure 5.6 500 MHz Imino Proton Spectrum of ca. 0.5 mM dC-RNA2
Figure 5.7  Native and Denaturing Polyacrylamide Gels of dC-RNA2
Figure 5.8  500 MHz Spectra of 0.4 mM and 0.2 mM dC-RNA2
Figure 5.9  Temperature Dependence of dC-RNA2, 20 mM MgCl₂
Figure 5.10  Half-Life Determination of rC-RNA2
Figure 6.1  Sequence of RNA3
Figure 6.2  500 MHz Imino Proton Spectra of RNA1 and RNA3 at 20°C
Figure 6.3  RNA3 Stem II NOE Difference Spectra
Figure 6.4  RNA3 Stem III NOE Difference Spectra
Figure 6.5  RNA3 NOE Difference Spectra Connecting Resonances F and L
Figure 6.6  RNA3 Stem I Difference Spectra
Figure 6.7  Temperature Dependence of RNA3
Figure 6.8  Imino Proton Assignments of RNA3
Figure 7.1  GU Wobble Base Pair
Figure 7.2  Sequences and 500 MHz Imino Proton Spectra of RNA3 and RNA4
Figure 7.3  RNA4 Stem II Difference Spectra
Figure 7.4  RNA4 Stem I Difference Spectra
Figure 7.5  Saturation of Resonance C at 15° and 30°C
Figure 7.6  RNA4 Stem III Difference Spectra
Figure 7.7  RNA4 Stem III Difference Spectra
Figure 7.8  Saturation of Resonances F and L at 15°C
Figure 7.9  Imino Proton Assignments of RNA4
Figure 7.10  Temperature Dependence of RNA4, No MgCl₂
Figure 7.11  Temperature Dependence of RNA4, 10 mM MgCl₂
Figure 7.12  500 MHz Imino Proton Spectra of RNA1, RNA3, and RNA4
Figure 7.13  500 MHz Imino Proton Spectra of RNA1 and RNA4
in 10 mM MgCl₂ and 30°C

Figure 8.1 Sequence of RNA6

Figure 8.2 Half-Life Determination of RNA6

Figure 8.3 Sequences and 500 MHz imino Proton Spectra
of RNA4 and RNA6 at 20°C

Figure 8.4 NOESY of RNA6 in H₂O at 20°C: Stem III Connectivities
Figure 8.5 NOESY of RNA6 in H₂O at 20°C: Stem I Connectivities
Figure 8.6 NOESY of RNA6 in H₂O at 20°C: Imino-Aromatic Region
Figure 8.7 NOESY of RNA6 in H₂O at 15°C: Stem I Connectivities
Figure 8.8 NOESY of RNA6 in H₂O at 15°C: Stem II Connectivities
Figure 8.9 NOESY of RNA6 in H₂O at 15°C: Stem III Connectivities
Figure 8.10 NOESY of RNA6 in H₂O at 20°C: Imino-Aromatic Region
Figure 8.11 Temperature Dependence of RNA6, No MgCl₂
Figure 8.12 500 MHz Imino Protons Spectra of RNA4 and RNA6
in 0 mM and 10 mM MgCl₂
Figure 8.13 Temperature Dependence of RNA6, 10 mM MgCl₂
Figure 8.14 Assignment of RNA6
Figure 8.15 Analytical Cleavage Reactions of RNA6
Figure 8.16 MgCl₂ Concentration Dependence of RNA6 Cleavage
in Phosphate Buffer
Figure 8.17 Spermidine Concentration Dependence of RNA6 Cleavage
Figure 8.18 500 MHz Imino Protons Spectrum of Cleaving RNA6
Figure 8.19 500 MHz Imino Proton Spectra of RNA4 and CL-RNA6
Figure 8.20 500 MHz Imino Proton Spectra of UC-RNA6 and CL-RNA6
Figure 8.21 Temperature Dependence of CL-RNA6
List of Tables

Table 1.1  NMR Properties of Selected Nuclei  4
Table 2.1  Parameters of A and B Double Helices  38
Table 2.2  Short Sequential Distances A and B DNA Helices  55
Table 3.1  Analytical EcoRI Digests  74
Table 4.1  FTNMR 1D Processing Macro  94
Table 8.1  Buffers for Analytical Cleavage Reactions of RNA6  230
Table 8.2  Analytical RNA6 Cleavage Reaction Conditions: pH and Buffer  231
Table 8.3  Analytical RNA6 Cleavage Reaction Conditions:
   MgCl₂ Dependence  235
Table 8.4  Analytical RNA6 Cleavage Reaction Conditions:
   Spermidine Dependence  238
Chapter 1

Introduction to NMR
1.1 What is Spectroscopy?

Spectroscopy in general can be described as the interaction between matter and electromagnetic radiation such that energy is absorbed or emitted according to the Bohr frequency relation $\Delta E = h\nu$, where $\Delta E$ is the energy difference between the final and initial states, $h$ is Planck's constant, and $\nu$ is the frequency of the electromagnetic radiation [Harris (1983)]. In nuclear magnetic resonance spectroscopy (NMR) a sample is placed in a static magnetic field, $H_0$, which serves to split the nuclear spin states. Transitions are then induced between the spin states by application of a radiofrequency (rf) field, $H_1$, which give rise to the NMR absorption signal.

The energy difference between the spin states is directly proportional to the strength of the applied magnetic field and the *gyromagnetic ratio*, $\gamma$:

$$\Delta E = \gamma H_0$$  \hspace{1cm} (1.1)

The gyromagnetic ratio is a nucleus-dependent constant which represents the sensitivity of the energy levels to the magnetic field. The greater the value of $\gamma$ for a particular nucleus, the more sensitive is its NMR signal at a constant $H_0$ field. Fortunately, the gyromagnetic ratio of the hydrogen nucleus is large making it a very sensitive NMR nucleus in addition to being very abundant in biological macromolecules. In fact, only the radioactive isotope $^3\text{H}$ has a larger $\gamma$-value than does $^1\text{H}$. The high sensitivity of $^1\text{H}$ relative to other nuclei has certainly contributed to its popularity as an NMR nucleus. This work is no exception, as all of the NMR spectra in this thesis are of the nucleus $^1\text{H}$.

1.2 Nuclear Spin

The high sensitivity of the hydrogen nucleus is not the only factor which makes
it particularly attractive to NMR spectroscopists. It is the simplest nucleus, comprised of only a single proton (and for this reason the hydrogen nucleus will subsequently be referred to as the "proton" as in "proton NMR"). Since a proton is a spin 1/2 particle, the hydrogen nucleus is a spin 1/2 nucleus. Nuclei with a total nuclear spin of 1/2 are "well-behaved" from an NMR standpoint for two reasons: (i) when placed in a magnetic field spin 1/2 nuclei split into only two energy levels, with an energy of separation described by eqn. 1.1, and (ii) spin 1/2 nuclei are spherically symmetric, and thus do not possess an electric quadrupole moment (which leads to undesirable spectroscopic features).

All other nuclei contain neutrons, also spin 1/2 particles, in addition to protons. The total nuclear spin of a nucleus which is comprised of p protons and n neutrons is then a vector combination of p+n spins, each of magnitude 1/2. Unfortunately the rules which govern the vector of addition of nuclear spin are complex, making it difficult to theoretically predict the total spin of a nucleus based on its atomic mass and number. The observed spins of some nuclei have been rationalized in this manner, however, and there exists a set of empirical rules for nuclear spin prediction. Table 1.1 lists the nuclear spin, γ-value, and natural abundance of the biologically important nuclei. It is important to note that not all isotopes of elements possess a nuclear spin. These nuclei are then unaffected by the presence of a magnetic field, and do not give rise to NMR spectra. \(^{12}\text{C}\) is an example of a nucleus which does not manifest an NMR spectrum.

Nuclear spin is really an angular momentum, and as such all the rules which apply to other types of angular momenta apply to it. For spectroscopy, the most important result is that the energy levels of nuclear spin are quantized. The total nuclear spin vector, \(\mathbf{S}\), can then be described in terms of a nuclear spin
<table>
<thead>
<tr>
<th>Isotope</th>
<th>Spin</th>
<th>Natural Abundance (%)</th>
<th>Magnetogyric Ratio $\mu/10^7 \text{ rad T}^{-1} \text{ s}^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1\text{H}$</td>
<td>$1/2$</td>
<td>99.985</td>
<td>26.7519</td>
</tr>
<tr>
<td>$^3\text{H}$</td>
<td>$1/2$</td>
<td>- - -</td>
<td>28.535</td>
</tr>
<tr>
<td>$^{13}\text{C}$</td>
<td>$1/2$</td>
<td>1.108</td>
<td>6.7283</td>
</tr>
<tr>
<td>$^{14}\text{N}$</td>
<td>1</td>
<td>99.63</td>
<td>1.9338</td>
</tr>
<tr>
<td>$^{15}\text{N}$</td>
<td>$1/2$</td>
<td>0.37</td>
<td>-2.712</td>
</tr>
<tr>
<td>$^{19}\text{F}$</td>
<td>$1/2$</td>
<td>100</td>
<td>25.181</td>
</tr>
<tr>
<td>$^{31}\text{P}$</td>
<td>$1/2$</td>
<td>100</td>
<td>10.841</td>
</tr>
</tbody>
</table>

quantum number, \( I \):

\[
\mathbf{S} = \frac{\hbar}{2\pi} \sqrt{I(I+1)}
\]  

where the convention of underlining a total vector is used. Because nuclear spin is a vector property it must be defined by a direction as well as a magnitude. Therefore, another quantum number, \( m_I \), can be specified such that the component of the angular momentum along a reference direction, say \( z \), is defined by:

\[
S_z = \frac{\hbar}{2\pi} m_I
\]  

Quantum mechanics shows that the allowed values of \( m_I \) are given by eqn. 1.4. There are thus \( 2I+1 \) allowed values of \( m_I \):

\[
m_I = I, \ (I-1), \ (I-2), \ ... , \ -I
\]  

For a proton there are only two allowed values of \( m_I \); \( m_I = +1/2 \) and \( m_I = -1/2 \). Figure 1.1 illustrates quantized nuclear spin.

1.3 Larmor Precession

A spinning charged particle is analogous to an electric current flowing through a loop, and therefore it has associated with it a magnetic dipole, \( \mu \). For a nucleus with spin \( I \) the size of the dipole (strength of the magnet) can be shown to be:

\[
\mu = -\gamma \frac{\hbar}{2\pi} \sqrt{I(I+1)}
\]  

When this spin is placed in a magnetic field, quantal rules state that \( \mu \) can be oriented only so that the magnitude of its component along the direction of the applied field is integral or half-integral. This can be shown mathematically by
Figure 1.1 Quantized Nuclear Spin. (a) The general case explicitly showing the use of quantum numbers $I$ and $m_I$. (b) The possible orientations for nuclear spin of $I=1/2$. Note that $I$ and $m_I$ do not specify direction in the xy-plane.
relating eqns. 1.2, 1.3, and 1.5:

\[ \mu_z = -\gamma \frac{h}{2\pi} m_i \]

where the field is applied along the z-direction. According to eqn. 1.4, the magnitude of \( \mu_z \) for a proton must then be half-integral. But if \( I \) is half integral, which it is for a proton, the magnitude of \( \mu \) (eqn. 1.5) \textit{cannot} be half-integral. The result is that \( \mu \) can never lie exactly in the direction of the applied field. It will instead lie at some angle \( \theta \) with respect to the field (fig. 1.2a). Because \( \mu \) lies across the field it experiences a torque tending to turn it into the direction of the field. The vector \( \mu \) responds to the torque by precessing about the applied field (fig. 2b). The frequency at which \( \mu \) precesses is called the Larmor frequency, and is given by (in units of radians):

\[ \omega_0 = \gamma H_0 \]

where \( \omega_0 \)(rad)\(=2\pi v_0 \)(Hz), and \( v_0 \) is the Larmor frequency in Hertz. Comparison of eqns. 1.7 and 1.1 shows that this frequency is exactly the energy separation between nuclear spin states upon placing the spin in the static field. This then describes a mode by which the nuclear spin states can exchange energy with their surroundings. Electromagnetic radiation applied at the Larmor frequency will interact coherently with the spin states and cause transitions between the energy levels. Radiation of a different frequency will not interact. This is a resonant phenomenon, and hence defines the "R" in NMR.
Figure 1.2 Behavior of \( \mu \) in a magnetic field. (a) \( \mu \) lies at an angle \( \theta \) with respect to the magnetic field direction. (b) The precessional cone traced by \( \mu \) about the applied field. For a spin 1/2 nucleus, \( \theta \) is 54° 44'.

\[ \mu_z = -\gamma \frac{h}{2\pi} m, \quad \mu = -\frac{\hbar}{2\pi} [1/(1-1)]^z \]

\[ \omega_c = \gamma H_c \]
1.4 NMR is an Inherently Insensitive Spectroscopy

One of the difficulties of applying NMR to biological molecules is the relatively high sample concentrations that are required for NMR spectroscopy. This is in contrast to other forms of spectroscopy, as it is typically a very sensitive technique. Having introduced the role $\gamma$ plays in the sensitivity of the NMR signal, it is important to stress that $\gamma$s reflect the relative sensitivity of different nuclei in the same magnetic field. The small energy separation between spin states makes NMR an inherently insensitive molecular spectroscopy for any nucleus at biological temperatures.

The NMR signal arises from transitions induced between nuclear spin energy levels. The probability of these transitions is directly proportional to the populations of the spin states. Therefore it is instructive to develop an expression for the population of the energy levels to understand the insensitivity of NMR spectroscopy. Experience with bar magnets tells us that when a magnetic moment is placed in a magnetic field it will tend to align itself with the field. This is equivalent to saying that the moment will assume the lowest energy orientation. When nuclei which possess a nuclear spin are placed in a magnetic field the magnetic moments also tend to align themselves with the field, but because of thermal motions, they all do not occupy the lowest energy level. Classical theory states that at a temperature $T$ (°K) the ratio of the spin state populations for a spin $I=1/2$ particle is given by the Boltzmann distribution:

$$
\frac{N_{\text{upper}}}{N_{\text{lower}}} = \exp \left( - \frac{\Delta E}{kT} \right)
$$

where $k$ is the Boltzmann constant. At any temperature above absolute zero the upper energy level is then populated. However, if the separation between energy
levels is sufficiently large, $N_{\text{upper}}$ may be negligible. Unfortunately, the energy difference between nuclear spin states is extremely small, even at the high field strengths available today. The net result is that at a field strength of 11.74 Tesla (this corresponds to a resonant proton frequency of 500 MHz) and a temperature of $300^\circ$K, the ratio of energy level populations is near unity. However, $N_{\text{lower}}$ is slightly larger than $N_{\text{upper}}$, so when electromagnetic radiation is applied to induce transitions between the levels, upward transitions slightly predominate over downward transitions, resulting in a net absorption of energy from the electromagnetic radiation. Because the net absorption for each individual spin is small, a large number of spins are required to give an appreciable signal. For biological spectroscopists this translates into sample concentrations of millimolar, and because the spectrometer geometry requires a sample volume of about 0.5 mL, the sample may require tens of milligrams of material.

1.5 The Simple Fourier Transform NMR Experiment

Finally, after a seemingly (too) lengthy introduction of nuclear spins, magnetic moments, and energy level populations, the basic NMR experiment is introduced. This will not only develop the nature of the interactions which give rise to the transitions between nuclear spin energy levels, but will also set the stage for a brief description and terminology of the parameters that are most useful for structural and dynamical studies of biological macromolecules by NMR.

Consider an NMR sample which contains many identical molecules, each possessing only one nucleus with a nuclear spin of $I=1/2$. When the sample is placed in the static field, $H_0$, the magnetic moments precess at an angle $\theta$ about the
field as previously described. Because the sample contains many identical spins, the \( \mu \)'s will be evenly distributed about a "precessional cone", a situation illustrated in fig. 1.3a. The vector sum of the precessing moments results in a net magnetization vector for the sample, \( \mathbf{M} \). Because the lower spin state is slightly more populated than the upper state, \( \mathbf{M} \) is aligned along the direction of \( H_0 \) (fig. 1.3b). In the simplest experiment a rf field, \( H_1 \), is applied along the x-axis. \( \mathbf{M} \) experiences a torque from \( H_1 \) and responds by precessing about \( H_1 \) with frequency \( \gamma H_1 \) (eqn. 1.7). If the rf field is applied for a suitable duration, \( \mathbf{M} \) is rotated by \( 90^\circ \) into the xy-plane (fig. 1.3c). In the most basic NMR spectrometer, the geometry of the coils is such that the rf coil is positioned along the x-axis and the receiver coil is positioned along the y-axis. Therefore the receiver coil detects fluctuations in the y-axis without being coupled to the rf coil because they are perpendicular to one another.

When \( \mathbf{M} \) is rotated into the xy-plane by application of the \( 90^\circ \) pulse, three things immediately begin to happen; (i) \( \mathbf{M} \) experiences a torque from the main field \( H_0 \), to which \( \mathbf{M} \) responds by precessing about \( H_0 \) in the xy-plane (fig. 1.3d), (ii) the precession of \( \mathbf{M} \) in the xy-plane results in a fluctuation along the y-axis, sending a signal to the receiver coil, and (iii) \( \mathbf{M} \) gradually relaxes back to its equilibrium position in the direction of \( H_0 \). The overall motion of \( \mathbf{M} \) after the application of the \( 90^\circ \) pulse can be visualized as a spiral (fig. 1.3e). The detector is then receiving a signal which is exponentially decaying with time, called the free induction decay (FID). Fourier transformation of the time-domain FID results in the frequency-domain NMR spectrum.

The relaxation of \( \mathbf{M} \) can be dissected into two components; that which affects decay along the z-axis (longitudinal), and that which affects decay in the xy-plane (transverse). The longitudinal relaxation time, \( T_1 \), involves interactions between
Figure 1.3 The Simple NMR experiment. (a) $\mu$'s of a group of spinning nuclei are evenly distributed about a precessional cone. The large arrows represent the magnitude of the vector sum of the $\mu$'s. (b) Resultant net magnetization vector M for the sample. (c) Application of $H_1$ for suitable duration along the -$x$ axis rotates M through 90° about the $x$ axis. (d) Immediately after turning off $H_1$ M begins to precesses in the xy-plane about $H_0$. (e) Relaxation of M back to its equilibrium position along the direction of $H_0$.

the spin system and its surroundings, called the lattice. For this reason it is also referred to as the spin-lattice relaxation time. The transverse relaxation time, $T_2$, is associated with interactions between spins, and is sometimes referred to as the spin-spin relaxation time. The mechanisms by which a spin system is relaxed are rather weak, so the spin state reverts slowly to a lower state. According to Heisenberg this translates into a small uncertainty of the excited state, and hence narrow lines. Therefore factors which contribute to increased relaxation rates will generally increase linewidths. More will be said about some modes of relaxation in the following section.

1.6 Parameters in Biological NMR

In order to correctly interpret NMR spectra, it is necessary to understand the types of interactions which manifest particular spectral features. This section has been reserved for the purpose of surveying the spectral features which are important for extracting structural and dynamical information.

1.6.1 Chemical Shift

The previous discussions of how a nuclear spin behaves in a magnetic field have concentrated on isolated spins. Real life samples, however, are not collections of bare nuclei; the nuclei have associated with them an electron cloud. When the sample is placed in a magnetic field the electron clouds of the nuclei circulate in a direction such that they produce a field opposing the applied field. The nucleus is thus shielded from the static field by an amount which is dependent on the local electronic environment of the nucleus. The actual field felt by each nucleus is called the effective field and can be expressed by:
The field induced by electron circulation is directly proportional to the applied magnetic field. Mathematically this can be described as:

\[ H_{\text{induced}} = \sigma H_{\text{applied}} \]  

(1.10)

where \( \sigma \) is a shielding constant for a particular nucleus. Because the extent of shielding will be constant for a given atom in isolation, but will vary with the electron density about a nucleus in a molecule, eqns. 1.9 and 1.10 can be generalized to:

\[ H_j = H_0 \left( 1 - \sigma_j \right) \]  

(1.11)

where \( H_j \) is the field experienced by a particular nucleus \( j \) whose shielding constant is \( \sigma_j \). Therefore different protons in the molecule will resonant at slightly different frequencies according to:

\[ v_j = \left| \frac{\gamma}{2\pi} \right| H_0 \left( 1 - \sigma_j \right) \]  

(1.12)

where \( v_j \) represents the resonant frequency of nucleus \( j \). This phenomenon is called the chemical shift, \( \delta \), and is illustrated in figure 1.4. The chemical shift is measured relative to a standard, and because it is dependent on the strength of the applied field it is reported in parts per million (ppm), a field-independent unit;

\[ \delta = 10^6 \left( \frac{v_{\text{sample}} - v_{\text{standard}}}{v_{\text{standard}}} \right) \]  

(1.13)

Because certain types of protons have characteristic chemical shifts (even in different molecules), it is a very powerful tool in determining the origin of a proton which manifests a particular resonance line.
Figure 1.4 Energy level diagrams for two isolated nuclei of spin $I=1/2$, A and X, demonstrating constant frequency and constant field conditions for resonance. If transition 1 is stimulated for nucleus X, then either the field must be increased at constant frequency (transition 2), or the frequency must be decreased at constant field (transition 3) to achieve resonance for A. Transitions 1 and 3 represent the chemical shift difference between spins A and X at constant field due to different local electronic environments of A and X.
1.6.2 Ring Currents

Basically a dramatic chemical shift effect, ring currents occur when a proton is in the proximity of a conjugated system such as a benzene ring. The electrons of the ring circulate in such a direction as to create the field illustrated in fig. 1.5. Protons which are situated above or below the plane of the ring feel an effective field which opposes the applied field, and are shifted to a higher field (to the right in the spectrum) due to the influence of the nearby ring. Protons which are located on the periphery of the ring experience an increased effective field, and are shifted down-field in the spectrum (to the left). In nucleic acids ring currents play a major role in the chemical shift dispersion that is observed in the NMR spectrum.

1.6.3 Spin-Spin Coupling Constants

If a molecule possesses only two nuclei with non-zero nuclear spin, \( I_1 \) and \( I_2 \), which are three or fewer bonds apart, it is found that the resonance of spin \( I_1 \) is split into \( 2|I_2|+1 \) lines of equal intensity, and that of \( I_2 \) is similarly split into \( 2|I_1|+1 \) lines. This splitting is a through bond effect, and is due to the different \( m_I \) values of the other spin. Because the energy levels of the spins are modified by through-bond interactions with other spins, spin-spin (or scalar) coupling constants, \( J \), are independent of the applied field and are reported in Hertz. Scalar coupling constants, like chemical shifts, are sensitive to changes in the local environment, and are therefore a useful probe of molecular structure. Scalar coupling constants are particularly important in the determination of bond angles.
Figure 1.5 Induced circulation of the delocalized π-electrons for a benzene ring oriented perpendicular to the applied magnetic field. The field induced by electron circulation is labelled $H_i$. Nuclei on the periphery of the molecular plane of the ring are high frequency shifted, those above and below the plane of the ring are low frequency shifted.
1.6.4 Intensity of an NMR Resonance

It is now clear that an NMR spectrum will consist of numerous lines at various chemical shifts, reflecting the local electronic environment of the individual protons. In a large molecule, or a molecule with high symmetry, the situation arises in which different protons resonate at the same chemical shift. This situation is called chemical shift degeneracy, and each resonance line in the spectrum will have an integrated intensity representing the number of protons which manifest that particular resonance frequency.

1.6.5 Relaxation

The longitudinal and transverse relaxation times, $T_1$ and $T_2$, were previously introduced. This section will describe some of the modes of relaxation which affect these relaxation times.

1.6.5.1 Dipolar Relaxation

A molecule containing many nuclei which possess a nuclear spin can be thought of as a collection of tiny magnets, due to the magnetic dipole moment supplied by each of the "spinning" nuclei. Because the molecules tumble in solution the net magnetic moment each spin induces at the site of another averages to zero. Therefore the dipolar interaction observed with solids is not observed in solution NMR. However each spin feels the presence of other spins in its vicinity at any particular time. Because of Brownian motion, the value of the field felt by a spin fluctuates, effectively giving rise to an oscillating magnetic field. Those components of this tiny oscillating field which are at the Larmor (resonant) frequency of nearby spins can induce transitions of those spins. This results in a mechanism whereby
the neighboring protons can dissipate energy, and thus relax. The effectiveness of
the oscillating field in bringing about relaxation depends on the magnitude of the
magnetic moment and on the distance between the interacting nuclei. Because of the
nature of the interaction, this type relaxation is termed dipolar relaxation. Dipolar
relaxation contributes to both $T_1$ and $T_2$ relaxation [Abragam (1961)].

1.6.5.2 Quadrupole Relaxation

It was stated earlier that nuclei which have a spin of $I \geq 1$ possess an electric
quadrupole moment in addition to a magnetic moment. The electric quadrupole
interacts strongly with an applied electric field gradient. The electric quadrupoles,
and hence also the nuclear moments will tend to lie in the direction of the field
gradient. If the gradient changes direction, the nuclear moments will try to follow
this change. The coupling between the quadrupole moment and the field gradient can
be so strong that it is sufficient to prevent the coherent alignment of the nuclear
magnetic moments in an applied magnetic field. The fluctuating electric field
gradient supplies a mechanism by which the magnetic spins can be relaxed. This
relaxation mechanism is very efficient, and when present is usually the dominant
mode of relaxation [Becker (1980)].

Quadrupole relaxation has two effects of NMR spectra. First, a nucleus which
possess a quadrupole manifests a very broad resonance line. $^{14}$N is an example of
such a nucleus. Second, the quadrupole nucleus may also broaden the resonances of
other nuclei to which it is coupled. At times these proton resonances are
sufficiently broad that they are difficult to resolve from the spectrum baseline.
1.6.5.3 Paramagnetic Relaxation

An atom is called a paramagnet if it has unpaired electrons. A paramagnet, then, also acts like a tiny magnet, but each electron's moment is 660 times that of a proton's. If the NMR sample contains a paramagnetic metal (such as Mn$^{2+}$), the electrons will relax the nuclear spins. The net result is that all of the NMR resonances will be severely broadened. This broadening, however, is not always an undesirable spectroscopic feature. If the molecule of interest contains a metal binding site, replacement of the non magnetic metal with a paramagnetic one can help identify those protons which are near the coordinated metal, as the nearby resonances will be broadened. These types of studies were done with tRNA to determine the binding site of magnesium by replacing it with manganese [for a review see Reid (1981)]. It should be noted that these types of studies depend heavily on the affinity of the binding site for the metal. If the binding is weak and the paramagnet is essentially free in solution, it will broaden the entire spectrum, not just those resonances near the binding site.

1.7 The Nuclear Overhauser Effect

The Nuclear Overhauser Effect (NOE) refers to changes in intensities of NMR signals on double resonance that result from dipolar relaxation [Becker (1980)]. Consider a two spin system, AX, where each nucleus has spin $I=1/2$, and where the only mechanism for relaxation is via dipolar interactions. The four energy levels of this system in a static magnetic field are shown in fig. 1.6a. Due to the selection rule of $\Delta m=\pm 1$ for NMR transitions, absorption of energy from an rf field can only induce the transitions depicted in fig. 1.6a. However, dipolar relaxation processes can cause transitions between all the energy levels, as shown in fig. 1.6b. In the
absence of rf radiation, relaxation proceeds along these various pathways until the Boltzmann distribution (eqn. 1.8) is reached. If the X spin is saturated by irradiation with an rf field, H2, at frequency \( v_X \), the population difference between connected energy levels becomes zero, making \( N_1 = N_3 \) and \( N_2 = N_4 \). Under these conditions relaxation still operates, but a non-Boltzmann distribution is reached; the populations \( N_2 \) and \( N_4 \) increase. The net result is that the intensity of the A transition changes from a relative value of unity to a value of:

\[
1 + \frac{\gamma_X}{2 \gamma_A}
\]  

(1.14)

The NOE, given the symbol \( \eta \), is defined as the difference between the saturated and non-saturated intensities, and is usually quoted in percent of the unperturbed resonance intensity. Thus the NOE is defined by:

\[
\eta = \frac{\gamma_X}{2 \gamma_A}
\]  

(1.15)

For homonuclear spectroscopy (the condition where \( \gamma_A = \gamma_X \)), the maximum NOE is 50%. Because the NOE is a dipolar interaction, it is dependent on the distance between the interacting nuclei. This dependence can be expressed as:

\[
\eta = \frac{f(\tau_c)}{6 r_{ij}}
\]  

(1.16)

where \( f(\tau_c) \) is a function of the rotational correlation time, \( \tau_c \), which accounts for motional averaging processes, and \( r_{ij} \) is the distance between the nuclei [Wüthrich (1986)]. Double resonance NOE spectra are usually presented as the difference between an unperturbed spectrum and a spectrum recorded with a pre-irradiating pulse saturating the resonance line of interest. The difference spectrum then
Figure 1.6 Energy Level Diagram for an AX Spin System. (a) Allowed transitions induced by absorption of radiation at the frequencies of A and X. (b) Allowed transitions under the influence of dipolar relaxation.
displays the fractional intensity changes, with the resultant peaks being called NOEs. For protons the NOE falls to a value of 5% at about 3.5Å.

This discussion of the NOE is based on three assumptions: (i) the product of the correlation time and Larmor frequency is much less than unity \((\omega_0 \tau_c << 1)\), called the extreme narrowing limit, (ii) a two spin system, and (iii) solely dipolar relaxation. When condition (i) is not satisfied, \(\eta\) depends on \(\tau_c\) and \(\omega_0\), a condition which will be discussed in the next section. When condition (ii) is not met and there exist three or more interacting nuclei, the observed NOE intensities may be modified by two or more subsequent cross relaxation steps [Hull and Sykes (1975); Kalk and Berendsen (1976); Solomon (1955)]. This is known as spin diffusion and is illustrated in fig. 1.7. If condition (iii) is not satisfied and other modes of relaxation in addition to dipolar relaxation the NOE will show a marked decrease in intensity, and can even be completely quenched. Because the NOE is sensitive to distance, protons which are far apart in primary sequence but are close in three-dimensional space will manifest an NOE. This makes the NOE a powerful tool for the elucidation of the three-dimensional structure of biological molecules.

Clearly the study of biological macromolecules does not fall within the extreme narrowing limit. For a proton resonant frequency of 500 MHz, \(\omega_0 \tau_c = 1\) at a molecular weight of about 2000 Daltons \((\tau_c = 1\text{nsec/kD})\). In fact, biological molecules are generally in the limit where \(\omega_0 \tau_c >> 1\). How then, does this affect the observation of an NOE? As stated above, when not in the extreme narrowing limit the NOE depends on \(\omega_0\) and \(\tau_c\). The take-home message here is that large molecules result in positive NOEs whereas small molecules manifest negative NOEs. In the limit where \(\omega_0 \tau_c = 1\) the intensity of the NOE goes through zero, and may be non-observable even for spins very close in space.
Figure 1.7 Direct cross-relaxation between spins 1 and 2 (solid arrow), where \( r_{1,2} \leq 5\text{Å} \). The dashed arrow indicates a spin diffusion pathway between spins 2 and 3 where \( r_{2,3} \leq 5\text{Å} \) and \( r_{1,3} > 5\text{Å} \). An apparent NOE between spins 1 and 3 is thus observed, even though the \( r_{1,3} \) distance is too great to manifest a direct NOE, and the intensity of the spin 2 NOE is decreased.
1.8 NMR Time Scales

What happens if the molecule under study is not rigid, or if it contains labile bonds? Certainly biological molecules fall in this class, as the amide protons of proteins, and the imino and amino protons of nucleic acids can exchange with solvent protons. In NMR jargon this is referred to as chemical exchange, although Harris, 1983, uses a better phraseology; *magnetic site exchange*. This phrase is preferable because it includes processes which do not necessarily involve bond breakage, as the term chemical exchange implies. Magnetic site exchange can be further broken down into *mutual* and *non-mutual* exchange. Mutual exchange occurs when two magnetic sites (i.e. two different magnetic environments for a given nuclide) are interchanged. Non-mutual exchange occurs when a single spin undergoes a change to a new magnetic environment.

NMR spectra recorded in the presence of magnetic site exchange depend on the rate of the relevant process. The rates need be compared to the "NMR timescale", which generally refers to lifetimes of the order of 1 sec. to 10^-6 sec. When the exchange lifetime greatly exceeds the NMR time scale, the system is in the slow-exchange regime. The fast-exchange regime results when the lifetime is substantially less than the NMR time scale. Slow-exchange spectra are the superposition of sub-spectra due to each species present, and fast-exchange spectra may be considered as being manifested by a single species whose NMR parameters (chemical shifts, coupling constants) are a populations *weighted average* of those for the individual species.

At first glance one may expect the lines observed in the rapid-exchange regime to be broad, but in fact this exchange regime results in sharp lines. Consider an FT experiment in which an FID is recorded. An exchange process in T_2 changes the
frequency of the FID. If the event occurs in a time which is short compared to
the chemical shift difference between the two frequencies, the FID will be
indistinguishable from that given by a system with the average frequency, and hence
a sharp line results.

What happens if the exchange regime is on the order of the NMR time scale?
Consider a nucleus in environment A with lifetime \( \tau \). The uncertainty in energy for
this system is given by:

\[
\Delta E = \frac{\hbar}{2\pi} \tau_A^{-1}
\]  

(1.17)

This directly influences the linewidths of the NMR lines, giving a broadening
contribution at half-height of:

\[
\Delta V_{1/2} = (\tau_A \pi)^{-1}
\]  

(1.18)

The effect of exchange on the spectrum will only be noticeable if \( \tau_A \leq T_2 \), and under
certain conditions, linewidths can be used to monitor lifetimes. The important
message here is that a slow-exchange process can be monitored as it approaches the
fast-exchange regime through this intermediate regime by NMR spectroscopy. In
the slow-exchange regime the spectrum will contain two distinct, sharp resonances.
As the process approaches the fast regime, the resonances will broaden. When the
fast-exchange regime is finally reached, the spectrum will contain a single, sharp
resonance at a chemical shift which is the populations weighted average of the
chemical shifts of the two species in the slow-exchange regime. The opening of
individual base pairs as a DNA double helix melts can be monitored by such a
process.
1.9 Two-Dimensional Spectroscopy

Although greater than 90% of the NMR spectra for this thesis are one-dimensional, a very brief description of two-dimensional NMR will be given, mainly to outline the advantages of 2D spectroscopy for studies of large molecules. In addition, this section will prove useful when the difficulties of RNA spectroscopy compared to that of DNA are discussed in the next chapter.

Wüthrich, 1986 presents an introduction to two-dimensional spectroscopy as an extension of the one-dimensional experiment. Because this method is particularly easy to understand, a similar approach will be taken here. Figure 1.8a presents the pulse scheme for the most basic 1D NMR experiment discussed in section 1.5. In this experiment the 90° pulse is non-selective in that it rotates the magnetization of all types of protons (regardless of their chemical shift) into the xy-plane. Immediately after this pulse the FID is recorded. During the course of an experiment many FIDs are recorded and added to one another to increase the signal-to-noise (S/N) of the experiment. Fourier transformation of the coadded FID’s then results in the NMR spectrum. As discussed in section 1.5, relaxation of the net magnetization vector back to its equilibrium position begins immediately after the application of the 90° pulse. The envelope of the FID is then determined by the inverse of the apparent transverse relaxation time, $1/T_2^*$, and the repetition rate for each scan is determined by the longitudinal relaxation time, $T_1$. The simple experiment in fig. 1.8a provides information on the chemical shifts and the spin-spin coupling of the individual resonances in the spectrum. This simple pulse scheme can be manipulated to detect other types of interactions by changing

---

1 The true transverse relaxation time can be decreased by such things as field inhomogeneity, resulting in an apparent transverse relaxation time, $T_2^*$. Therefore it is $T_2^*$ and not $T_2$ which is manifested in the NMR spectral features.
Figure 1.8 Schemes for FT Pulse Experiments. (a) A normal 1D FT NMR experiment. (b) A 1D NOE experiment. Resonance A is selectively saturated by irradiation during the time period $\tau$ immediately preceding the 90° observation pulse.
variables such as the flip angle of the pulse, the axis along which the pulse is applied (the phase of the pulse), and by selectively exciting only certain types of protons in the spectrum. In addition, a delay interval can be inserted between the pulse and the acquisition of the FID. Figure 1.8b illustrates one such manipulation; selective saturation of resonance A by irradiation during the delay $\tau$ immediately preceding the 90° pulse. This scheme represents a 1D NOE experiment, detecting for dipolar relaxation.

Although the one-dimensional experiment provides the chemical shifts and spin-spin couplings for all the protons in the sample, the information may be very difficult to extract. A biological sample may contain hundreds of protons, each which manifests an NMR resonance. These protons may in turn be scalar coupled to neighboring protons, resulting in an enormous number of spectral lines. As the chemical shift dispersion covers only about 15 ppm in proton NMR, resolving, and thus assigning single proton resonances in a one-dimensional spectrum is a very formidable task. In addition to the vast number of resonance lines the spectrum is further complicated by $T_2$ broadening (remember that $T_2$ decreases monotonically with size), resulting in further spectral overlap. A one dimensional spectrum may also not be very useful for extracting NOE data, as a single experiment corresponding to the presaturation of each resonance line must be obtained. So although one-dimensional NMR is a very powerful tool for structure elucidation of small organic molecules, it is, in general, not the approach taken by biological NMR spectroscopists.

A 2D NMR experiment can overcome the limitations large molecules impose on 1D NMR. It is analogous to a 1D experiment in that the FID is recorded during the detection period $t_2$. However, in the 2D experiment one or more additional pulses
are applied prior to the observation period. The influence of these pulses on the FID detected during \( t_2 \) depends on the length of the evolution period \( t_1 \). The second dimension is therefore created by incrementing \( t_1 \), collecting an FID for each \( t_1 \) value. Fourier transformation as a function of \( t_2 \) (with fixed \( t_1 \)) results in a modulated 1D spectrum. Further transformation with \( t_1 \) as the independent variable results in the 2D spectrum. A 2D spectrum is generally represented as a contour plot (fig. 1.9). Resonances which interact in some way are then manifested as cross peaks. The 2D experiment helps increase the spectral resolution because it correlates interacting pairs of spins. A spin which is close in chemical shift to another spin may not be coupled to that spin, but rather to one which is at a distinctly different chemical shift. The separation along the second axis helps increase the resolution of the chemical shift information. 2D NMR is also advantageous because information for all of the protons in the molecule can be obtained in a single experiment. An example is the 2D experiment which detects protons coupled via dipolar interactions (called a NOESY). In one dimension, an individual presaturation experiment for each resonance in the spectrum would be necessary to yield the same information obtained in single NOESY experiment. Manipulation of the pulse sequence allows the detection of other types of interactions. The most popular experiment is probably the COSY experiment, which detects through-bond scalar couplings. The pulse sequences for the COSY and NOESY experiments are illustrated in figure 1.10.
Figure 1.9 Contour Plot Representation of a 2D NOESY Experiment. The DNA sequence is shown across the top of the plot. The chemical shift information is contained along the diagonal. Peaks which are dipolar coupled manifest cross peaks. The representative plot clearly illustrates the increased spectral resolution of a 2D experiment, as two spins nearly degenerate in chemical shift are coupled to different partners far apart in chemical shift (indicated by lines), thus helping to resolve their degeneracy.
Figure 1.10  (a) Experimental scheme for the generic 2D experiment.  (b-c) Pulse schemes for the NOESY and COSY experiments respectively. The second dimension is introduced by incrementing the $t_1$ period through a range of values.
Chapter 2

Nucleic Acid Structure and NMR
2.1 Introduction

Nucleic acid structure can be examined on two levels; macroscopic and microscopic. Macroscopically the structure of a particular nucleic acid correlates with its function within the cell. DNA is fairly simple in that it is found in the nucleus of the cell as a long, double stranded helix. RNA, however, is structurally more diverse than DNA. It is also found as double stranded helices, but in addition can adopt a more globular form; short double helical domains connected by single stranded regions. However different the macroscopic structure of nucleic acids, all are assembled from essentially the same microscopic building blocks. This chapter will introduce the building blocks of nucleic acids, with particular attention paid to those properties which are useful to NMR spectroscopists.

2.2 The Anatomy of Nucleic Acids

Nucleic acids are polymeric chains of nucleotide units. For RNA these units consist of β-D-ribose (a cyclic furanose) substituted at the C1' position by one of four different heterocycles via a β-glycosyl C1'-N linkage. The heterocycles are the purine (pur) bases adenine (A) and guanine (G) and the pyrimidine (pyr) bases uracil (U) and cytosine (C). DNA differs from RNA in that the sugar is 2'-deoxyribose and the pyrimidine thymidine (T) replaces uracil (fig. 2.1). The absence of the 2'-hydroxyl group proves to be a major spectroscopic advantage, and will be discussed in a later section. The free bases contain a hydrogen at position 1 (pyr) and at position 9 (pur). It is at this position that the (deoxy)ribose attaches, resulting in a nucleoside unit. Phosphorylation at one of the sugar hydroxyls (usually the 5' position) then yields the nucleotide unit. The nucleotides are normally linked via a 3',5'-phosphodiester bond to produce a single
Figure 2.1 The basic components of RNA. The DNA base thymine has a methyl group at the C5 position, shown in parentheses. In DNA, the sugar is 2'-deoxyribose. The labile protons are shown as open circles, non-labile protons are black circles, and the hydrogen atoms of the sugar hydroxyl are shown as shaded circles. Carbon atoms are not labelled.
stranded nucleic acid chain. An example of a single stranded RNA is illustrated in figure 2.2. The single strands can further associate yielding the familiar right-handed double helix.

RNA and DNA double helices are similar in that they are comprised of essentially planar Watson-Crick base pairs (fig. 2.3) and a sugar phosphate backbone. It is here that the similarities end. When confined to the double helix, 2'-deoxyribose adopts a sugar conformation where the 2' carbon is displaced above the plane of the ring (in the direction of the 5' carbon). This conformation, or sugar pucker, is termed C2'-endo. Ribose, on the other hand, typically adopts the C3'-endo conformation. Surprisingly, the pucker of the sugar plays a major role in determining the overall double helical dimensions. DNA has been observed in many double helical conformations including A, B, C, and left-handed Z-form. Conversely, RNA appears to adopt mainly the A-form helix conformation. It has been observed in a conformation called A'-form, but the helical dimensions of the A' helix do not differ appreciable from those of A-form [Saenger (1984)]. Some properties of A- and B- form double helices are listed in table 2.1, and figure 2.4 illustrates these double helices. For a comprehensive review of nucleic acid structure see Saenger (1984).

2.3 Nucleic Acids and Nuclear Spin

Chapter 1.2 stated that a nucleus must possess a nuclear spin to manifest an NMR resonance. RNA (DNA) contains three such abundant nuclei: $^{14}$N, $^{31}$P, and $^1$H. $^{14}$N has a spin of $I=1$, making it a quadrupolar nucleus. As previously discussed, quadrupolar nuclei are not well-behaved for high resolution NMR, and the major contribution of $^{14}$N to RNA (DNA) spectroscopy is to broaden the resonance lines of
Figure 2.2 Diagramatic representation of an RNA single strand. Hydrogen atoms are shown as black circles, carbon atoms are not labelled.
Table 2.1 Parameters of A and B Double Helices

<table>
<thead>
<tr>
<th>Helix Parameter</th>
<th>A-Helix</th>
<th>B-Helix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handedness</td>
<td>right</td>
<td>right</td>
</tr>
<tr>
<td>Bases per Turn</td>
<td>10.9</td>
<td>10.0</td>
</tr>
<tr>
<td>Height per Base Pair</td>
<td>2.9 Å</td>
<td>3.4 Å</td>
</tr>
<tr>
<td>Pitch</td>
<td>31.6 Å</td>
<td>34.0 Å</td>
</tr>
<tr>
<td>Glycosyl Angle</td>
<td>anti</td>
<td>anti</td>
</tr>
<tr>
<td>Sugar Pucker</td>
<td>3'-endo</td>
<td>2'-endo</td>
</tr>
<tr>
<td>Major Groove</td>
<td>very deep</td>
<td>deep, wide</td>
</tr>
<tr>
<td>Minor Groove</td>
<td>shallow</td>
<td>deep, narrow</td>
</tr>
</tbody>
</table>

Figure 2.3 Watson-Crick base pairs found in duplex DNA and RNA. Hydrogen bonds are represented as dashed lines. Note that the GC base pairs contains three hydrogen bonds whereas the AU(T) has only two. The geometry of the base pairs is such that the hydrogen bonds are linear.
Figure 2.4  (a) Space filling representation of a B-form DNA helix. Note the deep major and minor grooves. (b) Space filling representation of an RNA A-form helix. The minor groove of the A-helix is very shallow, whereas the major groove is deep.
other nuclei in its vicinity. Although $^{31}$P is a spin $I=1/2$ nucleus with a natural abundance of 100% and a relatively large $\gamma$-value, its use as a probe of nucleic acid structure is fairly limited because of its isolated position in the backbone of the RNA (DNA) chain. We are thus left with $^1$H, a sensitive, well-behaved nucleus. The protons in a nucleic acid can be divided into two groups: exchangeable and non-exchangeable. This is illustrated in fig. 2.1; the non-exchangeable being represented by black circles and the exchangeable as open circles. Shaded circles represent the hydrogens of the sugar hydroxyls. These protons are not usually observed in proton NMR spectra of nucleic acids.

2.4 NMR Parameters and DNA/RNA Spectroscopy

Exhaustive studies of both deoxy- and ribonucleotides have resulted in a tabulation of the $^1$H and $^{31}$P chemical shifts, the $^1$H-$^1$H and $^1$H-$^{31}$P coupling constants [Cheng and Sarma (1977) and Cheng et al. (1984)], the $^{13}$C chemical shifts, and the $^{13}$C-$^{31}$P coupling constants [Lankhorst et al. (1984)]. Unfortunately this vast store of information does not make the assignment problem for a given nucleic acid easier, because the observed chemical shifts and coupling constants for nucleic acids differ appreciably from those of the nucleotide units. The chemical shifts are thus said to be sequence specific. Because of the nature of nucleic acid structure, certain types of protons reside in a relatively similar electronic environment irrespective of the species, and consequently a rough tabulation of proton chemical shift ranges for both single stranded and duplex RNA and DNA exists [Cheng & Sarma (1977); Clore et al. (1984); Hare et al. (1983); and Scheek et al. (1984)]. These chemical shift ranges are illustrated in figure 2.5.
Figure 2.5 Proton chemical shift ranges for RNA and DNA.

2.5 Spectroscopy of Exchangeable Protons

The labile (exchangeable) protons in DNA and RNA are comprised of the amino protons of cytosine, adenine, and guanine, and the imino protons of thymidine, uracil, and guanine. These protons, especially the imino protons, play a major role in the study of nucleic acids. They reveal not only structural information, but dynamical information as well. As discussed in section 1.8, labile protons will be observed only if their rate of exchange with solvent is slow on the NMR time scale (in the fast exchange regime the labile proton chemical shift becomes that of water protons). In an 11.74 Tesla spectrometer this corresponds to an exchange rate constant of less than $1 \times 10^3 \text{ min}^{-1}$ [Wüthrich (1986)].

Using the formalism of Eigen (1964), an expression for the intrinsic exchange rate of a labile proton can be developed. In this formalism the exchange is always acid or base catalyzed. The primary steps involve the diffusion controlled formation of a hydrogen-bonded complex between the labile proton and the catalyst, transfer of the proton to the catalyst while in the hydrogen-bonded complex, and dissociation of the complex. The occurrence of a net transfer then depends on the proton transfer equilibrium in the complex. The resulting forward rate for the proton transfer is given by:

$$k_{1r} = k_D \frac{10^{\Delta pK}}{10^{\Delta pK} + 1}$$

(2.1)

where $k_{1r}$ is the proton transfer rate constant, $k_D$ is the second-order rate constant for diffusion controlled complexation, and $\Delta pK$ is the difference between the acidity constants of the proton donor and the proton acceptor in the complex. At 298K, $k_D = 10^{10} \text{ M}^{-1}\text{s}^{-1}$. The rate constant for exchange is then $k_{1r}$ multiplied by the concentration of the catalyst. This can be written in a general form, accounting for
the presence of several catalysts:

\[ k_{intr} = \sum_i k_{tr,i} [\text{catalyst}]_i \]  

(2.2)

where \( k_{intr} \) is the intrinsic exchange rate constant, \( k_{tr,i} \) is the proton transfer rate constant when complexed with catalyst \( i \). In aqueous solution \( \text{OH}^- \) and \( \text{H}_2\text{O} \) are potential base catalysts, and \( \text{H}_3\text{O}^+ \) and \( \text{H}_2\text{O} \) are potential acid catalysts. Buffers may also provide additional exchange pathways. Values of \( k_{intr} \) have been estimated using eqn. 2.2 and literature data for the pK\(_a\)'s of the proton exchange sites in nucleotide units [Wüthrich (1986)], with the result that \( k_{intr} \geq 10^6 \text{ min}^{-1} \) at 25 °C for the labile base protons of oligonucleotides over the entire pH range. This value is three orders of magnitude greater than that necessary for the observation of these protons on the NMR time scale! Luckily, hydrogen bonding (and other) interactions protect some of these protons from solvent exchange in the native nucleic acid. This decreases the intrinsic rate of exchange sufficiently under certain circumstances such that some labile protons can be easily observed.

2.6 Imino Protons are a Probe of Secondary Structure

Labile protons which can be observed under the conditions of hydrogen bonding are exemplified by the imino protons of thymidine, uracil, and guanine. It is clear from figure 2.3 that these protons are involved in a hydrogen bond when G, T, and U are paired with their corresponding partners in the Watson-Crick geometry. Stated another way, the observation of imino protons usually requires that the bases which manifest those resonances are in a duplex region of the molecule. The imino protons of single stranded regions are generally not protected from exchange, and are not observed. Additional base pairing geometries are known to exist in tRNAs, and in
these cases additional labile protons are protected from exchange. Because of the
nature of tRNA structure, some of these base pairs involve tertiary interactions
(fig. 2.6). The observation of an imino proton resonance in a "globular" RNA may
therefore provide tertiary structural information as well.

Because of their extreme low-field chemical shift, imino protons are easily
identified in the NMR spectrum (fig. 2.5). This is a consequence of both hydrogen-
bonding interactions and the geometry of the imino proton with respect to the
stacked conjugated ring systems. In addition to being far removed from the other
resonances in the spectrum, each Watson-Crick base pair contains only a single
hydrogen-bonded imino proton. Each base pair then manifests a single imino proton
resonance with a relative intensity of unity. For a small DNA oligomer or an RNA
which contains relatively short duplex regions it is a straight forward task to count
the number of base pairs which participate in the double helix. The major weakness
of this type of spectroscopy is the intense signal of the water protons. The protons
of the oligonucleotide are present in only millimolar concentration whereas the
water proton concentration is 110 molar. Limitations of the spectrometer analog-
to-digital converter require that the water resonance be supressed by a substantial
amount (about a factor of 1000) to satisfactorily digitize the signal manifested by
the oligonucleotide. The method used to supress water for the work described here is
called the 1331 pulse train [Hore (1983)], and is illustrated in fig. 2.7.
Figure 2.6 Some of the non-Watson-Crick base pairs observed for yeast tRNA^{Phe}. 
(a) Pulse sequence. The tip angle $\alpha$ is $11.25^\circ$.

(b) Sphere showing the trajectory the tip of the magnetization vector $M$ sweeps out during the experiment at an offset $1/2\tau$. At the "north pole", $Z$, the magnetization is rotated by $\pm X$ pulses separated by 180° precession, which leaves $M$ along the $-Y$ axis at the end of the pulse train. Magnetization which is on-resonance does not precess, and is tipped back and forth in the $YZ$ plane before returning to the $Z$-axis at the end of the pulse train. Magnetization which is on-resonance is therefore not observed.

2.7 Assignment of DNA and RNA Imino Proton Resonances

Imino proton resonances provide maximum information when they are assigned to the specific base pair which manifests them. Although RNA and DNA helix dimensions differ, the assignment strategy for these protons for RNA and DNA is identical. To successfully assign the imino proton spectrum, the base pair type as well as the position of the pair within the helix must be determined. This is accomplished using NOE connectivities. A Watson-Crick AU (AT) base pair will manifest a sharp imino-to-aromatic CH2 NOE, whereas a GC pair will manifest a broad imino-to-amino NOE [Sanchez, et al. (1980)], allowing easy identification of the base pair type. Inter-base pair imino-imino NOEs are then used to determine nearest neighbors [Johnston & Redfield (1981)]. In cases of near degeneracy, second order NOE's, representing next nearest-neighbor interactions, can be used to facilitate the assignment procedure. This assignment strategy is illustrated in figure 2.8. It is important to realize that this method does not allow one to determine the sense of the helix. In other words, nearest neighbors can be easily identified, but it is impossible to distinguish which NOE represents the 3' neighbor, and which the 5' neighbor. Knowledge of the sequence can distinguish the sense of the helix only if the connectivity patterns are unique. If the sequence has symmetry, or contains two helical domains of identical sequence, it may be impossible using this type of information alone to unambiguously assign the imino proton spectrum.

The basic experimental scheme involves saturation of a single imino proton resonance by application of a selective rf pulse at its resonant frequency, resulting in the perturbed spectrum. A non-perturbed “reference” spectrum is also collected by moving the saturation frequency to a region of the spectrum which contains no
peaks (off resonance). The perturbed and non-perturbed spectra are collected in parallel by alternating scans between on and off resonance. The perturbed spectrum is stored in one block of memory, and the non-perturbed in a different block. Subtraction of the non-perturbed spectrum from the perturbed spectrum yields the NOE difference spectrum. A complete set of data then requires the collection of a perturbed and non-perturbed spectrum for every imino proton resonance.

There are times when the collection of 1D NOEs may be inconvenient, and a two-dimensional experiment is more practical. Such a situation arises when the imino proton spectrum is very crowded, and the increased spectral resolution introduced by two-dimensional techniques proves beneficial. A 2D experiment is also desirable in this situation because all of the NOE information can be obtained in a single experiment, vastly decreasing the spectrometer time necessary to collect the desired information.

At this point, one may wonder why a 2D experiment isn't preferrential under all circumstances. As mentioned earlier, spectra which are recorded in H₂O solution require significant suppression of the H₂O resonance to properly digitize the oligonucleotide signal. It turns out that the preferrential method of water suppression, the 1331 experiment, produces a rolling baseline when the spectra are phase corrected. This can be easily compensated for in 1D spectra using baseline correction routines, but is very problematic in 2D spectroscopy, as even minor baseline irregularities result in t₂ streaking. A similar supression scheme is known that does not introduce baseline distortions, the 11-echo experiment, but the level of water suppression is generally not as good as that obtained with the 1331 pulse scheme. There are additional factors in the 2D experiment which attenuate the level of solvent suppression achieved compared to the 1D experiment which will
Figure 2.8 Assignment strategy for a 1D NOE Experiment. (A) Reference or "off-resonance" spectrum. (B) Perturbed or "on-resonance" spectrum. The resonance which has disappeared is that which is being saturated (denoted by arrow). (C) NOE difference spectrum obtained from subtracting the perturbed spectrum from the reference spectrum. The large, truncated peak is the peak which was saturated. The other peaks, called NOE's, represent fractional intensity changes of nearby neighbors. (D) NOE difference spectrum of one of the NOE peaks in (c) showing the connectivity of the two resonances.
not be discussed. The final result is that a H$_2$O NOESY requires much higher sample concentrations than does an equivalent one-dimensional experiment. In addition, collection of a 2D spectrum with reasonable digital resolution requires a large amount of disk space. The observation of imino proton resonances requires a spectral window of about 12,000 Hz (at 500 MHz). Because of disk space limitations, the maximum number of complex points which can be collected in the t$_2$ dimension (with 512 t$_1$ datapoints) is approximately 2048, resulting in a digital resolution of only six Hertz per point, and a rather large two dimensional data matrix. In one-dimension, 8192 datapoints can easily be collected without taxing disk space, increasing the digital resolution of the spectrum two-fold. In this instance the number of datasets collected depends on the number of resonances in the imino proton spectrum, but will definitely be fewer than 512! So opting for a 2D dataset in H$_2$O generally results in a trade-off rather than a bonus; all of the NOE information is collected in a shorter time interval with the added benefit of increased spectral resolution, but there is a decrease in digital resolution, and higher sample concentrations are required.

2.8 Assignment of the Nonexchangeable Protons in DNA

Although the labile protons of DNA can give much insight into the structure and dynamics of an oligonucleotide, spectroscopy of the nonexchangeable protons yields by far the greatest amount of structural information. Because these protons are nonlabile, the spectroscopy can be performed in D$_2$O instead of H$_2$O, thereby eliminating the dynamic range problem discussed above. In addition, the residual HOD resonance can be removed from the spectrum by presaturation techniques. Given that a proton-proton NOE is observable to a distance of about 4 Å, a general
sequential assignment strategy utilizing both COSY and NOESY spectra has been
developed for deoxyribooligonucleotides [Hare et al. (1983)]. In this approach, the
pyrimidine residues are readily identifiable by their COSY cross peaks; cytosine
manifests a strong CH5-CH6 cross peak, and thymidine manifests a weak four bond
methyl-CH6 cross peak (fig. 2.9). Additional COSY peaks are attributed to the
protons of deoxyribose. The COSY spectrum then allows one to group the resonances
according to the base they are on, or to a single sugar, without yielding any
information about the position of these residues in the sequence. The sequential
information is revealed by a NOESY spectrum. For a right handed B-form double
helix, the purine CH8 or pyrimidine CH6 is close in space to the CH1' of its own
sugar and to the CH1' of the (n-1) sugar, using the convention of counting the bases
from the 5' end of each strand. This region of the spectrum is called the
"fingerprint" region. An internal check of the CH6/CH8(n)-CH1'(n-1)
connectivity is provided by the existence of a CH6/CH8(n) - C2"H(n-1) crosspeak.
Additional NOESY cross peaks exist which connect the aromatic CH5 of cytidine to the
CH6 or CH8 of an (n-1) base, the cytidine CH5 to the methyl group of an (n-1)
thymidine, and the CH5 of cytidine to an (n-1) CH1' proton (fig. 2.10). Table 2.2
gives a complete tabulation of the short sequential nonexchangeable proton distances
for both A and B double helices. It should be noted that a cytosine CH5-CH6 cross
peak is observed in both COSY and NOESY spectra, as is a thymidine methyl-CH6
cross peak (figs. 2.9 and 2.10). Therefore, the methyl and aromatic chemical shifts
of cytidine and thymidine can be discerned from a crowded NOESY spectrum by
comparison with a COSY spectrum. Using this method, one can then "walk"
sequentially along either strand of a B DNA helix. The power of this assignment
method originates from the geometry and chemical shift dispersion of B DNA.
Figure 2.9 COSY spectrum of a duplex DNA oligomer. Boxes 1 and 2 show the regions of the spectrum which contain the four bond methyl-CH6 crosspeaks of thymidine and the strong cytosine CH5-CH6 crosspeaks respectively. The DNA sequence is denoted in the lower right-hand corner. Box 1 contains no cross peaks, as no four bond couplings were observed for this sequence.
Figure 2.10 NOESY spectrum of the DNA oligomer shown in figure 2.9. Labelled boxes indicate the regions of the spectrum most likely to contain crosspeaks from those protons.
Table 2.2  Short Sequential Distances of non-labile protons in A and B DNA Helices

<table>
<thead>
<tr>
<th>Distance</th>
<th>A-DNA</th>
<th>B-DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1'; 5'</td>
<td>4.6 Å</td>
<td>3.0 Å</td>
</tr>
<tr>
<td>2'; 1'</td>
<td>4.0</td>
<td>-</td>
</tr>
<tr>
<td>2'; 3'</td>
<td>4.1</td>
<td>-</td>
</tr>
<tr>
<td>2'; 5'</td>
<td>2.7</td>
<td>-</td>
</tr>
<tr>
<td>2''; 2'</td>
<td>-</td>
<td>3.7</td>
</tr>
<tr>
<td>2''; 5'</td>
<td>2.4</td>
<td>4.0</td>
</tr>
<tr>
<td>3'; 2'</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3'; 3'</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4'; 5'</td>
<td>4.1</td>
<td>3.4</td>
</tr>
<tr>
<td>1'; 5</td>
<td>-</td>
<td>4.3</td>
</tr>
<tr>
<td>1'; 6</td>
<td>4.6</td>
<td>3.5</td>
</tr>
<tr>
<td>1'; 8</td>
<td>4.6</td>
<td>3.6</td>
</tr>
<tr>
<td>2'; 5</td>
<td>3.5</td>
<td>3.6</td>
</tr>
<tr>
<td>2'; M</td>
<td>3.8</td>
<td>3.4</td>
</tr>
<tr>
<td>2'; 6</td>
<td>2.0</td>
<td>4.0</td>
</tr>
<tr>
<td>2'; 8</td>
<td>2.1</td>
<td>3.8</td>
</tr>
<tr>
<td>2''; 5</td>
<td>-</td>
<td>2.7</td>
</tr>
<tr>
<td>2''; M</td>
<td>-</td>
<td>2.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distance</th>
<th>A-DNA</th>
<th>B-DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2''; 6</td>
<td>3.7 Å</td>
<td>2.2 Å</td>
</tr>
<tr>
<td>2''; 8</td>
<td>3.9</td>
<td>2.1</td>
</tr>
<tr>
<td>3'; 5</td>
<td>3.3</td>
<td>-</td>
</tr>
<tr>
<td>3'; M</td>
<td>3.0</td>
<td>-</td>
</tr>
<tr>
<td>3'; 6</td>
<td>3.1</td>
<td>4.6</td>
</tr>
<tr>
<td>3'; 8</td>
<td>3.0</td>
<td>4.6</td>
</tr>
<tr>
<td>4'</td>
<td>4.1</td>
<td>-</td>
</tr>
<tr>
<td>5; 5</td>
<td>3.7</td>
<td>4.7</td>
</tr>
<tr>
<td>5; M</td>
<td>3.7</td>
<td>4.7</td>
</tr>
<tr>
<td>6; M</td>
<td>3.5</td>
<td>3.8</td>
</tr>
<tr>
<td>6; 5</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td>6; 8</td>
<td>4.5</td>
<td>4.8</td>
</tr>
<tr>
<td>8; 5</td>
<td>3.8</td>
<td>3.9</td>
</tr>
<tr>
<td>8; M</td>
<td>3.4</td>
<td>3.8</td>
</tr>
<tr>
<td>8; 8</td>
<td>4.6</td>
<td>5.0</td>
</tr>
</tbody>
</table>

\(a\)Distances for A-RNA are similar to those listed for A-DNA.

2.9 Spectroscopy of RNA

RNA seems an almost ideal candidate for structural studies by NMR. The x-ray crystallographic structure of yeast tRNA^Phe certainly demonstrated that RNA is capable of structural diversity usually associated only with proteins [Kim (1975,1976); Quigley & Rich (1976)]. Although commercial chemical synthesis methods are not yet widely available, the development of in vitro transcription methods using DNA templates and T7 RNA polymerase allow one to produce milligram quantities of almost any RNA sequence [Milligan et al. (1987)]. Certainly then, the availability of interesting RNA species is not the rate-limiting step in the application of NMR spectroscopy to the elucidation of RNA structures.

Spectroscopy of the non-exchangeable protons of RNA is more problematic than is that of DNA. The sequential NOE connectivities used to assign DNA oligomers are not present in RNA. According to table 2.2, the CH6/CH8(n)-CH1'(n-1) distance in the A-form helix is 4.6Å, compared to 3.5Å for the B-form helix. This distance is just beyond that which manifests an NOE. There are short sequential distances in the A-form helix, the CH6/CH8(n)-CH2'(n-1) distance is 2Å, and CH6/CH8(n)-CH3'(n-1) distance is 3Å, but because of the presence of the 2'-hydroxyl in RNA, the 2' and 3' protons resonate at about the same chemical shift, making it difficult to sort out the different connectivity pathways. Additional ribose protons resonate in this region as well, further complicating the issue. Because the CH1' proton is fairly well-resolved from the other ribose protons, a possible experimental scheme for using this nucleus as a "reporter" nucleus for the CH8/CH6(n)-CH2'(n-1) crosspeak is possible if the magnetization from the CH2'(n) proton could be transferred to the CH1'(n) proton. An obvious possibility is to transfer the magnetization through scalar interactions, as CH1' and CH2' are bonded to
neighboring carbons. The experimental strategy would then involve the cross-
relaxation of CH8/CH6(n) with CH2'(n) and CH2'(n-1), followed by subsequent
magnetization transfer from the CH2'(n) proton to the CH1'(n) proton via J-
coupling. Although a seemingly plausible experimental scheme, this method is
thwarted by the geometry of the A-form helix. It turns out that the scalar coupling
constant between CH1' and CH2' in RNA is essentially zero [Altona & Sundaralingham
(1972)]1 At this point the clever spectroscopist may notice that the CH1'(n)-
CH2'(n) distance in the A-form helix is about 4 Å, which is just within the NOE
limit. Therefore another experimental scheme can be envisioned whereby CH1' is
used as reporter molecule, but this time the magnetization transfer from CH2' to
CH1' is via dipolar, not scalar, interactions. In this scheme the mixing time of the
NOESY is chosen to be sufficiently long such that spin diffusion pathways are
observed, and hence weak CH6/CH8(n)-CH1'(n) and CH6/CH8(n)-CH1'(n-1)
crosspeaks will be observed. This method has been successfully used for small RNA
oligomers [Puglisi (1989); Davis (1989)], but presents a problem for large
RNAs, as large molecules may relax quickly compared to the mixing time necessary
for spin diffusion processes occur. Even when the molecule is small enough such
that this approach is feasible, RNA is a more problematic species to study than is
DNA. In RNA uracil replaces thymidine, and as can be seen from fig. 2.3, uracil and
cytosine have identical spin systems. What this means is that the cytosine CH5(n)-
CH6(n) crosspeak, which is unique in DNA, is not unique in RNA. Uracil manifests
this cross peak as well. The net result is that it is impossible to distinguish between
adenine and guanine, and cytosine and uracil based on their cross peak patterns in a
NOESY spectrum of RNA. In addition, the chemical shift dispersion of the aromatic
resonances in RNA is generally decreased compared to that for DNA, meaning that
these resonances occupy a smaller chemical shift range. The absence of the thymidine methyl group may seem to be another hinderence, but this is only true in spectra of RNAs which do not contain any modified bases. Some naturally occurring RNAs contain methylated bases, and in these cases the methyl resonances can be used as assignment ”markers”. Because only a few bases will manifest methyl resonances, their assignment is relatively easy.

Yet another difficulty is the fact that sequential assignments are usually dependent on the presence of regular secondary structure. Even if one could selectively label the CH6, CH8, and CH2′ positions with 3H such that only the CH6/CH8(n)-CH2′(n) and CH6/CH8(n)-CH2′(n-1) connectivities would be observable in a 3H NOESY spectrum, RNAs usually contain single stranded regions which would most likely manifest NOEs between positions other than those which are labelled. If the crystal structure of the RNA has been determined, one can certainly decide which positions are advantageous for isotopic labelling. Unfortunately, there exists no general approach for solving the spatial arrangement of single stranded regions or of the tertiary arrangement of helical domains within an RNA molecule.

At this point it appears as though little structural information can be practically ascertained via NMR spectroscopy of RNA. This may be presently true for non-exchangeable proton spectroscopy of large RNAs, but is certainly not true of spectroscopy involving the labile protons. The vast store of information these protons supply is exemplified by the NMR studies which were done with different tRNAs and 5S RNA. As discussed earlier, the observation of an imino proton resonance in an RNA spectrum does not limit the origin of the resonance to a Watson-Crick base pair of a duplex region. Non-Watson-Crick tertiary base pairs as well as base triples in tRNA have been identified both in the crystal structure and from
NMR evidence [for a review see Reid (1981)]. In addition, NOE evidence supports the fact that two helical domains from the tRNA cloverleaf secondary structure are co-stacked in solution (as well as in the crystal) to form the familiar L-shaped three-dimensional structure [Hare & Reid (1982)]. NMR spectroscopy of the imino protons can therefore furnish information about the tertiary, as well as the secondary, structure of a ribonucleic acid. Certainly the availability of the tRNA crystal structure simplified the assignment task for the NMR spectroscopists. Armed with this information, the spectroscopists, in a sense, already knew what to look for. In the absence of a crystal structure, one must be able to assign secondary and tertiary interactions whose origins are unknown. Finally, the motivation for the work in this thesis becomes clear; to study the structure of an RNA, in the absence of crystallographic information, not only to learn about the particular species under study, but also to begin laying a foundation for the general application of NMR spectroscopy to the study of relatively large RNA molecules.
Chapter 3

Experimental Methods
3.1 Introduction

Most of the laboratory and spectroscopic techniques used to produce, purify, and study the different RNA species were standard techniques. To avoid confusion and clutter when discussing the results obtained for each RNA species, the preparatory methods will be described in this chapter. Also, a set of standard laboratory protocols is given in appendix A for those who wish to use these techniques. The handling of the pure sample for spectroscopy will be described in the text of the appropriate chapter, and the exact buffer and spectroscopic parameters will be given in the figure legends.

3.2 Isolation of RV/R/lplBl20 Vector DNA

The plasmid DNA necessary to transcribe the originally conceived RNA sequence, RNA1, was constructed by Dr. Rosalind Kim and given to me. Plasmid DNA which encoded for subsequent RNA's was constructed using pBl20 (IBI Chemicals) as the vector. The sequence and endonuclease restriction sites of pBl20 and the DNA oligomers necessary to construct the insert DNA for each RNA species are given in figure 3.1.

20 µg of pBl20 plasmid DNA was digested with 20 U EcoRI (BRL) in 300 µL of 50 mM NaCl, 90 mM tris·HCl, pH 7.5, 10 mM MgCl₂, and 0.1 µg/µL bovine serum albumin (BSA) for 1 hour at 37°C. 3 µL of the digestion reaction was removed, dried under vacuum, re-dissolved in 2 µL 1X native gel loading buffer (5% glycerol, 0.05% xylene cyanol, 0.05% bromophenol blue) and analyzed on a 0.8% agarose gel [Maniatis, et al. (1982) p. 163] to check for complete digestion. The EcoRI was removed by phenol/chloroform extraction [Maniatis et al. (1982) p. 458] and the DNA precipitated with ethanol [Maniatis, et al. (1982) p. 461]. The
Figure 3.1 Sequences of pIBI20, Insert DNA's and their respective RNA's. The EcoRV and EcoRI endonuclease restriction sites are denoted on the pIBI20 sequence. Only one strand of pIBI20 is shown. The +1 designates the +1 position of the T7 RNA polymerase promoter sequence. RNA1, RNA3, RNA4, and RNA6 insert DNA sequences were constructed from four oligomers, with the break denoted by the heavy line. The figure continues on the next five pages.
5' ATCGAATTTA TACGACTCAC TATAGGGGC TGATGAGTTG GGTGAGAAGA AACATGCG3'
3', TAGCTTTAATT ATGCTGAGTG ATATCCCTCG ACGACGCAAG CACTCTTGCT TTGTCGCTTT AA

3' EcoRV  transcription  EcoRI  5'

50- U U A A G
50- C G G C
50- G C U A
50- A U
50- C C
50- A
50- C
50- G
50- U

Chemical synthesis

Cleavage Site

Transcription

dC-RNA2
TACGAATTA  TACGACTCAC  TATAGGGAGC  CTTGTCAAA  G  GAGAGATC  T  TTCTGAG
ATGCTTAATT  ATGCTGAGTG  ATATCCCTCG  GAACAGTTT  C  CTCTCTAGG  A  AAGACTC

EcoRV

transcription →

TCTGTGAGGA  CGAAACAGGG  CTCCG
AGACAGTCCT  GCTTTGTCCC  GAGGGCTTA

EcoRI

66-UUAAGGAAGGGACUGUUCCU~

III

Cleavage Site

RNA4
ATCGAATTAA TACGACTCAC TATAGGGAGC CTTGTCAAA G GAGAGATC T TTCTGATGAG
TAGCTTAATT ATGCTGAGTG ATATCCCTCG GAACAGTTT C CTCCTCTAGG A AAGACTACTC

EcoR V

transcription

TCCGTGAGGA CAAACAGGG CTCCCG
AGGGCACTCT GTTTG7CCC GAGGGCTTTAA

EcoR I

66-U U A A G
C G C G C G
C G U A C G
G C G C G
G G C G U
A A C C U G

Cleavage Site

RNA6

AAUC

AA A A AC

AA A A
pellet was washed twice with 70% ethanol to remove residual salts, and dried under vacuum. EcoRI digested pIBl20 (RI/pIBl20) was digested with 20 U EcoRV (BRL) in a volume of 300 µL of 150 mM NaCl, 6 mM tris·HCl, pH 7.4, 6 mM MgCl₂, and 6 mM β-mercaptoethanol (BME) for 1 hour at 37°C. After confirming that the digestion was complete the sample was extracted with phenol/chloroform and precipitated with ethanol as described above. The pellet was re-dissolved in 20 µL TE buffer (10 mM tris·HCl and 1 mM EDTA), pH 7.6, 5 µL 5X native gel loading buffer was added, and the sample loaded on a 20 X 20 cm preparative 0.8% agarose gel containing 0.5 µg/mL ethidium bromide, 5 µL per well. Electrophoresis was carried out at 20 volts in 1X TAE buffer (40 mM tris-acetate, 1 mM EDTA) for 18 hours. The gel was analyzed by short-wave UV irradiation and the appropriate band excised. Pure RV/RI/pIBl20 DNA was recovered from the gel slice by electroelution into 14K molecular weight cut-off (MWCO) dialysis tubing [Maniatis et al. (1982) p. 164-165]. The solution inside the dialysis tubing was removed and concentrated to a volume of 50 µL by extraction with 1-butanol [Maniatis et al. (1982) p. 463], extracted twice with phenol/chloroform, and precipitated with ethanol. The pellet was washed twice with 70% ethanol, dried, re-dissolved in 50 µL TE, pH 7.6, and quantitated by agarose gel electrophoresis [Maniatis et al. (1982) p. 469]. The vector DNA was stored at 4°C.

3.3 Construction of Insert DNA

Synthesis and Purification of the DNA Oligomers. The DNA sequences shown in figure 3.1 were synthesized by David Koh using solid state phosphoramidite chemistry [Beaucage & Caruthers (1981)] and an Applied Biosystems DNA synthesizer. The 5'-dimethoxytrityl (DMT) group was left on at the last step of the
synthesis for subsequent reverse phase purification of the oligonucleotides. The resin was removed from each column and placed in a 1.5 mL plastic screw-top centrifuge tube. The resin was covered with 1 mL of concentrated ammonium hydroxide and incubated at 55°C for 18 hours in a heat block to deprotect the exocyclic amines of deoxyadenosine, deoxycytidine, and deoxyguanosine, and to remove the 5'-acetyl capping group of the failed sequences. The ammonium hydroxide was removed with a pasteur pipette and transferred to a fresh tube. The resin was washed with 1-2 mL conc. ammonium hydroxide, which was removed and pooled with the first fraction. The oligonucleotides were then purified on Applied Biosystems Oligonucleotide Purification Cartridges (OPC) using the modified OPC protocol outlined in Applied Biosystems User Bulletin No. 51. The pure oligos were evaporated to dryness. The 5'-DMT was removed by resuspending the oligos in 100 µL 80% (v/v) acetic acid and incubating at room temperature for 20 min. The solution was placed in a dry ice/ethanol bath until frozen and dried by lyophilization. This resulted in a thick yellow oil. The oil was resuspended in 100 µL H₂O and dried under vacuum to yield a white powder. This was resuspended in 100 µL TE, pH 7.6 [Maniatis, et al. (1982) p. 448] and the number of optical density (OD) units determined at an absorbance of 260 nm (A₂₆₀).

5'-Phosphorylation of the DNA Oligomers. 0.05 A₂₆₀ units of each oligonucleotide was 5'-phosphorylated with 5 units of T4 polynucleotide Kinase (BRL) in 30 µL of 50 mM tris·HCl, pH 7.5, 10 mM MgCl₂, 5 mM dithiothreitol (DTT), and 1 mM spermidine during a 1 hour incubation at 37°C. The reactions were extracted with an equal volume of 50:50 phenol/chloroform [Maniatis, et al. (1982) p. 458] and the aqueous phases of the appropriate strands pooled in a fresh
eppendorf microfuge tube. The organic phases were back-extracted with 30 µL TE, pH 7.6, and these aqueous phases pooled with the first fraction. The oligomers were precipitated with ethanol at -70°C [Maniatis, et al. (1982) p. 461], washed 2X with 70% ethanol, and dried. The DNA was resuspended in 20 µL TE, pH 7.6 and heated to 90°C for 1 min. The oligos were annealed by simply turning off the heat block and allowing it to return to room temperature. If the insert DNA was constructed from four DNA strands, the appropriate annealed duplexes were mixed in a single eppendorf microfuge tube and precipitated with ethanol. The pellet was resuspended in 20 µL 1X ligase buffer (40 mM tris·HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 1mM ATP, and 50 µg/mL BSA) containing 5U T4 DNA Ligase (BRL) and the reaction incubated at 37°C for 1 hour. If the insert DNA was constructed from only two oligos this ligation step was skipped. The duplexed, ligated, insert DNA was purified on a 10x20x0.07 cm 8% polyacrylamide gel [Maniatis, et al. (1982) p. 173-178]. The gel was visualized by staining with ethidium bromide [Maniatis, et al. (1982) p. 161], the appropriate band excised, and the insert DNA recovered by electroelution into a piece of 12-14K MWCO dialysis tubing [Maniatis, et al. (1982) p. 164]. The solution inside the dialysis membrane was removed and concentrated to a volume of about 50 µL by extracting with 1-butanol [Maniatis, et al. (1982) p. 463], extracted 2X with an equal volume of 50:50 phenol/chloroform, and precipitated with ethanol. The precipitate was washed 2X with 70% ethanol and dried. After resuspending in 20 µL TE, pH 7.6, the insert DNA was quantitated on a 8x10x0.07 cm 8% polyacrylamide gel by comparing the fluorescence of the ethidium bromide stained insert band with those of a Haelll/pBR322 DNA control lane. The insert DNA was stored at 4°C.
3.4 Construction and Propagation of Plasmid DNA

*Ligation of Vector and Insert DNA.* 0.014 µg of insert DNA was ligated to 0.1 µg of RV/RI/plBI20 vector DNA with 2U T4 DNA ligase (BRL) in a volume of 10 µL of 40 mM tris·HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, and 50 µg/mL BSA during a 4 hour incubation at 15°C. 3 µL of the ligation reaction was added directly to 100 µL of SubCloning Efficiency DH5α Competent Cells (BRL). Transformation was carried out according to the instructions enclosed with the cells. The final reaction was diluted 100:1 with LB media (10 g/L bacto-tryptone, 5 g/L bacto-yeast extract, 10 g/L NaCl, pH 7.0), and 100 µL of this was spread onto an LB agar plate containing 50 µg/mL ampicillin [Maniatis, et al.(1982) p. 70] and grown overnight at 37°C. Colonies were screened for the appropriate plasmid by transcription of EcoRI linearized DNA isolated from 5 mL minicultures [Maniatis, et al. (1982) p. 360-369].

3.5 Large Scale Purification of Plasmid DNA

This procedure is modified from methods outlined in the Promega Catalog, Birnboim & Doly (1979), and Maniatis et al. (1982). A three liter preparation typically yields 10-15 mg pure plasmid DNA.

3L of LB media containing 50 µg/mL of ampicillin [Maniatis, et al. (1982) p. 68] was inoculated with a 25 mL overnight culture of DH5α *E. Coli* containing the desired plasmid. The culture was grown overnight (16-18 hours) at 37°C with vigorous shaking. The cultures were cooled for 30 min. in an ice water bath and the cells collected by centrifugation at 5000 rpm in a GS-3 rotor for 10 min. at 4°C. The supernatant was decanted and the cells resuspended in 90 mL ice cold solution I (50 mM glucose, 25 mM tris, 15 mM HCl, 10 mM EDTA) containing 4 mg/mL
lysozyme (Sigma). After incubation for 30 min. in an ice water bath 175 mLs freshly prepared solution II (0.2 M NaOH, 1% SDS) was added to the cell suspension and the solution mixed thoroughly by inversion. This suspension was held in an ice water bath for 20 min., after which 130 mL solution III (5 M KOAc, 5M HOAc, pH 4.8 [Maniatis, et al. (1982) p. 447]) was added and the solution mixed thoroughly by inversion. This mixture was held in an ice water bath for 30 min., divided into two GSA centrifuge bottles, and spun at 8000 rpm for 30 min. at 4°C. The supernatant was carefully decanted through a Mr. Coffee coffee filter into a 1L bottle and the pellets discarded. 0.6 volumes of room temperature isopropanol were added to the supernatant and the suspension incubated at room temperature for 15 min, to precipitate the nucleic acids. The nucleic acids were pelleted by centrifugation at 8000 rpm and 21°C for 40 min. The supernatant was carefully decanted and discarded. The pellets were washed with 70% ethanol and dried by lyophilization. The dry pellets were dissolved in 8 mL of TE buffer, pH 8, transferred to a 50 mL plastic screw-top tube (Corning), and equilibrated to 37°C. The RNA was digested by addition of 50 µL 10 mg/mL DNase-free RNase A [Maniatis, et al. (1982) p. 451] and incubation at 37°C for 30 min. 100 µL 10 mg/mL Proteinase K [Maniatis, et al. (1982) p. 450], 100 µL (w/v) SDS, and 1.75 mL H2O were added to the RNase digestion reaction and incubated for 1 hour at 37°C to digest the proteins. The solution was transferred to a polyallomer SS-34 centrifuge tube and successively extracted with 10 mL phenol/chloroform [Maniatis, et al. (1982) p. 458], vortexing for 3-5 min. for each extraction, until no white precipitate appeared at the aqueous/organic interface. The aqueous phases were pooled, made to 0.1 M NaCl, and the DNA precipitated with ethanol. After incubation for 30 min. at -20°C the precipitate was pelleted by centrifugation at 12K rpm and 4°C for 30
min. The ethanol was decanted, the pellet washed twice with 70% ethanol, and dried by lyophilization. The pellet was dissolved in 3.36 mL H$_2$O. To this was added 0.64 mL 5M NaCl, and 13% PEG-8000 containing 0.02% NaN$_3$ which had been previously de-ionized by stirring a 13% PEG-8000, 8% BioRad AG 501-X8 20-50 mesh resin suspension for 18 hours at room temperature. The suspension was held in an ice water bath for approx. 1 hour to precipitate the plasmid DNA. The DNA was pelleted by centrifugation at 12K rpm and 4°C for 30 min. The supernatant was carefully decanted, the pellet washed twice with 70% ethanol, and dried by lyophilization. The pellet was resuspended in 1-2 mL TE, pH 8 and quantitated by both UV absorbance [Maniatis, et al. (1982) p. 468] and by 0.8% agarose gel electrophoresis [Maniatis, et al. (1982) p. 469].

3.6 Linearization of Plasmid DNA with EcoRI

Analytical EcoRI Digests. To avoid wasting EcoRI each batch of plasmid was titered analytically. Stock solutions which would be used for the large scale reaction were used for the analytical reactions to ensure complete success. The plasmid to be linearized was diluted to a concentration of 2 mg/mL with TE, pH 7.6 [Maniatis, et al. (1982) p. 440]. Analytical EcoRI digests were prepared according to Table 3.1. Each analytical reaction was incubated for 1 hour at 37°C. After incubation, 3 μL from each reaction was removed, diluted with 1 μL 5X native gel loading buffer (30% glycerol, 0.25% xylene cyanol, 0.25% bromophenol blue) and loaded onto a 0.8% agarose minigel containing 50 μg/mL ethidium bromide. Electrophoresis was carried out in 1X TAE buffer [Maniatis, et al. (1982) p. 156] at 15 V/cm for 30 min. and the stained DNA bands visualized by irradiation with short-wave UV light.
Table 3.1 Analytical EcoRI Digests

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X EcoRI(^1)</td>
<td>3.0 μL</td>
<td>3.0 μL</td>
<td>3.0 μL</td>
<td>3.0 μL</td>
</tr>
<tr>
<td>2 mg/mL BSA</td>
<td>1.5 μL</td>
<td>1.5 μL</td>
<td>1.5 μL</td>
<td>1.5 μL</td>
</tr>
<tr>
<td>Plasmid DNA</td>
<td>2.0 μL</td>
<td>2.0 μL</td>
<td>2.0 μL</td>
<td>2.0 μL</td>
</tr>
<tr>
<td>EcoRI (1U/μL)</td>
<td>0.5 μL</td>
<td>1.0 μL</td>
<td>1.5 μL</td>
<td>2.0 μL</td>
</tr>
<tr>
<td>Total Volume</td>
<td>30 μL</td>
<td>30 μL</td>
<td>30 μL</td>
<td>30 μL</td>
</tr>
</tbody>
</table>

\(^1\)10X EcoRI buffer is 500 mM NaCl, 900 mM tris·HCl, pH 7.5, 100 mM MgCl\(_2\).

[Promega Biotech Catalog]
Large Scale EcoRI Digests. The analytical reaction which required the least amount of EcoRI for complete digestion in a 1 hour incubation was scaled up to the volume required to linearize the desired amount of plasmid DNA. All of the reagents were mixed together in a plastic screw-top tube (Corning) except for the EcoRI. This mixture was equilibrated to 37°C, exactly one-half the number of calculated EcoRI units were added, and the solution mixed thoroughly by inversion. The digestion reaction was incubated at 37°C for 18-24 hours. The reaction was checked for completeness by 0.8% agarose gel electrophoresis. If the reaction was incomplete more EcoRI was added (1/4 as much as was added originally) and the incubation carried out for an additional 18-24 hours. The EcoRI was removed by phenol/chloroform extraction and the RI/plasmid precipitated with ethanol. The linearized plasmid DNA was pelleted by centrifugation at 10K rpm and 4°C for 15 min., washed twice with 70% ethanol, and dried by lyophilization. The dry pellet was dissolved in a small volume (3-8 mL) of TE buffer, pH 7.6, and quantitated by both UV absorbance [Maniatis, et al. (1982) p. 468] and 0.8% agarose gel electrophoresis [Maniatis, et al. (1982) p. 469].

3.7 Purification of T7 RNA Polymerase

The general purification procedure was given to me in protocol form by Jackie Wyatt of the Tinoco group. The method outlined here and in appendix A is a slightly modified version of her protocol. Plasmid containing the T7 RNA polymerase gene (pAR1219 in E. Coli BL21) was the gift of Jackie Wyatt. It is available from William Studier, Brookhaven National Laboratory, Upton, NY 11713.

A 4 L flask containing 1 L of M9 medium [Maniatis, et al. (1982) p. 440] containing 50 μg/mL ampicillin was equilibrated to 37°C and inoculated with a 25
mL overnight culture of pAR1219/E. Coli BL21 grown in the same medium. The 1 L
culture was grown at 37°C with vigorous shaking. After 2 hours, 0.238 g of
isopropylthiogalactoside (IPTG) was added, and the culture grown an additional 4
hours. The cells were harvested by centrifugation at 9,000 rpm and 4°C for 30
min. The supernatant was carefully decanted, the pellet washed with 50 mL
tris-sucrose buffer (50 mM tris·HCl, pH 8 at 4°C, 25% sucrose) and pelleted as
described above. The supernatant was decanted, the cells resuspended in tris-sucrose
buffer (10 g cell/100 mL buffer), and stored frozen at -20°C. When ready for
purification, 100 mL of the tris-sucrose slurry (10 g cells) was thawed at 4°C (12-
18 hours). The slurry was thoroughly mixed by inverting the tube several times.

170 mL buffer "L" (22 mM NH₄Cl, 20 mM tris·HCl, pH 8 at 4°C, 1 mM DTT, 3 mM
EDTA, 5% glycerol, 4 µg/mL phenylmethylsulfonyl fluoride, 0.4 µg/mL leupetine)
was added to the suspension and the slurry placed on a magnetic stir plate at 4°C. 30
mL buffer "L" containing 3 mg/mL lysozyme (Sigma) was added to the slurry and
the suspension stirred for 30 min. The cell membranes were disrupted by addition
of 0.15 g sodium deoxycholate dissolved in 1 mL buffer "L". After stirring an
additional 10 min. at 4°C, the suspension was pulled through an 18-20 gauge needle
and pushed through the needle into ten SS-34 centrifuge tubes. The tubes were spun
at 15K rpm and 4°C for 60 min. to pellet the cell membranes. The supernatant was
carefully transferred to a GS-3 centrifuge bottle and the nucleic acids and proteins
precipitated by addition of 80 g (NH₄)₂SO₄ with stirring over a 20 min. period at
4°C. The solution was stirred an additional 20 min. after the last addition, and spun
at 9K rpm and 4°C for 20 min. in the GS-3 rotor to pellet the white precipitate.
The supernatant was decanted and discarded. The pellet was resuspended in 10 mL
Buffer I (10 mM potassium phosphate, pH 8 at 4°C, 10 mM BME, 0.1 mM EDTA,
10% glycerol, 4 μg/mL PMSF, 0.4 μg/mL leupeptine) containing 0.15 M NaCl and dialyzed in 12-14K MWCO dialysis tubing against 2 L buffer I containing 0.15 M NaCl at 4°C for 18 hours. The suspension was removed from the dialysis tubing and centrifuged at 15K rpm and 4°C for 30 min. to remove any particulate matter. The clear supernatant was loaded onto a 1.5 X 30 inch column containing Whatman cellulose P-11 which had been previously prepared according to the manufacturers instructions and equilibrated with 1 L buffer I containing 0.15 M NaCl. The column was washed at 4°C with buffer I containing 0.15 M NaCl at a flow rate of 0.5 mL/min. until the A260 reached baseline to remove the nucleic acids. The proteins were eluted from the cellulose column at 4°C with a 200 mL linear gradient of 0.25 M to 0.6 M NaCl in buffer I at a flow rate of 0.5 mL/min. Those fractions with significant absorbance at 280 nm were checked for polymerase activity (20 μL transcription reaction with 2μL polymerase fraction), and those fractions which contained significant activity with no RNase contamination were pooled. The pure polymerase was dialyzed against 2 L of buffer II (100 mM NH₄Cl, 0.05 mM EDTA, 1 mM DTT, 10 mM tris-·HCl, pH 8 at 4°C) containing 10% glycerol for 4 hours at 4°C followed by 4L of buffer II containing 25% glycerol for 4 hours at 4°C. The polymerase was then dialyzed against 2 L of buffer II containing 50% glycerol for 18 hours at -20°C. This final dialysis concentrated the polymerase solution approximately two-fold. The polymerase was removed from the dialysis tubing and dispensed into 1 mL aliquots at -20°C, and stored at -20°C. The purified T7 RNA polymerase was not assayed for specific activity.
3.8 Large Scale Transcription Reactions

Analytical Transcription Reactions. To ensure success with the large scale transcriptions and to avoid wasting the T7 RNA polymerase, each batch of linearized plasmid and T7 RNA polymerase prepared were analyzed for optimal RNA yields. The analytical reactions were tested with the same batches of stock solutions that would be used for the large scale reactions. 20 μL transcription reactions (40 mM tris·HCl, pH 8.1 at 37°C, 6 mM MgCl₂, 5 mM NaCl₂, 10 mM DTT, 1 mM spermidine, 1 mM ATP, 1 mM CTP, 1 mM GTP, 1 mM UTP, 0.2 μg/μL RI/plasmid DNA), were prepared with varying amounts of T7 RNA polymerase (1-10 μL) and incubated for 2 hours at 37°C [Nielsen & Shapiro (1986); Promega Biotech Catalog]. The reactions were quenched by addition of 2 μL 60 mM EDTA and 10 μL 5X native gel loading buffer (30% glycerol, 0.25% xylene cyanol, 0.25% bromophenol blue). The reactions were heated to 90°C for 1 min. to denature the RNA, and immediately placed on ice. 5 μL from each reaction was loaded onto a 8x10x0.07 cm 10% polyacrylamide gel. Electrophoresis was carried out in 1X TBE buffer [Maniatis, et al. (1982) p. 156] at 300 volts for 20 min. The gel was stained with toluidine blue-O (Sigma) and destained with H₂O. The RNA yield was estimated by comparing the intensity of the stained product band with those of known quantities of yeast tRNA_Phe (Sigma) run on the same gel.

Large Scale Transcription Reactions, RNA1, 3/4-RNA2, RNA3, and RNA4. The analytical reaction with the greatest RNA polymerase yield was scaled up to produce the desired amount (usually 6-10 mg) of RNA. All of the reagents except for the T7 RNA polymerase were mixed in a 50 mL plastic screw-top tube (Corning) and equilibrated to 37°C. Once equilibrated, 5 μL/10 mL reaction mixture recombinant
RNasin (Promega) and 2 μL/10 mL inorganic pyrophosphatase (Boehringer Mannheim) were added and the solution mixed thoroughly by gently inverting the tube. The calculated quantity of T7 RNA polymerase was added, the solution mixed as above, and the reaction incubated at 37°C for 18 hours. The reaction was quenched by addition of EDTA to 10 mM, and the enzymes removed by extraction with 50:50 phenol chloroform [Maniatis, et al. (1982) p. 458]. After addition of NaCl to 0.1 M the nucleic acids were precipitated with 3 volumes of ethanol for 18 hours at -20°C. The precipitate was pelleted by centrifugation at 10K rpm and 4°C for 30 min. The pellets were washed twice with 70% ethanol and dried by lyophilization. The pellets were stored dry at -20°C.

Large Scale Transcription Reactions, RNA6. The methodology for transcription of RNA6 was identical to that for the other RNA species except that 40 mM potassium phosphate buffer, pH 8 at 37°C, was used instead of tris buffer. This buffer greatly attenuated the cleavage rate of RNA6 (for reasons to be discussed in chapter 8), and thus provided a higher yield of uncleaved RNA during an 18 hour transcription. If tris buffer is used, the transcription reaction should not proceed longer than 3-4 hours.

Large Scale Purification, RNA1, RNA3, and RNA4. The pellets from the large scale transcription were dissolved in a small volume of TE, pH 7.6 (the pellet from a 10 mL reaction usually dissolved in a volume of 1-2 mL). If the solution contained white particulate matter that did not dissolve it was centrifuged for 10 min. at 12K rpm in an eppendorf centrifuge and the clear supernatant transferred to a fresh tube. 100-200 μL 5X native gel loading buffer (30% glycerol, 0.25%
xylene cyanol, 0.25% bromophenol blue) was added and the solution and mixed by vortexing. This was loaded onto a 20x40x0.1 cm 12% preparative polyacrylamide native gel containing 5 wells, 100-200 μL per well. Electrophoresis was carried out at 15 W constant power (300V) for approximately 16 hours (or until the xylene cyanol dye marker had migrated half-way down the gel) in 1X TBE buffer [Maniatis, et al. (1982) p. 156]. The gel was transferred to a saran wrap-covered fluorescent screen and shadowed with short-wave UV light. The band which migrated with the xylene cyanol dye marker was excised, chopped into tiny pieces (ca. 1 mm on each edge) and transferred to a 50 mL plastic screw-top tube (Corning). The gel slices were just covered with gel elution buffer (0.1M tris-HCl, pH 7.4 at 37°C, 0.15 M NaCl, 15 mM EDTA, 1% SDS) and incubated at 37°C with gentle shaking for 24 hours. To determine if most of the RNA had soaked out of the slices, a few slices were removed, rinsed thoroughly with water, and UV shadowed as described above. If the slices cast a very dark shadow the incubation was continued an additional 24 hours. After incubation of suitable duration the gel elution buffer was carefully removed with a drawn out pasteur pipette (to avoid transferring flecks of acrylamide) and transferred to a fresh SS-34 centrifuge tube. The slices were washed with 1/2 volume of TE buffer, pH 7.4, which was carefully removed and pooled with the first fraction. The RNA was precipitated directly from the elution buffer by addition of three volumes of ethanol, followed by a 24 hour incubation at -20°C. The precipitate was pelleted by centrifugation at 10K rpm and 4°C for 30 min., and dried by lyophilization. The pellet was dissolved in 1 mL TE, pH 7.4 and extracted twice with 50:50 phenol/chloroform to remove residual acrylamide. The aqueous phase was transferred to a clean tube and the RNA precipitated with ethanol. The precipitate was pelleted as described above, washed
twice with 70% ethanol, and dried by lyophilization. The pure RNA was stored dry at -20°C. The RNA was exchanged into the appropriate buffer for NMR either by dialysis or Sephadex (Pharmacia) column chromatography as described in the text for each RNA species.

**Large Scale Purification 3/4-RNA2 and RNA6.** The methodology for purification of 3/4-RNA2 and RNA6 was identical as that described for RNA1, RNA3, and RNA4 except for the electrophoretic conditions. 3/4-RNA2 was loaded onto a 20x40x0.1 cm 10% polyacrylamide gel containing 7 M urea. Electrophoresis was carried out at 25 W constant power (2000 V) against 1X TBE buffer until the bromophenol blue dye marker had just migrated off the bottom of the gel (approx. 6 hours). Uncleaved RNA6 was separated from its cleavage products on a 20x40x0.1 cm 8% polyacrylamide gel containing 7 M urea. In this case electrophoresis was carried out at 25 W constant power against 1X TBE buffer until the xylene cyanol dye marker had migrated half-way down the gel (approx. 4 hours).

3.9 Demonstration of Cleavage at the Appropriate Site

The NMR studies presented in this thesis are not of naturally occurring hammerhead domains, but of domains whose sequences have been manipulated to facilitate the NMR assignment procedure. Therefore it is necessary to demonstrate that they cleave at the appropriate site. The cleavage of RNA2 and the cleavage kinetics of RNA6 are demonstrated in chapters 5 and 8 respectively. The cleavage of RNA1, RNA3, and RNA4, and the slow cleavage of RNA6 will be demonstrated in this section.
Purification of $^{32}$P-Body-Labelled RNA. $^{32}$P-body-labelled RNA was prepared by *in vitro* transcription with $\alpha^{-^{32}}$P-CTP. 4 μL of 5X transcription buffer (200 mM Tris·HCl, pH 8.1 at 37°C, 30 mM MgCl$_2$, 50 mM DTT, 25 mM NaCl, 5 mM spermidine) was combined with 1μg Rl/pRNA DNA, 8 μL of 2.5 mM each ATP, GTP, UTP, 2.5 μL 100 μM CTP, and enough H$_2$O to bring the volume to 18 μL. 0.5 μL $\alpha^{-^{32}}$P-CTP (Amersham 400 Ci/mmol) was added to the mixture and the contents of the tube thoroughly mixed by tapping the tube. Transcription was initiated by addition of 2 μL T7 RNA polymerase (activity unknown). The reaction was incubated for 2 hours at 37°C. After incubation the reaction was quenched by addition of 2 μL quench buffer (60 mM EDTA, 3.0 M sodium acetate pH 5.2), and the nucleic acids precipitated with ethanol. The precipitate was pelleted, washed twice with 70% ethanol, and the pellet dried. The pellet was resuspended in 20 μL TE, pH 7.6, 5 μL 5X native gel loading dye added (30% glycerol, 0.25% xylene cyanol, 0.25% bromophenol blue), and loaded immediately onto a 8x10x0.07 cm 10% polyacrylamide gel, 5 μL per well. Electrophoresis was carried out at 200 Volts for 1 hour against 1X TBE buffer. The gel was removed from the plates, wrapped in plastic film, and autoradiographed for 30 min. at room temperature. The appropriate bands were excised, crushed, and covered with 200 μL gel elution buffer (0.1 M Tris·HCl, pH 7.4 at 37°C, 0.15 M NaCl, 15 mM EDTA, 1% SDS) in an eppendorf microfuge tube. The covered gel slices were incubated at 37°C for 24 hours. After incubation the buffer was carefully removed with a drawn out pasteur pipette and transferred to a fresh tube. The slices were washed with 100 μL H$_2$O and the wash pooled with the first fraction. The RNA was precipitated with 900 μL ethanol at -70°C for 20 min. The precipitate was pelleted, washed twice with 70% ethanol, and dried.
**Demonstration of Cleavage.** The pellets of pure RNA were dissolved in 40 μL TE, pH 7.6. 10 μL of 5X denaturing gel loading dye (80% formamide, 0.25% xylene cyanol, 0.25% bromophenol blue) was added and the RNA denatured by heating to 90°C for 1 min. 5 μL was immediately loaded onto a 20x40x0.04 cm 20% polyacrylamide gel containing 7M urea. Electrophoresis was carried out at 25 W constant power for 4 hours against 1X TBE buffer. The gel was removed from the plates, wrapped in plastic film and autoradiographed for 24 hours at room temperature. A lane of 5'-32P-end labelled deoxythymidine (dT) homopolymers of specific lengths was run on the gel to help determine the size of the cleavage fragments.

**Results.** The autoradiogram of the 20% polyacrylamide denaturing gel is shown in figure 3.2. The sizes of the dT homopolymers are denoted on the figure. Lanes 2, 3, and 4 clearly contain two bands. The uppermost band is the large fragment, and the faint band which comigrates with a dT 13-mer is the small cleavage fragment. The small fragment, a 12-mer in each case, comigrates with the dT 13-mer due to sequence specific effects. The faintness of the band is due to the presence of only four C's in the sequence. The fact that the uppermost bands are those of the large fragment is determined by comparison with lane 5. This lane contains uncleaved RNA6, which clearly migrates more slowly than do the large fragments of cleaved RNA1, RNA3, and RNA4. Lane 5 also contains a very faint band which comigrates with the major bands of lanes 2, 3, and 4. This is due to the large fragment of cleaved RNA6, as a small percentage of RNA6 cleaves during a 2 hour transcription (chapter 8). The corresponding 12-mer band is too faint to see. Although the large fragments of RNA3 and RNA4 are identical in length, that of RNA4 migrates faster.
Figure 3.2 20% denaturing PAGE of RNA1, RNA3, RNA4, and RNA6. The experimental conditions are described in the text. The length of the major bands of the deoxythymidine homopolymers is denoted on the left hand side of the autoradiograph. Lanes: 1. oligo dT; 2. RNA1; 3. RNA3; 4. RNA4; 5. RNA6. LF= large fragment; SF= small fragment.
than that of RNA3. This is thought to be due to slight differences in secondary structure of the two fragments. These minor sequence specific anomalies aside, the fact that RNA1, RNA3, and RNA4 separate into two bands of the appropriate lengths demonstrates cleavage at the proper site. Similarly, these result clearly indicate that RNA6 is predominantly uncleaved after a 2 hour incubation in transcription buffer.
Chapter 4

NMR Investigation of RNA1
4.1 Introduction

Once thought of as merely the dumb slave of DNA, RNA has finally emerged in the biochemistry community as a distinct class of molecule. Not only a carrier of genetic material, but capable of catalysis as well, RNA appears to fill the intermediary gap between DNA and proteins. Certainly the first clue that RNA is a molecule to be reckoned with appeared when Rich and Kim solved the three-dimensional structure of yeast tRNA$^{\text{Phe}}$ by x-ray crystallography [Quigley & Rich (1976); Kim (1975, 1976)]. But it was not until the discovery that RNase P, the enzyme which processes the 5'-end of precursor tRNA, requires both an RNA and a protein component for its in vivo activity, that RNA stepped out from under the shadow of DNA and emerged as a distinct species [Kole, et al. (1980)]. The later discovery that the active site of RNase P lies completely within the RNA moiety [Guerrier-Takada, et al. (1983); Guerrier-Takada & Altman (1984)], coupled with the observation that a ribosomal precursor RNA in Tetrahymena pseudomonas processes in the absence of proteins [Cech, et al. (1981)], really drove the message home. RNA is certainly not an "extra" copy of genetic material present only to take the strain of translation off the almighty DNA. RNA represents a new class of molecule; a class exemplified by vast structural diversity and the ability to catalyze reactions on both itself and other nucleic acids.

The smallest and perhaps simplest catalytic RNAs are those associated with certain plant viruses, viroids, and virusoids. A site-specific self-cleavage reaction was first observed for the (+) and (-) strands of the satellite RNA associated with tobacco ringspot virus (STobRV RNA) which resulted in 5'-hydroxyl and 2',3'-cyclic phosphate termini [Prody, et al. (1986); Buzayan, et al. (1986a, 1986b)]. This RNA is believed to replicate via a rolling circle method, with the self-cleavage
of multimeric repeats into monomer units completing the replication cycle. It was found that when multimers were transcribed \textit{in vitro}, efficient cleavage occurred only if the RNA was incubated with magnesium ions [Prody, et al. (1986)]. Although incubation with divalent manganese resulted in some cleavage to monomer (as did incubation with other divalent metal ions), the most efficient cleavage was achieved with magnesium ions. Studies with deletion mutations showed that fewer than 100 nucleotides were necessary for efficient cleavage of (+)STobRV RNA [Buzayan, et al. (1986a)]. Similar cleavage reactions were observed for the (+) and (-) strands of avocado sunblotch viroid (ASBV) [Hutchins, et al. (1986)] and lucerne transient streak virusoid (LTSV) [Forster & Symons (1987b)]. Symons and coworkers suggested a base pairing scheme for these small cleavage domains called a hammerhead [Forster & Symons (1987a)]. Although this secondary structure cannot account for the unusual lability of the phosphodiester bond, it did correctly predict the cleavage site in a transcript of newt DNA [Epstein & Gall (1987)]. The hammerhead domain consists of thirteen unpaired nucleotides and three variable length double-stranded helices. Cleavage occurs only at a specific phosphodiester linkage and requires no cofactor other than magnesium ions. Although the nucleotides at and near the cleavage site may vary, the position of cleavage within the hammerhead domain is conserved. Because the hammerhead requires only divalent ions for cleavage, it is clear that to fully understand the cleavage reaction we must first understand the three dimensional structure of the hammerhead domain. The nature of the interactions giving rise to this folding are as yet unknown.

The next four chapters describe the design of several hammerhead domains, originally derived from the minimal length of STobRV RNA which was known to
cleave efficiently, and the structural information which was obtained via one-dimensional nuclear Overhauser (1D NOE) spectroscopy. It is in no way a complete solution of the three dimensional structure, but rather serves as a strong foundation for continued NMR investigation of this domain. In an effort to retain continuity in the discussion of experimental motivation and results, most standard laboratory procedures used for sample preparation have been confined to chapter 3. However, the methods used to handle the NMR sample (i.e. change buffers) are discussed in the text.

4.2 The Basic Strategy

The central task to begin this project was to design an RNA sequence which demonstrated the ability to cleave at the proper site and which manifested an imino proton spectrum well enough resolved to be completely, unambiguously assigned. The basic strategy to achieve this goal followed the course of reasoning that was taken for the NMR studies of tRNA; the imino proton spectrum would be assigned in the absence of magnesium ions with the rationale that these resonances would reflect secondary base pairs of the hammerhead structure. Magnesium ions would then be added to the sample, with the hope that new imino proton resonances would appear, reflecting the introduction of tertiary structure. These new resonances could then be more easily recognized and assigned, having already assigned those due to the double helical stems of the structure. Built into this strategy is the assumption that the cleavage of a single phosphodiester bond does not perturb the structure greatly from its uncleaved form. This was a necessary starting assumption because it turns out that the small cleavage domain is a much more efficient cleaver than is the full 359-nucleotide STobRV RNA sequence [Buzayan et al. (1986)]. Whereas the full-
length sequence cleaved slowly enough that the uncleaved form could be separated by polyacrylamide gel electrophoresis from the cleavage products, the highly efficient cleavage domain (min STobRV RNA) cleaved too quickly in the presence of magnesium ions for a similar separation. Unfortunately the *in vitro* transcription reaction used to produce the NMR samples requires magnesium ions, allowing the RNA to cleave completely before it could be purified. This assumption did not seem unreasonable as an initial starting point, because interfering with the cleavage reaction represents a "chicken-and-egg" argument anyway; because the three-dimensional structure appears to be the driving force for the cleavage reaction, anything which hinders the cleavage must be disrupting the native structure. Therefore the investigation of an already cleaved form is more reasonable than that of a similar domain which does not cleave. If cleavage does not result in a subsequent rearrangement of the native structure, then this assumption is a good system for study of the cleavage mechanism. A domain which does not possess even the propensity to cleave may provide no information about the mechanism. With this circular argument aside, then, an effort was made to produce a biologically active, unambiguously assignable self-cleaving RNA domain.

The path taken to achieve this goal was somewhat meandering. The system was redesigned several times to reflect new information in the literature, to overcome NMR assignment problems, and to try to create a system in which no assumptions needed to be made. However, this seemingly haphazard approach always had the major objective in mind, and the motivations for the changes made will be clearly discussed. For this reason I am forced to present my experiments chronologically, as the relevance of each system is lost if presented any other way.
4.3 RNA 1

Luckily, the Kim group here at Berkeley was interested in studying this cleavage domain by x-ray crystallography, and Rosie Kim constructed the plasmid DNA template necessary to produce RNA1. The sequences of min STobRV RNA and RNA1 are shown in figure 4.1. Stem III of RNA1 is elongated compared to that of min STobRV RNA to ensure that the cleavage products remained base paired. The 3'-overhang of RNA1 is due to the EcoRI restriction site used to linearize the plasmid DNA for run-off transcription.

4.4 NMR Spectroscopy of RNA1

Materials and Methods. Approximately 3 mg of pure, lyophilized RNA1 was dissolved in 2.5 mL of column buffer (1.4 mM phosphate buffer, pH 7.5, 14 μM EDTA). This was loaded onto a Pharmacia PD-10 column which was previously equilibrated with 25 mL of the same buffer. The sample was eluted with column buffer in a volume of 3.5 mL according to the enclosed Pharmacia instructions, and dried under vacuum (Savant speed-vac). The RNA was stored dry at -20°C. Just prior to NMR study, RNA1 was dissolved in 0.25 mL 90% H2O/10% D2O and transferred to a Wilmad 628--PP microcell sample tube using a PipetteMan with a piece of capillary tubing attached to the tip. The final sample conditions were ca. 0.5 mM RNA1, 20 mM phosphate buffer, pH 7.5, and 0.2 mM EDTA. The RNA1 sample was stored in the sample tube at 4°C between NMR experiments. Due to degradation, several RNA1 samples were prepared. All of the samples were handled as described above.

A temperature series was recorded on a GN500 spectrometer operating at a proton frequency of 500 MHz using the 1331 experiment (appendix B). 4000
Figure 4.1 Sequences of min STobRV RNA and RNA1 shown in the proposed hammerhead secondary folding pattern [Hutchins et al. (1986)]. The conserved nucleotides are boxed. The arrow denotes the cleavage site. RNA1 has stem III elongated compared to that of min STobRV RNA to ensure that the cleavage products remain base paired. The 3'-UUUAG overhang of RNA1 is due to the EcoR1 restriction site used to linearize the plasmid DNA prior to T7 RNA polymerase run-off transcription.
scans of 8K data points were averaged for each temperature increment. The temperature was increased in ten degree units between 20 and 60°C, allowing the sample to equilibrate for 5 min. at each temperature. The temperature was then lowered to 55°C, followed by ten degree increments to a final temperature of 25°C, again allowing the sample to equilibrate at the desired temperature for 5 minutes. The temperature data were taken in this manner to ensure that RNA1 renatured properly after unfolding. The FID's were transferred to a MicroVax workstation and manipulated using FTNMR software [Hare Research]. The raw data were left-shifted one point, added to the original FID (a trick which obliterates the H2O peak from the spectrum), and appodized using a 30° phase-shifted sinebell with a skew of 0.5. Fourier transformation resulted in the frequency domain spectrum. Each spectrum was indirectly referenced through the chemical shift of the water resonance. Additional temperature series were recorded in this same manner after addition of 10 mM MgCl2 and 20 mM MgCl2 respectively to the sample. The exact experimental parameters are given in the figure legends.

NOE difference spectra were recorded at several temperatures and buffer conditions on a GN500 spectrometer using the 1331NO experiment (appendix B). For each dataset the actual sample conditions and experimental parameters are given in the figure legends. The difference spectra were either appodized and Fourier transformed on the GN500 according to the scheme illustrated in figure 4.2, or on a MicroVax computer using the FTNMR software [Hare Research] macro listed in table 4.1. The spectra are indirectly referenced to the chemical shift of water at the appropriate temperature.
Table 4.1 FTNMR 1D NOE Processing Macros

<table>
<thead>
<tr>
<th>Reference Spectrum</th>
<th>Difference Spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. re &amp;file</td>
<td>1. re &amp;file</td>
</tr>
<tr>
<td>2. bc</td>
<td>2. bc</td>
</tr>
<tr>
<td>3. zrr X Y</td>
<td>3. zrr X Y</td>
</tr>
<tr>
<td>4. zi</td>
<td>4. zi</td>
</tr>
<tr>
<td>5. xbf 2</td>
<td>5. xbf 2</td>
</tr>
<tr>
<td>6. cl</td>
<td>6. cl</td>
</tr>
<tr>
<td>7. re &amp;file</td>
<td>7. re &amp;file</td>
</tr>
<tr>
<td>8. shl X</td>
<td>8. bc</td>
</tr>
<tr>
<td>9. zi</td>
<td>9. shl X</td>
</tr>
<tr>
<td>10. exc</td>
<td>10. zi</td>
</tr>
<tr>
<td>11. abf 2</td>
<td>11. exc</td>
</tr>
<tr>
<td>12. lbf 2</td>
<td>12. abf 2</td>
</tr>
<tr>
<td>13. si X</td>
<td>13. lbf 2</td>
</tr>
<tr>
<td>14. xbf 1</td>
<td>14. si X</td>
</tr>
<tr>
<td>15. lbf 1</td>
<td>15. xbf 1</td>
</tr>
<tr>
<td>16. shl 1</td>
<td>16. lbf 1</td>
</tr>
<tr>
<td>17. sbf 1</td>
<td>17. shl 1</td>
</tr>
<tr>
<td>18. lbf 1</td>
<td>18. sbf 1</td>
</tr>
<tr>
<td>19. sb X 30</td>
<td>19. lbf 1</td>
</tr>
<tr>
<td>20. ft</td>
<td>20. lb 2.5</td>
</tr>
<tr>
<td>21. rev</td>
<td>21. em</td>
</tr>
<tr>
<td>22. ph</td>
<td>22. ft</td>
</tr>
<tr>
<td>23. clr</td>
<td>23. rev</td>
</tr>
<tr>
<td>24. dr</td>
<td>24. ph</td>
</tr>
<tr>
<td>25. cl</td>
<td>25. wr filename.1</td>
</tr>
<tr>
<td>26. end</td>
<td>26. cl</td>
</tr>
<tr>
<td></td>
<td>27. re &amp;file</td>
</tr>
<tr>
<td></td>
<td>28. bc</td>
</tr>
<tr>
<td></td>
<td>29. exc</td>
</tr>
<tr>
<td></td>
<td>30. zrr X Y</td>
</tr>
<tr>
<td></td>
<td>31. zi</td>
</tr>
<tr>
<td></td>
<td>32. xbf 2</td>
</tr>
<tr>
<td></td>
<td>33. cl</td>
</tr>
<tr>
<td></td>
<td>34. re &amp;file</td>
</tr>
<tr>
<td></td>
<td>35. bc</td>
</tr>
<tr>
<td></td>
<td>36. exc</td>
</tr>
<tr>
<td></td>
<td>37. shl X</td>
</tr>
<tr>
<td></td>
<td>38. zi</td>
</tr>
<tr>
<td></td>
<td>39. exc</td>
</tr>
<tr>
<td></td>
<td>40. abf 2</td>
</tr>
<tr>
<td></td>
<td>41. lbf 2</td>
</tr>
<tr>
<td></td>
<td>42. si X</td>
</tr>
<tr>
<td></td>
<td>43. xbf 1</td>
</tr>
<tr>
<td></td>
<td>44. lbf 1</td>
</tr>
<tr>
<td></td>
<td>45. shl 1</td>
</tr>
<tr>
<td></td>
<td>46. sbf 1</td>
</tr>
<tr>
<td></td>
<td>47. lbf 1</td>
</tr>
<tr>
<td></td>
<td>48. lb 2.5</td>
</tr>
<tr>
<td></td>
<td>49. em</td>
</tr>
<tr>
<td></td>
<td>50. ft</td>
</tr>
<tr>
<td></td>
<td>51. rev</td>
</tr>
<tr>
<td></td>
<td>52. ph</td>
</tr>
<tr>
<td></td>
<td>53. wr filename.2</td>
</tr>
<tr>
<td></td>
<td>54. cl</td>
</tr>
<tr>
<td></td>
<td>55. re filename.1</td>
</tr>
<tr>
<td></td>
<td>56. xbf 5</td>
</tr>
<tr>
<td></td>
<td>57. cl</td>
</tr>
<tr>
<td></td>
<td>58. re filename.2</td>
</tr>
<tr>
<td></td>
<td>59. sbf 5</td>
</tr>
<tr>
<td></td>
<td>60. lbf 5</td>
</tr>
<tr>
<td></td>
<td>61. clr</td>
</tr>
<tr>
<td></td>
<td>62. dr</td>
</tr>
<tr>
<td></td>
<td>63. end</td>
</tr>
</tbody>
</table>

X and Y depend on the size of the dataset. For 8K data points X = 4096 and Y = 8192.
Figure 4.2  GN500 scheme for processing 1D NOE difference spectra
**NOE Difference Spectra, No MgCl₂.** Figure 4.3 shows the sequence and imino proton spectrum of RNA1 at 30°C with the unambiguous imino proton assignments labelled. The spectrum is labelled in a manner which will allow direct comparison with subsequent RNA species. The secondary folding model contains 21 base pairs, and if there is no end-fraying of the helices, this secondary structure alone should give rise to 21 imino proton resonances. However, studies with DNA oligomers show that terminal base pairs of helices do fray, and the imino protons of these pairs are not observed except at low temperatures. It is difficult to guess, based on the sequence alone, which pairs, if any, will exhibit fraying in RNA1, but it is likely that GC₁ and AU₁₁ exhibit this behavior, as GC₁ terminates stem III, and AU₁₁ may be similar to a terminal base pair following a cleavage event. AU₁₃ of stem I may not be observed for reasons similar to AU₁₁, but because nothing is known about the tertiary structure of RNA1 little more can be said about these pairs. Similarly, the junction GC of stem II may not manifest an imino proton resonance, but because this pair is near the conserved region, its status cannot be predicted. The base pairs closing the hairpin loops of stems I and II most likely manifest imino proton resonances at 30°C, as it has been shown that RNA hairpins of 4 and 7 bases are fairly stable when closed by GC base pairs. [Martin et al. (1972); Uhlenbeck et al. (1973); and Bloomfield et al. (1974)]. Therefore an imino proton spectrum which contains between 17-21 resonances would be consistent with the hammerhead folding pattern.

It was stated in section 2.7 that because each Watson-Crick base pair manifests a single imino proton resonance it is a straight forward task to count the number of base pairs in a sequence by counting the number of resonances in the imino proton spectrum. Although straight forward in theory, it is not actually quite
Figure 4.3 Sequence and 500 MHz imino proton spectrum of RNA 1 in 20 mM phosphate buffer, pH 7.5 and 0.2 mM EDTA taken at 20°C. The unambiguous assignments are labelled on the sequence.
so easy in practice. First, solvent suppression results in uneven excitation of the imino protons [Hore, P. J. (1983)]. In figure 4.3 it is clear that resonances N, O, and P have integrated intensities which are greater than those observed for the other resonances. Because the excitation maximum is adjusted for this region of the spectrum it is difficult to determine if the apparent intensities are due to the number of protons which give rise to the resonances, or if this is an artifact of the excitation profile. In addition, resonances C and G' are each due to a single proton (data presented later will show this), but they certainly do not have equal integrated intensity, even though their linewidths are similar. This is due to uneven excitation of these resonances. Second, the imino proton spectrum is very crowded between 12.5-12.0 ppm, making the task of counting individual resonances even more difficult. Therefore, in addition to the 17 imino proton resonances labelled on the figure, there is a strong possibility that there are additional resonances between 12.5 and 12.0 ppm which have not been labelled. However, there is nothing in the spectrum of fig. 4.3 to suggest that it is inconsistent with the hammerhead folding model.

To determine if this folding pattern is a correct model for RNA1 it is necessary to assign the base pairs of the helical stems to particular resonances in the spectrum. As previously discussed in section 2.8 this is accomplished by NOE difference spectroscopy. NOE difference spectra of RNA1 are plotted in figures 4.4 - 4.10. The exact experimental conditions are given in the figure legends. Figures 4.4 and 4.5 establish the connectivity of peaks M-A-I at 25° and 30°C respectively. Clearly irradiation of resonance A produces NOEs to peaks I and M at both temperatures. This resonance also produces a sharp NOE to the aromatic region of the spectrum, identifying it as an AU base pair. Irradiation of resonance I produces
Figure 4.4  NOE difference spectra of ca. 1 mM RNA1 in 20 mM phosphate buffer, pH 7.6, and 0.2 mM EDTA at 25°C. 2048 scans of 8K data points were averaged for each spectrum. The reference spectrum was appodized using a 30° phase-shifted sinebell. Difference spectra were appodized using an exponential multiplication with a linebroadening of 2.5. The shoulders on peaks A, G, and H are due to sample degradation. Peaks due to spill-over irradiation are labelled with an "X".
Figure 4.5  NOE difference spectra of ca. 0.5 mM RNA1 in 20 mM phosphate buffer, pH 7.03, and 0.2 mM EDTA at 30°C. 4000 scans of 8K data points were averaged for each spectrum. The spectra were appodized as described for fig. 4.4. The large aromatic resonance in some of the 30°C spectra an artifact (labelled with an X). Some of the data do not contain this artifact, as all of the spectra were not acquired in a single experiment.
a very weak NOE to A at 30°C and no others (irradiation at 25°C produces no NOE's). This means that I is either a terminal base pair, or is the nearest neighbor of a base pair which is too broad to show an NOE. Should the latter be the case, it is not unreasonable to expect irradiation of this broad neighbor to give an NOE at I, because irradiation of A produces a relatively strong NOE to I. Examination of all the NOEs at both 25° and 30°C shows this to not be the case. Only irradiation of A results in an NOE to I. This, then, identifies I as a base pair terminating a helix. In addition, the absence of a sharp aromatic NOE identifies I as a GC base pair.

Irradiation of M produces a weak NOE at F/G at 30°C (this NOE is very weak at 25°C), but none at A. Based on the very weak NOE at M upon irradiation of peak I, the fact that M lacks an NOE to A is not unreasonable, especially if M is broader, and hence more difficult to saturate, than I. Unfortunately spectral crowding in the region of M makes it difficult to say much about its linewidth. This crowding also makes it difficult to saturate M cleanly, and the NOE to F/G is most likely due to spill-over irradiation of L, M', or M" (similar to that observed for saturation of I at 25°C). Unfortunately saturating other resonances in the vicinity of M does not produce this same NOE, nor does saturation of F/G produce an NOE back to M (fig. 4.8). This leaves the possible M-F/G connectivity ambiguous.

M may also be producing NOEs to its nearly degenerate neighbors which cannot be resolved. This is highly probable, as classification of M (in addition to I) as a terminal GC would require that RNA1 contain a three base pair stem of sequence G/C-A/U-G/C, which it does not. It is also clear that M must show an NOE to a GC in addition to its connectivity to resonance A, because A/U-G/C-A/U-G/C occurs only once in the RNA1 sequence, and assignment of M-A-I to this sequence requires that either I or M manifest an NOE to the sharp resonance J/K, which they clearly do not.
This evidence requires that M-A-I define the sequence G/C(?)-G/C(M)-A/U(A)-G/C(I), with the assignment of each pair given in parentheses. Because the sense of the helix cannot be ascertained from NOE difference spectroscopy, this sequence can be situated in RNA1 with four possible arrangements; defining the four base pairs of stem II and three arrangements within stem III. The arrangement in stem III requiring I or M to NOE to J/K has already been discussed. The other two possible stem III arrangements require that the unknown peak (?) show an NOE to M and to another GC base pair which is in the center of a stable helix. Because M is in a very crowded region, one could argue that this is indeed the case, and that all of the central GC base pairs in stem III resonate in this crowded region. However, as will be demonstrated below, a different connectivity pathway can be clearly traced for stem III, defining the M-A-I connectivity as being that of stem II. Melting data will be presented which corroborate these assignments.

Figures 4.6 and 4.7 show the NOE difference spectra which establish the connectivity of peaks O-C-G-H-J/K at 25°C and 30°C respectively. The spectrum corresponding to the saturation of resonance O at 30°C is missing in figure 4.7 (the original datafile became corrupt, but its NOE difference spectrum looked similar to that taken at 25°C when originally plotted). Saturation of resonance O at 25°C clearly results in an NOE to C, saturation of C produces an NOE back to O/P and also to G, and saturation of G results in an NOE back to C and what appears to be an NOE to J/K. However, the J/K NOE has an intensity which is less than that of C. This is unlikely if the G-J/K connectivity represents nearest neighbor interactions, as resonances C and J/K are equally sharp, and should therefore produce NOE's of equal intensity. The J/K NOE most likely represents a second-order next-nearest neighbor interaction. Saturation of J/K produces an NOE to H, and a second order
Figure 4.6 NOE difference spectra which establish the connectivity of peaks O-C-H-G-J/K of RNA1 at 25°C. The experimental conditions are given in the legend of fig. 4.4.
Figure 4.7 NOE difference spectra which establish the connectivity of peaks C-H-G-J/K of RNA1 at 30°C. The experimental parameters are given in the legend of fig. 4.5. The spectrum corresponding to saturation of resonance O is missing, since the original datafile became corrupt.
NOE to G. Therefore it can be established that what appears to be a spill-over shoulder to H upon saturating resonance G is really an NOE. Saturation of resonance H further confirms this assignment, as it produces first order NOEs to G and J/K, and the expected second order NOE to C (very weak). It is clear that saturation of H produces an NOE to G and not to G', as the intensity of G is greater than that of G'. Since G' is closer to H in chemical shift, one would expect it to be of greater intensity than G if the presence of G and G' are due to spill-over irradiation.

Examination of the aromatic region of the spectrum reveals that resonance C arises from an AU base pair, H and G are GC base pairs, and O/P and J/K are each two GC base pairs. Although it is difficult to determine the exact number of protons which contribute to a particular resonance line when using 1331 solvent suppression as discussed above, the fact that saturating O/P and J/K each results in two amino proton NOE's proves the degeneracy in these resonances. The fact that it is O which gives an NOE to C and not P can be determined from figure 4.9. The difference spectrum which corresponds to saturation of P produces an NOE to C as does that for saturation of O (fig. 4.7). However, the fact that this NOE has a greater intensity upon saturation of O than it does for P proves the O-C connectivity. The NOE observed in the difference spectrum corresponding to the saturation of P is due to spill-over irradiation. Unfortunately, the J/K degeneracy cannot be similarly broken, as it produces only two first-order NOE's. It is apparent in figures 4.9 and 4.10 that saturation of the resonances in the upfield-most region of the spectrum produce NOE's to J/K, but whether the NOE is that of J or K, or even due to spill-over cannot be determined. The only connectivity pathway which can be made with confidence is thus G/C(O)-A/U(C)-G/C(G)-G/C(H)-G/C(J/K), where the assignments are given in parentheses. This connectivity is clearly that of stem III.
because the sequence G/C-A/U-G/C-G/C-G/C is unique to that stem.

Having determined the connectivity to be that of stem III, it is necessary to obtain the orientation of this sequence within the helix. This sequence can be positioned in stem III with two orientations; resonance O being that of GC₃ or GC₁₀. Unfortunately, the NOE pathway terminates with peaks J/K. Because of the degeneracy of J/K it is possible that this resonance produces an NOE to itself. If this is the case, irradiation of J/K should produce an NOE to H, an NOE to itself (which cannot be resolved), and an NOE to another peak. Regardless of the orientation of the sequences within stem III this additional NOE must be to an AU base pair. The only other AU base pair observed in the NOE data is F (fig. 4.8). There is some evidence that irradiation of J/K produces a weak NOE to F (fig. 4.6), but because F is nearly degenerate with G (which is a next-nearest neighbor of J/K) it is impossible to unambiguously sort out the data. Irradiation of F does produce an NOE to J/K (fig. 4.8), but again this connectivity cannot be proven because of its near degeneracy with G. In addition, spill-over to H (from saturation of G) is clearly evident in figure 4.7, as is G's NOE to C, further challenging the J/K-F connectivity. The orientation of the stem III assignments in figure 4.3 were obtained from the melting data. If resonance O is GC₁₀, it should melt at a fairly low temperature, as it is near the end of a helix terminated by an AU base pair. The melting data will show that in fact O is stable to 60°C.

Unfortunately no unambiguous stem I assignments were made, due mostly to the crowded upfield region of the imino proton spectrum (remember that GC base pairs generally resonate between 13-12 ppm, and AU pairs resonate between 15-13 ppm). As can be seen in figure 4.9, irradiation of G' at 25°C results in very weak NOE's to C, J/K, and N. The C and J/K NOE's can certainly be explained by
Figure 4.8 (a) Difference spectrum obtained upon irradiation of resonance F at 25°C. The sharp aromatic NOE suggests that F is an AU base pair. The experimental parameters are given in the legend of fig. 4.4. (b) Difference spectrum obtained upon irradiation of resonance F at 30°C. The sharp aromatic NOE suggests that F is an AU base pair. The experimental parameters are given in the legend of figure 4.5.
Figure 4.9 (a) The rest of the NOE difference spectra recorded at 25°C. Experimental parameters are given in the legend of fig. 4.4.
Figure 4.10 The rest of the NOE difference spectra recorded at 30°C. The experimental parameters are given in the legend of fig. 4.5.
irradiation spill-over, but this leaves a clear NOE to N. Irradiation of this same peak at 30°C (fig. 4.10) also results in a weak NOE to an upfield peak, but it cannot be determined if it is to the same peak as in the 25°C spectrum, because the peaks in this crowded region of the spectrum may be rearranging as the temperature increases. These peaks may also be at slightly different chemical shifts in the two spectra because the data were obtained from different RNA preparations, and although they were handled similarly, the sample conditions are not identical. However, irradiation of G' at 30°C probably also give an NOE at N. Because irradiation of N at both 25°C and 30°C produces an NOE to J/K but not to G', this connectivity cannot be proven. In addition, saturation of J/K at either temperature does not result in an NOE to N, eliminating the possibility of connecting resonance N to other resonances. The only other possible connection that can be made from the NOE data is of peaks M' and P, as irradiation of both of these resonances results in an NOE to the other. However, spectral overlap makes it difficult to determine if in fact this connectivity is real.

The very broad resonance at ca. 14 ppm was saturated at 10°C (labelled A' in fig. 4.3), but the S/N was too poor to determine the nature of this base pair (data not shown). Based on its chemical shift it is most likely an AU base pair, and its broad linewidth suggests that it terminates a helix or is from a relatively unstable tertiary pair. There are three unassigned AU's in RNA1; the terminal AU of stem I, and the two AU's of stem II. Unfortunately, without any NOE connectivities this resonance cannot be assigned. It is also clear that there are too few AU resonances to account for these three base pairs. It is reasonable that the helices are end-fraying, and these AU pairs are exchanging too quickly with water protons give visible resonances.
The broad resonances between 12.5 and 10 ppm were also saturated at 10°C, but again the S/N was so poor that no information was obtained (data not shown). The imino protons of thymidine in DNA hairpin loops have been shown to resonate in this chemical shift region [Hare & Reid (1986)]. The rationale is that interactions in the loop protect the imino protons somewhat from exchange with solvent H2O so that resonances from these non hydrogen-bonded imino protons can be seen. They resonate upfield of imino protons involved in base pairs because there is no hydrogen bond interaction to further deshield them. These broad resonances in the RNA1 spectrum are most likely due to similarly looped U bases in the hairpin loops of stems I and II, or in the single stranded regions comprised of the conserved bases.

2D NOESY Spectroscopy, No MgCl2. A two dimensional NOESY spectrum in H2O was collected at 30°C to help corroborate the assignments obtained from 1D difference spectroscopy, and to try and establish connectivities within the crowded 12.5-12.0 ppm region of the spectrum. In addition, the 2D spectrum was used to make a more careful count of the exact number of resonances in the imino proton spectrum, as it is sometimes useful to take advantage of the increased spectral resolution of a 2D experiment to help resolve questions about degeneracy. Water supresson was achieved by substituting two 11 pulses for the hard 90° pulses of the normal NOESY pulse sequence. A 11 echo read pulse was used to prevent baseline distortions [Bax, A. (1988)]. The carrier frequency was adjusted to the resonant frequency of water. An offset delay of 155 µsec was used to set the excitation maximum to the center of the imino proton resonances. 256 scans of 2K data points using a recycle delay of 0.8 sec. were averaged for each of 512 T1 increments. The mixing time was 100 msec. The imino-aromatic region of the H2O NOESY of RNA1
is shown in figure 4.11. Figure 4.12 shows the imino-imino region of the same spectrum. Exact experimental conditions are given in the figure legends. In the contour plot of figure 4.11 the horizontal axis is the chemical shift of the imino protons, the vertical axis contains the chemical shift of the aromatic and amino protons. As discussed in section 2.7, the base pair type is determined by the aromatic NOE's. A Watson-Crick (W-C) AU pair shows a sharp NOE between the imino proton of uracil and the aromatic CH2 of the adenine, and a very broad NOE to the amino protons of the adenine. A W-C GC pair shows NOEs between the imino proton of guanosine and the amino protons of cytosine (broad) and guanosine (very broad). Therefore one can expect the imino-aromatic region of the H2O NOESY to contain a very sharp (and possibly a broad) crosspeak for each AU pair, and two broad crosspeaks for each GC pair. Even though the NOESY spectrum has low digital resolution (12 Hz/point) and S/N, the resolution is greatly improved compared to that of the one dimensional spectrum. The sharp aromatic cross peak of A is greatly attenuated due to the uneven excitation profile of the experiment. The imino-aromatic crosspeaks of C and F are very intense, but the weak imino-amino crosspeaks of these basepairs are so broad that they cannot be readily distinguished from the baseline. The two imino-amino crosspeaks of G, G', and H are very intense, whereas those of I are very weak. J/K has three intense imino-amino crosspeaks, verifying the presence of two peaks. The crosspeak at (12.7,7.2) ppm appears to be twice the intensity of the those at (12.7,8.7) and (12.7,8.9) ppm, suggesting that it is due to degenerate amino protons. This, then explains the absence of a fourth crosspeak for the peaks J/K. There are fourteen crosspeaks in the vertical slice between 12.5 and 12.3 ppm, verifying that there are 7 base pairs in this crowded region as labelled. Therefore the NOESY has determined that the
Figure 4.11 Imino-aromatic region of RNA1 H$_2$O NOESY at 30°C. Sample conditions were ca 0.8 mM RNA, 20 mM phosphate, pH 7.03, and 0.2 mM EDTA. Solvent suppression was achieved with a 1-1 pulse scheme [Bax, A. (1988)]. An offset delay of 155 µsec was used to set the excitation maximum to the center of the imino proton spectrum. This delay was too long, as resonance A is greatly attenuated. 256 scans of 2K data points using a recycle delay of 0.8 sec for each of 512 T$_1$ increments were collected. The mixing time was 100 msec., and the spectral width was 11363.63 Hz in both dimensions. The spectrum is referenced indirectly through the chemical shift of H$_2$O. The imino proton chemical shifts are labelled according to figure 4.3 across to top of the plot.
Figure 4.12 Imino-imino region of RNA1 H₂O NOESY. Experimental parameters and conditions are given in the legend of fig. 4.11. The diagonal peaks are labelled according to the convention used in fig. 4.3. The C-G-H-J/K connectivity is drawn on the plot.
number of base pairs were correctly counted in the one dimensional imino proton spectrum. This may not seem to be important, but to get a clear picture of the secondary folding of RNA1 (and to fully assign the imino proton spectrum) it is necessary to know exactly how many protons each resonance represents. This simple resonance identification has also demonstrated the power of 2D spectroscopy for increasing spectral resolution.

The imino-imino region of the H$_2$O NOESY is shown in figure 4.12. Just as fig. 4.11 demonstrated one of the benefits of 2D spectroscopy, figure 4.12 demonstrates one of the drawbacks; poor signal to noise. The sample for this experiment was ca. 0.8 mM in RNA, and even though this concentration is greater than used that for the 30°C 1D NOE data, the 2D NOESY has only marginal S/N. In addition, the poor digital resolution makes it difficult to determine if there are any NOE's close to the diagonal between peaks L, M', M", M, N, O, and P. To determine if crosspeaks were real or noise, the criterion that they must be present on both sides of the diagonal was used, as each half of the spectrum should contain identical information. This criterion proved sufficiently stringent that only three connectivities could be made with certainty. These connectivities are indicated in fig. 4.12. Clearly the use of 2D spectroscopy for labile protons is limited by the amount of sample which can be produced.

*Temperature Data, No MgCl$_2$. As mentioned earlier, melting data was used to corroborate the orientation of the stem III assignments, and the stem II assignments. The observation of base pair imino proton requires that the lifetime of the base pair be long compared to the inverse of the chemical shift difference between the hydrogen bonded form and the solvent (section 1.8). As the temperature is
increased, base pairs begin to open allowing the imino proton to exchange more efficiently with $H_2O$. This is observed in the NMR spectrum as a broadening, and eventual disappearance, of the imino proton resonances. Although the imino proton exchange rate is not limited by base pair opening [Leroy, J-L. et al. (1985)], the order of onset of resonance broadening generally does follow the order of melting. Therefore the opening of individual base pairs can be monitored as a function of temperature as each helical stem melts. Figure 4.13 shows the imino proton spectrum of RNA1 plotted as a function of temperature. Based solely on helical length, stem II is expected to melt before stems I and III. Because stems I and III both contain stretches of GC base pairs, and exchange broadening can be seen at internal pairs, the order in which these two stems melt cannot be predicted.

As the temperature is decreased from $20^\circ$ to $0^\circ C$, three things begin to happen: (i) resonance A' and the resonances in the 11.2-10.0 ppm region of the spectrum increase in intensity, (ii) the entire spectrum suffers a general decrease in resolution, and (iii) resonances A, I, and L broaden severely. The increase in intensity of A' can be easily rationalized, as decreasing the temperature decreases the rate at which imino protons exchange with water protons, thus increasing their resonance intensity. It has been speculated that the resonances between 11.2 and 10.0 ppm are due to imino protons in hairpin loop or single stranded regions of the RNA which are somewhat protected from solvent exchange by base stacking interactions. The fact that these resonances increase in intensity with decreasing temperature can be rationalized with same argument used for resonance A'. The overall loss of resolution may be due to the increase of the rotational correlation time of the RNA in solution at low temperatures, molecular aggregation, or a combination of both. However, resonances A, I, and L clearly suffer greater
broadening than the overall spectrum as the temperature decreases. In fact, at 0°C resonance I has completely disappeared. The behavior of resonance I is interesting, as it is a junction GC base pair. Its behavior is contrary to that expected for exchange broadening, and therefore suggests that the junction becomes conformationally stressed with decreasing temperature. The fact that A broadens less severely than I substantiates this hypothesis because A is the AU neighbor of I, and stress in the junction should affect I more than A. This line of reasoning suggests that L is also near, but not directly at, the junction of the three helices. Possible assignments are GC₁₀ of stem III or GC₁₄ of stem I. Unfortunately there is no NOE evidence to support these plausible assignments.

Upon increasing the temperature from 20°C to 30°C, resonance A' completely melts, resonances A, I, and L begin to broaden, and the 13-12 ppm region becomes more resolved as resonance O moves out from under P, and M'' moves out from under N. At 40°C, resonances A, F, and I have broadened severely, peak L has completely melted, and a peak appears to be moving out from resonance O. Because of the crowding in the 13.5-12 ppm region, it is difficult to determine the fate of resonance M (it is between M' and M'' in 40°C spectrum), but it appears as though this resonance has also broadened. At 45°C it is clear that A, I, and M have completely disappeared, which corroborates their assignment to stem II. In addition, resonance F has completely disappeared and G' has slightly broadened.

Increasing the temperature to 60°C results in the broadening of resonances G', M', M'', N, and P. Only resonances C, G, H, J/K, M', and O remain sharp at this temperature. This melting behavior corroborates the orientation of the stem III assignments. If O were GC₁₀ instead of GC₃, the expected melting temperature of resonance O would be lower than 60°C. Only resonance M' is not assigned to stem III.
Figure 4.13  500 MHz imino proton spectrum of RNA plotted as a function of temperature. The sample was 0.5 mM RNA1, 20 mM phosphate, pH 7.5, and 0.2 mM EDTA. 4000 scans of 8K datapoints were averaged for each temperature increment using a recycle delay of 1.0 sec. The temperature of each spectrum is labelled on the right margin in degrees Celsius. The figure continues on the next page.
It has been speculated that this resonance is a GC base pair within stem I. Based on the sequence of this stem, this assignment is certainly possible.

**Discussion and Conclusions, No MgCl_2.** The 500 MHz imino proton spectrum and unambiguous assignments obtained by NOE difference spectroscopy are summarized in figure 4.3. These assignments were corroborated by the behavior of RNA1 as a function of temperature. The stem of RNA1 was elongated compared to that of min STobRV RNA to ensure that the cleavage products remained base paired. This design strategy certainly served its intended purpose, as the resonances of the central five base pairs of this elongated stem remain sharp at 60°C. It was estimated that 17-21 imino proton resonances would be consistent with the proposed hammerhead folding pattern. The presence of 17 resonances was confirmed by 2D NOESY spectroscopy. Although few unambiguous assignments were obtained, there is no evidence to suggest that the imino proton spectrum is inconsistent with the folding model. There is slight evidence which suggests that the junction of RNA1 becomes strained as the temperature decreases, but a complete assignment of these resonances must be obtained to say anything further about the stability of the junction base pairs. Clearly a sequence which manifests a more resolved imino proton spectrum was necessary to completely understand the secondary structure of this class of self-cleaving RNA's.

**NOE Difference Spectroscopy, 10 mM MgCl_2.** The goal of this project was to identify tertiary interactions which are necessary for a cleavage event to occur. Although a complete assignment of the secondary structure is crucial to fully understand this domain, overwhelming curiosity about the fate of the spectrum
prompted addition of MgCl$_2$ (which is required for cleavage) to the sample. The 500 MHz imino proton spectra of RNA1 taken at 30°C and several MgCl$_2$ concentrations are shown in figure 4.14. In 10 mM MgCl$_2$, resonance G' has doubled in intensity compared to the 0 mM MgCl$_2$ spectrum. Either G' and G have become degenerate, or F has sharpened and moved under G'. Peaks J and K are no longer degenerate in chemical shift, and the resolution between 12.5 and 12.0 ppm has also increased. There is no further gain of spectral resolution upon increasing the MgCl$_2$ concentration to 20 mM. Although addition of MgCl$_2$ did not result in blatantly obvious additional resonances, as was hoped, the increased spectral resolution prompted the collection of more NOE difference spectra, as further secondary base pair assignments were desired.

Figures 4.15 and 4.16 show the NOE difference spectra of RNA1 in 10 mM MgCl$_2$ at a temperature of 30°C. Spectra which corroborate assignments already obtained will not be discussed. The spectra corresponding to saturation of resonances A, C, and H are missing because the data files became corrupt, however, their respective difference spectra looked very similar to those taken in the absence of MgCl$_2$. Saturation of resonance P results in a spill-over NOE to C (from O), and an NOE to M'. Saturation of M' produces the expected NOE back to P and a weak NOE to J/K. Although the M'-P connectivity was suggested from difference spectroscopy in the absence of MgCl$_2$, this data displays this connectivity more clearly, as the 12.5-12.0 ppm region of the spectrum is better resolved. Due the increased spectral resolution of peaks J and K, these peaks were saturated independently of one another. Irradiation of J results in NOE's to peaks H and G. Saturation of K produces these same NOE's, however the NOE intensity is reduced compared to those obtained upon saturation of J. This implies that it is J, and not K, which is
Figure 4.14 Comparison of the 500 MHz imino proton spectra of RNA1 at 30°C in (A) 0 mM MgCl₂, (B) 10 mM MgCl₂, and (C) 20 mM MgCl₂. The spectral parameters for each spectrum are given in the figure legends of figures 4.13, 4.17, and 4.18 respectively.
Figure 4.15 NOE difference spectra of 0.5 mM RNA1 in 20 mM phosphate, pH 7.0, and 10 mM MgCl$_2$ at 30°C. Spectra corresponding to the saturation of peaks A, C, and H are missing because the original data files became corrupt.
Figure 4.16 More NOE's of RNA1 in 10 mM MgCl₂ at 30°C. Sample conditions are given in the legend of fig. 4.15.
connected to G and H, and hence it is J which assigns to stem III. In addition, saturation of K produces a weak NOE to N, which was not observed in the absence of MgCl₂. It is unlikely that the weak J/K NOE from irradiation of M' is due to spill-over irradiation from N because the difference spectrum which corresponds to the saturation of resonances M" and M lack this NOE, and these resonances are closer to N in chemical shift than is M'. If the NOE from M' is to J and not K, then either resonance M' or P would have to be an AU base pair because it is J which assigns to stem III. Both the chemical shift of these resonances and the aromatic region of their difference spectra refute this possibility. M' must therefore be connected to K. Peaks M', P, and N can all be identified as GC base pairs, and must therefore assign to stem I. Unfortunately the orientation of the P-M'-K-N assignment cannot be ascertained. Even if there were NOE data connecting peaks out to the terminal AU pairs, the orientation would still be unknown due to the dyadic symmetry axis of the sequence. One other additional NOE can be observed in the 10 mM MgCl₂ data. Saturation of peak G' produces spill-over NOE's to C and H, and an NOE to M". This connectivity is difficult to rationalize. If the P-M'-K-N connectivity is that of stem I, this G'-M" connectivity must be that of stem III, because under these circumstances the sequence G/C-G/C doesn't occur anywhere else! If this is the case, the orientation of the sequence must be such that G' is GC₁, because no O-G' NOE is observed. It is, however, possible that O manifests an NOE to M" which is not resolved due to spectral overlap. As plausible as this arguments appears, I do not believe that this assignment is correct. If G' is the terminal GC base pair of stem I it should be broader, due to end effects, than resonance O, which is an interior GC pair. Clearly it is not. A more plausible explanation is that resonances M and M' have rearranged upon the addition of MgCl₂. This means that G' may be the GC pair
which closes the hairpin loop of stem II. There is evidence to support this hypothesis. Saturation of M and not $M''$ produces an NOE to $G'$. Unfortunately this cannot be explicitly resolved, as neither M nor M' has an NOE to A.

Figure 4.15 also corroborates the fact that there is an AU resonance under peak G, as saturation of the left side of G produces a sharp aromatic NOE. The NOE to K may be due to spill-over irradiation of G, as G is connected to J. Because these data suggest this NOE is really K, and this assignment is plausible, it is labelled as K. However, there is no further evidence to substantiate this assignment, as saturation of K will produce an NOE to G from spill-over to J which will mask the possible K-F connectivity.

**Temperature Data, 10 and 20 mM MgCl$_2$.** Figure 4.17 shows the temperature dependence of RNA1 in 10mM MgCl$_2$. It is interesting that resonances A, I, and L sharpen before they begin to melt out as the temperature is increased from 25 to 80°C. The broadness of these resonances at 25°C corresponds to the low temperature behavior that was observed in the absence of MgCl$_2$. Perhaps the native structure is more open in this junction region than the secondary folding pattern predicts, and addition of MgCl$_2$ is driving the conformation of the domain toward its true cleavage conformation by stabilizing as yet unidentified tertiary interactions in the native structure.

Again, peaks C, G, H, J/K, M', and O remain at 60°C, corroborating the stem III assignments already obtained. In fact G, H, J/K, and O are still fairly sharp at 80°C. The addition of 10 mM MgCl$_2$ has raised the melting temperature ($T_m$) of each helix by approximately 5°C, as is evident by comparing figs. 4.6 and 4.7. The increased helical stability is due to increasing the ionic strength of the sample with a divalent
Figure 4.17 500 MHz imino proton spectrum of RNA1 as a function of temperature. The sample was ca. 0.5 mM RNA1, 20 mM phosphate buffer, pH 7.0, 10 mM MgCl₂, and 0.2 mM EDTA. 1500 scans of 8K data points were averaged for each spectrum. The temperature at which each spectrum was acquired is denoted in the right margin in degrees Celsius. The figure continues on the next page.
Figure 4.18 500 MHz imino proton spectrum of RNA1 as a function of temperature. The sample was ca. 0.5 mM RNA1, 20 mM phosphate buffer, pH 7.0, 20 mM MgCl$_2$, and 0.2 mM EDTA. 1500 scans of 8K data points were averaged for each spectrum. The temperature at which each spectrum was acquired is denoted in the right margin in degrees Celsius. The figure continues on the next page.
Figure 4.19 Sequence and 500 MHz imino proton spectrum of RNA1 taken at 30°C. The unambiguous assignments are denoted in upper case letter, possible assignments are denoted in lower case letters. The conserved bases are boxed. The arrow denotes the cleavage site.
cation, as the stability of the helix generally increases with increasing polyvalent ion concentration.

Although addition of MgCl₂ did change the imino proton spectrum, it is difficult to determine, if additional resonances appeared because of spectral overlap. The MgCl₂ concentration was increased to 20 mM with the hope that this higher concentration would result in even better spectral resolution. These spectra are plotted in figure 4.18 as a function of temperature. It is clear that this higher magnesium concentration had the opposite effect; the spectral resolution has actually decreased. This is most likely due to aggregation of the sample, as all of the resonances appear to be equally affected. The order in which the resonances melt is the same as that observed for the 10 mM MgCl₂ sample, as are the Tₘ's of the individual helices.

_Discussion and Conclusions, 10 and 20 mM MgCl₂._ Addition of 10 mM MgCl₂ to the RNA1 sample increased the spectral resolution between 12 and 13 ppm. However, further increasing the MgCl₂ concentration to 20 mM decreased the resolution of the entire imino proton spectrum, which implies that RNA1 aggregates at this MgCl₂ concentration. The melting data at both magnesium ion concentrations corroborate the previously obtained RNA1 assignments. NOE difference spectra recorded in 10 mM MgCl₂ broke the degeneracy of the J/K assignment, and introduced some other possible connectivities which could not be unambiguously resolved. The final RNA1 assignments are shown in figure 4.19.

It was hoped that addition of magnesium would result in extra imino proton resonances. It can be said with confidence that no extra resonances appeared in the 14.5-13 ppm region of the spectrum, but relatively little can be said about the 13-
12 ppm region. This region is just crowded enough at 500 MHz to make counting the number of resonances difficult. Additional resonances could therefore go undetected, especially if they were somewhat broad.

Although some ground was broken toward the elucidation of RNA1's secondary structure, a species which manifested a more well-resolved imino proton spectrum was necessary to completely assign these secondary resonances, and thus to begin identifying those resonances which may reflect tertiary contacts in the molecule.
Chapter 5

Investigation of RNA2
5.1 Introduction

It was clear from the spectra of RNA1 that an RNA which manifested a better resolved imino proton spectrum was necessary to yield a complete assignment. At this point, one of two approaches could have been taken to pursue this goal; the sequence of RNA1 could have been altered, keeping the overall design of the molecule the same, or the molecule could have been drastically redesigned, taking advantage of the "ribozyme" technology of Cech and Uhlenbeck [Zaug & Cech (1986); Uhlenbeck (1987)]. In these studies, Cech and Uhlenbeck demonstrated that RNA's which catalyze intramolecular cleavage reactions could catalyze the cleavage of a different RNA strand (intermolecular) when the two strands could associate to form the proper cleavage structure. In addition, once the "substrate" strand was cleaved, the "enzyme" strand dissociated and associated with another substrate strand, causing it to cleave too. Because this behavior is similar to the turnover associated with protein enzymes, the strands which remained unchanged during the reactions were dubbed "ribozymes." Since there were some clear advantages to a bimolecular construct for the NMR studies, this was the approach taken to redesign the hammerhead domain.

5.2 Motivation for the RNA2 Construct

Figure 5.1 shows the sequence of RNA2. As is marked on the figure, the "3/4" strand was enzymatically synthesized from a plasmid DNA template using T7 RNA polymerase run-off transcription, whereas the "1/4" strand was chemically synthesized. One of the problems with the RNA1 construct was that it completely cleaved during the course of transcription, and hence was already cleaved before purification. Because the assumption had to be made that cleavage does not perturb
Figure 5.1 The sequence of RNA2. The "3/4" strand was enzymatically synthesized from pRNA2, while the "1/4" strand was chemically synthesized. The 3'UUAAG of the 3/4 strand is an artifact of the EcoRI restriction site necessary to linearize the plasmid DNA for run-off transcription. The conserved bases are boxed. The arrow denotes the cleavage site of the fully ribo construct.
the structure, a system in which this assumption need not be made was desired. This 
3/4:1/4 construct overcomes this obstacle. The cleavage products, a 5'-OH and a 
2',3'-cyclic phosphate, imply that the 2'-OH is a requisite for cleavage. Since the 
1/4 strand is chemically synthesized, a deoxycytidine can easily be incorporated at 
the cleavage site, unobtrusively preventing the species from undergoing cleavage. 
This RNA with a single deoxy "point mutation" can then be studied in the presence of 
magnesium without undergoing cleavage, and presumably without disrupting the 
true cleavage conformation. As a control, a fully ribo 1/4 strand was synthesized 
and annealed with the 3/4 strand in the presence of magnesium ions, demonstrating 
the ability of this construct to cleave.

5.3 Analytical Cleavage reactions of Ribo-C and Deoxy-C RNA2

Materials and Methods. 100 µg of pure 3/4 RNA2 was dialyzed in a volume of 1 
mL against 2L of 10 µM tris·HCl, pH 7.5, and 0.1 mM EDTA at 4°C for 18 hours. 
The RNA solution was then dispensed into eppendorf tubes, 32 µL (3.2 µg) per tube, 
and dried under vacuum. 0.2 µg of pure ribo-1/4 (rC12-mer) RNA and 0.2 µg of 
deoxy-C 1/4 RNA (dC12-mer) were 5'-32P end-labelled with 1U T4 
polynucleotide kinase (BRL) in 10 µL of 50 mM tris·HCl, pH 7.5, 10 mM MgCl2, 5 
mM DTT, and 1 mM spermidine (the synthetic oligomers were the gift of Dr. Patrick 
Chou). After incubation for 1 hour at 37°C the reactions were quenched with 10 µL 
30 mM EDTA. Each reaction was then transferred to a tube containing a 3.2 µg 
 aliquot of dry 3/4 RNA2. The 3/4 and 1/4 strands were annealed by heating the 
solution to 60°C and allowing it to cool slowly. After annealing, the T4 
polynucleotide kinase was removed by phenol/chloroform extraction, and the 
nucleic acids precipitated with ethanol at -70°C. The pellets were washed twice
with 70% ethanol, dried, and resuspended in 20 μL TE, pH 7.6. The rC-RNA2 was dispensed into ten eppendorf tubes, 2 μL per tube. The dC-RNA2 was similarly aliquoted. To each aliquot was added 2 μL of 10 mM X, where X=EDTA (control), MgCl₂, MnCl₂, Spermidine, ZnSO₄, CdCl₂, and Pb(OAc)₂. The tubes were incubated for 18 hours at 37°C, and the cleavage reactions quenched after this time by addition of 4 μL "quench buffer" (10 mM EDTA, 80% formamide, 0.02% xylene cyanol, 0.02% bromophenol blue). The RNA was denatured by heating for 2 min. at 90°C. 6 μL from each aliquot was loaded onto a 10x12x0.07 cm 20% polyacrylamide gel containing 7 M urea. The gel was run at 600V (11 mA) for 30 min. using 1X TBE [Maniatis et al. (1982) p. 454] as the reservoir buffer. The gel was visualized by autoradiography (30 min. exposure).

This experiment was repeated for rC12-mer with a cleavage incubation time of 1 hour.

Results and Discussion. The assay for cleavage of RNA2 is quite simple. If no cleavage occurs, a band corresponding to a nucleotide of length 12 should be the only band visible on the gel. If cleavage occurs, and the reaction has gone to completion, the only band should be that of a 7-mer, as is evident from the position of the cleavage site diagramed in figure 5.1. In cleavage reactions which have not gone to completion both bands will be visible, with the intensity of the 12-mer band being less than that of the control lane. Figure 5.2 shows the autoradiogram of the 18 hour cleavage reactions, figure 5.3 the 1 hour reactions.

Magnesium (II), manganese (II), and zinc (II) clearly catalyze cleavage of rC-RNA2. While Zn²⁺ does catalyze cleavage, the rate of the reaction is slower than that
Figure 5.2 18 hour analytical cleavage reactions of rC-RNA2 and dC-RNA2. The experimental conditions are described in the text. Lanes: 1. rC/EDTA; 2. dC/EDTA; 3. rC/MgCl₂; 4. dC/MgCl₂; 5. rC/MnCl₂; 6. dC/MnCl₂; 7. rC/spermidine; 8. dC/spermidine; 9. rC/ZnSO₄; 10. dC/ZnSO₄; 11. rC/CdCl₂; 12. dC/CdCl₂; 13. rC/Pb(OAc)₂; 14. dC/Pb(OAc)₂.
Figure 5.3 1 hour analytical cleavage reactions of rC-RNA2. The experimental conditions are described in the text. Lanes: 1. EDTA; 2. MgCl₂; 3. MnCl₂; 4. spermidine; 5. ZnSO₄; 6. CdCl₂; 7. Pb(OAc)₂.
catalyzed by either Mg$^{2+}$ or Mn$^{2+}$, as is evident from the presence of both 12-mer and 7-mer bands for the 18 hour incubation and the lack of cleavage for the 1 hour incubation. Spermidine, Cd$^{2+}$, and Pb$^{2+}$ did not catalyze cleavage of rC-RNA2, although Pb$^{2+}$ did catalyze the degradation of both rC12-mer and dC12-mer. dC-RNA2 did not cleave under in any conditions. Both Mg$^{2+}$ and Mn$^{2+}$ catalyzed complete cleavage of the RNA during the 1 hour incubation. Again, Pb$^{2+}$ appears to have catalyzed RNA degradation.

**Conclusions.** It was necessary to prove that rC-RNA2 possessed the ability to undergo magnesium catalyzed cleavage to avoid questions about whether this species approximates a true cleavage conformation. This was demonstrated, as was the fact that divalent manganese and zinc also catalyze this cleavage reaction, although Zn$^{2+}$ catalyzed catalysis is slow compared to that of either Mg$^{2+}$ or Mn$^{2+}$. However, these results differ from those obtained for the biologically occurring 359-nucleotide STobRV RNA cleavage domain, in that spermidine alone did not catalyze cleavage even after 18 hours of incubation. In studies of full-length STobRV RNA, 23% of the molecules had undergone cleavage after a 15 min. incubation in 1 mM spermidine at 30°C [Prody et al. (1987)]. This spermidine induced cleavage may in fact be due to insufficient removal of Mg$^{2+}$ from the purified STobRV RNA. If the Mg$^{2+}$ binds tightly to the RNA, it may be difficult to remove. It is easier to control the environment of RNA2, as divalent ions are scavenged before creation of the cleavage domain, and hence the metal binding site(s). There is evidence which supports this hypothesis. In the StobRV RNA studies, 6-8% cleavage was observed after incubation in 2 mM EDTA, whereas no cleavage of ribo-RNA2 was observed after an 18 hour incubation in 5 mM EDTA.
It was hoped that divalent cadmium would also catalyze cleavage, as $^{113}\text{Cd}$ is a spin $\frac{1}{2}$ nucleus, and could be used as an NMR nucleus to probe the metal binding site of the RNA. This did not prove to be the case, as even an 18 hour incubation of rC-RNA2 with divalent cadmium ions resulted in no cleavage of the RNA. Divalent lead was also tried, as tRNAs are known to undergo lead catalyzed cleavage, but the only apparent effect lead had on the RNA was to catalyze its degradation. This result again differs from that obtained with STobRV RNA, as divalent lead was reported to catalyze cleavage [Prody et al. (1987)].

The fact that dC-RNA2 did not cleave in any conditions demonstrated that cleavage does in fact result from attack of the 2'-hydroxyl of this residue on the 5'-phosphate of its 3'-neighbor. As mentioned earlier, this was thought to be the case, as the cleavage reaction yields 2',3'-cyclic phosphate and 5'-hydroxyl termini.

The net result of these studies was that dC-RNA2 could be studied under cleavage conditions without undergoing cleavage (and thus introducing possible structural rearrangements), with confidence that this system approximates the true cleavage conformation. It should be noted that this type of bimolecular construct provided a means by which the kinetics of the cleavage reaction could be studied, and in fact the Uhlenbeck lab in Colorado is currently probing this topic with this type of system.

5.4 Spectroscopy of Deoxy-C RNA2

Materials and Methods. 3/4-RNA2. Approximately 1 mg pure lyophilized 3/4 RNA2 was dissolved in 250 μL sample buffer (20 mM phosphate buffer, pH 7.2, 0.1 M NaCl, 0.2 mM EDTA, and 0.02% (w/v) NaN3 in 90%H2O/10%D2O) and transferred to a Wilmad 628-PP microcell sample tube. The conc. of RNA2 was ca. 0.2 mM.
**dC-RNA2.** 340 μL of 1 mg/mL dC12-mer RNA was added to the 3/4 RNA2 NMR sample. The two RNA strands were annealed by heating the sample to 60°C and allowing it to cool slowly to room temperature over a period of 3-4 hours. dC-RNA2 was then purified from any excess single strands by preparative native polyacrylamide gel electrophoresis (chapter 3). Just prior to spectroscopy, approximately 3 mg of pure, lyophilized deoxy-C RNA2 was dissolved in 200 μL sample buffer (20 mM phosphate buffer, pH 7.2, 0.1 M NaCl, 0.2 mM EDTA, and 0.02% (w/v) NaN₃ in 90%H₂O/10%D₂O) and transferred to a Wilmad 428-PP NMR sample tube.

NMR spectra were collected on a GN 500 spectrometer using a 1331 pulse [Hore, P. J. (1983)] with the carrier frequency set at the resonance of water. For each spectrum the exact experimental parameters are given in the figure legends.

**Results, 3/4-RNA2.** The 500 MHz imino proton spectrum of the 3/4 fragment of RNA2 is shown in figure 5.4. (It should be noted that the floppy disk containing the RNA2 data became corrupt and the original data could not be recovered. All of the figures were salvaged from "working-copy" spectra which were plotted at the time the data was collected.) Clearly from the sequence this fragment should have only four base pairs (those of stem II). This is indeed the case, as the spectrum contains only four imino proton resonances. Because the sequence of stem II has a diadic symmetry axis, the resonances could not be assigned by NOE difference spectroscopy. However, provided these resonances do not change drastically in chemical shift upon formation of the bimolecular complex, this spectrum can be compared with that of dC-RNA2 to help identify the resonances due to these four base pairs in the spectrum of dC-RNA2. This is not an unreasonable assumption, as stem
II can be considered a sub-domain of the hammerhead structure, and there is certainly precedence for studying subdomains.

*Results, dC-RNA2.* As can be seen from comparison of figures 4.1 and 5.1, the sequence of RNA2 differs dramatically from RNA1, as does its size. Stem II has been shortened to 6 base pairs, stem I to 5 base pairs, and stem I does not contain a hairpin loop. The sequence of stem I is determined indirectly by T7 RNA polymerase, as the enzyme requires certain sequences from the +1 position of the promoter sequence to transcribe efficiently, and the sequence 5'-GGGAG is one of the most efficient transcription start sequences [Milligan et al. (1987)]. The sequence of stem II was not changed in length compared to that of RNA1, but the internal GC base pair was changed to an AU base pair. The reason for this change was two-fold. First, this GC resonance in RNA1 was in a very crowded region of the imino proton spectrum, and changing the base pair to an AU base pair would remove an upfield resonance and introduce a resonance in the less crowded region between 15-13 ppm. Second, the introduction of an A/U-A/U connectivity, although not unique to this stem, should be relatively easy to assign since there is only one other A/U-A/U connectivity in the entire sequence. An A/U-A/U connectivity was also added to stem III in addition to reducing its length. This connectivity was added for the same reasons as discussed above. To prevent degeneracy of the AU resonances with those of stem II, the bases of stem III were stacked 5'-A-U and 5'-A-A in stem II. In retrospect, the sequence of stem II should have been designed differently, because the sense of the helix cannot be determined from the NOE connectivity pattern, as discussed above.
Figure 5.4 500 MHz imino proton spectrum of 3/4 RNA2 at 25°C. The sample was ca. 0.2 mM RNA, 20 mM phosphate, pH 7.2, 0.1 M NaCl, 0.2 mM EDTA, 0.02% NaN₃. 8000 scans of 8K data points using a recycle delay of 1 sec. were averaged. The spectrum is referenced indirectly through the chemical shift of water.
Based on the end-fraying that is observed for short DNA oligomers in solution, and the lack of NMR evidence for the existence of the AU base pair terminating stem III in RNA1, it was expected that the spectrum of dC-RNA2 contain between 11 and 12 resonances, with the GC base pairs on the ends of stems I and III and the AU base pair terminating stem III melted out. Although all four resonances of stem II were observed in the spectrum of 3/4 RNA2, status of this resonance cannot readily be determined because of the evidence from RNA1 that this base pair melts out as the RNA approaches a true cleavage conformation.

Figure 5.5 shows the 500 MHz imino proton spectrum of dC-RNA2 compared to that of 3/4 RNA2. The overlap of the spectra is reasonable, and it is fairly easy to identify the resonances of stem II in the spectrum of dC-RNA2. Although the spectrum is not quite properly phased, there appear to be between 11 and 12 resonance, a number which is consistent with the predicted secondary folding model. Unfortunately the sample concentration was too low for difference spectroscopy, and additional sample had to be prepared.

Additional dC-RNA2 was prepared by mixing stoichiometric amounts of 3/4 RNA2 and dC12-mer together in a volume of 1 mL 10 mM phosphate, pH 7.0, 0.1 mM EDTA. The sample was annealed by raising the temperature to 80°C for 1 min. and allowing the solution to return to room temperature over a 3-4 hour period. The sample was dried by lyophilization and redissolved in 200 μL 90% H2O/10% D2O and transferred to a Wilmad 428-pp NMR sample tube, yielding a 0.5 mM dC-RNA2, 20 mM phosphate, pH 7.0, and 0.2 mM EDTA NMR sample. The 500 MHz imino proton spectrum of this sample at 25°C is shown in figure 5.6. Clearly this spectrum differs greatly from the spectrum of 0.2 mM dC-RNA2 (fig. 5.5).
Figure 5.5 500 MHz imino proton spectra of dC-RNA2 and 3/4-RNA2 at 25°C. The sample was ca. 0.2 mM RNA, 20 mM phosphate, pH 7.2, 0.1 M NaCl, 0.2 mM EDTA, 0.02% NaN₃. 8000 scans of 4K data points were averaged using a recycle delay of 1 sec. The spectra are indirectly referenced through the chemical shift of water.
Figure 5.6 500 MHz imino proton spectrum of ca. 0.5 mM dC-RNA2, 20 mM phosphate, pH 7.0, 0.1 M NaCl, 0.2 mM EDTA, 0.2% NaN₃ at 25°C. 4000 scans of 4K data points were averaged using a recycle delay of 1 sec. The spectra are indirectly referenced through the chemical shift of water.
most obvious difference is the fact that the spectrum contains too many peaks to be a single dC-RNA2 structure. A less obvious difference is the very large intensity of the central peak in the 0.5 mM spectrum. Comparison of this spectrum with that of 3/4 RNA2 implies that the 0.5 mM sample contains an excess of 3/4 RNA2. Because this sample was not purified by gel electrophoresis this certainly seemed plausible, although it did not explain the extra resonances in the NMR spectrum.

The NMR sample was purified by preparative polyacrylamide gel electrophoresis under non-denaturing conditions. The purified product was verified to be the proper species by analytical polyacrylamide gel electrophoresis under denaturing conditions (fig. 5.7). The 500 MHz imino proton spectrum of the newly purified dC-RNA2 sample at a concentration of ca. 0.4 mM is shown in figure 5.8. Although this spectrum does not appear to be contaminated with excess 3/4 RNA2, there are still more peaks than were originally observed for dC-RNA2! A possible explanation is that the additional resonances are from increased sample concentration, as this sample is twice as concentrated as the original sample. To test this hypothesis, the dC-RNA2 sample was diluted to ca. 0.1 mM and another spectrum acquired (fig. 5.8). This spectrum is not much different than that of 0.4 mM dC-RNA2, and they are both similar to that of fig. 5.3. The difference between the spectra of fig. 5.8 and that of fig. 5.3 was never resolved. One hypothesis is that the buffer conditions of the two samples were not identical, and the latter spectra represent an equilibrium of structures in solution.
Figure 5.7  (A) 12% polyacrylamide native gel of gel-purified dC-RNA2 NMR sample. Lanes: 1. dC-RNA2 NMR sample; 2. pure 3/4-RNA2. Each lane contained 2 μg of material. (B) 20% polyacrylamide/7M urea denaturing gel of gel-purified dC-RNA2 NMR sample. Lanes: 1. pure dC-12mer; 2. pure 3/4-RNA2; 3. dC-RNA2 NMR sample. Each lane contained 4 μg of material.
Figure 5.8 (A) 500 MHz imino proton spectrum of gel-purified dC-RNA2 at 25°C. The sample was ca. 0.4 mM RNA, 20 mM phosphate, pH 7.0, 0.1 M NaCl, and 0.2 mM EDTA. (B) 500 MHz imino proton spectrum of a 2-fold dilution of the 0.4 mM RNA sample. The spectra are indirectly referenced through the chemical shift of water.
It was postulated that because the structure requires magnesium (or other
divalent ions) for cleavage, addition of MgCl₂ to the sample might drive this
equilibrium to a single structure. The spectra of dC-RNA2 in 20 mM MgCl₂ at
various temperatures are shown in figure 5.9. Clearly the addition of magnesium
has done little to stabilize the structure. In fact, there may be even more
resonances in these spectra than in those taken without MgCl₂, but this is difficult to
determine because of the poor S/N of the spectra. About the time these disappointing
spectra were acquired, Olke C. Uhlenbeck presented an informal seminar to discuss
the kinetic results his lab obtained with similar 3/4-1/4 constructs.
Interestingly, the rate of cleavage of these domains appeared to be sequence
dependent. Some constructs cleaved with a half-life of less than 2 min., whereas
others had a half-life of about 30 min. in the same conditions. Unfortunately there
was nothing obvious about the sequences of the quickly cleaving RNA's to distinguish
them from the slowly cleaving ones. This prompted an experiment to determine the
half-life of rC-RNA2 cleavage, as a slow cleavage rate would add credence to the
multiple structure hypothesis.

_Determination of the Cleavage Half-Life._ 0.2 µg of pure rC12-mer RNA and
was 5'.3²P end-labelled as described in section 5.3. After incubation for 1 hour at
37°C the reaction was quenched with 10 µL 30 mM EDTA. The reaction was
transferred to a tube containing a 3.2 µg aliquot of dry 3/4 RNA2. The 3/4 and 1/4
strands were annealed by heating the solution to 60°C and allowing it to cool slowly.
After annealing, the T4 polynucleotide kinase was removed by phenol/chloroform
extraction, and the nucleic acids precipitated with ethanol at -70°C. The pellet was
washed twice with 70% ethanol, dried, and resuspended in 20 µL TE, pH 7.6. rC-
Figure 5.9 Temperature dependence dC-RNA2. The sample was ca. 0.2 mM RNA, 20 mM phosphate buffer, pH 7.0, 20 mM MgCl₂, 0.1 M NaCl, and 0.2 mM EDTA. 8000 scans of 4K data points using a recycle delay of 0.5 sec. were averaged for each temperature. The temperature at which each spectrum was collected is denoted in the right-hand margin in degrees Celsius. The spectra are indirectly referenced through the chemical shift of water.
RNA2 was dispensed into ten eppendorf tubes, 2 μL per tube. At specific time points 2 μL of 10 mM MgCl₂ was added to the rC-RNA2 aliquot and incubated at 37°C. The reactions were quenched by addition of 4 μL "quench buffer" (10 mM EDTA, 80% formamide, 0.02% xylene cyanol, 0.02% bromophenol blue), and the RNA denatured by heating for 1 min. at 90°C. The denatured samples were immediately frozen with dry ice. A control reaction was prepared by diluting 2 μL of 5'³²P-rC12-mer with 6 μL quench buffer and denaturing as described above. 6 μL from each aliquot was loaded onto a 10x12x0.07 cm 20% polyacrylamide gel containing 7 M urea. The gel was run at 600V (11 mA) for 30 min. using 1X TBE [Maniatis et al. (1982) p. 454] as the reservoir buffer. The gel was visualized by autoradiography (fig. 5.10), from which the half-life was estimated to be approximately 5 minutes.

**Conclusion and Discussion.** RNA1 suffered from two major flaws; first, the RNA was completely cleaved before purification, casting doubt on whether the structure is relevant, and second, there were too many base pairs to obtain a complete, unambiguous assignment of the imino proton spectrum. To overcome these shortcomings, RNA2 was conceived. The ability of rC-RNA2 to cleave was demonstrated by analytical gel electrophoresis, and it was found that divalent manganese and zinc, in addition to magnesium, catalyzed cleavage. Divalent cadmium and lead, however, did not catalyze cleavage of RNA2. It was also demonstrated that dC-RNA2 did not cleave after an 18 hour incubation in the presence of Mg²⁺, Mn²⁺, Zn²⁺, Pb²⁺, Cd²⁺, or spermidine, proving that the 2'-OH of the base in the cleavage site is necessary for cleavage.

Because it is difficult to imagine the substitution of the ribonucleotide in the
Figure 5.10  Half-life Determination of rC-RNA2 in 5 mM MgCl$_2$ at 37°C. Lanes: 1. control.; 2. 5 min.; 3. 10 min.; 4. 20 min.; 5. 30 min.; 6. 40 min. The cleavage reaction has gone essentially to completion in 20 min.
cleavage site by a deoxynucleotide altering the structure required for cleavage, a system was designed in which the cleavage domain could be studied in cleavage conditions, without perturbing that structure by undergoing a cleavage event. To this end, the structure was designed with a sequence which should have manifested an imino proton spectrum well enough resolved to be completely, albeit not quite unambiguously, assigned. Although a good concept on paper, in practice this sequence was not too helpful, because the imino proton spectrum contained too many lines, implying a dynamic equilibrium between two or more structures. Addition of 20 mM MgCl₂ did not drive the equilibrium toward a single structure. Uhlenbeck's lab noticed that the cleavage half-life of similar domains appeared to be very sequence dependent, although a correlation between the sequence and fast cleaving was not apparent. The cleavage half-life of rC-RNA2 was determined to be about 10 min. at 37°C. Because this half-life is not unusually slow, it is difficult to rationalize the NMR spectra from this standpoint. In addition, the NMR spectrum of 3/4 RNA2 contains only four resonances, implying that this large fragment does not adopt additional conformations which would hinder the ability of the two strands to anneal quickly. It is possible that the NMR data is a consequence of concentration, as the analytical studies were performed at concentrations about 1000 times more dilute than were the NMR studies. In retrospect, the analytical cleavage reactions probably should have been performed at the same concentrations used for the NMR studies. The end result is that the discrepancy between the NMR data cannot be fully resolved. The latter data suggest that dC-RNA2 exists in at least two conformations in both the absence and presence of MgCl₂. Therefore, once again, the RNA sequence needed to be redesigned to study its structure further by NMR.
Chapter 6

NMR Investigation of RNA3
6.1 Introduction

Once again, two paths could have been taken toward designing a new RNA cleavage domain; the sequences of either RNA1 or RNA2 could have been altered. An RNA2-type system had obvious advantages, but there were disadvantages to this type of construct, as well. First, the results of Uhlenbeck implied that experimentation with numerous sequences would be necessary to find one that cleaved quickly, and thus had a greater propensity of existing as a single structure in solution. Second, the synthetic RNA had to be obtained from an outside source, as solid state RNA synthesis with commercially available ribonucleotide phosphoramidites did not work well in my own hands. For these reasons, and because RNA1 proved to exist as a single species in solution, it was decided that the sequence of RNA1 be altered. Another driving force behind this decision was the discovery that replacement of the 3'-5' phosphodiester linkage in the cleavage site by a 5'-3'-phosphorothioate diester linkage greatly slowed the rate of cleavage [Buzayan et al. (1988)]. This discovery made possible a source of slowly cleaving RNA by simply incorporating the proper nucleotide phosphorothioate during transcription. Once an RNA was designed which manifested an imino proton spectrum which could be unambiguously assigned, the phosphorothioate analog could be prepared and studied, thus alleviating any question about the relevancy of the structure. Ergo, RNA3 was born (figure 6.1).

6.2 Spectroscopy of RNA3

Approximately 8 mg of pure, lyophilized RNA3 was dissolved in 2.5 mL of column buffer (1.4 mM phosphate buffer, pH 7.2, 14 μM EDTA, and 0.003% NaN₃). This was loaded onto a Pharmacia PD-10 column which was previously equilibrated with 25 mL of the same buffer. The sample was eluted with column buffer in a
volume of 3.5 mL according to the enclosed Pharmacia instructions, and dried under vacuum (Savant speed-vac). The RNA was stored dry at -20°C. Just prior to NMR, RNA3 was dissolved in 0.25 mL 90% H2O/10% D2O and transferred to a Wilmad 428-PP sample tube. The final sample conditions were ca. 1.7 mM RNA3, 20 mM phosphate buffer, pH 7.2, 0.2 mM EDTA, and 0.02% NaN3. Between experiments the sample was remove from the tube, dried, and stored at -20°C.

A temperature series was recorded on a GN 500 spectrometer operating at 500 MHz using 1331 solvent suppression [Hore, P. J. (1983)] with the carrier frequency set at the resonance of H2O. 512 scans of 8K data points using a recycle delay of 1 sec. were averaged for each temperature increment. The temperature was increased in ten degree increments between 10 and 40°C, allowing the sample to equilibrate for 5 min. at each temperature. The temperature was then lowered to 35°C, followed by ten degree increments to a final temperature of 5°C, again allowing the sample to equilibrate at the desired temperature for 5 minutes. The FID's were transferred to a MicroVax workstation and manipulated using FTNMR software [Hare Research, unpublished]. The raw data were left-shifted one point, subtracted from the original FID, and appodized using a 30° phase-shifted sinebell with a skew of 0.3. Fourier transformation resulted in the frequency domain spectrum. Each spectrum was indirectly referenced through the chemical shift of the water resonance.

NOE difference spectra were recorded at 15 and 25°C on a GN 500 spectrometer using the 1331NO experiment (appendix B) with the carrier frequency set at the resonance of H2O. 1024 scans of 8K data points using a recycle delay of 0.5 sec. were averaged for each spectrum. The reference and saturated spectra were recorded in parallel by alternating the decoupler frequency between on and off.
Figure 6.1 Sequence of RNA3 shown in the proposed hammerhead secondary structure. The conserved nucleotides are boxed. The arrow denotes the cleavage site.
resonance at every scan. In cases of near degeneracy, the peaks were irradiated off exact top, in the direction away from the degenerate neighbor. The spectra were either appodized and Fourier transformed on the GN 500 spectrometer according to the scheme illustrated in figure 4.2, or on a MicroVax computer using the FTNMR software [Hare Research, unpublished] macro listed in table 4.1. The spectra are indirectly referenced to the chemical shift of water at the appropriate temperature.

Results. Figure 6.2 shows the 500 Mhz imino proton spectra of RNA1 and RNA3 taken at 20°C, and the sequences of the RNA's. The assignments obtained for RNA1 have been included in the figure. No unambiguous assignments of stem I for RNA1 were made, therefore this sequence was changed in RNA3. The central four GC base pairs caused spectral overlap in the 13-12 ppm range, so two of these were changed to AU base pairs in RNA3. Theoretically this should have reduced the number of resonances in the 13-12 ppm region of the spectrum by two, and added two additional resonances between 14.5 and 13.5 ppm. Comparison of the imino proton spectra in figure 6.2 shows that this is indeed the case; two extra resonances appear at ca. 14 ppm in the spectrum of RNA3, and a few peaks have clearly disappeared in the 12.6-12.2 ppm region, but it is difficult to determine exactly how many have disappeared. In addition, the sequential AU's of RNA3 represent a unique connectivity pathway in the molecule, even if the AU pair in the junction of the molecule is too broad to observe (as evidence from RNA1 suggests it may be). The sequences of stems II and III remain unchanged from those of RNA1, as these stems were already partially assigned for RNA1, and the changes to stem I should help facilitate further assignments of these stems by reducing the number of peaks in the 13-12 ppm region of the spectrum.
Figure 6.2  500 MHz imino proton spectra of RNA1 and RNA3 taken at 20°C with their respective sequences. The RNA1 sequence is labelled with the unambiguous imino proton assignments previously obtained. The experimental parameters of the RNA1 spectrum are given in the legend of figure 4.15. 512 scans of 8K data points using a recycle delay of 1 sec. were averaged for the RNA3 spectrum. Each spectrum is indirectly referenced through the chemical shift of water.
Figure 6.3 shows the NOE difference spectra which establish the connectivity of peaks M-A-I at 15°C. The arguments used to assign these resonances to stem II with the orientation depicted in figure 6.7 were identical to those for RNA1. As in the difference spectra of RNA1, resonance I gives an NOE at A. However, irradiation of resonance M for RNA3 does show an NOE at A, whereas for RNA1 this NOE was absent. This is most probably due to the temperature at which the RNA3 NOE data was collected rather than something particular to the sequence. Therefore either I or M could be the terminal base pair of this stem. However, resonance M is in a spectrally crowded region, whereas I is well resolved from its neighbors. Therefore I must be the terminal GC base pair as depicted in figure 6.7, because there is no possibility that I manifests an NOE to a nearly degenerate neighbor, as there is for resonance M.

Figure 6.4 shows the NOE difference spectra which establish the connectivity of peaks O-C-G-H-J-G' at 15°C. Again, arguments similar to those used to assign these resonances in RNA1 were used. During the course of this experiment the decoupler did not work properly, and there is no data for direct saturation of peak G'. In the difference spectra for the saturation of G, the nearly degenerate neighbor of G', there is a second order NOE to J. Saturation of J results in a first order NOE to G'. These two bodies of evidence confirm the assignment of G'. In addition, saturation of F (which is under G') at 25°C (fig. 6.5) establishes the J-G' connectivity directly. Evidence will be presented later which supports the assignment of resonance J to stem III, and K to stem I.

The difference spectra obtained upon saturation of resonances F and L 15°C and 25°C are shown in figure 6.5. Saturation of F at both temperatures produces a rather weak (being attenuated by the cut off in the 1331 excitation profile), but
Figure 6.3 500 MHZ NOE difference spectra which establish the connectivity of peaks M-A-I at 15°C. The sample was 1.7 mM RNA3, 20 mM phosphate buffer, pH 7.2, 0.2 mM EDTA, and 0.02% NaN3. 1024 scans of 8K data points were averaged using a recycle delay of 0.5 sec. were averaged for each spectrum. NOE's due to spill-over irradiation are labelled with an "X".
Figure 6.4 500 MHz NOE difference spectra which establish the connectivity of peaks O-C-G-H-J-G' at 15°C. Experimental parameters are given in the legend of figure 6.3. NOE's due to spill-over irradiation are labelled with an "X".
Figure 6.5 500 MHz NOE difference spectra connecting resonances L and F. (A) Difference spectrum collected at 15°C. (B) Difference spectrum collected at 25°C. The experimental parameters are given in the legend of figure 6.3. NOE's due to spill-over irradiation are labelled with an "X".
sharp NOE to the aromatic region of the spectrum, revealing an AU base pair beneath resonance G'. Based on the intensity of this aromatic NOE, the peak must be rather broad, and thus it will be difficult to observe NOE's to F upon irradiation of its neighbors. At 15°C, saturation of F results in a very broad NOE to L, but L does NOE back to F, which is not surprising. At 25°C saturation of F does not produce an NOE to L (it does, however, connect G' to J), but saturation of L at this same temperature appears to NOE to F. It is very unlikely for L to produce an NOE to F at 25 and not 15°C, since one expects F to become increasingly broad as the temperature increases. This apparent NOE must be due to spill-over irradiation to J, as peaks J and G' are connected, and F is nearly degenerate with G'. In addition, peaks J and G' are relatively sharp and produce strong NOEs, and it is not unreasonable that spill-over NOEs from these peaks be observed. This leaves F-L as an isolated A/U-G/C connectivity. The degeneracy of F and G' further complicate the issue, as there may be an NOE between F and G', which would assign both F and L into stem III, as shown in figure 6.7. However, there are two A/U-G/C sequences in stem I, and without a clear NOE between F and G' or a complete assignment of stem I the stem III assignment cannot be made with confidence.

One final connectivity pathway could be determined from the NOE data. Figure 6.6 shows the NOE difference spectra of resonances D-K-N at 25°C. The D-K connectivity was hinted at in figure 6.4, as saturation of J produces a very weak NOE to D. Saturation of D clearly produces an NOE to K and a sharp aromatic NOE, identifying D as an AU base pair. Because D produces only one NOE it is either an AU pair terminating a helix, or neighbors a base pair which has a resonance too broad to NOE. Clearly from the sequence of RNA3, either assignment is possible for D. Saturation of K produces NOEs to H, G', and N. The H and G' NOE's have already been
Figure 6.6 500 MHz NOE difference spectra which establish the connectivity of peaks D-K-N at 25°C. The experimental parameters are given in the legend of figure 6.3. NOE's due to spill-over irradiation are labelled with an "X".
identified as those of J, and hence N must be the nearest neighbor to K. These NOE patterns describe the connectivity D-K-N, with the sequence being A/U-G/C-G/C. The only place this sequence occurs which has not already been confidently assigned is in stem I. However, because no A/U-A/U NOE was observed upon irradiation of D, the orientation of the sequence within the helix cannot be determined. Both resonances D and E were saturated 10°C (data not shown), but the S/N of the spectra were too poor to observe NOE's, and the orientation of these assignments remains ambiguous.

As was observed for RNA1, there are too few resonances in the 14.5-13.5 ppm region of the spectrum to account for all of the AU base pairs in the sequence. It is not unreasonable that the AU at the cleavage site exchange with solvent protons too rapidly to manifest an imino proton resonance, as this resonance is subject to end-fraying after cleavage of the RNA. It was also previously suggested that the AU of stem III in the junction of the molecule also be in fast exchange with solvent, but clearly there are two additional AU's which have not been assigned, namely the AU closing the hairpin loop of stem I, and the AU adjacent to that in the cleavage site.

_Melting Data._ Figure 6.7 shows the imino proton spectrum of RNA3 plotted as function of temperature. The resonances are labelled on the 15°C spectrum for clarity. As the temperature decreases from 15° to 5°C, E grows in, and it also appears that a resonance which is degenerate with C grows in, as the intensity of C increases compared to that of D, and there is a slight shoulder on the left side of C at 15°C. O and P also split apart as the temperature is decreased. At 5°C, the entire spectrum has suffered a loss of resolution, but there is no evidence that resonances I and M have broadened more than the others, as was the case for RNA1. Also
Figure 6.7 500 MHz imino proton spectrum of RNA3 as a function of temperature. 512 scans of 8K data points using a recycle delay of 1 sec. were averaged for each temperature increment. Each spectrum is indirectly referenced through the chemical shift of water. The temperature at which each spectrum was recorded is indicated in the right margin in degrees Celsius.
appearing in the spectrum as the temperature decreases are two resonances at ca. 10.5 ppm. As discussed for RNA1, these resonances most likely arise from nucleotides in the hairpin loops or single stranded regions of the RNA.

Upon increasing the temperature to 20°C, resonance E almost completely disappears. It was speculated that this resonance is one of the stem I AU base pairs near the cleavage site, and low melting temperature corroborates this speculation. Increasing to 30°C results in substantial broadening of peaks A, D, and L, and some broadening of peaks F/G', I, and M'. At 35°C, resonances D and L have completely disappeared, and A, I, G', and M' have further broadened. The only sharp resonances remaining at 40°C are C, G, J, K, M, N, M", and O. From the labeling of the peaks in the 15°C spectrum it appears as though O has broadened and P remains sharp. However, as the helix melts, some bases can "swing-out" from their stacked position within the double helix. Imino protons attached to bases which are adjacent to those which no longer stack properly are no longer ring current shifted to a lower field, and thus move upfield in the spectrum. For this reason, resonance O has moved up-field of P, and it is indeed O that remains sharp at 40°C.

Based on sequence, the predicted stability of stem II is ? = M > A > I. Although the exact location of ? is unknown, it is clear from the other assignments that this resonance is either M', M", or P. The order of melting of these resonances corroborates this assignment. It was speculated that resonances G', F, and L are assigned to stem I as shown in figure 6.8. If this assignment pathway is correct, the resonances should melt in the order L > F > G'. Clearly resonance L melts before G', but because F is concealed by G' it is difficult to say anything about the status of F at this temperature. The central base pairs of stem III and the GC base pairs of stem I are sharp at 40°C, as expected.
Conclusion. The assignments of RNA3 are summarized in figure 6.8. The sequence of RNA1 was altered to give a better resolved imino proton spectrum. The sequence of stem I was changed to reduce the number of resonances in the crowded 12-12 ppm region of the spectrum. In addition, this sequence change should have introduced two resonances between 14.5 and 13.5 ppm. This was demonstrated in figure 6.2. The specific change was chosen not only to relieve the crowding observed in the 13-12 ppm region of the RNA1 spectrum, but also to introduce a unique connectivity pathway to facilitate the assignment procedure. The M-A-I and O-C-G-H-J connectivities which were established for RNA1 were corroborated by the RNA3 spectra. In addition, the stem III connectivities were extended to include O-C-G-H-J-G' and possibly F-L in RNA3. In addition, RNA3 provided some stem I assignments, which were not obtained for RNA1. This was clearly due to the sequence design of RNA3, as RNA1 contained too many GC pairs in this stem to easily assigned. However, this sequence did not lead to a complete imino proton assignment, and hence yet another RNA sequence was born.
Figure 6.8  Sequence and 500 MHz imino proton spectrum of RNA3 at 20°C. The unambiguous assignments are shown in upper case letters, possible assignments in lower case letters. The conserved nucleotides are boxed, the cleavage site denoted by the arrow.
Chapter 7

NMR Investigation of RNA4
7.1 Introduction

Although much progress toward a complete secondary imino proton assignment was made with the RNA3 sequence, a few more sequence changes were needed to fully assign the spectrum. The GC closing the hairpin loop of stem II was not assigned for either RNA1 or RNA3, but it can be said with confidence that this base pair gives a resonance in the 12.5-12.0 ppm region of the spectrum. In addition, resonances F, G, and G' are very crowded in the RNA3 spectrum. Although resonances G and G' have been assigned, relieving the F/G near degeneracy would help facilitate the assignment of F. Unfortunately, nucleic acids are composed of only four major bases, so the availability of possible mutations is greatly reduced compared to proteins. In addition, the bases pair with only two geometries, A with U and G with C. For this reason it is difficult to manipulate nucleic acid sequences in a way which will most benefit the spectroscopist. Replacing a GC pair with an AU does remove a resonance from the upfield region of the imino proton spectrum, but this type of change can be undesirable because it also introduces a resonance in the 15-13 ppm region. Although the order in which the bases are stacked affects the chemical shift of neighboring resonances, these changes are not as great as those observed for base pair replacement, and the direction and extent of the changes cannot be readily predicted. A change which relieves spectral crowding in one region may increase crowding in another. Also, one has to keep in mind the biological activity of the molecule; arbitrary changes which inactivate the structure are not very useful. How then, could the sequence of RNA3 be manipulated to exhibit the desired spectral changes without affecting the "cleavability" of the domain? The answer is through incorporation of GU wobble base pairs. The use of wobble pairs has distinct spectroscopic advantages. First, each GU wobble pair contains two imino protons.
and thus gives rise to two imino proton resonances, allowing for two independent nearest neighbor assignment pathways. Second, the two imino protons of the GU wobble generally resonate upfield of Watson-Crick imino protons, reducing chemical shift degeneracies. Third, the two imino protons produce strong mutual NOE's (30-40%) due to their close spatial proximity (<3 Å), allowing for easy identification of the GU wobble pair. In fact, GU4 of yeast tRNA\textsuperscript{Phe} [Johnston & Redfield (1978)], and GU50 of both E. \textit{Coli} tRNA\textsuperscript{fmet} [Hurd & Reid (1979)] and E. \textit{Coli} tRNA\textsubscript{1}Val [Johnston & Redfield (1979)] were the first reliably assigned resonances in their respective low-field NMR spectra [Reid (1981)]. In addition, x-ray-crystallographic studies of GC vs. GT pairing in DNA have shown minimal perturbation in backbone torsion angles on replacing a Watson-Crick base pair with a GT wobble base pair [Kneal, et al. (1985); Ho, et al. (1985)], making the use of GU wobble pairs even more attractive.

### 7.2 Spectroscopy of RNA4

Approximately 12 mg of pure, lyophilized RNA4 was dissolved in 2 mL of 5 mM potassium phosphate buffer, pH 7.0 and dialyzed against 4 liters of 5 mM potassium phosphate buffer, pH 7.0, 250 mM NaCl, 0.05 mM EDTA, and 0.005% NaN\textsubscript{3} for 24 hours at 4°C. After removal from the dialysis tubing the RNA was dried by lyophilization and stored at -20°C. Just prior to spectroscopy the RNA was dissolved in 0.5 mL of 90% H\textsubscript{2}O/10% D\textsubscript{2}O to give a 1 mM RNA, 20 mM phosphate, pH 7.0, 0.1 M NaCl, 0.2 mM EDTA, 0.02% NaN\textsubscript{3} NMR sample. This was transferred directly to a Wilmad 528pp NMR tube. Between NMR experiments the solution was removed from the NMR tube and the RNA stored as a lyophilized powder at -20°C. For spectra recorded with magnesium ions, the NMR sample was made to
Figure 7.1 GU wobble base pair. Note that this pair differs from Watson-Crick base pairs in that it contains two hydrogen-bonded imino protons.
10 mM by addition of 5 µL 1 M MgCl₂ directly to the sample.

Melting data was collected on a GN 500 spectrometer using solvent suppression [Hore, P. J. (1983)] with the carrier frequency set at the resonance of water. 1024 scans of 8K data points using a recycle delay of 0.5 sec. were averaged for each temperature increment. The spectra were recorded by increasing the temperature in ten degree increments from 15° to 35°C, allowing the sample to equilibrate for 5 min. at each temperature. The temperature was then raised to 40°C, and lowered in ten degree increments to a final temperature of 10°C. Again a 5 min. equilibration time was used at each temperature. The spectra were recorded in this manner to ensure that the RNA renatured properly after unfolding. The raw data were left-shifted one point, stored, and subtracted from the original FID to obliterate the residual H₂O resonance. This FID was then appodized using a DM=8, Fourier transformed, and phase corrected using GN 500 software. The frequency domain spectra were then transferred to a MicroVax Workstation III and plotted using FTNMR software [Hare, D. unpublished]. The spectra are indirectly referenced through the chemical shift of water at each temperature.

A complete set of NOE experiments was collected at both 15° and 20°C as previously described for RNA3. 1024 scans of 8K data points were averaged for each experiment. The off-resonance FID was subtracted from the on-resonance FID using GN 500 software, and the resultant FID transferred to a MicroVax III workstation. Fourier transformation, phase correction, and baseline correction using FTNMR software [Hare, D. unpublished] generated the difference spectra. The reference spectrum was appodized using a 30° phase-shifted sinebell with a skew of 0.3.
Figure 7.2 Sequences and 500 MHz imino proton spectra of RNA3 and RNA4 at 15°C. The conserved nucleotides are boxed, cleavage site denoted by the arrow. The assignments obtained from difference spectroscopy for RNA3 are labelled on the sequence; upper case denotes confident assignments, lower case tentative assignments. Note that RNA4 is identical to RNA3 with the exception of two GU wobble base pairs, shown in bold print.
Results, NOE Difference Spectroscopy, No MgCl$_2$. Figure 7.2 shows the sequences and 500 MHz imino proton spectra of RNA 3 and RNA4 at 15°C. The RNA3 assignments are labelled on the sequence. The RNA4 sequence is identical to that of RNA3 with the exception of two GU wobble base pairs, shown in bold print. Based on the RNA3 assignments, the resonance labelled G' in the RNA3 spectrum should have disappeared in the RNA4 spectrum, with the introduction of two new upfield resonances. Similarly the resonance from the GC base pair closing the loop of stem II of RNA3 should have disappeared, again with the introduction of two upfield resonances. This resonance could not be assigned in RNA3 due to spectral crowding, but is known to resonate between 12.5 and 12.0 ppm. Examination of the RNA4 spectrum reveals that indeed G' has disappeared. In addition, peak N has moved out from under M and M', and based on the intensity of these resonances, it appears as though RNA4 contains one fewer resonance than does RNA3 in this region of the spectrum. Four resonances between 12 and 10 ppm are clearly visible in RNA4 that were absent in RNA3. These resonances are presumably due to the two GU wobble base pairs of RNA4. Resonance F shifted down-field due to the replacement of its GC neighbor with a GU wobble pair. Comparison of the spectra reveal other differences as well. Resonances B, E, and F are quite sharp in RNA4, whereas these resonances were somewhat broad in RNA3. In fact, B was a mere shoulder on C, and F was not clearly visible at all. This was bothersome at first, because the sequences are identical save for the GU wobbles, but close scrutiny of my lab notes revealed that the RNA4 NMR sample contained 0.1 M NaCl, and the RNA3 NMR sample did not. Therefore the increased stability of these resonances can be attributed to the increased ionic strength of the solution. The fact that all of the resonances in the RNA4 spectrum are sharper than those of RNA3 substantiates this argument, as do
the melting data to be presented later; the helical stems of RNA4 in the presence of 0.1 M NaCl have slightly higher T_m's than do their RNA3 counterparts. One other noticeable difference between the two spectra is the splitting of the J/K degeneracy of RNA4. This is due to the GU wobble of stem III, as it slightly alters the electronic environment of resonance J in RNA4 relative to its electronic environment in RNA3. The spectral changes observed upon replacement of two RNA3 GC base pairs with GU wobble base pairs were not only consistent with the predicted spectral changes, but also looked very promising, as the imino proton spectrum of RNA4 is very well resolved.

Figure 7.3 shows the NOE difference spectra which establish the sequential connectivity of peaks R/T-M-A-I. The exact experimental parameters are given in the figure legend. Irradiation of T, which is broad even at 15°C produces a strong NOE to R and a weak NOE to M. This NOE pattern is indicative of a GU wobble pair, and because R and T have only one nearest neighbor NOE to M, the GU wobble base pair must terminate a helix. The only terminal GU wobble is that which closes the hairpin loop of stem II. The M-A-I connectivity has already been established for RNA1 and RNA3, and is clearly illustrated in fig. 7.3. Irradiation of M produces a weak NOE to A. The expected NOE's to R and T are not visible because these resonances are too broad to manifest NOE's. However, the fact that they produce NOE's to M upon being irradiated clearly defines the R/T-M connectivity. Saturation of A produces NOE's to both I and M, and a sharp aromatic adenine CH2 NOE, identifying A as an AU base pair with M and I as neighbors. Saturation of I produces only an NOE to A, identifying it as a terminal base pair.

Figure 7.4 shows the imino-imino NOE's sequentially connecting resonances E-D-K-N-B. The near degeneracy of resonances B and C and the lack of NOE to B upon
Figure 7.3 NOE difference spectra of RNA4 which establish the connectivity of peaks R/T-M-A-I. The sample was 1 mM RNA4, 20 mM phosphate buffer, pH 7.0, 0.1 M NaCl, 0.2 mM EDTA, and 0.02% NaN₃. NOE's are labelled according to the reference spectrum. Spill-over NOE's are labelled with an "X". Peaks M, A, and I were saturated at 20°C, R and T were saturated at 15°C.
Figure 7.4 NOE difference spectra of RNA4 which establish the connectivity of peaks E-D-K-N-B. Sample conditions are given if the legend of figure 7.3. All spectra were recorded at 20°C. Spill-over NOE's are labelled with an "X".
Figure 7.5 Saturation of resonance C at both 15° and 30°C. Sample conditions are given in the legend of figure 7.3. At 30°C resonance B has broadened severely such that only NOE's due to saturation of C remain.
irradiation of N makes the N-B connectivity slightly ambiguous. However, two lines of evidence support this assignment. The first is that although irradiation of peak N does not result in an NOE to B, it does result in and NOE to the same aromatic CH2 that B produces an NOE to when it is saturated. Secondly, as the temperature is raised to 30°C resonance B broadens and disappears, and saturation of C at 30°C results in NOE's only to peaks O and G (fig. 7.5), meaning that it must be saturation of resonance B which manifests the NOE to N. The imino-imino NOE connecting E and D is also weak due to the near degeneracy of these two peaks, but again NOE's to a common aromatic CH2 confirm this assignment. The sharp NOE's to aromatic CH2's identify resonances B, D, and E as AU base pairs. Because sequential AU pairs are unique to stem I, determination of the E-D connectivity assigns resonances E-D-K-N-B as those of stem I.

Figures 7.6 and 7.7 show the NOE difference spectra which define stem III. Saturation of P results in a weak spill-over NOE to C and an NOE to M'. The fact that the M' NOE is of greater intensity than that of N proves that the M' NOE is real. If M' were due to spill-over irradiation, resonance N would have greater intensity than M' because it is closer in chemical shift to P than M' is. Saturation of M' results in a weak spill-over NOE to A and the expected NOE to P. The connectivity O-C-G-H-J has been discussed for both RNA1 and RNA3, and is clearly shown in figure 7.6. The neighbor of J in RNA4 is a GU wobble base pair. As discussed earlier a wobble pair has two imino protons, and thus saturation of J should result in NOE's to H and two other resonances. In addition, these two other resonances should manifest strong mutual NOE's. Figure 7.7 shows that saturation of J produces NOE's to H, Q, and S. Saturation of Q results in a very strong NOE to S, and a relatively weak NOE to J. Saturation of S produces a similarly strong NOE to Q and the same weak NOE to J.
Figure 7.6 NOE difference spectra showing the connectivity of peaks M'/P-M'/P-O-C-G-H at 20°C. The sample conditions are given in the legend of figure 7.3. Spill-over NOE's are labelled with an "X".
Figure 7.7 Noe difference spectra showing the connectivity of peaks J-Q/S-F-L at 20°C. The sample conditions are given in the legend of figure 7.3. Spill-over NOE's are labelled with an "X".
This is the NOE pattern expected for J and its neighboring GU wobble pair. This then establishes that resonances Q and S are due to the GU wobble of stem III. Continuing with the stem III assignments, saturation of F produces a very weak NOE to L and a sharp aromatic CH2 NOE identifying F as an AU base pair. Saturation of L, however, does not result in an NOE back to F to confirm F's assignment as being that of stem III, nor does F produce the expected NOEs to peaks Q and S at 20°C. The assignment of F can be confirmed at 15°C, as is shown in figure 7.8. Saturation of L at this temperature results in a very weak, albeit real, NOE to F, further substantiating these assignments. Saturation of U at both 15° and 20°C resulted in no NOE's (data not shown), and again it is presumed that this resonance is due to a nucleotide in either a hairpin loop or single stranded region of the RNA. The complete RNA4 imino proton NOE's are summarized in figure 7.9. It should be noted that the exact order of the M'-P connectivity cannot be determined because of the near degeneracy of peaks O and P. However, these base pairs are not needed for cleavage, and this slight ambiguity in no way diminishes the usefulness of these secondary assignments.

As was previously observed for RNA1 and RNA3 there is no NMR evidence to substantiate the existence of the AU base pairs which terminate stems I and III. The studies with RNA1 provided evidence which suggested that the junction of the three helices is stressed in as the temperature decreases, and also upon addition of magnesium to the sample. The fact that these AU's do not pair may be due to the structure of the molecule, but the absence of these base pairs could also be due to end effects, as discussed for RNA1. More will be said about the base pairs in the junction of the helices in a subsequent section.
Figure 7.8 Saturation of resonances F and L at 15°C. The sample conditions are given in the legend of figure 7.3.
Figure 7.9 Sequence and 500 MHz imino proton spectrum of RNA4 at 20°C. The sample conditions are given in the legend of figure 7.3. The conserved nucleotides are boxed, the cleavage site denoted by the arrow. The resonance assignments are labelled on the sequence.
Results, Temperature Studies, No MgCl₂. Figure 7.10 shows the 500 MHz imino proton spectrum of RNA4 plotted as a function of temperature. Similar to RNA1 and RNA3, the melting behavior of the helical stems was used to corroborate the assignments obtained by difference spectroscopy. As the temperature is increased from 20° to 25°C, the terminal base pairs and the base pairs closing the loops of stems II and III (resonances B, E, R, and T) begin to open and exchange with H₂O. Upon increasing to 30°C, resonance E has completely disappeared and peaks A, I, R, and T of stem II have broadened significantly. In addition, the terminal base pairs of stem III (resonances F, L, M', and P) begin to open and exchange with H₂O. Increasing the temperature to 35°C results in a further broadening of these resonances, and finally at 40°C only resonances M (the GC base pair of stem II), N and K (the two GC base pairs of stem I), and C, O, G, H, Q, and S (the center five base pairs from stem II) remain stable. This uniform "unzipping" from the ends of the individual helical sub-domains is in agreement with the assignments of figure 7.9, and suggests that there is no cooperativity between the helical stems upon denaturation. It should be noted that for RNA1 and RNA3 the GC pair closing the loop of stem II appeared to be stable at 40°C, as there was an additional resonance in the region of M, M', and N at this temperature. The fact that this base pair is not stable at 40°C in RNA4 is not surprising, as GU wobble pairs are generally less stable than their Watson-Crick GC counterparts.

If the RNA tumbles isotropically in solution, imino proton resonances corresponding to base pairs with long lifetimes on the NMR time scale should have similar linewidths. It is clear from figure 7.8 that this is not the case. The central base pair resonances of stem I (D, N, and K) exhibit slightly narrower linewidths than do similarly stable base pair resonances associated with stems II and III.
Figure 7.10 500 MHz imino proton spectrum of RNA4 plotted as a function of temperature. The sample conditions are given in the legend of figure 7.3. The temperature at which each spectrum was recorded is indicated in the right margin in degrees Celsius.
Because narrow linewidths are indicative of increased molecular motion, this suggests that stem I possesses motion, possibly due to the nick introduced during cleavage, which is independent of the global structure tumbling in solution. Stem I can, in essence, "flap around" independently of the rest of the structure. It is interesting that the central base pairs of stem III do not exhibit similarly narrow lines, as this stem is also terminated at the cleavage site. It is conceivable that the observed bases are interacting in some way to inhibit additional motion of stem III, as this stem contains four conserved bases. Stem I on the other hand contains no conserved residues, and it is possible that the tertiary folding of the domain allows additional motion of this stem.

Results, Temperature Studies, 10 mM MgCl₂. Magnesium chloride was added to the sample with the hope that it would stabilize tertiary base pairing interactions necessary for cleavage, and thus introduce new resonances into the imino proton spectrum. Figure 7.11 shows the imino proton spectrum of RNA4 in 10 mM MgCl₂ plotted as a function of temperature. Comparing the 20°C spectrum with that taken in the absence of MgCl₂ (fig. 7.10) reveals that there are no additional resonances. But this is not to say that addition of magnesium has not perturbed the structure. There is a general decrease of chemical shift dispersion, especially in the regions containing peaks A-F and L-N. Also, resonances B and C as well as T and U have become completely degenerate, suggesting some structural rearrangement. It is interesting that resonance I has disappeared, and that A, E, and L have broadened significantly more than the other resonances upon decreasing the temperature. It was suggested for RNA1 that the base pairs in the junction of the helices destabilize as the structure approaches a cleavage conformation. However, all of the junction
Figure 7.11 500 MHz imino proton spectrum of RNA4 plotted as a function of temperature. The sample was ca. 1 mM RNA4, 20 mM phosphate buffer, pH 7.0, 10 mM MgCl\(_2\), 0.1 M NaCl, 0.2 mM EDTA, and 0.02% NaN\(_3\). The temperature at which each spectrum was recorded is indicated in the right margin in degrees Celsius.
base pairs were not assigned for this RNA. All of the base pairs have been assigned for RNA4, and its behavior is consistent with that observed for RNA1; the resonances corresponding to the junction base pairs broaden upon addition of MgCl$_2$ to the sample. As was speculated for RNA1, this effect is probably due to a conformational change rather than solvent exchange. Along these lines, resonances D, K, and N no longer exhibit unusually narrow lines in the 10 mM MgCl$_2$ spectrum. Although resonances N and D are no longer well enough resolved for linewidth comparison, it is clear that K, which is still well resolved in 10 mM MgCl$_2$, is not narrower than J. In fact, below 25°C resonance K is much broader than J. The addition of MgCl$_2$ has apparently "tacked down" the once independent stem I, thereby eliminating its independent motion. Unfortunately the exact interactions which are responsible for eliminating the independent motion of stem I cannot be identified.

Addition of MgCl$_2$ has also increased the stability of the helices. At 40°C resonances A, I, and M of stem II remain, whereas only resonance M remains at this temperature in the absence of MgCl$_2$. The T$_m$'s of the terminal resonances of stem III have also increased. This behavior is most likely due to the increased ionic strength of the solution rather than to magnesium ion-specific interactions. In fact the cleaved form may not have a "good" magnesium binding site, but without further data nothing more specific can be said.

Ultimately the goal of these RNA studies was to identify the tertiary interactions, in the form of base pairings between conserved nucleotides, which drive the cleavage conformation of a hammerhead RNA domain. To achieve this goal it was necessary to design an RNA which not only cleaved, but which also exhibited an imino proton spectrum well enough resolved to be completely, unambiguously assigned. This intermediate goal was achieved by logical sequence design. The
originally conceived RNA1 resulted in only a partial imino proton assignment. However, the RNA1 data seemed to be consistent with the hammerhead folding model, as the spectrum contained the predicted number of resonances for the model. The RNA1 spectrum suffered from many chemical shift near degeneracies, and thus the sequence was changed, resulting in RNA3. Spectroscopy of RNA3 resulted in additional, albeit not complete, imino proton assignments. A complete assignment of the imino proton spectrum was finally achieved upon the replacement of two Watson-Crick GC base pairs in RNA3 with GU wobble pairs. This assignment substantiated, with the exception of two terminal AU pairs, the hammerhead folding model predicted by Symons et al. (1987). The absence of imino resonances for the AU pairs does not necessarily suggest that these residues are not base paired. The NMR melting of a base pair does not require that the pair is "open" in the sense that a pair opens for an optical melt. A base pair melts in the NMR spectrum when the imino proton exchanges with solvent sufficiently fast on the NMR time scale that its resonance line moves to a chemical shift which is at the populations weighted average of the bonded and non-bonded forms. Because the imino proton exchanges with protons from water, which are at a concentration of 110 M, the resonance line moves under the water resonance, thus disappearing from the imino proton spectrum. However, the proton can exchange without fully unstacking the base pair. Therefore these AU pairs may in fact be in the stacked geometry associated with A double helices. There was evidence for RNA1 that the AU terminating stem III may be present, although the possible resonance (A') was not assigned. Figure 7.12 shows the imino proton spectra of RNA1, RNA3, and RNA4 at 20°C. Except for the introduction of a GU wobble pair in stem III of RNA4, this stem is identical in the three RNA's. Therefore resonance F, which was unassigned for RNA1, must be due to
Figure 7.12 500 MHz imino proton spectra of RNA1, RNA3, and RNA4 at 20°C. The sample conditions are given in the legends of figures 7.3 (RNA4), 6.3 (RNA3), and 4.4 (RNA1).
AU₉ of stem III (F has shifted out from under G in the RNA4 spectrum because of the neighboring GU wobble pair). F is, as is expected, at the same chemical shift for both RNA1 and RNA3. Resonances D and E are not present in the RNA1 spectrum because the RNA1 sequence does not contain these base pairs within stem I. This leaves the two AU base pairs terminating stems I and III, and the AU closing the hairpin loop of stem I as possible assignments for A' in RNA1. Because the loop sequence is different for RNA1 than it is for RNA3 and RNA4, the base pair closing this loop probably resonates at a different frequency for RNA1 than for RNA3 and RNA4. Therefore I believe that the resonance previously labelled A' on the RNA1 spectrum is the AU pair closing the hairpin loop of stem I, and to keep consistency with the labelling of RNA3 and RNA4, it has been re-labelled as resonance B in figure 7.12. The only rationale for this assignment is that neither the RNA3 nor the RNA4 imino proton spectra contain evidence of base pairing for the AU's terminating stems I and III. This reasoning then refutes the possibility that the junction AU's are paired. The fact that no additional low field resonances grow in with decreasing temperature for RNA1, RNA3, or RNA4 further supports this hypothesis, as one would expect this for bases which are capable of pairing.

While the imino proton spectrum has been successfully assigned, the final goal was not achieved. Although the spectra of both RNA1 and RNA4 changed somewhat upon the addition of magnesium ions, no new resonances were introduced (figure 7.13). It is difficult to rationalize the conservation of thirteen nucleotides in such a small structure unless they interact in a sequence specific manner. It is therefore evident that cleavage of the RNA results in a subsequent rearrangement of the structure, and this cleaved form does not represent a true cleavage conformation. This should not detract from the fact that the cleaved domain provided evidence that
Figure 7.13 500 MHz imino proton spectra of RNA1 and RNA4 in 10 mM MgCl₂ taken at 30°C. The sample conditions are given in the legends of figures 7.11 (RNA4) and 4.19 (RNA1).
this class of self-cleaving RNA does in fact fold into the predicted hammerhead secondary structure. To further study a truly relevant cleavage domain it was necessary to isolate slowly cleaving RNA in its uncleaved form and study the domain under cleavage conditions. The next chapter describes the design of such an RNA and the spectroscopic results obtained.
Chapter 8

Investigation of RNA6
8.1 Introduction

The results of the spectroscopic studies of RNA4 were concomitantly exciting and disappointing. Although a complete secondary imino proton assignment was successfully obtained, little insight into the tertiary folding of the domain was revealed, even in the presence of magnesium ions. In fact, these studies cast some doubt on the relevance of studying the cleaved form of the RNA. The most obvious solution to this problem was to resurrect the RNA2 concept and devise another 3/4:1/4 RNA domain. However, this did not seem a particularly attractive solution, as RNA2-type constructs suffered from sequence dependent cleavage rates, and the requisite of a chemically synthesized 1/4 RNA strand. The fact that the secondary assignment of RNA4 was already obtained, and changing to a bimolecular construct meant literally "going back to the drawing board," made this approach even less satisfactory. A complete secondary imino proton assignment could not be guaranteed even if a suitable sequence could be found and the 1/4 strand could be chemically synthesized in high yield. It was therefore much more desirable to mutate RNA4 in some way which would decrease the cleavage rate such that uncleaved RNA could be purified and studied under cleavage conditions.

One possibility, mentioned briefly already, involved incorporation of adenosine-5'-O-(1-thiotriphosphate) during the transcription reaction. This would replace every 3'Ap-N linkage in the RNA with a phosphorothioate linkage. Most importantly, the 3'Ap-C linkage in the cleavage site would be replaced with a phosphorothioate, greatly reducing the cleavage rate of the RNA [Buzayan et al. (1988)]. This method was tried by Mark Kubinec in the Wemmer lab, but unfortunately the NMR spectrum of this RNA suffered from unusually broad lines, and very little information was obtained.
Due to the unexpected results of the phosphorothioate studies it was thought that the domain would have to be completely redesigned as a bimolecular construct. However, new light was shed on the problem by Dr. Rosie Kim. While generating mutants of RNA1 for crystallographic purposes she created an RNA sequence which did not cleave during the course of a two hour transcription reaction. After sequencing the plasmid DNA template, she determined that the mutation occurred at one of the conserved nucleotides. This mutation was incorporated into the RNA4 sequence, resulting in RNA6 (fig. 8.1). Aside from the mutation (shown in bold type) the sequence differs from RNA4 in that it contains only one GU wobble base pair (stem II).

The cleavage rate of RNA6 was studied under a variety of conditions, and it turned out that the cleavage was not in fact stopped, but the rate significantly attenuated by this mutation. This was a far more desirable result for spectroscopic studies, as a slowly cleaving RNA must be adopting a relevant cleavage conformation.

8.2 Determination of the Cleavage Half-life of RNA6

Analytical Transcription Reactions. Because of the time required to collect NOE difference or NOESY spectra, a rough estimate of the cleavage half-life of RNA6 was desired. 20 μL transcription reactions (40 mM tris·HCl, pH 8.1 at 37°C, 6 mM MgCl₂, 5 mM NaCl, 1 mM DTT, 1 mM spermidine, 1 mM ATP, 1 mM GTP, 1 mM CTP, 1 mM UTP) containing 2 μg RV/RI/pRNA6 and 2 μL pure T7 RNA polymerase (activity unknown) were initiated at appropriate times and incubated at 37°C. All of the reactions were quenched at the same time by addition of 1 μL 100 mM EDTA.
Figure 8.1 Sequence of RNA6 shown in the proposed hammerhead structure. The mutation is shown in bold type, the conserved nucleotides are boxed, and the cleavage site is denoted by the dashed arrow.
and 5 μL 5X denaturing gel loading dye (80% formamide, 0.25% xylene cyanol, 0.25% bromophenol blue). The samples were heated to 90°C for 1 min. and transferred immediately to an ice water bath. 5 μL from each reaction was loaded onto a 8x10x0.07 cm 10% polyacrylamide gel containing 7 M urea. Electrophoresis was carried out against 1X TBE buffer [Maniatis et al. (1982) p. 156] at 600 Volts for 20 min. The gel was stained with 0.02% toluidine blue-O (Sigma) for approx. 2 min., destained with water, and photographed.

Results. The results of the cleavage reactions are shown in figure 8.2. The upper most band on the gel corresponds to uncleaved RNA6. The cleavage products are labelled on the gel. The cleavage half-life is between 8-20 hours, and based on the extent of cleavage observed for the 8 and 20 hour incubations, 12 hours is a reasonable estimate of the actual half-life. The extra bands which begin to appear at the 20 hour time point and increase in intensity with increasing time are degradation products.

8.3 Spectroscopy of RNA6

Preparation of the Sample. Approximately 6 mg of pure, lyophilized RNA6 was dissolved in 2 mL of dialysis buffer (2 mM phosphate buffer, pH 7.0, 0.02 mM EDTA, 0.01 M NaCl, 0.002% NaN₃) and dialyzed (1000 MWCO tubing).against 2 L of the same buffer at 4°C for 24 hours. The solution was removed from the tubing, transferred to a fresh tube, and dried by lyophilization. RNA6 was stored dry at -20°C. Just prior to NMR spectroscopy, RNA6 was dissolved in 220 μL 90% H₂O/10% D₂O and transferred to a Wilmad 428-PP NMR tube. For spectra recorded in the presence of magnesium ions, a small volume of 1M MgCl₂ was added
Figure 8.2 Half-Life Determination of RNA6. Reaction conditions are given in the text. Lanes: 1. 0.5 hour; 2. 1 hour; 3. 2 hour; 4. 4 hour; 5. 8 hour; 6. 20 hour; 7. 26 hour; 8. 44 hour.

UC = uncleaved RNA; LF = large fragment; sf = small fragment
directly to the NMR sample. The exact buffer conditions are indicated in the figure legends.

**NOE Difference Spectroscopy.** NOE difference spectra were recorded at 15°C using the 1331NO experiment (appendix B) with the carrier frequency set at the resonance of water. 1024 scans of 8K data points using a recycle delay of 0.5 sec. were averaged for each spectrum. The FID's were transformed using GN 500 software according to the scheme depicted in figure 4.2.

**NOESY Spectroscopy.** Two-dimensional NOESY spectra were recorded at both 15° and 20°C using a NOESY experiment modified for H2O solvent suppression, called 13NSY (appendix B). The modification replaces the first two 90° pulses with 11 pulses, and uses a 1331 read pulse. The exact experimental parameters are given in the figure legends.

**Melting Spectra.** Melting data were collected on a GN 500 spectrometer using the 1331 experiment (appendix B) [Hore, P. J. (1983)] with the carrier frequency set at the resonance of water. The temperature increased in ten degree increments from 20° to 40°C, lowered to 35°C, and decreased in ten degree increments to 15°C. The final temperature point was collected at 10°C. The sample equilibrated for 5 min. at each temperature. 1024 scans of 8K data points using a recycle delay of 0.5 sec. were averaged for each spectrum. The raw data were left-shifted one point and stored. The left-shifted FID was subtracted from the raw FID, and the result appodized using a DM=8, Fourier transformed, and phased using GN 500 software. The transformed spectra were transferred to a MicroVax workstation.
and plotted using FTNMR software [Hare Research].

Results, NOESY Spectroscopy, No MgCl₂. Figure 8.3 shows the sequences and imino proton spectra of RNA4 and RNA6 taken at 20°C. The RNA4 assignments are labelled on the sequence. RNA6 contains one fewer GU wobble base pair than RNA4 (stem II). The imino proton of this GC base pair is known to resonate in the 12.5-12.0 ppm region of the spectrum (results from RNA1 and RNA3). Accordingly, the RNA6 spectrum should contain one additional resonance in this region of the spectrum compared to that of RNA4, and two fewer resonances between 12.0 and 10.0 ppm (R and T) due to the loss of the GU wobble pair. The loss of R and T is clearly confirmed in fig. 8.3. Although the 12.5-12.0 ppm region of the spectrum is very crowded, an additional resonance can be confirmed by comparing the intensity of the M/N resonance with those of O and P. In RNA4 resonances O, P, and N have comparable integrated intensities, whereas in the RNA6 spectrum the resonance labelled M, M", N has a much larger integrated intensity that peaks O and P, verifying the existence of an additional resonance in the RNA6 spectrum. There are additional differences between the spectra as well. Resonance L has completely disappeared, and A, B, D, E, and F have broadened significantly. Because resonance K has also broadened this data suggests that stem I (E-D-K-N-B) of RNA6 does not exhibit the additional motion for RNA4 in the same buffer conditions. The RNA6 spectrum is more consistent with that of RNA4 in the presence of 10 mM MgCl₂, the independent motion of stem I was lost upon addition of MgCl₂ to the sample. The broadness of resonances A, E and F and the disappearance of L are also consistent with the 10 mM MgCl₂ RNA4 data as well as the low temperature RNA1 data, as the helical junction was thought to undergo a conformation change as the structure
Figure 8.3 Sequences and 500 MHz imino proton spectra of RNA4 and RNA6 recorded at 20°C. The experimental conditions for RNA4 are given in the legend of figure 7.3, those of RNA6 in the legend of figure 8.7.
folded into a true cleavage conformation. Resonance I has been labelled in fig. 8.3 such that it is degenerate with peak J. Although it is not readily apparent in this 1D spectrum, NOE data to be presented later will confirm its position. The fact that I has changed chemical shift in RNA6 compared to RNA4 is due to the replacement of its 3'-guanosine neighbor with an adenosine. If resonance I were sharp, resonance J should have increased in intensity compared to K due to its degeneracy with I. Because resonance J has not increased in intensity it can be presumed that I is fairly broad at 20°C. This is consistent with the broadness observed for the other junction resonances, as discussed above.

The NOESY spectrum taken in H₂O at 20°C is shown in figures 8.4, 8.5, and 8.6. The imino proton assignments are divided between two plots to reduce crowding. The diagonal peaks are labelled according to the convention used in figure 8.3. Figure 8.4 shows the connectivities which determine stem III. The intense cross peak between resonances Q and S is indicative of a GU wobble base pair, and as there is only one GU wobble pair in the sequence, it must be that of stem III. The J-S cross peak is very weak. For RNA4 this NOE was fairly intense. This indicates that the GU wobble pair is less stable in RNA6 than in RNA4. Apparently the native structure destabilizes base pairs in addition to those directly at the junction of the three helices. This is consistent with the RNA4 data taken in 10 mM MgCl₂. Comparison of the RNA4 melting data with and without 10 mM MgCl₂ showed that resonances Q and S have slightly broader lines in the spectra acquired in MgCl₂. The remainder of the stem III connectivities have already been discussed for RNA1, RNA3, and RNA4, and are indicated on the figure.

Figure 8.5 shows the connectivity D-K-N. There are no cross peaks connecting resonances N and B or E and D. One dimensional NOE difference spectra recorded at
Figure 8.4  Imino-imino region of H₂O NOESY showing the stem III connectivities of RNA6. The diagonal peaks are labelled with the convention used in figure 8.3. The sample was ca. 1 mM RNA6, 20 mM phosphate buffer, pH 7.0, 0.2 mM EDTA, 0.1 M NaCl, and 0.02% NaN₃. 192 scans of 2K data points were averaged using a recycle delay of 0.7 sec for each of 485 t₁ increments. The offset delay was 110 μsec., the mixing time 100 msec.
Figure 8.5 Imino-imino region of H₂O NOESY showing the stem I D-K-N connectivity of RNA6. The sample conditions and experimental parameters are given in the legend of figure 8.4.
Figure 8.6 Imino-aromatic region of RNA6 H_2O NOESY. The sample conditions and experimental parameters are given in the legend of figure 8.4. The imino proton resonance chemical shifts are labelled across the top of the plot according to the convention used in figure 8.3.
20°C also do not show these NOE's (data not shown). However, peaks D, E, and B are relatively broad, and it is not surprising that NOEs involving these resonances are not observed.

It is interesting that no cross peak between resonances A and M was observed at 20°C, as this connectivity was clearly established for RNA1, RNA3, and RNA4. It was though that resonance A was being attenuated by the excitation profile of the 13NSY experiment, but the 1D data also did not show NOE's between these resonances (data not shown). This connectivity had to be determined at 15°C (data to be shown in a later section). Resonance M cannot be connected to its GC neighbor (M') due to the near chemical shift degeneracy of these peaks.

The imino-aromatic region of the NOESY spectrum was used to obtain a more accurate count of the number of base pairs which resonate between 12.5 and 12.0 ppm (figure 8.6). Although an intensity argument has been used to verify an additional resonance in this region of the spectrum, an exact count was desired. Resonances A, B, C, and D manifest very intense imino-aromatic crosspeaks verifying the fact that these resonances are due to AU base pairs. Resonances E and F do not exhibit a sharp imino-aromatic crosspeaks. Presumably this is due to the broadness of these resonances. The region of the spectrum containing the imino-amino crosspeaks of resonances M, M', M'', N, O, and P is sufficiently crowded that these resonances are very difficult to count. However, there do not appear to be more than twelve crosspeaks in this region, which is consistent with the six base pairs labelled across the top of the figure.

A NOESY spectrum was also collected in H₂O at 15°C. The assignments obtained for stems I, II, and III are shown in figures 8.7, 8.8, and 8.9 respectively. The imino-aromatic region of the spectrum is shown in figure 8.10. This spectrum
Figure 8.7 Imino-imino region of the RNA6 H2O NOESY taken at 15°C showing the assignments of stem I. The sample conditions and spectral parameters are given in the legend of figure 8.4.
Figure 8.8  Imino-imino region of the RNA6 H₂O NOESY taken at 15°C showing the assignments of stem II. The sample conditions and spectral parameters are given in the legend of figure 8.4.
Figure 8.9 Imino-imino region of the RNA6 H₂O NOESY taken at 15°C showing the assignments of stem III. The sample conditions and spectral parameters are given in the legend of figure 8.4.
Figure 8.10 Imino-aromatic region of the RNA6 H2O NOESY taken at 15°C. The resonances are labelled across the top of the plot using the convention of figure 8.3. The sample conditions and spectral parameters are given in the legend of figure 8.4. Assignments obtained from the 15°C data were those of stem II (figure 8.8).
clearly contains a crosspeak between resonances A and M, and in addition there is a cross peak between resonance A and what appears to be J. It is this crosspeak that identifies resonance I as being degenerate with J, as no other assignment is consistent with the RNA sequence. The change in chemical shift of I is due to the replacement of its 3'-guanosine neighbor in RNA4 to a 3'-adenine in RNA6.

Even at 15°C there is no data connecting resonance M to its GC (M") neighbor. However, this peak can be unambiguously assigned by indirect evidence. First, the stem II sequence of RNA3 is identical to that of RNA6, and both of these spectra contain resonance M". Secondly, the fact that it is resonance M" which appeared in the RNA6 spectrum upon the replacement of the GU wobble with a W-C GC pair means that it must be M" which neighbors resonance M. Due to the near degeneracy in chemical shift of M and M" no cross peak connecting these resonance can be observed. In addition, there are no other cross peaks involving resonance M", indicating that it is indeed a terminal base pair.

The imino-aromatic region of the spectrum (figure 8.10) is similar to that obtained at 20°C, except that this spectrum contains a cross peak involving resonance U. Based on its chemical shift and lack of NOE connectivities, resonance U has been previously assigned to a uridine in a hairpin loop region of the molecule. The observation of this imino-amino cross peak does not refute this assignment, but unfortunately it does not help assign U to a particular base in the sequence.

Results, Melting Data, No MgCl₂. Figure 8.11 shows the imino proton spectrum of RNA6 as a function of temperature. As the temperature is decreased from 20°C to 10°C resonances A, B, and C become degenerate, as do resonances I/J and K. The entire spectrum broadens somewhat, but in this case, all of the resonances are
equally affected. In the spectra of the previous RNA's those resonances assigned (or speculated to be) to junction base pairs were more affected than the rest of the spectrum. The fact that all of the resonances broaden equally with decreasing temperature for RNA6 is not surprising, as the junction resonances are broad or missing even at 20°C, as previously discussed. If these pairs were exchanging rapidly with water protons due to end-fraying type effects, they should "grow in" with decreasing temperature. The fact that they do not implies that these residues are oriented in such a way that they cannot easily pair. In other words, the hammerhead domain is more "open" than the folding model predicts. The bases in the junction do not easily form pairs as the model implies.

Increasing the temperature to 30°C results in a broadening of resonances B, D, E, Q, and S. Surprisingly, resonances A and K show signs of becoming sharper as the temperature increases, and further sharpen as the temperature increases to 40°C. If all of the resonances sharpened equally a rational explanation would be that the sharpening is due to a decrease in the rotational correlation time of the RNA with increasing temperature. However, the fact that A and K are independently sharp implies that increasing the temperature disrupts tertiary interactions, which in turn stabilizes these secondary base pairs. A stronger argument could be made if the other resonances of stems I and II also became increasingly sharp with increasing temperature, but most of these base pairs have begun to exchange broaden by 30°C. Only the GC base pair assigned to resonance N is fairly stable at 30°C, but it is in too crowded a region of the spectrum for linewidth analysis. Nonetheless, the independent motion observed for stem I of RNA4 does corroborate this hypothesis.
Figure 8.11 500 MHz imino proton spectrum of RNA6 as a function of temperature. The sample was 1 mM RNA6, 20 mM phosphate buffer, pH 7.2, 0.1 M NaCl, 0.2 mM EDTA, and 0.02% NaN₃. 1024 scans of 8K data points using a recycle delay of 0.5 sec. were averaged for each temperature increment. The offset delay was 125 μsec. The temperature at which each spectrum was recorded is given in the right margin in degrees Celsius.
Comparison of the RNA4 melting data (figure 4.7) with that of RNA6 in the absence of MgCl₂ reveals very few differences. The general order of base pair melting as well as the resonances which remain stable at 40°C is essentially the same for both RNA's. The major difference between RNA4 and RNA6 is the melting temperature of resonances A, Q, and S. In RNA6 resonance A is fairly stable at 40°C, whereas in RNA4 it has completely melted out at this temperature. The opposite is observed for resonances Q and S; they are more stable in RNA4 than in RNA6. The decreased stability of A observed for RNA4 compare to RNA6 is most likely due to the neighboring GU wobble (R/T), as GU wobbles are less stable than their Watson-Crick GC counterparts. Because the melting of a double helix is a cooperative process, the instability of one base pair then affects the stability of its neighbors. This hypothesis can be verified by comparing the melting temperature for resonance A of RNA3 with that of RNA6 as these two RNA's share a common stem II sequence. For RNA3 resonances A, M, M*, and I are still fairly stable at 40°C (figure 6.7). The fact that the RNA3 spectra were recorded at a lower ionic strength than the RNA4 spectra further strengthens the argument that the instability of resonance A for RNA4 is due to its neighboring GU wobble base pair.

The behavior of Q and S cannot be similarly explained. The stem III sequence of RNA4 is identical to that of RNA6. The NOESY spectrum contained very weak J-Q and J-S crosspeaks, indicating that the imino protons of the GU wobble were exchanging with fairly rapidly with water protons even at 20°C. The low melting temperature of Q and S corroborates this NOE data. As before, it can only be speculated that this instability is due to an "open" junction structure.
Results, NOE Difference Spectroscopy, 10 mM MgCl₂. Figure 8.12 shows the 500 MHz imino proton spectra of RNA6 in the absence and presence of 10 mM MgCl₂ taken at 20°C. There are some obvious differences in the low field region of the spectrum. Resonance A appears to have broadened or moved under resonance B/C, an additional resonance appeared as a shoulder of peak C, and an additional resonance appeared between peaks D and E. In addition, 10 mM MgCl₂ increased the spectral resolution of the region between 12.5 and 12.0 ppm, similar to the results obtained for RNA1 and RNA4.

Figure 8.12 also shows the 500 MHz imino proton spectra of RNA4 and RNA6 in 10 mM MgCl₂ taken at 25°C. The studies with RNA4 seemed to indicate that the cleaved form of the RNA may not approximate a true cleavage conformation because no additional peaks appeared upon addition of 10 mM MgCl₂. The similarity between the RNA4 and RNA6 spectra show that this is not the case. The RNA6 spectrum contains one more resonance between 12.5 and 12.0 ppm, accounting for the replacement of the GU wobble of stem II with a Watson-Crick GC as previously. The only other major difference between the two spectra is the loss of two resonances between 12.0 and 10.0 ppm, again due to the replacement of the stem II GU wobble base pair. Aside from these sequence specific differences, the two spectra overlay almost exactly, demonstrating the relevance of the RNA4 studies.

NOE difference spectra recorded at 35°C in 10 mM MgCl₂ result in the same connectivities already obtained in the absence of MgCl₂ (data not shown). NOE difference spectra of the 14.5-13.5 ppm region of the spectra were collected at 15°C. These data did not result in any additional assignments (data not shown).

Figure 8.13 shows the temperature dependence of RNA6 in 10 mM MgCl₂. The order of base pair melting is the same as that observed in the absence of MgCl₂.
Figure 8.12 Comparison of the 500 MHz imino proton spectra of (A) RNA6, 0mM MgCl₂ (B) RNA6, 10 mM MgCl₂ (C) RNA4, 20 mM MgCl₂. All spectra were recorded at 20°C. Sample conditions and spectral parameters are given in the legends of figures 8.7, 8.9, and 7.3 respectively.
Figure 8.13 500 MHz imino proton spectrum of RNA6 plotted as a function of temperature. The sample was 1 mM RNA6, 20 mM phosphate buffer, pH 7.2, 0.1 M NaCl, 10 mM MgCl₂, 0.2 mM EDTA, and 0.02% NaN₃. 1024 scans of 8K data points using a recycle delay of 0.5 sec. were averaged for each temperature increment. The offset delay was 125 μsec. The temperature at which each spectrum was recorded is denoted in the right margin in degrees Celsius.
Conclusions. Figure 8.14 shows the sequence and 500 MHz imino proton spectrum of RNA6 taken at 20°C labelled with the assignments obtained from NOESY spectroscopy in upper case letters, and those inferred from the RNA4 assignments in lower case letters. The spectra of RNA1 and RNA4 indicated that the base pairs in the junction of the three helices become stressed as the RNA folds into a relevant cleavage conformation. The fact that resonances corresponding to these base pairs are absent in spectra of uncleaved RNA6 both in the absence and presence of MgCl₂ substantiates this hypothesis. Decreasing the temperature does not result in stabilization of these pairs, further supporting the theory that this effect is due to the tertiary folding of the RNA. Although the imino proton spectrum of RNA6 recorded 10 mM MgCl₂ differs from that recorded in the absence of MgCl₂, the confirmation (and assignment) of additional resonances could not be ascertained. The order of base pair melting of RNA6 in 10 mM MgCl₂ is the same as that observed in the absence of magnesium, and is consistent with the assignments of figure 8.10. For reasons to be discussed below, the relevance of the 10 mM MgCl₂ RNA6 spectra were questionable. A more thorough discussion of the spectroscopic results will be given after proving their relevance.
Figure 8.14 Sequence and 500 MHz imino proton spectrum of RNA6 taken at 20°C. The assignments obtained from NOE difference and NOESY spectroscopy are labelled in upper case. The sequence is shown with the helical junction in a more "open" conformation, as this models the NMR data more appropriately. The assignments which were inferred from the RNA4 data are denoted in lower case letters. The sample conditions and experimental parameters are given in the legend of figure 8.7.
8.4 Determination of the Cleavage Conditions of RNA6 for NMR

After collecting the NOE difference spectra of RNA6 in 10 mM MgCl₂ approximately 2 μg was removed from the NMR sample and analyzed for cleavage on a 8x10x0.07 cm 10% polyacrylamide gel containing 7 M urea. Surprisingly, none of the RNA cleaved (data not shown). This result seemed very strange since cleavage was observed for RNA6 in transcription buffer. One possible explanation involved the Mg²⁺/RNA ratio. In the NMR sample this ratio was 10 whereas in a 20 μL transcription reaction (yielding approx. 5 μg of RNA) this ratio is about 550. Another possible explanation was that the phosphate buffer was competing with the RNA for the Mg²⁺, and cleavage would occur in this buffer only if the concentration of MgCl₂ exceeded that of the phosphate buffer. If this were true, RNA transcribed in phosphate buffer instead of tris buffer would not cleave, which would be very advantageous for large scale purification of uncleaved RNA6. The only other obvious difference between the NMR sample and the cleavage reactions was the pH. A series of analytical cleavage studies which mimicked the NMR sample conditions was done to determine which, if any, of these hypotheses was correct, and also to determine the sample conditions necessary to study cleaving RNA6.

Preparation of the RNA. 1 mg of pure, uncleaved RNA6 was dissolved in 1 mL of 10 μM phosphate buffer, pH 7.0, and dialyzed against 2 L of the same buffer at 4°C for 24 hours. The RNA solution was removed from the dialysis tubing and quantitated by UV spectrophotometry [Maniatis et al. (1982) p. 468]. Pure, uncleaved RNA6 was then dispensed into several eppendorf tubes, 2 μg per tube, and dried under vacuum (Savant Speed-Vac). The unused stock solution was also dried, and the RNA stored dry at -20°C.
Determination of Buffer and pH Dependence of RNA6 Cleavage. To determine if RNA6 cleavage was buffer or pH dependent, several buffers were prepared which differed slightly in composition and pH from transcription buffer. An additional buffer was prepared which mimicked the NMR sample in that all of the concentration ratios were identical to those of the NMR sample (table 8.1).

Five 2 μg aliquots of pure, uncleaved RNA6 were dissolved in 4 μL of H₂O. 4 μL of one of the buffers from table 8.1 was added to each aliquot and the RNA incubated at 25°C (actual reaction conditions are given in table 8.2). At 4, 21, and 46 hours 2 μL was removed from each tube, 2 μL quench buffer (3 mM EDTA, 0.6 M sodium acetate, pH 5.2) added, and the RNA precipitated with 100 μL ethanol at -20°C. The early time points were stored in ethanol for the duration of the experiment. The RNA was pelleted by centrifugation in an eppendorf centrifuge at 4°C for 10 min., the pellet washed twice with 70% ethanol, and dried by lyophilization. The pellets were dissolved in 5 μL 1X denaturing gel loading buffer (20% formamide, 0.05% xylene cyanol, 0.05% bromophenol blue), heated to 90°C for 1 min. and loaded onto a 8x10x0.07 cm 10% polyacrylamide gel containing 7 M urea. The control lane contained 0.5 μg pure, uncleaved RNA6 dissolved in 5 μL 1X denaturing gel loading buffer which had been heat denatured as above. Electrophoresis was carried out at 600 Volts for 20 min. against 1X TBE buffer [Maniatis et al. (1982) p. 156]. The gel was stained for 2 min. with 0.02% toluidine blue-O (Sigma), destained with H₂O, and photographed.

Results. Figure 8.15 shows the photograph of the gel from the above experiment. The upper most band is uncleaved RNA6. The cleavage products and bromophenol blue dye marker are labelled on the figure. Only lanes 15, 16, and 17
## Table 8.1 Buffers for Analytical Cleavage Reactions

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E$^1$</th>
<th>F$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[phosphate], mM</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>-</td>
<td>0.44</td>
</tr>
<tr>
<td>[tris·HCl], mM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>40</td>
<td>-</td>
</tr>
<tr>
<td>[MgCl$_2$], mM</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>1.1</td>
</tr>
<tr>
<td>[spermidine], mM</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>[NaCl], mM</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>10</td>
<td>2.2</td>
</tr>
<tr>
<td>pH</td>
<td>7.2</td>
<td>8.0</td>
<td>7.2</td>
<td>8.0</td>
<td>8.2</td>
<td>7.0</td>
</tr>
</tbody>
</table>

$^1$ E = transcription buffer  
$^2$ F = NMR sample concentration ratios
### Table 8.2 Analytical Cleavage Reaction Conditions

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E₁</th>
<th>F₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>[RNA6], mM</td>
<td>0.022</td>
<td>0.022</td>
<td>0.022</td>
<td>0.022</td>
<td>0.022</td>
<td>0.022</td>
</tr>
<tr>
<td>[phosphate], mM</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>-</td>
<td>0.22</td>
</tr>
<tr>
<td>[tris·HCl], mM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>[MgCl₂], mM</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.55</td>
</tr>
<tr>
<td>[spermidine], mM</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>[NaCl], mM</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>5</td>
<td>1.1</td>
</tr>
<tr>
<td>pH</td>
<td>7.2</td>
<td>8.0</td>
<td>7.2</td>
<td>8.0</td>
<td>8.2</td>
<td>7.0</td>
</tr>
</tbody>
</table>

1 E = transcription buffer
2 F = NMR sample concentration ratios
show appreciable cleavage. These lanes contain RNA6 which was incubated in buffer E (transcription buffer). The cleavage half-life can be estimated to be slightly longer than 21 hours, whereas the half life was measured to be approximately 12 hours at 37°C. A 46 hour incubation at 25°C resulted in virtually complete cleavage, which is comparable to the 44 hours required at 37°C (figure 8.2). In lane 11 the band corresponding to uncleaved RNA6 is very faint, but there is no band corresponding to the large fragment of cleaved RNA6. This faint uncleaved band is therefore due to the quantity of RNA loaded onto the gel rather than cleavage. Lanes 18, 19, and 20, which represent NMR sample conditions do not show appreciable cleavage even after a 46 hour incubation. These results agree with those obtained for the NMR sample, as no cleavage was observed at the end of a 40 hour NMR experiment.

If only the Mg\(^{2+}\)/RNA6 ratio were important for cleavage of RNA6, RNA incubated in buffer D (lanes 12-14) should have cleaved, as the only difference between buffers D and E is that D is buffered with phosphate and E with tris·HCl. The fact that cleavage is not observed after incubation in buffer D indicates that it is the Mg\(^{2+}\)/phosphate ratio and not the Mg\(^{2+}\)/RNA6 ratio which is important for cleavage.

**Determination of MgCl\(_2\) Concentration Required for Phosphate Buffered Cleavage of RNA6.** Additional analytical cleavage reactions were done to further test the hypothesis that the Mg\(^{2+}\)/phosphate ratio is important for RNA6 cleavage. Six 2 μg aliquots of pure, uncleaved RNA6 were dissolved in 2 μL H\(_2\)O. 2 μL buffer A, C, E, G, H, or I was added to each tube. The final reaction conditions are given in table 8.3. After a 49 hour incubation at 25°C 3 μL 5X denaturing gel loading dye (80%
formamide. 0.25% xylene cyanol, 0.25% bromophenol blue) was added to each tube. A control was prepared by dissolving 2 μg of pure, uncleaved RNA6 in 4 μL TE, pH 7.4 [Maniatis et al. (1982) p. 448] and 3 μL 5X denaturing gel loading dye. All of the tubes were heated to 90°C for 1 min. and loaded onto a 8x10x0.07 cm 10% polyacrylamide gel containing 7 M urea (3 μL per well). Electrophoresis conditions were identical to those described above.

Results. Figure 8.16 shows the results of these incubations. After a 49 hour incubation at 25°C only lanes 4 and 6 contained cleaved RNA6. Lane 4 contained RNA6 incubated in transcription buffer and was the "cleavage control" lane. Lane 6 contained RNA6 incubated in buffer H. The fact that RNA6 cleaved in buffer H but not buffer C demonstrates that cleavage occurs in phosphate buffer only if the concentration of magnesium ions exceeds the concentration of the buffer. In addition, the fact that lane 5 (buffer G) contained only uncleaved RNA6 demonstrates that spermidine is also a requisite for cleavage, as the only difference between buffers G and H was spermidine. Lane 7 contained RNA6 which was incubated in buffer I. This reaction was analogous to an NMR sample containing 2 mM RNA6, 20 mM tris·HCl, pH 7.6, 25 mM MgCl2, and 0.1 M NaCl. This reaction was done to determine if merely exchanging the NMR sample buffer to tris·HCl would result in cleavage of the RNA. The fact that no cleavage occurred after a 49 hour incubation at 25°C (lane 7) proves this to not be the case, and implies that spermidine is required for cleavage in tris buffer as well as phosphate buffer.

Determination of the Spermidine Concentration Required to Cleave RNA6. The previous studies indicated that cleavage of RNA6 occurs in phosphate buffered
Table 8.3 Conditions for Analytical Cleavage Reactions

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>E$^1$</th>
<th>G</th>
<th>H</th>
<th>I$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[RNA6], mM</td>
<td>.022</td>
<td>.022</td>
<td>.022</td>
<td>.022</td>
<td>.022</td>
<td>.022</td>
</tr>
<tr>
<td>[phosphate], mM</td>
<td>20</td>
<td>20</td>
<td>-</td>
<td>0.5</td>
<td>0.5</td>
<td>-</td>
</tr>
<tr>
<td>[tris·HCl], mM</td>
<td>-</td>
<td>-</td>
<td>20</td>
<td>-</td>
<td>-</td>
<td>0.22</td>
</tr>
<tr>
<td>[MgCl$_2$], mM</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.28</td>
</tr>
<tr>
<td>[spermidine], mM</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>[NaCl], mM</td>
<td>60</td>
<td>60</td>
<td>5</td>
<td>60</td>
<td>60</td>
<td>1.1</td>
</tr>
<tr>
<td>pH</td>
<td>7.2</td>
<td>7.2</td>
<td>8.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.6</td>
</tr>
</tbody>
</table>

1 E= transcription buffer

2 I= Equivalent to an NMR sample containing 2 mM RNA6, 20 mM tris·HCl, pH 7.6, 25 mM MgCl$_2$, and 0.1 M NaCl
Figure 8.16 Determination of the MgCl₂ Concentration Necessary for RNA6 Cleavage in Phosphate Buffer. The experimental conditions are given in the text.

Lanes: 1. control; 2. buffer A; 3. buffer C; 4. buffer E; 5. buffer G; 6. buffer H; 7. buffer I.

UC = uncleaved RNA; LF = large fragment; SF = small fragment
solutions only when the concentration of MgCl$_2$ exceeds that of the phosphate. Although this was not true for tris·HCl buffered RNA6 cleavage, cleavage was observed in both buffers only when both MgCl$_2$ and spermidine were present. It was therefore necessary to determine the spermidine concentration dependence of RNA6 cleavage in these buffers. For these reactions, buffers were prepared which mimicked the concentration ratios of an NMR sample containing 2 mM RNA6, 20 mM buffer, pH 7.5, 25 mM MgCl$_2$, and 0.1 mM NaCl. 2 &mu;g aliquots of pure, uncleaved RNA6 were dissolved in 5 &mu;L of the desired buffer (table 8.4) and incubated at 25°C for 50 hours. At the end of the incubation period the reactions were quenched by addition of 5 &mu;L quench buffer (0.5 mM EDTA, 0.05% xylene cyanol, 0.05% bromophenol blue, 10% formamide), and the RNA denatured by heating to 90°C for 1 min. The reactions were placed immediately into an ice water bath and 2 &mu;L from each tube loaded onto a 8x10x0.07 cm 10% polyacrylamide gel containing 7 M urea. Electrophoresis was carried out at 600 Volts for 20 min. against 1X TBE buffer. The gel was stained with 0.02% toluidine blue-O (Sigma), destained with H$_2$O, and photographed. The "cleavage control" reaction was 2 &mu;g RNA6 dissolved in 5 &mu;L 1X transcription buffer (table 8.3) and incubated as described above. A non-cleavage control was prepared just prior to electrophoresis by dissolving 2 &mu;g of pure, uncleaved RNA6 in 7 &mu;L TE, pH 7.4 containing 0.05% xylene cyanol and 0.05% bromophenol, blue and treating as described for the other samples.

Results, Spermidine Cleavage Reactions. Figure 8.17 shows the gel of the cleavage reactions with spermidine. Bands corresponding to uncleaved RNA6 and the large fragment of cleaved RNA6 are labelled on the gel. Lane 1 is the non-cleavage control, lane 2 the cleavage control. The only lanes which exhibited cleavage were
Table 8.4 Cleavage Reactions with Spermidine

<table>
<thead>
<tr>
<th></th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>[RNA6], mM</td>
<td>.018</td>
<td>.018</td>
<td>.018</td>
<td>.018</td>
<td>.018</td>
<td>.018</td>
<td>.018</td>
<td>.018</td>
</tr>
<tr>
<td>[phosphate], mM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
</tr>
<tr>
<td>[tris·HCl], mM</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[MgCl$_2$], mM</td>
<td>.225</td>
<td>.225</td>
<td>.225</td>
<td>-</td>
<td>.225</td>
<td>.225</td>
<td>.225</td>
<td>-</td>
</tr>
<tr>
<td>[NaCl], mM</td>
<td>.9</td>
<td>.9</td>
<td>.9</td>
<td>.9</td>
<td>.9</td>
<td>.9</td>
<td>.9</td>
<td>.9</td>
</tr>
<tr>
<td>[spermidine]</td>
<td>.009</td>
<td>.018</td>
<td>.036</td>
<td>1</td>
<td>.009</td>
<td>.018</td>
<td>.036</td>
<td>1</td>
</tr>
<tr>
<td>pH</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
</tr>
</tbody>
</table>
2, 5, and 9. Lane 2 was expected to show cleavage, as transcription buffer was used as the cleavage control reaction. Lane 5 contained RNA6 incubated in tris·HCl buffer with a MgCl₂/RNA molar ratio of 12.5 and a spermidine/RNA molar ratio of 2. Lane 9 contained RNA6 incubated in phosphate buffer with the same MgCl₂/RNA and spermidine/RNA ratios as lane 5. No cleavage was observed for RNA6 incubated in either tris or phosphate buffer with a spermidine/RNA ratio of 0.5 (lanes 3 and 7) or 1.0 (lanes 4 and 8). The requirement for MgCl₂ was demonstrated for both buffers, as RNA6 incubated in 1 mM spermidine in the absence of magnesium ions resulted in no cleavage (lanes 6 and 10).

Conclusion and Discussion. The observation that the RNA6 NMR sample did not cleave in a 10-fold excess of MgCl₂ after a 48 hour NMR experiment prompted an investigation into the conditions required to cleave the RNA. RNA6 was found to cleave in phosphate buffer only when the MgCl₂ concentration exceeded that of the phosphate. Although this was not observed for cleavage in tris·HCl, a spermidine/RNA ratio of >2 was required for cleavage in either buffer. These results indicate that RNA6 has a poor magnesium binding site compared to the wild-type hammerhead domain sequence. So poor, in fact, that interactions between phosphate buffer and Mg²⁺ compete with the RNA6 magnesium binding site. Since the replacement of a conserved G with an A in RNA4 converts a fast-cleaving RNA4 into slowly-cleaving RNA6, it is not unreasonable to speculate that this residue plays a key role in the binding of Mg²⁺. The finding that RNA6 cleavage is spermidine concentration dependent seems to indicate that Mg²⁺ binding forces several phosphate groups within close proximity of one another. Because RNA6 has a poor Mg²⁺ binding site, additional positive charges are required to overcome the
Figure 8.17 Determination of the Spermidine Concentration Required for RNA6 Cleavage. The experimental conditions are given in the text. Lanes: 1. non-cleavage control; 2. cleavage control; 3. buffer J; 4. buffer K; 5. buffer L; 6. buffer M; 7. buffer N; 8. buffer O; 9. buffer P; 10. buffer Q.

UC = uncleaved RNA6; LF = large fragment; SF = small fragment.
repulsive interphosphate forces. Spermidine supplies the additional charges. Presumably the wild-type sequence binds the Mg$^{2+}$ sufficiently strongly that spermidine is not required to overcome the interphosphate repulsion.

Strange as these results seem, the requirement that the MgCl$_2$ concentration exceed that of the phosphate buffer for cleavage was actually an advantage. Transcription reactions could be carried out in phosphate buffered solutions rather than tris.HCl buffered solutions with no cleavage of the RNA. This proved to be the method of choice for isolating pure, uncleaved RNA$_6$ in high yield (chapter 3).

8.5 NMR Spectroscopy of Cleaving RNA$_6$

Although the results of the analytical cleavage reactions proved to be advantageous from an RNA production standpoint, they were disadvantageous from a spectroscopic standpoint. First, the information already obtained for RNA$_6$ in phosphate buffered solution containing 25 mM MgCl$_2$ may be irrelevant due to the fact that RNA$_6$ does not cleave in these conditions. Second, spermidine-dependent RNA$_6$ cleavage translates into the addition of a two-fold molar excess of a proton rich compound to the NMR sample. Although these extra protons do not resonate in the low field region, and thus will not affect imino proton spectroscopy, they will affect future NMR studies of the non-exchangeable protons. An additional problem is introduced when one considers that RNA$_6$ has a tendency to degrade in cleavage buffer.

The addition of spermidine to the NMR sample can be circumvented if it can be shown that RNA$_6$ adopts similar conformations in when a two-fold molar excess of spermidine is added to the NMR sample. Figure 8.18 shows the results of such an experiment. The spectra are nearly identical, suggesting that the addition of
Figure 8.18 Spermidine titration of RNA6. The sample was 1 mM RNA6, 20 mM phosphate, pH 7.2, 0.1 M NaCl, 25 mM MgCl₂, 0.2 mM EDTA, and 0.02% NaN₃. The spermidine concentration is indicated in the right-hand margin in millimolar. The "peak" labelled with an "X" is a spectrometer glitch. All spectra were recorded at 25°C.
spermidine has not significantly changed the conformation of the RNA. Spectra recorded at other temperatures were also identical to those taken without spermidine (data not shown). A spectrum was similarly recorded in tris·HCl buffer, and again no differences were observed between the imino proton spectra (data not shown). These results indicate that the NMR sample need not contain spermidine, as no conformational change is observed. Therefore the results summarized in figure 8.14 can be presumed to be of the appropriate structure, and future NMR studies can be done in the presence of MgCl$_2$ alone, without sacrificing the sample to cleavage.

8.6 Spectroscopy of Cleaved RNA6

One final RNA6 spectrum was desired. It was argued for RNA1 and RNA4 that those resonances which severely broadened upon the addition of MgCl$_2$ to the sample did so because of the tertiary folding of the RNA. Spectra of uncleaved RNA6 confirmed this hypothesis, as these resonances were either broad or missing from its NMR spectrum. If indeed cleavage allows these bases to pair, the spectrum of cleaved RNA6 should clearly contain resonance L. In addition, resonances A, D, E, and F should exhibit linewidths similar to those observed for RNA4.

Several attempts were made to purify 100% cleaved RNA6, but each time the RNA degraded. However, a sample of partially cleaved RNA6 was obtained. The percent cleavage was estimated by denaturing polyacrylamide gel electrophoresis to be approximately 80% (data not shown). Figure 8.19 shows the spectra of 80% cleaved RNA6 (CL-RNA6) and RNA4 at 15°C, figure 8.20 the spectra of uncleaved RNA6 (UC-RNA6) and CL-RNA6 at the same temperature. The buffer conditions are indicated in the figure legend. The peak labelled with an "X" on the CL-RNA6
Figure 8.19 (A) Approximately 80% cleaved RNA6. The sample was 1 mM RNA6, 20 mM phosphate buffer, pH 7.2, 0.1 M NaCl, 0.2 mM EDTA, and 0.02% NaN3. 1024 scans of 8K data points were averaged using a recycle delay of 0.5 sec.

(B) RNA4. Sample conditions and spectral parameters are given in the legend of figure 7.3. Both spectra were acquired at 15°C.
Figure 8.20 (A) 500 MHz imino proton spectrum of CL-RNA6. The sample was 1 mM RNA6, 20 mM phosphate buffer, pH 0.0, 0.1 M NaCl, 0.2 mM EDTA, and 0.02% NaN3. 1024 scans of 8K data points using a recycle delay of 0.5 sec. were averaged. (B) 500 MHz imino proton spectrum of UC-RNA6. Sample conditions and spectral parameters are given in the legend of figure 8.11. Both spectra were acquired at 15°C.
spectrum is a spectrometer glitch. Aside from the sequence-dependent spectral changes previously discussed, there is little difference between the CL-RNA6 spectrum and that of RNA4. Resonance A has sharpened and moved down-field from its position in the UC-RNA6 spectrum to a chemical shift identical to that found in RNA4. In addition resonances D, E, and F have sharpened significantly. Resonance L is also clearly visible in the CL-RNA6 spectrum. It was postulated that cleavage would result in the "growing in" of this resonance, and clearly this is the case. The chemical shift of L is almost identical for UC-RNA6 and RNA4. This is expected, because aside from any possible tertiary interactions, the stem III sequences of these two RNA's are identical. There is a slight shoulder on the left-hand side of resonance B in the spectrum of CL-RNA6. This is most likely due to the presence of ca. 20% uncleaved RNA in the sample, as this shoulder aligns with resonance A in the UC-RNA6 spectrum. The shoulders on resonances D, E, and F appear to be caused by the UC-RNA6 "contaminant" as well.

The most striking difference between the three spectra involves resonance I. Resonance I is not expected to resonate at the same chemical shift in RNA4 and RNA6 because of nearest neighbor differences in sequence. In UC-RNA6, resonance I was degenerate with J, and postulated to be somewhat broad. In the spectrum of CL-RNA6, resonance I is not only as relatively sharp, but has moved up-field such that it is nearly degenerate with resonance L. Clearly I is not a shoulder of L due to the presence of 20% UC-RNA6 in the sample. The spectrum of 100% UC-RNA6 does not have a resonance at this chemical shift, and L and I are of equal intensity, which would not be the case were I truly a shoulder of L. It was speculated that the resonance broadening observed for the junction base pairs for RNA1 and RNA4 in the presence of MgCl₂ and for UC-RNA6 was due to a conformational change in the
junction. The fact that I changes in chemical shift upon cleavage of RNA6 strongly substantiates this argument, as a change in chemical shift is associated with a change in electronic environment. Clearly then, the environment of resonance I has changed upon cleavage of the domain. This environmental change must be due to a conformational rearrangement of the domain. Because resonance I is affected, the change must be in the vicinity of this resonance, however, the nature of this structural change cannot be ascertained from this data. The near overlap of the RNA4, CL-RNA6, and UC-RNA6 spectra verify the validity of the cleaved domain as a secondary folding model.

The 500 MHz imino proton spectrum as a function of temperature is shown in figure 8.21. The melting behavior of resonances A, I, and M helps corroborate the fact that resonance I is nearly degenerate with L, as labelled. As the temperature increases from 25°C to 40°C the order of broadening is I>A>M>M'. This is exactly the expected order of broadening for the stem II resonances as assigned in figure 8.11. The order of melting of the other resonances is similar to that observed for RNA4.

It appears as though resonances D and K of CL-RNA6 are slightly sharper than the other resonances of the spectrum (fig. 8.19), and clearly D is sharper in CL-RNA6 than in UC-RNA6. This evidence suggests, as was previously suggested for RNA4, that stem I has additional motion after a cleavage event. In essence, cleavage results in the disruption of tertiary interactions which hold stem I immobile in the uncleaved RNA. Spectral crowding in the region of resonance N in the RNA6 spectra prevents a similar analysis of this stem I resonance.
Figure 8.21 500 MHz imino proton spectrum of CL-RNA6 as a function of temperature. The temperature at which each spectrum was recorded is indicated in the right-hand margin in degrees Celsius. The sample conditions and spectral parameters are given in the legend of figure 8.20.
8.7 Final Conclusions and Discussion

Although a complete secondary assignment was obtained for RNA4, the spectroscopic results with MgCl₂ seemed to indicate that cleavage results in a subsequent rearrangement of the structure. In the absence of MgCl₂ resonances were observed for all of the base pairs except the two AU pairs directly in the junction of the three helices. Addition of MgCl₂ did not introduce extra resonances as expected, but it did result in selective broadening of resonances assigned to base pairs directly at and near this junction region. This behavior paralleled that of RNA1, as resonances which were speculated to be those of junction base pairs for RNA1 broadened under similar conditions. However, the lack of additional resonances cast doubt on the relevancy of studying cleaved RNA, as it is difficult to rationalize the conservation of thirteen nucleotides which do not interact in a sequence-specific manner. Thus the ability to study uncleaved RNA was desired.

This problem was solved when Rosie Kim discovered a mutant of RNA1 which did not cleave during a two hour transcription reaction. This mutation, which occurred at one of the conserved nucleotides, was incorporated into the RNA4 sequence to yield RNA6. Investigation of the cleavage kinetics of RNA6 showed that its cleavage was not completely stopped, but rather the rate significantly attenuated by this mutation. This proved to be a very desirable result because slowly cleaving RNA must be adopting a true cleavage conformation. The spectrum of RNA6 in the absence of MgCl₂ resembled that of RNA4, with a few very important differences. Resonance L was absent in the RNA6 spectrum, and peaks E, D, and F were very broad. These were the very resonances which selectively broadened upon addition of MgCl₂ to the RNA1 and RNA4 samples. These results indicated that uncleaved RNA is more "open" in the helical junction than the hammerhead folding model predicts, as speculated
for RNA1 and RNA4. Two dimensional NOESY and 1D NOE spectroscopy provided the secondary assignments given in figure 8.10. This figure illustrates RNA6 in the more open conformation consistent with the NMR data.

MgCl$_2$ was added to the RNA6 NMR sample with the hope that additional resonances which represented tertiary contacts between the conserved nucleotides would be introduced. Although there were spectral changes, no additional peaks from stable interactions appeared. There were some possible additional peaks in the 15-13.5 ppm region of the spectrum, but no NOEs were observed upon saturation of these resonances. The fact that the RNA6 NMR sample did not cleave after a 48 hour experiment in 10 mM MgCl$_2$ also cast doubt on the relevance of these spectra.

A series of analytical-scale cleavage reactions was done to determine the exact NMR buffer conditions which cause cleavage of the RNA. It was discovered that cleavage occurred in phosphate buffered solutions only if the MgCl$_2$ concentration exceeded that of the buffer. This was not a requisite for tris·HCl buffered cleavage, but a 2:1 molar ratio of spermidine to RNA was required for both buffers. Spectra of cleaving RNA6 were acquired in both phosphate and tris·HCl buffers. Comparison of these cleaving spectra with those acquired in 10 mM MgCl$_2$ (non-cleaving) contained no differences. These results indicated that either the global structure of the domain does not change upon the addition of spermidine to the sample, or that the spermidine affects those bases which are not observed in the imino proton region of the spectrum, namely those which do not participate in base pairs. In either case, the fact that the spectra did not change validated the imino proton results previously obtained.

The observation that RNA6 cleaved in phosphate buffered solution only when the MgCl$_2$ was greater than that of the phosphate implied that the residue which was
mutated is important for Mg\(^{2+}\) binding, as Mg\(^{2+}\) appears to have a greater affinity for free phosphate than for the magnesium binding site(s) of RNA6. This single point mutation turned a "good" Mg\(^{2+}\) binding site (RNA4) into a "poor" binding site (RNA6). The fact that RNA6 cleavage was spermidine dependent implied that magnesium juxtaposes several phosphate groups. Because RNA6 has a poor Mg\(^{2+}\) binding site polyvalent spermidine is required to overcome interphosphate repulsions.

To prove that cleavage does in fact result in a subsequent rearrangement of the structure, as was speculated for RNA4, a spectrum of completely cleaved RNA6 was desired. Unfortunately incubating RNA6 in cleavage conditions for the time required for complete cleavage causes degradation for large-scale reactions. Several attempts were made with both phosphate and tris·HCl buffers at different temperatures with similar results; the RNA degraded. The RNA was incubated for a period of time which resulted in approximately 80% cleavage and no degradation, and a spectrum acquired. Aside from sequence-specific differences, the spectrum of CL-RNA6 was almost identical to that of RNA4 at the same temperature and buffer conditions. Resonance L appeared, resonance A moved down-field to a chemical shift similar to that of RNA4, and resonances D, E, and F sharpened significantly. Based on the conclusions from the RNA4 data all of these changes were expected. The one unexpected change involved resonance I. Although this resonance appeared to be fairly sharp in the CL-RNA6 spectrum, as expected, its chemical shift was different than that observed for UC-RNA6. Because chemical shift is sensitive to electronic environment, it was argued that cleavage must result in a subsequent rearrangement of the structure such that resonance I resides in different environments in UC- and CL-RNA6. This was exactly the result speculated for RNA4. In addition, the
linewidths of the stem I resonances of UC- and CL-RNA6 indicated that cleavage also allows additional mobility of stem I. Again this was identical to the behavior observed for RNA4.

The results of these NMR studies are somewhat disappointing, as no tertiary contacts were identified in the imino region of the spectrum for the slowly cleaving RNA domain. However, it can be said with confidence that the secondary structure of the uncleaved domain more closely approximates that given in figure 8.10 than the model of Hutchins et al. (1987), and that the studies done with an already cleaved domain are certainly relevant to the secondary folding of the domain. It is still difficult to rationalize the conservation of thirteen nucleotides which do not interact in a sequence specific manner, but clearly there appears to be little more to be learned from spectroscopy of exchangeable protons of similar domains. To fully elucidate this difficult problem, assignments of the non-exchangeable protons will have to be ascertained.

There is one final possibility which can be explored. Because RNA6 is slowly cleaving it may not be in a true cleavage conformation at all times. In fact, only a very small population of RNA molecules may be in the cleavage conformation at any given time. Because NMR is a probe of the average structure, possible tertiary interactions would not be observed. For this reason I believe that it is necessary to re-design the sequence as a bimolecular construct similar to RNA2 once again. However, care must be taken that the new sequence exhibit a fast cleavage rate to avoid the structure equilibria observed for RNA2. The cleavage, as before, can be stopped by incorporation of a deoxynucleotide at the cleavage site. Only with this type of construct does one have the ability to stop cleavage without affecting the conserved residues, and hence the ability to study uncleaved RNA in a non-perturbed structure.
Appendix A

Biochemical Protocols
Purification of Plasmid DNA


1. In each of three 2 L flasks prepare 1 L of LB media [Maniatis, et al. (1982) p. 68], plug the flask with cotton and autoclave. At the same time autoclave a cotton plugged 125 mL flask.

2. When the media has cooled, transfer 30 mL of it to the 125 mL flask using a sterile pipette. Add the appropriate antibiotic to a final concentration of 50 μg/mL [Maniatis, et al. (1982) p. 71-72]. Inoculate this 30 mL culture with cells containing the desired plasmid DNA and grow overnight at 37°C on a shaker plate set to 200 rev/min.

3. Equilibrate the rest of the media at 37°C, and add the appropriate antibiotic to a final concentration of 50 μg/mL. Inoculate each 1 L culture with 10 mL of the overnight culture and grow overnight (16-18 hrs.) at 37°C with shaking.

4. Cool the cultures on ice for approx. 30 min. Collect the cells by centrifugation at 5000 rpm in the GS-3 rotor for 10 min. at 4°C.

5. Resuspend the cells in 90 mL ice cold solution I (50 mM glucose, 25 mM Tris,
15 mM HCl, 10 mM EDTA) containing 4 mg/mL lysozyme [Maniatis, et al. (1982) p. 451].

6. Hold the suspension on ice for 30 min.

7. Add 175 mL freshly prepared solution II (0.2 M NaOH, 1% (w/v) SDS).

8. Mix by inversion and hold the suspension on ice for 20 min.

9. Add 130 mL ice cold solution III (5 M KOAc, 5 M HOAc, pH 4.8) [Maniatis, et al. (1982) p. 477] and mix by inversion.

10. Hold the suspension on ice for 30 min.

11. Divide into two GSA centrifuge bottles and spin at 8000 rpm for 30 min. at 4°C.

12. Carefully pour the supernatant through a funnel lined with a disposable coffee filter into a 1 L bottle. Avoid transferring white flecks. Filter again if necessary.

13. Add 2 volumes of room temperature ethanol, or 0.6 volumes of room temperature isopropanol. Let the suspension stand at room temperature for 15 minutes, no longer.

14. Spin the suspension (it should be cloudy) at 8000 rpm in a warm GSA rotor for 40 min. at room temperature. If a warm rotor is used it is not necessary to warm the centrifuge chamber before spinning.
15. Carefully decant off the supernatant. Wash the pellets with 10 mL 70% ethanol and spin at 5000 rpm in the GSA rotor for 10 min.

16. Carefully decant off the supernatant and invert the bottles on a kimwipe to drain. The pellets can either be left inverted overnight to dry, or dried by lyophilization.

17. Resuspend the pellets in a total volume of 8 mL TE buffer, pH 8 [Maniatis, et al. (1982) p. 448]. This is best done at 37°C as the pellets do not dissolve easily.

18. Transfer the suspension to a 50 mL plastic screw-top tube and add 50 μL 10 mg/mL DNase-free RNase A [Maniatis, et al. (1982) p. 451]. Incubate at 37°C for 30 min.

19. Add 100 μL 10 mg/mL Proteinase K [Maniatis, et al. (1982) p. 450], 100 μL 10% (w/v) SDS, and 1.75 mL H₂O. Incubate at 37°C for 1 hour.

20. Transfer to a polyallomer SS-34 centrifuge tube, add 10 mL phenol [Maniatis, et al. (1982) p. 458], and vortex vigorously for 3-5 min. Spin at 10,000 rpm for 5 min. Carefully remove the aqueous phase (top) with a plastic pipette and transfer to a fresh SS-34 tube. Do not discard the phenol.

21. Add 10 mL 1:1 phenol/chloroform [Maniatis, et al. (1982) p. 458], vortex and spin as above. Repeat the phenol/chloroform extraction until there is no white solid at the organic/aqueous interface. Usually this requires two extractions.
22. Pool the phenol and phenol/chloroform from the extractions and add an equal volume of TE, pH 8. Vortex and spin as in step 18. Remove the aqueous phase and pool with the other aqueous fraction. Discard the phenol/chloroform.

23. Add NaCl to 0.1 M and 2 volumes of ethanol. Mix thoroughly. A white precipitate should be visible. Hold the suspension at -20°C for 30 min.

24. Spin at 12,000 rpm in the SS-34 rotor for 30 min. at 4°C. The pellet should be large and white. Wash the pellet with 70% ethanol and dry by lyophilization.

25. Dissolve the pellet in 3.36 mL H2O. Add 0.64 mL 5 M NaCl and 4 mL 13% deionized PEG-8000\(^1\). Hold the suspension in an ice/water bath until a precipitate is visible (at least one hour).

26. Spin in the SS-34 rotor for 30 min. at 12,000 rpm and 4°C. Decant off the supernatant, wash the pellet with 70% ethanol, and dry. Resuspend the pellet in 1-2 mL TE, pH 8 and quantitate by UV absorbance [Maniatis, et al. (1982) p. 468]. The purity of the DNA should be checked by agarose gel electrophoresis [Maniatis, et al. (1982) p. 150-163]. A typical yield for a 3 L prep is 10-15 mg pure plasmid DNA.

\(^1\)To deionize PEG-8000, add 20-50 mesh BioRad AG 501-X8 resin to a concentration of 8% and stir for 18 hours at room temperature. Remove the resin by filtration. Add sodium azide to 0.02% (w/v) and store the PEG-8000 at 4°C.
Linearization of Plasmid DNA

Adapted from protocols in Maniatis, T. Fritsch, E. F. & Sambrook, J. (1982)
Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

1. Prepare 10 mL of 10X EcoRI buffer (0.1 M Tris·HCl, pH 7.5 at 37°C, 0.5 M NaCl, 0.1 M MgCl₂, 10 mM DTT) [Maniatis, et al. (1982) p. 453].

2. Store the 10X EcoRI buffer at -20°C in 2 mL aliquots.

3. Prepare 5 mL of 2 mg/mL aqueous BSA, dispense into 1 mL aliquots and store at -20°C.

4. Dilute the plasmid to be linearized to an approximate concentration of 2 mg/mL with TE, pH 7.6 [Maniatis, et al. (1982) p.440]. Prepare analytical EcoRI digests according to the chart below, using the stock of plasmid, buffer, and enzyme which will be used for the large scale linearization.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X EcoRI</td>
<td>3.0 μL</td>
<td>3.0 μL</td>
<td>3.0 μL</td>
<td>3.0 μL</td>
</tr>
<tr>
<td>2 mg/mL BSA</td>
<td>1.5 μL</td>
<td>1.5 μL</td>
<td>1.5 μL</td>
<td>1.5 μL</td>
</tr>
<tr>
<td>Plasmid DNA</td>
<td>2.0 μL</td>
<td>2.0 μL</td>
<td>2.0 μL</td>
<td>2.0 μL</td>
</tr>
<tr>
<td>EcoRI (1U/μL)</td>
<td>0.5 μL</td>
<td>1.0 μL</td>
<td>1.5 μL</td>
<td>2.0 μL</td>
</tr>
<tr>
<td>Total Volume</td>
<td>30.0 μL</td>
<td>30.0 μL</td>
<td>30.0 μL</td>
<td>30.0 μL</td>
</tr>
</tbody>
</table>
5. Incubate the reactions for 1 hour at 37°C.

6. Run 3 μL from each reaction on a 0.8% agarose gel [Maniatis et al. (1982) p.150-163] to determine the which lanes were completely linearized.

7. For the large scale linearization, scale up the reaction which used the least amount of enzyme which yielded complete digestion, except use one half as much EcoRI as calculated from the analytical digest.

8. Mix all of the reagents for the large scale reaction in a plastic tube except for the EcoRI and equilibrate this mixture to 37°C.

9. Add the EcoRI to the reaction and incubate for 18 hours at 37°C.

10. Verify that the linearization has gone to completion by running 3 μL from the reaction on a 0.8% agarose gel. If the reaction has not gone to completion, add more EcoRI (about 1/4 as much as was originally added), and continue the incubation.


thoroughly.

13. Precipitate the DNA with 2-3 vols. of ethanol at -20°C for 18-24 hours.

14. Spin at 10K rpm and 4°C for 15 min. to pellet the precipitate. Wash the pellet twice with 70% ethanol, and dry by lyophilization.

15. Resuspend the pellet in a few milliliters of TE, pH 7.6 and quantitate by UV spectroscopy at both 260 and 280 nm. Pure plasmid DNA will have an absorbance ratio of $A_{260}/A_{280} = 1.8$ [Maniatis, et al. (1982) p. 468]. If the ratio is lower than this value, the preparation is contaminated with either protein or phenol.

16. To remove phenol from the sample, extract with an equal volume of water saturated ether [Maniatis, et al. (1982) p. 460], and repeat steps 11-14. If the $A_{260}/A_{280}$ ratio is still low, there is a protein contaminant and steps 10-14 should be repeated.
Purification of T7 RNA Polymerase


2. Inoculate the media with about 25 mL of an overnight culture of *E. Coli* BL21 containing pAR1219 grown in the same media. Divide the culture into two sterile, cotton plugged 4 L flasks and shake vigorously at 37°C for 2 hours.

3. Add IPTG to 1 mM (0.119 g/500 mL) and shake and additional 4 hours at 37°C.

4. Centrifuge for 30 min. at 9000 rpm to harvest the cells. Decant the supernatant, wash the pellet once with tris·sucrose buffer (50 mM tris·HCl, pH 8.0 at 4°C, 25% (w/v) sucrose), and pellet as before. Decant the supernatant and discard.

5. Carefully dissolve the cells in tris·sucrose buffer (10 g cells/100 mL buffer) and store at -20°C.
The following steps describe the purification of the protein from 10 g of *E. Coli* BL21 cells and preparation of the Cellulose P-11 (Whatman) column. All manipulations should be performed at 4°C unless stated otherwise. The solutions do not need to be autoclaved if care is taken in their preparation. If autoclaving is desired, add dithiothreitol (DTT), β-mercaptoethanol (BME), phenylmethylsulfonyl fluoride (PMSF), and leupeptine after autoclaving.

1. Prepare a 20 mg/mL stock solution of PMSF in isopropanol. Dispense into conveniently sized aliquots and store at -20°C.

2. Prepare a 5 mg/mL stock solution of aqueous leupeptine. Dispense into conveniently sized aliquots and store at -20°C.

3. Prepare 500 mL buffer L (22 mM NH₄Cl, 20 mM tris·HCl, pH 8.0 at 4°C, 1 mM DTT, 3 mM EDTA, 5% glycerol, 2 μg/mL PMSF, 0.2 μg/mL leupeptine), 5 L buffer I+0.15 M NaCl (10 mM potassium phosphate, 10 mM BME, 0.1 mM EDTA, 10% glycerol, 2 μg/mL PMSF, 0.15 M NaCl), 500 mL buffer I+0.25 M NaCl (10 mM potassium phosphate, 10 mM BME, 0.1 mM EDTA, 10% glycerol, 2 μg/mL PMSF, 0.25 M NaCl), and 1.5 L buffer I+0.6 M NaCl (10 mM potassium phosphate, 10 mM BME, 0.1 mM EDTA, 10% glycerol, 2 μg/mL PMSF, 0.6 M NaCl). Equilibrate the buffers to 4°C.

4. Prepare 2 L of buffer II+10% glycerol (10 mM tris·HCl, pH 8.0 at 4°C, 100 mM NH₄Cl, 0.05 mM EDTA, 1 mM DTT, 10% glycerol), 2 L of buffer II+25% glycerol (10 mM tris·HCl, pH 8.0 at 4°C, 100 mM NH₄Cl, 0.05 mM EDTA, 1 mM
DTT, 25% glycerol), and 2 L of buffer I+50% glycerol (10 mM tris·HCl, pH 8.0 at 4°C, 100 mM NH₄Cl, 0.05 mM EDTA, 1 mM DTT, 50% glycerol). Equilibrate these buffers at 4°C.

5. Prepare 70 mL dry Cellulose P-11 column material (Whatman) according to the manufacturers instructions. Suspend the resin in 1 L 100 mM potassium phosphate, pH 8.0. Adjust the pH to 8.0 with conc. KOH. Remove the buffer and fines. Suspend the resin in 500 mL of buffer I+0.15 M NaCl and adjust the pH to 8.0 with conc. KOH. Allow the resin to settle and decant the buffer and fines. Repeat suspending the P-11 in buffer I+0.15 M NaCl until the suspension reaches pH 8.0 without further adjustment. Pour the suspension into a 3 x 30 cm column (Pharmacia) and equilibrate the resin with 500 mL buffer I+0.15 M NaCl at a flow rate of 0.5 mL/min.

6. Thaw the E. Coli BL21 cell/tris-sucrose slurry at 4°C (about 10-12 hours). Add 170 mL buffer L and mix by inversion.

7. With stirring, add 90 mg lysozyme (dissolved in 30 mL buffer L). Stir for 30 min.

8. Add 0.15 g sodium deoxycholate dissolve in a small volume of buffer L and stir an additional 10 min.

9. Pull the suspension through an 18-20 gauge needle and push through the needle into SS-34 centrifuge tubes. Spin at 15,000 rpm for 60 min. in the SS-34 rotor.
Carefully transfer the supernatant to a GS-3 centrifuge bottle (the pellet is loose and slimy).

10. With stirring, slowly add 80 g (NH₄)₂SO₄ to the solution. Stir an additional 20 min. after the final addition. A white precipitate should be visible.

11. Spin for 20 min. at 9000 rpm in the GS-3 rotor to pellet the precipitate. Carefully decant the supernatant (it may contain white particulate matter) and discard.

12. Resuspend the pellet in 10-15 mL buffer I+0.15 M NaCl. Dialyze against 1 L of the same buffer for 18-24 hours in 12-14 MWCO dialysis tubing.

13. If a precipitate formed during dialysis, spin the suspension at 10,000 rpm for 15 min. and transfer the clear supernatant to a fresh tube. Load onto the Cellulose P-11 column by carefully pipetting the solution into the void volume above the resin.

14. Wash the column with buffer I+0.15 M NaCl at a flow rate of 0.25 mL/min. until the A₂₆₀ returns to base line. The DNA elutes during this time.

15. Carefully remove the buffer in the void above the resin and replace it with an equal volume of buffer I+0.25 M NaCl. Elute the protein with a 200 mL linear gradient of 0.25 M NaCl to 0.6 M NaCl in buffer I at a flow rate of 0.25 mL/min. Collect 5-10 mL fractions until the A₂₈₀ returns to baseline.
16. Check fractions with significant absorbance at 280 nm for polymerase activity in a 20 \( \mu \)L transcription reaction with 2 \( \mu \)L from each fraction.

17. Check those fractions with significant polymerase activity for RNase activity by incubating 2 \( \mu \)L from each fraction with 1 \( \mu \)g of \( ^{32} \text{P} \)-labelled RNA oligonucleotide in 20 \( \mu \)L 1X transcription buffer. Pool those fractions with significant polymerase activity and minimal RNase activity.

18. Dialyze the pooled fractions against 2 L of buffer II+10% glycerol for 8 hours at 4\(^{\circ}\)C. Change the dialysis buffer to 2 L of buffer II+25% glycerol and dialyze an additional 8 hours at 4\(^{\circ}\)C, followed by a 12 hour dialysis against buffer II+50% glycerol at -20\(^{\circ}\)C. This final step should concentrate the enzyme approximately two-fold.

19. Dispense the T7 RNA polymerase into 1 mL aliquots at -20\(^{\circ}\)C, and store at -20\(^{\circ}\)C.
Transcription Buffers and Conditions


1. All solutions should be prepared from glass-distilled or DEPC-treated water [Maniatis, et al. (1982) p. ] to avoid contamination with RNase.

2. When dispensing solutions into aliquots, pass through a 0.45 μm filter to sterilize.

3. Prepare stocks of 5X transcription buffer (200 mM tris·HCl, pH 8.1 at 37°C, 32 mM MgCl₂, 50 mM NaCl, 5 mM spermidine), 10X DTT (100 mM DTT), and 10X NTP (20 mM ATP, CTP, GTP, and UTP in 40 mM tris·HCl, pH 8.1 at 37°C). Filter sterilize the stocks and dispense into conveniently sized aliquots. Store the stocks at -20°C.

4. Dilute the linearized plasmid DNA to an approximate concentration of 2 mg/mL with TE, pH 8.0 [Maniatis, et al. (1982) p. 440]. Prepare analytical transcription reactions according to the chart on the next page, using the stock of plasmid, transcription buffer, DTT, NTP, and T7 RNA polymerase that will be used for the large-scale transcription.
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>5X Trans. Buffer</td>
<td>4 µL</td>
<td>4 µL</td>
<td>4 µL</td>
<td>4 µL</td>
<td>4 µL</td>
</tr>
<tr>
<td>10X DTT</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
</tr>
<tr>
<td>20X NTP</td>
<td>1 µL</td>
<td>1 µL</td>
<td>1 µL</td>
<td>1 µL</td>
<td>1 µL</td>
</tr>
<tr>
<td>RI/plasmid DNA</td>
<td>1 µL</td>
<td>1 µL</td>
<td>1 µL</td>
<td>1 µL</td>
<td>1 µL</td>
</tr>
<tr>
<td>T7 RNA polym.</td>
<td>1 µL</td>
<td>3 µL</td>
<td>5 µL</td>
<td>7 µL</td>
<td>9 µL</td>
</tr>
<tr>
<td>Total Volume</td>
<td>20 µL</td>
<td>20 µL</td>
<td>20 µL</td>
<td>20 µL</td>
<td>20 µL</td>
</tr>
</tbody>
</table>

5. Incubate the reactions at 37°C for 1 hour.

6. Quench the reactions by addition of 1 µL of 0.5 M EDTA.

7. Run 3 µL from each reaction on a 8 x 10 x 0.7 cm 12% native polyacrylamide minigel [Maniatis, et al. (1982) p.]. On this same gel run a few lanes containing known quantities of pure yeast tRNA^Phe^ (Sigma).

8. Stain the gel with 0.02% (w/v) toluidine blue-O to visualize the RNA. Quantitate the RNA yields by comparing the stain intensity of the RNA product bands with those of the tRNA standards.
Large Scale RNA Transcription

1. From the analytical reactions, calculate the volumes of reagents necessary to produce the desired amount of RNA.

2. In a 50 mL Corning plastic centrifuge tube mix together the appropriate volumes of 5X transcription buffer, linearized plasmid DNA, 10X DTT, 20X NTP, and H₂O. Equilibrate this solution to 37°C.

3. Add the appropriate volume of T7 RNA polymerase and mix gently, but thoroughly, by inversion. To this mixture add 10 U RNasin (Promega) and 1 μL Inorganic Pyrophosphatase (Boerrhinger-Mannheim) for every 10 mL of reaction volume. Mix thoroughly by inversion.

4. Incubate the reaction at 37°C for 4-18 hours. If a white precipitate appears during the course of the reaction, do not be alarmed! It is inorganic pyrophosphate, and although it makes the sample more difficult to handle in subsequent steps, it is not harmful.

5. Quench the reaction by adding EDTA to 6.5 mM.

7. Carefully transfer the aqueous phase to a fresh SS-34 tube. Repeat step 6 until the organic/aqueous interface is clear.

8. Back-extract the phenol/chloroform with an equal volume of TE, pH 8.0, and pool the aqueous phase with the first fraction. Discard the phenol/chloroform.

9. Add NaCl to 0.1 M and precipitate the nucleic acids for 18-24 hours in 3 volumes of ethanol at -20°C.

10. Pellet the precipitate by centrifugation at 10,000 rpm and 4°C for 20 min. Wash the pellet twice with 2 mL 70% Ethanol, and dry the pellet under vacuum.

11. The pellet can be stored dry at -20°C until ready for purification.
Large Scale Purification of RNA

1. Prepare a 20 x 40 x 0.1 cm 12% native polyacrylamide preparatory slab gel (for slowly leaving RNA prepare 12% polyacrylamide containing 7 M urea). Use a comb that has only 5 teeth.

2. Pre-run the native gel at 15 W constant power (approx. 300 V/12 mA) in 1X TBE [Maniatis, et al. (1982) p. 454] for 2-4 hours (the denaturing gel should be pre-run at 25 W constant power).

3. Dissolve the pellet from the transcription reaction in 1 mL of TE, pH 7.6. (If the transcription had a white precipitate the pellet will not dissolve completely. Pellet the pyrophosphate by spinning in an eppendorf centrifuge for 5 min. and carefully transfer the clear supernatant to a fresh eppendorf tube.) Add 200 μL glycerol dye [Maniatis, et al. (1982) p. 455] and vortex briefly.

4. Load onto the prep. gel, 240 μL per well, and run the gel at 15 W constant power until the xylene cyanol dye marker (top marker) has migrated half-way down the gel (10-16 hours). [The denaturing gel should be run at 25 W constant power until the bromophenol blue dye marker has just run off the gel (6-10 hours).]

5. Pry the plates apart and transfer the gel to a piece of Saran Wrap®. Place the gel on a fluorescent TLC plate and shine short-wave UV light on the gel. The RNA will cast a dark shadow.
6. Excise the appropriate band with a sharp razor blade. Cut the gel slice into tiny pieces (2 x 2 mm) and transfer them to a 50 mL Corning plastic centrifuge tube. Add enough gel extraction buffer (10 mM tris·HCl, pH 7.6 at 37°C, 0.15 M NaCl, 10 mM EDTA, 0.1% (w/v) SDS) to just cover the slices.

7. Soak the covered gel slices overnight at 37°C with shaking at approx. 200 rpm.

8. Remove the buffer with a drawn out pasteur pipette (being careful not to transfer pieces of polyacrylamide) and transfer to a fresh tube. Rinse the slices with extraction buffer and pool the rinse with the first fraction.

9. Add 0.5 vol. of H$_2$O to the recovered buffer to lower the NaCl conc. to 0.1 M, and precipitate the RNA for 18-24 hours at -20°C in 3 volumes of ethanol.

10. Pellet the precipitate by centrifugation at 10,000 rpm and 4°C for 20-30 min. Decant the Ethanol, wash the pellet twice with 2 mL 70% Ethanol, and dry the pellet under vacuum.

11. Exchange buffer conditions by passing across a Sephadex G-25 (Pharmacia) column, or by dialysis at 4°C.

12. If the sample contains polyacrylamide, it can be removed by extracting with an equal volume of 50:50 phenol/chloroform. Clean the sample by ethanol precipitation.
Appendix B

GN-500 Pulse Sequences
1331 Solvent Suppression

1. D5
2. P1, G
   XMTR PHASE: A
3. D7, G
4. P3, G
   XMTR PHASE: B
5. D7, G 6. P3, G
   XMTR PHASE: A
7. D7, G
8. P1, G
   XMTR PHASE: B
9. A, G
   RCVR PHASE: A
10. D6  JUMP TO #1

PHASE  A = 2*S  MODULO 4
PHASE  B = (2*S)+2  MODULO 4

QPD + Y = 0, AB-
D7 = 1/[2(\nu_{\text{max}} - \nu_{\text{null}})]
P2 = 1/4 \times 90^\circ  pulse
OC = 57
GN = 25

Solvent supression with presaturation during D10

1. D5
2. D10, D, C, L1, G
3. D12, G
4. P1, G
   XMTR PHASE: A
5. D7, G
6. P3, G
   XMTR PHASE: B
7. D7, G
8. P3, G
   XMTR PHASE: A
9. D7, G
10. P1, G
    XMTR PHASE: B
11. A, G
    RCVR PHASE: A
12. D6
13. D5
14. D10, G
15. D12, G
16. P1, G
   XMTR PHASE: A
17. D7, G
18. P3, G
   XMTR PHASE: B
19. D7, G
20. P3, G
   XMTR PHASE: A
21. D7, G
22. P1, G
   XMTR PHASE: B
23. A, G
    RCVR PHASE: A
24. D6 JUMP TO #1

PHASE A = 2*S MODULO: 4
PHASE B = (2*S)+2 MODULO: 4

QPD + Y = 0, AB-
D7 = 1/[2 vmax - v null)]
P2 = 1/4 * 90° pulse
OC = 57
GN = 25

MB = 2
D12 = 1 msec.
L1 = 37-40
11NOE1
11 Jump Return NOESY with 11-Echo Detect

1. D5, G
2. P2, G
   XMTR PHASE: A
3. D2, G
4. P2, G
   XMTR PHASE: A+2
5. D+8, G
6. D+8, G
7. D+8, G
8. D+8, G
9. P2, G
   XMTR PHASE: B
10. D2, G
11. P2, G
    XMTR PHASE: B+2
12. D9, G
13. D+8, G
14. D9, G
15. P2, G
16. D2, G
17. P2, G
18. D10, G
19. P2, G
20. D4, G
21. P2, G
22. D10, G
23. A, G
    RCVR PHASE: E
24. D6 JUMP TO #1

PHASE A = 3+[2*(S/4)]+(S/16)+# MODULO: 4
PHASE B = 3+(S/16) MODULO: 4
PHASE C = 3+[2*(S/8)]+[3·(S/16)] MODULO: 4
PHASE D = S MODULO: 4
PHASE E = (2*S)+{2*[(S/4)+1]+(S/16)]} MODULO: 4

QPD + Y = 0, AB-
D2 = 1/[2(V_{max} - V_{null})]
P2 = 90° pulse

D9 = \tau_{mixing}/2
D4 = 2*D2
OC = 57

13NSY

11-NOESY with 1331 Detect

1. D5, G
2. P4, G
   XMTR PHASE: A
3. D2, G
4. P4, G
   XMTR PHASE: A+2
5. D+8, G
6. D+8, G
7. D+8, G
8. D+8, G
9. P4, G
   XMTR PHASE: B
10. D2, G
11. P4, G
   XMTR PHASE: B+2
12. D9, G
13. D+8, G
14. D9, G
15. P2, G
   XMTR PHASE: C
16. D2, G
17. P3, G
   XMTR PHASE: C+2
18. D2, G
19. P3, G
   XMTR PHASE: C
20. D2, G
21. P1, G
22. A, G
   RCVR PHASE: D
23. D6 JUMP TO #1

PHASE A = [2*(S/4)]+(S/8)+# MODULO: 4
PHASE B = (S/8) MODULO: 4
PHASE C = S+(S/8) MODULO: 4
PHASE D = {S+[2*(S/4)]+(S/8)} MODULO: 4

NS = 32*
D9 = \tau_{\text{mixing}}/ 2
D2 = 1/[2(V_{\text{max}} - V_{\text{null}})]
OC = 57
QP D+ Y = 0, AB-
P2 = 1/4 * 90° pulse
GN = 25

References


