INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Bacillus megaterium ribosomal 5S RNA structure from proton nuclear magnetic resonance spectroscopy and molecular dynamics simulations

Kim, Jong Hwa, Ph.D.

The Ohio State University, 1990
BACILLUS MEGATERIUM RIBOSOMAL 5S RNA STRUCTURE FROM PROTON NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY AND MOLECULAR DYNAMICS SIMULATIONS

DISSertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Jong Hwa Kim, B.S., M.S.

* * * * *

The Ohio State University
1990

Dissertation Committee:
Professor Alan G. Marshall
Professor Daniel L. Leussing
Professor Anthony W. Czarnik

Approved by

Adviser
Department of Chemistry
ACKNOWLEDGMENTS

A number of people have made important contributions to this dissertation. First of all, I would like to express my sincere appreciation to Professor Alan G. Marshall for his kind advice and guidance throughout my coursework and research. Without him, nothing might have been possible. I would also like to extend my appreciation to Professors Daniel L. Leussing and Anthony W. Czarnik for helpful suggestions that brought this dissertation to its present form.

I owe a special debt of gratitude to Dr. Charles E. Cottrell and Dr. Ruth Hsu for advice on NMR experiments and computer simulations, and to Jane Tolley, who synthesized RNA fragments in the course of this research, although that could not be included in this dissertation. The Fermentation Lab of The Ohio State University extended valuable assistance by providing *Bacillus megaterium* cells for this research. The U.S. Public Health Service provided grants N.I.H. GM-29274 and N.I.H. RR-01458. I am very grateful to my colleagues George Alber, Ruidan Chen, Pete Grosshans, Nick Hill, Hsiau-Wen Huang, Zhenmin Liang, Mike May, Chuck Ross, Chris Williams, Troy Wood, Jane Xiang, Winnie Yin, and Pat Limbach for helpful suggestions and discussion. Special thanks to Susan Holleran for her consistent encouragement.

Acknowledgments to my family are of a different sort. Each of them has contributed to this work in ways that are difficult to articulate. My mother has provided constant and unselfish support throughout the course of my education, which has kept me far from home for so long.
VITA

April 28, 1956 ............... Born, Taegu, Republic of Korea
Feb. 26, 1980 ............... B.S., Department of Chemistry, Seoul National University, Republic of Korea
1980-1982 .................. Teaching Assistant, Department of Chemistry, Seoul National University
Feb. 26, 1982 ............... M.S., Department of Chemistry, Seoul National University
1982-1982 .................. Military Service (Second Lieutenant in the Korean Army)
1984-Present ............... Graduate Teaching Associate and Graduate Research Associate, Department of Chemistry, The Ohio State University

PUBLICATIONS


iii
FIELDS OF STUDY

Major Field: Chemistry

Analytical Chemistry, Biophysical Chemistry
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER I: INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>A. Objectives</td>
<td>1</td>
</tr>
<tr>
<td>B. Structures of RNA Molecules</td>
<td>5</td>
</tr>
<tr>
<td>1. Chemical Components of RNA and Nomenclature</td>
<td>5</td>
</tr>
<tr>
<td>2. Base Pairs and RNA Helices</td>
<td>7</td>
</tr>
<tr>
<td>C. Proposed Secondary Structure Models of 5S rRNA</td>
<td>13</td>
</tr>
<tr>
<td>D. Investigations of Three-Dimensional Structure of 5S rRNA</td>
<td>15</td>
</tr>
<tr>
<td>E. Methods for Structural Determination of 5S rRNA</td>
<td>18</td>
</tr>
<tr>
<td>1. Comparative Sequence Analysis</td>
<td>20</td>
</tr>
<tr>
<td>2. Enzymatic Accessibilities to 5S rRNA</td>
<td>23</td>
</tr>
<tr>
<td>3. Chemical Modifications</td>
<td>24</td>
</tr>
<tr>
<td>4. Binding of Oligonucleotides to RNA Molecules</td>
<td>28</td>
</tr>
<tr>
<td>5. Hydrodynamics and X-ray Scattering Studies</td>
<td>29</td>
</tr>
<tr>
<td>6. UV, IR, and Raman Spectroscopy</td>
<td>31</td>
</tr>
</tbody>
</table>
CHAPTER II: APPLICATION OF $^1$H NMR SPECTROSCOPY TO STRUCTURAL STUDIES OF RIBOSOMAL 5S RNA

A. Resonances of Hydrogen-Bonded (Base-Paired) Imino Protons

B. Water Suppression in $^1$H FT NMR Spectroscopy
   1. Saturation of the Water Signal
   2. Water Eliminated Fourier Transform (WEFT) Method
   3. Redfield 2-1-4 Pulse Sequence
   4. Binomial 1-3-3-1 Pulse Sequence
   5. Data Shift Accumulation (DSA) and Alternate Delay Accumulation (ADA)

C. Nuclear Overhauser Enhancements (NOE) — The Base Pair Assignment Strategy
   1. The Nuclear Overhauser Effect (NOE)
   2. The Resonance Assignment by means of NOE
   3. Two-Dimensional NOE Spectroscopy (NOESY)

D. Ring Current Shift Estimation

CHAPTER III INVESTIGATION OF THE STRUCTURE OF BACILLUS MEGATERIUM RIBOSOMAL 5S RNA

A. B. megaterium Cell Growth and Isolation and Purification of 5S rRNA
1. Production of RNase T1-Cleaved Fragments .................. 105
2. Purifications of the Three RNase T1-Cleaved Fragments ....................................................................................... 105
3. Purity Check of the Three Fragments .............................. 106
B. Confirmation of the Primary Nucleotide Sequences of the Three Fragments by RNA T1 Mapping ............................................. 109
   1. Experimental Methods for RNA T1 Mapping ....................... 109
      a. Preparation of Buffer Solutions ............................................ 109
      b. 5'-End Dephosphorylation of RNA with CAP ................ 110
      c. 3'-End Labeling of RNA with [5'-32P] pCp ...................... 111
      d. Purification of the Labeled RNA ........................................... 111
      e. T1 Digestion and Alkaline Hydrolysis ............................... 112
      f. Sequencing Gel .......................................................................... 113
   2. Sequencing Results ................................................................. 114
C. Identification and Assignment of Base Pairs from Fragment A by Proton Homonuclear Overhauser Enhancements ........... 114
   1. Mg+2-Depleted NMR Sample Preparation ............................... 114
   2. NMR Spectroscopy .......................................................................... 116
   3. Base Pairs in Fragment A (Helices II and III) ...................... 116
   4. Conformations of RNase T1-Cleaved Fragment A and the Corresponding Segment of Intact 5S rRNA .................. 124
D. Dynamic Structure of Fragment A ....................................................... 125
   1. NMR Spectroscopy .......................................................................... 127
   2. Conformational Change of Fragment A .............................. 127
   3. Evidence against Two Conformations of the Helix II–III Segment .................................................................................. 137
viii
CHAPTER V  FRAGMENT B AND FRAGMENT C: BASE-PAIRINGS IN HELICES IV AND I OF B. MEGATERIUM RIBOSOMAL 5S RNA

A. Base-Pairings in the prokaryotic Loop (Helix IV) of B. megaterium 5S rRNA (NMR Analysis of Fragment B) ................ 147
   1. NMR Spectroscopy ................................................................. 147
   2. Assignment of Base Pair Sequence in Fragment B (Helix IV) ................................................................. 148

B. Base-Pairings in Helix I of B. megaterium 5S rRNA (NMR Analysis of Fragment C) ..................................................... 158
   1. NMR Spectroscopy ................................................................. 158
   2. Base Pairs in Fragment C (Helix I) ....................................... 159

C. Conformations of Fragments B and C and the Corresponding Segments of Intact 5S rRNA ................................................................. 163

D. Base Pairs in Other Structural Segments ........................................ 164

E. Conclusion .................................................................................... 166

CHAPTER VI  MOLECULAR MODELING OF RNASE T1-CLEAVED FRAGMENT A AND FRAGMENT B OF B. MEGATERIUM RIBOSOMAL 5S RNA ........................................ 167

A. Basic Theory of Molecular Dynamics ............................................. 168
   1. Force Fields ................................................................. 168
   2. Molecular Dynamics (MD) ............................................. 169
B. Procedures for Molecular Dynamics Simulations ......................... 171
  1. Building of RNA Molecules using BIOGRAF ........................... 171
  2. Minimization ............................................................................. 172
  3. Molecular Dynamics .................................................................. 173
  4. Parameters for Molecular Dynamics ........................................ 176
     a. Dynamics Variables ......................................................... 176
     b. Temperature Variables .................................................. 176
     c. Running Annealed Dynamics ........................................... 177
C. Structure of the Helices II-III Region (Fragment A) .................... 178
D. Structure of the Helix IV Region (Fragment B) .......................... 186
E. Concluding Remarks .................................................................. 193

REFERENCES .......................................................................................... 196
LIST OF TABLES

Table 1. Size and composition of the ribosomes from various species. From Wittmann (1986). .......................................................... 2

Table 2. Approximate $^1$H NMR chemical shift ranges of various RNA protons. From Wüthrich (1986). .......................... 42

Table 3. Summary of NOE connectivities obtained by irradiation of imino proton resonances of B. megaterium 5S rRNA. ... 102

Table 4. Identification and assignment of base pair imino proton resonances in B. megaterium 5S rRNA Fragment A by means of NOE connectivities. ................................................ 121

Table 5. NOE connectivities of B. megaterium 5S rRNA Fragment A obtained under three different conditions. ................. 133

Table 6. Observed and ring current-calculated base pair imino proton chemical shifts for Fragment B of B. megaterium 5S rRNA. ................................................................. 150

Table 7. Identification and assignment of base pair imino proton resonances in Fragment B by means of NOE connectivities. ................................................................. 155

Table 8. Distance constraints used for dynamics calculations of Fragment A. ................................................................. 174

Table 9. Distance constraints used for dynamics calculations of Fragment B. ................................................................. 175

Table 10. Distances between neighboring base-paired imino protons in the helices II–III region of B. megaterium 5S rRNA. Calculated by molecular dynamics. ....................... 184

Table 11. Distances between neighboring imino protons in the helix IV region of B. megaterium 5S rRNA. Calculated by molecular dynamics without constraints. ............... 192
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Components of prokaryotic and eukaryotic ribosomes. Numbers in brackets are for eukaryotic ribosomes. From Erdmann et al. (1980).</td>
</tr>
<tr>
<td>2</td>
<td>Ribonucleic acid (RNA) consists of adenosine (A), guanosine (G), cytidine (C), and uridine (U), which are connected together via 3', 5'-phosphodiester bonds. The chain direction is from 5' to 3' as shown. From Saenger (1984).</td>
</tr>
<tr>
<td>4</td>
<td>(Left) Sugar puckering modes in RNA. (a) Flat five-membered sugar, which was never observed. (b) C₃-endo. (c) C₂-endo. (d) C₂-exo-C₃-endo. (e) Major C₃-endo and minor C₂-exo. (Right) Definition of major and minor grooves in A•U and G•C base pairs. From Saenger (1984).</td>
</tr>
<tr>
<td>5</td>
<td>(Top) Definition of propeller twist θₚ, which is dihedral angle between two base planes. θₚ is defined positive if the colser base is rotated clockwise with respect to the far one when we look down the long axis of a base pair. (Bottom) Definition of twist t, tilt θₜ, and roll θᵣ. The pseudodyad is reference line for t and θₜ, and a vertical to the pseudodyad passing approximately through C₆ (pyrimidine) and C₈ (purine) is for θᵣ. Positive signs for these angles are clockwise rotations as shown. From Saenger (1984).</td>
</tr>
<tr>
<td>6</td>
<td>Molecular structure of A-RNA described by ball-and-stick model. (Left) Viewed from perpendicular (top) and...</td>
</tr>
</tbody>
</table>
parallel (bottom) to the helix axis. (Right) The double helix is tilted by 32° to show the major (M) and the minor (m) grooves. From Saenger (1984).

Figure 7. Derivation of the most popular proposed models of 5S rRNA secondary structure.

Figure 8. The minimal model of eubacterial 5S rRNA secondary structure. Squares indicate conserved base-pairings, while circles indicate unpaired nucleotides. From Erdmann & Wolters (1986).

Figure 9. (Top) A tertiary Y-shaped 5S rRNA structure proposed based on a small-angle X-ray scattering study. From Osterberg et al. (1976). (Bottom) A three-dimensional 5S rRNA structure proposed based on hydrodynamic studies. From Fox and Wong (1979).

Figure 10. (Top) A three-dimensional 5S rRNA structure proposed based on a study with chemical reagent phenylglyoxylic acid (PGD). From Hancock & Wagner (1982). (Bottom) A three-dimensional 5S rRNA structure proposed based on chemical reactivity studies with monoperphthalic acid. Interaction sites for tRNA and other ribosomal RNAs are shown in (a), and the binding sites for ribosomal proteins are shown in (b). From Luck & Erdmann (1983).

Figure 11. Various experimental techniques used to investigate the structure of 5S rRNA.

Figure 12. One typical example of comparative sequence analysis. The sequence of 5S rRNAs from ten different organisms are aligned and compared with each other to obtain maximum sequence homology. The sequences in boxes are presumably base-pairing regions. Highly phylogenetically conserved regions are shown in brackets. From Fox & Woese (1975).

Figure 13. (Top) The modified Fox and Woese secondary structural model of E. coli 5S rRNA. Arrows show the primary cutting positions of ribonucleases A, T1 and T2, which are single-stranded specific. The arrow size represents the degree of cutting. (Bottom) The modified Fox and Woese secondary structure model of B. stearothermophilus 5S rRNA. Arrows indicate the cutting positions of ribonucleases A, T1, T2, and cobra venom.
Figure 14. (A) Modification of uridine 5'-phosphate with carbodiimide (I), where one of the R groups contains a closely situated quaternary ammonium group. The predicted structure of the product is II. From Gilham (1962). (B) The selective modification of cytidine by O-methylhydroxylamine. From Kochetkov et al. (1963). (C) Selective 1-N-oxidation of adenine by monoper-phthalic acid. Strong inhibition of oxidation was observed when adenines were base-paired. From Cramer & Seidel (1964).

Figure 15. (A) Different accessibilities of native and A form of E. coli 5S rRNA. The secondary structure of E. coli 5S rRNA is drawn according to the Fox and Woese model. From Lewis & Doty (1977). (B) Proposed oligonucleotide binding sites for prokaryotic 5S rRNAs based on nucleotide binding studies. From Wrede et al. (1978).


Figure 17. Proposed secondary structure (the Fox and Woese model with an additional helix V) adapted to the primary base sequence of B. megaterium 5S rRNA. Heterogeneity of bases in the sequence is denoted as follows: R is G or A, K is G or U, and Y is C or U.

Figure 18. Watson-Crick base pairs (A•U and G•C) and the "wobble" G•U base pair. Helical regions in 5S rRNAs comprise these three types of base pairs.

Figure 19. Example of the 1H NMR spectrum of 5S rRNA. Protons of each category (imino, amino, aromatic, or ribose protons) have characteristic chemical shifts.

Figure 20. The reduction in the intensity of the base-paired imino proton resonances with the increase of saturation pulse length is demonstrated in the downfield 1H NMR spectrum of 1.5 mM tRNA. (a) 10 msec, (b) 20 msec, (c) 50 msec, and (d) 100 msec. From Campbell et al. (1977).
Figure 21. (Top) Pulse sequence for modified WEFT, which consists of a recovery time PD (~5T1), a long weak nulling pulse T, a waiting time to null, and acquisition time (AT). From Gupta (1976). (Bottom) (a) Normal WEFT, (b) DASWEFT, and (c) DASWEFT with NOE preirradiation pulse. The "DANTE" π-pulse comprises one hundred 1 μsec pulses with a repetition rate of 1 msec. From Haasnoot (1983).

Figure 22. Construction of the Redfield 2–1–4 pulse sequence. (A) A single soft long pulse in the time-domain and its frequency-domain counterpart obtained by FT operation of the pulse. (B) Two 180° phase-shifted pulses and the resulting frequency power spectrum. (C) The addition of the pulses in (A) and (B), and its frequency-domain power spectrum obtained by FT operation. (D) Comparison of the Redfield 2–1–4 pulse and a single soft long pulse.

Figure 23. Transverse magnetization excited by the sequences 11, 121, 1331, and 14641 as a function of offset from the transmitter frequency. Calculations were done with τ = 500 μsec, γB1/2π = 5 kHz, and α1 (the flip angle of the first pulse) equal to 45, 22.5, 11.25, and 5.625°, respectively. From Hore (1983b).

Figure 24. (Top) The initial part of 1H FID (a), and after it is left-shifted by two data points (half a wavelength of water frequency) (b). The addition of (a) and (b) is (c). (Bottom) The noise spectrum obtained without a sample. (a) Before DSA operation, (b) after shifting FID by one data point and adding to the unshifted FID data. (c) and (d) are obtained in the same way as (b) except that two data points shift (c) and four data points shift (d) instead of one. From Roth et al. (1980).

Figure 25. Comparison of 1H NMR spectra of B. megaterium 5S rRNA Fragment A before (bottom) and after (top) DSA treatment. The significant water suppression achieved by a 1331 pulse sequence is greatly enhanced by DSA operation in such a way that the water peak (4.78 ppm) has almost disappeared.

Figure 26. (A) Energy levels and population of a homonuclear AX system. (B) Connections between the energy levels of
the system. (C) Populations of the energy levels
immediately after the saturation of the S transitions.
(D) Initial direction of cross-relaxation after saturation

Figure 27. The plot of NOE in a two-spin system (f_i(S)) versus
log ωτ_c, where ω is the Larmor frequency and τ_c is
the molecular correlation time. From Bothner-By
(1979). ............................................................................................... 69

Figure 28. Identification of base pair types by NOE. The NOE
difference spectrum from each base pair (A•U, G•C,
or G•U) has a characteristic feature described in the
text. .................................................................................................. 71

Figure 29. NOESY spectrum of 1 mM B. megaterium 5S rRNA
Fragment A in 10 mM sodium phosphate, 100 mM
NaCl, 1 mM EDTA, pH 7.0 at 23 °C. A 1331 pulse
sequence was used for water suppression.
Experimental conditions to obtain this spectrum are
given in Chapter 4. Some cross-peaks between imino
and amino or aromatic protons are visible. Cross-peaks
between imino protons and water arising from proton
exchange are clearly visible. .......................................................... 74

Figure 30. Isolation of Ribosomal RNAs from B. megaterium
cells. .................................................................................................. 79

Figure 31. Elution profile obtained from DE-32 ion-exchange
chromatography of B. megaterium crude RNA extract.
Experimental conditions, column size, and buffer
solution are described in the text. .................................................. 80

Figure 32. Elution profile obtained from Sephadex G-75 gel-
filtration chromatography of B. megaterium RNAs.
Fractions of large RNAs, 5S rRNA, and tRNA are
specified in the figure. ..................................................................... 82

Figure 33. Purification of B. megaterium 5S rRNA by DE-32 ion-
exchange chromatography and Sephadex G-75 gel-
filtration chromatography. ......................................................... 84

Figure 34. (A) B. megaterium RNAs before purification (left) and
purified B. megaterium 5S rRNA (right) visualized on
dried gels. (B) Purified B. megaterium 5S rRNA shown
by a gel scan. ............................................................................... 88
Figure 35. UV absorption spectrum (absorbance versus wavelength) of *B. megaterium* 5S rRNA. ................................. 90

Figure 36. Normalized A$_{260}$ UV hyperchromism for *B. megaterium* 5S rRNA in the absence and presence of 10 mM Mg$^{2+}$. ............................................................................. 92

Figure 37. Normalized A$_{260}$ hyperchromism for *B. megaterium* 5S rRNA in the absence and presence of 10 mM Mg$^{2+}$. .................................................................................................. 94

Figure 38. 500-MHz downfield proton NMR spectrum of *B. megaterium* 5S rRNA at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95 %:5 % H$_2$O:D$_2$O. The peaks are labeled from A to O for convenience. ..................................................................................... 96

Figure 39. 500 MHz proton NOE difference spectra of resonances J, E, K$_1$, and M of *B. megaterium* 5S rRNA. The G·U base pair (resonances K$_1$/M) is connected to a G·C base pair (E) which is connected to another G·C pair (J). This segment is assigned as G$_4$·C$_{112}$ - G$_5$·C$_{111}$ - U$_6$·G$_{110}$ in helix I. .......................................................................................................... 98

Figure 40. 500 MHz downfield proton NMR spectra of *B. megaterium* 5S rRNA in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 at various temperatures. Peaks J, and E are stable to ~50 °C, and have about the same melting onset temperature as peaks K$_1$ and M. .................................................................................................... 99

Figure 41. 500-MHz proton NOE difference spectra of resonances A, F, K, and L of *B. megaterium* 5S rRNA. Resonances K$_2$ and L arise from another G·U base pair. ......................... 101

Figure 42. Elution profile from Sephadex G-50 gel-filtration chromatography of RNase T1-digested fragments of *B. megaterium* 5S rRNA. Fractions of Fragments A, B, and C are specified. ................................................................. 107

Figure 43. (A) Visualized bands of RNase T1-cleaved fragments on a nondenaturing electrophoresis gel obtained by gel scan. The second peak is Fragment A and the third one represents Fragments B and C which cannot be resolved in this condition. (B) RNase T1-cleaved fragments are
shown on a dried gel obtained by a gel dryer. A, B, and C denote Fragments A, B, and C, respectively. ............... 108

**Figure 44.** Proposed secondary structure segments of RNase T1-cleaved Fragment A (top), B (middle), and C (bottom) of *B. megaterium* 5S rRNA obtained by RNA T1 mapping. While Fragment A consists of one homogeneous single chain, Fragments B and C show heterogeneity. ................................................................. 115

**Figure 45.** The downfield proton NMR spectrum of Fragment A at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95%:5% H2O/D2O. .................................................. 117

**Figure 46.** NOE difference spectra of resonances A1, F1, I", and B of Fragment A obtained at 23 °C. ................................. 119

**Figure 47.** NOE difference spectra of resonances I, J1, H, and D1 of Fragment at 23 °C. .......................................................... 120

**Figure 48.** Heat-induced melting pattern of the proton 500-MHz NMR spectrum of Fragment A in 10 mM cacodylic acid, 0.1 M NaCl, and 1 mM EDTA at pH 7.0. ......................... 123

**Figure 49.** Comparison of the NOE connectivities of intact 5S rRNA with those of the corresponding enzyme-cleaved Fragment A. These homologous NOE connectivities in intact 5S rRNA and Fragment A strongly support the retention of secondary structure on enzymatic cleavage. ......................................................... 126

**Figure 50.** Two possible conformations for *B. megaterium* 5S rRNA Fragment A. Switch between these two conformations is not observed in the present experiments; rather Fragment A remains as "conformation 1" under all solution conditions. ......................................................... 128

**Figure 51.** Variable temperature experiments of Fragment A with a 500-MHz proton NMR spectrometer. Fragment A in 0.01 M cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 was cooled down to specified temperatures to monitor its structural change. ...................... 129

**Figure 52.** Mg²⁺ titration experiments of Fragment A in 0.01 M sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0. The NMR spectra were obtained at 23 °C with various concentrations of Mg²⁺. ......................................................... 130
Figure 53. Variation with temperature of relative peak heights of base-paired imino proton resonances in the 500-MHz ¹H NMR spectrum of Fragment A. Vertical scaling has been chosen for convenience in discussion. .......................... 131

Figure 54. Variation with temperature of chemical shifts of base-paired imino proton resonances in the 500-MHz ¹H NMR spectrum of Fragment A. .......................................... 132

Figure 55. Variation with Mg²⁺ concentration of relative peak heights of base-paired imino proton resonances in the 500-MHz ¹H NMR spectrum of Fragment A. ............................. 135

Figure 56. Variation with Mg²⁺ concentration of chemical shifts of base-paired imino proton resonances in the 500-MHz ¹H NMR spectrum of Fragment A. ..................... 136

Figure 57. NOE difference spectra of resonances A₁, B, D, and F₁ in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with 15 mM Mg²⁺ obtained at 23 °C. ... 139

Figure 58. NOE difference spectra of resonances H, I', I, and I" in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with 15 mM Mg²⁺ obtained at 23 °C. ... 140

Figure 59. NOE difference spectra of resonances A₁, B, D, and F₁ in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 at 3 °C. ................................................................. 141

Figure 60. NOE difference spectra of resonances H, I, and I" in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with 15 mM Mg²⁺ obtained at 3 °C. ............................. 142

Figure 61. Cross-peaks between imino and amino or aromatic protons observed in NOESY spectrum of 1 mM B. megaterium 5S rRNA Fragment A in 10 mM sodium phosphate, 100 mM NaCl, 1 mM EDTA, 15 mM MgCl₂, pH 7.0 at 23 °C. A 1331 pulse sequence was used for water suppression. Experimental conditions are given in the text. ................................................................. 144

Figure 62. The downfield proton NMR spectrum of Fragment B at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95%:5% H₂O/D₂O. ............................. 149

Figure 63. The NOE difference spectra of resonances A₂, D₂,
and $M_2$ of Fragment B. It is evident that $D_2 (G\cdot C)$ is connected to $A_2 (A\cdot U)$ and $M_2$ (unpaired U), and $M_2$ is connected to $D_2$ and $O$ (unpaired U). Irradiation of $A_2$ shows NOE connectivity to $D_2$ and I" (G•C).

**Figure 64.** The NOE difference spectra of resonances $O$ and $M_1$ of Fragment B. From Figures 63 and 64, the sequence, I" (G•C) - $A_2$ (A•U) - $D_2$ (G•C) - $M_2$ (U) - $O$ (U) - $M_1$ (U), can be confidently assigned.

**Figure 65.** The NOE difference spectra of resonances I'$_2$, I, and I" of Fragment B. Though it is difficult to tell NOE connectivities between the three resonances from these spectra, it is evident that I" (G•C) is connected to $A_2$ (A•U), as can be seen here. NOE connectivities observed in Figures 63-65 can be sufficient evidence to establish complete helical segment of the prokaryotic loop.

**Figure 66.** Variable temperature experiments of Fragment B in 0.01 M cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with a 500-MHz proton NMR spectrometer. As can be seen in the figure, the resonances melt in the order I'$_2$ (G$_{80}$•C$_{92}$), I (G$_{81}$•C$_{91}$), I" (G$_{82}$•C$_{90}$) and $D_2$ (C$_{84}$•G$_{88}$). This is a good indication that the unwinding of helix IV begins with the breakage of the base pair G$_{80}$•C$_{82}$.

**Figure 67.** The downfield proton NMR spectrum of Fragment C at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95 %:5 % H$_2$O/D$_2$O.

**Figure 68.** The NOE difference spectra of resonances $E$, $J$, $K_1$, and $M$ of Fragment C. The sequence of base pairs $G_4$•$C_{112}$ - $G_5$•$C_{111}$ - $U_6$•$G_{110}$, previously inferred from NOE difference spectra of intact 5S rRNA, can now be confirmed with greater confidence from these more highly resolved NOE difference spectra of Fragment C.

**Figure 69.** Variable temperature experiments of Fragment C in 0.01 M cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with a 500-MHz proton NMR spectrometer. As can be seen here, it is evident that resonances $E$,
J, K, M (G₄·C₁₁₂ − G₅·C₁₁₁ − U₆·G₁₁₀) are the most stable resonances of all presumably due to their locations (middle of the helix).

Figure 70. 500 MHz ¹H NMR spectrum of intact 5S rRNA (bottom) and composite spectrum obtained by the addition of the spectra of Fragments A, B, and C (top). The close match in chemical shifts between the two spectra (taking into account the broader lines for the larger intact 5S rRNA molecule) offers further evidence for high conformational similarity between intact 5S rRNA and its fragments. Some peaks present in intact 5S rRNA are absent in the combined spectrum (e.g. peak C). The missing peaks in the combined spectrum are believed to result from secondary base-pairs in helix V and from tertiary base pairs.

Figure 71. Stereoscopic view of the helices II-III region of B. megaterium 5S rRNA after the structure is minimized. Hydrogens are omitted. Bases in hairpin loop area (lowest part of the structure) are well-stacked. However, bases in helix III are not base-stacked.

Figure 72. Superposition of structures (the helices II-III region) from a number of trajectory files during dynamics run. The helix III region experiences the most significant change as shown.

Figure 73. Stereoscopic view of the helices II-III region of B. megaterium 5S rRNA after dynamics run. Compared to the structure before dynamics (Figure 71), bases in the helix II region are stacked well in the shape of A-helix.

Figure 74. All-atom structure (all hydrogens included) of the helix II-III region of B. megaterium 5S rRNA.

Figure 75. Susceptibility of the helix III region to structural change in the absence of constraints. The helix II region is very stable.

Figure 76. All-atom stereoscopic view of the helix IV region of B. megaterium 5S rRNA after the structure is minimized. Initially 14 constraints were included for a short time, then minimization was continued without constraints. Five base pairs in the helix are well-preserved.
suggesting the stability of the helix. U$_{87}$ is sticking out of the loop. ................................................................. 188

**Figure 77.** Superposition of structures (helix IV) from a number of trajectory files during dynamics run. Convergence to more stable conformation is monitored. .......................... 190

**Figure 78.** All-atom stereoscopic view of helix IV region of *B. megaterium* 5S rRNA after dynamics run. No significant change from the previous view can be found, suggesting a very stable conformation of the structure. ................................................................. 191
A. Objectives

Ribosomes are molecular complexes consisting of a number of proteins and RNAs (ribonucleic acids). They play an essential role in the biosynthesis of proteins, and are the only cell organelles that occur in all organisms. A necessary prerequisite for understanding the function of ribosomes at the molecular level is knowledge of their structures. Therefore, investigation of the structure of ribosomes is very important for elucidation of the mechanism of biosynthesis of proteins at the molecular and cellular level.

Due to their biological importance, ribosomes have been extensively studied for more than two decades. Various chemical, physical, immunological, and genetic methods have been used to investigate the structure of ribosomes and their contents. These efforts have contributed to the understanding of ribosomal structure. Table 1 summarizes size and composition of the ribosomes from various species. Prokaryotic ribosomes can be divided into two subunits, 50S and 30S (Tissieres & Watson, 1958). Each subunit
Table 1. Size and composition of the ribosomes from various species. From Wittmann (1986).

<table>
<thead>
<tr>
<th>Organism</th>
<th>Organelle</th>
<th>Ribosome</th>
<th>Proteins</th>
<th>RNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td></td>
<td>70S</td>
<td>50-60</td>
<td>5S; 16S; 23S</td>
</tr>
<tr>
<td>Plants</td>
<td>Chloroplasts</td>
<td>70S</td>
<td>50-60</td>
<td>(4.5S); 5S; 16S; 23S</td>
</tr>
<tr>
<td>Plants</td>
<td>Mitochondria</td>
<td>75S</td>
<td>Not known</td>
<td>5S; 18S; 26S</td>
</tr>
<tr>
<td>Protozoans and Fungi</td>
<td>Mitochondria</td>
<td>75S</td>
<td>Not known</td>
<td>15-17S; 21-24S</td>
</tr>
<tr>
<td>Mammals</td>
<td>Mitochondria</td>
<td>55S</td>
<td>80-90</td>
<td>12S; 16S</td>
</tr>
<tr>
<td>Eukaryotes (Cytoplasm)</td>
<td>Mitochondria</td>
<td>80S</td>
<td>75-90</td>
<td>5S; 5.8S; 17-18S; 26-28S</td>
</tr>
</tbody>
</table>
consists of many ribosomal proteins and several ribosomal RNAs, as can be seen in Figure 1.

Ribosomal RNAs, which comprise over 50% of the mass of the ribosome, were long believed simply to provide a structural framework for assembling ribosomal proteins into their appropriate location in the ribosome. However, several lines of evidence obtained from recent studies have led researchers to believe that ribosomal RNAs play central roles in ribosomal functions (Noller, 1984). In spite of the wide acceptance of their essential roles in protein synthesis, the precise functions of ribosomal RNAs are uncertain.

As mentioned above, insight into the structures of ribosomal RNAs is necessary for a better understanding of protein synthesis in ribosomes. 5S rRNA, an important structural and functional component of the large subunit of all ribosomes, is the smallest ribosomal RNA molecule. Because of its small size and its essential role in protein synthesis in the ribosome (Pieler et al., 1984), understanding the structure and function of 5S rRNA seems to be the first step in the elucidation of ribosomal function. The interpretation of the structure of 5S rRNA molecules can be facilitated from tRNA structural information obtained from X-ray crystal structure determination. Moreover, some information about the secondary structure of Escherichia coli (a Gram-negative bacterium) 5S rRNA is already available. For this reason, studies of the structure of 5S rRNA from Bacillus megaterium (a Gram-positive bacterium) are necessary for an understanding of the structure and function of prokaryotic 5S rRNAs and for a better description of the mechanism of protein synthesis.
Figure 1. Components of prokaryotic and eukaryotic ribosomes. Numbers in brackets are for eukaryotic ribosomes. From Erdmann et al. (1980).
B. Structures of RNA Molecules

The structural features of RNA molecules and nomenclature have been thoroughly reviewed by Saenger (1984). The contents and figures in this section are largely derived from his summary.

1. Chemical Components of RNA and Nomenclature

All kinds of RNAs (transfer RNA, messenger RNA, and ribosomal RNA) consist of a repeating basic component, the nucleotide. It is composed of a cyclic, furanoside-type sugar (β-D-ribose) which is phosphorylated in the 5' position. One of the four bases (adenine, guanine, cytosine, or uracil) is connected to C₁ of the sugar through a β-glycosyl C₁-N linkage (Figure 2). When a base is combined with a sugar ring, it is called a nucleoside (adenosine (A), guanosine (G), cytidine (C), or uridine (U)). Besides these nucleosides (A, G, C, and U), a number of modified nucleosides have been found to exist naturally and have been chemically synthesized. However, these additional nucleosides are seldom found in ribosomal RNA molecules.

When a nucleoside is phosphorylated at one of the free sugar hydroxyls, it becomes a nucleotide. Several isomers with phosphate at the 2', 3', and 5' positions are possible. These nucleotides can be linked together by phosphodiester bonds to form a ribonucleic acid. In naturally occurring RNAs, the individual nucleotides are generally linked via 3', 5'-phosphodiester bonds (see Figure 2).

The atomic numbering scheme for purine (adenine and guanine) and pyrimidine (cytosine and uracil) bases is shown in Figure 2. Sugar
Figure 2. Ribonucleic acid (RNA) consists of adenosine (A), guanosine (G), cytidine (C), and uridine (U), which are connected together via 3', 5'-phosphodiester bonds. The chain direction is from 5' to 3' as shown in the figure. From Saenger (1984).
atoms are distinguished from base atoms by a prime. As can be seen from Figure 3, the counting direction of atoms of a polynucleotide backbone is P - O₅' - C₅' - C₄' - C₃' - O₃' - P, and the numbering sequence in a sugar ring is C₁' - C₂' - C₃' - C₄' - O₄' - C₁'. The conformation of the sugar-phosphate backbone is described by torsion angles α, β, γ, δ, ε, ζ in alphabetical order. The endocyclic torsion angles of the sugar are named υ₀ to υ₄, and the torsion angle between a base and a sugar is given by χ (O₄' - C₁' - N₁ - C₂ for pyrimidines and O₄' - C₁' - N₉ - C₄ for purines).

The sugar ring is usually not planar; rather it can be puckered in an envelope form (E) or in a twist form (T). When atoms are displaced from these three- or four-atom planes and on the same side as C₅', the sugar conformation is called endo; when the atoms are opposite C₅', the sugar conformation is called exo. For ease of notation, superscripts for endo atoms and subscripts for exo atoms usually precede or follow the letter E or T, for major or minor puckering. For example, ²E is a C₂'-endo envelope, ³T₂ is an unsymmetrical C₃'-endo- C₂'-exo twist with major C₃' and minor C₂' puckering (Figure 4).

2. Base Pairs and RNA Helices

Within RNA molecules, hydrogen bonding between A and U, or G and C (Watson-Crick base-pairing), or sometimes between G and U ("wobble") is possible. Therefore, a double helix is produced by complementary base-pairings between the two polynucleotide chains. In general, RNA double helices display pseudodyad symmetry (Figure 4). Base pairs are not centered but rather displaced from the helix by
Figure 3. Atomic numbering scheme and description of torsion angles in a polyribonucleotide chain. From Saenger (1984).
Figure 4. (Left) Sugar puckering modes in RNA. (a) Flat five-membered sugar, which was never observed. (b) C$_3$-endo. (c) C$_2$-endo. (d) C$_2$-exo-C$_3$-endo. (e) Major C$_3$-endo and minor C$_2$-exo. (Right) Definition of major and minor grooves in A•U and G•C base pairs. From Saenger (1984).
distance $D$ (Figure 5). In addition, they are inclined at tilt angle $\theta_T$ and roll angle $\theta_R$ from the plane perpendicular to the helix axis. The two bases in a base pair usually have a propeller twist $\theta_P$.

The outer envelope of the double helix is not cylindrically smooth, but rather displays two grooves of different width and depth (see Figure 4), caused by the branching off of the two glycosyl bonds from one side of the base pairs and the displacement of base pairs (by $D$) from the helix axis. The O$_2$ side of a pyrimidine or N$_3$ side of a purine of base pairs is named the minor groove, while the opposite side is called the major groove.

According to the salt concentration, RNA double helices can form two major, structurally similar conformations. The A-RNA, which is favored at low ionic strength, has 11 base pairs per helix turn. At higher salt concentration, the helix is transformed into A'-RNA which has a 12-fold helix. A third conformation, A''-RNA, has been proposed by Arnott et al. (1968), but has not been thoroughly analyzed. The RNA molecule can also form a B-helix corresponding to the typical 10-fold DNA helix.

A-RNA is recognized as the standard conformation of RNA helices (Figure 6). This type is a right-handed double helix with antiparallel arrangement of the two polynucleotide chains. Base pairs are displaced by 4.4 Å from the helix axis with a very deep major groove and a shallow minor groove. The pitch height is about 30 Å, 11 nucleotides per turn. The axial rise per residue is around 2.73 to 2.81 Å, and the tilt angle of the base pair is between 16° to 19°. The sugar conformation of the A-RNA is usually 3'-endo.
Figure 5. (Top) Definition of propeller twist $\theta_p$, which is dihedral angle between two base planes. $\theta_p$ is defined positive if the closer base is rotated clockwise with respect to the far one when we look down the long axis of a base pair. (Bottom) Definition of twist $t$, tilt $\theta_T$, and roll $\theta_R$. The pseudodyad is reference line for $t$ and $\theta_T$, and a vertical to the pseudodyad passing approximately through $C_6$ (pyrimidine) and $C_8$ (purine) is for $\theta_R$. Positive signs for these angles are clockwise rotations as shown in the figure. From Saenger (1984).
Figure 6. Molecular structure of A-RNA described by ball-and-stick model. (Left) Viewed from perpendicular (top) and parallel (bottom) to the helix axis. (Right) The double helix is tilted by 32° to show the major (M) and the minor (m) grooves. From Saenger (1984).
C. Proposed Secondary Structure Models of 5S rRNA

Since the late 1960s, a large number of 5S rRNAs from many different prokaryotic and eukaryotic ribosomes has been sequenced. So far, complete base sequences of 5S rRNAs from more than 500 organisms have been compiled (Specht et al., 1990; Wolters & Erdmann, 1988). However, unlike the tRNA molecule, whose cloverleaf model was obvious since its first primary sequence was determined, no unique base-pairing scheme for 5S rRNA with computer programs was possible on the basis of the primary sequence alone (Richards, 1969; Jordan, 1971). Difficulty in the establishment of consensus secondary structural model for 5S rRNA was primarily due to its larger number of base-pairing possibilities, arising from longer chain length, as well as the uncertain functional role of 5S rRNA in the ribosome.

In spite of these difficulties, there has been a general belief that there should be a universal secondary structure for 5S rRNAs. Many universal secondary structure models have been proposed on the basis of comparative sequence analysis (Fox & Woese, 1975; Luehrsen & Fox, 1981; Studnicka et al., 1981; de Wachter et al., 1982; Delilhas & Anderson, 1982), enzymatic cleavage and chemical modification (Nishikawa & Takemura, 1974a; Pieler & Erdmann, 1982), and physical measurements (Kearns & Wong, 1974; Osterberg et al., 1976; Luoma & Marshall, 1978a,b; Chang et al., 1984). As can be seen in Figure 7, all secondary structure models except one are based on the Fox and Woese model (Fox & Woese, 1975), which has three stems.
Figure 7. Derivation of the most popular proposed models of 5S rRNA secondary structure.
Most of the subsequent models have been built by adding just a few extra base-pairings to this basic model. Only the secondary structure model proposed by Luoma and Marshall (1978a,b) has a four-stem cloverleaf structure with extensive base-pairings.

The minimal model of 5S rRNA secondary structure, on which most researchers agree, was first proposed by Nishikawa and Takemura (1974b) for eukaryotes, and by Fox and Woese (1975) for prokaryotes. Based on investigation by Studnicka et al. (1981), MacKay et al. (1982), Delihas and Anderson (1982), de Wachter et al. (1982), Kuntzel et al. (1983), a common secondary structure of all 5S rRNAs has been proposed (Erdmann & Wolters, 1986) (Figure 8). This minimal model consists of five helices (named A to E), connected by loops named a to e in three stems.

D. Investigations of Three-Dimensional Structure of 5S rRNA

There also have been many reports on the tertiary structure of 5S RNA. Small-angle X-ray scattering of *E. coli* 5S rRNA was interpreted to show that *E. coli* 5S rRNA consists of one large and two small double helices arranged in the form of the letter Y with tertiary interaction between G20 and C71, as shown in Figure 9 (top) (Osterberg *et al.*, 1976). Fox and Wong (1979) used hydrodynamic measurements, sedimentation equilibrium, and UV absorption to study the three-dimensional structure of 5S rRNA. According to their report, 5S rRNA consists of three double helices in the shape of a prolate ellipsoid, with two of the double helical regions at one end of the molecule (Figure 9, bottom). Intramolecular crosslinking of *E. coli*
Figure 8. The minimal model of eubacterial 5S rRNA secondary structure. Squares indicate conserved base-pairings, while circles indicate unpaired nucleotides. From Erdmann & Wolters (1986).
Figure 9. (Top) A tertiary Y-shaped 5S rRNA structure proposed based on a small-angle X-ray scattering study. From Osterberg et al. (1976). (Bottom) A three-dimensional 5S rRNA structure proposed based on hydrodynamic studies. From Fox and Wong (1979).
5S rRNA was investigated by Hancock and Wagner (1982) by using a chemical reagent. They reported that 5S rRNA has four helices (I-IV) (Figure 10, top). McDougall and Nazar (1983) studied the tertiary structure of eukaryotic 5S rRNA by using ethylnitrosourea reactivity as a probe for phosphodiester bonds. Luck and Erdmann (1983) showed that *E. coli* 5S rRNA *in situ* in the 50S ribosomal subunit reacted with monoperphthalic acid and was isolated from the subunit to measure chemical reactivity of *E. coli* 5S rRNA. They proposed a three-dimensional structure model for *E. coli* 5S rRNA (Figure 10, bottom), which shows tertiary interaction between nucleotides around 43 and 75.

In spite of the amount of work done thus far, none of the models for the three-dimensional structure of 5S rRNAs has strong evidence or supporting data. As a matter of fact, it has proved difficult even to conceive a complete *secondary* structure model.

**E. Methods for Structural Determination of 5S rRNA**

Various methods have been applied to investigate the structure of ribosomal 5S rRNA and to establish a universal secondary structure model. Spectroscopic techniques like ultraviolet (UV) absorption and circular dichroism (CD), infrared (IR) absorption, Raman scattering, and NMR have been used to identify base pairs, stacking of bases, and to estimate the number of base pairs. Other techniques such as limited enzymatic hydrolysis and chemical modifications have been applied to identify single-stranded bases or segments. The gross shape of ribosomal RNA molecules has been ascertained through
Figure 10. (Top) A three-dimensional 5S rRNA structure proposed based on a study with chemical reagent phenyldiglyoxal (PDG). From Hancock & Wagner (1982). (Bottom) A three-dimensional 5S rRNA structure proposed based on chemical reactivity studies with monoperphthalic acid. Interaction sites for tRNA and other ribosomal RNAs are shown in (a), and the binding sites for ribosomal proteins are shown in (b). From Luck & Erdmann (1983).
hydrodynamics and small-angle X-ray scattering, as discussed in the previous section. Figure 11 describes various techniques used for the investigation of the secondary structure of 5S rRNA. The approaches which have contributed significantly to the establishment of the structure of 5S rRNA will be discussed below.

1. Comparative Sequence Analysis

The primary sequences of 5S rRNAs from various organisms can be aligned and compared with each other to locate common helical segments. Once they are well-aligned on the basis of sequence homology, base changes between the compared sequences can be analyzed. If compensating base changes maintain complementarity (A·U, G·C, or G·U) between two potential pairing regions, that position is considered to be within the helix. However, if the compensating base changes are mismatches, that position is not part of a helix (Figure 12).

Precise sequence alignment is not easy, and has often been aided by computers. Secondary structure models have been built on the basis of free energy minimization algorithms (Tinoco et al., 1971; Pipas & McMahon, 1975; Studnicka et al., 1978; Nussinov & Jacobson, 1980; Zuker & Stiegler, 1981; Studnicka et al., 1981; de Wachter et al., 1982). The values for nearest-neighbor stacking energies of unpaired loops, hairpin loops, bulges, and mismatches have been empirically estimated based on helical RNA oligomers, and added for free energy minimization calculations (Tinoco et al., 1973).
Figure 11. Various experimental techniques used to investigate the structure of 5S rRNA.
**Figure 12.** One typical example of comparative sequence analysis. The sequence of 5S rRNAs from ten different organisms are aligned and compared with each other to obtain maximum sequence homology. The sequences in boxes are presumably base-pairing regions. Highly phylogenetically conserved regions are shown in brackets. From Fox & Woese (1975).
Comparative sequence analysis is the approach which is the basis of most of the proposed secondary structure models. This approach can be used only under the assumption that the 5S rRNA molecules under comparison have the same secondary structure, which is the basic idea of the universality of the secondary structure of 5S rRNAs. This approach has produced a large number of secondary structure models, but other experimental techniques are necessary to support the validity of these models.

2. **Enzymatic Accessibilities to 5S rRNA**

5S rRNA can be digested partially by a number of ribonucleases specific to single-stranded or double-stranded regions of RNA molecules. Fragments produced by limited enzymatic hydrolysis can be analyzed by RNA sequencing methods (Douthwaite & Garrett; 1981). Investigation of accessibilities of ribonucleases into 5S rRNA under various conditions can give valuable information about the structure of the RNA molecules. The ribonucleases used as structure probes are as follows: RNase A specific for C or U residues, RNase T1 specific for G residues, RNase T2 specific for G or A residues (Vigne et al., 1973; Douthwaite & Garrett, 1981), RNase S1 specific for any single-stranded regions (Ross & Brimacombe, 1979; Pieler & Erdmann, 1982), and cobra venom RNase for double-stranded regions (Douthwaite & Garrett, 1981; Vassilenko et al., 1981).

Cleavage of the sugar phosphate backbone during enzymatic digestion, however, may induce significant conformational change and rearrangement of RNA structure, which then becomes susceptible to
ribonuclease accessibility different from the original. Some researchers distinguish this secondary cleavage from primary cleavage (Douthwaite & Garrett, 1981); only primary cleavage sites should be used to interpret the original structure of the molecule. Unfortunately there is sometimes uncertainty in distinguishing primary cuts from secondary ones. This remains a potential problem for the use of this approach as a structural probe although there is no evidence for occurrence of conformational change during hydrolysis.

Figure 13 shows one example of enzymatic accessibility experiments, and demonstrates that single-stranded and base-paired regions conjectured by ribonuclease hydrolysis experiments are consistent with those from the secondary structure model (Fox-Woese model) based on comparative sequence analysis. Data obtained from enzymatic accessibilities to 5S rRNA can elaborate the secondary structure of 5S rRNA, and can even predict tertiary interaction regions and the folding pattern of the molecule.

3. Chemical Modifications

In addition to limited enzymatic hydrolysis, chemical modifications of RNA molecules have been used to investigate single- and double-stranded regions. Gilham (1962) has reported that uridine-5' phosphate modified by carbodiimide could block pancreatic ribonuclease activity of uridine 3'-phosphoryl bond cleavage (Figure 14, A). Yeast RNA was treated with carbodiimide, then pancreatic ribonuclease, and after removal of the enzyme the blocking groups were also removed. When the obtained fragments were compared
Figure 13. (Top) The modified Fox and Woese secondary structural model of *E. coli* 5S rRNA. Arrows show the primary cutting positions of ribonucleases A, T1 and T2, which are single-stranded specific. The arrow size represents the degree of cutting. (Bottom) The modified Fox and Woese secondary structure model of *B. stearothermophilus* 5S rRNA. Arrows indicate the cutting positions of ribonucleases A, T1, T2, and cobra venom. From Douthwaite & Garrett (1981).
Figure 14. (A) Modification of uridine 5'-phosphate with carbodiimide (I), where one of the R groups contains a closely situated quaternary ammonium group. The predicted structure of the product is II. From Gilham (1962). (B) The selective modification of cytidine by O-methylhydroxylamine. From Kochetkov et al. (1963). (C) Selective 1-N-oxidation of adenine by monoperphthalic acid. Strong inhibition of oxidation was observed when adenines were base-paired. From Cramer & Seidel (1964).
with those from RNA with no carbodiimide treatment, uridine 2' (3')-phosphate present in the hydrolysate was found to be reduced by 93% with the modification of uridines with carbodiimide. The protection of the majority of uridine bases from ribonucleolytic attack in this way can permit the isolation of larger fragments from the enzymatic degradation of RNA.

Kochetkov et al. (1963) have reported that cytidine can be selectively modified by O-methylhydroxylamine by adjusting the pH of the reaction mixture (Figure 14, B). Selective 1-N-oxidation of adenine by monoperphthalic acid has been accomplished by Cramer and Seidel (1964). They also observed strong inhibition of 1-N-oxidation of adenine units after the formation of the poly-(A+U) double-stranded helix due to hydrogen bond formation between uracil 3-N positions and adenine 1-N positions. The time dependence of the oxidation of adenine derivatives in a double-strand thus can introduce the possibility of estimating the portion of nucleotides not bound as a complex; such an approach could be applicable to the elucidation of the secondary structure of different RNAs (Figure 14, C). Silberkang et al. (1983) also have reported selective oxidation of adenine by monoperphthalic acid. Woese et al. (1980) have used m-chloroperbenzoate for selective modification of adenine.

Guanine bases of ribosomal RNA have been selectively modified by kethoxal and analyzed by a diagonal electrophoresis method (Noller, 1974). Protection of some of these sites from reaction with kethoxal by bound mRNA or tRNA would provide evidence for the possible interaction of these ligands with specific RNA sequences. Dimethyl
sulfate alkylates the N-7 position of guanosines (Fedoroff et al., 1977) and the N-3 position of cytosines (Agliano et al., 1979); diethyl pyrocarbonate carbethoxylates the N-7 of adenosines (Jeffreys & Flavell, 1977). The diethyl pyrocarbonate reaction is believed to be especially sensitive to the stacking of adenosines. Therefore these reagents effectively evaluate the conformation by providing structure-specific data as well as base-specific conditions. These chemical reagents have been used to probe the secondary and tertiary interactions of RNA molecules in solution (Peattie & Gilbert, 1980).

A number of chemical reagents has provided sensitive probes of the conformation of RNAs in solution. With base-specific reagents and terminally labeled polynucleotides, even a single experiment can determine the properties of a base wherever it occurs in the molecule.

4. Binding of Oligonucleotides to RNA Molecules

The specific binding of complementary oligonucleotides to RNA molecules of known sequence has been used to locate stacked single-stranded regions in partially double-stranded RNA. This method has been applied to investigate structures of uncharged tRNAs (Hogenauer, 1970; Uhlenbeck et al., 1970, 1974; Pongs et al., 1973; Freier & Tinoco, 1975), amino-acylated tRNA (Danchin & Grunberg-Manago, 1970; Pongs et al., 1976; Schwarz et al., 1976), and 5S rRNAs (Lewis & Doty, 1970; Erdmann et al., 1973; Mankin et al., 1981).

Lewis and Doty (1970) have asserted that specific sequences in E. coli 5S rRNA can be identified as single-stranded from oligonucleotide binding. From studies of oligonucleotides binding to E. coli 5S rRNA,
they have identified five single-stranded regions in the native conformation of the molecule. Those regions are bases 10-14, 28-31, 39-49, 60-62, and 78-82 (as shown in Figure 15, A), in support of the Fox and Woese model (Fox & Woese, 1975).

Wrede et al. (1978) have described similar experiments on E. coli 5S rRNA and Bacillus stearothermophilus 5S rRNA. Sequences around positions 10, 30, 60, 70, 85, and 95 have been reported to be in a single-stranded conformation in both 5S rRNAs (Figure 15, B), thus providing additional supporting evidence for the conservation of the overall structure of bacterial 5S rRNA during evolution. However, Wrede et al. concluded that the oligonucleotide binding data for E. coli 5S rRNA did not fully support any previously proposed structural model.

One potential problem in this method may be the perturbation of the RNA structure by oligomers, especially when the stability of an oligomer-RNA complex is significantly greater than the original structure of the RNA molecule. Generally, there has been inconsistency between the suggested locations of single-stranded regions of RNA molecules, which are based on oligonucleotide binding and chemical modification or enzymatic hydrolysis.

5. Hydrodynamic and X-ray Scattering Studies

As mentioned previously, hydrodynamic and X-ray scattering techniques have been applied to structural studies of 5S rRNA. Fox and Wong (1979) have proposed a three-dimensional structure of E. coli 5S rRNA based on several physical/biophysical techniques,
Figure 15. (A) Different accessibilities of native and A form of *E. coli* 5S rRNA. The secondary structure of *E. coli* 5S rRNA is drawn according to the Fox and Woese model. From Lewis & Doty (1977). (B) Proposed oligonucleotide binding sites for prokaryotic 5S rRNAs based on nucleotide binding studies. From Wrede et al. (1978).
including sedimentation velocity, sedimentation equilibrium, viscosity, circular dichroism, UV absorption spectroscopy, and melting studies. According to Fox and Wong, 5S rRNA has the shape of a prolate ellipsoid 160 Å in length and 32 Å wide, which is consistent with Osterberg et al.'s (1976) previously proposed three-dimensional model of 5S rRNA based on small-angle X-ray scattering studies.

6. UV, IR, and Raman Spectroscopy

Ultraviolet (UV) absorption spectroscopy has been used to approximate the number of base pairs present in 5S rRNA (Fresco et al., 1963), and the percentages of G·C and A·U base pairs are calculated by using ultraviolet absorption difference spectra (Fox & Wong, 1979). Infrared (IR) spectroscopy has also been applied to the quantitation of base pairs (Stulz et al., 1981; Bohm et al., 1981). With the introduction of FT-IR, it was possible to approximate the number of base pairs in 5S rRNA more precisely by high resolution FT-IR spectra (Burkey et al., 1983; Li et al., 1984; Chang et al., 1984).

Raman spectroscopy has been recognized as a sensitive technique to probe the secondary and tertiary structure of RNA molecules in solution (Peticolas, 1972; Thomas, 1975) for the determination of the types and extent of base stacking, the percentage of U residues in base-paired versus single-stranded regions, the degree of order at the backbone phosphate linkages (Thomas & Hartman, 1973; Chen & Thomas, 1974), and reflection of changes in RNA secondary or tertiary conformation (Chen et al., 1975). E. coli 5S rRNA has been examined by Raman spectroscopy (Chen et al., 1978), and Luoma and Marshall
have proposed three-stem cloverleaf secondary structure models for prokaryotic (Luoma & Marshall, 1978a) and eukaryotic (Luoma & Marshall, 1978b) 5S rRNAs based on laser Raman results (Figure 16).

7. Nuclear Magnetic Resonance (NMR) Spectroscopy

Of the many spectroscopic techniques, NMR spectroscopy has been most useful for investigation of the solution structure of nucleic acids. The first proton NMR spectrum of 5S RNA was obtained from yeast 5S rRNA at 200-MHz field strength (Wong et al., 1972). Soon a number of papers were published about 5S rRNAs: E.coli 5S rRNA by 300-MHz $^1$H NMR (Kearns & Wong, 1974) and by 400-MHz $^1$H NMR (Burns et al., 1980), yeast 5S rRNA by 360-MHz $^1$H NMR (Luoma et al., 1980), and Bacillus licheniformis 5S rRNA by 360-MHz $^1$H NMR (Salemink et al., 1981).

Proton nuclear Overhauser enhancement (NOE) has become the most powerful experiment for the determination of base-paired sequences in nucleic acids through application to tRNA (Johnson & Redfield, 1978). Supplemented by the complete structural determination of tRNA by X-ray crystallography (Kim, 1976; Rich, 1977), NOE experiments have successfully identified (as A·U, G·C, or G·U) and assigned (to specific primary sequence bases) virtually all of the secondary and tertiary base pairs in several tRNAs in aqueous solution (Schimmel & Redfield, 1980; Reid, 1981; Johnston & Redfield, 1981; Heerschap et al., 1983a,b; Roy & Redfield, 1983).

Application of NOE techniques to the larger 5S rRNA molecule has presented several problems, as noted in a recent review of $^1$H,
$^{13}$C, $^{15}$N, $^{19}$F, and $^{31}$P NMR of 5S rRNA's (Marshall & Wu, 1990). The main obstacles have been broader peak widths and more overlap of the proton resonances in the downfield region resulting from more base-paired imino protons than in tRNA. Moreover, the secondary structure and the tertiary folding patterns of 5S rRNA have not been established from X-ray diffraction. Although some of the imino proton resonances in the downfield proton NMR spectrum of 5S rRNA have been assigned to specific base pairs in the universal secondary structure models (e.g., Chang & Marshall, 1986a), NOE experiments have had to be supported by other experiments that shifted or simplified the proton NMR spectrum. Among such methods are site-specific spin labeling (Lee & Marshall, 1987) and the use of enzymatic cleavage fragments (Kime & Moore, 1983a,b; Kime et al., 1984; Li & Marshall, 1986; Chen & Marshall, 1986; Li et al., 1987; Kim & Marshall, 1990a,b; Wu & Marshall, 1990a,b). This enzymatic digestion of 5S rRNA into smaller fragments is one of the strategies used in this research, and will be discussed later.

NMR of nuclei other than protons has also been applied for the structural elucidation of 5S rRNA. $^{31}$P NMR has been used to monitor the phosphate backbone (Salemink et al., 1981; Wu & Marshall, 1990b), and $^{19}$F NMR has been applied to the simultaneous monitoring of all of the uracil sites of 5S rRNA, where 5-fluorouracil (5FU) was incorporated to all uracil positions (Marshall & Smith, 1977, 1980). *E. coli* 5S rRNA has been labeled with 4-$^{13}$C-uracil and examined by $^{13}$C NMR spectrum to study the hydrogen bonding of uracils (Hamill et al., 1978). $^{15}$N NMR has been very useful for base
type identification in proton NMR spectra (Gonnella et al., 1982), and has made it possible to identify all of the base-paired (hydrogen-bonded) imino proton resonances of a RNase A-cleaved fragment of *E. coli* 5S rRNA (Kime, 1984a). The $^{15}$N-$^1$H shift correlation method has been useful for identifying new resonances specific to the RNA:protein complex formed between *E. coli* 5S rRNA fragment 1 and ribosomal protein L25 (Kime, 1984b).

**F. Bacillus megaterium 5S rRNA**

*Bacillus megaterium* strain KM (ATCC 13632) is a Gram-positive bacterium. The nucleotide sequence of *B. megaterium* 5S rRNA was first determined by Pribula *et al.* (1974), and comparative sequence analysis of the bacterium relative to other *Bacillus* bacteria has been reported (Woese *et al.*, 1976; Marotta *et al.*, 1976). Even though the Fox-Woese model (Fox & Woese, 1975) has been generally accepted as the universal secondary structure of 5S rRNA, there have been some models which offer slightly different secondary structures for separate phylogenetic groups (e.g., Gram-negative vs. Gram-positive) (Hori & Osawa, 1979; Studnicka *et al.*, 1981). Therefore, comparison of the secondary structures of 5S rRNAs from Gram-negative and Gram-positive bacteria seems to be very important for understanding of the molecules. The secondary structure of prokaryotic 5S rRNA has been studied in *E. coli* (Gram-negative) and *B. subtilis* (Gram-positive) bacteria by NMR techniques. Comparison of the imino proton NMR spectra of two or more 5S rRNAs which have similar base sequences can facilitate interpretation of the NMR spectra, and thus help to
Figure 17. Proposed secondary structure (the Fox and Woese model with an additional helix V) adapted to the primary base sequence of *B. megaterium* 5S rRNA. Heterogeneity of bases in the sequence is denoted as follows: R is G or A, K is G or U, and Y is C or U.
establish the secondary structure (Chen & Marshall, 1986). NMR analysis of \textit{B. megaterium} 5S rRNA is particularly useful, because its primary base sequence (and probably its secondary structure) is very similar to that of previously studied \textit{B. subtilis}. The secondary structure model of \textit{B. megaterium} 5S rRNA adapted to the Fox-Woese model is shown in Figure 17.
A. Resonances of Hydrogen-Bonded (Base-Paired) Imino Protons

5S rRNA consists of 116-120 nucleotides with molecular weight of approximately 37,000 daltons. Because of the huge number of resonances arising from about 1300 protons, it is very difficult to perform $^1$H NMR studies for this molecule. Moreover, since the broad linewidths of the resonances caused by its long reorientational tumbling time make all resonances overlap each other, it is virtually impossible to resolve or assign any resonance in the $^1$H NMR spectrum.

However, a number of partially resolved resonances has been observed in the downfield region (10 to 15 ppm) of the 220-MHz $^1$H NMR spectrum of tRNA in water (Kearns et al., 1971). These resonances arise from hydrogen-bonded exchangeable imino protons from base-paired guanines or uracils, which cannot be observed in D$_2$O (Figure 18). Without hydrogen bonding, imino protons exchange with water molecules so rapidly that their resonances cannot be observed.
Figure 18. Watson-Crick base pairs (A•U and G•C) and the "wobble" G•U base pair. Helical regions in 5S rRNAs comprise these three types of base pairs.
Base-paired imino protons, however, exchange relatively slowly (>5 msec), hence detection of their resonances is possible.

Chemical shift ranges for the various protons in RNAs have been established from the inventory of synthetic oligonucleotides and tRNAs (Reid, 1981; Wüthrich, 1986). As can be seen from Figure 19 and Table 2, ribose protons range from 3.7-6.3 ppm, non-labile base protons from 5.3-8.4 ppm, base amino protons from 6.6-9.0 ppm, and base imino protons from 10 and 15 ppm relative to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). Among these protons, hydrogen-bonded imino protons have been very useful for the investigation of the structural features of RNA molecules for the reasons described below.

The imino protons resonate in the 10-15 ppm range, which is not a crowded region. Therefore, it is possible to resolve some of the resonances at higher magnetic fields (500-MHz or 600-MHz 1H NMR). In addition, base pair type (A·U, G·C, or G·U) can be determined from the resonance of each hydrogen-bonded imino proton by its chemical shift range and NOE spectrum. Imino protons from A·U base pairs typically resonate around 14.5 to 13.5 ppm, whereas those from G·C base pairs resonate from 13.5 to 12 ppm (Figure 19). While G·C and A·U base pairs have one hydrogen-bonded imino proton per base pair (one resonance), each G·U base pair generates two resonances from its two hydrogen-bonded imino protons. These two G·U resonances are usually located around 12 to 11 ppm. Base pair type can be identified by the characteristic NOE spectra of A·U, G·C, and G·U base pairs, as will be discussed later. Moreover, base pair sequencing is possible from NOE connectivities obtained from NOE spectra of the imino
Figure 19. Example of the $^1$H NMR spectrum of 5S rRNA. Protons of each category (imino, amino, aromatic, or ribose protons) have characteristic chemical shifts.
Table 2. Approximate $^1$H NMR chemical shift ranges of various RNA protons. From Wüthrich (1986).

<table>
<thead>
<tr>
<th>Protons</th>
<th>Chemical Shifts (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2'H$, $3'H$, $4'H$, $5'H$, $5''H$</td>
<td>3.7 - 5.2</td>
</tr>
<tr>
<td>$1'H$</td>
<td>5.3 - 6.3</td>
</tr>
<tr>
<td>$5H$ of C and U</td>
<td>5.3 - 6.0</td>
</tr>
<tr>
<td>$6H$ of C and U</td>
<td>7.1 - 7.6</td>
</tr>
<tr>
<td>$8H$ of A and G, $2H$ of A</td>
<td>7.3 - 8.4</td>
</tr>
<tr>
<td>$NH_2$ of A, C, and G</td>
<td>6.6 - 9.0</td>
</tr>
<tr>
<td>Ring $NH$ of G and U</td>
<td>10 - 15</td>
</tr>
</tbody>
</table>
proton resonances, since the imino protons from two adjacent base pairs are spatially close enough (<5 Å) that their dipolar coupling can be observed through NOE experiments.

Extremely deshielded aromatic CH protons have been reported to resonate between 10 and 15 ppm (Reid, 1981; Robillard & Reid, 1979). However, these protons are very rare, and exchangeable imino protons can be distinguished from non-exchangeable aromatic protons by obtaining 1H NMR spectrum of the molecule in D2O, since the imino proton resonance disappears in D2O (Wong et al., 1972; Kearns & Wong, 1974). Even though the identification and assignment of these proton resonances are desirable, application of these techniques to 5S rRNA molecule requires assistance from auxiliary experiments, like production of smaller RNA fragments of 5S rRNA by ribonucleases, or site-specific spin labeling. In the following section, water suppression technique, the remedy of the first problem encountered in the 1H NMR experiments of 5S rRNA in water (dynamic range problem) will be discussed.

B. Water Suppression in 1H FT NMR Spectroscopy

Since the imino proton resonances of A•U, G•C, and G•U base pairs cannot be observed in D2O due to the exchange of those protons with the solvent, it is necessary to conduct the experiment in H2O (with a small amount of D2O (~5% v/v) for the purpose of a field-frequency lock). While the concentration of the protons in water is about 110 M, the concentration of protons in 5S rRNA is usually ~1 mM. Therefore, the resulting dynamic range (~100,000:1) is beyond not only the
linear range of the analog detector but also the dynamic range of the analog-to-digital converter (ADC) required to sample the time-domain free induction decay (FID). For a typical 16 bit ADC, 1 bit is used for sign and the other 15 bits (only 32 K) are not enough to cover the broad dynamic range. This huge dynamic range also causes other problems like spinning sidebands, curved spectral baseline and difficult phasing.

The dynamic range problem in H$_2$O is especially serious in Fourier transform (FT) NMR spectroscopy, since the whole spectrum including the H$_2$O signal is excited and detected at once. For continuous-wave (CW) NMR spectroscopy, the huge H$_2$O peak can be eliminated by turning off the radiofrequency (rf) frequency-sweep generator when the excitation frequency is around the H$_2$O resonance frequency. While the earlier NMR studies of 5S rRNA were done by this frequency sweep mode (Wong et al., 1972; Kearns & Wong, 1974), the use of FT NMR is more advantageous, in spite of the severe dynamic range problem, due to its higher sensitivity and simpler double-resonance capability.

In an FT NMR spectrometer, all of the signals are excited simultaneously by a short (2~400 µs) rf magnetic field pulse. The induced rf response voltage is detected, amplified, heterodyned to audiofrequency bandwidth, sampled by an ADC, and accumulated in computer memory. The amplitude of the FID should be adjusted so that the maximum magnitude of the FID is slightly smaller than the dynamic range of the ADC. Therefore, the maximum obtainable number of scans from a signal without any noise can be calculated by
subtracting the ADC dynamic range from the length of the computer word. For a Bruker Aspect 3000 computer (24 bit word) with a 16 bit ADC, the maximum scan number is 8 bits (24 bits minus 16 bits), which corresponds to $2^8 (=256)$ scans. However, the maximum number of scans obtainable can be significantly greater than this in the presence of noise, because random noise accumulates as the square root of the number of FIDs. In reality, the signal-to-noise (S/N) ratio continues to increase up to the point where 10,000 scans are possible.

To overcome the dynamic range problem in FT NMR, numerous solvent suppression methods have been proposed (Meier & Marshall, 1990). These methods can be classified into two categories. The first group comprises those that force the H$_2$O magnetization to an extremely small level at the time the observation pulse is applied. Saturation (Schaefer, 1972; Jesson et al., 1973; Bleich & Glasel, 1975; Hoult, 1976; Campbell et al., 1977; Cunneal et al., 1980; Wider et al., 1983) and inversion recovery (Patt & Sykes, 1972; Benz et al., 1972; Mooberry & Krugh, 1975; Krugh & Schaeffer, 1975; Gupta, 1976; Gupta & Gupta, 1979; Inubushi & Becker, 1983; Lanterwein & Gerothannssis, 1983; Haasnoot, 1983) fall into this category. The second strategy is not to excite the intense H$_2$O signal. Methods which belong to this category include rapid scan correlation spectroscopy (Dadok & Sprecher, 1974; Gupta et al., 1974; Arata & Ozawa, 1976) and selective excitation (Alexander, 1961; Redfield & Gupta, 1971; Tomlinson & Hill, 1973; Redfield et al., 1975; Morris & Freeman, 1978; Wright et al., 1981; Sklenar & Starcuk, 1982; Hore,
1983a,b; Turner, 1983; Clore et al., 1983; Bleich & Wilde, 1984; Starck & Sklenar, 1985).

1. Saturation of the Water Signal

This strategy is to suppress the H$_2$O signal by saturating the H$_2$O resonance with a long, selective gated pulse at the H$_2$O resonant frequency immediately before the nonselective observation pulse (Schaefer, 1972). This H$_2$O suppression method, used in the studies of the exchangeable NH proton resonances in 90% H$_2$O of bacitracin (Campbell et al., 1974), has produced good spectra comparable to those obtained in D$_2$O, with a minimum of distortion. However, if the length of the presaturation pulse and the longitudinal relaxation time (T$_1$) of an exchangeable proton are long relative to the chemical exchange rate of the proton, transfer of saturation from irradiated water molecules will reduce the intensity of the exchangeable resonances (Glickson et al., 1974). Imino proton resonances of RNA molecules usually show decreased intensity or sometimes even disappear with saturation of the H$_2$O signal. Figure 20 shows the reduction of the intensity of the base-paired imino proton resonances of yeast tRNA$^{Phe}$ with saturation of the H$_2$O signal.

2. Water Eliminated Fourier Transform (WEFT) Method

The water eliminated Fourier transform (WEFT) technique, based on the inversion recovery principle, takes advantage of the large difference in T$_1$'s between the resonance of the large molecule (<1.0 sec) and the resonance of H$_2$O or HDO in D$_2$O solution (5-15 sec). A pulse sequence of the form T - $\pi$ - t$_n$ - $\pi$/2, where T is a relatively long
Figure 20. The reduction in the intensity of the base-paired imino proton resonances with the increase of saturation pulse length is demonstrated in the downfield $^1$H NMR spectrum of 1.5 mM tRNA. (a) 10 msec, (b) 20 msec, (c) 50 msec, and (d) 100 msec. From Campbell et al. (1977).
initial waiting time and $t_n$ is the time required for the H$_2$O resonance to reach zero longitudinal magnetization, is used to eliminate the H$_2$O signal from the spectrum (Patt & Sykes, 1972). At the point when H$_2$O attains zero longitudinal magnetization, the main observation pulse ($\pi/2$) is applied. In this way, only the signals from the RNA molecule are recorded. A major advantage of the WEFT sequence for H$_2$O suppression is that it is easily adaptable to modern commercial FT NMR spectrometers. However, the relative intensities of the resonances can be significantly distorted by the use of WEFT. In addition, WEFT allows the recording of only partially relaxed spectra as opposed to fully relaxed ones, leading to reduction in the overall sensitivity. This is an especially serious problem for the RNA samples because the sample quantities are limited and optimum S/N is required. Moreover, the intensities of exchangeable protons will be further distorted by cross-relaxation and chemical exchange effects.

A modified WEFT pulse sequence (Gupta, 1976) has been suggested to offset the above disadvantages (Figure 21). The modified WEFT has been reported to be independent of the differential relaxation behavior of H$_2$O and the protons of interest and hence is applicable to systems with any $T_1$ values for water, especially in working with paramagnetic biochemical systems where water $T_1$ may be considerably shortened by electron-nuclear interaction. DASWEFT (DAnte pulse used in Selective Water Eliminated Fourier Transform) has been another pulse sequence proposed to improve the WEFT method.
Figure 21. (Top) Pulse sequence for modified WEFT, which consists of a recovery time PD (~5T₁), a long weak nulling pulse T, a waiting time t_null, and acquisition time (AT). From Gupta (1976). (Bottom) (a) Normal WEFT, (b) DASWEFT, and (c) DASWEFT with NOE preirradiation pulse. The "DANTE" π-pulse comprises one hundred 1 μsec pulses with a repetition rate of 1 msec. From Haasnoot (1983).
3. Redfield 2-1-4 Pulse Sequence

The duration of a pulse in FT NMR determines the width of excitation in frequency units. An FT operation of an infinitely short band in the time-domain gives an infinitely wide band in frequency-domain. Likewise, a short pulse excites signals over the wide region in the spectrum. On the other hand, a long pulse excites only a small frequency-domain range but with higher energy in that region since the frequency-domain excitation peak magnitude is proportional to the time-domain waveform area. This long pulse is desirable for the observation of imino protons in the RNA sample. Ideally, a long pulse centered at 12.2 ppm with a correct pulse duration can give a zero excitation of the H$_2$O signal. In practice, the H$_2$O signal is still observable after suppression; thus a non-flat baseline is obtained.

Alexander (1961) first employed a long, weak observation pulse which excited signals of interest to a flip-angle of between 45° and 90° without flipping the spins of solvent molecules. Dadok and Sprecher (1974) have proposed an important way to flip over protons resonating in a band of frequencies while leaving the solvent signal relatively unaffected. However, a more effective and simpler method has been proposed by Redfield et al. (1975).

A long 90° pulse of amplitude $H_1$, frequency $f_0$ Hz, and length $\tau (= (\pi/2)\gamma H_1$, where $\gamma$ is the gyromagnetic ratio) is converted into the excitation curve of $\sin[2\pi(f-f_0)\tau]/(f-f_0)$ by a Fourier transform. This transformed curve (Figure 22A) has a null at frequencies $\pm 1/\tau$ away from the carrier frequency, while the null in the effectiveness of the pulse for flipping nuclei occurs $\sim 0.97/\tau$ away (Redfield & Gupta,
Figure 22. Construction of the Redfield 2–1–4 pulse sequence. (A) A single soft long pulse in the time-domain and its frequency-domain counterpart obtained by FT operation of the pulse. (B) Two 180° phase-shifted pulses and the resulting frequency power spectrum. (C) The addition of the pulses in (A) and (B), and its frequency-domain power spectrum obtained by FT operation. (D) Comparison of the Redfield 2–1–4 pulse and a single soft long pulse.
Another pulse sequence, two very narrow pulses with \((1/2)\tau\) space between them, turns into a curve of \(\cos[\pi(f-f_0)\tau]\) with a first null at the same frequency as the long pulse of length \(\tau\) (Figure 22B). Subtraction of the second pulse sequence from the first with proper relative strength and phase followed by an FT operation generates an excitation power spectrum of a broad null (Figure 22C). Actually the long pulse is divided into ten segments with a ratio of 2-1-4-1-2, and the third and seventh segments are phase-shifted by 180° as can be seen in Figure 22. This, the so-called Redfield 2-1-4 pulse (Redfield et al., 1975), is comparatively straightforward to implement on modern FT spectrometers, and does not cause the problem of saturation transfer by cross relaxation or chemical exchange from the solvent to the spins under investigation. Moreover, it does not depend on the solvent and solute having substantially different \(T_1\)'s, as required by WEFT and its derivatives. This pulse sequence has been used to obtain some of the NMR spectra of \textit{B. megaterium} 5S rRNA in this research.

However, the H\(_2\)O signal can be suppressed by the Redfield 2-1-4 pulse only up to a few hundred-fold, leaving an H\(_2\)O signal which is still significantly greater than that from other proton resonances from the RNA sample. In practice, errors resulting from inaccurate 180° phase-shifted pulses, inhomogeneity of the static \(B_0\) field and \(B_1\) field, etc., can prevent effective water suppression. In addition, while resonances at the carrier-frequency are excited by 90° flip-angle and the water resonance signal (which is set at 0.97\(\tau\) away from the carrier-frequency) is not excited at all, other resonance signals
between the two are excited to various degrees, producing not only a non-linear excitation problem but the necessity of a phase correction of the spectrum. Phase correction cannot eliminate the non-flat phase line problem resulting from phase shift of the intense residual water peak. Dispersion components of peaks far from the carrier-frequency cannot be corrected either. By modifying the "4" and the last "2" pulse lengths, phase shifts, delays, and transmitter frequency, these problems can be reduced, thus making possible the achievement of improved water suppression.

4. Binomial 1-3-3-1 Pulse Sequence

In principle, selective pulse sequences, which leave the solvent magnetization directed along the magnetic field axis (Z-axis), are ideal for solvent suppression and should cause no problem. In practice, however, a wide variety of small instrumental imperfections foils such expectations. Therefore, painstaking optimization of a variety of experimental parameters is necessary for reduction of such defects. The desirable properties for a solvent suppression sequence are: insensitivity to $B_0$ inhomogeneity and small errors in the choice of transmitter frequency; wide band excitation, preferably on both sides of the solvent; insensitivity to pulse imperfections such as $B_1$ inhomogeneity, non-ideal pulse shapes, off-resonance effects, and phase-shift errors; the requirement for linear phase correction only; simple modification to obtain a 180° pulse; and ease of programming and use. Design of such pulse sequences has been discussed by Hore (1983a,b).
The criteria for such pulse sequences are such that the excited XY magnetization as a function of offset $v$ (the "excitation spectrum") and its gradient and curvature (1st and 2nd derivatives) are all equal to zero at the offset frequency of the solvent. The approach here is based on the approximate proportionality of the excitation spectrum to the Fourier transform of the pulse sequence. An excitation spectrum (excited transverse magnetization as a function of offset frequency $v$) can be expressed as

$$S_n(v) = \sin^n(\pi v \tau)$$  \hspace{1cm} [1]$$

with $n$ a positive integer. All its derivatives with respect to $v$ up to and including the $(n-1)$th vanish at $v=0$ (and at $2\pi n \tau = \pm 2\pi, \pm 4\pi, ...$). The inverse Fourier transform of $S_n(v)$,

$$\sum_{k=0}^{k=n} (-1)^k \binom{n}{k} \delta(t + [k - \frac{n}{2}]\tau)$$  \hspace{1cm} [2]$$

consists of $n+1$ equally spaced delta functions (separation $\tau$) with alternating signs and amplitudes given by the binomial coefficients $\binom{n}{k}$. This expression suggests the family of pulse sequences $1\underline{1}$, $1\underline{2}$, $1\underline{3}$, $1\underline{4}$, $1\underline{6}$, ..., where the numerals give the relative pulse lengths, the underlines indicate a $180^\circ$ phase-shifted pulse, and equal delays between pulses are understood. For example, $1\underline{3}$ corresponds to

$$\alpha(X) - \tau - 3\alpha(-X) - \tau - 3\alpha(X) - \tau - \alpha(-X)$$  \hspace{1cm} [3]$$

where $\alpha$ is the relative flip angle rather than the amplitude of a constituent pulse given by $\binom{n}{k}$. For reasonably strong pulses and small offsets, the above argument is not too misleading. The corresponding cosine functions

$$C_n(v) = \cos^n(\pi v \tau)$$  \hspace{1cm} [4]$$
lead to the sequences without phase shifts, 11, 121, 1331, 14641, ... which have \((n-1)^\text{th}\) order nulls at \(2\pi \nu t = \pm\pi, \pm 3\pi, \ldots\).

Before Hore investigated these pulse sequences for water suppression, there had been a few binomial pulses proposed based on different arguments -- the Bloch equations. Plateau and Gueron (1982) used "JR" ("Jump and Return", 1\text{1} pulse) for yeast tRNA\text{Phe} to achieve water suppression by a factor of 300. Sklenar and Starcuk (1982) reported water suppression (factor of 300-400) by use of a 1\text{21} sequence. The 1331 was introduced by Turner (1983).

Whereas the Fourier transform approximation is a conceptual approach, an exact mathematical calculation of the excitation spectrum of a general pulse sequence is possible by using rotation operators. Through the rotation operator treatment, a large number of pulse sequences (with equal delays between pulses and only two phases, 180° apart) can be deduced which satisfy the conditions for a broad region of zero excitation around \(v=0\). The sensitivity of these sequences to instrumental imperfections (\(B_0\) inhomogeneity, \(B_1\) inhomogeneity, non-ideal phase shapes, phase shifts, phase corrections) has been tested experimentally (Hore, 1983b), and the results have shown that the antisymmetric phase alternating binomial sequences of hard pulses (like 1\text{21}, 1\text{331}, etc., with the exception of 1\text{1}, denoted as the S series) were more successful (Figure 23) than those without phase shifts (11, 121, 1331, etc., denoted as the C series). The two most successful sequences are reportedly to be 1\text{331} and 1\text{5} 10 1\text{0} 5 1. Although the 1\text{5} 10 1\text{0} 5 1 pulse is more efficient, 1\text{331} is simpler, requires only small first order phase correction of
Figure 23. Transverse magnetization excited by the sequences 11, 121, 1331, and 14641 as a function of offset from the transmitter frequency. Calculations were done with $\tau = 500 \mu\text{sec}$, $\gamma B_1/2\pi = 5 \text{ kHz}$, and $\alpha_1$ (the flip angle of the first pulse) equal to 45, 22.5, 11.25, and 5.625°, respectively. From Hore (1983b).
the resulting spectrum, and gives wider band with excitation. The $^{133}$ pulse sequence also has shown better water suppression than other selective pulses in experiments done (by the present author) with *B. megaterium* 5S rRNA samples by using a Bruker AM-500 FT-NMR spectrometer. Therefore, the $^{133}$ pulse sequence has been chosen to obtain most of the spectra of *B. megaterium* 5S rRNA and its RNase T1-digested fragments in this research.

5. Data Shift Accumulation (DSA) and Alternate Delay Accumulation (ADA)

Thus far the discussion has concerned detection of weak signals in the presence of a huge water signal by suppressing the large water signal prior to detection by selective saturation or more efficient selective excitation of solute protons. There are a couple of ways to achieve a similar goal by reducing the size of the FID after collection, which removes information relating to the large water signal while retaining information from the small proton signals from RNA samples.

For optimum operation of the ADC, the input signals must be adjusted to fill as much of its word length as possible. However, computers have limited word length. Therefore, continuous accumulation of the FID in the time-domain will cause computer memory overflow, resulting in truncation of any part of the FID, which adversely influences all the signals in the transformed spectrum. This memory overflow is prevented by the computer's sensing the time when the accumulated signal is approaching the full word length of
the memory and then scaling down the data accordingly. Then the new input data from the ADC must be scaled down by the same extent for proper averaging of the data. This scaling-down is usually achieved by discarding the contents of the appropriate least significant bits in the ADC. Unfortunately the reduction of the ADC resolution in this way decreases the S/N ratio, hence continued averaging under these conditions will not improve the overall S/N ratio. Increase of the computer word length certainly will remove the necessity for the scaling-down and maintain the use of the maximum ADC, but this is not a realistic solution.

Roth et al. (1980) have proposed a very useful approach to overcoming the memory overflow problem -- the reduction of the size of the signals being accumulated in the memory -- by removing information about the solvent signals while retaining that from the weak signals from the solute. In RNA samples, the water signal is responsible for most of the intensity in the FID. The FID has a sine wave shape with an exponentially decreasing amplitude and a frequency $v$ determined by the difference between the carrier and the solvent frequency. If the FID is shifted to the left by one half wavelength of this sine wave and then added to the unshifted FID, most of the signal corresponding to $v$ will be removed (Figure 24, top).

This so-called "Data Shift Accumulation (DSA)" (Roth et al., 1980) can substantially reduce signal being stored in the computer memory. This process will also remove noise components at the solvent frequency because the noise components are correlated in the shifted and unshifted FID signals. If we neglect all dc components, the
Figure 24. (Top) The initial part of 1H FID (a), and after it is left-shifted by two data points (half a wavelength of water frequency) (b). The addition of (a) and (b) is (c). (Bottom) The noise spectrum obtained without a sample. (a) Before DSA operation, (b) after shifting FID by one data point and adding to the unshifted FID data. (c) and (d) are obtained in the same way as (b) except that two data points shift (c) and four data points shift (d) instead of one. From Roth et al. (1980).
content $A(j)$ of each memory location can be expressed as the sum of
the intensities of all frequency components at time $t(j)$:

$$A(j) = \sum_{k=1}^{N/2} \left[ a_k \cos \left( \frac{2\pi jk}{N} \right) + b_k \sin \left( \frac{2\pi jk}{N} \right) \right]$$

[5]

where $a_k$ and $b_k$ are the Fourier coefficients and $N$ is the total number
of data points. After addition of two memory locations ($j$) and ($j+1$),

$$A(j) + A(j+1) = \sum_{k=1}^{N/2} \{ 2 \cos \left( \frac{\pi k}{N} \right) \left[ a_k \cos \left( \frac{\pi k (2j+1)}{N} \right) + b_k \sin \left( \frac{\pi k (2j+1)}{N} \right) \right] \}$$

[6]

where we have neglected the memory locations $(N-1) < j < N$. As can
be seen from the above equation, the amplitudes of the Fourier
components will be modulated by the term $2 \cos(\pi l k / N)$, which can be
rewritten as $2 \cos(\pi f_k / 2 f_N)$ where $f_N$ is the Nyquist frequency and $f_k$ is
the frequency of the Fourier component given by $f_k = (2 k / N) f_N$. An
FID treated by a DSA followed by Fourier transform shows this cosine
modulation (Figure 24, bottom). The spectrum resulting from a DSA
with a left-shift of two data points has a region centrally located in the
spectral width where the intensities of the frequency components
have been substantially decreased. Hence, when a water signal is
located at the center of that spectral region (e.g., with a 1331 pulse,
the carrier-frequency is set at the water resonance), a shift of two data
points in the FID followed by addition to an unshifted FID will
eliminate most of the intensity of the huge water signal.

A variation of the DSA method, in which the shifted FID data are
subtracted from rather than added to the unshifted data, is also
possible. Figure 25 shows two NMR spectra of RNase T1-digested
Fragment A of *B. megaterium* 5S rRNA (to be discussed later) obtained
by a 1331 pulse sequence. One has been treated with the DSA prior to
Figure 25. Comparison of $^1$H NMR spectra of *B. megaterium* 5S rRNA Fragment A before (bottom) and after (top) DSA treatment. The significant water suppression achieved by a 1331 pulse sequence is greatly enhanced by DSA operation in such a way that the water peak (4.78 ppm) has almost disappeared.
Fourier transform while the other has not. As can be seen, the water signal is almost undetectable in the spectrum obtained with the DSA, but with sacrifice of the signals close to the water resonance. When water suppression is used, it is usually necessary to apply the DSA to the acquired data of imino protons of RNA samples with NOESY before 2D-FT operation. Since water suppression produces a non-flat baseline due to the immense water signal, the DSA can ease phasing problems.

Another method of overcoming the dynamic range problem is "Alternate Delay Accumulation (ADA)" (Roth et al., 1980). A 180° phase-shifted FID can be obtained by introducing a delay of half a cycle of the heterodyne frequency of the solvent signal prior to data collection on alternate scans, rather than left-shifting existing data. Unlike the DSA where the "holes" are generated at fixed positions within the spectral range, the ADA can produce a "hole" at any frequency within the spectral width by addition of the shifted and unshifted FIDs. Furthermore, the ADA requires only half the memory required by the DSA method. However, the RMS noise level remains constant across the spectrum after ADA treatment, thus, compared to the DSA, the S/N ratio of the ADA progressively decreases as one approaches the frequency at the center of the "hole". This is the main reason why the DSA method is usually preferred.
C. Nuclear Overhauser Enhancements (NOE) -- The Base Pair Assignment Strategy

Since the proton homonuclear Overhauser enhancement (NOE) experiment can determine not only the base pair type (A·U, G·C, or G·U) but also provide valuable "base pair sequencing", it has been recognized as one of the most powerful techniques for structural determination of nucleic acids. The NOE method has been employed in the investigation of solution structure of tRNAs (Johnston & Redfield, 1977, 1981; Schimmel & Redfield, 1980; Roy & Redfield, 1981, 1983; Hare & Reid, 1982a,b; Heerschapel et al., 1983a,b), prokaryotic 5S rRNAs (Kime & Moore, 1983a,b; Kime et al., 1984; Chang & Marshall, 1986; Leontis & Moore, 1986; Leontis et al., 1986; Zhang & Moore, 1989; Kim & Marshall, 1990), eukaryotic 5S rRNAs (Li & Marshall, 1986; Chen & Marshall, 1986; Lee & Marshall, 1987; Li et al., 1987; Wu & Marshall, 1990a,b) and 5.8S rRNA (Lee & Marshall, 1986).

1. The Nuclear Overhauser Effect (NOE)

The NOE is a measure of the change in the integrated NMR intensity of a nuclear spin when another spin is saturated. In other words, when we eliminate the population difference for a given transition (by irradiating with a weak resonant rf field) a change in the intensity of other signals occurs to compensate that change (saturation) to regain thermal equilibrium of the total system. This phenomenon is well-illustrated by a two-spin system (Noggle & Schirmer, 1971).
Figure 26 shows the energy level diagram for two spin-$\frac{1}{2}$ nuclei, I and S, of the same nuclear species but different chemical shifts (not J coupled). At thermal equilibrium, according to the Boltzmann distribution, the state $\alpha\alpha$ (lower energy) will have an excess of nuclei over states $\alpha\beta$ or $\beta\alpha$ (the difference between $\alpha\beta$ and $\beta\alpha$ is negligible) while the state $\beta\beta$ (higher energy) will contain fewer nuclei than $\alpha\beta$ or $\beta\alpha$ by an equal number. If we call the excess or deficiency $\delta$, the population in each energy level can be described as shown in Figure 26A. The population difference between states can be summarized as follows:

\[
\begin{align*}
\text{I transitions} & \quad \delta \quad \alpha\alpha - \alpha\beta \\
& \quad \beta\alpha - \beta\beta \\
\text{S transitions} & \quad \delta \quad \alpha\alpha - \beta\alpha \\
& \quad \alpha\beta - \beta\beta \\
\Delta M=0 \text{ transition} & \quad 0 \quad \beta\alpha - \alpha\beta \\
\Delta M=2 \text{ transition} & \quad 2\delta \quad \alpha\alpha - \beta\beta
\end{align*}
\]

where $M$ is the nuclear magnetic quantum number. The first four transitions (single quantum transitions) are responsible for the NMR lines, since they satisfy the selection rules.

We will assume that the relaxation for a single transition is a first order process. The rate constants for the various processes will be denoted as $W$, with a superscript to indicate the change in $M$ involved ($W_0$ for the zero quantum transition, $\beta\alpha - \alpha\beta$; $W_1$ for the single quantum transitions, $\alpha\alpha - \alpha\beta$, etc.; $W_2$ for the two quantum transition, $\alpha\alpha - \beta\beta$) and a superscript to indicate to which nucleus the transition belongs ($W_1^I$ or $W_1^S$) (Figure 26B). Immediately after the nucleus S has been
Figure 26. (A) Energy levels and population of a homonuclear AX system. (B) Connections between the energy levels of the system. (C) Populations of the energy levels immediately after the saturation of the S transitions. (D) Initial direction of cross-relaxation after saturation of the S transitions. From Derome (1987).
saturated, a new population of each energy level is established since the populations of $\alpha\alpha$ and $\beta\alpha$ should become equal, likewise the populations of $\alpha\beta$ and $\beta\beta$ (Figure 26C). The new population differences are as follows:

$I$ transitions

- $\delta \quad \alpha\alpha - \alpha\beta$
- $\beta\alpha - \beta\beta$

$S$ transitions

- $0 \quad \alpha\alpha - \beta\alpha$
- $\alpha\beta - \beta\beta$

$\Delta M=0$ transition

- $\delta \quad \beta\alpha - \alpha\beta$

$\Delta M=2$ transition

- $\delta \quad \alpha\alpha - \beta\beta$

The system is no longer at equilibrium, and there is a tendency to establish a new equilibrium through possible relaxation pathways. $W_F$ is not possible because of the saturation of the resonance. The $I$ transition still maintains its population difference ($\delta$), therefore no change is necessary. In other words, saturation of $S$ nucleus does not accompany the intensity change of nuclei $I$ as a result of the single quantum transitions only. However, a population difference between $\alpha\beta$ and $\beta\alpha$ emerges after saturation of nucleus $S$. Hence, there is a relaxation tendency of $W_0$ to transfer population from the state $\beta\alpha$ to the state $\alpha\beta$ to try to restore population equality between the two states. This, in turn, increases the population of the top of one $I$ transition, and decreases the population of the lower level of the other one, thus decreasing the total intensity of signals due to $I$ (Figure 26D). This tendency is counteracted by $W_I$ because the $I$ transitions are already at equilibrium. Therefore, the net result will depend on the balance of $W_I$ and $W_0$. As we have seen so far, if $W_0$ is the
dominant relaxation pathway, a negative NOE (see below) at I will result by saturation of S. For macromolecules or molecules in very viscous solutions, $W_0$ dominates, so negative NOE occurs. Alternatively, if $W_2$ dominates, which is the case for small molecules in non-viscous solution, a positive NOE is obtained.

NOE is defined as

$$\eta_l(S) = \frac{(I - I_0)}{I_0}$$

where $I_0$ is the normal intensity of resonance I, and I is the intensity observed when nucleus S is saturated. By solving differential equations for this system (Noggle & Schirmer, 1971), the NOE is given by

$$\eta_l(S) = \frac{W_2 - W_0}{2W_1 + W_2 + W_0} \frac{\gamma_S}{\gamma_I}$$

where $\gamma_S$ and $\gamma_I$ are the gyromagnetic ratios of nuclei S and I. Since nuclei of the same nuclear species are assumed (spin-$\frac{1}{2}$) in this case, $\gamma_S = \gamma_I$. Then the equation becomes:

$$\eta_l(S) = \frac{W_2 - W_0}{2W_1 + W_2 + W_0}$$

The numerator $(W_2 - W_0)$ is called the cross-relaxation term. The change in the intensity is the result of cross-relaxation between the observed and the irradiated spins. The NOE is inversely proportional to the 6th power of the distance between the two nuclei, and the first-order NOE usually occurs within 5Å distance.

The NOE can be expressed in terms of molecular correlation time, $\tau_c$:

$$\eta_l(S) = \frac{5 + \omega^2 + \tau_c^2 - 4\omega^4\tau_c^4}{10 + 23\omega^2\tau_c^2 + 4\omega^4\tau_c^4}$$

For 5S rRNA molecules, $\omega^4\tau_c^4 > 1$ (slow molecular tumbling time), hence the NOE is negative. This can be better understood from a plot
of NOE intensity versus $\log \tau_c$ (Figure 27). The Larmor frequency ($\omega$) is about $3 \times 10^9$ rad/sec at 500 MHz field strength, and the correlation time ($\tau$) for 5S rRNA molecules is $\sim 10^{-8}$ sec, thus $\log \tau_c$ is $-1.5$. It is evident from the plot that the NOE from 5S rRNA is negative, in the absence of internal rotation of the dipole-dipole axis with respect to the macromolecular rotational diffusion axis.

2. Resonance Assignment by means of NOE

As previously mentioned, the NOE method has been extensively used for solution structure of tRNAs (Johnston & Redfield, 1978, 1979, 1981; Reid, 1981) because of its capability of gathering valuable information about water-exchangeable imino protons which resonate in the far downfield region (10-15 ppm from DSS) in $^1$H NMR spectrum. There are several advantages of NOE that are worth mentioning.

First, NOE enables us to identify the type of the base pair (A•U, G•C, or G•U) responsible for each resonance on the basis of its characteristic NOE difference spectrum. The resonances from two hydrogen-bonded imino protons in a G•U ("wobble") base pair (N1-H of guanine and N3-H of uracil) usually range between 12 and 11 ppm. Since the two imino protons are spatially very close ($\sim 2.5\text{Å}$), they produce strong mutual NOEs (5-10%) to each other (Figure 28, bottom). As can be seen in the figure, the enhancement produced by irradiation of the guanine imino proton is larger than that produced by irradiation of the uracil proton (Hurd & Reid, 1979: Johnston & Redfield, 1981). These characteristic NOE difference spectra of the
Figure 27. The plot of NOE in a two-spin system ($f_l(S)$) versus $\log \omega \tau_c$, where $\omega$ is the Larmor frequency and $\tau_c$ is the molecular correlation time. From Bothner-By (1979).
imino protons of a G•U base pair enable us to successfully distinguish the resonances caused by the G•U base pairs from those by other types of base pairs.

The imino proton resonances from A•U base pairs usually range between 14.5 and 13.5 ppm. The imino proton in an A•U base pair (N3-H of uracil) can easily be distinguished from those in G•C or G•U base pairs by its chemical shift and characteristic NOE of C2-H from adenine. Since the distance between N3-H of uracil and C2-H of adenine in an A•U base pair is very close, the irradiation of the imino proton (N3-H of uracil) gives a sharp, strong NOE difference peak of adenine C2-H, which is located around the 7 to 8 ppm region, as can be seen from the NOE difference spectrum of an A•U base pair from *B. megaterium* 5S rRNA (Figure 28, top). This C2-H NOE can be distinguished from NOEs from amino protons by its narrow linewidth and strong intensity.

Between these two types of imino proton resonances (those from A•U and G•U base pairs) lie imino proton resonances from G•C base pairs (13.5-12 ppm) (Figure 28, middle). Since G•C base pairs are most common in 5S rRNAs, this region is most crowded, thus the resonances are hard to resolve.

Another beneficial feature of the NOE experiment is "base pair sequencing". The imino protons from neighboring base pairs immediately above or below the base pair whose imino proton is being irradiated are close enough to generate cross-relaxation. Hence, imino proton NOEs of a base pair can also identify the imino protons of the neighboring base pairs. Therefore, it is possible to establish a
Figure 28. Identification of base pair types by NOE. The NOE difference spectrum from each base pair (A•U, G•C, or G•U) has a characteristic feature described in the text.
linear sequence of the base pairs by sequentially irradiating the imino protons that are connected by NOEs. When a clear assignment of a resonance is chosen as a starting point, ideally the complete assignment of each resonance in the downfield region of the $^1$H NMR spectrum for 5S rRNA can be achieved.

3. Two-Dimensional NOE Spectroscopy (NOESY)

Obviously, two-dimensional NMR spectroscopy has several advantages over its one-dimensional counterpart. NOESY is especially attractive since it can provide all NOE information without having to go through a tedious sequence of one-dimensional experiments. NOESY experiments have already become popular for structural investigation of small synthetic nucleic acids.

The basic pulse sequence for NOESY is

$$\frac{\pi}{2} - t_1 - \frac{\pi}{2} - \tau_m - \frac{\pi}{2} - t_2$$

[13]

where $\tau_m$ is the mixing time. In the presence of cross-relaxation, changes in the $z$ magnetization of one nucleus during $\tau_m$ will lead to variation in the $z$ magnetization of another, and the pairs of nuclei which would show an NOE in a 1D-experiment may show cross peaks in the 2D-experiment. However, unlike one-dimensional NOE, NOESY also shows cross peaks between nuclei undergoing chemical exchange, since a nucleus whose $z$ magnetization was modulated by one chemical shift during $t_1$ may have the opportunity to migrate to another site during $\tau_m$ in the presence of exchange. Unfortunately, these two types of cross peaks cannot be distinguished from each other. Various
means have been suggested for suppressing chemical exchange cross peaks.

NOESY is not of great use for large RNA molecules since there are too many protons in those molecules. Study of water-exchangeable imino protons is a good choice for large molecules like tRNA or 5S rRNA. There is, however, also a problem to overcome -- water suppression. Higher concentration of the sample and optimization of parameters are more important than for 1D-experiments. There have been a couple of reports of achievement of successful water suppression to obtain imino proton connectivities in NOESY experiments with tRNA samples (Heerschap et al., 1985; Hare et al., 1985).

Figure 29 shows the NOESY spectrum of RNase T1-digested Fragment A of B. megaterium 5S rRNA. Due to the very low concentration (~1mM) compared to those cases of tRNA, imino protons cannot be clearly observed. In addition, DSA treatment before 2D-FT operation caused dissymmetry with respect to the diagonal peaks. Only cross peaks between water and imino protons by exchange are clearly visible.

D. Ring Current Shift Estimation

The aromatic ring current of electrons in a base can influence the static field (B₀) felt by the imino protons of neighboring bases in RNA by means of an induced magnetic field. This will result in either an upfield shift or a downfield shift of the imino proton resonances in the ¹H NMR spectrum, depending on the direction of the induced
Figure 29. NOESY spectrum of 1 mM *B. megaterium* 5S rRNA Fragment A in 10 mM sodium phosphate, 100 mM NaCl, 1 mM EDTA, pH 7.0 at 23 °C. A 133 pulse sequence was used for water suppression. Experimental conditions to obtain this spectrum are given in Chapter 4. Some cross-peaks between imino and amino or aromatic protons are visible. Cross-peaks between imino protons and water arising from proton exchange are clearly visible.
magnetic field with respect to the static field. Theoretical estimation of the chemical shifts of the imino proton resonances has been established previously (Giessner-Prette & Pullman, 1970; Arter & Schmidt, 1976), and has been tested and modified through comparison of the theoretical values with the experimental values obtained from the NMR experiments with tRNAs (Reid et al., 1979; Robillard & Reid, 1979; Johnston & Redfield, 1981).

The chemical shifts of imino protons of base pairs obtained by ring current shift calculations alone are no longer convincing, since significant differences have been sometimes observed between calculated values and the values obtained from NMR spectra. Although ring current shift rules for the G•U wobble base pairs have been proposed (Geerdes & Hilbers, 1979), the estimation of imino proton resonances from a G•U base pair is known to be especially inaccurate due to its different base stacking interaction compared to the Watson-Crick base pairs.

In the studies of B. megaterium 5S rRNA and its fragments, the chemical shifts of some resonances have been estimated from ring currents on the basis of an 11-fold RNA A-helix (Arter & Schmidt, 1976), with assumed intrinsic A•U and G•C positions of 14.35 and 13.45 ppm, respectively (Reid et al., 1979). These estimations are made merely to support the assigned resonances by the comparison of calculated values and experimental values.
CHAPTER III
INVESTIGATION OF THE STRUCTURE OF BACILLUS
MEGATERIUM RIBOSOMAL 5S RNA

A. B. megaterium Cell Growth and Isolation and Purification of 5S rRNA

Bacillus megaterium strain KM (a Gram-positive bacterium, ATCC 13632) was purchased from American Type Culture Collection. The cells were grown in 20L or 30L of Trypticase Soy Broth media at room temperature, and provided by the Fermentation Laboratory of The Ohio State University.

So far a great number of procedures for isolation of 5S rRNA has been published. Isolation of ribosomal RNAs is generally achieved by ultracentrifugation, with subsequent extraction with phenol (Geroch et al., 1968). The 5S rRNA can be further purified by ion-exchange chromatography and gel-filtration chromatography. Alternatively, ribosomal RNAs can be directly extracted from cells with phenol and purified by gel electrophoresis (Rubin, 1975). Most of these procedures have been developed for binding studies of 5S rRNA to ribosomal proteins, nuclease digestion, oligonucleotide binding, or primary nucleotide sequencing (5S rRNAs from over 500 organisms have been
compiled recently by Specht et al., 1990). These procedures are not suitable for producing the large quantity of 5S rRNA necessary for NMR experiments.

The preparative-scale technique for isolation of 5S rRNA developed by Li et al. (1984) was employed for mass production of B. megaterium 5S rRNA. This method consists of phenol-sodium dodecyl sulfate (SDS) extraction of RNAs from cells, removal of proteins by DE-32 ion-exchange chromatography, and purification of 5S rRNA by Sephadex G-75 gel-filtration chromatography.

1. Isolation of Ribosomal RNAs from B. megaterium Cells

Each batch of B. megaterium cells grown in 20 L media yielded ~200 g of cells upon centrifugation at 11,000 X g for 20 min. The cells were then suspended in ~1 L of solution containing 0.01 M sodium acetate and 0.1 M NaCl, pH 5.0, and were centrifuged at 11,000 X g for 15 min. The resulting pellets were suspended in ~900 mL of 0.01 M sodium acetate, 0.1 M NaCl, 0.25% SDS, pH 5.0, and an equal volume of liquified phenol was added. The mixture was then stirred at room temperature for 2 h. The aqueous layer could be separated from the phenol layer by centrifugation at 11,000 X g for 15 min, with subsequent withdrawal of the supernatant by a pipette or a syringe. RNAs were allowed to precipitate overnight at −20 °C after 2.5 volumes of cold 95% ethanol were added to the aqueous layer.

Most of the solution was decanted, and the remaining suspension was centrifuged at 10,000 X g for 15 min. The precipitate was then dissolved in a solution of 0.01 M Tris (tris[hydroxymethyl]aminoethane)
and 0.3 M NaCl, pH 7.5. Undissolvable materials were eliminated by passing the solution through a filter (~0.5 μm of diameter) or by centrifugation. The crude extract was then ready for ion-exchange column chromatography. The procedure to this point is outlined in Figure 30.

2. DE-32 Ion-Exchange Chromatography

Crude RNA sample at this stage contains proteins, polysaccharides from cell membrane, and phenol. These unwanted materials can be eliminated from RNAs by ion-exchange chromatography, since RNA binds tightly to anion-exchange resins (e.g., DEAE (diethyl aminoethyl cellulose) due to its highly negative-charged phosphate groups.

A DE-32 (Whatman) ion-exchange column of 5.0 cm in diameter and 10 cm in length was packed, and pre-equilibrated at room temperature with 0.01 M Tris, 0.3 M NaCl, pH 7.5 until the pH of collected buffer was the same as that of the fresh buffer. Usually 2–3 L of buffer was required for equilibration of the column. Then the crude RNA sample in 0.01 M Tris, 0.3 M NaCl, pH 7.5 was applied to the column, with subsequent elution of the column with the same buffer. At this point, all materials except strongly bound RNAs will be eluted. The column was washed with the buffer until the absorption at 260 nm ($A_{260}$) was below 0.1. By this time the yellow color in the crude RNA solution completely disappeared.

The remaining RNAs in the column can be released by eluting the column with a buffer solution of higher salt concentration- 0.01 M Tris, 1.0 M NaCl, pH 7.5. The elution profile of DE-32 column chromatography is shown in Figure 31. The collected RNAs in solution
Grown cells in 20 L of Trypticase Soy Broth medium

Centrifuged at 11,000 X g for 20 min

200 g of \textit{B. megaterium} cells  
Supernatant (discarded)

Suspended in 0.01 M sodium acetate, 0.1 M NaCl, pH 5.0
Centrifuged at 11,000 X g for 15 min

Washed \textit{B. megaterium} Cells  
Supernatant (discarded)

900 mL of 0.01 M sodium acetate, 0.1 M NaCl, 0.25% SDS, pH 5.0 and 900 mL of liquified phenol are added
Stirred at room temperature for 2 h
Centrifuged at 11,000 X g for 15 min

Aqueous layer  
Phenol layer (discarded)

2.5 volumes of cold (-20 °C) 95% ethanol are added
Stored at -20 °C overnight
Centrifuged at 10,000 X g for 15 min

Precipitate (Crude RNAs)  
Supernatant (discarded)

Dissolved in 100 mL of 0.01 M Tris, 0.3 M NaCl, pH 7.5
Removal of any undissolvable material by centrifugation or applying the sample through micropore filters

DE-32 ion-exchange chromatography

\textbf{Figure 30.} Isolation of ribosomal RNAs from \textit{B. megaterium} cells
Figure 31. Elution profile obtained from DE-32 ion-exchange chromatography of *B. megaterium* crude RNA extract. Experimental conditions, column size, and buffer solution are described in the text.
(clear solution) were again combined with 2.5 volumes of cold 95% ethanol and stored in the freezer overnight to precipitate RNAs.

3. Purification of *B. megaterium* 5S rRNA by Sephadex G-75 Gel-Filtration Chromatography

The RNA precipitate was again collected by centrifugation, and dried by flowing dry N₂ gas. The mixture then was redissolved in ~15 mL of 0.01 M Tris, 1.0 M NaCl, pH 7.5.

The packed Sephadex G-75 column (5.0 X 150 cm) was pre-equilibrated with three column volumes of the same buffer solution. About three times the packed column volume of the buffer solution was needed. The RNA solution was carefully loaded on the top of the gel without disturbing the surface of the gel. Then the column was eluted with the buffer solution. The eluted buffer solution passed through the monitor of a UV detector (set at 254 nm) and was collected by a fraction collector. The flow rate was usually 20–30 mL/h (equivalent to 5–7 sec/drop), and 7 mL were collected in each tube.

Figure 32 shows the elution profile of the Sephadex G-75 column chromatography. As can be seen in the figure, RNAs were separated into three major components. The first peak arises from large RNA molecules, whose molecular weights are far greater than that of 5S rRNA. The second peak is from 5S rRNA (molecular weight of ~37,000 daltons), and is followed by a huge tRNA (molecular weight of ~25,000 daltons) peak. The fractions which contain 5S rRNA only were collected and again combined with 2.5 volumes of 95% cold ethanol. The purity of the collected 5S rRNA fractions was examined by polyacrylamide gel
Figure 32. Elution profile obtained from Sephadex G-75 gel-filtration chromatography of *B. megaterium* RNAs. Fractions of large RNAs, 5S rRNA, and tRNA are specified in the figure.
electrophoresis (PAGE). Two batches of grown cells (~400 g of cells) yielded about 30 mg of 5S rRNA of purity greater than 95%. Purification of *B. megaterium* 5S rRNA by DE-32 ion-exchange chromatography and Sephadex G-75 gel-filtration chromatography is outlined as a schematic flow chart in Figure 33.

**B. Polyacrylamide Gel Electrophoresis (PAGE)**

The purity of the isolated 5S rRNA can be analyzed by the polyacrylamide gel electrophoresis (PAGE) technique. Acrylamide and N,N'-methylene-bisacrylamide (BIS) turn into polyacrylamide gel by polymerization of the monomers in the presence of a free radical initiator (ammonium persulfate) with a catalyst N,N,N',N'-tetramethyl-ethylenediamine (TEMED). RNAs of different sizes have different mobility in the gel upon application of a suitable voltage across the gel. Since RNAs are negatively charged, they move toward the positive electrode. In the presence of urea, which is a denaturing agent, RNAs can be separated according to their size on the gel plate. Usually 8-10% denaturing gels of polyacrylamide are optimal for the purity check of 5S rRNA.

**1. Gel Preparation**

TBE 10X solution (1.0 M Tris-borate, pH 8.3, 20 mM EDTA) is prepared by dissolving 1 mole of Trizma base, 1 mole of boric acid, and 20 mmole of disodium EDTA in distilled water. The total volume is set to 1 L and the pH is adjusted to 8.3.
Crude RNA (from Figure 30)

Applied to the DE-32 ion-exchange column (5.0 X 10 cm) pre-equilibrated with 0.01 M Tris, 0.3M NaCl, pH 7.5

Washed with the same buffer until A_{260} < 1

Eluted with 0.01 M Tris, 1.0 M NaCl, pH 7.5

RNA peak is collected

RNAs are precipitated with 2.5 volumes of cold (-20 °C) 95% ethanol with subsequent storage at -20 °C overnight

Centrifuged, dried over dry flowing N\textsubscript{2} gas

Dissolved in 0.01 M Tris, 1.0 M NaCl, pH 7.5

Applied to the Sephadex G-75 column (5.0 X 150 cm) pre-equilibrated with 0.01 M Tris, 1.0 M NaCl, pH 7.5

Eluted with the same buffer

Collected fractions are combined

Mixed with 2.5 volumes of cold 95% ethanol and stored at -20 °C

~15 mg of purified 5S rRNA

**Figure 33.** Purification of *B. megaterium* 5S rRNA by DE-32 ion-exchange chromatography and Sephadex G-75 gel-filtration chromatography
For 10% denaturing gels, 4.75 g of acrylamide and 0.25 g of bisacrylamide are dissolved in 5 mL of TBE 10X and a few mL of water, with subsequent addition of 25 g of urea. More distilled water is added to give a final volume of 50 mL. Since the dissolution of urea in water is an endothermic process, warming up the solution is necessary for rapid and complete dissolution of urea. The solution should be cooled down to room temperature and degassed under vacuum. The final concentration of the solution is 0.1 M Tris-borate, 2 mM EDTA, 10% acrylamide, and 50% urea.

When the solution is cool enough, a small amount of ammonium persulfate (10–20 mg) is dissolved in the solution followed by ~40 µL of TEMED. The resulting solution is quickly pipetted into a slab gel apparatus (0.75 mm in thickness, 140 mm in width, and 160 mm in length), and a comb is inserted between the two glass plates holding the gel to establish sample wells before polymerization. When the gel is completely polymerized, the comb is removed from the gel, and the sample wells should be rinsed with distilled water and the electrophoresis buffer solution (TBE 1X) to eliminate remaining urea in the wells.

2. Sample Preparation for Electrophoresis

A small amount of purified 5S rRNA (~30 µg) from cold ethanol precipitate is transferred to a 1.5 mL polypropylene centrifuge tube by using a Pasteur pipette, and centrifuged in a high-speed microcentrifuge (Fisher Model 59A) for 5 min to recover 5S rRNA. The precipitate is separated from the supernatant by decanting, followed by subsequent
drying with filtered N\textsubscript{2} gas. About 10 \(\mu\)L of electrophoresis buffer (80% (v/v) formamide, 10 mM NaOH, 1 mM EDTA, 0.1% (w/v) bromophenol blue) are added and vortexed, followed by heating in the water-bath of 65 \(^\circ\)C for 5 min to induce complete denaturation of RNA. The cooled sample is then applied to the well of the electrophoresis gel with a Hamilton syringe (usually 5 \(\mu\)L are enough for each well). The electrophoresis gel, which is sandwiched between two glass plates, is placed in the electrophoresis apparatus filled with the running buffer solution (TBE 1X).

3. Running the Gel

When constant voltage (~270 V) across the slab gel is applied (positive at the bottom of the gel), the movement of RNA bands can be monitored by bromophenol marker dye which migrates faster than any RNA molecule. The gel can be run at constant temperature by connecting the heat-exchanger to a water tap, but this is not necessary. The power source is turned off when the tracking dye band advances nearly to the bottom of the gel. The running time is ~3 h at room temperature.

4. Staining and Destaining of the Gel

After the power is turned off, the slab gel is removed from the electrophoresis apparatus. One of the glass plates is first removed, and the other glass plate can be removed from the gel after immersion in the fixing solution (5% acetic acid). The gel should be soaked in the fixing solution for 20 min. Then the solution is carefully decanted, and the gel is immersed in the staining solution (0.2% (w/v) methylene blue in 0.4 M
sodium acetate, 0.4 M acetic acid, pH 4.7). The staining solution is decanted after 20 min, with subsequent destaining of the gel in water. The blue RNA bands should be visible after the gel is rinsed with water several times.

5. Gel Scanning and Gel Drying

The RNA bands in the gel disappear when the gel stays in water for a prolonged period. In order to keep the RNA bands in the gel for a record, it is necessary to dry the gel on a slab gel dryer. The gel is taken out of the water and placed on a filter paper. This filter paper with a gel is again placed on a Hoefer gel dryer (Model SE 540). The gel dryer dries the gel, retaining the image of the bands on the gel by removing moisture with heat under vacuum. It takes about an hour for the blue RNA bands of the gel to be fixed onto the filter paper with a gel dryer.

The gel, on the other hand, can be scanned by a Beckman DU-8 UV-visible spectrometer. Since RNA bands visualized by the staining solution have blue color, their intensities can be monitored at 550 nm wavelength. The gel is cut into strips ~1 cm in width corresponding to each sample well, and the strip which is going to be scanned is placed in a special container for gel scanning with distilled water. Since the spectrophotometer is equipped with a scan program (COMPUSET™), the scanning can be easily done. Figure 34 shows the comparison of the *B. megaterium* 5S rRNA sample before purification (after DE-32 ion-exchange chromatography) and after purification (after Sephadex G-75 gel-filtration chromatography). Figure 34 A shows RNA bands of dried
Figure 34. (A) *B. megaterium* RNAs before purification (left) and purified *B. megaterium* 5S rRNA (right) visualized on dried gels. (B) Purified *B. megaterium* 5S rRNA shown by a gel scan.
gels obtained by a gel dryer, while Figure 34 B shows the purified RNA band obtained by gel scanning.

C. UV Absorption Spectroscopy of *B. megaterium* 5S rRNA

As previously discussed, UV spectroscopy has been used for estimating the total number of base pairs in 5S rRNAs and the percentages of G·C or A·U base pairs among them (Fresco *et al.*, 1963; Boedtker, 1967; Boedtker & Kelling, 1967; Fox & Wong, 1979; Chang *et al.*, 1984). The concentration of any 5S rRNA in solution can be determined based on the UV absorbance at 260 nm, since RNA shows the strongest absorption at around 260 nm, as can be seen from the absorption spectrum of *B. megaterium* 5S rRNA (Figure 35).

1. Determination of the Concentration of *B. megaterium* 5S rRNA

Although the exact concentration is difficult to determine, the estimation of 5S rRNA concentration in solution for this work has been performed by assuming that 1 mg of RNA in 1 mL of buffer with appropriate salt concentration (10 mM cacodylic acid or sodium phosphate, and 0.1 M NaCl) results in A$_{260}$ of 22.4. This conversion factor is generally accepted, and has been obtained from the calculation for *Bacillus subtilis* 5S rRNA (Chang, 1985), the base components of which are very similar to those of *Bacillus megaterium* 5S rRNA. The calculated value (A$_{260}$ = 22.4 for 1 mg/mL of 5S rRNA in solution) was experimentally obtained from the total absorbance of 116 mononucleotides of *B. subtilis* 5S rRNA after complete hydrolysis of the molecule under alkaline condition.
Figure 35. UV absorption spectrum (absorbance versus wavelength) of *B. megaterium* 5S rRNA.
2. UV Melting Analysis of *B. megaterium* 5S rRNA

About 30 mg of purified *B. megaterium* 5S rRNA were dissolved in 2 mL of 10 mM cacodylic acid, 10 mM EDTA, 100 mM NaCl, and dialyzed against 1 L of the same buffer at 4 °C overnight with one change of the buffer solution. Then, a small amount (~100 μL) of 5S rRNA in the solution was diluted with the same buffer to give A$_{260}$ of ~1.0. The resulting 5S rRNA sample was used for UV melting analysis with a Beckman DU-8 Spectrometer with the "T$_{m}$ COMPUSET™" accessory. Absorbances at 260 nm or 280 nm were taken at every 1 °C increase from 25 °C to 95 °C with a temperature increase rate of ~1 °C/min. For the 5S rRNA samples with Mg$^{+2}$, a small portion of 0.1 M MgCl$_2$ solution was added to give Mg$^{+2}$ concentration of 10 mM.

Figure 36 shows the normalized A$_{260}$ hyperchromism profiles of *B. megaterium* 5S rRNA, which are very similar to those of *B. subtilis* 5S rRNA (Chang et al., 1984). The melting curves in the absence and in the presence of Mg$^{+2}$ also show, though not as obviously, diphasic melting patterns just like *B. subtilis* (Chang, 1985). In addition, the melting temperatures determined from *B. megaterium* 5S rRNA are nearly identical to those of *B. subtilis* 5S rRNA (54 °C and 68 °C in the absence of Mg$^{+2}$, broad diphasic transitions between 65 °C and 75 °C with 10 mM Mg$^{+2}$). According to Chang (1985), the upper limit of the number of bases involved in base-pairings for *B. subtilis* 5S rRNA is 66% of total bases (~38 base pairs) in the presence of Mg$^{+2}$. The calculations for *B. megaterium* 5S rRNA will give similar results, except that *B. megaterium* 5S rRNA contains fewer G•C base pairs than *B. subtilis* 5S rRNA.
Figure 36. Normalized $A_{260}$ UV hyperchromism for *B. megaterium* 5S rRNA in the absence and presence of 10 mM Mg$^{+2}$. 
Figure 37 shows the $A_{280}$ hyperchromism of $B. megaterium$ 5S rRNA. As can be seen in the figure, the diphasic melting pattern in the absence of $Mg^{+2}$ changes to monophasic with addition of $Mg^{+2}$, which has been observed in $B. subtilis$ 5S rRNA (Chang, 1985).

D. Investigation of $B. megaterium$ 5S rRNA Secondary Structure by $^1H$ NMR Spectroscopy

As previously mentioned, $^1H$ NMR, especially the NOE technique, is effective for investigation of secondary structure of 5S rRNAs in solution. Comparison of the imino proton NMR spectra of two or more 5S rRNAs which have similar base sequences can facilitate interpretation of the spectra, and thus help to establish the secondary structure (Chen & Marshall, 1986). NMR analysis of $B. megaterium$ 5S rRNA is particularly useful, because its primary base sequence (and probably its secondary structure) is very similar to that of the previously studied $B. subtilis$.

1. Preparation of NMR Samples

Purified $B. megaterium$ 5S rRNA was dissolved in 10 mM EDTA, 100 mM NaCl, and 10 mM cacodylic acid, pH 7.0, and then dialyzed against 500 volumes of 1 mM EDTA, 100 mM NaCl, 10 mM cacodylic acid, pH 7.0 at 5 °C for 3 h. The sample was then dialyzed against fresh buffer overnight. The sample was concentrated by ultrafiltration by use of an Amicon Centricon apparatus. $D_2O$ was added to a final 5% v/v concentration to provide a $^2H$ field-frequency lock signal. The final 5S rRNA concentration of the NMR sample was 28.7 mg/mL, with at least 95% purity, as estimated from the electrophoresis gel scan.
Figure 37. Normalized $A_{280}$ hyperchromism for *B. megaterium* 5S rRNA in the absence and presence of 10 mM Mg$^{2+}$. 
2. NMR Spectroscopy

All $^1$H NMR spectra were obtained with a Bruker AM-500 FT/NMR spectrometer, with phase-cycled quadrature detection, without sample spinning. Water suppression in NMR spectra was achieved by use of a 1331 hard-pulse sequence (Hore, 1983b) with the carrier frequency centered at the $H_2O$ resonance. Chemical shifts were measured relative to the $H_2O$ resonance, which was taken as 4.78 ppm relative to DSS (3-(trimethylsilyl)-1-propane-sulfonic acid) at all temperatures. Downfield shifts are defined as positive.

The nuclear Overhauser enhancement (NOE) difference spectra of 5S rRNA were obtained by use of a modified Redfield 214 pulse sequence (Redfield et al., 1975; Chang & Marshall, 1986) with the radiofrequency carrier centered at 15 ppm, with an acquisition period of 0.17 s (4 K time-domain data, typically 17,600 time-domain transients (8,800 on-resonance and 8,800 off-resonance)), 0.5 s preirradiation (0.2 or 0.5 mW decoupler power) of the resonance of interest, and off-resonance irradiation at 18 ppm. NOE experiments were carried out at 23 °C.

3. Peak Identification in 5S rRNA

The downfield (9-15 ppm) 500 MHz proton NMR spectrum of B. megaterium 5S rRNA (Figure 38) exhibits ~20 resolved peaks corresponding to hydrogen-bonded ring imino proton resonances from Watson-Crick (A·U and G·C) base pairs, "wobble" G·U base pairs (Reid, 1981), and some unpaired U and G imino protons which are shielded from the solvent with a sufficiently slow exchange rate with $H_2O$ (Hare & Reid, 1982b). However, ~30 separate imino proton resonances can be
Figure 38. 500-MHz downfield proton NMR spectrum of *B. megaterium* 5S rRNA at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95 %:5 % H₂O/D₂O. The peaks are labeled from A to O for convenience.
identified in the downfield proton NMR spectrum of 5S rRNA from the higher-resolution spectra obtained from the three RNA fragments to be discussed below.

As can be seen from Figure 38, the assignment of particular resonances to specific base-pair protons in helical RNA segments is a problem. However, two G•U base pairs (peaks K1/M and K2/L) can immediately be identified from their characteristic NOE difference spectra. Starting from the G•U base pair, K1/M, the short NOE-connected segment, G•U – G•C – G•C (K1/M, E, J), can be established (see Figure 39). From the universal secondary structure model (Figure 38), such a segment occurs only in the terminal stem (helix I) of B. megaterium 5S rRNA: either G4•C112 – G5•C111 – U6•G110 or G8•U108 – C9•G107 – G10•C106. If the resonances K1/M, E, J are G8•U18 – C9•G107 – G10•C106, then resonance J would be expected to melt quickly at elevated temperature because base pair G10•C106 is located at the end of the helix. For example, Gewirth et al. (1987) have reported that the base pairs G10•C110 and G9•U111 in helix I of E. coli 5S rRNA are unstable both in the fragment and in native 5S rRNA. As can seen from the variable temperature experiment shown in Figure 40, resonances J and E are in fact highly stable and do not melt even at 50 °C. Thus, we exclude the assignments of resonances J and E as G10•C106– C9•G107, and we assign resonances J, E, K1/M as G4•C112 – G5•C111 – U6•G110. With that assignment, peaks J (G4•C112) and E (G5•C111) are expected to be very stable at elevated temperature because they are located in the middle of the most stable helix (helix I).
Figure 39. 500 MHz proton NOE difference spectra of resonances J, E, K₁, and M of *B. megaterium* 5S rRNA. The G·U base pair (resonances K₁/M) is connected to a G·C base pair (E) which is connected to another G·C pair (J). This segment is assigned as G₄·C₁₁₂ – G₅·C₁₁₁ – U₆·G₁₁₀ in helix I.
Figure 40. 500 MHz downfield proton NMR spectra of *B. megaterium* 5S rRNA in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 at various temperatures. Peaks J, and E are stable to ~50 °C, and have about the same melting onset temperature as peaks $K_1$ and M.
It is interesting to compare the $G_4\cdot C_{112} - G_5\cdot C_{111} - U_6\cdot G_{110}$ segment ($J$, $E$, $K_1$/$M$) in *B. megaterium* with the assigned segment $G_4\cdot C_{112} - G_5\cdot C_{111} - U_6\cdot G_{110}$ in *B. subtilis*, another Gram-positive bacterium, as previously reported by Chang and Marshall (1986a). The two bacterial 5S rRNAs exhibit high homology in base-pair sequence, so that the $G_4\cdot C_{112} - G_5\cdot C_{111} - U_6\cdot G_{110}$ segments in both 5S rRNAs are located at the same place in the secondary structure, as evidenced by the virtually identical chemical shifts for resonances $J$, $E$, and $K_1$/$M$ (12.10, 13.30, and 11.84/10.84 ppm, respectively) in the spectrum of *B. megaterium* 5S rRNA and resonances $O$, $G$, and $Q/T$ (12.12, 13.38, and 11.7/10.9 ppm, respectively) in *B. subtilis* 5S RNA (Chang & Marshall, 1986a). (Resonance $K_1$ in *B. megaterium* 5S rRNA is shifted slightly (0.14 ppm) from resonance $Q$ in *B. subtilis* 5S rRNA, probably because $G_7\cdot U_{109}$ in *B. megaterium* 5S rRNA is replaced by $G_7\cdot C_{109}$ in *B. subtilis* 5S rRNA.) The virtually identical chemical environment for the base-pairs $G_4\cdot C_{112} - G_5\cdot C_{111} - U_6\cdot G_{110}$ in both bacteria offers additional evidence for a common secondary structure for 5S rRNAs from all prokaryotic ribosomes.

The assignment of the segment $G_4\cdot C_{112} - G_5\cdot C_{111} - U_6\cdot G_{110}$ in helix I was reconfirmed by NOE experiments with Fragment C (to be discussed in Chapter 5). Although numerous NOE connectivities were observed between the various resonances of intact 5S rRNA (Figure 41 shows NOE difference spectra from some of other resonances), the extensive peak overlap precluded definitive assignment of base-paired segments. Table 3 summarizes NOE experiments of *B. megaterium* 5S rRNA. Due to irradiation spillover problems, NOE connectivities were
Figure 41. 500-MHz proton NOE difference spectra of resonances A, F, K, and L of *B. megaterium* 5S rRNA. Resonances K2 and L arise from another G·U base pair.
Table 3. Summary of NOE connectivities obtained by irradiation of imino proton resonances of *B. megaterium* 5S rRNA.

<table>
<thead>
<tr>
<th>Resonances</th>
<th>ppm</th>
<th>Peak Type</th>
<th>Connectivity to</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A₁, A₂</td>
<td>14.2</td>
<td>A•U</td>
<td>F, D, I, I&quot;</td>
</tr>
<tr>
<td>B</td>
<td>13.9</td>
<td>A•U</td>
<td>I&quot;</td>
</tr>
<tr>
<td>C</td>
<td>13.6</td>
<td>A•U (?)</td>
<td>Spillover</td>
</tr>
<tr>
<td>D (D₁, D₂)</td>
<td>13.5</td>
<td>A•U and G•C</td>
<td>Spillover</td>
</tr>
<tr>
<td>E</td>
<td>13.3</td>
<td>G•C</td>
<td>J, K₁, M</td>
</tr>
<tr>
<td>F</td>
<td>13.2</td>
<td>G•C</td>
<td>A₁</td>
</tr>
<tr>
<td>G</td>
<td>13.1</td>
<td>G•C</td>
<td>I (?)</td>
</tr>
<tr>
<td>H</td>
<td>12.8</td>
<td>G•C</td>
<td>Spillover</td>
</tr>
<tr>
<td>I'</td>
<td>12.6</td>
<td>G•C</td>
<td>Spillover</td>
</tr>
<tr>
<td>I</td>
<td>12.5</td>
<td>G•C</td>
<td>A (?)</td>
</tr>
<tr>
<td>I&quot;</td>
<td>12.4</td>
<td>G•C</td>
<td>A₁</td>
</tr>
<tr>
<td>J</td>
<td>12.2</td>
<td>G•C</td>
<td>E</td>
</tr>
<tr>
<td>K₁, K₂</td>
<td>11.9</td>
<td>G•U</td>
<td>E, M, L</td>
</tr>
<tr>
<td>L</td>
<td>11.7</td>
<td>G•U</td>
<td>K₂</td>
</tr>
<tr>
<td>M</td>
<td>10.9</td>
<td>G•U</td>
<td>E, K, L</td>
</tr>
<tr>
<td>N</td>
<td>10.4</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>O</td>
<td>9.7</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
hard to determine accurately in many cases. Moreover, since most peaks in the spectrum represent more than one resonance (judged from the total number of peaks and their intensities, etc.), base pair sequences cannot be established unequivocally.
CHAPTER IV

STRUCTURAL INVESTIGATION OF RNASE T1-DIGESTED
FRAGMENT A OF B. MEGATERIUM RIBOSOMAL 5S RNA BY
$^1$H NMR SPECTROSCOPY

A. Production and Purification of RNase T1-Cleaved Fragments of B. megaterium 5S rRNA

As shown in Chapter 3, NMR analysis alone cannot provide deep insight into the structural feature of the 5S rRNA molecule because of broad peak widths and extensive overlap of the proton resonances in the downfield region. Therefore, additional methods are needed to simplify the proton spectrum. Among such methods are site-specific spin labeling (Lee & Marshall, 1987) and the use of enzymatic cleavage fragments (Kime & Moore, 1983a,b; Kime et al., 1984; Li & Marshall, 1986; Chen & Marshall, 1986; Li et al., 1987). Proton NMR studies of enzyme-cleaved fragments of 5S rRNA are especially attractive because of greatly reduced peak overlap resulting from fewer, narrower proton resonances. Three different fragments (denoted as A, B, and C) of B. megaterium 5S rRNA, which contain helices II and III, IV, and I, respectively, were produced by enzymatic cleavage with RNase T1, as described in this section. The primary sequences of the three fragments were determined by an RNA
sequencing procedure, described in the next section. The structural investigation of Fragment A will be discussed in the remainder of this chapter.

1. Production of RNase T1-Cleaved Fragments

Purified *B. megaterium* 5S rRNA in cold ethanol was precipitated by centrifugation at 10,000 X g for 15 min, with subsequent decanting of the supernatant. The remaining precipitate (~60 mg of 5S rRNA) was dissolved in 65 mL of a solution containing 0.01 M tris(hydroxymethyl) aminomethane (Tris) base, 0.3 M NaCl, pH 7.5 (at a concentration of $A_{260} = 20$/mL). RNase T1 (Boehringer Mannheim) was added at 2.5 µL (= 250 units) per milliliter of 5S rRNA, and the reaction was allowed to proceed for 25 min at room temperature (~24 °C) with occasional mixings. The reaction was stopped by the addition of an equal volume (65 mL) of 5% SDS. Another equal volume of phenol (65 mL) was added, and the mixture was mixed thoroughly for 2 min. The aqueous layer containing RNA was separated from the phenol layer by centrifugation at 10,000 X g for 10 min. Phenol extraction was done twice to completely remove RNase T1. 2.5 volumes of cold (~20 °C) ethanol were added to the aqueous layer, and the resulting solution was stored in a freezer (~20 °C) for at least two days to precipitate the RNA fragments.

2. Purifications of the Three RNase T1-Cleaved Fragments

A Sephadex G-50 column (2.5 X 150 cm) was prepared and equilibrated with at least three column volumes of freshly prepared, degassed 0.01 M Tris buffer, pH 7.5, containing 1.0 M NaCl. The RNA fragments were recovered from ethanol precipitate by centrifugation,
followed by drying over filtered N\textsubscript{2} gas. The RNA sample was then dissolved in 4 mL of 0.01 M Tris, 1.0 M NaCl, pH 7.5, and applied to the column with subsequent elution of the column with the same buffer. The flow rate was adjusted to be around 13 mL/h. After 250 mL was eluted to the waste container, the eluent was collected by a fraction collector with each fraction of 6 mL/tube. The first peak (Fragment A) appeared by about 300 mL elution volume, followed by peaks of Fragment B and Fragment C. The elution profile is shown in Figure 42. Fractions containing Fragment A, Fragment B, and Fragment C were collected separately, and 2.5 volumes of cold (-20 °C) 95% ethanol were added to each of the collections. They were stored in the freezer until use.

3. Purity Check of the Three Fragments

The characterization and purity test of the RNase T1 digestion products were done by polyacrylamide gel electrophoresis under nondenaturing conditions. The gel consisted of 15% acrylamide including 0.75% bis(acrylamide), 0.08% TEMED, and 0.5% ammonium sulfate. The gels were run at 3V/cm for 18 h at room temperature with 0.1 M KCl, 5 mM MgCl\textsubscript{2}, 50 mM Tris-borate, pH 7.8 as the electrophoresis buffer; RNA was visualized with methylene blue. Figure 43 shows the RNA bands of the three RNase T1-cleaved fragments visualized on the nondenaturing gel before and after purification. As can be seen in the figure, all three fragments show >95% purity on the nondenaturing gel.
Figure 42. Elution profile from Sephadex G-50 gel-filtration chromatography of RNase T1-digested fragments of *B. megaterium* 5S rRNA. Fractions of Fragments A, B, and C are specified.
Figure 43. (A) Visualized bands of RNase T1-cleaved fragments on a nondenaturing electrophoresis gel obtained by gel scan. The second peak is Fragment A and the third one represents Fragments B and C which cannot be resolved in this condition. (B) RNase T1-cleaved fragments are shown on a dried gel obtained by a gel dryer. A, B, and C denote Fragments A, B, and C, respectively.
On the other hand, the three purified fragments visualized on the denaturing gel show different patterns. Fragment A shows a single band, which means that it consists of one pure fragment without a nick. However, Fragment B and Fragment C not only exhibit minor bands but also show a much decreased chain length. Therefore, it is quite certain that Fragment B and Fragment C each either consists of more than one chain combined together or has a nick in the chain. Anyway, compared to the homogeneity of Fragment A, the compositions of the other two fragments seem to be rather complicated.

B. Confirmation of the Primary Nucleotide Sequences of the Three Fragments by RNA T1 Mapping

The primary nucleotide sequences of RNase T1-digested fragments A, B, and C of *B. megaterium* 5S rRNA were confirmed by an RNA sequencing procedure (BRL RNA sequencing system, Bethesda Research Laboratories). This procedure consisted of initial 3'-dephosphorylation with calf intestine alkaline phosphatase followed by 3'-end labeling with [5'-32P] pCp with T4 RNA ligase, purification of the end-labeled fragments, RNA T1 partial digestion, and autoradiography of a subsequent sequencing electrophoresis gel.

1. Experimental Methods for RNA T1 Mapping

a. Preparation of Buffer Solutions

Buffer solutions for RNA T1 mapping were prepared as follows.

10 X dephosphate buffer: 0.5 M Tris-HCl, 10 mM MgCl₂, 0.1 mM ZnCl₂, pH 8.0
10 X ligation buffer: 0.5 M HEPES, 180 mM MgCl₂, 33 mM DTT, pH 7.5
Elution buffer: 0.5 M NH₄OAc, 0.1% SDS, 1 mM EDTA
Alkaline hydrolysis buffer: 50 µL of 0.5 M NaHCO₃/Na₂CO₃, pH 9.2 (from BRL), 1 µL of 0.1 M EDTA, 5 µL of tRNA (5 µg/µL), 44 µL of water (total volume 100 µL)
Sample buffer: 8M urea, 2 X TBE, 10% glycerol, 0.01% bromophenol blue (BPB), 0.01% xylene cyanol (XC)
Solution A: 0.25 M sodium citrate, pH 5.0
Solution D: 10 M urea, 1.5 mM EDTA, 0.05% BPB, 0.05% XC
T1 digestion buffer: 2 µL of Solution A + 14 µL of Solution D + 3 µL of water
2 X EGTA: 0.1 M EGTA, pH 8.0 (pH was adjusted by Tris)
[5'-³²P] pCp: 250 µCi in 25 µL

All buffer solutions except buffers containing DTT, HEPES, ATP, glycerol, BPB, XC, and acrylamide were autoclaved. 1.5 mL Eppendorf tubes were silicaranized by filling them with chloroform containing dimethyldichlorosilane in a dessicator (vacuumed) and leaving them overnight with subsequent rinsing with water. The silicaranized tubes, pipette tips, and other glasswares were also autoclaved.

b. 5'-End Dephosphorylation of RNA with CAP

5 µg of RNA, 5 µL of 10 X dephosphate buffer, and 1.5 U of CAP were placed in a 1.5 mL silicaranized Eppendorf tube, and water was added to give the final volume of 50 µL. After incubation at 50 °C for 1 h, 50 mL of 2 X EGTA was added, followed by reincubation at 65 °C for 45 min. Then 75 µL of phenol was added, vortexed for 5 min, centrifuged for 5 min, and
the aqueous layer was separated into another tube. 50 μL of water was added to the phenol layer, vortexed, centrifuged, and the resulting aqueous layer was combined with the previous aqueous layer. To the aqueous layer 80 μL of phenol was added, vortexed, and centrifuged, and the aqueous layer was separated again. 5 M NaCl (3% the volume of the aqueous layer) was added, followed by 3 volumes of cold 95% ethanol, and the resulting mixture was stored at -80 °C for 1.5h. Then it was centrifuged for 10 min, and after the addition of 150 mL of 75% cold ethanol, was put on ice with subsequent washing for 30 min (vortexed every 10 min). 5'-end dephosphorylated RNA was recovered by centrifugation, and dried under vacuum for 20 min.

c. 3'-End Labeling of RNA with [5'-32P] pCp

The dried 5'-end dephosphorylated RNA, 3 μL of 10 X ligation buffer, 3 μL of DMSO, 1 μL of ATP (1 mM/μL), 25 μCi [5'-32P]pCp (=2.5 μL), 10 U T4 ligation (1 μL of diluted T4 ligase) were placed in a tube, and water was added to give the final volume of 30 μL. It was stored at 4 °C overnight (12–16 h). 1 μL of 5 M NaCl and 3 volumes (~93 μL) of ethanol were added, and the mixture was stored at -80 °C for 1 h. It was centrifuged at 16,000 X g for 20 min, with subsequent removal of the supernatant. The precipitate was dried under vacuum for 20 min.

d. Purification of the Labeled RNA

A 12% polyacrylamide gel with 7 M urea was prepared by mixing 11.4 g of acrylamide, 0.6 g of bisacrylamide, 10 mL of 10 X TBE, 0.3 mL of 10% ammonium persulfate solution, 48 g of urea, 80 mL of TEMED in water (total volume of 100 mL). The labeled RNA was dissolved in 8 μL of
sample buffer. After being heated in the waterbath of 90 °C, RNA was applied to the gel. The gel was run at 300 V for 2 h.

The gel was removed from the electrophoresis apparatus, and a radioactive marker was made on the gel. The gel, covered with Saran Wrap, was exposed to an X-ray film for 5 min in the darkroom. The exposed film was developed for 5 min, and fixed for another 5 min. By matching the marker in the gel with its exposure on the film, the band containing the labeled RNA was located and sliced. The sliced gel was smashed in a homogenizer with 1 mL of elution buffer. The resulting buffer was transferred into a tube. Another 1 mL of elution buffer was added to the homogenizer. The rinsed buffer was then combined with the previous portion of the buffer solution, with subsequent vortexing for 1 min. The buffer solution was left at room temperature overnight. The supernatant was separated from the gel fragments by centrifugation. To the supernatant, ~5 μg of tRNA and 3 volumes of ethanol were added and stored at -80 °C for 1 h.

The purified labeled RNA was divided into several portions so that T1 digestion and alkaline hydrolysis could be done. The RNA portion for alkaline hydrolysis needs three times the radioactivity (in cpm) of the RNA portion for T1 digestion. Each portion was centrifuged for 20 min to get RNA precipitate. The precipitate was washed with 150 mL of 75% ethanol (on ice, vortexed every 10 min, centrifuged to recover the RNA precipitate) and dried for 20 min under vacuum.

e. T1 Digestion and Alkaline Hydrolysis

T1 digestion was accomplished as follows. 5 μg of tRNA (~1 μL) and 5 μL of T1 digestion buffer was added to the dried RNA, and pre-
equilibrated in a waterbath of 37 °C for 20 min. 1 µL of RNase T1 was added (the concentration of RNase T1 had to be adjusted so that it could produce a homogeneous digestion pattern) followed by incubation at 37 °C for 10 min. The digested RNA was frozen at -80 °C immediately and stored at that temperature until use.

Alkaline hydrolysis was performed as follows. 7 µL of alkaline hydrolysis buffer was added to the dried RNA. The RNA in the buffer was drawn into a capillary, with subsequent sealing of both ends of the capillary. The capillary was immersed in a waterbath of 90 °C for 15 min. The RNA was then drawn out of the capillary into a tube. 7 mL of sample buffer was added, and the RNA was stored at -80 °C until use.

f. Sequencing Gel

The gel plates were cleaned thoroughly with detergent and water. The plates were then dried with acetone, and coated with silicon. A 12% sequencing gel was prepared by dissolving 11.4 g of acrylamide monomer, 0.6 g of bisacrylamide, and 50 g of urea in water containing 10 mL of 10 X TBE. The total volume was adjusted to 100 mL by adding water. 0.3 mL of 10% ammonium persulfate solution and 75 µL of TEMED were added to polymerize the gel. The gel was polymerized for several hours. The gel was then pre-run for half an hour before samples were applied. The T1-digested and the alkaline-hydrolyzed RNA samples were heated in a waterbath of 90 °C for 30 sec, and applied to the gel. The gel was run at 2000 V (30-40 mA) for 2-3 h depending on the RNA size. The gel was then fixed in a solution containing 5% methanol and 5% acetic acid, placed on a filter paper, and vacuum dried by a gel drier for 2 h. A film was exposed to the gel, developed, and fixed.
2. Sequencing Results

By RNA sequence analysis, RNase T1-cleaved fragment A was confirmed to consist of residues 15-65 of 5S rRNA (Figure 44, top). Fragment B showed a rather complicated pattern. Under denaturing conditions, fragment B exhibited a few fragments of different sizes consisting of bases between residues 66-98, but consisted mainly of residues 78-98 (Figure 44, middle). Fragment C is believed to correspond to mainly residues 1-14 and 99-116 (Figure 44, bottom) with a small amount comprising a variety of short fragments.

C. Identification and Assignment of Base Pairs from Fragment A by Proton Homonuclear Overhauser Enhancements

Fragment A consists of helices II and III of *B. megarhizium* 5S rRNA. Because of its homogeneity and because it is largest in size of the three fragments, Fragment A is of primary interest. Therefore, the structure of Fragment A was investigated by $^1$H NMR spectroscopy.

1. Mg$^{2+}$-Depleted NMR Sample Preparation

A Mg$^{2+}$-depleted NMR sample of Fragment A was prepared as follows. Fragment A was recovered from ethanol by centrifugation, and dried thoroughly with the blowing of filtered N$_2$ gas. The dried sample was dissolved in 2 mL of 10 mM EDTA, 100 mM NaCl, 10 mM cacodylic acid, pH 7.0, and dialyzed against 1000 volumes of 1 mM EDTA, 100 mM NaCl, 10 mM cacodylic acid, pH 7.0 at 4 °C for 5h. The sample was then dialyzed against 1000 volumes of freshly prepared buffer overnight, and concentrated by ultrafiltration using an Amicon Centricon apparatus.
Figure 44. Proposed secondary structure segments of RNase T1-cleaved Fragment A (top), B (middle), and C (bottom) of *B. megaterium* 5S rRNA obtained by RNA T1 mapping. While Fragment A consists of one homogeneous single chain, Fragments B and C show heterogeneity.
The concentration of fragment A in the NMR sample (in 95% H₂O:5% D₂O) was 225 A₂₆₀ units in 0.5 mL.

2. NMR Spectroscopy

Proton FT/NMR spectra were obtained with a Bruker AM-500 FT-NMR spectrometer, by use of phase-cycled quadrature detection, without sample spinning. Water suppression was achieved by use of a 1331 hard pulse sequence (Hore, 1983b). Chemical shifts were measured relative to the water resonance, which was taken as 4.78 ppm relative to DSS [3-(trimethylsilyl)-1-propanesulfonic acid] at all temperatures.

The nuclear Overhauser enhancement (NOE) difference spectra of Fragment A were produced with a 1331 hard-pulse sequence, with an acquisition period of 0.67 s [16 K time-domain data, typically 16,000 time-domain transients (8000 on-resonance and 8000 off-resonance)], 0.5 s preirradiation (0.2 mW decoupler power) of the resonance of interest, with off-resonance irradiation at 18 ppm. All NOE experiments for Fragment A were carried out at 23 °C, unless stated otherwise below. Exponential apodization equivalent to frequency-domain line broadening of 1-5 Hz was usually applied to improve the signal-to-noise ratio in the final spectrum, and the spectrum was plotted without baseline correction.

3. Base Pairs in Fragment A (Helices II and III)

The RNase T1-cleaved Fragment A of *B. megaterium* 5S rRNA contains helices II and III (see Figures 45). Eight distinct peaks (which contain more than 10 imino proton resonances) are observed in the downfield proton NMR spectrum of Fragment A, and many of these
Figure 45. The downfield proton NMR spectrum of Fragment A at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95%:5% H₂O/D₂O.
resonances in this fragment are well resolved (at the same chemical shifts) in the spectrum of intact 5S rRNA, indicating that the secondary structure in the intact 5S rRNA is preserved in Fragment A.

NOE connectivities obtained by successive irradiation of these resonances are summarized in Table 4, and shown in NOE difference spectra of the resonances in Figures 46 and 47. Peak I" (G·C) actually consists of two separate peaks (two imino proton resonances) whose chemical shifts are too close to be irradiated selectively. Peak A₁ (A·U) also consists of two resonances, one of which melts at a slightly elevated temperature (Figure 48). The other resonance A₁ (A·U), together with F₁ (G·C) and I" (G·C), is quite stable at high temperature. From the NOE connectivities, the base pair sequence information can be inferred as follows. Resonance F₁ is NOE-connected to A₁, and resonance A₁ to F₁ and I" (Figure 46). Peak I" (G·C) is connected to A, B (A·U), and J₁ (G·C) (Figure 46), and peak J₁ to I" and I (Figure 47). At this stage, there are several possible combinations of resonances for the establishment of base pair sequences, since peaks A₁ and I" each contain two resonances. Although unequivocal assignment is not possible, it is very likely that F₁ – A₁ – I" – B is G₂₁·C₅₈ – A₂₀·U₅₉ – G₁₉·C₆₀ – A₁₈·U₆₁ and A₁ – I" – J₁ – I is U₃₂·A₄₆ – G₃₁·C₄₇ – C₃₀·G₄₈ – C₂₉·G₄₉, since those assignments fully account for the observed NOE connectivities manifested in the NOE difference spectra produced by successive irradiation at peaks A₁, B, F₁, I, I", and J₁. The high stability of one of the A₁ peaks (A₂₀·U₅₉) observed in the variable temperature experiment (Figure 48) can be explained by its sandwiched location between two G·C pairs (G₂₁·C₅₈ and G₁₉·C₆₀) (Table 4).
Figure 46. NOE difference spectra of resonances $A_1$, $F_1$, $I''$, and $B$ of Fragment A obtained at 23 °C.
Figure 47. NOE difference spectra of resonances I, J₁, H, and D₁ of Fragment at 23 °C.
Table 4. Identification and assignment of base pair imino proton resonances in *B. megaterium* 5S rRNA Fragment A by means of NOE connectivities.

<table>
<thead>
<tr>
<th>Peak</th>
<th>Chemical Shift</th>
<th>Base Pair</th>
<th>NOE connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>14.18</td>
<td>A₂₀•U₅₉ and U₃₂•A₄₆</td>
<td>F₁, I''</td>
</tr>
<tr>
<td>B</td>
<td>13.90</td>
<td>A₁₈•U₆₁</td>
<td>I''</td>
</tr>
<tr>
<td>D₁</td>
<td>13.44</td>
<td>unassigned</td>
<td>H</td>
</tr>
<tr>
<td>F₁</td>
<td>13.18</td>
<td>G₂₁•C₅₈</td>
<td>A₁</td>
</tr>
<tr>
<td>H</td>
<td>12.73</td>
<td>unassigned</td>
<td>D₁</td>
</tr>
<tr>
<td>I</td>
<td>12.52</td>
<td>C₂₉•G₄₉</td>
<td>J₁</td>
</tr>
<tr>
<td>I''</td>
<td>12.44</td>
<td>G₁₉•C₆₀ and G₃₁•C₄₇</td>
<td>A₁, B, J₁</td>
</tr>
<tr>
<td>J₁</td>
<td>12.19</td>
<td>C₃₀•G₄₈</td>
<td>I, I''</td>
</tr>
</tbody>
</table>
A few secondary structural questions remain. One G-C and one A-U resonance (H and D₁), which show strong mutual NOE connectivity (Figure 47), are not assigned. It is interesting to note that an A-U pair which melts at slightly higher temperature (D₁) is spatially adjacent to a G-C pair which is very stable at high temperature (H) (see Figure 48). Another peak I', which is weak in intensity and hidden between two intense signals (H and I), cannot be assigned at this juncture. The possible sources of these extra resonances are intra-loop base pairs and the base pairs from the region between helices II and III. While intra-loop base pairs have been proposed for the "common arm" fragment of wheat germ 5S rRNA (Li et al., 1987), those base pairs have been reported to be unlikely in chloroplast 5S rRNA from spinach (Romby et al., 1988). The extra resonances in Fragment A of B. megaterium 5S rRNA do not seem to correspond to the intra-loop base pairs in the light of their NOE connectivities; they more likely result from possible base pairs in the region between helices II and III. Imino protons located near the open end of the fragment appear to exchange rapidly with solvent, since no resonances for A₁₁₋₁₇U₆₂ or G₁₆₋₁₆U₆₃ are observed and peak B (A₁₈₋₁₈U₆₁) melts at relatively lower temperature than most peaks. A very broad, very weak signal (which melts out completely at 35 °C) between 10.5-11 ppm is believed to arise from the resonances of the imino protons of unpaired uridines or guanidines inside the loop, whose exchange rates with water are not slow enough to show detectable signals.

Unlike "fragment 2" of E. coli 5S rRNA (Leontis & Moore, 1986), the helix II and III region of B. megaterium does not show as many imino resonances—no resonances exist in 9-12 ppm of Fragment A, while
**Figure 48.** Heat-induced melting pattern of the proton 500-MHz NMR spectrum of Fragment A in 10 mM cacodylic acid, 0.1 M NaCl, and 1 mM EDTA at pH 7.0.
several resonances can be found in this region of "fragment 2"—suggesting that Fragment A of *B. megaterium* 5S rRNA may not be so extensively base-paired in those helical regions as the corresponding segments of *E. coli* 5S rRNA. Melting of Fragment A appears to be quite reversible. Virtually all imino proton resonances disappeared by \(-45^\circ C\), but the original resonances reappeared upon decrease in temperature.

Although these assignments of resonances from Fragment A are not complete, they do establish that helices II and III are present in 5S rRNA of Gram-positive bacteria. As base pairs in helices II and III have been already identified and assigned in 5S rRNA from a Gram-negative bacterium, *E. coli* (Leontis & Moore, 1986) and a eukaryote, *Triticum aestivum* (wheat germ) (Li & Marshall, 1986; Li *et al.*, 1987), the helical segments II and III appear to occur universally among prokaryotic (both Gram-positive and Gram-negative bacteria) and eukaryotic 5S rRNA’s.

4. Conformations of RNase T1-Cleaved Fragment A and the Corresponding Segment of Intact 5S rRNA

Comparison of the proton NMR spectrum of the enzyme-cleaved Fragment A with the proton NMR spectrum of intact 5S rRNA serves to establish whether or not the fragments retain the original conformation of the intact 5S rRNA molecule. Since virtually all of the imino proton resonances observed in the spectrum of Fragment A (Figure 45) can be found at the same chemical shifts in intact 5S rRNA, the base-paired regions of the fragment are likely to be highly homologous in secondary structure to those of intact 5S rRNA. Moreover, demonstration of the same NOE connectivity for an imino proton resonance of a fragment and
the corresponding resonance in intact 5S rRNA further corroborates the similarity of conformations between them (Li & Marshall, 1986). For example, irradiation of peak A of intact 5S rRNA shows NOE connectivities to resonances I"', F₁, D₂, which are observed separately by irradiation of peak A₁ in Fragment A (NOE connectivity to F₁ and I'') and peak A₂ in Fragment B (NOE connectivity to D₂ and I'', to be discussed in Chapter 5), as can be seen in Figure 49. These near-identical NOE patterns for Fragment A and intact 5S rRNA strongly support the retention of conformation on cleavage of 5S rRNA into Fragment A.

D. Dynamic Structure of Fragment A

The primary nucleotide sequence spanning helices II and III, to which ribosomal protein L18 binds, is recognized as important to the ribosomal function of 5S rRNA (Peattie et al., 1981). The existence of more than one conformation of E. coli 5S rRNA in this region has been demonstrated (Kao & Crothers, 1980), and it has been suggested that some conformational states are biologically active whereas others are not (Christensen et al., 1985). Two different conformations of the helix II-III segment of 5S rRNA proposed by de Wachter et al. (1982, 1984) have been reported to interconvert at a rate which depends on temperature and Mg²⁺ concentration, based on ¹H NMR studies of the RNase-digested fragment of E. coli (a Gram-negative bacterium) 5S rRNA containing helices II and III (Leontis & Moore, 1986).

Fragment A has been shown to retain the original conformation of the corresponding segment of intact B. megaterium 5S rRNA, and the downfield (10-15 ppm) resonances of 500-MHz proton NMR spectrum
Figure 49. Comparison of the NOE connectivities of intact 5S rRNA with those of the corresponding enzyme-cleaved Fragment A. These homologous NOE connectivities in intact 5S rRNA and Fragment A strongly support the retention of secondary structure on enzymatic cleavage.
have recently been assigned to the corresponding base pairs in helices II-III. This section describes the conformational flexibility of the helix II-III segment of *B. megaterium* 5S rRNA (see Figure 50) investigated by 500-MHz proton NMR spectroscopy. Two resonances, D1 and H, which had not previously been assigned, can now be assigned as A27•U53 and C28•G52. The possibility of two interconvertible conformations was examined by proton NMR monitoring of conformational changes induced by variation in temperature and Mg2+ concentration.

1. **NMR Spectroscopy**

   A Mg2+-depleted NMR sample of *B. megaterium* 5S rRNA Fragment A was prepared as described above. The concentration of Fragment A in the NMR sample (in 10 mM sodium phosphate, 1 mM EDTA, 100 mM NaCl, pH 7.0 with 95% H2O: 5% D2O) was ~1 mM. For Mg2+ titration, aliquots of 0.5 M MgCl2 solution were added to produce the desired Mg2+ concentration. Proton FT/NMR spectra were obtained with a Bruker AM-500 FT-NMR spectrometer in the same way as before.

2. **Conformational Change of Fragment A**

   Figure 51 shows the change in the resonance intensity of Fragment A as temperature decreases, and Figure 52 shows Mg2+ titration of Fragment A monitored in the imino proton range. As can be seen from Figures 51 and 53, the relative intensity of resonance B (assigned as A18•U61—see Table 5) decreases relatively rapidly with increase of temperature while the intensity of F1 (assigned as G21•C58) does not change appreciably. Also, resonance B exhibits a smaller temperature-induced chemical shift than resonance F1 (Figure 54). It is well-known
Figure 50. Two possible conformations for *B. megaterium* 5S rRNA Fragment A. Switch between these two conformations is not observed in the present experiments; rather Fragment A remains as "conformation 1" under all solution conditions.
Figure 51. Variable temperature experiments of Fragment A with a 500-MHz proton NMR spectrometer. Fragment A in 0.01 M cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 was cooled down to specified temperatures to monitor its structural change.
Figure 52. Mg$^{2+}$ titration experiments of Fragment A in 0.01 M sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0. The NMR spectra were obtained at 23 °C with various concentrations of Mg$^{2+}$. 
Figure 53. Variation with temperature of relative peak heights of base-paired imino proton resonances in the 500-MHz $^1$H NMR spectrum of Fragment A. Vertical scaling has been chosen for convenience in discussion.
Figure 54. Variation with temperature of chemical shifts of base-paired imino proton resonances in the 500-MHz $^1H$ NMR spectrum of Fragment A.
Table 5. NOE connectivities of *B. megaterium* 5S rRNA Fragment A obtained under three different conditions.

<table>
<thead>
<tr>
<th>Resonances</th>
<th>Assignments</th>
<th>NOE connectivity to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>I&lt;sup&gt;a&lt;/sup&gt;</td>
<td>I&lt;sup&gt;b&lt;/sup&gt;</td>
<td>II&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>A&lt;sub&gt;1&lt;/sub&gt; A&lt;sub&gt;1&lt;/sub&gt; A&lt;sub&gt;1&lt;/sub&gt;</td>
<td>A&lt;sub&gt;20&lt;/sub&gt;•U&lt;sub&gt;59&lt;/sub&gt;, U&lt;sub&gt;32&lt;/sub&gt;•A&lt;sub&gt;46&lt;/sub&gt;</td>
<td>F&lt;sub&gt;1&lt;/sub&gt;, I'&lt;sub&gt;1&lt;/sub&gt;, I''&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
<tr>
<td>A&lt;sub&gt;3&lt;/sub&gt;</td>
<td>(U&lt;sub&gt;32&lt;/sub&gt;•A&lt;sub&gt;46&lt;/sub&gt;)</td>
<td></td>
</tr>
<tr>
<td>B B B</td>
<td>A&lt;sub&gt;18&lt;/sub&gt;•U&lt;sub&gt;61&lt;/sub&gt;</td>
<td>I''&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>D&lt;sub&gt;1&lt;/sub&gt; D&lt;sub&gt;1&lt;/sub&gt; D&lt;sub&gt;1&lt;/sub&gt;</td>
<td>A&lt;sub&gt;27&lt;/sub&gt;•U&lt;sub&gt;53&lt;/sub&gt;</td>
<td>H</td>
</tr>
<tr>
<td>F&lt;sub&gt;1&lt;/sub&gt; F&lt;sub&gt;1&lt;/sub&gt; F&lt;sub&gt;1&lt;/sub&gt;</td>
<td>G&lt;sub&gt;21&lt;/sub&gt;•C&lt;sub&gt;58&lt;/sub&gt;</td>
<td>A&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>H H H</td>
<td>C&lt;sub&gt;28&lt;/sub&gt;•G&lt;sub&gt;52&lt;/sub&gt;</td>
<td>D&lt;sub&gt;1&lt;/sub&gt;, I, J&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>I I I</td>
<td>C&lt;sub&gt;29&lt;/sub&gt;•G&lt;sub&gt;49&lt;/sub&gt;</td>
<td>J&lt;sub&gt;1&lt;/sub&gt;, H</td>
</tr>
<tr>
<td>I''&lt;sub&gt;1&lt;/sub&gt; I''&lt;sub&gt;1&lt;/sub&gt; I''&lt;sub&gt;1&lt;/sub&gt;</td>
<td>G&lt;sub&gt;19&lt;/sub&gt;•C&lt;sub&gt;60&lt;/sub&gt;</td>
<td>J&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>I''&lt;sub&gt;2&lt;/sub&gt; I''&lt;sub&gt;2&lt;/sub&gt; I''&lt;sub&gt;2&lt;/sub&gt;</td>
<td>G&lt;sub&gt;31&lt;/sub&gt;•C&lt;sub&gt;47&lt;/sub&gt;</td>
<td>J&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>J&lt;sub&gt;1&lt;/sub&gt; J&lt;sub&gt;1&lt;/sub&gt; J&lt;sub&gt;1&lt;/sub&gt;</td>
<td>C&lt;sub&gt;30&lt;/sub&gt;•G&lt;sub&gt;48&lt;/sub&gt;</td>
<td>I, I''&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
<tr>
<td>J&lt;sub&gt;1&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Condition I: 3 °C, [Mg<sup>2+</sup>] = 0

<sup>b</sup> Condition II: 23 °C, [Mg<sup>2+</sup>] = 0

<sup>c</sup> Condition III: 23 °C, [Mg<sup>2+</sup>] = 15 mM

Note: Resonances I''<sub>1</sub> and I''<sub>2</sub> cannot be irradiated selectively.
that imino proton resonances of helix-terminal and unpaired imino proton resonances show smaller temperature-induced chemical shifts and more rapid broadening than resonances of imino protons well-protected from solvent, such as imino protons in the middle of a base-paired helical segment. From the above results, it is possible to conclude that the open end of helix II is more exposed to solvent and is less thermally stable than the inter-helical region. The absence of observable NMR resonances corresponding to G₁₆\textit{•}U₆₃ and A₁₇\textit{•}U₆₂ at room temperature further supports this interpretation. As temperature increases, helix II opens from the open end of the helix. However, helix II appears to be more stable than helix III, since resonances F₁, A₁, and I" (G₂₁\textit{•}C₅₈, A₂₀\textit{•}U₅₉, and G₁₉\textit{•}C₆₀) show relatively greater high-temperature stability than do resonances J₁, I (C₃₀\textit{•}G₄₈ and C₂₉\textit{•}G₄₉).

Increase in Mg²⁺ concentration also induces a more significant change in conformation of helix III than of helix II, as seen from the greater change in chemical shift (see Figures 52, 55, and 56) for base pair imino protons from helix III than from helix II. More specifically, helix III seems to tighten slightly upon increase in [Mg²⁺], as judged from the downfield shifts of resonances I, J₁, and I". The significant decrease in the intensity of resonance F₁ with increase of Mg²⁺ concentration implies that Mg²⁺ acts to increase solvent access to G₂₁\textit{•}C₅₈ in helix II, leading to faster exchange of the G₂₁ imino proton with water.

Interestingly, the broad signal (labeled as M₃), which is thought to consist of superimposed resonances from unpaired imino protons of guanine and uracils in the loop region, shows no noticeable change in
Figure 55. Variation with Mg$^{2+}$ concentration of relative peak heights of base-paired imino proton resonances in the 500-MHz $^1$H NMR spectrum of Fragment A.
Figure 56. Variation with Mg$^{2+}$ concentration of chemical shifts of base-paired imino proton resonances in the 500-MHz $^1$H NMR spectrum of Fragment A.
chemical shifts and no intensity change on addition of MgCl₂. Thus, the loop region appears to retain its original shape upon increase in [Mg²⁺].

Resonances D₁ and H, which could not be assigned to specific base pairs in the previous chapter, can now be assigned as A₂⁷•U₅₃ and C₂⁸•G₅₂, based on their strong mutual homonuclear Overhauser enhancements (NOE's) under the wide range of temperature and [MgCl₂] of the present experiments. In particular (see Table 5), resonance H shows NOE connectivity to resonance I (C₂⁹•G₄₉) and a smaller but measurable NOE to resonance J₁ (C₃₀•G₄₈) for both [Mg²⁺] = 0 and [Mg²⁺] = 15 mM.

3. Evidence against Two Conformations of the Helix II-III Segment

The present results offer several arguments that the two-conformation switch proposed by de Wachter et al. (1982, 1984) and supported by Leontis and Moore's experiments (Leontis & Moore, 1986) for E. coli 5S rRNA does not occur in B. megaterium 5S RNA (see Figure 50). First (see Figures 51 and 53), resonances D₁ and H (A₂⁷•U₅₃ and C₂⁸•G₅₂) do not broaden and decrease in intensity quickly upon increase in temperature; in fact, resonance H is quite stable at high temperature. Also, the failure to observe any resonances which decrease in linewidth and increase in intensity with increase in temperature strongly suggests that no new base pairs are being formed at higher temperature. Although the intensities of D₁ and H increase slightly with increase in Mg²⁺ concentration (see Figure 55), their chemical shifts do not vary strongly, suggesting that the base pairs A₂⁷•U₅₃ and C₂⁸•G₅₂ remain intact from [Mg²⁺] = 0 – 15 mM.
Additional strong evidence against two interconverting conformations is provided by the NOE experiments shown in Figures 57-60, and summarized in Table 5. There are no (or at most relatively minor) differences between NOE connectivities obtained under three conditions: 3 °C with [Mg$^{2+}$] = 0, 23 °C with [Mg$^{2+}$] = 0, and 23 °C with [Mg$^{2+}$]=15 mM. Thus, it appears that the same base-pairing pattern for the helix II-III segment of *B. megaterium* 5S rRNA is conserved on change in temperature or [Mg$^{2+}$].

It is not obvious why 5S rRNAs from Gram-negative bacteria (e.g., *E. coli*) should exhibit two conformations (Christensen *et al*., 1985; Leontis & Moore, 1986) while 5S rRNA's from Gram-positive bacteria (e.g., *B. megaterium* in the present work) do not. It would seem more likely that all bacterial 5S rRNAs should have similar conformations, since (for example) the ribosomal protein L18 from *Bacillus stearothermophilus* can bind to 5S rRNA from *E. coli* (Hartmann *et al*., 1988). Perhaps the *B. megaterium* 5S rRNA achieves sufficient conformational change by tightening or loosening its helix II-III region that it is not necessary to break and re-form base pairs (as in *E. coli* 5S rRNA). In any case, since it is difficult to generalize from two examples, it may be worth investigating the conformations of a few more 5S rRNAs (both Gram-positive and Gram-negative) to see if either of the above results represents a general behavior.
Figure 57. NOE difference spectra of resonances A₁, B, D, and F₁ in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with 15 mM Mg²⁺ obtained at 23 °C.
Figure 58. NOE difference spectra of resonances H, I', I, and I'' in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with 15 mM Mg$^{2+}$ obtained at 23 °C.
Figure 59. NOE difference spectra of resonances $A_1$, $B$, $D$, and $F_1$ in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 at 3 °C.
**Figure 60.** NOE difference spectra of resonances H, I, and I'' in 10 mM sodium phosphate, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with 15 mM Mg^{2+} obtained at 3 °C.
E. Two-Dimensional NOE Spectroscopy (NOESY) of Fragment A in Water

Due to low concentration (~1 mM) and high molecular weight, two-dimensional NMR spectroscopy of Fragment A in water cannot provide much structural information. However, it showed NOEs between imino protons and amino protons by their cross-peaks, which could not be observed clearly in one-dimensional NOE experiments.

NOESY spectra in water were obtained with phase-sensitive mode, pure absorption line shape by using the time-proportional phase increment (TPPI) method (Redfield & Kunz, 1975; Marion & Wüthrich, 1983; Wu & Marshall, 1990a). The pulse sequence was $[\text{delay} - 90° - t_1 - 90° - t_m - \text{pulse} - \text{FID}(t_2)]_n$, where 1331 pulse sequence was used for effective water suppression. Transmitter power was attenuated to give 90° pulse width of 20 μs. 4K time-domain data in the $t_2$ dimension and 256 words in the $t_1$ dimension were acquired. Both the delay and the mixing time ($t_m$) were 0.3 s. The mixing time was varied randomly within 10% range to suppress scalar coupling correlations. For each $t_1$ value, 1024 scans were accumulated. Before FT operation, each time-domain data set was treated with a DSA (Haasnoot & Hilbers, 1983; Heerschap et al., 1985), zero-filled to 2K in the $t_1$ dimension to give better digital resolution.

The NOESY spectrum of Fragment A in the Mg$^{2+}$-depleted buffer solution is shown in Figure 29. Figure 61 is the NOESY spectrum of Fragment A in the buffer solution containing 15 mM Mg$^{2+}$, which shows imino-amino cross-peaks better than the one without Mg$^{2+}$. From the
**Figure 61.** Cross-peaks between imino and amino or aromatic protons observed in NOESY spectrum of 1 mM *B. megaterium* 5S rRNA Fragment A in 10 mM sodium phosphate, 100 mM NaCl, 1 mM EDTA, 15 mM MgCl₂, pH 7.0 at 23 °C. A 1331 pulse sequence was used for water suppression. Experimental conditions are given in the text.
spectra, it is evident that most peaks consist of more than one resonance.
CHAPTER V
FRAGMENT B AND FRAGMENT C: BASE-PAIRINGS IN HELICES IV AND I OF B. MEGATERIUM RIBOSOMAL 5S RNA

Of the three fragments of *Bacillus megaterium* ribosomal 5S RNA produced by enzymatic cleavage with ribonuclease T1, Fragment A was used to investigate the helices II-III region of 5S rRNA in the previous chapter. In this chapter, Fragment B and Fragment C will be studied to obtain deeper knowledge of base-pairing patterns in other helical regions (helices IV and I) of 5S rRNA. As will be shown in the following section, all (eight) imino proton resonances in the downfield region (9-15 ppm) of the 500 MHz proton FT/NMR spectrum of Fragment B have been identified and assigned as G80*C92 - G81*C91 - G82*C90 - A83*U89 - C84*G88 and three unpaired U's (U85, U86, and U87) in helix IV by proton homonuclear Overhauser enhancement connectivities. The secondary structure of helix IV of the prokaryotic loop is completely demonstrated spectroscopically for the first time in any native or enzyme-cleaved 5S rRNA. In addition, the identification and assignment of the G4*C112 - G5*C111 - U6*G110 segment in the terminal stem (helix I) of *B. megaterium* 5S rRNA (in Chapter 3) has been confirmed by means of NOE
experiments and variable temperature experiments on Fragment C. Base pairs in helices I and IV of the universal secondary structure of B. megaterium 5S RNA will be discussed.

A. Base-Pairings in the Prokaryotic Loop (Helix IV) of B. megaterium 5S rRNA (NMR Analysis of Fragment B)

Production and purification of Fragment B has been discussed in the previous chapter. According to RNA T1 mapping results, Fragment B mainly consists of the prokaryotic loop (helix IV) of B. megaterium 5S rRNA. Although base pairs from this helix have previously been identified by NOE connectivities in E. coli 5S rRNA, those assignments were not complete.

1. NMR Spectroscopy

The NMR sample of RNase T1-cleaved Fragment B of B. megaterium 5S rRNA was prepared in the same way as described for Fragment A in the previous chapter. The final Fragment B concentration of the NMR sample was 144 A_{260} units in 0.5 mL of 1 mM EDTA, 100 mM NaCl, 10 mM cacodylic acid, pH 7.0 (95% H_{2}O: 5% D_{2}O). The NMR sample exhibited at least 95% purity as estimated from electrophoresis gel scan obtained under non-denaturing condition.

All 1H NMR spectra were obtained with a Bruker AM-500 FT/NMR spectrometer, with phase-cycled quadrature detection, without sample spinning under the same condition as Fragment A. Water suppression in NMR spectra was achieved by use of a 1331 hard-pulse sequence (Hore, 1983b) with carrier frequency centered at the H_{2}O resonance.
2. Assignment of Base-Pair Sequence in Fragment B (Helix IV)

RNase T1-cleaved Fragment B contains helix IV (the prokaryotic loop) of *B. megaterium* 5S rRNA (see Figure 44 in the previous chapter). As can be seen in Figure 62, the imino proton resonances observed in this fragment are located in the most crowded region of the downfield spectrum of 5S rRNA. Hence, Fragment B offers an optimal window for study of the secondary structure in helix IV. Helix IV has been reported to be the part of 5S rRNA whose proton NMR spectrum is most strongly affected by binding of ribosomal protein L25 (Kime & Moore, 1983b). It is therefore believed that the binding site for ribosomal protein L25 on 5S rRNA includes helix IV.

Eight imino proton resonances were observed in the downfield spectrum of Fragment B. Five gave sharp peaks, and the other three upfield peaks were weaker. It appears likely that the five strong resonances (A₂, D₂, I₂, I, and I") are due to the base-pair sequence, G₈₀•C₉₂ – G₈₁•C₉₁ – G₈₂•C₉₀ – A₈₃•U₈₉ – C₈₄•G₈₈. The chemical shifts of the resonances can be estimated from ring currents on the basis of an 11-fold RNA A-helix (Arter & Schmidt, 1976), from assumed intrinsic A•U and G•C positions of 14.35 and 13.45 ppm, respectively (Reid et al., 1979). Although ring-current estimates of chemical shifts are far from perfect, they were nevertheless quite helpful in establishing a self-consistent set of assignments of all of these base pair imino proton resonances in tRNAPhe (Johnston & Redfield, 1981). The calculated chemical shift values for G₈₀•C₉₂ – G₈₁•C₉₁ – G₈₂•C₉₀ – A₈₃•U₈₉ – C₈₄•G₈₈ match reasonably well with the experimental values obtained from the proposed assignments of resonances, I₂, I, I", A₂, and D₂ (Table
Figure 62. The downfield proton NMR spectrum of Fragment B at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95 %:5 % H₂O/D₂O.
Table 6. Observed and ring current-calculated base pair imino proton chemical shifts for Fragment B of *B. megaterium* 5S rRNA.

<table>
<thead>
<tr>
<th>Peak&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Base Pair</th>
<th>Calculated chemical shifts&lt;sup&gt;b&lt;/sup&gt; (ppm)</th>
<th>Observed chemical shifts (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D&lt;sub&gt;2&lt;/sub&gt;</td>
<td>C&lt;sub&gt;84&lt;/sub&gt;•G&lt;sub&gt;88&lt;/sub&gt;</td>
<td>13.26</td>
<td>13.39</td>
</tr>
<tr>
<td>A&lt;sub&gt;2&lt;/sub&gt;</td>
<td>A&lt;sub&gt;83&lt;/sub&gt;•U&lt;sub&gt;89&lt;/sub&gt;</td>
<td>13.97</td>
<td>14.10</td>
</tr>
<tr>
<td>I''</td>
<td>G&lt;sub&gt;82&lt;/sub&gt;•C&lt;sub&gt;90&lt;/sub&gt;</td>
<td>12.33</td>
<td>12.44</td>
</tr>
<tr>
<td>I</td>
<td>G&lt;sub&gt;81&lt;/sub&gt;•C&lt;sub&gt;91&lt;/sub&gt;</td>
<td>12.76</td>
<td>12.51</td>
</tr>
<tr>
<td>I'2</td>
<td>G&lt;sub&gt;80&lt;/sub&gt;•C&lt;sub&gt;92&lt;/sub&gt;</td>
<td>12.97</td>
<td>12.60</td>
</tr>
</tbody>
</table>

<sup>a</sup> See Figure 62.

<sup>b</sup> Calculated for an 11-fold RNA A-helix (Arter & Schmidt, 1976) assuming intrinsic A•U and G•C positions of 14.35 and 13.45 ppm, respectively (Reid *et al.*, 1979).
The other three upfield resonances could be due to the N3-H protons of three unpaired uridines in the loop (U₈₅, U₈₆, and U₈₇). It has been reported that the chemical shifts for N3-H of uridine and N1-H of guanosine dissolved separately in dry Me₂SO-d₆ occur at 11.4 and 10.7 ppm (Hurd & Reid, 1979a,b). Some of the N3-H U and N1-H G protons in RNA may be shielded from exchange with H₂O. Those unpaired imino protons can thus give observably narrow proton NMR signals between 9 and 11 ppm (Hare & Reid, 1982; Chen & Marshall, 1986).

The assignment of the resonances I′₂, I, I″, A₂, D₂, M₁, O, and M₂ to G₈₀·C₉₂ − G₈₁·C₉₁ − G₈₂·C₉₀ − A₈₃·U₈₉ − C₈₄·G₈₈ and U₈₅, U₈₆, U₈₇ can be confirmed more convincingly from the NOE difference spectra of eight irradiated resonances (Figures 63-65). For example, resonance A₂ can easily be identified as an A·U base pair by its chemical shift (14.1ppm) and the sharp NOE difference peak from adenine C2-H at ~7-8 ppm. As predicted above, A₂ (A₈₃·U₈₉) is NOE-connected to D₂ (C₈₄·G₈₈) and I″ (G₈₂·C₉₀). D₂ is connected to A₂ and M₂ (U₈₇), M₂ to D₂ and O (U₈₆), O to M₁ (U₈₅) and M₂, and M₁ to O. Irradiation of peak I″ gives an NOE difference signal at A₂, but NOE of I (G₈₂·C₉₀) cannot be observed because of spillover. At this stage, the definitive NOE connectivities are sufficient to establish the base pair sequences G₈₀·C₉₂ − G₈₁·C₉₁ − G₈₂·C₉₀ − A₈₃·U₈₉ − C₈₄·G₈₈ (resonances I′₂, I, I″, A₂, D₂) and U₈₅, U₈₆, U₈₇ (resonances M₁, O, M₂) (Table 7). [The NOE difference signals corresponding to resonances I′₂ (G₈₀·C₉₂) and I (G₈₁·C₉₁) cannot be observed because any NOE effect is obscured by the pronounced spillover from irradiation.]
Figure 63. The NOE difference spectra of resonances $A_2$, $D_2$, and $M_2$ of Fragment B. It is evident that $D_2$ (G·C) is connected to $A_2$ (A·U) and $M_2$ (unpaired U), and $M_2$ is connected to $D_2$ and O (unpaired U). Irradiation of $A_2$ shows NOE connectivity to $D_2$ and I" (G·C).
Figure 64. The NOE difference spectra of resonances O and M₁ of Fragment B. From Figures 63 and 64, the sequence, I'' (G·C) - A₂ (A·U) - D₂ (G·C) - M₂ (U) - O (U) - M₁ (U), can be confidently assigned.
Figure 65. The NOE difference spectra of resonances I'2, I, and I'' of Fragment B. Though it is difficult to tell NOE connectivities between the three resonances from these spectra, it is evident that I'' (G·C) is connected to A2 (A·U), as can be seen here. NOE connectivities observed in Figures 63-65 can be sufficient evidence to establish complete helical segment of the prokaryotic loop.
Table 7. Identification and assignment of base pair imino proton resonances in Fragment B by means of NOE connectivities.

<table>
<thead>
<tr>
<th>Peak&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Chemical Shift (ppm)</th>
<th>Base Pair</th>
<th>NOE connectivity to</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>11.13</td>
<td>unpaired $U_{85}$</td>
<td>$O$</td>
</tr>
<tr>
<td>$O$</td>
<td>10.10</td>
<td>unpaired $U_{86}$</td>
<td>$M_1$, $M_2$</td>
</tr>
<tr>
<td>$M_2$</td>
<td>11.28</td>
<td>unpaired $U_{87}$</td>
<td>$O$, $D_2$</td>
</tr>
<tr>
<td>$D_2$</td>
<td>13.39</td>
<td>$C_{84}$,$G_{88}$</td>
<td>$M_2$, $A_2$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>14.10</td>
<td>$A_{83}$,$U_{89}$</td>
<td>$D_2$, $I''$</td>
</tr>
<tr>
<td>$I''$</td>
<td>12.44</td>
<td>$G_{82}$,$C_{90}$</td>
<td>$A_2$, spillover</td>
</tr>
<tr>
<td>$I$</td>
<td>12.51</td>
<td>$G_{81}$,$C_{91}$</td>
<td>spillover</td>
</tr>
<tr>
<td>$I'_2$</td>
<td>12.60</td>
<td>$G_{80}$,$C_{92}$</td>
<td>spillover</td>
</tr>
</tbody>
</table>

<sup>a</sup> See Figures 63-65.
Although prior studies of the base-pairing pattern in the prokaryotic loop of *E. coli* achieved much in the identification and assignment of the base pairs in that region (Kime & Moore, 1983a), ambiguity still remained, and three unpaired imino protons in the loop were never observed. The assignments of \((G\cdot C)_3(A\cdot U)(G\cdot C)\) and three U's presented here provide the first complete proton NMR description of the secondary structure in this helical region of any 5S rRNA and its fragments. In particular, the G•C base pair adjacent to the three unpaired pyrimidines (G\textsubscript{86}•C\textsubscript{90} in *E. coli* 5S rRNA) which could not be identified in *E. coli* 5S rRNA is observed in *B. megaterium* 5S rRNA (C\textsubscript{84}•G\textsubscript{88}). The putative G•U pair (U\textsubscript{78}•G\textsubscript{94}) is not observed in this fragment, probably because of too-fast exchange of N1-H of G\textsubscript{94} with water. N3-H of U\textsubscript{85} seems to be spatially farther than N3-H of U\textsubscript{87} from N1-H of G\textsubscript{88} in the loop, because NOE connectivities are observed between resonances D\textsubscript{2} (N1-H of G\textsubscript{88}) and M\textsubscript{2} (N3-H of U\textsubscript{87}), but not between D\textsubscript{2} and M\textsubscript{1} (N3-H of U\textsubscript{85}) (Figure 63). The G\textsubscript{80}•C\textsubscript{92} - G\textsubscript{81}•C\textsubscript{91} - G\textsubscript{82}•C\textsubscript{90} - A\textsubscript{83}•U\textsubscript{89} - C\textsubscript{84}•G\textsubscript{88} segment in helix IV is so stable that even at 40 °C the resonances I'\textsubscript{2}, I, I'', A\textsubscript{2}, and D\textsubscript{2} show strong signals in the proton NMR spectrum (Figure 66). As can be seen in Figure 66, the resonances melt, however, in the order, I'\textsubscript{2} (G\textsubscript{80}•C\textsubscript{92}), I (G\textsubscript{81}•C\textsubscript{91}), I'' (G\textsubscript{82}•C\textsubscript{90}) and D\textsubscript{2} (C\textsubscript{84}•G\textsubscript{88}). This melting pattern is a good indication that the unwinding of helix IV begins with the breakage of the base pair G\textsubscript{80}•C\textsubscript{92} at slightly elevated temperature. As temperature increases, base pairs melt in succession, ending with the base pair adjacent to the loop. The imino protons of U\textsubscript{85}, U\textsubscript{86}, and U\textsubscript{87} in the loop are also well protected from rapid exchange with water. Overall, the prokaryotic loop is very stable.
Figure 66. Variable temperature experiments of Fragment B in 0.01 M cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with a 500-MHz proton NMR spectrometer. As can be seen in the figure, the resonances melt in the order $I_2'$ ($G_{80} \cdot C_{92}$), $I$ ($G_{81} \cdot C_{91}$), $I''$ ($G_{82} \cdot C_{90}$) and $D_2$ ($G_{84} \cdot G_{88}$). This is a good indication that the unwinding of helix IV begins with the breakage of the base pair $G_{80} \cdot C_{92}$. 
However, unlike Fragment A, Fragment C, or intact 5S rRNA, the melting caused loss of the original conformation of Fragment B. After heating up to 45–50 °C and cooling to room temperature, Fragment B seems to form different base pairs in view of the different chemical shifts of the imino proton resonances. It is not easy to analyze the change of the conformation of Fragment B at this point due to its impure nature compared to that of Fragment A. At any rate, the irreversibility of the melting in this fragment suggests that the natural conformation of helix IV may not be the most stable one, even though helix IV can be very stable together with the rest of 5S rRNA.

B. Base-Pairings in Helix I of *B. megaterium* 5S rRNA (NMR Analysis of Fragment C)

Production and purification of RNase T1-cleaved Fragment C has been discussed, together with other fragments, in the previous chapter. Fragment C is believed to consist of mainly the terminal stem (helix I) of *B. megaterium* 5S rRNA on the basis of RNA T1 mapping experiment. Since helix I is the longest of the five helices, a complicated downfield $^1$H NMR spectrum is expected due to many (at least ten) base-paired imino proton resonances. In this section, Fragment C will be analyzed by $^1$H NMR spectroscopy to confirm the $\text{G}_4\cdot\text{C}_{112} - \text{G}_5\cdot\text{C}_{111} - \text{U}_6\cdot\text{G}_{110}$ segment previously assigned from the $^1$H NMR study of *B. megaterium* 5S rRNA.

1. NMR Spectroscopy

The NMR sample of RNase T1-cleaved Fragment C of *B. megaterium* 5S rRNA was prepared as described for the NMR sample of Fragment A.
The final concentration of Fragment C in the NMR sample was 140 A$_{260}$ units in 0.5 mL of 1 mM EDTA, 100 mM NaCl, 10 mM cacodylic acid, pH 7.0. The NMR sample showed at least 95% purity as estimated from electrophoresis gel scan obtained under non-denaturing condition.

All $^1$H NMR spectra including NOE spectra were obtained with a Bruker AM-500 FT/NMR spectrometer in the same manner as previously explained.

2. Base Pairs in Fragment C (Helix I)

The downfield proton NMR spectrum of RNase T1-cleaved Fragment C (mainly the helix I region) of B. megaterium 5S rRNA shows more than a dozen peaks arising from many imino proton resonances in the terminal stem (Figure 67). The sequence of base pairs G$_4$·C$_{112}$ – G$_5$·C$_{111}$ – U$_6$·G$_{110}$, previously inferred from NOE difference spectra of intact 5S rRNA, can now be confirmed with greater confidence from NOE difference spectra obtained by irradiating resonances J, E, K$_1$, and M in the more highly resolved spectrum of Fragment C (Figure 68), which gave the same NOE connectivities as observed in intact 5S rRNA. Peaks K$_1$ and K$_2$, which could not be resolved in the spectrum of intact 5S rRNA, gave two separate signals (very close in chemical shift) in a spectrum of Fragment C computed from a 32K time-domain data set. Hence, K$_1$ can be assigned as the N3-H resonance of the U$_6$ paired with G$_{110}$.

Figure 69 shows the melting pattern of the imino proton resonances arising from the base pairs in Fragment C according to temperature change. As was expected, it is evident that resonances E, J, K$_1$, M (G$_4$·C$_{112}$ – G$_5$·C$_{111}$ – U$_6$·G$_{110}$) are the most stable resonances of all. As
Figure 67. The downfield proton NMR spectrum of Fragment C at 23 °C in 10 mM cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0, and 95 %:5 % H₂O/D₂O.
Figure 68. The NOE difference spectra of resonances E, J, K₁, and M of Fragment C. The sequence of base pairs G₄·C₁₁₂ – G₅·C₁₁₁ – U₆·G₁₁₀, previously inferred from NOE difference spectra of intact 5S rRNA, can now be confirmed with greater confidence from these more highly resolved NOE difference spectra of Fragment C.
Figure 69. Variable temperature experiments of Fragment C in 0.01 M cacodylic acid, 0.1 M NaCl, 1 mM EDTA, pH 7.0 with a 500-MHz proton NMR spectrometer. As can be seen here, it is evident that resonances E, J, K₁, M (G₄·C₁₁₂ − G₅·C₁₁₁ − U₆·G₁₁₀) are the most stable resonances of all presumably due to their locations (middle of the helix).
can be seen from the figure, E, J, K₁, M are the only four resonances that show significant intensities at 45 °C, implying that the base pair segment G₄₋C₁₁₂ – G₅₋C₁₁₁ – U₆₋G₁₁₀ is the most stable part in helix I due to its location (middle of the helix). Melting of Fragment C, like that of intact 5S rRNA, was reversible, suggesting thermodynamic stability of the terminal helix. The melting pattern observed in Fragment C is exactly the same as that observed in intact 5S rRNA.

C. Conformations of Fragments B and C and the Corresponding Segments of Intact 5S rRNA

As discussed in the previous chapter, whether or not the Fragments B and C retain the original conformation of the intact 5S rRNA molecule is an important question. This retention of the original conformation in the fragments can be demonstrated by comparing the ¹H NMR spectrum of 5S rRNA with the ¹H NMR spectra of the fragments. In addition, comparison of the NOE difference spectra between them can verify that question. As can be seen from Figures 38, 62, and 67, virtually all of the imino proton resonances observed in the spectra of Fragments B and C can be found at the same chemical shifts in intact 5S rRNA. Moreover, the NOE difference spectra obtained by irradiating resonances K₁/M, E, J of Fragment C (Figure 68) and those of intact 5S rRNA (Figure 39) are identical. Furthermore, irradiation of peak A of intact 5S rRNA shows NOE connectivities to resonances I′, F₁, D₂, which are observed separately by irradiation of peak A₁ in Fragment A (NOE connectivity to F₁ and I′) and peak A₂ in Fragment B (NOE connectivity to D₂ and I′), as seen in Figures 49 and 63. The results shown above are probably
sufficient to prove the retention of conformation upon cleavage of 5S rRNA into fragments.

**D. Base Pairs in Other Structural Segments**

The existence of helix V in 5S rRNA has been demonstrated from assignment of base pairs in that region for both prokaryotic (Leontis et al., 1986; Chang & Marshall, 1986a; Zhang & Moore, 1989) and eukaryotic (Li & Marshall, 1986; Chen & Marshall, 1986) 5S rRNAs. Unfortunately, it is not possible to identify helix V base pairs from *B. megaterium* 5S rRNA and its fragments here. Helix V is known to be the least stable among all helical regions in 5S rRNA, whereas helices I and IV are the most stable (Chang & Marshall, 1986b). The existence of loop E (the internal loop between helices IV and V) has been confirmed in chloroplast 5S rRNA from spinach (Romby et al., 1988) and "fragment 1" of *E. coli* 5S rRNA (Zhang & Moore, 1989). Zhang and Moore have reported that an extensively base-paired model for loop E is incompatible with the NMR data obtained, suggesting that many of the bases in the loop are only slightly protected from solvent exchange, and that loop E may well not contain any conventional base pairs, thereby resulting in a shortened helix V segment. In *B. megaterium* 5S rRNA, helix V not only is relatively short, but also contains a noncanonical base pair (A102•R70) in the middle of the helix, in addition to the weak G100•U72 base pair (see Figure 17), resulting in a relatively unstable helix V which evidently does not survive the enzymatic cleavage process. Comparison of the spectrum of intact 5S rRNA with the spectrum obtained by the addition of the spectra of Fragments A, B, and C (Figure 70) reveals some peaks which
Figure 70. 500 MHz $^1$H NMR spectrum of intact 5S rRNA (bottom) and composite spectrum obtained by the addition of the spectra of Fragments A, B, and C (top). The close match in chemical shifts between the two spectra (taking into account the broader lines for the larger intact 5S rRNA molecule) offers further evidence for high conformational similarity between intact 5S rRNA and its fragments. Some peaks present in intact 5S rRNA are absent in the combined spectrum (e.g., peak C). The missing peaks in the combined spectrum are believed to result from secondary base-pairs in helix V and from tertiary base-pairs.
are present in intact 5S rRNA but not in those fragments. For example, peak C is clearly absent in the fragments. The missing resonances not observed in the fragments probably arise from base-paired imino protons in helix V and/or tertiary base pairs (Pieler & Erdmann, 1982; Chang & Marshall, 1986b). As one might expect for helix V or tertiary base pairs, these resonances disappear quickly as temperature increases (see Figure 40), leaving the more stable base-pairs assigned above.

**E. Conclusion**

So far, the complete secondary structure of the prokaryotic loop (helix IV) of *B. megaterium* ribosomal 5S rRNA, G\textsubscript{80}·C\textsubscript{92} – G\textsubscript{81}·C\textsubscript{91} – G\textsubscript{82}·C\textsubscript{90} – A\textsubscript{83}·U\textsubscript{89} – C\textsubscript{84}·G\textsubscript{88}, U\textsubscript{85}, U\textsubscript{86}, and U\textsubscript{87}, has been described. Also, U\textsubscript{32}·A\textsubscript{46} – G\textsubscript{31}·C\textsubscript{47} – C\textsubscript{30}·G\textsubscript{48} – C\textsubscript{29}·G\textsubscript{49} in helix III, G\textsubscript{21}·C\textsubscript{58} – A\textsubscript{20}·U\textsubscript{59} – G\textsubscript{19}·C\textsubscript{60} – A\textsubscript{18}·U\textsubscript{61} in helix II, A\textsubscript{27}·U\textsubscript{53} – C\textsubscript{28}·G\textsubscript{52} in the inter-helical region between helices II and III, and G\textsubscript{4}·C\textsubscript{111} – G\textsubscript{5}·C\textsubscript{111} – U\textsubscript{6}·G\textsubscript{110} in helix I have been assigned. Although the assignments of base pairs in helices I, II, and III are not complete, the present NOE-based conclusions strongly support the *B. megaterium* ribosomal 5S rRNA secondary structure adapted from the Fox and Woese model (Figure 17), offering additional evidence for the universality of the secondary structure of all prokaryotic and eukaryotic 5S rRNAs.
CHAPTER VI
MOLECULAR MODELING OF RNASE T1-CLEAVED FRAGMENT A AND FRAGMENT B OF B. MEGATERIUM RIBOSOMAL 5S RNA

Molecular dynamics (MD) simulations have become popular recently for the structural studies of nucleic acids (Woodson & Crothers, 1988; Baleja et al., 1990; Rao & Kollman, 1990), since these methods have the potential capability of simulating the structure and dynamics of nucleic acids at the atomic level. From the structure of a molecule constructed by dynamics calculations, NOE cross peak intensities have been predicted (Keeper & James, 1984; Suzuki et al., 1986; Gupta et al., 1988 Boelens et al., 1989). Most of these studies were performed with small DNA fragments. Only a few RNA fragments have been used for molecular mechanics or molecular dynamics studies. However, molecular dynamics simulations have potential for the structural investigation of RNA molecules, as the methods have proven feasible for DNA fragments.

So far the structure of B. megaterium 5S rRNA has been investigated by proton NMR spectroscopy. In this chapter, molecular dynamics simulations will be applied to test the structural features of the 5S rRNA molecule thus obtained by NMR and to achieve deeper understanding of
the molecule. RNase T1-cleaved Fragments A and B of *B. megaterium* 5S rRNA have been built, and their structures have been constructed by molecular dynamics.

**A. Basic Theory of Molecular Dynamics**

**1. Force Fields**

A molecule consists of a number of atoms held together by elastic or harmonic forces. The energy of an arbitrary geometry for the molecule is usually expressed as a superposition of various interactions. The forces on all particles of the molecule can be evaluated rapidly from such interactions. Based on the forces, the optimum geometry (by conjugate minimization, for example) or the dynamics of the motion (Newton's equation) can be obtained. These methods are called molecular mechanics or molecular dynamics.

The forces are described by potential energy functions of structural features (e.g., bond lengths, bond angles, nonbonded interactions, etc.). From the combination of these potential energy functions, one derives the so-called *force field* (Boyd & Lipkowitz, 1982). The energy, *E*, of the molecule in the force field arises from deviations from ideal structures. This energy can be approximated by the sum of various energy contributions, as in the following equation:

$$ E = E_s + E_b + E_w + E_{nb} + ... $$  \[14\]

where *E* (the so-called steric energy) is the difference in energy between an ideal and the real molecule. *E_s*, *E_b*, *E_w*, and *E_{nb}* stand for energy contributions from bond stretching, bond angle bending, dihedral angle torsion, and nonbonded interactions.
Since there are no exact rules about the types and the number of potential functions in the total energy \( E \), many different force fields have been developed. Some of the most popular ones include CHARMM (Nilsson & Karplus, 1986), AMBER (Weiner et al., 1984: Weiner et al., 1986), MM2 (Sprague et al., 1987), and MM3 (Allinger & Liu, 1987). BIOGRAF, which has been used for Fragments A and B of \( B. \ megaterium \) 5S rRNA here, incorporates several standard force fields to form a general energy expression as in the following equation:

\[
E = E_{\text{bonded}} + E_{\text{nonbonded}} + E_c + E_u
\]

where \( E_{\text{bonded}} = E_s + E_b + E_w + E_t \), \( E_{\text{nonbonded}} = E_{\text{vdw}} + E_{\text{el}} + E_{\text{hb}} \), and \( E_t \), \( E_{\text{vdw}}, E_{\text{el}}, E_{\text{hb}}, E_c \), and \( E_u \) stand for energy contributions from inversion, van der Waals interaction, electrostatic interaction, hydrogen bond, optional constraint added, and optional potential added by the user, respectively.

2. Molecular Dynamics (MD)

With the recent increase in understanding of the dynamic behavior of nucleic acids benefiting from a number of spectroscopic techniques together with X-ray and neutron diffraction crystallography, theoretical studies like molecular dynamics simulations have become of great interest. In molecular dynamics, the change of the acceleration of the atoms in a molecular structure with respect to time \( t \) is described by Newton's equations (Newton's second law of motion).

\[
F_{\text{sj}} = M_s a = - \frac{\delta E}{\delta R_{\text{sj}}} = M_s \left( \frac{\delta^2 R}{\delta t^2} \right)
\]

or

\[
\frac{\delta v_{\text{sj}}}{\delta t} = \frac{1}{M_s} F_{\text{sj}} = - \frac{1}{M_s} \frac{\delta E}{\delta R_{\text{sj}}}
\]
in which $E$ is the total energy, and the subscript $j$ refers to $x$, $y$, or $z$, the components of the position vector for atom $s$. $F_{sj}$ refers to the $j$-th component of the force, $v_{sj}$ to the $j$-th component of the velocity of atom $s$, $R_{sj}$ to the magnitude of the $j$-th component of the position vector (Cartesian coordinates), $M_s$ to the mass, and $\alpha$ to acceleration.

The starting geometry (or conformation) is defined by the $3N$ values of $R_{sj}$, where $N$ is the total number of atoms, and each atom position vector has an $x$, $y$, and $z$ component. From the starting geometry, the computer program (BIOGRAF) calculates the $3N$ forces given by $\frac{\delta E}{\delta R_{sj}}$, and solves the equation for each degree of freedom for all atoms. Based on these changes in velocity and position at regular time increments for every atom in the molecular structure, the change in the geometry of the structure with time is determined by the computer. This is the basic procedure of molecular dynamics.

For a system of $N$ atoms in equilibrium at temperature $T$, the velocities of the atoms satisfy the relationship

$$\sum (1/2) M_s <v^2_{sj}> = (3/2) N k T,$$  \[18\]

where $T$ is the temperature, $k$ the Boltzmann constant, $N$ the number of moles, and the brackets represent an average. With the start of a calculation at a specified temperature $T_i$, the computer program (BIOGRAF) generates an initial velocity for each atom such that the overall distribution is Maxwell-Boltzmann at temperature $T_i$. As the dynamics proceeds, velocity fluctuates and the instantaneous values for $\sum (M_s/3k) v^2_{sj}$ differ from $T_i$. Since the total energy remains the same, the value of $T$ in the equation is high at a low energy (a favorable conformation), and the value of $T$ is low when the system is at a high
energy (an unfavorable conformation). These instantaneous values of $T$ are not the true temperature. The true temperature is defined as an average over time. The computer program (BIOGRAF) calculates the temperature by averaging the equation over a number of steps in the dynamics calculations.

B. Procedures for Molecular Dynamics Simulations

Molecular dynamics simulations of nucleic acids have been performed under three different conditions: nucleic acids without any explicit inclusion of environment (Levitt, 1983; Singh et al., 1985; Rao et al., 1985), nucleic acids in the solvent water (Seibel et al., 1985), and nucleic acids without any explicit inclusion of solvent but with constraints obtained from NMR studies (Hare et al., 1986; Nilges et al., 1987a; Nilges et al., 1987b). In this research, molecular dynamics of RNase T1-cleaved Fragments A and B of B. megaterium 5S rRNA was performed without any inclusion of solvent molecules. BIOGRAF was used on a VAX 11/750 computer with an Evans & Sutherland PS 300 monitor.

1. Building of RNA Molecules Using BIOGRAF

Since building an RNA molecule directly by BIOGRAF is not possible, the corresponding A-DNA fragment was first built, with subsequent modifications of the DNA molecule into an A-RNA fragment. The modifications were as follows. Thymines (T) in the fragment were converted to uracils (U) by deleting the 5'-methyl group from each thymine. Deoxyribose sugar rings were changed into ribose sugar rings
by replacing 2"-H by 2"-OH. The single chain of the molecule (for practical purposes, a single-stranded DNA instead of a double-stranded DNA was initially built) was divided into two fragments by breaking a phosphodiester bond, and one of the resulting fragments was moved to a best spatial location by a "docking" operation. In this way, helical regions and hairpin loops could be created. Since RNase T1-cleaved fragments have 5'-end OH and 3'-end phosphate, a phosphate group was added to the 3'-side OH of the molecule. Finally, atom types of some atoms were readjusted in such ways that the atoms could be involved in hydrogen bonds.

2. Minimization

Before subjected to a dynamics simulation, the energy of the built molecule needs to be minimized to find a favorable conformation in the local region of the potential energy surface. In this way, considerable computer time can be saved before the molecule approaches the global minimum on the potential energy surface. The minimization was performed as a batch job on the computer with BIOGRAF (conjugate gradient minimization).

In order to reach a local energy minimum quickly, the minimization was executed with a constraint file. The constraint file consisted of a number of distance constraints obtained from crystal structure of base pairs (Saenger, 1984). Each A•U base pair had two distance constraints: 2.95 Å between N6 of A and O4 of U, and 2.82 Å between N1 of A and N3 of U. Each G•C base pair had three distant constraints: 2.91 Å between O6 of G and N4 of C, 2.95 Å between N1 of G and N3 of C, and 2.86 Å
between N2 of G and O2 of C. The base pairs in the RNase T1-cleaved fragments of *B. megaterium* 5S rRNA were chosen for constraints based on the common secondary structure of *B. megaterium* 5S rRNA and NMR data. Constraint files used for Fragment A and Fragment B are shown in Table 8 and Table 9, respectively.

3. Molecular Dynamics

The minimization process can find a favorable conformation only in a local minimum. Since macromolecules like RNA have a number of atoms with a great number of coordinates, they can have many favorable conformations, with an astronomical number of possible conformations. If the energy of each conformation is to be calculated one by one, it will take infinite time for a computer to find the optimal conformation. Fortunately, molecular dynamics has the capability of efficiently searching for low energy geometries of a system without going through every one of its conformations.

Four types of dynamics programs are available from BIOGRAF: straight dynamics, quenched dynamics, annealed dynamics, and impulse dynamics. In annealed dynamics, the temperature starts at $T_i$ (initial temperature) and increments by an amount $\Delta T$ (heat increment) at intervals of 0.100 ps until it reaches $T_f$ (final temperature). The temperature then increments back to $T_i$. Annealed dynamics repeats this procedure, back and forth between $T_i$ and $T_f$. When the low temperature is small (~0 °K) and the number of steps is large, the system may cool slowly enough to find the best minimum in the potential energy surface without getting trapped in a high energy local minimum. Annealed
Table 8. Distance constraints used for dynamics calculations of Fragment A.

<table>
<thead>
<tr>
<th>Base #</th>
<th>Atom</th>
<th>Base #</th>
<th>Atom</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A17</td>
<td>N6</td>
<td>U62</td>
<td>O4</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.82</td>
</tr>
<tr>
<td>A18</td>
<td>N6</td>
<td>U61</td>
<td>O4</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.82</td>
</tr>
<tr>
<td>G19</td>
<td>O6</td>
<td>C60</td>
<td>N4</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td></td>
<td>O2</td>
<td>2.86</td>
</tr>
<tr>
<td>A20</td>
<td>N6</td>
<td>U59</td>
<td>O4</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.82</td>
</tr>
<tr>
<td>G21</td>
<td>O6</td>
<td>C58</td>
<td>N4</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td></td>
<td>O2</td>
<td>2.86</td>
</tr>
<tr>
<td>A27</td>
<td>N6</td>
<td>U53</td>
<td>O4</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.82</td>
</tr>
<tr>
<td>C28</td>
<td>N4</td>
<td>G52</td>
<td>O6</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N3</td>
<td></td>
<td>N1</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td></td>
<td>N2</td>
<td>2.86</td>
</tr>
<tr>
<td>C29</td>
<td>N4</td>
<td>G49</td>
<td>O6</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N3</td>
<td></td>
<td>N1</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td></td>
<td>N2</td>
<td>2.86</td>
</tr>
<tr>
<td>C30</td>
<td>N4</td>
<td>G48</td>
<td>O6</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N3</td>
<td></td>
<td>N1</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td></td>
<td>N2</td>
<td>2.86</td>
</tr>
<tr>
<td>G31</td>
<td>O6</td>
<td>C47</td>
<td>N4</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td></td>
<td>O2</td>
<td>2.86</td>
</tr>
<tr>
<td>U32</td>
<td>O4</td>
<td>A46</td>
<td>N6</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N3</td>
<td></td>
<td>N1</td>
<td>2.82</td>
</tr>
</tbody>
</table>
Table 9. Distance constraints used for dynamics calculations of Fragment B.

<table>
<thead>
<tr>
<th>Base #</th>
<th>Atom</th>
<th>Base #</th>
<th>Atom</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G80</td>
<td>O6</td>
<td>C92</td>
<td>N4</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td></td>
<td>O2</td>
<td>2.86</td>
</tr>
<tr>
<td>G81</td>
<td>O6</td>
<td>C91</td>
<td>N4</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td></td>
<td>O2</td>
<td>2.86</td>
</tr>
<tr>
<td>G82</td>
<td>O6</td>
<td>C90</td>
<td>N4</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td></td>
<td>O2</td>
<td>2.86</td>
</tr>
<tr>
<td>A83</td>
<td>N6</td>
<td>U89</td>
<td>O4</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td></td>
<td>N3</td>
<td>2.82</td>
</tr>
<tr>
<td>C84</td>
<td>N4</td>
<td>G88</td>
<td>O6</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>N3</td>
<td></td>
<td>N1</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td></td>
<td>N2</td>
<td>2.86</td>
</tr>
</tbody>
</table>
dynamics were employed here for the structural studies of *B. megaterium* 5S rRNA.

4. **Parameters for Molecular Dynamics**

There are several parameters to be considered for annealed dynamics calculations. The value shown for each parameter is the one used for the dynamics calculations of Fragment A.

a. **Dynamics Variables**

- **Draw frequency (= 1):** The frequency at which the structure is updated on the monitor. The value of 1 means that the structure was redrawn for every step of dynamics.
- **Time step (= 0.001 ps):** The interval of time (Δt) between velocity and the coordinate is readjusted (in ps). This is the actual length of time step for every dynamics calculation. This should be much smaller than the time required for the fastest local motion in the molecule.
- **Annealed frequency (= 0.01 ps):** The interval between every annealing temperature readjustment. The number of steps of dynamics (= annealed frequency/time step) was 10.
- **Heat equilibration frequency (= 0.05 ps):** The interval between heating or cooling adjustment.
- **Write trajectory frequency (= 0.1 ps):** The interval at which the current coordinates are written on the trajectory file.

b. **Temperature Variables**

- **Initial temperature (= 1000 °K):** Initial temperature of annealed dynamics (T₀).
Final temperature (= 0 °K): Final temperature of annealed dynamics ($T_f$).

Heat increment (= 50 °K): Temperature increment ($\Delta T$).

Temperature window (= 50 °K): During each dynamics run, temperature of the system fluctuates. If temperature deviates more than the amount specified by the temperature window (either below or above), velocities of atoms are rescaled to force the temperature to fall within this permissible temperature range.

c. Running Annealed Dynamics

From the initial temperature, velocity and coordinates of each atom are assigned. Dynamics is run during time step ($\Delta t$). New velocity and new coordinates will be obtained. This process is repeated for the number of times given by (annealed frequency/time step). After a given cycle is over at the initial temperature, the temperature is raised by $\Delta T$. At this temperature, the same number of cycles is repeated. This procedure continues until the final temperature is reached. Then the temperature increment is $-\Delta T$, and the cycle of dynamics continues until it returns to the initial temperature. The total period described so far is one annealed cycle. The dynamics continues in the same manner, going back and forth between the initial and the final temperatures for the number of times specified by the annealed cycles. Annealed dynamics for Fragment A were run for several hundred annealed cycles to reach a relatively stable conformation of the molecule. From time to time between dynamics runs, minimization can be helpful.
C. Structure of the Helices II-III Region (Fragment A)

When the helices II-III region of B. megaterium 5S rRNA was built by BIOGRAF based on A-helix structure, refinement of the helical regions showed promising results with successive dynamics runs. However, the structure in the hairpin loop (loop C) was far from a stable conformation, showing no base-stacking at all. That tight conformation in the loop region seemed to be in a deep local energy minimum, such that the problem could not be alleviated at all with dozens of dynamics runs with modified parameters (hundreds of annealed cycles). Later, the loop area of the structure was deleted, and the remaining part (helices II, III, and the inter-helical region) was combined with a new loop borrowed from crystal tRNA structure. The anticodon loop of tRNA crystal structure (available from Brookhaven Data Bank) was cut off, and the bases in the anticodon loop were replaced by bases of loop C of B. megaterium 5S rRNA. It is believed that the starting geometries can be very important for obtaining a nice structure from the molecular dynamics simulations, especially when a large macromolecule is involved. Since the newly built RNA structure is closer to the global minimum in the potential energy surface, molecular dynamics is beneficial to the refinement of this structure.

When the energy of the molecule was minimized with a distance constraint file by conjugate gradient, it showed an improvement over the first structure built (Figure 71). Base stackings in the loop area were especially noticeable. Besides these, the helix II region showed a regular A-helix. However, bases in helix III were unstacked because the borrowed loop was combined with the helical regions through helix III.
Figure 71. Stereoscopic view of the helices II-III region of *B. megaterium* 5S rRNA after the structure is minimized. Hydrogens are omitted. Bases in hairpin loop area (lowest part of the structure) are well-stacked. However, bases in helix III are not base-stacked.
Annealed dynamics was run for more than a hundred annealed cycles, and the structure began to converge (Figure 72). After convergence was observed, the molecule was subjected to minimization, and the structure in Figures 73 and 74 was obtained. First, the helix III region has turned into a smooth helical region with well-stacked bases. The distances between neighboring base-paired imino protons in helices II and III were measured from the structure (Table 10). As can be seen from the table, the distances between the neighboring imino protons are between 3–4 Å, therefore observable NOEs are predicted between these protons in the 1H NMR spectrum. Indeed all NOEs between these imino protons were observed in NMR experiments previously discussed, showing the compatibility of the constructed structure by molecular dynamics and the solution structure observed by NMR.

The anticipated bulge (A50 and A51) does not show sufficient bulging from the helices. While strong mutual NOEs were observed between resonances A27•U53 and C28•G52 in NMR, the distance between imino protons of U53 and G52 obtained from dynamics is 5.36 Å, which is beyond the normal NOE observation range. In addition, the distance between imino protons of G52 and G49 is 7.05 Å, whereas NOEs between the two imino protons were observed in NMR. These facts are definitely caused by lack of bulging from the helix in the structure obtained by dynamics, even though A27•U53 and C28•G52 are base-paired. This structure is assumed to be in another deep local minimum in the potential energy surface. Since the bulge is not protruding enough, the tilt angles of bases in helix II are greater than the tilt angles of bases on helix III. The overall helix axis is warped rather than straight.
Figure 72. Superposition of structures (the helices II-III region) from a number of trajectory files during dynamics run. The helix III region experiences the most significant change as shown.
Figure 73. Stereoscopic view of the helices II-III region of *B. megaterium* 5S rRNA after dynamics run. Compared to the structure before dynamics (Figure 71), bases in the helix II region are stacked well in the shape of A-helix.
Figure 74. All-atom structure (all hydrogens included) of the helix II-III region of *B. megaterium* 5S rRNA.
Table 10. Distances between neighboring base-paired imino protons in the helices II–III region of *B. megaterium* 5S rRNA. Calculated by molecular dynamics.

<table>
<thead>
<tr>
<th>Imino proton from</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_{17}·U_{62}</td>
<td>4.26</td>
</tr>
<tr>
<td>A_{18}·U_{61}</td>
<td>3.78</td>
</tr>
<tr>
<td>G_{19}·C_{60}</td>
<td>3.92</td>
</tr>
<tr>
<td>A_{20}·U_{59}</td>
<td>4.35</td>
</tr>
<tr>
<td>G_{21}·C_{58}</td>
<td></td>
</tr>
<tr>
<td>A_{27}·U_{53}</td>
<td>5.36</td>
</tr>
<tr>
<td>C_{28}·G_{52}</td>
<td>7.05</td>
</tr>
<tr>
<td>C_{29}·G_{49}</td>
<td>3.54</td>
</tr>
<tr>
<td>C_{30}·G_{48}</td>
<td>3.89</td>
</tr>
<tr>
<td>G_{31}·C_{47}</td>
<td>3.42</td>
</tr>
<tr>
<td>U_{32}·A_{46}</td>
<td></td>
</tr>
</tbody>
</table>
The formation of a C•U base pair has been proposed in helix III of wheat germ 5S rRNA (Wu & Marshall, 1990a,b). The helices II-III region of *B. megaterium* 5S rRNA also shows the possibility of a C•U base pair formation, since bases in U33 and C45 are stacked parallel to the other bases of the helix III region with similar tilt angles. However, the distance between N4 of C45 and O4 of U33 (or N3 of C45 and N3 of U33) is more than 5 Å, which is far greater than the ~3 Å distance necessary for base pair formation. Therefore, a C•U base pair formation is a possibility, but difficult to ascertain from the structure obtained from dynamics at this point.

Extra base pair formation in the hairpin loop (loop C) does not seem to be possible from the geometries of atoms calculated from dynamics. The same is true for the inter-helical loop region (loop B). Some bases in loop B are slightly out of stack, and A25 and A56 are spatially very close. G16•U63 base pair formation seems to be difficult due to highly distorted base planes of the two bases. This distortion influences somewhat the A17•U62 base pair such that the distance between imino protons of U62 and U61 is unusually large, 4.26 Å. Since the imino protons from these two base pairs could not be observed from Fragment A in the NMR spectrum, probably due to their locations near the open terminal, no comparison can be made between the results obtained by the two methods.

Dynamics runs of the molecule without constraints were tried in a couple of different ways. After the dynamics run with all distance constraints between bases including A27•U53 and C28•G52, the dynamics was continued without any constraint. What occurred was a structural
change in the helix III region leading to the disruption of base pairs in that helical region, with the preservation of base pairs in helix II, \( A_{27} \cdot U_{53}, \) and \( C_{28} \cdot G_{52} \) (Figure 75). When dynamics without constraints was started with an initial structure of a small difference, the helices II and III remained intact while the bulge was destroyed. Since the disruption of helix III and the bulge never happened together in any dynamics run, bulge migration does not seem to be possible, but rather the structure prefers one conformation as discussed in Chapter IV.

At any rate, helix II is considered to be more stable than helix III based on these results, which is consistent with the NMR results previously discussed. However, helix III was usually very stable together with helix II without constraints in most cases. The problem was a strong steric energy caused by the formation of \( A_{27} \cdot U_{53} \) and \( C_{28} \cdot G_{52} \) without bulging out of the presumed bulge \( (A_{50} \) and \( A_{51}). \)

**D. Structure of the Helix IV Region (Fragment B)**

The helix IV segment spanning \( U_{78} \) to \( G_{94} \) of *B. megaterium* 5S rRNA was built by BIOGRAF in the same way as the helices II-III segment. The energy of the system was minimized by the conjugate gradient method initially with 14 distance constraints (Table 9). Then all hydrogens were added to the structure, and minimization continued without any constraint. After minimization was over, five base pairs in the helix were preserved in spite of the absence of distant constraints, suggesting the stability of the helix (Figure 76). In fact, the five base pairs constructed a short segment of A-helix.
Figure 75. Susceptibility of the helix III region to structural change in the absence of constraints. The helix II region is very stable.
Figure 76. All-atom stereoscopic view of the helix IV region of *B. megaterium* 5S rRNA after the structure is minimized. Initially 14 constraints were included for a short time, then minimization was continued without constraints. Five base pairs in the helix are well-preserved, suggesting the stability of the helix. U₈₇ is sticking out of the loop.
Annealed dynamics was run with the initial temperature of 600 °K and the final temperature of 0 °K. The all-atom model was used throughout. Within fifty annealed cycles, no virtual decrease of the total energy in the system was observed, although occasional energy fluctuation occurred. As can be seen in Figure 77, no significant change in any part of the structure could be noted during the dynamics run. The structure of the molecule after the dynamics run is shown in Figure 78.

Distances between neighboring imino protons are measured on the basis of the structure calculated by dynamics, and summarized in Table 11. The calculated distances agree with the results obtained from NMR experiments. Stability of the five base pairs is confirmed by the preservation of the region without any constraint. The helical region constructed from the dynamics calculations, therefore, seems to be quite consistent with the solution structure of the molecule. In addition to the five base pairs, the distance between U79 and G94 shows a spatial proximity with parallel base planes, suggesting a possible base pair formation. However, the distance is still larger than that of a usual base pair.

One unexpected feature in the structure is that U87 is protruding from the loop, thus reducing the crowding in the loop. Hence, the distances between N3-H of U87 and imino protons from the neighboring bases (G88, U85, and U86) are unusually large, such that NOEs between them are virtually impossible, in contradiction to the observed NOE connectivities between those imino protons in the NMR experiments previously conducted. The structure based on dynamics is a
Figure 77. Superposition of structures (helix IV) from a number of trajectory files during dynamics run. Convergence to more stable conformation is monitored.
Figure 78. All-atom stereoscopic view of helix IV region of *B. megaterium* 5S rRNA after dynamics run. No significant change from the previous view can be found, suggesting a very stable conformation of the structure.
Table 11. Distances between neighboring imino protons in the helix IV region of *B. megaterium* 5S rRNA. Calculated by molecular dynamics without constraints.

<table>
<thead>
<tr>
<th>Imino proton from</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G80*C92</td>
<td>4.22</td>
</tr>
<tr>
<td>G81*C91</td>
<td>4.20</td>
</tr>
<tr>
<td>G82*C90</td>
<td>3.95</td>
</tr>
<tr>
<td>A83*U89</td>
<td>3.71</td>
</tr>
<tr>
<td>C84*G88</td>
<td></td>
</tr>
<tr>
<td>C84*G88 and U85</td>
<td>3.64</td>
</tr>
<tr>
<td>C84*G88 and U86</td>
<td>7.06</td>
</tr>
<tr>
<td>C84*G88 and U87</td>
<td>5.51</td>
</tr>
<tr>
<td>U85 and U86</td>
<td>4.13</td>
</tr>
<tr>
<td>U85 and U87</td>
<td>6.00</td>
</tr>
<tr>
<td>U86 and U87</td>
<td>8.38</td>
</tr>
</tbody>
</table>
thermodynamically stable conformation. On the other hand, the natural structure of the molecule derived from intact 5S rRNA may not be energetically the most stable one. Local instability in the molecule can be compensated by stabilization caused by a variety of interactions between the relatively unstable part or fragment and the remaining part in intact 5S rRNA. This explanation is, of course, only speculation, due to lack of knowledge about those interactions. At any rate, it is probable that even though the helix is stable, the hairpin loop (loop D) is susceptible to any change with a reasonable activation energy barrier height. Obvious structural change in this region was observed in the NMR spectrum when Fragment B was heated over 45 °C and cooled down (Chapter V), whereas Fragment A and Fragment C returned to the original conformation after heating and cooling. However, intact 5S rRNA showed thermal reversibility. The presumed rearrangement of the molecular structure of the fragment containing helix IV (but not of the corresponding helix in intact 5S rRNA) after heating could be a line of evidence to support this idea.

E. Concluding Remarks

Molecular structure derived from molecular dynamics calculations can provide a deeper knowledge of molecular conformation. Structures of the helices II-III segment and the helix IV segment obtained by molecular dynamics were generally consistent with the solution structures monitored by NMR spectroscopy. Only the bulge (A_{50} and A_{51}) in the first segment and the arrangement of bases in the hairpin
loop of the helix IV segment (loop D) were slightly different from the solution structure observed by NMR.

There are a few possible reasons for the minor differences in the structure acquired by dynamics calculations from that observed by NMR. First, the dynamics studies were conducted without inclusion of solvent molecules (namely in vacuum state). Therefore, interactions between the molecules and water or cations from salts (e.g., Na+) were neglected. This difference in the environment of the molecule in dynamics from that in solution may contribute to minor structural differences between the two. Alternatively, the currently adopted structure of the molecule may rearrange into another one of lower total energy (higher stability) with continuous running of dynamics calculations. The scale of the potential energy surface for such a macromolecule is so vast that finding the global minimum is extremely difficult. Therefore, dynamics calculations have been occasionally performed with distance constraints obtained by two dimensional NOE experiments (for example, Hare et al., 1986), thus allowing construction of structures that both have low energy and fit the experimental data. For that strategy, however, there is always a danger of overinterpreting the NMR data to force agreement (Rao & Kollman, 1990). Another explanation for the difference is, as stated in the previous section, that the natural conformation of intact 5S rRNA may be stable overall because of numerous interactions within the molecule, but each segment of the whole molecule may not be in the most stable shape.

At any rate, molecular dynamics can be useful for gathering valuable knowledge about the structure of a molecule in addition to reconfirming some structural features of the molecule obtained by NMR experiments.
The stability of helices II, III, and IV (and of course helix I) of \textit{B. megaterium} 5S rRNA is a common conclusion drawn from both NMR and dynamics studies. Therefore, once more, the modified Fox-Woese model can certainly be believed to best describe the secondary structure of \textit{B. megaterium} 5S rRNA.
REFERENCES


Hare, D. R., & Reid, B. R. (1982b) *Biochemistry* 21, 5129-5131.


Tinoco, I. Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C.,

Tinoco, I., Uhlenbeck, O. C., & Levine, M. D. (1971) *Nature (London)* 230,
362-367.


510.

495-504.

Mol. Biol.* 152, 699-721.


784.


135.

Wittmann, H. G. (1986) in *Structure, Function, and Genetics of Ribosomes*

Evol.* 8, 143-153.

Woese, C. R., Magrum, L. J., Gupta, R., Siegel, R. B., & Stahl, D. A.


Biol.* 72, 741-749.


