NMR of Nucleic Acids

K.V.R. Chary
chary@tifr.res.in

Workshop on “NMR and it’s applications in Biological Systems”
TIFR
November 26, 2009
Nucleic Acids are Polymers of Nucleotides

Each nucleotide consists of a nitrogenous base, a sugar and a phosphate group.
chain reaction
Nucleic Acids are Polymers of Nucleotides

Each nucleotide consists of a nitrogenous base, a sugar and a phosphate group.
The two types of base

Pyrimidine Ring

Purine Ring
Common pyrimidine bases

- Cytosine
- Uracil (RNA)
- Thymine (DNA)
Common purine bases

Adenine

Guanine
Pentose sugars

Ribose

2-Deoxyribose
The Nomenclature in Nucleic acids
Deoxy D-ribose
D-Ribose
Adenine (A)
Guanine (G)
Inosine (I)
Cytosine (C)

Sugar

\( \text{Cytosine (C)} \)
Thymine (T)

Sugar
Uracil (U)

Sugar
Watson-Crick A:T base pair
Watson-Crick G:C base pair
Base-pairings in Nucleic Acids

- Watson-Crick base pairing
- Reverse Watson-Crick base pairing
- Hoogsteen base pairing
Pairing sites in A:T base-pair

Hoogsteen base pairing sites

Watson-Crick base pairing sites
Pairing sites in G:C base pair

Hoogsteen base pairing sites

Watson-Crick base pairing sites
Hoogsteen A.T base-pair
Hoogsteen G.C\(^+\) base-pair
T.A:T Triad
$C^+.G:C$ Triad
Why study Nucleic Acids Structures?
Structures of nucleic acids and their complexes represent around 10% of the PDB.

<table>
<thead>
<tr>
<th>PDB Holdings List: 20-Jan-2004</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Molecule Type</td>
</tr>
<tr>
<td>Proteins, Peptides, and Viruses</td>
</tr>
<tr>
<td>Exp.</td>
</tr>
<tr>
<td>Tech.</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Please note that theoretical models have been removed, effective July 02, 2002, as per PDB policy.

11656 Structure Factor Files
1805 NMR Restraint Files
NMR Spectroscopy is an Important Method for Structural Studies of Nucleic Acids

PDB Holdings List: 20-Jan-2004

<table>
<thead>
<tr>
<th>Exp.</th>
<th>X-ray Diffraction and other</th>
<th>18788</th>
<th>892</th>
<th>716</th>
<th>14</th>
<th>20410</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tech.</td>
<td>NMR</td>
<td>2923</td>
<td>95</td>
<td>565</td>
<td>4</td>
<td>3587</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>21711</td>
<td>987</td>
<td>1281</td>
<td>18</td>
<td>23997</td>
</tr>
</tbody>
</table>

Please note that theoretical models have been removed, effective July 02, 2002, as per PDB policy.

11656 Structure Factor Files
1805 NMR Restraint Files

<table>
<thead>
<tr>
<th>X-ray/NMR</th>
<th>6.4</th>
<th>9.4</th>
<th>1.4</th>
<th>7</th>
<th>5.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray/NMR(2001)</td>
<td>6.1</td>
<td>7.9</td>
<td>1.5</td>
<td>3.5</td>
<td>5.4</td>
</tr>
</tbody>
</table>
# HELICAL PARAMETERS OF NUCLEIC ACIDS

<table>
<thead>
<tr>
<th>Helix Type</th>
<th>Pitch $\text{Å}^\circ$</th>
<th>Axial Rise $\text{Å}^\circ$</th>
<th>Rotation $\text{Å}^\circ$</th>
<th>Sugar Pucker</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>28.1</td>
<td>2.6</td>
<td>33</td>
<td>C3’ endo</td>
</tr>
<tr>
<td>B</td>
<td>33.7</td>
<td>3.4</td>
<td>36</td>
<td>C2’ endo</td>
</tr>
<tr>
<td>C</td>
<td>30.9</td>
<td>3.3</td>
<td>39</td>
<td>C3’ endo</td>
</tr>
<tr>
<td>D</td>
<td>27.2</td>
<td>3.4</td>
<td>45</td>
<td>C3’ endo</td>
</tr>
<tr>
<td>$Z_1$</td>
<td>44.6</td>
<td>7.4*</td>
<td>-60*</td>
<td>alternating C3’ endo C2’ endo</td>
</tr>
</tbody>
</table>

* (Per dinucleotide)
Figure 1  H-DNA model. Bold line, homopurine strand; thin line, homopyrimidine strand; dashed line, the half of the homopyrimidine strand donated to the triplex.

- homopurine strand
- homopyrimidine strand
- part of the homopyrimidine strand
A TETRAMERIC DNA WITH PROTONATED CYTOSINE-CYTOSINE BASE PAIR
Statistical distribution of various torsion angles in B-DNA derived from PDB files.
Various backbone ($\alpha, \beta, \gamma, \delta, \varepsilon, \zeta$) and glycosidic ($\chi$) torsion angles for purine (R) and pyrimidine (Y) bases.

<table>
<thead>
<tr>
<th>Code</th>
<th>Torsion angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>{O3$\ '_{i-1}$,P$_i$,O5$\ '_i$,C5$\ '_i$}</td>
</tr>
<tr>
<td>$\beta$</td>
<td>{P$_i$,O5$\ '_i$,C5$\ '_i$,C4$\ '_i$}</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>{O5$\ '_i$,C5$\ '_i$,C4$\ '_i$,C3$\ '_i$}</td>
</tr>
<tr>
<td>$\delta$</td>
<td>{C5$\ '_i$,C4$\ '_i$,C3$\ '_i$,O3$\ '_i$}</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>{C4$\ '_i$,C3$\ '_i$,O3$\ '<em>i$,P$</em>{i+1}$}</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>{C3$\ '<em>i$,O3$\ '<em>i$,P$</em>{i+1}$,O5$\ '</em>{i+1}$}</td>
</tr>
<tr>
<td>$\chi$(Py)</td>
<td>{O4$\ '_i$,C1$\ '_i$,N1$\ '_i$,C2$\ '_i$}</td>
</tr>
<tr>
<td>$\chi$(Pu)</td>
<td>{O4$\ '_i$,C1$\ '_i$,N9$\ '_i$,C4$\ '_i$}</td>
</tr>
</tbody>
</table>
Proteins vs Nucleic acids

• On an average protons per residue is same in nucleotides and amino acids.

• The average molecular weight is three times for nucleotides (~350 Da) as compared to peptides (110 Da).

• This results in a three-fold lower proton density in nucleic acids.
NMR of Nucleic acids

• Predominantly rod like structures

• No long-range interactions

• Results in smaller chemical shift dispersion, except for non-canonical structural elements such as hairpins, bulges or internal loops.
NMR of Nucleic acids

• Seven of the protons in nucleotides (H1’, H2’, H2’’, H3’, H4’, H5’, and H5’’) are contributed by the sugar rings, resulting in fairly complex spin-coupled system.

• The 100% abundant hetero-nucleus $^{31}$P shows J-coupling to H3’ on one end, and H5’/H5’’ at the other.

• All these factors lead to overlapping and strongly coupled spectra.
NMR of Nucleic acids

• Information on the backbone structure can be obtained from a limited number of nuclei, namely, H3′, H4′, H5′, and H5″ and 31P.

• These protons and 31P suffers from relatively small chemical shift dispersions.

• This results in fewer experimental data for constraining the six backbone torsion angles.

• For these reasons, the structure determination of DNA and RNA from 1H NMR has been limited to fragments containing less than 50 nucleotides.
NMR structure determination of nucleic acids is primarily based on:

• Estimation of short $^1\text{H-}^1\text{H}$ distances detected by NOEs

• Dihedral angles that can be derived from the through bond couplings

• Characterisation of base-pairing.
Resonance Assignments in DNA

(a) Classification of individual chemical shifts in 1D NMR spectrum

(b) Identification of networks of coupled spin systems contained within the individual nucleotide units using COSY, E.COSY, TOCSY etc., and

(c) Sequential linking of the identified nucleotide units by NOESY spectrum recorded in H₂O or ²H₂O, where one observes inter-nucleotide NOE connectivities, such as

\[(TH3/G1H) \rightarrow \text{imino (TH3/G1H)} \]
\[(H6/H8)_{i+1} \rightarrow (H1'/H2'/H2'')_i \]
NMR of Nucleic acids

d-GGTACIAGTACC
CCATGAICATGG-d
Resonance Assignments in DNA

A) Exchangeable protons: 1D $^1$H, 2D NOESY

B) Non-exchangeable protons
• Aromatic Spin Systems: 2D DQF-COSY (H5-H6), 2D NOESY
• Sugar Spin Systems: 2D DQF-COSY 2D TOCSY
• Sequential Assignment: 2D NOESY 2D ($^{31}$P, $^1$H) HELCO

C) Correlation of exchangeable and non-exchangeable protons: 2D NOESY
Identification of non-exchangeable protons: C(H5-H6)/T(CH₃-H6)
Correlations in DQF-COSY

A)
Identification of non-exchangeable sugar protons: H1'-H2'/H2'' Correlations in 2D-TOCSY/2D-DQF-COSY
Intranucleotide and Internucleotide connectivities in a NOESY spectrum

Intranucleotide peaks seen in NOESY spectrum

$\{\text{Base}(H6/H8)\}_i \rightarrow \{\text{Sugar}(H1'/H2'/H2''/H3')\}_i$

These distances are less than 50nm

Internucleotide peaks seen in NOESY spectrum

$\{\text{Base}(H6/H8)\}_i \rightarrow \{\text{Base}(H6/H8)\}_{i-1}$

$\{\text{Sugar}(H1'/H2'/H2''/H3')\}_i \rightarrow \{\text{Sugar}(H1'/H2'/H2''/H3')\}_{i-1}$
NOESY spectrum of a 12mer ds DNA with a Hemimethylated –GATC-tract

5’ G G C A G A T C C G T A 3’
3’ C C G T C T A G G C A T 5’
Intranucleotide and Internucleotide connectivities in a NOESY spectrum

Intranucleotide peaks seen in NOESY spectrum

\{\text{Base(H6/H8)}\}_i \rightarrow \{\text{Sugar(H1’/H2’/H2”/H3’)}\}_i

These distances are less than 50nm

Internucleotide peaks seen in NOESY spectrum

\{\text{Base(H6/H8)}\}_i \rightarrow \{\text{Base(H6/H8)}\}_{i-1}

\{\text{Base(H6/H8)}\}_i \rightarrow \{\text{Sugar(H1’/H2’/H2”/H3’)}\}_{i-1}
SEQUENCE SPECIFIC RESONANCE ASSIGNMENTS OF NUCLEIC ACIDS

1. (Base)_{i+1} \xrightarrow{NOESY} (Sugar)_{i} \xrightarrow{NOESY} (Base)_{i} \xrightarrow{NOESY} (Base)_{i+1} \xrightarrow{NOESY} (Base)_{i}

2. (P)_{i+1} \xrightarrow{HELCO} (H3’)_{i} \xrightarrow{TOCSY} (H4’)_{i} \xrightarrow{HELCO} (P)_{i} \xrightarrow{----- J – Correlation}

3. (imino protein)_{i+1} \xrightarrow{NOESY} (imino protein)_{i}

- A T \xrightarrow{\text{NOESY}} T3 NH
- G C \xrightarrow{\text{NOESY}} G1 NH

NOESY in 90% H2O + 10% D2O
Typical NOEs seen in DNA
Typical NOEs seen in DNA
NOESY spectrum of a 12mer ds DNA with a Hemimethylated –GATC-tract

5’ G G C A G A T C C G T A 3’
3’ C C G T C T A G G C A T 5’
Assignments of non-exchangeable protons: Sequential Assignments
Correlations between exchangeable and non-exchangeable protons
Correlations between exchangeable and non-exchangeable protons
(A) I (anti) – A (anti)

\[ \text{I1NH} - \text{AH2} \]
\[ \text{IH2} - \text{AH2} \]

(B) I (syn) – A (anti)

\[ \text{IH8} - \text{AH2} \]

(C) I (anti) – A (syn)

\[ \text{I1NH} - \text{AH8} \]
\[ \text{IH2} - \text{AH8} \]
NOESY IN 90% $\text{H}_2\text{O} + 10% \text{ }^2\text{H}_2\text{O}$
Correlations between exchangeable and non-exchangeable protons
Correlations between exchangeable and non-exchangeable protons
Sugar Pucker
D-Ribose
Sugar Pucker

H1′-H2″  H2′-H3′

H2″-H3′

H3′-H4′

H3′-H4′

H2″-H3′

H1′-H2′
DETERMINATION OF ‘J’
<table>
<thead>
<tr>
<th>EXPERIMENTAL</th>
<th>SIMULATED</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>(b)</td>
</tr>
</tbody>
</table>

H1' (A5)

<table>
<thead>
<tr>
<th>EXPERIMENTAL</th>
<th>SIMULATED</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>(b)</td>
</tr>
</tbody>
</table>

H2', H2''

H1'
Fixing $\chi$

The glycosidic torsion angle
Nucleoside geometry. On the left, drawings to show the two major furanose conformations encountered in structural studies and the effect each has on the phosphorus-phosphorus distances. On the right, the syn and anti conformation of purines with respect to the furanose rings.
NOESY DECAMER \([T_m \ 80 \ ms]\)
Simulated isodistance contours in the P, \( \chi \) space (for pyrimididine)
$^{31}$P – $^1$H Correlation

Small molecules: Fine, No problem

Large molecules: - Small chemical shift differences

(H3’,H4’,H5’,H5” and $^{31}$P)
- Complex multiplet structures
- Short $T_2$’s of H5’ & H5”.
- Frequency dependent $T_2$’s of $^{31}$P

$^{31}$P

Spin – $\frac{1}{2}$

100% Natural Abundance
(P)\textsubscript{i} \rightarrow (H3')\textsubscript{i-1} \rightarrow (H4')\textsubscript{i-1} \rightarrow (P)\textsubscript{i-1}
$^{31}\text{P} - \text{H}3'$ Active Coupling

$\text{H}3' - \text{H}2'$ Passive couplings

$\text{H}3' - \text{H}2''$ couplings

$\text{H}4'$
C Exp Sim

(\textsuperscript{31}P - H3\textsuperscript{1}) cross peaks.

D Exp Sim

(\textsuperscript{31}P - H4\textsuperscript{1}) cross peaks.
Figure 4. (a) Variation of H3'-P coupling constant as a function of $\epsilon$ (redrawn from Ref. 15). The Newman projection diagram is also shown to indicate the relation between $\epsilon$ and the H3'-C3'-O3'-P torsion angle. (b) Energy profiles for a mononucleotide having 3' phosphate as a function of $\epsilon$ torsion angle. The sugar geometry is fixed in 04'-endo geometry. The curves a-c correspond to different sets of backbone torsion angles. (a) $\alpha = 303^\circ$, $\beta = 179^\circ$, $\gamma = 52^\circ$, $\delta = 97^\circ$. (b) $\alpha = 90^\circ$, $\beta = 101^\circ$, $\gamma = 5^\circ$, $\delta = 92^\circ$. ...
New and Novel Constraints

- Hydrogen-bond constraints from $J$ couplings across hydrogen bonds
- Sugar pucker and backbone angles from cross-correlated relaxation
- Long-range orientation constraints from residual dipolar couplings
Trans Hydrogen bond
Coupling constants
Trans hydrogen-bond J-correlations

$2^h J(\text{NN})$

$J_{\text{NN}} = 6.3$ Hz

$J_{\text{NN}} = 6.7$ Hz

Trans hydrogen-bond $^{2h}J(\text{NN})$ in non-Watson-Crick base pairs:
Trans hydrogen-bond $^{2h}J(\text{NN})$ via Remote Non-Exchangeable Protons

Sheared GA in G(CA) triads ($J_{\text{NN}} = 1.7 \text{ Hz}$)

AU Hoogsteen in AUA Triple ($J_{\text{NN}} = 5.2 \text{ Hz}$)
Nucleic Acids Polymorphism

Examples of non-duplex structures are
Single stranded hair-pins
Triplexes,
Tetraplexes
i-motif structures

Such nucleic acid polymorphism is principally
 governed by the factors such as sequence, concentration,
temperature, pH, and other persistent solvent conditions
Nucleic Acids Polymorphism
Nucleic Acids Polymorphism
The Triplex DNA (H-DNA)

Purine strand
Pyrimidine strand
Part of the purine strand
NMR Spectroscopy is also important for Structural Studies of RNA

<table>
<thead>
<tr>
<th>Methods</th>
<th>X-ray</th>
<th>NMR</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA</td>
<td>77</td>
<td>128</td>
<td>14</td>
<td>219</td>
</tr>
<tr>
<td>RNA-Protein Complexes</td>
<td>104</td>
<td>21</td>
<td>9</td>
<td>134</td>
</tr>
<tr>
<td>DNA</td>
<td>431</td>
<td>236</td>
<td>14</td>
<td>681</td>
</tr>
<tr>
<td>DNA/RNA Hybrids &amp; Complexes</td>
<td>63</td>
<td>23</td>
<td>0</td>
<td>86</td>
</tr>
</tbody>
</table>
Resonance Assignment of RNA

Exchangeable protons:

1D  1H,
2D NOESY
2D ( 15N, 1H) HSQC
3D  15N-edited NOESY-HSQC

Non-exchangeable protons

2D ( 13C, 1H) HSQC
3D HCCH-COSY
3D HCCH-TOCSY
2D ( 15N, 1H) HCN
Resonance Assignment of RNA

Sequential Assignment:
3D/4D  13C-edited NOESY-HSQC
HCP-type only for small RNA

Correlation of exchangeable and non-exchangeable protons:
2D G-specific H(NC)-TOCSY-(C)-H
2D A-specific (H)N(C)-TOCSY-(C)-H
2D C-specific H(NCCC)H
2D U-specific H(NCCC)H
3D/4D  13C-edited NOESY-HSQC
Conclusions

• Conformational freedom of nucleotides
• Structure of DNA, Watson & Crick Model
• Base pairing and base stacking
• NMR spectroscopy of DNA
• Assignment Strategies
• Use of NMR parameters to fix sugar pucker, glycosidic dihedral angle and backbone torsion angles
• Base-pairing and stacking information from NMR
• Structure Simulations of Nucleic acids
• Variations in DNA structure, polymorphism, unusual DNA structures
• Higher order structures
• Structure of RNA: Ribosomal and T-RNA