Nucleic Acid NMR

Yuan-Chao Lou
羅元超

Institute of Biomedical Sciences
Academia Sinica

2011/04/27
Outline

I. Nucleic acids hold diverse structures and functions
   a. In vitro SELEX
   b. Diverse structures of nucleic acids

II. NMR spectroscopy for nucleic acid assignment
   a. The building blocks of nucleic acids
   b. Resonance assignment in nucleic acids

III. Advanced developments in NMR spectroscopy
   a. Residual Dipolar Coupling (RDC)
   b. Transverse Relaxation-Optimized Spectroscopy (TROSY)
   c. Paramagnetic spin labeling

IV. Cases of protein/nucleic acids complexes
I. Nucleic acids hold diverse structures and functions
Nucleic acids have diverse functions

* It serves multiple functions:
  
  * Carrier of heritable information
  * Catalysis: nucleic acid enzymes
  * Aptamer: nucleic acid monoclonal antibody (SELEX)

* Targets for drug interaction
SELEX Methodology

• Systematic Evolution of Ligands by EXponential enrichment

• The ligands that emerge from SELEX have been called aptamers

SELEX Methodology

- SELEX is a technology for the identification of high affinity oligonucleotide ligands. Large pool of random sequence single-stranded oligonucleotides, whether RNA or DNA, can be thought of conformationally not as short strings but rather as sequence-dependent folded structures. This conformational complexity means that such a library is a source of high affinity ligands for a surprising variety of molecular targets, including nucleic acid binding proteins such as polymerases and transcription factors, non-nucleic acid binding proteins such as cytokins and growth factors, as well as small organic molecules such as ATP. The range of applications of this technology extends from basic research reagents to the identification of novel diagnostic and therapeutic reagents, such as thrombin and HIV-1 integrase inhibitors.
SELEX Methodology

First prepare a DNA oligonucleotide with random sequences in the central region. Then use PCR to amplify the oligonucleotides.

Use a biotin-labeled DNA primer to separate DNA double strands.

If RNA, then a reverse PCR was performed to amplify the few oligonucleotides.
**Consensus Sequences**

The consensus sequences must form a conserved secondary structure.
DNA Aptamers as Potential anti-HIV Agents

a. Recently several DNA aptamers have been screened using SELEX to exhibit nanomolar activity as anti-HIV1-integrase inhibitors

b. The aptamers adopt novel G-quadruplex structures


Possible Multi-stranded G-Quadruplex structures

1. G-tetrad H-bonding
2. G-tetrad hydrophobic stacking
A Novel G-Quadruplex structures for 5’-GGGGTGAGGGAGGT-3’

1. Interlocked dimer
2. Three reversal loops
3. Contain G-G-G-G-A pentad
4. Four hanging out bases
A Novel G-Quadruplex structures for 5'-GGGGTGGGGAGGAGGGT-3'
G-Quadruplex and HIV1-Integrase Complex

G-Quadruplex is situated in the center to block the active site of HIV1-IN
Nucleic acids have diverse structures

* Triplex, Quadruplex, i-motif multi-stranded structures in the promoter, centromere, or telomere region
* Left-handed duplex structures
Human Telomere, Telomerase, and Cancer Chemotherapy

a. Human telomere contains \textit{tandem repeats} of TTAGGG sequence

b. DNA polymerase \textit{cannot} synthesize end of chromosome

c. The “end” problem needs to be fixed, or a cell will “\textit{age}” or “\textit{die}” due to the loss of its chromosome content

d. Telomerase comes as a rescue. It is quite active in cancer cell, but not in normal cell

e. The substrate of a telomerase is the \textit{single-stranded} (TTAGGG)n hanging. Since it can form a \textit{G-quadruplex}, hence an agent that can \textit{stabilize} such a structure can serve a drug for cancer therapy.
Human Telomere, Telomerase, and Cancer Chemotherapy

A simple mechanism of telomerase action

A reverse transcriptase
Nucleic Acids with Canonical GC and AT Paired Duplex under Neutral Condition
Nucleic Acids with G-rich sequences in a duplex may form separated multiple-stranded G-quadruplex and i-motif structures at acidic (pH 4.5) condition.
Nucleic Acids with G-rich sequences in a promoter may form separated multiple-stranded G-quadruplex and i-motif structures.

RNA structure

Secondary structure
- single strand
- A-form double helix
- Double helix with 5'-dangling end
- single nucleotide bulge
- three nucleotide bulge
- hairpin loop
- mismatch pair or symmetric internal loop of 2 nucleotides
- symmetric internal loop
- asymmetric internal loop
- two-stem junction or coaxial stack
- three-stem junction
- four-stem junction

Tertiary structure
- a) STEM 1
- b) Pseudoknot
- c) Kissing hairpins

Hairpin loop - bulge contact

By Michael Sattler
Proteins recognize unusual RNA structural elements

RNA structure: protein recognition

By stabilizing an adjacent interaction surface, bulges can participate in complex protein binding sites.
Nucleic Acids are Special Biomolecules

* It serves not only as targets for cancer chemotherapy, but also as inhibitors (aptamers) toward other targets such as HIV-1 integrase.
II.

NMR Spectroscopy for Nucleic Acid Assignment
The building blocks of nucleic acids

Nomenclature, structures, and atom numbering for the sugars contained in common nucleotides.
Nomenclature, structures, and atom numbering for the bases contained in common nucleotides.
1D $^1$H NMR spectrum in Nucleic Acids (in D$_2$O)
1D $^1$H NMR spectrum in Nucleic Acids (in H$_2$O)

**Canonical G:C and A:T base pairs**

*Watson-Crick G:C Pairing*

G imino proton
(12-13 ppm)

*Watson-Crick A:T Pairing*

T imino proton
(13-14 ppm)
Why Imino Protons have Very Downfielded Resonances

• Ring-current shifting effect
  • $A/G > C/T$
• Inductive effect
• H-bonded effect
NOE-based assignment in unlabeled nucleic acids
NMR of canonical DNA duplexes

1) Systematic assignment by 2D NOESY in H₂O

2) Systematic assignment by 2D NOESY in D₂O


Regular B-form DNA Duplex Structure

* Watson-Crick G • C and A • T H-bonding

* *Intra-strand* base-base hydrophobic stacking
Typical 1D and 2D-NOESY NMR Spectra In D₂O
Flowcharts for resonance assignment in nucleic acids

A. NOE-based assignment in unlabeled nucleic acids

I (H₂O)
Assignment of imino (and amino) resonances to establish base pairing
NOESY imino-imino, amino-imino

II (H₂O)
Partial resonance assignment of non-exchangeable protons via NOE connectivities to amino and/or imino protons
NOESY imino-H₂/H₆/H₈/H₅/H₁'

III (D₂O)
Identification of sugar proton spin systems
(mainly H₁'/H₂'/H₂''/H₃') (¹H, ¹H) COSY/TOCSY

Identification of aromatic spin systems
(Cytosine/Thymine H₅/H₆) (¹H, ¹H) COSY/TOCSY

Sequential resonance assignment
NOESY H₆/H₈-H₁', H₆/H₈-H₂'H₂''

IV (D₂O)
Assignment of ³¹P resonances and confirm/extend H₃’,H₄’,H₅’,H₅” assignments (¹H, ³¹P) HETCOR/HETTOC
Imino Proton to Amino to Aromatic Protons

- **a. Imino to imino**
- **b. Imino to amino/aromatic**
- **c. Amino to aromatic**
- **d. Aromatic to H2'/H2''**
I. Assignment of imino (and amino) resonances in H$_2$O

Imino Proton Assignments by 2D NOESY spectrum

5'–CGACGATGACGTCATCGTCG–3'
3'–GCTGCTACTGCAGTAGCAGC–5'

2G-18T-17G-5G-15T-7T-8G-12T-11G

II. NOESY imino-H2/H6/H8/H5/H1’ in H2O

b. Imino to amino/aromatic

Only intra-strand aromatic to aromatic connectivities

III. Identification of aromatic spin systems in D₂O

III. Sequential resonance assignment in D_{2}O

NOESY H6/H8-H1', H6/H8-H2'H2''

Cytosine: CH5-CH6

Duplex-hairpin 5'-CGCGTATACGCG-3'


III. Sequential resonance assignment in D$_2$O

**NOESY H6/H8-H1’**

Only intra-residue cross peaks were marked. a-f. are the six big CH5-CH6 cross peaks.

III. Sequential resonance assignment in D$_2$O

NOESY H6/H8-H2’H2”

Only intra-residue cross peaks were marked.
\(^1H-^{31}P\) Correlation Spectrum

\[
\begin{align*}
\end{align*}
\]

\[(n-1) \ H^3' - (n) \ P \quad (n) \ P - (n) \ H^4'
\]

NOE-based and via through-bond coherence transfer assignment in labeled nucleic acids
Flowcharts for resonance assignment in nucleic acids

B. NOE-based assignment in labeled nucleic acids

I (H₂O)

- Exchangeable proton/nitrogen correlation
  - 2D $^{15}$N-HMQC imino $^1$H optimized G N1H, U N3H
  - amino $^1$H optimized C N4H₂, G N2H₂, A N6H₂

- Exchangeable proton/nitrogen sequential assignment
  - 3D $^{15}$N-NOESY-HMQC (imino $^{15}$N edited NOESY)
    - imino-imino, amino-imino
  - 3D $^{15}$N-NOESY-HMQC (amino $^{15}$N edited NOESY)
    - amino-imino

II (H₂O)

- Partial resonance assignment of non-exchangeable proton from NOE connectivities with amino and/or imino protons
  - 3D $^{15}$N-NOESY-HMQC (imino $^{15}$N edited NOESY)
    - aromatic-imino

Flowcharts for resonance assignment in nucleic acids

III (D₂O)

Identification of **sugar proton spin systems**
- 3D HCCH-COSY H1’-H2’
- 3D HCCH-RELAY H1’-H2’/H3’
- 3D HCCH-TOCSY

Identification of **sugar carbon spin systems**
- 2D 13C-CT-HSQC/HMQC
- 3D HCCH-COSY H1’-C2’
- 3D HCCH-RELAY H1’-C2’/C3’
- 3D HCCH-TOCSY H1’-C2’/C3’/C4’/C5’

Identification of proton/carbon **aromatic spin systems**
- 2D ¹³C-CT-HSQC/HMQC H6-C6, H8-C8, H5-C5, H2-C2
- 2D/3D HCCH-COSY H6-H5, H6-C6/ C5, H5-C6/ C5

**Sequential resonance assignment**
- 3D ¹³C-NOESY-HMQC H6/H8-H1’, H6/H8-H2’H2”

IV (D₂O)

Assignment of ³¹P resonances e.g. (¹H, ³¹P) HETCOR/HETTOC

(continued)

Flowcharts for resonance assignment in nucleic acids

C. Assignment via through-bond coherence transfer in labeled nucleic acids

I (H₂O)
Exchangeable proton/nitrogen correlation
2D $^{15}\text{N}$-HMQC imino $^{1}\text{H}$ optimized G N1H, U N3H
amino $^{1}\text{H}$ optimized C N4H₂, G N2H₂, A N6H₂

II (H₂O)
Through-bond amino/imino to non-exchangeable base proton correlations
HNCCCH/HCCNH

III (D₂O)
1. Through-bond H2-H8 correlations \{HCCH-TOCSY/($^{1}\text{H},^{13}\text{C}$) HMBC\}
2. Through-bond base-sugar correlations
   \{HCN (base) with HCN (sugar), HCNCH,
   HCNH, \{HCN (sugar) with H8N9(H8)C8H8\},
   \{HCN (sugar) with (Hb,Hb) HSQC\},
   \{(H1’, C8/6) HSQC with (H8/6, C8/6) HSQC\}
3. Through-bond sugar correlations
   \{HCCH-COSY/ HCCH-TOCSY\}
4. Sequential resonance assignment via through-bond sugar-phosphate backbone correlations ($^{1}\text{H},^{13}\text{C},^{31}\text{P}$)
   HCP/ PCH/ PCCH-TOCSY/ HPHCH
RNA synthesis by in vitro transcription

RNA samples at natural isotopic abundance and enriched in $^{15}\text{N}$ and $^{13}\text{C}$ can be prepared with T7 RNA polymerase.
# Heteronuclear Chemical Shifts in Nucleotides

<table>
<thead>
<tr>
<th>Atom(s)</th>
<th>Chemical shift range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Carbon</strong></td>
<td></td>
</tr>
<tr>
<td>Purine C5</td>
<td>~120 ppm</td>
</tr>
<tr>
<td>Purine C8</td>
<td>~140 ppm</td>
</tr>
<tr>
<td>Purine C2,C4,C6</td>
<td>~150-160 ppm</td>
</tr>
<tr>
<td>Pyrimidine C5</td>
<td>~100 ppm</td>
</tr>
<tr>
<td>Pyrimidine C6</td>
<td>~145 ppm</td>
</tr>
<tr>
<td>Pyrimidine C2,C4</td>
<td>~155-170 ppm</td>
</tr>
<tr>
<td>Ribose carbons</td>
<td>~70-90 ppm</td>
</tr>
<tr>
<td><strong>Nitrogen</strong></td>
<td></td>
</tr>
<tr>
<td>Purine N9</td>
<td>~170 ppm</td>
</tr>
<tr>
<td>Purine N7</td>
<td>~220-240 ppm</td>
</tr>
<tr>
<td>Guanosine N1</td>
<td>~145-150 ppm</td>
</tr>
<tr>
<td>Guanosine N2</td>
<td>~70-80 ppm</td>
</tr>
<tr>
<td>Guanosine N3</td>
<td>~160-165 ppm</td>
</tr>
<tr>
<td>Adenosine N1</td>
<td>~220 ppm</td>
</tr>
<tr>
<td>Adenosine N3</td>
<td>~215 ppm</td>
</tr>
<tr>
<td>Adenosine N6</td>
<td>~80-90 ppm</td>
</tr>
<tr>
<td>Pyrimidine N1</td>
<td>~145-150 ppm</td>
</tr>
<tr>
<td>Uridine N3</td>
<td>~150-160 ppm</td>
</tr>
<tr>
<td>Cytosine N3</td>
<td>~195-200 ppm</td>
</tr>
<tr>
<td>Cytosine N4</td>
<td>~95 ppm</td>
</tr>
</tbody>
</table>

---

2D $^{15}$N–$^1$H HMQC spectra of RNA imino resonances at different conditions

The $^1$H-$^{13}$C HSQC spectra of labeled nucleic acids
(A) H6/H8-C6/C8, (B) H1'-C1', (C) H2'/H2''-C2', and (D) H3'-C3'
Intraresidue correlation
via through-bond coherence transfer NMR experiments

H1’, H2’, H3’, H4’, H5’, H5” correlations, HCCH-TOCSY

H2-H8 or H5-H6 correlation

H8-H1’ correlation, HCN

Nucleotide spin system
Correlation in the base-sugar: HCN 3D spectra

2D and 3D TROSY-HCN for obtaining ribose base and intra-base correlations in the nucleotides of DNA and RNA.

Dotted arrows indicate the intra-base transfers and solid arrows the ribose-base transfers.
3D HCCH-TOCSY

Interresidue correlation through bond (HCP)

2 spin systems can be linked

H3’-C3’-P(n-1)
Residue n

H5’,H5’’(n-1)-C5’(n-1)-P(n-1)

Direct observation of H-bonds in nucleic acid base pairs by inter-nucleotide $^2J_{NN}$ couplings

The connectivity between $^1H_3(U329)$ and the $^1H_2-^{13}C_2(A32)$

$J_{NN}$ HNN-COSY

$J_{HN}$ HSQC

3D$^{13}$C-NOESY

$^1H_2 (A)$ to $^{15}N_1(A)$ and $^{15}N_3(A)$

Structural Determination of Nucleic Acids by NMR

1) Similar to those used in protein
2) First build a nucleic acid sequence template
3) Input H-bonded constraints
4) Input all exchangeable and non-exchangeable distance constraints and/or dihedral constraints
5) Use Distance Geometry calculation to get some initial structures
6) Use Molecular Dynamics method to refine the structures
III.

Advanced developments in NMR Spectroscopy
New Techniques in NMR Spectroscopy

(1). Residual Dipolar Coupling (RDC)
   NOE, dihedral angle and H-bond are short-range restraints and have limitations for some structure determination, like extended structures or multiple-domain structure. RDC is a novel restraint and provides global structure information.

(2). Transverse Relaxation-Optimized Spectroscopy (TROSY)
   TROSY, which was developed by K. Wüthrich, can select one fourth of the signals that relax more slowly than the others. The utilization of TROSY techniques push the size limit of NMR spectroscopy to 30~50 kDa.

(3). Other Applications
   Paramagnetic Spin Labeling.
Residual Dipolar Coupling (RDC)

Residual dipolar couplings arise from dipole-dipole interactions between nuclei. In aqueous solution, the isotropic orientation of the molecules average out the dipolar couplings. However, in oriented media, the molecular tumbled anisotropically. The order of $10^{-4}$ to $10^{-3}$ of anisotropy tuned the dipolar coupling constant to be a residual value of few Hz, which are well detectable by NMR spectroscopy.

Values of static dipolar coupling constant of two-spin systems in protein backbone:

- $D_{\text{CH}} = 47.9$ kHz
- $D_{\text{CC}} = 4.9$ kHz
- $D_{\text{NH}} = -24$ kHz
- $D_{\text{CN}} = -2.0$ kHz
- $D_{\text{HH}} = 10 - 22$ kHz
Residual Dipolar Coupling (RDC)

The residual dipolar coupling between two spins A and B are given by:

\[
\langle D_{AB} \rangle = -C(B_0) \left[ \chi_a (3\cos^2\theta - 1) + \frac{3}{2} \chi_r (\sin^2\theta \cos^2\phi) \right]
\]

where

\[
C(B_0) = S(B_0^2/15kT)[\gamma_A \gamma_B \hbar/(4\pi^2 r_{AB}^3)].
\]

\(\gamma_A\) and \(\gamma_B\) are gyromagnetic ratios of A and B.

\(r_{AB}\) is the distance between A and B.

So, \(B_0 \uparrow, D_{AB} \uparrow\)

S (order parameter) \(\uparrow, D_{AB} \uparrow\)
Alignment Media

- Phages (Pf1, fd, TMV) (Zweckstetter, JBNMR)
- Bicelles (Sanders & Schwonek, Biochemistry, 1992; Ottiger&Bax, JBNMR 1998)
- Polyacrylamide gels (Tycko, JBNMR; Grzesiek JBNMR; Chou, JBNMR)
- Paramagnetic tagging (Opella, Griesinger, Byrd)
- C12E5/hexanol (Ruckert&Otting, JACS 2000)
- Cellulose crystallites (Matthews, JACS, ~2000)
Alignment of Molecules in Anisotropic Solutions

The most-used media for RDC measurement are:
(a). Phospholipid bicelles and (b). Filamentous phage
$^{15}$N-IPAP HSQC for HN RDC values

Isotropic solution + 5.3 mg/ml Pf1

- $J_{NH}$
  - $-94.57$ Hz

- $G81$
  - $-94.45$ Hz

- $T69$
  - $-92.50$ Hz
  - $-95.12$ Hz

- $G20$
  - $-93.30$ Hz

- $T71$
  - $-116.83$ Hz
  - $-121.33$ Hz

- $J_{NH} + D_{NH}$
  - $-91.71$ Hz
3D HNCO for C’N RDC values

Isotropic solution + 5.3 mg/ml Pf1

- T71: $J_{C'N}$ -54.33 Hz
- K63: -51.58 Hz
- F4: -54.59 Hz
- I44: -52.82 Hz

- T71: $J_{C'N} + D_{C'N}$ -52.69 Hz
- K63: -56.45 Hz
- F4: -52.73 Hz
- I44: -50.30 Hz
Structure Refinement with RDC Restraints

\[
<D_{AB}>(\theta,\phi) = D_a \left[ \chi_a(3\cos^2\theta -1) + \frac{3}{2} D_r(\sin^2\theta \cos 2\phi) \right]
\]
Problems with higher molecular weights

- slower tumbling in solution → fast decay of NMR signal → poor signal-to-noise
- larger number of signals → signal overlap in NMR spectra

Linewidth $\Delta v_{1/2} = 1/\pi T_2$

<table>
<thead>
<tr>
<th>$\tau_c$</th>
<th>4 ns</th>
<th>8 ns</th>
<th>12 ns</th>
<th>25 ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW</td>
<td>8 kDa</td>
<td>16 kDa</td>
<td>24 kDa</td>
<td>50 kDa</td>
</tr>
</tbody>
</table>

![NMR spectra graphs](image)
1. Main relaxation source for $^1$H and $^{15}$N: dipole-dipole (DD) coupling and, at high magnetic fields, chemical shift anisotropy (CSA).

2. Different relaxation rates (line width) for each of the four components of $^{15}$N-$^1$H correlation.

3. The narrowest peak (the blue peak) is due to the constructive canceling of transverse relaxation caused by chemical shift anisotropy (CSA) and by dipole-dipole coupling at high magnetic field.

4. TROSY selectively detect only the narrowest component (1 out of 4).
Interference between DD and CSA Relaxation

- DD relaxation is field-independent. However, CSA relaxation \( \propto B_0^2 \), therefore at high magnetic fields, CSA relaxation can be comparable to DD relaxation, and the interference effect on relaxation can be observed.

(A) At High Magnetic Field
- (TROSY line-narrowing effect)

(B) At Low Magnetic Field
- (almost no TROSY line-narrowing effect)
TROSY Effect is Field Dependent and Motion Dependent

- Optimal field strength: 1 GHz for amide NH; 600 MHz for CH in aromatic moieties (500-800 MHz applicable).
Deuteration- NMR structural study of larger proteins

Deuteration is also an important techniques for NMR study of larger proteins (> 20kDa). It is achieved by raising the *E. coli.* in D₂O medium (NT$ 10,000 / 1L D₂O).

Because of the significantly lower gyromagnetic ratio of ²H compared to ¹H (γ[²H] / γ[¹H] = 0.15), replacement of protons with deuterons removes contributions to proton linewidths from proton-proton dipolar relaxation and ¹H-¹H scalar couplings.

The effect of deuteration is similar with that of TROSY and both techniques are frequently used for NMR study of larger proteins.
The Sensitivity and Resolution Gain by TROSY and Deuteration

$^2\text{H}^1\text{N}$-Gyrase-45 (45 kDa), 750 MHz

Paramagnetic Relaxation Enhancement (PRE)

Paramagnetic spin labeling

\[
MTSL = S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate.
\]

\[ R2^* = R2 + R2^\text{sp} \]

\[ I_{\text{ox}} \approx \frac{1}{R2^*} \quad \text{and} \quad I_{\text{red}} \approx \frac{1}{R2} \]

\[
\frac{I_{\text{ox}}}{I_{\text{red}}} = \frac{R2 \exp(-R2^\text{sp}t)}{R2 + R2^\text{sp}}
\]

\[
r = \left[ \frac{K}{R2^\text{sp}} \left( 4\tau_c + \frac{3\tau_c}{1 + \omega_h^2\tau_c^2} \right) \right]^{1/6}
\]

MTSL= S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate.
Spin Labeling on DNA by EDTA-Derivatized Deoxythymididine

![Diagram of DNA structure and sequence](image)
PRE Effects as $\text{Mn}^{2+}$-chelated
PRE Effects Observed as Spin Labeling at Either End of the DNA

Oligo1
5' d-CCTGCACAACACCATG
3' d-GGACGTGTTTGTGGAC

Oligo2
5' d-CACCTGCACAACACACC
3' d-GGGACGTGTTTGTGG

dT-EDTA is depicted in red for oligo1 and green for oligo2.
Specific SRY/DNA Interactions
IV. Cases of protein/nucleic acids complexes
The common DNA recognition motifs

Helix-turn-helix (HTH) domain

Zinc-finger (ZF) domain

By Dr. Song Tan
The common DNA recognition motifs

Winged-helix (WH) domain

Chemical shift change plot based on NMR titration data

Chemical shift differences ($\Delta \delta_{\text{binding}}$) were calculated for each individual amide group using the formula:

$$\Delta \delta_{\text{binding}} = \frac{\sqrt{\Delta \delta_{\text{HN}}^2 + (\Delta \delta_{\text{N}}/5)^2}}{2}$$

Residue
The common RNA recognition motifs

RNP/RRM domain  
RNP: ribonucleoprotein  
RRM: RNA-recognition motif
The common RNA recognition motifs

The KH domains

KH: K-homology motif

(a) A view of the β-sheet face of the KH domain.
(b) A 90° rotation from the view shown in (a).

The invariant Gly–X–X–Gly segment is shown in yellow, and the variable loop in red.
NMR spectroscopy as a tool for secondary structure determination of large RNAs.

three intermolecular helices

four-way helical junction
1 G-U
2 U-U
21 Watson-Crick
The structure of HCV IRES domain II

Dependence of RDC values on the orientation of the interdipolar vector (C-H) and the alignment tensor

\[ D_{\text{CH}}(\theta, \phi) = D_{\text{a,CH}} \left\{ (3 \cos^2(\theta) - 1) + \frac{3}{2} R(\sin^2(\theta) \cos^2(\phi)) \right\} \]

Refinement of the HCV IRES domain II structures calculated with by the use of different sets of RDCs

Rmsd = 7.48 Å
Rmsd = 5.79 Å
Rmsd = 2.18 Å

Structure of the HIV-1 Nucleocapsid protein with SL3 Ψ-RNA recognition element

Fig. 1. (A) Amino acid sequence of the HIV-1<sub>N</sub>NC protein showing the zinc-binding modes of the two CCHC-type zinc knuckles. Residues that contact the RNA in the NC-SL3 complex are denoted by open letters; asterisks denote residues involved in intermolecular hydrogen bonding. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; T, Thr; V, Val; and W, Trp. (B) Nucleotide sequence and secondary structure of the HIV-1<sub>N</sub>Ψ-sequence (12). The dimer initiation and major splice donor sites are labeled DIS and SD, respectively, and the gag initiation codon (AUG) is given in open letters. (C) Sequence of the RNA construct used in our studies.
Structure of the HIV-1 Nucleocapsid protein with SL3 \( \Psi \)-RNA recognition element

Fig. 2. (A) Portion of the 800-MHz 2D NOE spectrum obtained for the NC-SL3 complex in \( D_2O \) solution. Intermolecular cross-peaks involving the aromatic protons of Phe\(^{16}\) (F1 knuckle) and Trp\(^{37}\) (F2 knuckle) are labeled. Strong intramolecular Trp\(^{37}\)-Phe\(^{16}\) cross-peaks (labeled) are indicative of interknuckle packing in the complex. (B) Selected strips from the 800-MHz 3D \( ^{13}C \)-filtered, \( ^{12}C \)-detected HMQC-NOESY data obtained for the NC-SL3 complex in \( D_2O \) showing unambiguously assigned intermolecular NOE cross-peaks associated with the Ile\(^{24}\)-\( \delta CH_3 \), Ala\(^{25}\)-CH\(_3\), and Lys\(^{26}\)-H\( \alpha \) protons.
The best fit superposition and space-filling representation of the SL3 RNA in the NC-SL3 complex.
Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d

(a) The ensemble of best 20 structures superposed on backbone heavy atoms in ordered regions of the protein and RNA. (b) Ribbon representation of a single structure with the addition of green side chains for the zinc-coordinating ligands. (c) Backbone superposition of the structure ensembles of fingers 1 and 2. Finger 1 (Arg153-Phe180) is dark blue (backbone), green (zinc coordinating side chains) and red (intercalating aromatic rings); the bound RNA (U6, A7, U8, U9) is orange. The corresponding colors for finger 2 are light blue, yellow, pink and yellow.

5’-UUUUUUUUUUU-3’
Comparison between nucleolin RBD12-sNRE complex and the other RBD-RNA complexes

RBD: RNA binding domain
NRE: nucleolin recognition element
Thank you