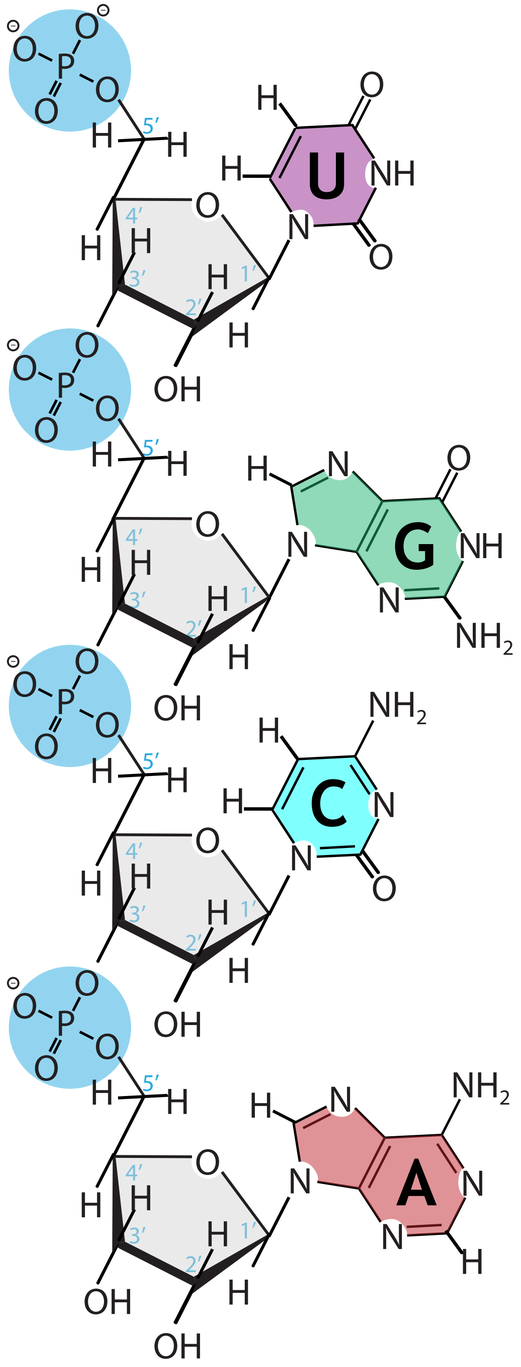
Structural Characterization of Large RNAs from HIV-1 using NMR

Jan Marchant

University of Maryland, Baltimore County

RNA



Polymeric chain of nucleotides

Pentose sugar ribose

Connected 3′ -> 5′ by negatively charged phosphate backbone

One of four nitrogenous bases attached at the 1′ position

Isotopic labeling is easy if precursors are available

- Residue-specific via solid phase synthesis from phosphoramidite monomers (size-limited)
- Uniform labeling by residue type by in vitro transcription from NTPs

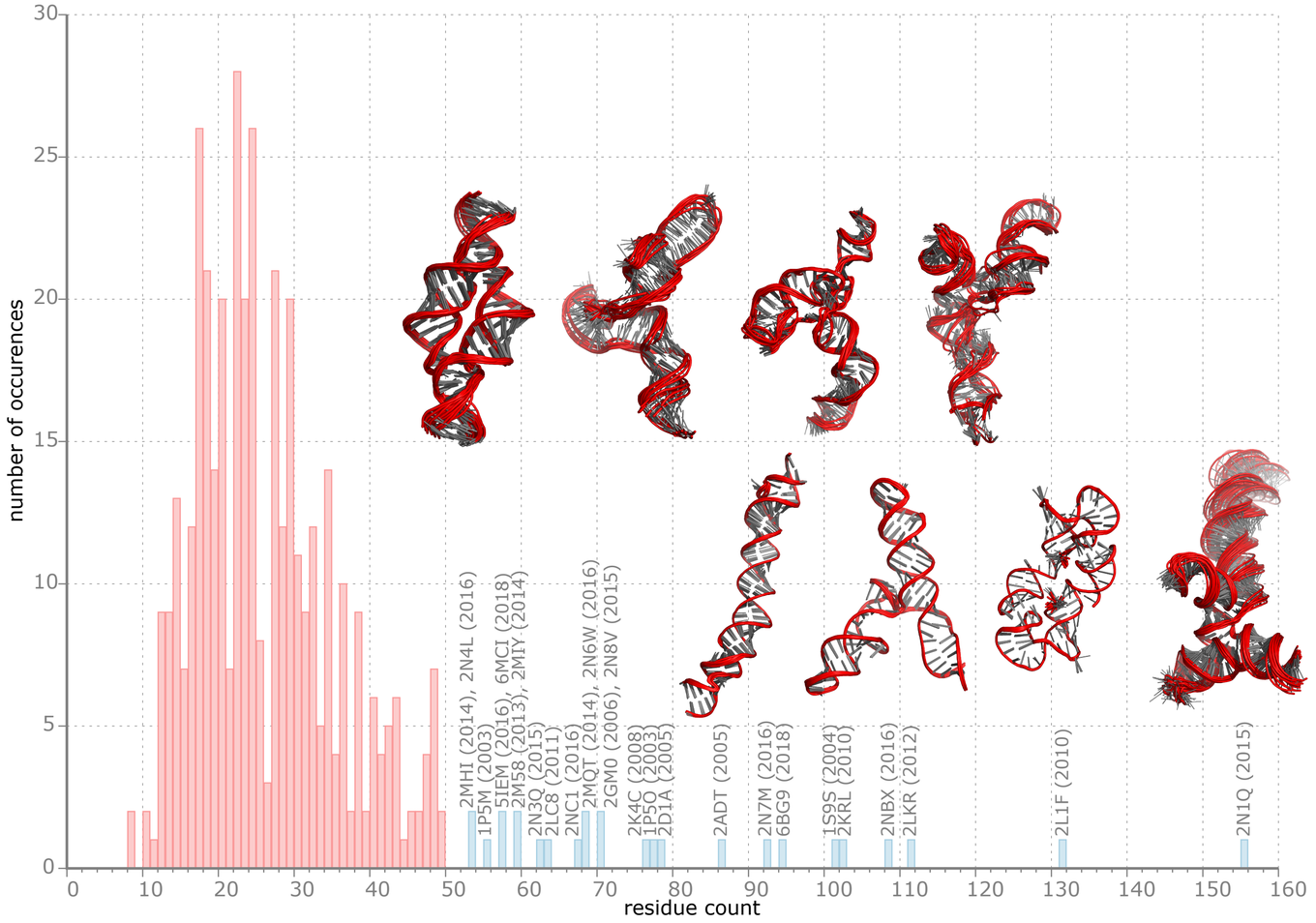
Large suite of homonuclear and heteronuclear NMR experiments allowing for sequential chemical shift assignment and acquisition of constraints for structure generation

General paucity of RNA structural information

In the PDB: 133,000 protein structures, 2200 protein:RNA complexes, 1400 RNA structures

NMR well-represented in RNA structures - 35% of depositions vs. 8% for proteins

Few NMR structures over ~50 nucleotides - median 24 nucleotides, average ~28 nucleotides



As of January 2020

Major challenges

Limited chemical shift dispersion / overlap

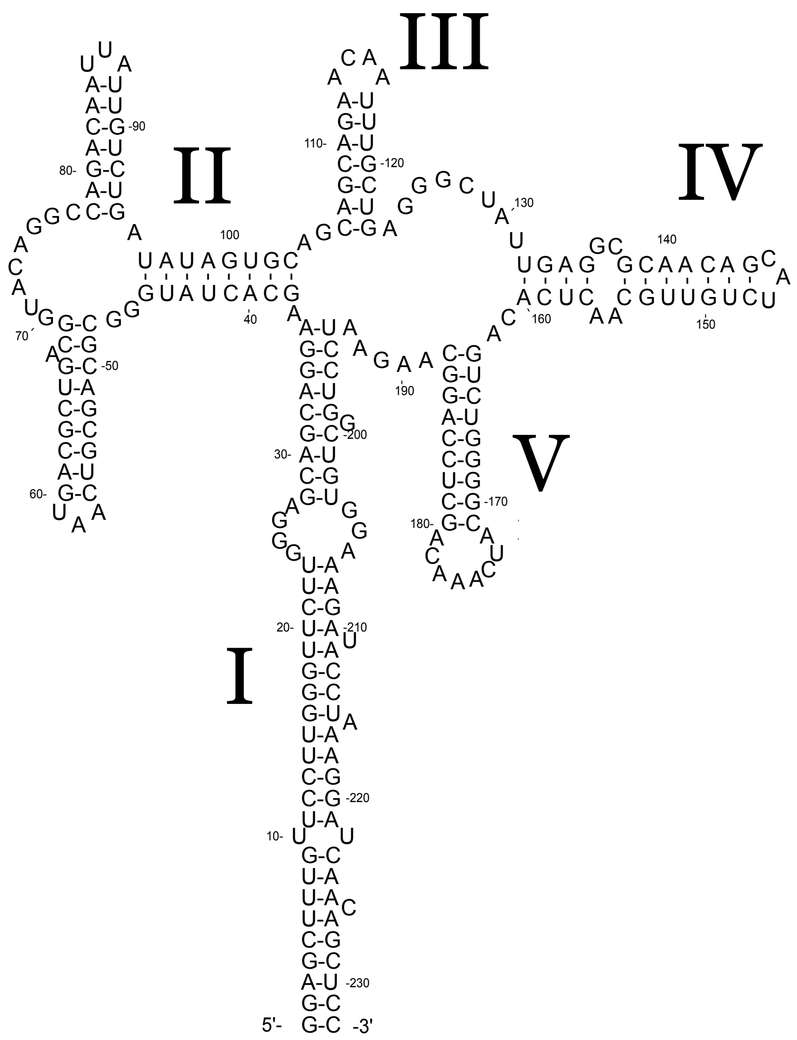
Difficulty applying heteronuclear correlation experiments to larger systems

Lack of NOEs between secondary structure elements

Isotope labeling is essential for addressing these challenges

Model system - HIV-1 RRE

232 nucleotide, ~75 kDa



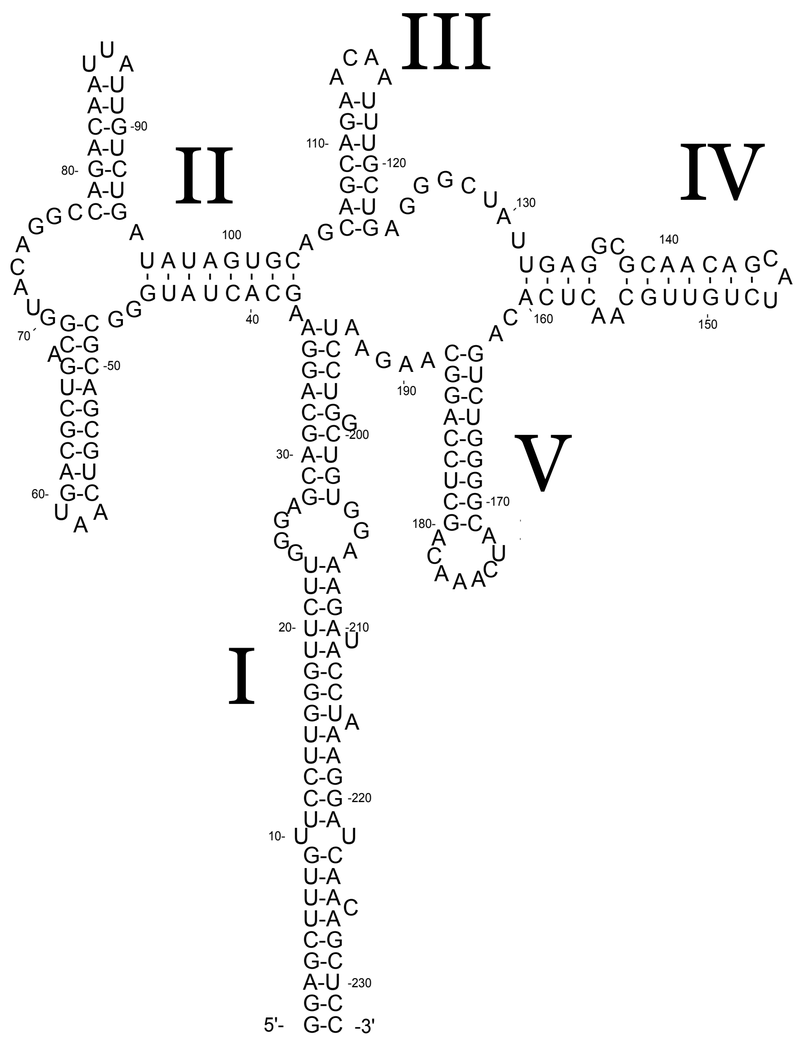
Generally single stranded

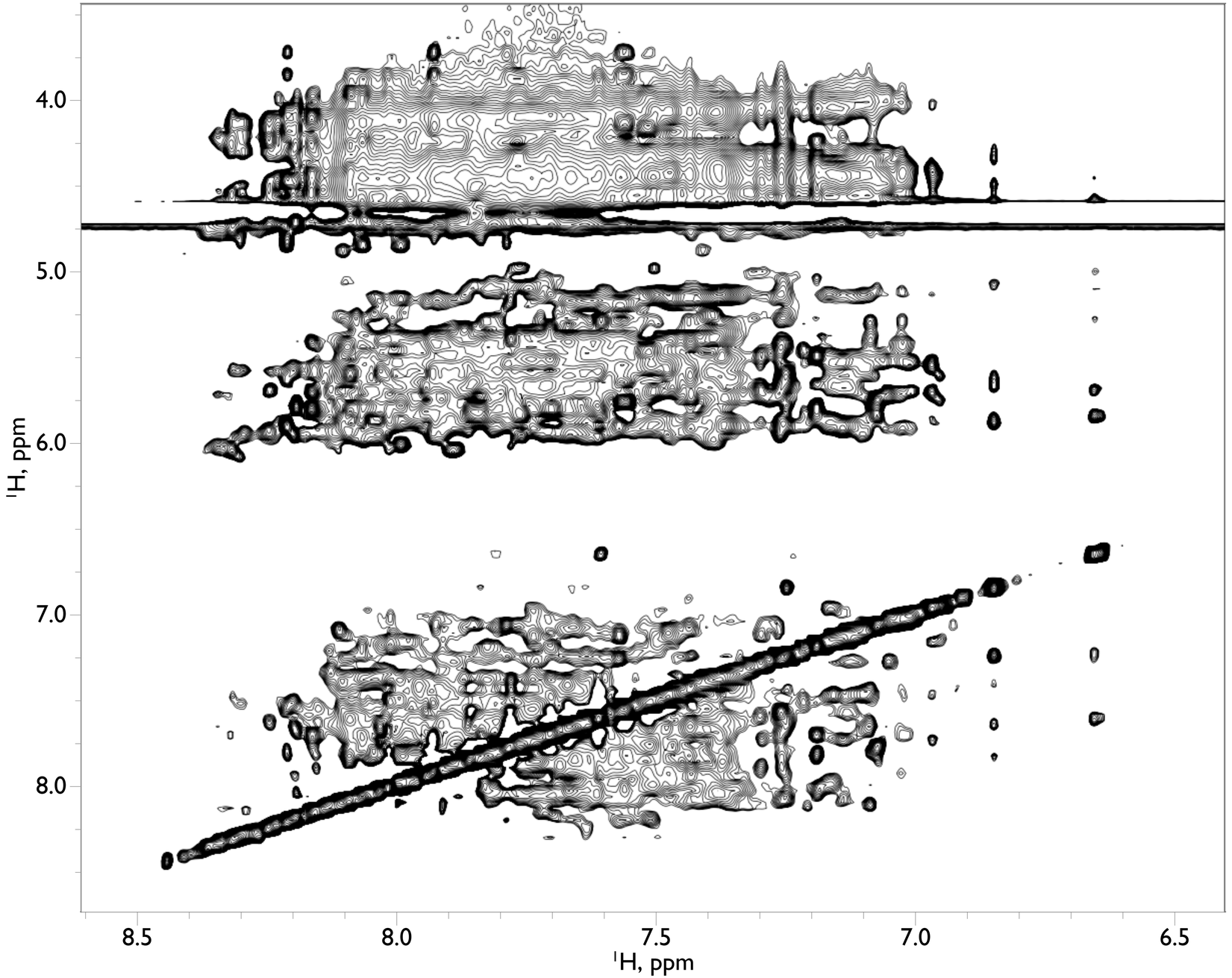
Hydrogen bonding between basaes leads to stable base pairing

Self-complementary sequences allow formation of A-form double helices

Self-complementary sequences allow formation of A-form double helices separated by bulges and large internal loops

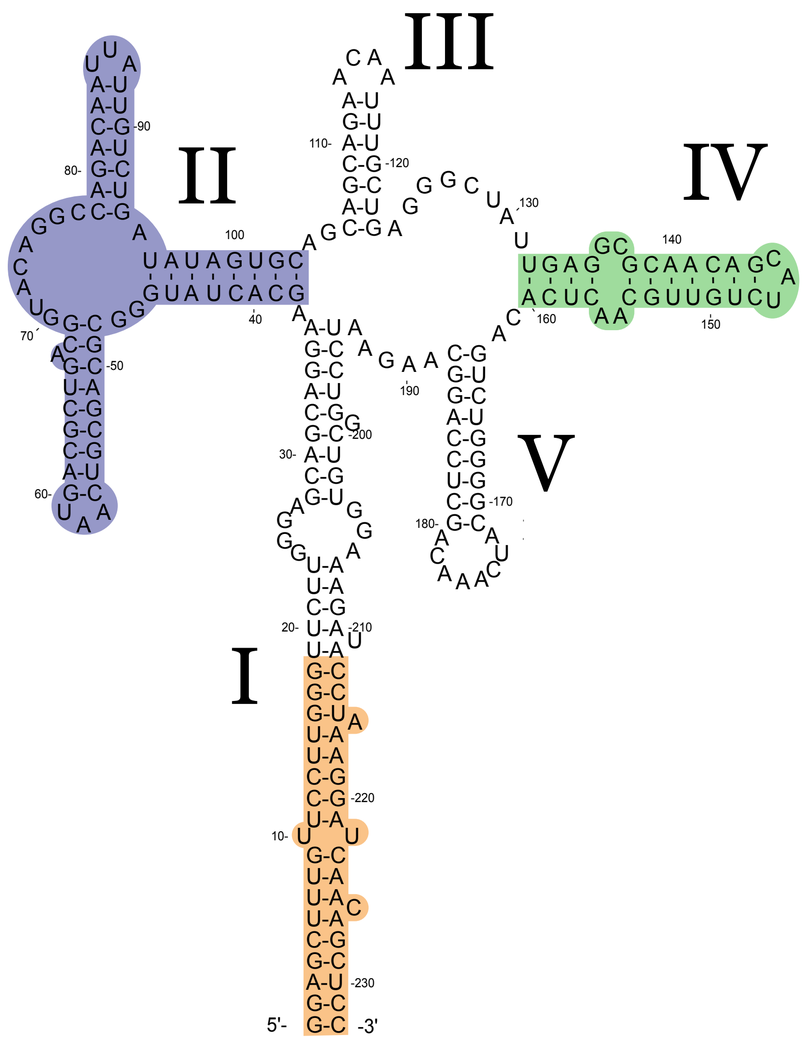
Challenge 1: signal overlap

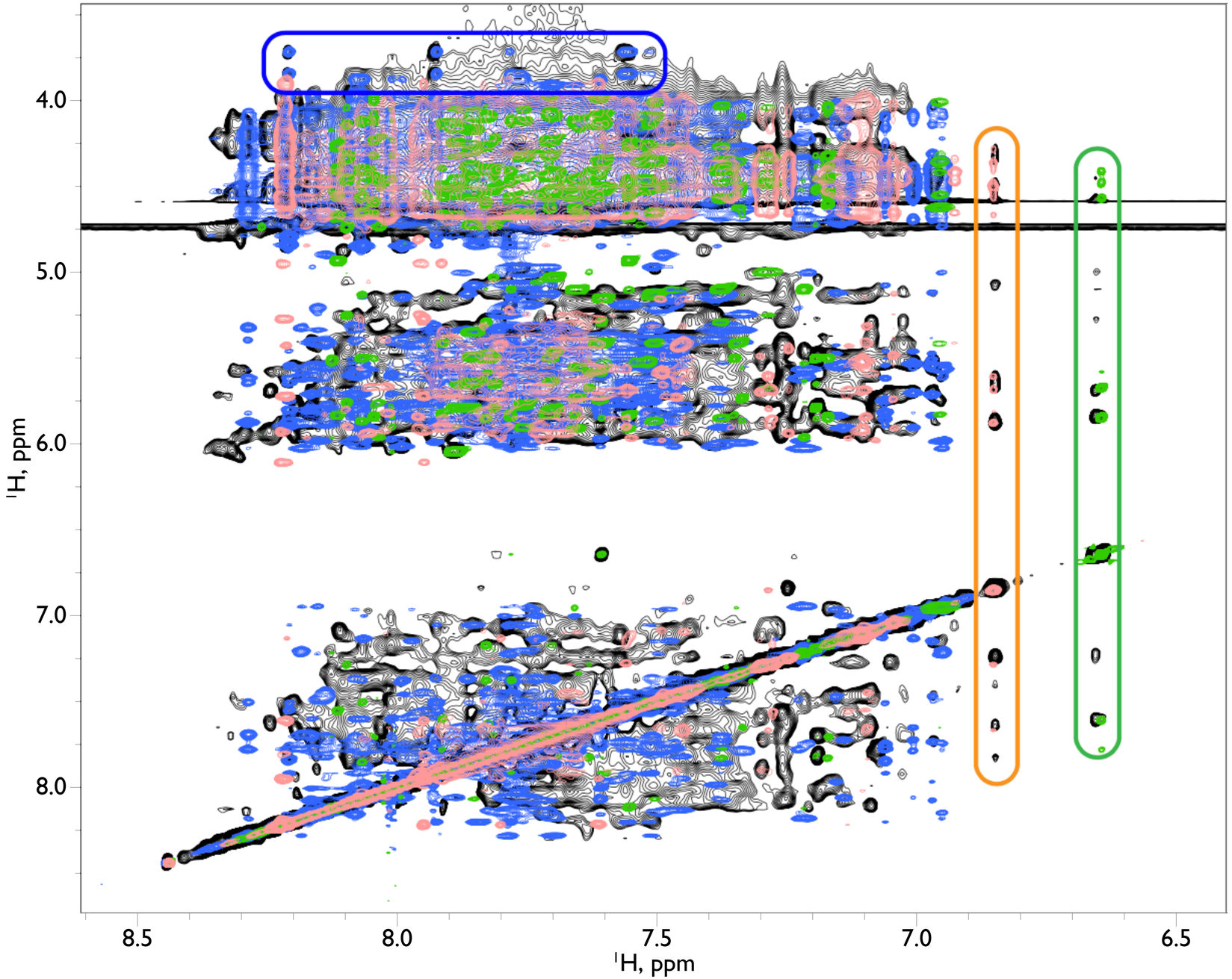




Challenge 1: signal overlap

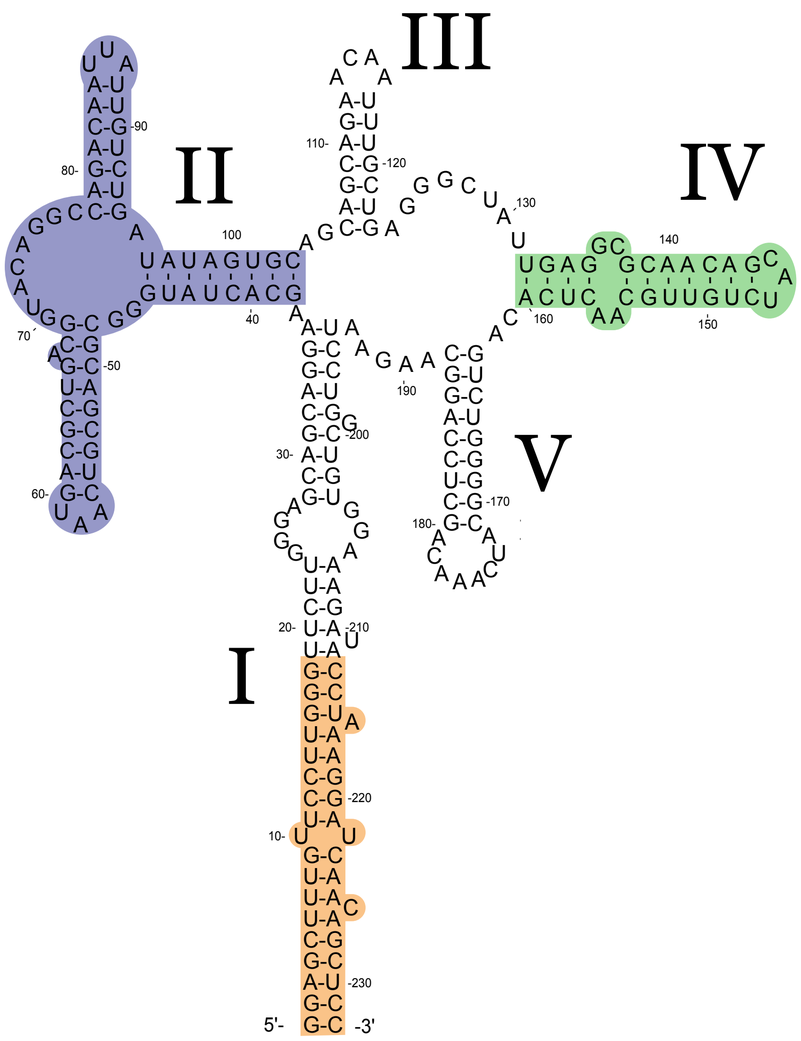
"Divide-and-conquer"

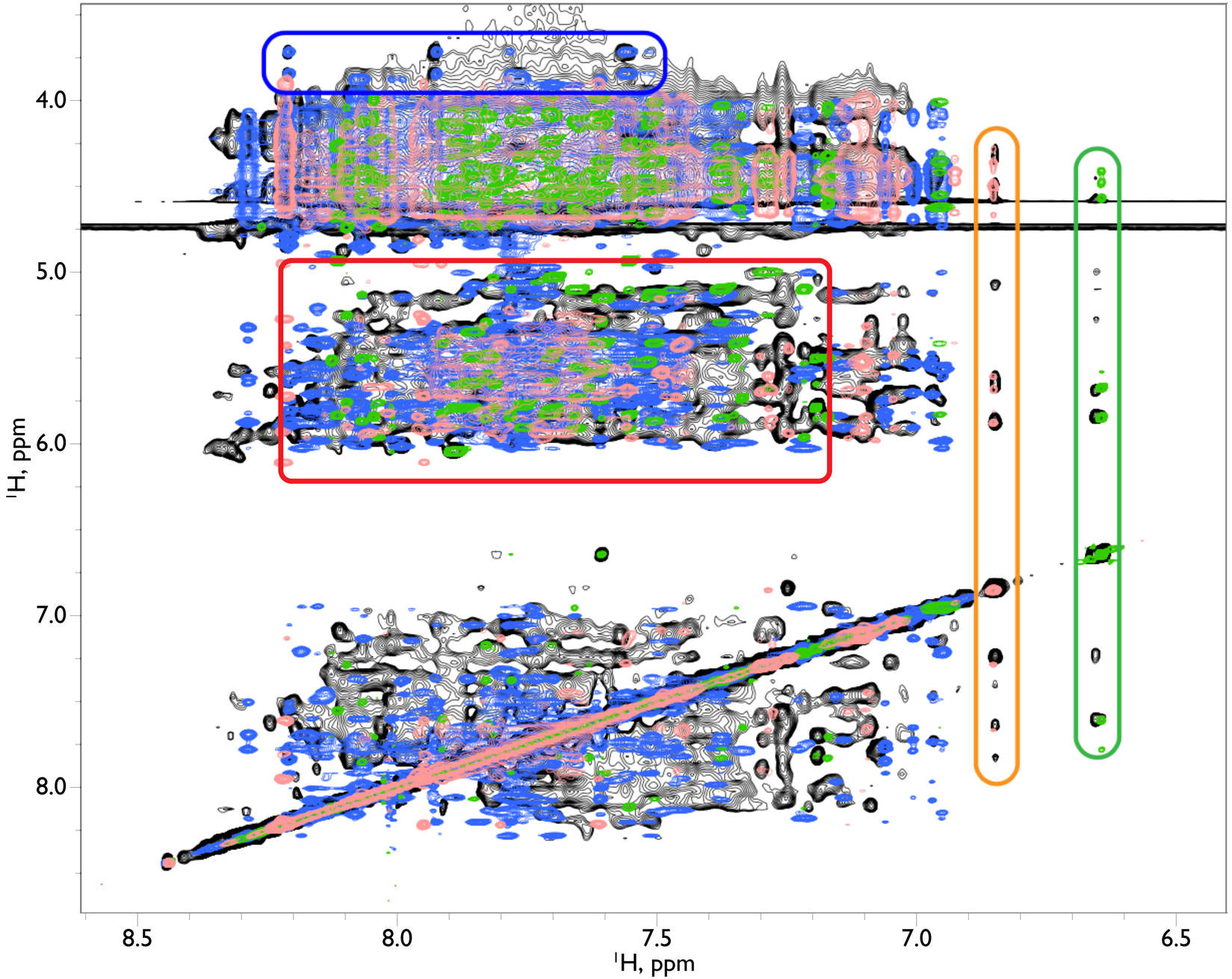




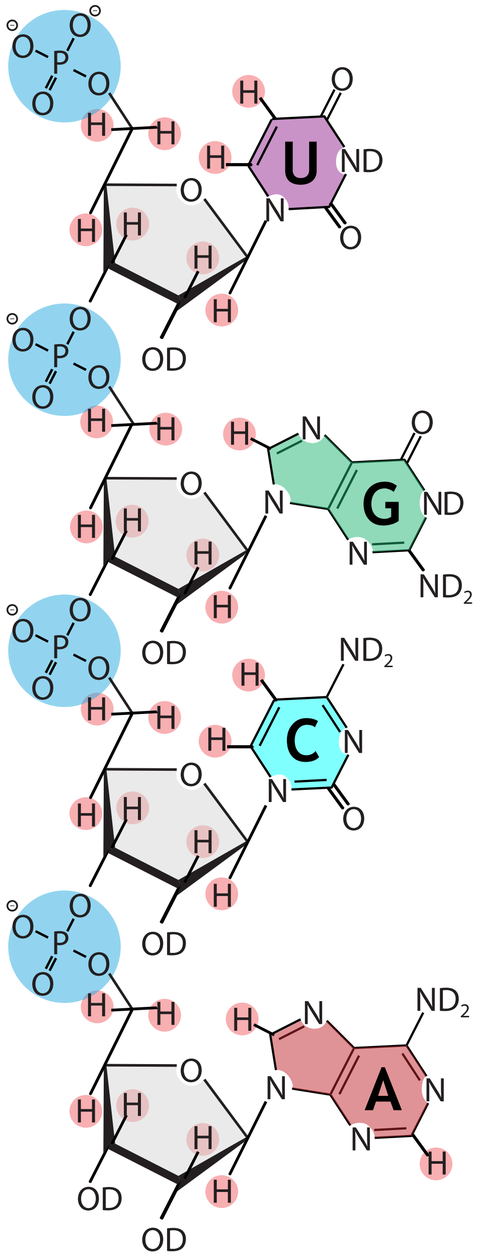
Challenge 1: signal overlap

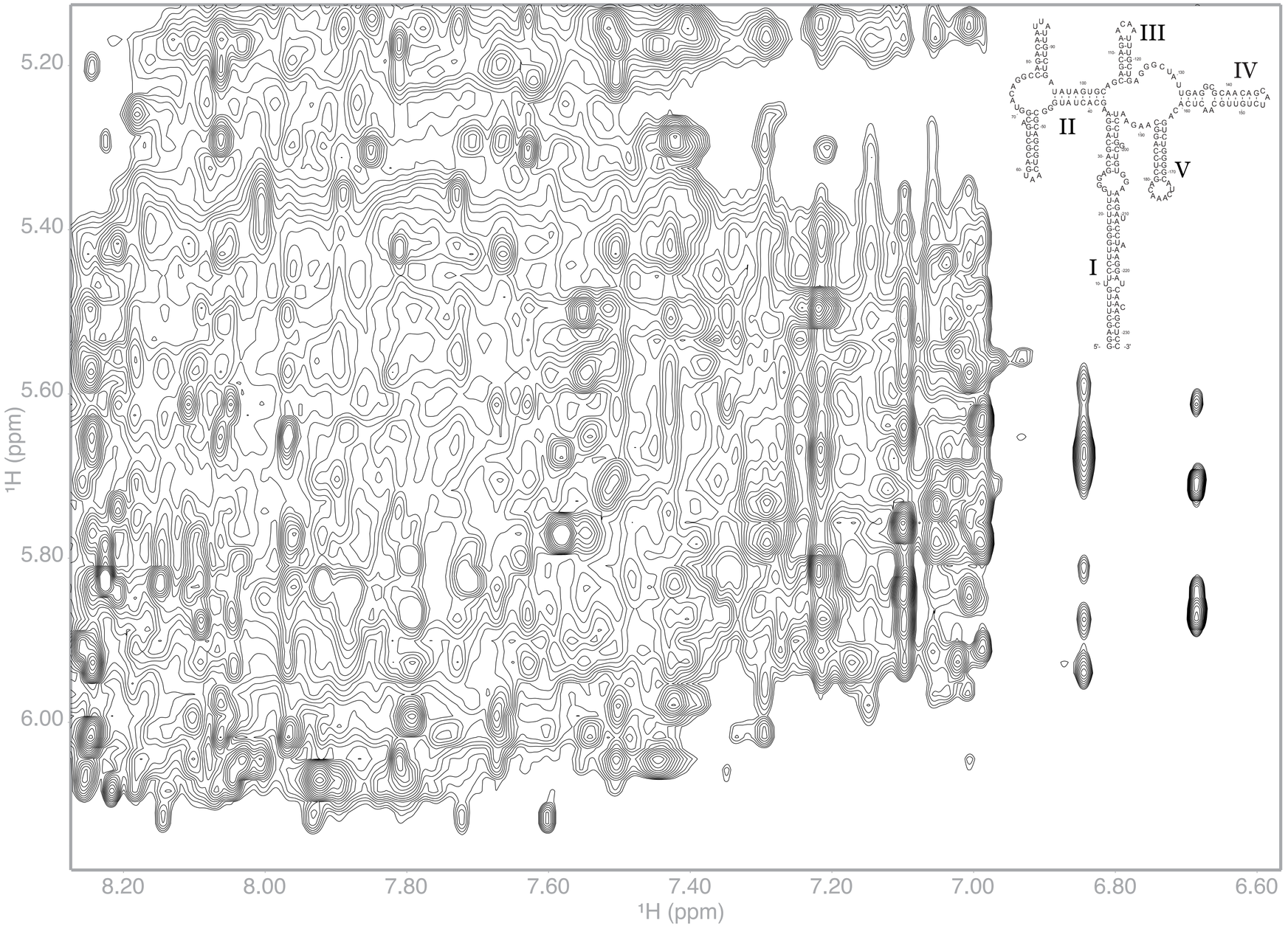
"Divide-and-conquer"





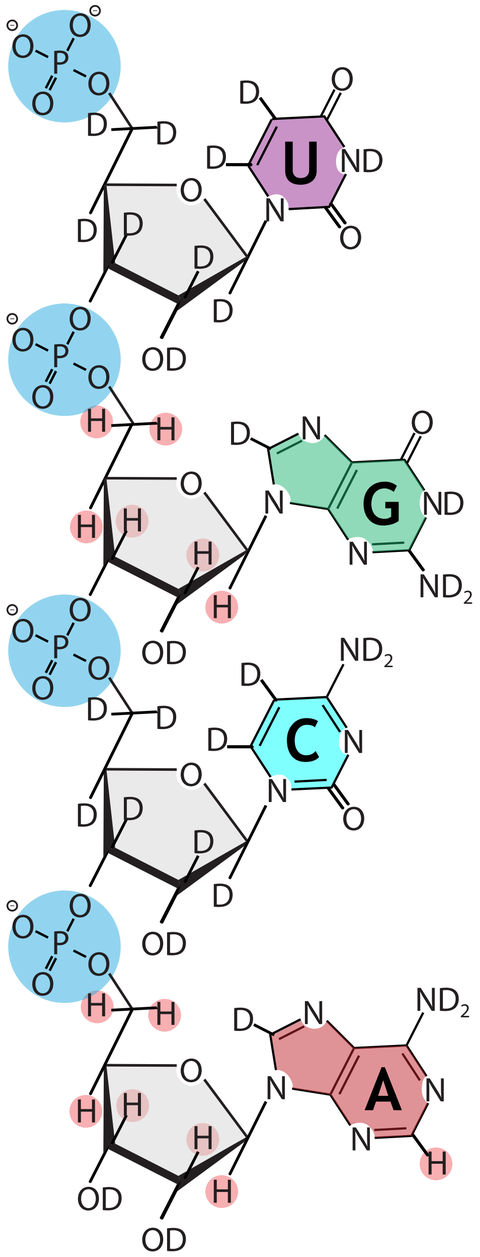
Challenge 1: signal overlap

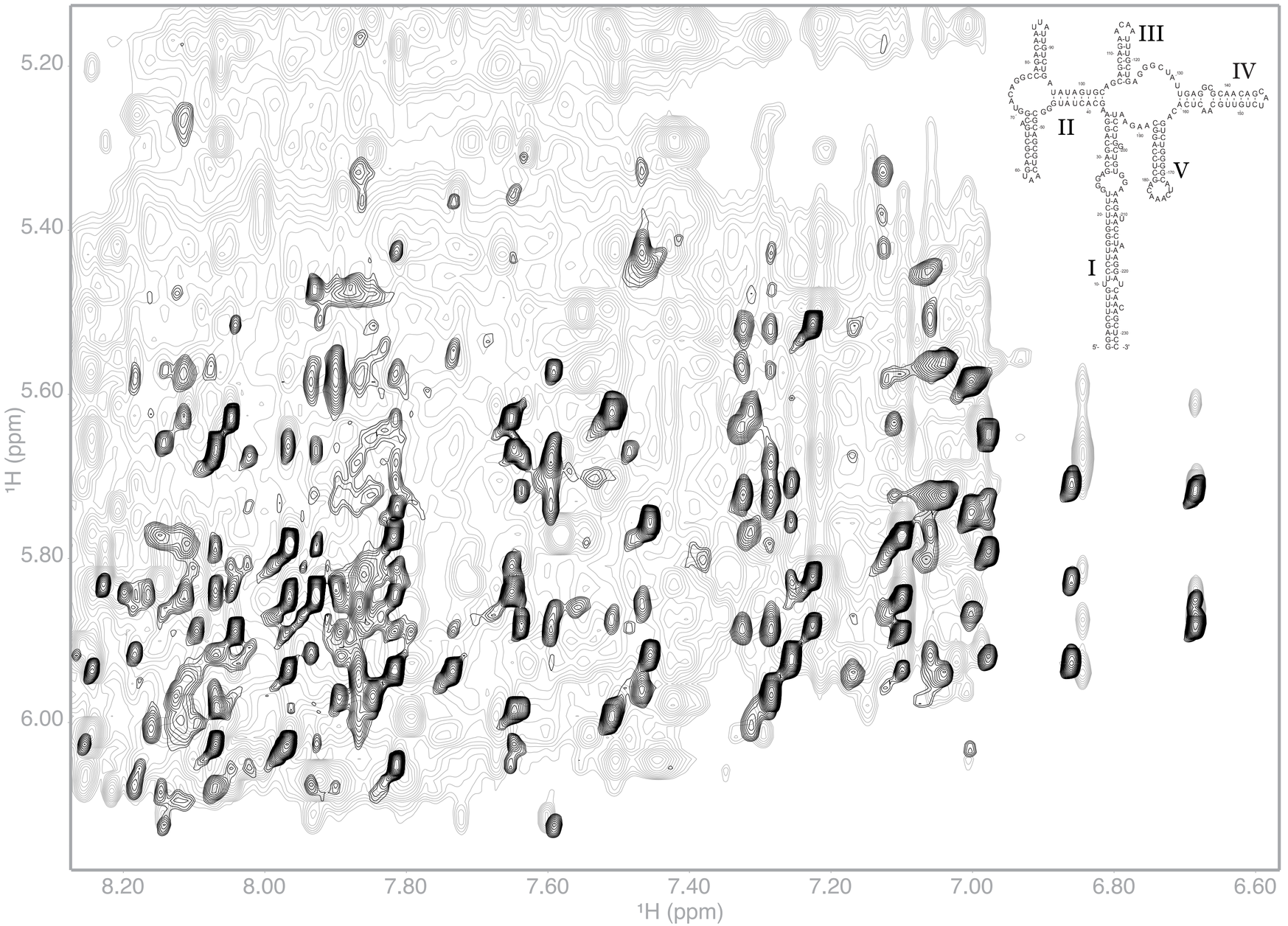




Challenge 1: signal overlap

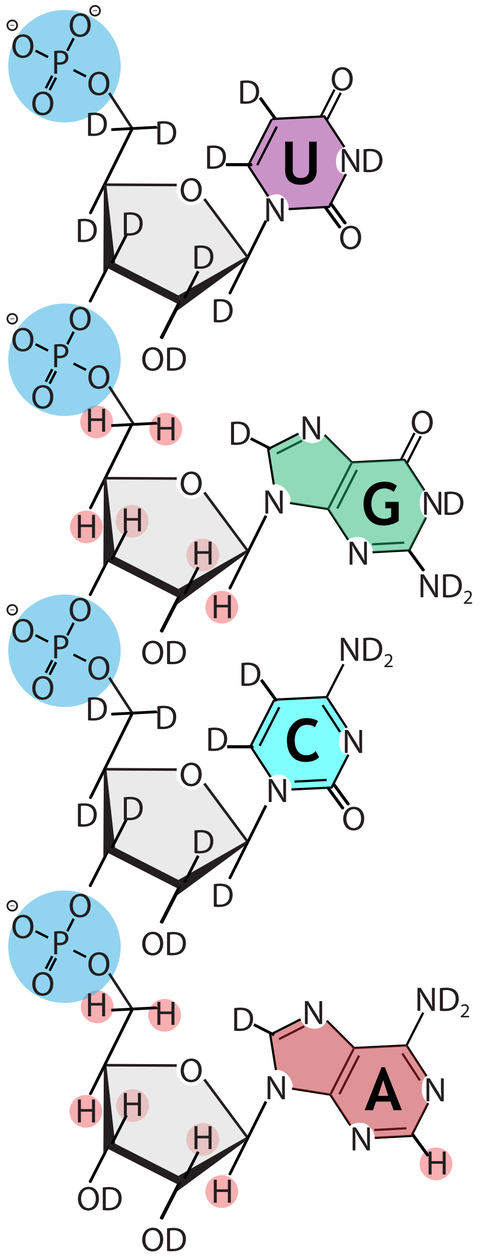
Deuteration

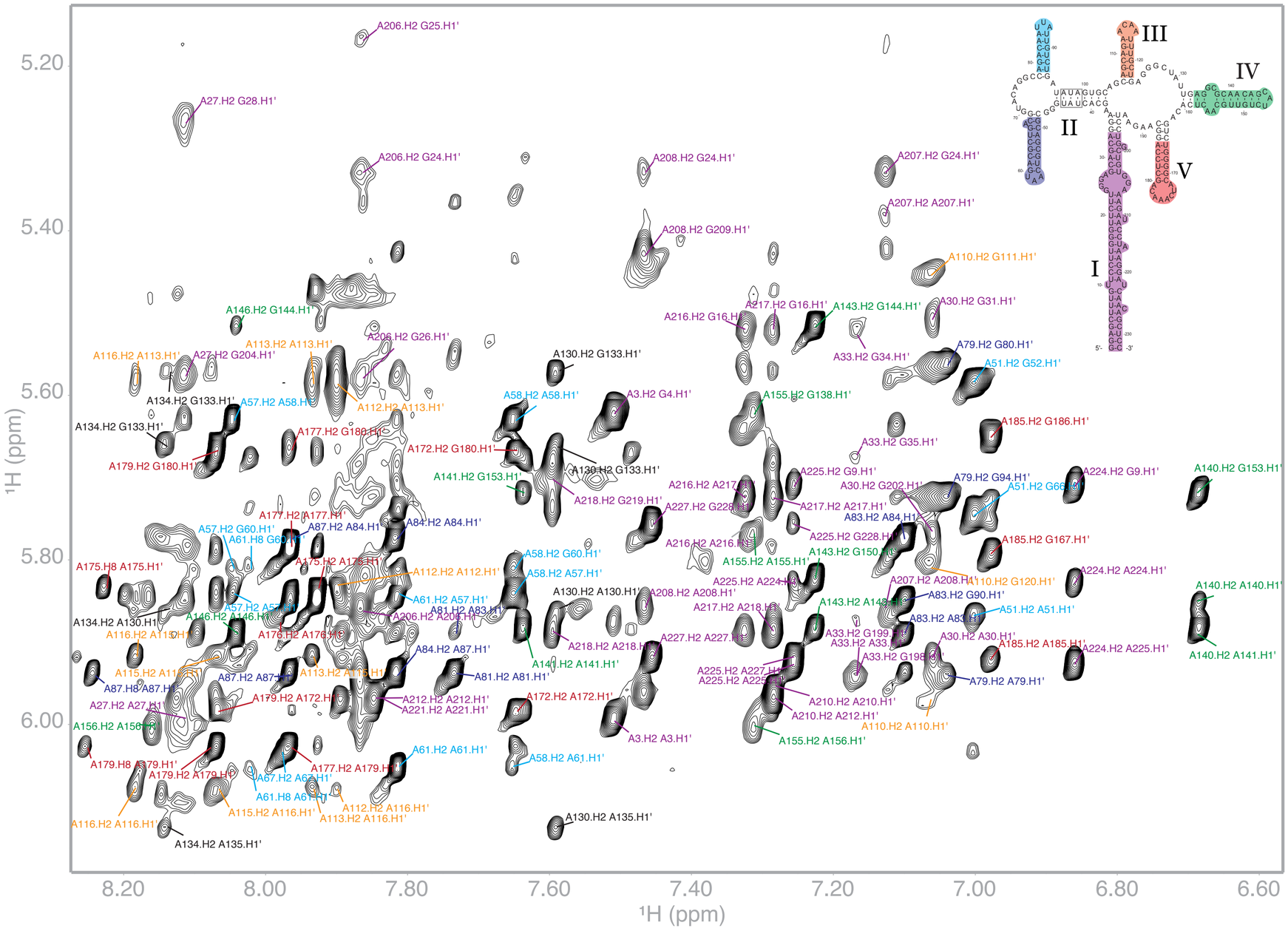




Challenge 1: signal overlap

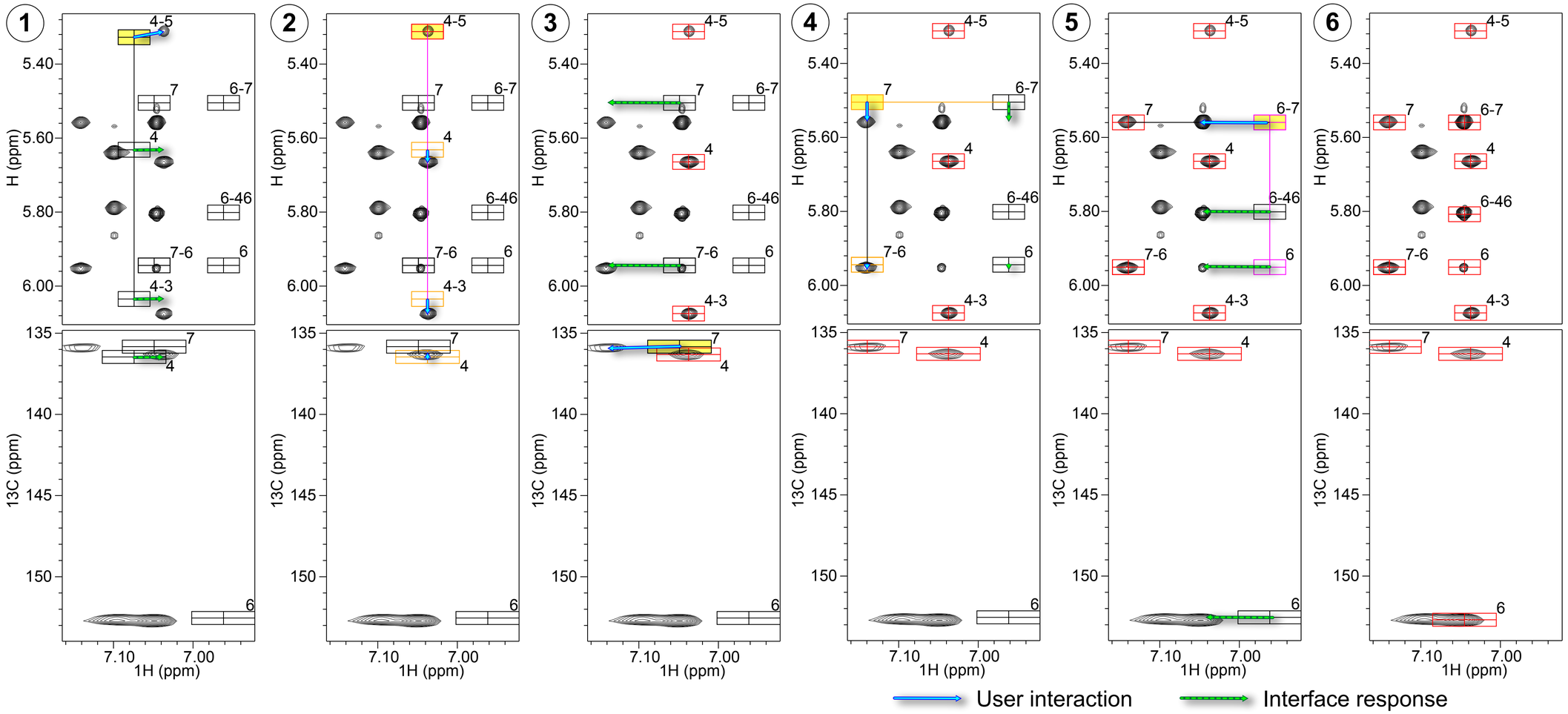
Deuteration





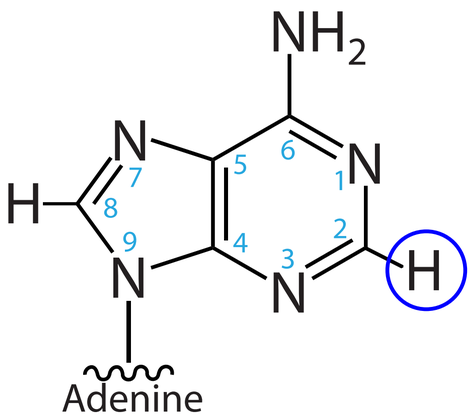
Challenge 1: signal overlap

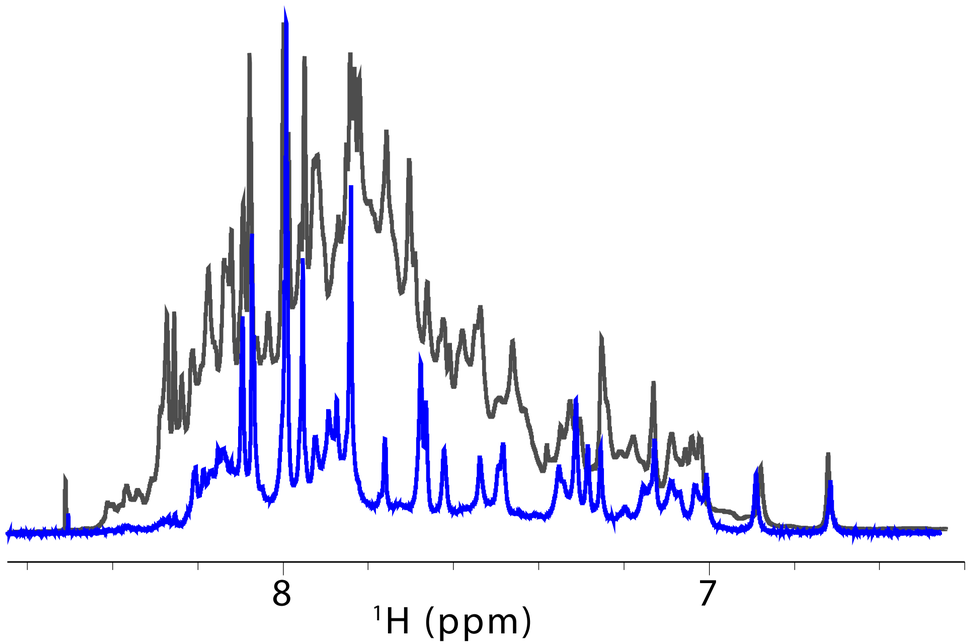
Computational tools greatly facilitate analysis

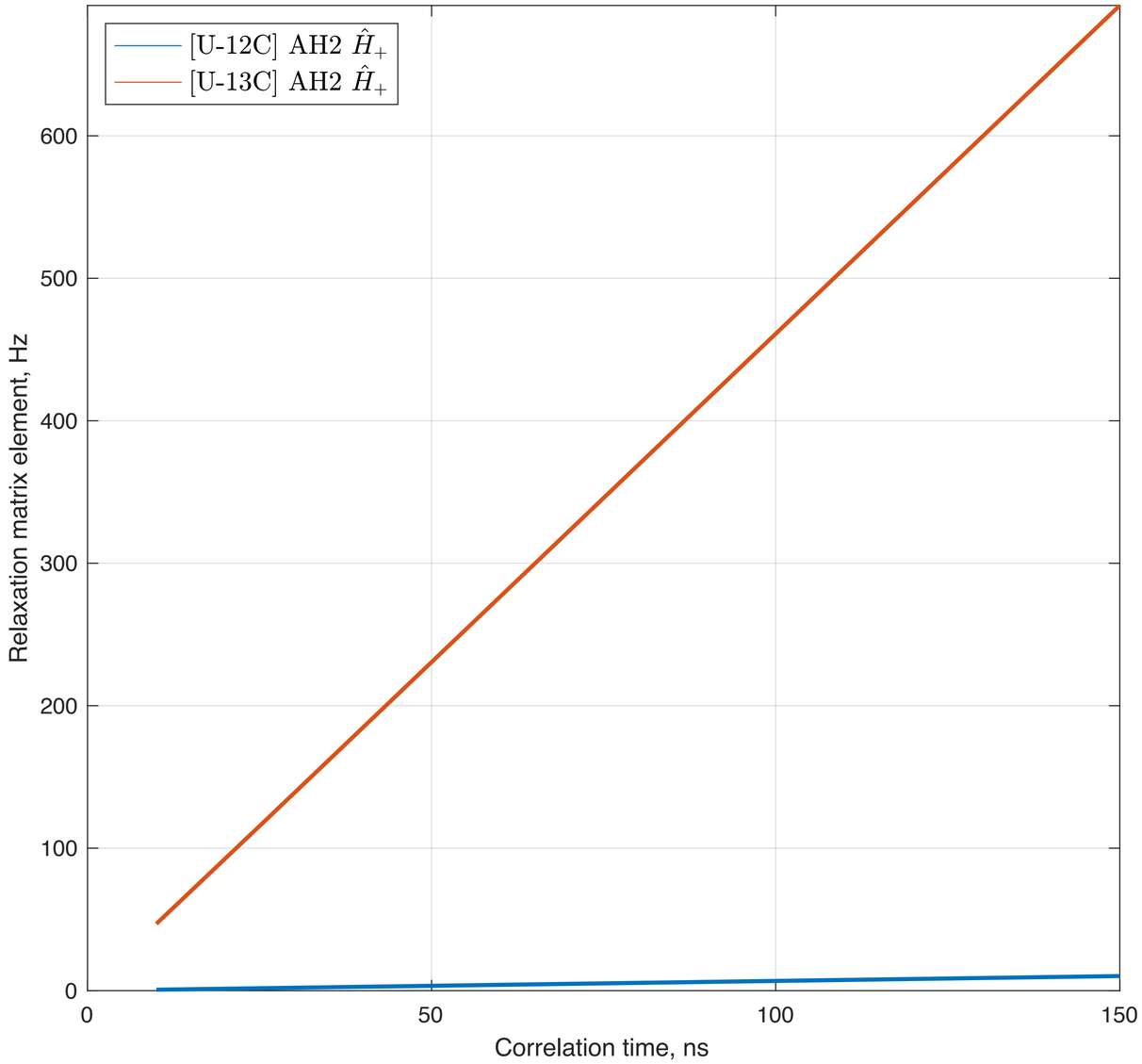


Marchant, Summers, Johnson, J. Biomol. NMR 2019

Challenge 2: heteronuclear labeling

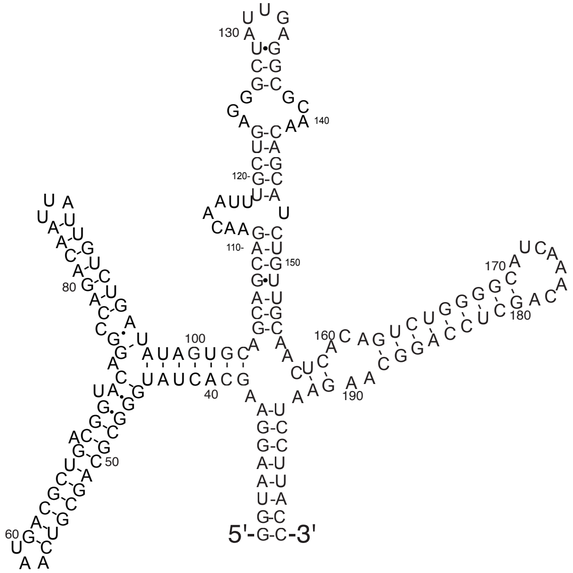






All simulations courtesy of Spinach: Hogben et al., JMR 2011

Challenge 2: heteronuclear labeling



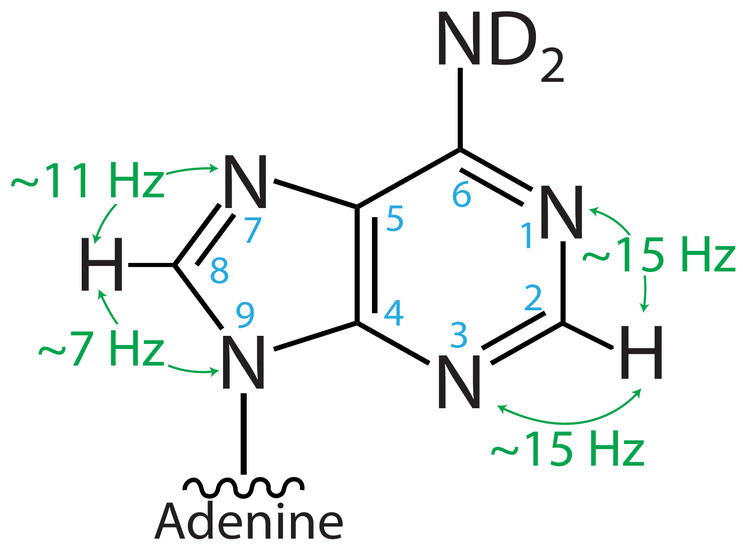
173 nt

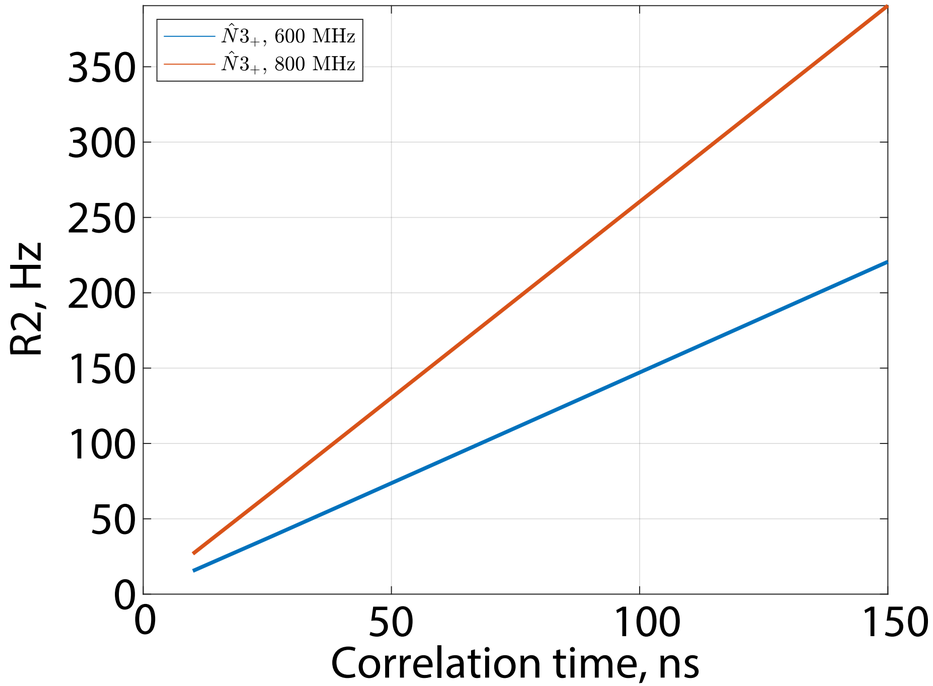
A28Cr

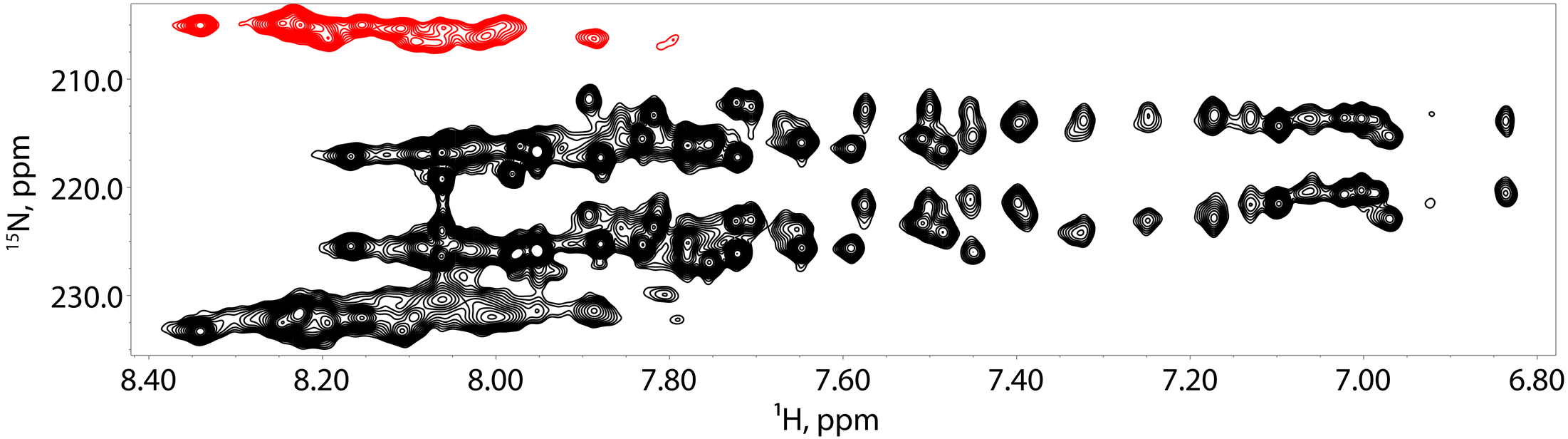
[U-13C] A

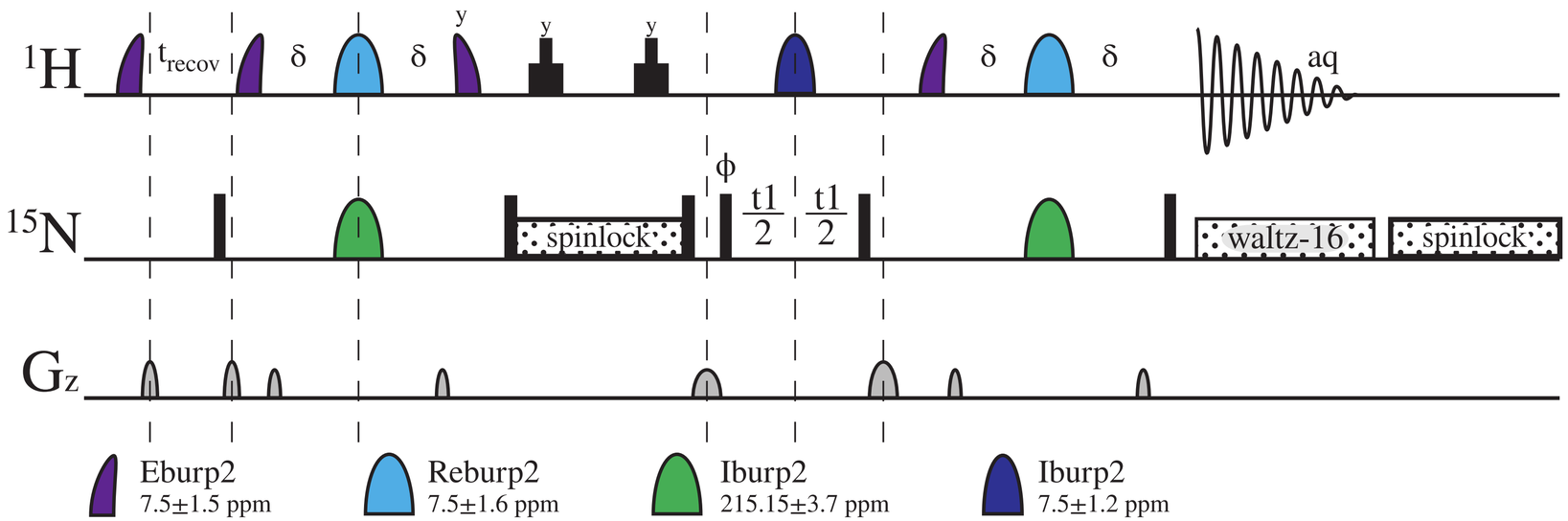
8.5 8.0 7.5 7.0 [ppm]

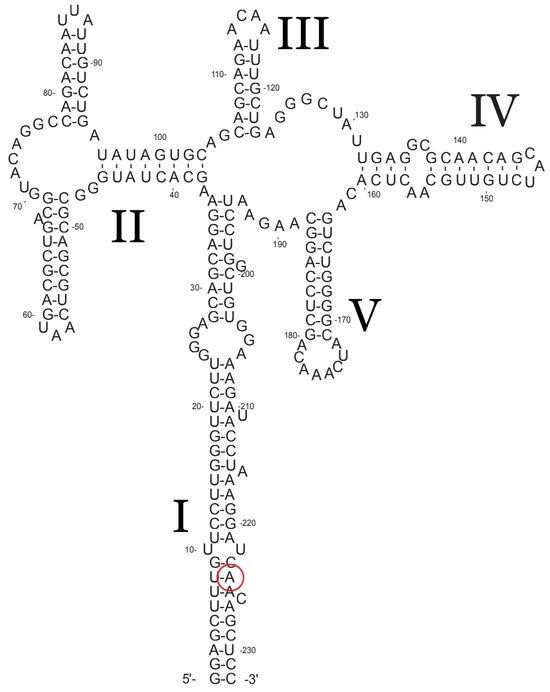
Challenge 2: heteronuclear labeling

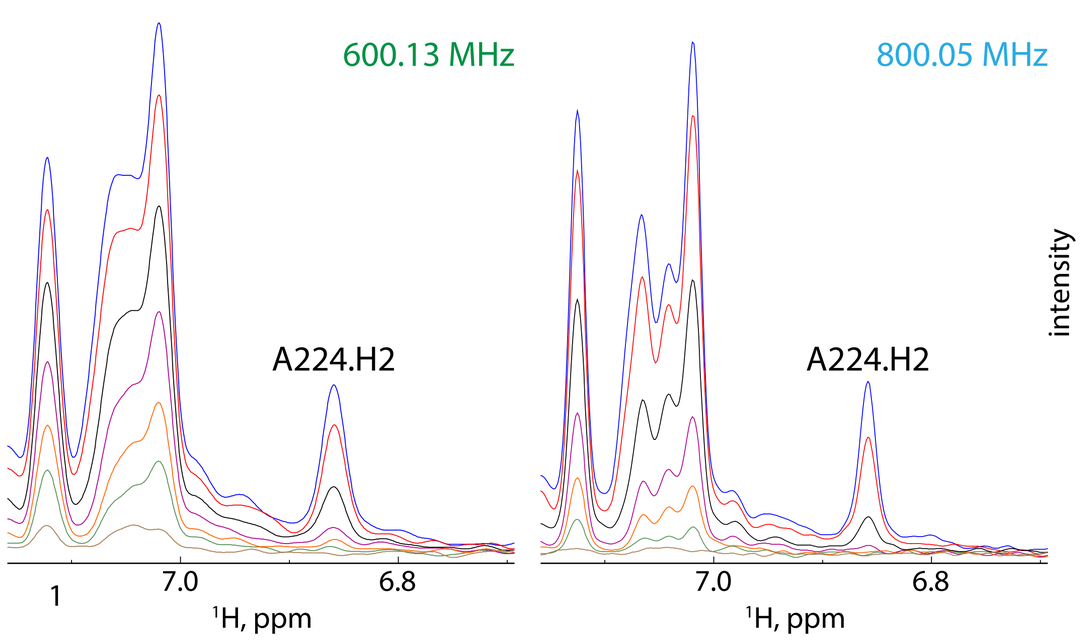


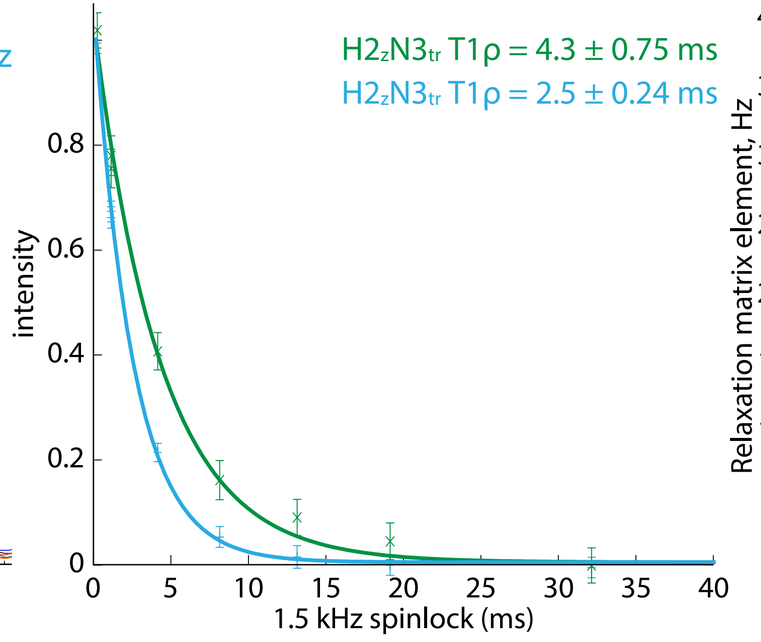


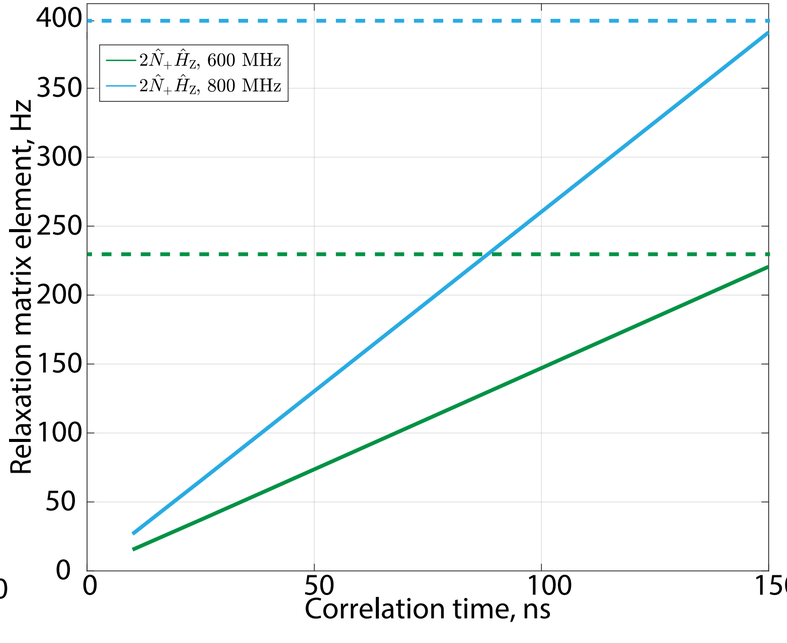


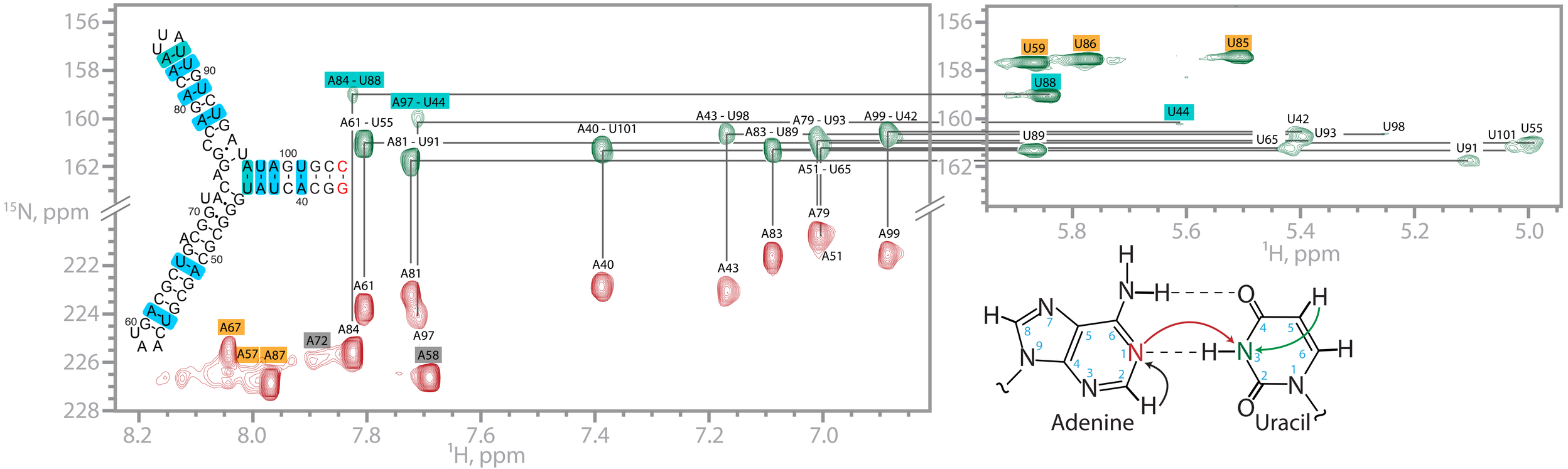

^{1}H
t_{recov}
δ
δ
y
y
y
δ
δ
aq
ϕ
^{15}N
spinlock
t1/2
t1/2
waltz-16
spinlock
Gz
Eburp2
7.5±1.5 ppm
Reburp2
7.5±1.6 ppm
Iburp2
215.15±3.7 ppm
Iburp2
7.5±1.2 ppm


I
II
III
IV
V


600.13 MHz
800.05 MHz
A224.H2
A224.H2
7.0
6.8
7.0
6.8
^{1}H, ppm
^{1}H, ppm


H2zN3tr T1ρ = 4.3 ± 0.75 ms
H2zN3tr T1ρ = 2.5 ± 0.24 ms
intensity
0
0.2
0.4
0.6
0.8
0
5
10
15
20
25
30
35
40
1.5 kHz spinlock (ms)

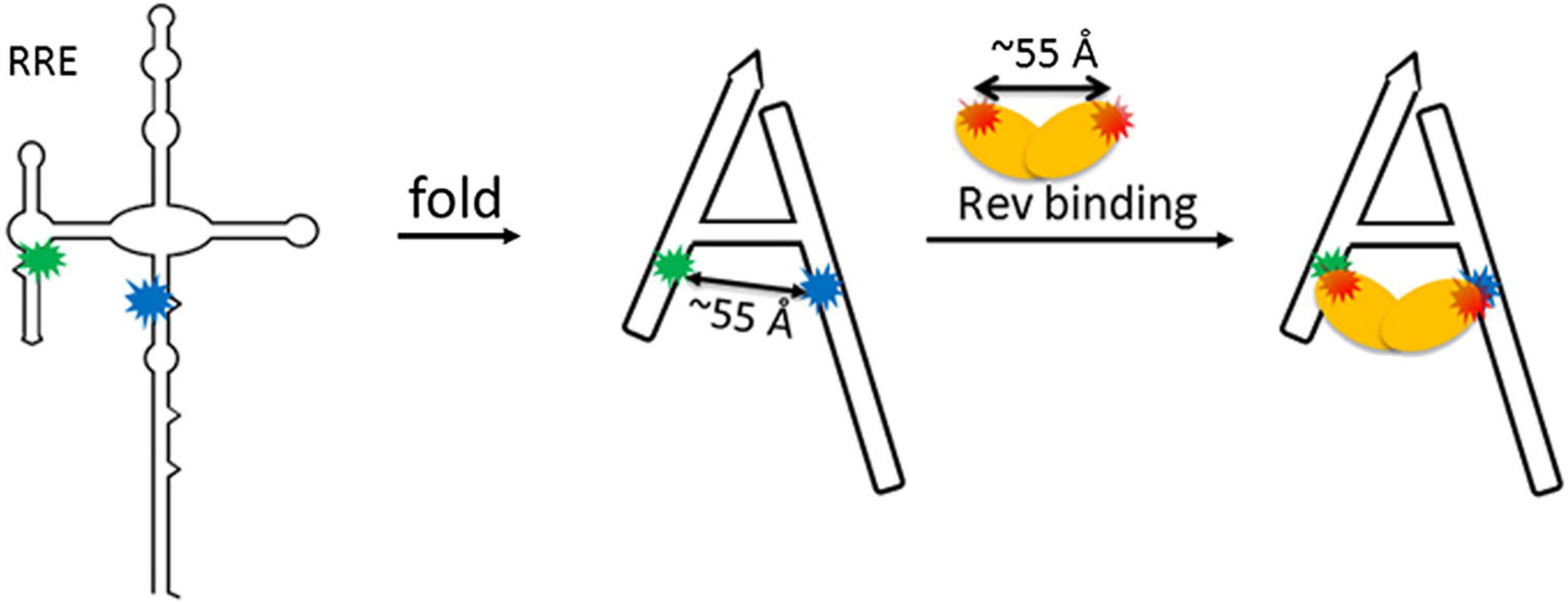

$2\hat{N}_+\hat{H}_Z$, 600 MHz
$2\hat{N}_+\hat{H}_Z$, 800 MHz
Relaxation matrix element, Hz
0
50
100
150
200
250
300
350
400
0
50
100
150
Correlation time, ns



Dallmann et al., Angew. Chem. 2013

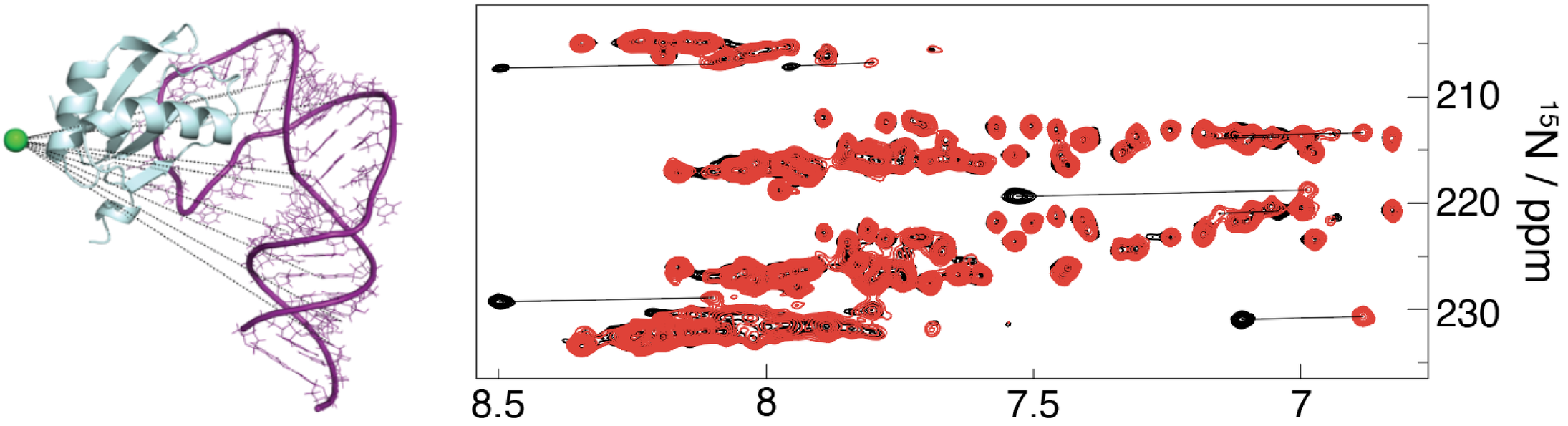
Challenge 3: relative orientation of secondary structure elements

Global features such as the relative orientation of distinct secondary structure elements remain challenging to obtain



Challenge 3: relative orientation of secondary structure elements

Pseudocontact shifts via interaction with paramgnetically tagged U1A protein:



Strickland et al., J. Am. Chem. Soc. 2019

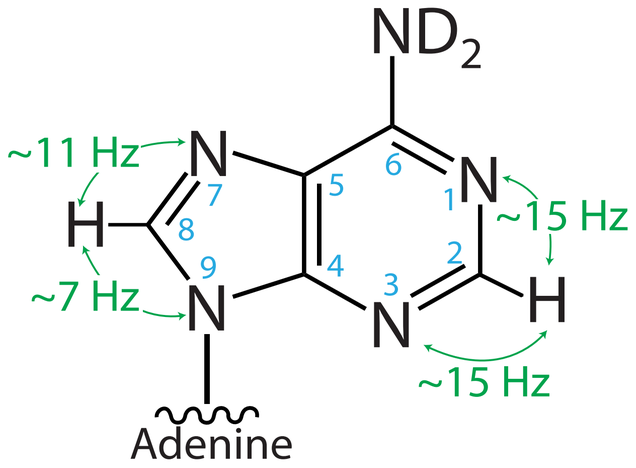
Challenge 3: relative orientation of secondary structure elements

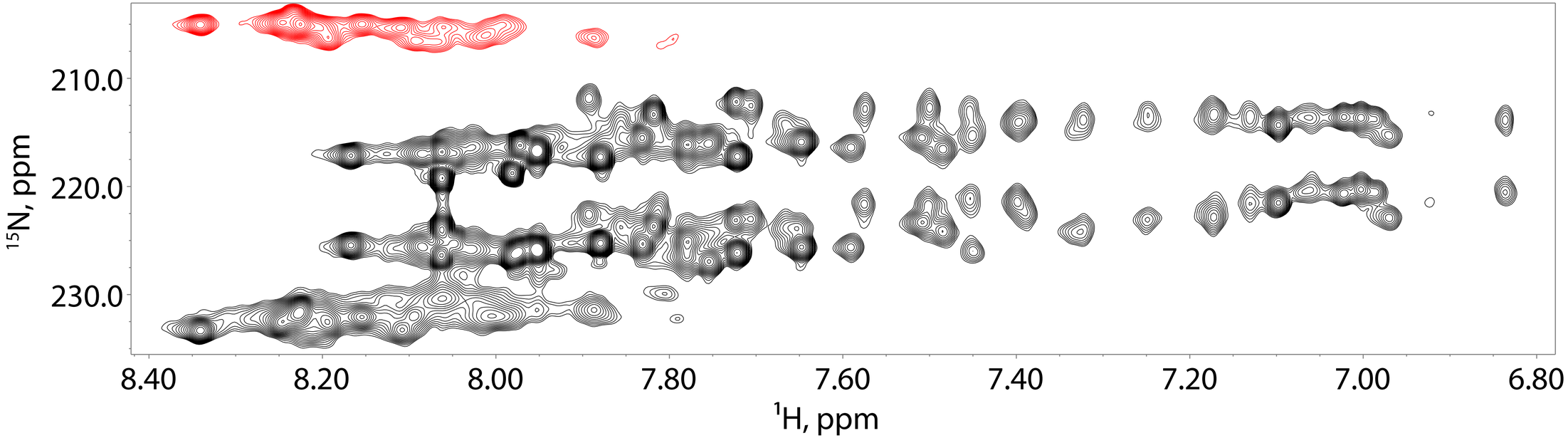
Residual dipolar couplings

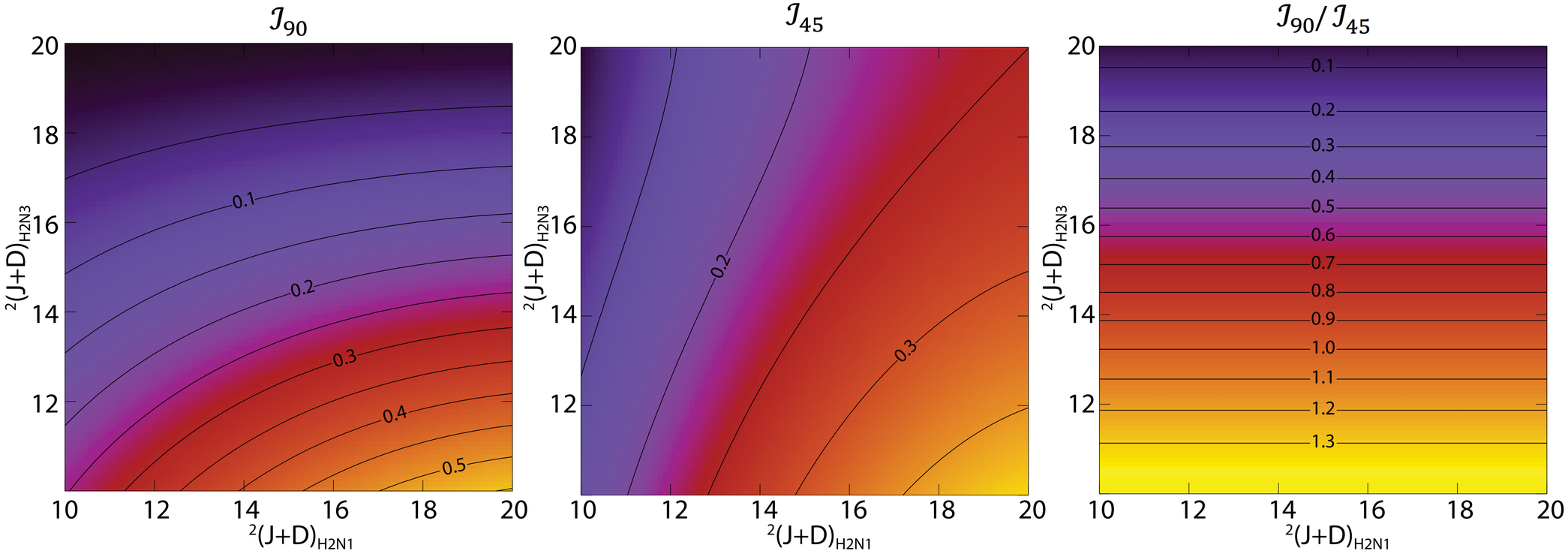
Large dependence on passive spin $\mathcal{I}_{IS} \propto \sin^2(\pi J_{IS}\tau)\ \cos^2(\pi J_{IT}\tau)$

Can fit if both signals resolved

Long selective nitrogen pulses



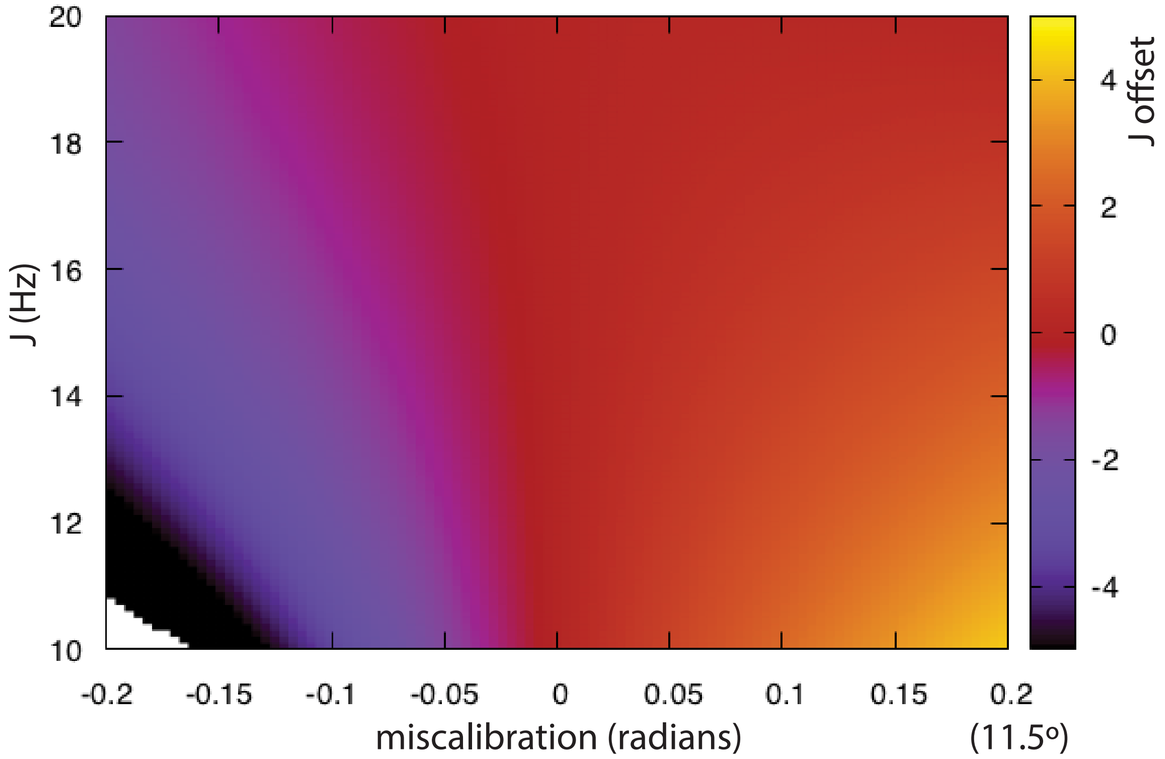




$$\mathcal{I}_\phi = \sin^2(\phi)\sin^2(\pi J_{IS}\tau)[\cos^2(\pi J_{IT}\tau) + \sin^2(\pi J_{IT}\tau)\cos^2(\phi)]$$

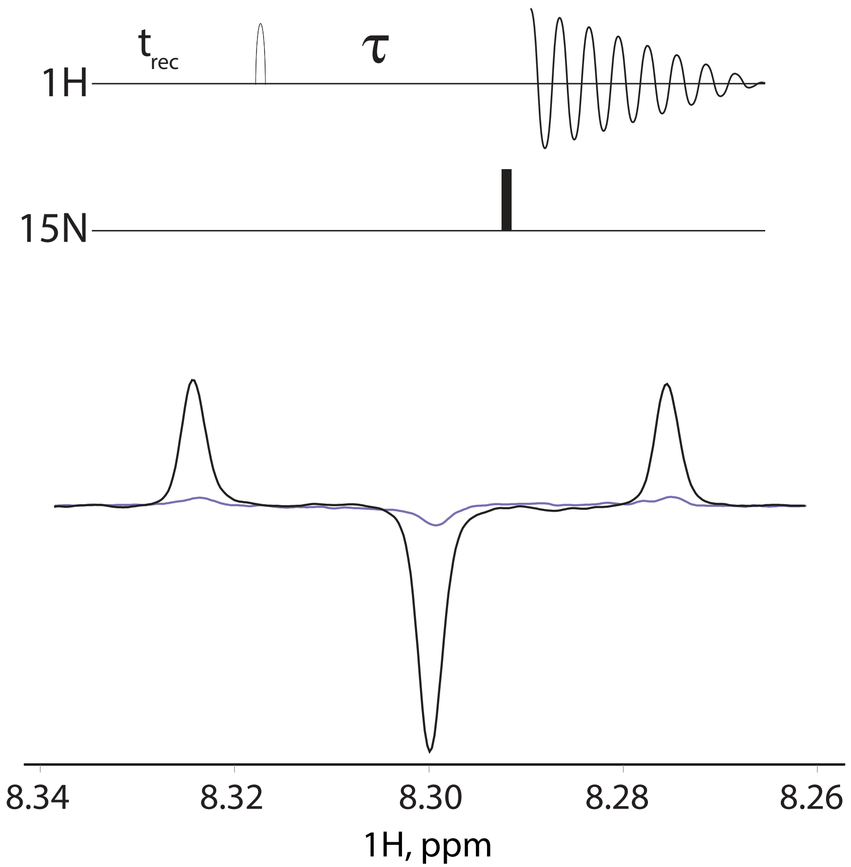
$$\frac{\mathcal{I}_{45}}{\mathcal{I}_{90}} = \frac{\tan^2(J_{IT}\pi\tau) + 2}{4}$$

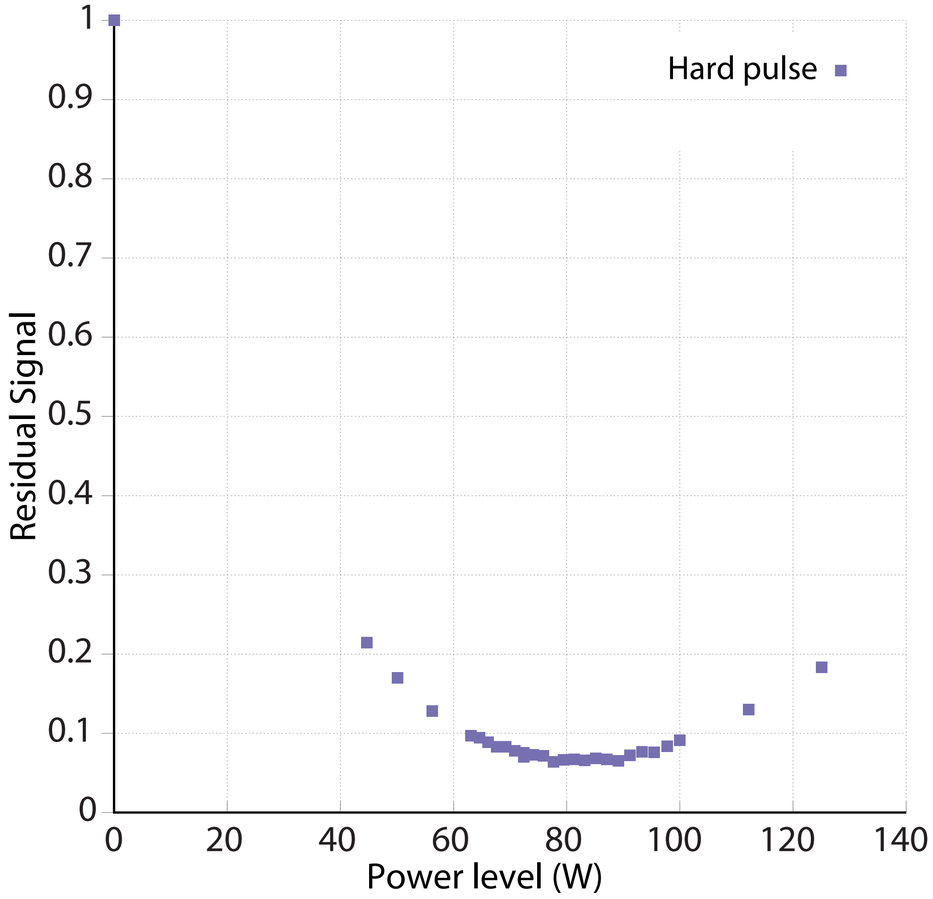
Pulse calibration essential



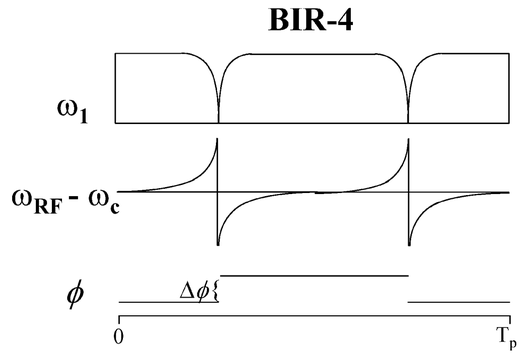
Difficult to accurately calibrate heteronuclear pulses

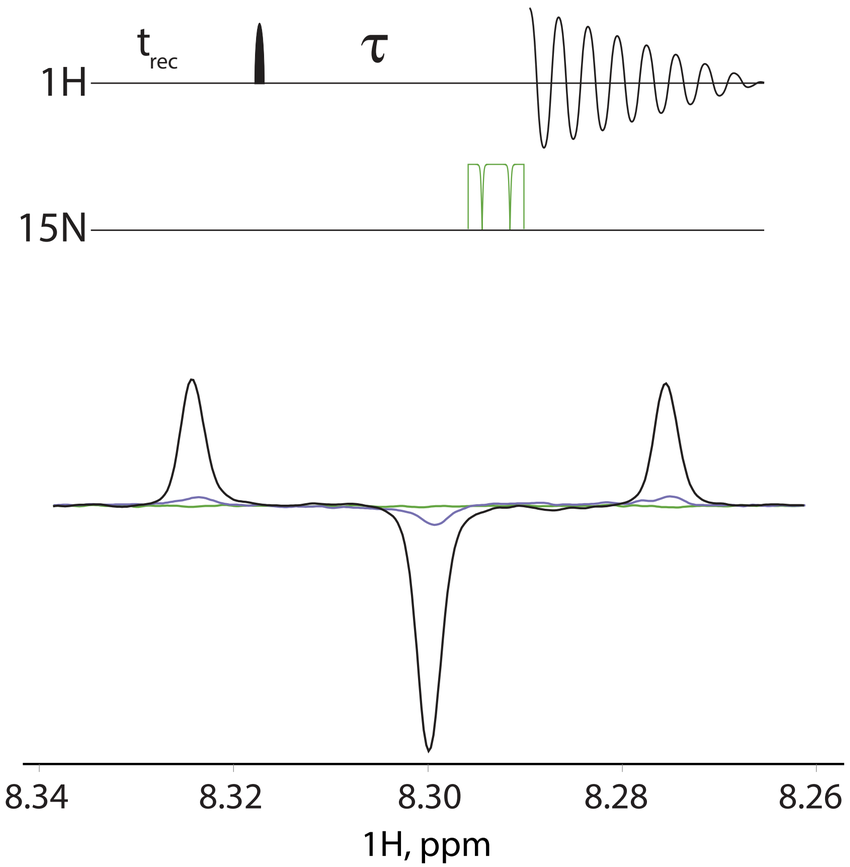
Probe power output is inhomogenous across the sample

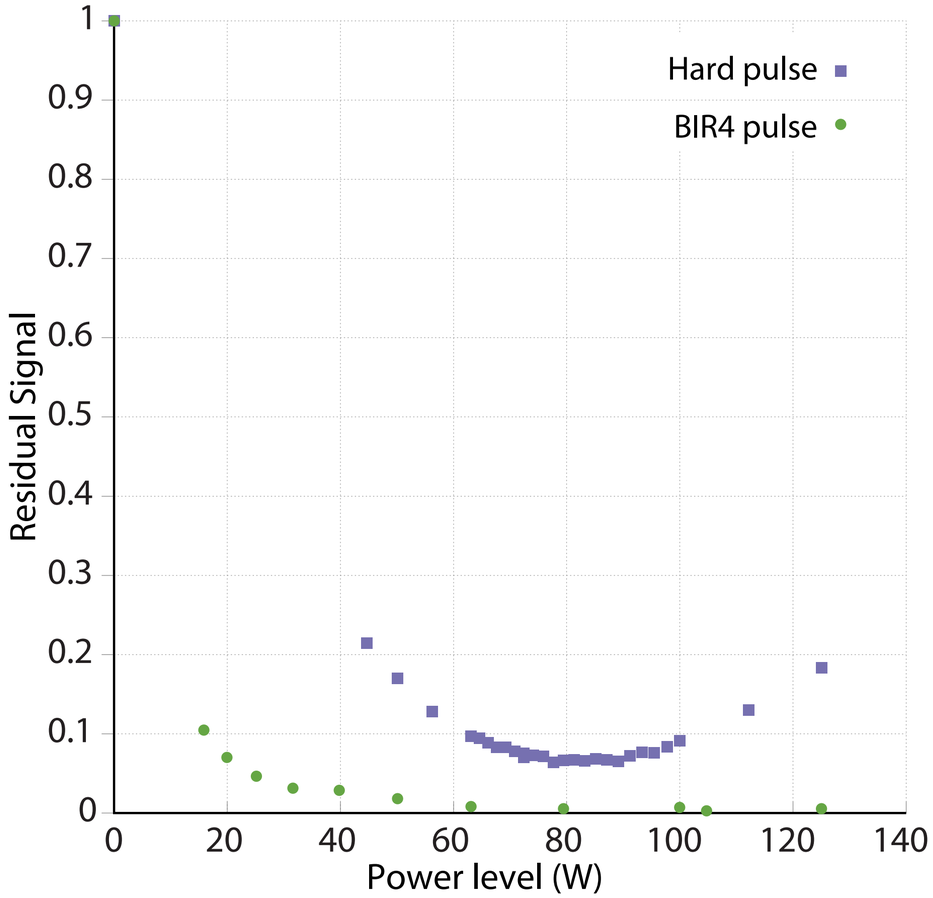


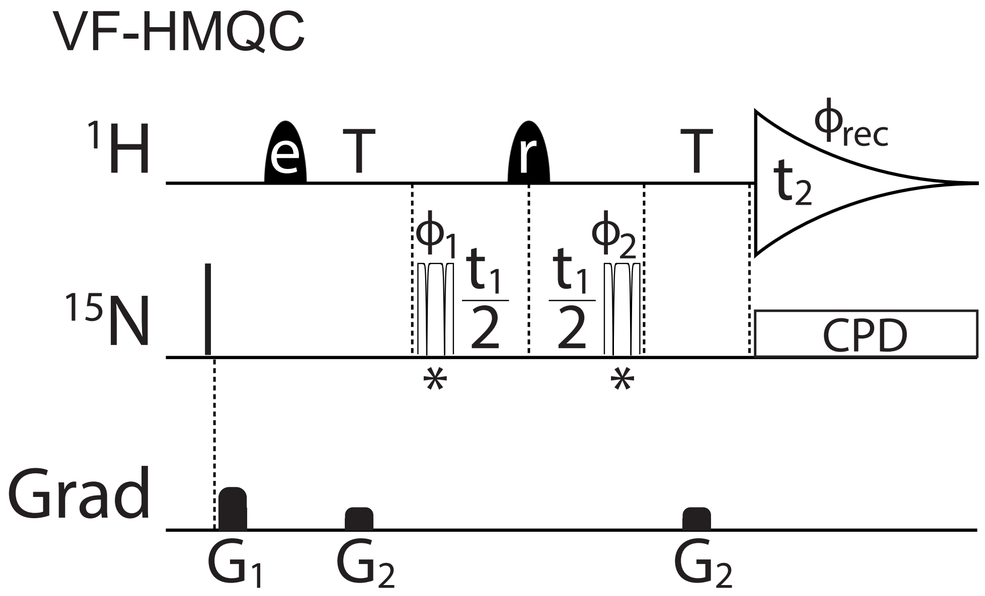


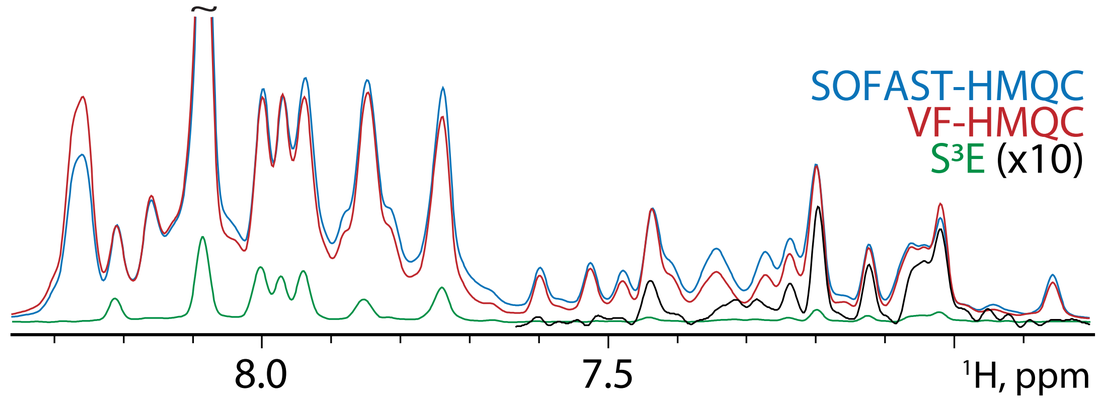
Adiabatic pulses designed for insensitivity to B1
(Garwood, M.; Ke, Y. J. Magn. Reson. 1991, 94, 511–525)

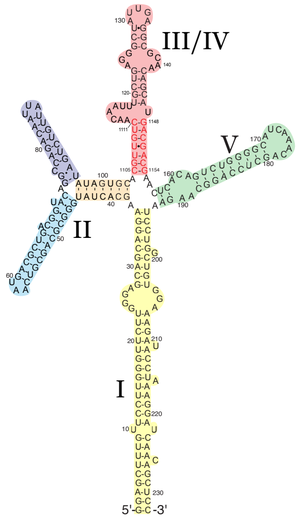


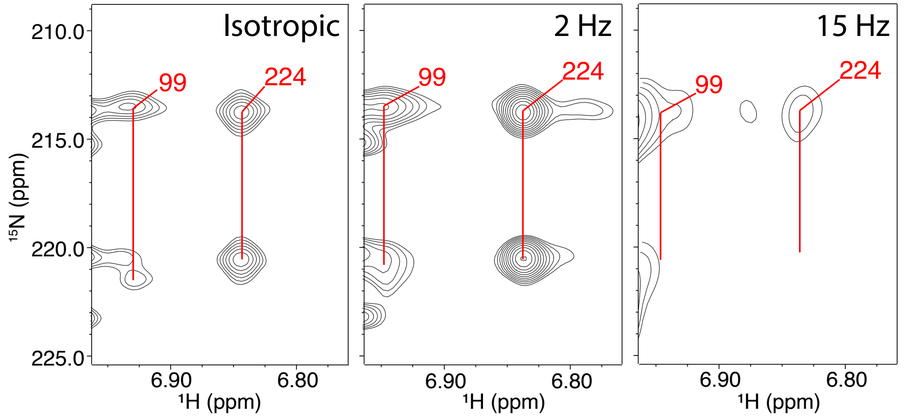


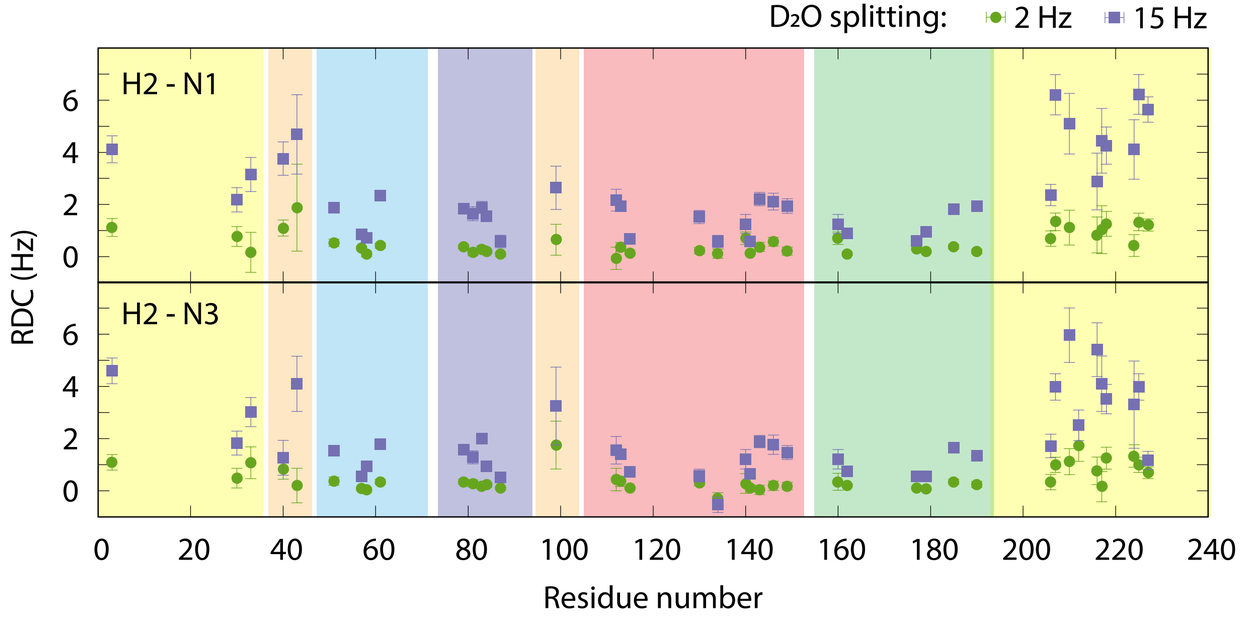








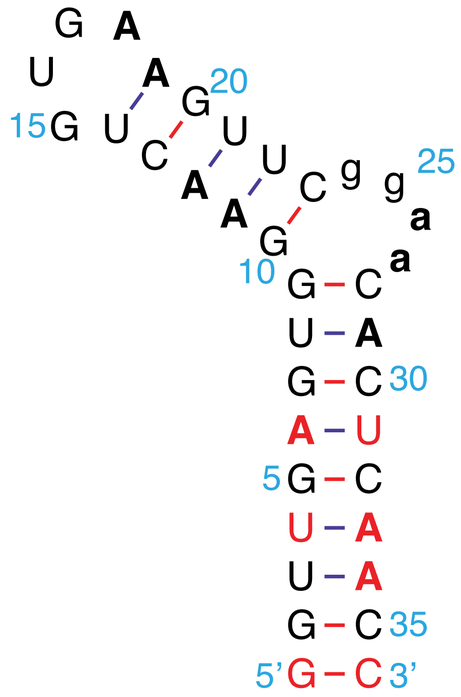




- 59 RDCs with uncertainty < 0.3 Hz measured for RRE
- No comparable method for measuring these couplings in large RNAs

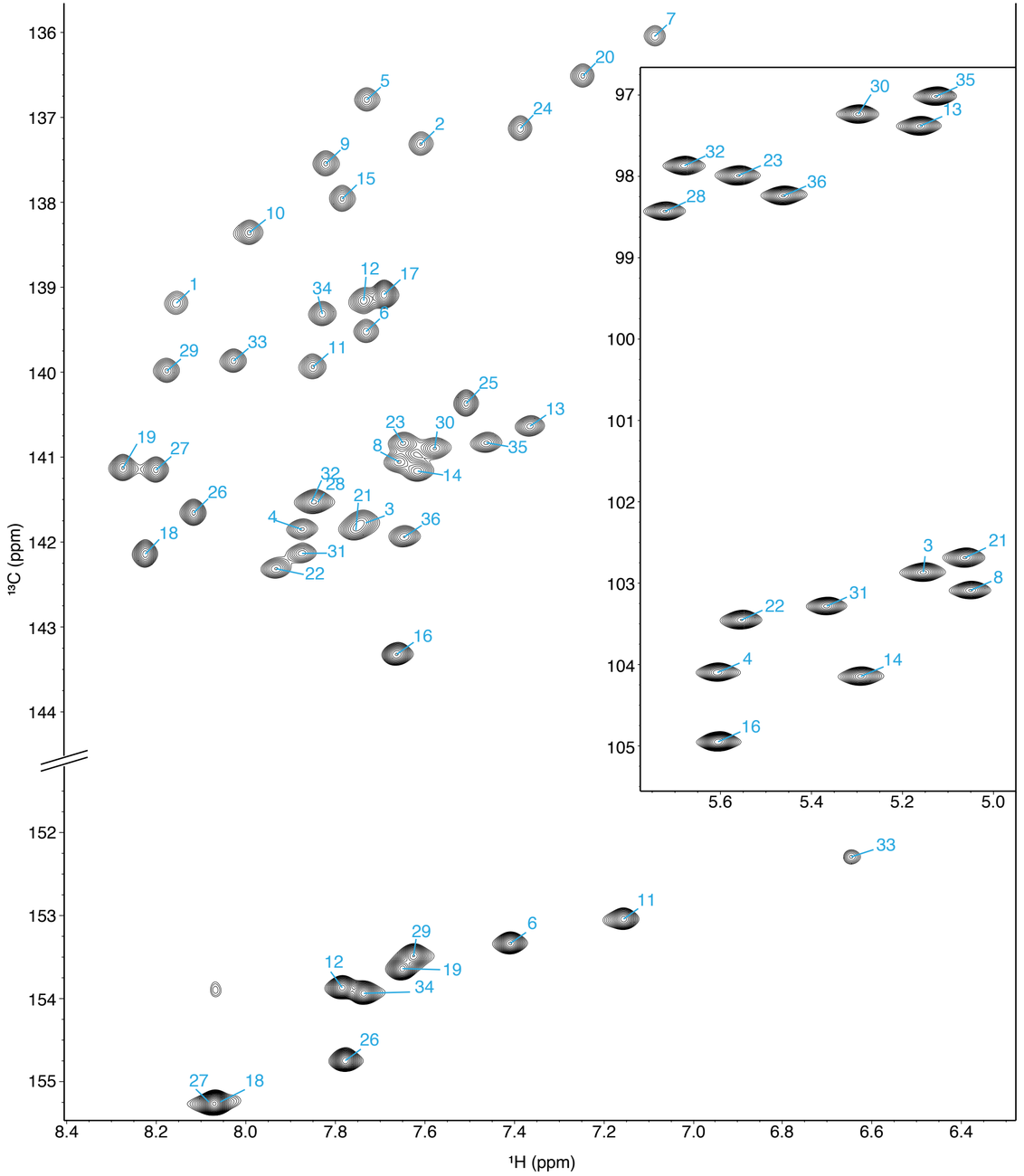
Marchant, Bax, Summers, J. Am. Chem. Soc. 2018

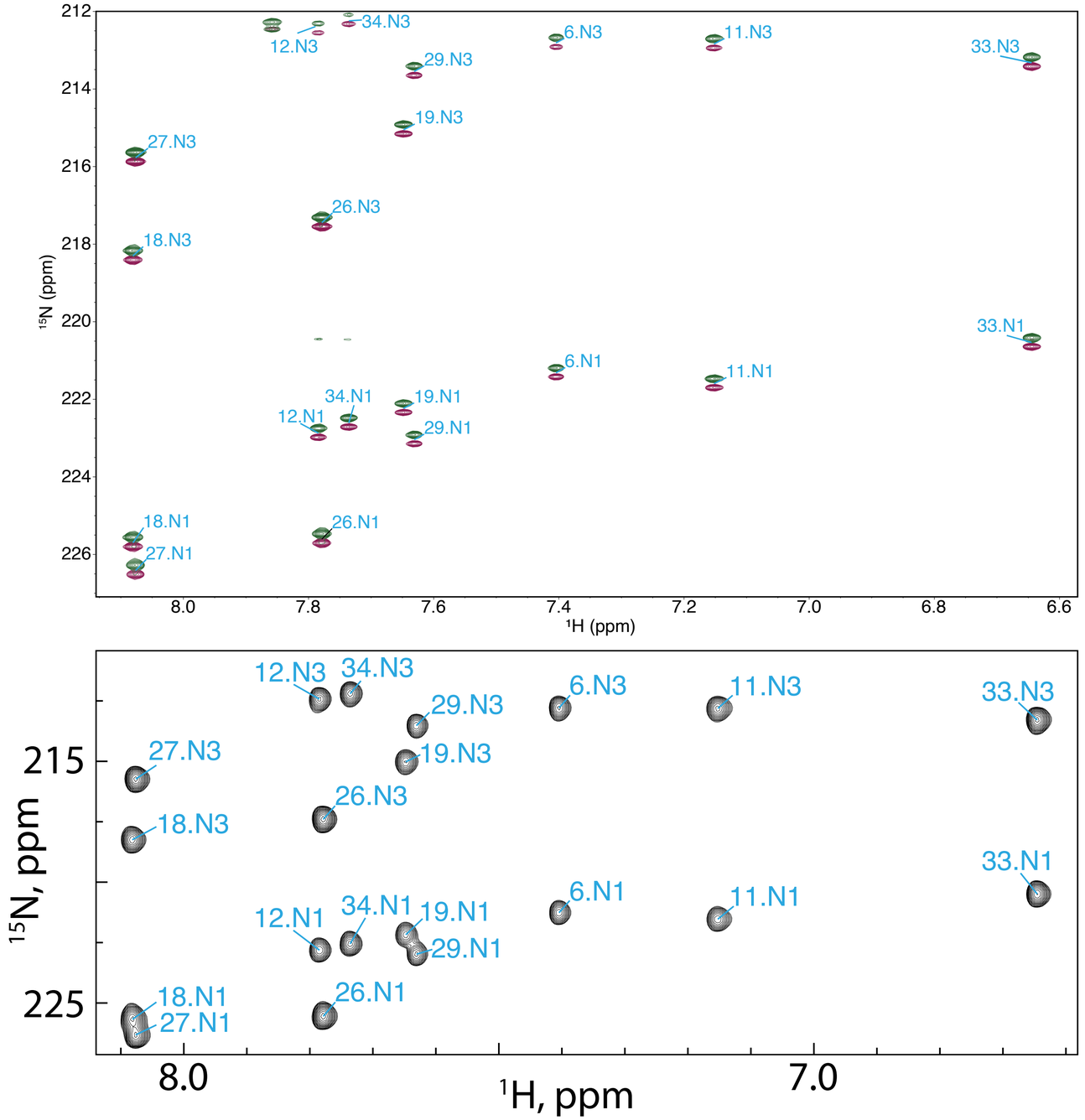
Model system

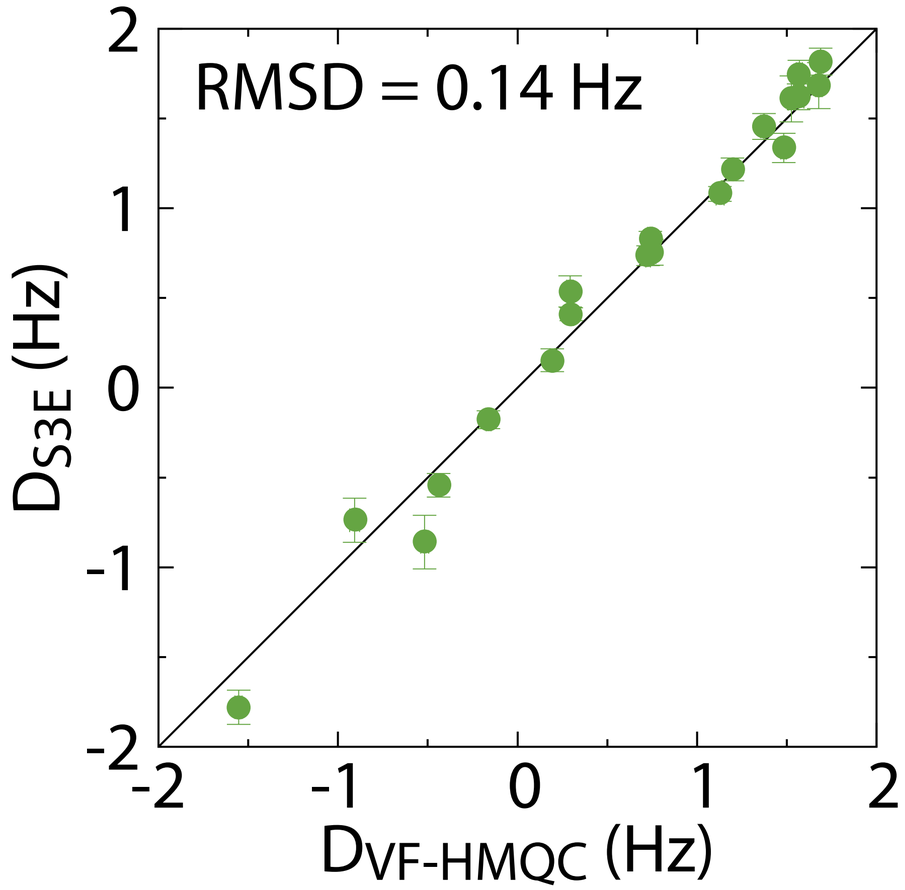


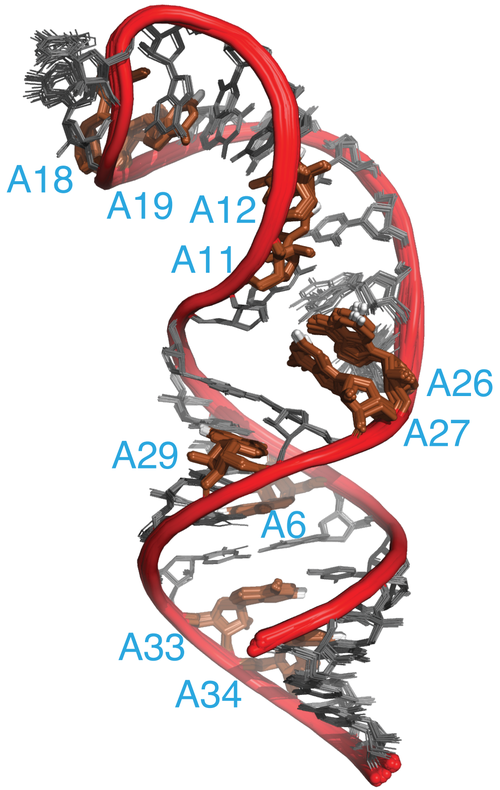
Derived from MMLV 5’-L SLC

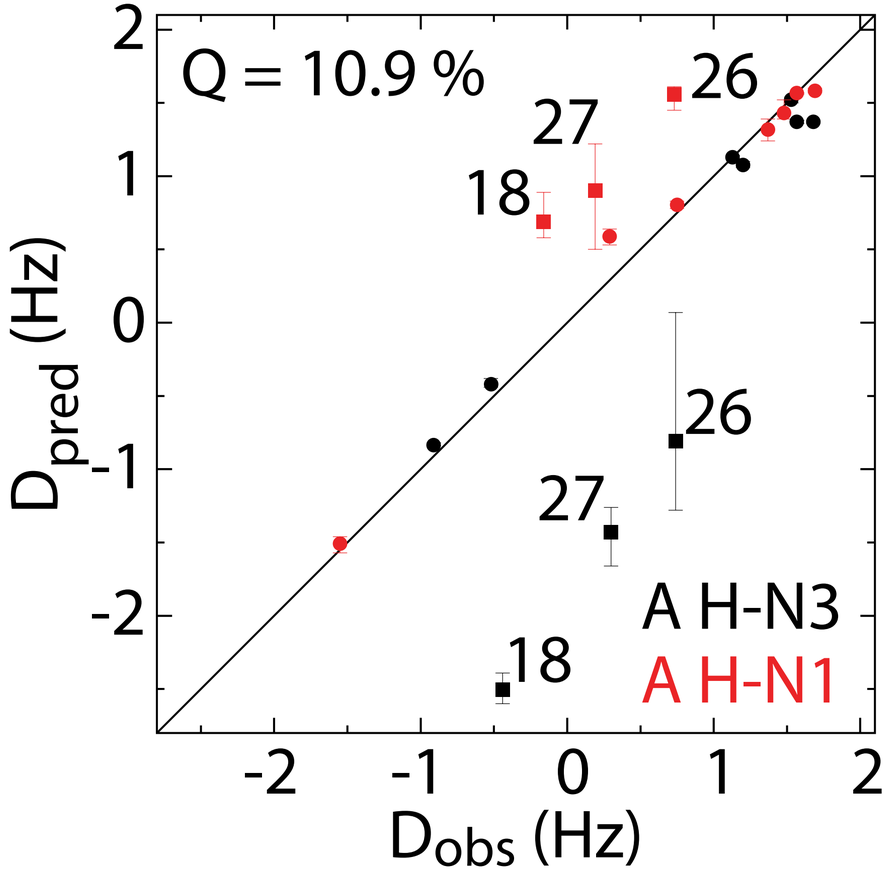
36 nt, 10 Adenosines



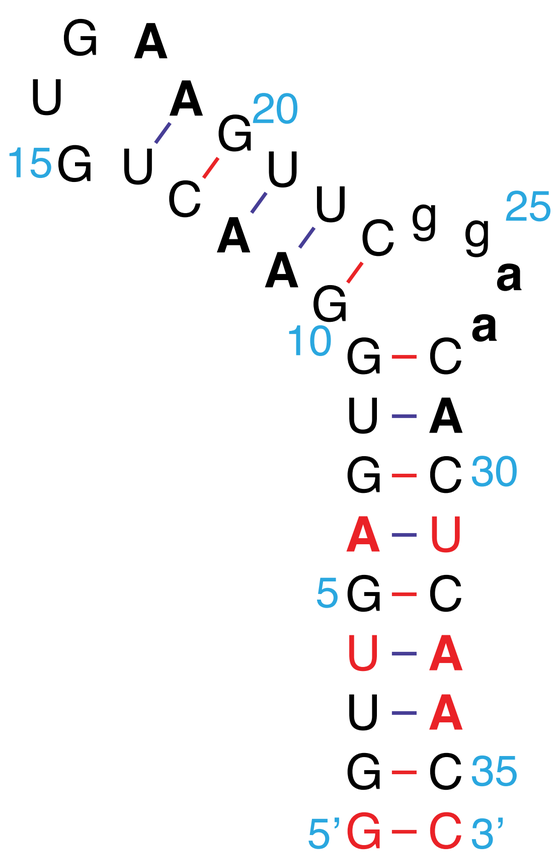

212
214
216
218
220
222
224
226
^{15}N (ppm)
8.0
7.8
7.6
7.4
7.2
7.0
6.8
6.6
^{1}H (ppm)
12.N3
34.N3
6.N3
11.N3
33.N3
29.N3
19.N3
27.N3
26.N3
18.N3
33.N1
6.N1
11.N1
34.N1
19.N1
12.N1
29.N1
18.N1
26.N1
27.N1
215
225
^{15}N, ppm
8.0
7.0
^{1}H, ppm

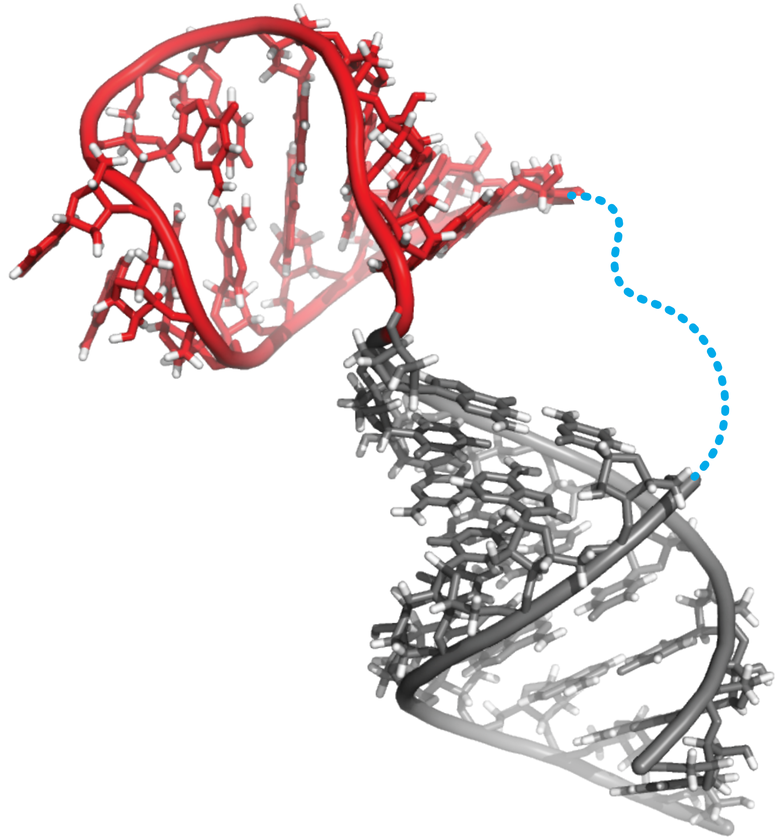

RMSD = 0.14 Hz
D_{S3E} (Hz)
$D_{VF-HMQC}$ (Hz)
2
1
0
-1
-2
-2
-1
0
1
2

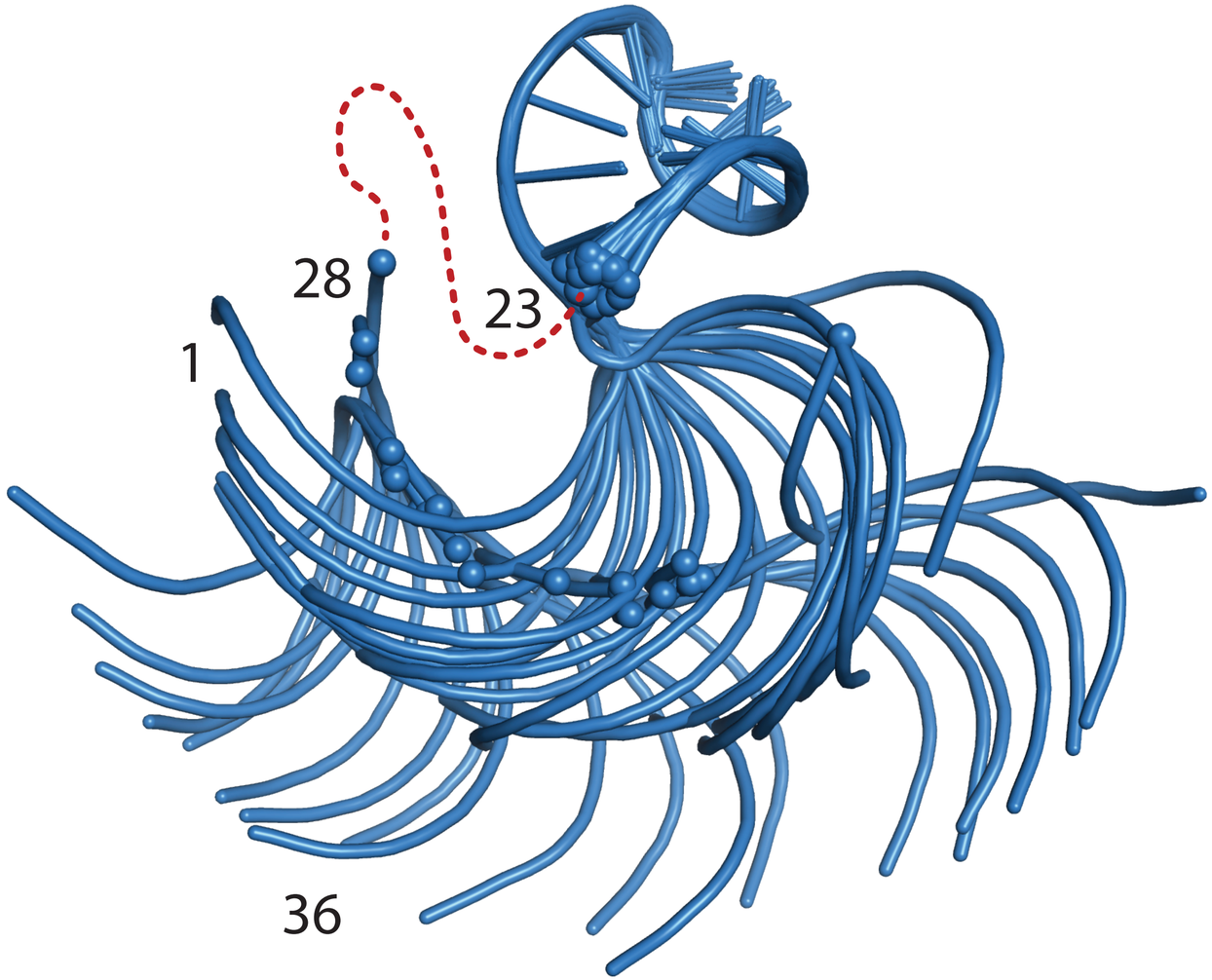

A18
A19
A12
A11
A26
A27
A29
A6
A33
A34

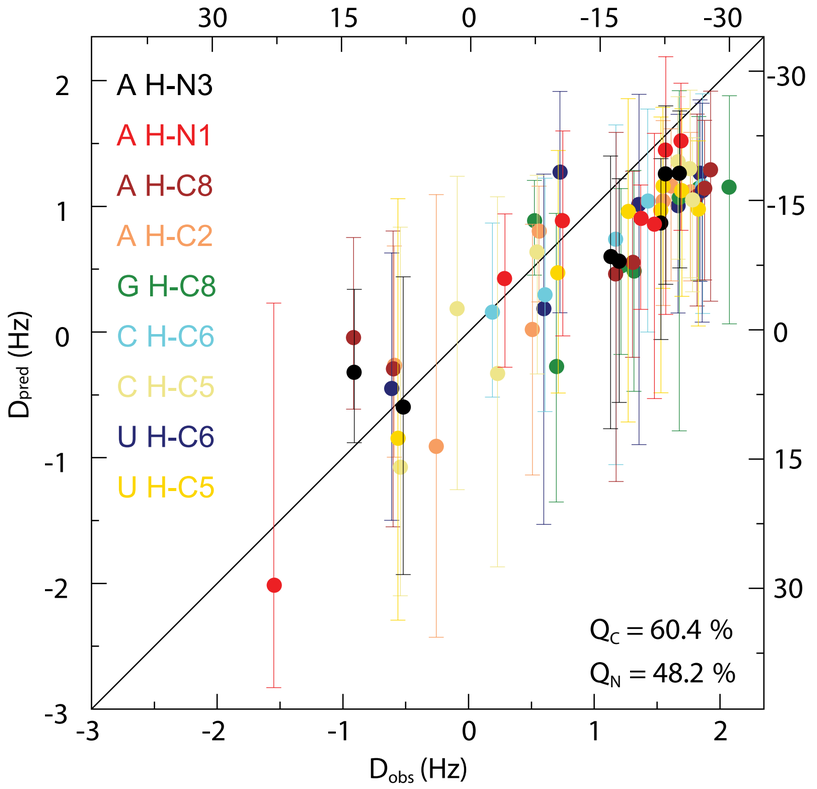

Q = 10.9 %
D_{pred} (Hz)
D_{obs} (Hz)
26
27
18
26
27
18
A H-N3
A H-N1
2
1
0
-1
-2
-2
-1
0
1
2

Removed helix:helix NOEs and relaxed geometry

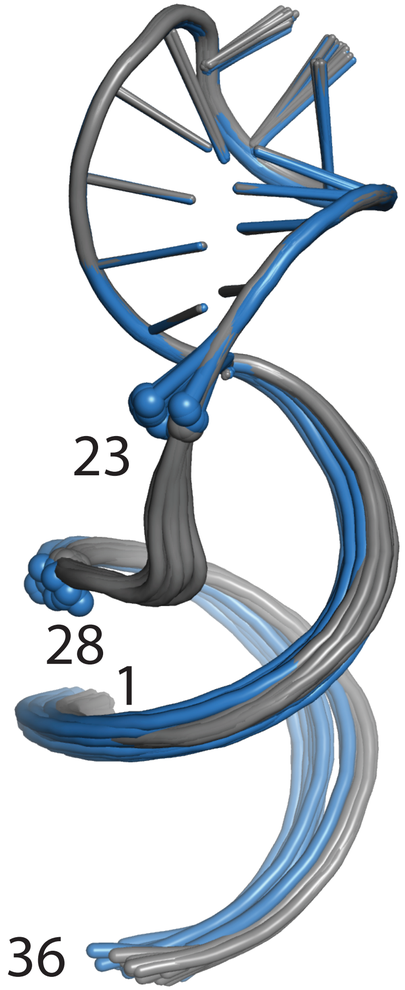


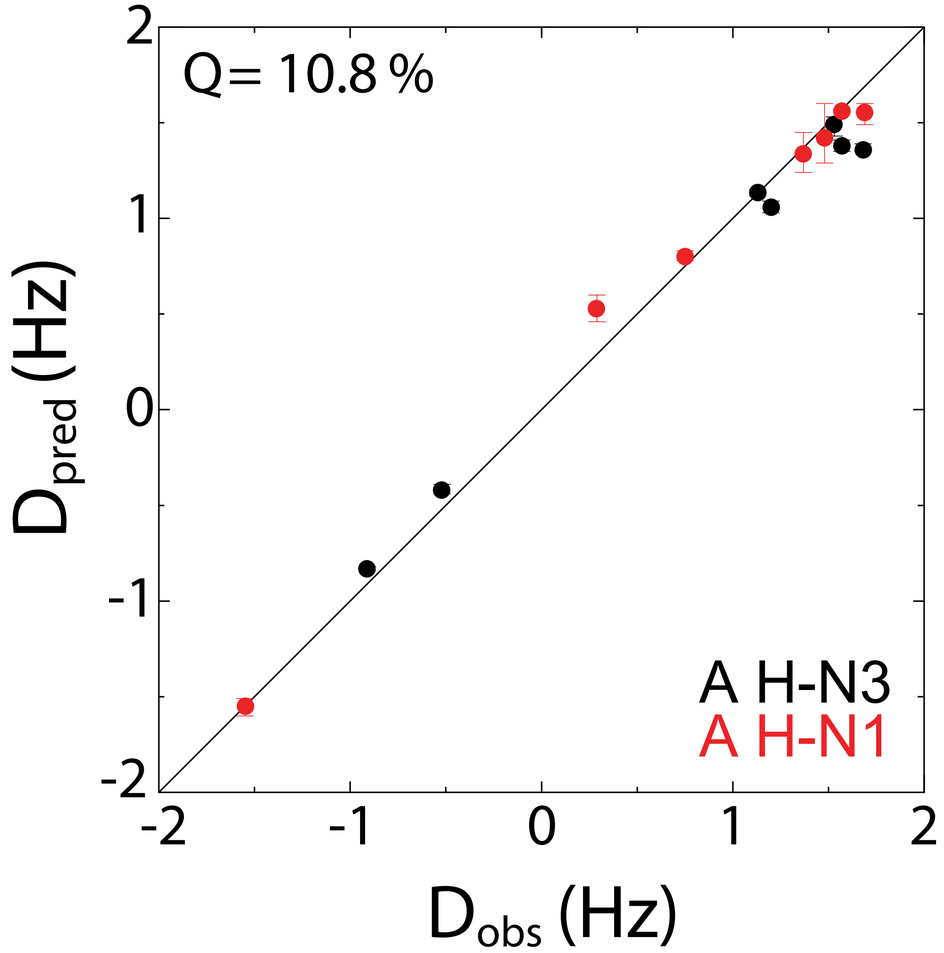



28
23
1
36

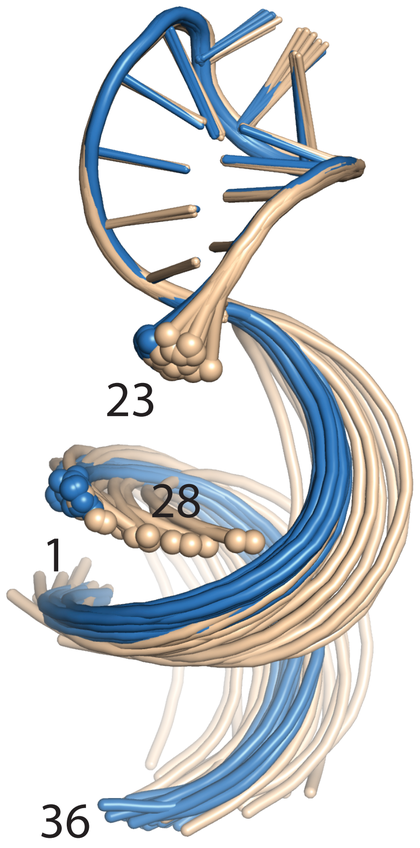

30
15
0
-15
-30
A H-N3
A H-N1
A H-C8
A H-C2
G H-C8
C H-C6
C H-C5
U H-C6
U H-C5
Q_C = 60.4 %
Q_N = 48.2 %
D_{pred} (Hz)
D_{obs} (Hz)

−83 - 110°

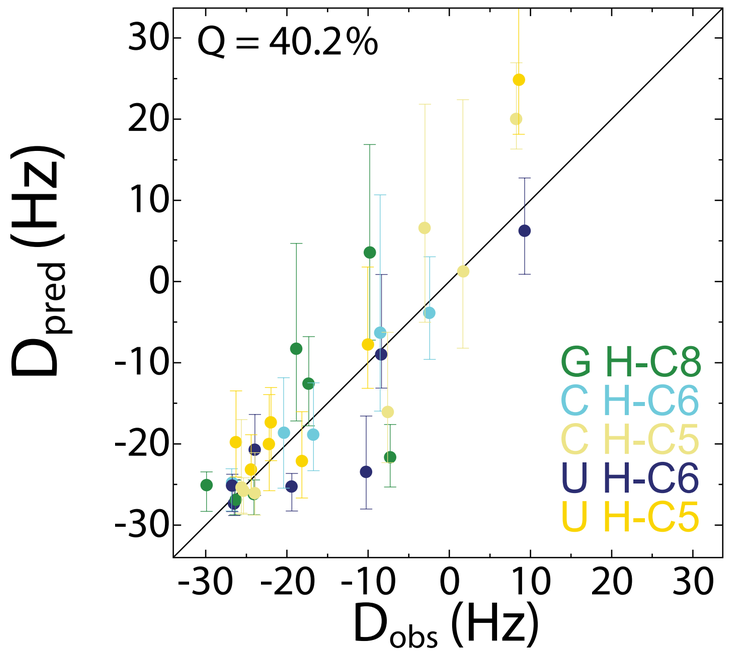


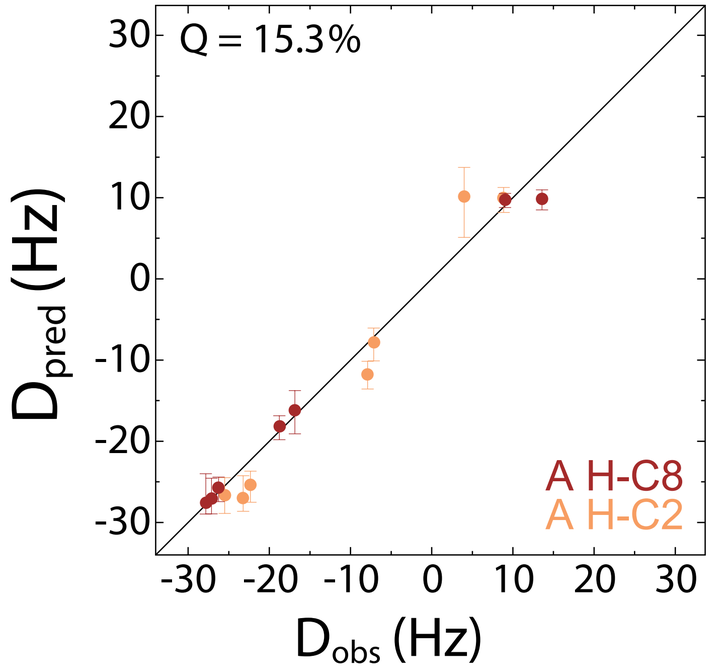


76 - 84° (M = 81.5°, SD = 2.6°)



77 - 94° (M = 84°, SD = 4.1°)





Significantly extends the size limit for RDC measurement in RNA

Consistent with existing methods

Sufficient to determine interhelical orientation with as few as three adenosines per element

Applicable to RNAs on the order of hundreds of nucleotides, providing access to higher quality structures of larger RNAs than previously possible

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Mike Summers
Danny Morris
Colin O'Hern
Lindsay Glang
Roald Teuben
The rest of the lab

NIH National Institutes of Health

Ad Bax
Nico Tjandra

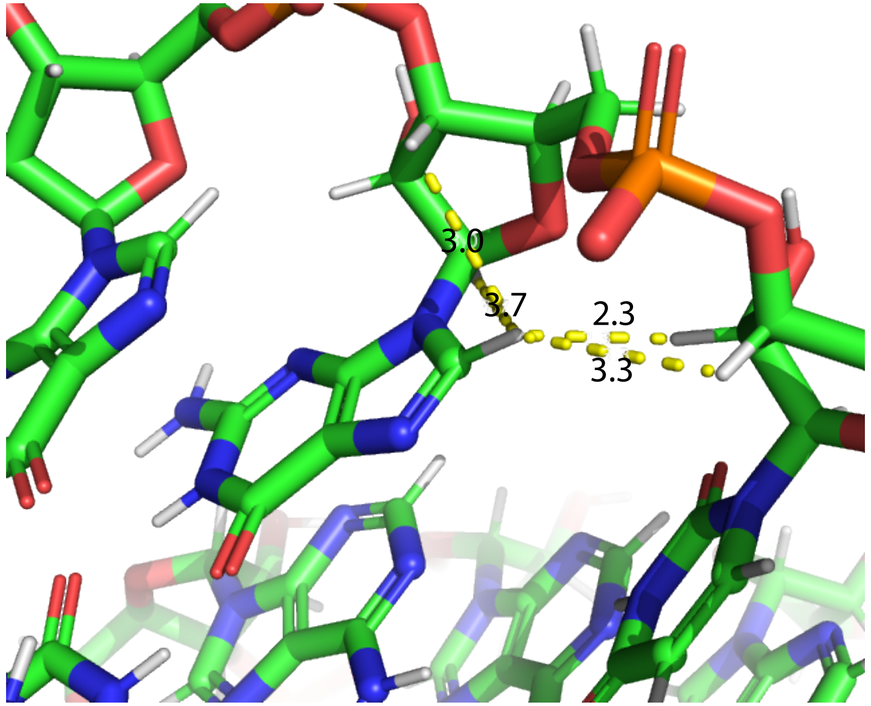
ADVANCED SCIENCE RESEARCH CENTER
THE GRADUATE CENTER
CITY UNIVERSITY OF NEW YORK

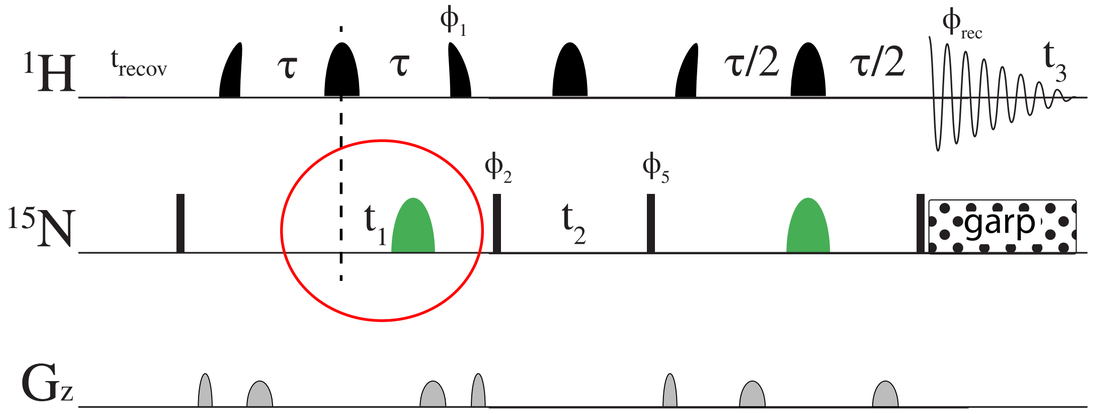
Bruce Johnson

CRNA

hhmi

NIH National Institutes of Health



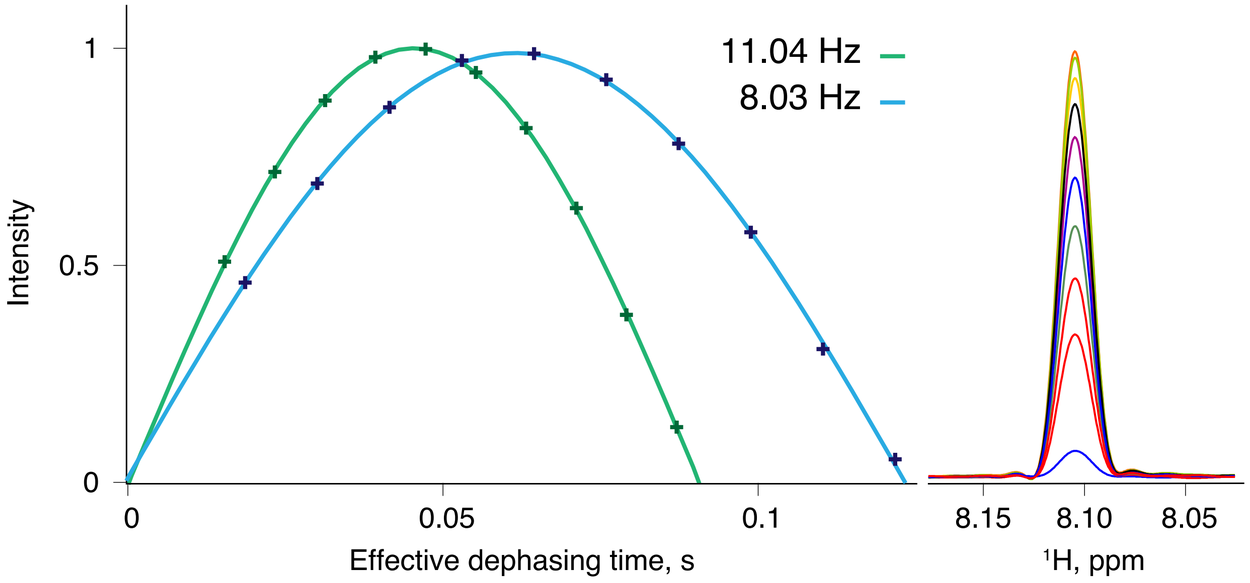


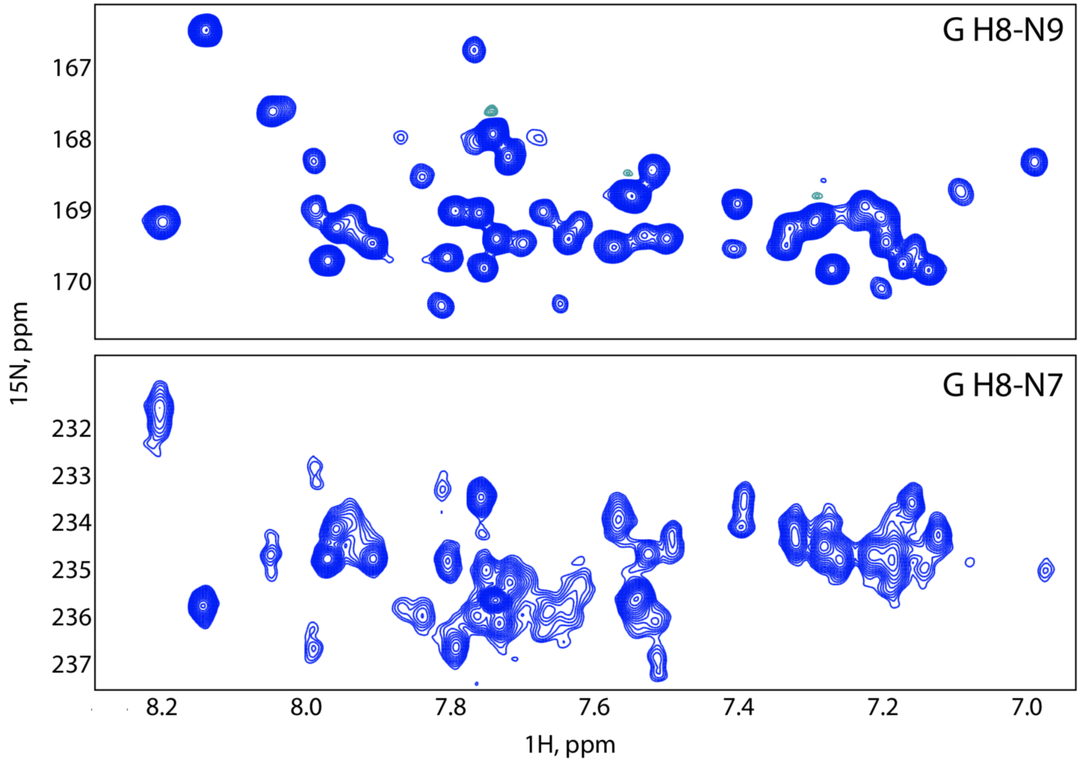
$\tau = 1/(2J)$

EBURP2, 6.4 - 8 ppm

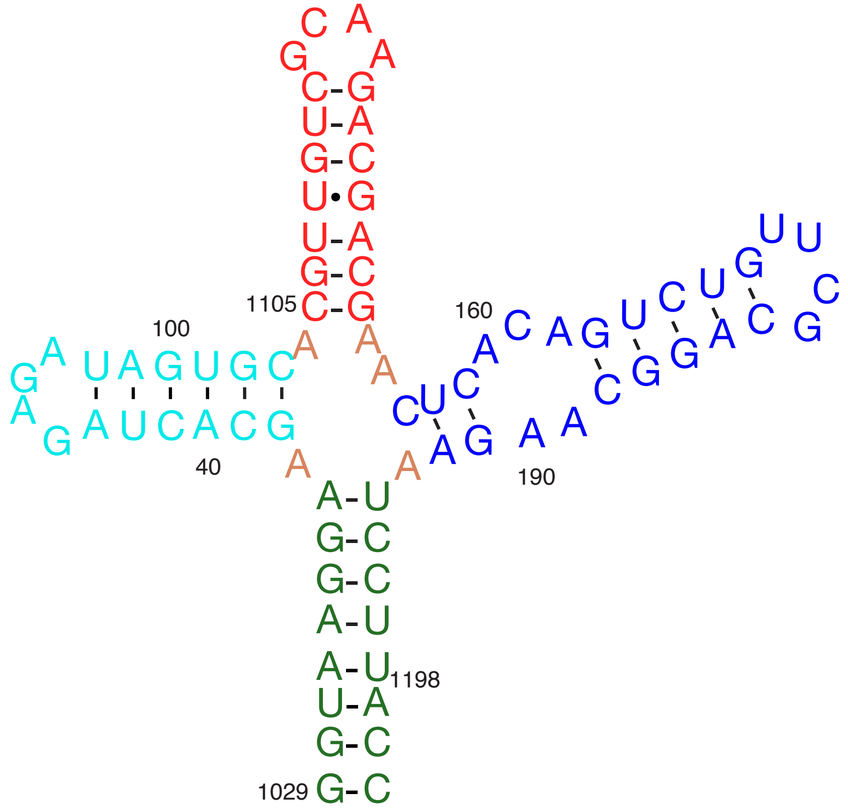
REBURP, 6.4-8 ppm

REBURP, 209-229 ppm

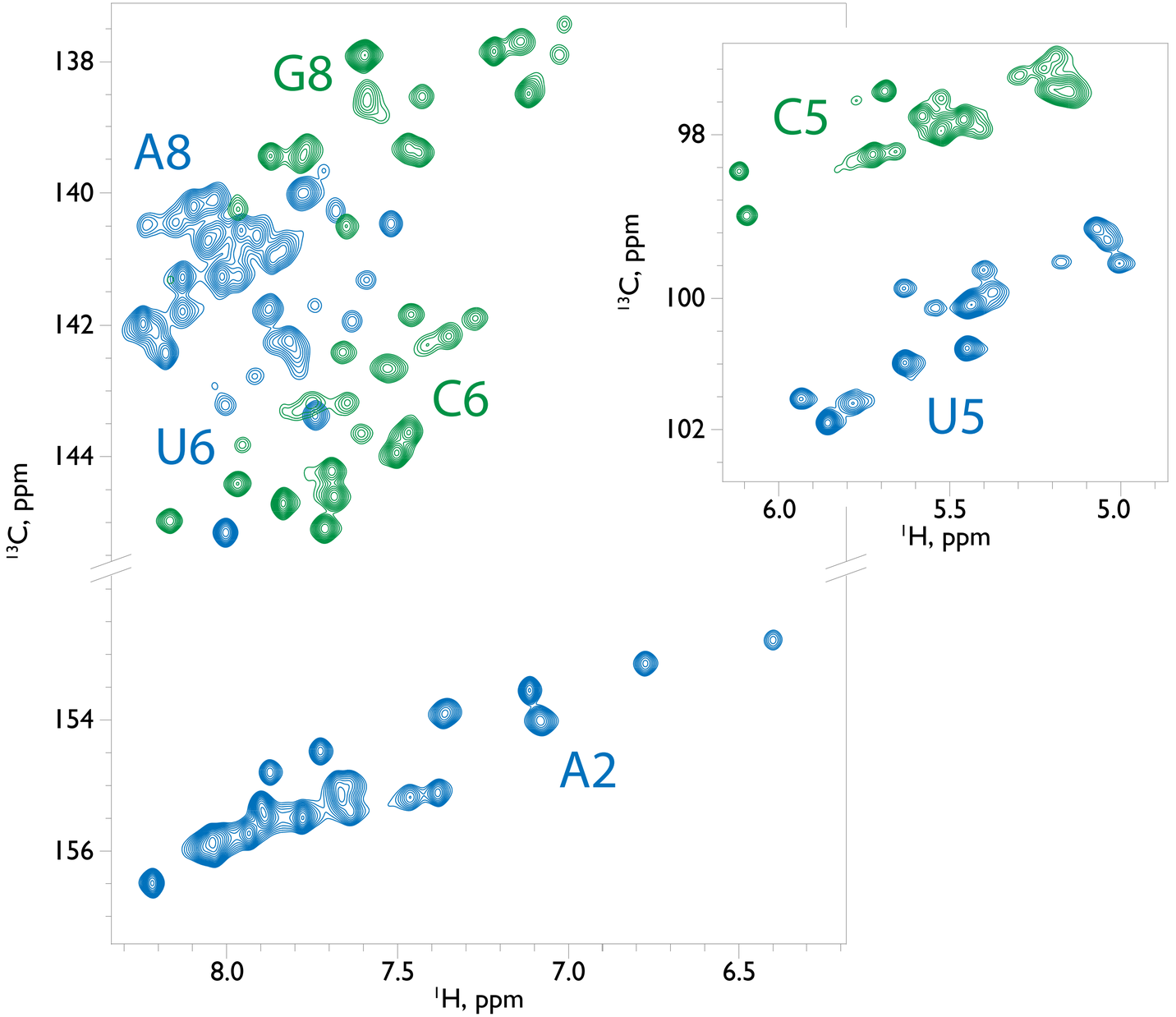




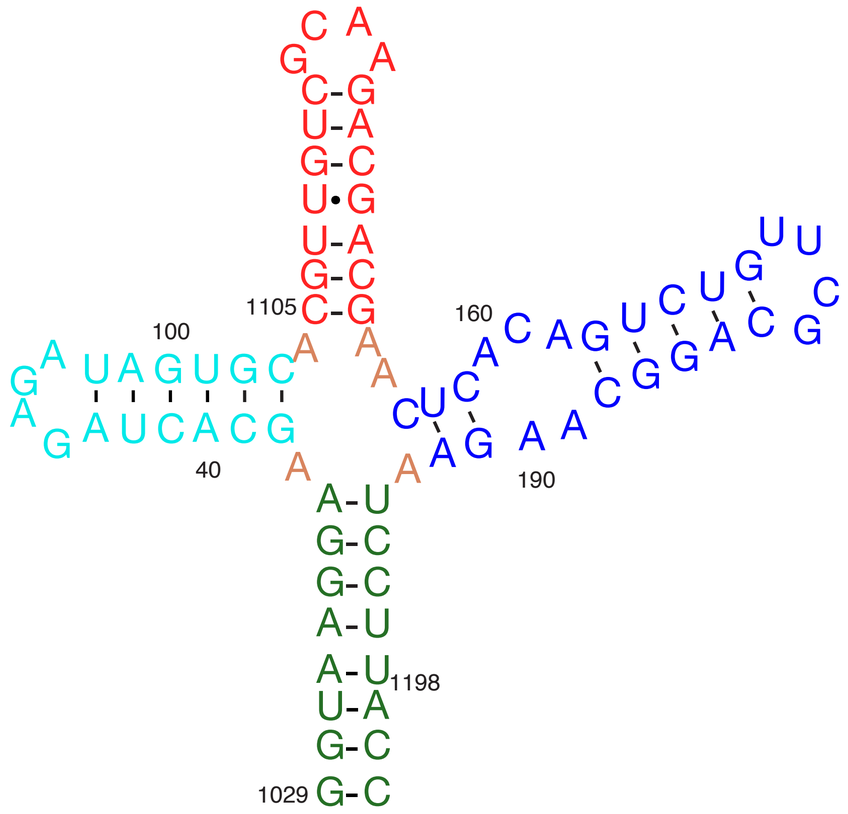
ARTSY



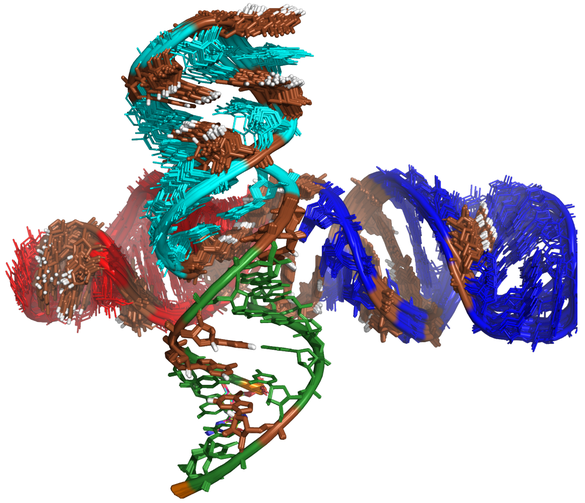
79 nt

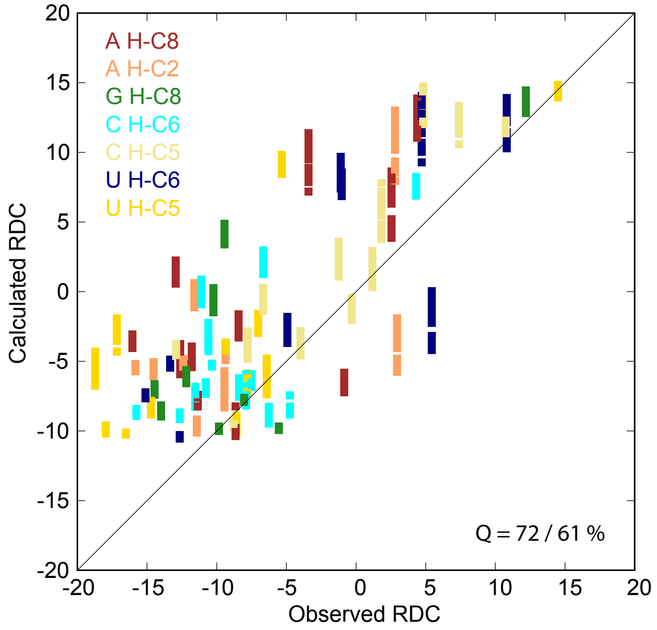


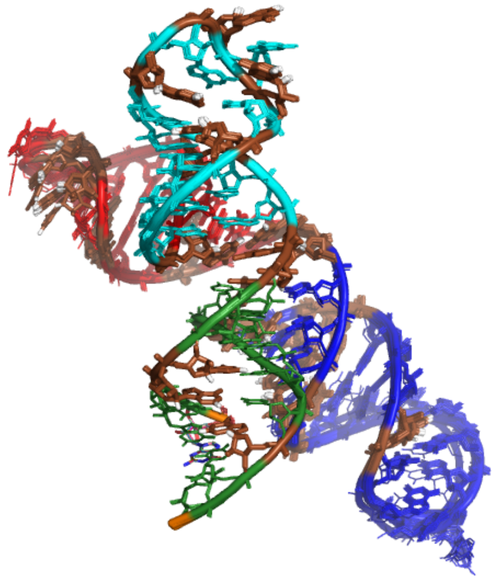
ARTSY

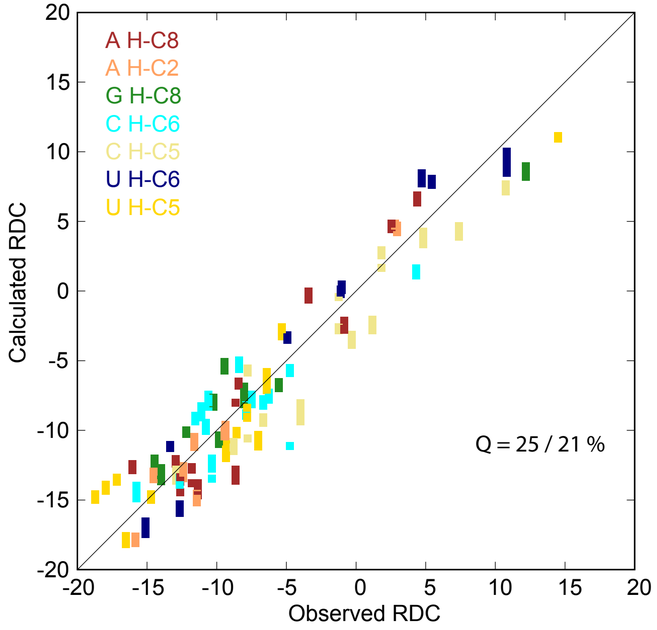

100
40
160
190
1105
1198
1029

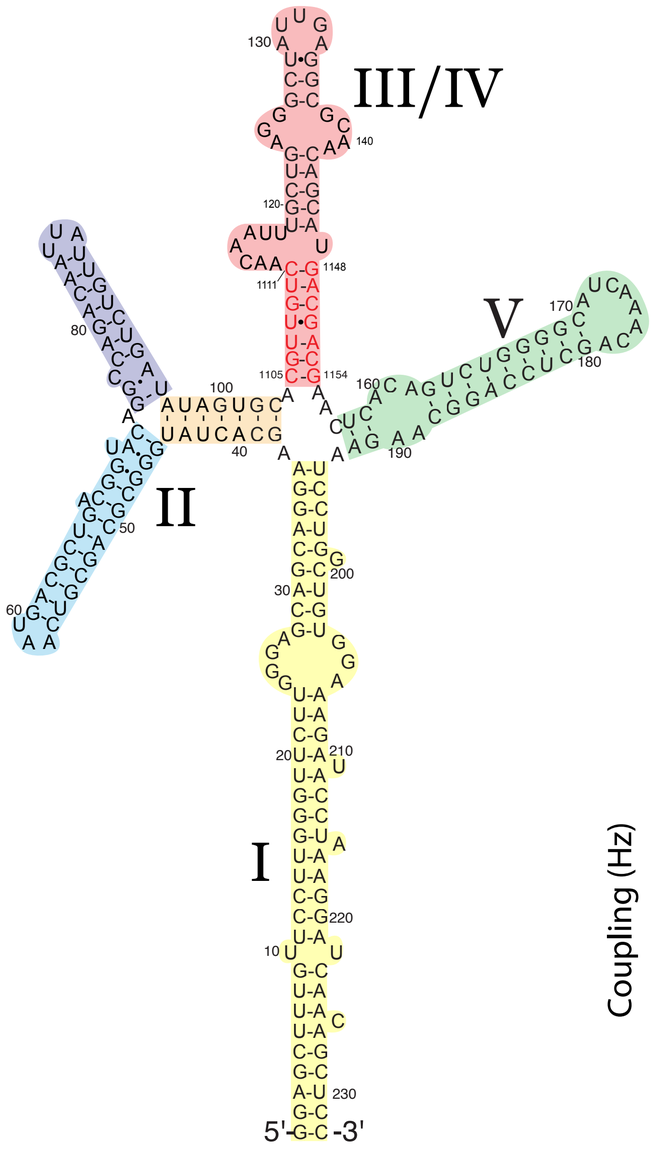
79 nt

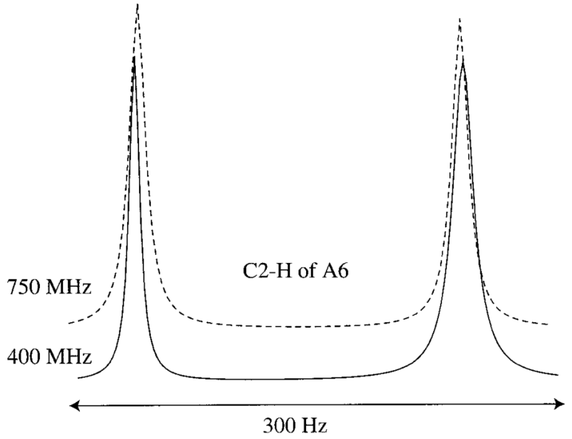



A H-C8
A H-C2
G H-C8
C H-C6
C H-C5
U H-C6
U H-C5
Calculated RDC
Observed RDC
Q = 72 / 61 %

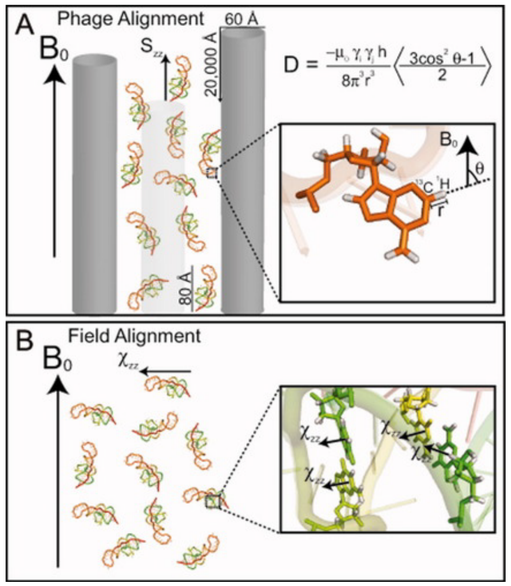


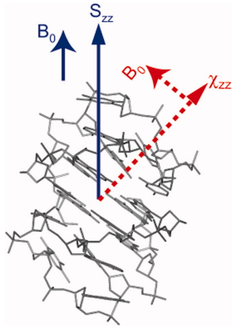

A H-C8
A H-C2
G H-C8
C H-C6
C H-C5
U H-C6
U H-C5
Calculated RDC
Observed RDC
Q = 25 / 21 %



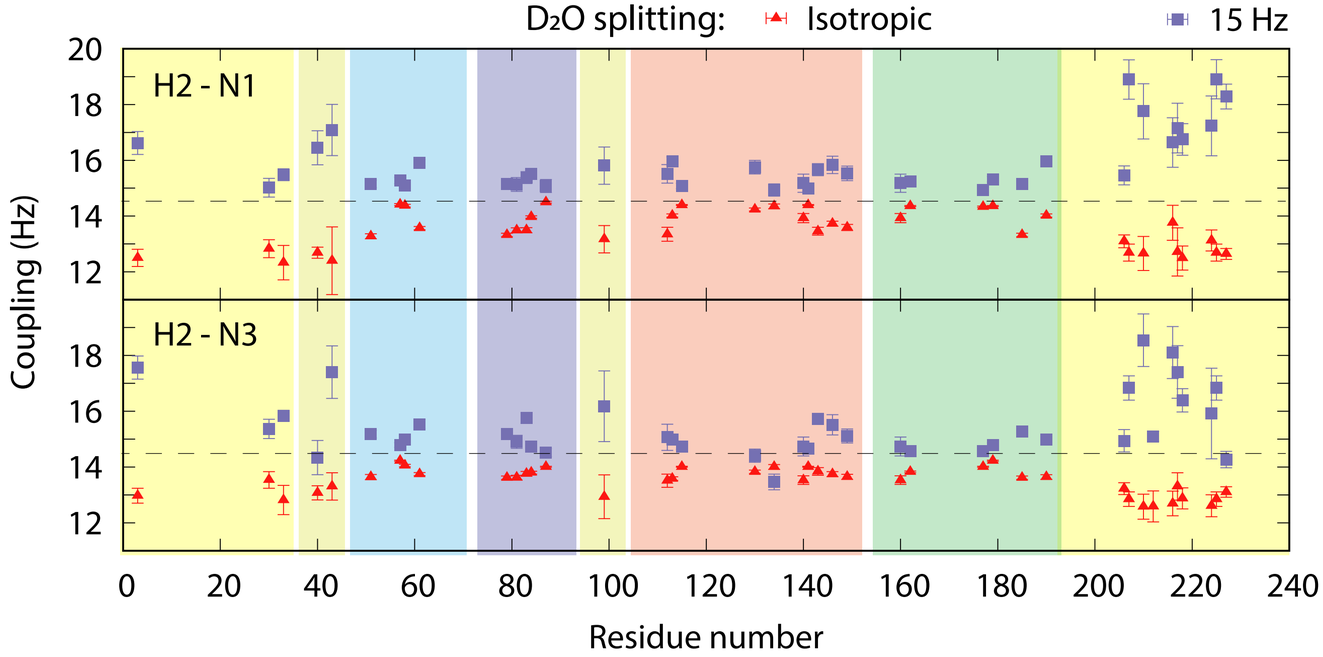


Kung et al., JMR Series B 109, 323–325 (1995)



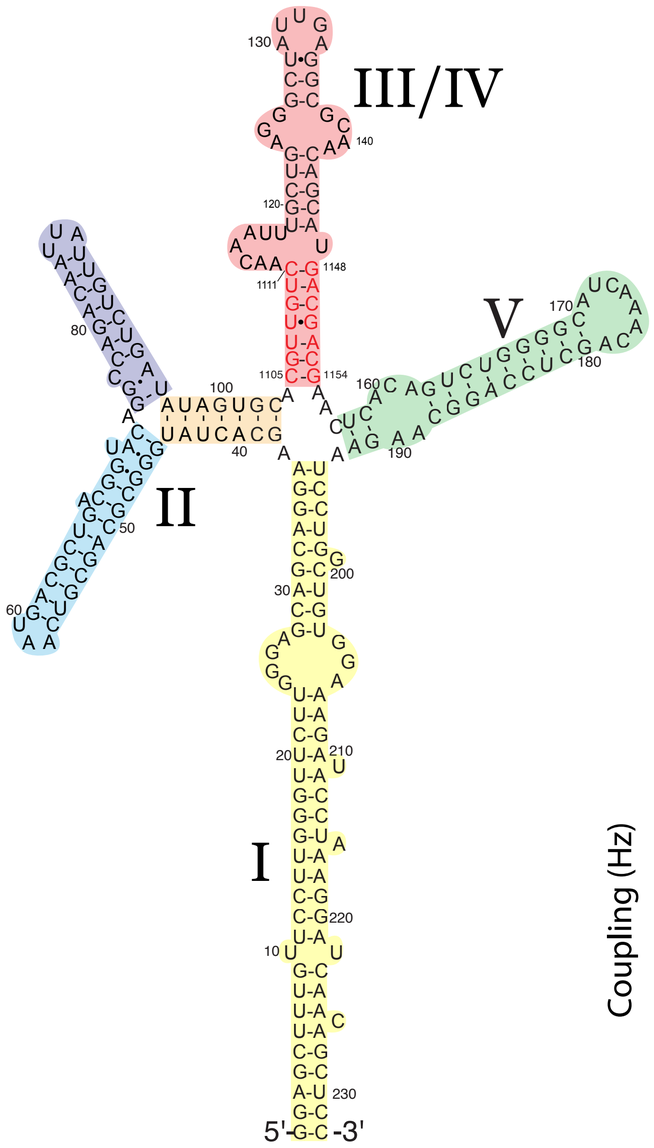


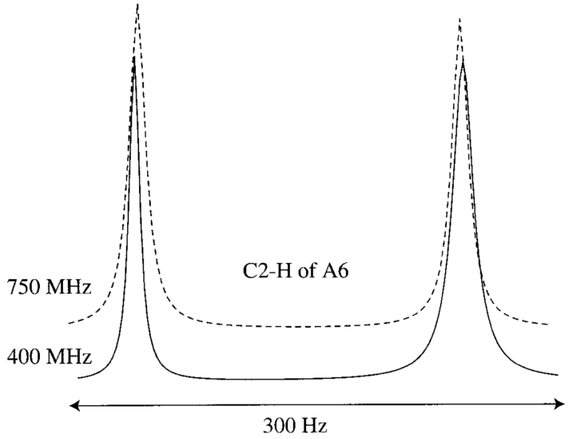
Getz, et al., Al-Hashimi, Biopolymers 86: 384-402 (2007).



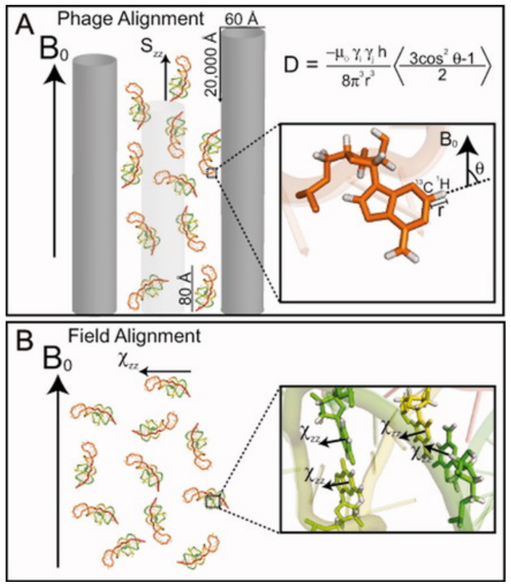
N.B. Calculations by Sklenár and colleagues suggest ~1.5 Hz decrease in $^{2}J_{H2N1}$ when base paired

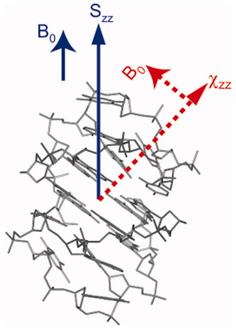
Fiala, R., Munzarova, M.L. & Sklenár , V. J Biomol NMR (2004) 29: 477



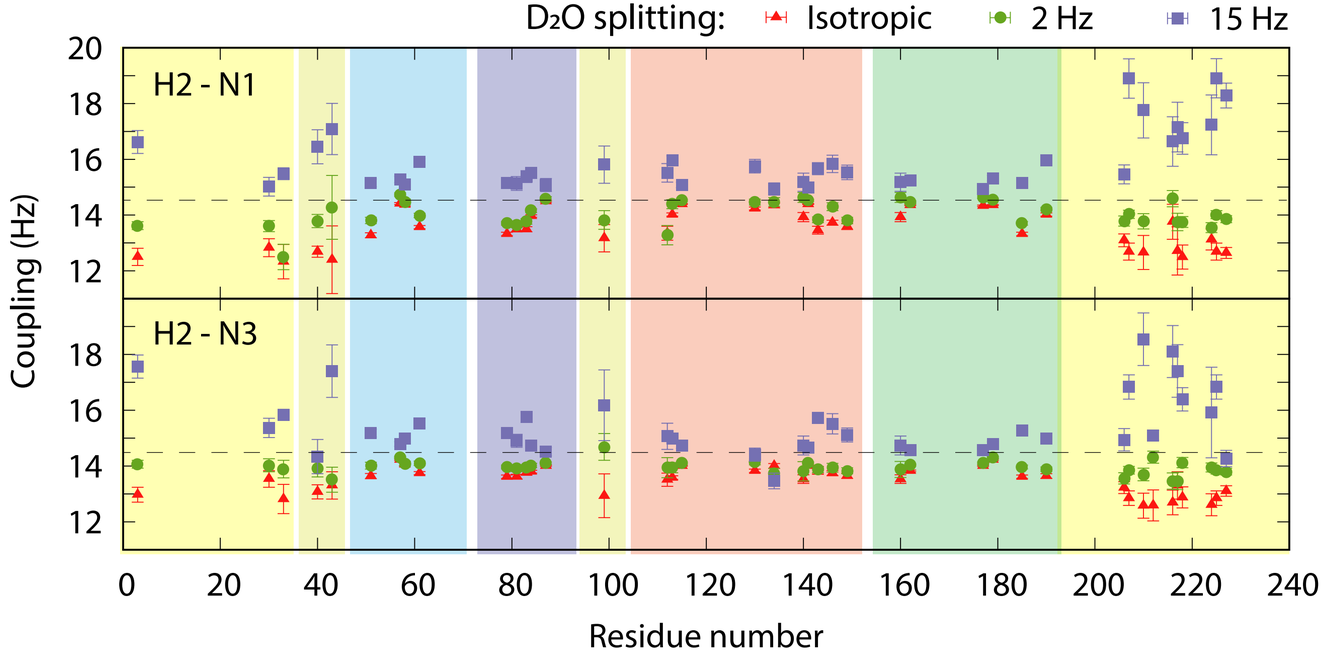


Kung et al., JMR Series B 109, 323–325 (1995)





Getz, et al., Al-Hashimi, Biopolymers 86: 384-402 (2007).



N.B. Calculations by Sklenár and colleagues suggest ~1.5 Hz decrease in $^{2}J_{H2N1}$ when base paired

Fiala, R., Munzarova, M.L. & Sklenár , V. J Biomol NMR (2004) 29: 477