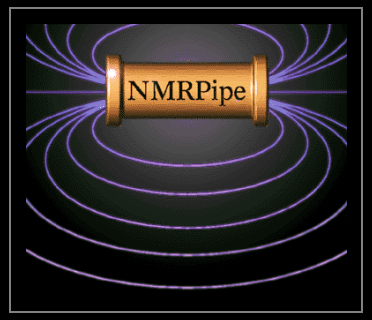
Spectral Processing and Analysis with NMRPipe

frank_delaglio@agilent.com

Frank Delaglio

July 2012



Spectrometer Format Conversion

1D-4D Fourier Transform and Signal Enhancement

Spectral Visualization

1D-4D Peak Detection and Quantification: Position, Amplitude, Width and Modulation/Evolution

Spectral Assignment

Extraction of Structural Parameters

Molecular Structure Calculation

Molecular Display and Structure Verification

Exploitation of Structure

Spectral Imaging

Automation, Batch Analysis, and Screening

Web-Based Server Implementations

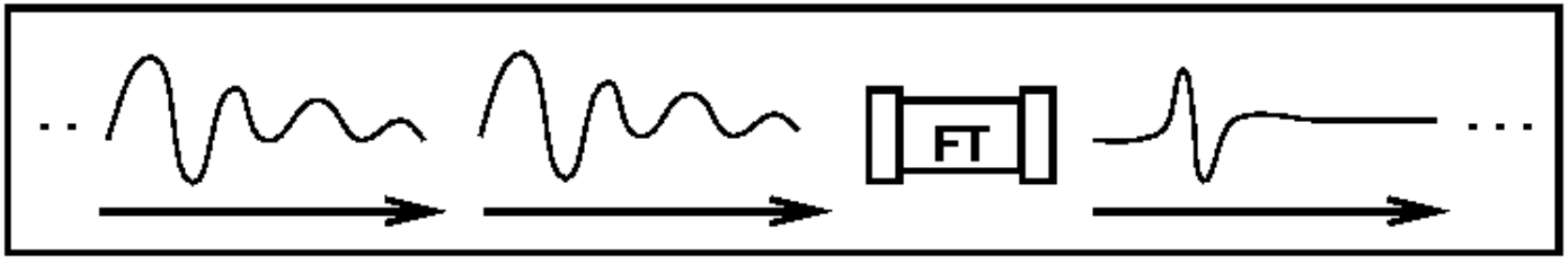
Ad Bax • Joeseph Barchi • James Chou • Gabriel Cornilescu • George Gray • Alex Grishaev
Stephan Grzesiek • Georg Kontaxis • John Kuszewski • Ryan McKay • John Pfeiffer
Ben Ramirez • Michael Shapiro • Tobias Ulmer • Gerteen Vuister • Justin Wu • Jinfa Ying
Shen Yang • Guang Zhu • Edward Zartler

July 2012

frank_delaglio@agilent.com

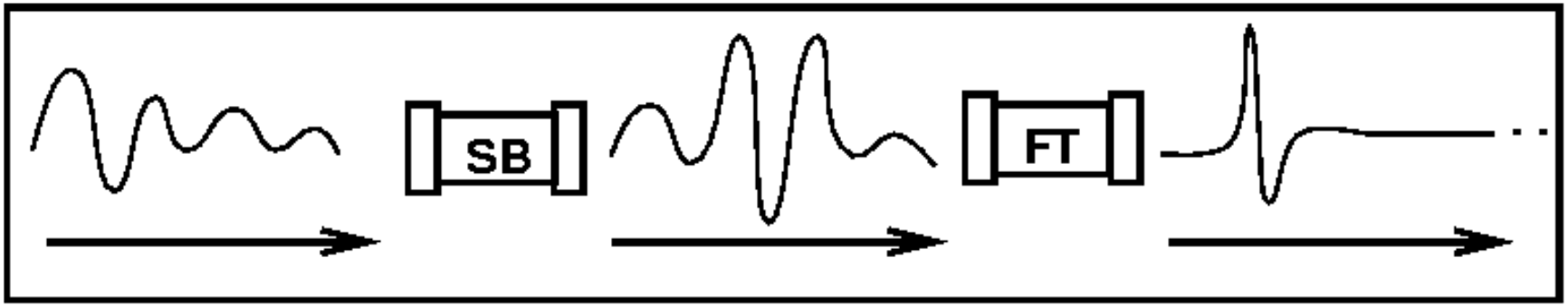
Spectral Processing Function as a UNIX Filter

fourierTransform < in.fid > out.ft

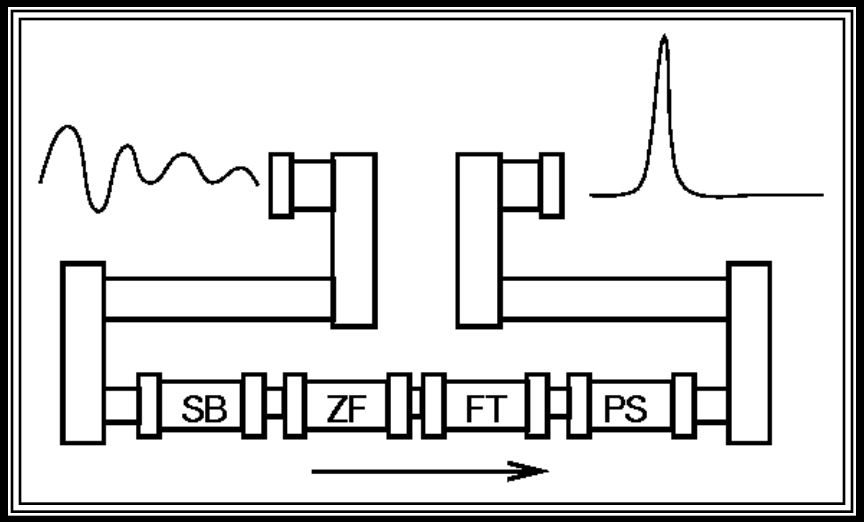


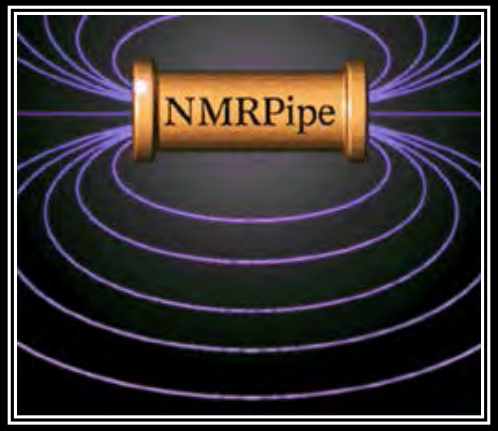
Spectral Processing Scheme as a UNIX Pipeline

sineBell < in.fid | fourierTransform > out.ft



NMRPipe: Spectral Processing as a UNIX Pipeline





1D-4D FT, LP, MEM, ML, PCA

Parallel Processing

1D-4D Peak Detection and Quantification

Spectral Graphics, Strips, Projections

Extensively Customizable

Molecular Structure Calculation

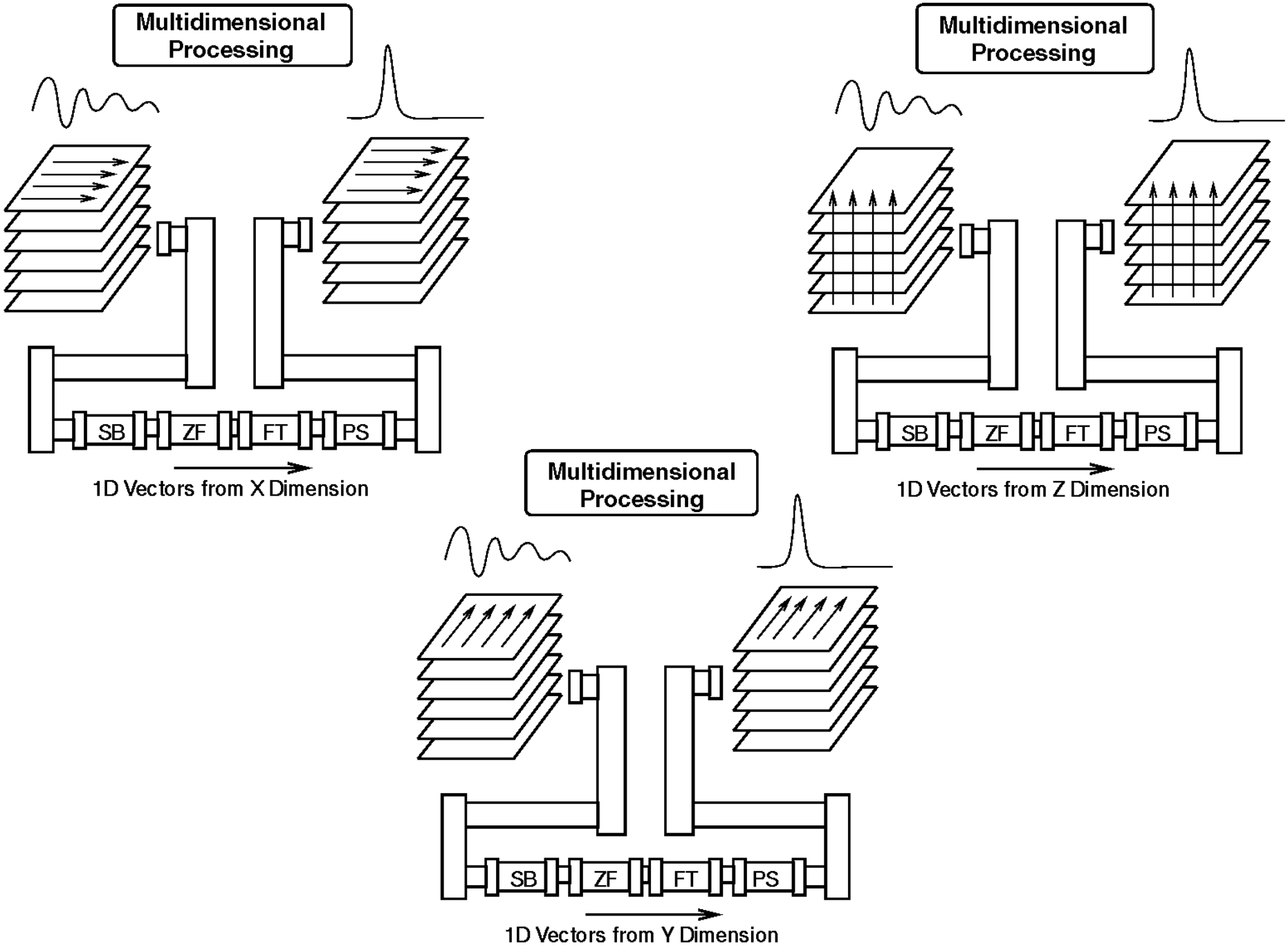
NMR Parameter Calculation (Shifts, Dipolar Couplings, PCS, etc)

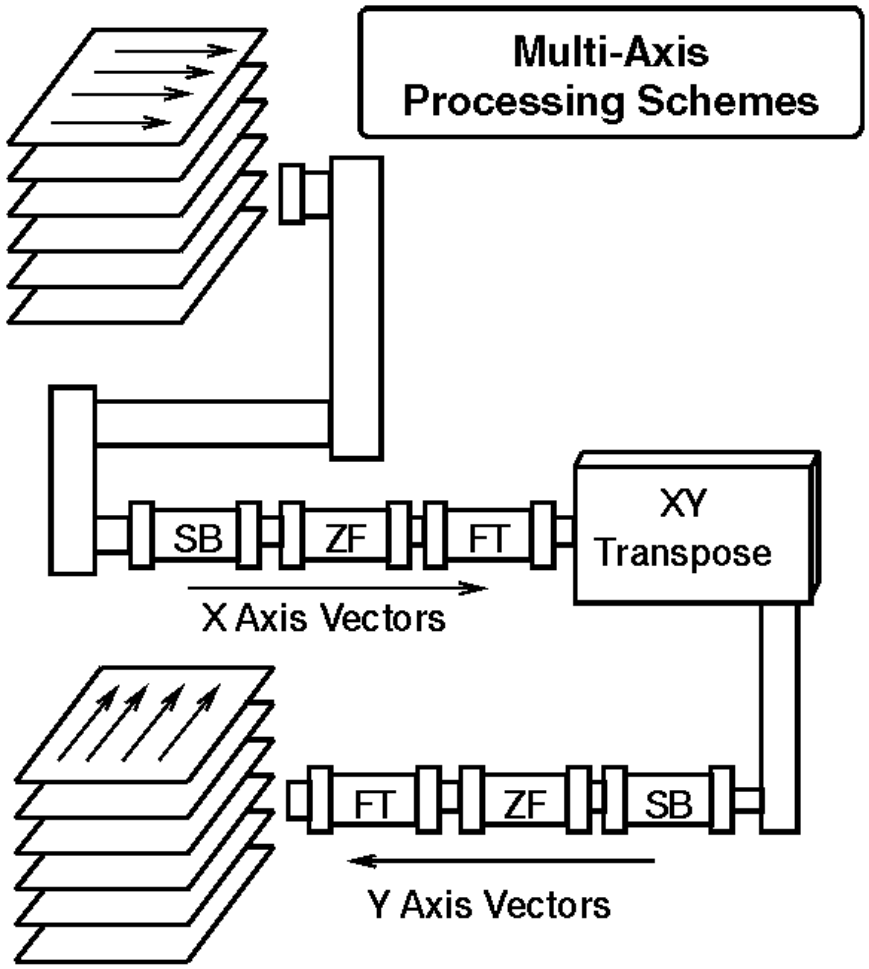
Customization is via standard scripting languages (C-shell, TCL)

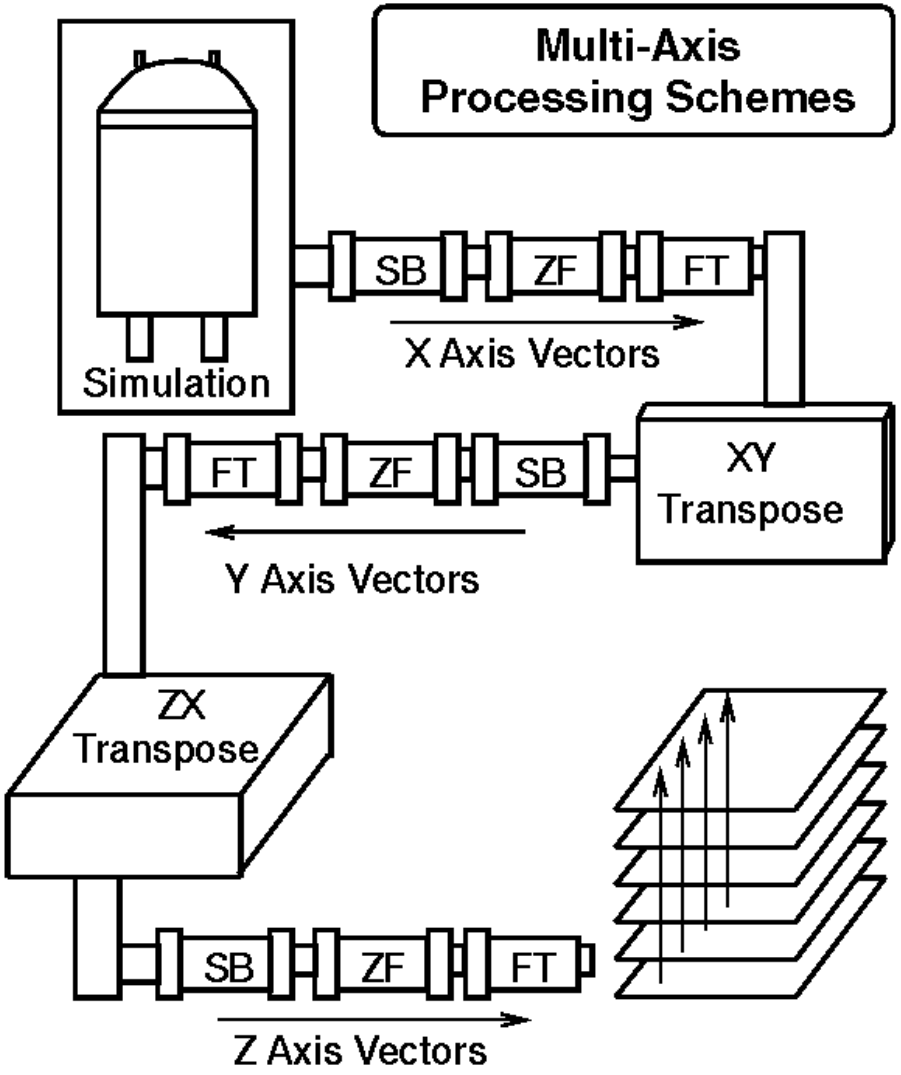
Created and Maintained by one developer, with contributed modules

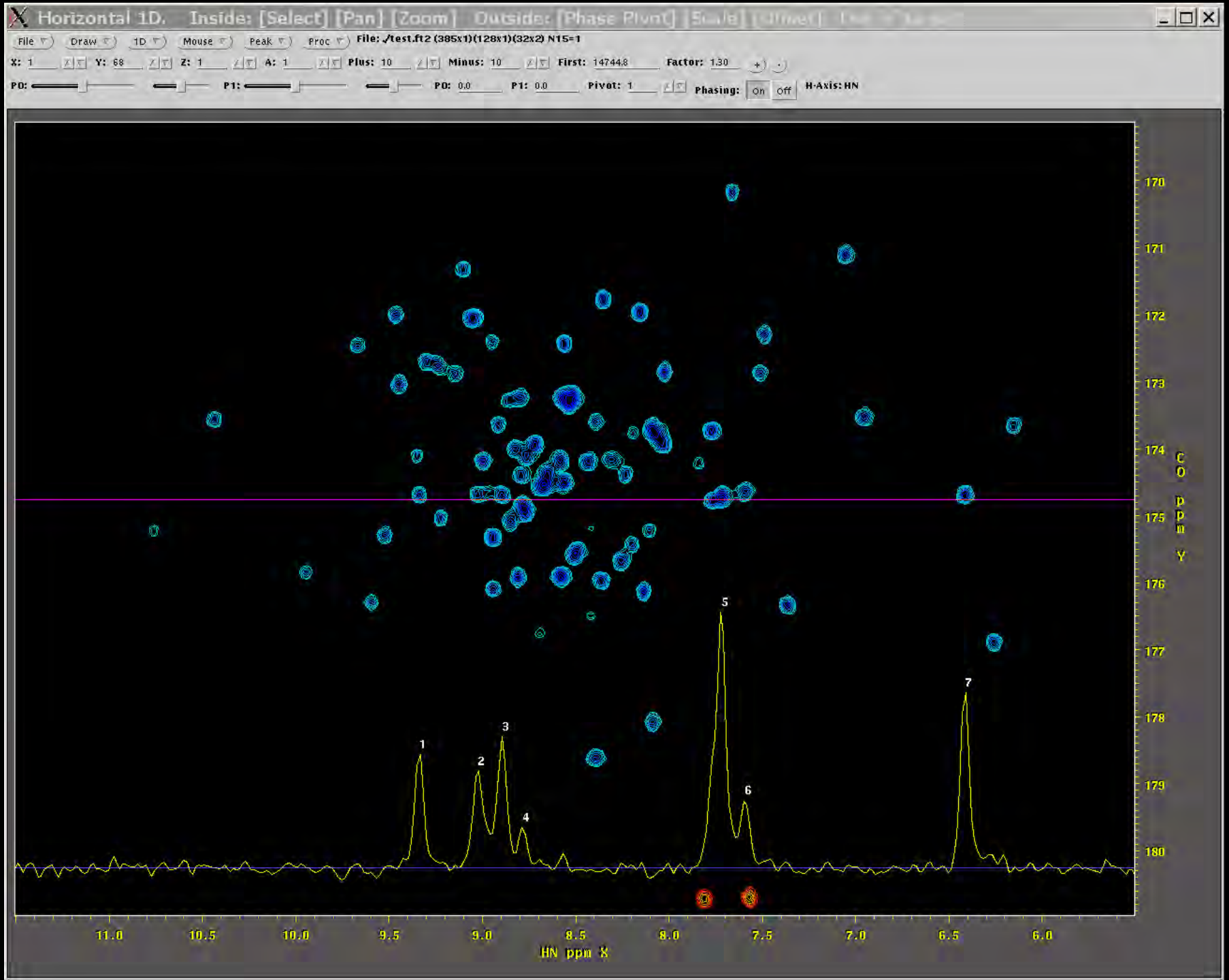
Solaris, IRIX, HP/UX, DEC OSF, IBM AIX, IBM Blue Gene, Convex OS, Cray OS, Mac OS X, Linux, WindowsXP Interix, VMWare Player

Bottom-up Software Design


Multidimensional Processing
SB
ZF
FT
PS
1D Vectors from X Dimension
Multidimensional Processing
SB
ZF
FT
PS
1D Vectors from Z Dimension
Multidimensional Processing
SB
ZF
FT
PS
1D Vectors from Y Dimension

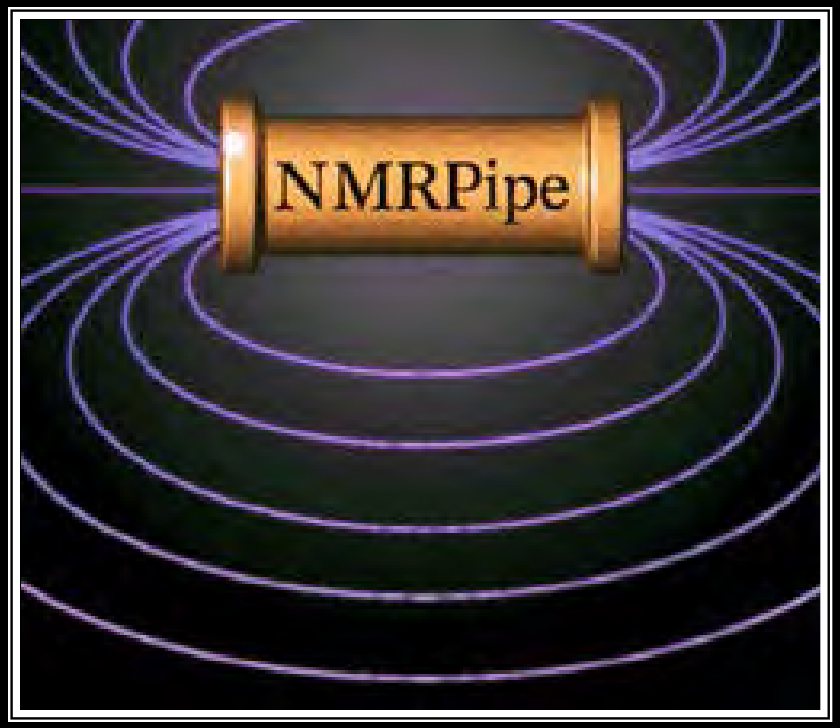

Multi-Axis
Processing Schemes
SB
ZF
FT
XY
Transpose
X Axis Vectors
FT
ZF
SB
Y Axis Vectors

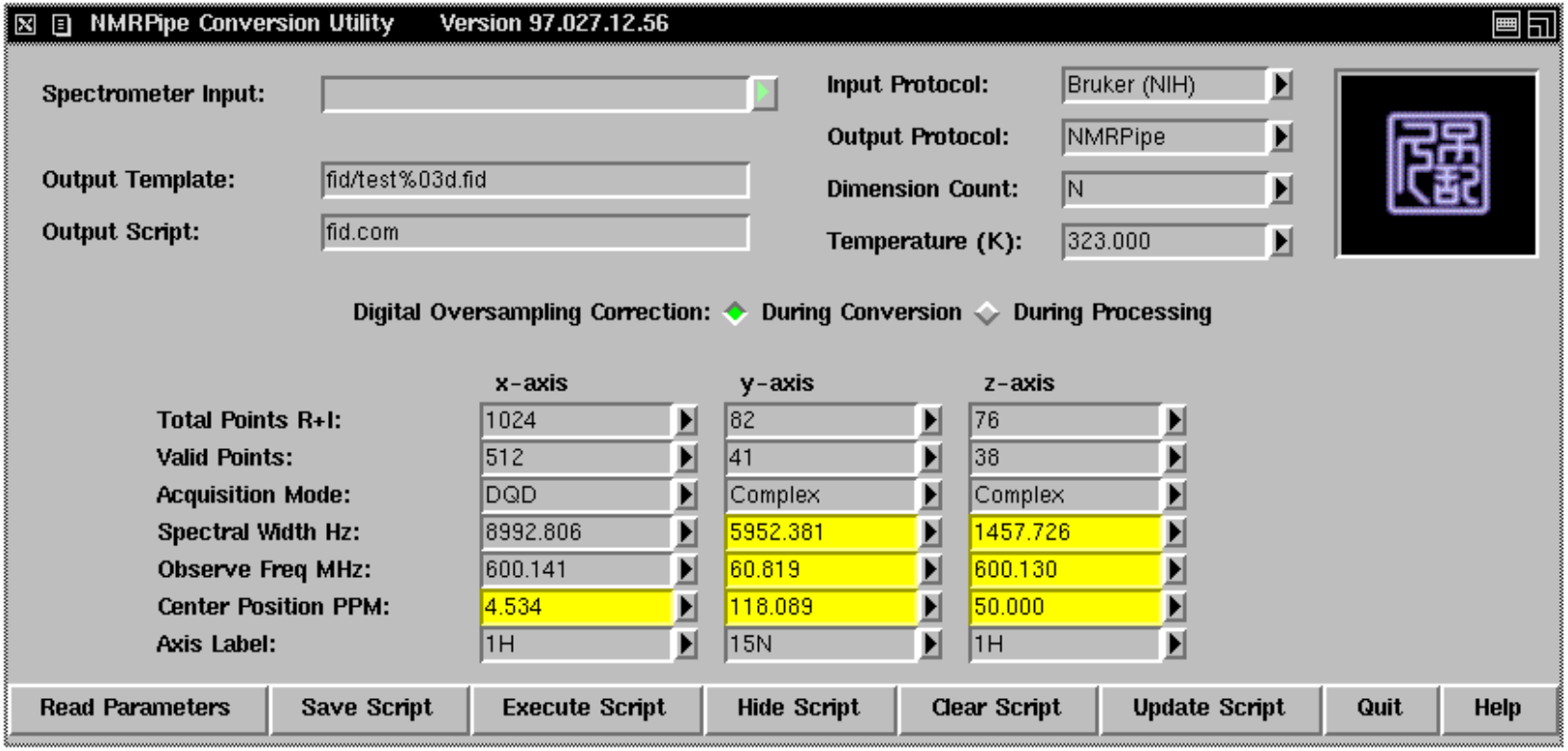

Multi-Axis
Processing Schemes
Simulation
SB
ZF
FT
X Axis Vectors
FT
ZF
SB
XY
Transpose
Y Axis Vectors
ZX
Transpose
SB
ZF
FT
Z Axis Vectors

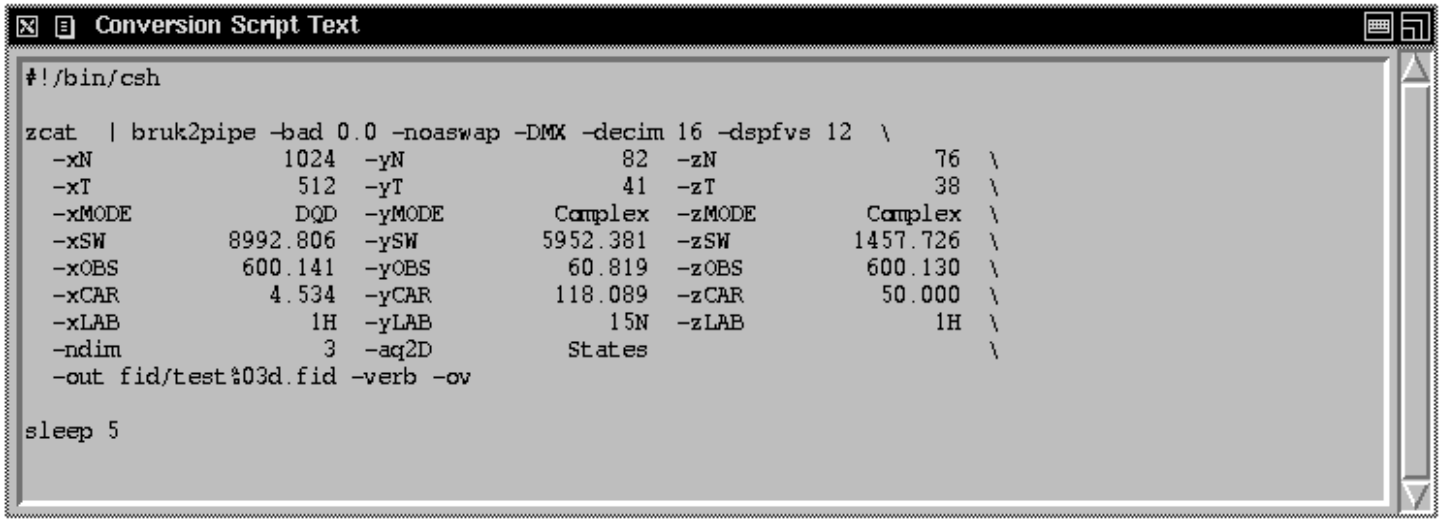

Horizontal 1D. Inside: [Select] [Pan] [Zoom] Outside: [Phase Pivot] [Scale] [Draw]
File
Draw
1D
Mouse
Peak
Proc
File: ./test.ft2 (385x1)(128x1)(32x2) N15=1
X: 1
Y: 68
Z: 1
A: 1
Plus: 10
Minus: 10
First: 14744.8
Factor: 1.30
P0:
P1:
P0: 0.0
P1: 0.0
Pivot: 1
Phasing:
On
Off
H-Axis: HN
1
2
3
4
5
6
7
170
171
172
173
174
175
176
177
178
179
180
C O ppm Y
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
HN ppm X

NMRWish

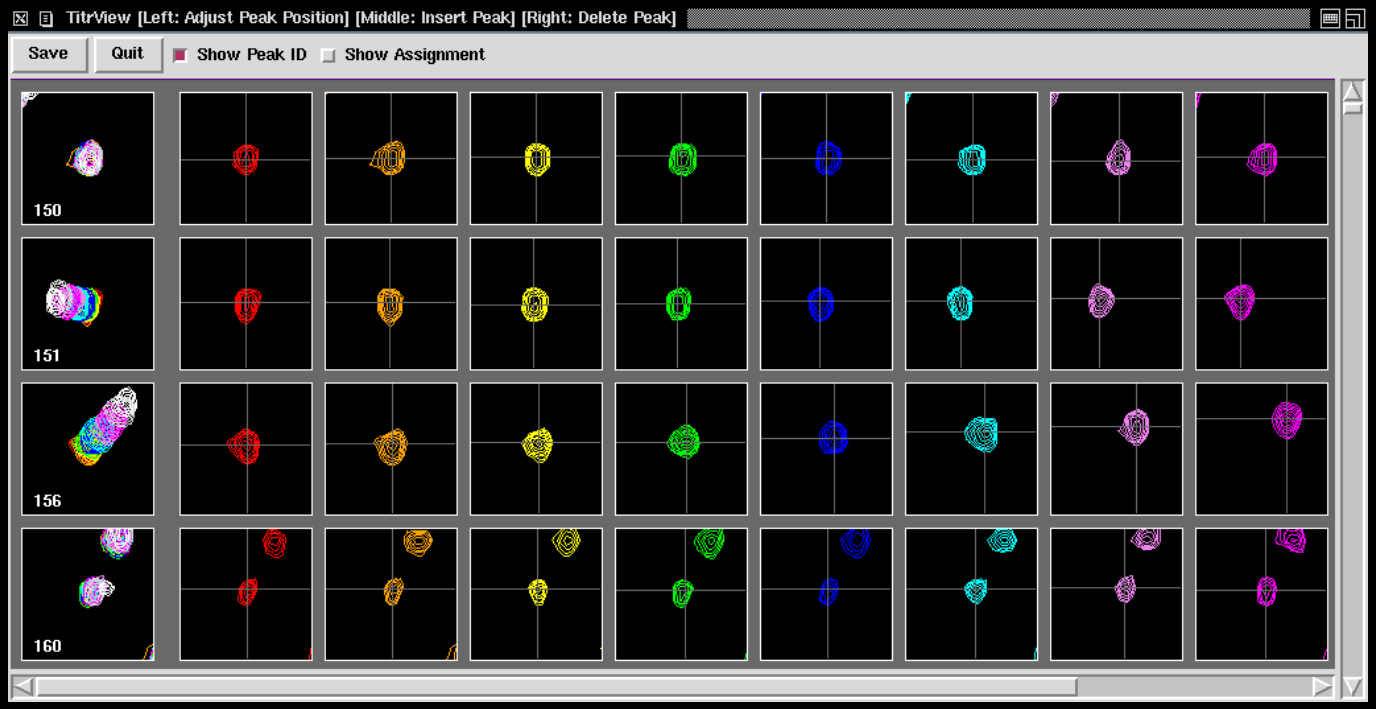
- Customized version of TCL/TK “wish” interpreter
- Script-based Interactive Spectral graphics (multi-window and PostScript)
- Generic Database Engine (GDB)
- Manipulate Peak Data, Assignments, NMR Parameters, and Molecular Structure

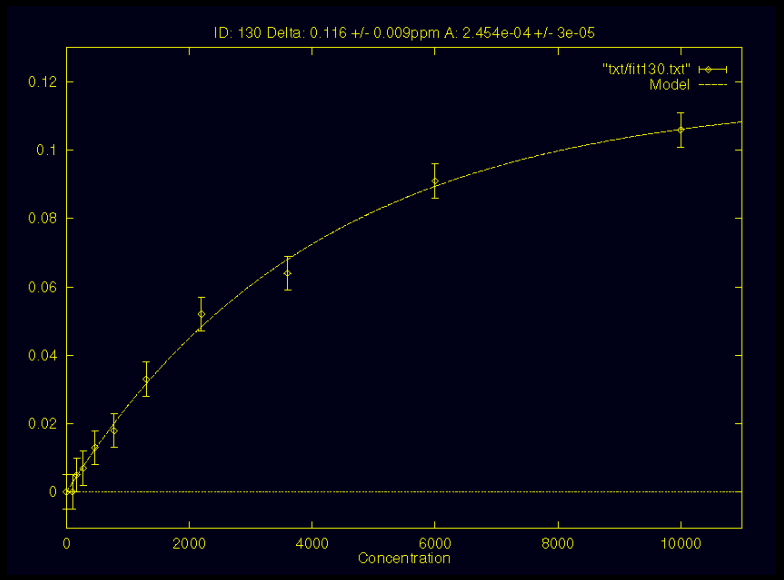


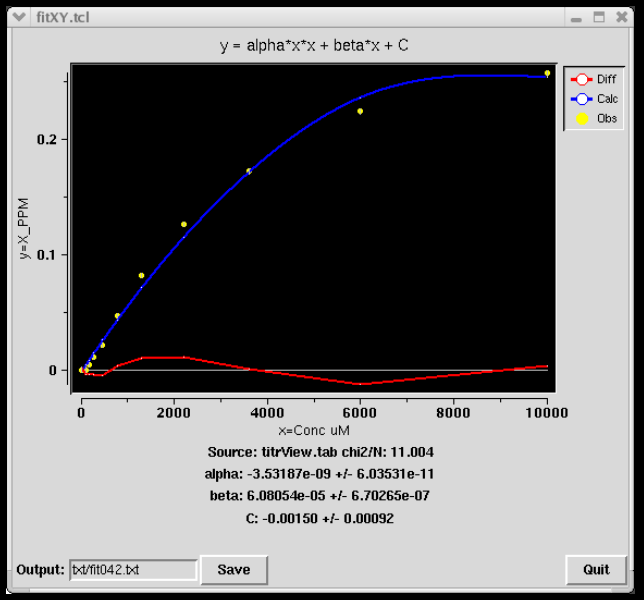

NMRPipe Conversion Utility Version 97.027.12.56
Spectrometer Input:
Output Template: fid/test%03d.fid
Output Script: fid.com
Input Protocol: Bruker (NIH)
Output Protocol: NMRPipe
Dimension Count: N
Temperature (K): 323.000
Digital Oversampling Correction: During Conversion During Processing
x-axis y-axis z-axis
Total Points R+I: 1024 82 76
Valid Points: 512 41 38
Acquisition Mode: DQD Complex Complex
Spectral Width Hz: 8992.806 5952.381 1457.726
Observe Freq MHz: 600.141 60.819 600.130
Center Position PPM: 4.534 118.089 50.000
Axis Label: 1H 15N 1H
Read Parameters
Save Script
Execute Script
Hide Script
Clear Script
Update Script
Quit
Help


Conversion Script Text
#!/bin/csh
zcat | bruk2pipe -bad 0.0 -noaswap -DMX -decim 16 -dspfvs 12 \
-xN 1024 -yN 82 -zN 76 \
-xT 512 -yT 41 -zT 38 \
-xMODE DQD -yMODE Complex -zMODE Complex \
-xSW 8992.806 -ySW 5952.381 -zSW 1457.726 \
-xOBS 600.141 -yOBS 60.819 -zOBS 600.130 \
-xCAR 4.534 -yCAR 118.089 -zCAR 50.000 \
-xLAB 1H -yLAB 15N -zLAB 1H \
-ndim 3 -aq2D States \
-out fid/test%03d.fid -verb -ov
sleep 5

Analyze Titration Curve to Estimate K_d

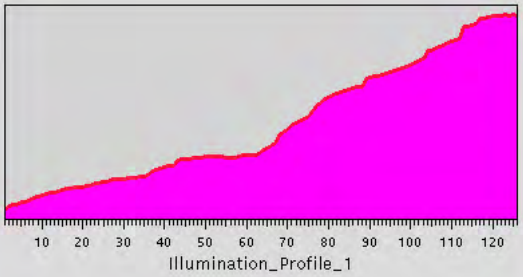


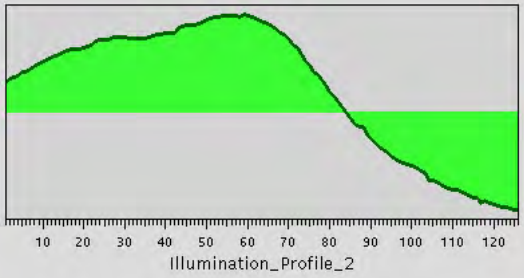


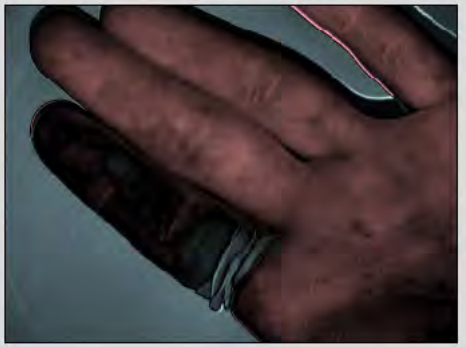


X=288 Y=146 Z=76 76.000 Pts I=747.000

File Acquire Process Help



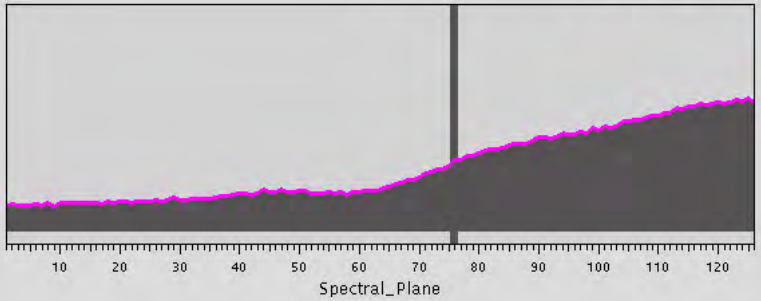


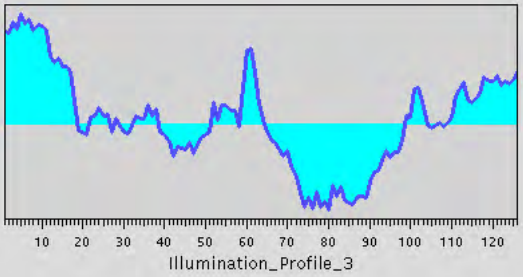


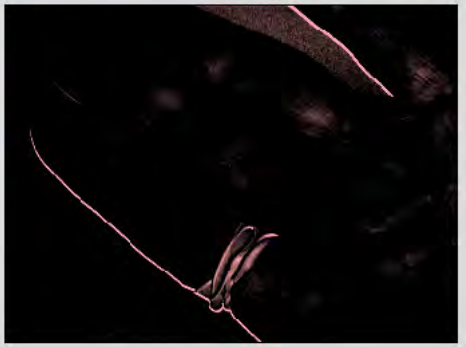
0.0 Brightness Level Min

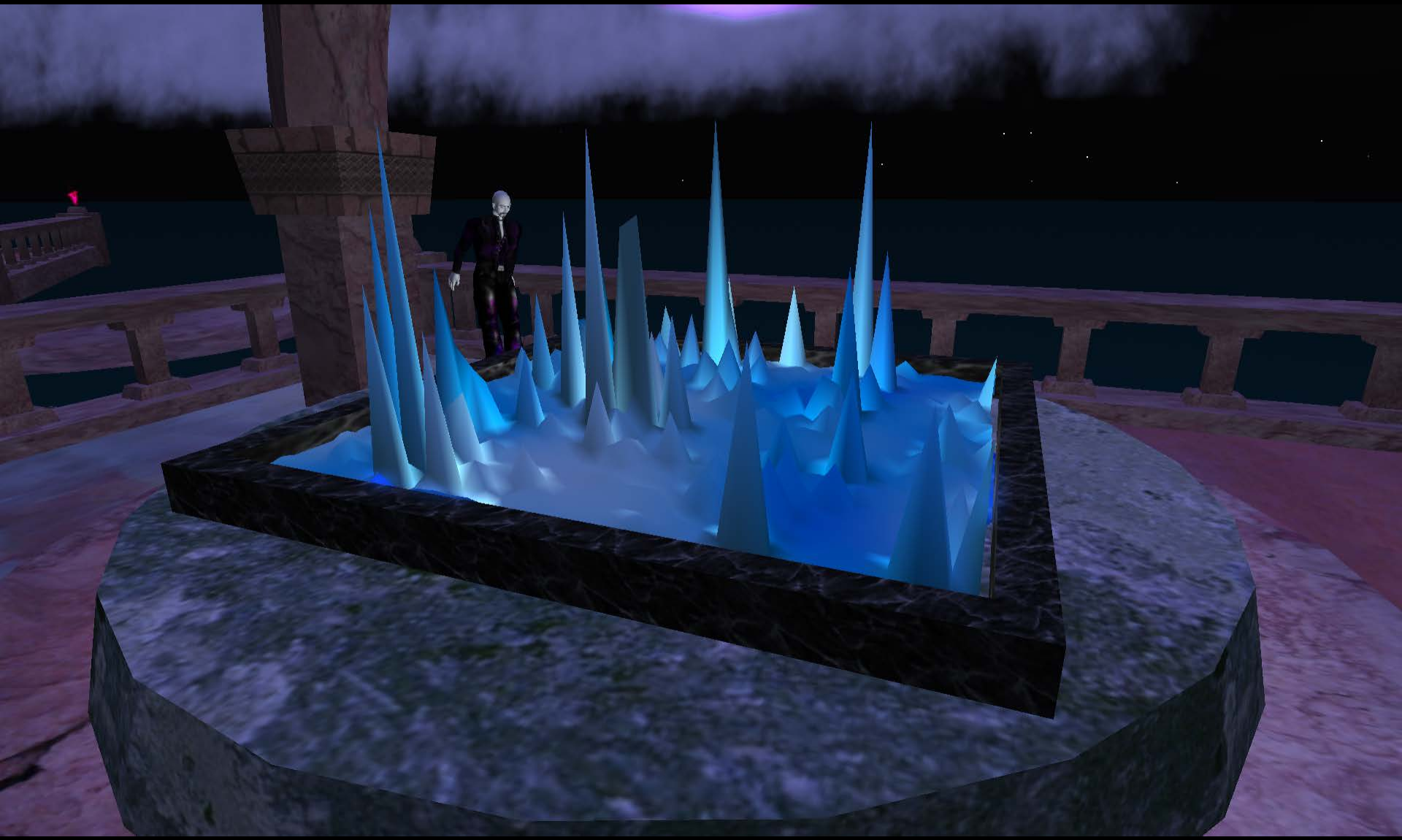
100.0 Brightness Level Max

76 Spectral Plane





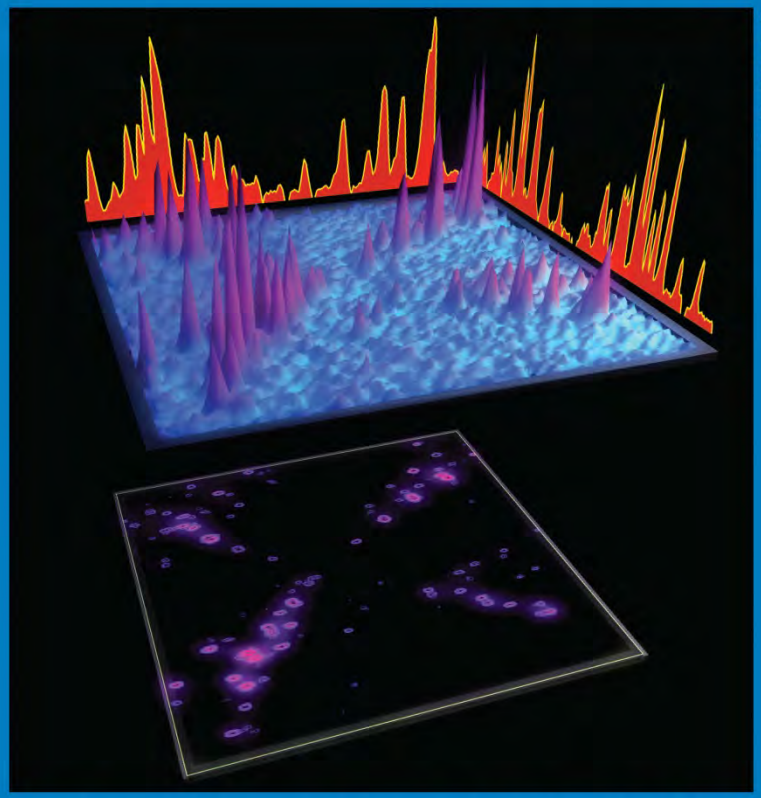




JPCAFH

NOVEMBER 4, 2010
VOLUME 114
NUMBER 43
pubs.acs.org/JPCA

THE JOURNAL OF PHYSICAL CHEMISTRY A



Coherent 2D Spectrum of Nitrogen Dioxide Showing Peaks That Form an X-Shaped Cluster (see page 8A)

DYNAMICS, KINETICS, ENVIRONMENTAL CHEMISTRY, SPECTROSCOPY, STRUCTURE, THEORY

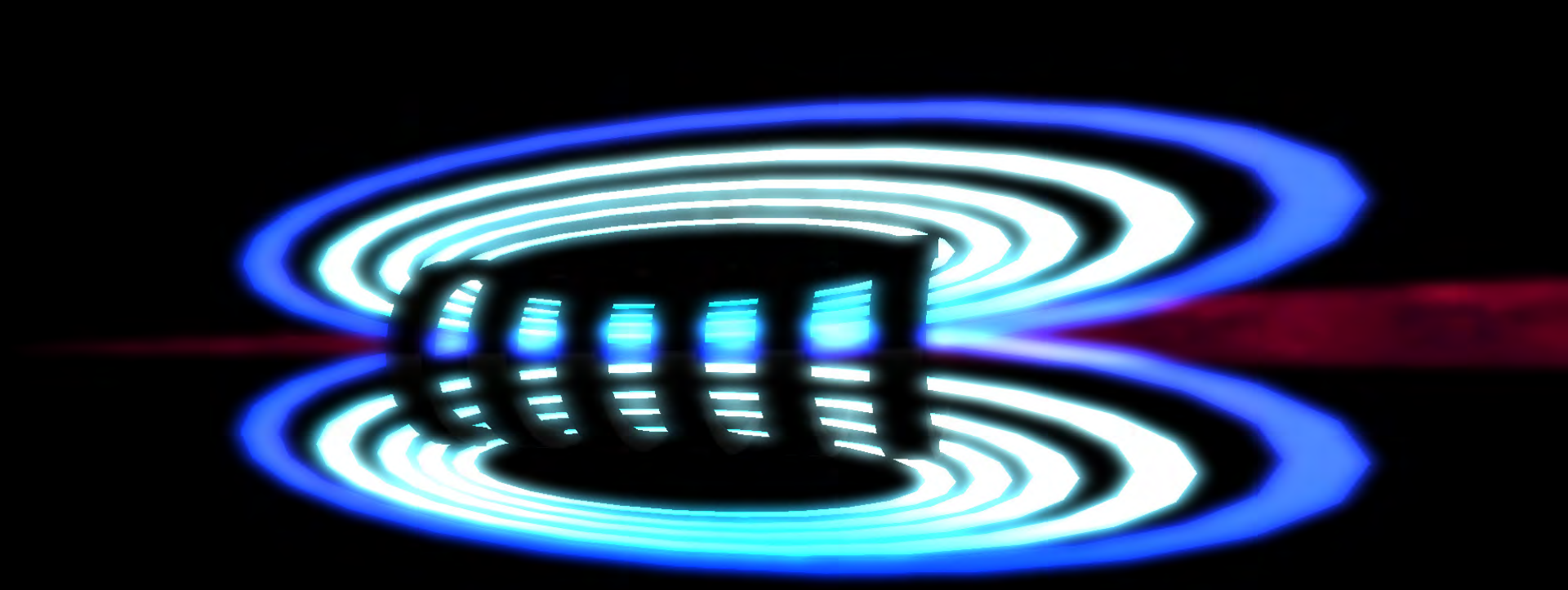
ACS Publications
High quality. High impact.

www.acs.org

About the Cover

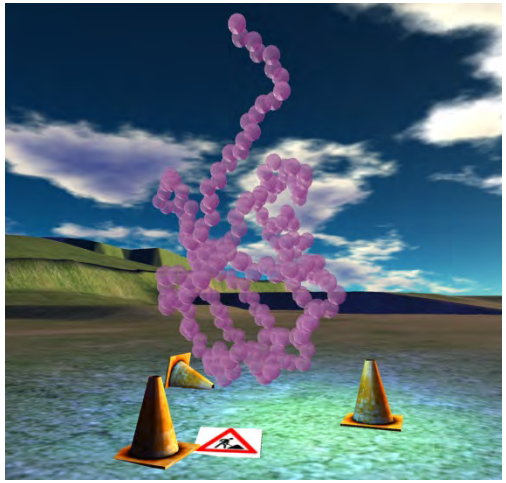
November 4, 2010: Vol. 114, Iss. 43

Conventional (1D) electronic spectra of NO_2 are largely patternless, but coherent 2D spectra of NO_2 show numerous X-shaped clusters such as that shown here. This spectrum is displayed in 2D contour form (below, in purple and black) and as a corresponding 3D intensity surface (above, in purple and blue). Orthogonal 1D projections (orange) of the surface resemble conventional 1D spectra. The spectral data was preprocessed using a pipeline-based software system for multidimensional spectroscopy (Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277—293), which was also used to generate images used to build the 3D scene. The figure was produced by Frank Delaglio (www.nmrscience.com), who used the interactive virtual reality environment Second Life to render this image (see page 11365).



Bax Group NMR Calculation Servers

http://spin.niddk.nih.gov/bax/nmrserver



NMRPipe: Related Programs and Features

TALOS+: prediction of protein backbone phi,psi torsions From chemical shifts.

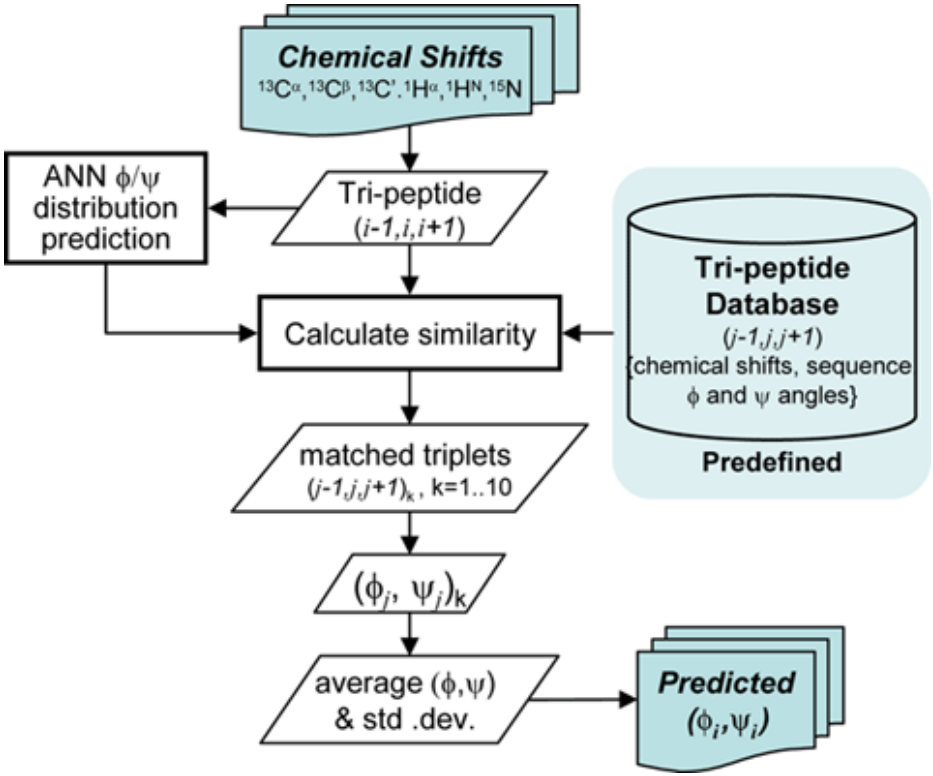
SPARTA+: prediction of protein backbone chemical shifts from structure.

PROMEGA: prediction of PRO cis-trans conformation from backbone chemical shifts.

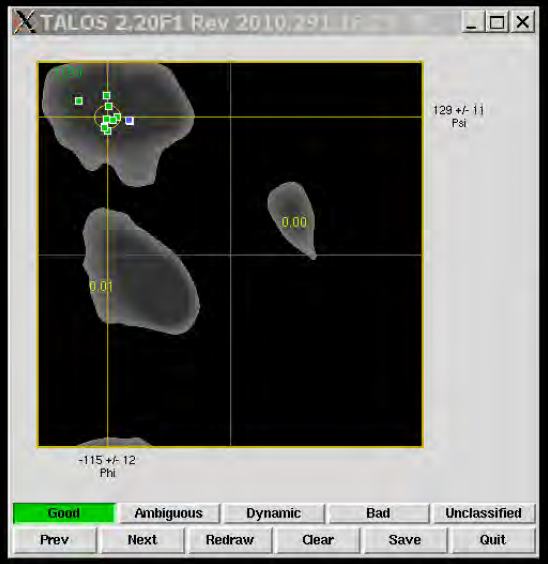
DYNAMO/PDBUTIL: simple structure calcuation, utilities to add protons, create extended structures, etc.

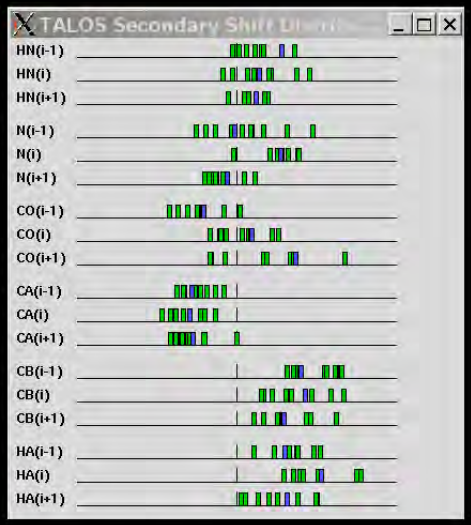
DC: manipulation of Dipolar Couplings and NMR Homology Search (MFR).

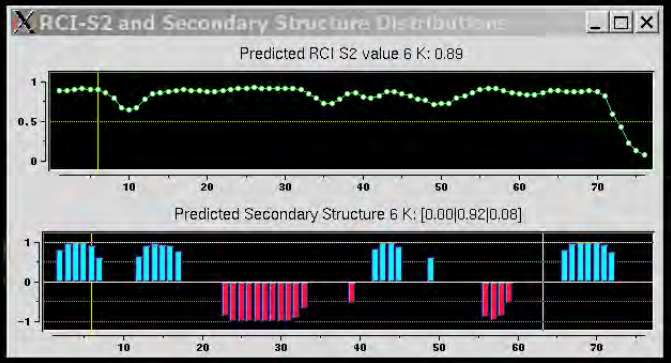
AXES: prediction of Small-Angle X-Ray Scattering from PDB

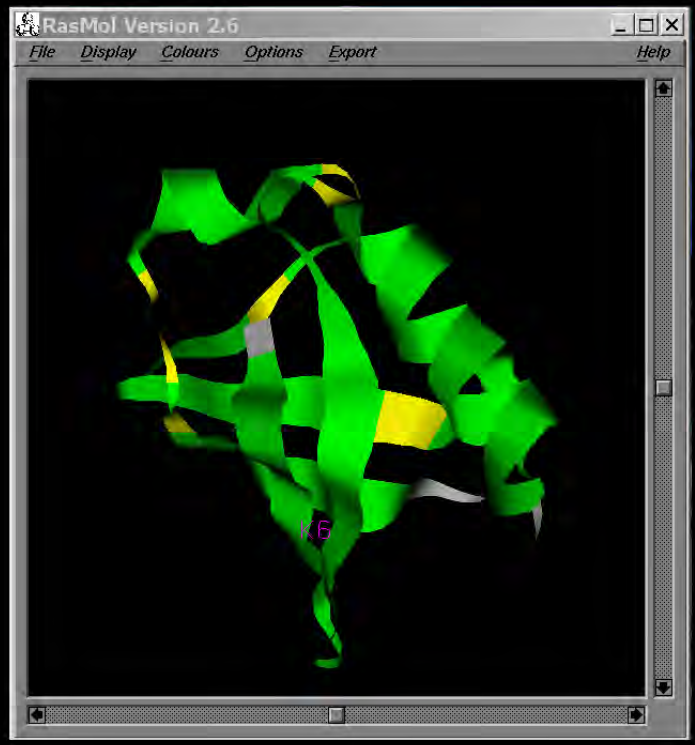


***TALOS+* predicts protein backbone torsions from chemical shifts. It can form predictions for 88% of residues on average, with an RMS of better than 15 degrees.**



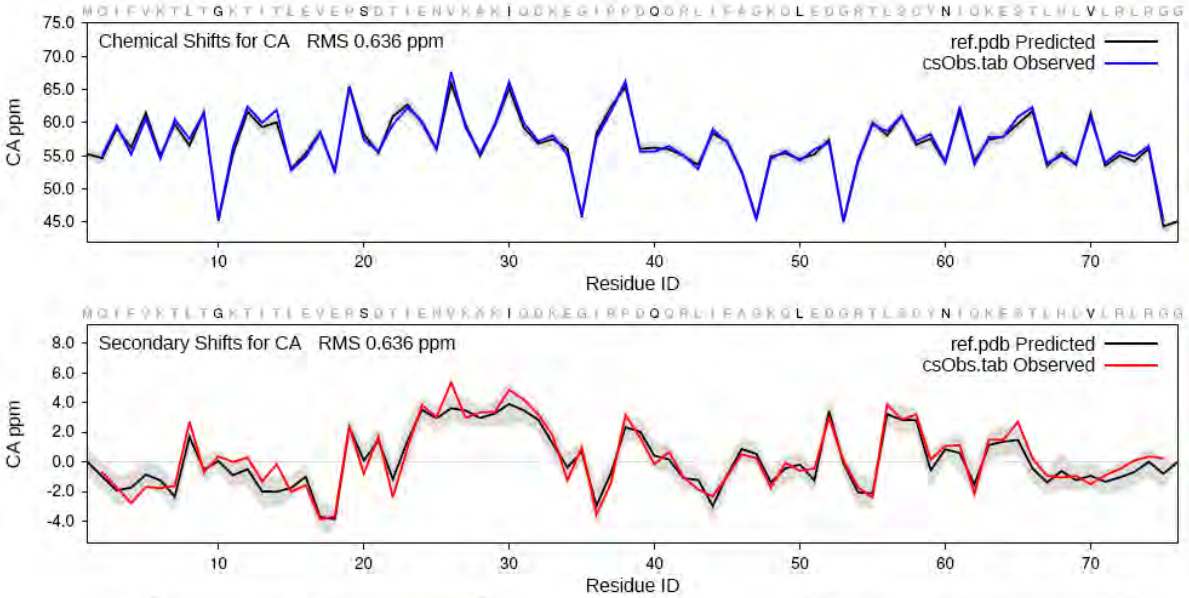


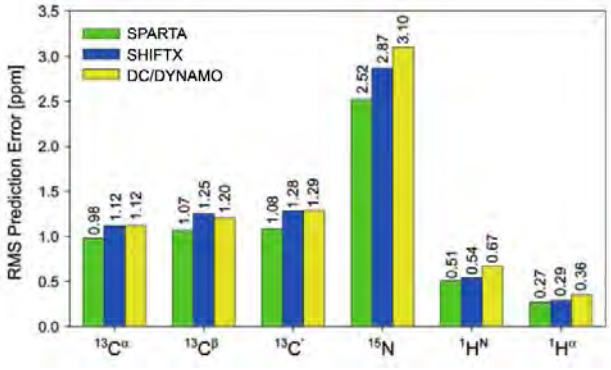


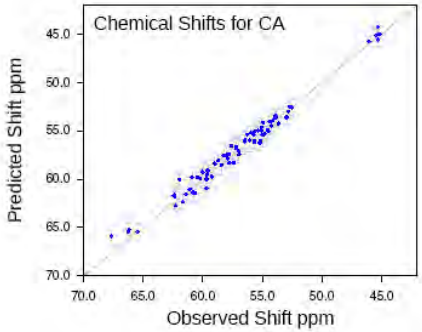


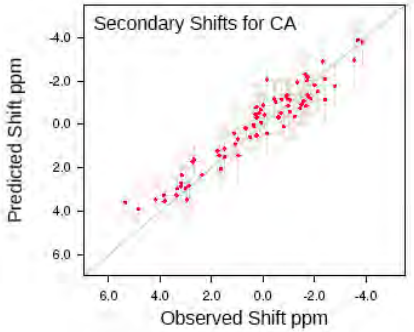
Agilent Technologies

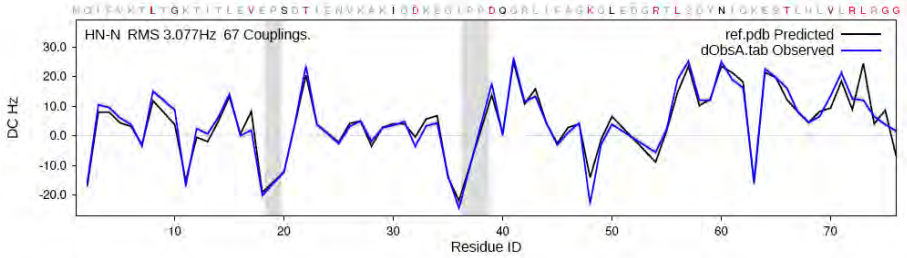

ΣΠΑΡΤΑ+

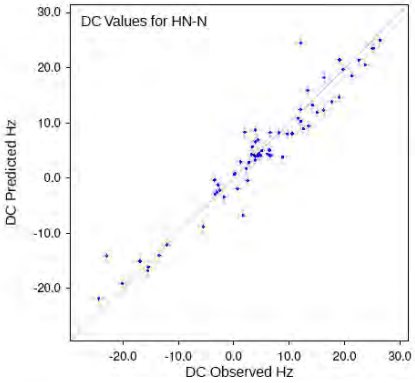

Chemical Shifts for CA RMS 0.636 ppm
ref.pdb Predicted
csObs.tab Observed
CA ppm
Residue ID
Secondary Shifts for CA RMS 0.636 ppm
ref.pdb Predicted
csObs.tab Observed
CA ppm
Residue ID

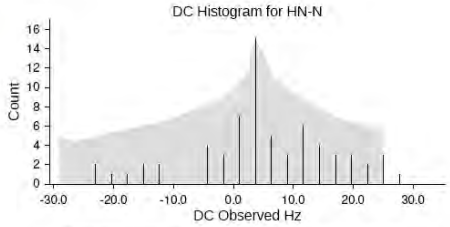

SPARTA
SHIFTX
DC/DYNAMO
RMS Prediction Error [ppm]
0.98
1.12
1.12
1.07
1.25
1.20
1.08
1.28
1.29
2.52
2.87
3.10
0.51
0.54
0.67
0.27
0.29
0.36
13Cα
13Cβ
13C'
15N
1HN
1Hα


Chemical Shifts for CA
Predicted Shift ppm
Observed Shift ppm


Secondary Shifts for CA
Predicted Shift ppm
Observed Shift ppm



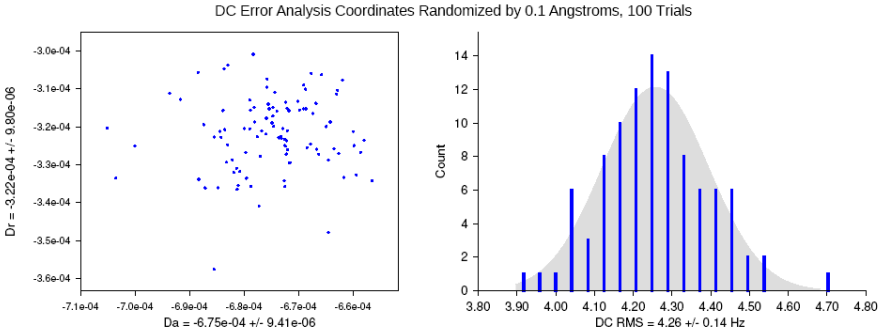


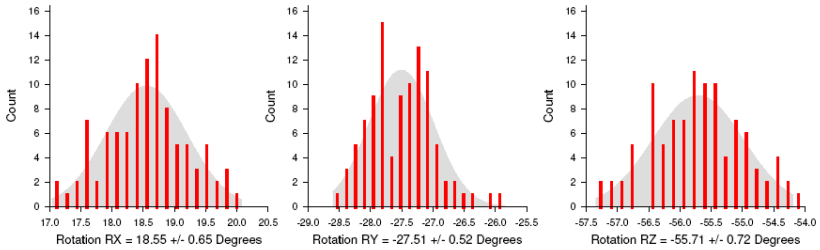


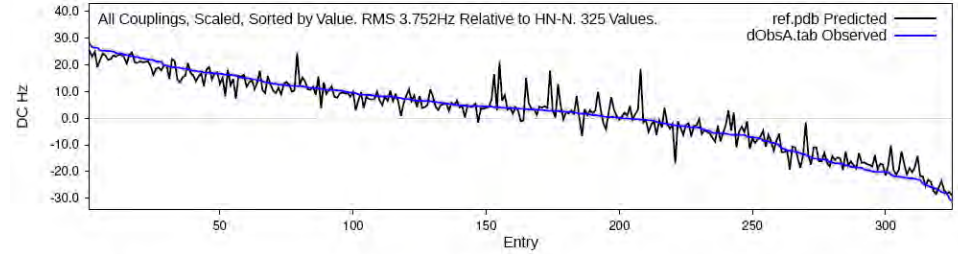
Tensor Parameters:
Da: -6.8589e-04 Dr: -3.2939e-04 GMag: 2.1011e-03 Rh: 0.480

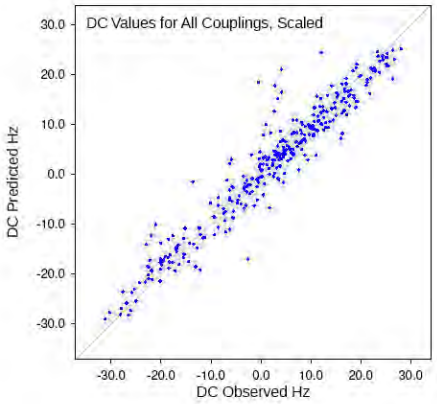
Euler Rotations, Degrees: X: 18.53 Y: -27.60 Z": -55.64

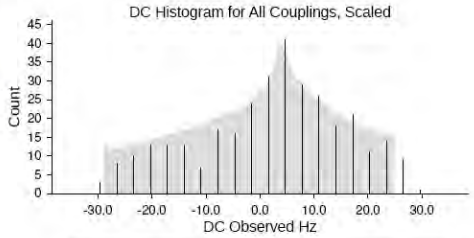
Order Matrix Values:
-8.7887e-04 -1.0887e-04 -6.7871e-04 -7.4130e-04 -2.5099e-04
sXX: 1.9180e-04 sYY: 1.1800e-03 sZZ: -1.3718e-03











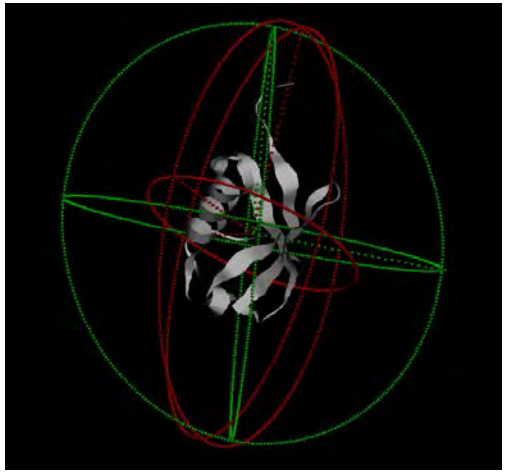
Tensor Parameters:
Da: -6.8589e-04 Dr: -3.2939e-04 GMag: 2.1011e-03 Rh: 0.480

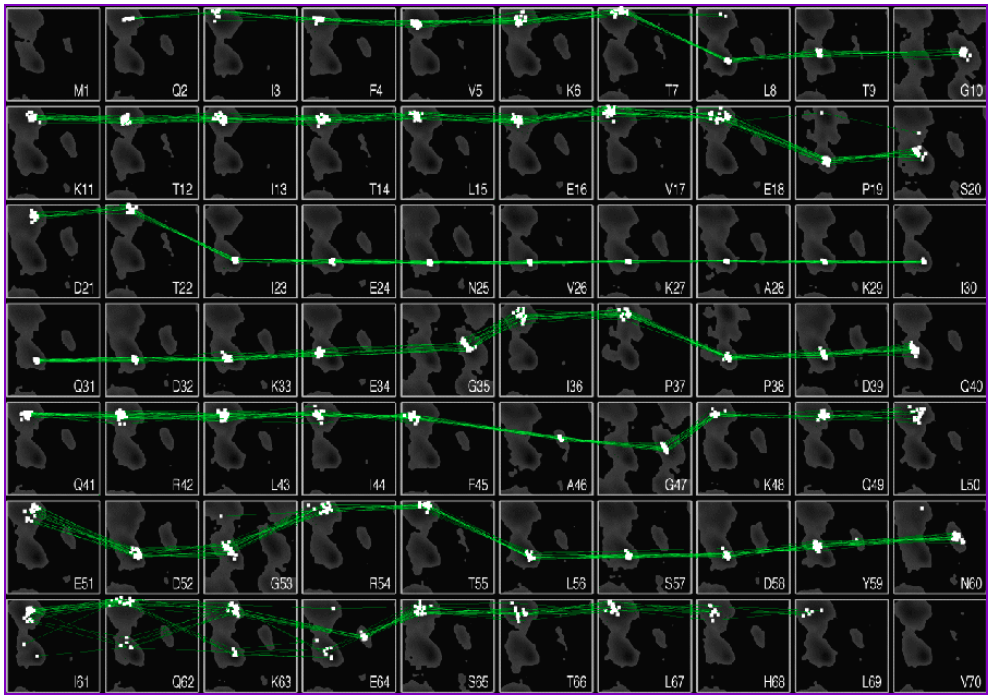
Euler Rotations, Degrees: X: 18.53 Y: -27.60 Z": -55.64

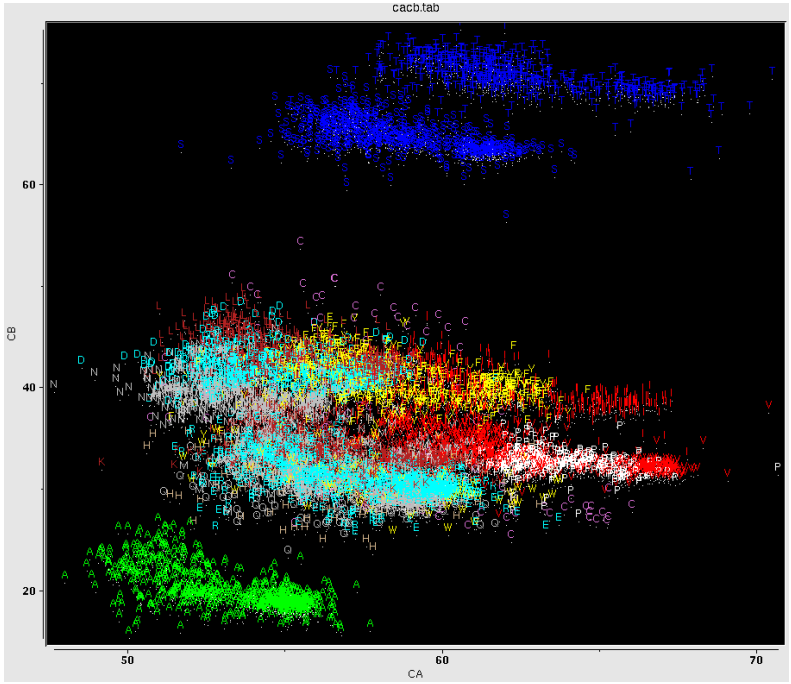
Order Matrix Values:
-8.7887e-04 -1.0887e-04 -6.7871e-04 -7.4130e-04 -2.5099e-04
sXX: 1.9180e-04 sYY: 1.1800e-03 sZZ: -1.3718e-03

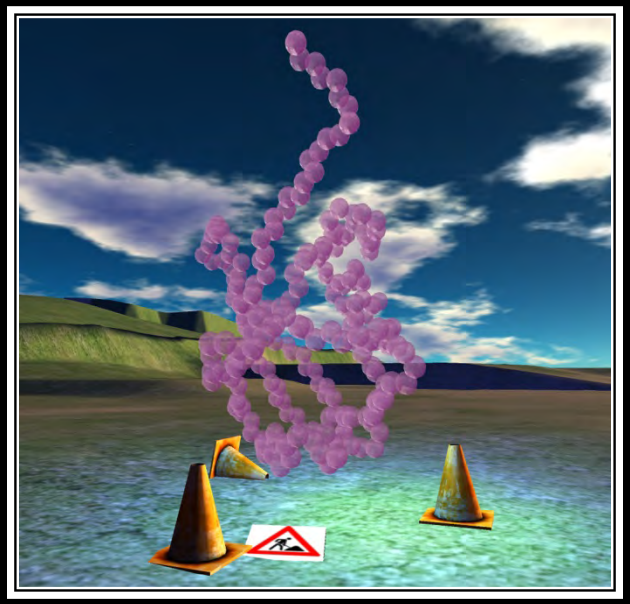
Tensor and NMR Homology Search Utilities

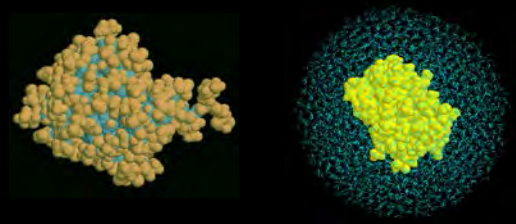
- Rotate PDB onto Tensor Frame
- Molecular Fragment Replacement
- Amino Acid Type by Chemical Shift



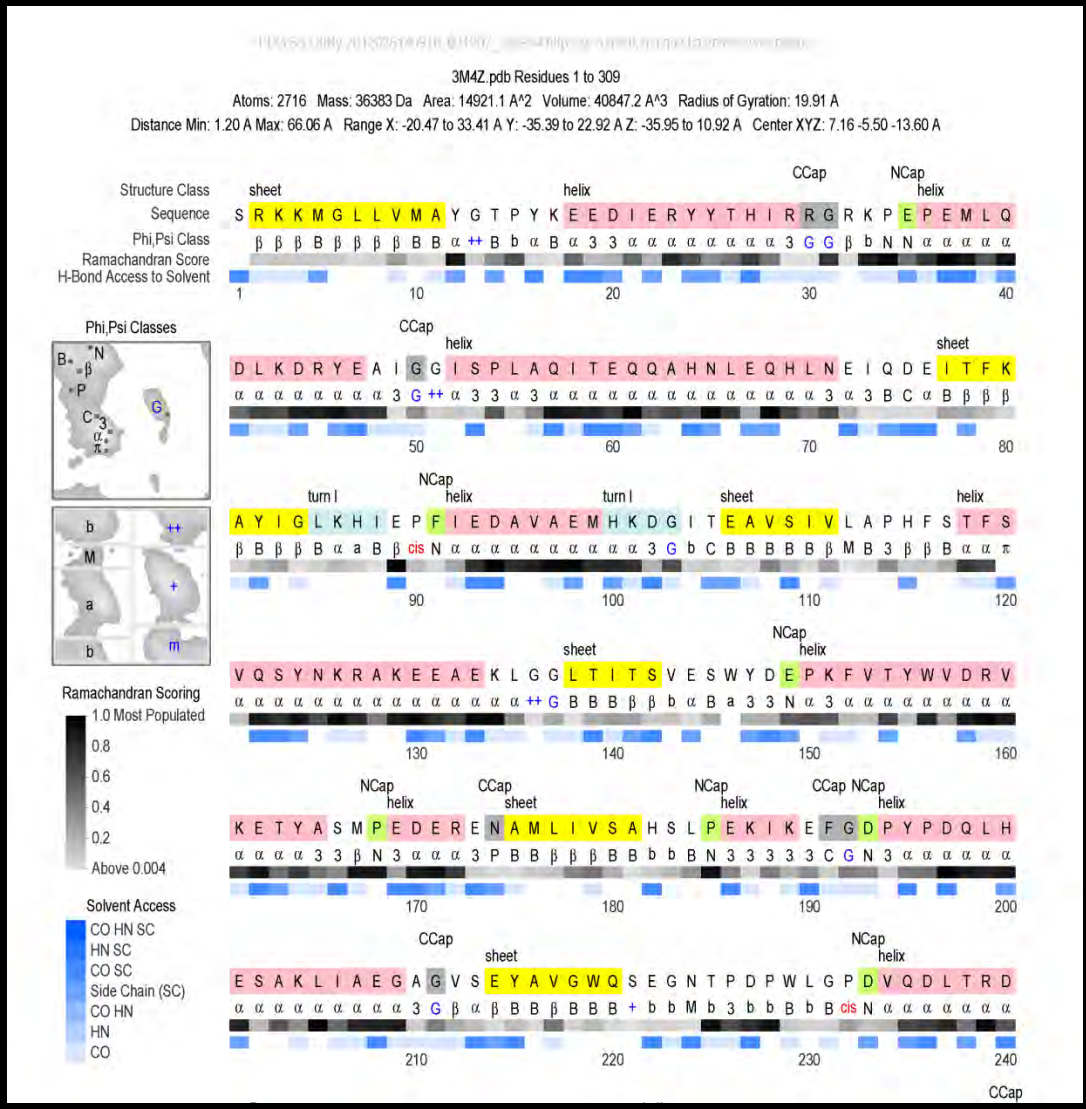






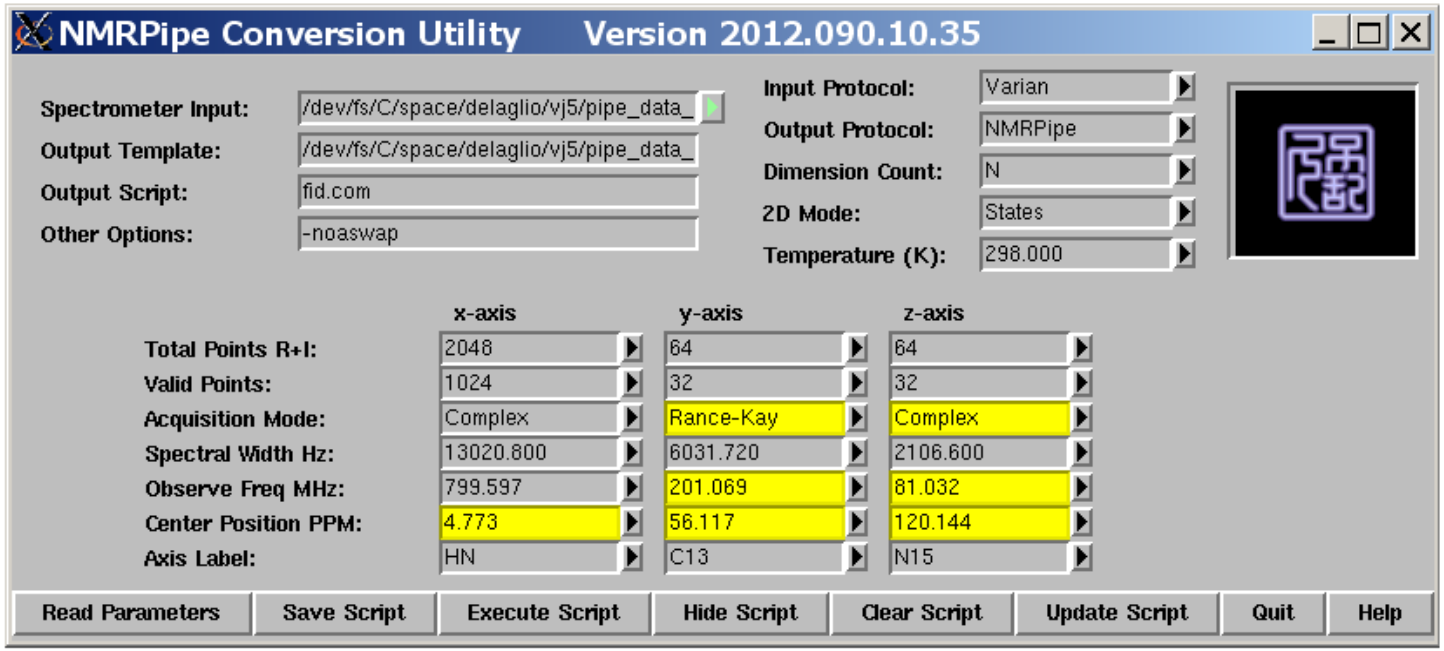


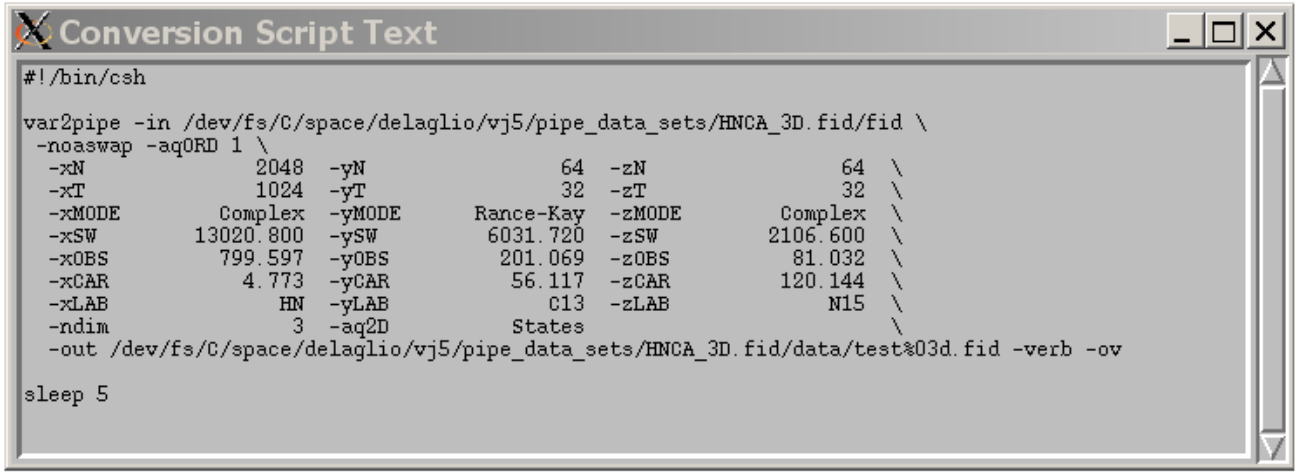
- Create Extended Structure
- Add Protons
- Transformations of PDB Coordinates
- List Secondary Structure, H-Bonds
- Mass, Volume, Surface Area
- Simple simulated annealing



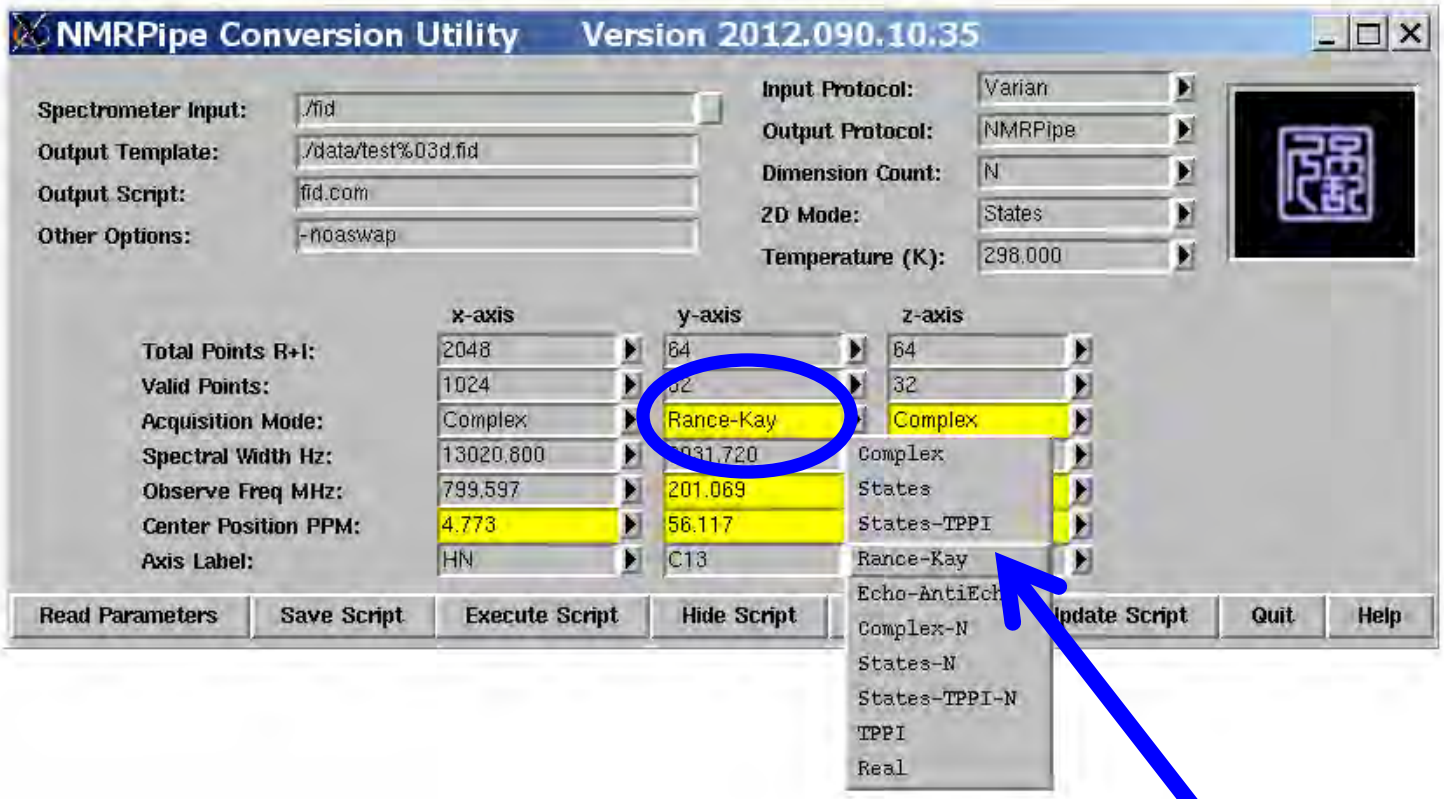
NMRPipe and VnmrJ

NMRPipe – the Old Way



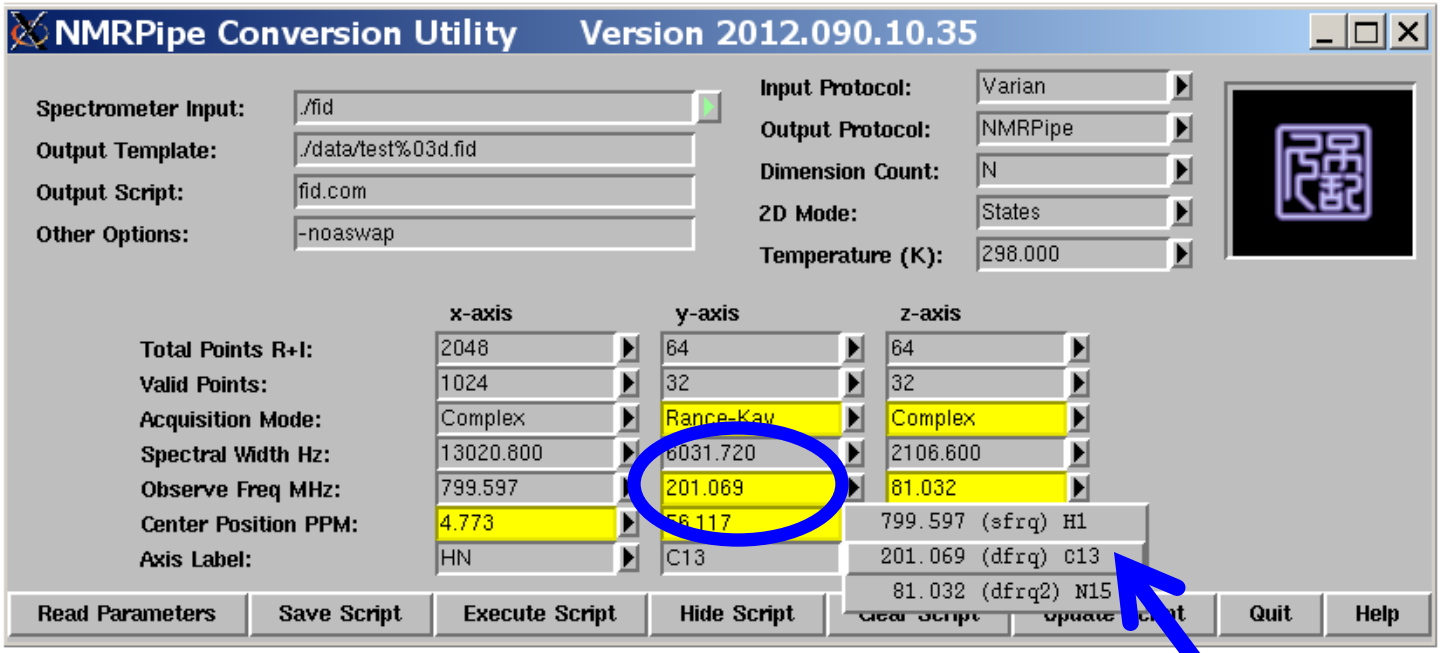


NMRPipe – the Old Way



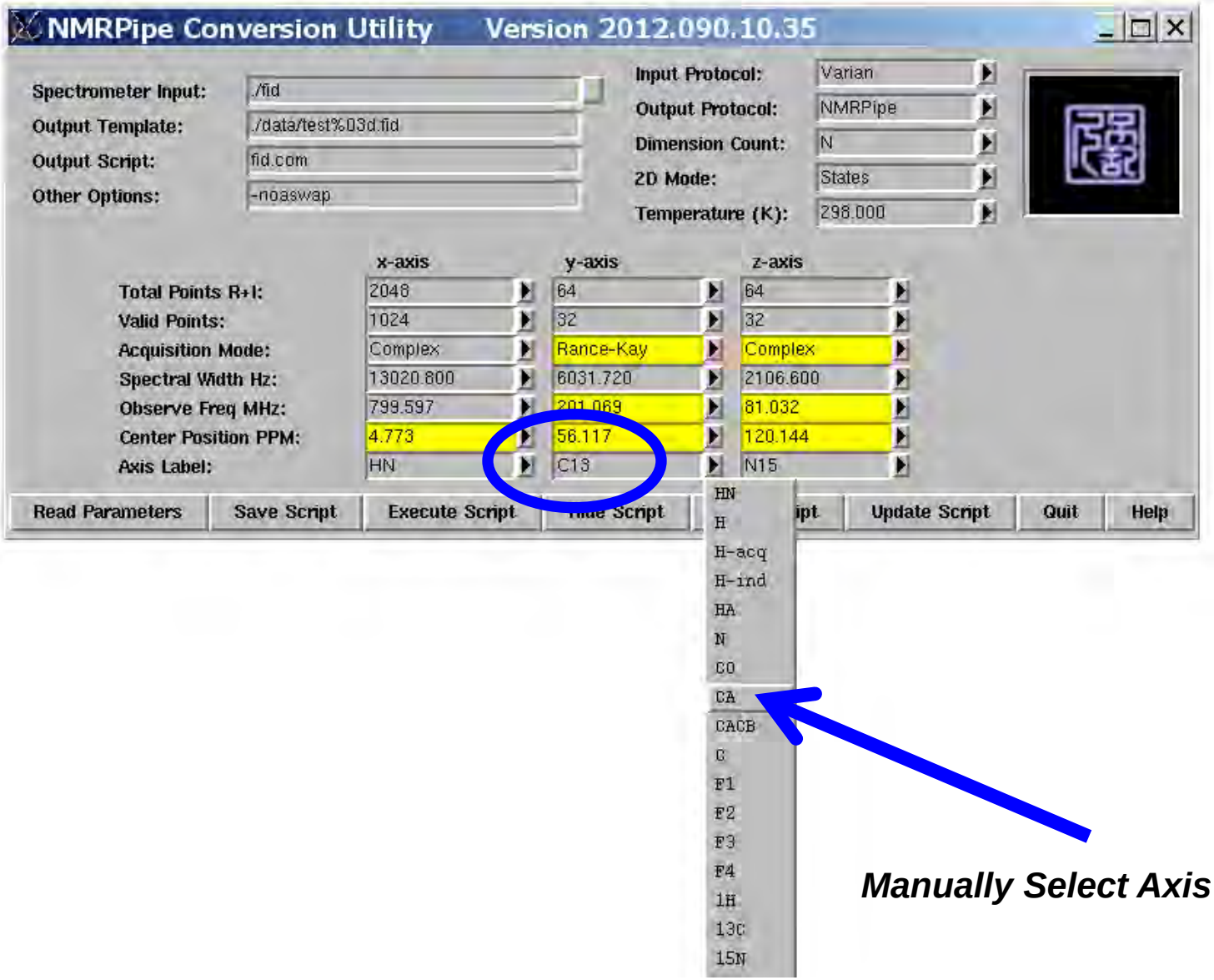
Manually Select Acquisition Mode

NMRPipe – the Old Way



Manually Select Nucleus

NMRPipe – the Old Way



Manually Select Axis Label

NMRPipe – the Old Way

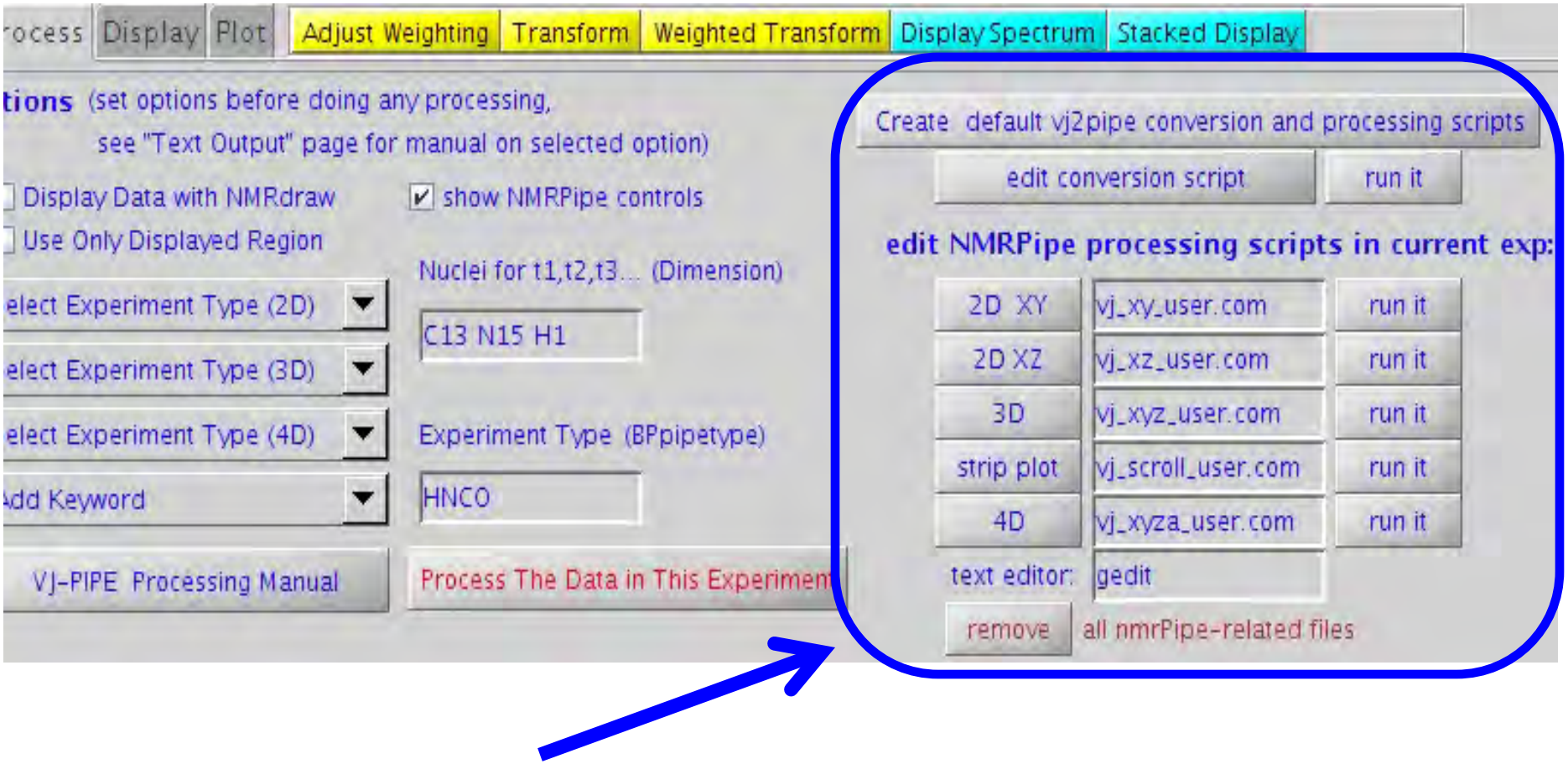
```
#!/bin/csh

xyz2pipe -in fid/test%03d.fid -x -verb \
| nmrPipe  -fn SOL                                  \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 2 -c 0.5  \
| nmrPipe  -fn ZF                                   \
| nmrPipe  -fn FT                                   \
| nmrPipe  -fn PS -p0 43.0 -p1 0.0 -di              \
| nmrPipe  -fn EXT -x1 10.5ppm -xn 5.7ppm -sw       \
| nmrPipe  -fn TP                                   \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 1 -c 1.0  \
| nmrPipe  -fn ZF                                   \
| nmrPipe  -fn FT                                   \
| nmrPipe  -fn PS -p0 -90.0 -p1 180.0 -di           \
| nmrPipe  -fn TP                                   \
| nmrPipe  -fn POLY -auto                           \
| pipe2xyz -out ft/test%03d.ft2 -x -to 0

xyz2pipe -in ft/test%03d.ft2 -z -verb              \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 1 -c 0.5  \
| nmrPipe  -fn ZF                                   \
| nmrPipe  -fn FT                                   \
| nmrPipe  -fn PS -p0 0.0 -p1 0.0 -di               \
| pipe2xyz -out ft/test%03d.ft3 -z
```

Traditionally, NMRPipe scripts are manually edited to set parameter values

```
#!/bin/csh

xyz2pipe -in _inName_ -x -verb \
| nmrPipe  -fn SOL                                    \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 2 -c _xC1_ \
| nmrPipe  -fn ZF                                    \
| nmrPipe  -fn FT                                    \
| nmrPipe  -fn PS -p0 _xP0_ -p1 _xP1_ -di            \
| nmrPipe  -fn EXT -x1 _xEXTX1_ -xn _xEXTXN_ -sw     \
| nmrPipe  -fn TP                                    \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 1 -c _yC1_ \
| nmrPipe  -fn ZF                                    \
| nmrPipe  -fn FT                                    \
| nmrPipe  -fn PS -p0 _yP0_ -p1 _yP1_ -di            \
| nmrPipe  -fn TP                                    \
| nmrPipe  -fn POLY -auto                            \
| pipe2xyz -out _auxName_ -x -to 0

xyz2pipe -in _auxName_ -z -verb                      \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 1 -c _zC1_ \
| nmrPipe  -fn ZF                                    \
| nmrPipe  -fn FT                                    \
| nmrPipe  -fn PS -p0 _zP0_ -p1 _zP1_ -di            \
| pipe2xyz -out _outName_ -z
```

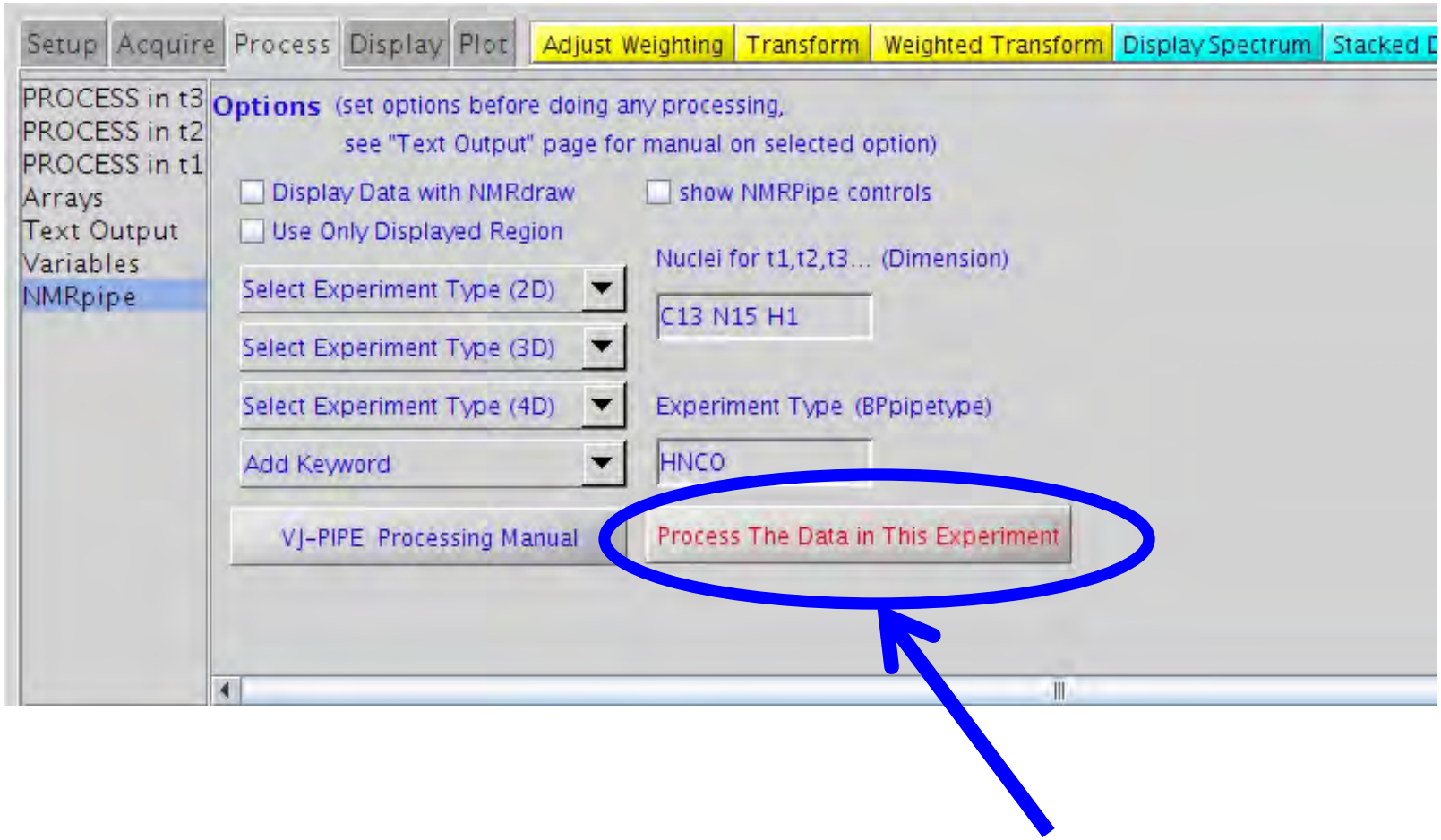
Now, values determined in VnmrJ can be automatically inserted into any NMRPipe script

NMRPipe in VnmrJ



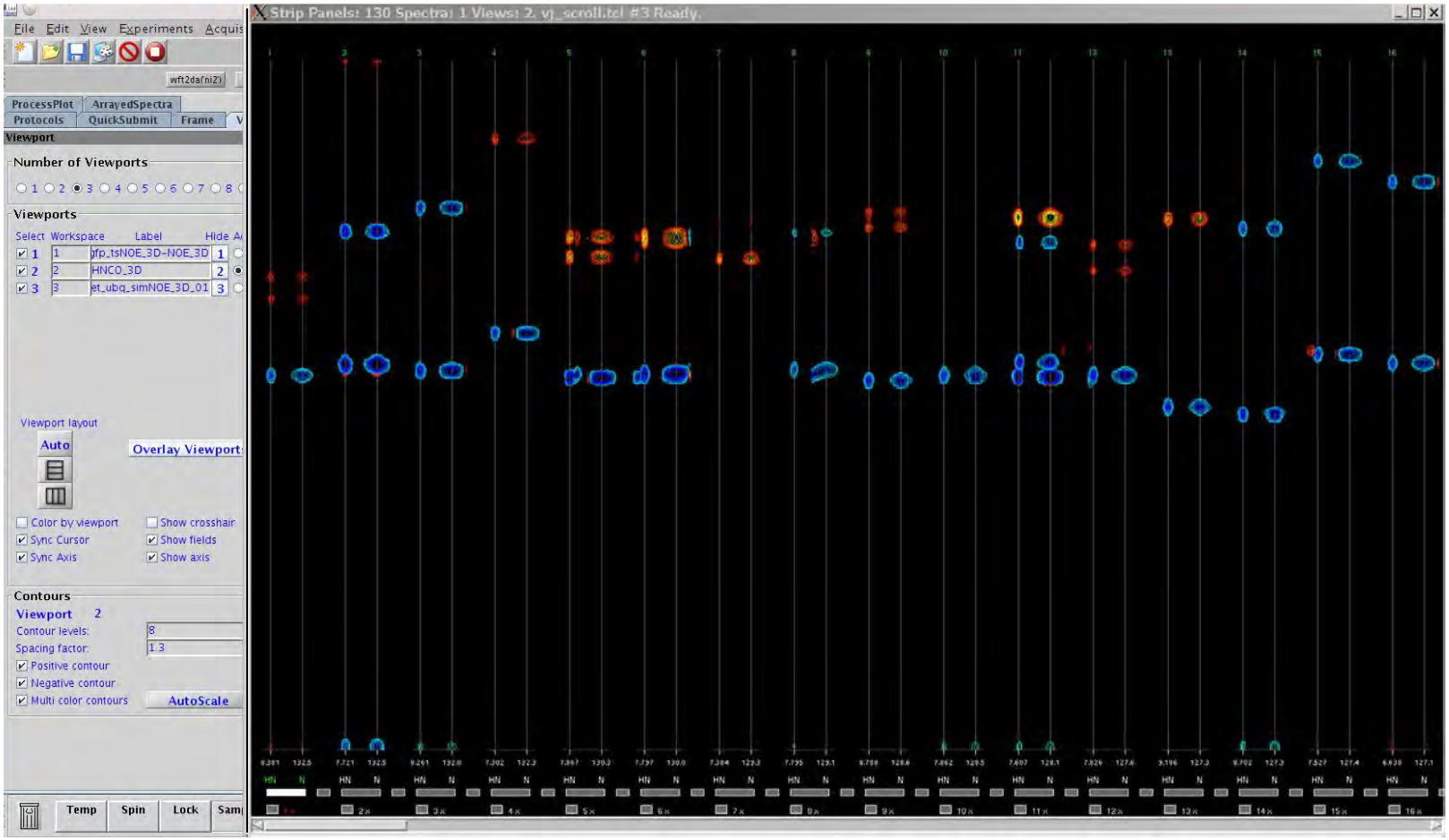
Automatically Create and Run NMRPipe Scripts

NMRPipe in VnmrJ



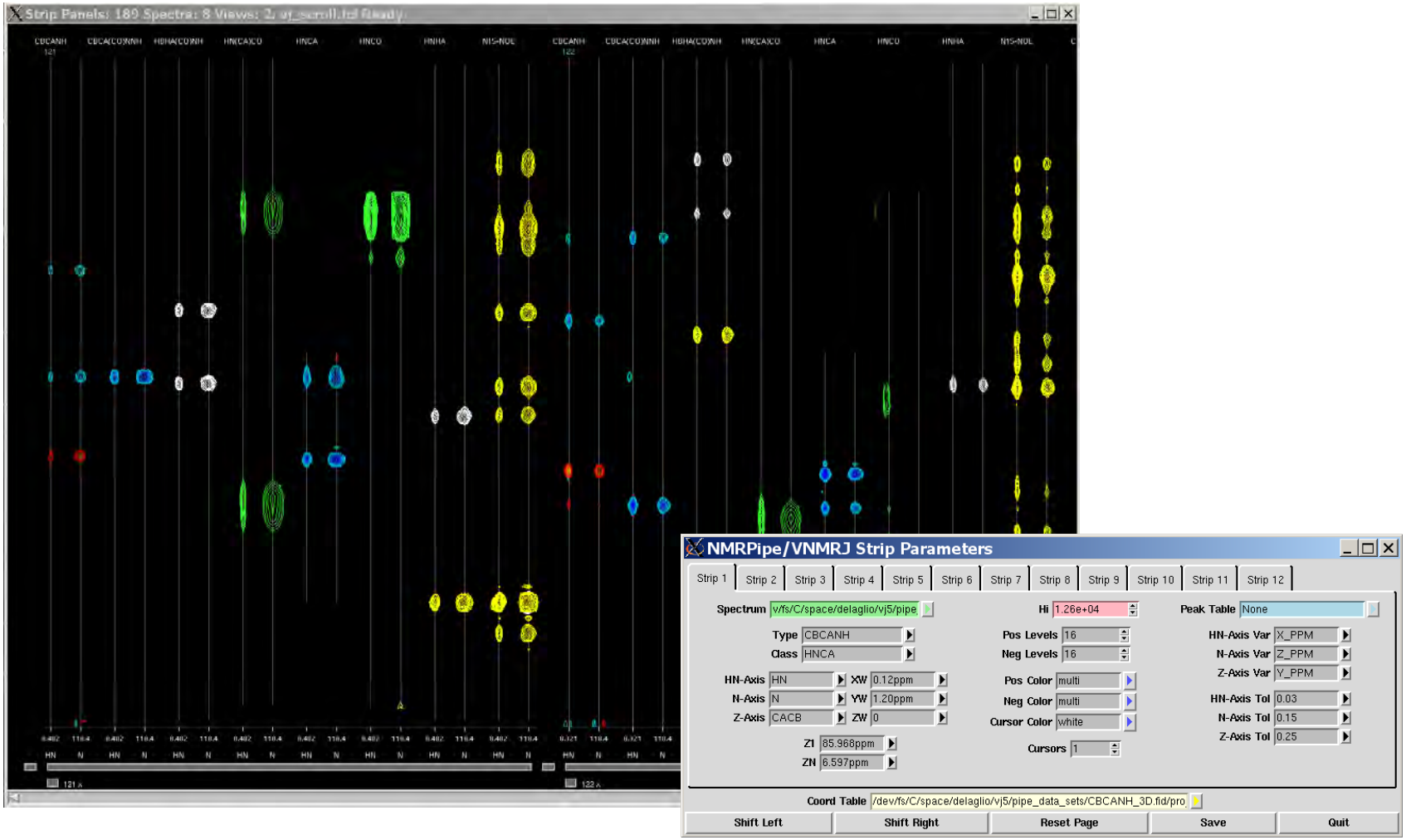
The VnmrJ "Do it All" Button

NMRPipe in VnmrJ



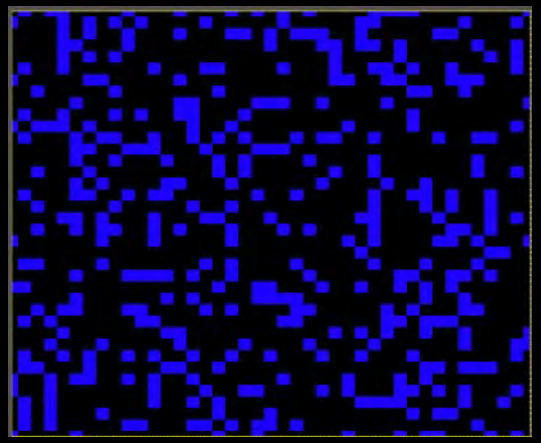
Strip display overview of processed result is created automatically

NMRPipe in VnmrJ



Strip display schemes for multiple spectra can be created automatically

CBCA(CO)NH
NUS Sampling Schedule
25% Density

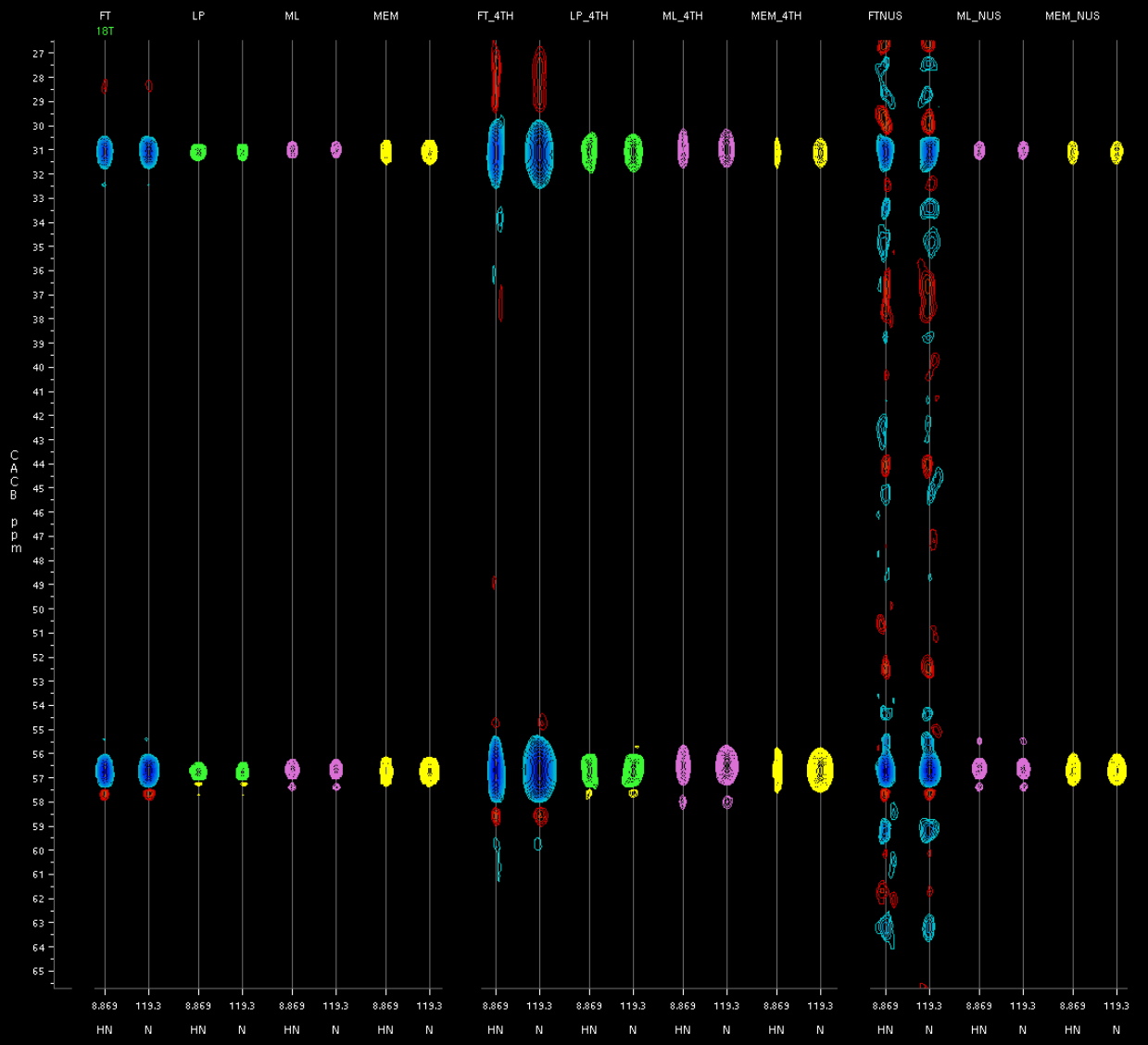


^{15}N (Uniform)

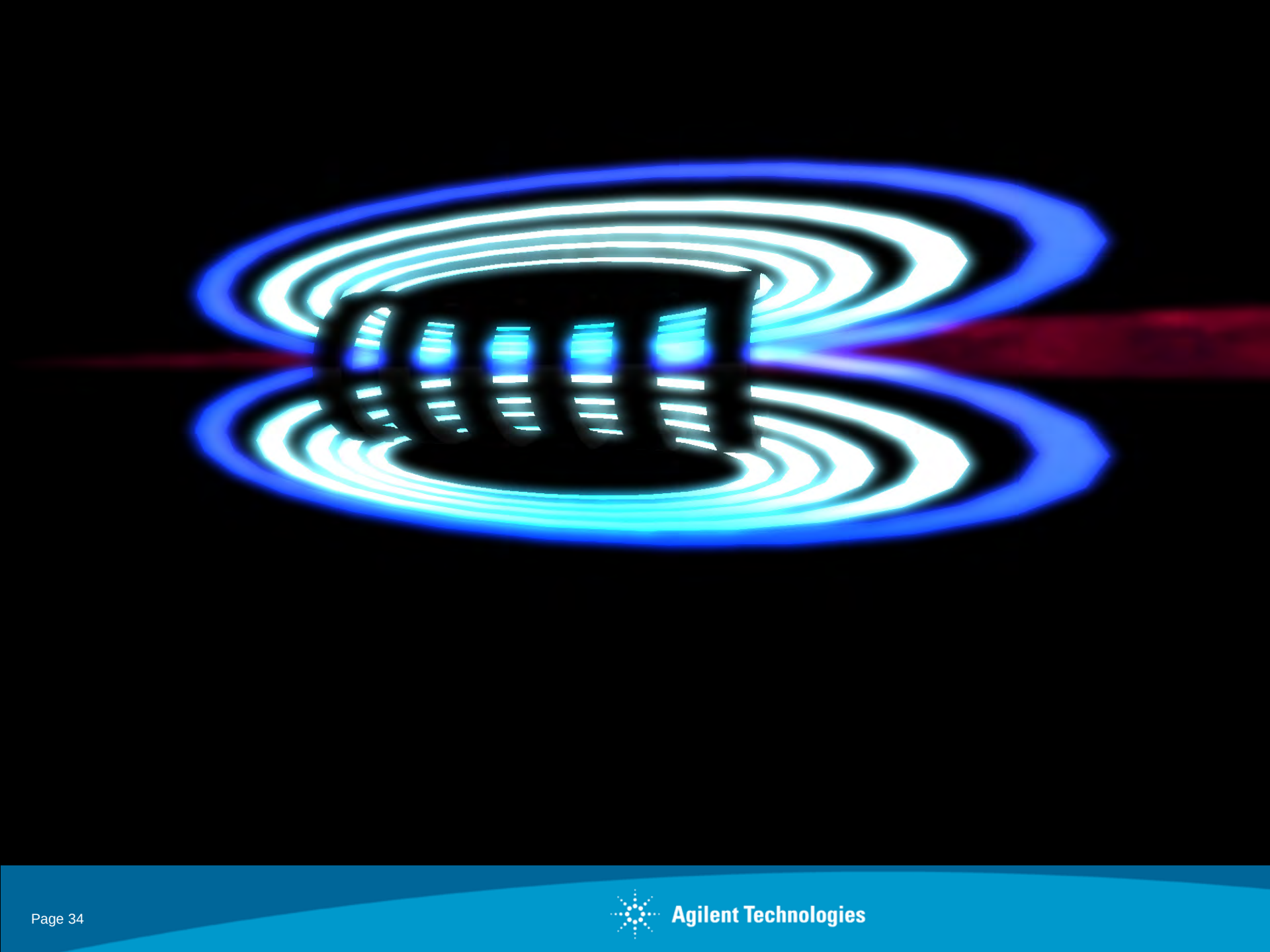
^{13}C (10% Decay)

Original Data | ¼ Resolution | NUS

FT LP ML MEM FT LP ML MEM FT ML MEM



Agilent Technologies

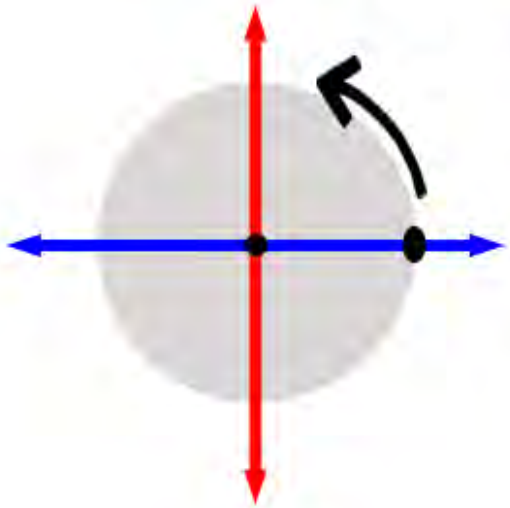


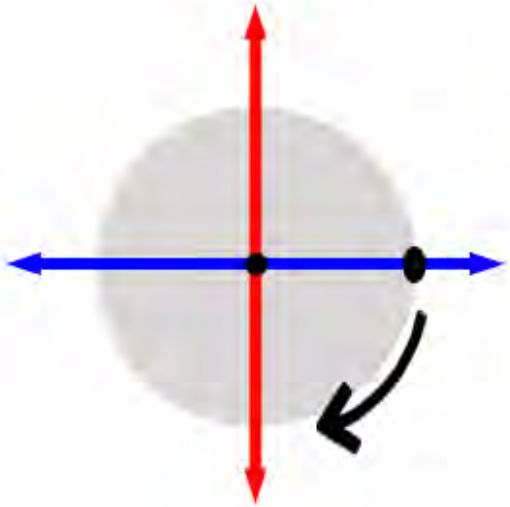
Signal Processing and NMRPipe

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi f t}\, dt,$$

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \qquad k = 0, \ldots, N-1$$

x	i^x
1	i
2	−1
3	−i
4	1
5	i
6	−1
7	−i
8	1

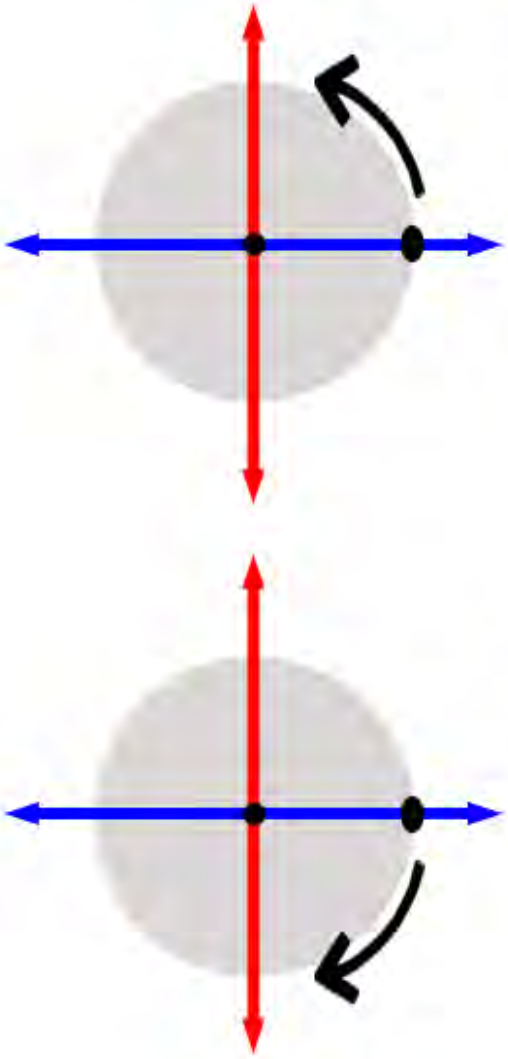


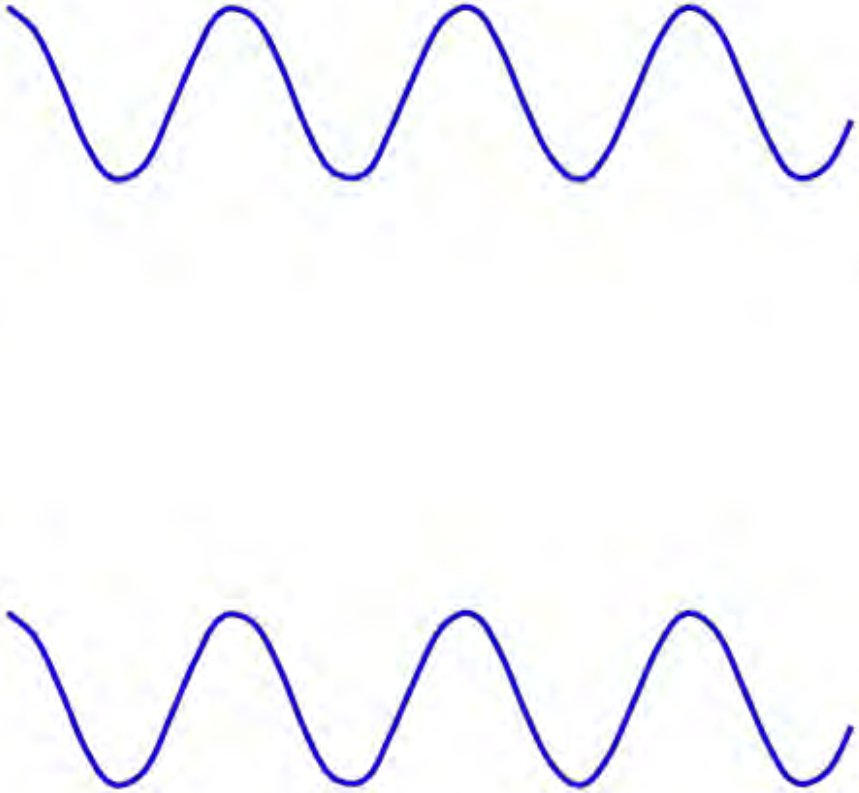


$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi f t}\, dt,$$

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \qquad k = 0, \ldots, N-1$$

x	i^x
1	i
2	−1
3	−i
4	1
5	i
6	−1
7	−i
8	1

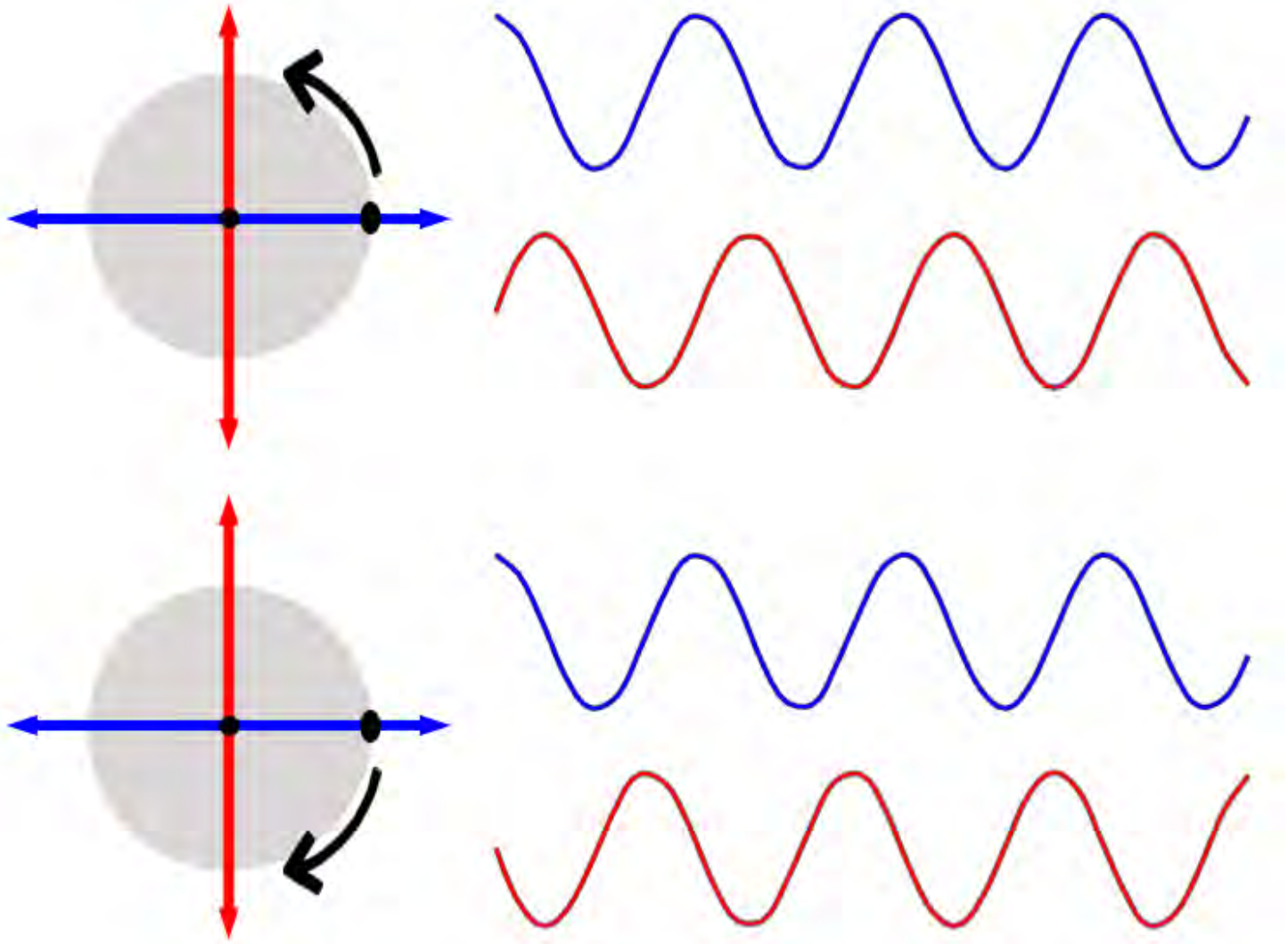




exp(*-i* 2πft) = cos(2πft) - *i* sin(2πft) … exp(*i* 2πft) = cos(2πft) + *i* sin(2πft)

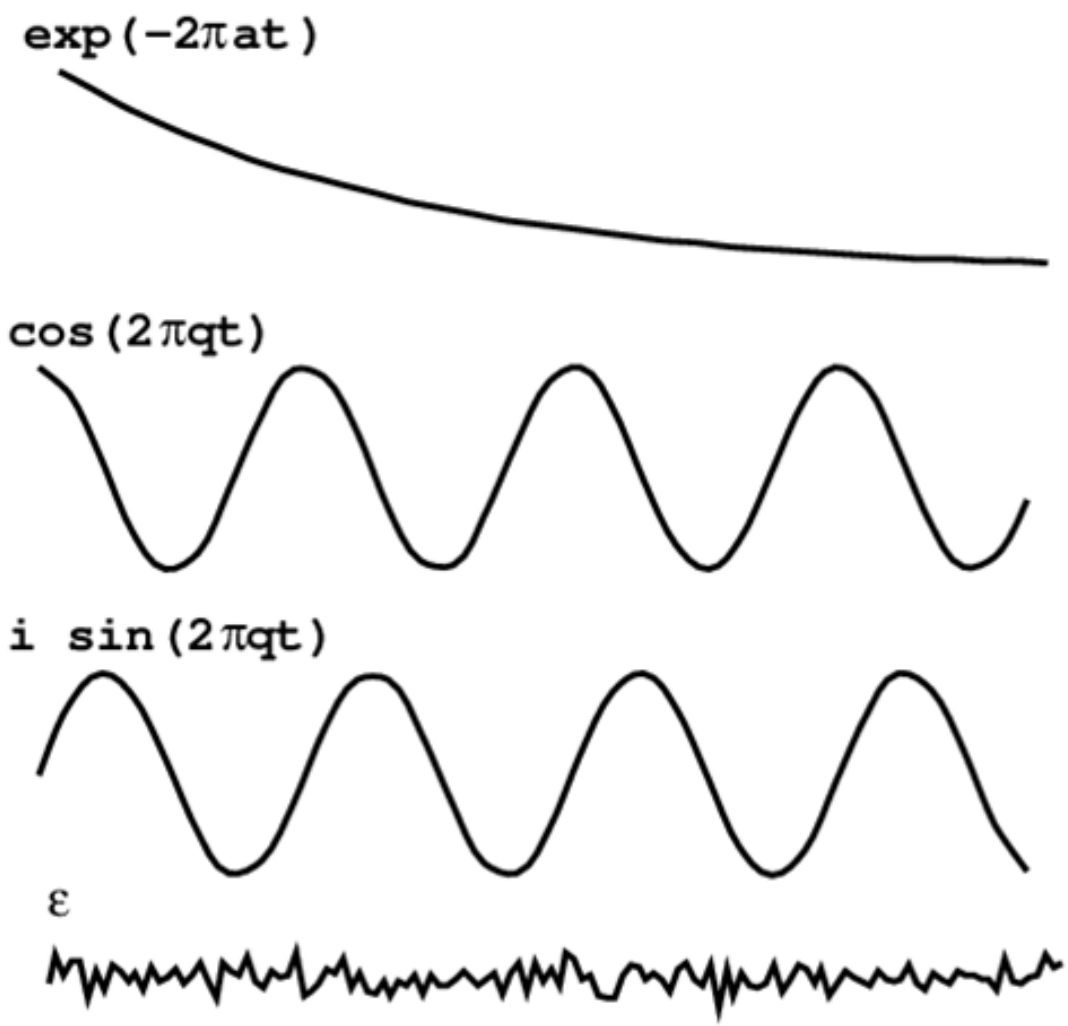
$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi ft}\, dt,$$

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \qquad k = 0, \ldots, N-1$$



A One-Dimensional
Absorbtion-Mode Time-Domain Signal

$$x(t) = c\ w(t)\ \exp(2\pi i q t) + \varepsilon$$

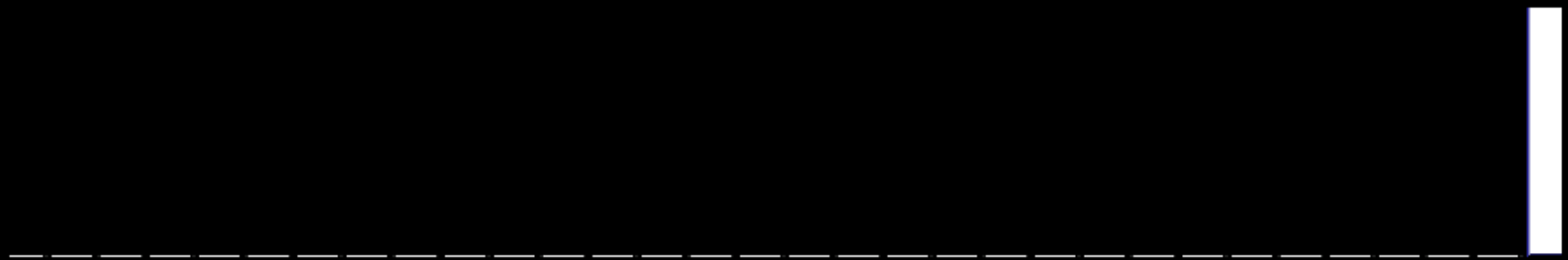


$$\mathbf{X}(f) = \sum x(t)\,[\cos(2\pi ft / N) - i\sin(2\pi ft / N)]$$

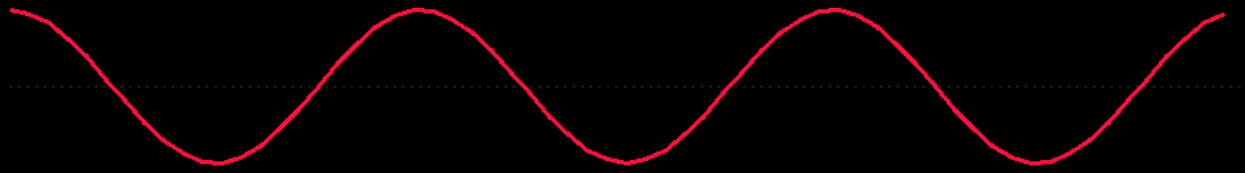
Fourier Term - Time Domain - Real Part

Fourier Term - Time Domain - Imaginary Part

Fourier Term - Frequency Domain - Real Part



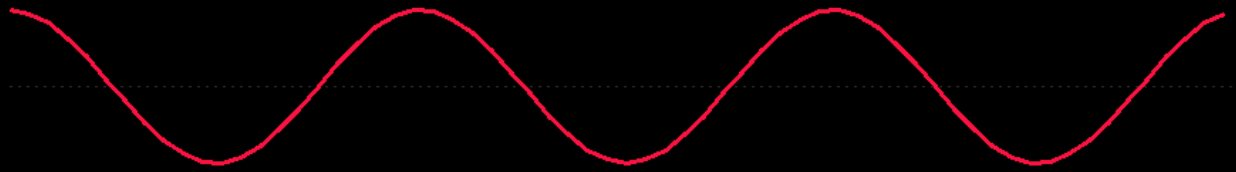
Time Domain Data to Fourier Transform



Time Domain (Real Part)

Fourier Term

Multiplied with Time Domain Data



Forms the Product

Sum Over Product to Form a Frequency Point:

Frequency Domain (Real Part)

```
/* Fourier transform of complex data tR,tI to produce fR,fI. */

void ft( float *tR, float *tI, float *fR, float *fI, int size )
{
    float vR, vI, twoPI;
    int   mid, k, n;

    twoPI = 4.0*acos( 0.0 );
    mid   = size/2;

    for( k = 0; k < size; k++ )           /* For every output freq point … */
       {
        fR[k] = 0.0;
        fI[k] = 0.0;

        for( n = 0; n < size; n++ )       /* Sum over input times sinusoid. */
           {
            f  = twoPI*(k - mid)*n/size;

            vR = cos( f );
            vI = sin( f );

            fR[k] += tR[n]*vR - tI[n]*vI;
            fI[k] += tR[n]*vI + tI[n]*vR;
           }
       }
}
```

Truncated Time Domain Data to Fourier Transform

Time Domain (Real Part)

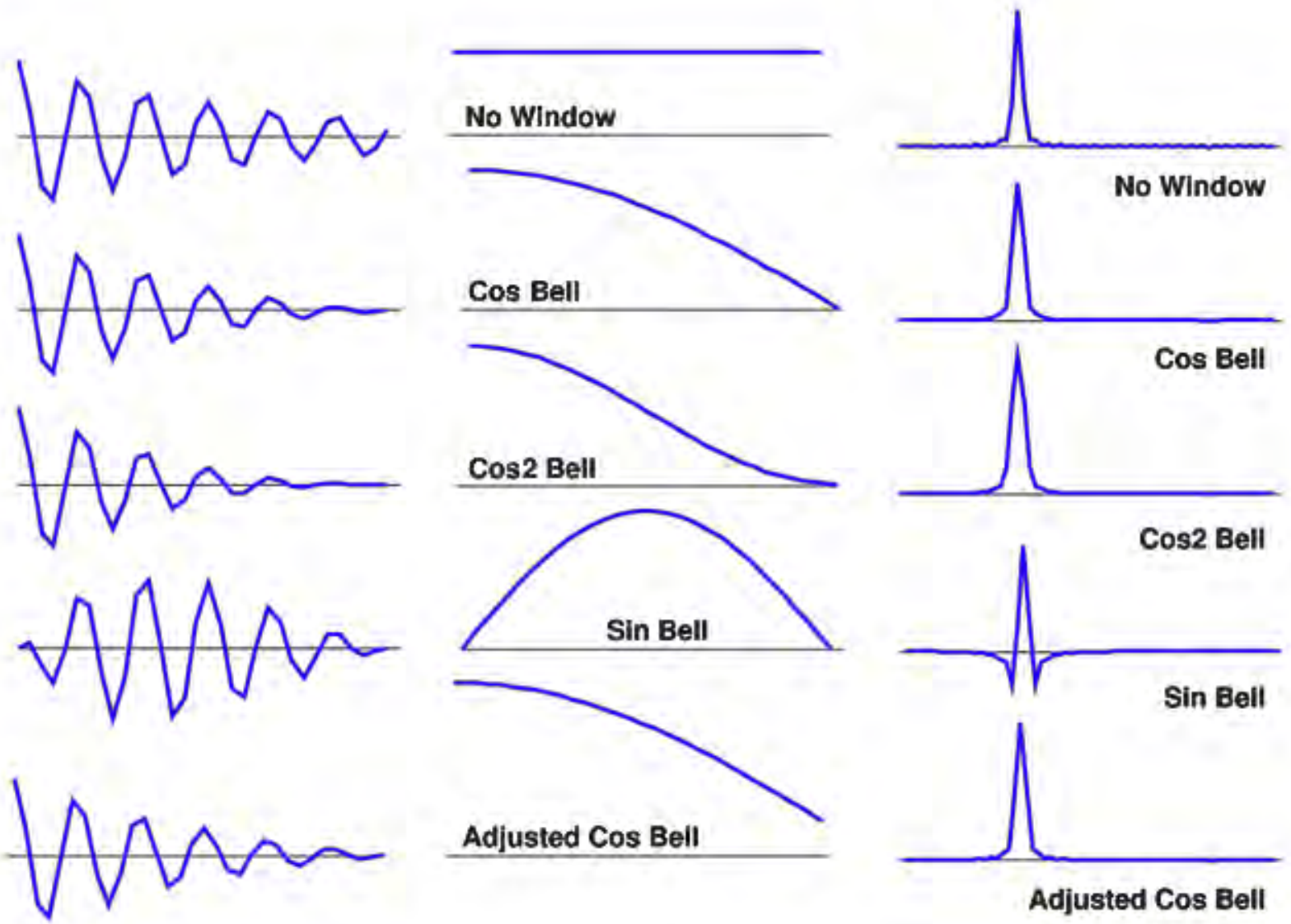
Fourier Term

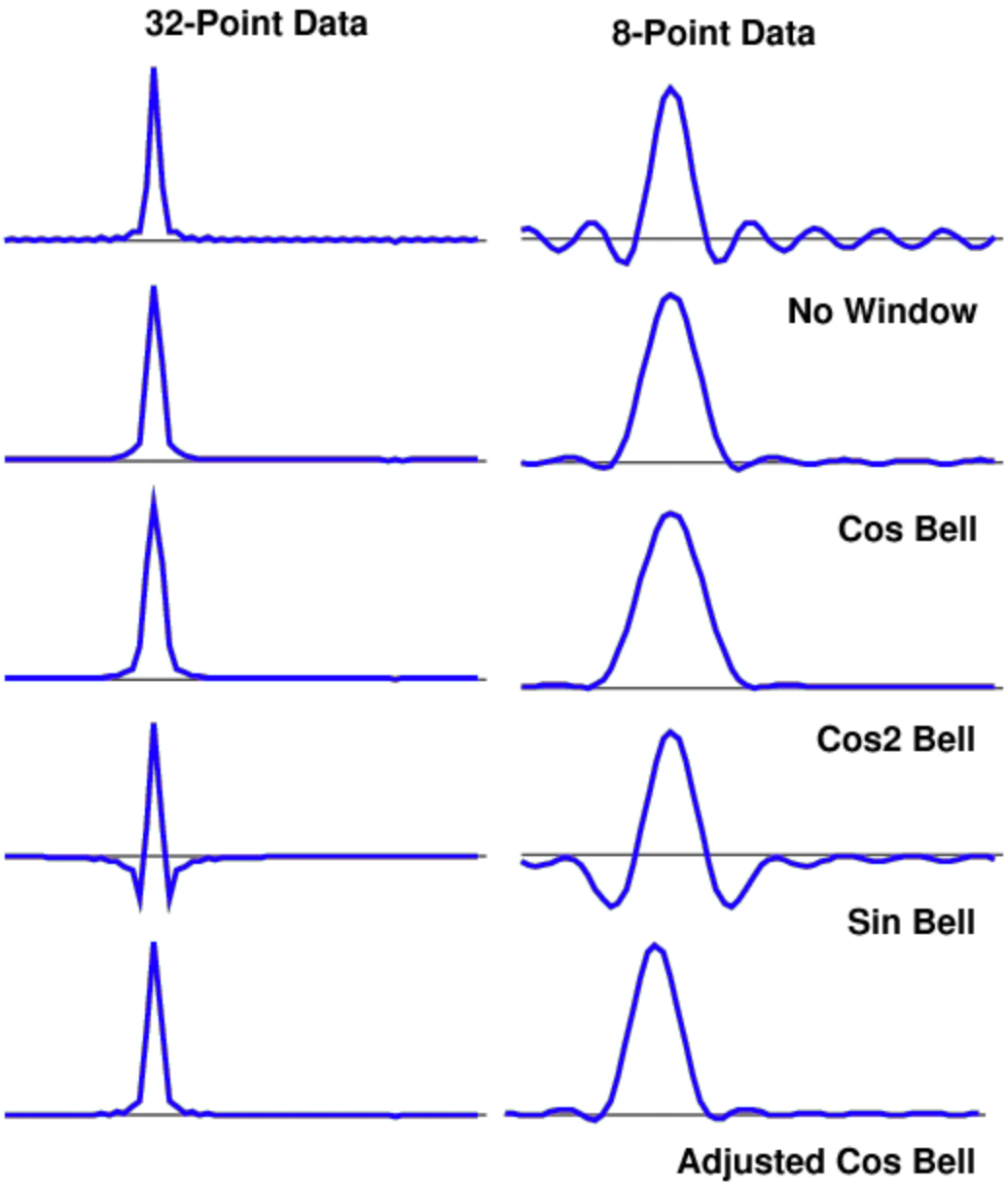
Multiplied with Truncated Time Domain Data

Forms the Product

Sum Over Product to Form a Frequency Point:

Frequency Domain (Real Part)

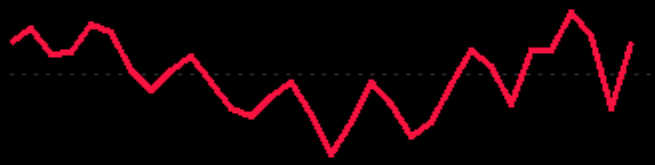

No Window
Cos Bell
Cos2 Bell
Sin Bell
Adjusted Cos Bell
No Window
Cos Bell
Cos2 Bell
Sin Bell
Adjusted Cos Bell


32-Point Data
8-Point Data
No Window
Cos Bell
Cos2 Bell
Sin Bell
Adjusted Cos Bell

Time Domain

Original Time-Domain

Inverse Exponential Applied

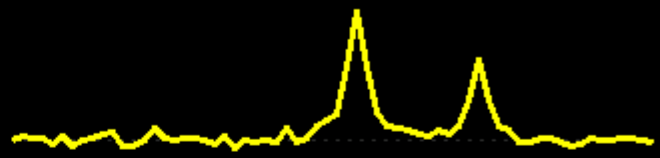


Original Noise

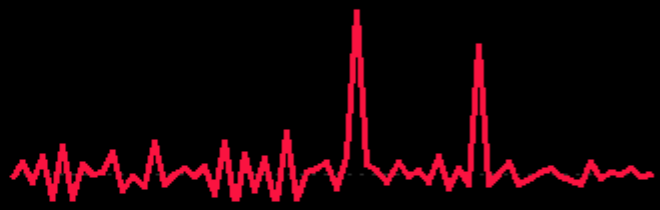
Amplified Noise

Frequency Domain

Original Freq-Domain



Inverse Exponential Applied



Original Noise

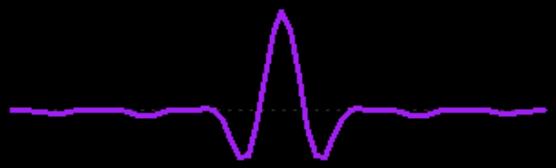
Amplified Noise

Time Domain

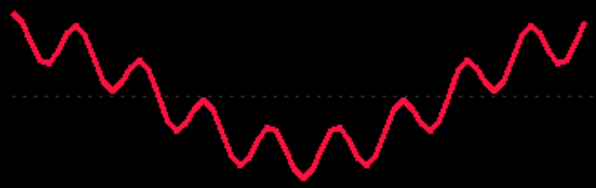
Shifted Convolution Kernel

Frequency Domain

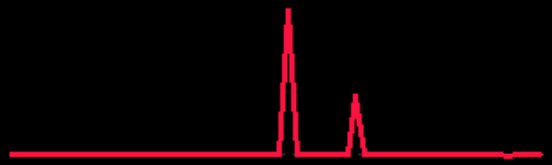
Fourier Transform of Kernel



Multiplied With



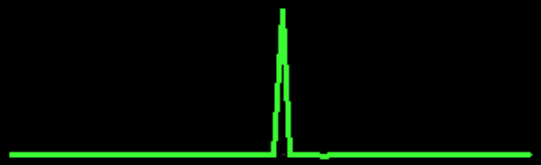
Multiplied With



Forms the Product

Sum Over Product to Form a Convolution Point:

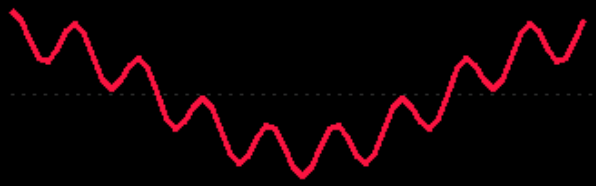
Forms the Product



Time Domain

Shifted Convolution Kernel

Multiplied With

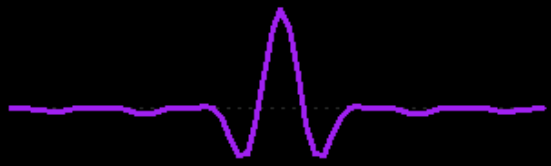


Forms the Product

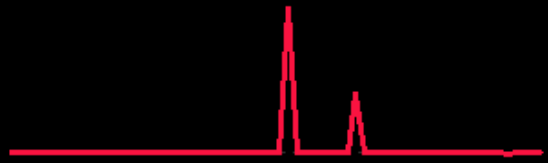
Sum Over Product to Form a Convolution Point:

Frequency Domain

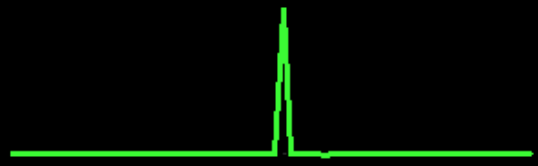
Fourier Transform of Kernel



Multiplied With



Forms the Product



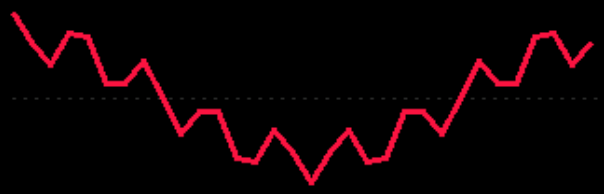
Initial Domain

Shifted Convolution Kernel

Inverse Domain

Fourier Transform of Kernel

Multiplied With

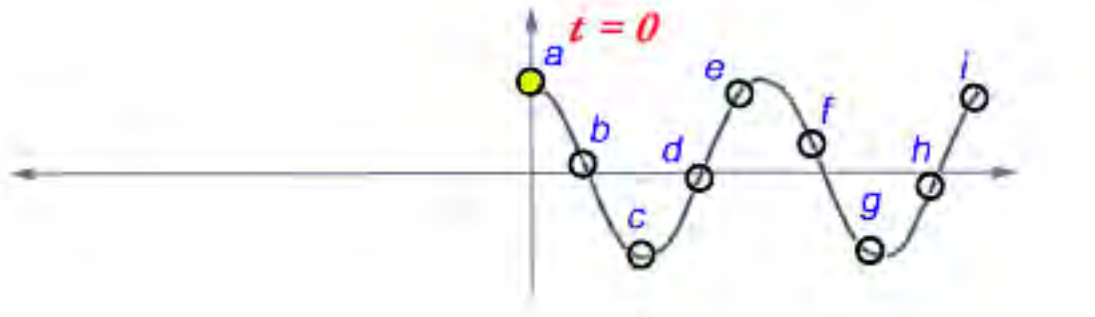


Forms the Product

Sum Over Product to Form a Convolution Point:

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi f t}\, dt, \qquad X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \qquad k = 0, \ldots, N-1$$

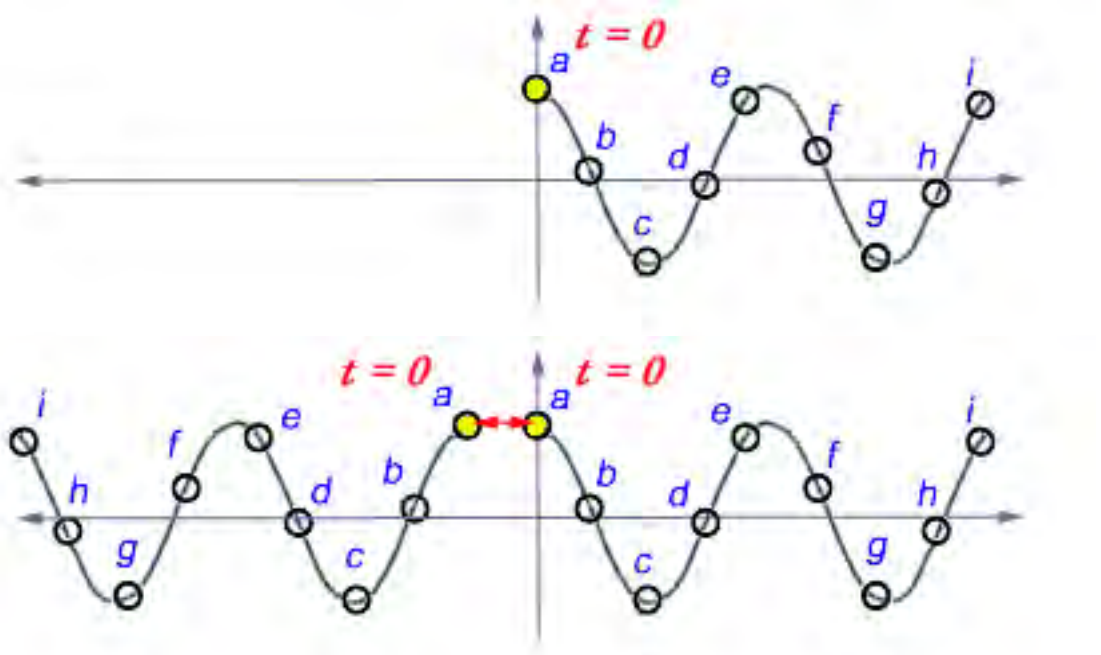
exp(*-i* 2πft) = cos(2πft) - *i* sin(2πft)



First Point Acquisition Delay

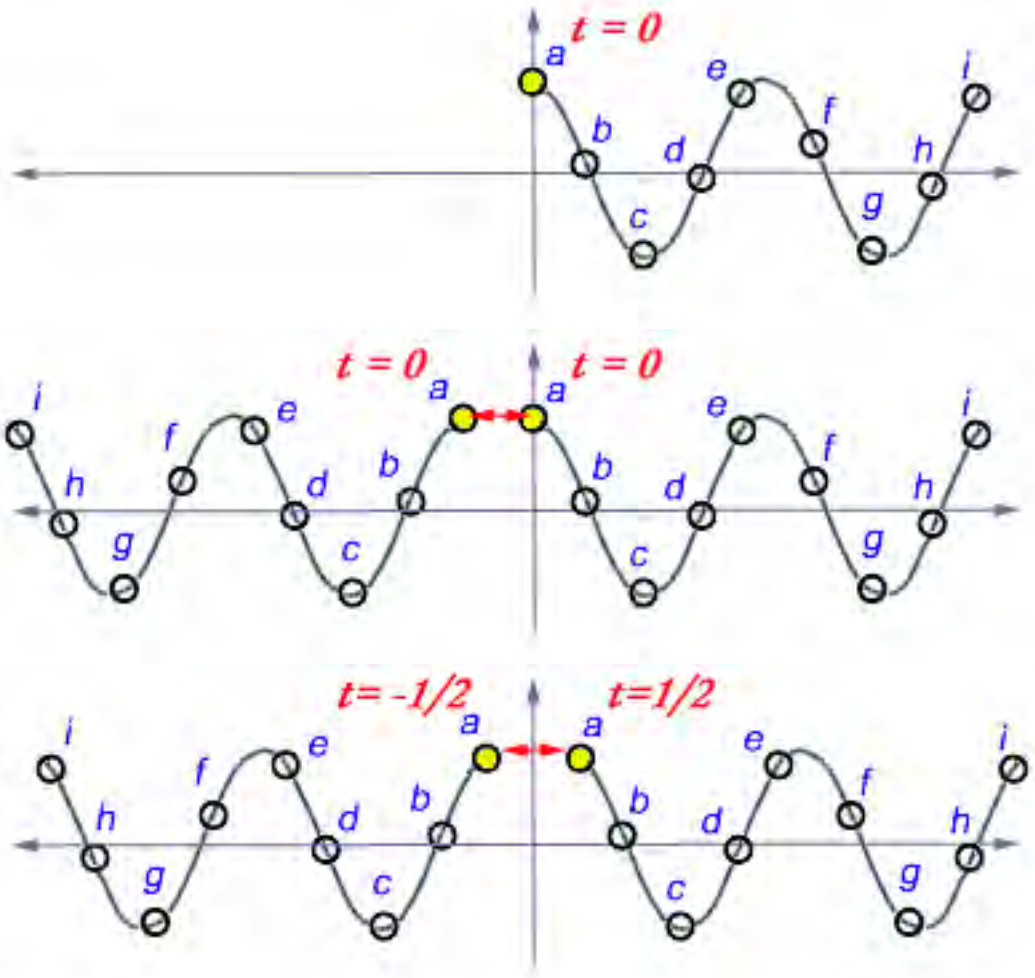
$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi f t}\, dt, \qquad X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \qquad k = 0, \ldots, N-1$$

exp(*-i* 2πft) = cos(2πft) - *i* sin(2πft)



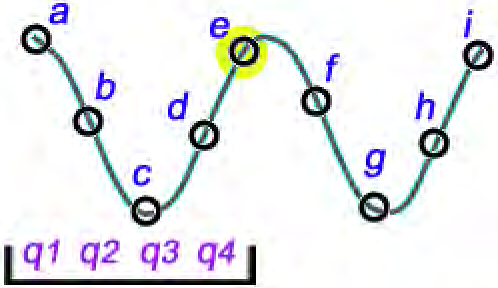
$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi f t}\, dt, \qquad X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \qquad k = 0, \ldots, N-1$$

exp(*-i* 2πft) = cos(2πft) - *i* sin(2πft)

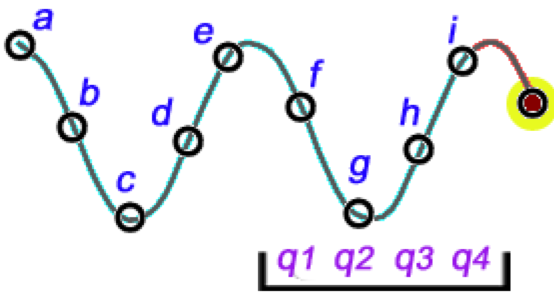


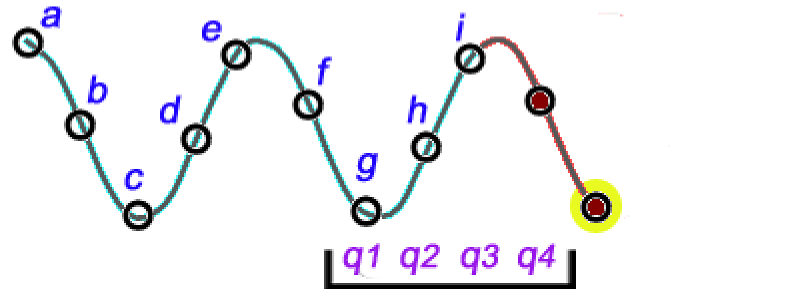
Special Methods in Signal Processing

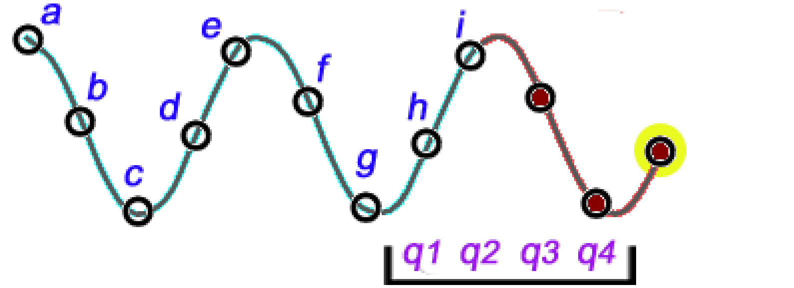
Linear Prediction

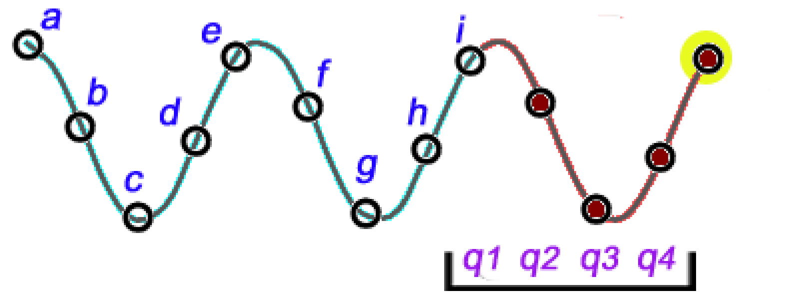


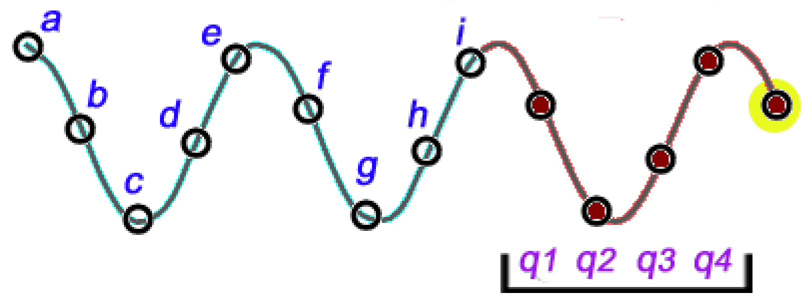
$$
\begin{aligned}
q1*a + q2*b + q3*c + q4*d &= e \\
q1*b + q2*c + q3*d + q4*e &= f \\
q1*c + q2*d + q3*e + q4*f &= g \\
q1*d + q2*e + q3*f + q4*g &= h \\
q1*e + q2*f + q3*g + q4*h &= i
\end{aligned}
$$

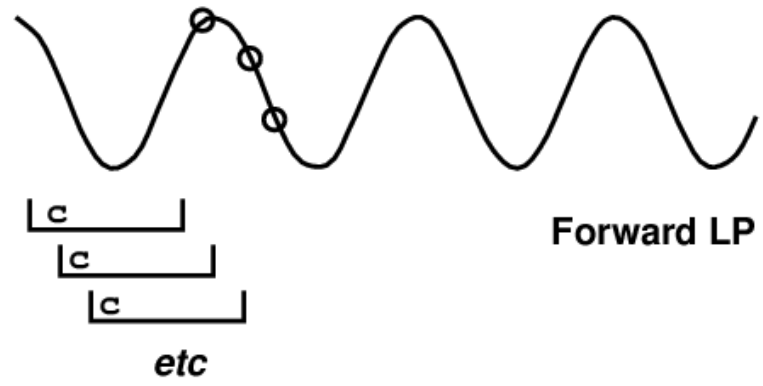


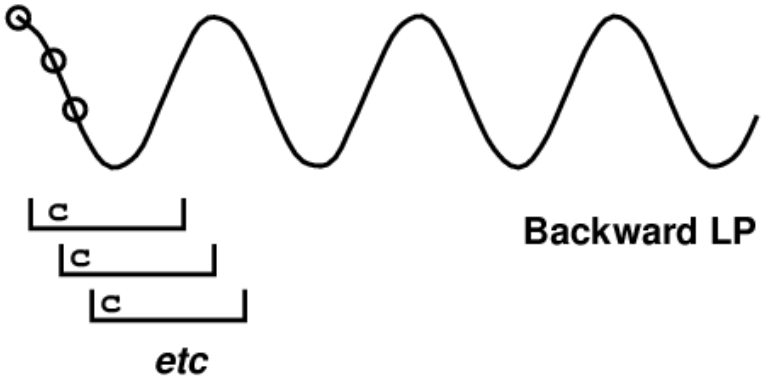






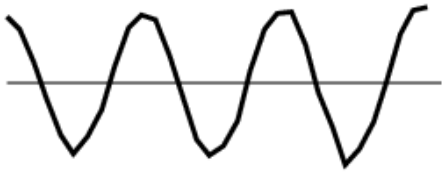


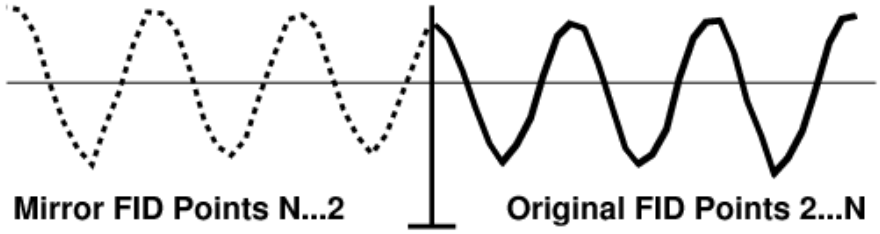


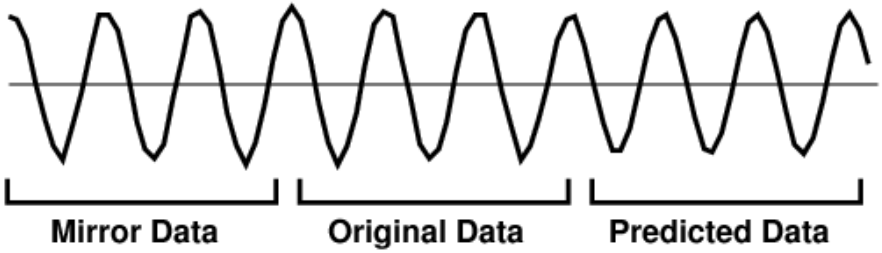


Forward and Backward Coefficients Can be Averaged for Stability

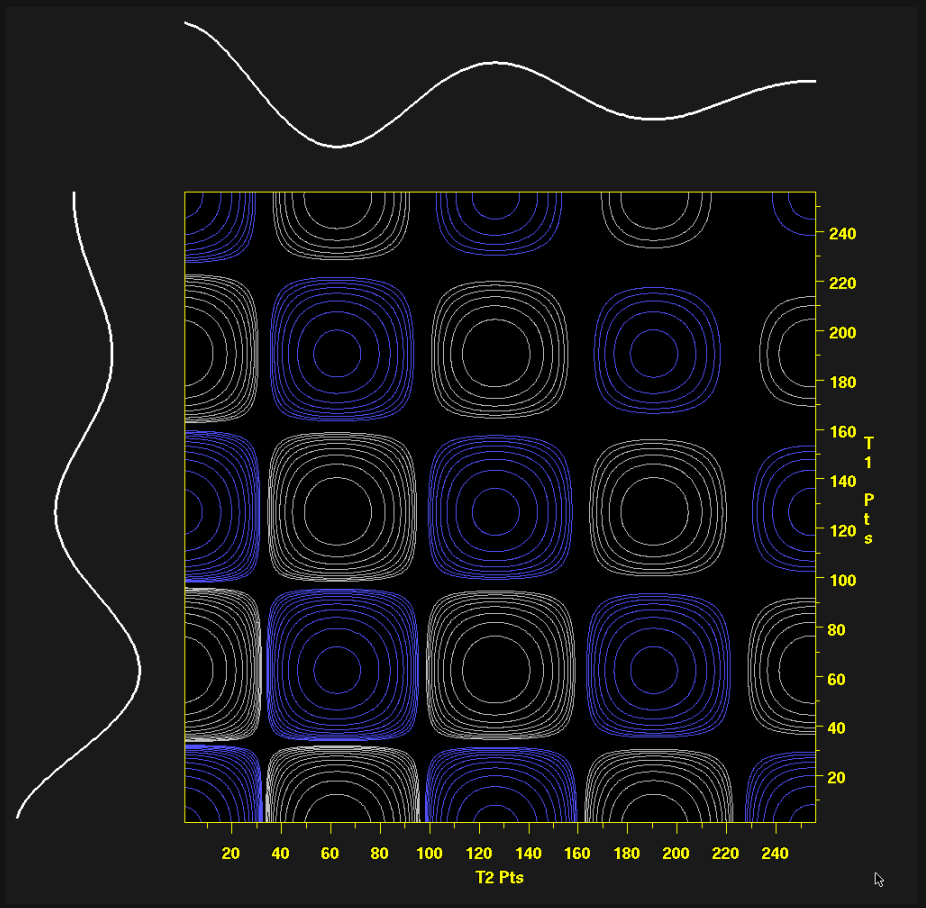
Mirror Image LP - Zero Delay Time

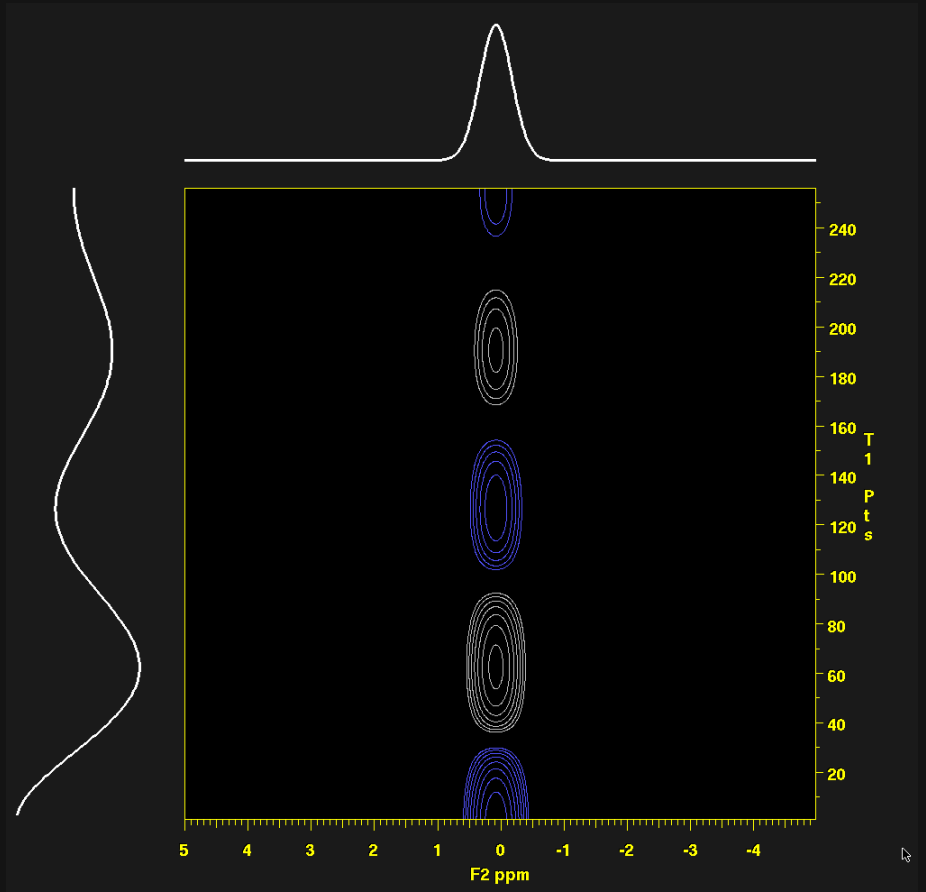






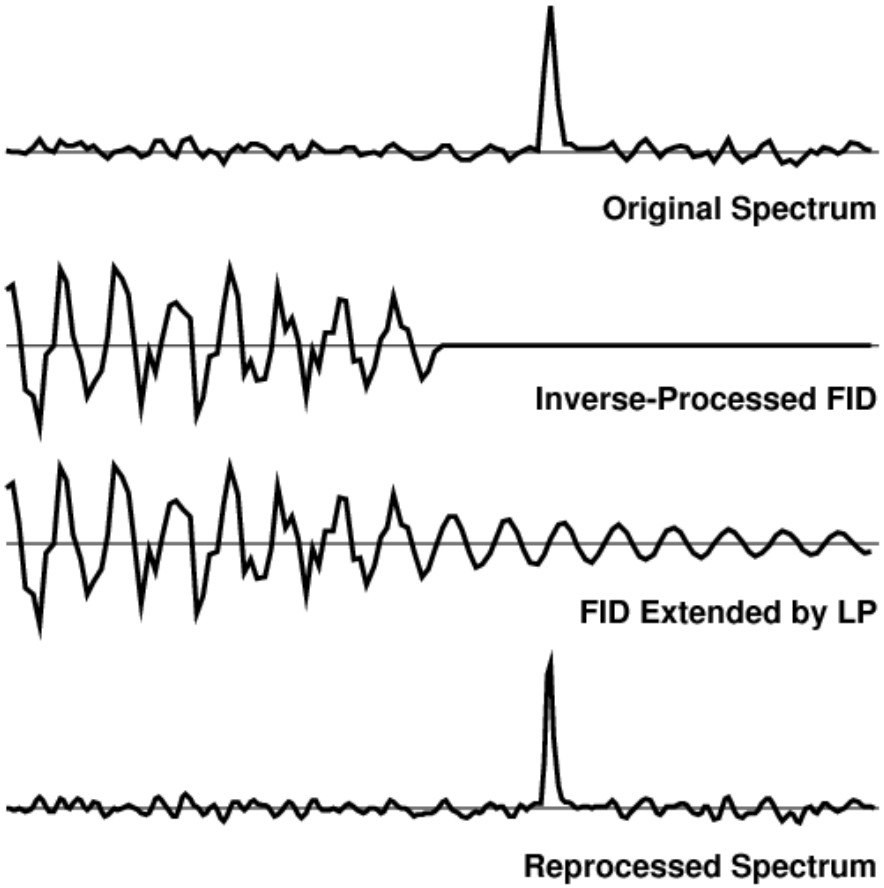
Extending the Length of the Input Data Allows More LP Coefficients, Hence More Signals

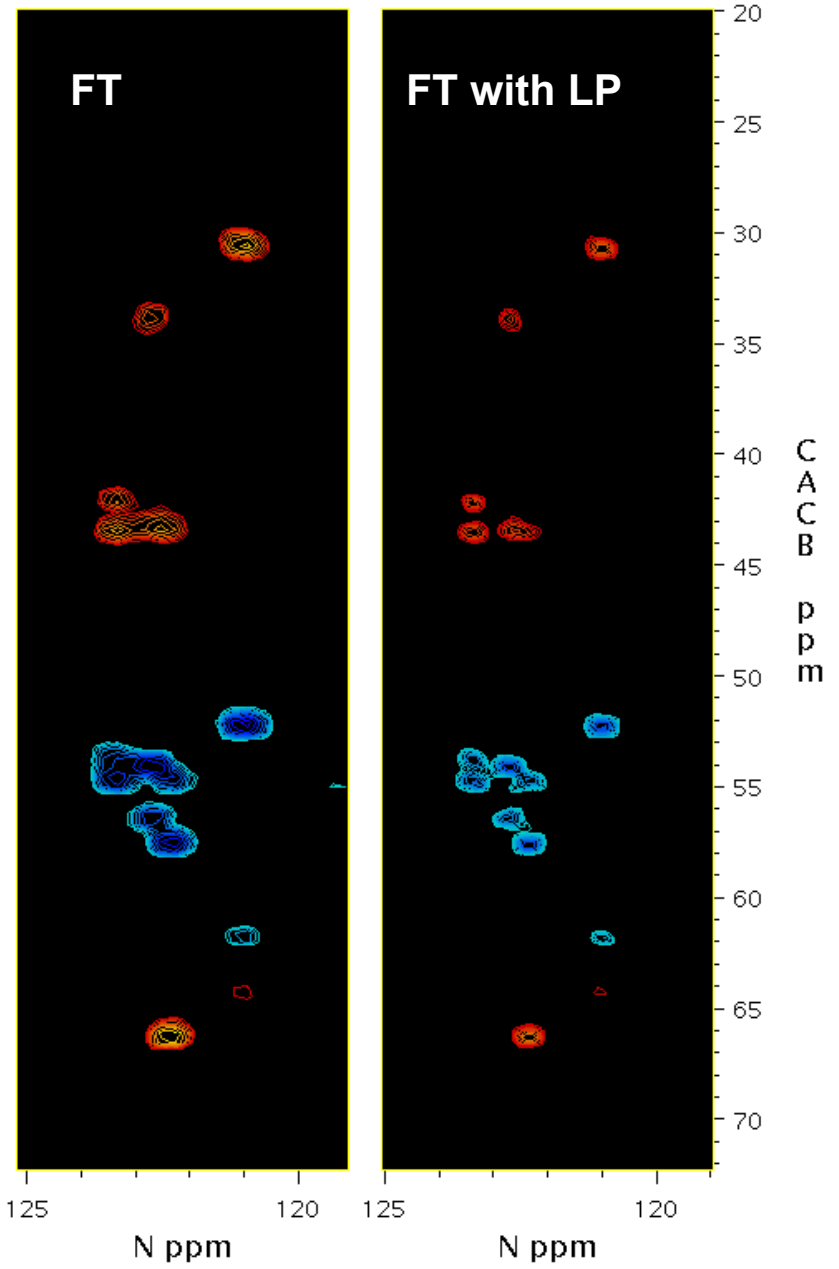

240
220
200
180
160
140
120
100
80
60
40
20
T1 Pts
20 40 60 80 100 120 140 160 180 200 220 240
T2 Pts


240
220
200
180
160
140
120
100
80
60
40
20
T1 Pts
5 4 3 2 1 0 -1 -2 -3 -4
F2 ppm

```
xyz2pipe -in fid/test%03d.fid -x  -verb                \  # Process the Directly-detected X-Axis.
| nmrPipe  -fn SOL                                     \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 2 -c 0.5     \
| nmrPipe  -fn ZF -auto                                \
| nmrPipe  -fn FT                                      \
| nmrPipe  -fn PS -p0 43 -p1 0.0 -di                   \
| nmrPipe  -fn EXT -x1 11.5ppm -xn 5.5ppm -sw          \
| pipe2xyz -out lp/x%03d.ft1 -x

xyz2pipe -in lp/x%03d.ft1 -z -verb                    \ # Process the Indirectly-detected Z-Axis.
| nmrPipe  -fn SP -off 0.5 -end 0.95 -pow 1 -c 0.5     \
| nmrPipe  -fn ZF -auto                                \
| nmrPipe  -fn FT                                      \
| nmrPipe  -fn PS -p0 0.0 -p1 0.0 -di                  \
| pipe2xyz -out lp/xz%03d.ft2 -z

xyz2pipe -in lp/xz%03d.ft2 -y -verb                   \ # Linear Predict and Process the
| nmrPipe  -fn LP -fb -ord 12                          \ # Indirectly-Detected Z-Axis.
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 1 -c 1.0     \
| nmrPipe  -fn ZF -auto                                \
| nmrPipe  -fn FT                                      \
| nmrPipe  -fn PS -p0 -135 -p1 180 -di                 \
| pipe2xyz -out lp/xyz%03d.ft3 -y

xyz2pipe -in lp/xyz%03d.ft3 -z -verb                  \ # Inverse Process, Linear Predict,
| nmrPipe  -fn HT  -auto                               \ # and Re-Process the Z-Axis
| nmrPipe  -fn PS  -inv -hdr                           \
| nmrPipe  -fn FT  -inv                               \
| nmrPipe  -fn ZF  -inv                               \
| nmrPipe  -fn SP  -inv -hdr                          \
| nmrPipe  -fn LP  -fb                                \
| nmrPipe  -fn SP  -off 0.5 -end 0.98 -pow 1 -c 0.5   \
| nmrPipe  -fn ZF  -auto                               \
| nmrPipe  -fn FT                                      \
| nmrPipe  -fn PS  -hdr -di                            \
| pipe2xyz -out lp/test%03d.ft3 -z
```

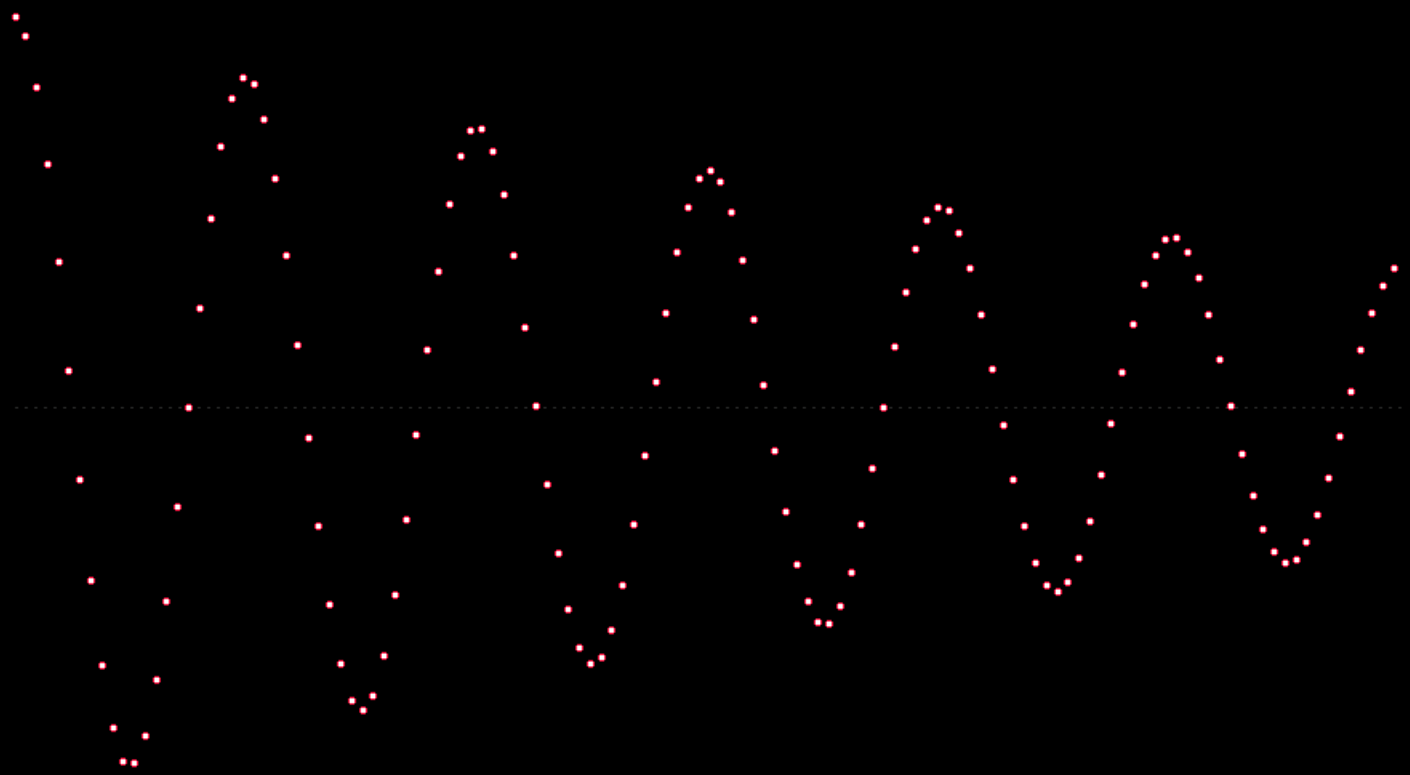
Linear Prediction



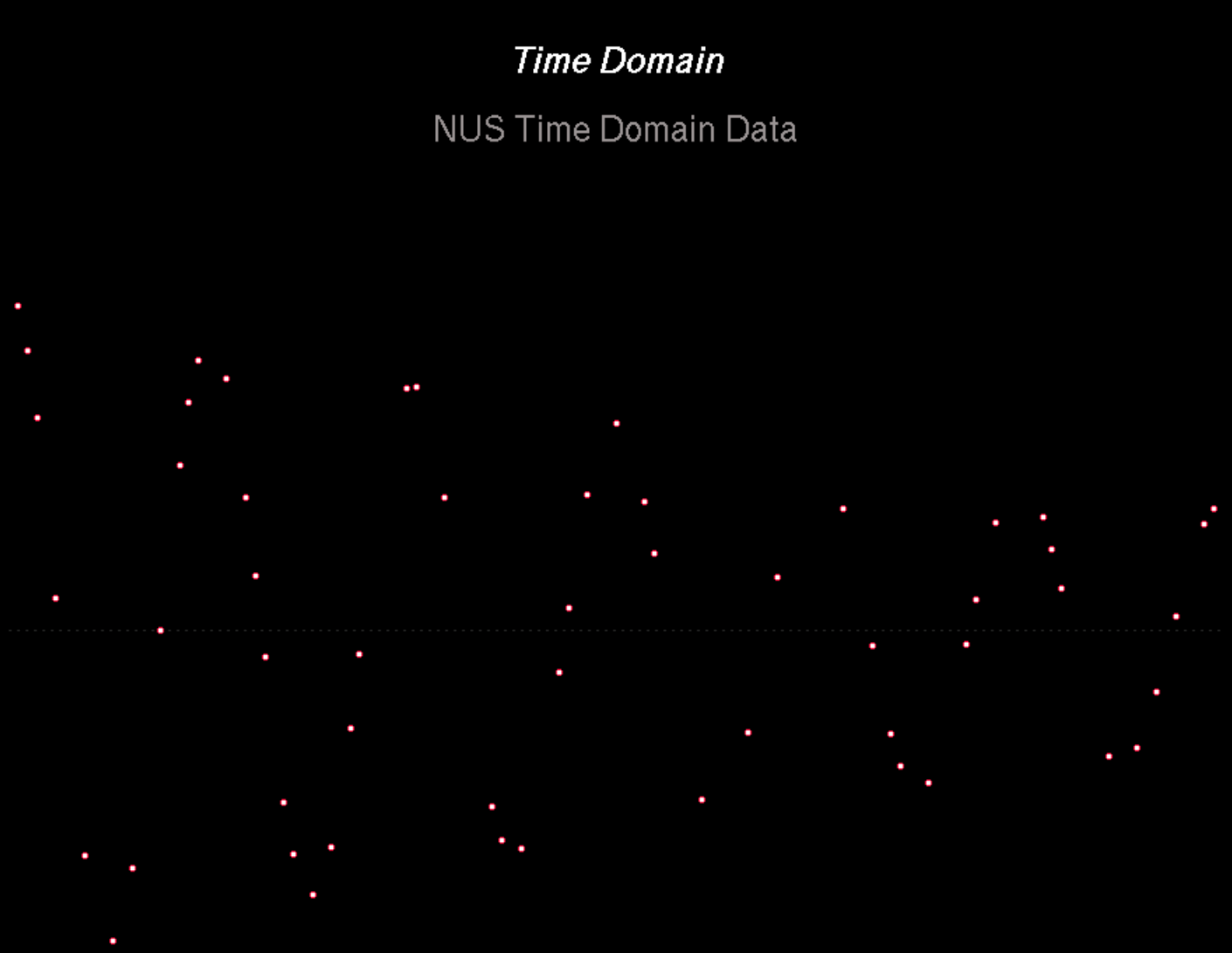


Time Domain

Full Time Domain Data



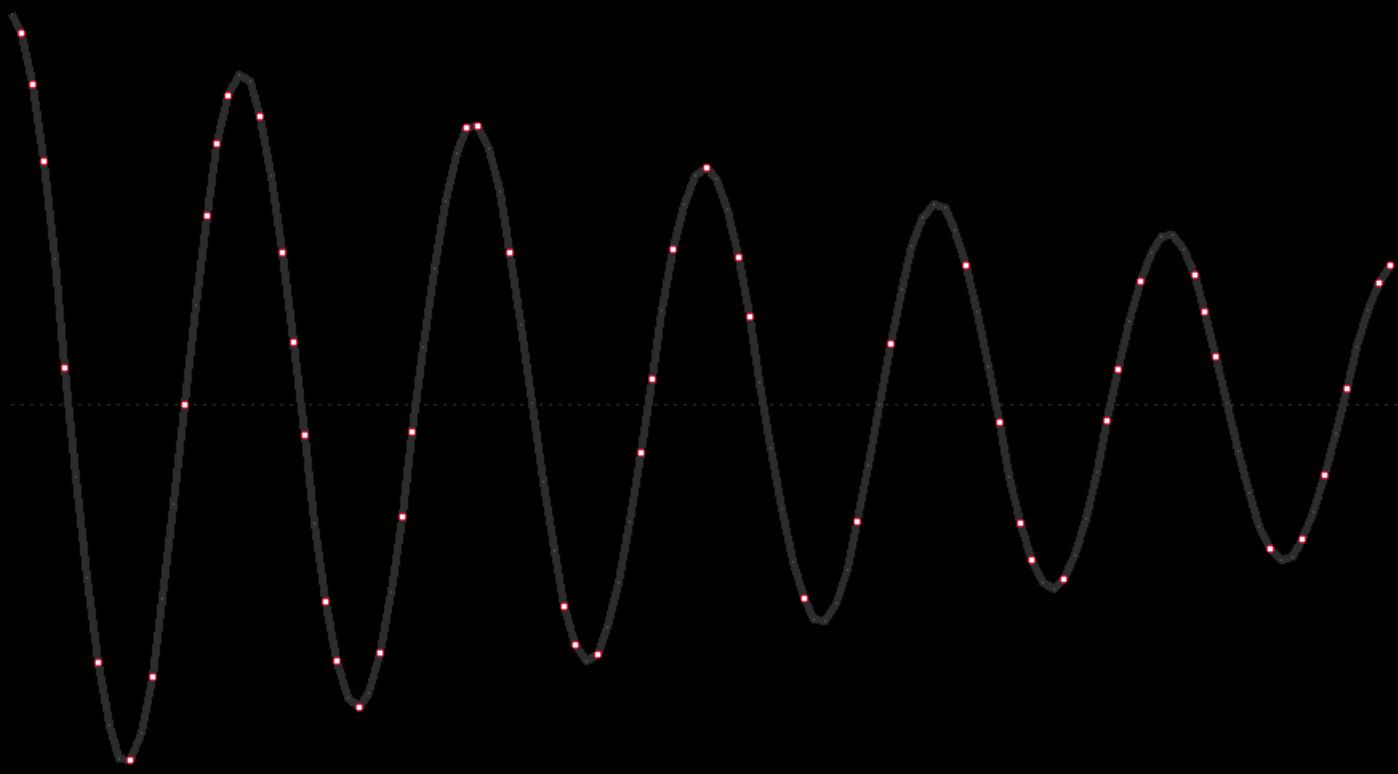
Non-Uniform Sampling on a Uniform Grid



Non-Uniform Sampling: Skip a Fraction of the Points

Time Domain

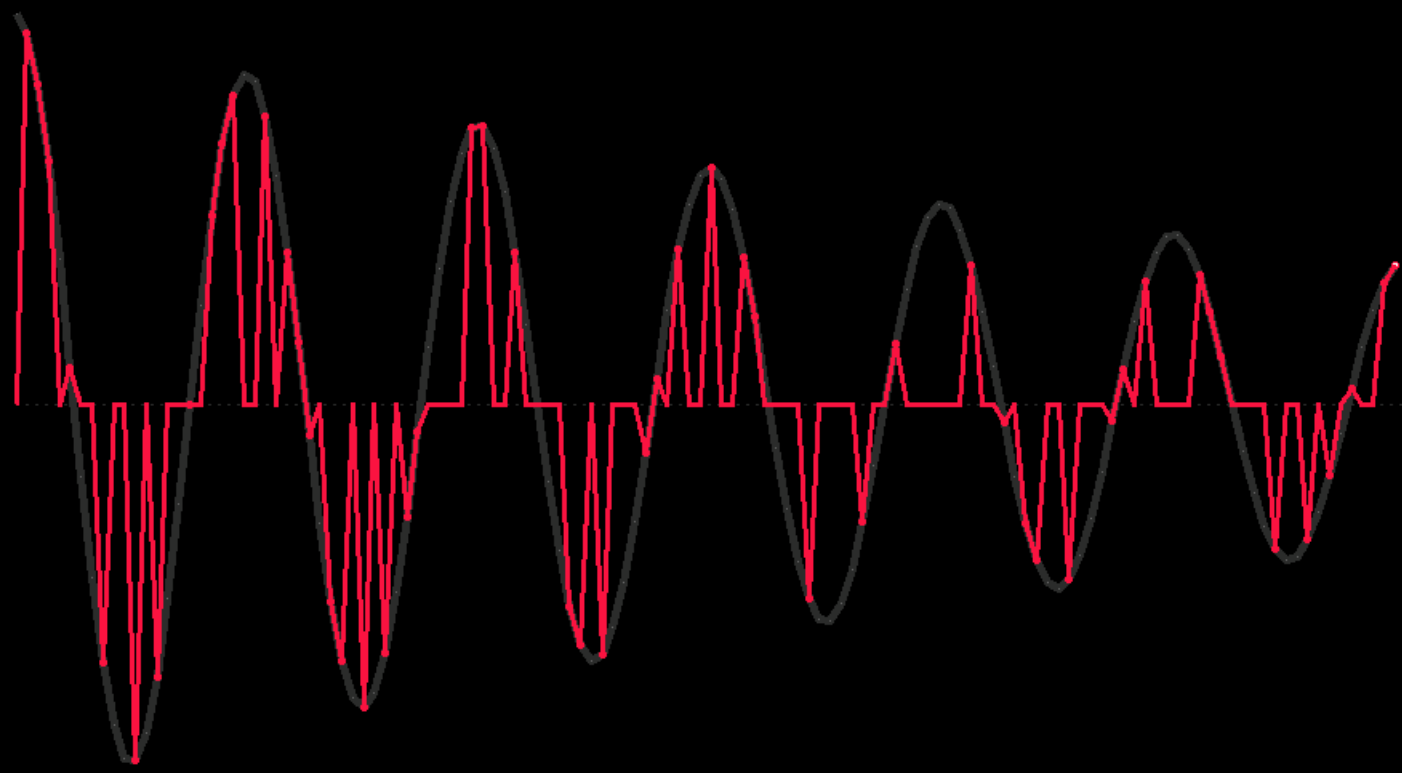
NUS Time Domain Data



Non-Uniform Sampling

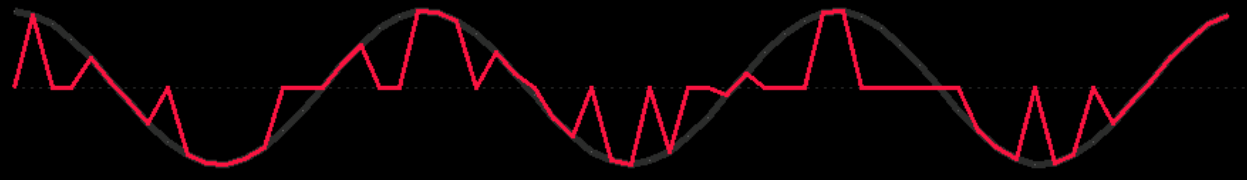
Time Domain

NUS Time Domain Data



Non-Uniform Sampling:
for Fourier Transform, Replace Missing Points with Zeros

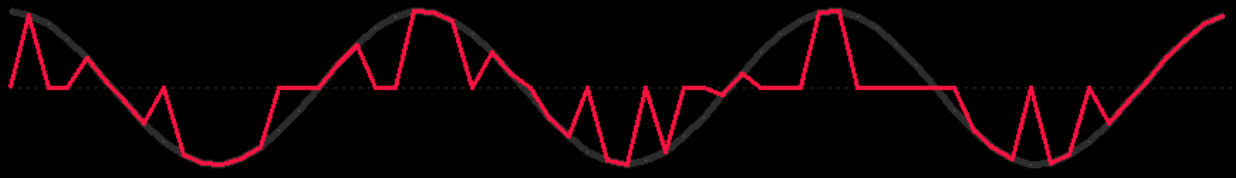
NUS Time Domain Data to Fourier Transform



Time Domain

Fourier Term

Multiplied with NUS Time Domain Data

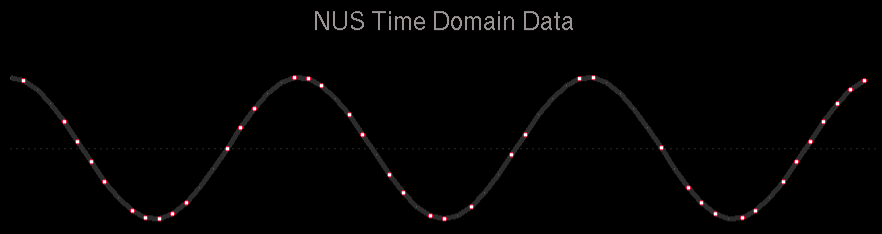


Forms the Product

Sum Over Product to Form a Frequency Point:

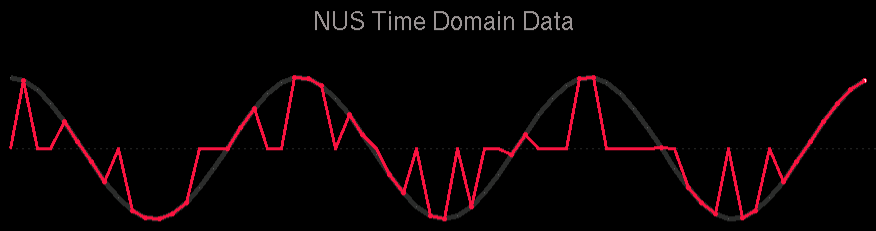
Frequency Domain

Time Domain



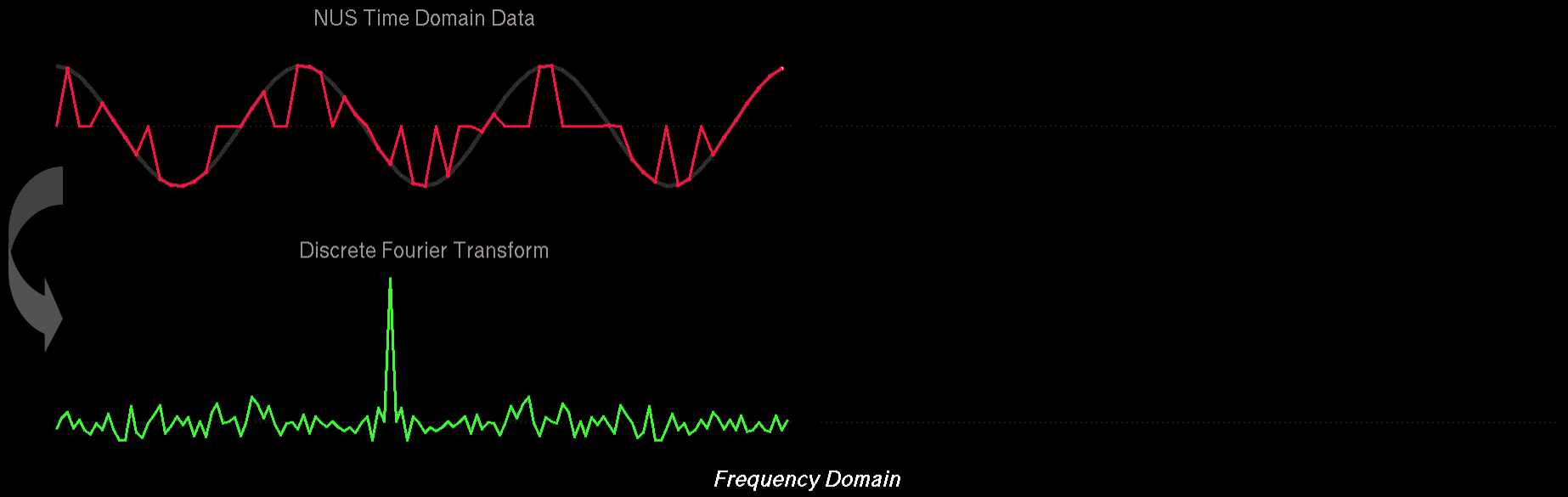
Frequency Domain

Time Domain



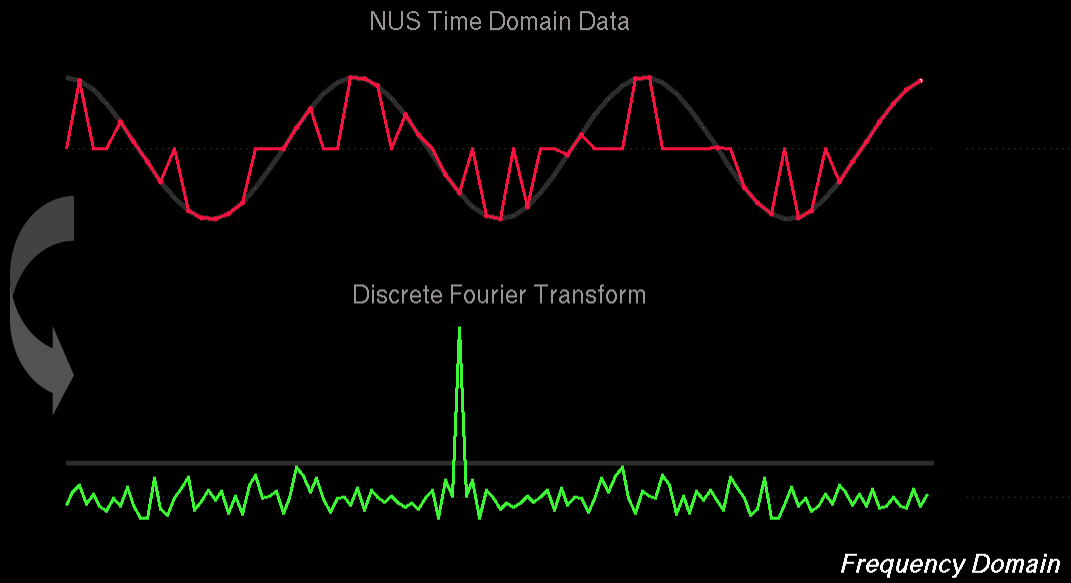
Frequency Domain

Time Domain

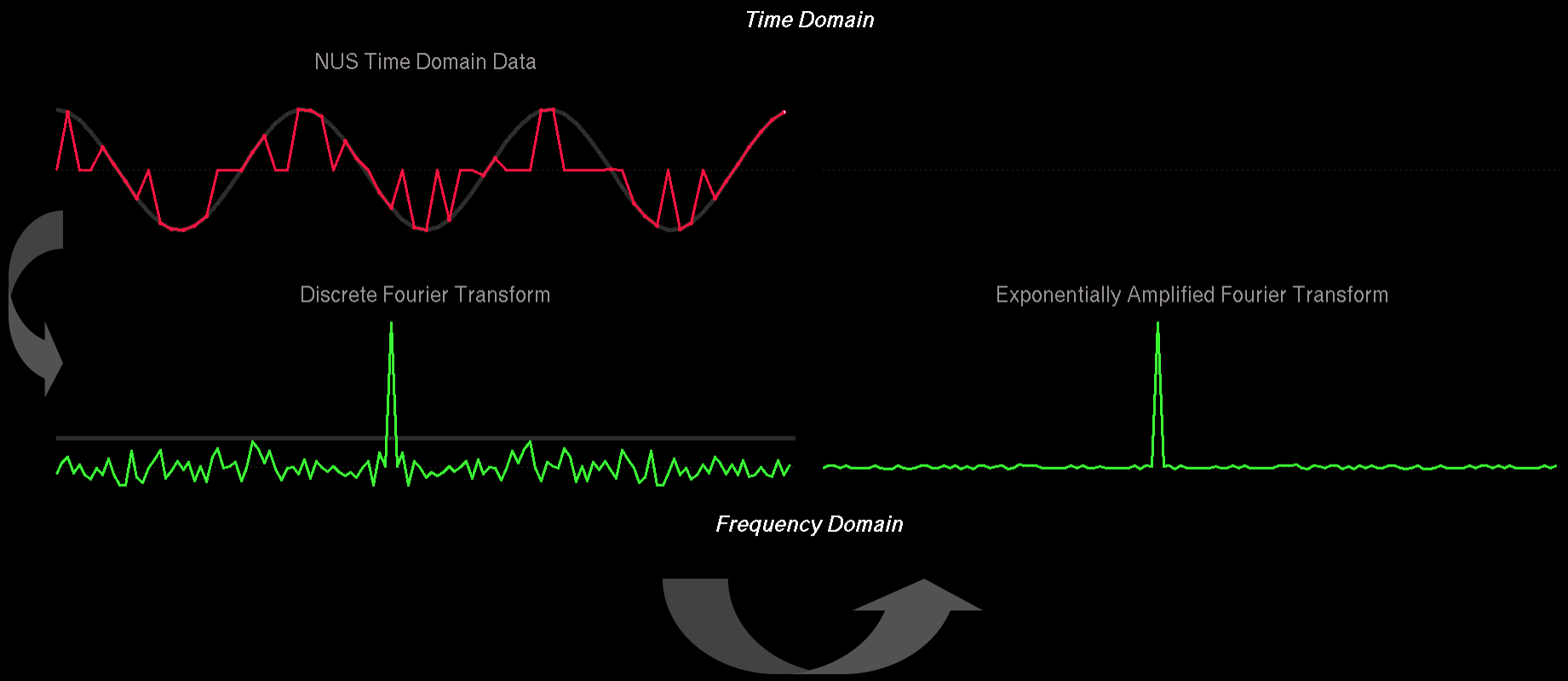


Frequency Domain

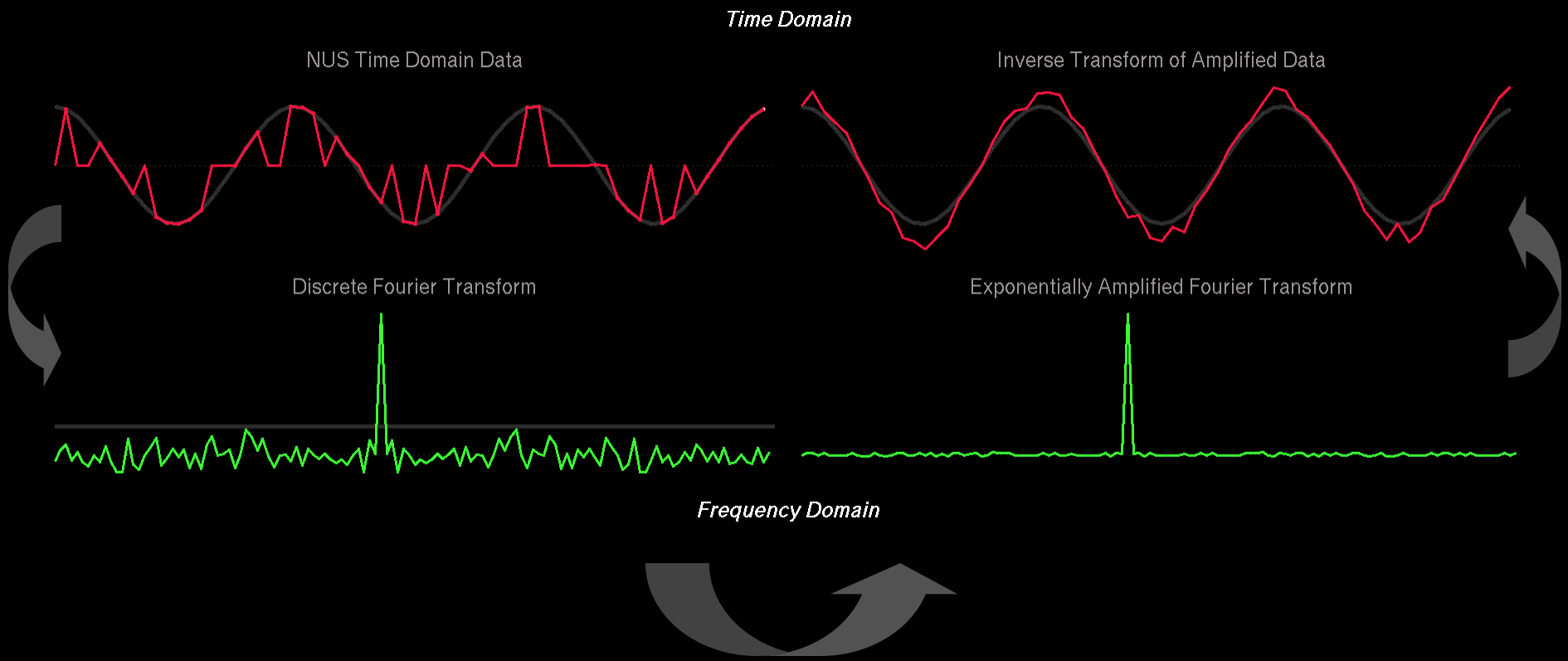
Time Domain



Frequency Domain

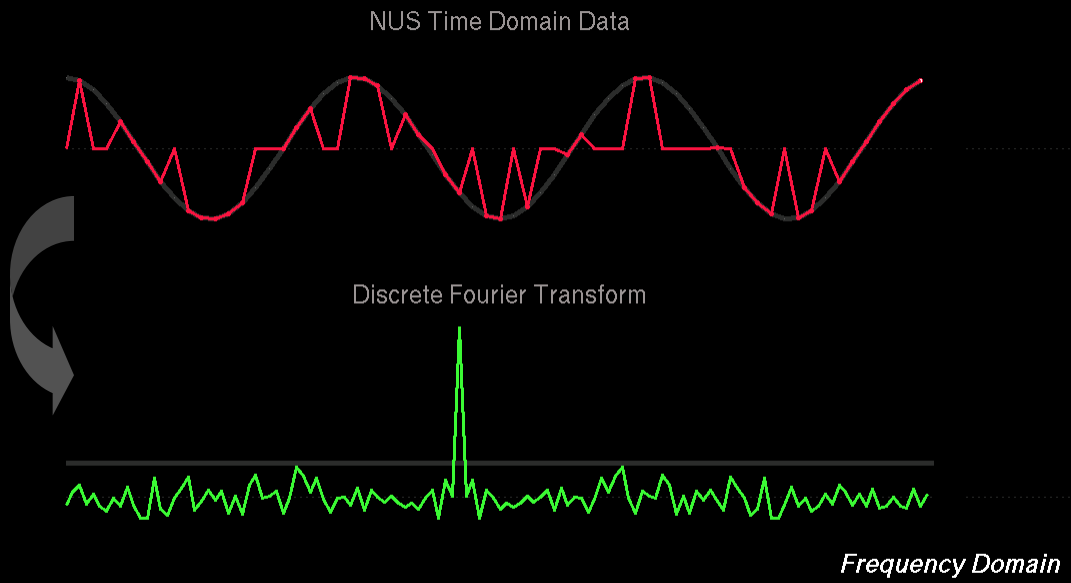


Maximum Entropy Methods

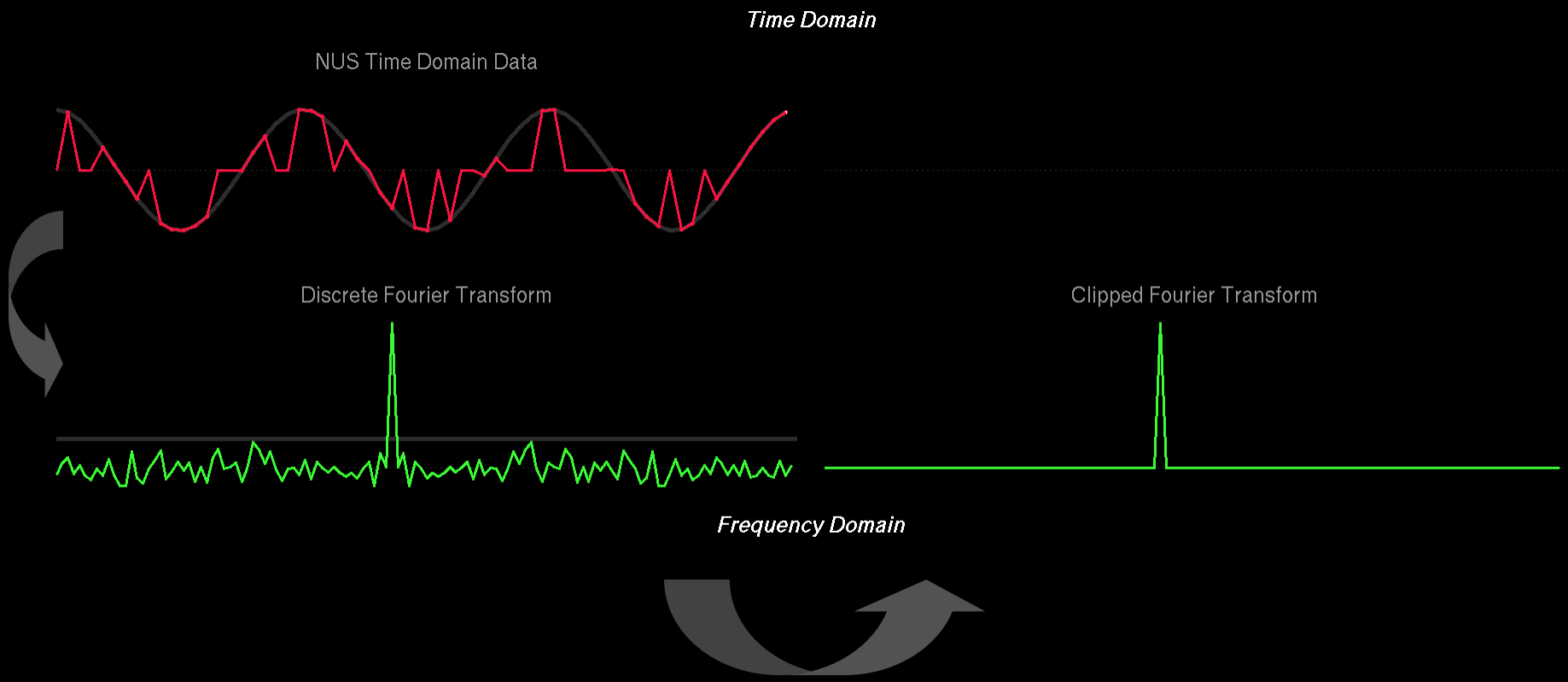


Maximum Entropy Methods

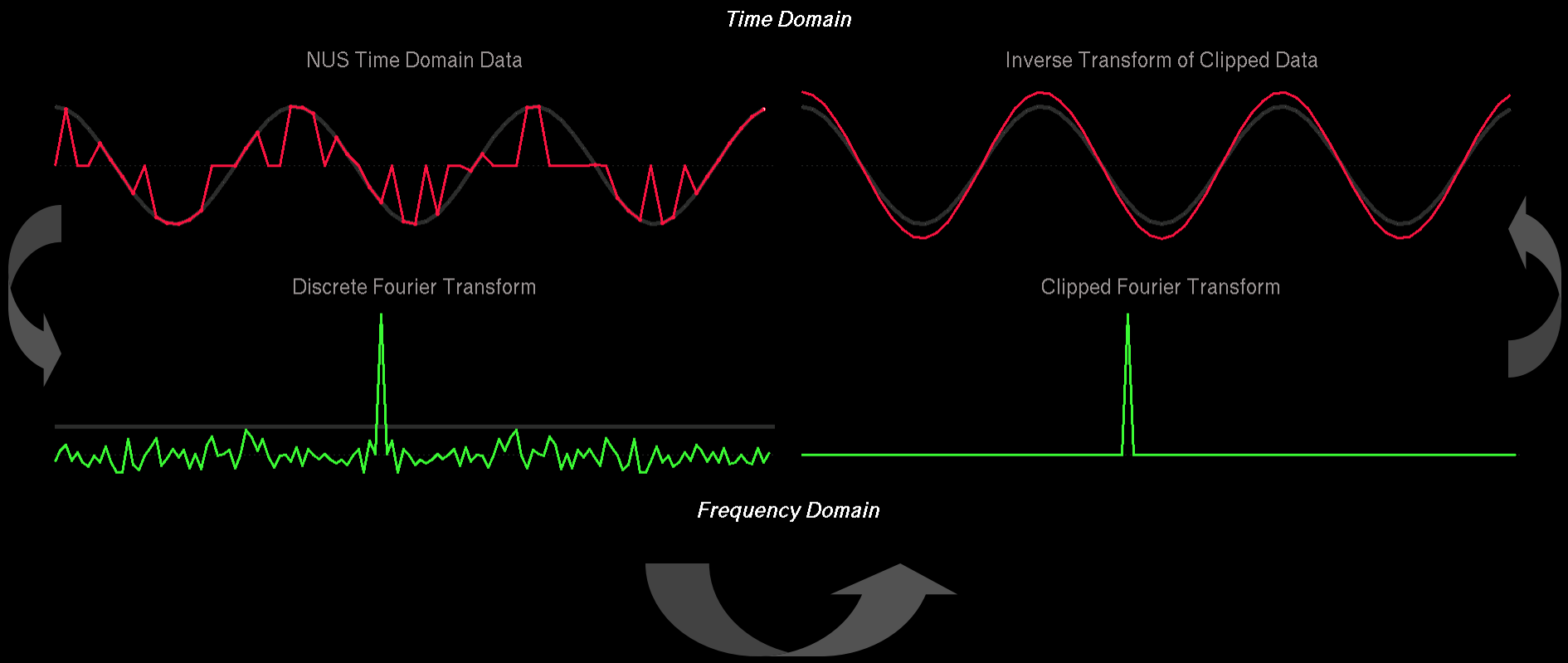
Time Domain



Frequency Domain



Threshold Methods



Threshold Methods

Maximum Entropy Method

Time Domain

Inverse-Transformed MEM

Multiplied by Distortion

Original Time-Domain

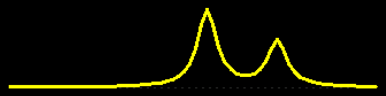
Original - Distorted MEM

Frequency Domain

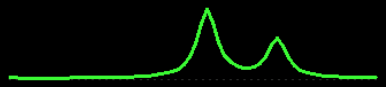
Proposed MEM Spectrum

MEM Spectrum with Distortion

Original Spectrum

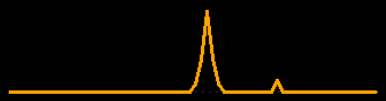


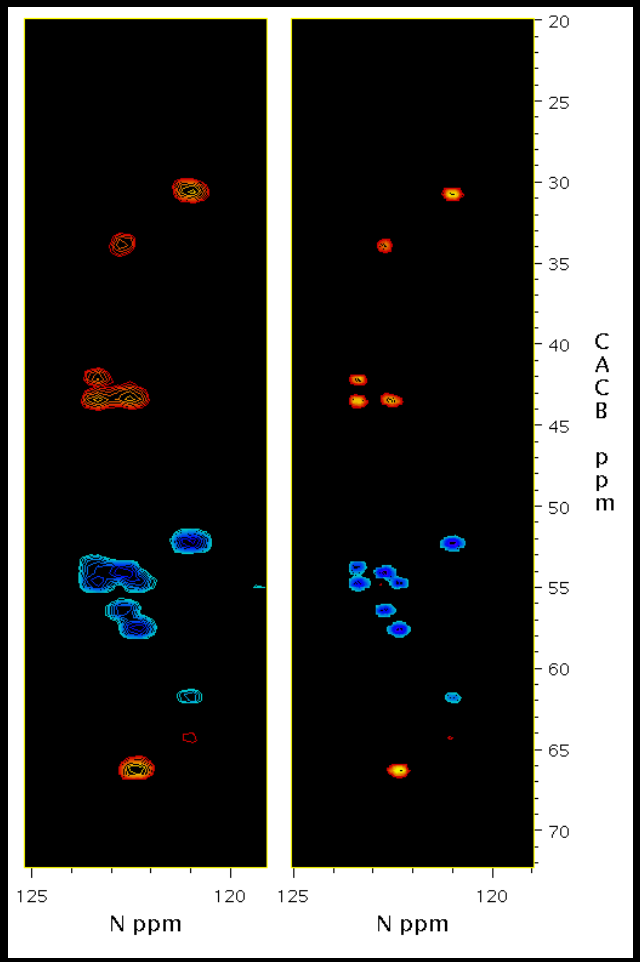
Transformed Residual



Residual with Threshold

Exponential Amplification



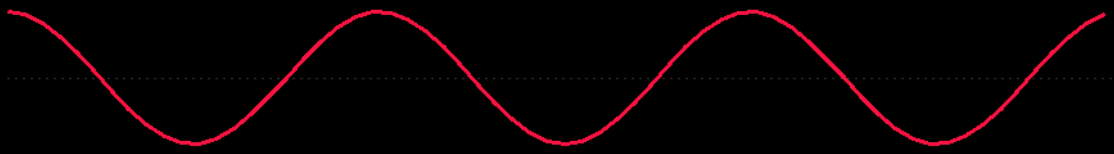


Maximum Likelihood Frequency Map

Model: Fourier Term (A)

Model Scaled by c to Minimize RMS (scale c = $\sum \frac{A\,B}{A\,A}$)

Subtracted from this Time Domain Data (B)



Forms the Difference (B – c A)

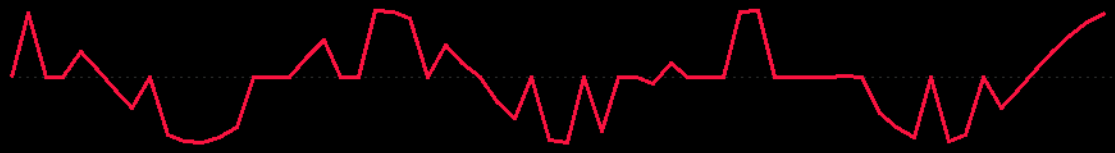
Find RMS to Form a Frequency Map Point:

Maximum Likelihood Frequency Map of NUS Data

Model: NUS-Adjusted Fourier Term (A)

Model Scaled by c to Minimize RMS (scale c = $\sum \frac{A\,B}{A\,A}$)

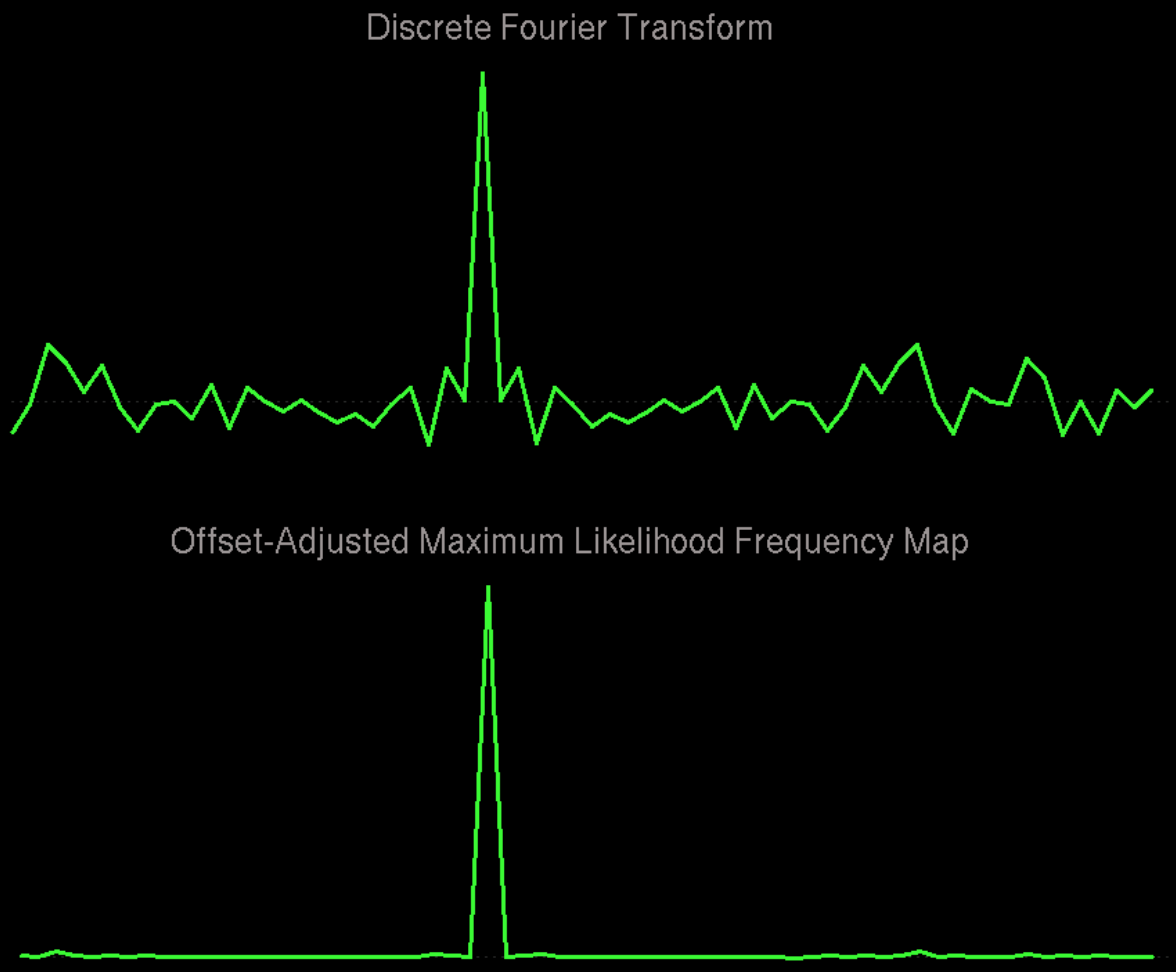
Subtracted from NUS Time Domain Data (B)



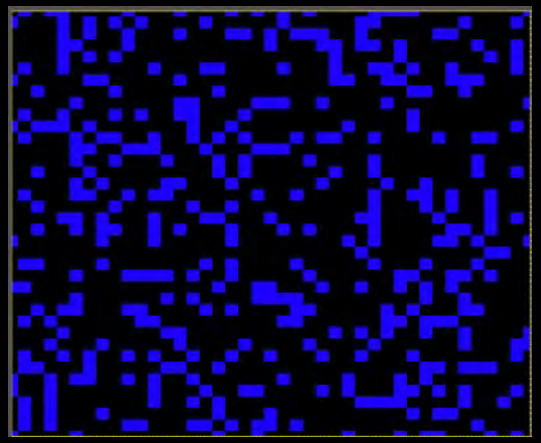
Forms the Difference (B – c A)

Find RMS to Form a Frequency Map Point:

Fourier Transform vs Maximum Likelihood Frequency Map



CBCA(CO)NH
NUS Sampling Schedule
25% Density

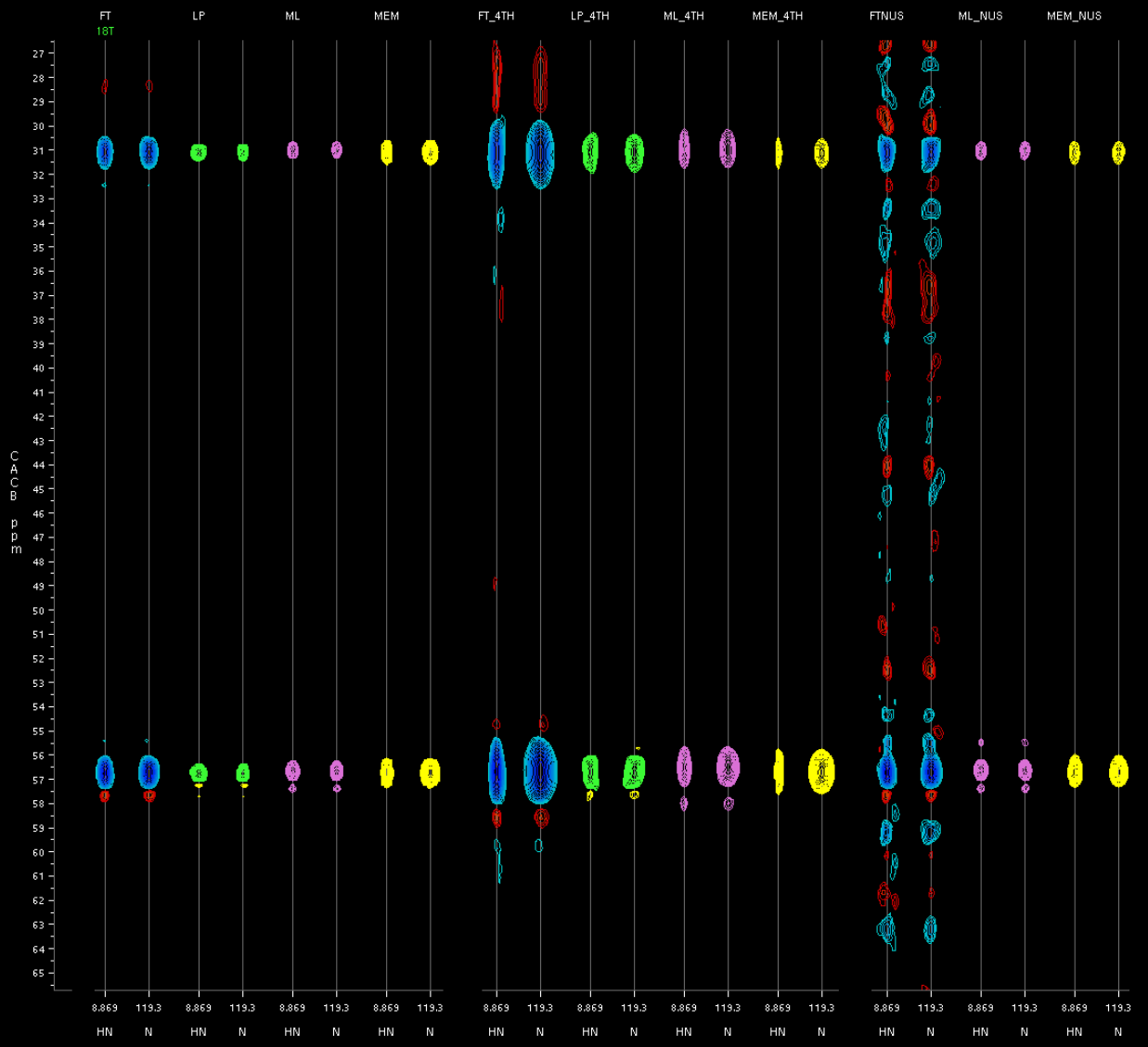


^{15}N (Uniform)

^{13}C (10% Decay)

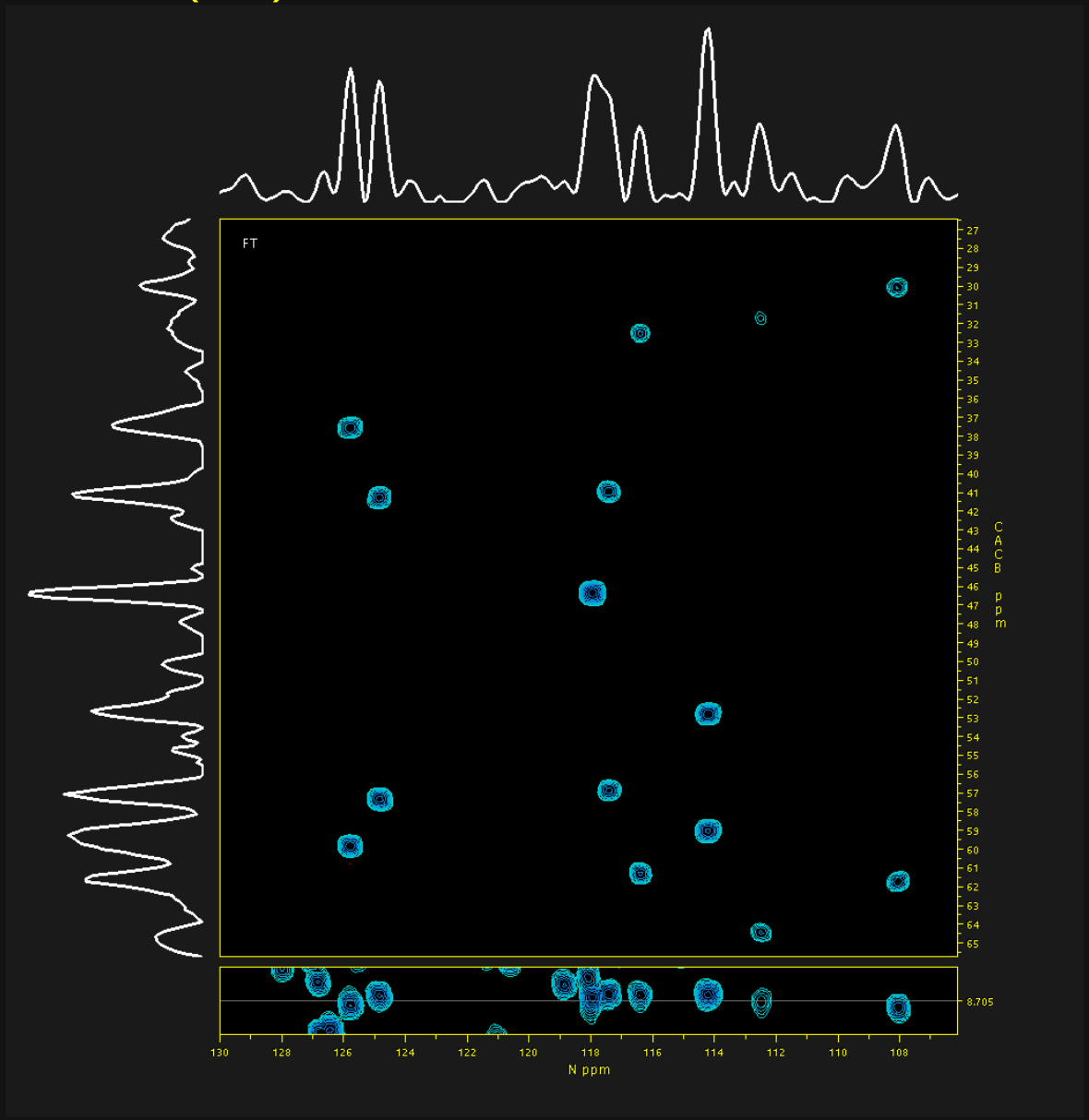
Original Data | ¼ Resolution | NUS

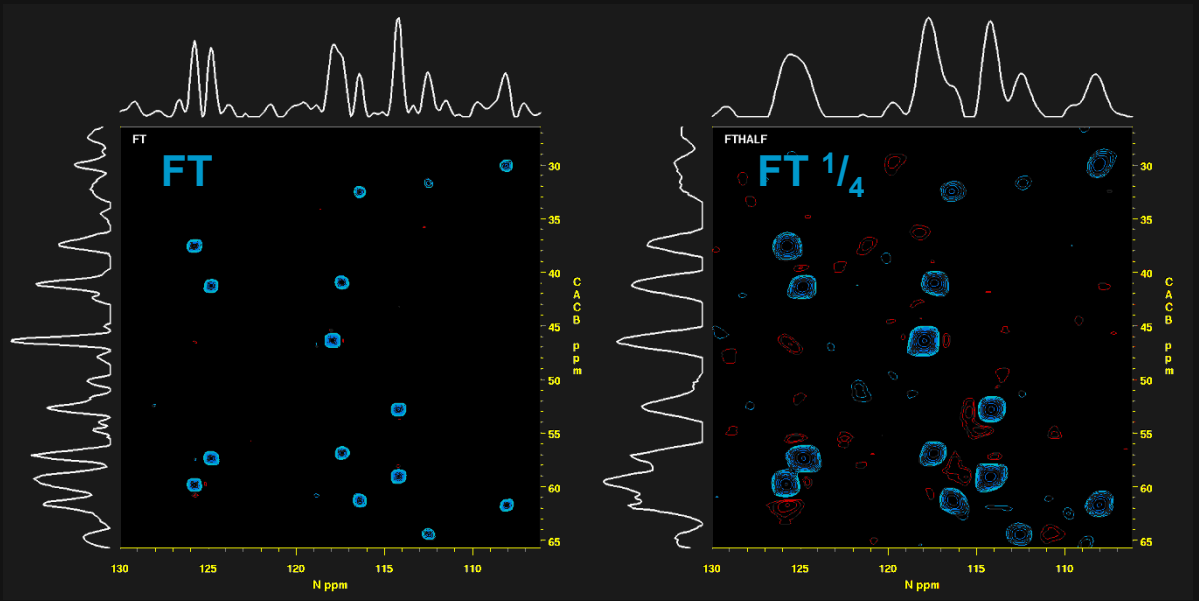
FT LP ML MEM FT LP ML MEM FT ML MEM

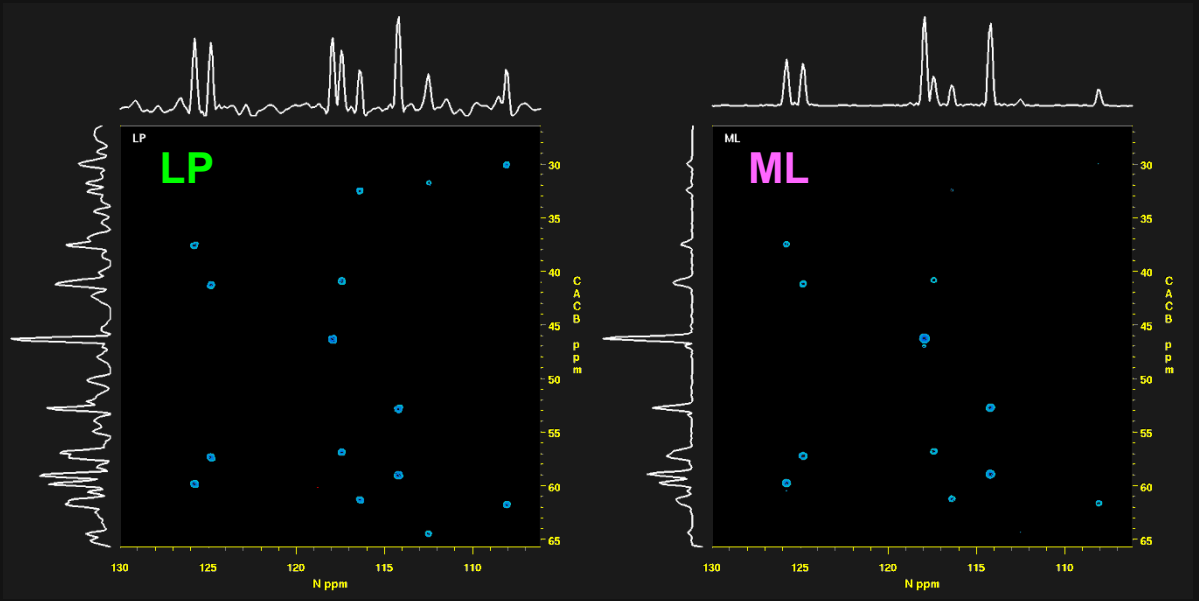


Agilent Technologies

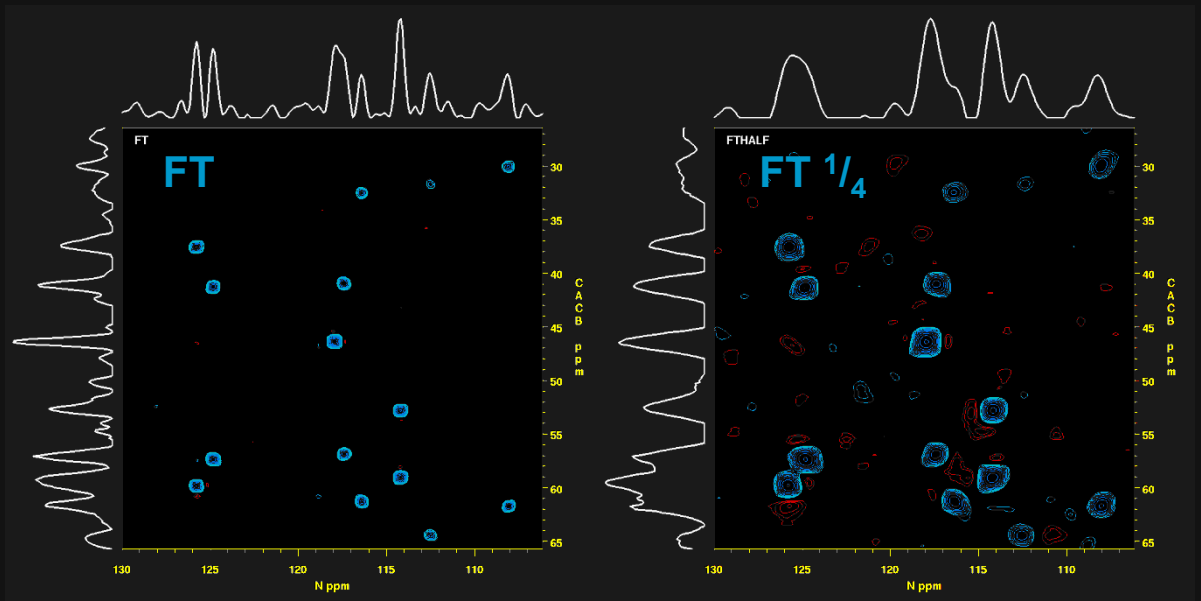
CBCA(CO)NH 15N 38* Points, 13C 41* Points

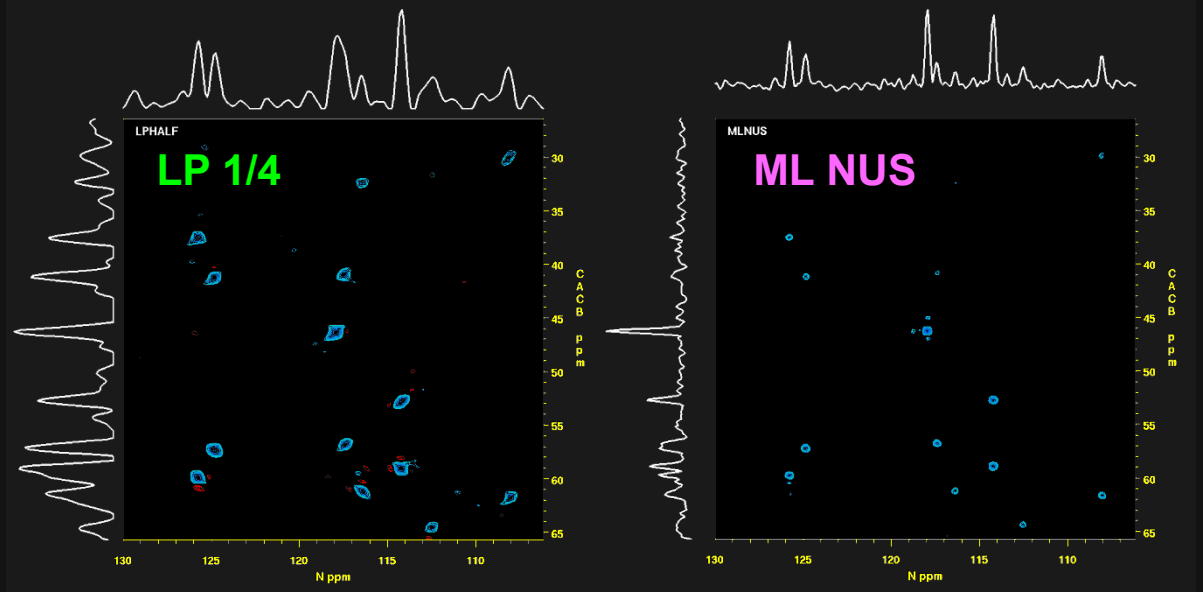






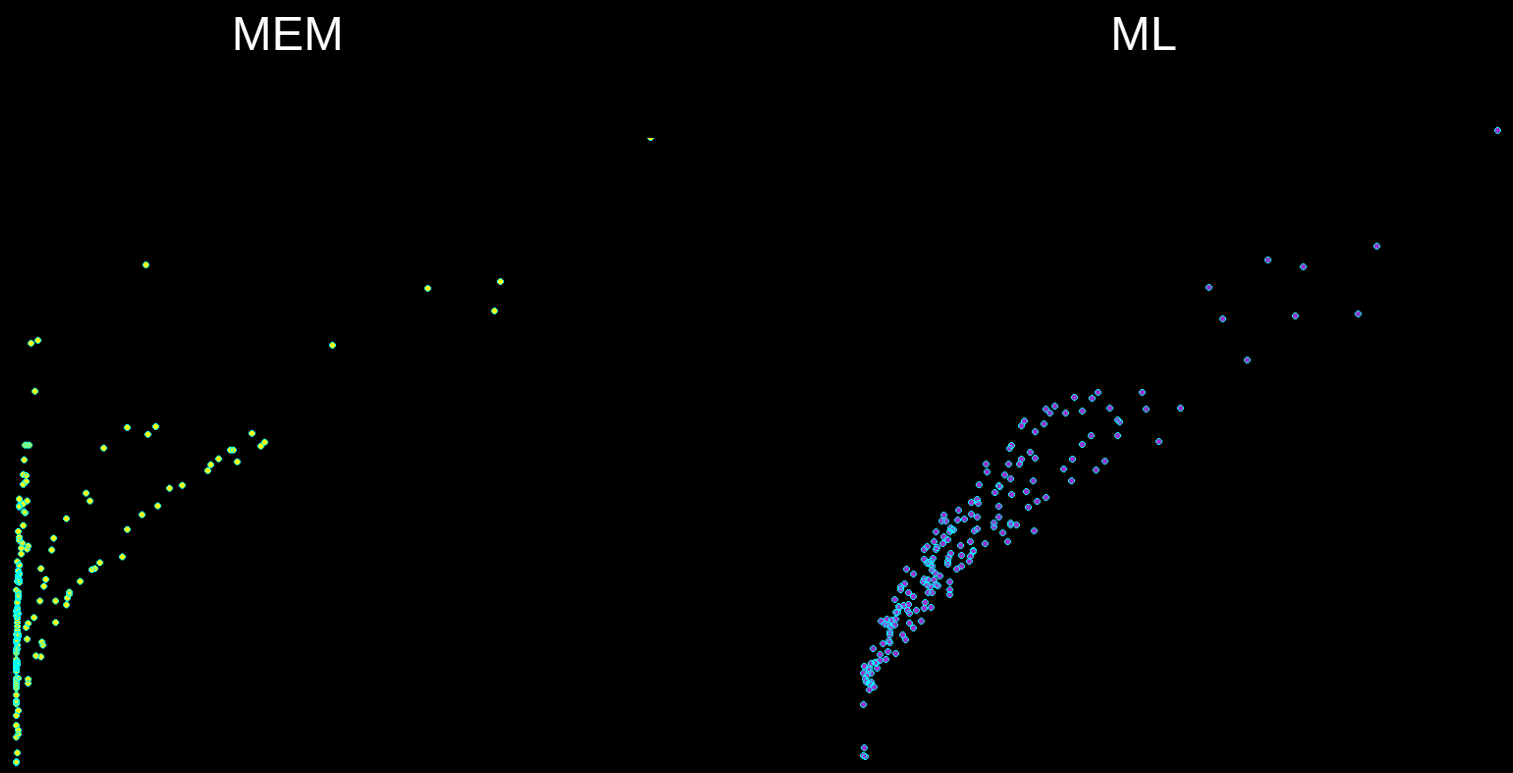
Agilent Technologies





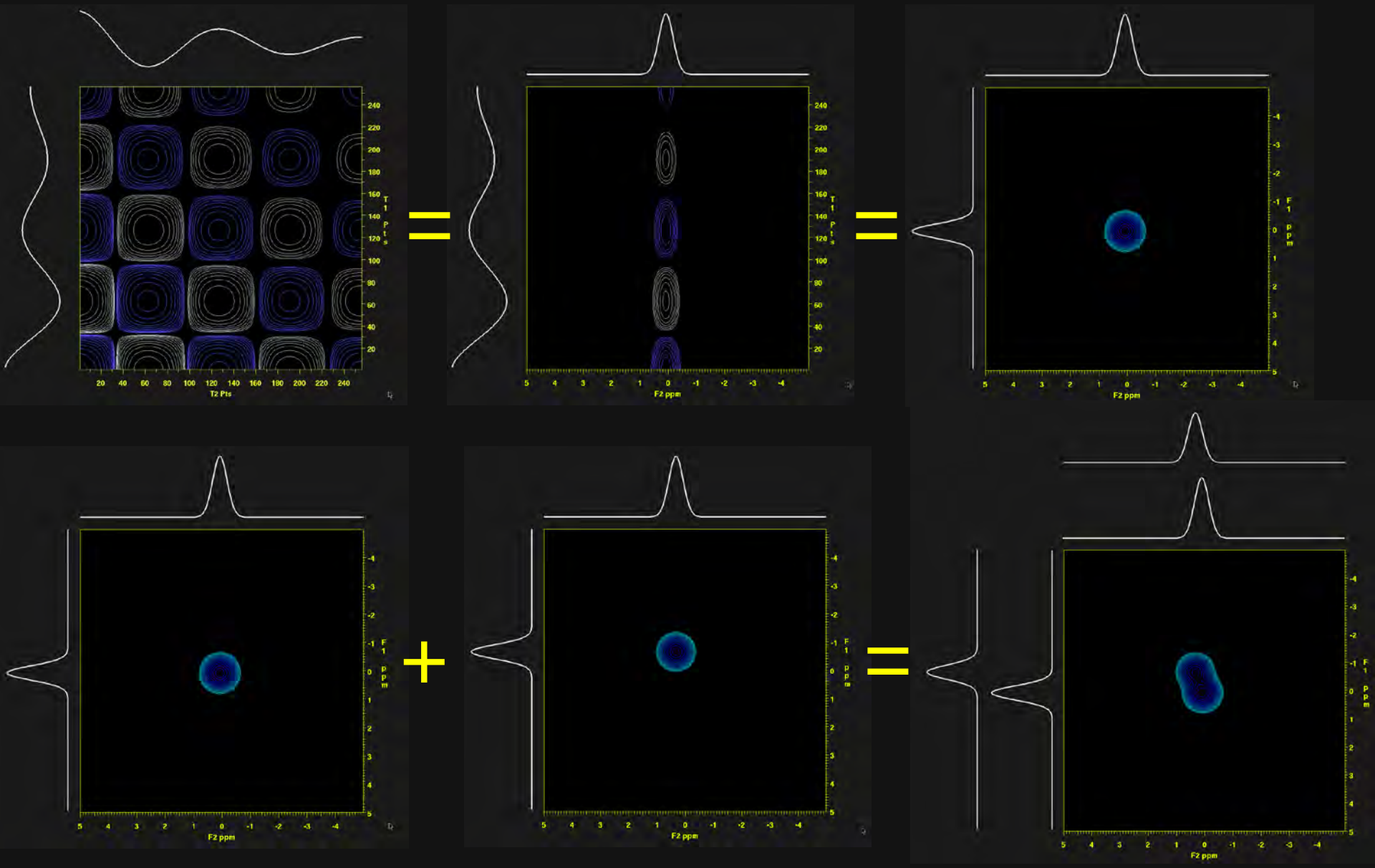
Agilent Technologies

Linearity – Reconstructed Peak Height vs Peak Height in FT Spectrum



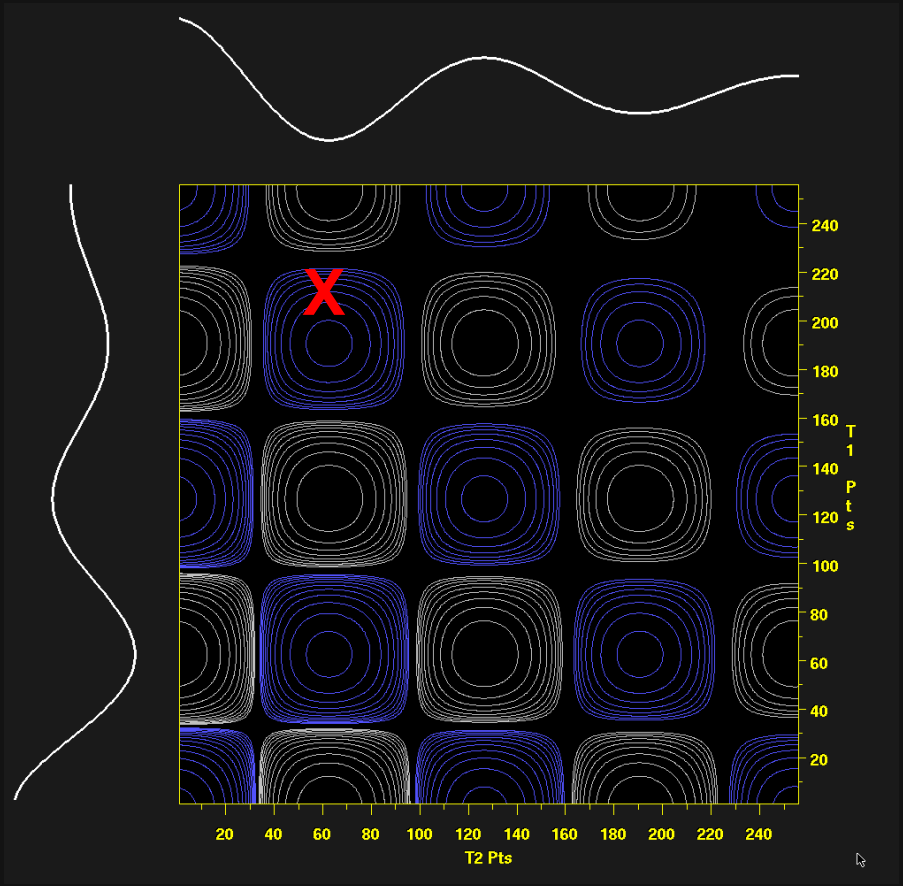
Spectral Matrix Decomposition

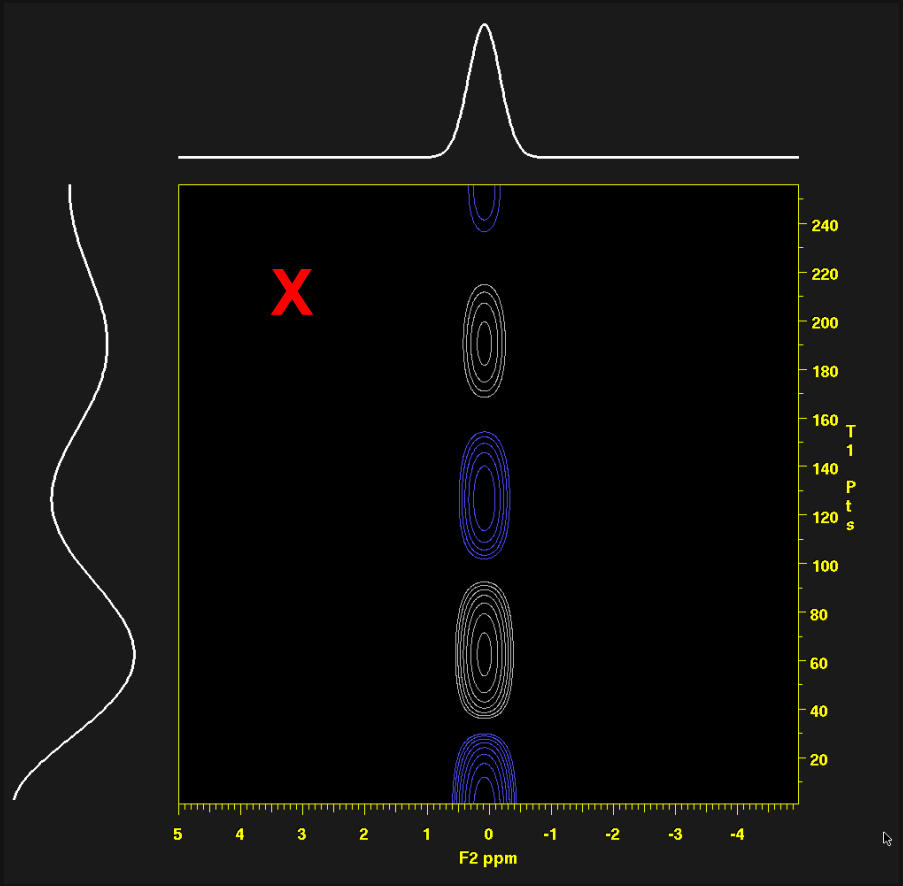
Matrix Decomposition: Each Multidimensional NMR Signal is the Product of 1D Vectors



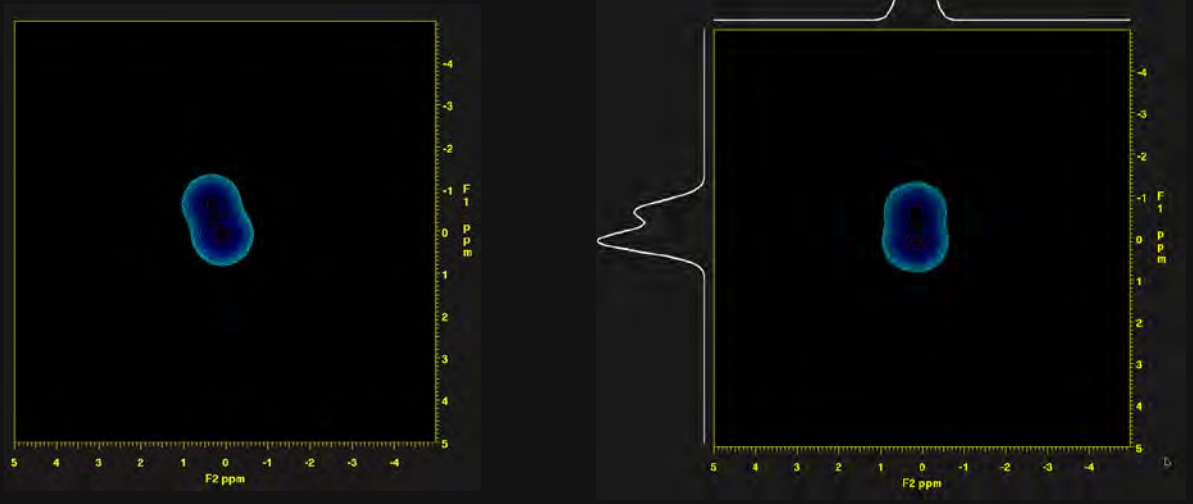
Agilent Technologies

NUS – Missing Information in a Deleted Point is Also Contained in the Same Row or Column



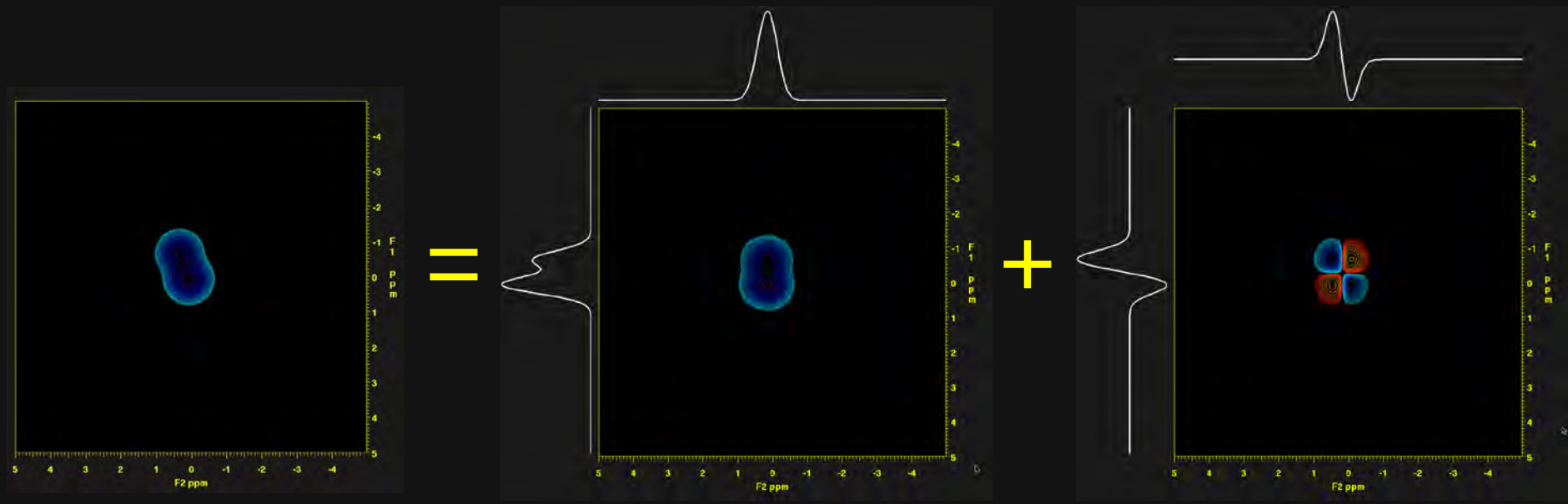


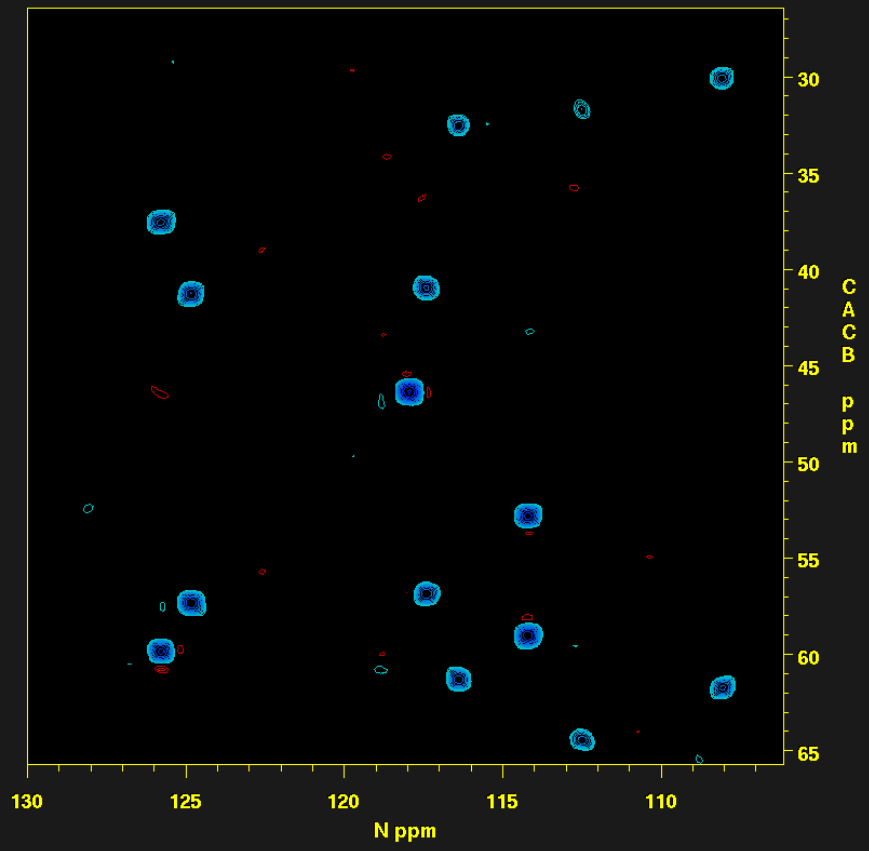
Spectral Matrix Decomposition by Principal Component Analysis (PCA)

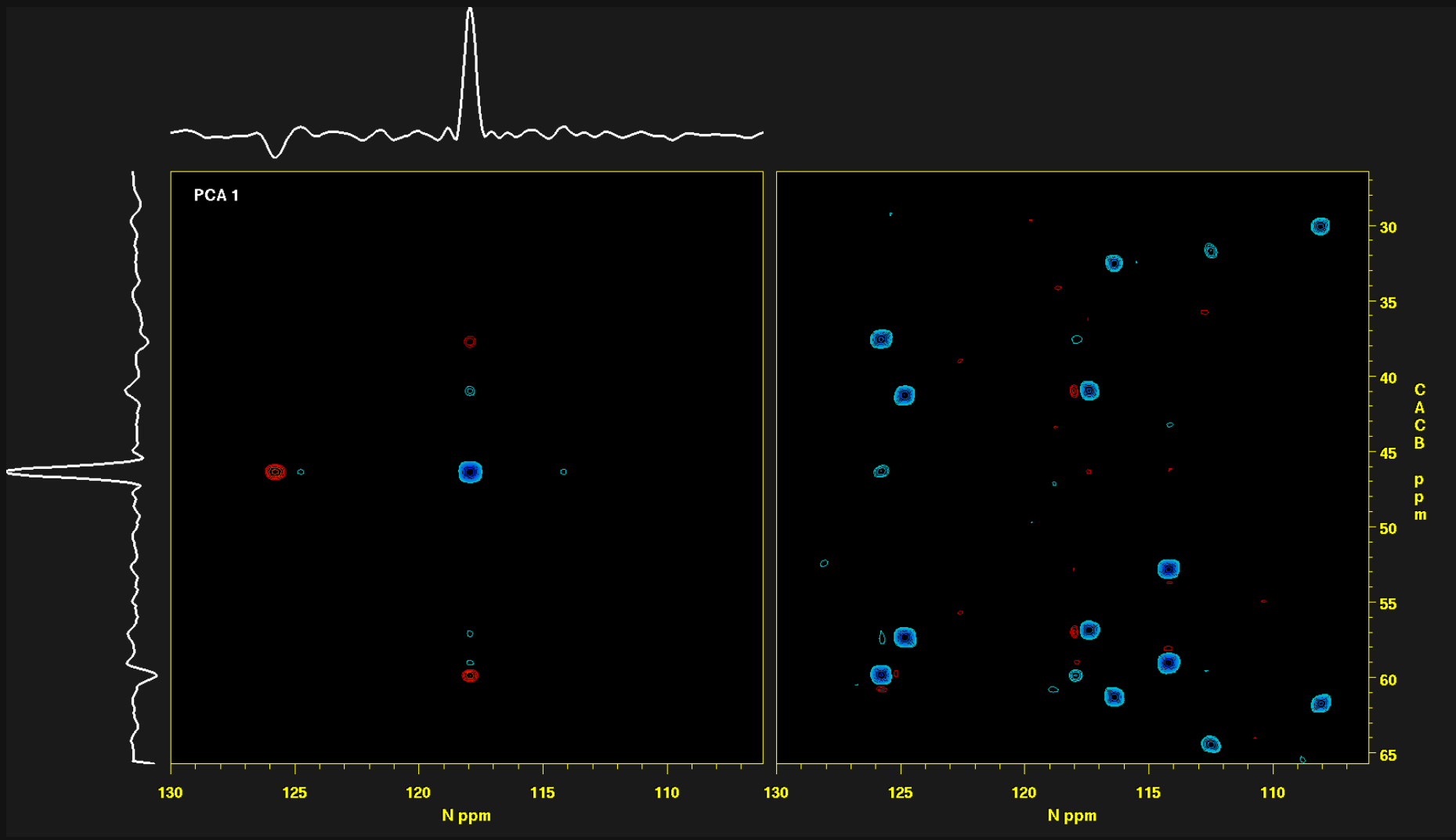


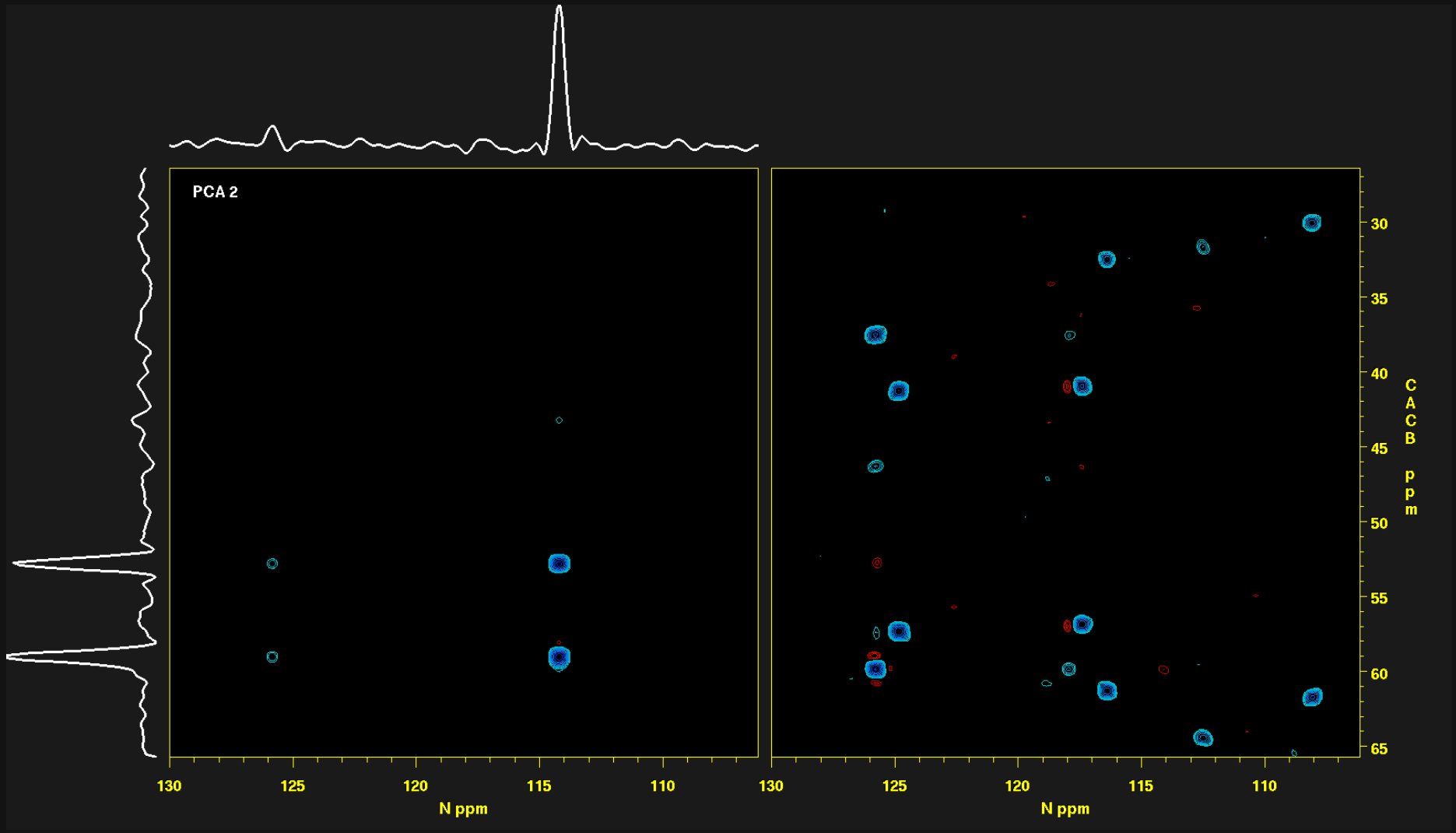
Agilent Technologies

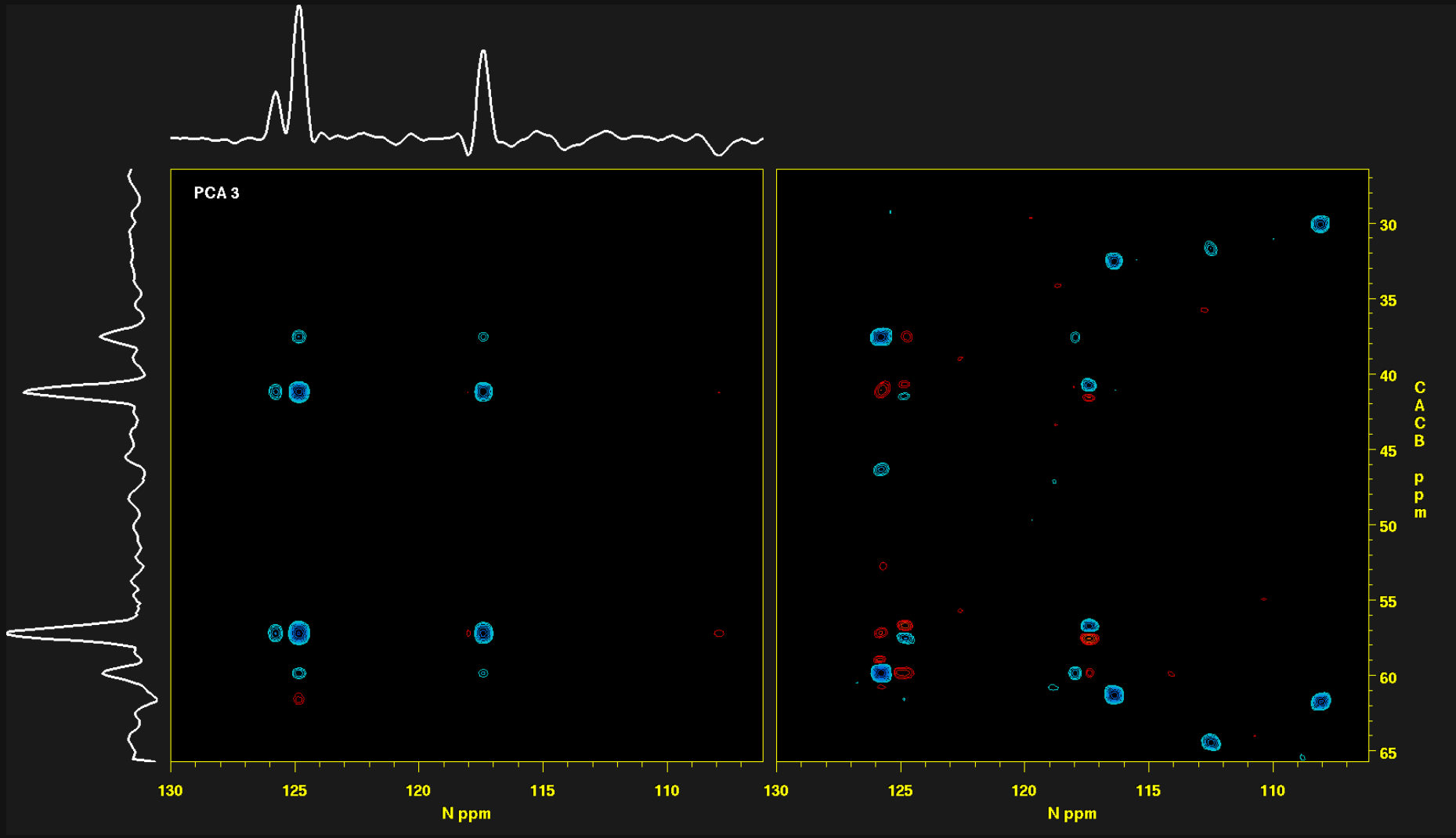
Spectral Matrix Decomposition by Principal Component Analysis (PCA)

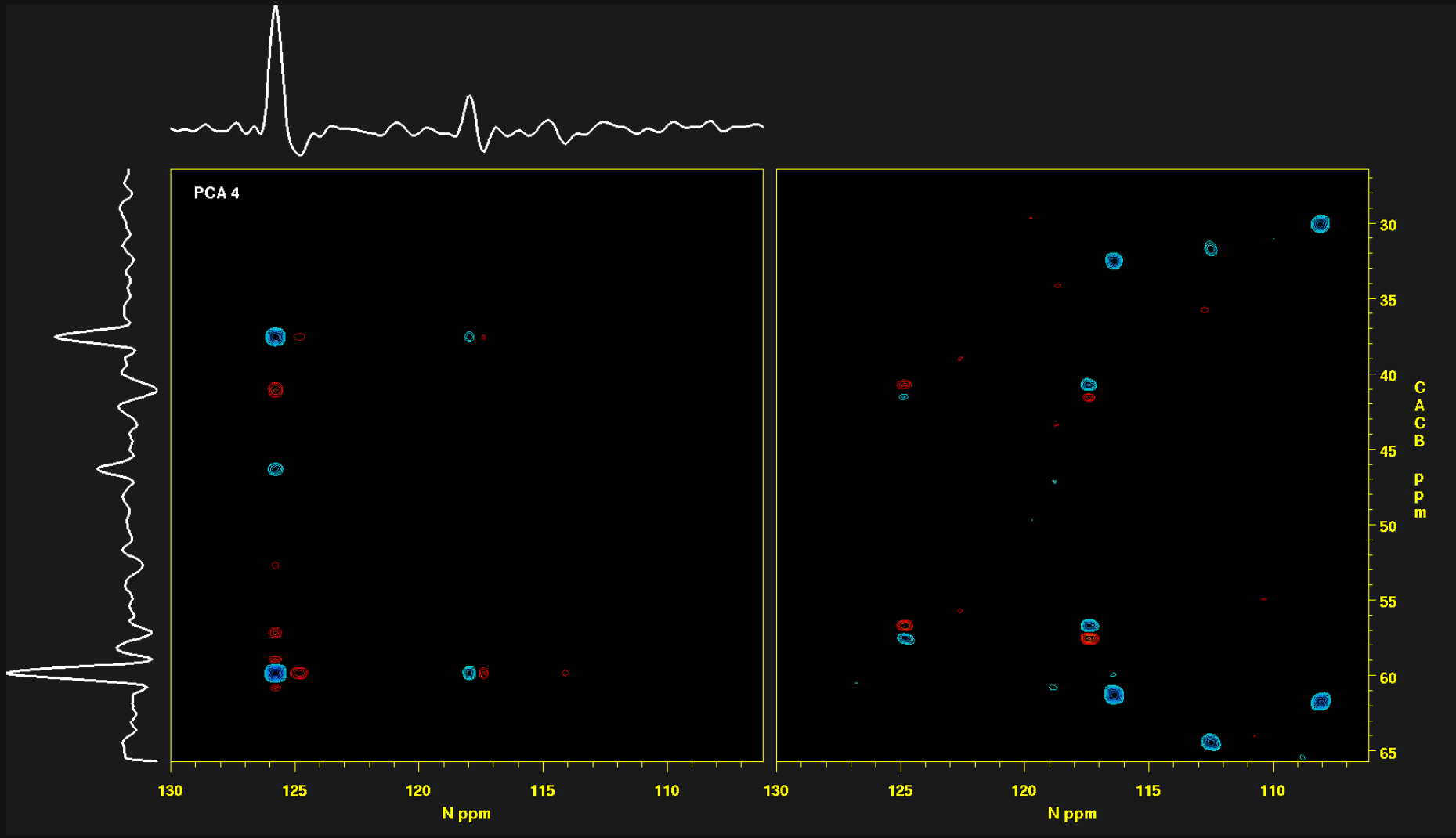


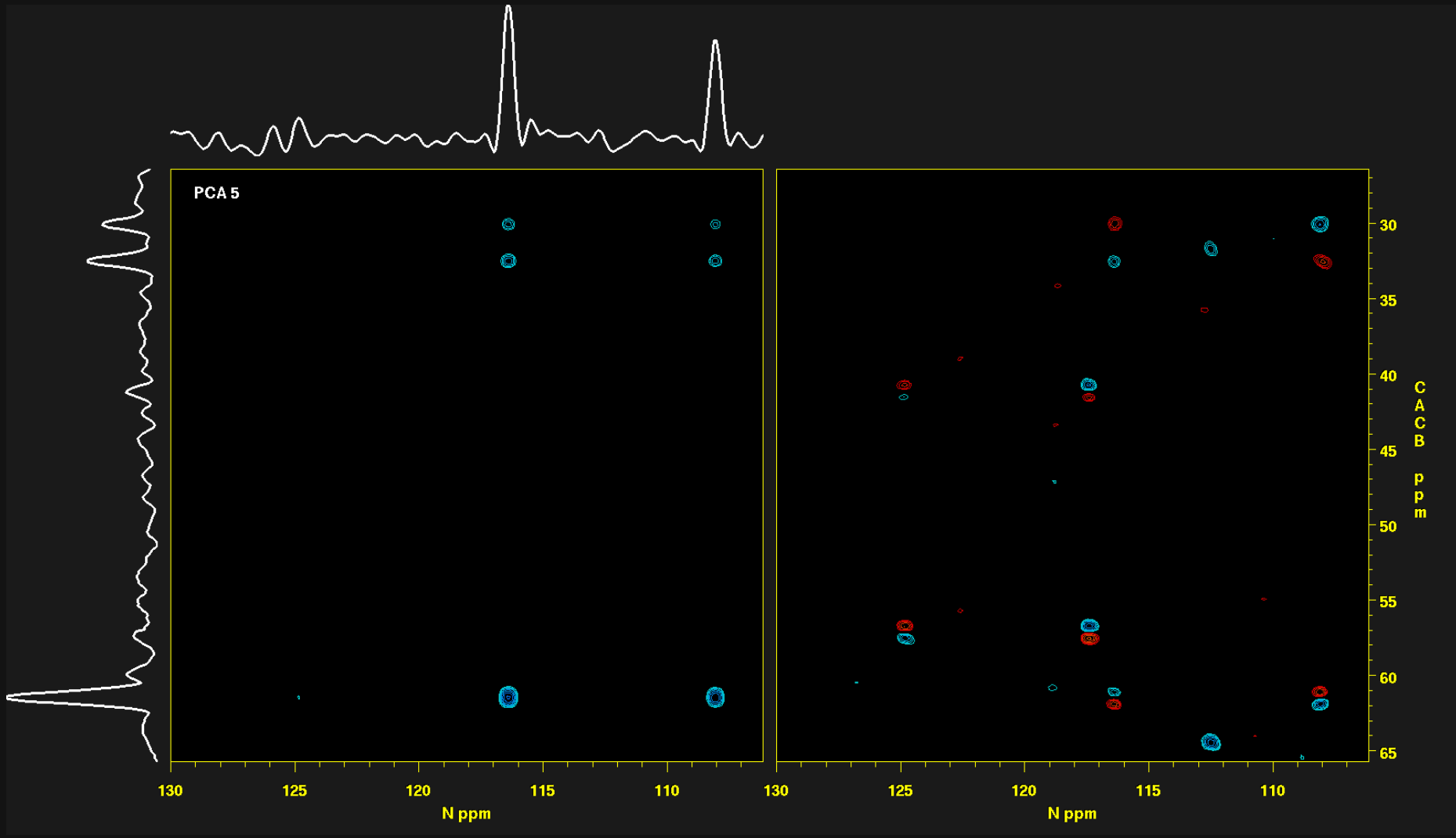

30
35
40
45
50
55
60
65
CA CB ppm
130
125
120
115
110
N ppm

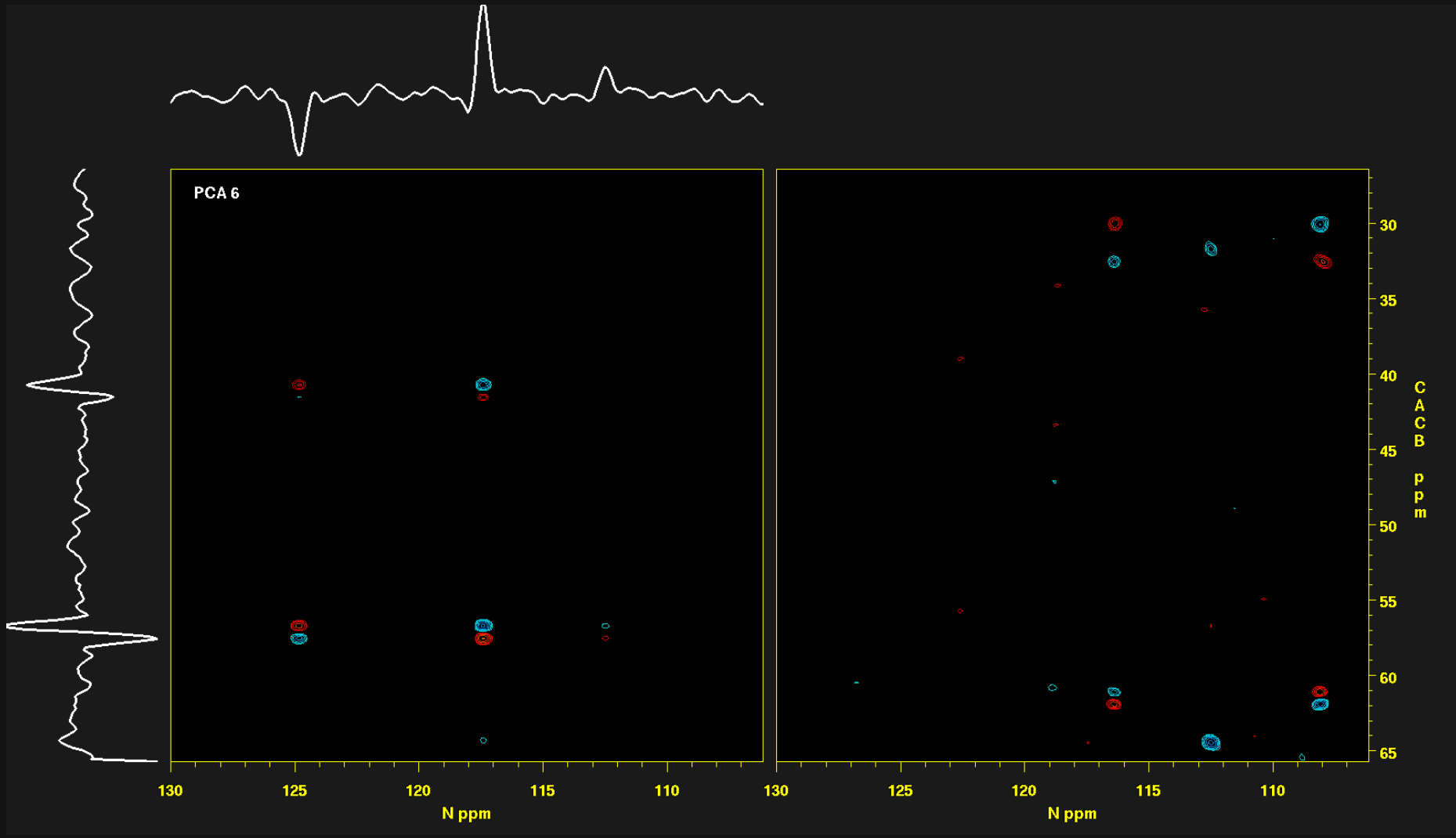


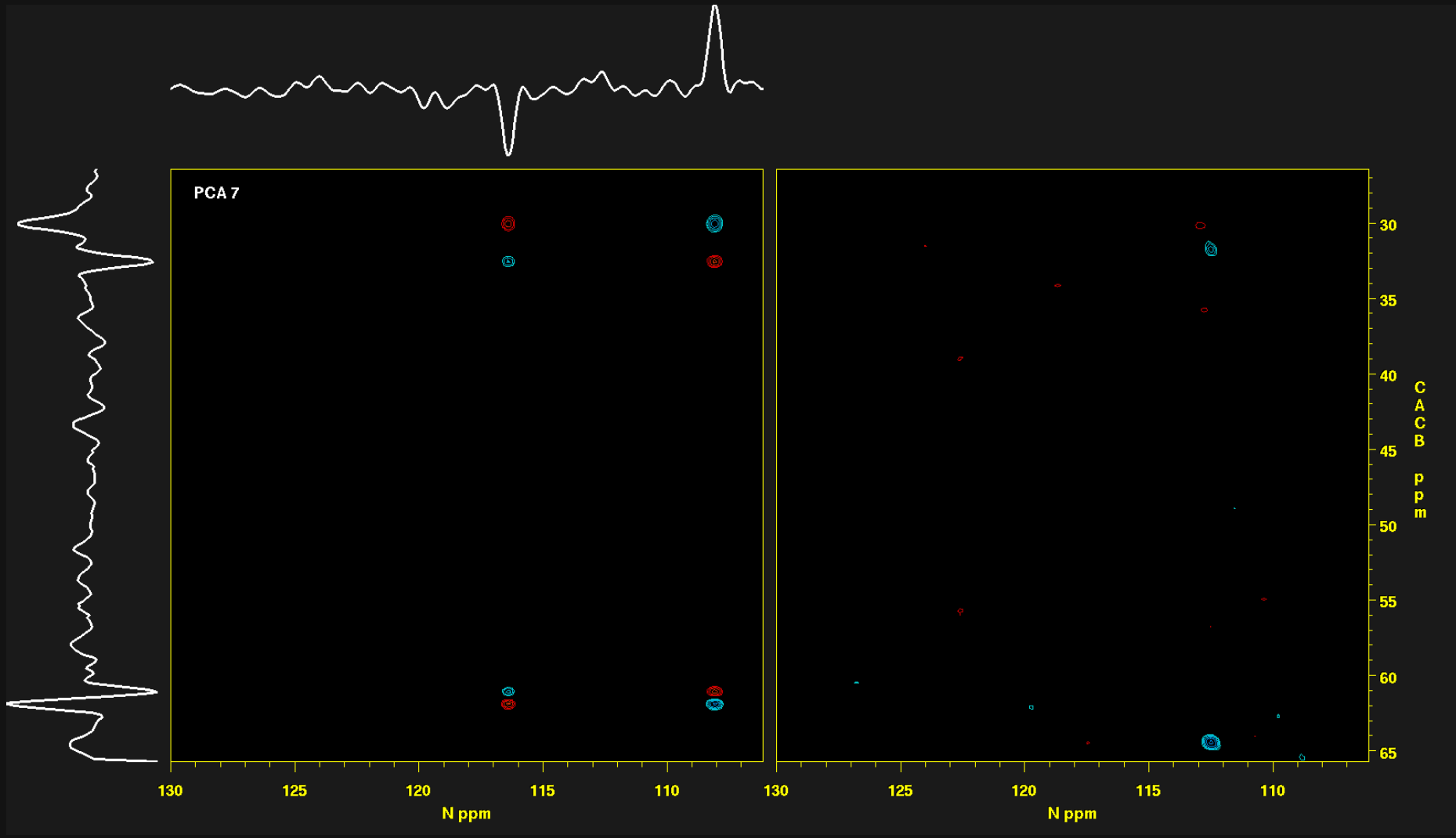


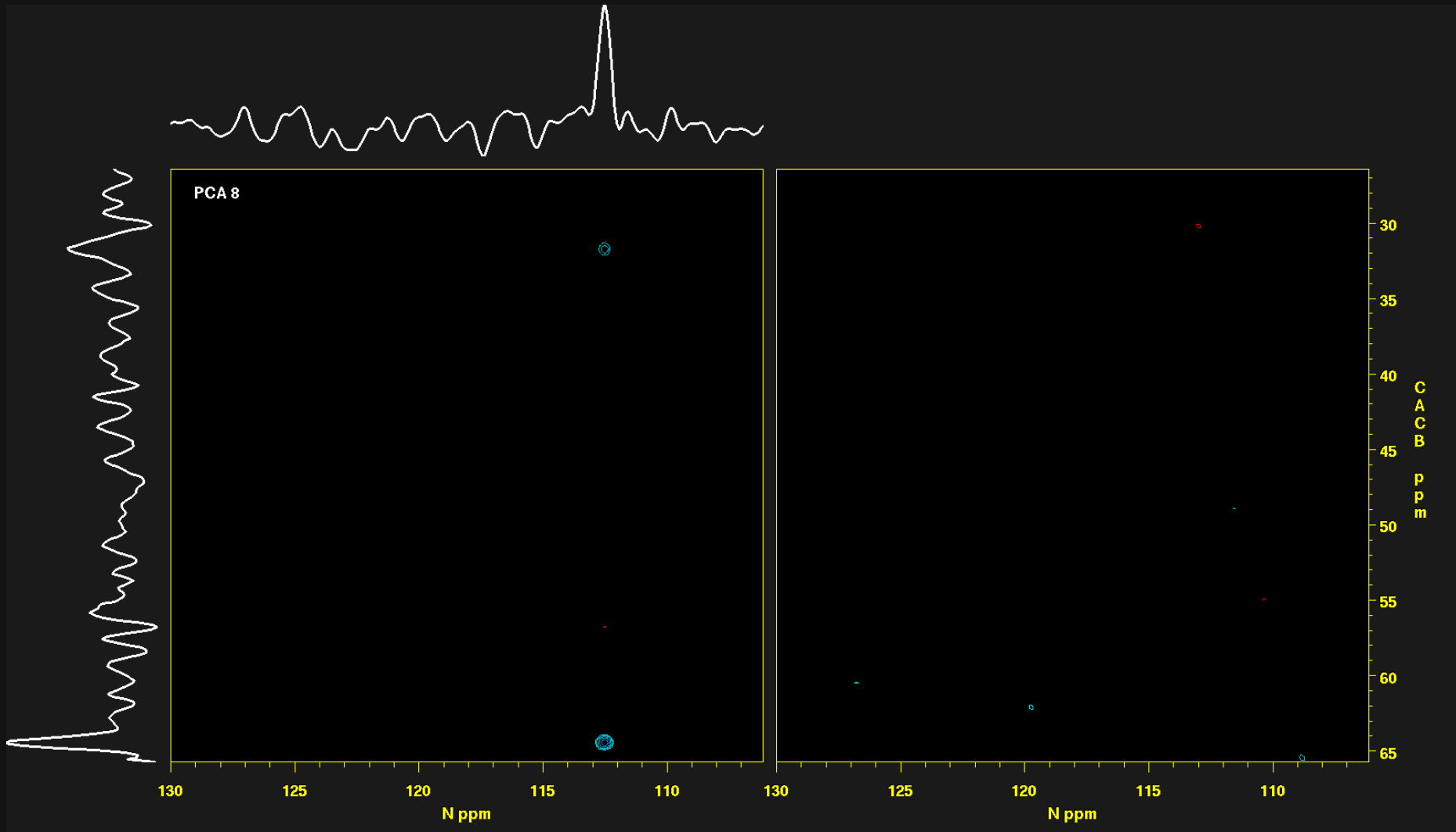


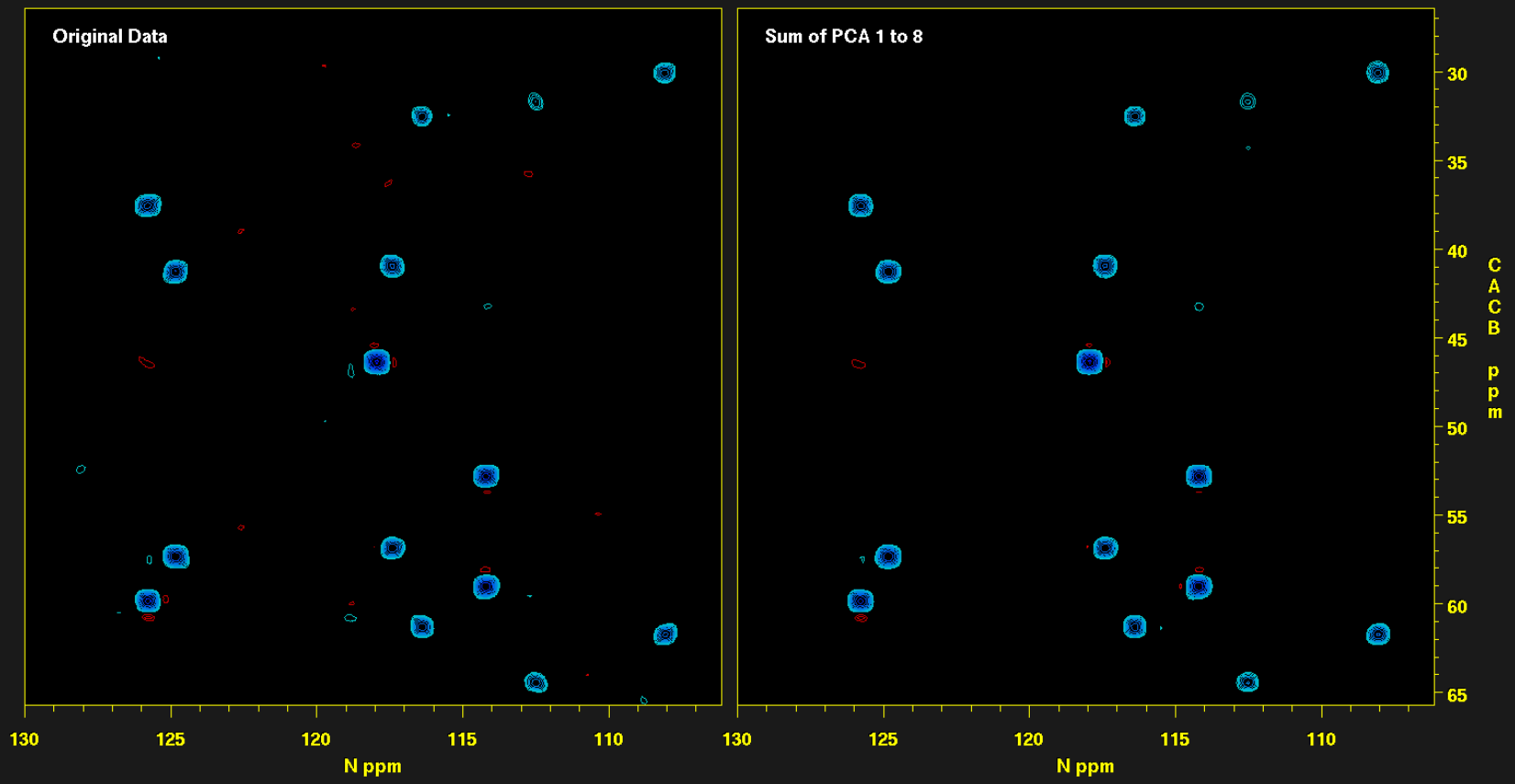


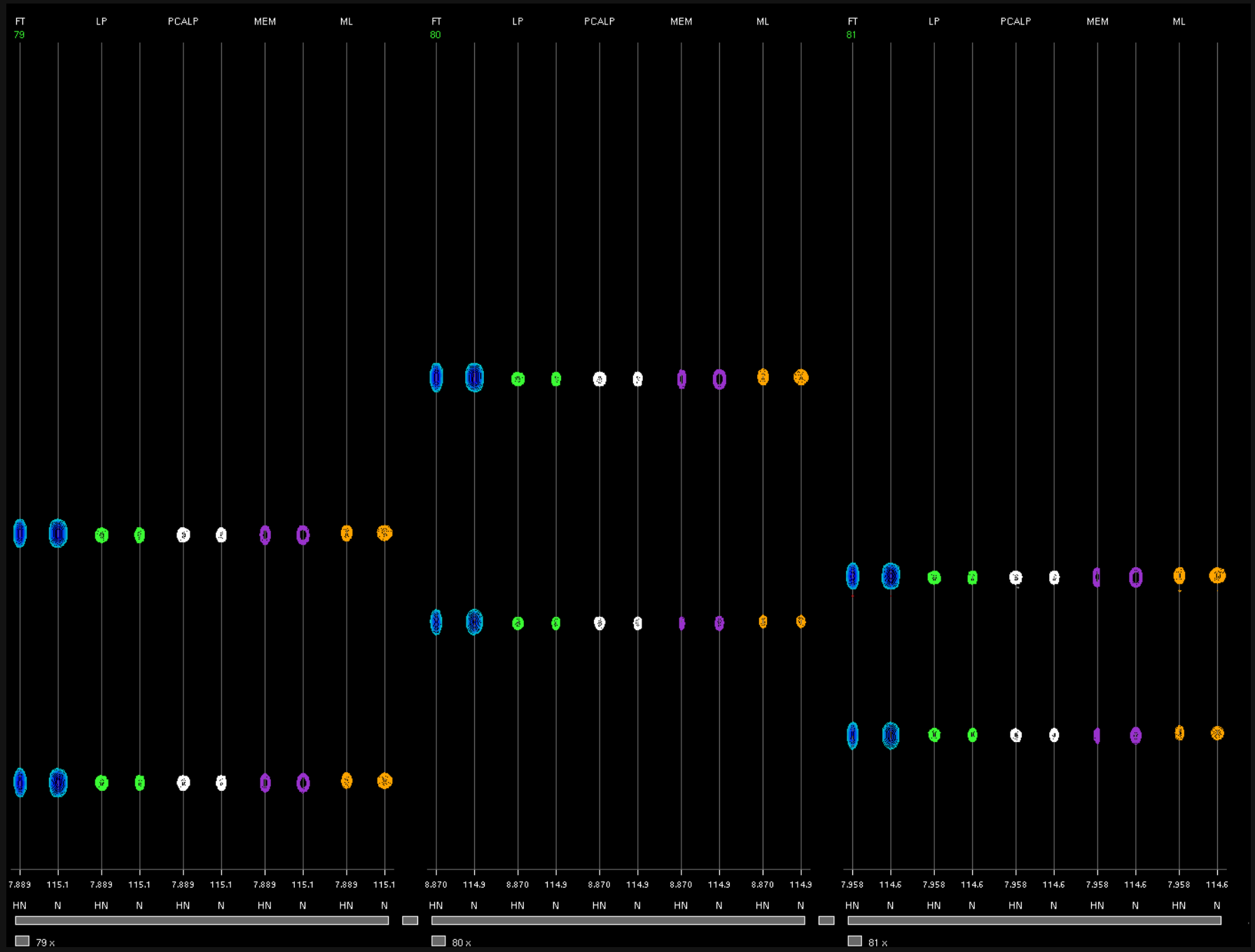





PCA 7
130
125
120
115
110
N ppm
130
125
120
115
110
N ppm
30
35
40
45
50
55
60
65
CACB ppm


PCA 8
130
125
120
115
110
N ppm
130
125
120
115
110
N ppm
30
35
40
45
50
55
60
65
CACB ppm

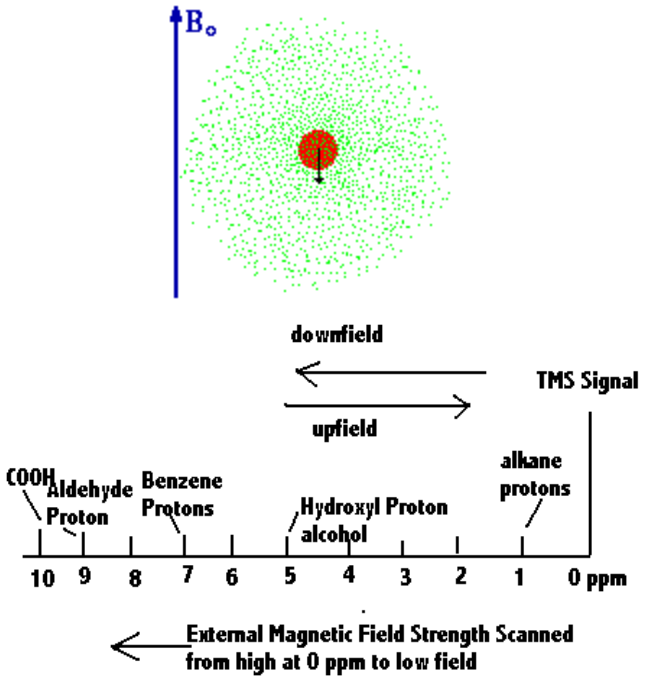



FT
LP
PCALP
MEM
ML
79
80
81
7.889
115.1
8.870
114.9
7.958
114.6
HN
N
79 x
80 x
81 x

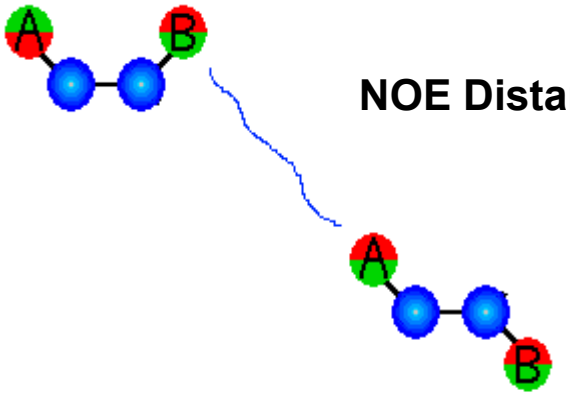
Structural Data from NMR

Structural Data from NMR

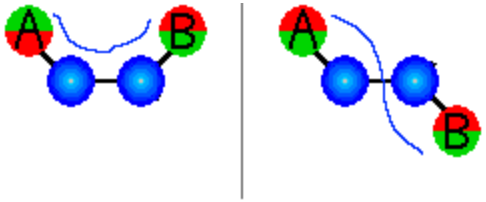
Chemical Shift

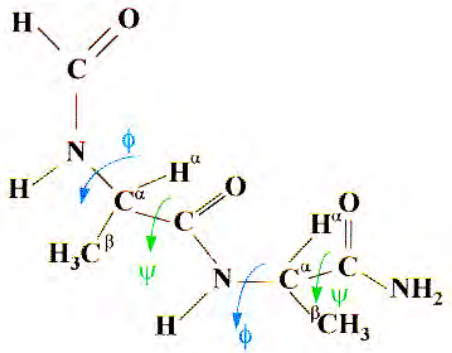


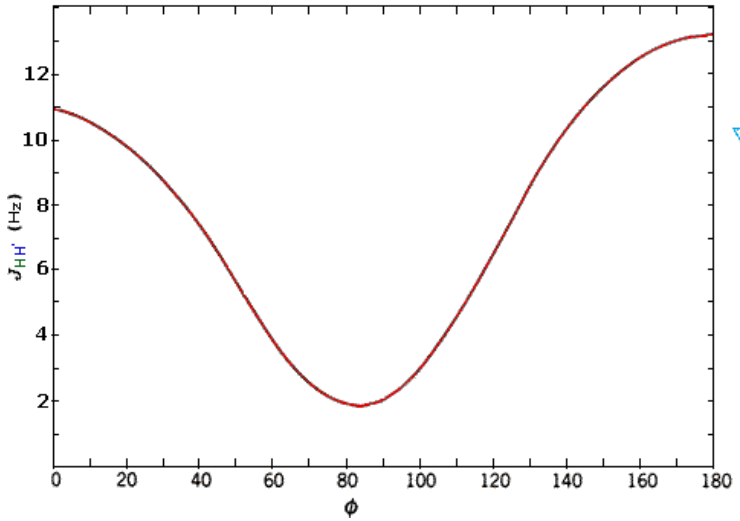
NOE Distance

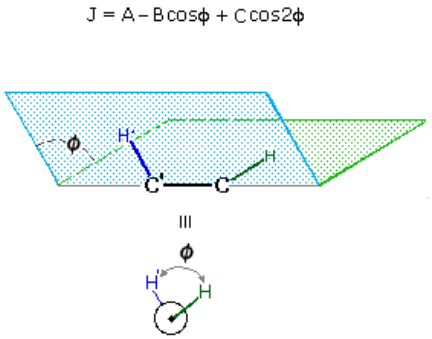


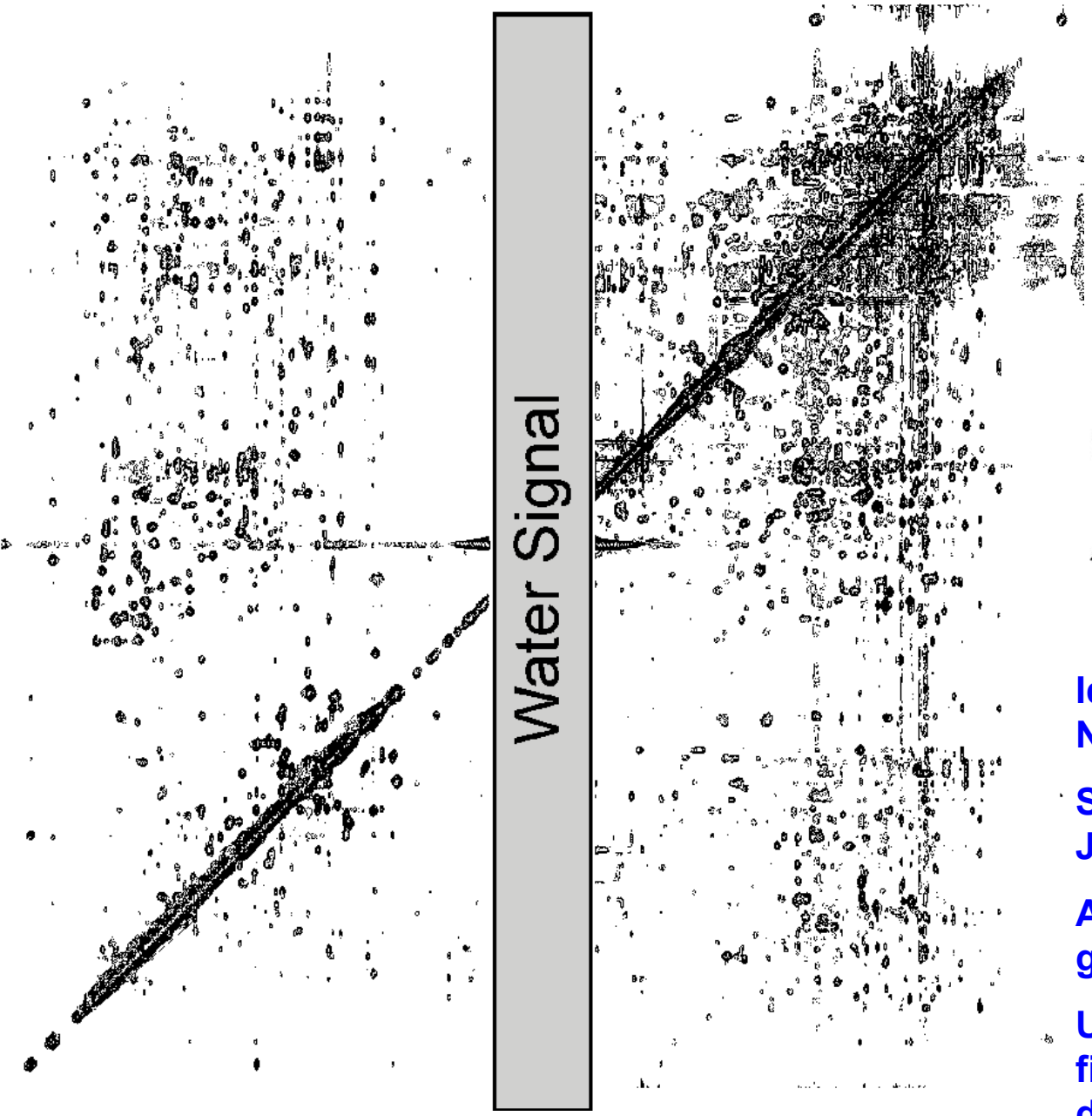
J-Coupling

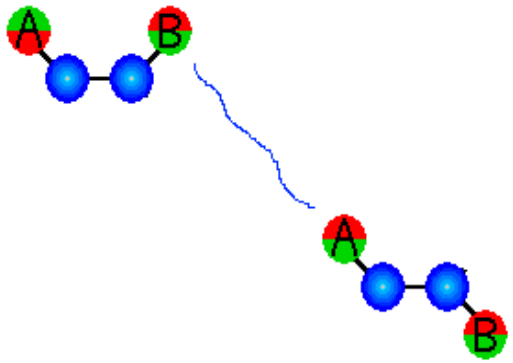




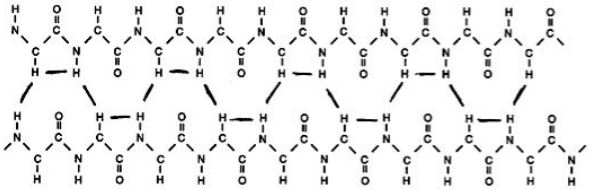








NOE cross peaks in parallel beta strands

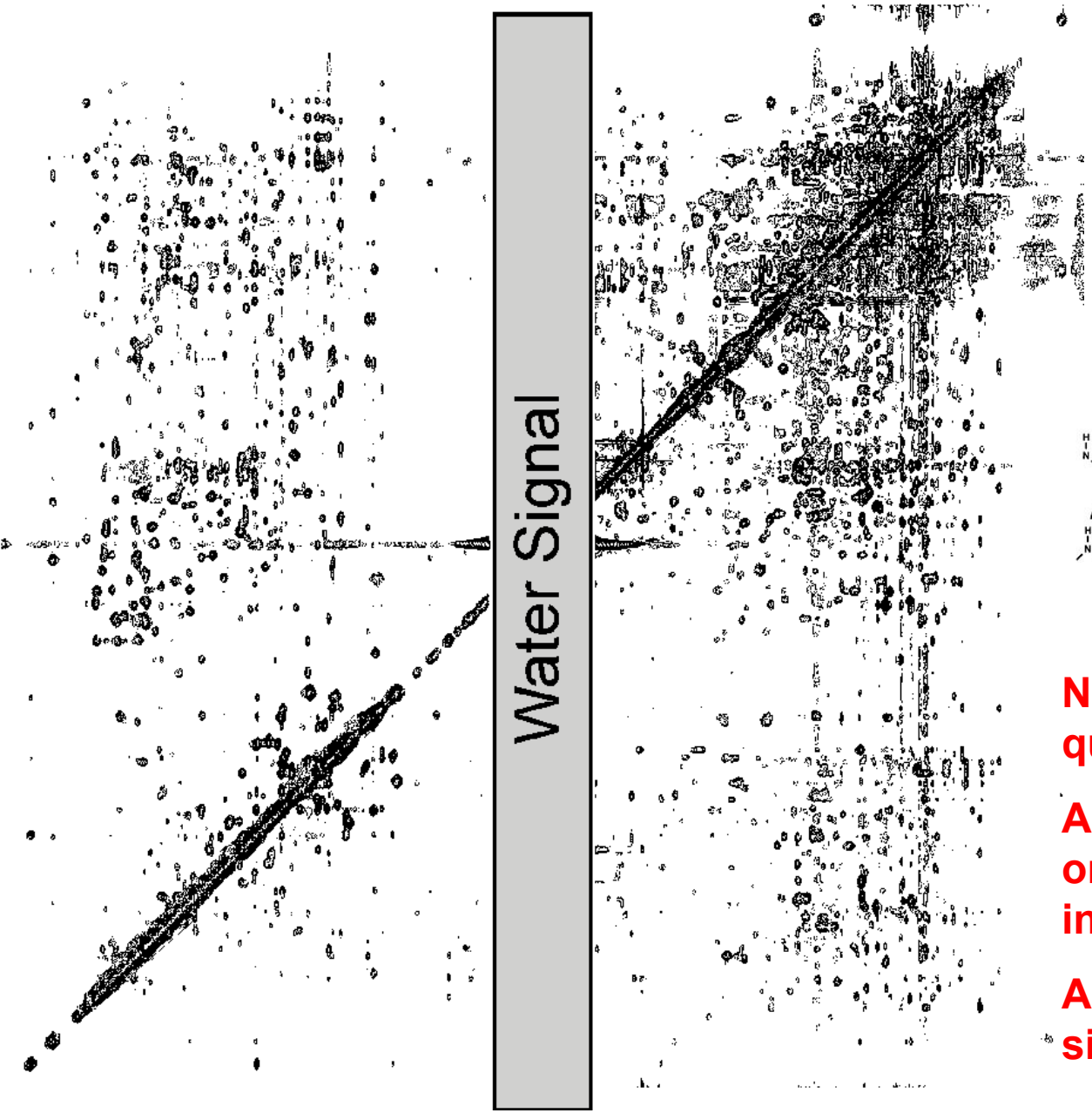


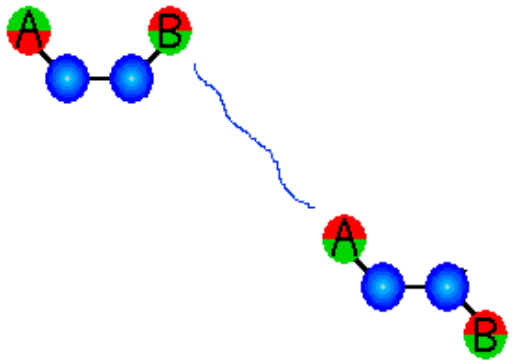
Identify many H-H short range NOE distances

Supplement with torsions from J-Coupling values

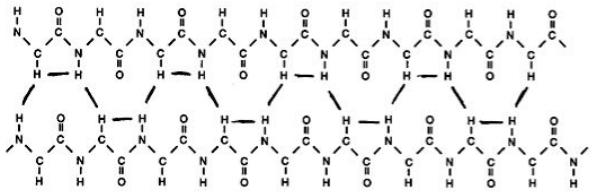
Assume standard peptide geometry

Use simulated annealing to find a structure which matches distances





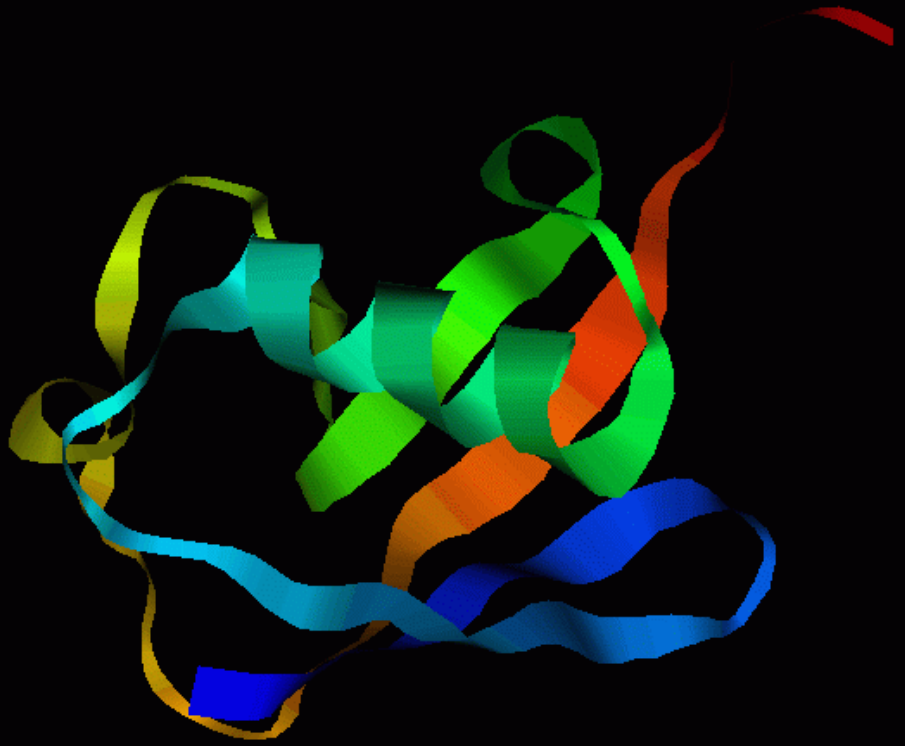
NOE cross peaks in parallel beta strands

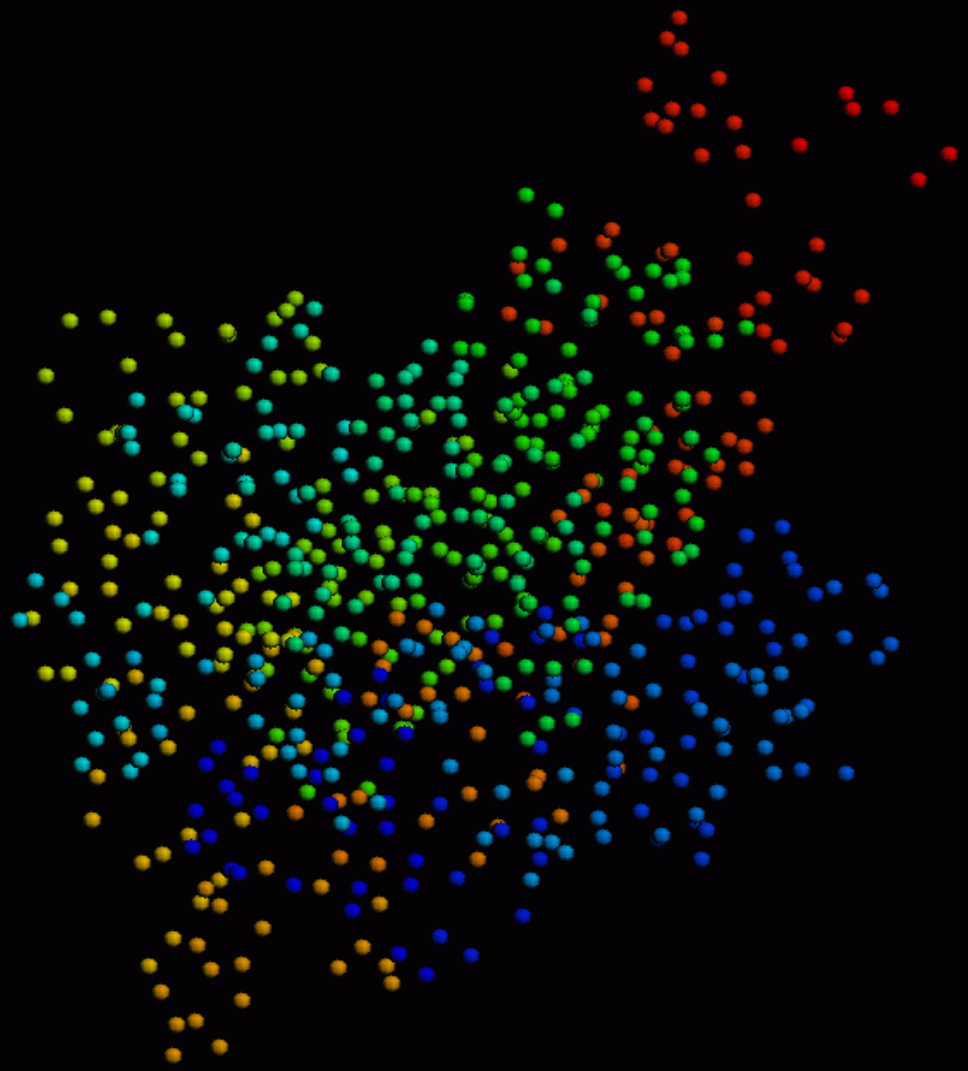


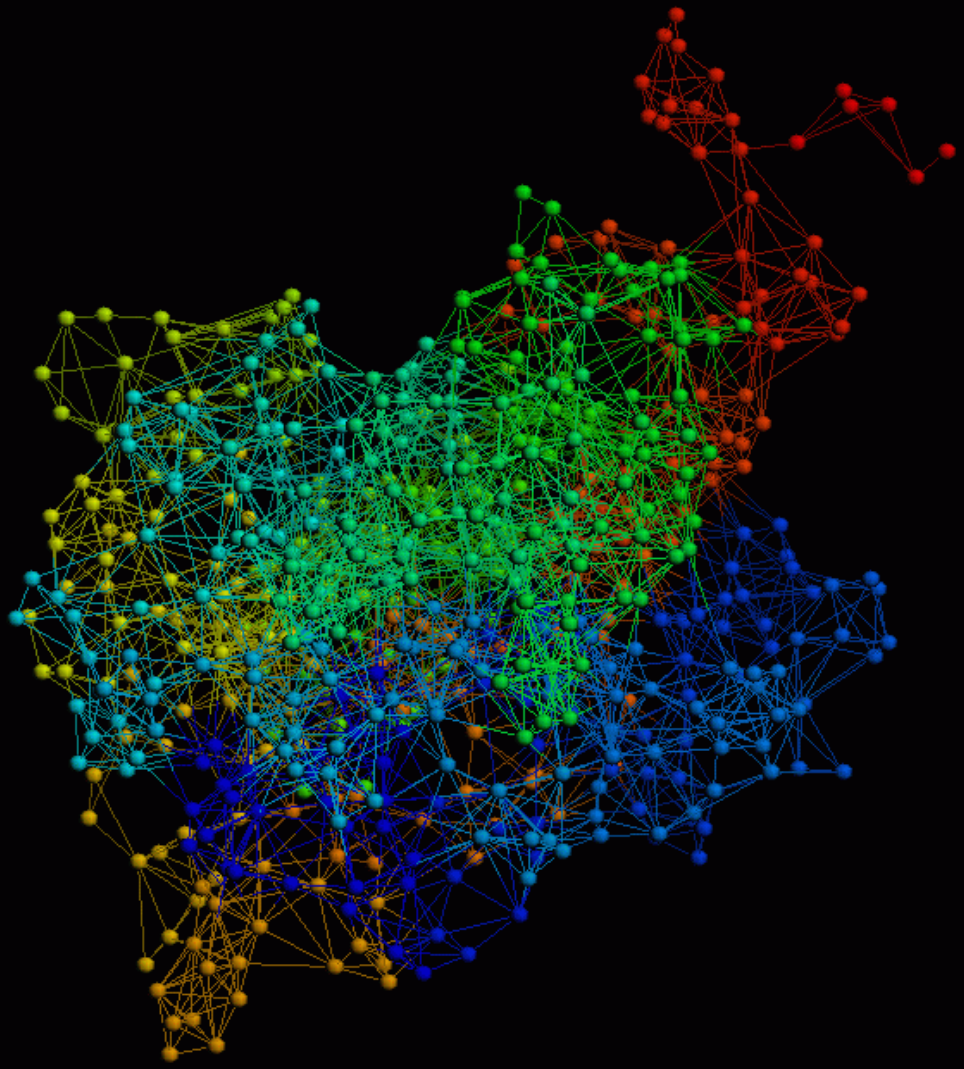
NOE distances are only qualitative

A given peak might be the only evidence of an interaction

A mis-assigned peak can be similarly problematic

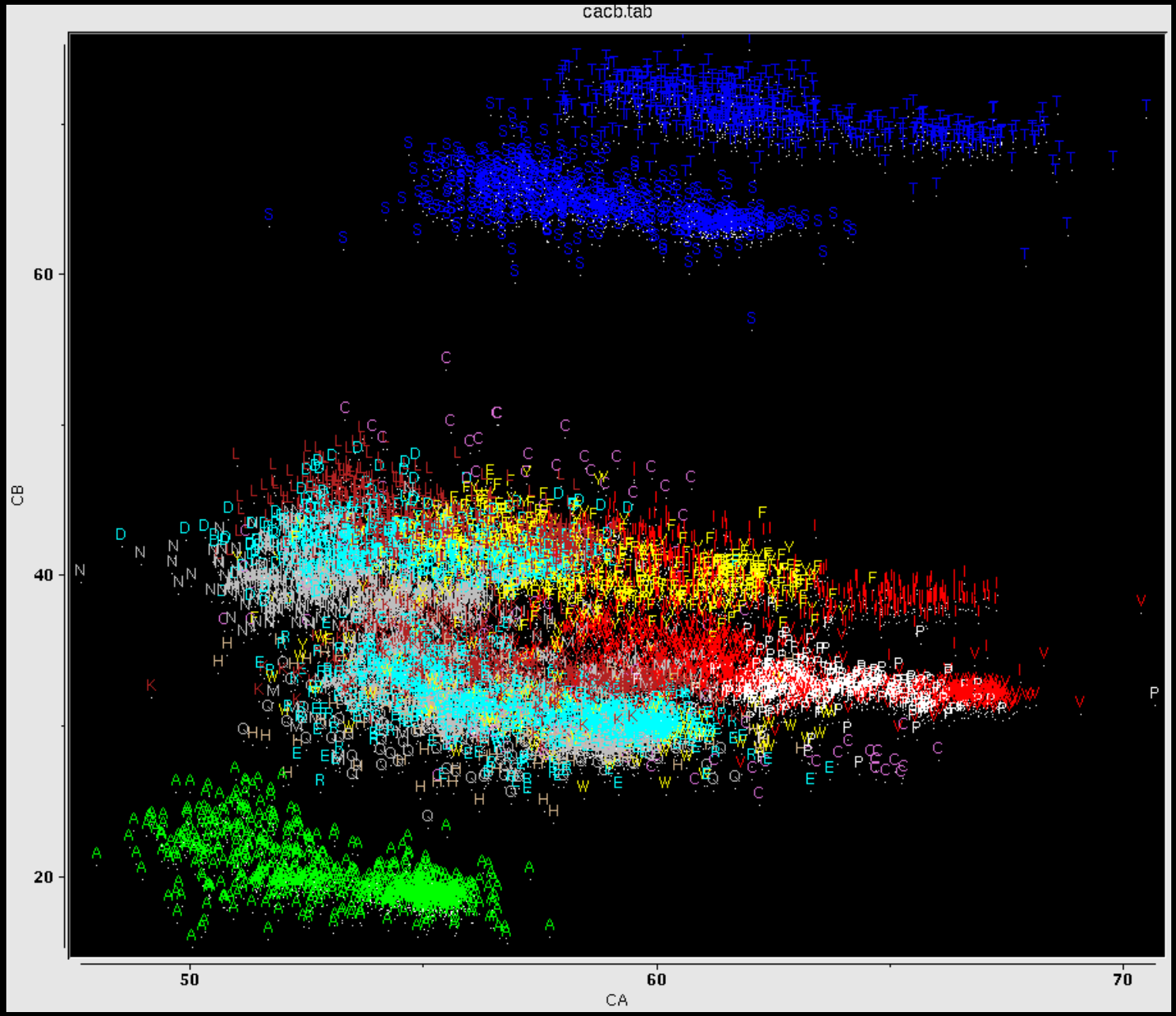






Alternate Approaches to NMR Structure

Chemical Shifts


cacb.tab
CB
CA
60
40
20
50
60
70

Residue Type Probability from Chemical Shifts

Use chemical shifts directly rather than secondary shifts (SS).

For a given residue in the target:

For a given residue type in the database:

Compute chemical shift distance.

Use Gaussian to estimate a P value

Find avg P value for residue type

Normalize over all residue types

DINI SS Method D81 Residues

50	14	4	2	3	(8)
61.7%	17.3%	4.9%	2.4%	3.7%	9.9%

DINI Direct Method 81 Residues:

64	11	3	2	1	0
79.0%	13.6%	3.7%	2.5%	1.2%	0%

FBP 170 residues:

129	20	10	6	2	(3)
75.9%	11.8%	5.9%	3.5%	1.2%	(1.8%)

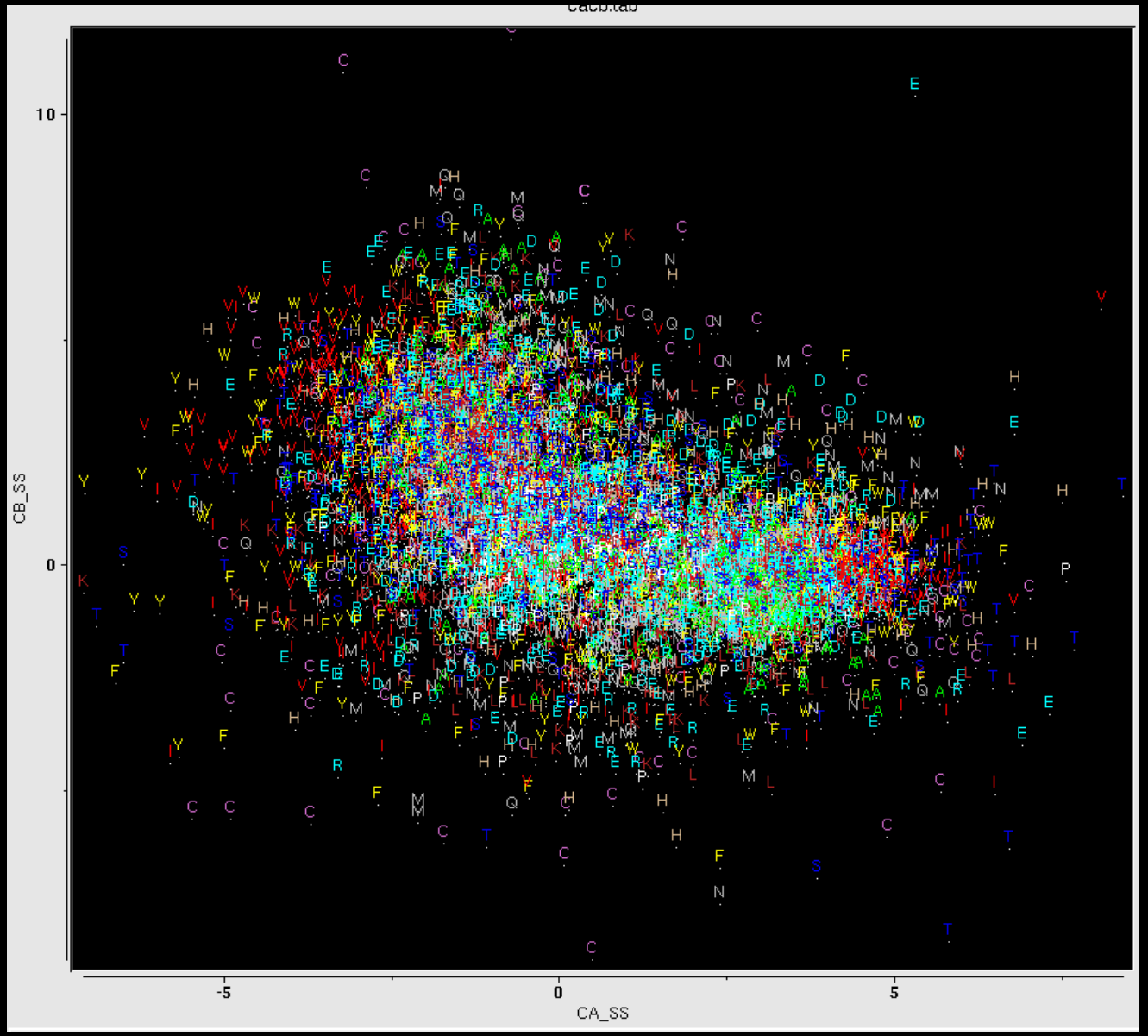
V-Alpha 114 residues:

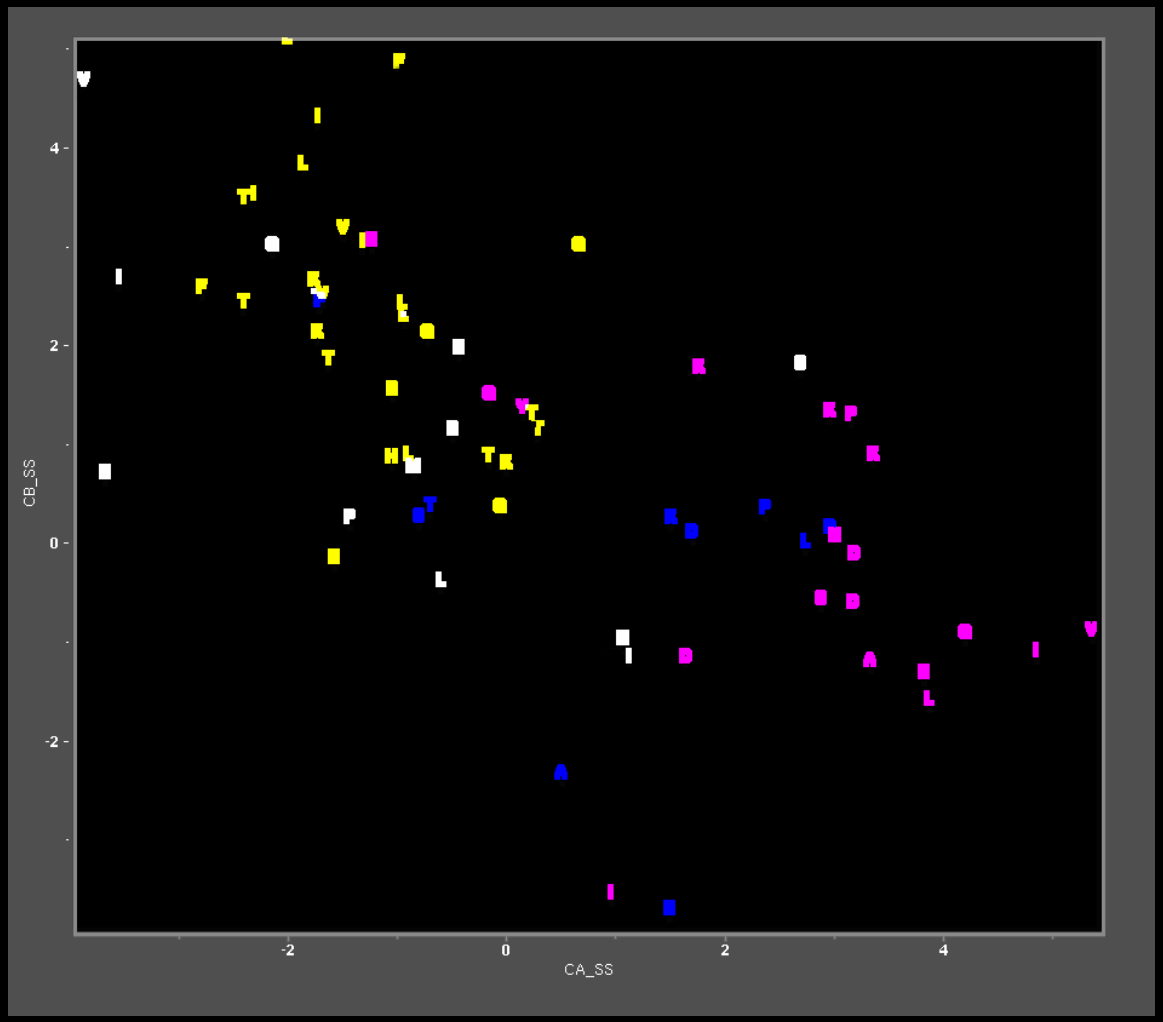
66	18	14	2	3	(11)
57.9%	15.8%	12.3%	1.8%	2.6%	(9.6%)

Gamma Crystalin 170 residues:

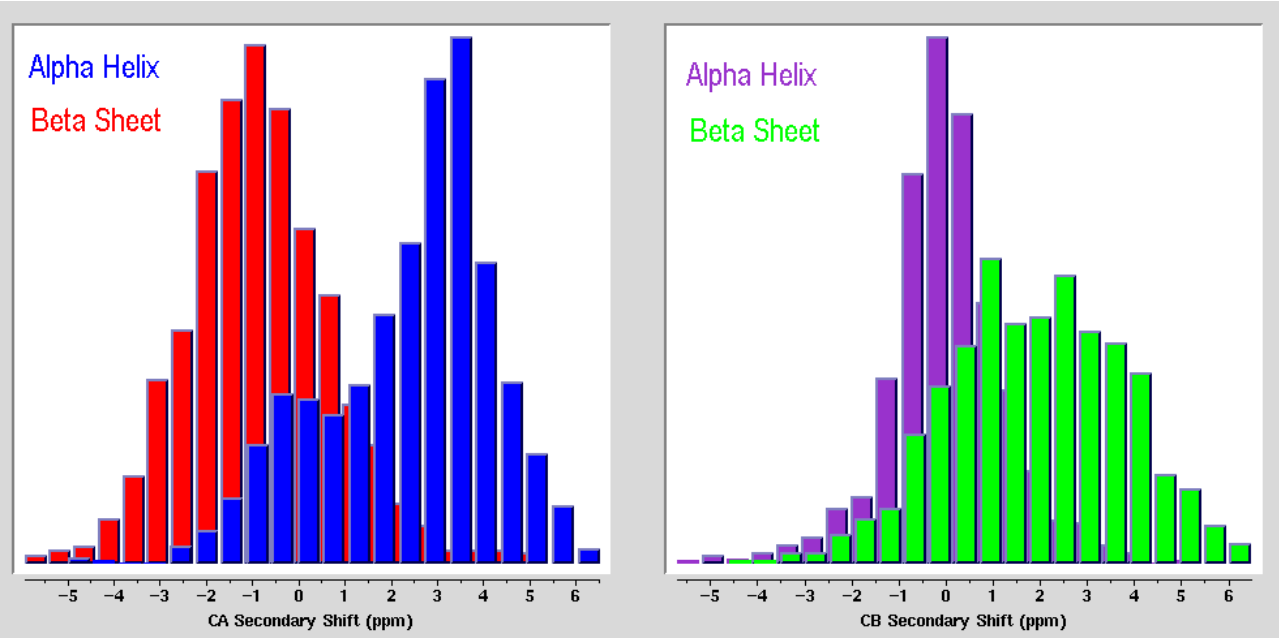
87	24	20	12	8	(19)
51.2%	14.1%	11.8%	7.1%	4.7%	11.2%

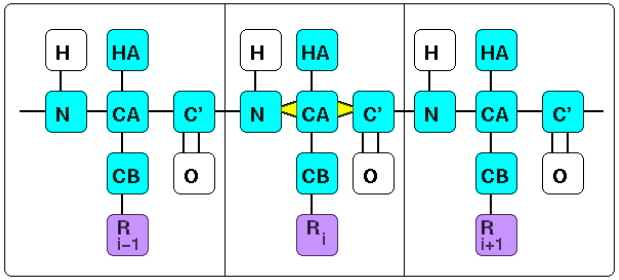
Subtract Residue-Specific *Random Coil Shift* to form *Secondary Shift*




CB_SS
CA_SS
4
2
0
-2
-2
0
2
4

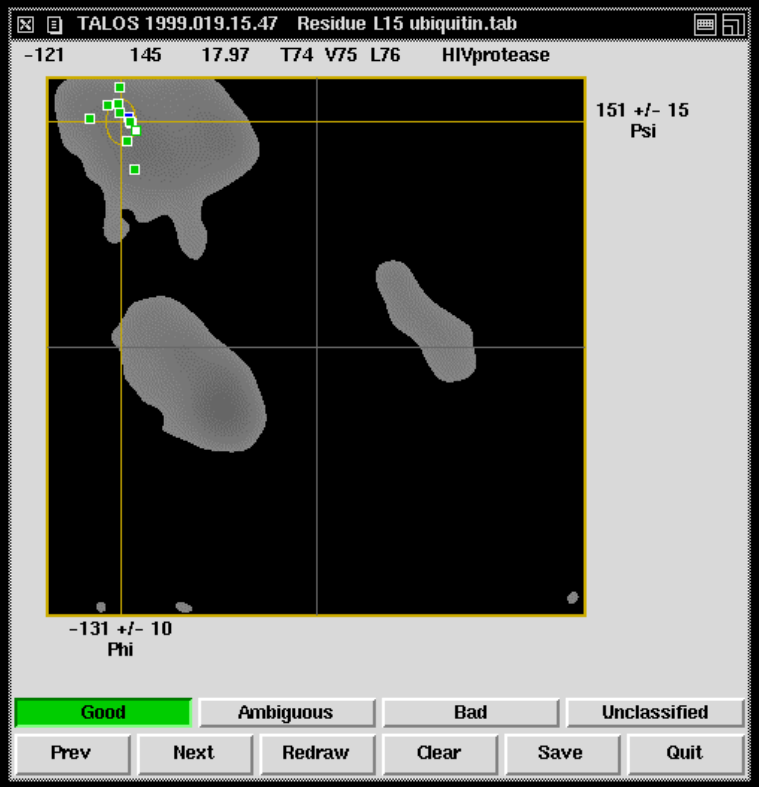
Chemical Shift and Backbone Structure Motif



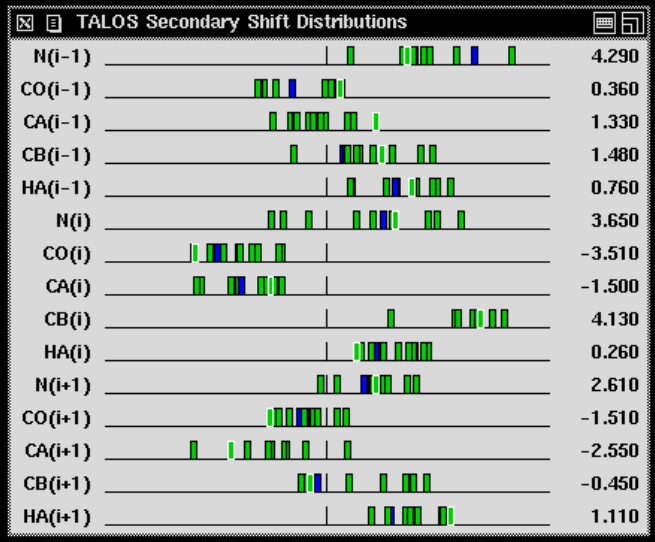


Match database triplet with target, based on sum-of squares difference in chemical shifts, plus residue type homology term.

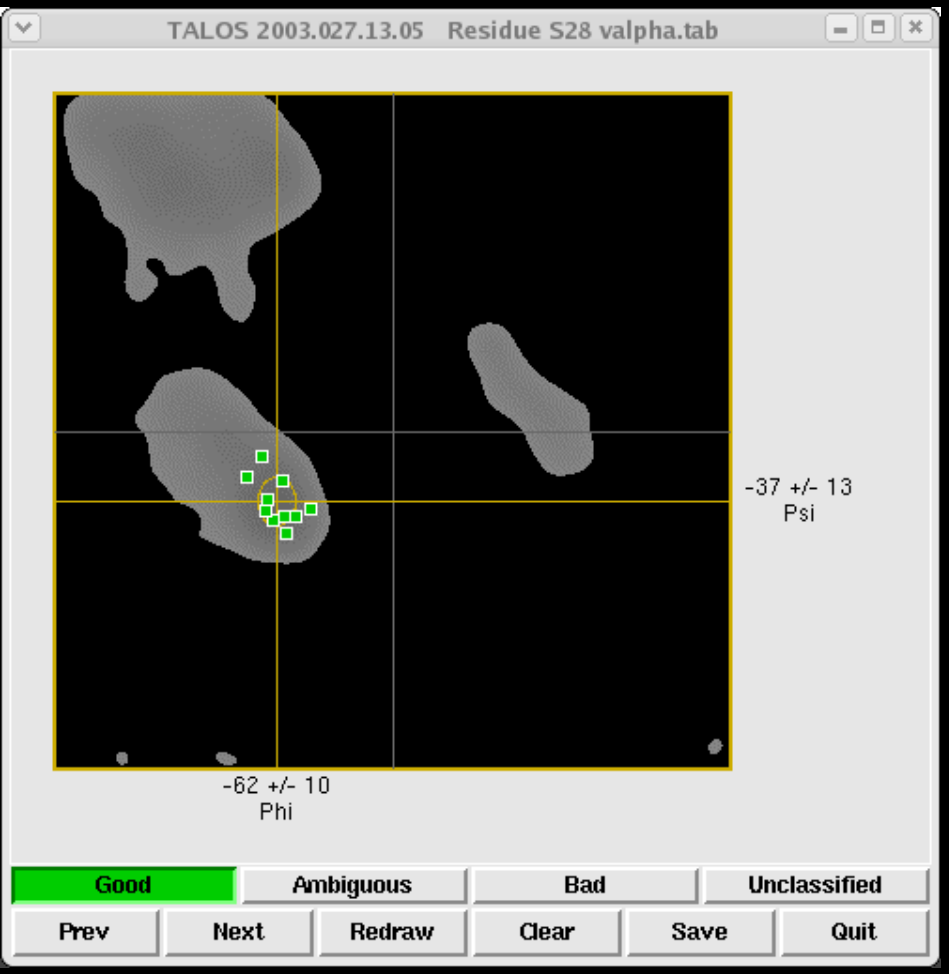
Use central residue as predictor of phi and psi.

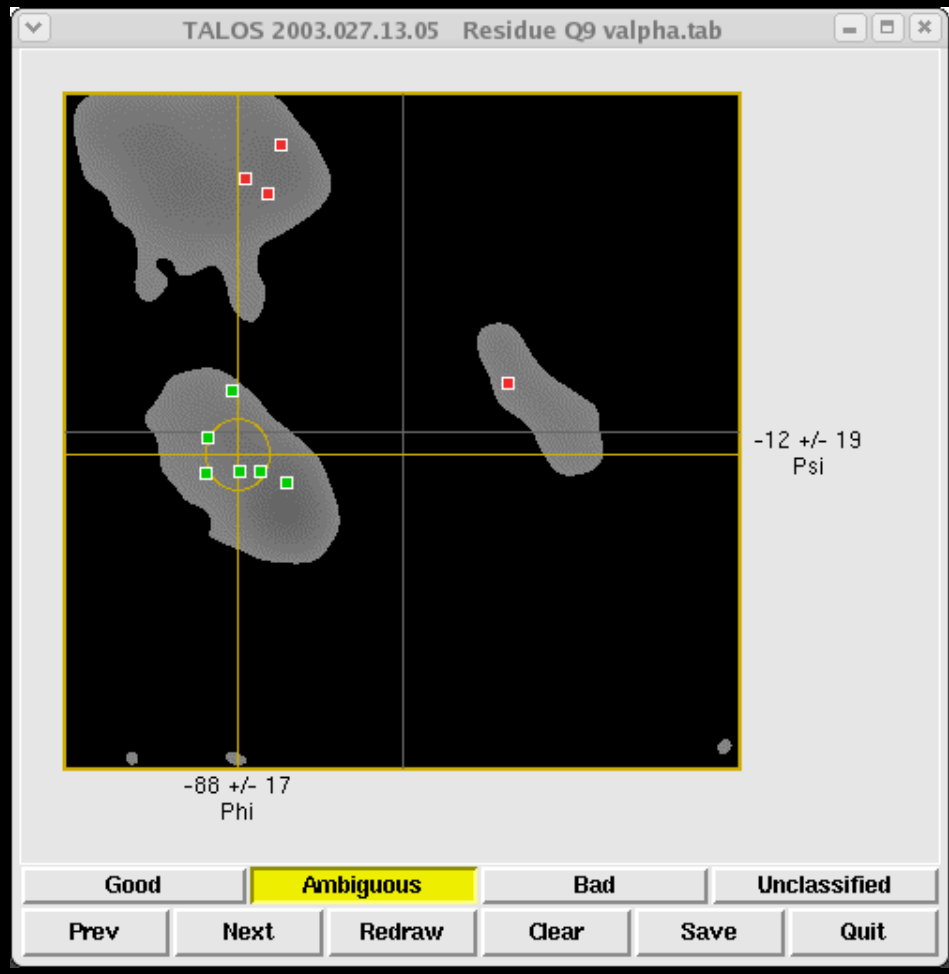

TALOS 1999.019.15.47 Residue L15 ubiquitin.tab
-121 145 17.97 T74 V75 L76 HIVprotease
151 +/- 15
Psi
-131 +/- 10
Phi
Good
Ambiguous
Bad
Unclassified
Prev
Next
Redraw
Clear
Save
Quit


TALOS ubiquitin.tab 76 Residues
M1 Q2 I3 F4 V5 K6 T7 L8 T9 G10
K11 T12 I13 T14 L15 E16 V17 E18 P19 S20
D21 T22 I23 E24 N25 V26 K27 A28 K29 I30
Q31 D32 K33 E34 G35 I36 P37 P38 D39 Q40
Q41 R42 L43 I44 F45 A46 G47 K48 Q49 L50
E51 D52 G53 R54 T55 L56 S57 D58 Y59 N60
I61 Q62 K63 E64 S65 T66 L67 H68 L69 V70
L71 R72 L73 R74 G75 G76

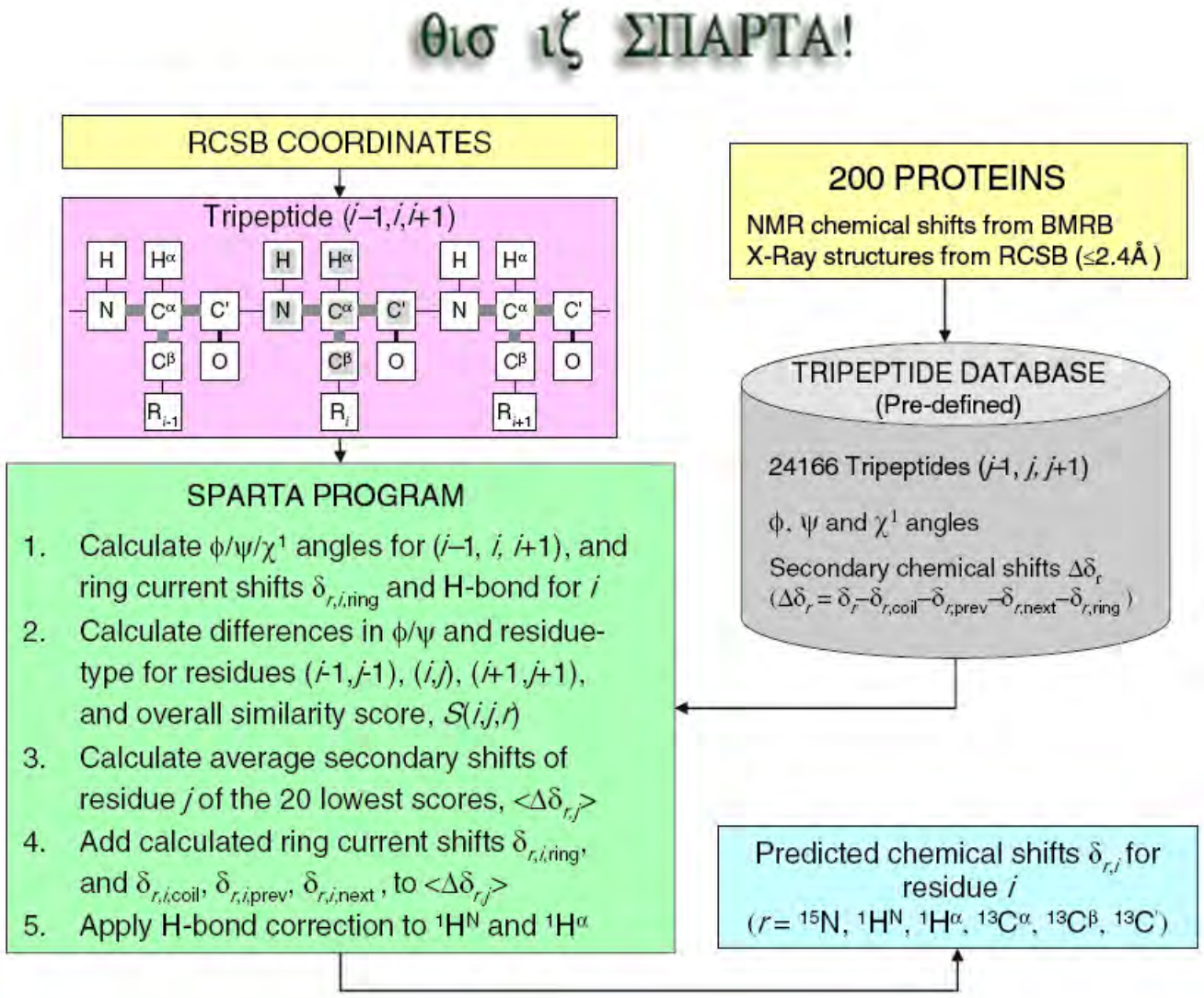

TALOS Secondary Shift Distributions
N(i-1) 4.290
CO(i-1) 0.360
CA(i-1) 1.330
CB(i-1) 1.480
HA(i-1) 0.760
N(i) 3.650
CO(i) -3.510
CA(i) -1.500
CB(i) 4.130
HA(i) 0.260
N(i+1) 2.610
CO(i+1) -1.510
CA(i+1) -2.550
CB(i+1) -0.450
HA(i+1) 1.110

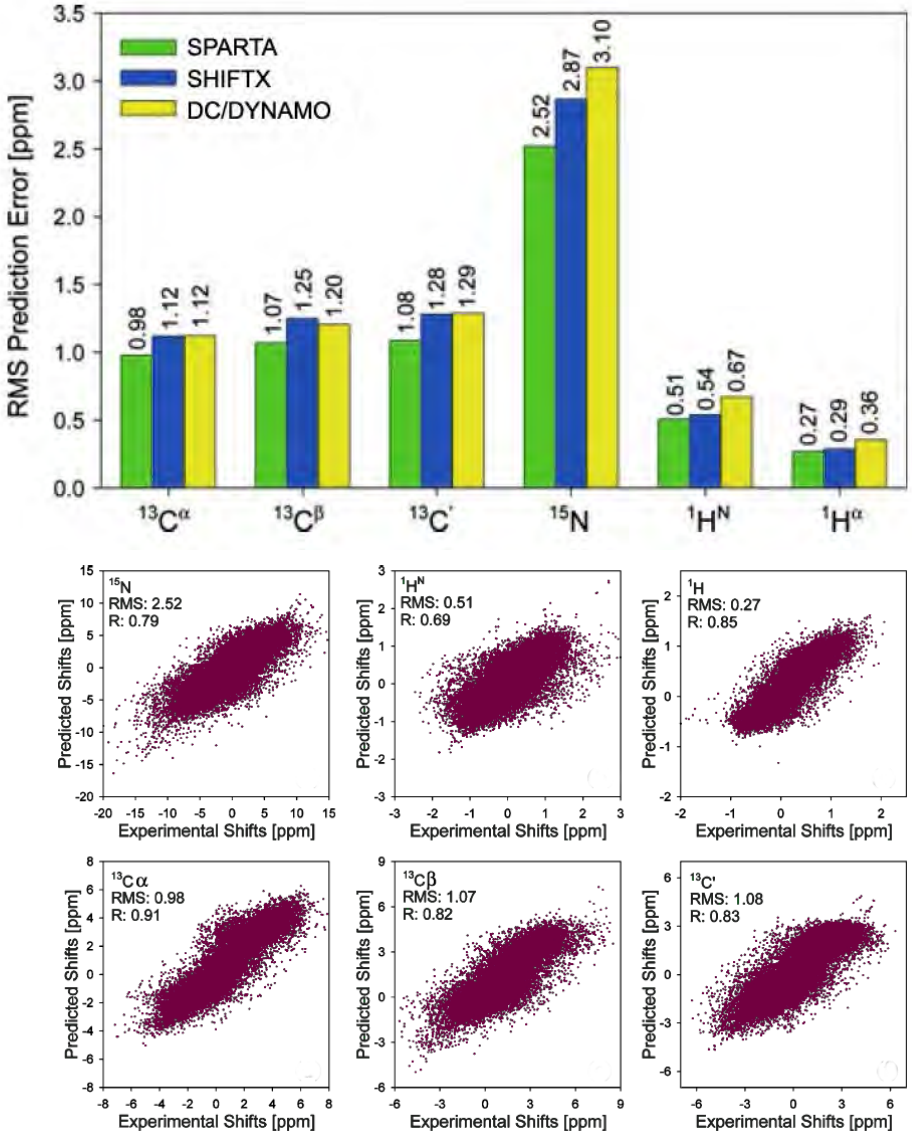

Residue L15, Triplet T14 L15 E16
-127 138 15.34 L63 I64 E65 HIVprotease
-132 157 17.05 K24 L25 M26 snase
-121 145 17.97 T74 V75 L76 HIVprotease
-140 162 18.55 L19 K20 E21 HIVprotease
-125 151 19.54 V138 V139 I140 IIIglc
-152 153 23.97 K70 A71 I72 HIVprotease
-133 163 25.21 L10 V11 T12 HIVprotease
-124 150 25.80 R165 I166 K167 IIIglc
-122 119 25.90 K27 V28 A29 maxacal
-132 174 26.19 R112 V113 K114 IIIglc
-131 151 21.55 Average

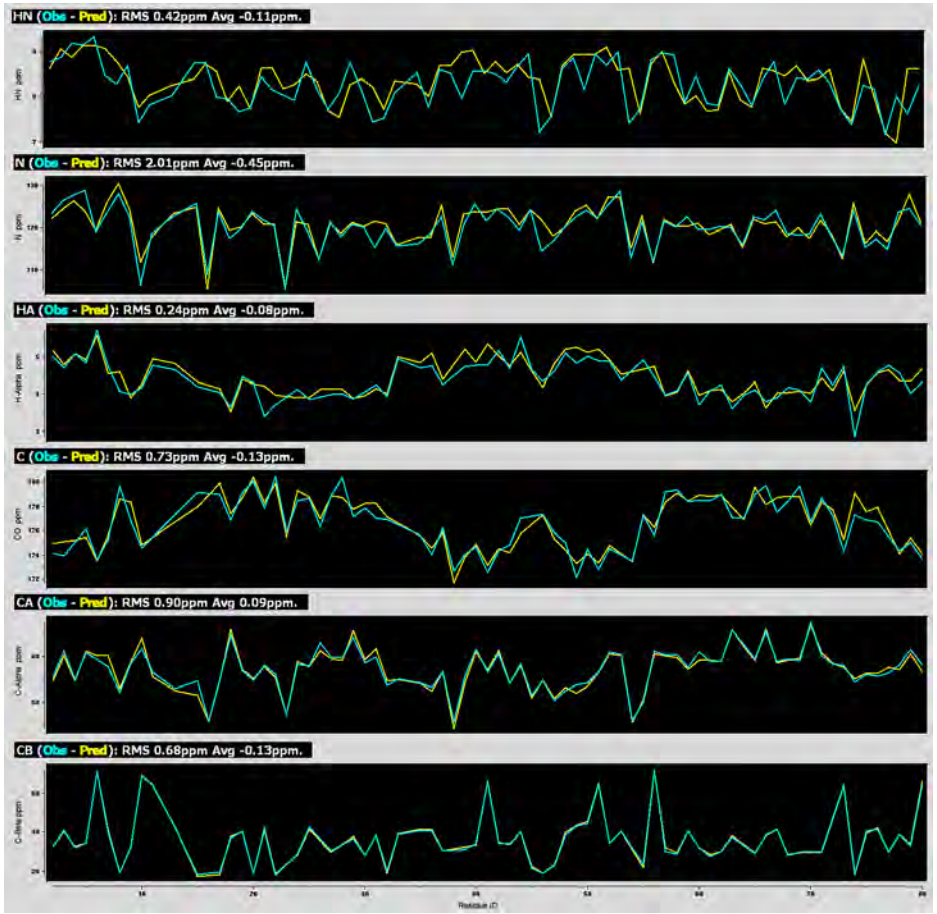

TALOS 2003.027.13.05 Residue S28 valpha.tab
-37 +/- 13
Psi
-62 +/- 10
Phi
Good
Ambiguous
Bad
Unclassified
Prev
Next
Redraw
Clear
Save
Quit


TALOS 2003.027.13.05 Residue Q9 valpha.tab
-12 +/- 19
Psi
-88 +/- 17
Phi
Good
Ambiguous
Bad
Unclassified
Prev
Next
Redraw
Clear
Save
Quit

The SPARTA Program of Shen and Bax …



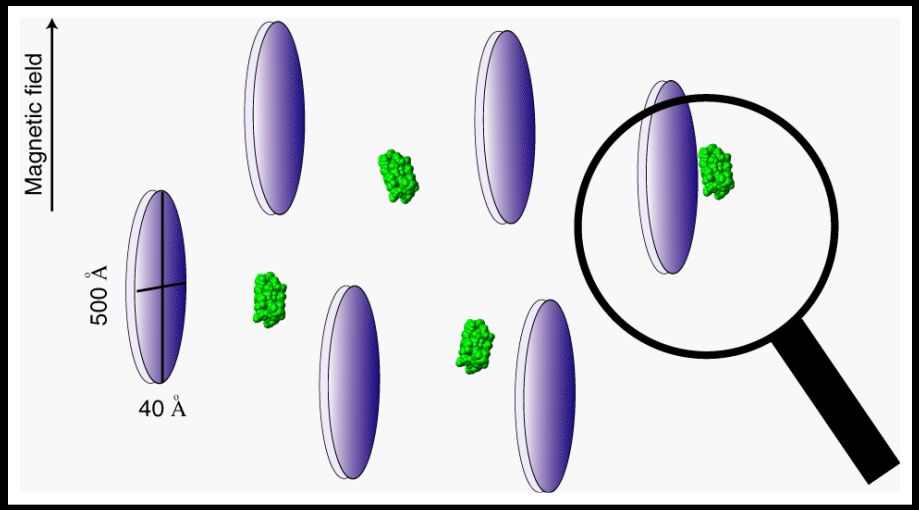

RMS Prediction Error [ppm]
SPARTA
SHIFTX
DC/DYNAMO
0.98
1.12
1.12
1.07
1.25
1.20
1.08
1.28
1.29
2.52
2.87
3.10
0.51
0.54
0.67
0.27
0.29
0.36
$^{13}C^{\alpha}$
$^{13}C^{\beta}$
$^{13}C'$
^{15}N
$^{1}H^{N}$
$^{1}H^{\alpha}$
^{15}N RMS: 2.52 R: 0.79
$^{1}H^{N}$ RMS: 0.51 R: 0.69
^{1}H RMS: 0.27 R: 0.85
$^{13}C\alpha$ RMS: 0.98 R: 0.91
$^{13}C\beta$ RMS: 1.07 R: 0.82
$^{13}C'$ RMS: 1.08 R: 0.83
Predicted Shifts [ppm]
Experimental Shifts [ppm]

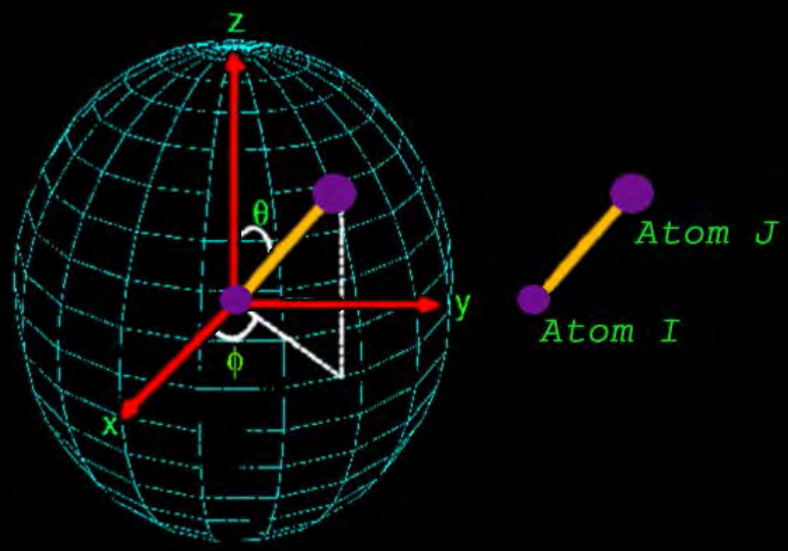

HN (Obs - Pred): RMS 0.42ppm Avg -0.11ppm.
N (Obs - Pred): RMS 2.01ppm Avg -0.45ppm.
HA (Obs - Pred): RMS 0.24ppm Avg -0.08ppm.
C (Obs - Pred): RMS 0.73ppm Avg -0.13ppm.
CA (Obs - Pred): RMS 0.90ppm Avg 0.09ppm.
CB (Obs - Pred): RMS 0.68ppm Avg -0.13ppm.

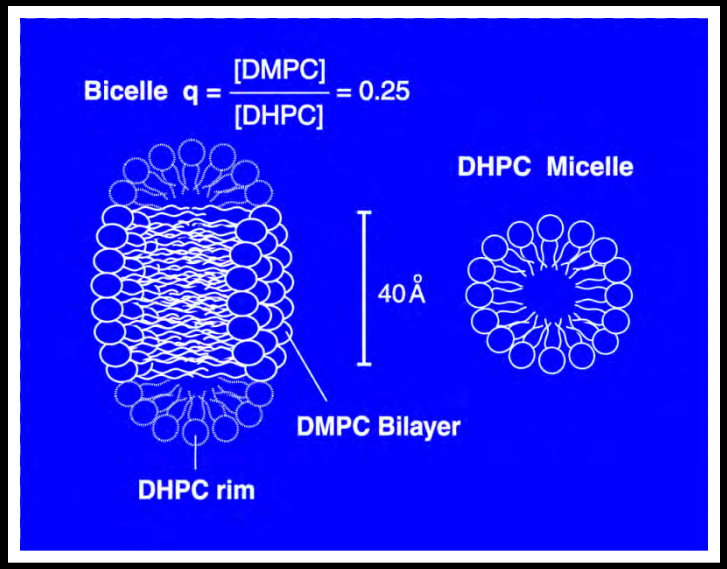
Alternate Approaches to NMR Structure

Residual Dipolar Couplings

Alignment by Liquid Crystal



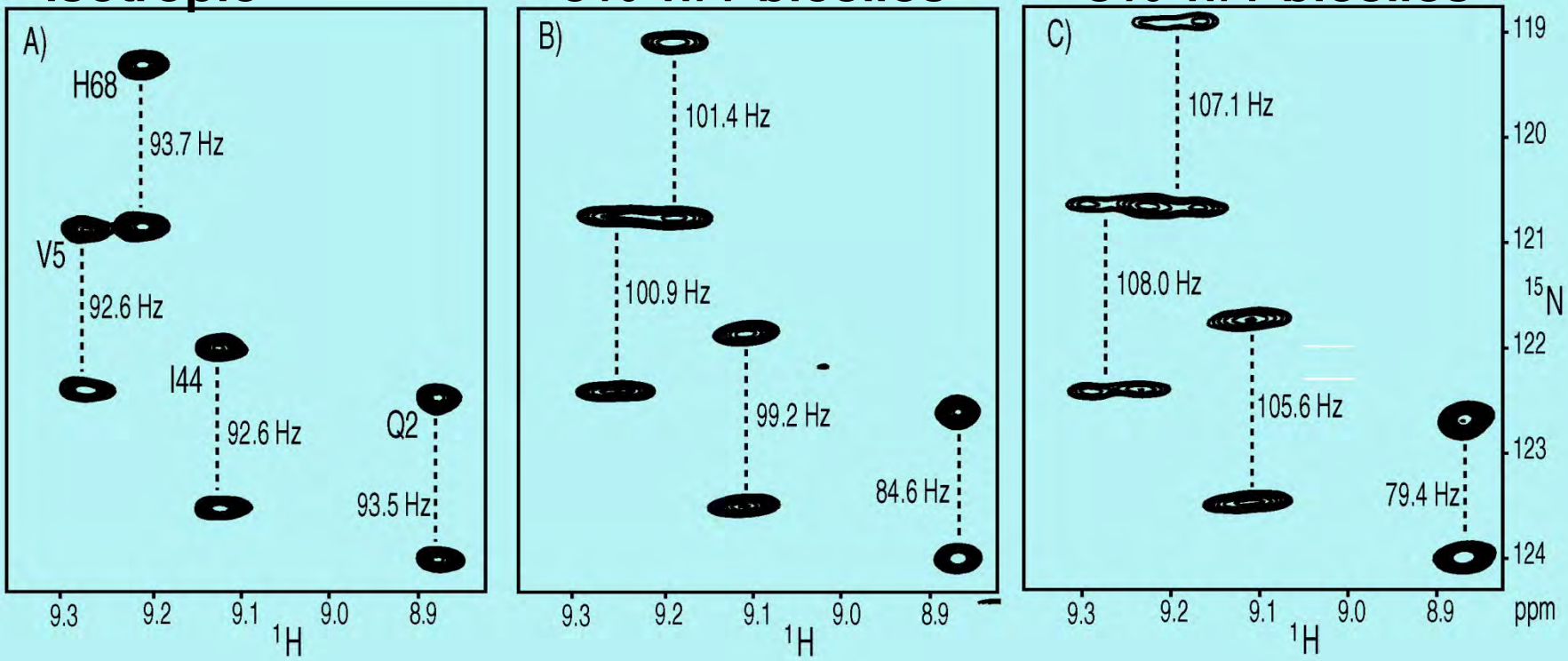




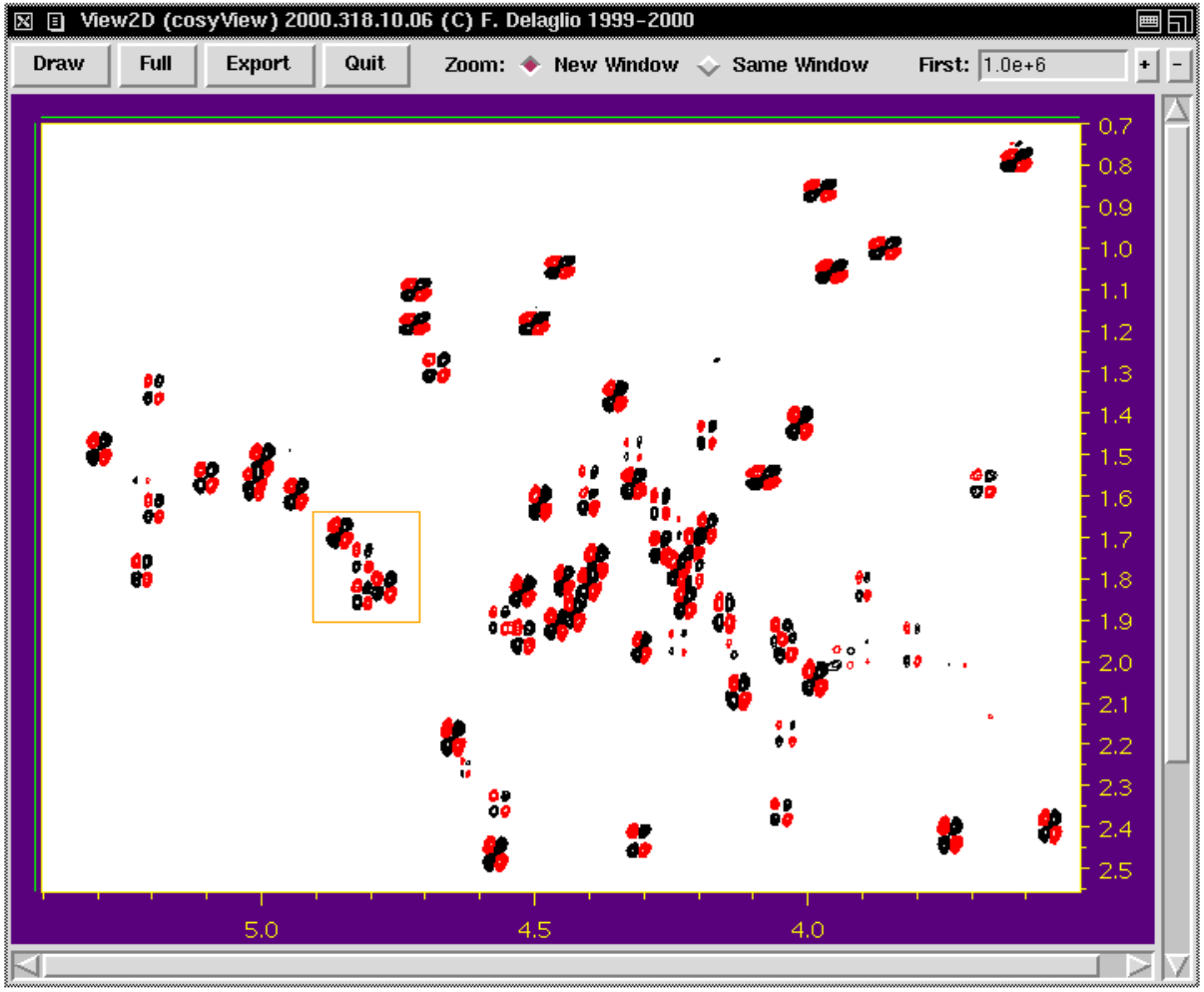
$$D_{ij} = D_{\max}\left[D_a(3\cos^2\theta - 1) + \frac{3}{2} D_r \sin^2\theta \cos 2\phi \right]$$

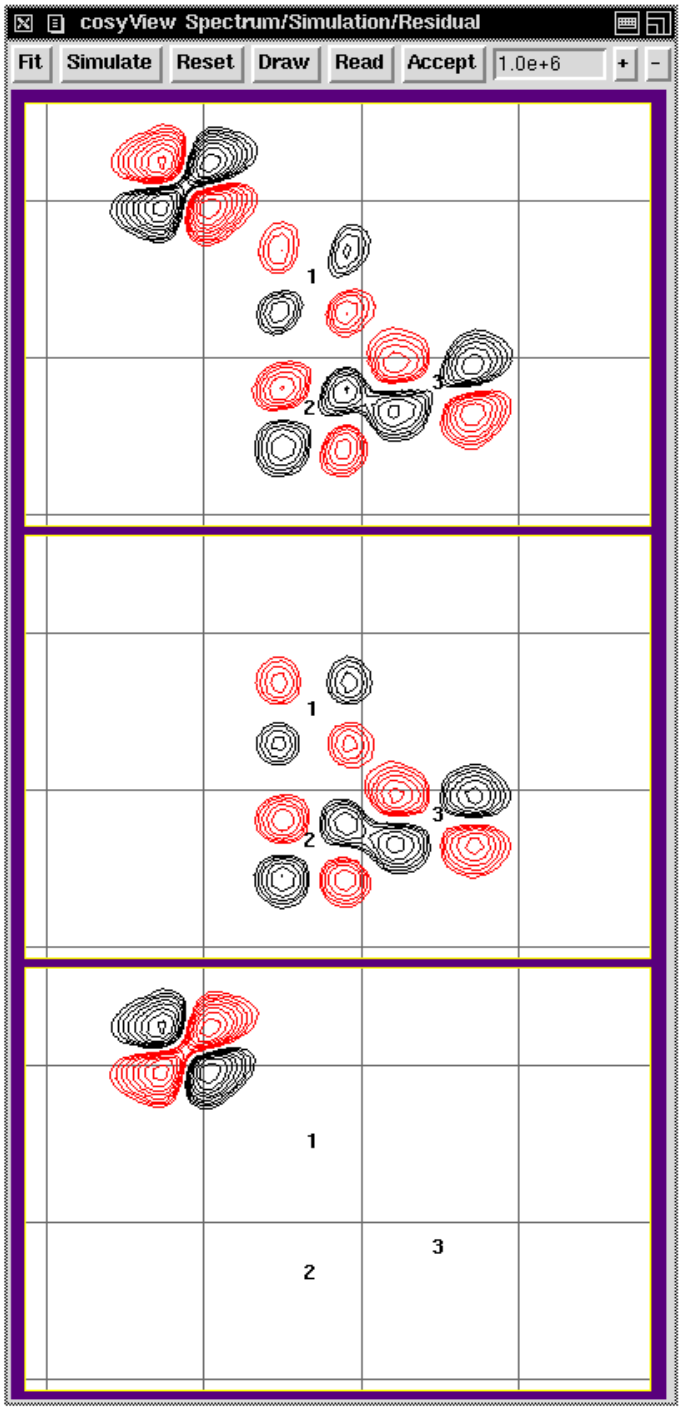
$$D_{max} = -\mu_0 (h/2\pi)\, \gamma_i \gamma_j / (4\pi r_{2r} 3_{ij})$$

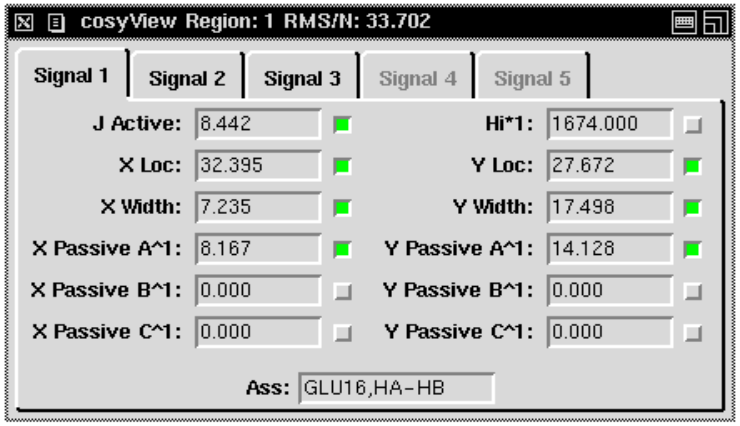
^{1}H-^{15}N HSQC spectra of ubiquitin

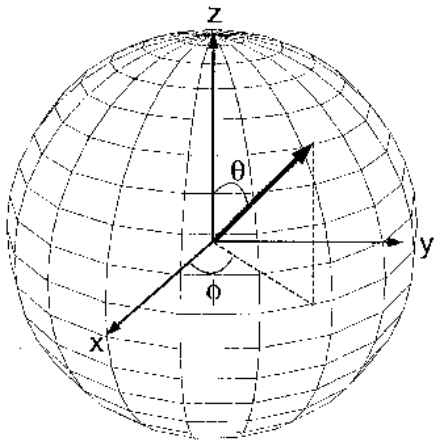


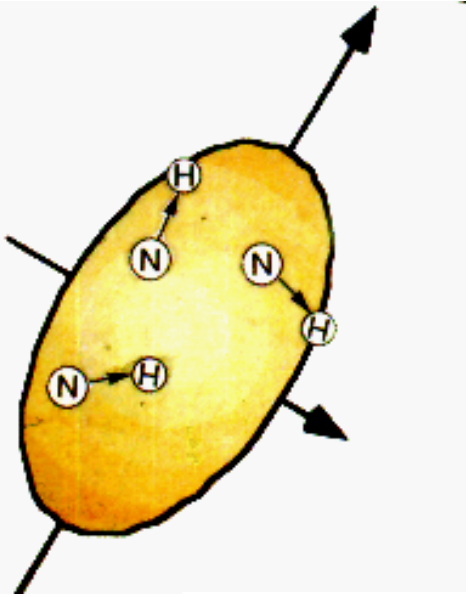
Alignment: 0 % 0.1% 0.16%

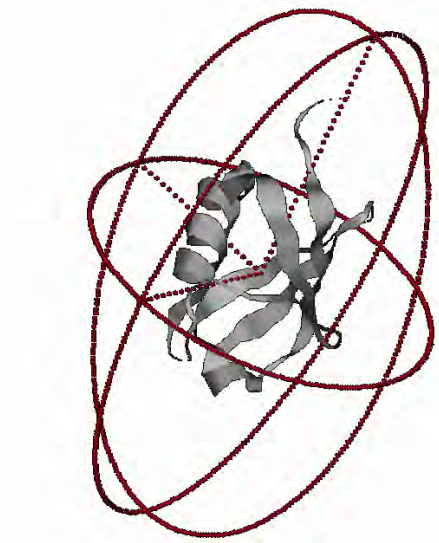

View2D (cosyView) 2000.318.10.06 (C) F. Delaglio 1999-2000
Draw
Full
Export
Quit
Zoom:
New Window
Same Window
First:
1.0e+6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
2.4
2.5
5.0
4.5
4.0


cosyView Spectrum/Simulation/Residual
Fit
Simulate
Reset
Draw
Read
Accept
1.0e+6
1
2
3
1
2
3
1
2
3

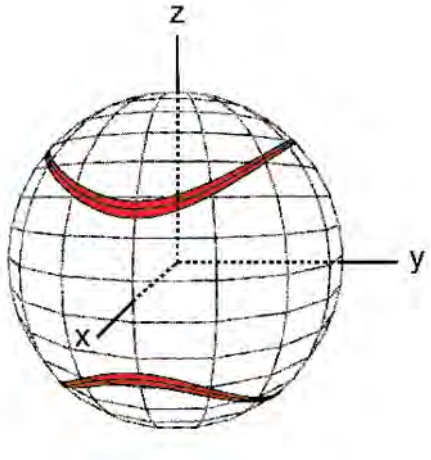

cosyView Region: 1 RMS/N: 33.702
Signal 1
Signal 2
Signal 3
Signal 4
Signal 5
J Active: 8.442
Hi*1: 1674.000
X Loc: 32.395
Y Loc: 27.672
X Width: 7.235
Y Width: 17.498
X Passive A^1: 8.167
Y Passive A^1: 14.128
X Passive B^1: 0.000
Y Passive B^1: 0.000
X Passive C^1: 0.000
Y Passive C^1: 0.000
Ass: GLU16,HA-HB

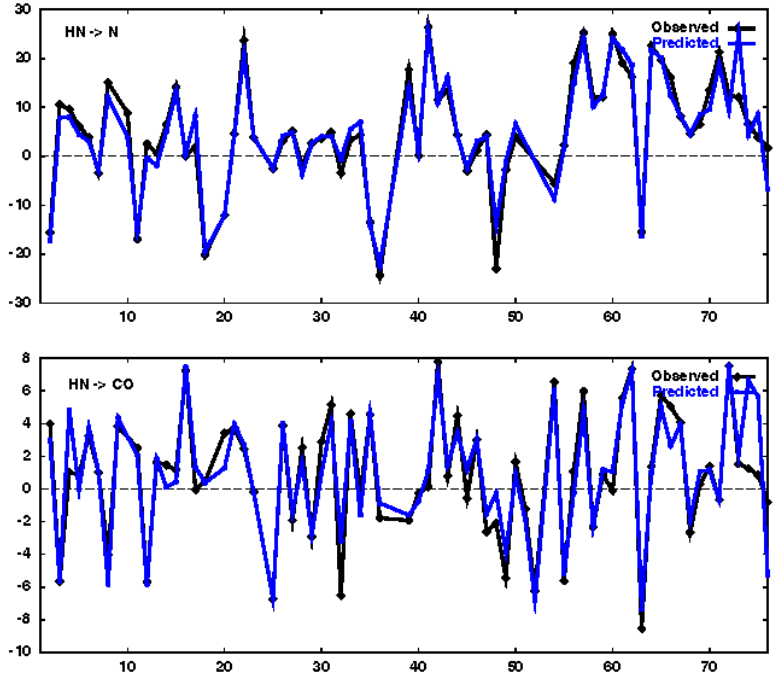






$$D_{ij} = D_{\max}\left[D_a(3\cos^2\theta - 1) + \frac{3}{2} D_r \sin^2\theta \cos 2\phi \right]$$





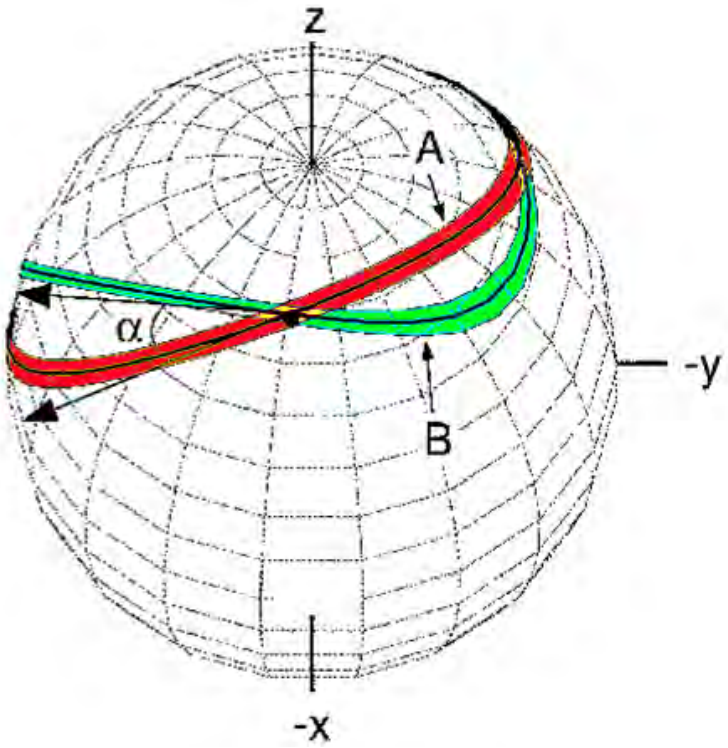
Dab = Dmax [(s1)*0.5*(3.0*zIJ*zIJ - 1.0)

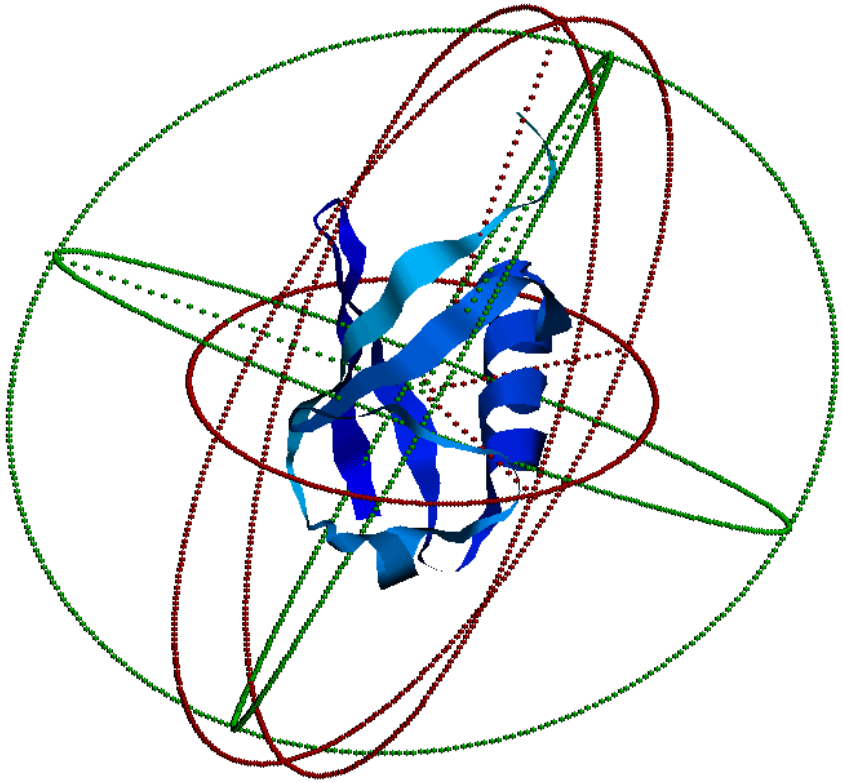
+ (s2)*0.5*(xIJ*xIJ - yIJ*yIJ)

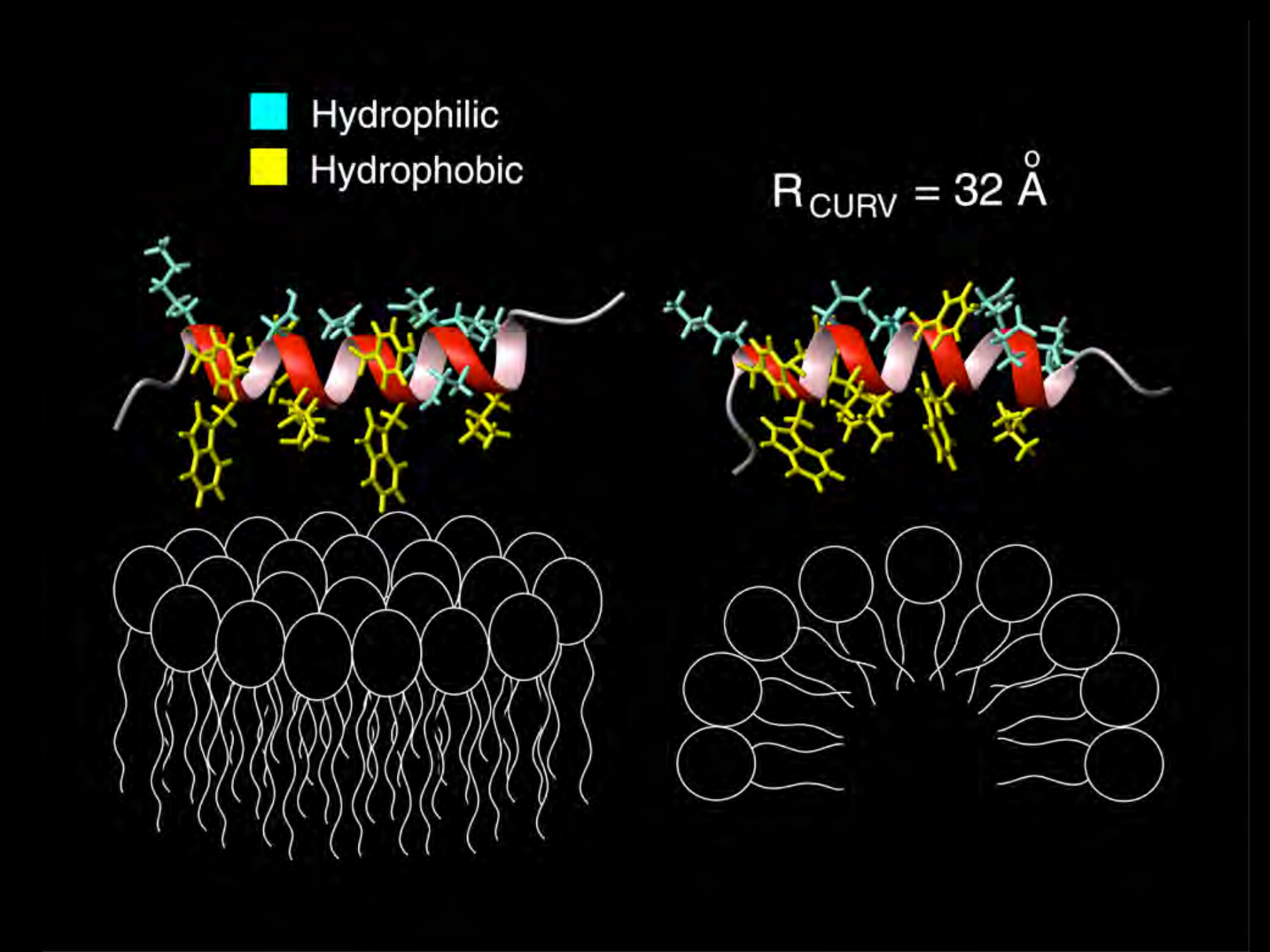
+ (s3)*2.0*xIJ*yIJ

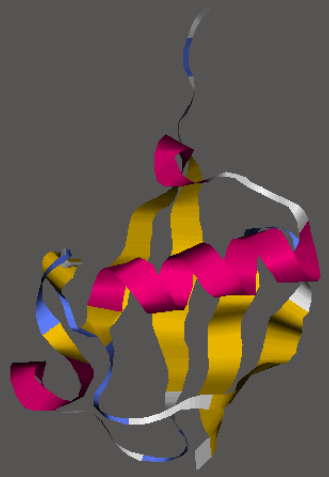
+ (s4)*2.0*xIJ*zIJ

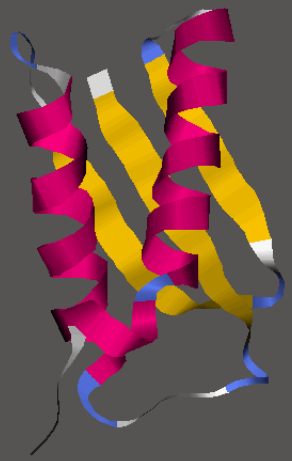
+ (s5)*2.0*yIJ*zIJ]

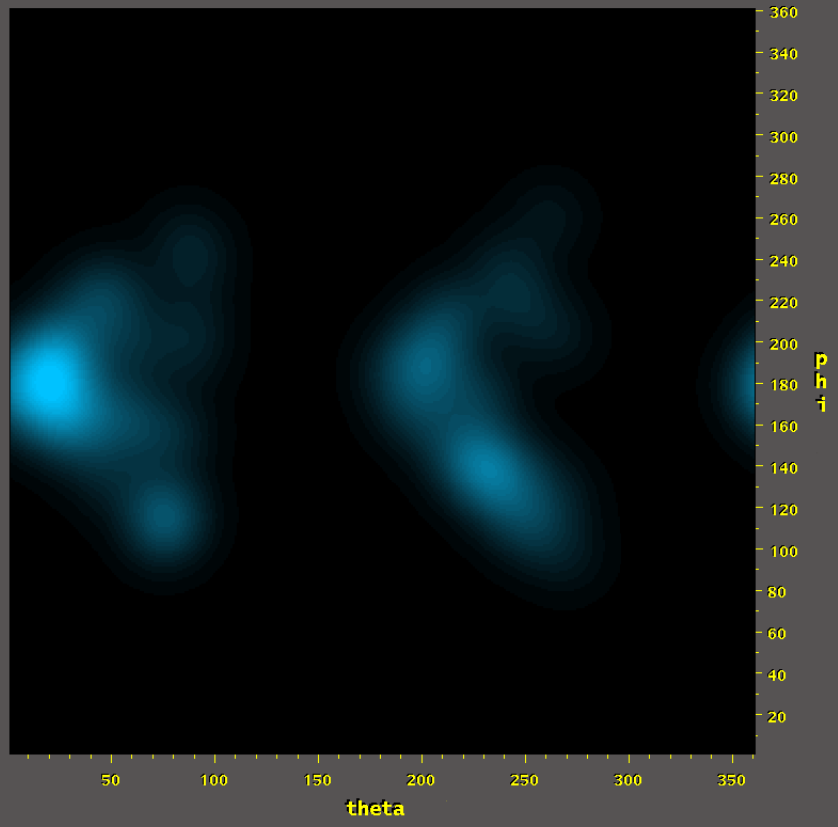

z
A
α
-y
B
-x

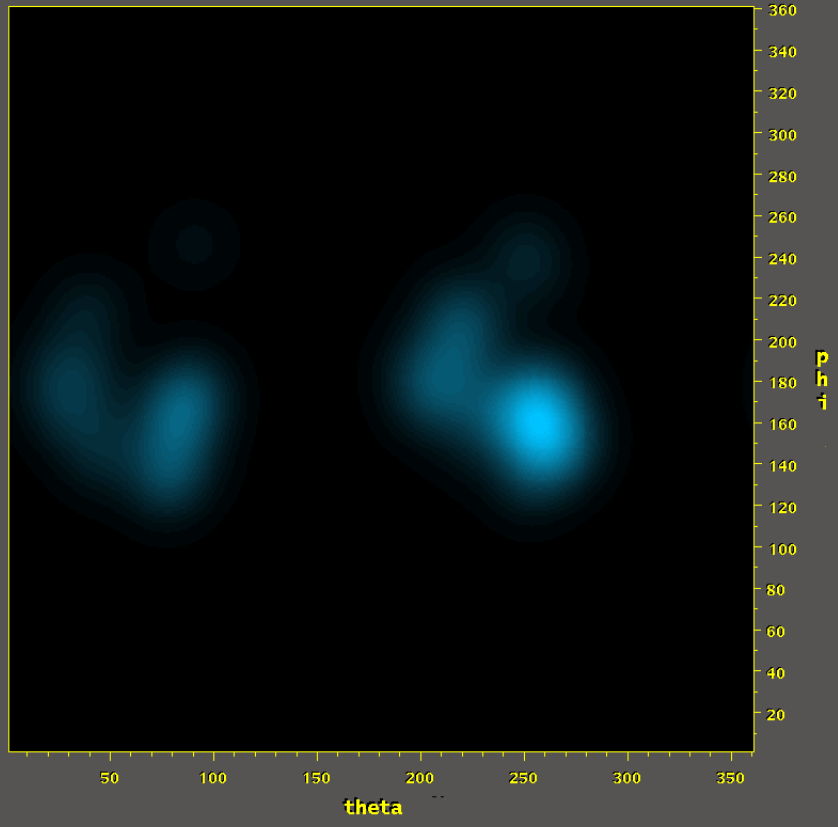


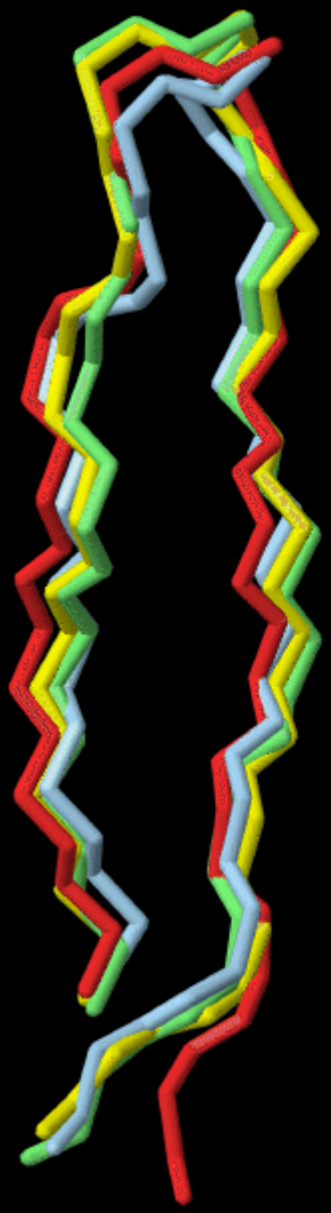

Hydrophilic
Hydrophobic
R_{CURV} = 32 Å






360
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
phi
50
100
150
200
250
300
350
theta


360
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
phi
50
100
150
200
250
300
350
theta



Molecular Fragment Replacement (MFR)

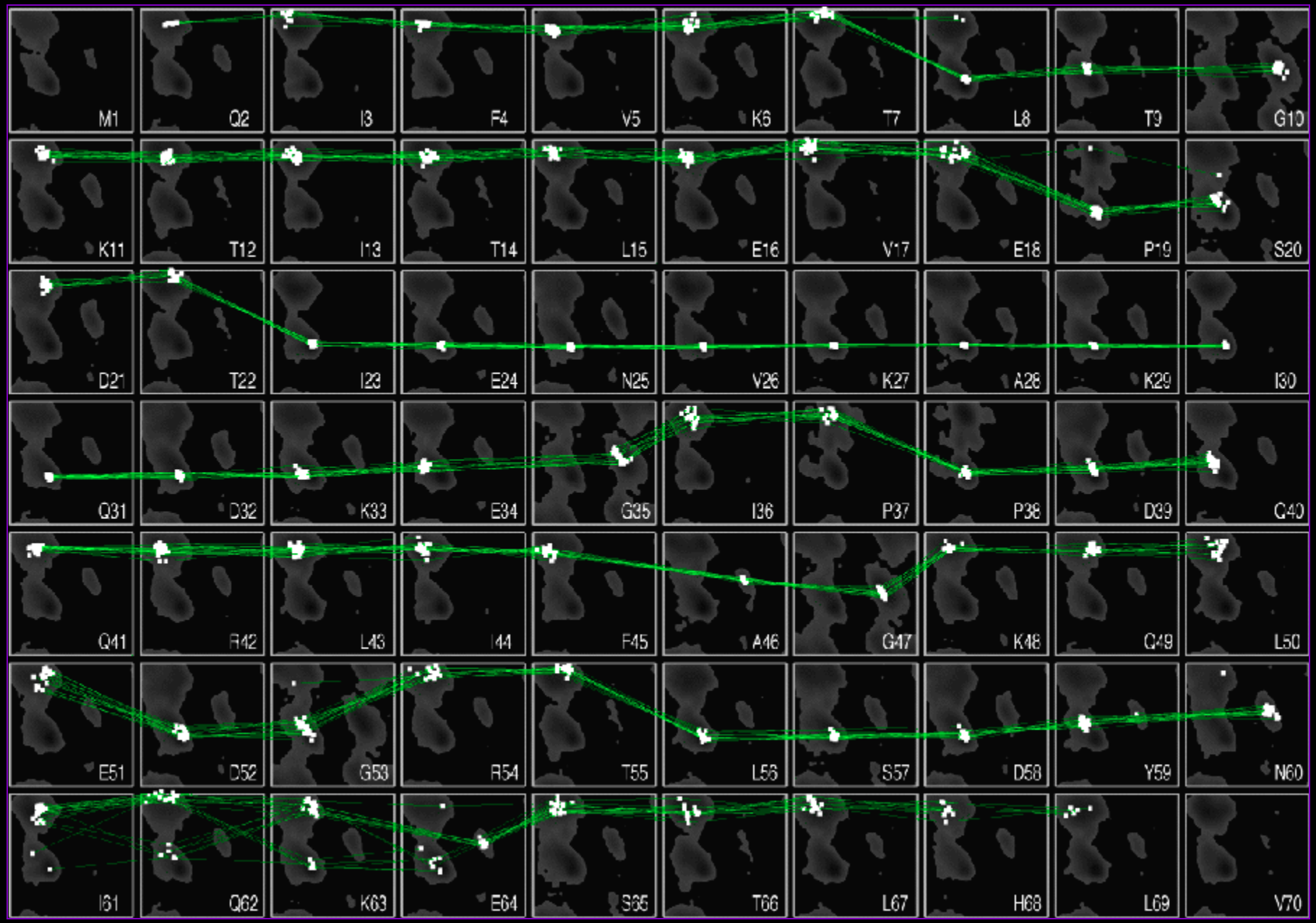
- Search PDB for small fragments whose simulated dipolar couplings and shifts match the observed values.
- Use the fragment information to reconstitute larger structural elements.
- Also: Sequential NOEs, J values, etc
- Nucleic Acid Applications

1ubq.pdb 2 - 17

1bii.pdb 189 - 204

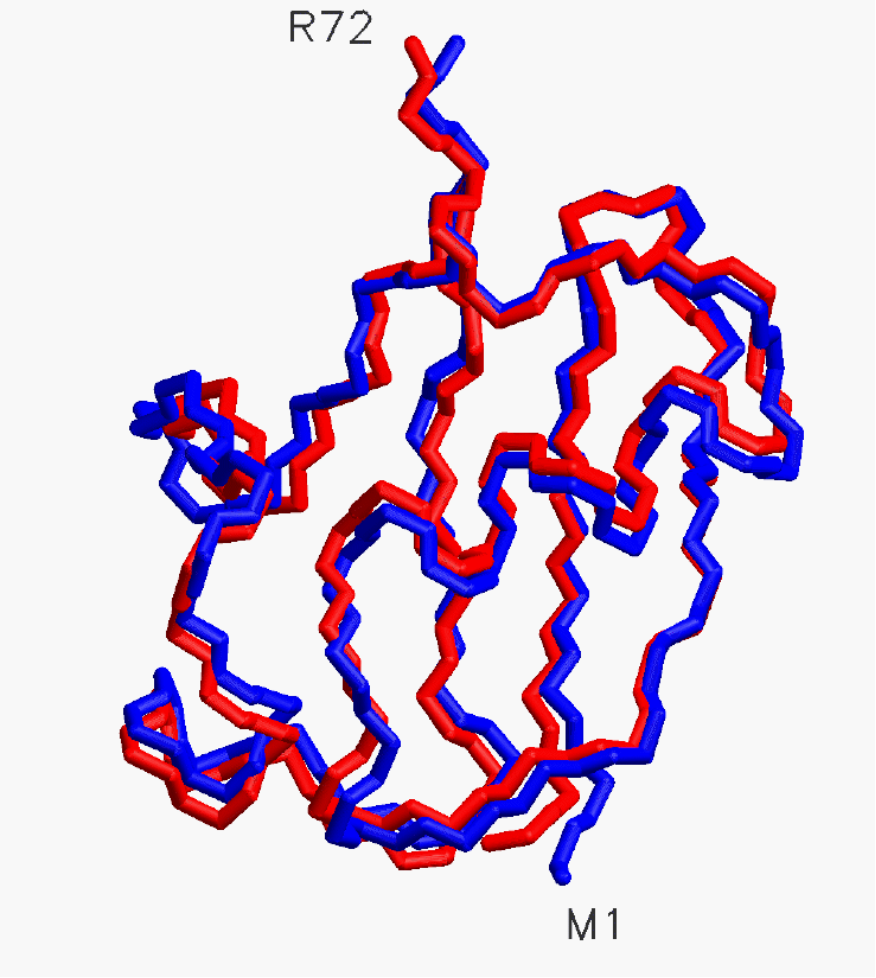
1cel.pdb 15 - 30

1gtm.pdb 40 - 55


M1 Q2 I3 F4 V5 K6 T7 L8 T9 G10
K11 T12 I13 T14 L15 E16 V17 E18 P19 S20
D21 T22 I23 E24 N25 V26 K27 A28 K29 I30
Q31 D32 K33 E34 G35 I36 P37 P38 D39 C40
Q41 R42 L43 I44 F45 A46 G47 K48 C49 L50
E51 D52 G53 R54 T55 L56 S57 D58 Y59 N60
I61 Q62 K63 E64 S65 T66 L67 H68 L69 V70

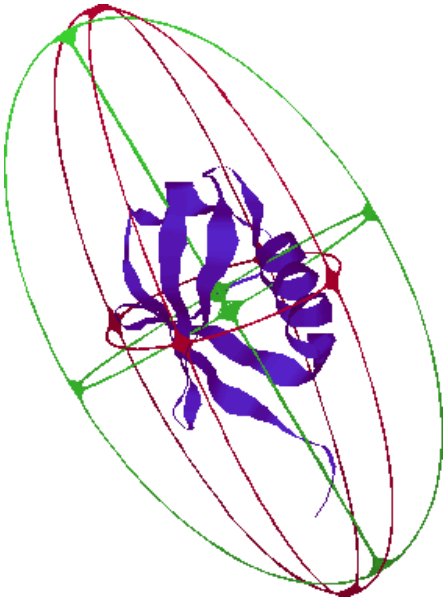
Initial Structure from Average Phi and Psi of Fragment Ensemble

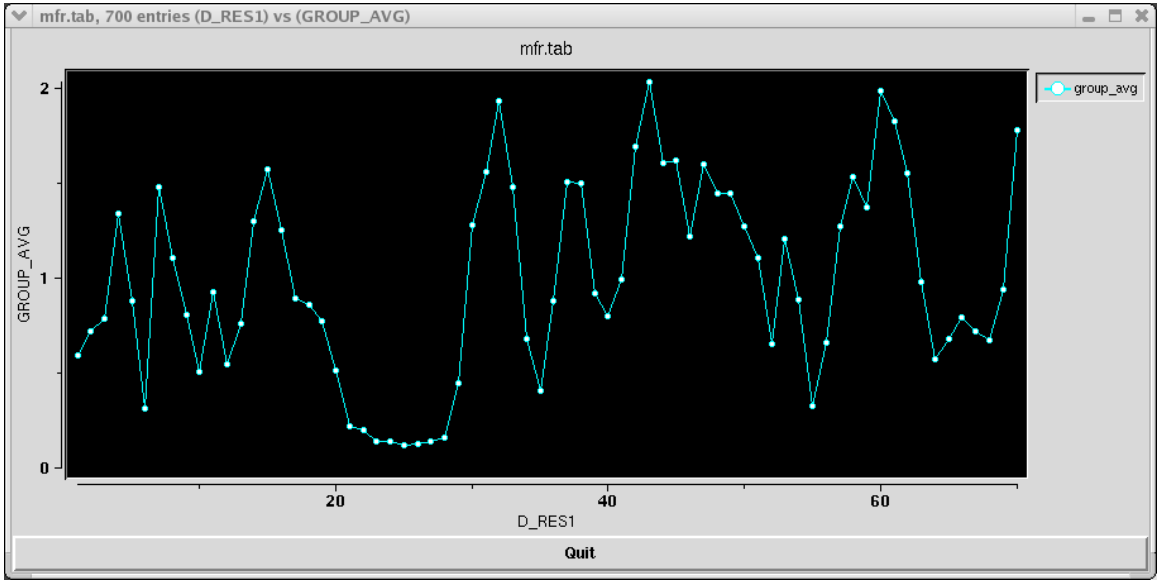
1ubq vs MFR phi/psi refined structure

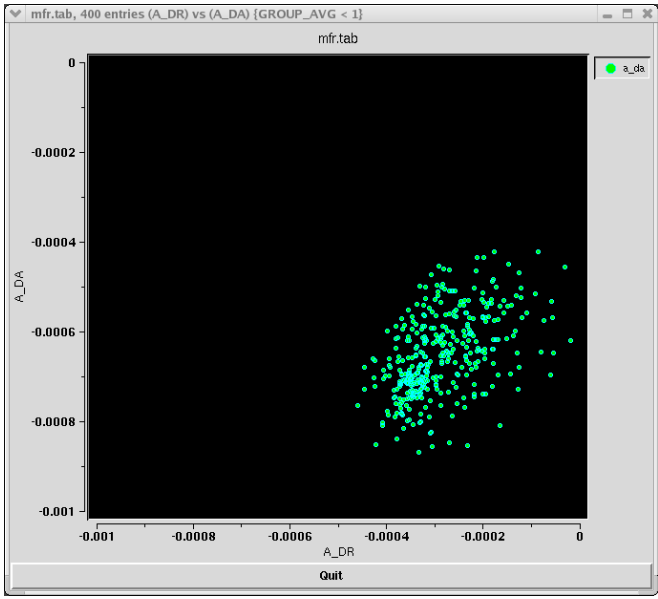


MFR Estimation of Tensor Parameters

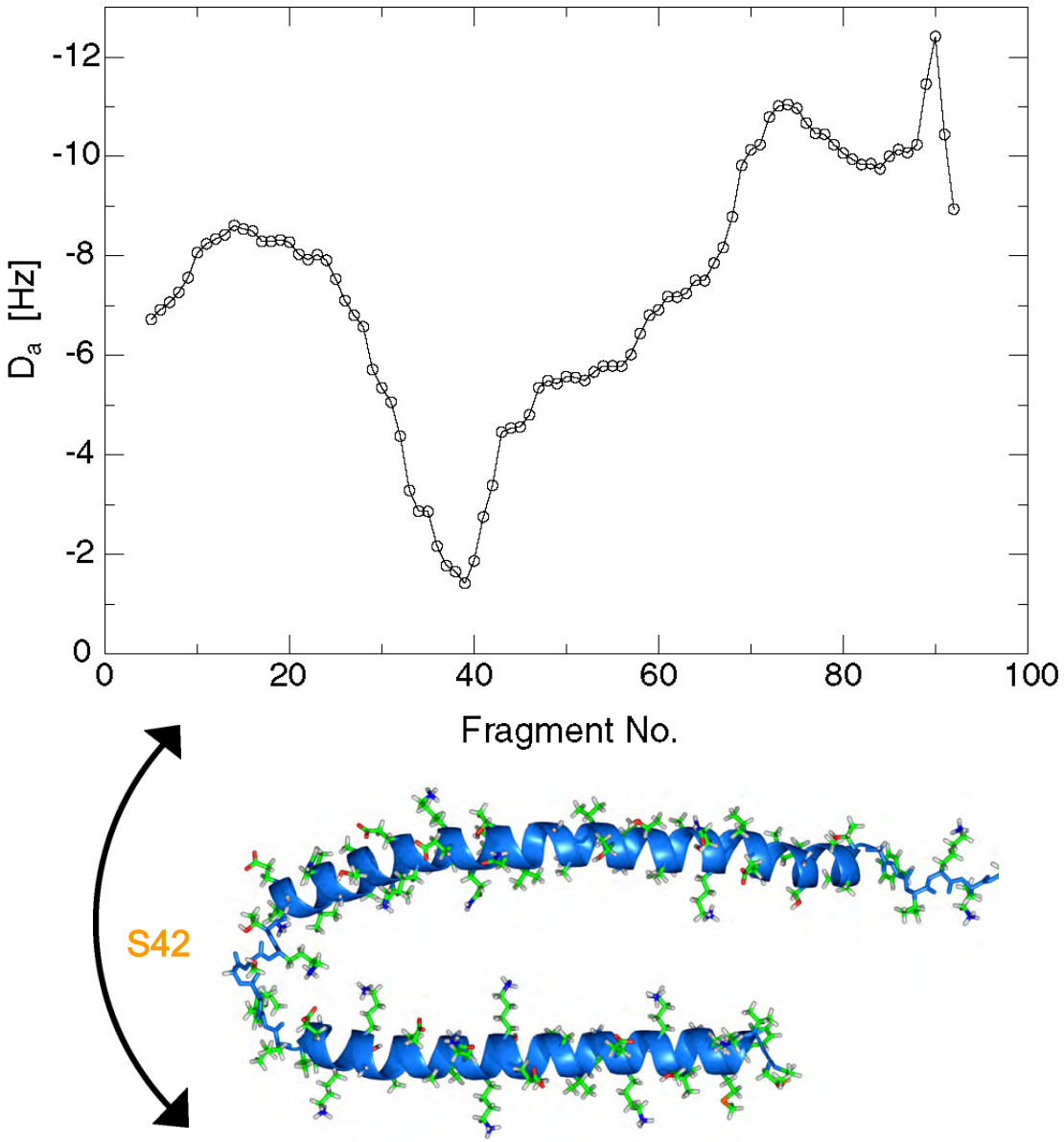
- **Magnitude**
- **Rhombicity**
- **Orientation (Euler Angles)**







MFR Fragment Tensor Magnitudes Reveal Dynamics



Gamma S

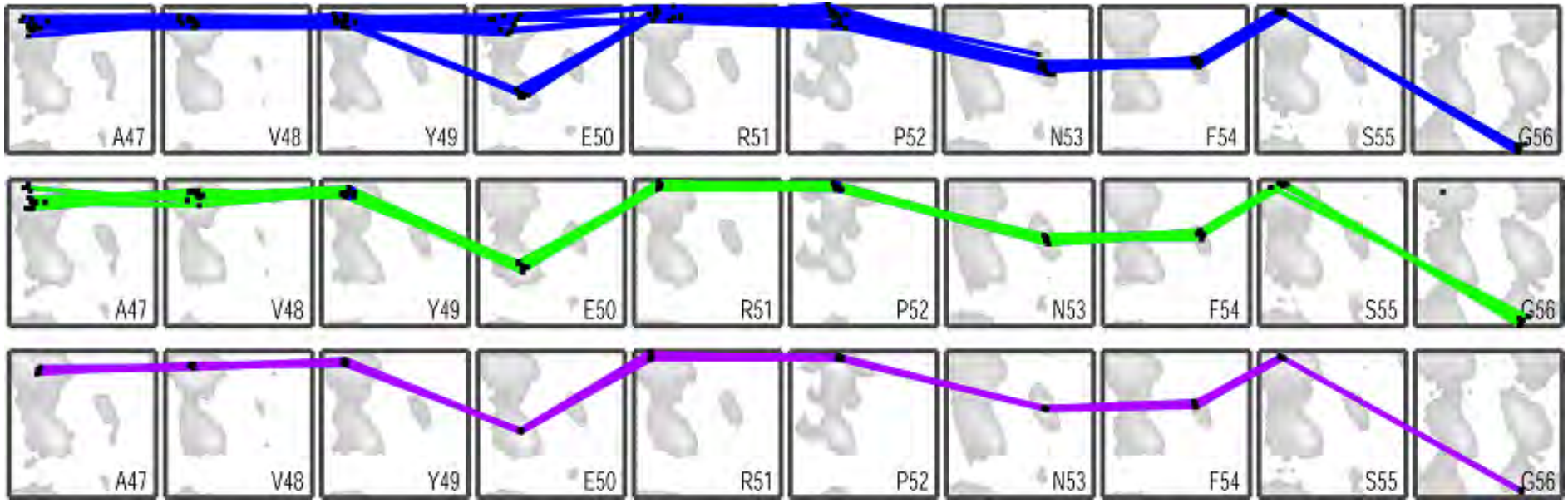
177 Residues, two similar domains, homologous structure is known.

179 Amide-Amide NOEs, 70 Methyl-Methyl NOEs, including 6 inter-domain

DC Medium 1: 144 HN-N, 111 CA-CB, 150 CA-C', 134 N-C'

DC Medium 2: 147 HN-N, 135 CA-CB, 153 CA-C', 139 N-C'

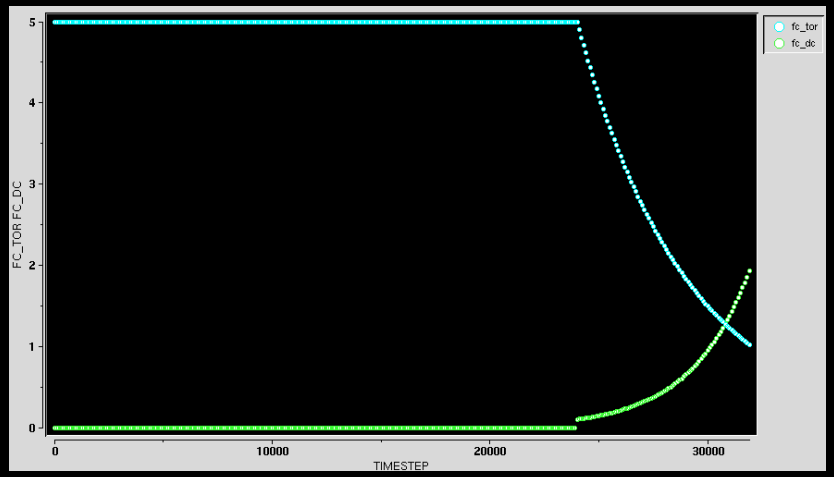
Side-chain $\chi 1$ angles from $^{3}JNC\gamma$ and $^{3}JC'C\gamma$ couplings, $\chi 2$ from $^{3}JC\gamma C\delta$



- **Conduct MFR Search with SVD (free tensor)**
- **Conduct second MFR Search with fixed tensor Da, Rh, and relative orientation**
- **Refine all fragments with fixed tensor Da, Rh to yield Phi and Psi for 90% of residues; 50% have better than 5 degree RMS consenus; 33% are 3 degree RMS or better.**

- MFR Torsions Preserve Secondary Structure During High Temperature Phase
- During Cooling, MFR Torsion Restraint Force Constant is Decreased
- DC Force Constant is Increased as Ideal Fold is Approached

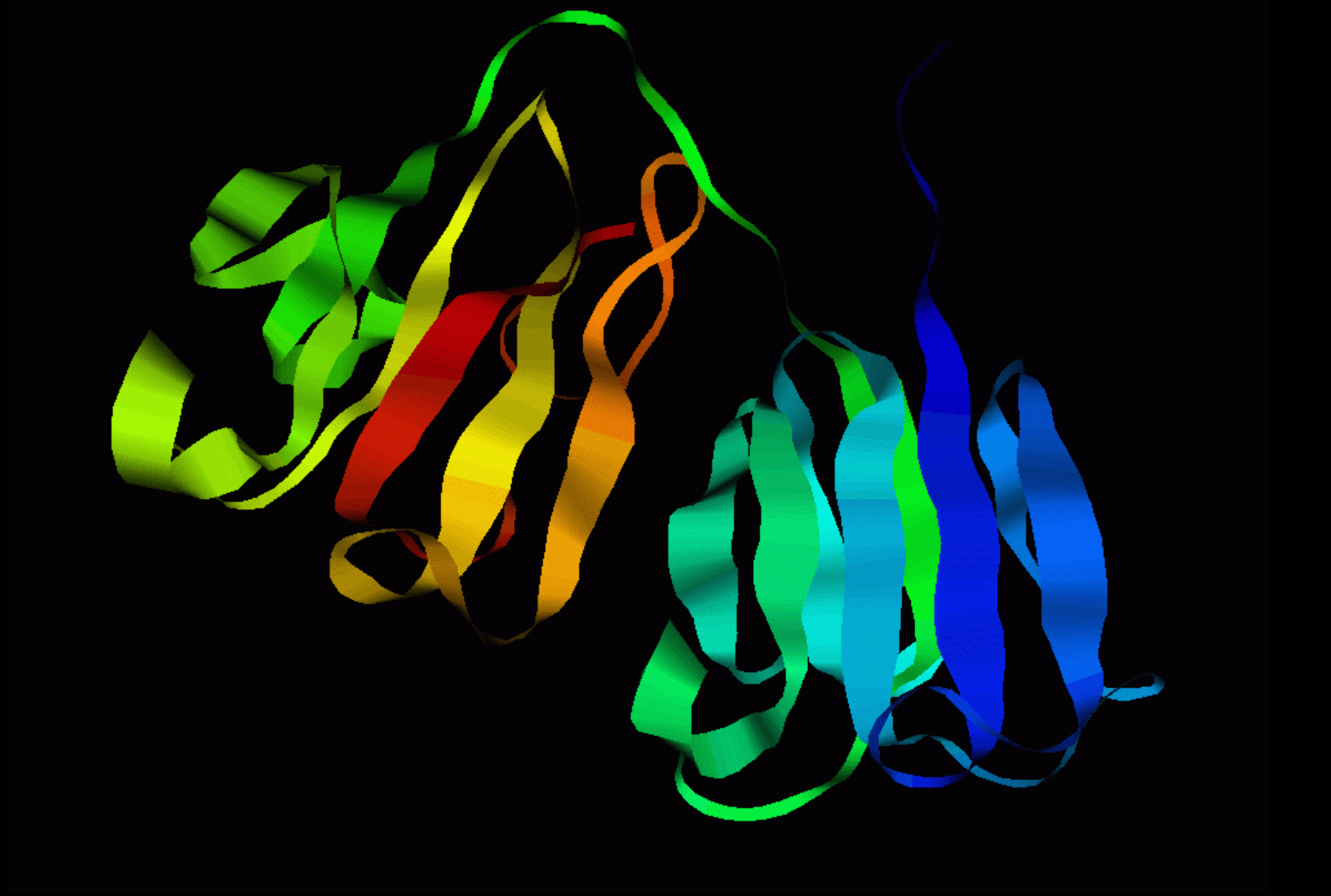
```
dynReadGMC -gmc $gmcDir -pdb $pdbName

for {set i 1} {$i <= $count} {incr i} \
   {
    dynSimulateAnnealing -graph -print 50 -rasmol 500 \
        -sa    stepCount   init       100 \
               stepCount   high     24000 \
               stepCount   cool      8000 \
               timeStep    all          3 \
               temperature all       4000 \
               temperature coolEnd      0 \
         -fc   dc          coolEnd    2.0 \
               torsion     all         50 \
               torsion     coolEnd     10 \
               noe         all         25 \
               noe         coolEnd    100 \
               radGyr      all        0.0

    set outName [format $outTemplate $i]

    dynWrite -pdb -src $dynInfo(gmc,pdb) -out $outName -rem $dynInfo(energyText)
    dynRead  -pdb -src $dynInfo(gmc,pdb) -in  $pdbName

    incr iseed 111
    srand $iseed
   }
```



Backbone RMSD γS(NMR) and γB-crystallin (X-ray)

C-terminal domain: 1.09 A

N-terminal domain: 0.63 A



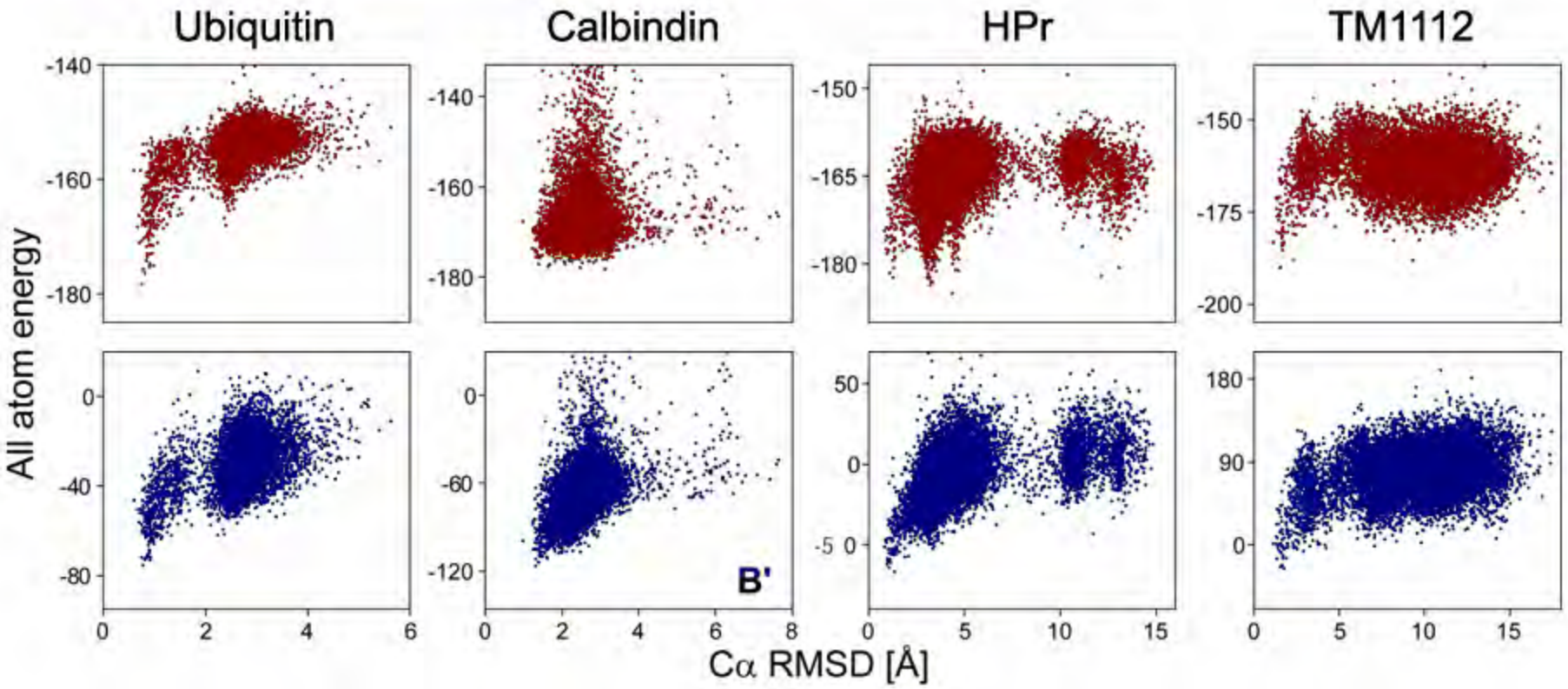
Consistent blind protein structure generation from NMR chemical shift data

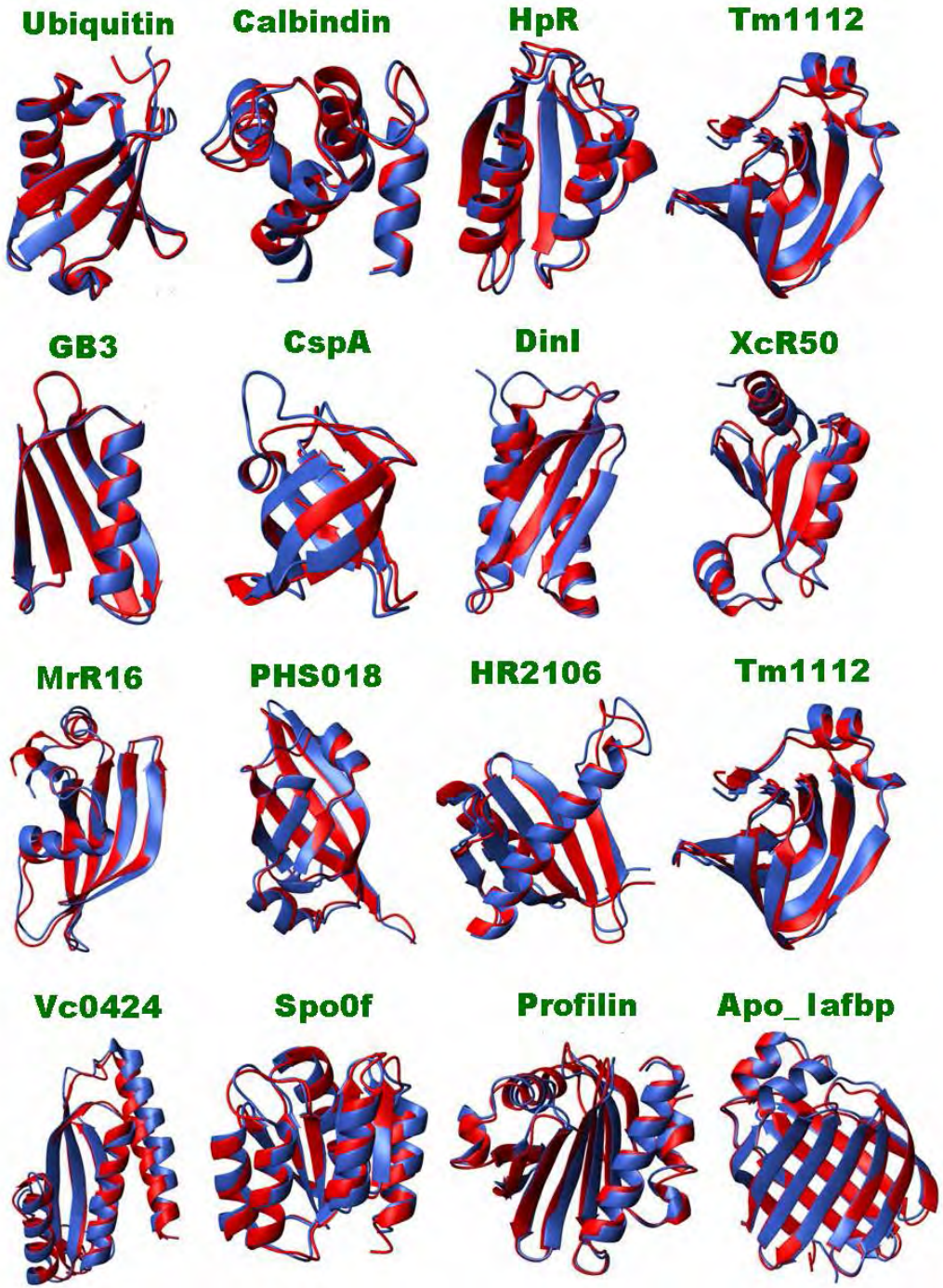
Proc Natl Acad Sci USA, (2008) 105, 4685-4690

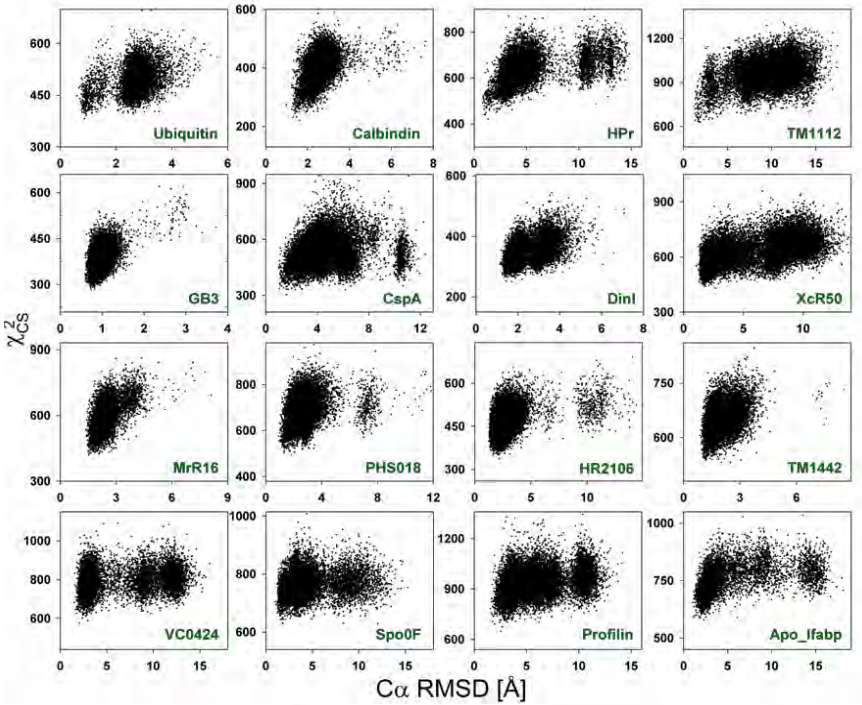
Yang Shen
Oliver Lange
Frank Delaglio
Paolo Rossi
James M. Aramini
Gaohua Liu
Alexander Eletsky
Yibing Wu
Kiran K. Singarapu
Alexander Lemak
Alexandr Ignatchenko
Cheryl H. Arrowsmith
Thomas Szyperski
Gaetano T. Montelione
David Baker
Ad Bax

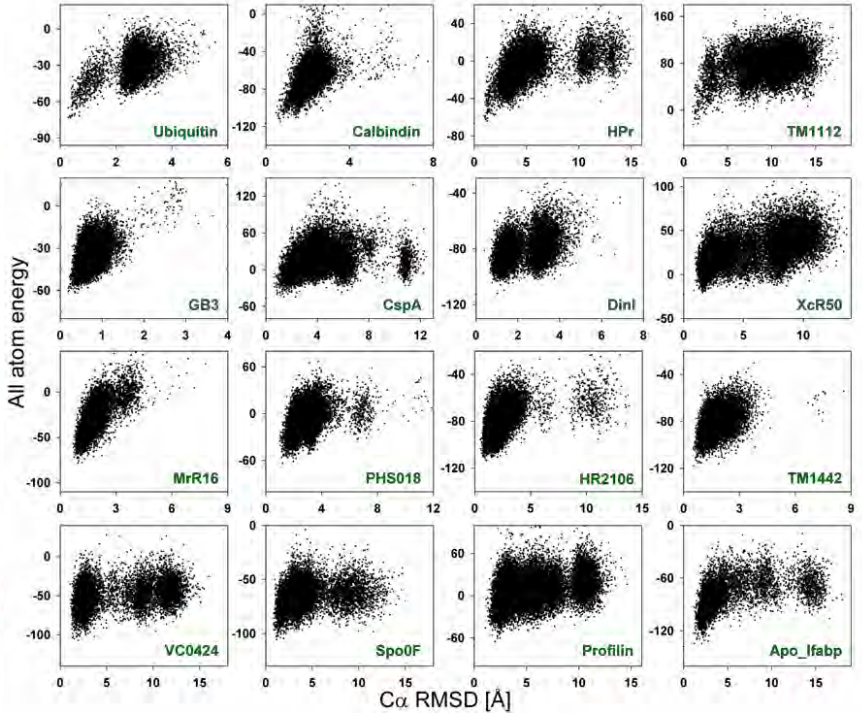
Using SPARTA Chemical Shift Prediction to Improve ROSETTA Scoring Function

- Original ROSETTA Energy Score
- CS-ROSETTA Energy Score



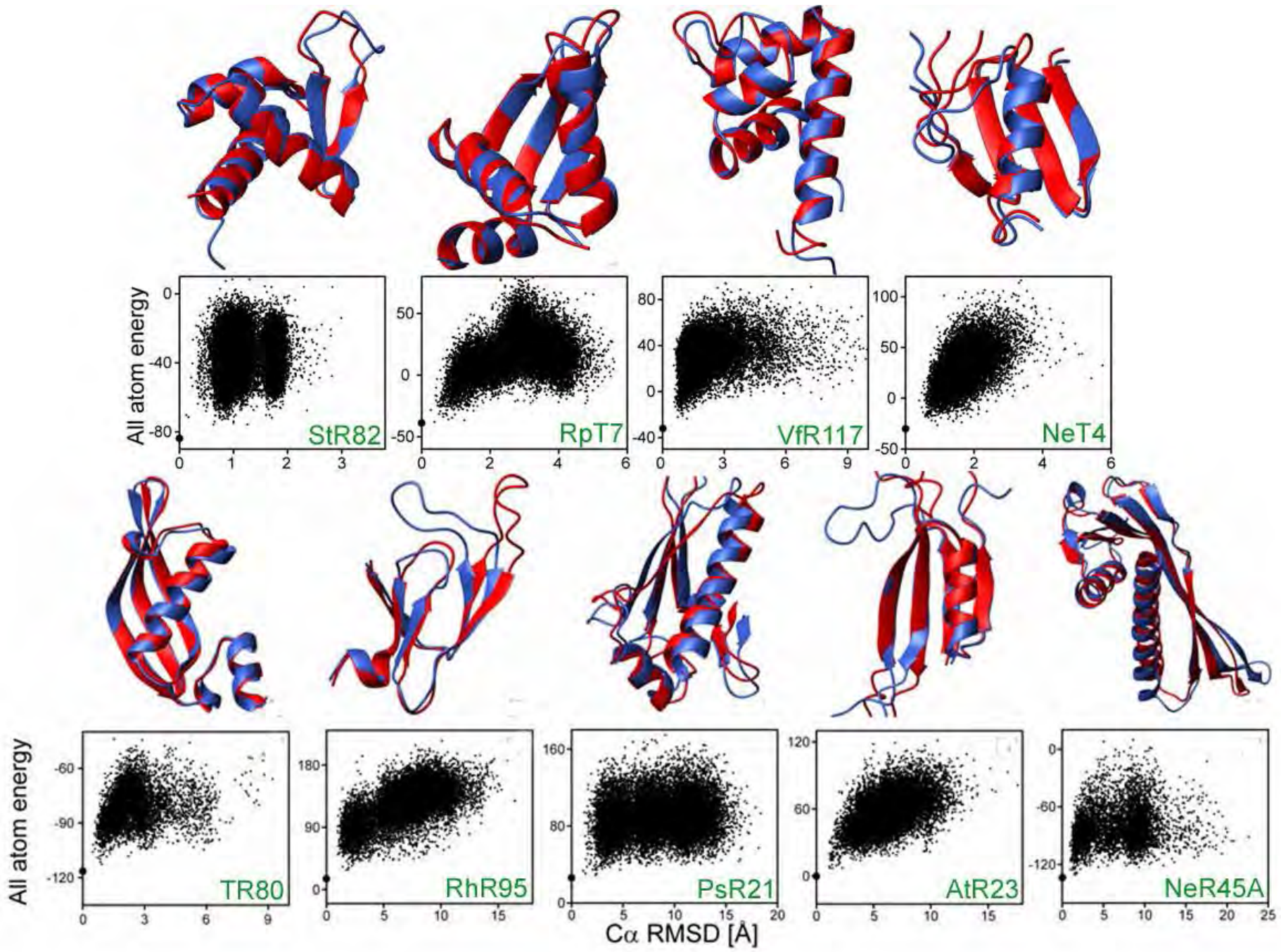

Ubiquitin
Calbindin
HpR
Tm1112
GB3
CspA
DinI
XcR50
MrR16
PHS018
HR2106
Tm1112
Vc0424
Spo0f
Profilin
Apo_ Iafbp


Ubiquitin
Calbindin
HPr
TM1112
GB3
CspA
DinI
XcR50
MrR16
PHS018
HR2106
TM1442
VC0424
Spo0F
Profilin
Apo_Ifabp
χ^2_{CS}
Cα RMSD [Å]

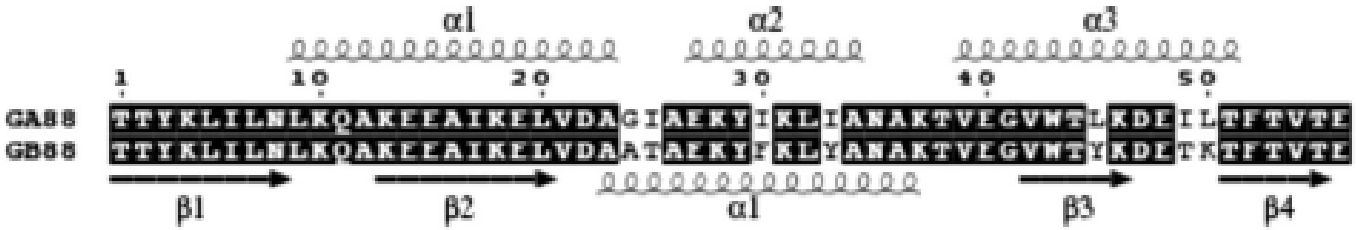

Ubiquitin
Calbindin
HPr
TM1112
GB3
CspA
DinI
XcR50
MrR16
PHS018
HR2106
TM1442
VC0424
Spo0F
Profilin
Apo_Ifabp
All atom energy
Cα RMSD [Å]

CS-ROSETTA performance on nine structural genomics proteins

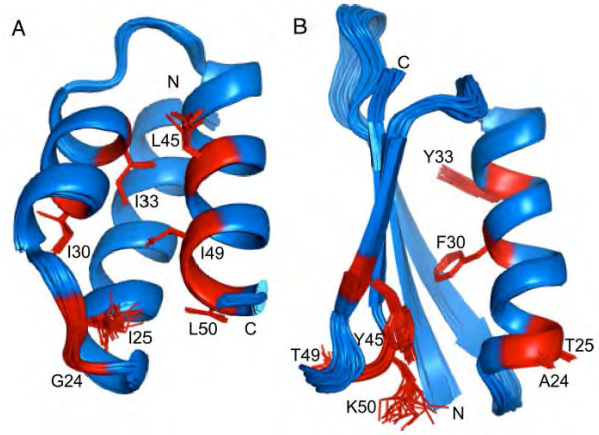
Protein	Number of Residues	PDB ID	RMSD Å (backbone)	RMSD Å (all)	% NOE Peak Agreement
RpT7	65	2jtv	**0.64**	**1.29**	69
StR82	69	2jt1	**0.57**	**1.14**	65
RhR95	72	2jvm	**0.66**	**1.18**	55
NeT4	73	2jv8	**0.70**	**1.42**	57
TR80	78	2jxt	**0.69**	**1.27**	67
VfR117	80	2jvw	**0.60**	**1.40**	37
PsR211	100	2jva	**2.07**	**2.34**	57
AtR23	101	2jya	**1.10**	**1.81**	60
NeR45A	147	2jxn	**2.03**	**2.85**	53


All atom energy
StR82
RpT7
VfR117
NeT4
TR80
RhR95
PsR21
AtR23
NeR45A
Cα RMSD [Å]

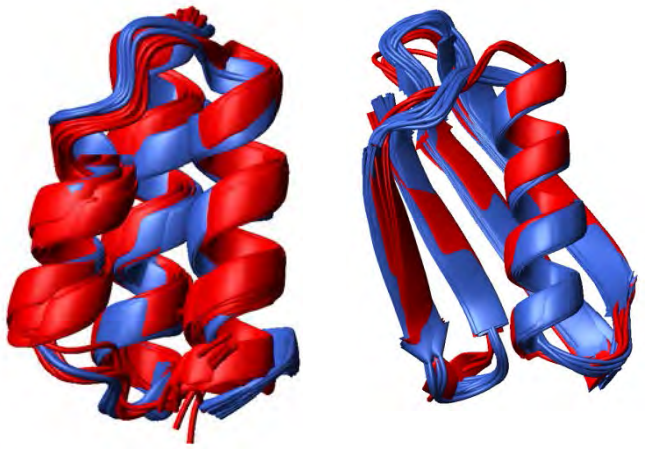
Structures of two designed proteins with high sequence identity



NMR structures of Ga88 and Gb88



NMR structures *vs* csRosetta models



Mean-to-mean backbone RMSD

1.07A

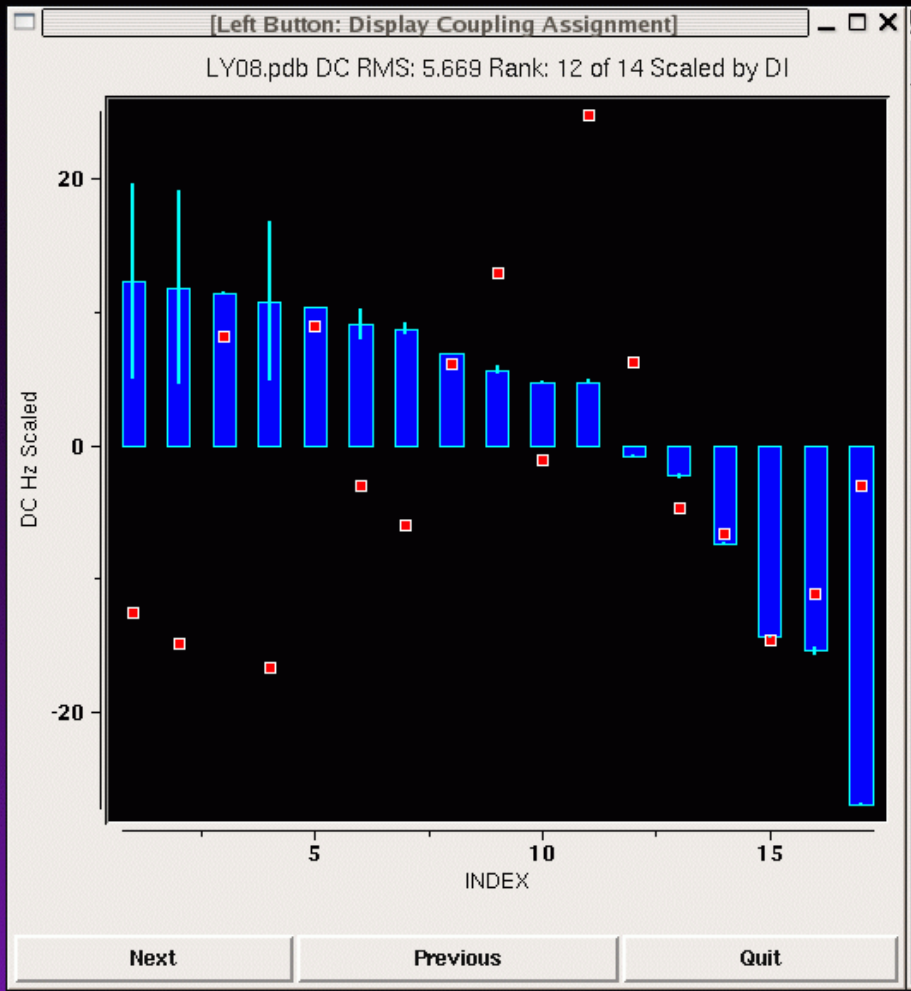
1.31A

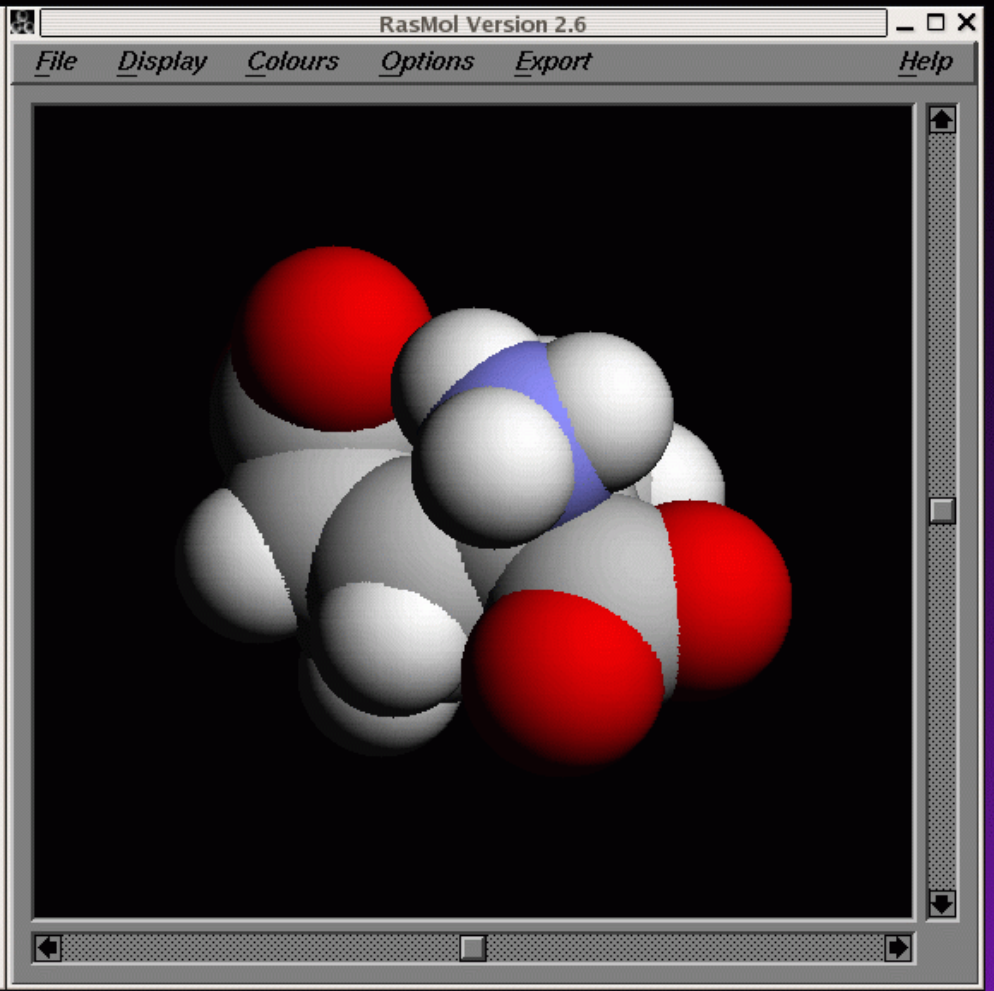
Patrick A. Alexander, Yanan He, Yihong Chen, John Orban, and Philip N. Bryan

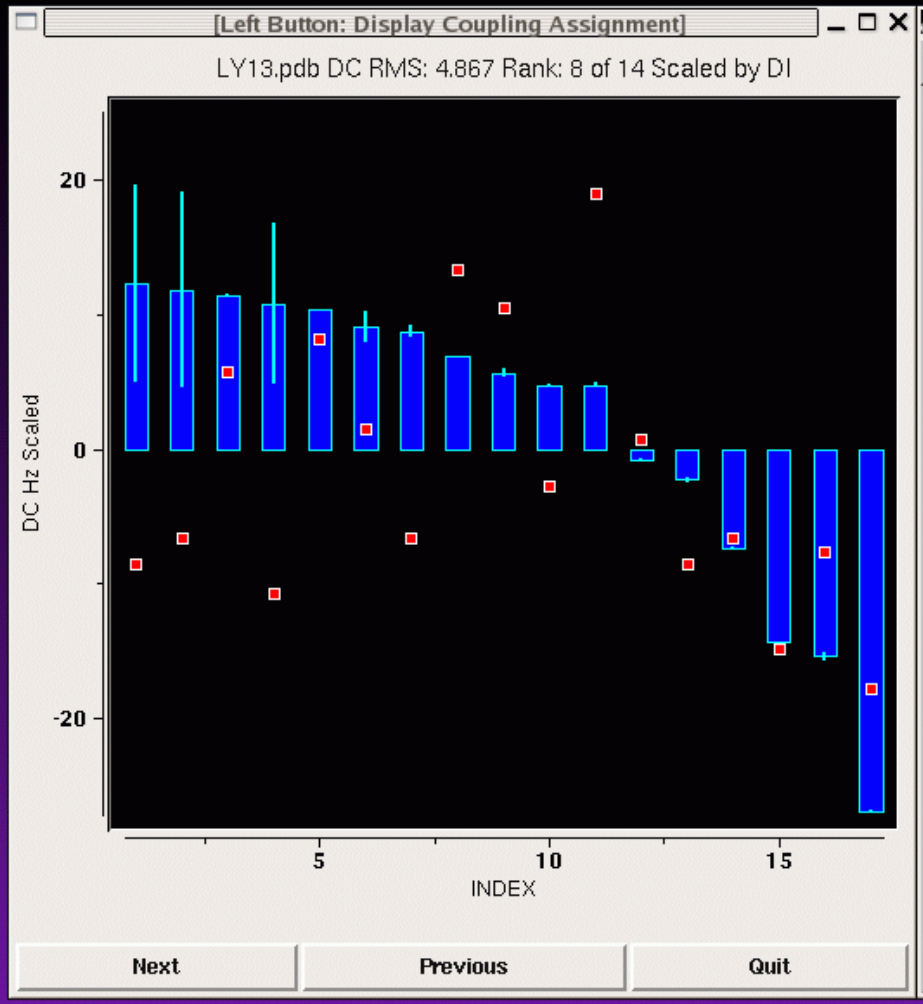
PNAS, 2007, 104:11963-11968

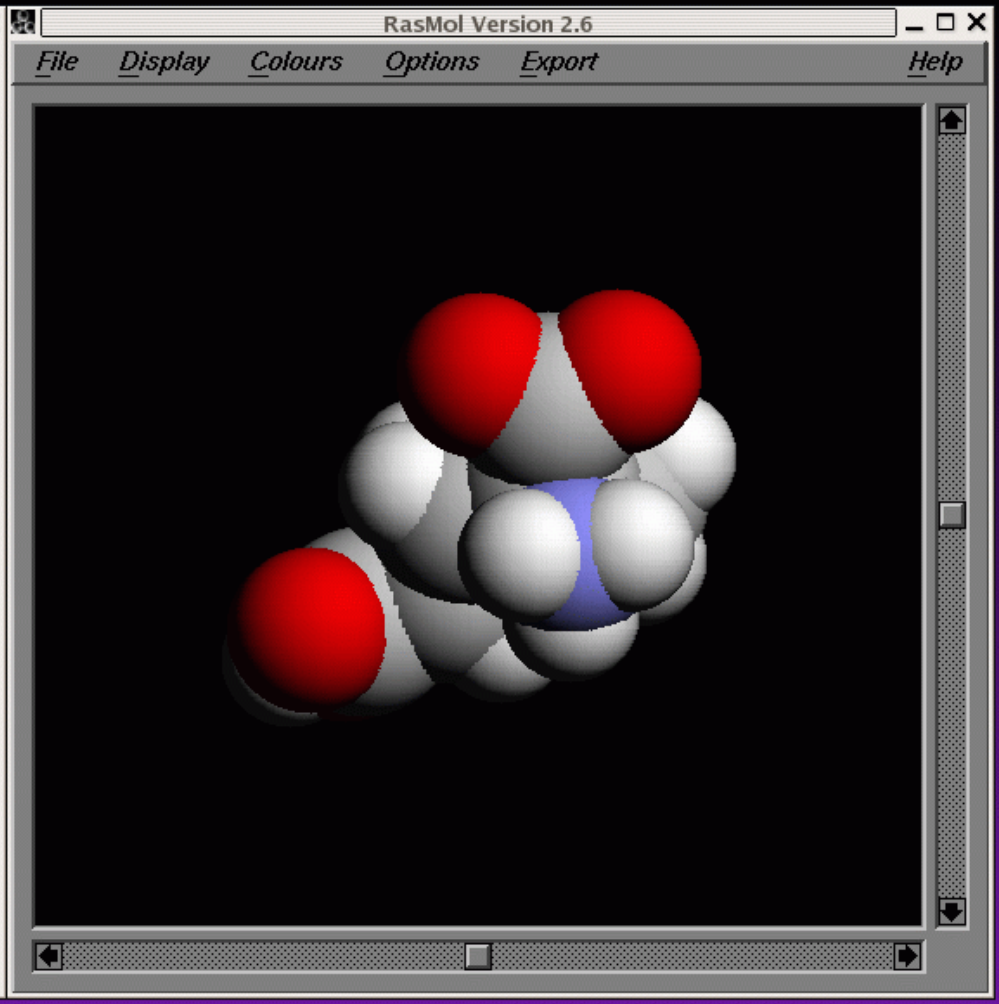
PNAS, 2008, 105:14412-14417

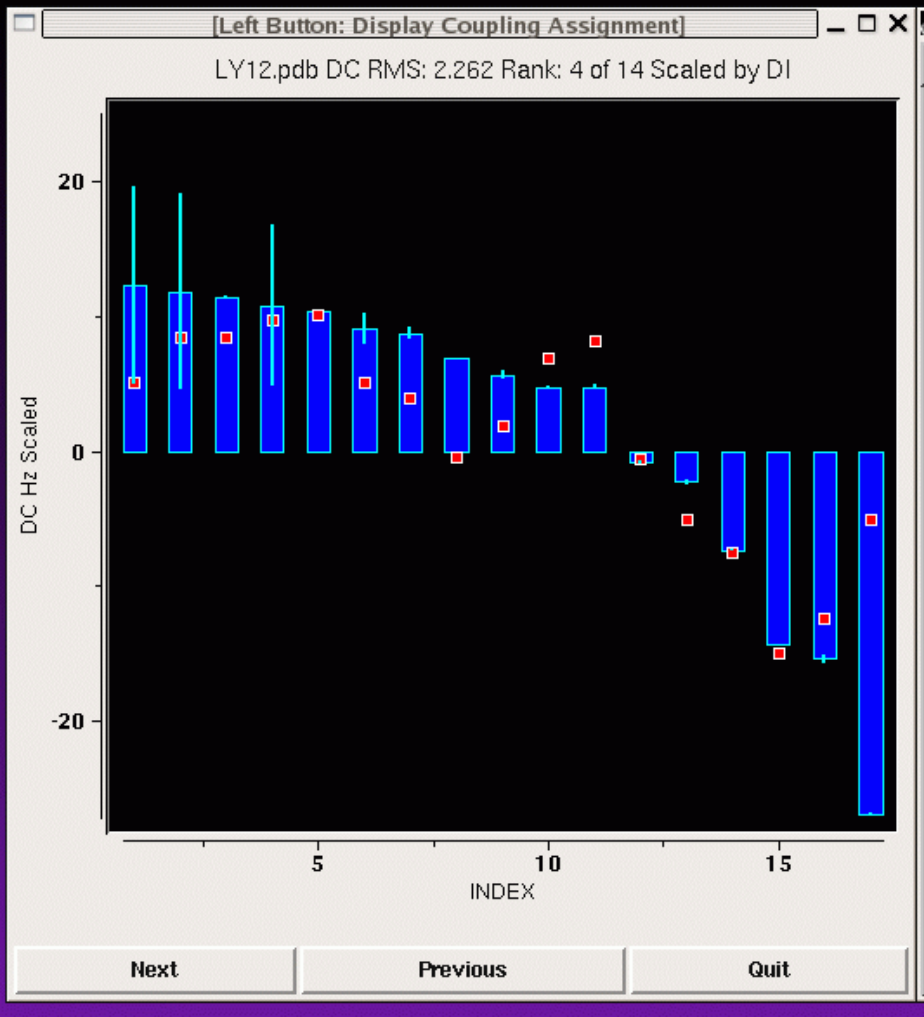
NMR Applications in Drug Discovery

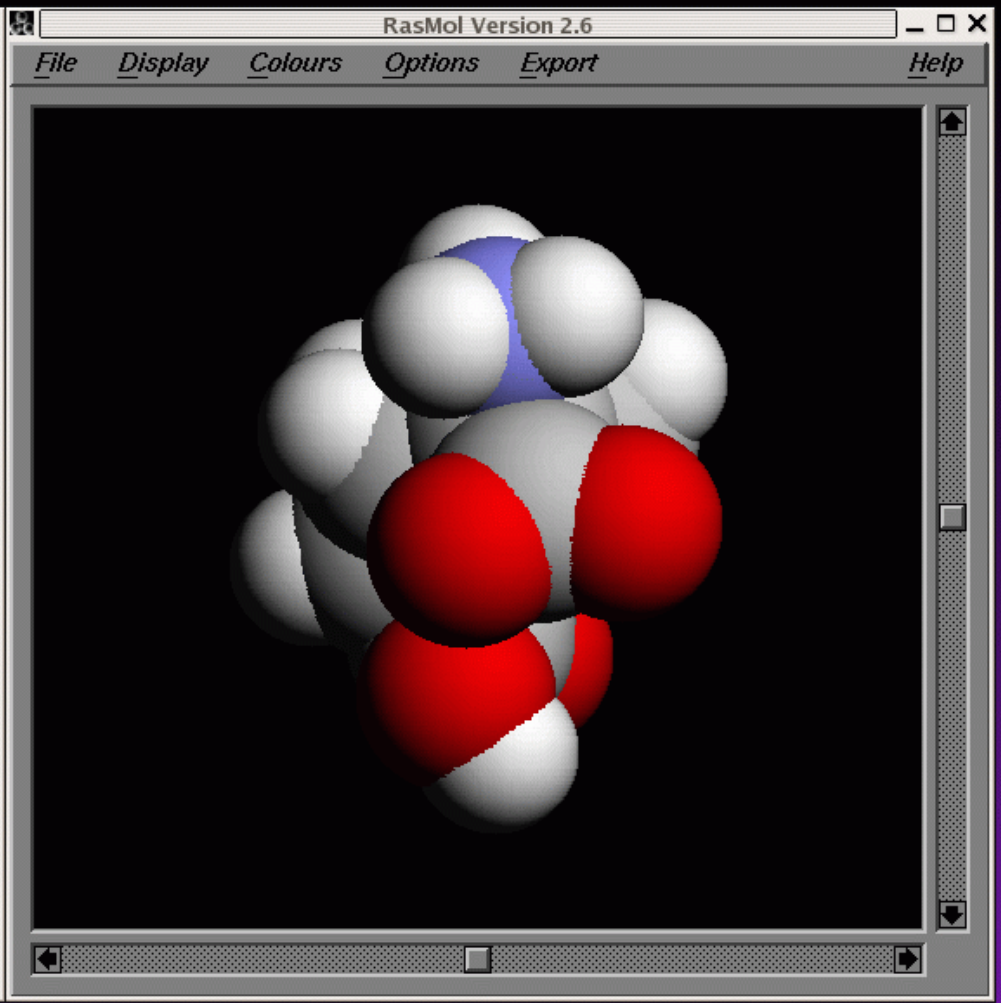

[Left Button: Display Coupling Assignment]
LY08.pdb DC RMS: 5.669 Rank: 12 of 14 Scaled by DI
20
0
-20
DC Hz Scaled
5
10
15
INDEX
Next
Previous
Quit

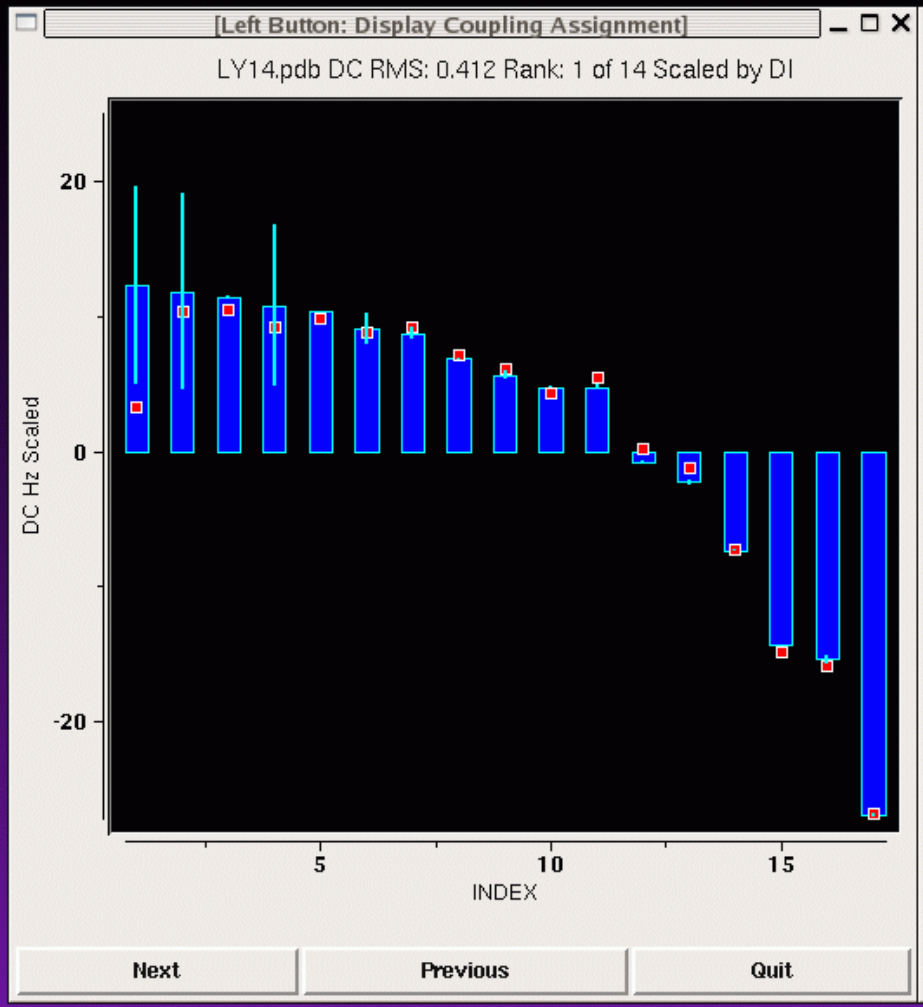

RasMol Version 2.6
File
Display
Colours
Options
Export
Help

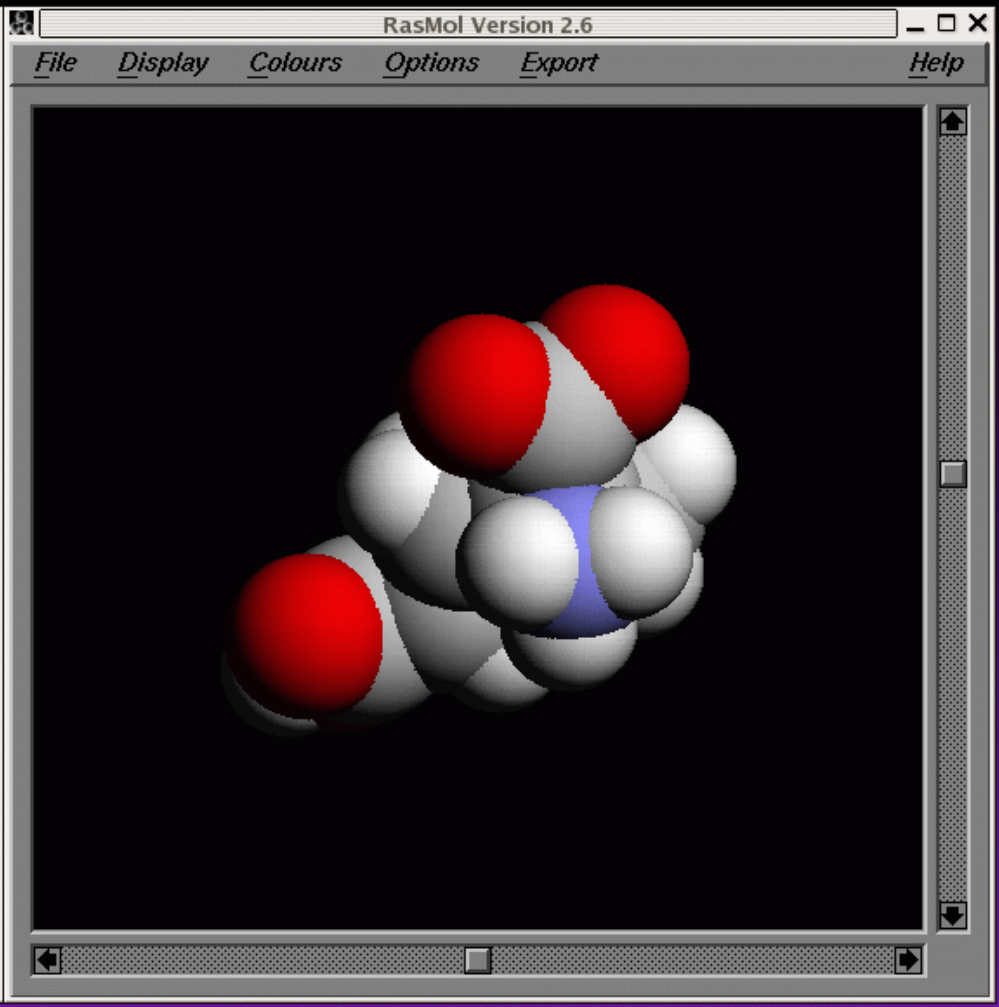

[Left Button: Display Coupling Assignment]
LY13.pdb DC RMS: 4.867 Rank: 8 of 14 Scaled by DI
20
0
-20
DC Hz Scaled
5
10
15
INDEX
Next
Previous
Quit


RasMol Version 2.6
File
Display
Colours
Options
Export
Help


[Left Button: Display Coupling Assignment]
LY12.pdb DC RMS: 2.262 Rank: 4 of 14 Scaled by DI
20
0
-20
DC Hz Scaled
5
10
15
INDEX
Next
Previous
Quit


RasMol Version 2.6
File
Display
Colours
Options
Export
Help

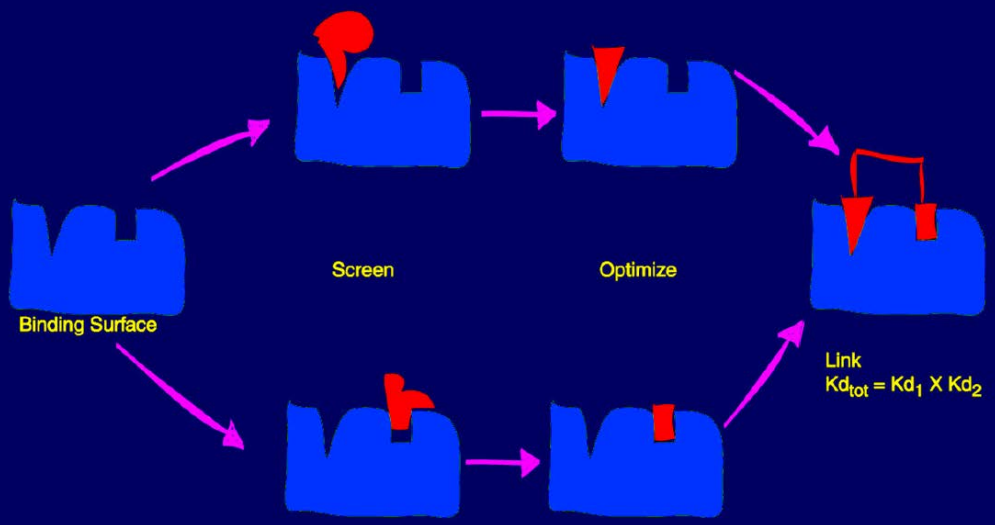

[Left Button: Display Coupling Assignment]
LY14.pdb DC RMS: 0.412 Rank: 1 of 14 Scaled by DI
20
0
-20
DC Hz Scaled
5
10
15
INDEX
Next
Previous
Quit


RasMol Version 2.6
File
Display
Colours
Options
Export
Help

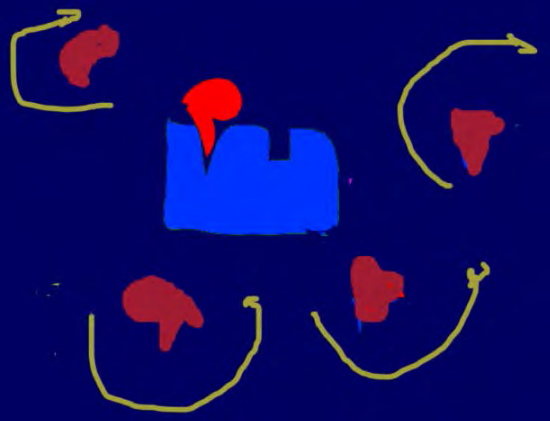
Applications of NMR in the Drug Discovery Process

NMR Spectral Series: Two Approaches

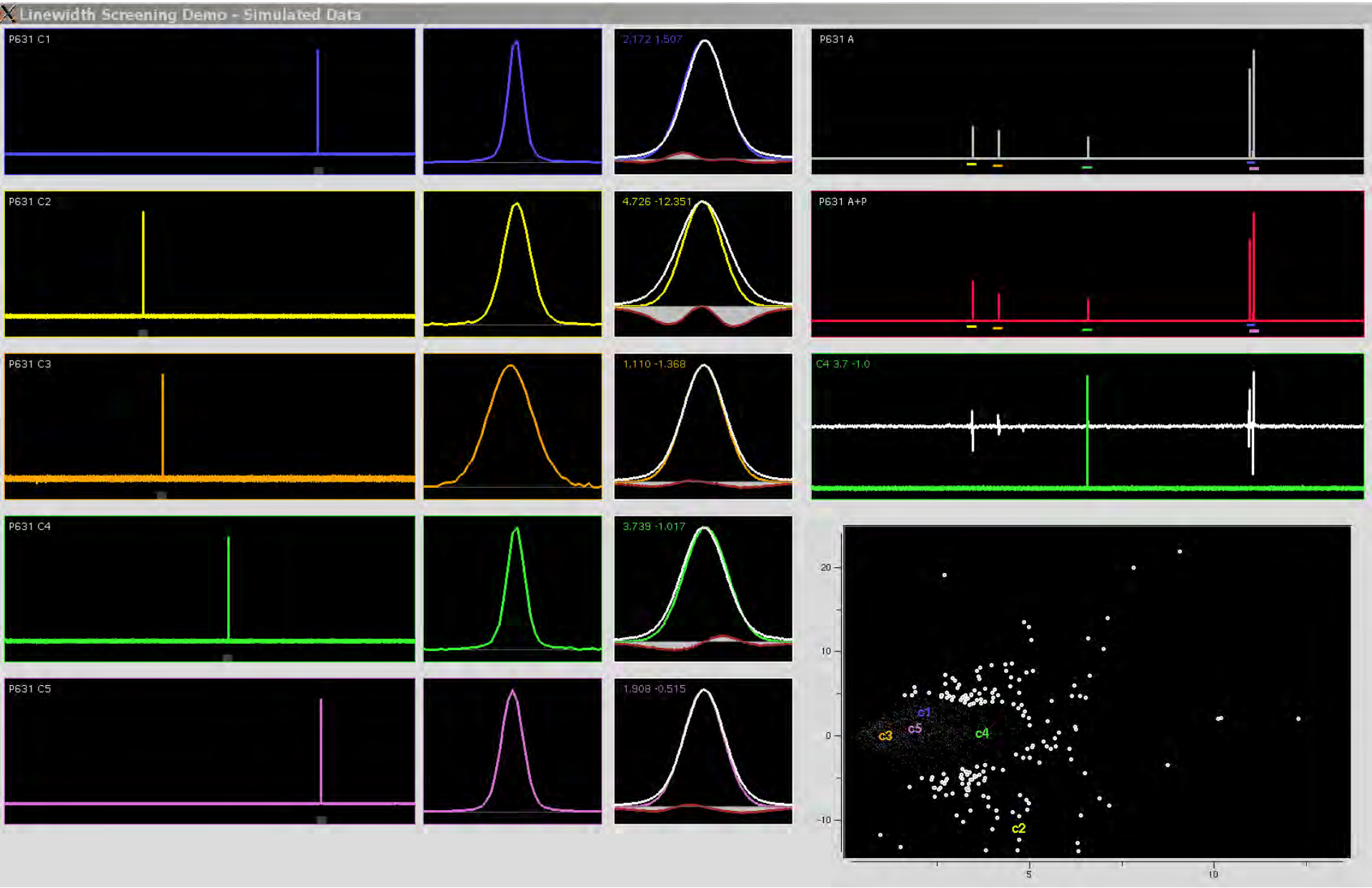
SAR by NMR (Abbott Labs)

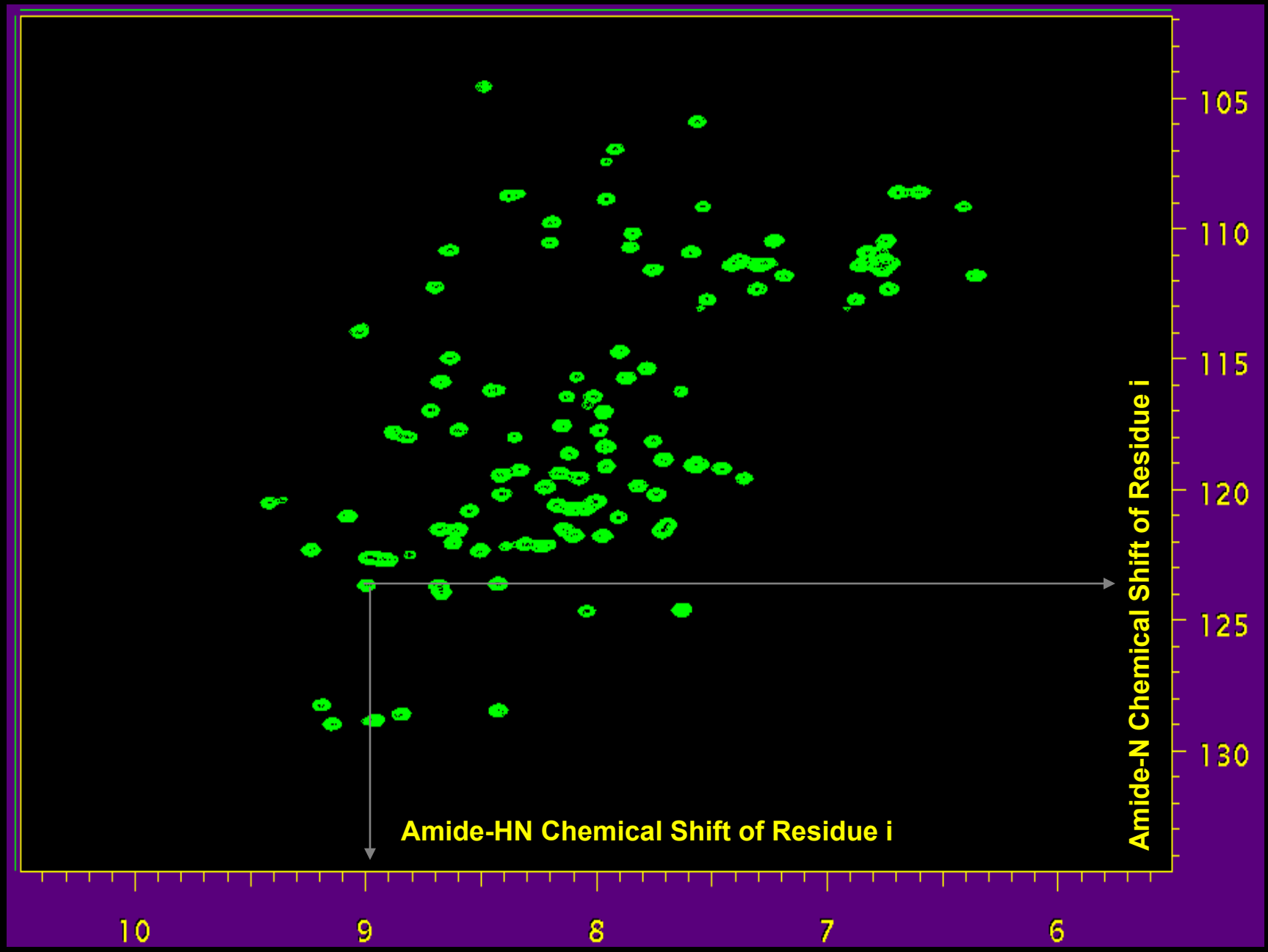


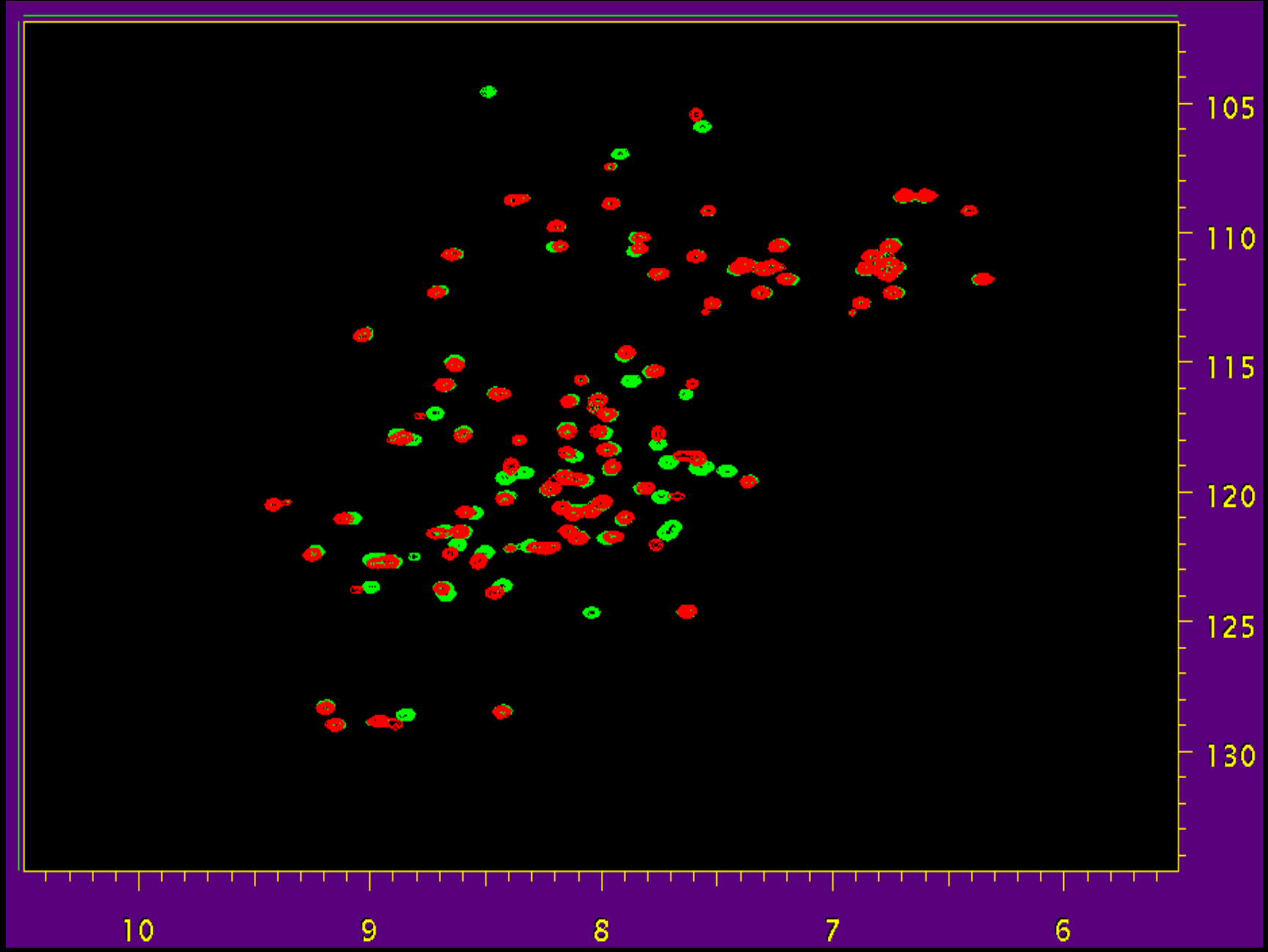
Observe Protein Signals

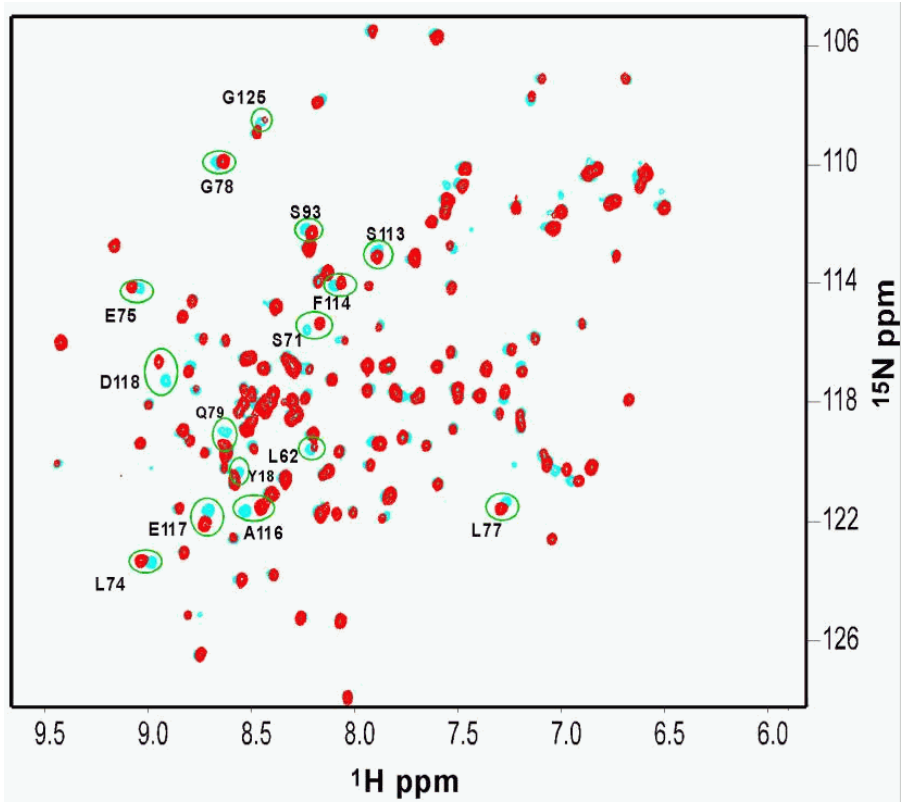


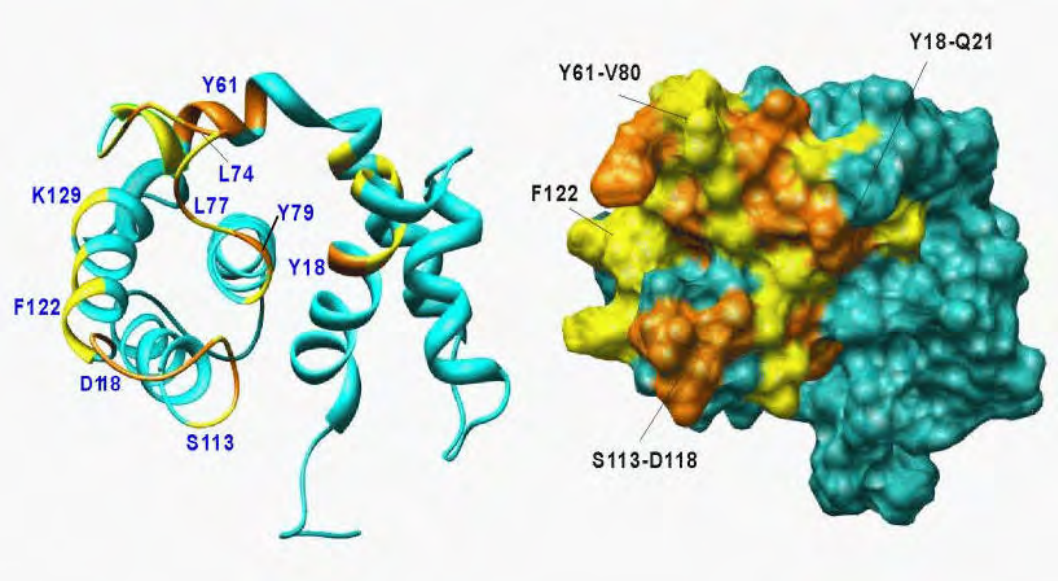
Observe Ligand Signals


Linewidth Screening Demo - Simulated Data
P631 C1
2.172 1.507
P631 A
P631 C2
4.726 -12.351
P631 A+P
P631 C3
1.110 -1.368
C4 3.7 -1.0
P631 C4
3.739 -1.017
P631 C5
1.908 -0.515
20
10
0
-10
5
10
c1
c5
c3
c4
c2

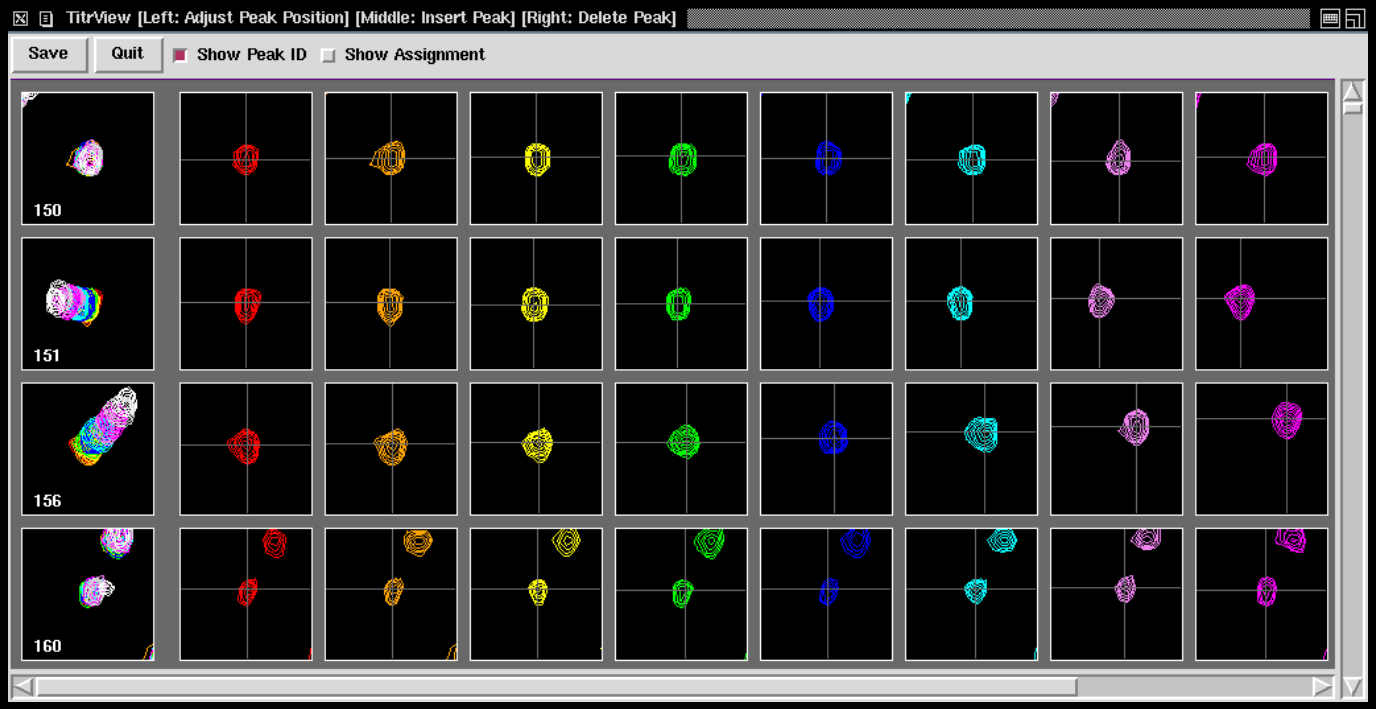

105
110
115
120
125
130
Amide-N Chemical Shift of Residue i
Amide-HN Chemical Shift of Residue i
10
9
8
7
6

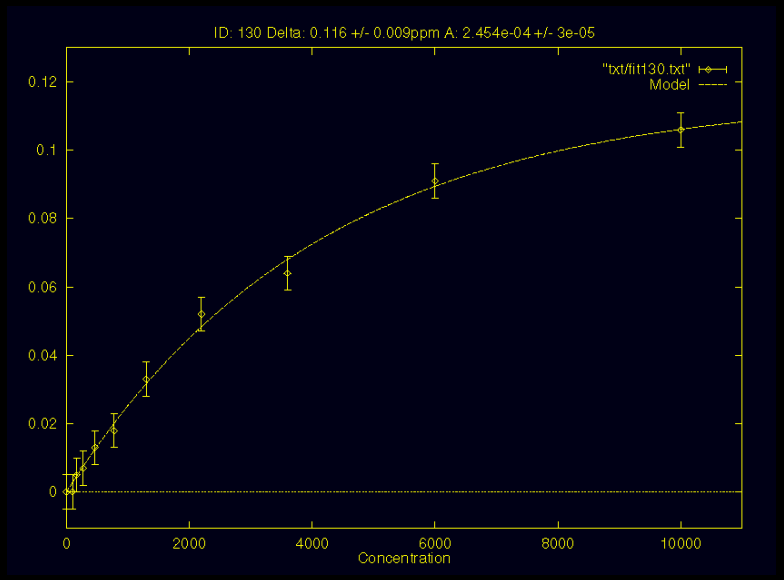

105
110
115
120
125
130
10
9
8
7
6

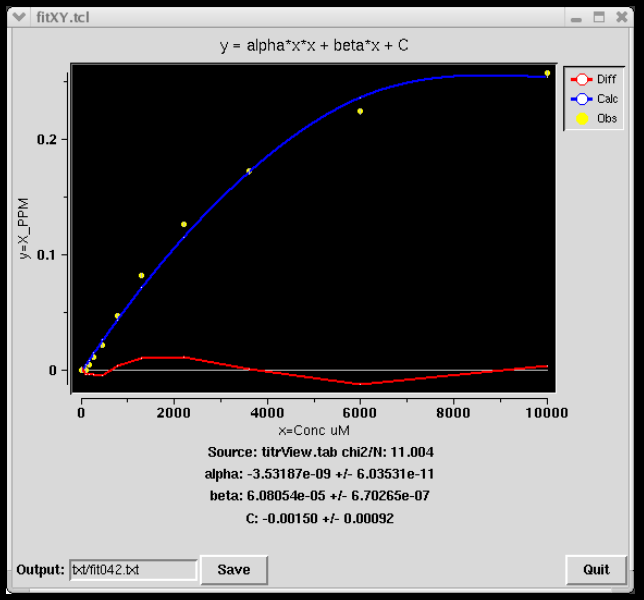

G125
G78
S93
S113
F114
E75
S71
D118
Q79
L62
Y18
E117
A116
L77
L74
106
110
114
118
122
126
15N ppm
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
1H ppm


Y61
L74
K129
L77
Y79
Y18
F122
D118
S113
Y18-Q21
Y61-V80
F122
S113-D118

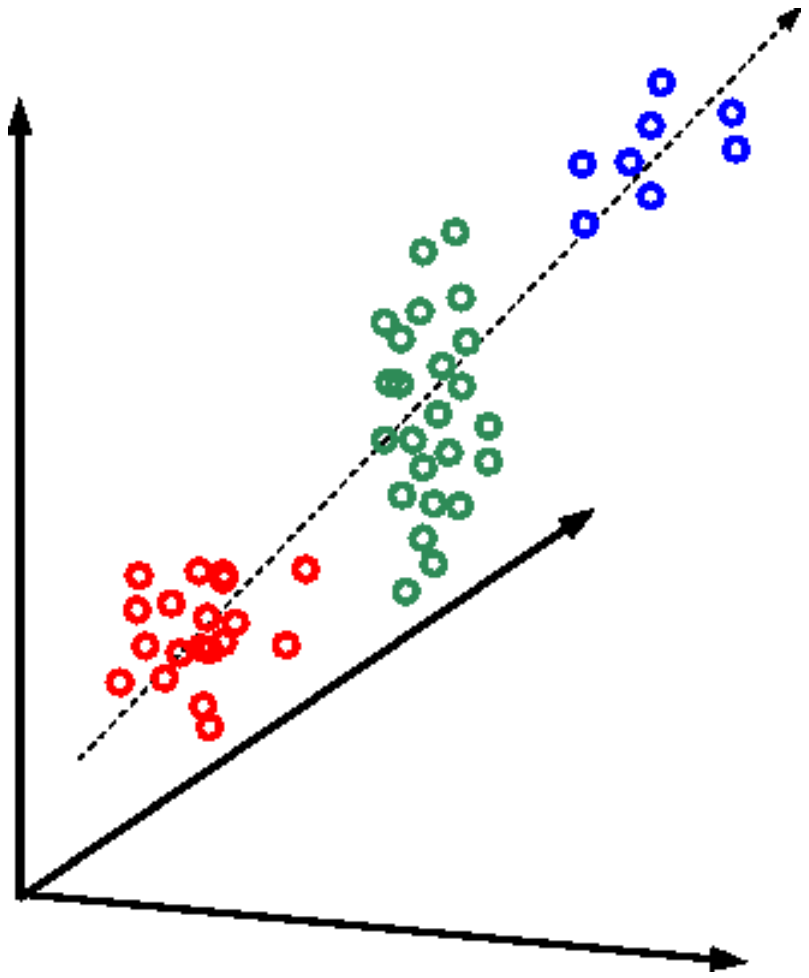
Analyze Titration Curve to Estimate K_d







Multivariate Navigation by Principal Component Analysis (PCA)

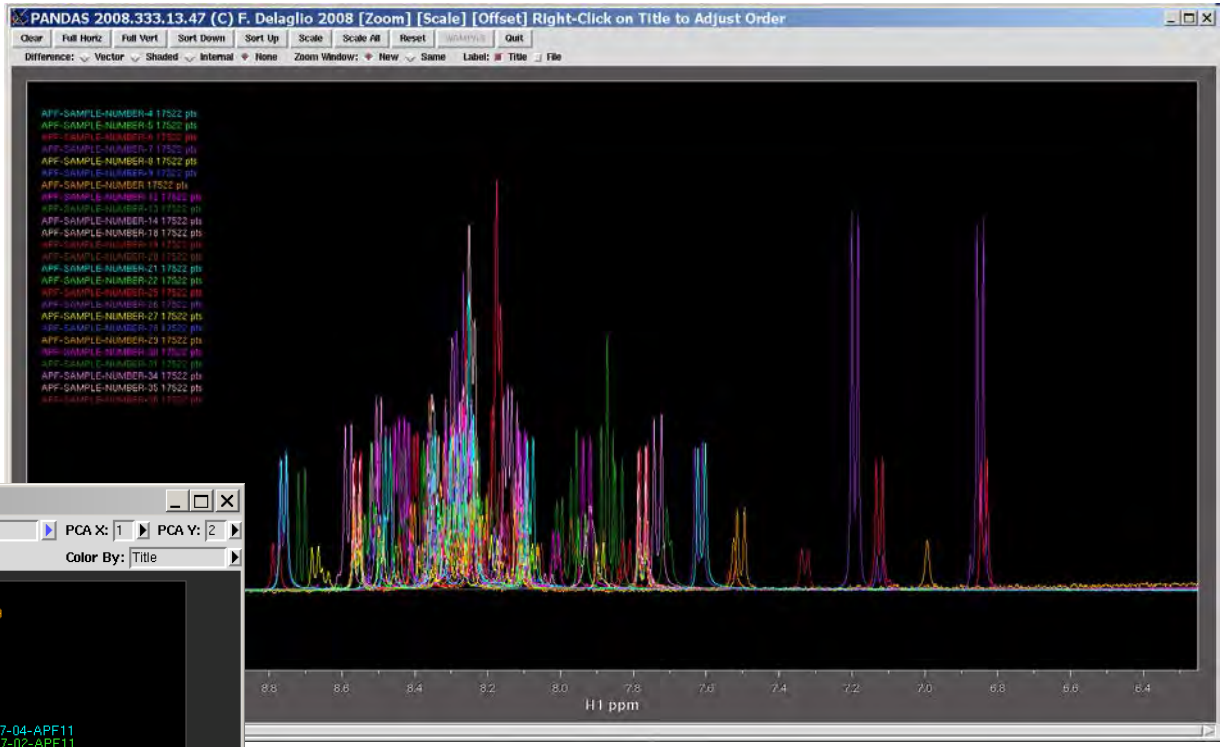


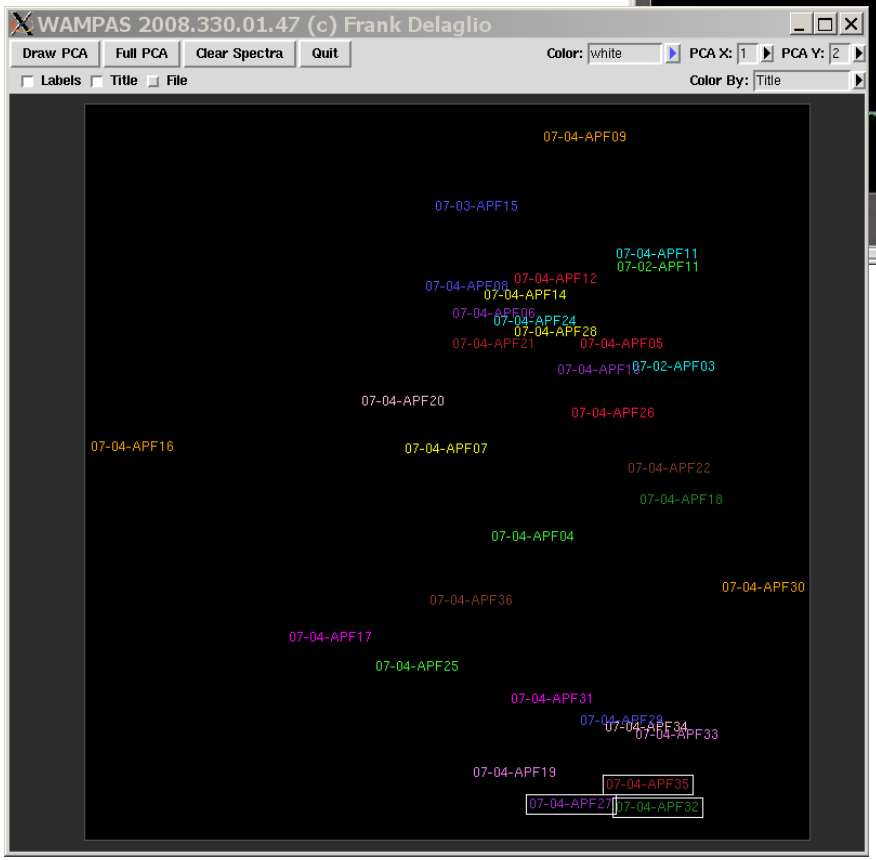
Entire spectrum is a single object in multdimensional space.

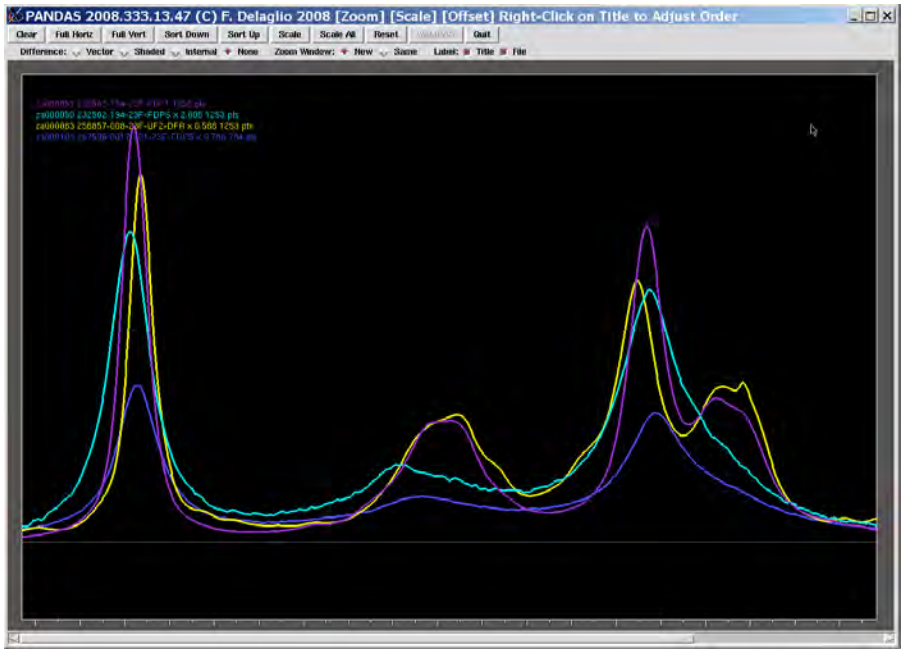
Coordinates of the object are the spectral intensities.

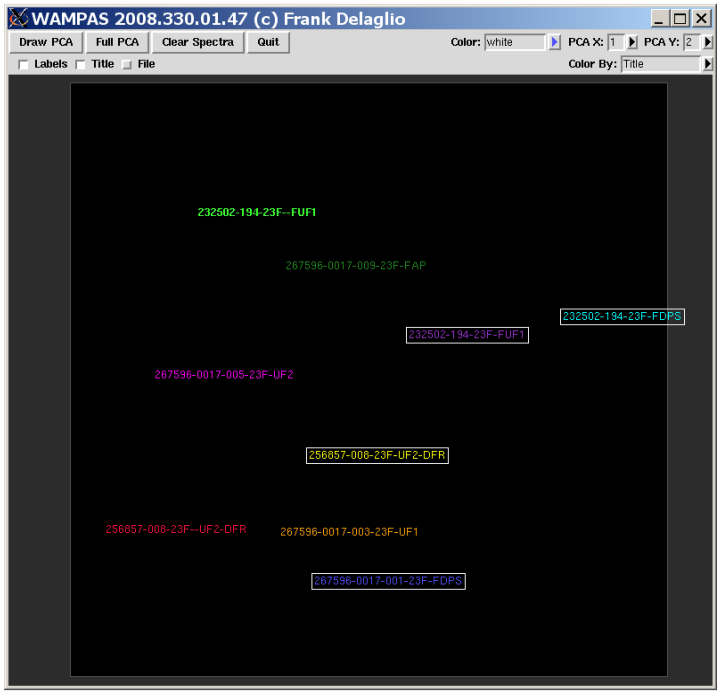
Similar spectra cluster together.

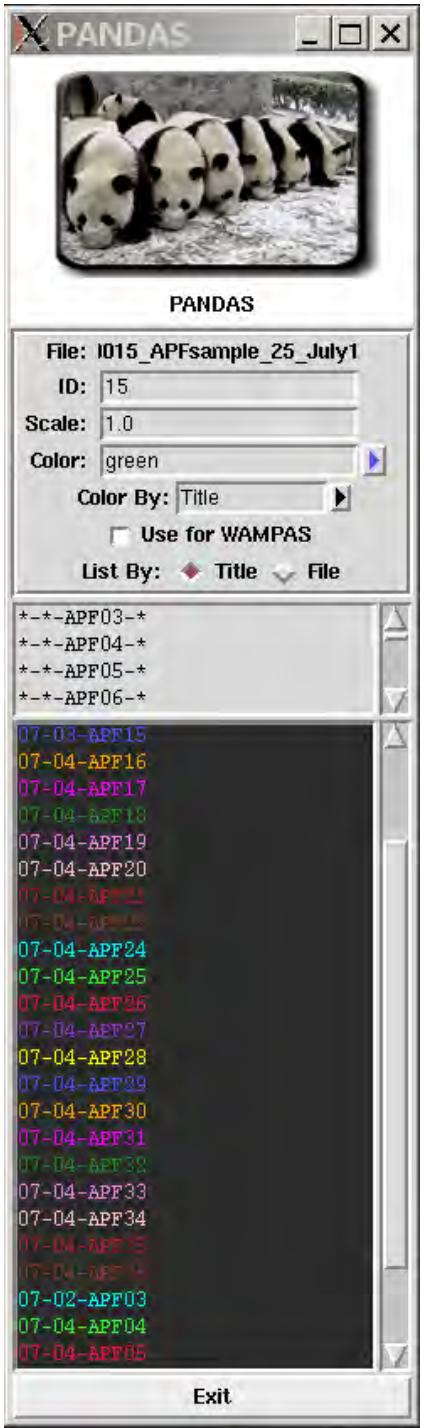
Spectra with similar features lie along lines and curves

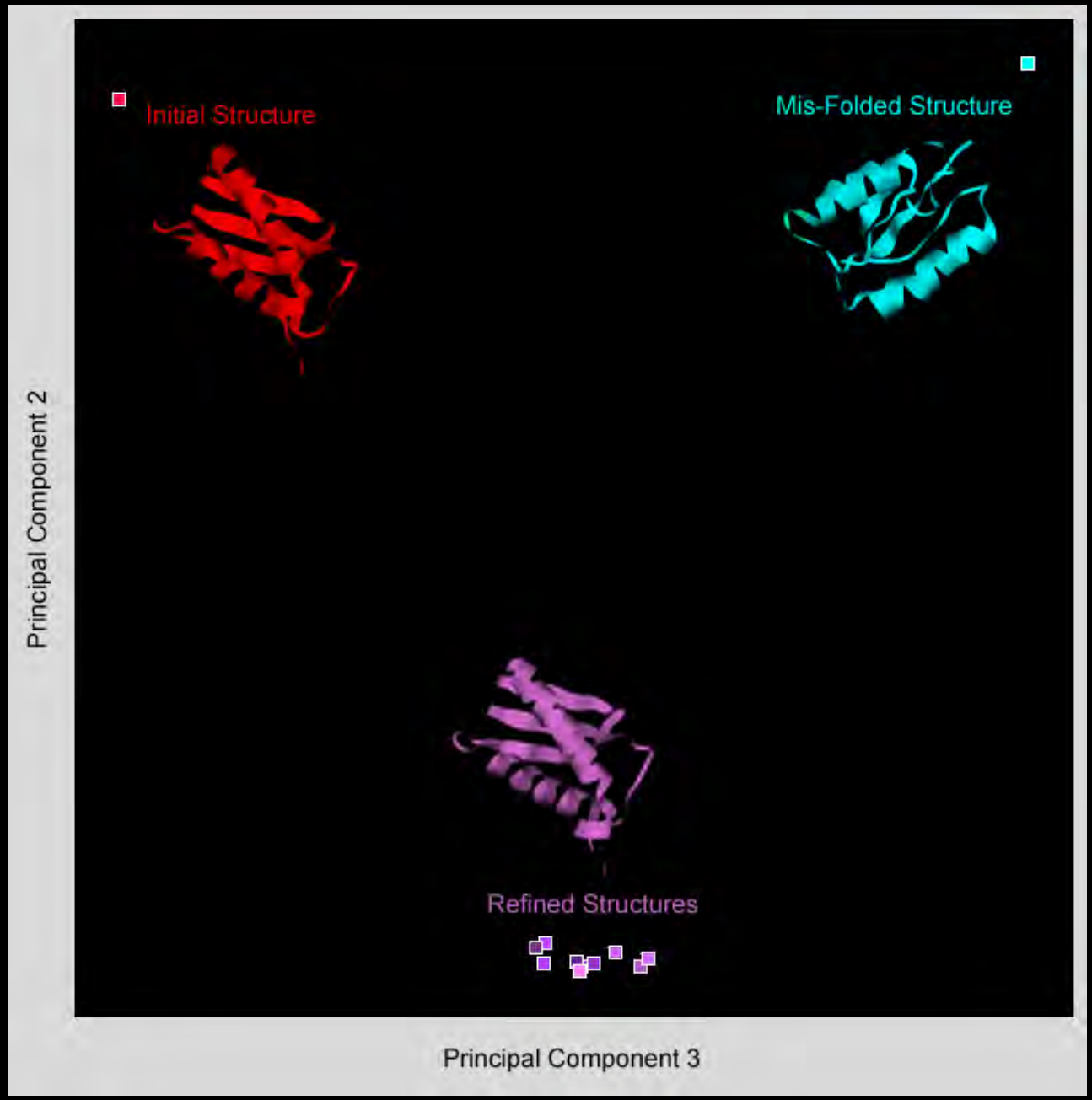

PANDAS 2008.333.13.47 (C) F. Delaglio 2008 [Zoom] [Scale] [Offset] Right-Click on Title to Adjust Order
Clear
Full Horiz
Full Vert
Sort Down
Sort Up
Scale
Scale All
Reset
Quit
Difference: Vector Shaded Internal None
Zoom Window: New Same
Label: Title File
APF-SAMPLE-NUMBER-4 17522 pts
APF-SAMPLE-NUMBER-5 17522 pts
APF-SAMPLE-NUMBER-7 17522 pts
APF-SAMPLE-NUMBER-8 17522 pts
APF-SAMPLE-NUMBER 17522 pts
APF-SAMPLE-NUMBER-14 17522 pts
APF-SAMPLE-NUMBER-18 17522 pts
APF-SAMPLE-NUMBER-21 17522 pts
APF-SAMPLE-NUMBER-22 17522 pts
APF-SAMPLE-NUMBER-27 17522 pts
APF-SAMPLE-NUMBER-29 17522 pts
APF-SAMPLE-NUMBER-34 17522 pts
APF-SAMPLE-NUMBER-35 17522 pts
8.8
8.6
8.4
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
6.6
6.4
H1 ppm


WAMPAS 2008.330.01.47 (c) Frank Delaglio
Draw PCA
Full PCA
Clear Spectra
Quit
Color: white
PCA X: 1
PCA Y: 2
Labels
Title
File
Color By: Title
07-04-APF09
07-03-APF15
07-04-APF11
07-02-APF11
07-04-APF12
07-04-APF08
07-04-APF14
07-04-APF06
07-04-APF24
07-04-APF28
07-04-APF21
07-04-APF05
07-04-APF10
07-02-APF03
07-04-APF20
07-04-APF26
07-04-APF16
07-04-APF07
07-04-APF22
07-04-APF18
07-04-APF04
07-04-APF30
07-04-APF36
07-04-APF17
07-04-APF25
07-04-APF31
07-04-APF28
07-04-APF34
07-04-APF33
07-04-APF19
07-04-APF35
07-04-APF27
07-04-APF32


PANDAS 2008.333.13.47 (C) F. Delaglio 2008 [Zoom] [Scale] [Offset] Right-Click on Title to Adjust Order
Clear
Full Horiz
Full Vert
Sort Down
Sort Up
Scale
Scale All
Reset
Quit
Difference:
Vector
Shaded
Internal
None
Zoom Window:
New
Zoom
Label:
Title
File

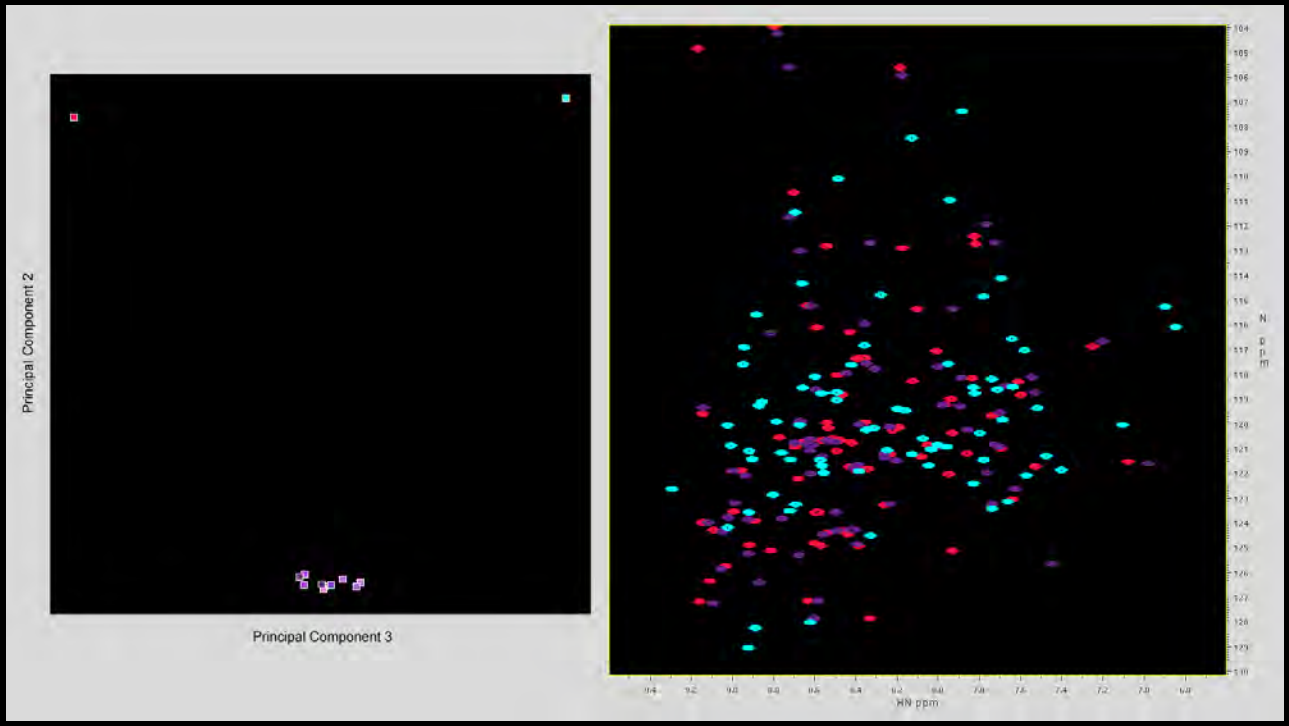

WAMPAS 2008.330.01.47 (c) Frank Delaglio
Draw PCA
Full PCA
Clear Spectra
Quit
Color:
white
PCA X:
1
PCA Y:
2
Labels
Title
File
Color By:
Title
232502-194-23F--FUF1
267596-0017-009-23F-FAP
232502-194-23F-FDPS
232502-194-23F-FUF1
267596-0017-005-23F-UF2
256857-008-23F-UF2-DFR
256857-008-23F--UF2-DFR
267596-0017-003-23F-UF1
267596-0017-001-23F-FDPS

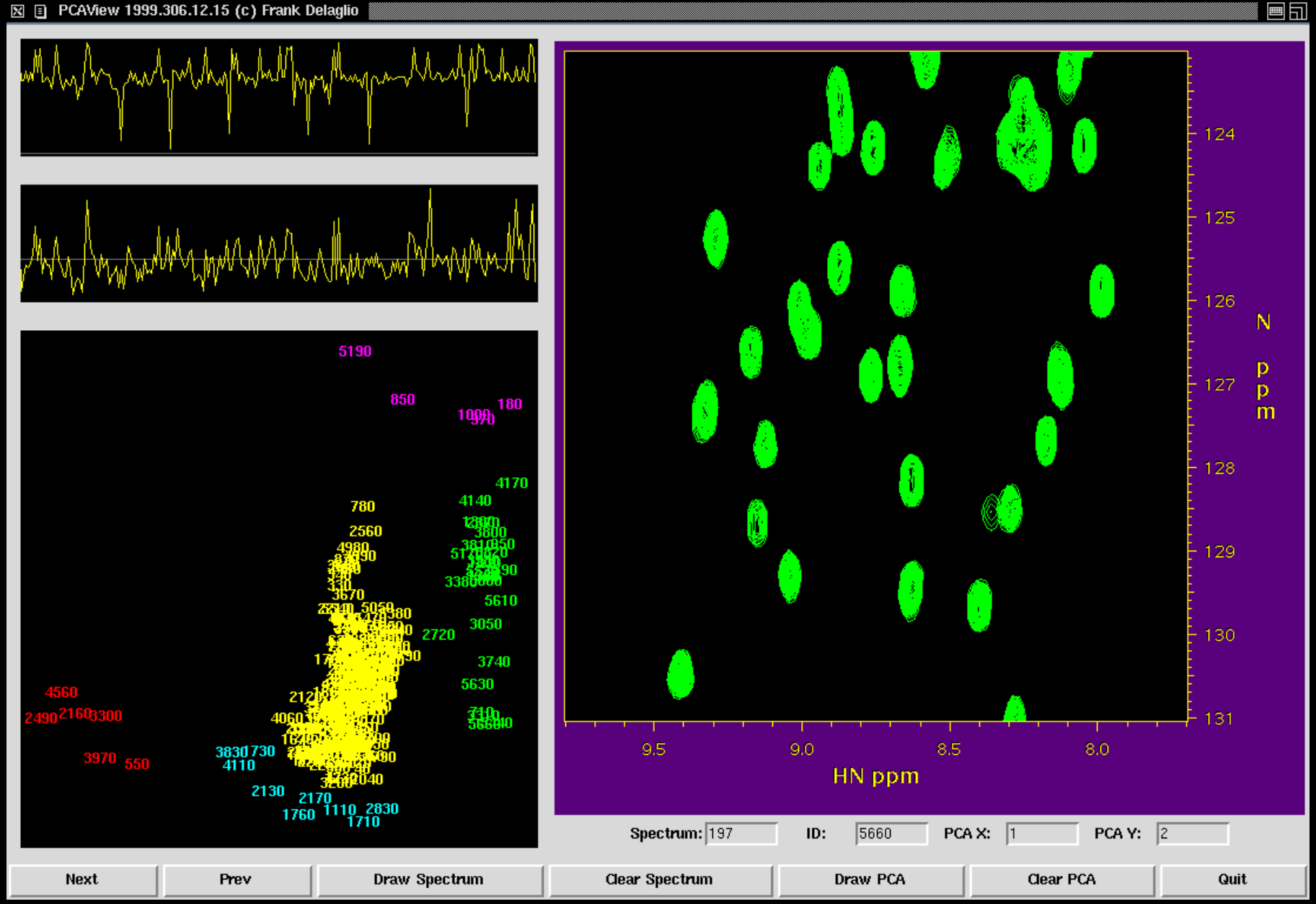

PANDAS
PANDAS
File: l015_APFsample_25_July1
ID:
15
Scale:
1.0
Color:
green
Color By:
Title
Use for WAMPAS
List By:
Title
File
--APF03-*
--APF04-*
--APF05-*
--APF06-*
07-03-APF15
07-04-APF16
07-04-APF17
07-04-APF18
07-04-APF19
07-04-APF20
07-04-APF24
07-04-APF25
07-04-APF26
07-04-APF27
07-04-APF28
07-04-APF29
07-04-APF30
07-04-APF31
07-04-APF33
07-04-APF34
07-02-APF03
07-04-APF04
07-04-APF05
Exit

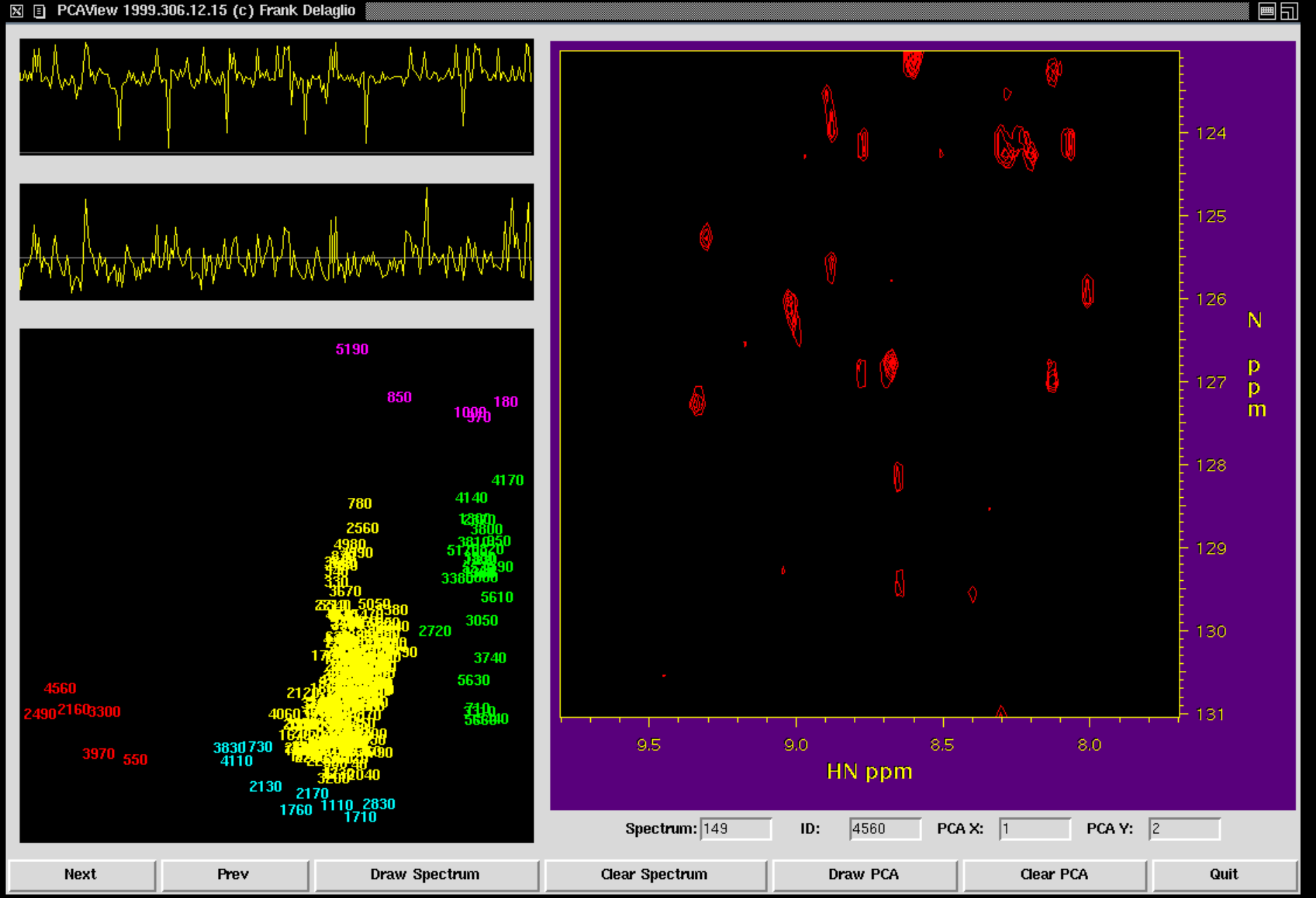

Initial Structure
Mis-Folded Structure
Refined Structures
Principal Component 2
Principal Component 3

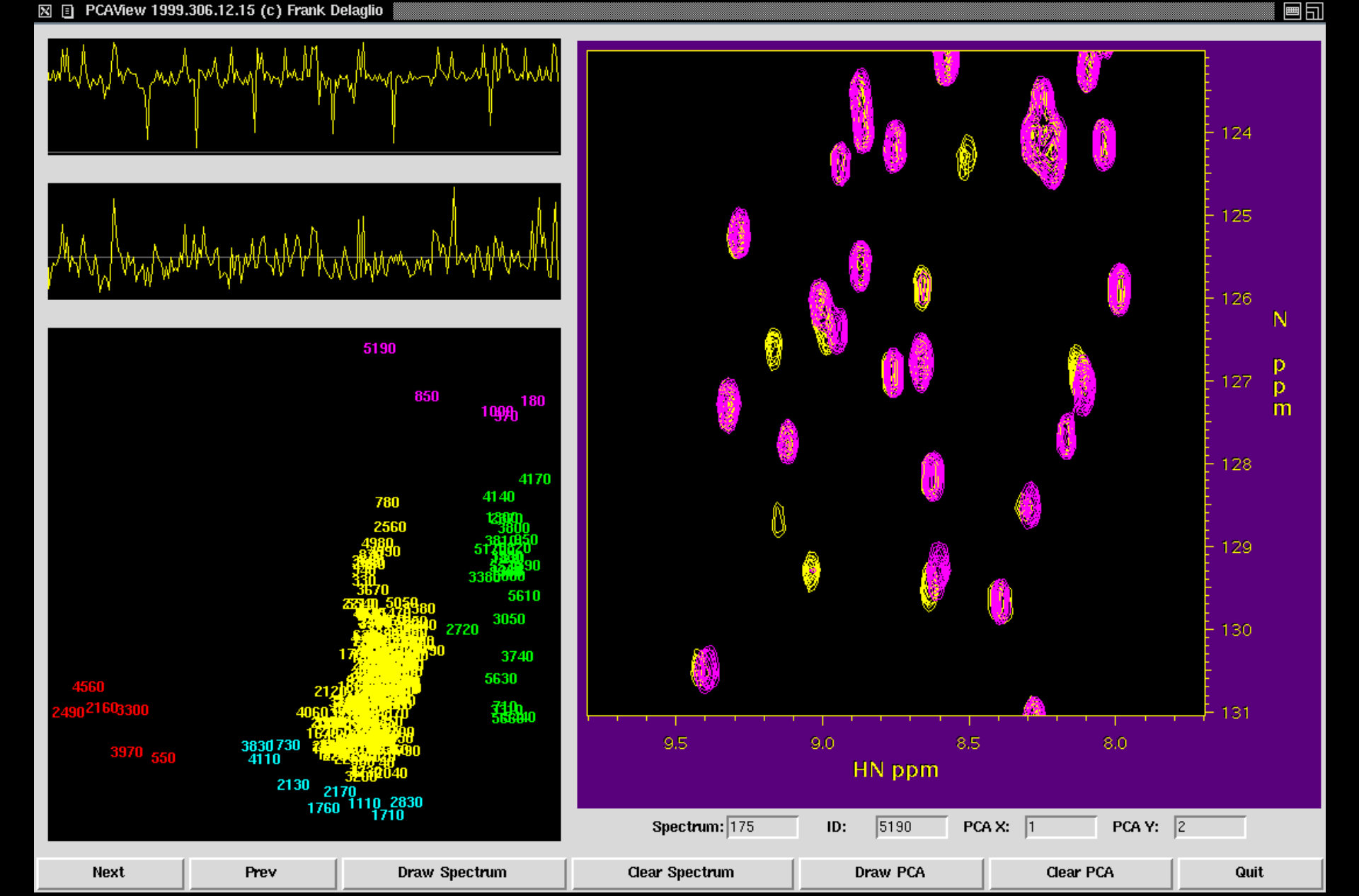
Correct Structure

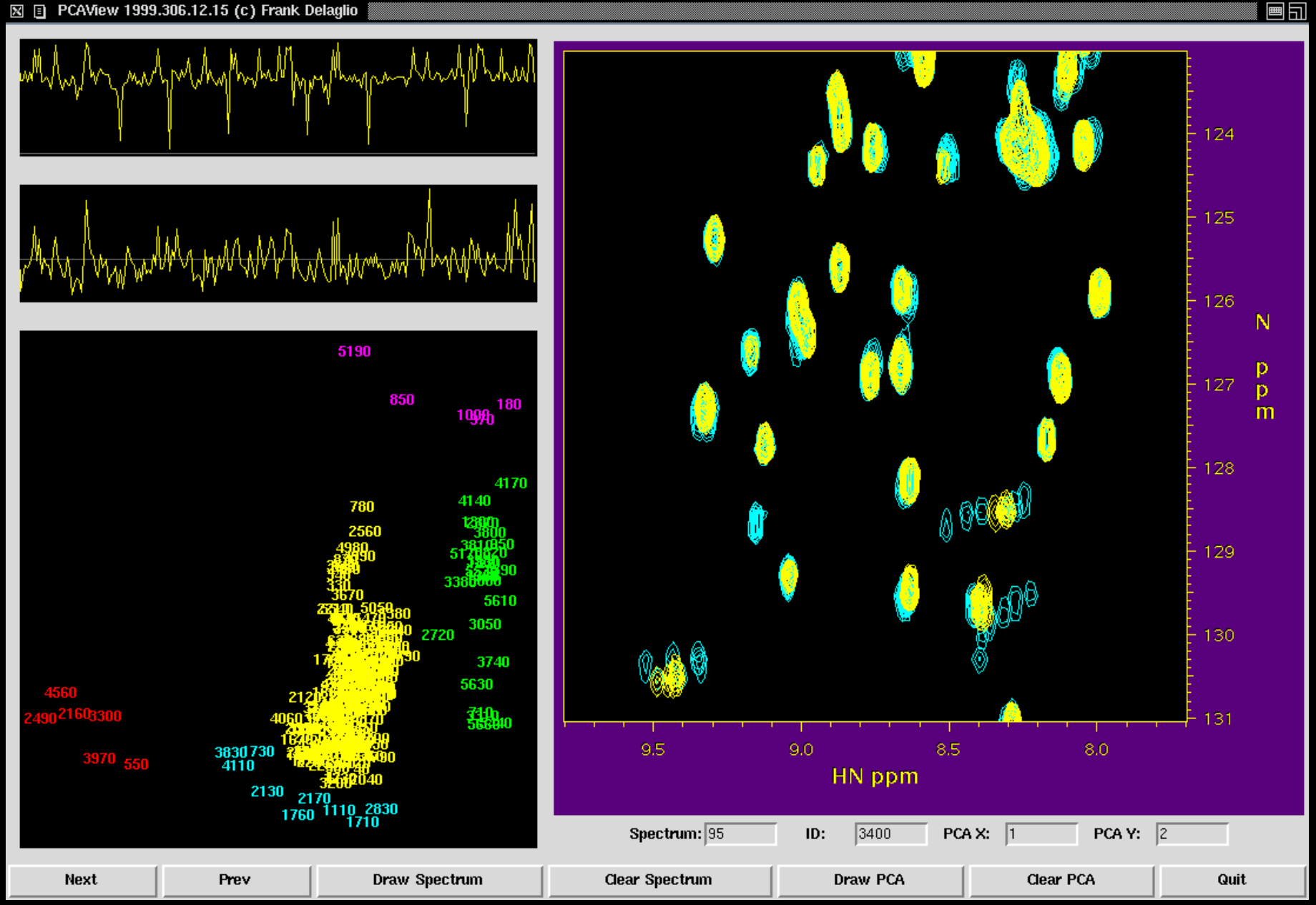
Mis-Folded Structure with
Secondary Structure Conserved



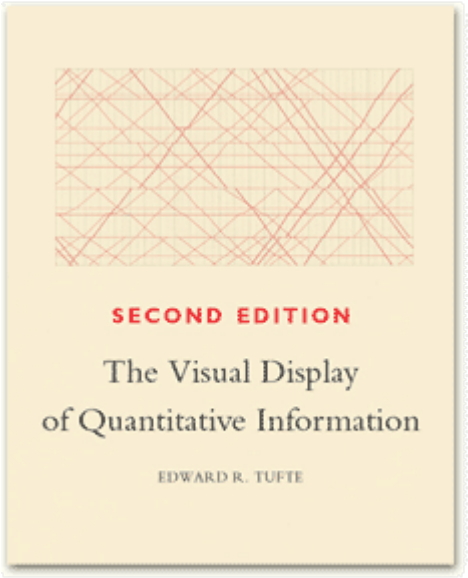

PCAView 1999.306.12.15 (c) Frank Delaglio
5190
850
180
4170
4140
780
2560
4980
3670
5610
3050
2720
3740
5630
4560
2490
2160
3300
3970
550
3830
730
4110
2130
2170
1760
1110
2830
1710
124
125
126
127
128
129
130
131
N
p
p
m
9.5
9.0
8.5
8.0
HN ppm
Spectrum: 197
ID: 5660
PCA X: 1
PCA Y: 2
Next
Prev
Draw Spectrum
Clear Spectrum
Draw PCA
Clear PCA
Quit

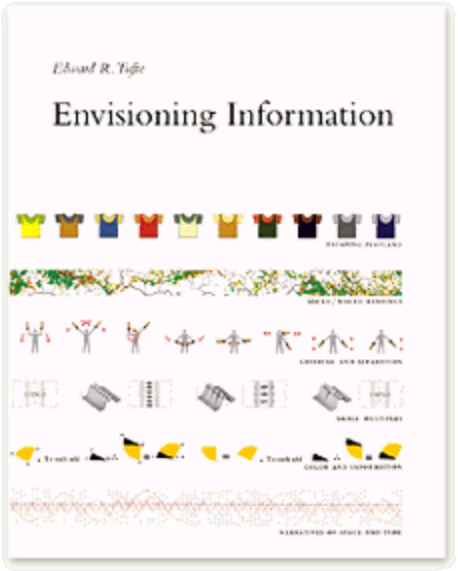

PCAView 1999.306.12.15 (c) Frank Delaglio
5190
850
180
4170
4140
780
2560
4980
3670
5610
3050
2720
3740
5630
4560
2490
2160
3300
3970
550
3830
730
4110
2130
2170
1760
1110
2830
1710
124
125
126
127
128
129
130
131
N
p
p
m
9.5
9.0
8.5
8.0
HN ppm
Spectrum: 149
ID: 4560
PCA X: 1
PCA Y: 2
Next
Prev
Draw Spectrum
Clear Spectrum
Draw PCA
Clear PCA
Quit

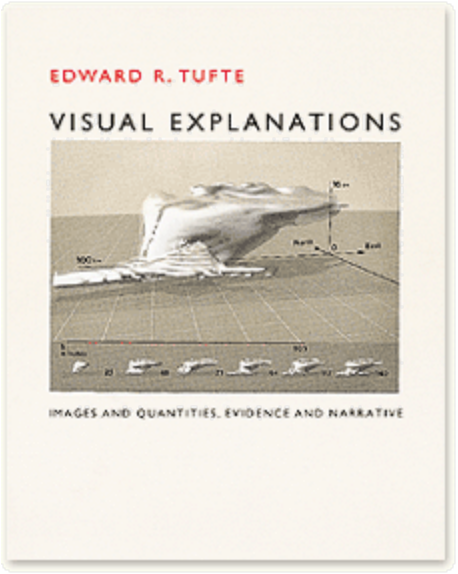

PCAView 1999.306.12.15 (c) Frank Delaglio
5190
850
180
4170
4140
780
2560
3800
3670
5610
2720
3050
3740
5630
4560
2490
2160
3300
3970
550
3830
730
4110
4060
2130
2170
1760
1110
2830
1710
124
125
126
127
128
129
130
131
N ppm
9.5
9.0
8.5
8.0
HN ppm
Spectrum: 175
ID: 5190
PCA X: 1
PCA Y: 2
Next
Prev
Draw Spectrum
Clear Spectrum
Draw PCA
Clear PCA
Quit


PCAView 1999.306.12.15 (c) Frank Delaglio
5190
850
180
4170
4140
780
2560
3800
3670
5610
3050
2720
3740
5630
4560
2490
2160
3300
3970
550
3830
730
4110
2130
2170
1760
1110
2830
1710
124
125
126
127
128
129
130
131
N ppm
9.5
9.0
8.5
8.0
HN ppm
Spectrum: 95
ID: 3400
PCA X: 1
PCA Y: 2
Next
Prev
Draw Spectrum
Clear Spectrum
Draw PCA
Clear PCA
Quit

Useful Graphics Strategies







Graphics Strategies of Edward Tufte

www.edwardtufte.com

Above all, Show the Data

Show Cause and Effect

Represent Data and Scale Faithfully

Maximize “Data Ink” and Data Density, Minimize “Chart Junk”

Shrink Graphics - Integrate Text, Values, and Graphics

Be Multivariate

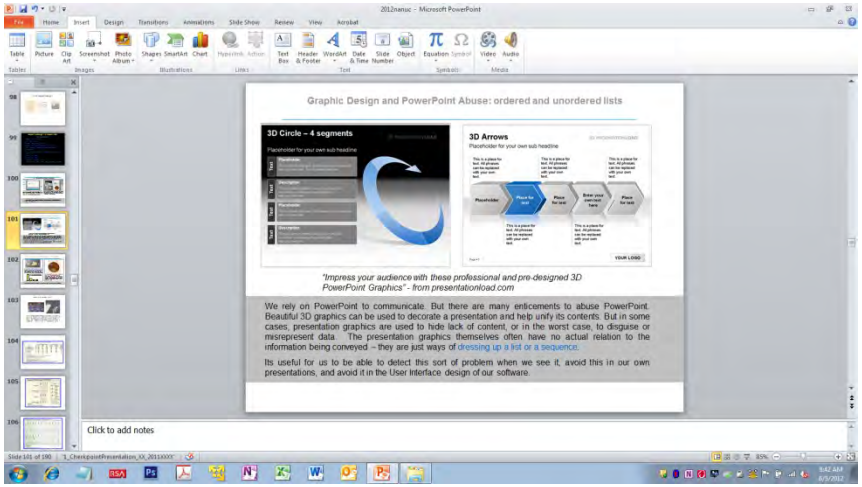
Use Layers – Use Macro and Micro Interpretations - Clarify by Adding Detail

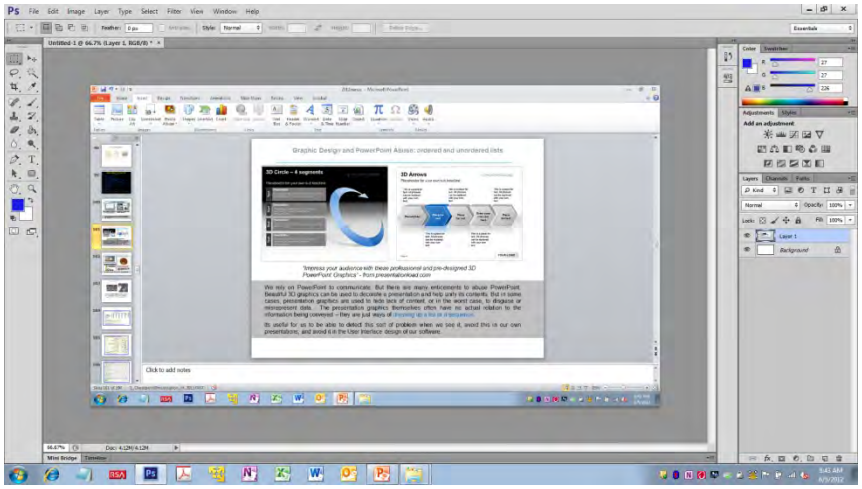
Conserve Color Space

Use Small Multiples

Find Ways to Show All of the Data

Treat Design as a Solved Problem, then Find the Best Examples



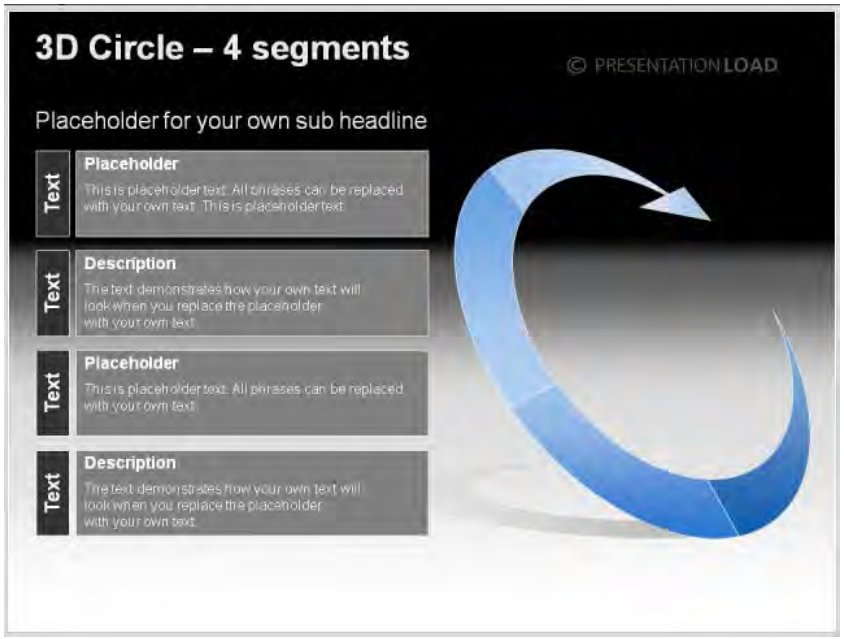


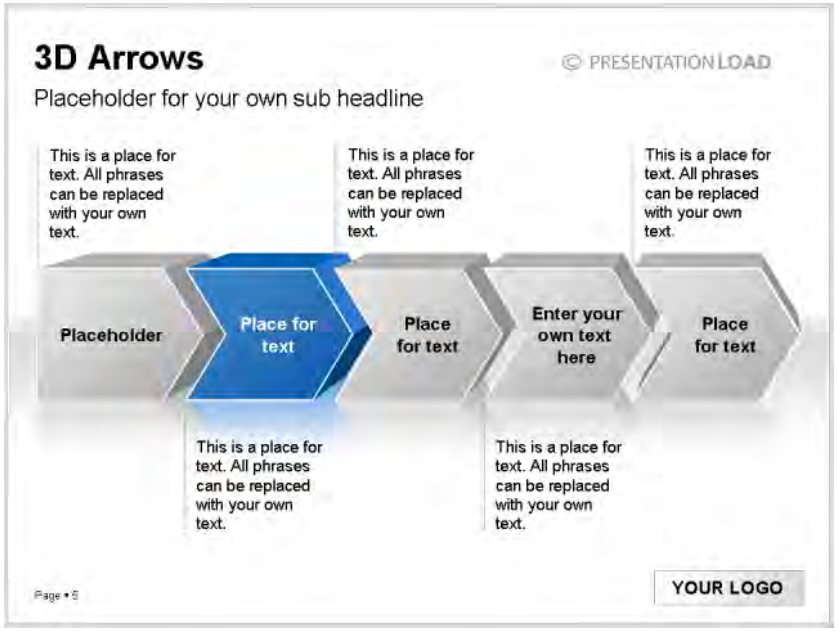
Microsoft PowerPoint "with new and improved tools for video and photo editing, dramatic new transitions, and realistic animation, you can add polish to presentations that will captivate your audience." - *from microsoft.com*

Adobe Photoshop "redefines digital imaging with powerful new photography tools and breakthrough capabilities for complex image selections, realistic painting, and intelligent retouching." - *from adobe.com*

In the content-creation arena, amazon.com lists Microsoft Office and Adobe Photoshop as highest-sellers. It could be claimed, Photoshop is about the content, while PowerPoint is often about the scaffold …

Graphic Design and PowerPoint Abuse: ordered and unordered lists



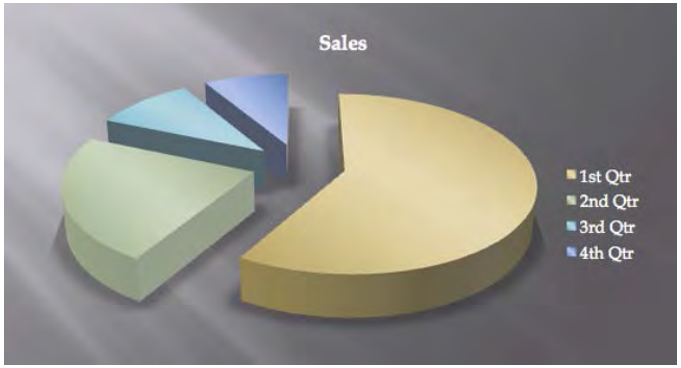


"Impress your audience with these professional and pre-designed 3D PowerPoint Graphics" - from presentationload.com

We rely on PowerPoint to communicate. But there are many enticements to abuse PowerPoint. Beautiful 3D graphics can be used to decorate a presentation and help unify its contents. But in some cases, presentation graphics are used to hide lack of content, or in the worst case, to disguise or misrepresent data. The presentation graphics themselves often have no actual relation to the information being conveyed – they are just ways of dressing up a list or a sequence.

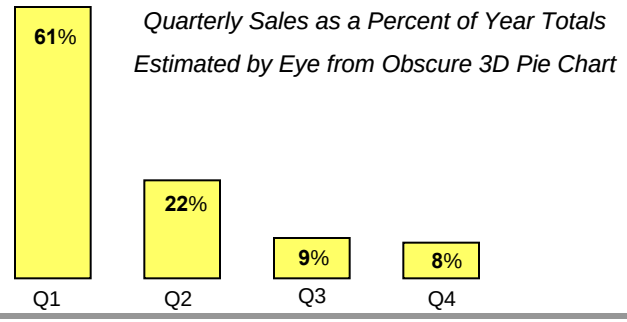
Its useful for us to be able to detect this sort of problem when we see it, avoid this in our own presentations, and avoid it in the User Interface design of our software.

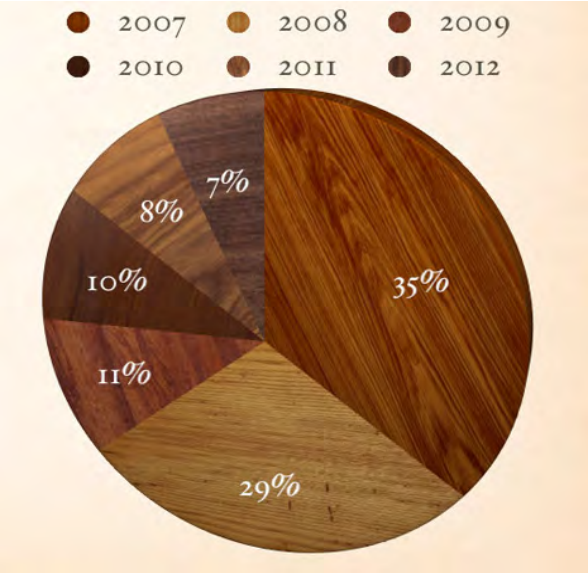
Avoiding inappropriate presentation modes and unneeded 3D effects



Microsoft PowerPoint – "as you can see, the blue-greenish quarter was about the same as the blue one."

There are no numbers given, and the trend over time, which shows falling sales, is hidden. Also, the 3D perspective might actually distort the apparent values:



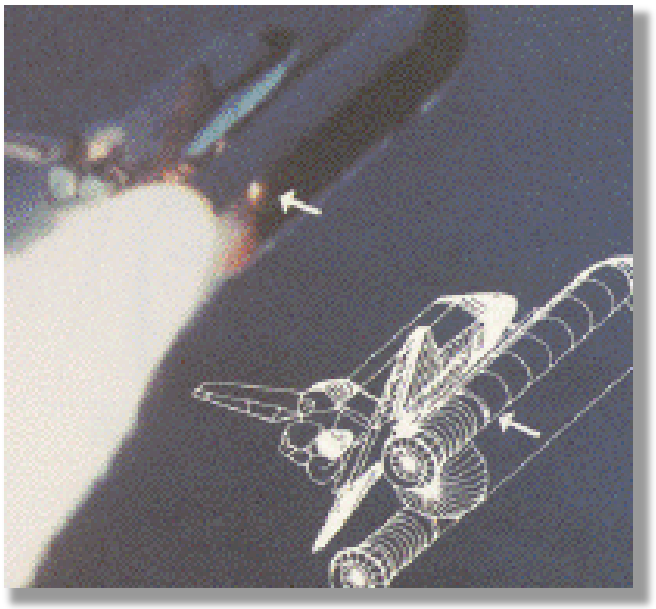


Apple Keynote – "as you can see, our maple year was only 1% different than our teak year."

The segments are at least nicely labeled with numbers, but the wood textures, while really cool, are impossible to decipher. Likewise, the trend with time is hidden, and 3D perspective potentially distorts the apparent values.

The wrong graphical paradigm can make data hard to interpret, or even misleading.

Debate on the Challenger Launch



Morton Thiokol engineers debated the problem of O-ring failure due to low temperature for several hours the night before the launch, and made the company's only no-launch request in 12 years. Their presentation of evidence did not convince NASA management. The shuttle blew up 73 seconds after ignition.

HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

Joint	Sqm No.	Cross Sectional View: Erosion Depth (in.)	Cross Sectional View: Perimeter Affected (deg)	Cross Sectional View: Nominal Dia. (in.)	Top View: Length of Max Erosion (in.)	Top View: Total Heat Affected Length (in.)	Clocking location (deg)
61A LH Center Field**	22A	None	None	0.280	None	None	36°-66°
61A LH ~~Center~~ FIELD**	22A	NONE	NONE	0.280	NONE	NONE	334°-18°
51C LH Forward Field***	15A	0.010	154.0	0.280	4.25	5.25	353
51C RH Center Field (prim)***	15B	0.038	330.0	0.280	12.50	58.75	354
51C RH Center Field (sec)***	15B	None	45.0	0.280	None	21.50	354
41C RH Forward Field	17B	0.028	112.0	0.280	3.00	None	275
41C LH Aft Field*	11A	None	None	0.280	None	None	--
41C LH Forward Field	10A	0.040	217.0	0.280	3.00	14.50	353
51C-2 RH Aft Field	2B	0.053	115.0	0.280	--	--	80

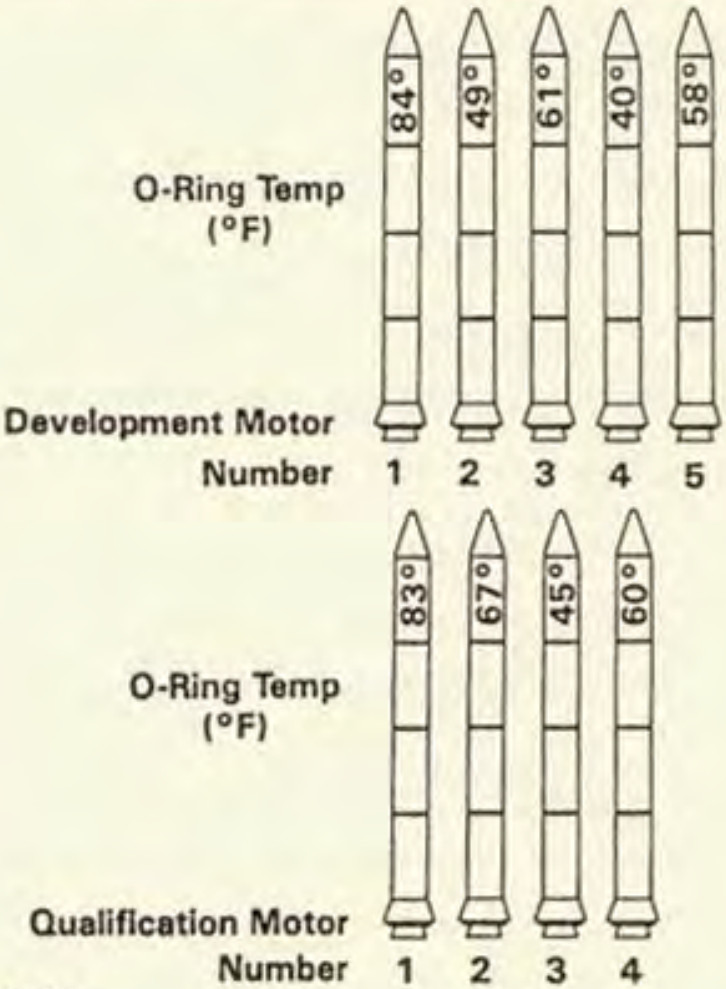
*Hot gas path detected in putty. Indication of heat on O-ring, but no damage.
**Soot behind primary O-ring.
***Soot behind primary O-ring, heat affected secondary O-ring.

Clocking location of leak check port = 0 deg.

OTHER SRM-15 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY AND NO SOOT NEAR OR BEYOND THE PRIMARY O-RING.

SRM-22 FORWARD FIELD JOINT HAD PUTTY PATH TO PRIMARY O-RING, BUT NO O-RING EROSION AND NO SOOT BLOWBY. OTHER SRM-22 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY.

History of O-Ring Damage in Field Joints



Code

S = Heating of Secondary O-Ring

B = Primary O-Ring Blowby

E = Primary O-Ring Erosion

H = Heating of Primary O-Ring

= No Damage

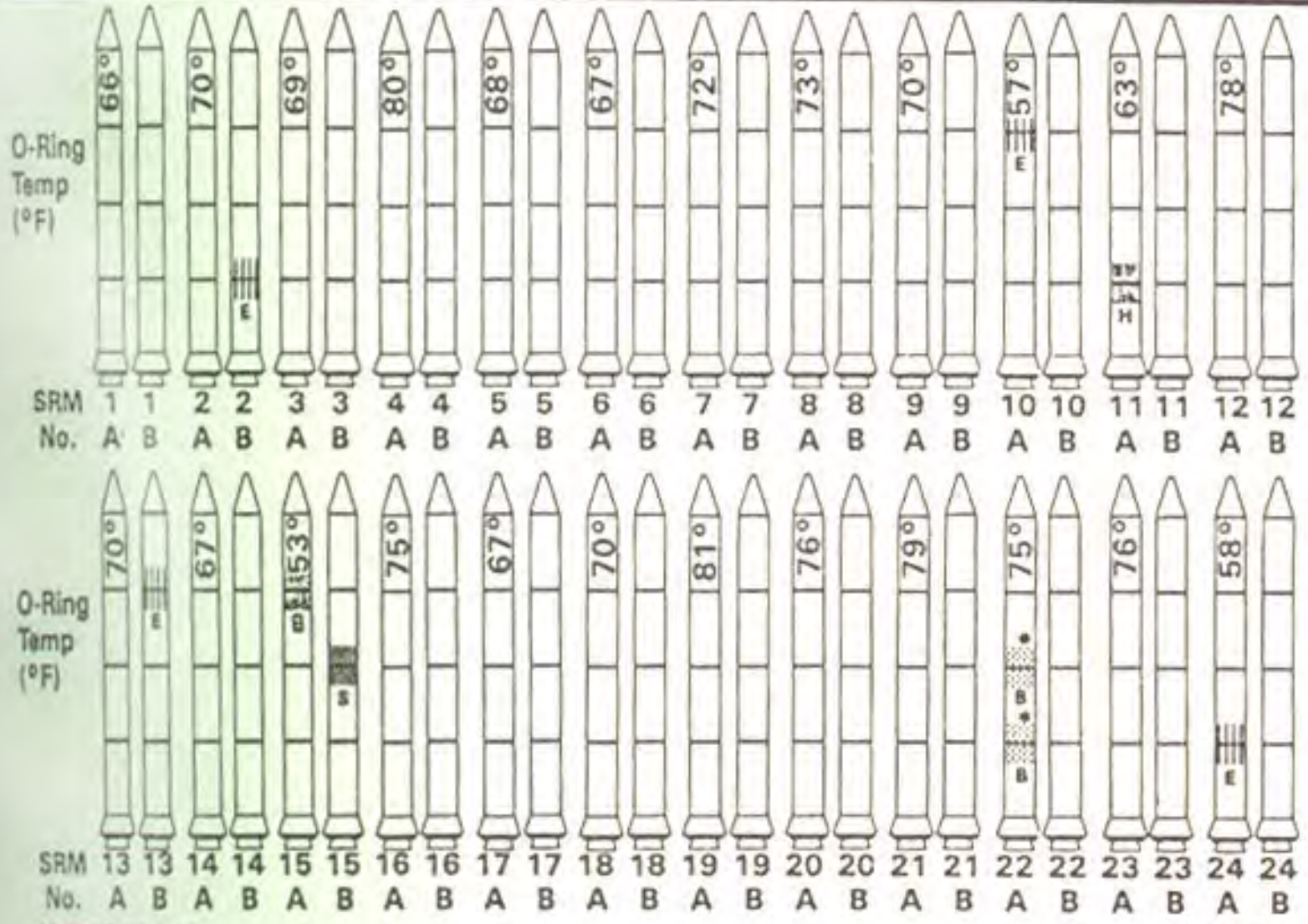
STATIC TEST MOTORS

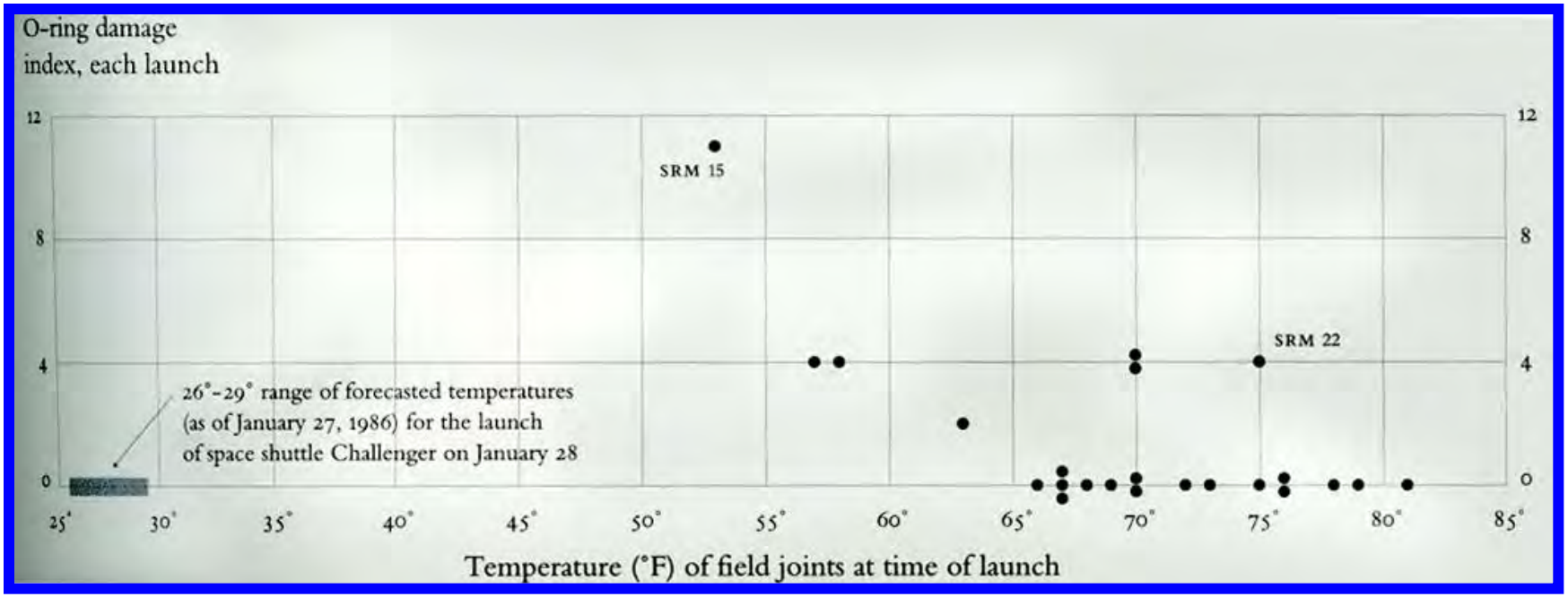
- HORIZONTAL ASSEMBLY
- SOME PUTTY REPAIRED

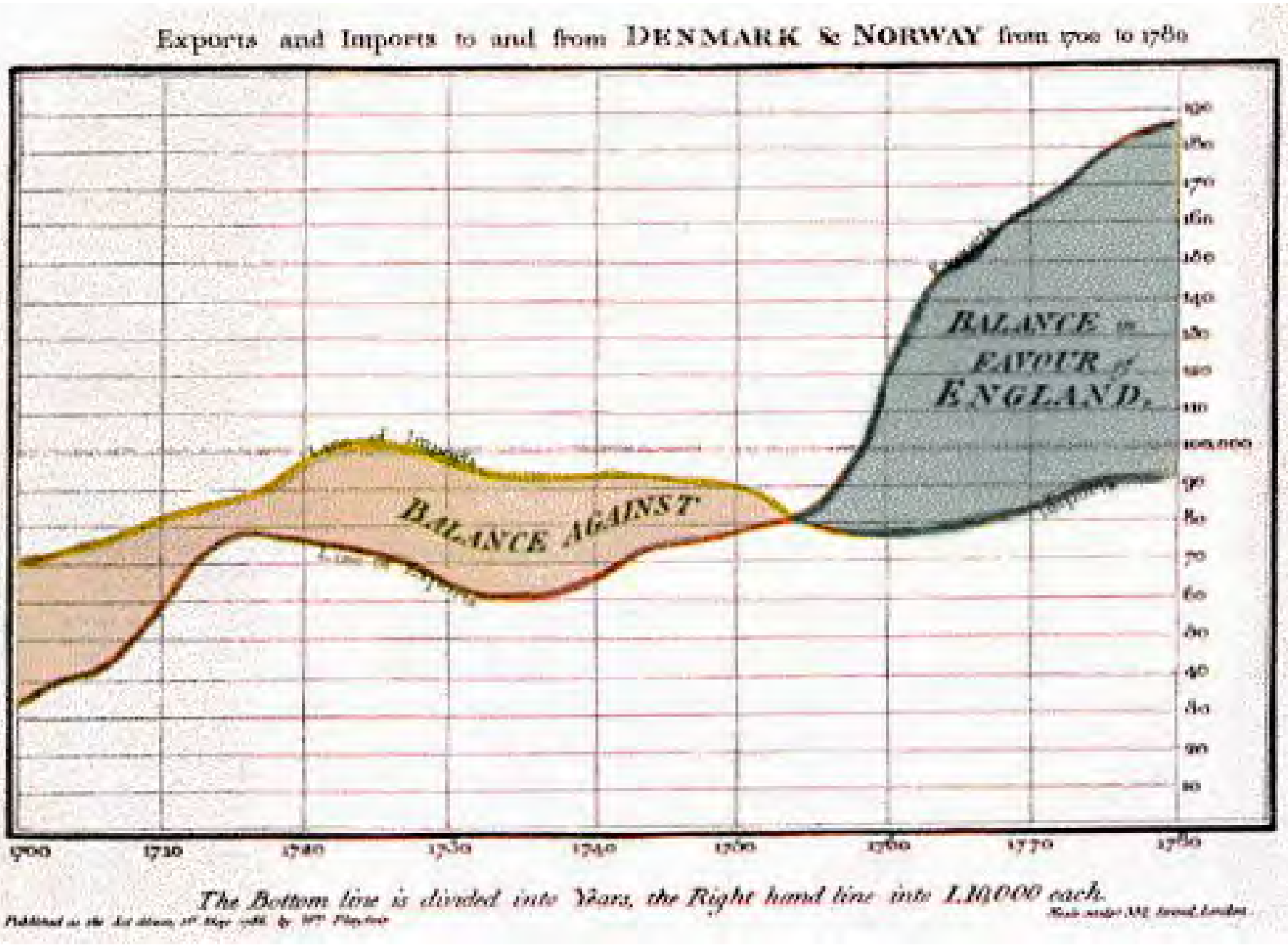
MORTON THIOKOL, INC.
Wasatch Operations

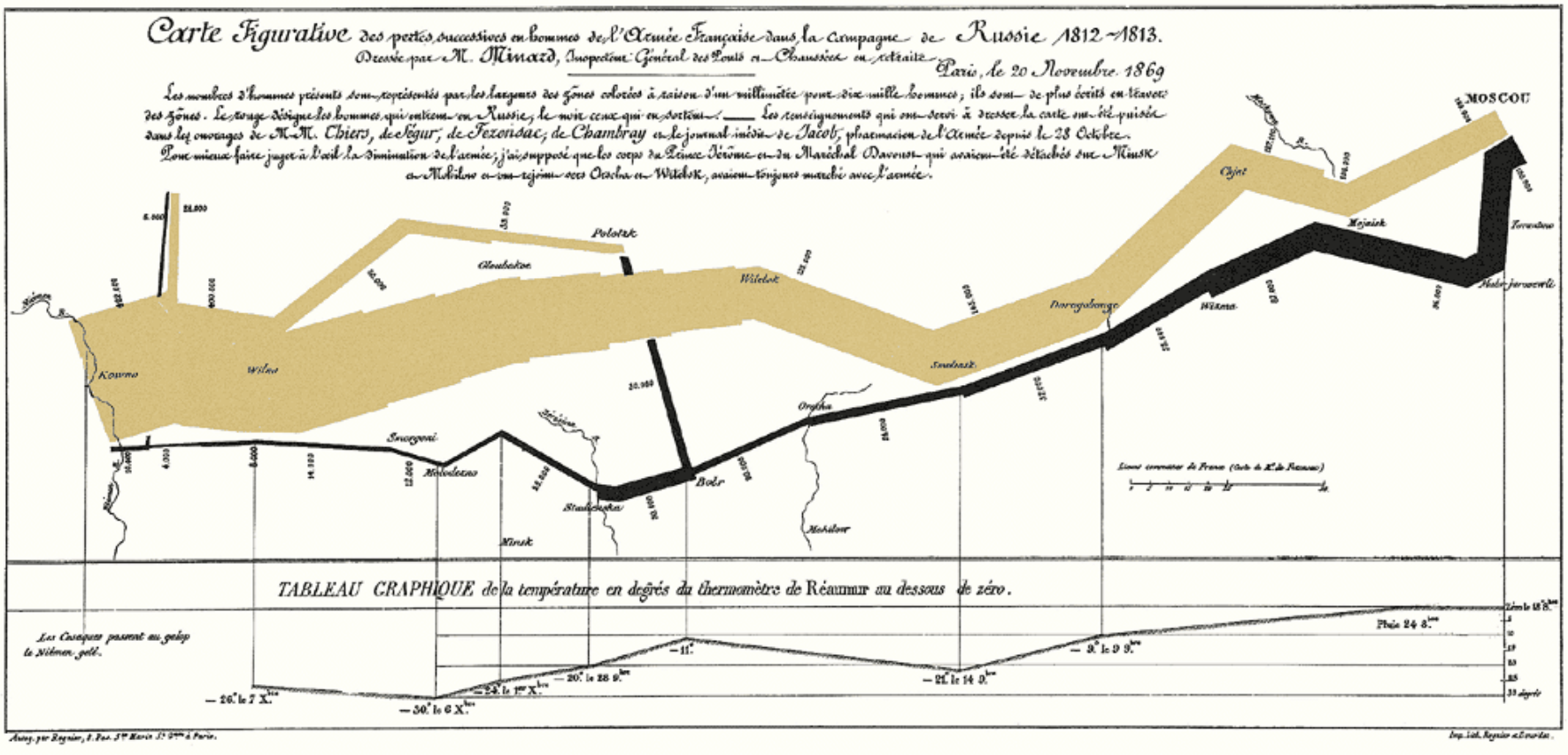
INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION

History of O-Ring Damage in Field Joints (Cont)

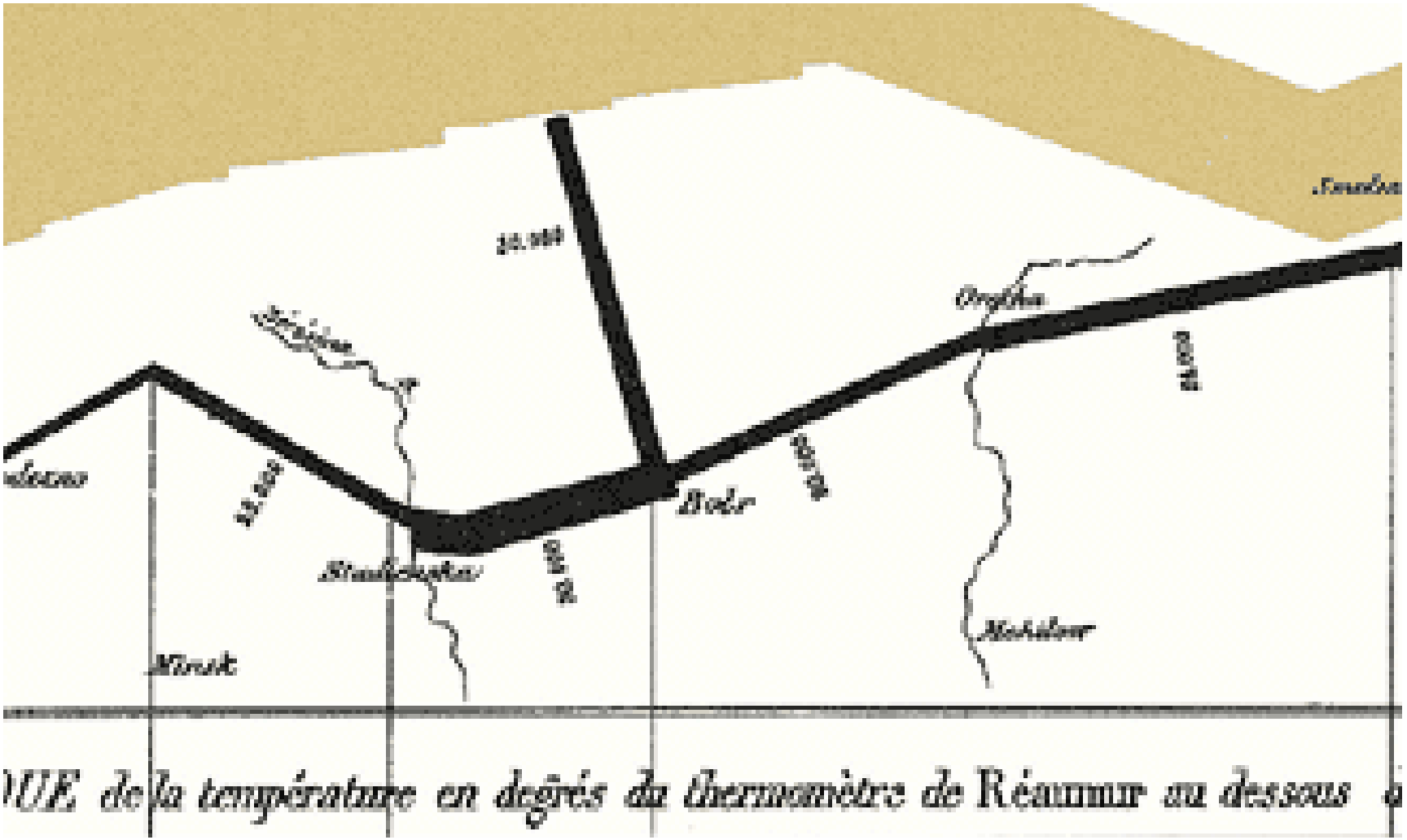



O-ring damage
index, each launch
12
8
4
0
SRM 15
SRM 22
26°-29° range of forecasted temperatures
(as of January 27, 1986) for the launch
of space shuttle Challenger on January 28
25°
30°
35°
40°
45°
50°
55°
60°
65°
70°
75°
80°
85°
Temperature (°F) of field joints at time of launch

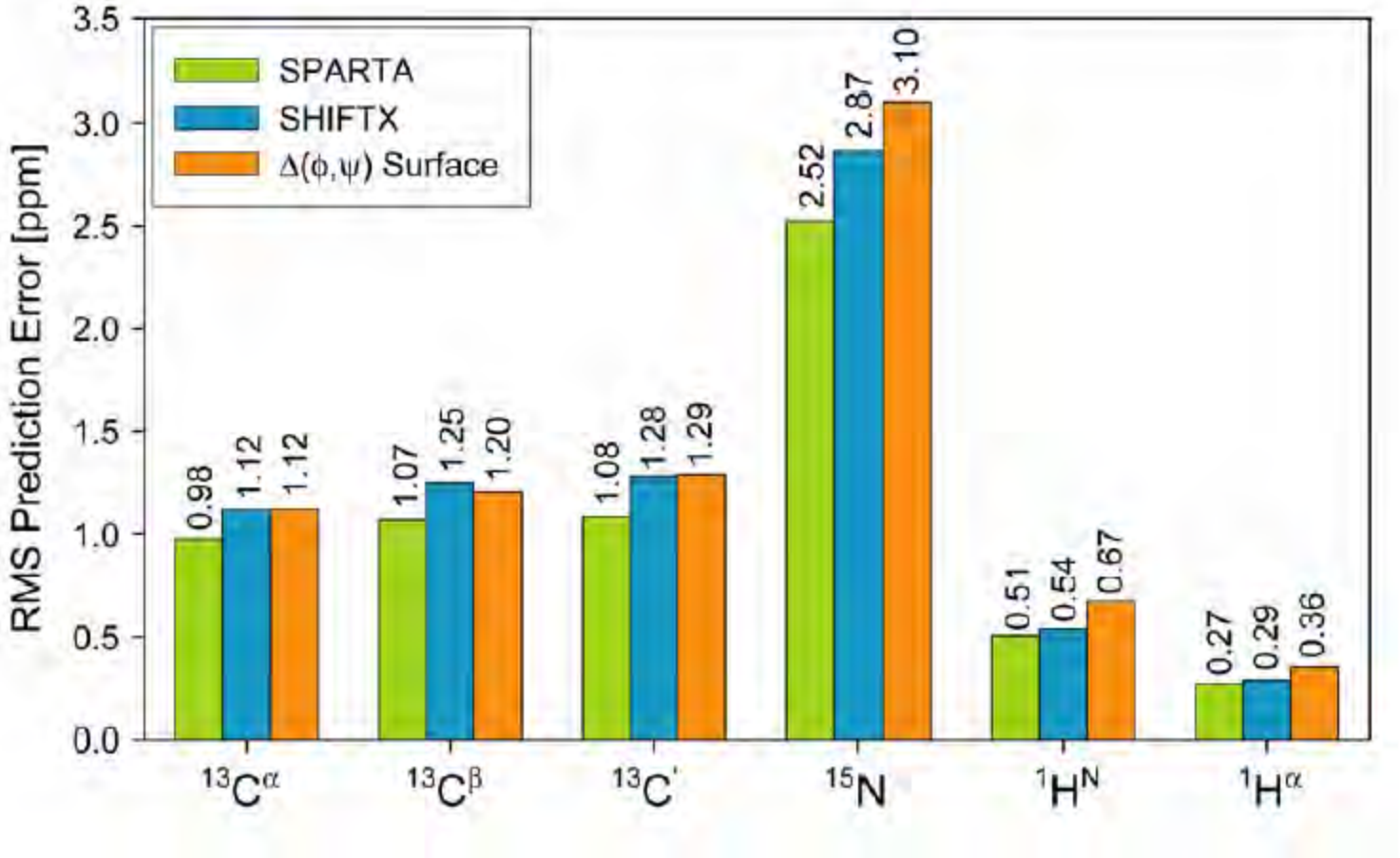

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780
BALANCE IN FAVOUR of ENGLAND.
BALANCE AGAINST
190
180
170
160
150
140
130
120
110
100,000
90
80
70
60
50
40
30
20
10
1700
1710
1720
1730
1740
1750
1760
1770
1780
The Bottom line is divided into Years, the Right hand line into £100000 each.
Published as the Act directs, 1st May 1786, by Wm Playfair
Neele sculpt 352 Strand London



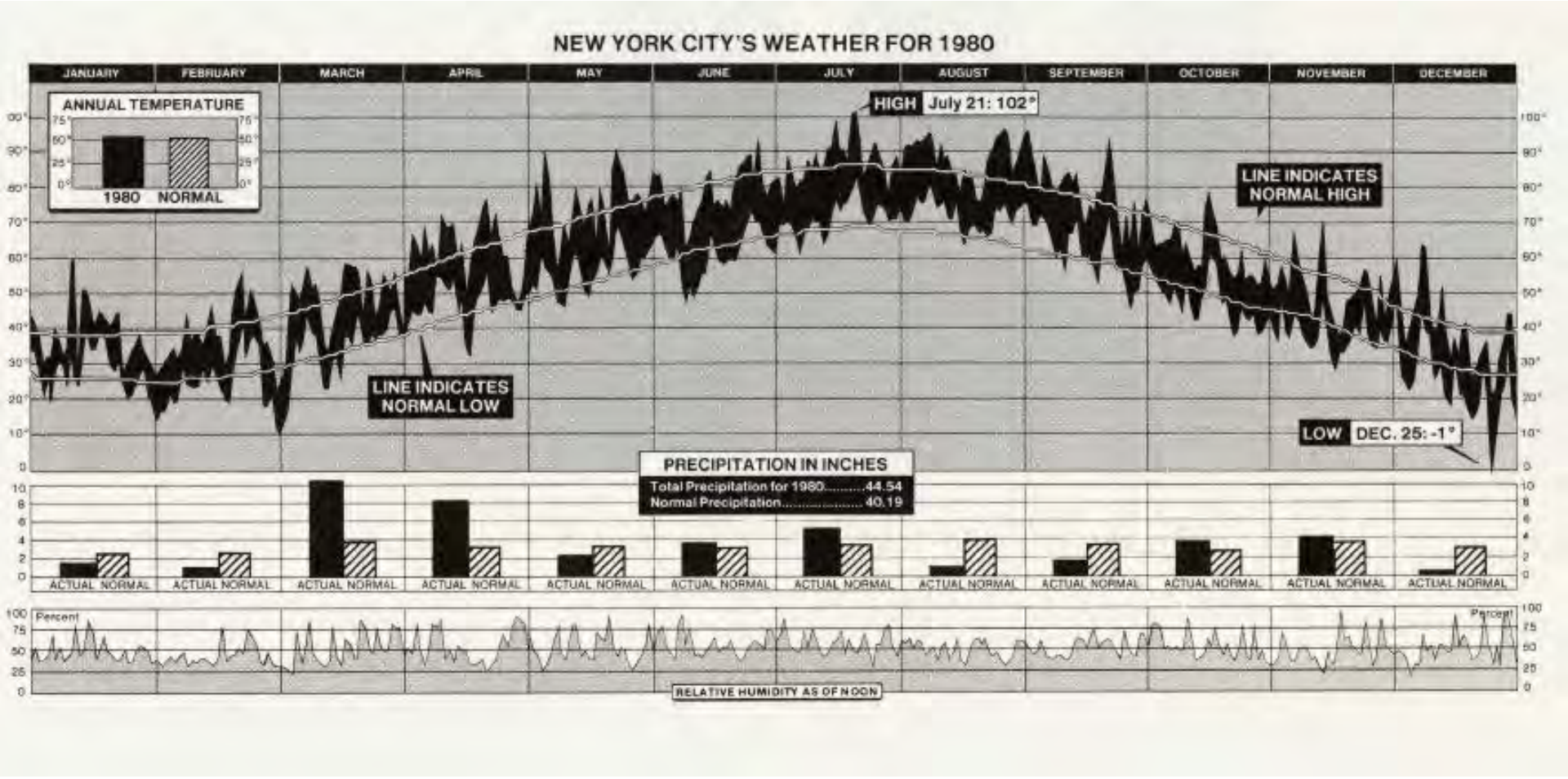
Napolean's Route: 422,000 Men to 10,000 Men, Five Dimensions



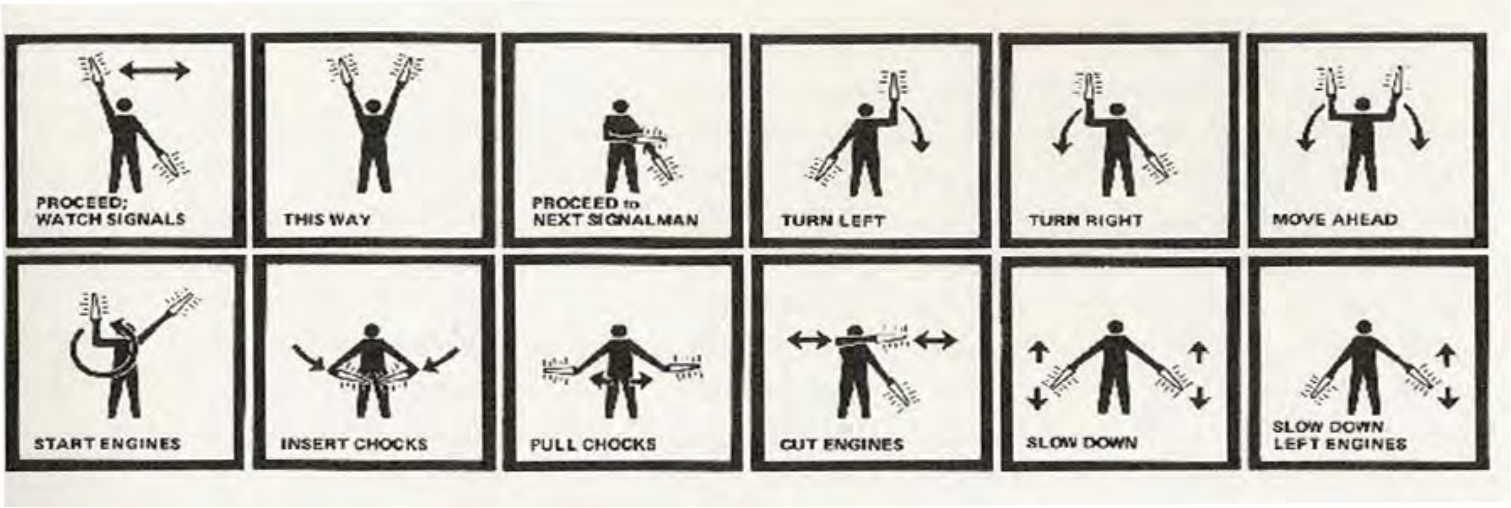
VUE de la température en degrés du thermomètre de Réaumur au dessous

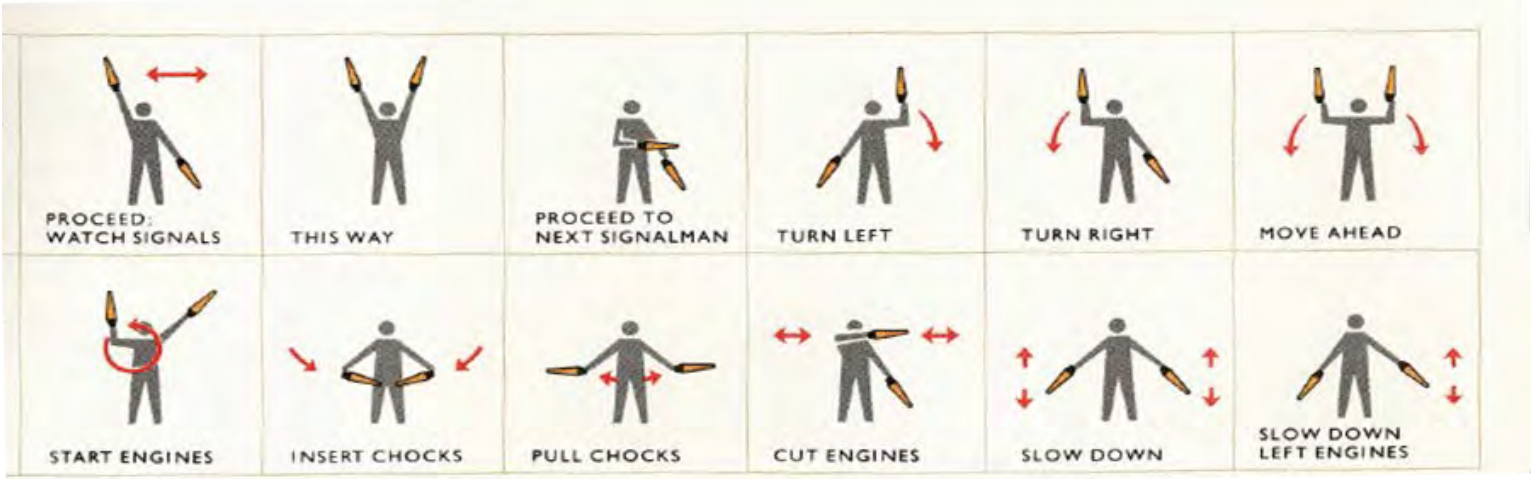


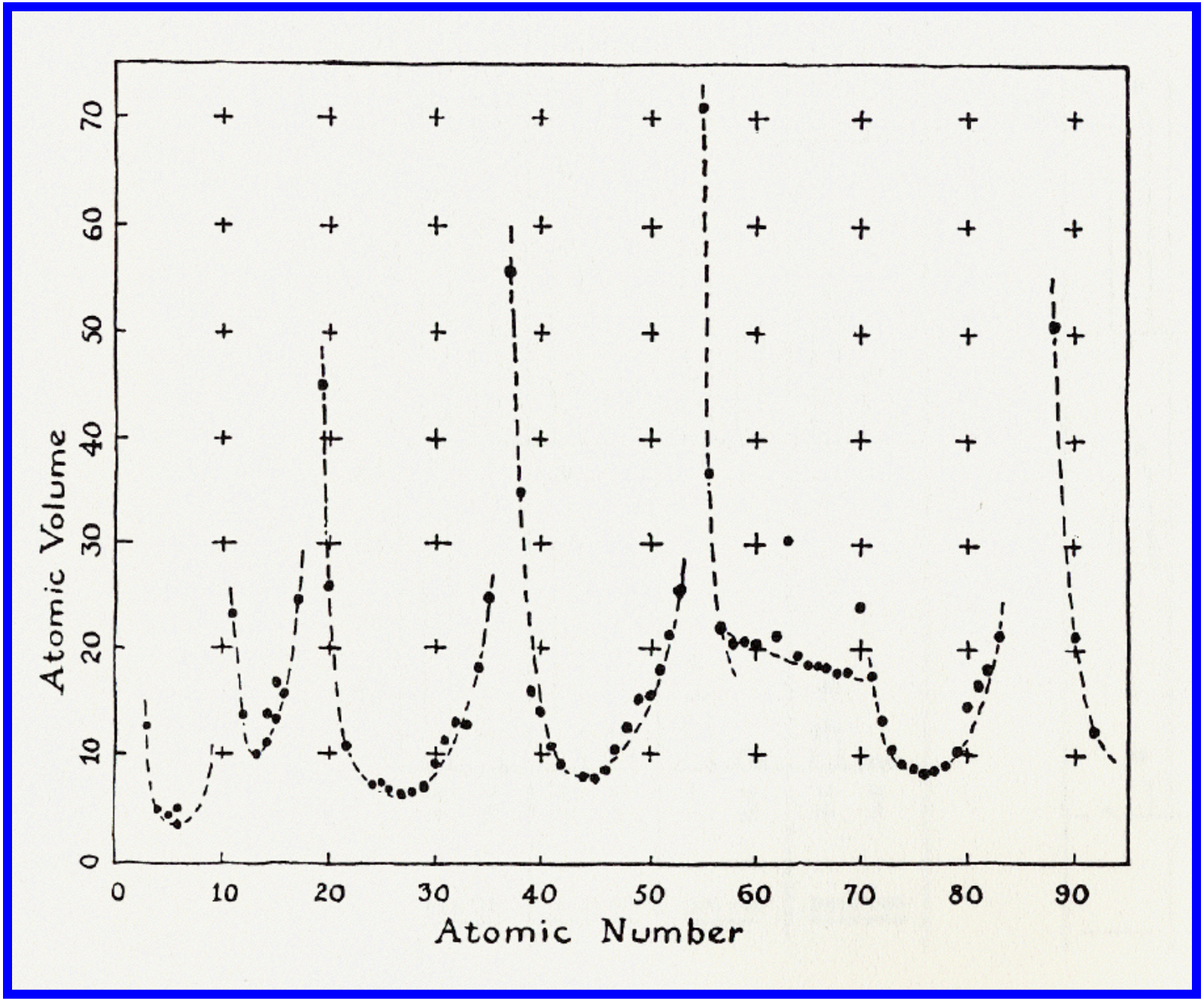
Bax Group Figure: 18 values

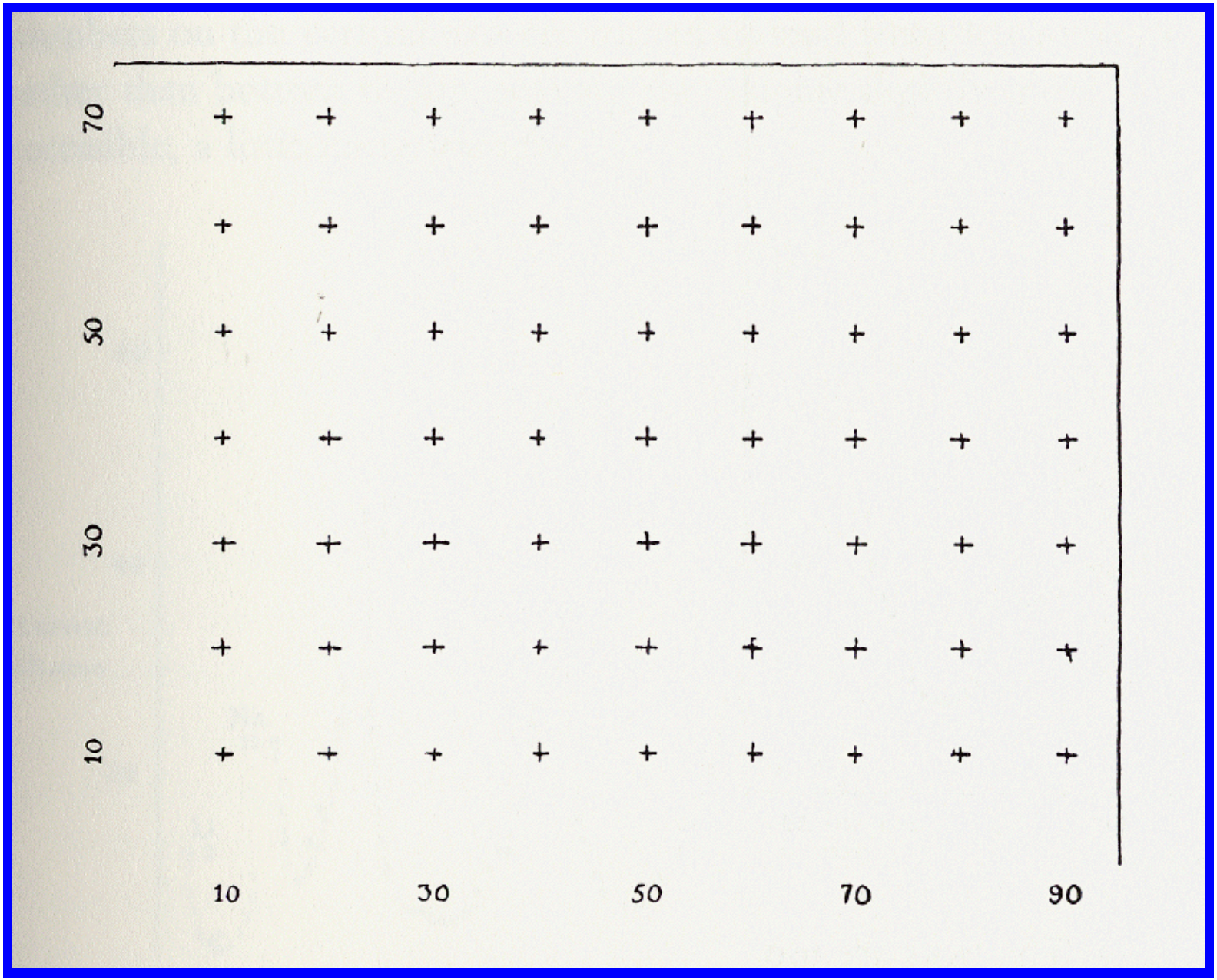


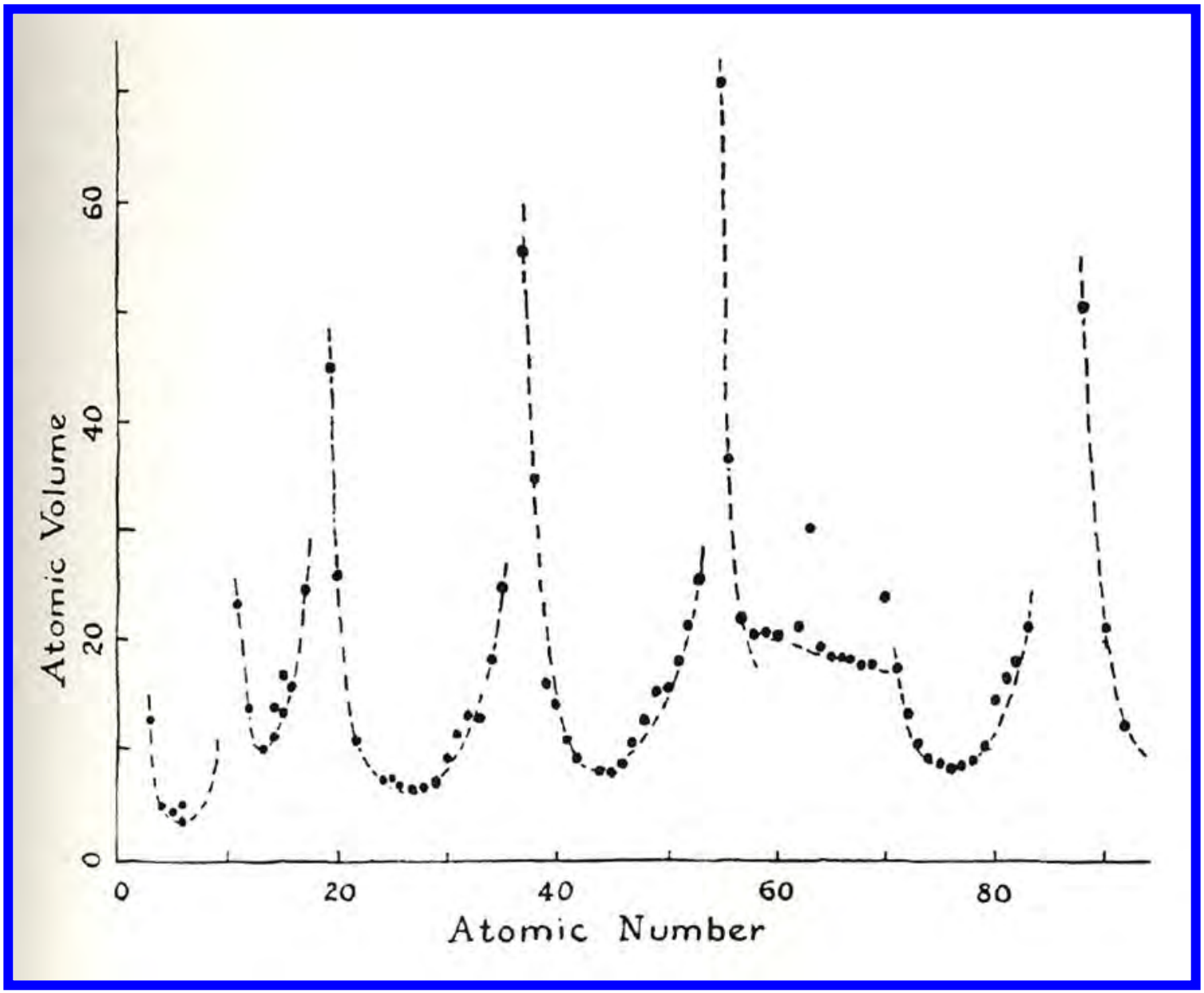
Weather Statistics: 1,800+ Values, Four Variables, Notations

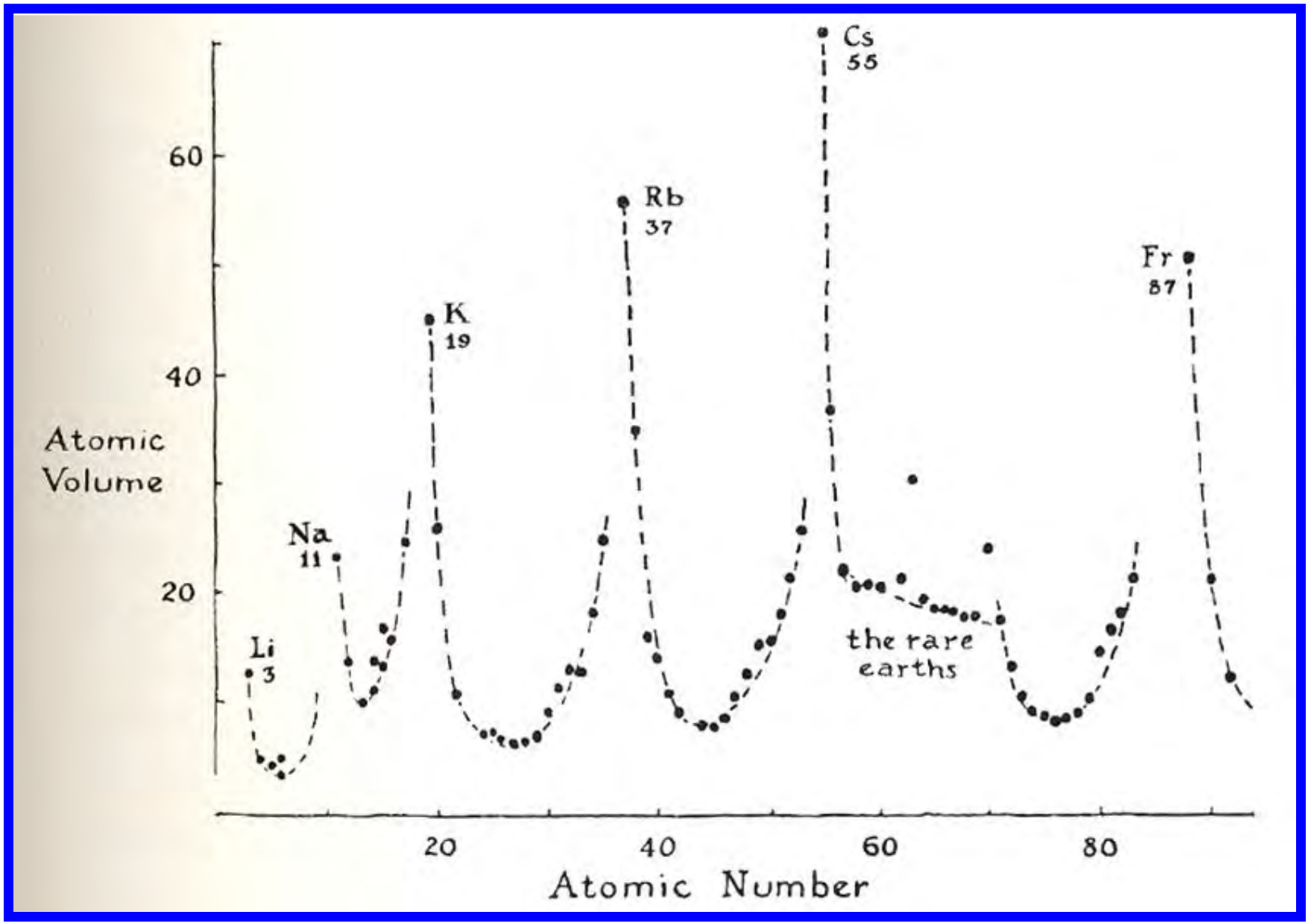

PROCEED;
WATCH SIGNALS
THIS WAY
PROCEED to
NEXT SIGNALMAN
TURN LEFT
TURN RIGHT
MOVE AHEAD
START ENGINES
INSERT CHOCKS
PULL CHOCKS
CUT ENGINES
SLOW DOWN
SLOW DOWN
LEFT ENGINES

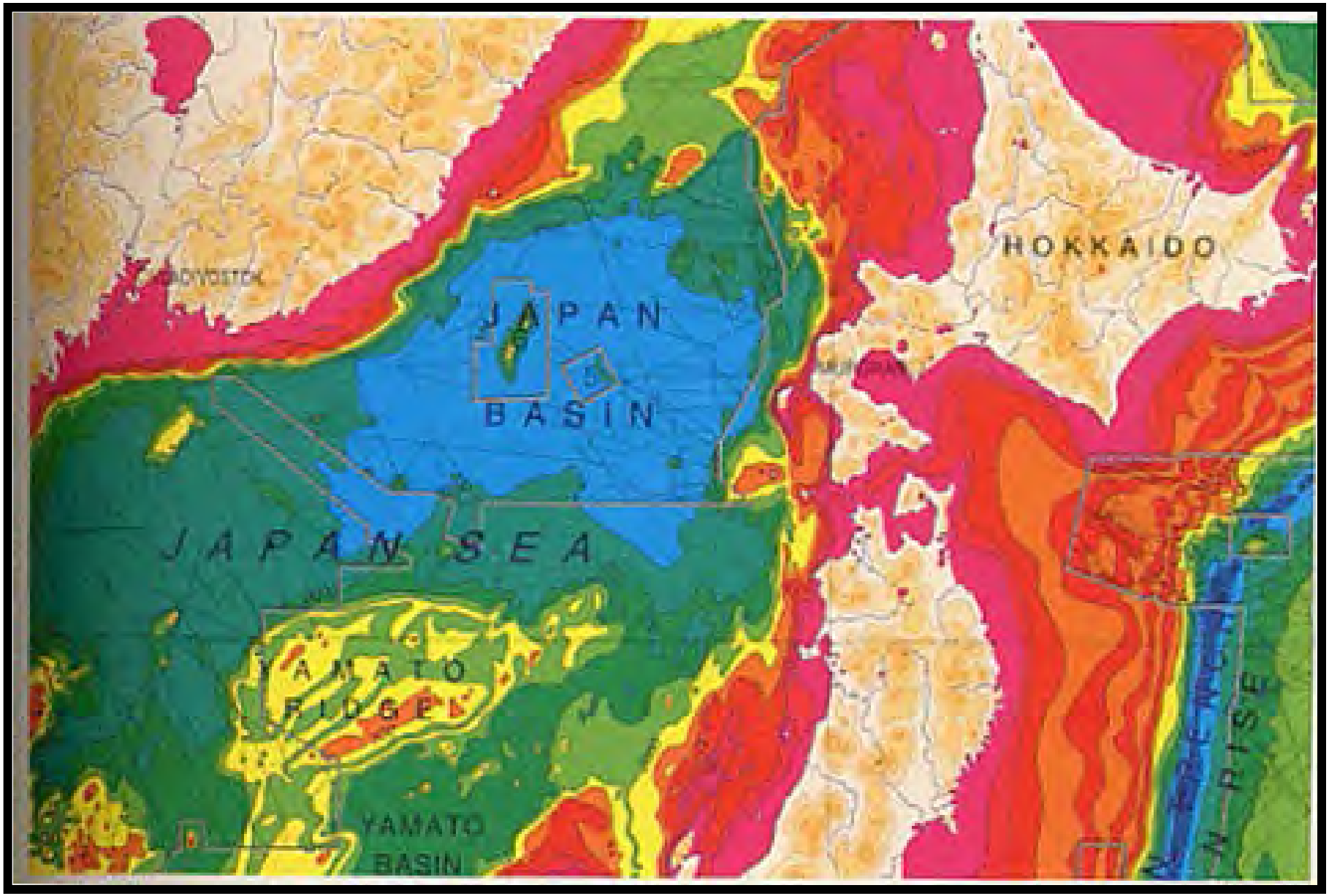

PROCEED:
WATCH SIGNALS
THIS WAY
PROCEED TO
NEXT SIGNALMAN
TURN LEFT
TURN RIGHT
MOVE AHEAD
START ENGINES
INSERT CHOCKS
PULL CHOCKS
CUT ENGINES
SLOW DOWN
SLOW DOWN
LEFT ENGINES

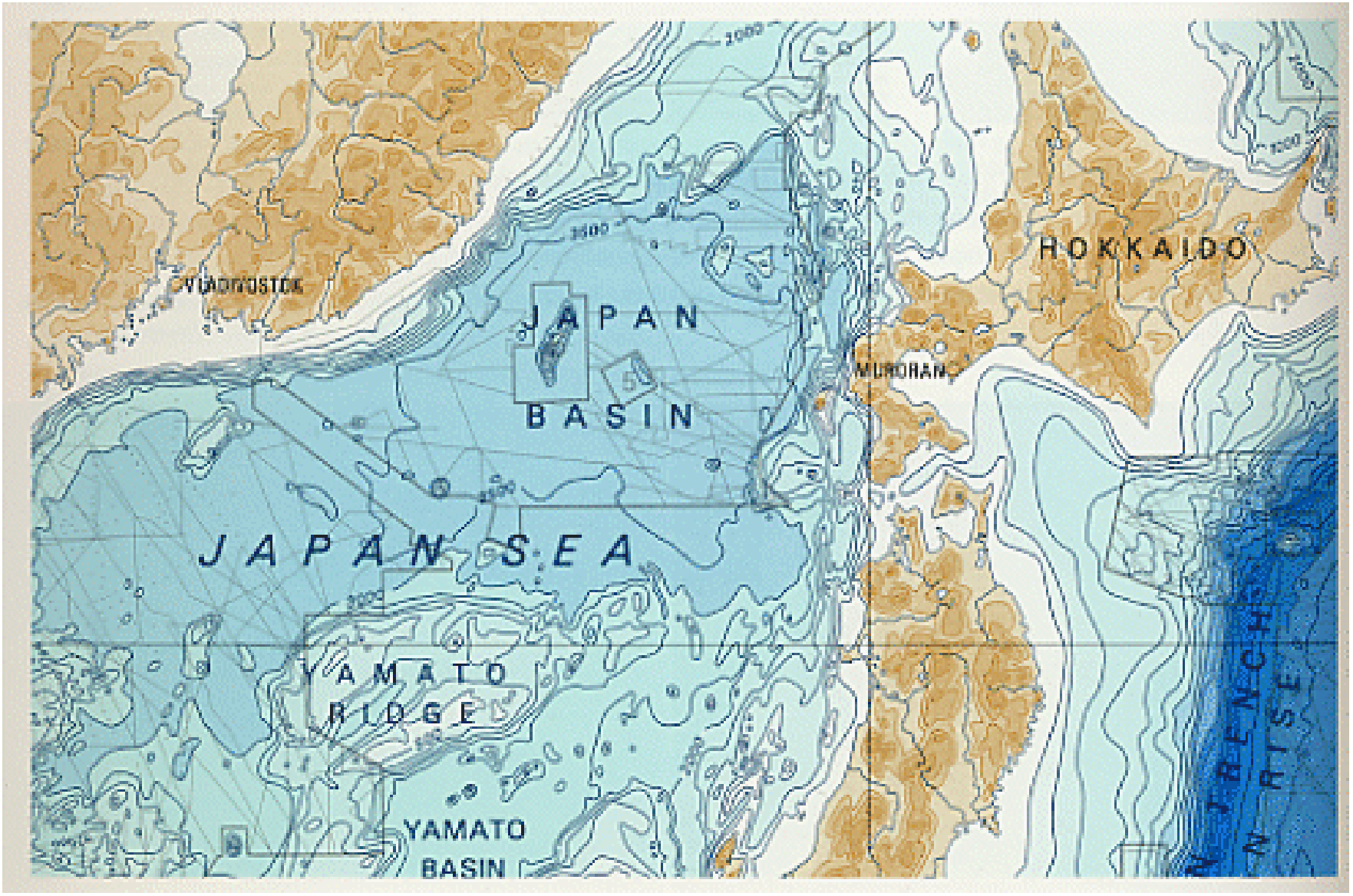

Atomic Volume
0
10
20
30
40
50
60
70
0
10
20
30
40
50
60
70
80
90
Atomic Number

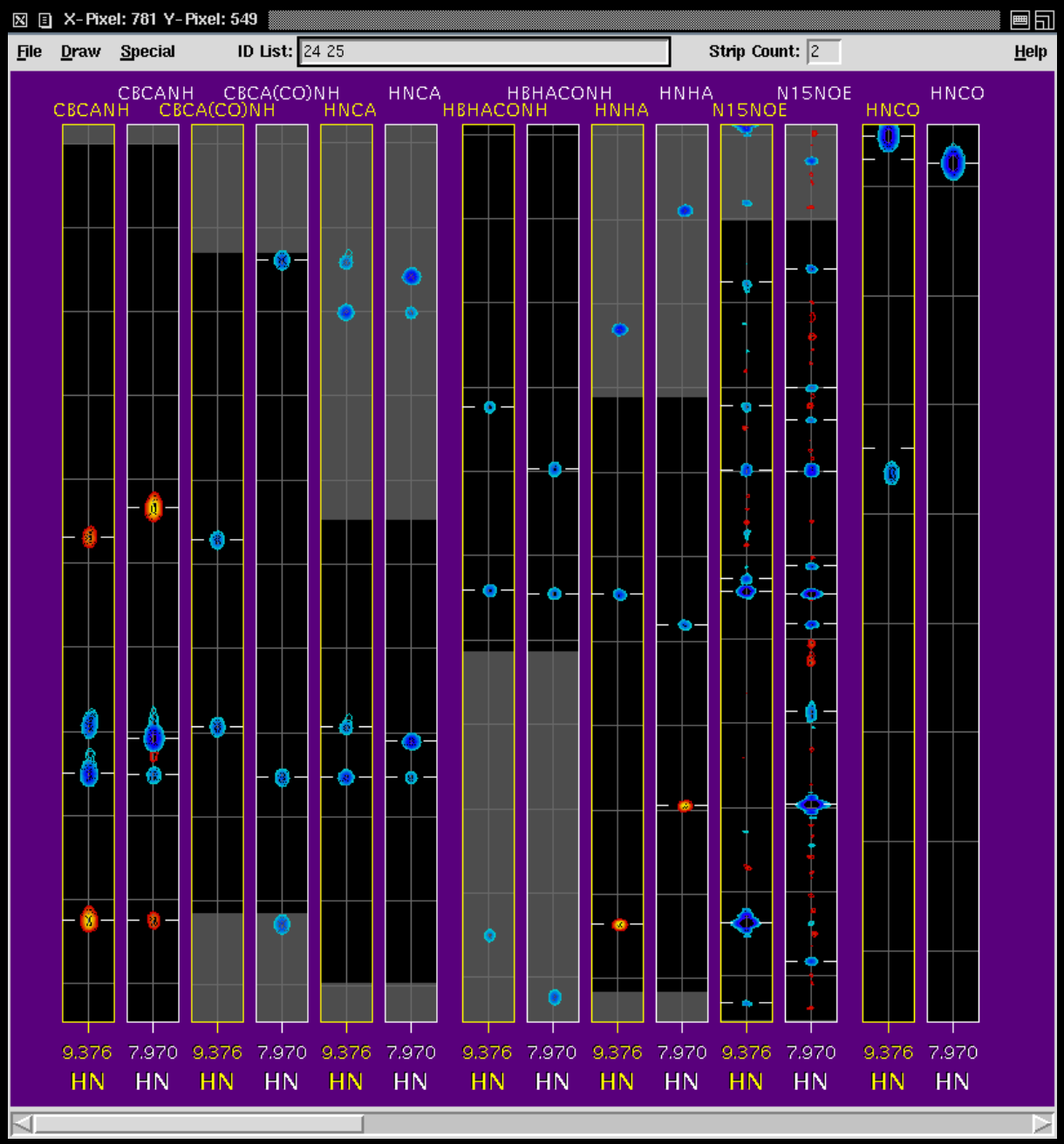

70
50
30
10
10
30
50
70
90

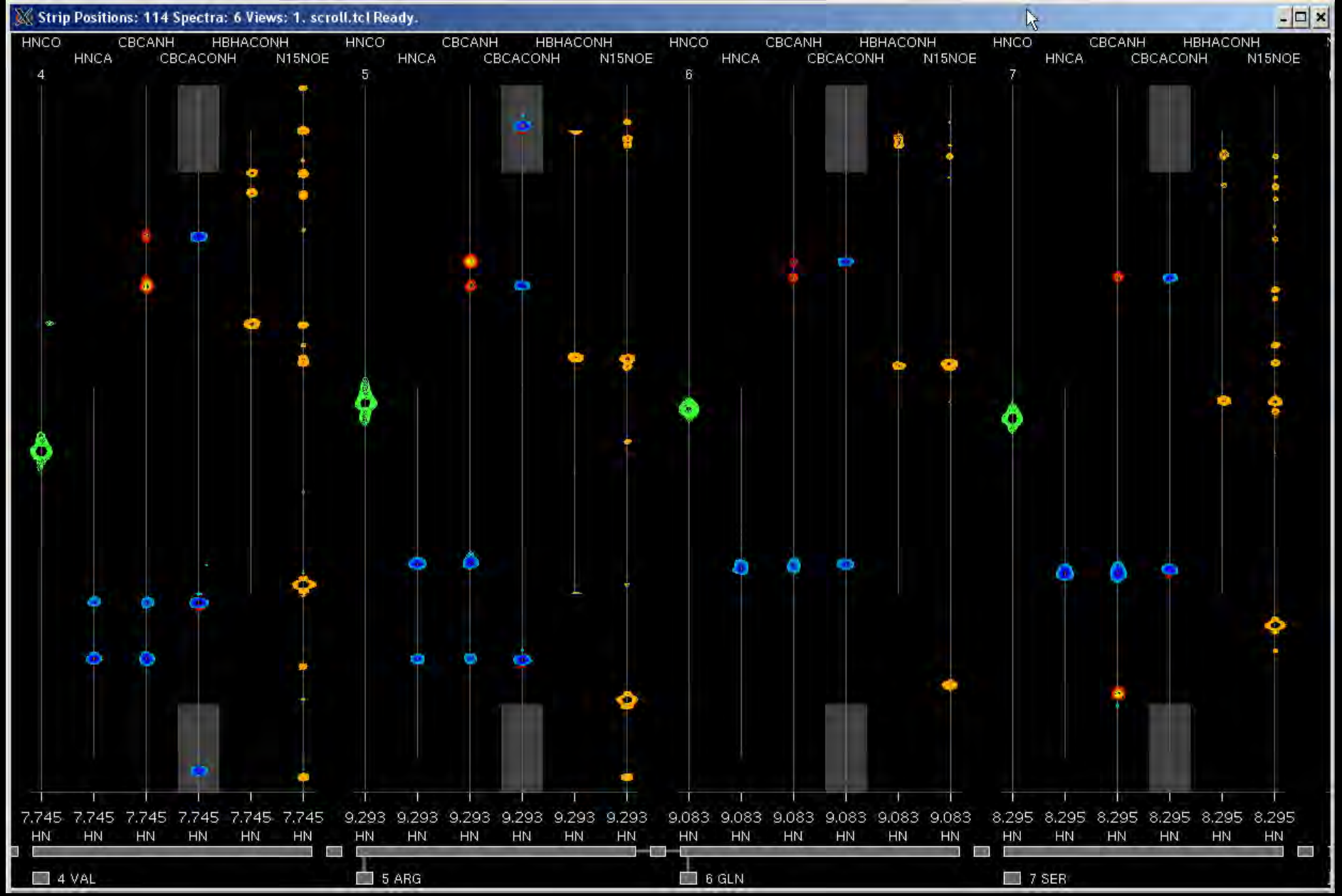

Atomic Volume
0
20
40
60
0
20
40
60
80
Atomic Number


Cs
55
Rb
37
Fr
87
K
19
Na
11
Li
3
60
40
20
Atomic
Volume
the rare
earths
20
40
60
80
Atomic Number


HOKKAIDO
JAPAN
BASIN
JAPAN SEA
YAMATO
RIDGE
YAMATO
BASIN


VLADIVOSTOK
JAPAN
BASIN
HOKKAIDO
MURORAN
JAPAN SEA
YAMATO
RIDGE
YAMATO
BASIN
1000
2000
3000


X-Pixel: 781 Y-Pixel: 549
File
Draw
Special
ID List: 24 25
Strip Count: 2
Help
CBCANH
CBCA(CO)NH
HNCA
HBHACONH
HNHA
N15NOE
HNCO
9.376
7.970
HN


Strip Positions: 114 Spectra: 6 Views: 1. scroll.tcl Ready.
HNCO
HNCA
CBCANH
CBCACONH
HBHACONH
N15NOE
4
5
6
7
7.745 7.745 7.745 7.745 7.745 7.745
9.293 9.293 9.293 9.293 9.293 9.293
9.083 9.083 9.083 9.083 9.083 9.083
8.295 8.295 8.295 8.295 8.295 8.295
HN
4 VAL
5 ARG
6 GLN
7 SER