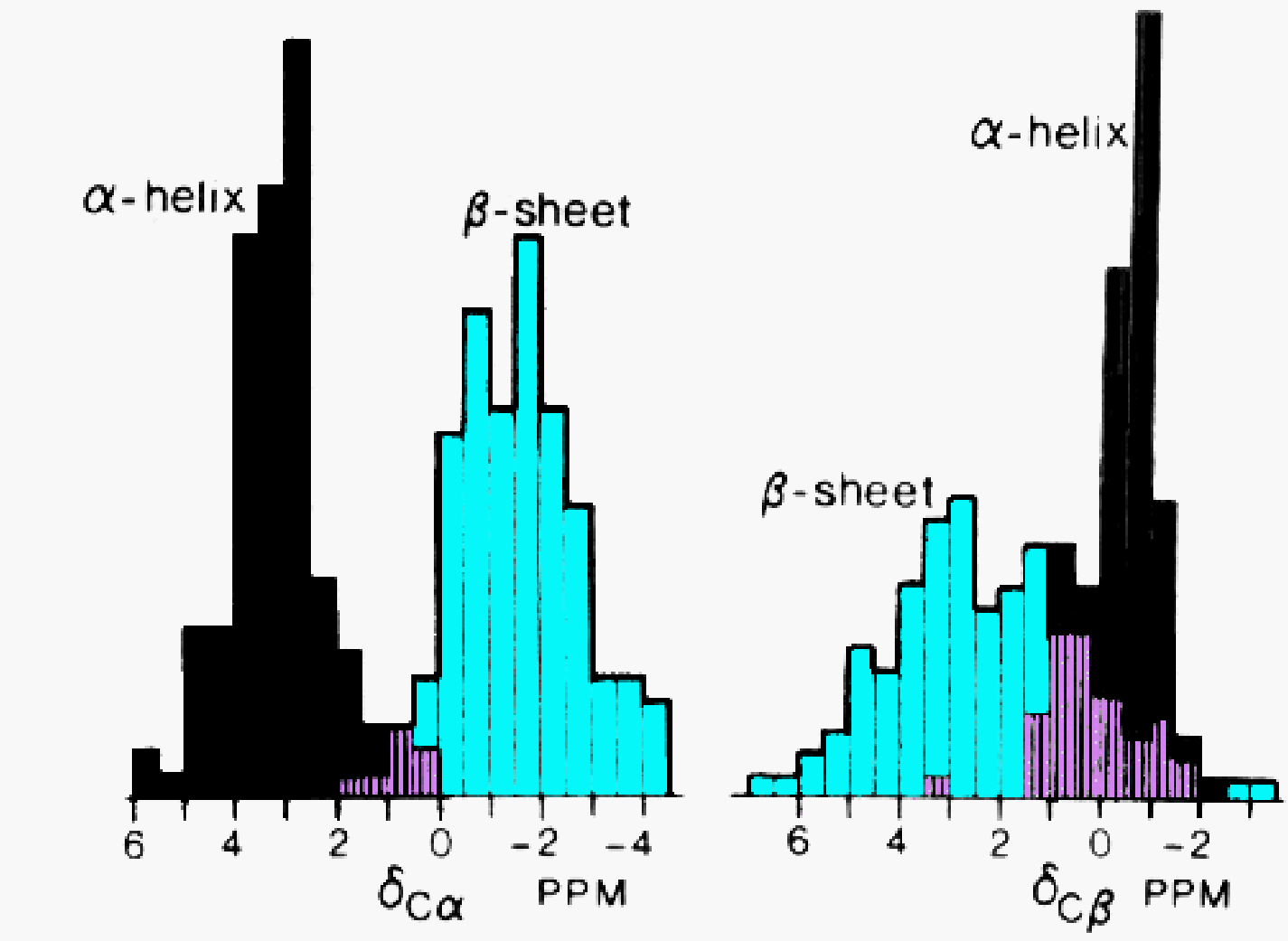
Summary Lecture 10

Secondary structure, distance restraints and preparation of calculation

Secondary chemical shifts encode secondary structure information



δC_α (exp) – δC_α(RC) δC_β (exp) – δC_β(RC)

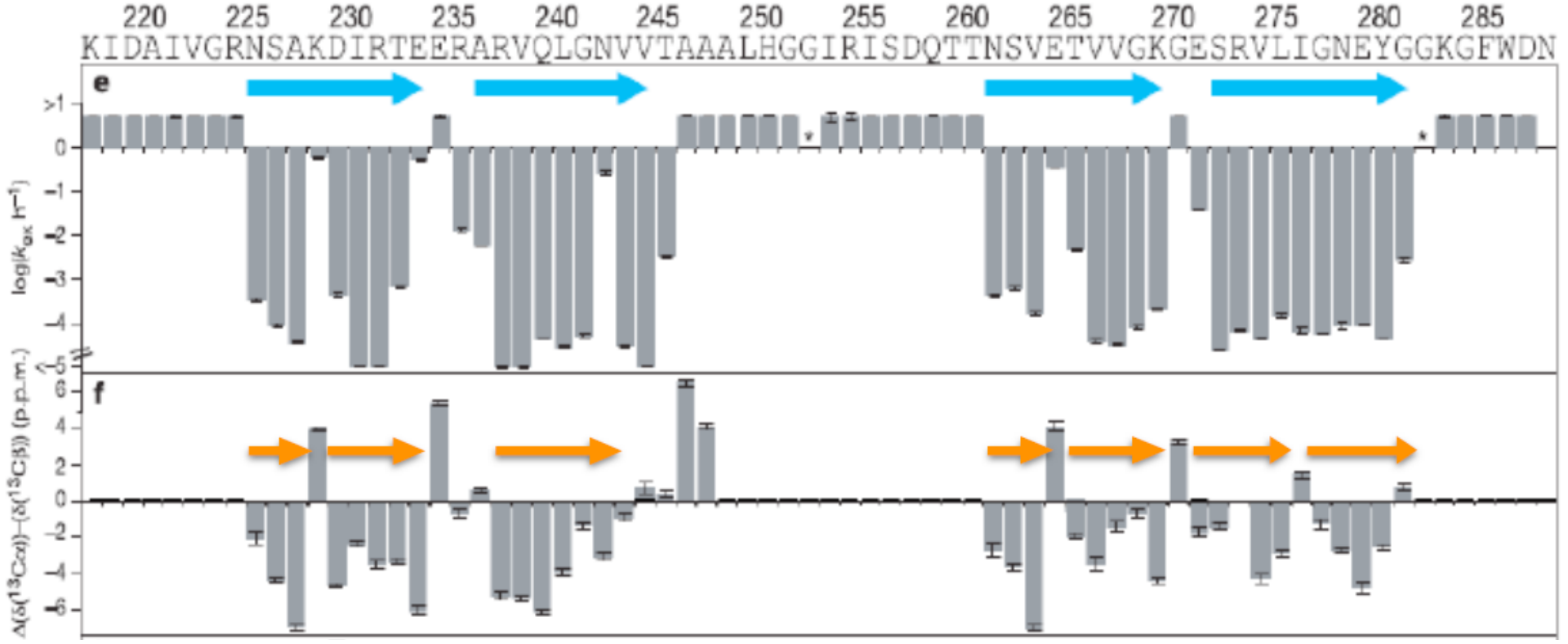
Bax et al., *J. Am. Chem. Soc.*, **1991**, *113*, 5490.

Chemical shifts encode secondary structure information

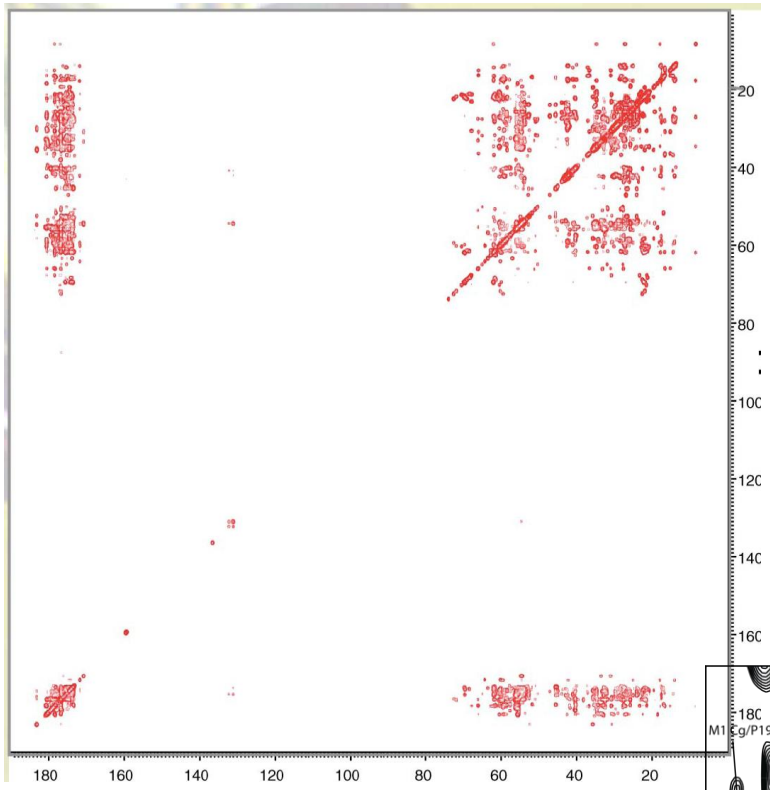
Definitions

a) **α-helix**: At least 4 residues in a row with a positive secondary chemical shift
b) **β-sheet**: At least 3 residues in a row with a negative secondary chemical shift

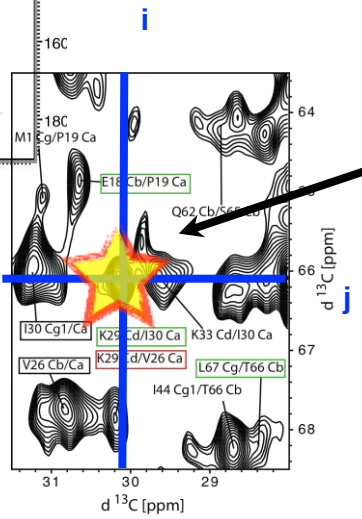
Example: HET-s $(\delta C_\alpha \text{ (exp)} - \delta C_\alpha \text{(RC)}) - (\delta C_\beta \text{ (exp)} - \delta C_\beta \text{(RC)})$



Cross-Peak Intensities somehow represent the Dipolar Interaction



19 kHz MAS, 200 ms mix DARR, 258 K, 850 MHz, U-^{13}C-^{15}N Ubiquitin



Kinetic Master Equation for Populations (through perturbation theory)

$$\frac{d}{dt}\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ \vdots \\ p_N \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} & \vdots & \vdots & W_{1N} \\ W_{21} & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ W_{N1} & \vdots & \vdots & \vdots & W_{NN} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ \vdots \\ p_N \end{bmatrix}$$

W_{ij}: Rate constants for Polarization Transfer

Cross-Peak Intensity encodes dipole interaction d_{ij}

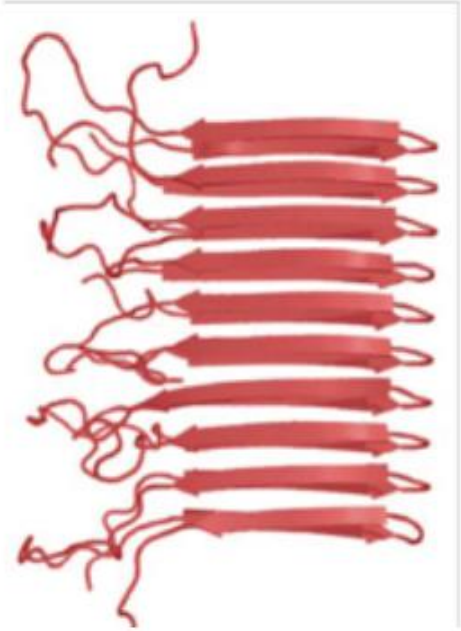
$$W_{ij} = \frac{\pi}{2}(d^{(i,j)})^2 f_{ij}(0)$$

Distance Restraints

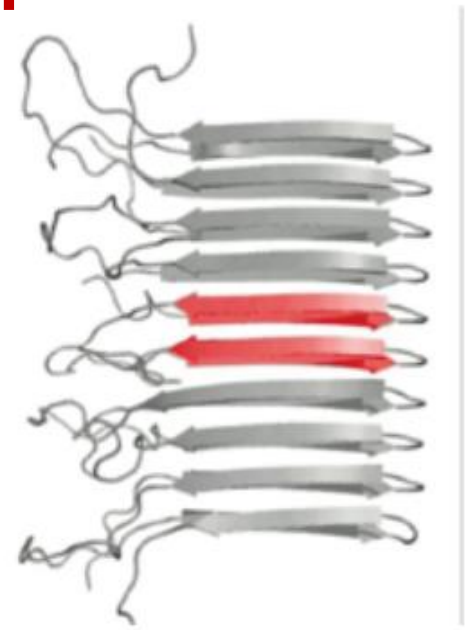
C-C	PDSD/DARR/MIRROR/ PAR
N-C	PAIN
H-H	CHHC, NHHC

On which samples do we record these experiments?

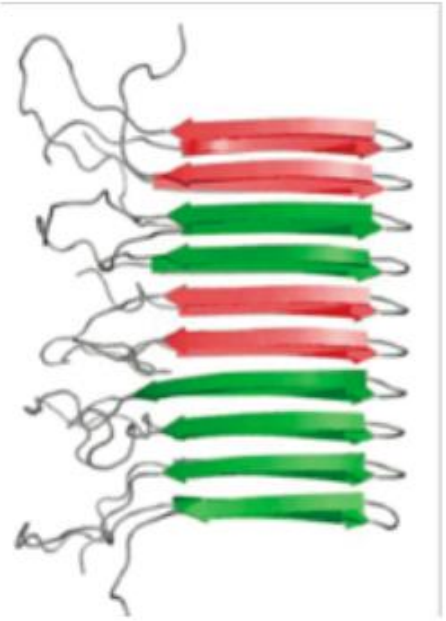
Isotope labelling allows to disentangle the intra/inter problem



All monomers labelled (^{15}N, ^{13}C)
intra and inter constraints in PDSD / CHHC / NHHC spectra
+ PAR



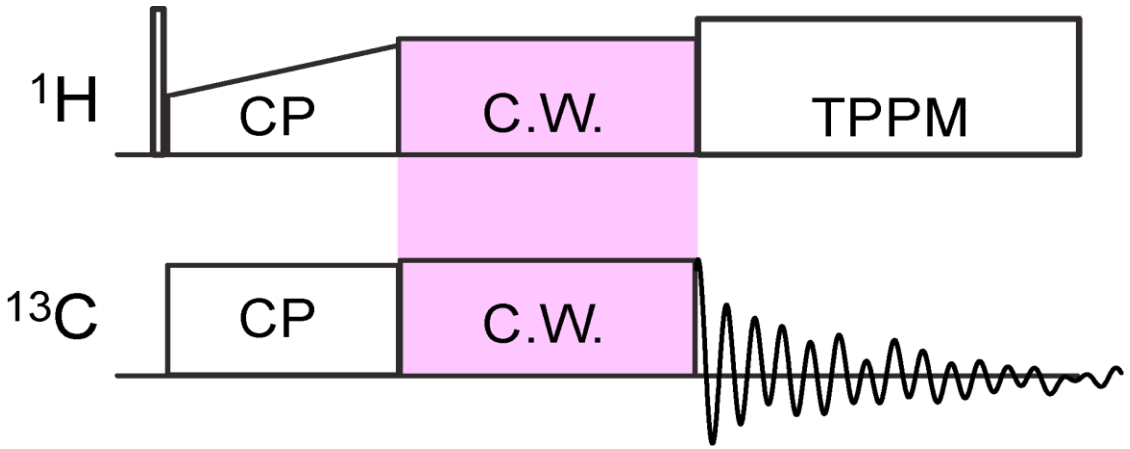
Labelled (^{15}N, ^{13}C) monomers dilluted in natural abundance.
Only intra constraints in PDSD / CHHC / NHHC spectra **+ PAR**

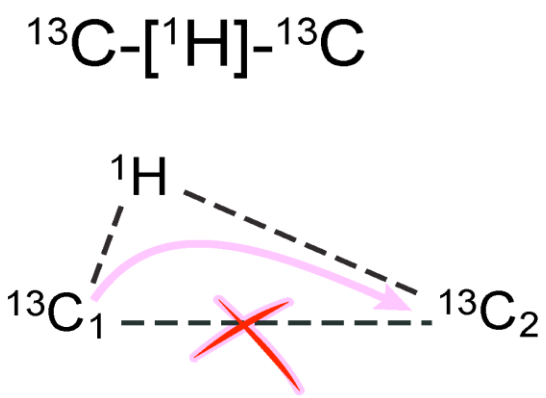


^{13}C labelled monomers and ^{15}N labelled monomers mixed.
Only inter constraints (in NHHC spectra)
+ PAIN

Proton assisted recoupling (PAR)

Non-resonant homonuclear correlation experiment.

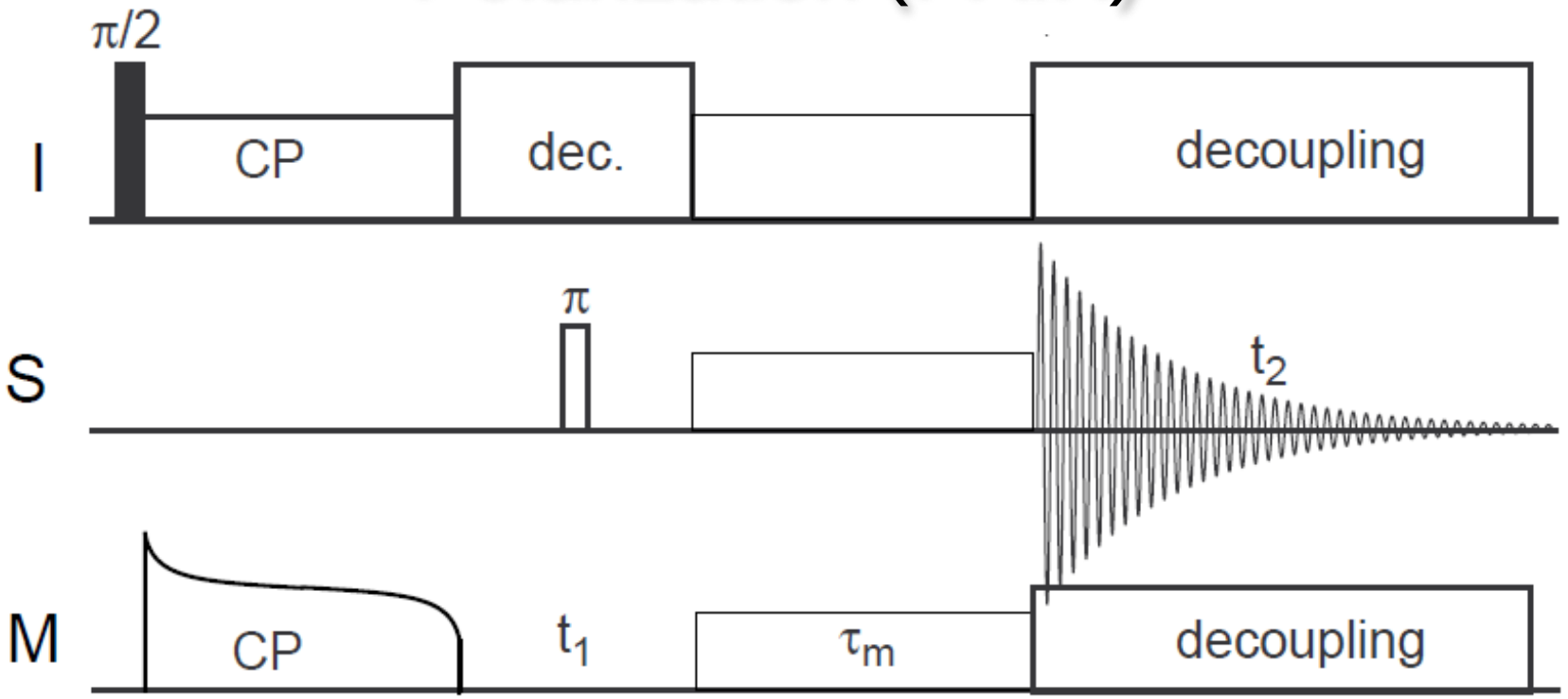




Experimental aspects:

- Avoid Hartmann-Hahn matching conditions
- Avoid rotary resonance conditions

Proton Assisted Insensitive Nuclei Cross Polarization (PAIN)



- Heteronuclear version of PAR
- Resonant polarization transfer
- Resonance condition $\omega_{1S} - \omega_{1M} = n_r \omega_r$

PAIN to study fibril stacking

NCA

UL ^{13}C, ^{15}N

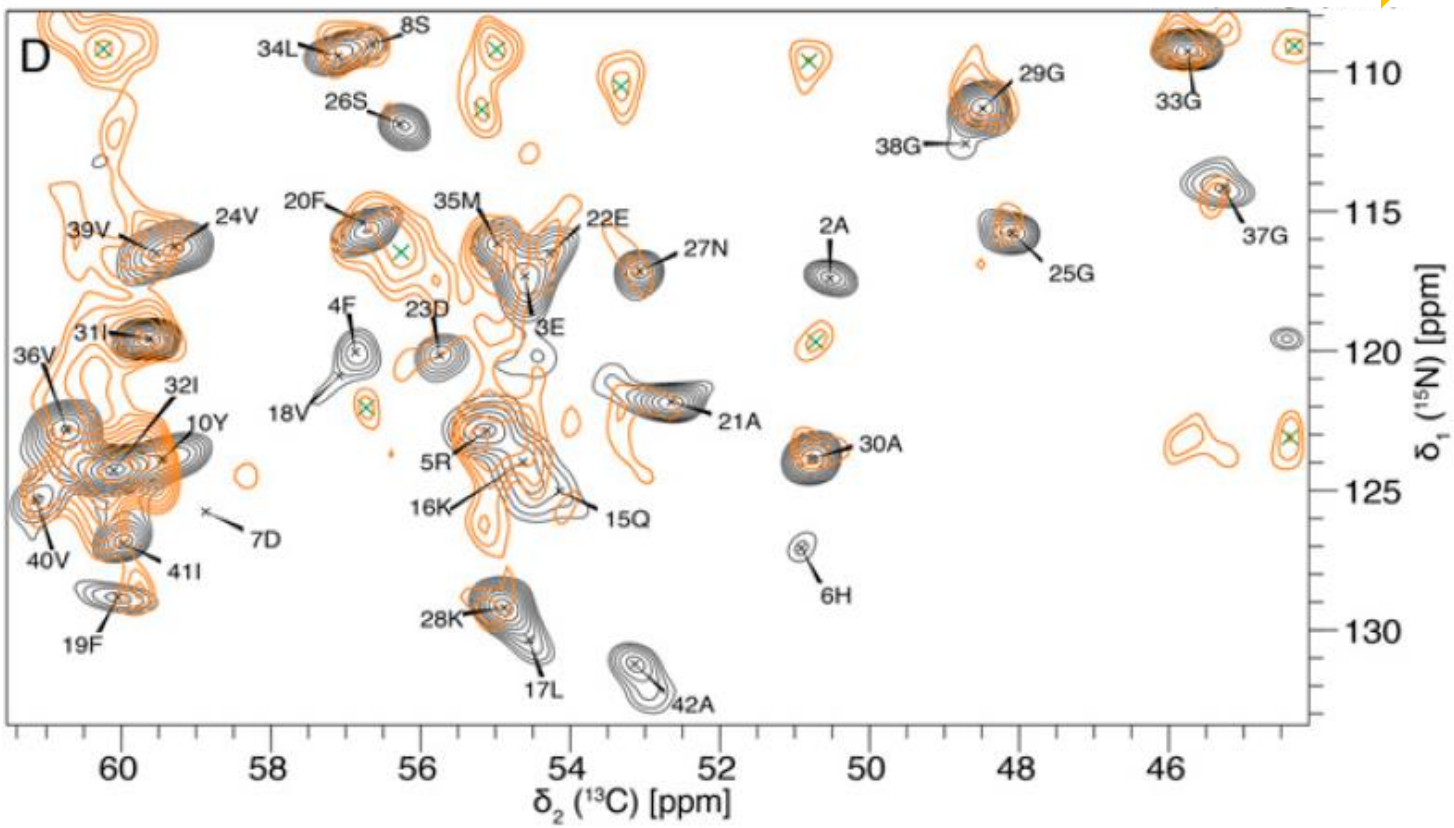
15N

13C

15N

PAIN

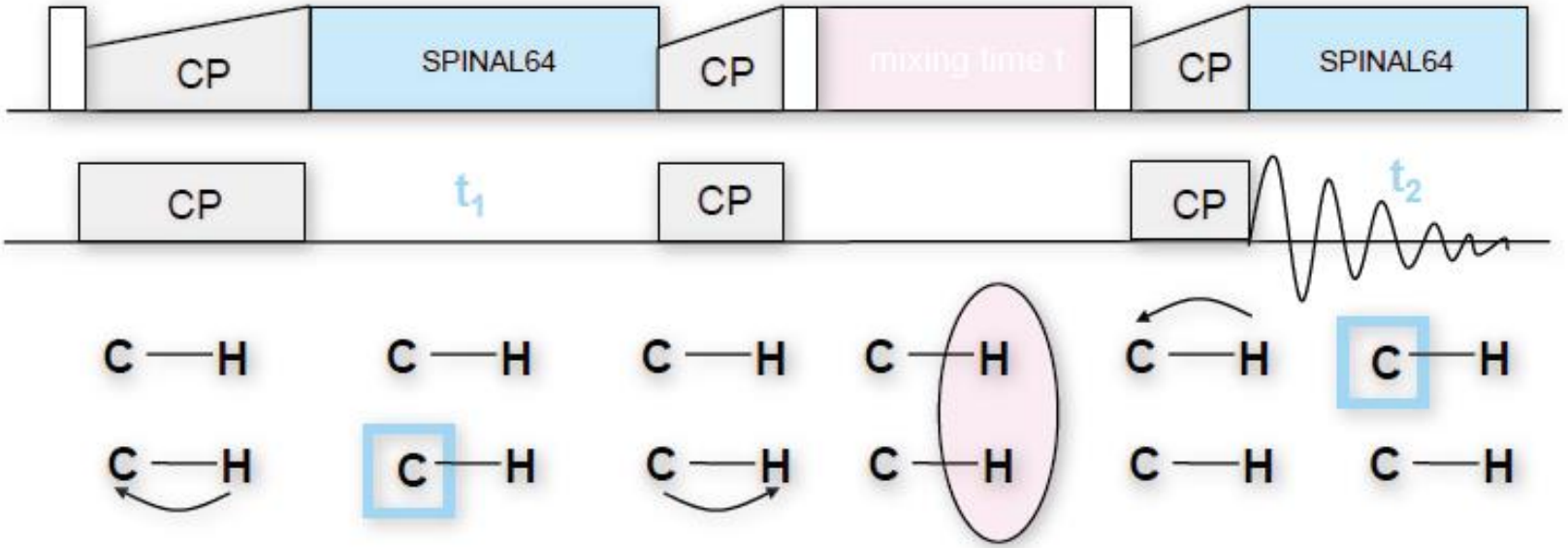
Mixed labelled



In-register parallel

Wälti et. al., *PNAS*, **2016**.

The CHHC/NHHC experiments



^{13}C-^{13}C or ^{15}N-^{13}C correlation experiment

Lecture 11

Structure calculation

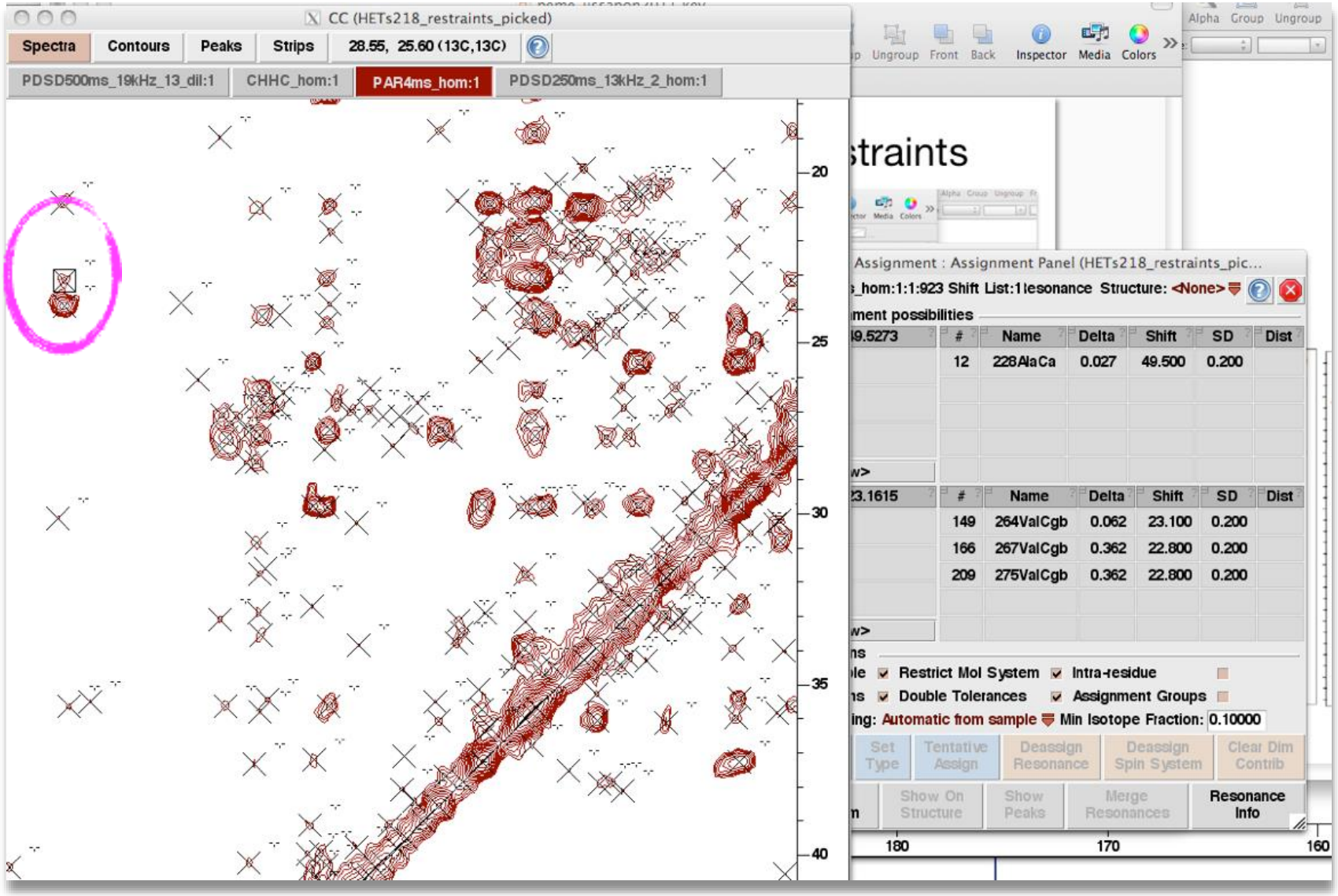
CYANA

www.cyana.org
cyana wiki

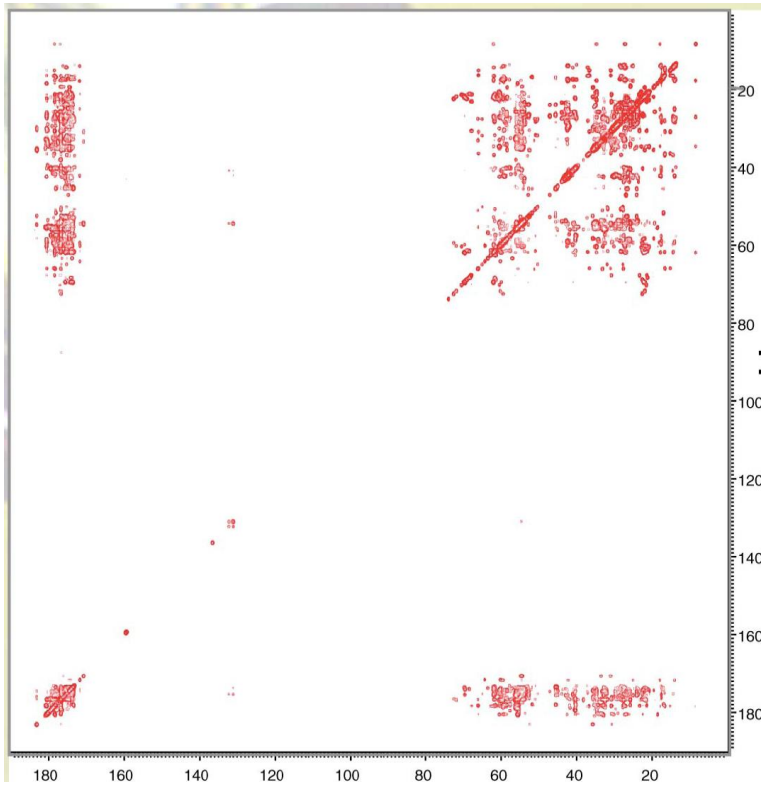
Peter Güntert

Institute for Biophysical Chemistry,
Biomolecular Magnetic Resonance Center,
and Frankfurt Institute for Advanced Studies
Goethe University Frankfurt am Main

Long-range restraints: assignment ambiguities

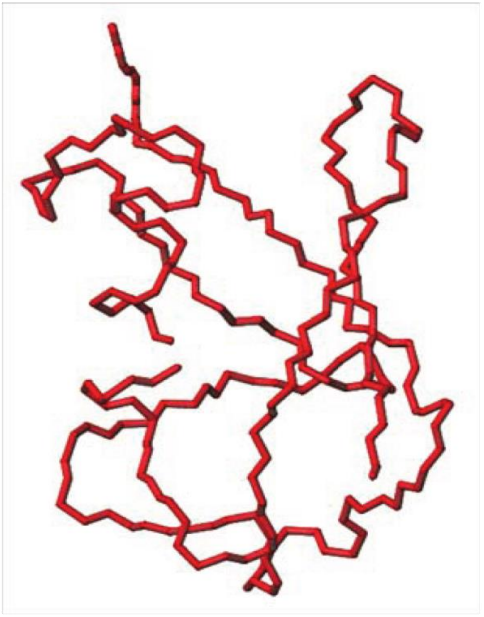


Structure determination is an inverse problem



calculate

???



...but even the forward problem is not trivial

Forward calculation in a simple approach

- Rate approach for NOESY

$$p_i = \langle S_{iz} \rangle$$

$$\frac{d}{dt}\begin{bmatrix} p_1 \\ p_2 \\ : \\ : \\ p_N \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} & : & : & W_{1N} \\ W_{21} & : & : & : & : \\ : & : & : & : & : \\ : & : & : & : & : \\ W_{N1} & : & : & : & W_{NN} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ : \\ : \\ p_N \end{bmatrix}$$

Initial rate approx.

$$W_{ij} \sim r_{ij}^{-6}$$

$$I_{ij} \sim r_{ij}^{-6}$$

- Solids static (spin-diffusion concept by Bloembergen)

$$W_{ij} = \frac{(\mu_0 \hbar \gamma_i \gamma_j)^2}{32\pi} \frac{1}{r_{ij}^6} \cdot P_2(\cos\theta_{ij})^2 f_{ij}(0)$$

$$W_{ij} \cong \frac{k}{r_{ij}^6}$$ *Crude approximation*

- MAS more complex!!! (but time dependent Liouville-von-Neumann equation works, of course)

Forward calculation in a simple approach

- Complex relation between cross-peak intensity and local geometry
- BUT: no mechanism can lead to cross-peaks between nuclei if they are separated by more than an *upper distance* u_{ij}
- u_{ij} depends on the experiment performed
- many „unprecise“ *distance restraints* are defined

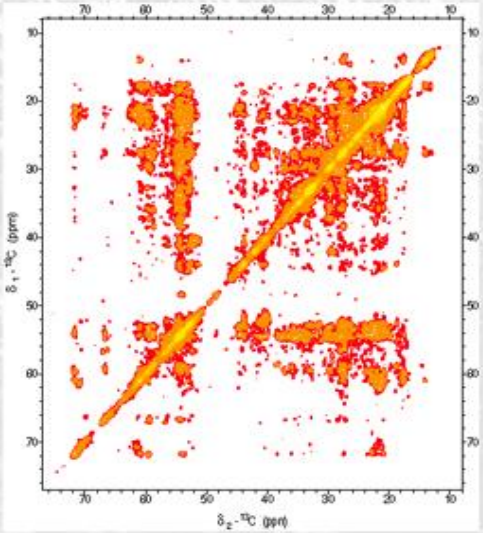
Manolikas et al., *J. Am. Chem. Soc.*, **2008**, *130*, 3959.

Structure determination schemes

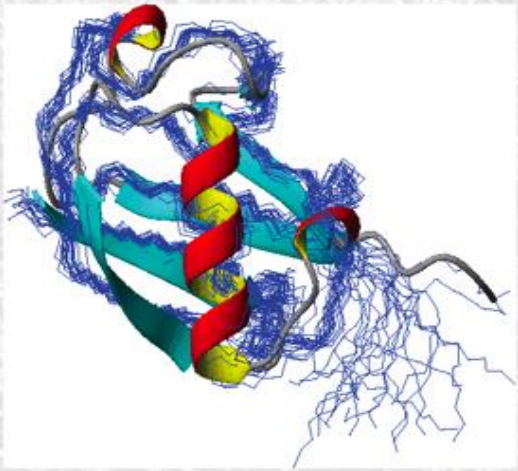
We acknowledge that $I_{ab} = c\tau / r_{ab}^{6}$ is certainly wrong but believe that it is not too wrong -> we reduce the data evaluation to the NOESY problem.

automatic/manual

data



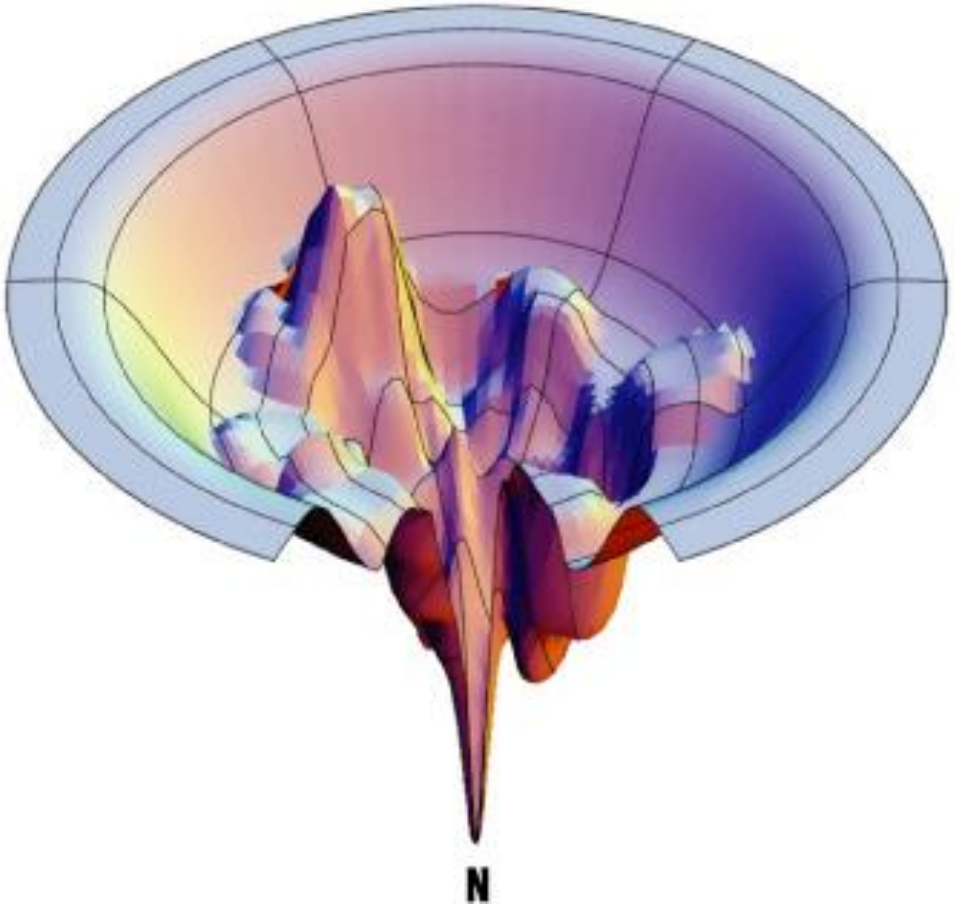
structure



Structure calculations

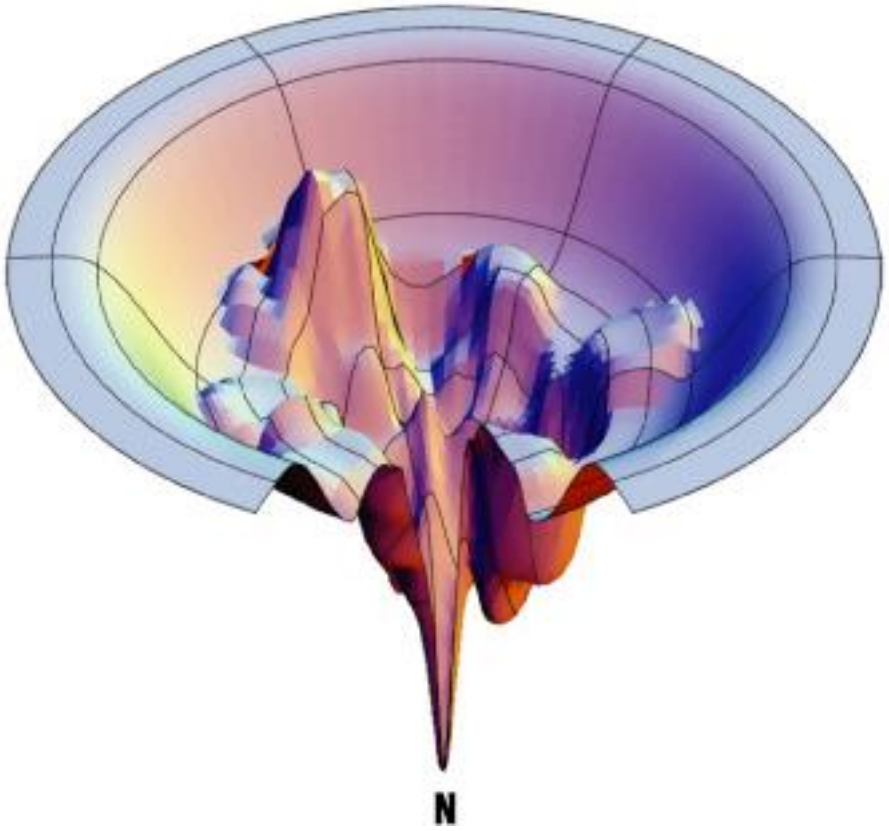
- Structure calculation programs try to **“fold” a protein** into a three dimensional structure that agrees with the measured data.
- Differences between measured data and the structure are manifested as **violations**.
- Violations cause forces that act on the molecule, driving it towards **minimal (pseudo)energy** and optimal agreement with the measured data.
- The **target function** (pseudoenergy) is the sum of squares (or similar) of the violations.
- The **energy landscape** of this target function is complex and has many local minima.

This is like folding of a protein



Finding the global minimum is difficult, but MD has developed concepts, *e.g.* simulated annealing.

...but the surface is not a real energy but a pseudoenergy (**Target Function**)

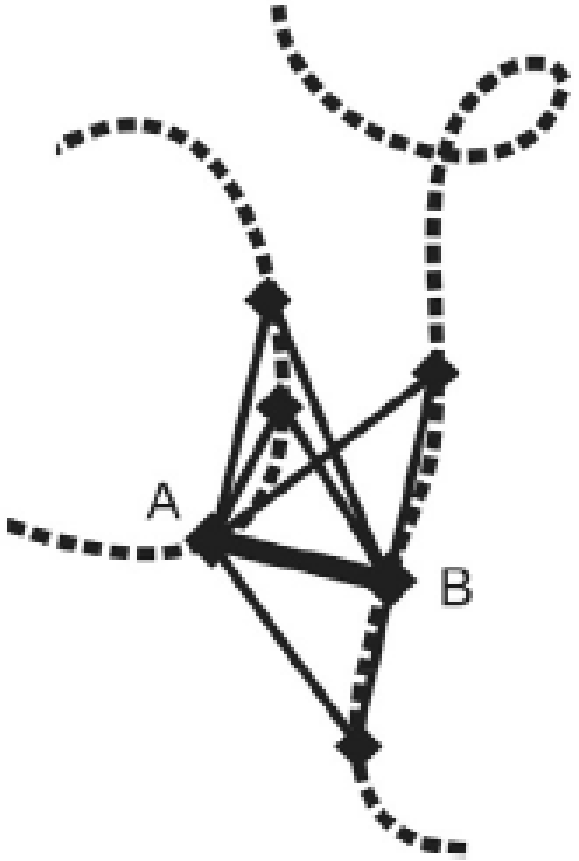


Structure calculation using CYANA

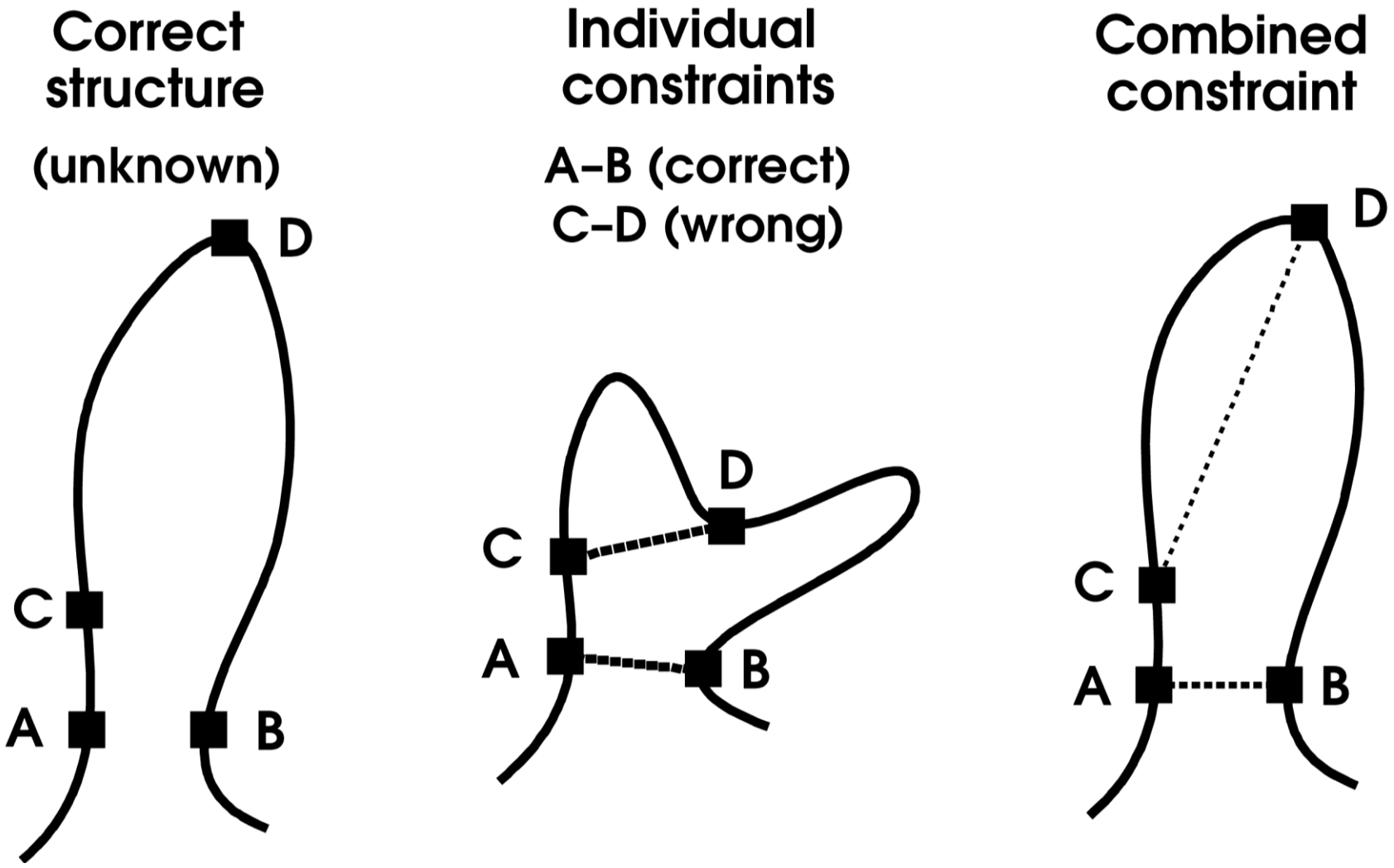
Please take notes, not all slides will be distributed because some are taken (with permission) from a talk from Prof. Peter Güntert and will NOT be distributed.

Automated Cross-peak assignment

Network-anchoring

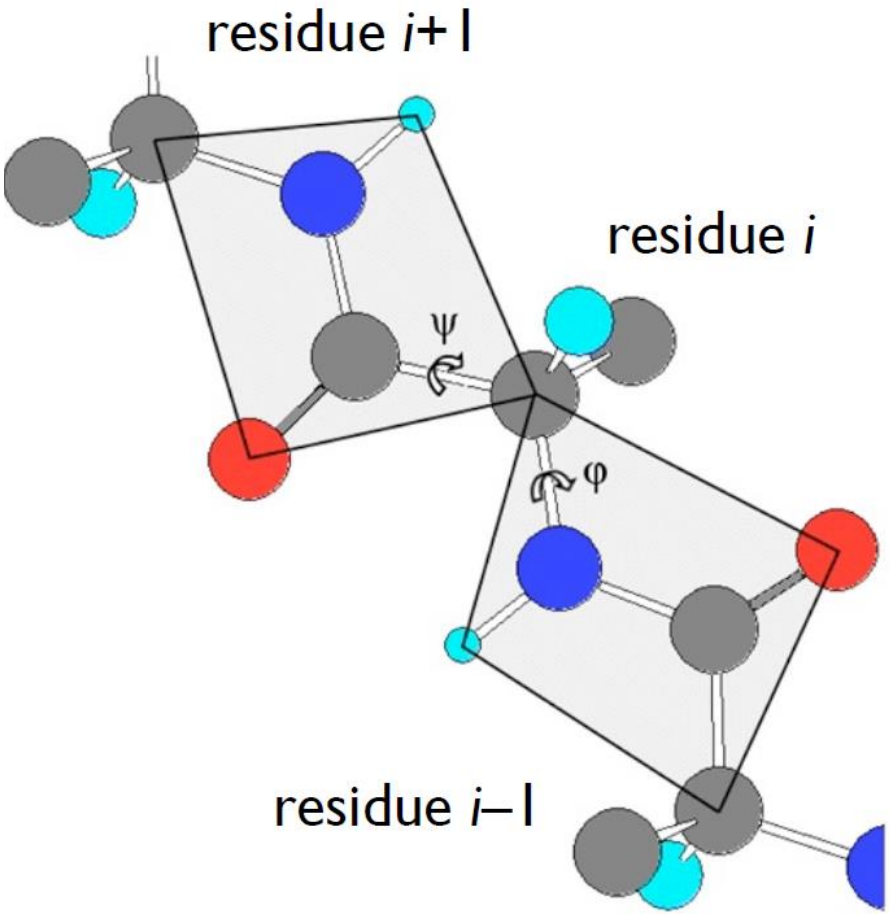


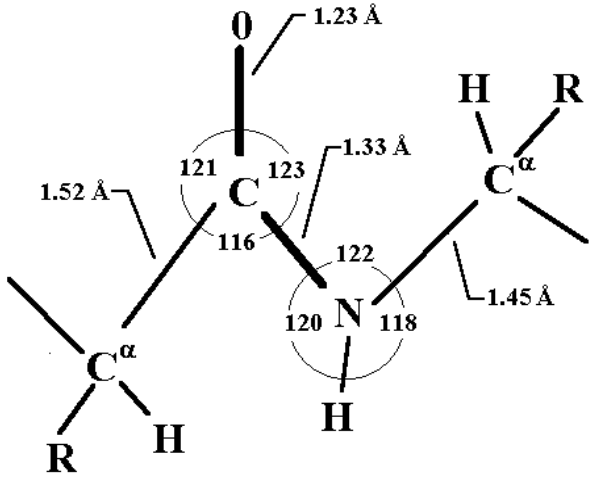
Restraint combination (first two cycles)



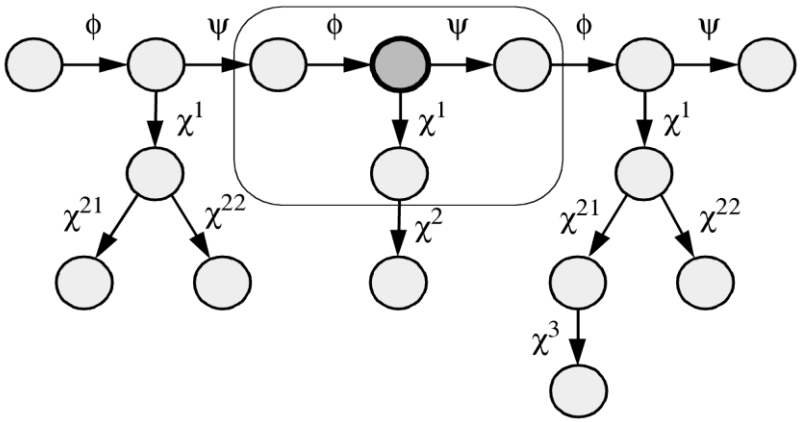
Structure calculation using CYANA

Torsion angles



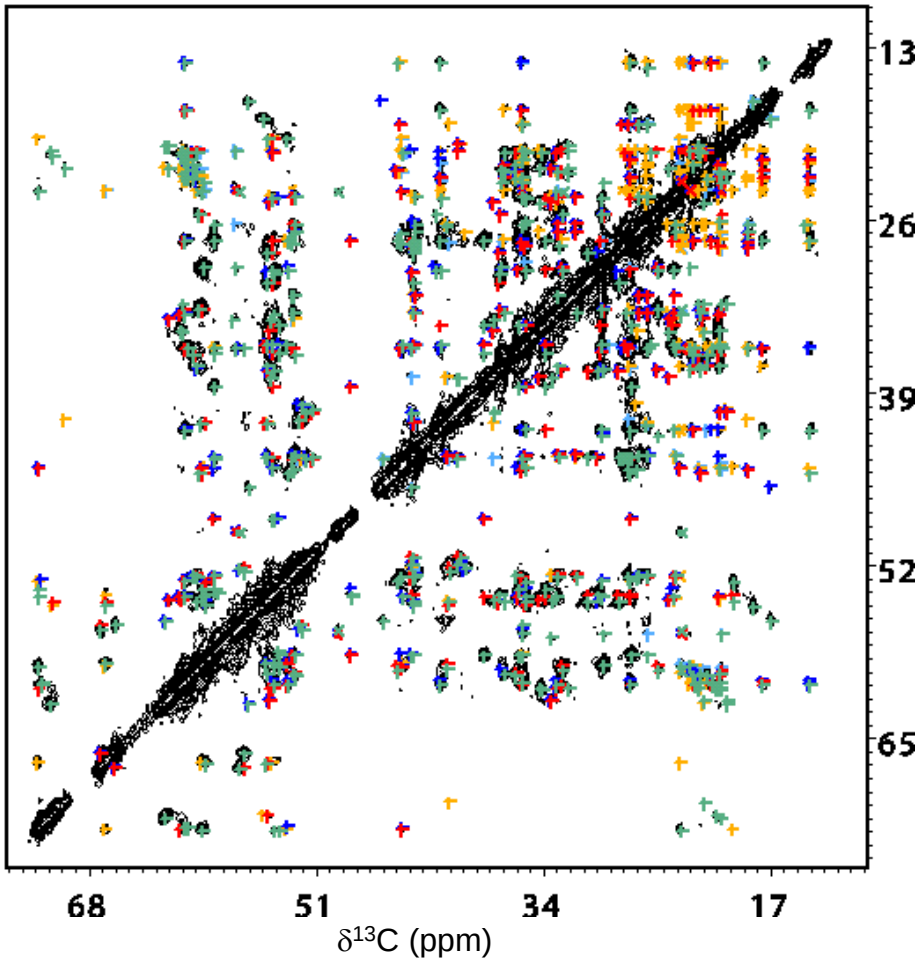


Other distances are fixed



Structure explains NMR spectra (back-calculation matches experiment)

CHHC spectrum, U-^{15}N,^{13}C labelled HET-s(218-289),
150 μs mix, 9.5 kHz MAS, 850 MHz



Restraints classification:

Short-range
Intramolecular β-sheet
Intramolecular middle and long-range
Intermolecular β-sheet
Intermolecular middle and long-range

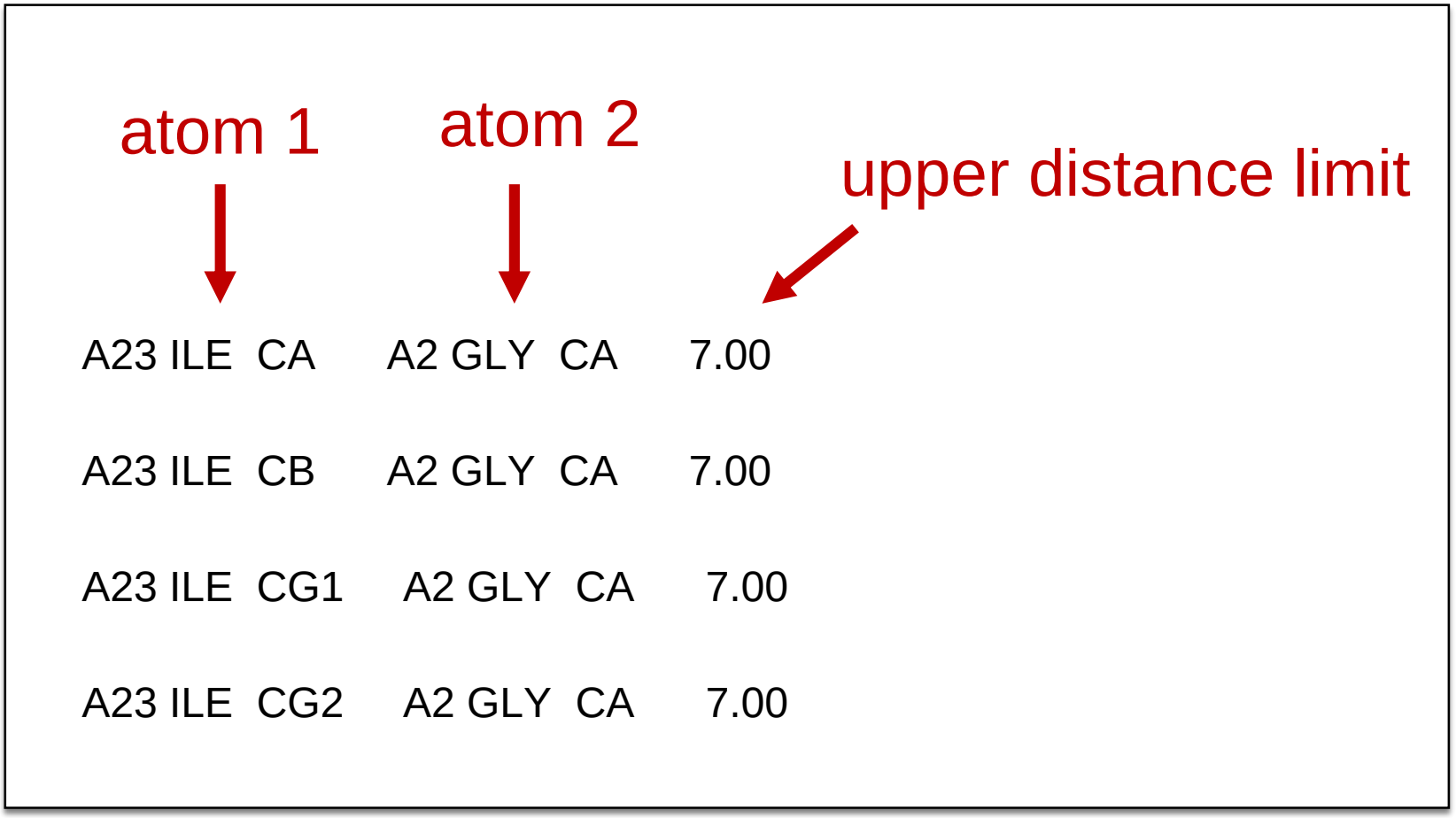
All peaks explained

Preparation of the structure calculation

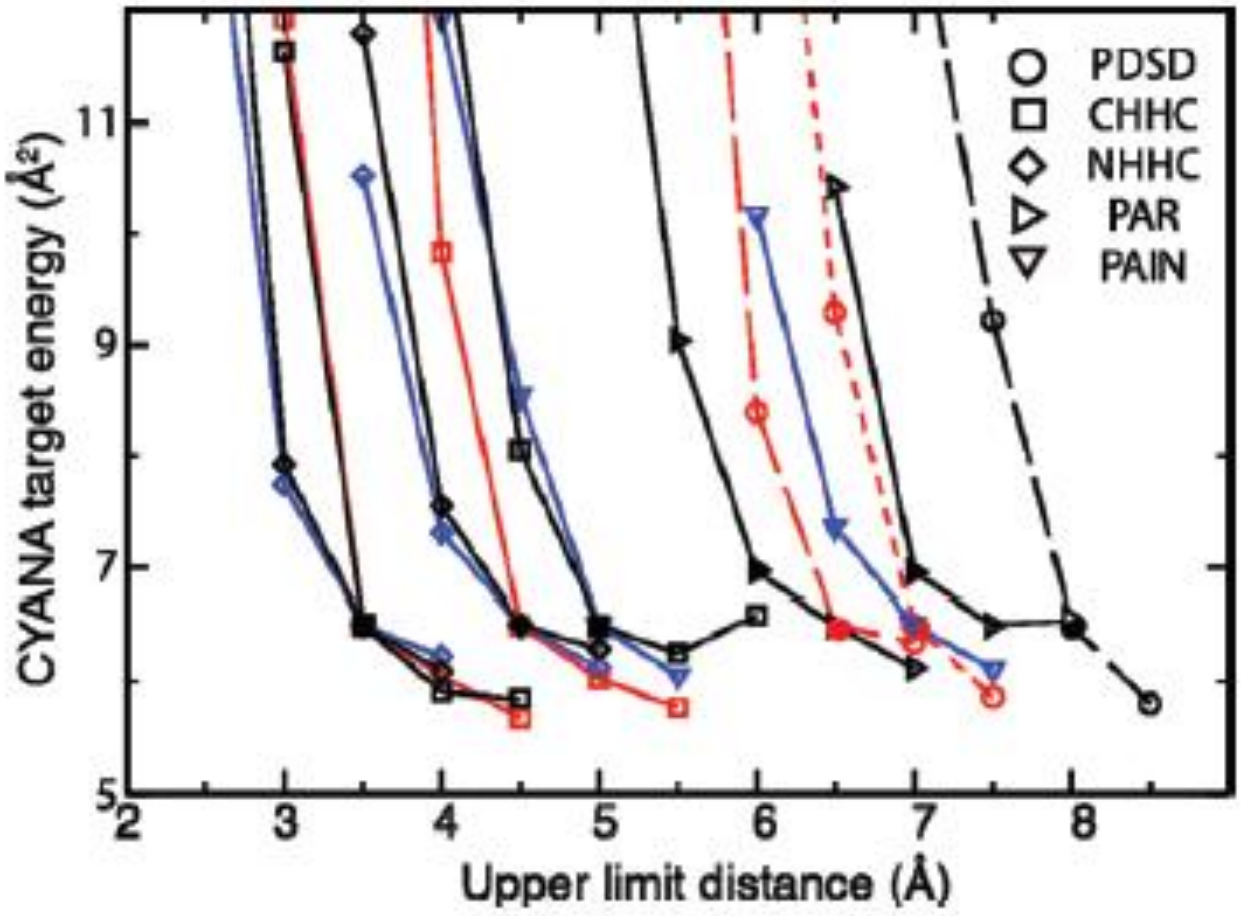
Three ways to prepare distance restraints for structure calculation

- Pick a region automatically and export list (filename.peaks)
- Pick peaks by hand and export list for CYANA (filename.peaks)
- Pick peaks and assign them manually (and create a restraint list for CYANA manually) (filename.upl)

Manual unambiguous restraints



Adapted upper distance limits...



Lines: uniformly labeled samples are continuous; those for Glycerol labeled samples are dashed (2-^{13}C-Glycerol, long, 1,3-^{13}C-Glycerol, short).

Required files for structure calculation

- HETs_trimer.seq (sequence)
- HETs_trimer.prot (assignment)
- HET-s_talos_trimer.aco (TALOS+ data)
- *.peaks for all experiments
- HET-s_hbonds.lol / upl for H-bonds
- init.cya (cyana symmetries etc.)
- Auto.cya (structure calculation details)

Sequence file (.seq)

```
MET    A217
LYS    A218
ILE    A219
ASP    A220
ALA    A221
ILE    A222
VAL    A223
...
HIS    A295

PL     L01
LL2    L02
LL2    L03
...
LL2    L48
LP     L49

MET    B217
LYS    B218
ILE    B219
ASP    B220
ALA    B221
ILE    B222
...
```

HET-s(218-289)
from CCPN

Flexible linker

HET-s(218-289)
from CCPN

Chemical shift files (.prot)

```
   1  175.070 0.000       C A222
   2   62.000 0.000      CA A222
   3   38.262 0.000      CB A222
   4   13.014 0.000     CD1 A222
   5   27.492 0.000     CG1 A222
   6   18.689 0.000     CG2 A222
   7  128.938 0.000       N A222
...
 328  175.070 0.000       C  B222
 329   62.000 0.000      CA  B222
 330   38.262 0.000      CB  B222
 331   13.014 0.000     CD1  B222
 332   27.492 0.000     CG1  B222
 333   18.689 0.000     CG2  B222
 334  128.938 0.000       N  B222
....
 1180  62.000 0.000      HA  A222
 1181  38.262 0.000      HB  A222
 1182  13.014 0.000     QD1  A222
 1183  27.492 0.000     QG1  A222
 1184  18.689 0.000     QG2  A222
  982 128.938 0.000       H  A222
...
```

chemical shifts as exported from CCPN (xeasy format)

identical shifts for chains B and C

translation to attached protons for CHHC and NHHC

Dihedral angle restraints from TALOS (*.aco)

```
A222      ILE     PHI    -116.0  -84.0
A222      ILE     PSI     116.0  138.0
A223      VAL     PHI    -139.0 -101.0
A223      VAL     PSI     120.0  164.0
A225      ARG     PHI    -142.0  -90.0
A225      ARG     PSI     122.0  154.0
A226      ASN     PHI    -116.0  -82.0
A226      ASN     PSI     115.0  149.0
A227      SER     PHI    -140.0 -122.0
A227      SER     PSI     138.0  170.0
...
B222      ILE     PHI    -116.0  -84.0
B222      ILE     PSI     116.0  138.0
B223      VAL     PHI    -139.0 -101.0
B223      VAL     PSI     120.0  164.0
...
```

See Exercise hour.

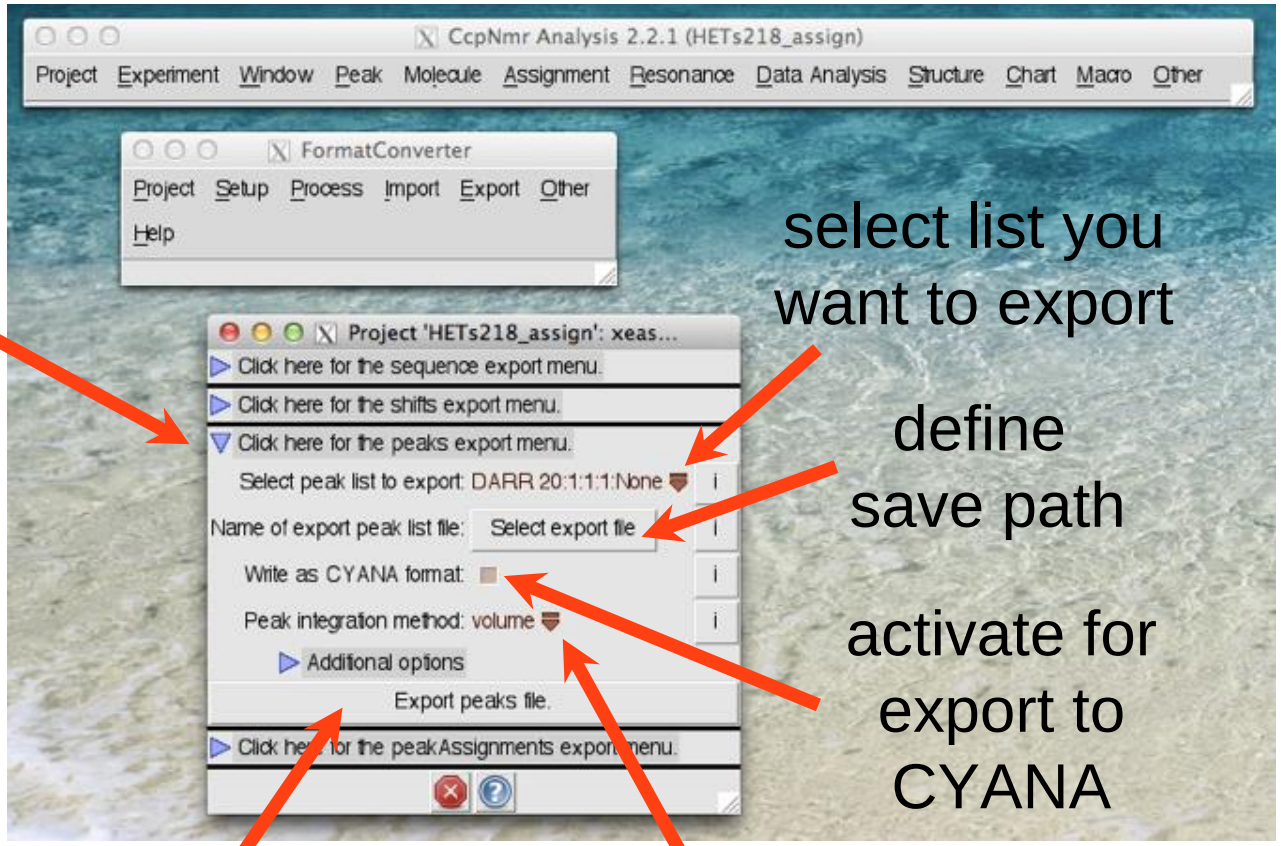
Peak files for all spectra (.peaks)

PAIN_mixed.peaks

```
# Number of dimensions 2
# INAME 1 N
# INAME 2 C
#CYANAFORMAT NC
#TOLERANCE 0.55 0.25
   1 135.022  67.146 +9 T      3.727e+08  0.00e+00 a    0     0     0 0 (9: intermolecular)
   2 131.784  61.415 1 T       4.263e+08  0.00e+00 a    0     0     0 0 (8: intra, 1 ambiguous)
   3 130.491  52.789 1 T       6.357e+08  0.00e+00 a    0     0     0 0
   4 130.411  35.365 1 T       4.183e+08  0.00e+00 a    0     0     0 0
   5 129.779  20.938 1 T       4.401e+08  0.00e+00 a    0     0     0 0
   6 128.896  41.830 1 T       3.873e+08  0.00e+00 a    0     0     0 0
   7 128.814  27.950 1 T       5.204e+08  0.00e+00 a    0     0     0 0
   8 128.215  60.864 1 T       4.883e+08  0.00e+00 a    0     0     0 0
   9 128.504  54.278 1 T       6.229e+08  0.00e+00 a    0     0     0 0
  10 128.381  20.887 1 T       4.145e+08  0.00e+00 a    0     0     0 0
  11 127.509  52.213 1 T       3.776e+08  0.00e+00 a    0     0     0 0
  12 126.424  67.270 1 T       3.706e+08  0.00e+00 a    0     0     0 0
  13 126.643  59.764 1 T       6.177e+08  0.00e+00 a    0     0     0 0
  14 126.686  43.821 1 T       4.019e+08  0.00e+00 a    0     0     0 0
  15 125.981  53.566 1 T       3.968e+08  0.00e+00 a    0     0     0 0
  16 124.036  22.097 1 T       5.249e+08  0.00e+00 a    0     0     0 0
  17 122.955  60.715 1 T       8.797e+08  0.00e+00 a    0     0     0 0
```

Exporting Peak Lists in CCPN

change to the peaks export menu



select list you want to export

define save path

activate for export to CYANA

export height/volume?

click to export list

Proton-restraints for heteroatom spectra and mixed vs. complete labelling

NHHC_**mixed**.peaks

.peaks

```
# Number of dimensions 2
# INAME 1 N H
# INAME 2 C H
#CYANAFORMAT NC HH
#TOLERANCE 0.8 0.4
   1 135.022  67.146 19 T       3.727e+08  0.00e+00 a   0    0     0 0 (9: intermolecular)
   2 131.784  61.415 1 T        4.263e+08  0.00e+00 a   0    0     0 0 (8: intra, 1 ambiguous)
   3 130.491  52.789 1 T        6.357e+08  0.00e+00 a   0    0     0 0
   4 130.411  35.365 1 T        4.183e+08  0.00e+00 a   0    0     0 0
   5 129.779  20.938 1 T        4.401e+08  0.00e+00 a   0    0     0 0
   6 128.896  41.830 1 T        3.873e+08  0.00e+00 a   0    0     0 0
   7 128.814  27.950 1 T        5.204e+08  0.00e+00 a   0    0     0 0
   8 128.215  60.864 1 T        4.883e+08  0.00e+00 a   0    0     0 0
   9 128.504  54.278 1 T        6.229e+08  0.00e+00 a   0    0     0 0
  10 128.381  20.887 1 T        4.145e+08  0.00e+00 a   0    0     0 0
  11 127.509  52.213 1 T        3.776e+08  0.00e+00 a   0    0     0 0
  12 126.424  67.270 1 T        3.706e+08  0.00e+00 a   0    0     0 0
  13 126.643  59.764 1 T        6.177e+08  0.00e+00 a   0    0     0 0
  14 126.686  43.821 1 T        4.019e+08  0.00e+00 a   0    0     0 0
  15 125.981  53.566 1 T        3.968e+08  0.00e+00 a   0    0     0 0
  16 124.036  22.097 1 T        5.249e+08  0.00e+00 a   0    0     0 0
  17 122.955  60.715 1 T        8.797e+08  0.00e+00 a   0    0     0 0
```

H-bonds (hbonds)

```
B226 ASN  H     B261 THR  O      2.00  10
B262 ASN  H     B225 ARG  O      0.00
B226 ASN  N     B261 THR  O      3.00  10
B262 ASN  N     B225 ARG  O      0.00

A226 ASN  H     B261 THR  O      2.00  10
B262 ASN  H     A225 ARG  O      0.00
A226 ASN  N     B261 THR  O      3.00  10
B262 ASN  N     A225 ARG  O      0.00

B226 ASN  H     C261 THR  O      2.00  10
C262 ASN  H     B225 ARG  O      0.00
B226 ASN  N     C261 THR  O      3.00  10
C262 ASN  N     B225 ARG  O      0.00

....
```

intramolecular

intermolecular

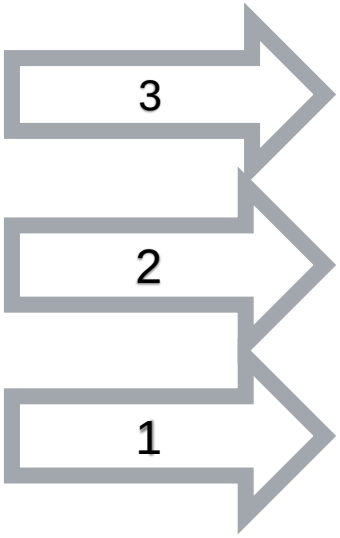
.upl

```
B226 ASN  H     B261 THR  O      1.80  10
B262 ASN  H     B225 ARG  O      0.00
B226 ASN  N     B261 THR  O      2.70  10
B262 ASN  N     B225 ARG  O      0.00

A226 ASN  H     B261 THR  O      1.80  10
B262 ASN  H     A225 ARG  O      0.00
A226 ASN  N     B261 THR  O      2.70  10
B262 ASN  N     A225 ARG  O      0.00

B226 ASN  H     C261 THR  O      1.80  10
C262 ASN  H     B225 ARG  O      0.00
B226 ASN  N     C261 THR  O      2.70  10
C262 ASN  N     B225 ARG  O      0.00

....
```

.lol

init.cya

```
name:=HETs218-289
rmsdrange:=B226-B245,B262-B281

welldefined:="A226-A245,A262-A281,B226-B245,B262-B281,C226-C245,C262-C281"

cyanalib
read seq HETs_trimer.seq

molecules define A218-A289 B218-B289 C218-C289

# Symmetry-related distances
molecules contacts 1,1=2,2=3,3 \
            1,2=2,3 \
            2,1=3,2
molecules identity "* $welldefined" #info=full
weight_ide=0.15
molecules symdist "CA $welldefined" "CA $welldefined" number=3000 #info=full
weight_sym=0.05

# Allowed contacts for distance restraint assignment
molecules contacts 1,1=2,2=3,3 \
1,2=2,3 \
2,1=3,2 \
1,3=3,1=0
```



dihedral angles in “welldefined” regions have to be equal in all monomers

identical interfaces between monomers, restraints are applied between all of them

=0 : no restraints allowed

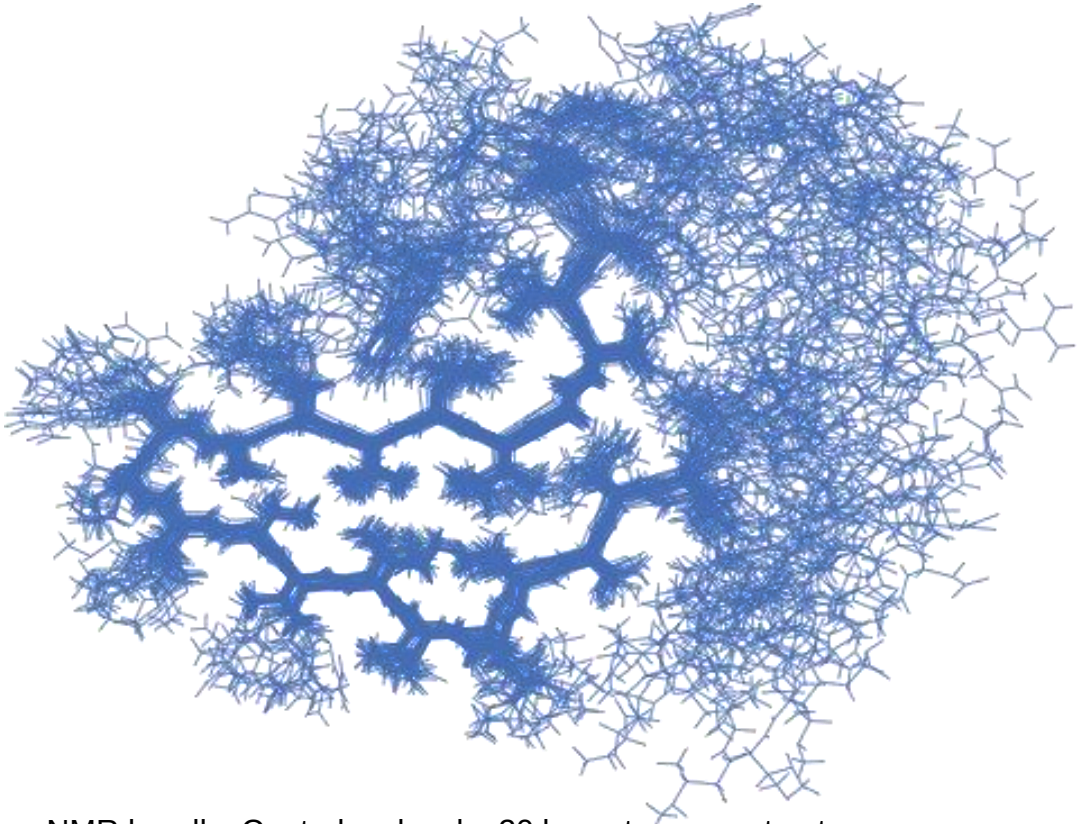
Auto.cya

```
peaks       := PAR4_U_hom.peaks,PAIN5_U_mix.peaks,CHHC_U_hom.peaks,NHHC_U_hom.peaks,NHHC_U_mix.peaks
# names of peak lists

prot        := HETs_trimer.prot
# names of chemical shift lists

restraints  :=
SIhbonds_trimer.lol,SIhbonds_trimer.upl,HETs_talos_trimer_2std.aco,CHHC_U13C_diluted_trimer.upl,CHHC_U1
3C_trimer.upl,NHHC_U13C15N_trimer.upl,PDSD_1_3C_diluted_trimer.upl,PDSD_2C_diluted_trimer.upl
# additional constraints

#tolerance   := 0.3,0.3,0.3
#chemical shift tolerances
# order: 1H(a), 1H(b), 13C/15N(a), 13C/15N(b)

#calibration_constant := 2E+10
# NOE calibration parameters

calibration_upl := 7.5,7.5,5.5,5.0,4.0
# larger numbers lead to longer constraints

upl_values := 1.0,12.0
structures  := 100,10
steps       := 20000
rmsdrange   := 226..245,262..281
randomseed  := 434726

./noeassign peaks=$peaks prot=$prot autoaco
read final.pdb
atoms select "* :A* :B* :C*"
write HETs.pdb all selected
```

Here are the u_{ij} defined!

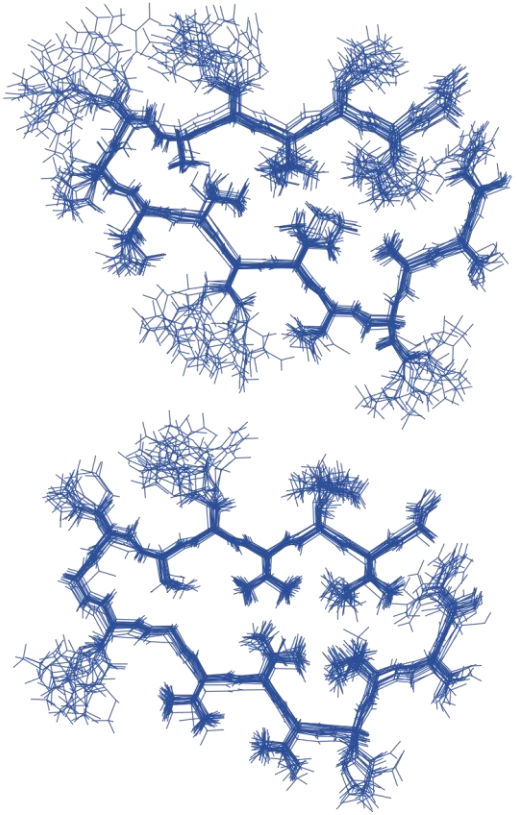
Result of structure calculation for HET-s(218-289)

Overall shape and hydrophobic core



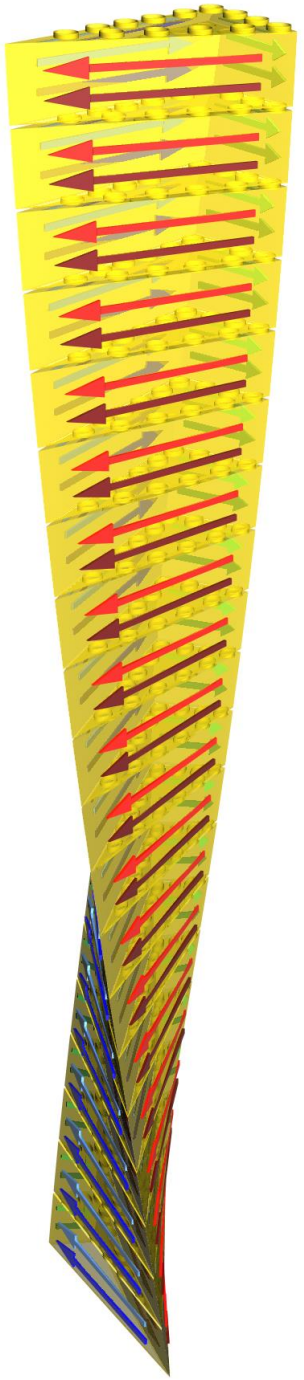
NMR bundle: Central molecule, 20 lowest energy structures

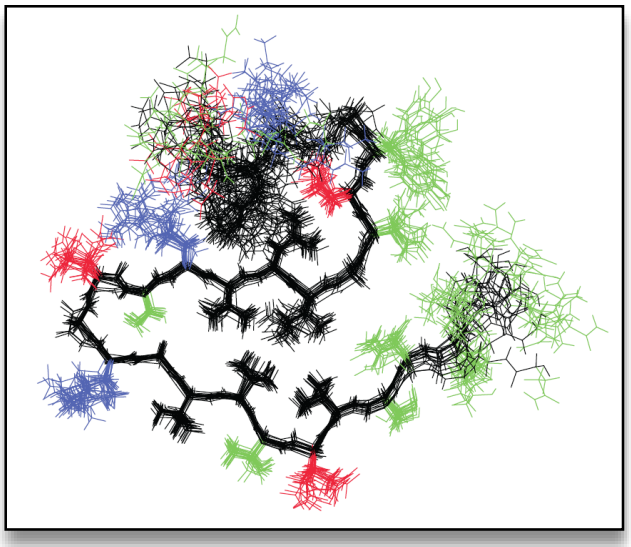
rmsd (for assigned residues:
S226-A248, T260-G282, F286-W287) :
Backbone 0.63 A, All heavy atoms 1.16 A

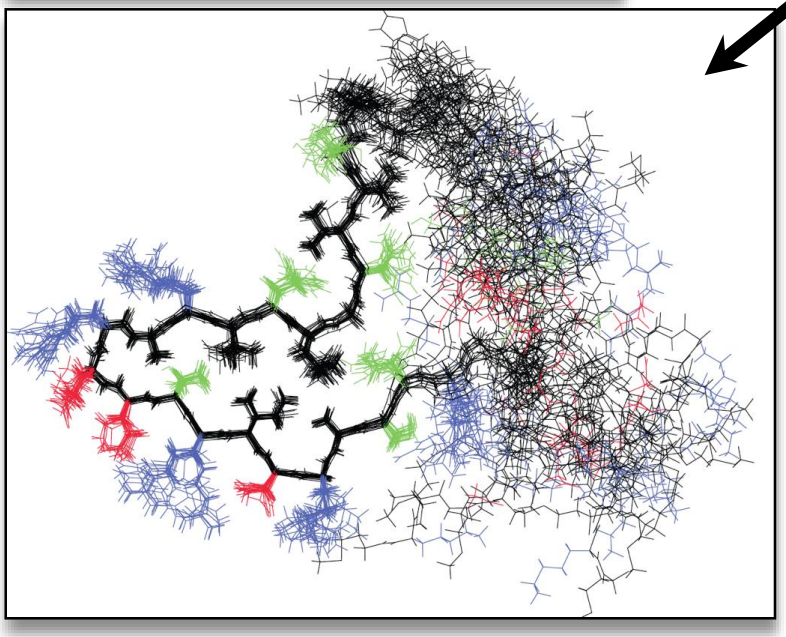


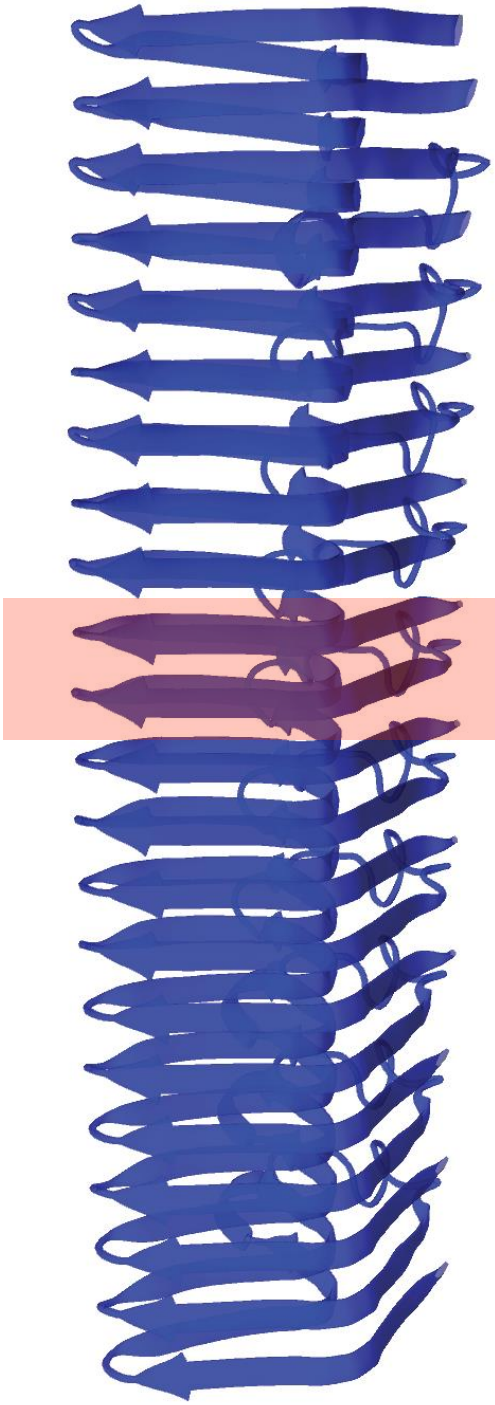
Separate bundles for the two layers

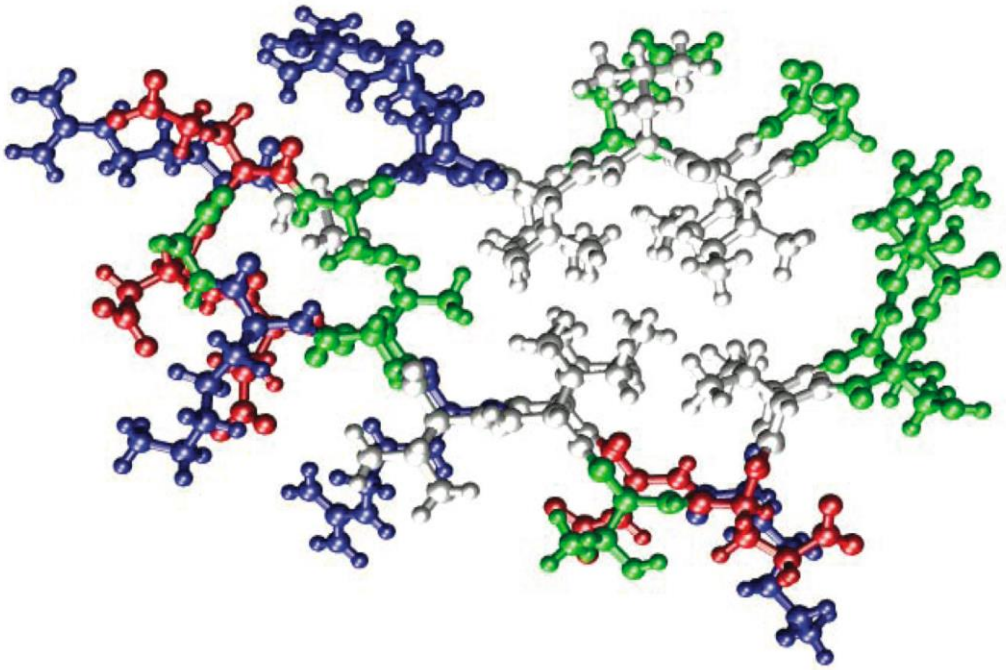
Structure of the entire Object, e.g. a Fibril



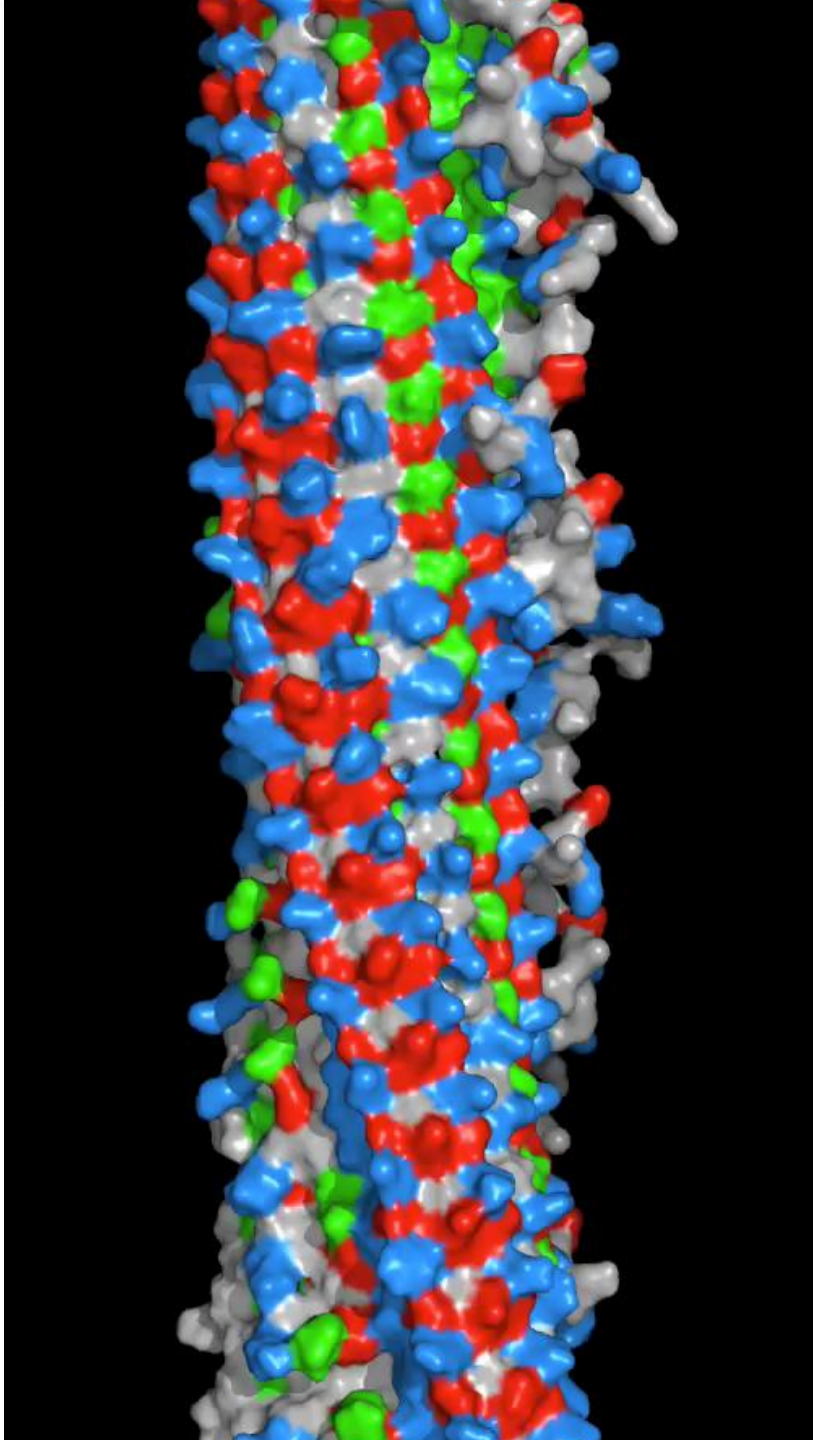








- **At least 23 H-bonds per monomer**
- **A hydrophobic core**
- **Polar and charged residues point “outside”.**
- **Charge compensation between the two layers of a monomer (which are pseudorepeats).**
- **Aspargine ladders**



Take Home Messages

- CYANA handles assignment ambiguities by network anchoring, chemical shift matching and restraint combination
- CYANA calculations are performed in Torsion Angle Space
- CYANA calculations use simulated annealing like in MD
- CYANA is based on the minimization of a target function (pseudoenergy)
- Validation of the obtained structure is very important!!!
- Exercise: How is the structure calculation performed in detail?

Recommended literature on CYANA

Güntert, P. & Buchner, L. Combined automated NOE assignment and structure calculation with CYANA. *Journal of Biomolecular NMR* **62**, 453-471 (2015)

Güntert, P. Automated NMR Structure Calculation With CYANA. in *Protein NMR Techniques* (ed. Downing, A.K.) 353-378 (Humana Press, Totowa, NJ, 2004).