From NMR spectra to structures

- find conditions under which the protein does not aggregate and is reasonably stable
- measure NMR spectra
- sequence-specific sequential resonance assignment
- identify spin systems
- link spin systems (sequential assignment)
- (stereospecific assignment of diasterotopic protons)
- fully interpret NOESY spectrum
- convert NOESY peak amplitudes into distances
- calculate 3D structure and refine the output
Chemical shifts (frequencies) and line intensities

- Backbone HN
- Sidechain HN
- Aromatic
- H\(\alpha\)
- Aliphatic
- Me groups
NOE information is used to introduce distance restraints into the structure calculations.

\[ \text{NOE} \sim B_{\text{loc}}^2 \]

\[ B_{\text{loc}} \sim \frac{\gamma_I\gamma_S(3\cos^2\phi-1)}{r^3} \]
From distances to 3D structures (II)

- Restrained molecular dynamics:

\[ F = m \frac{\partial^2 r}{\partial t^2} \]

\[ F = \frac{\partial U_{pot}}{\partial r} \]
\[ U_{pot} = U_{bond} + U_{angle} + U_{dihedral} + U_{chiral} + U_{v.d.Waals} \]

\[ + U_{coulomb} + U_{NMR} \]

\[ U_{NMR} = U_{NOE} + U_{J} + \ldots \]
The origin of the NOE is dipolar (through-space) coupling of protons

\[ B_{\text{loc}} \sim \frac{\gamma_I \gamma_S (3 \cos^2 \phi - 1)}{r^3} \]

\[ R_{2,1} \sim B_{\text{loc}}^2 \]
During the structure calculation only rotations about dihedrals are made.
Protein sequence
Chemical shift list
Positions and volumes of NOESY cross peaks

Find new assignments
Evaluate assignments
Structure calculation

NOESY assignments
3D structures
Automated calculation of NMR structures

(a) Chemical shift agreement

\[ |\omega_1 - \omega_A| < \Delta \omega_{tol} \]
\[ |\omega_2 - \omega_B| < \Delta \omega_{tol} \]

(b) Network-anchoring

(c) Consistency with preliminary structure

\[ d_{AB} < d_{max} \]

Güntert, Quarterly Reviews of Biophysics 31 (1998), 145-237
Methods for assigning larger proteins
$^{15}\text{N},^{1}\text{H}$ HSQC spectra are fingerprints of proteins
$^{15}$N spectra of proteins: Sidechain peaks
Sequence-specific resonance assignment in $^{13}$C,$^{15}$N labelled proteins
Use of 3-dimensional (tripleresonance) experiments
Backbone Assignment - 3D Experiments

HNCACB / HN(CO)CACB

HNCO / HN(CA)CO

HN(CACO)NH

15N NOESY

residue i-1   residue i    residue i+1
Backbone Assignment - 3D Experiments

HNCACB / HN(CO)CACB

HNCO / HN(CA)CO

HN(CACO)NH

15N NOESY

residue i-1  residue i  residue i+1  residue j-1  residue j  residue j+1
Backbone Assignment - 3D Experiments

HNCACB / HN(CO)CACB

HNCO / HN(CA)CO

HN(CACO)NH

15N NOESY

residue i-1 residue i residue i+1 residue j-1 residue j residue j+1
sequential assignment strategy: building of fragments

1. picking of HN-C peaks -> spin systems (numbers 1-125)
2. alignment of $^{13}\text{C}$ dimension for linking of correct successors/ predecessors

CACB / COCACB
CG / COCG
NN

3. try to match fragments on amino acid sequence
view of the 3D spectrum in CARA
2D $^1\text{H}-^{15}\text{N}$ HSQC is the root experiment of most of the standard *triple-resonance* ($^1\text{H}$, $^{13}\text{C}$, $^{15}\text{N}$) NMR experiments used for backbone assignment.

All the 3D triple resonance experiments are related by the common $^1\text{H}$, $^{15}\text{N}$ chemical shifts of the HSQC spectra: **AMIDE STRIP**
Backbone Assignment - 3D Experiments

HNCACB / HN(CO)CACB

HNCO / HN(CA)CO

HN(CACO)NH

15N NOESY
Standard Carbon Chemical shifts

Biological Magnetic Resonance Databank (Diamagnetic Shifts - 03/09/2007)

http://www.bmrbr.wisc.edu/ref_info/statful.htm
Standard Carbon Chemical shifts

Biological Magnetic Resonance Databank (Diamagnetic Shifts - 03/09/2007)
Identifying Residue type from chemical shifts

Can easily identify Gly, Ser/Thr and Ala from CB,CA shifts: Gly $C_\alpha \sim 45$ ppm; Ser and Thr can be distinguished by $C_\beta$ shifts: Thr $C_\beta \sim 70$ ppm; Ser $C_\beta \sim 63$ ppm. Ala $C_\beta$ chemical shifts is around $\sim 18$ ppm.

Can group Leu, Tyr, Phe, Asn, Ile and Asp based on their $C_\beta$ shifts $\sim \sim 35$ ppm.

Differentiate between the residues having two (Asp, Asn, Trp, Tyr, Cys, His, Phe) carbons sidechains and those having 3 or more carbons in the sidechain by using $CC(CO)NH$.

Among the residues having three carbon sidechain: Val, Met, Thr, Glu and Gln, Val has most upfield $C_g$ chemical shifts. Ser/Thr can be distinguished $C_g$ shift. Glu and Gln can be identified by their $C_g$ shifts, Glu $C_\gamma > 35$ ppm and Gln $C_\gamma < 35$ ppm.

Residues with four carbons sidechain: Pro and Arg can easily be distinguished by their $C_\delta$ shifts, for Pro $C_\delta \sim 50$ ppm whereas for Arg $C_\delta \sim 43$ ppm.

Among the residues having five carbons side chain: Leu, Lys and Ile. Ile has the most upfield $C_\delta$ shifts $\sim 10$ ppm whereas Leu has $C_\beta \sim 43$ ppm and Lys will have $C_\varepsilon \sim 43$ ppm.
Standard Proton Chemical shifts

Biological Magnetic Resonance Databank (Diamagnetic Shifts - 03/09/2007)

- Val
- Tyr
- Trp
- Thr
- Ser
- Pro
- Phe
- Met
- Lys
- Leu
- Ile
- His
- Gly
- Glu
- Gln
- Cys
- Asp
- Asn
- Arg
- Ala

$^1\text{H}$ Chemical Shifts (ppm)
Side Chain Assignment Strategies

Identification of backbone protons:

\[ \text{HBHA} (\text{CACBCO}) \text{NH} \]

\[ \text{HN} (\text{COCA}) \text{HA} \]

\[ \text{HACACO} \]

Side chain assignment:

\[ \text{hCCH} \]

\[ \text{HCcH} \]
$^{13}$C dimension strips of single spin systems for linking
Prediction of secondary structure using chemical shifts using TALOS
The magnitude of the $^{1}H\{^{15}N\}$-NOE depends on the motional properties
NMR of metallothionineins
Metallothioneins

- Small proteins: ~60 aa with ~30% cysteine residues
- Coordinate metal ions
- No secondary structure elements
- Two metal-thiolate clusters per protein:
  - α-domain: 11 Cys coordinating 4 divalent metal ions
  - β-domain: 9 Cys coordinating 3 divalent metal ions

Crystal structure of MT-2 isolated from rat liver. (Romero-Isart N, Vasák M. J Inorg Biochem. Feb 2002)
Structures of Littorina littorea MT

N-term

center

C-term

center and C-term