# Residual Dipolar Couplings

BCMB/CHEM 8190

# Long-Range Structural NMR Restraints

- Traditional NOE-based protein structure determination methods suffer from the lack of long-range structural restraints
  - in this context, "long-range" means parts of the protein not close in space

- For instance, for two- or multi-domain proteins where there is limited contact between domains, it is often difficult to properly orient domains with respect to one another (using only NOE-based restraints)

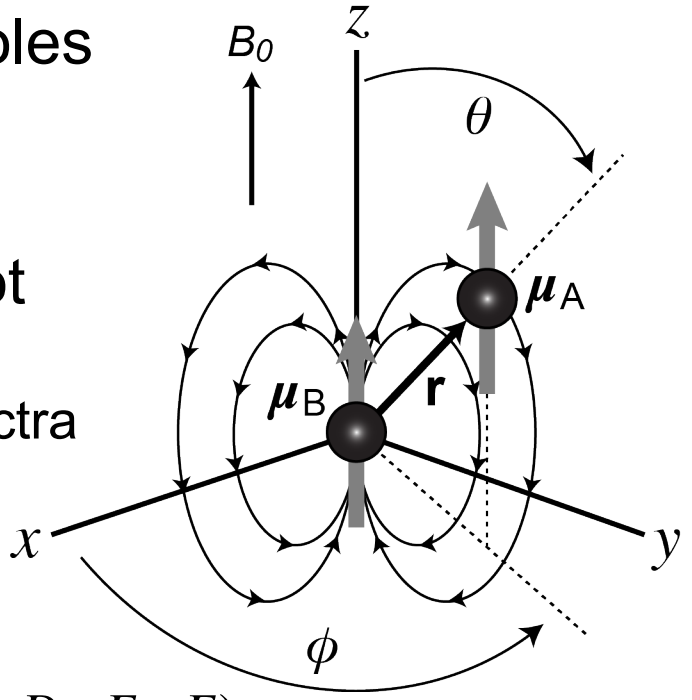
- Residual dipolar couplings (RDCs) afford a route to effective long range orientational (not translational) restraints
  - RDCs allow orientation of bond vectors with respect to a reference axis (typically *z*-axis, $B_0$ direction in the laboratory frame)
  - thus, these restraints effectively constrain bond vectors relative to one another, which amounts to long-range conformational/orientational restraints

# Reviews


- Prestegard, A-Hashimi & Tolman, *Quart. Reviews Biophys.* **33**, 371-424 (2000).
- Bax, Kontaxis & Tjandra, *Methods in Enzymology*, **339**, 127-174 (2001)
- Prestegard, Bougault & Kishore, *Chemical Reviews*, **104**, 3519-3540 (2004)
- Lipsitz & Tjandra, *Ann. Rev. Biophys. Biomol. Struct*., **33**, 387-413 (2004)
- Fushman et al., *Prog. NMR Spect*. **44**, 189-214 (2004)
- Hu & Wang, Ann Rpts NMR Spect, **58**,231-303 (2006)
- Bailor et al., Nature Protocols, **2**, 1536-1546 (2007)

# The Dipole-Dipole Interaction

- Through space interaction between dipoles
- Angular dependence ($\theta$ and $\phi$)
- Distance dependence (r)
- Dipolar splittings average in solution (not observed), but are present in solid state
  - are complicated, and dominate solid state spectra




$$\hat{\boldsymbol{H}}_{\mathrm{D}} = \frac{\mu_0}{4\pi}\frac{(\vec{\boldsymbol{\mu}}_I \cdot \vec{\boldsymbol{\mu}}_S)}{r^3} - \frac{3(\vec{\boldsymbol{\mu}}_I \cdot \vec{\boldsymbol{r}})(\vec{\boldsymbol{\mu}}_S \cdot \vec{\boldsymbol{r}})}{r^5} = \frac{\mu_0 \gamma_I \gamma_S h^2}{(16\pi^3 r^3)}(A + B + C + D + E + F)$$

$$A = -\hat{\boldsymbol{I}}_z\hat{\boldsymbol{S}}_z(3\cos^2\theta - 1),\ B = (1/4)(\hat{\boldsymbol{I}}^+\hat{\boldsymbol{S}}^- + \hat{\boldsymbol{I}}^-\hat{\boldsymbol{S}}^+)(3\cos^2\theta - 1),\ C = -(3/2)(\hat{\boldsymbol{I}}^+\hat{\boldsymbol{S}}_z + \hat{\boldsymbol{I}}_z\hat{\boldsymbol{S}}^+)(\sin\theta\cos\theta e^{-i\phi})$$

$$D = -(3/2)(\hat{\boldsymbol{I}}^-\hat{\boldsymbol{S}}_z + \hat{\boldsymbol{I}}_z\hat{\boldsymbol{S}}^-)(\sin\theta\cos\theta e^{i\phi}),\ E = -(3/4)\hat{\boldsymbol{I}}^+\hat{\boldsymbol{S}}^+\sin 2\theta e^{-2i\phi},\ F = -(3/4)\hat{\boldsymbol{I}}^-\hat{\boldsymbol{S}}^-\sin 2\theta e^{2i\phi}$$

$$\vec{\boldsymbol{r}} = \vec{\mathbf{i}}\, r_x + \vec{\mathbf{j}}\, r_y + \vec{\mathbf{k}}\, r_z = \vec{\mathbf{i}}\, r\sin\theta\cos\phi + \vec{\mathbf{j}}\, r\sin\theta\sin\phi + \vec{\mathbf{k}}\, r\cos\theta$$

- Important parts: $1/r^3$ dependence, angular dependence

# First Order QM Term ($I_{z1}I_{z2}$) is Most Important

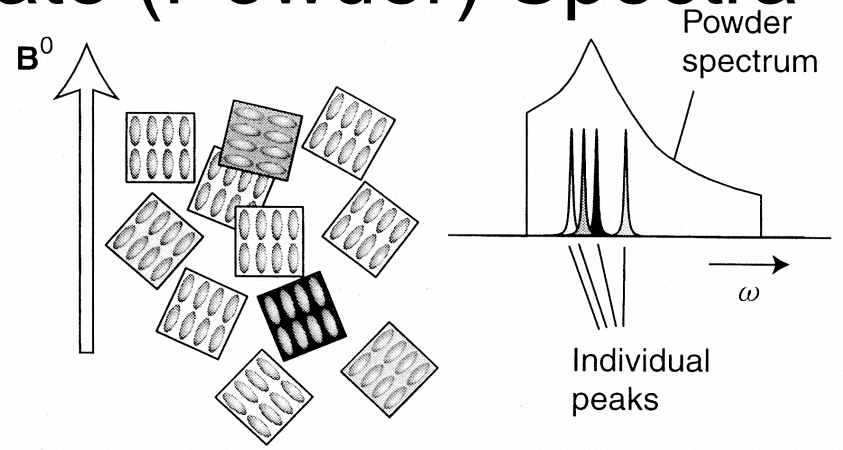
- Consider signal splitting due to dipolar coupling (solid state)
  - if we assume interaction of dipoles between two nuclear dipoles is small compared to the interaction of the dipoles with $B_0$ (good assumption), then can ignore all terms except 'A' term
  - we can then use the simple two-spin wavefunction basis for a Hamiltonian that includes Zeeman interactions with $B_0$ plus the (simplified) dipolar term

$$\hat{\boldsymbol{H}}_{\mathrm{D}} = \frac{\mu_0 \gamma_I \gamma_S h^2}{(16\pi^3 r^3)}(-\hat{\boldsymbol{I}}_z \hat{\boldsymbol{S}}_z (3\cos^2\theta - 1))$$

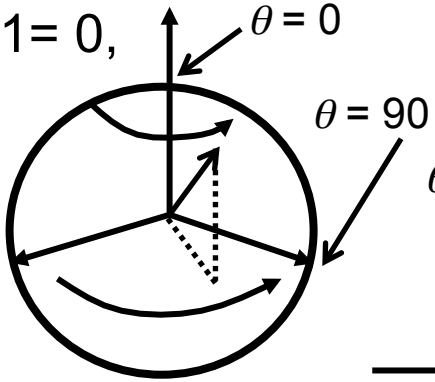
- In solid state, result is doublet with $1/r^3$ and $(3\cos^2\theta - 1)$ dependencies
  - in solid state (powder), all possible values of $\theta$ are present, so result is a superposition of all possible signals with all possible couplings
  - splittings are large (~60 kHz for $^{13}C$-$^{1}H$, ~250 kHz for $^{1}H$-$^{1}H$)
  - splittings are angle dependent (-60 kHz to +30 kHz for $^{13}C$-$^{1}H$)
- In solution, $\theta$ is averaged, so no splitting is observed

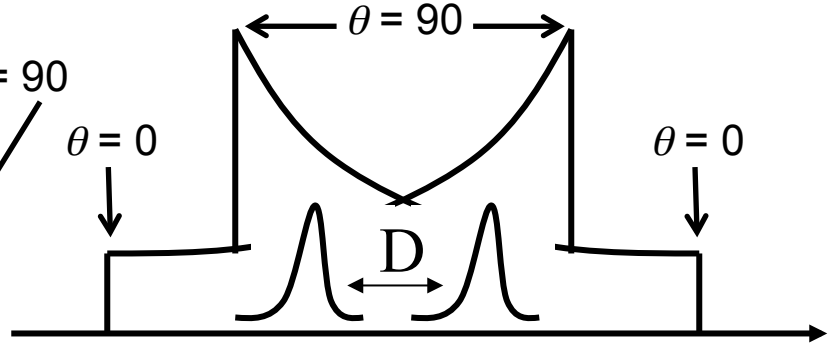
# Dipolar Splitting in Solid-State (Powder) Spectra

- In the solid state, chemical shift anisotropy leads to broad signals due to presence of signals from all possible molecular orientations
- Dipolar splittings result in very broad doublets




- Consider various values of $\theta$ (and all possible values of $\phi$)
  - values of $\theta$ and $\phi$ define a sphere (for some fixed r, i.e. C-H bond))
  - *largest splittings* (most positive) *are when* $\theta = 0$ ($[3\cos^2\theta - 1]$ is largest when $\theta = 0°$): this is a single point on the sphere, so, not highly populated (low intensity, but largest splitting, ~60 kHz for C-H)
  - *smallest splittings* (most negative) *are when* $\theta = 90°$ ($[3\cos^2\theta - 1]$ is smallest when $\theta = 90°$): this is around the equator of the sphere, so, highly populated (high intensity, but smallest (most negative) splitting)
  - when $\theta = 54.7°$, $3\cos^2\theta - 1 = 0$, so no splittings at that value of $\theta$



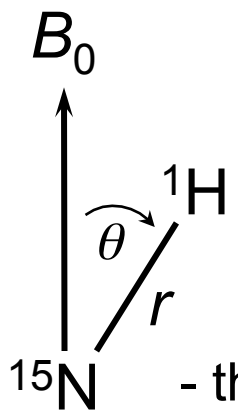



# The Dipolar Interaction Between Two Spins: Observing Dipolar Couplings in Solution

- In solution, $\theta$ is averaged, so no splitting is observed
  - molecular tumbling is very fast, and strict averaging ($\langle\ \rangle$) precludes observation of dipolar couplings in solution

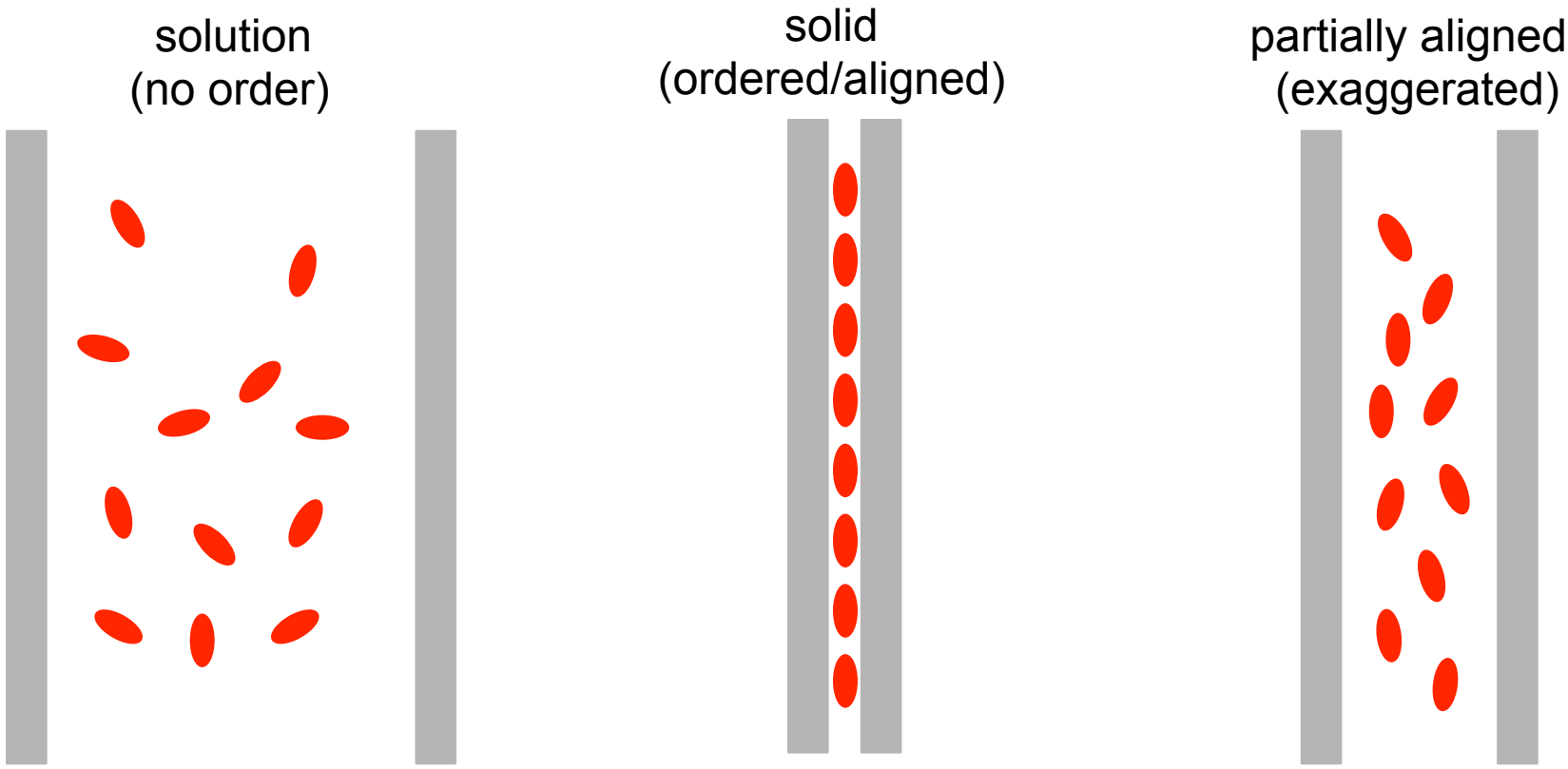
$$D = \frac{C}{r^3}\left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle \hat{\boldsymbol{I}}_{Hz}\hat{\boldsymbol{I}}_{Nz}$$

- Consider the advantages of "reintroducing" a *small* degree of motional anisotropy ("partial alignment") of the protein molecules in the magnetic field
  - each amino acid has a N-H bond, so could get angle between each vector and $B_0$ (essentially excellent long-distance conformational restraint information
  - a *small* alignment of the protein should NOT result in significantly increased linewidths, and observed coupling constants would be much smaller (on the order of *J* couplings) than the huge couplings observed in solids
  - the simplified dipolar coupling contribution to the Hamiltonian (just $\boldsymbol{I}_z\boldsymbol{I}_z$, like *J* coupling), indicates that the dipolar coupling would just add to the scalar coupling (measured as change in scalar coupling)
  - so, all we have to do is reintroduce a small degree of motional anisotropy



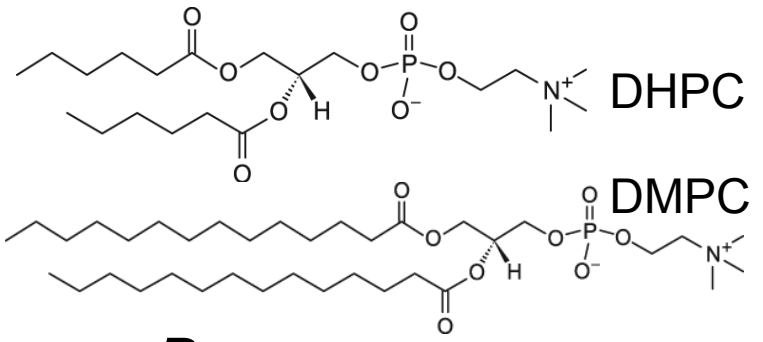
# Inducing Partial Alignment

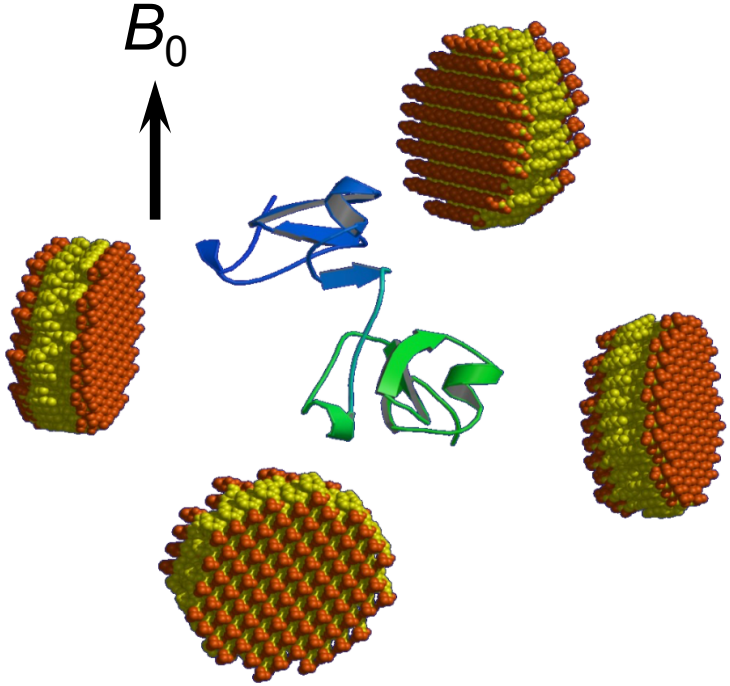
- A small degree of alignment (1/10$^{th}$ of 1% deviation from isotropic) of the proteins in the sample allows measurement of dipolar couplings (1/10$^{th}$ of 1% of 10 kHz is 10 Hz)
  - other mechanisms (electrostatics) can also contribute to alignment
  - dipolar coupling observed as change in scalar (*J*) coupling in the isotropic versus partially aligned state



# Inducing Order Using Liquid Crystalline Media

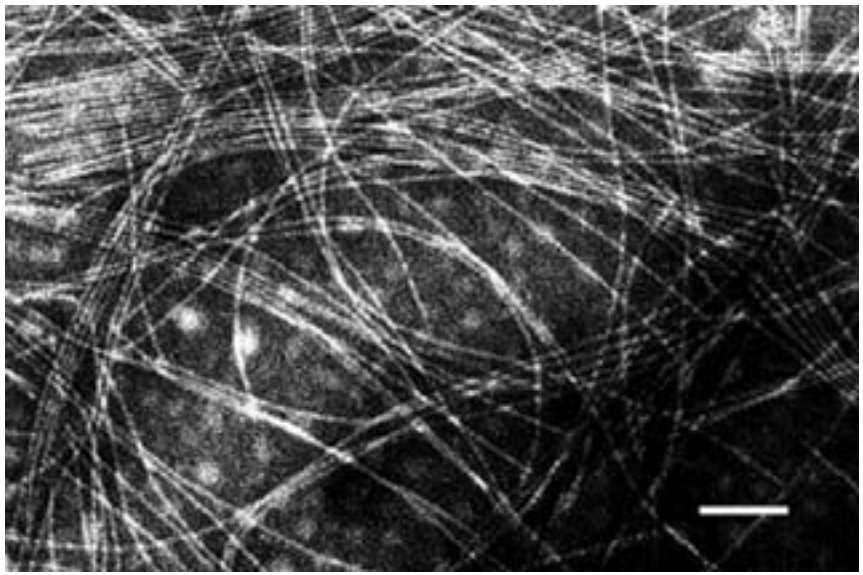
- Partial alignment is induced by adding protein to any of a number of types of media that promote alignment
  - liquid crystalline media, such as bicelles formed by lipids (DHPC/DMPC) or filamentous bacteriophage are some of the first and still commonly used



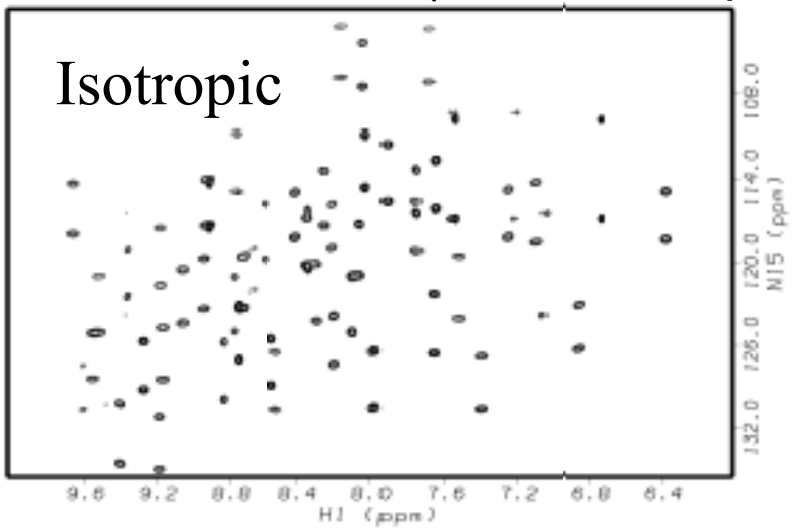

- these large have a large net induced magnetic moment in a magnetic field that causes them to align in one direction in a magnetic field
- mechanical interactions of the proteins with these particles promote the small net alignment of the proteins

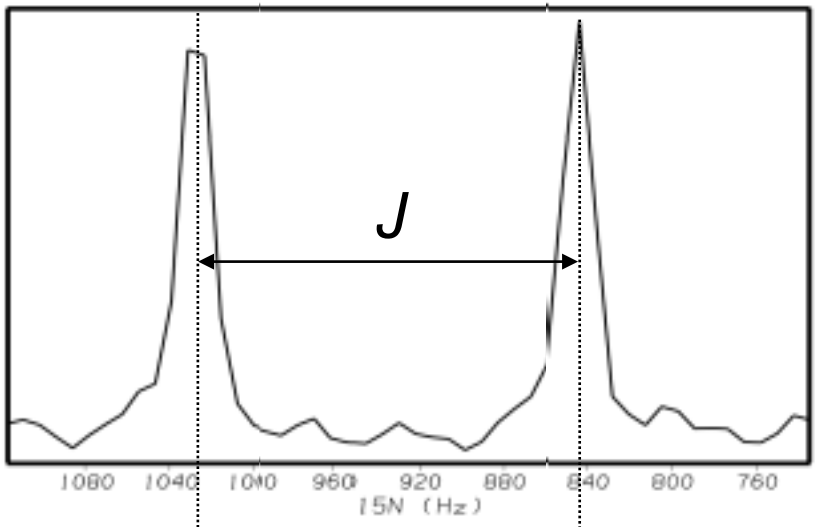
filamentous bacteriophage (Pf1)

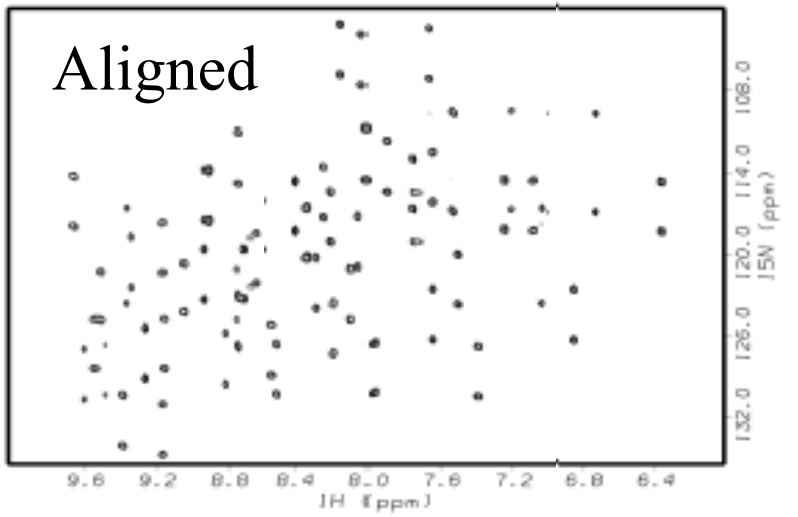


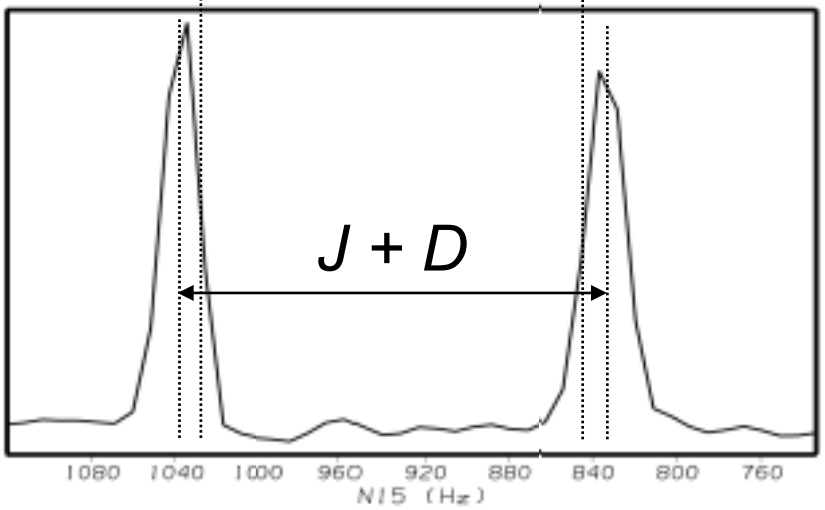
# Measuring Dipolar Couplings - Coupled HSQC

- Are many ways to measure the residual dipolar couplings
  - simplest is just to record HSQC spectra without $^{1}H$ decoupling during $^{15}N$ evolution, under both isotropic and aligned conditions
  - difference in measured splitting is dipolar contribution
  - better methods (IPAP, decoupled HSQC plus TROSY)

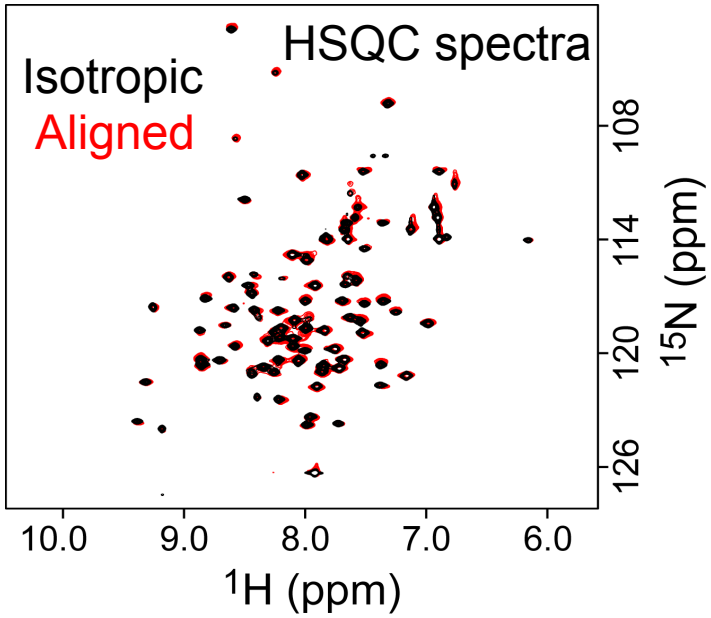


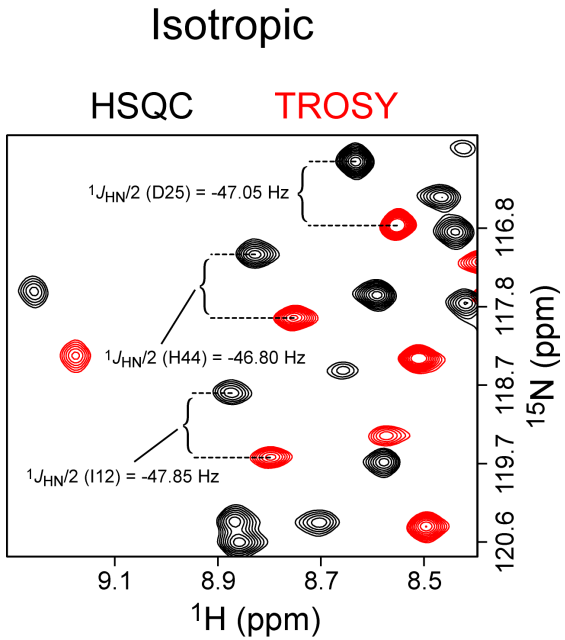



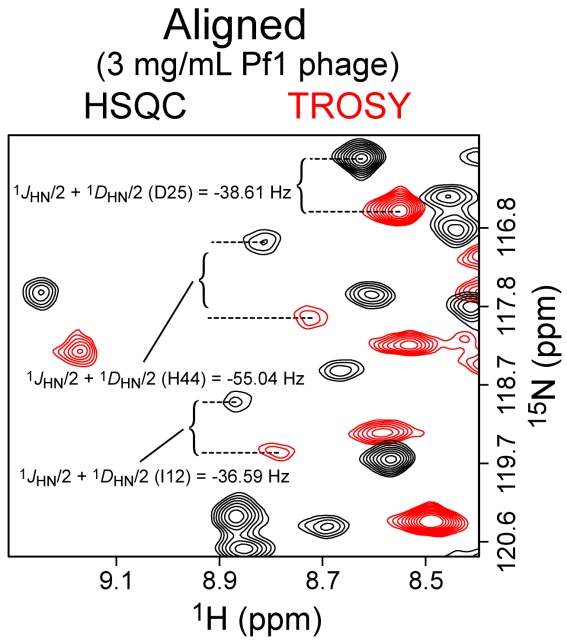



# Measuring Dipolar Couplings - Coupled HSQC




- Important to remember:
  - cannot trust measurements if alignment medium changes the protein structure (chemical shifts for each signal in HSQC spectrum for isotropic state must be identical to those in aligned state)
  - $^{1}$H-$^{15}$N scalar coupling constants are *negative* ($^{1}$H-$^{15}$N dipolar couplings are both positive and negative)



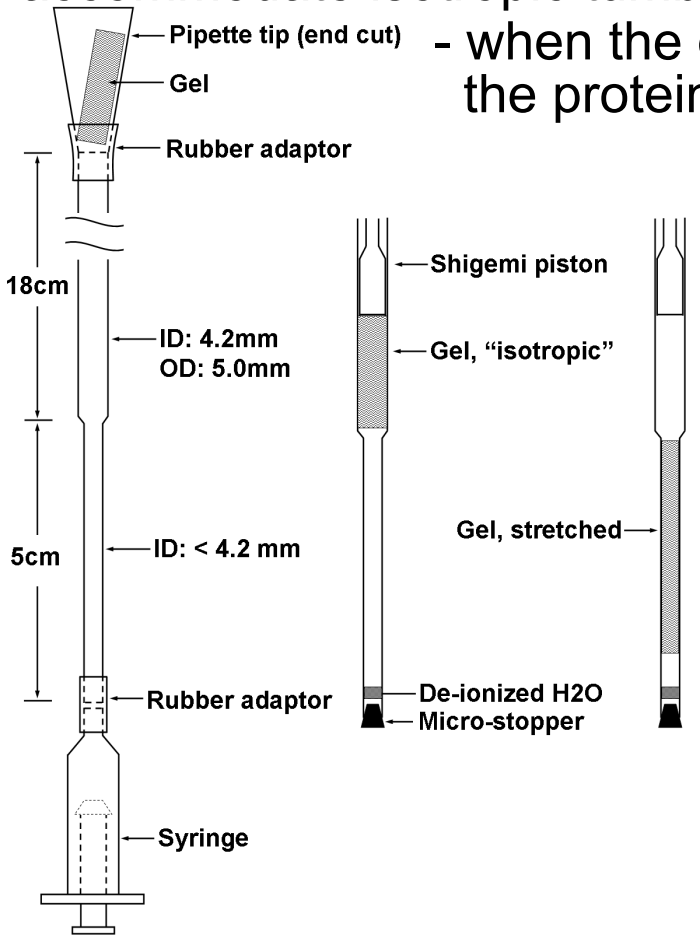

$^{1}D_{HN}$ (D25) = 16.88 Hz

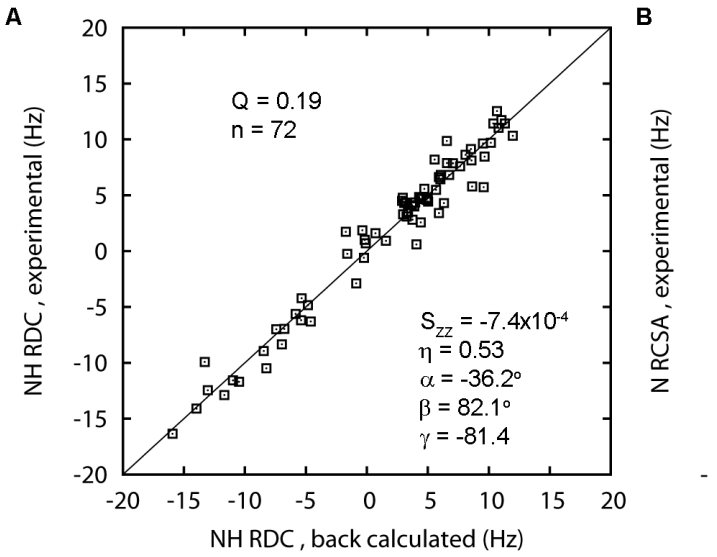
$^{1}D_{HN}$ (H44) = -16.48 Hz

$^{1}D_{HN}$ (I12) = 22.52 Hz

# Polyacrylamide Gels (Another Alignment Medium)

- Liquid crystalline media can be complicated
  - concentration, temperature dependencies, some very viscous, sometimes structural changes, or tight interactions (disappearing signals), etc.
- Nice alternative is stretched polyacrylamide gels
  - at correct percentage of crosslinking (etcetera), proteins tumble isotropically in these gels (spaces in the gels large enough to accommodate isotropic tumbling)
  - when the gels are stretched, the spaces get smaller and the proteins are mechanically aligned



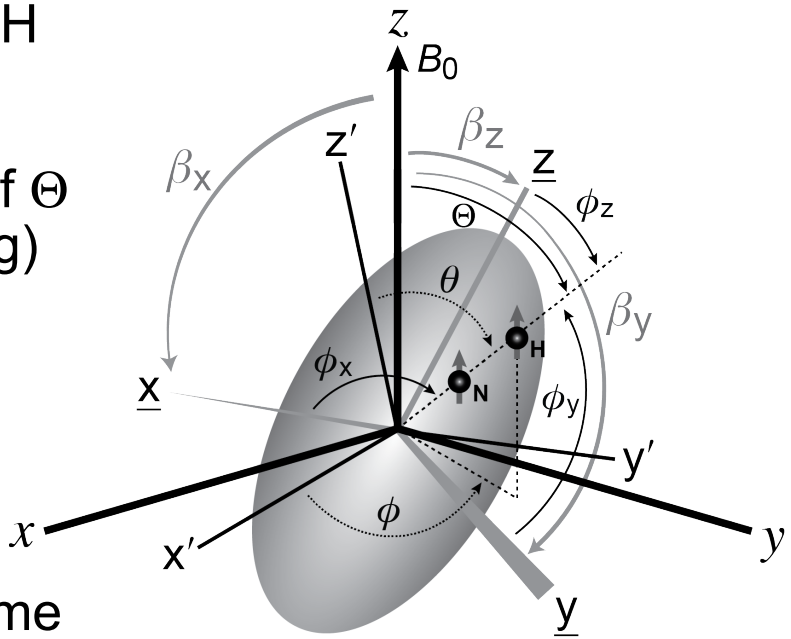

Yizhou Liu, J. Prestegard (2010)
*J. Biomol NMR*, **47**: 249-258.

# Order Matrix Analysis

- How do we get the angular restraint information from the dipolar coupling constants?
  - the alignment media introduces a very small degree of anisotropic tumbling of the protein molecule
  - have to take into account averaging resulting from molecular tumbling, so the result is an average of a very small degree of rotational anisotropy
- The equation for the residual dipolar coupling is as follows

$$D = \frac{-\mu_0 \gamma_I \gamma_S h}{(8\pi^3 r_{IS}^3)} \left\langle \frac{3\cos^2\Theta - 1}{2} \right\rangle = D_{\max} \left\langle \frac{3\cos^2\Theta - 1}{2} \right\rangle$$

  - here $\Theta$ is the angle between $B_0$ and the N-H bond vector, which varies with time as the molecule tumbles
  - for isotropic tumbling, all possible values of $\Theta$ are equally probable (no observed coupling)
- The N-H bond vector is fixed with respect to a molecular frame (x, y, and z axes, light colored)
  - molecular tumbling in the laboratory frame ($x$, $y$, and $z$ = $B_0$ axes, wide dark lines) is equivalent to tumbling of the molecular frame in the laboratory frame



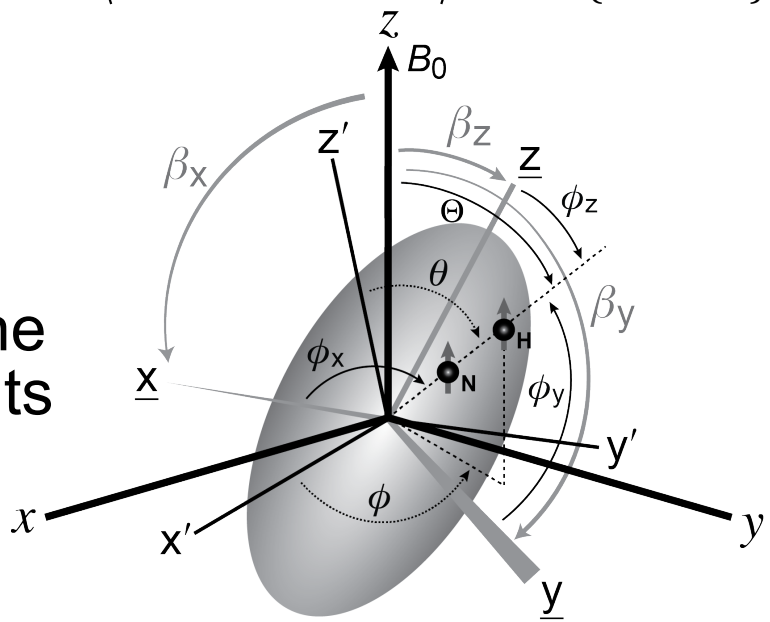
# Order Matrix Analysis

- The RDC equation can be recast in terms of 3×3 matrix ***S*** ("Saupe matrix", "order matrix", "alignment tensor")
  - the anisotropic averaging is represented in this matrix, and structural and motional parameters can be extracted using this approach
  - the elements of ***S*** ($S_{ij}$) describe the projection of the axes of the molecular frame on the $B_0$ ($z$) axis of the laboratory frame
  - the $\cos\phi_i\cos\phi_j$ terms are directional cosines describing the orientation of the N-H bond vector in the molecular frame ($\phi_x$, $\phi_y$, $\phi_z$)

$$D = D_{\max}(x\ y\ z)\begin{bmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = D_{\max}\sum_{i,j=\{x,y,z\}} S_{ij}\cos\phi_i\cos\phi_j \quad S_{ij} = \left\langle \frac{3\cos\beta_i\cos\beta_j - \delta_{ij}}{2} \right\rangle \quad \delta_{ij} = \begin{cases} 1\ (i=j) \\ 0\ (i\neq j) \end{cases}$$

  - the order matrix is symmetric and traceless ($S_{xx}+S_{yy}+S_{zz}=0$) so contains only *five independent elements*

- With enough measured RDCs ($D$), one can solve for the order matrix elements

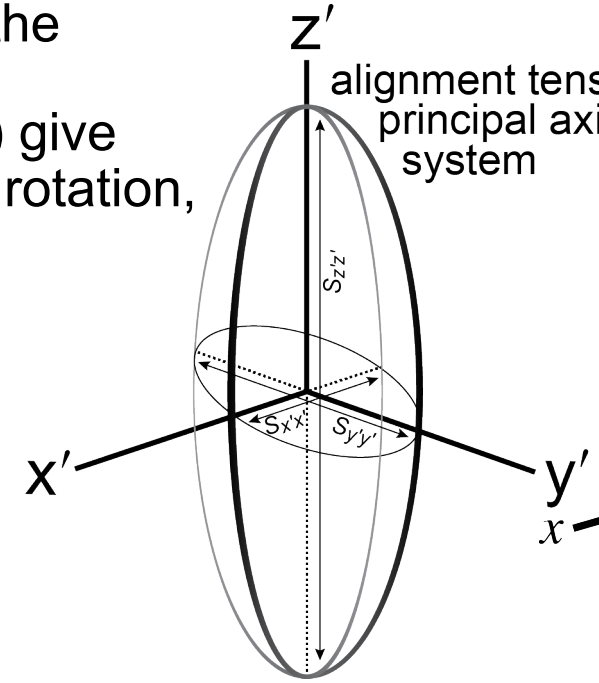


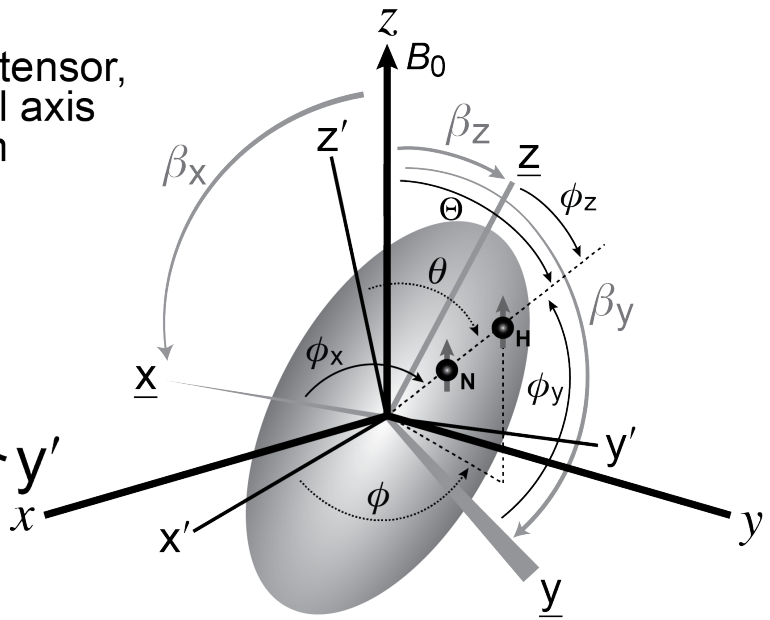
# Order Matrix Analysis

- It is typically convenient to work in the principal axis frame
  - the molecular frame (x, y, z) axes system can be rotated to diagonalize the order (***S***) matrix
  - the new axes system (x', y', z') referred to as the principal axis frame
  - the *eigenvalues* define the axial ($S_{z'z'}$) and rhombic ($S_{x'x'}$, $S_{y'y'}$) components of the alignment tensor ( $|S_{z'z'}| > |S_{y'y'}| > |S_{x'x'}|$ )
  - the diagonal S matrix is also traceless, so $S_{x'x'}$, $S_{y'y'}$ and $S_{z'z'}$ are not independent

$$\begin{bmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{bmatrix} = [A] \begin{bmatrix} S_{x'x'} & & \\ & S_{y'y'} & \\ & & S_{z'z'} \end{bmatrix} [A^{-1}]$$

- the eigenvectors (from the matrix ***A*** that effects the rotation/diagonalization) give the Euler angles for the rotation, and therefore relate the molecular frame to the principal axis system



# Order Matrix Analysis

- The orientation of the (N-H) bond vector to the principal axis frame can be defined using spherical coordinates
  - the angle $\theta$ is the angle from the z axis (z') of the principal axis system to the internuclear bond vector (polar angle)
  - the angle $\phi$ is the angle from the x axis (x') of the principal axis system to the internuclear bond vector projection on the x'-y' plane (azimuth angle)
- The equation for RDCs can then be written in a usable form, in the principal alignment axis frame as:

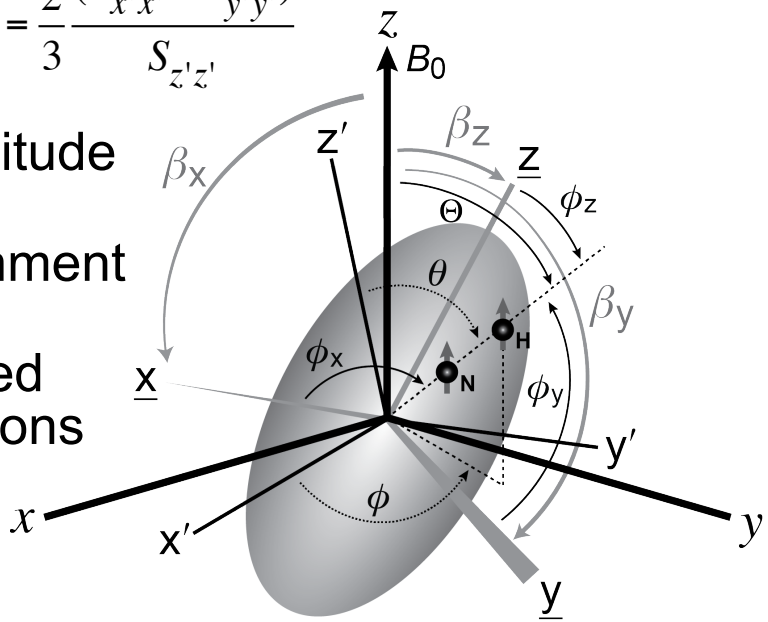
$$D = D_{\max}\left\{S_{z'z'}\frac{1}{2}(3\cos^2\theta - 1) + (S_{x'x'} - S_{y'y'})\frac{1}{2}\sin^2\theta\cos 2\phi\right\} \qquad D_{\max} = -\frac{\mu_0\gamma_I\gamma_S h}{8\pi^3 r_{IS}^3}$$

  - this is typically rewritten as:

$$D = D_a\left\{(3\cos^2\theta - 1) + \frac{3}{2}R\sin^2\theta\cos 2\phi\right\} \qquad D_a = \frac{D_{\max}S_{z'z'}}{2} \qquad R = \frac{2}{3}\frac{(S_{x'x'} - S_{y'y'})}{S_{z'z'}}$$

  - $D_a$ is the axial component of the alignment tensor, and is often referred to as the magnitude of the tensor (subsumes constants in $D_{\max}$)
  - $R$ is the rhombicity (asymmetry) of the alignment tensor (deviation from axial symmetry)
  - $D$ is usually written to include the generalized order parameter to account for fast fluctuations of the bond vector (when necessary)

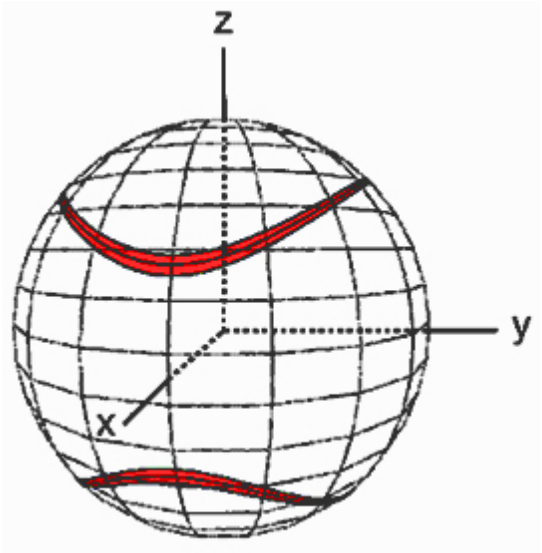
$$D = D_a S\left\{(3\cos^2\theta - 1) + \frac{3}{2}R\sin^2\theta\cos 2\phi\right\}$$

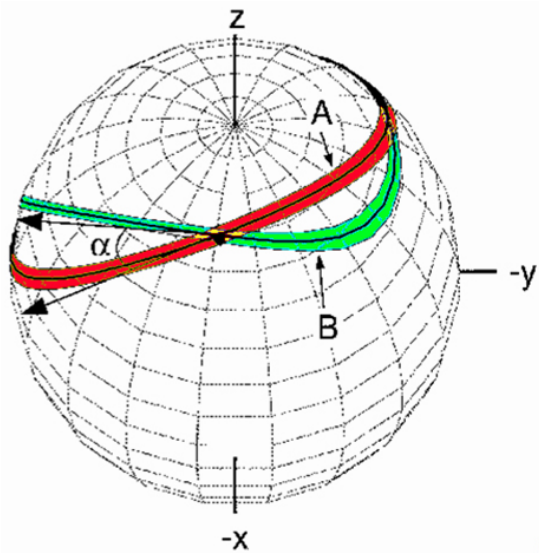


# Order Matrix Analysis

$$D = D_a S\left\{(3\cos^2\theta - 1) + \frac{3}{2}R\sin^2\theta\cos 2\phi\right\}$$

- Dipolar couplings from a single alignment do not yield unique solutions
  - the angular dependencies give rise to a distorted, continuous "cone" of solutions and its inverse (two ranges of solutions, below left)
  - RDCs from a second alignment can restrict the possible solutions (provided the second alignment is significantly different than the first)

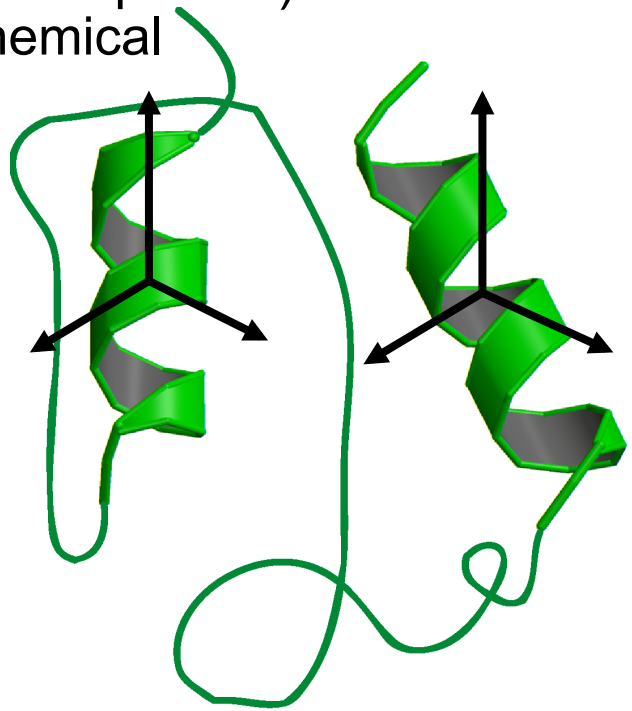



*https://spin.niddk.nih.gov/bax/software/NMRPipe/mfrintro/*

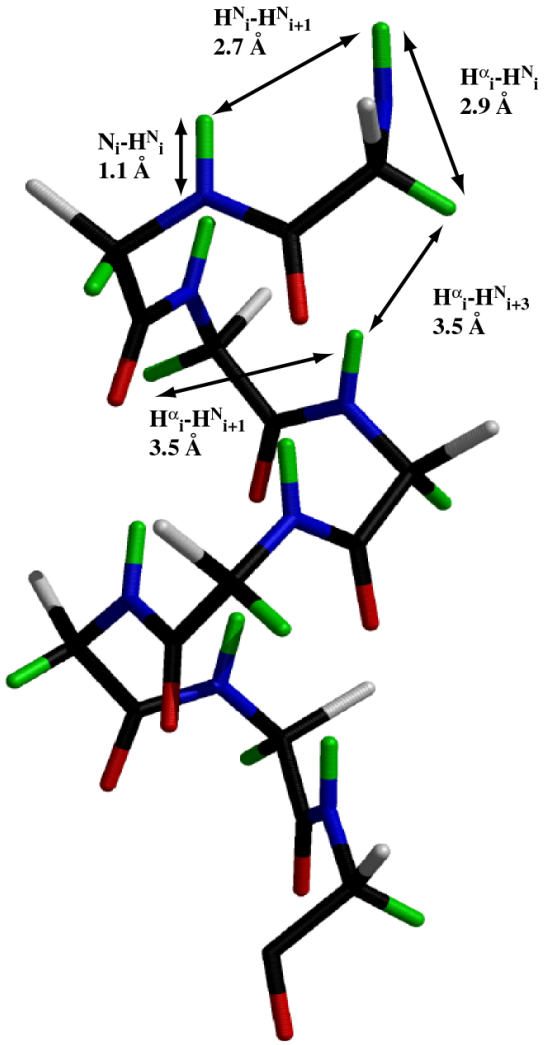
# Strategy for Protein Fold Determination

- Example: properly orient the two helices in a small protein
  - if secondary structural elements (helices) can be identified (chemical shifts), and the protein is rigid, the helices will have the same alignment frame (principal axis frame)
  - if enough RDCs (~20) can be collected for each helix, then can superimpose the alignment frames and orient the helices with respect to one another
- Protocol
  - express $^{15}$N-labeled protein
  - assign resonances (probably need $^{13}$C, $^{15}$N-labeled protein)
  - identify secondary structure elements (using chemical shifts, $^{1}H^{N}$-$^{1}H^{N}$ NOEs, $^{1}H^{\alpha}_{i}$-$^{1}H^{N}_{i+3}$ NOEs)
  - assume ideal geometry for helices
  - orient protein (perhaps liquid crystal medium)
  - collect RDCs
  - orient individual elements
  - assemble protein fold



# Measurable Dipolar Couplings

- RDCs involving $^{15}N$ (i.e. $^{1}H^{N}_{i}$-$^{15}N_{i}$) are easy to measure and provide useful information

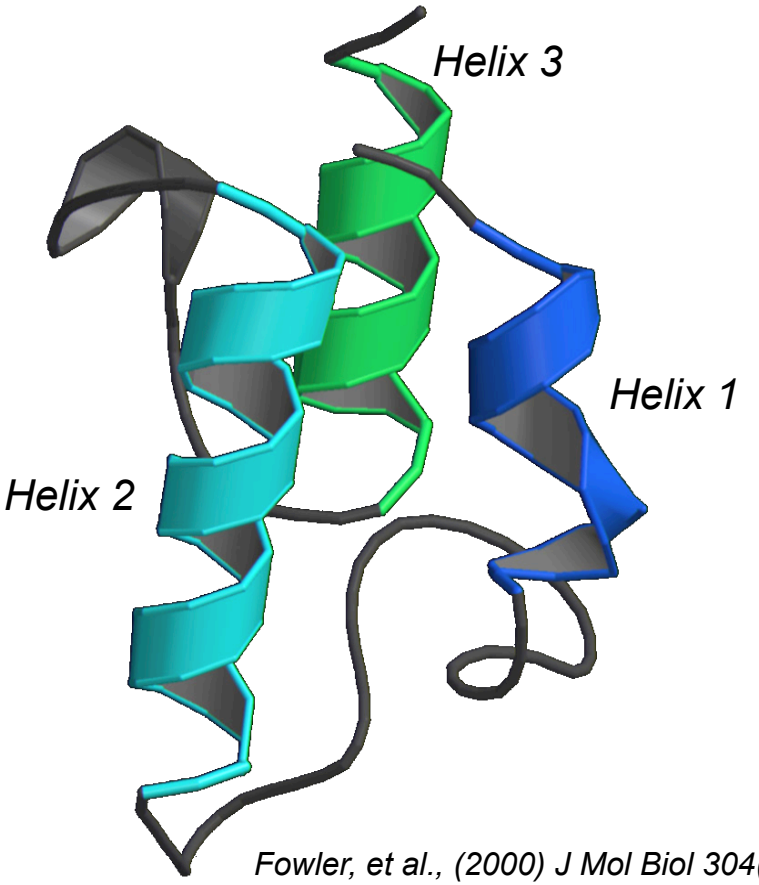



- In order to get ~20 or more RDCs per helical element, need more than just $^{1}H^{N}_{i}$-$^{15}N_{i}$ RDCs
- Are also many other couplings that are useful and can be measured
  - $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i}$, $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i+1}$, $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i-1}$, $^{1}H^{N}_{i}$-$^{1}H^{N}_{i+1}$, $^{1}H^{\alpha}_{i}$-$^{13}C^{\alpha}_{i}$, $^{1}H^{N}_{i}$-$^{13}C{=}O_{i-1}$, $^{15}N_{i}$-$^{13}C{=}O_{i-1}$

# Acyl Carrier Protein Structure using RDCs

- Test system: Acyl Carrier Protein (ACP), 77 residues
- Goal is to properly orient the three alpha helices with RDCs

  - ACP aligned in 20 mg/mL Pf1 bacteriophage
  - $^{1}H^{N}_{i}$-$^{15}N_{i}$, $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i}$, $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i+1}$, $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i-1}$, $^{1}H^{N}_{i}$-$^{1}H^{N}_{i+1}$ RDCs measured




*Fowler, et al., (2000) J Mol Biol 304(3):447-460*

# Data Used in ACP Fold Determination

- Table of values for RDCs measured for the three helices of ACP
  - note: as mentioned previously, values of RDCs can be positive or negative
  - have to be careful of signs of couplings: constants include gyromagnetic ratios of the nuclei (can be positive or negative)

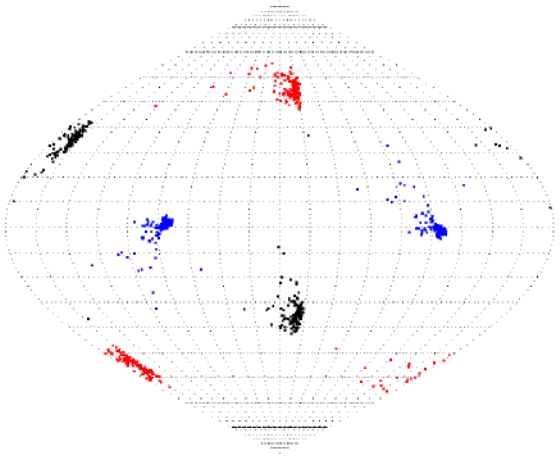
**Table 1.** Data used to determine the backbone fold of ACP *Fowler, et al., (2000) J Mol Biol 304(3):447-460*

| | Couplings (Hz) Helix 1 | | Helix 2 | | Helix 3 | |
|---|---|---|---|---|---|---|
| A. *Dipolar couplings in α-helices* | | | | | | |
| Amide couplings | I3 | 1.4 | L37 | −2.6 | Q66 | 8.2 |
| ($N_i$-$H_i^N$) | E4 | 0.4 | D38 | 1.6 | A67 | 7.7 |
| | E5 | 3.4 | T39 | −0.3 | I69 | 6.8 |
| | V7 | −1.0 | V40 | −2.6 | D70 | 6.1 |
| | K8 | 0.8 | | | N73 | 7.4 |
| | I10 | 2.0 | | | G74 | 5.5 |
| | I11 | −0.3 | | | H75 | 7.7 |
| | G12 | 2.1 | | | | |
| | E13 | 2.1 | | | | |
| | Q14 | 0.0 | | | | |
| | L15 | −1.9 | | | | |
| Amide-alpha couplings | R6 | 3.0 | L37 | 0.4 | Q66 | 0.0 |
| ($H_i^N$-$H_{\alpha i}$) | V7 | −3.5 | L42 | 0.0 | A68 | −8.5 |
| | K9 | 4.5 | V43 | 3.5 | H75 | −9.0 |
| | L15 | 0.0 | M44 | 2.0 | | |
| | | | L46 | −2.0 | | |
| ($H_i^N$-$H_{\alpha i \pm 1}$) | | | $V43_N$-$L42_\alpha$ | 2.0 | | |
| | | | $M44_N$-$A45_\alpha$ | −2.0 | | |
| Amide-amide couplings | | | D38 | 2.0 | N73 | 2.0 |
| ($H_i^N$-$H_{i+1}^N$) | | | T39 | 2.0 | | |
| | | | V40 | 2.0 | | |
| B. *NOEs and distances used to position helices in POSE* | | | | | | |
| I3 $H^N$ -F50 $H^\alpha$, V7 $H^N$-F50 $H_\zeta$, V7 $H^\alpha$-F50 $H_\zeta$, V7 methyl-A68 $H^N$, I3 methyl-N73 $H^N$ | | | | | | |
| 6 Å 8 Å 8 Å 8 Å 8 Å | | | | | | |

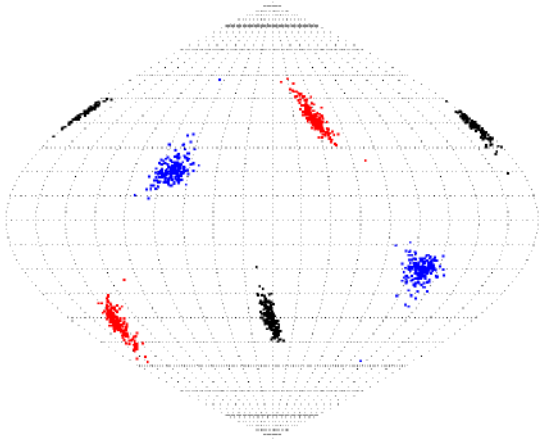
# Orientation Maps for Three ACP Helices

- Sauson-Flamsteed plots/projections are a convenient way to visualize how each helix is oriented relative to the principal alignment frame
  - for each helix, the red, black and blue spots indicate how the x, y, and z axes of the principal alignment frame are oriented relative to the individual helices
  - for each there is a range of values
  - the relative orientation of the z-axes of the three helices appears very similar, with x- and y-axes somewhat similar (suggest three helix bundle)

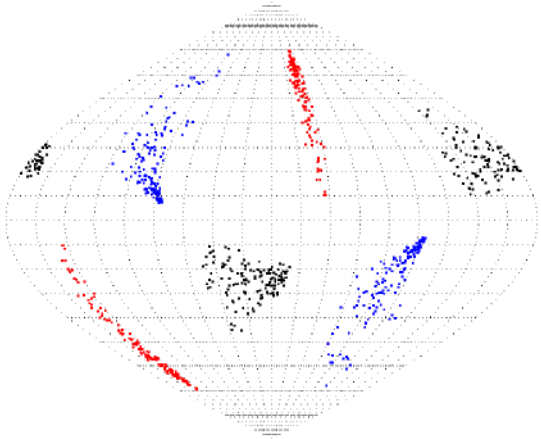
Helix 1



Helix 2



Helix 3

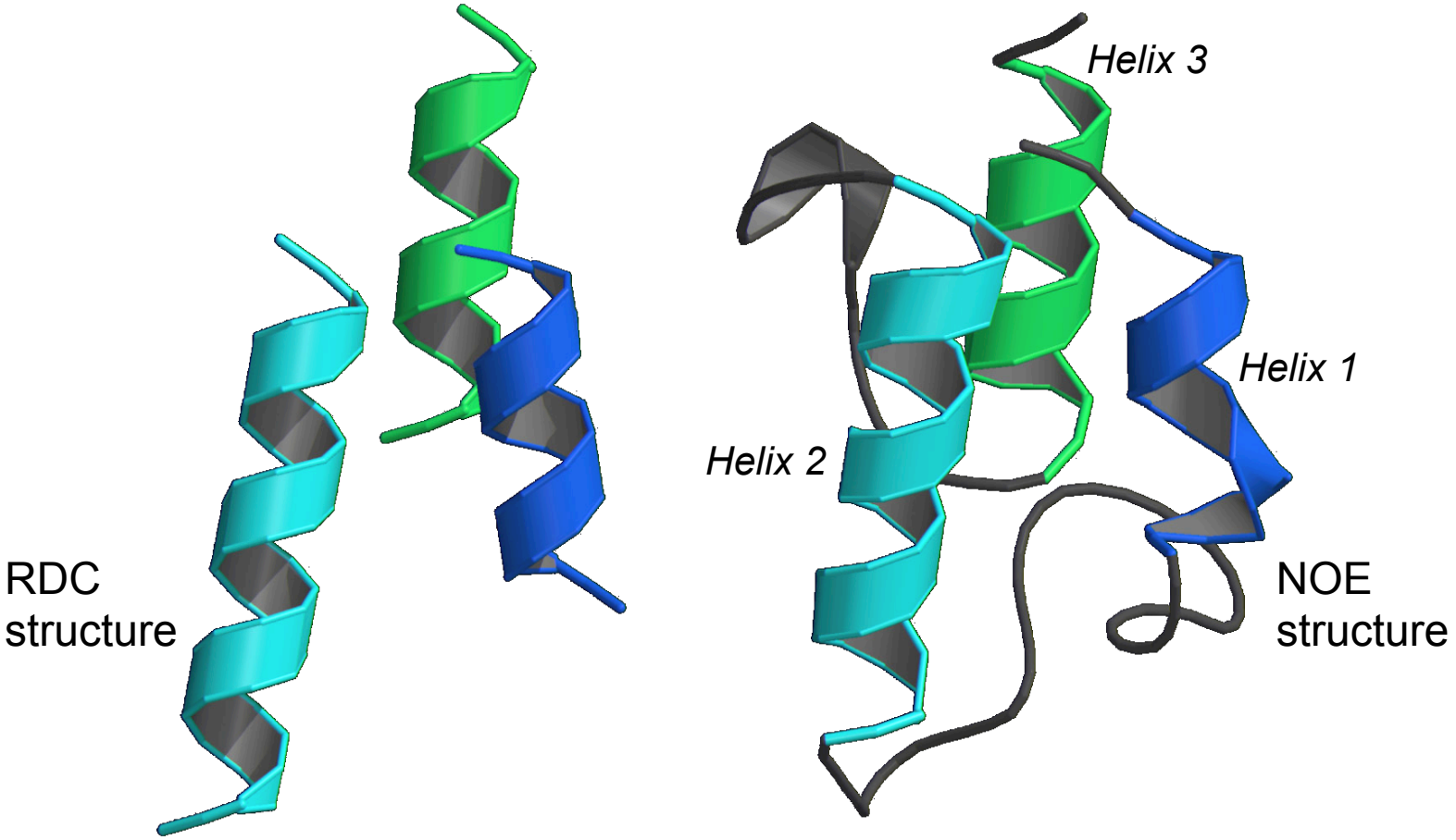


Red = Szz; Black = Syy; Blue = Sxx

Sauson-Flamsteed plots/projections

# ACP Dipolar Fold vs. NOE Structure

- Results show helix orientations nearly identical to those in NOE-based structure
  - important to remember that RDCs do NOT provide translational information (only orientation information), so it is necessary to provide restraints (NOEs for instance) to restrict/define positions of helices




*Fowler, et al., (2000) J Mol Biol 304(3):447-460.*

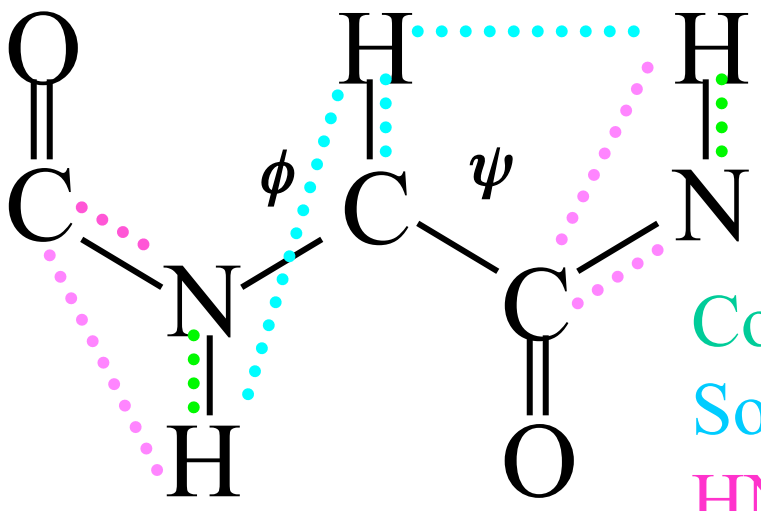
# Some Experiments for RDC Data Acquisition

- List of some references that describe experiments used often to measure RDCs


- Tolman JR, Prestegard JH: Measurement of one-bond amide N-15-H-1 couplings. *JMR,* 1996, 112:245-252
- Ottiger M, Delaglio F, Bax A, Measurement of couplings using IPAP *JMR* 1998,131: 373-378.
- Wang YX, Marquardt JL, Wingfield P, Stahl SJ, Lee-Huang S, Torchia D, Bax A: Measurement of H-1-N-15, H-1-C-13 ', and N-15-C-13 ' dipolar couplings. *JACS,* 1998, 120:7385-7386.
- Yang DW, Venters RA, Mueller GA, Choy WY, Kay LE: TROSY-based HNCO pulse sequences. *JBNMR* 1999, 14:333-343.
- Liu, Y. and J.H, Prestegard. Measurement of one and two bond N-C couplings in large proteins by TROSY-based *J*-modulation experiments. JMR 2009, 200:109-118.
- Delaglio F, Wu ZR, Bax A: Homonuclear proton couplings from regular 2D COSY spectra. *JMR* 2001, 149:276-281.

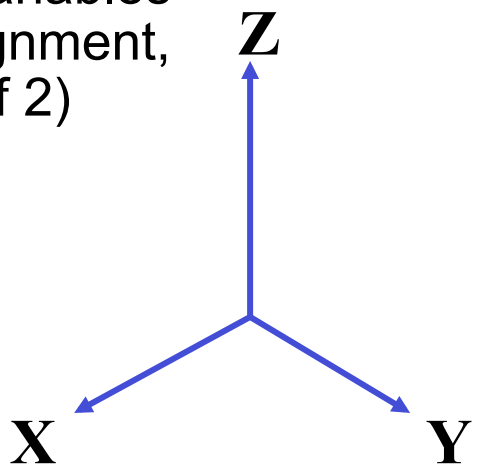
# Measurable Dipolar Couplings in a Dipeptide Define an Order Frame

- Not simple, but not impossible, to determine a complete protein structure using RDCs alone
  - need to be able to collect enough RDCs for each amino acid and between adjacent amino acids to restrain main chain angles ($\phi$ and $\psi$) and determine the alignment
  - $^{1}H^{N}_{i}$-$^{15}N_{i}$, coupled HSQC, IPAP, HSQC/TROSY
  - $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i}$, $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i-1}$, soft HNCA-E.COSY
  - $^{1}H^{\alpha}_{i}$-$^{13}C^{\alpha}_{i}$, soft HNCA-E.COSY
  - $^{1}H^{N}_{i}$-$^{13}C{=}O_{i-1}$, $^{15}N_{i}$-$^{13}C{=}O_{i-1}$, HNCO




- here shown 9 RDCs that can be measured for two variables ($\phi$ and $\psi$) plus 5 for alignment, so 7 needed (excess of 2)

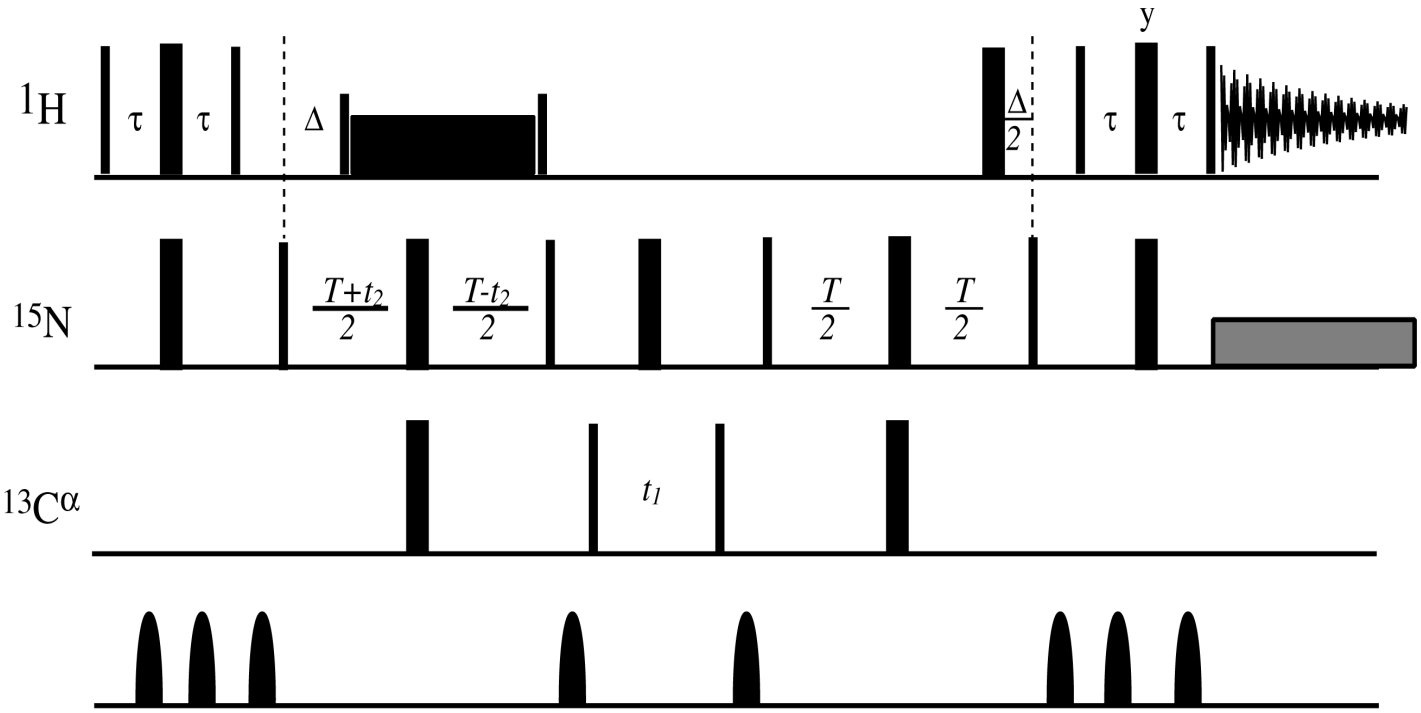
Coupled HSQC
Soft HNCA-E.COSY
HNCO



# Soft HNCA – E.COSY

- HNCA experiment with E. COSY mixing of $H^N$ and $H^\alpha$
  - in $H^N$ dimension ($t_3$) get two of the four peaks of a COSY crosspeak (the two on the diagonal, so same phase), so get $^3J_{HNHA}$ ($^1H^N_i$-$^1H^\alpha_i$ RDC)
  - HNCA gives correlation to previous CA also, so get $^1H^N_i$-$^1H^\alpha_{i-1}$ RDC
  - no $^1H$ decoupling during $^{13}C^\alpha$ evolution ($t_1$), so get $^1J_{CAHA}$ ($^1H^\alpha_i$-$^{13}C^\alpha_i$ RDC)

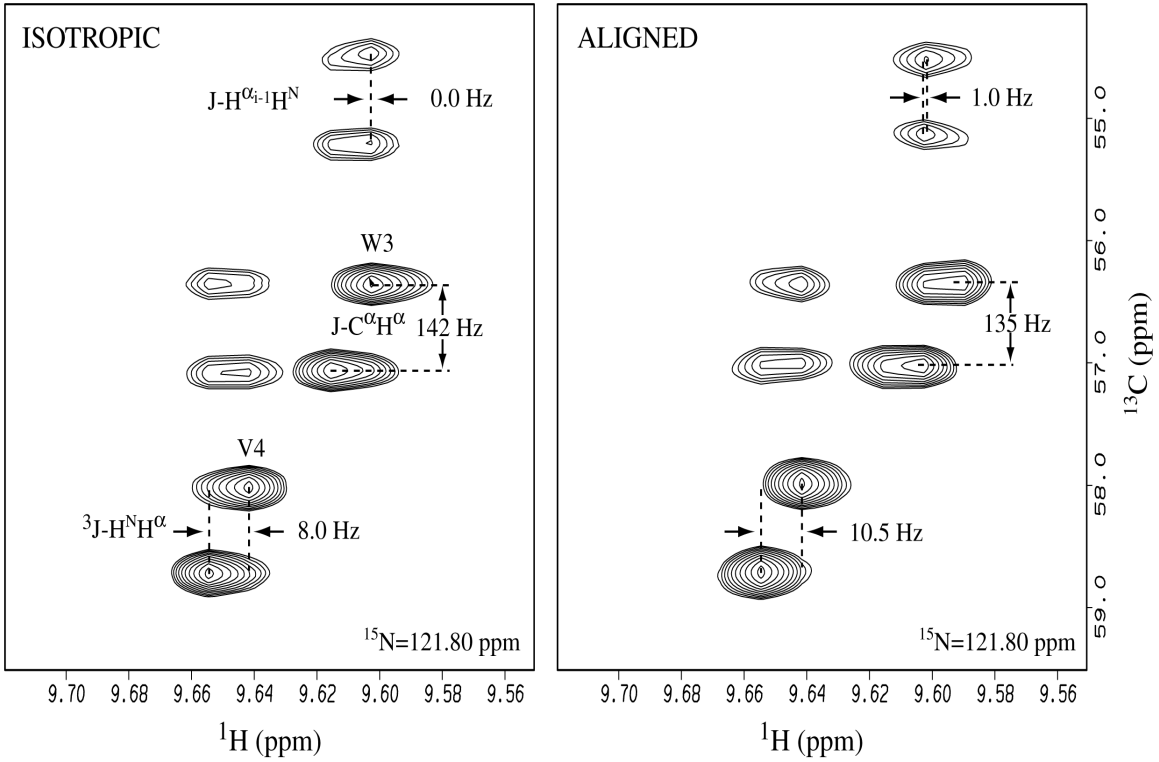
- $^1H^N_i$-$^1H^\alpha_i$, $^1H^N_i$-$^1H^\alpha_{i-1}$, $^1H^\alpha_i$-$^{13}C^\alpha_i$, soft HNCA-E.COSY




*Weisemann, Ruterhans, Schwalbe, Schleucher, Bermel, Griesinger, J. Biomol. NMR, 4, 231-240, 1994*

# Soft HNCA E-COSY Spectra of $^{15}$N-Labeled $^{13}$C Natural Abundance Rubredoxin

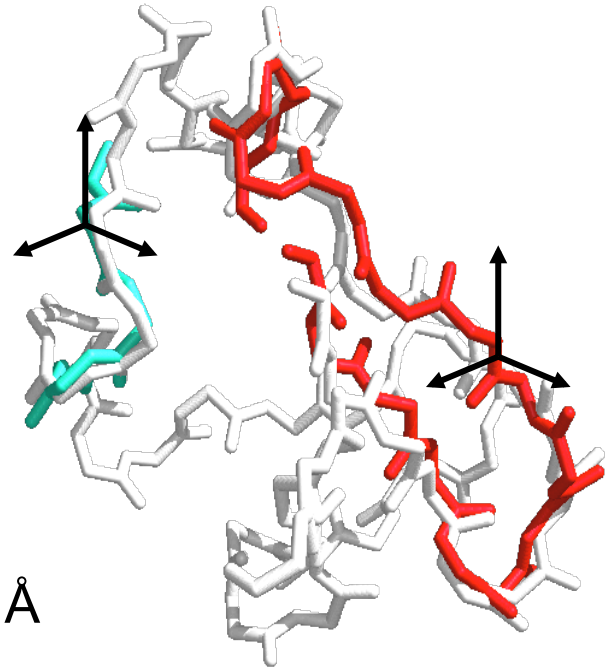
- Get typical correlations seen in HNCA
  - no $^{1}$H decoupling during $^{13}C^{\alpha}$ evolution ($t_1$), so get $^{1}J_{CAHA}$ ($^{1}H^{\alpha}_{i}$-$^{13}C^{\alpha}_{i}$ RDC)
  - in $H^N$ dimension ($t_3$) get two of the four peaks of a COSY crosspeak (the two on the diagonal, so same phase), so get $^{3}J_{HNHA}$ ($^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i}$ RDC)
  - HNCA gives correlation to previous CA also, so get $^{1}H^{N}_{i}$-$^{1}H^{\alpha}_{i-1}$ RDC




$C^{\alpha}$ chemical shift, $C^{\alpha}_{i}$ to $C^{\alpha}_{i-1}$ connectivity, $^{3}J$-$H^{N}H^{\alpha}$ coupling, $C^{\alpha}$-$H^{\alpha}$, $H^{N}H^{\alpha}$ and $H^{\alpha}_{i-1}H^{N}$ dipolar coupling

# Multiple Peptide Segments Oriented to Superimpose Order Frames Yield Structures

- Example: determination of the structure of regions of the protein rubredoxin
  - RDC data were used to determine the structures of fragments of the protein rubredoxin
  - these were aligned and superimposed onto the X-ray structure of rubredoxin
  - excellent agreement with the X-ray structure



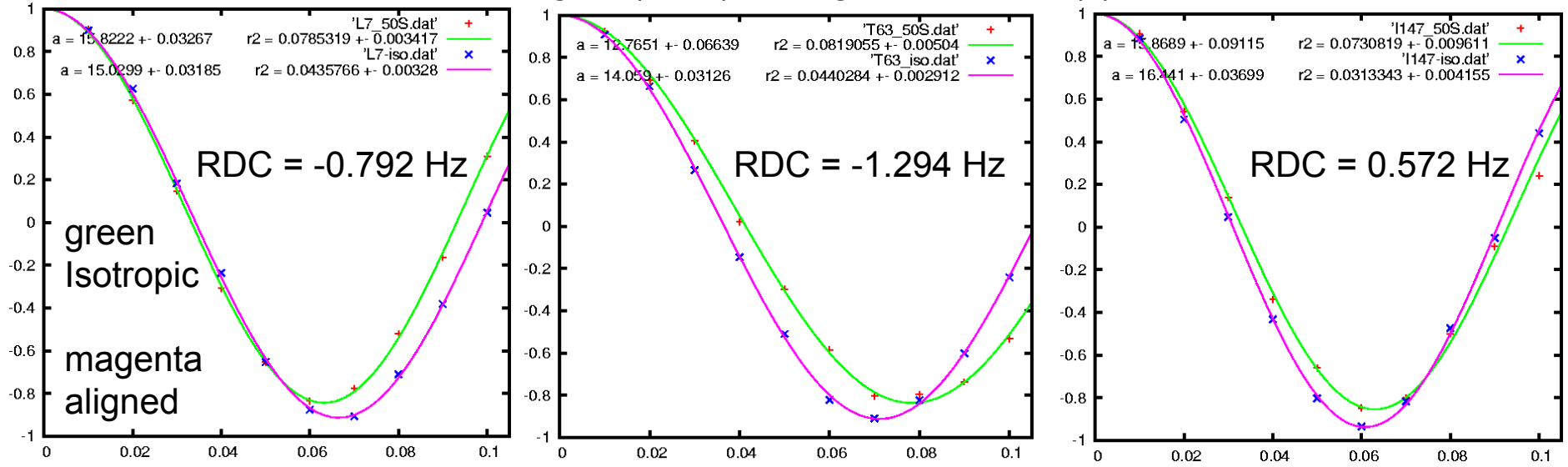
  - overlays X-ray structure (PDB ID 1BRF) to 1.6 Å

*Tian, Valafar, & Prestegard, (2001)*
*J. Am. Chem. Soc.* ***123****, 11791-11796*

# More Recent $^{15}N$-$^{1}H$ Depositions use a *J*-modulation Experiment: Also can be used for $^{15}N$-$^{13}C'$ , $^{15}N$-$^{13}C\alpha$

- Are new methods for RDC measurement that work well for even very large proteins and for small couplings as well
  - here, *J*-modulation experiment (HSQC with *J*-modulation in third dimension)
  - data below are for 70 kDa protein (lines are broad, and would be difficult to measure couplings using peak splitting)

*Liu, Y. and J.H, Prestegard (2009) J. Magn. Reson.* ***200****(1): 109-118*




- Cross-peaks overlap HSQC peaks exactly
- Time requirements are similar to TROSY/HSQC
- Based on TROSY detection for application to larger proteins
- Fit gives $T_2$ estimate – used to eliminate data on loops

# Analysis of Residual Dipolar Couplings

- Much of the software used routinely to determine or refine protein structures now incorporates RDC routines
  - some (Rosetta) use RDCs and routines for "de novo" protein structure calculation and database information (chemical shifts) to calculate structures
  - with larger proteins, becomes more difficult to get NOEs for full structures, but still "easy" to get chemical shifts and RDCs, so, methods that use these data, combined with computation, become more attractive for large proteins

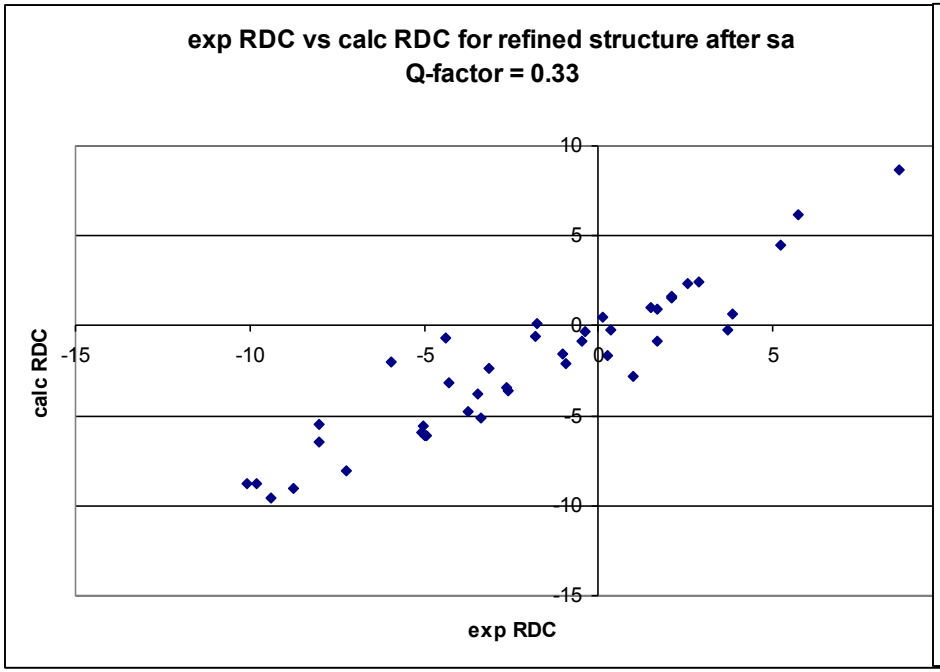

- Schwieters CD, Kuszewski JJ, Tjandra N, et al. XPLOR-NIH, J. Magn. Res. 160 (1): 65-73 JAN 2003
- Meiler J, Blomberg N, Nilges M, Griesinger C, Residual dipolar couplings as restraints in structure elucidation *JBNMR* 2000, 16: 245-252.
- Rohl CA, Baker D: Backbone structure from residual dipolar couplings using Rosetta. *JACS,* 2002, 124:2723-2729.
- Delaglio F, Kontaxis G, Bax A: Structure from molecular fragment replacement and dipolar couplings. *JACS,* 2000, 122:2142-2143.
- Valafar H, Prestegard J. 2004, J. Mag. Res., **167**, 228-241 http://www.ccrc.uga.edu/web/CarbResource/Software/
- Dosset, Hus, Marion & Blackledge (2001), JBNMR, 20: 223-231

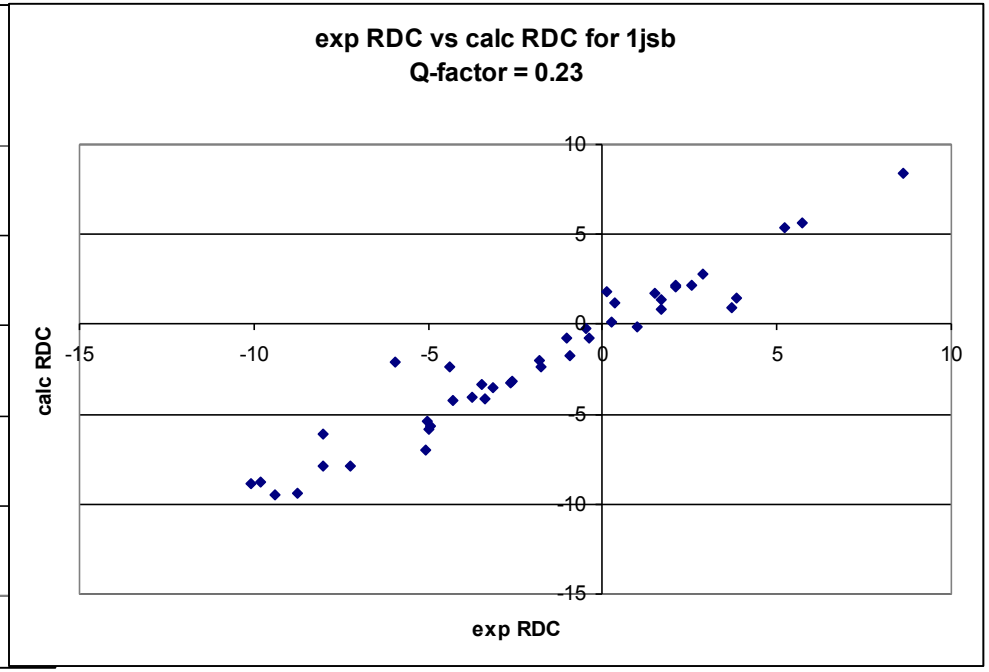
# Example of Validation and Refinement (MTH1743)

- RDCs commonly used simply to help validate and refine structures
  - a measure commonly used, the '*Q*' factor, measures how well data fit an experimental model
  - smaller values of *Q* indicate better agreement
  - interesting, RDCs fit X-ray structures better than NOE-based NMR structures not refined using RDCs

$$Q = \sqrt{\frac{\left(\sum (D_{obs} - D_{calc})^2\right)}{\left(\sum D_{obs}^2\right)}}$$

- Below: some examples of comparisons of experimental versus calculated RDCs



# X-ray Structures fit RDCs Better than NOE-Based NMR Structures

- Agreement of measured RDCs with X-ray versus NMR (NOE) structures shows they typically agree better with X-ray structures
  - typically *Q* values smaller for X-ray structures
  - also note Q for NMR structures not correlated with RMSD (RMSD is a good measurement to make, but is not necessarily a measure of how 'good' a structure is

| neSG | bmrb | pdb nmr | pdb xray | alignment media | #residues | nmr Q | xray Q | RMSD* |
|---|---|---|---|---|---|---|---|---|
| BeR31 | 15702 | 2k2e | 3cpk | phage | 150 | 0.52 | 0.28 | 1.39 |
| CsR4 | 15317 | 2jr2 | 2ota | peg (and peg+ctab | 68 | 0.37 | 0.32 | 0.52 |
| CtR107 | 16097 | 2kcu | 3e0h | phage (and peg) | 158 | 0.44 | 0.30 | 1.84 |
| GmR137 | 15844 | 2k5p | 3cwi | peg | 70 | 0.38 | 0.21 | 1.37 |
| HR3646E** | 16250 | 2khn | 3fia | polyacrylamide gel | 110 | 0.53 | 0.29 | 1.06 |
| MbR242E | 16368 | 2kko | 3gw2 | peg | 100 | 0.36 | 0.29 | 1.05 |
| PfR193A | 16385 | 2kl6 | 3idu | phage | 114 | 0.36 | 0.30 | 0.86 |
| SgR42 | 15604 | 2jz2 | 3c4s | peg | 58 | 0.42 | 0.23 | 0.58 |
| SoR77 | 15456 | 2juw | 2qti | polyacrylamide gel | 72 | 0.26 | 0.21 | 0.91 |

* PSVS analysis listed structured regions (obtained via PROCHECK)

1st NMR model compared to X-Ray structure for all analysis

** It was difficult to compare the xray and nmr structures for this protein.

$$Q = \frac{\left(\sum (D_{obs} - D_{calc})^2\right)^{1/2}}{\left(\sum D_{obs}^2\right)^{1/2}}$$

# Structure Refinement Using RDCs

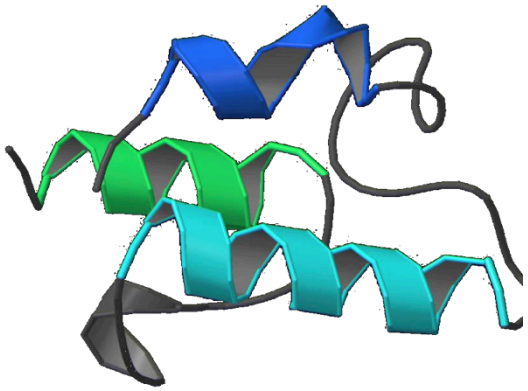
Write RDCs in principal alignment frame:

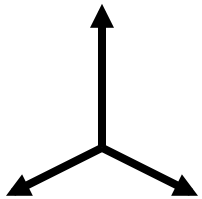
$$D = (D_a/r^3)\{(3\cos^2\theta - 1)/r^3 + (3/2)R\sin^2\theta\cos(2\phi)\}$$

Write error function in terms of $D_{meas}$ and $D_{calc}$

$$E_{RDC} = (D_{meas} - D_{calc})^2$$

Seek minimum in $E_{RDC}$ to refine structure –
Need to float alignment axes during search

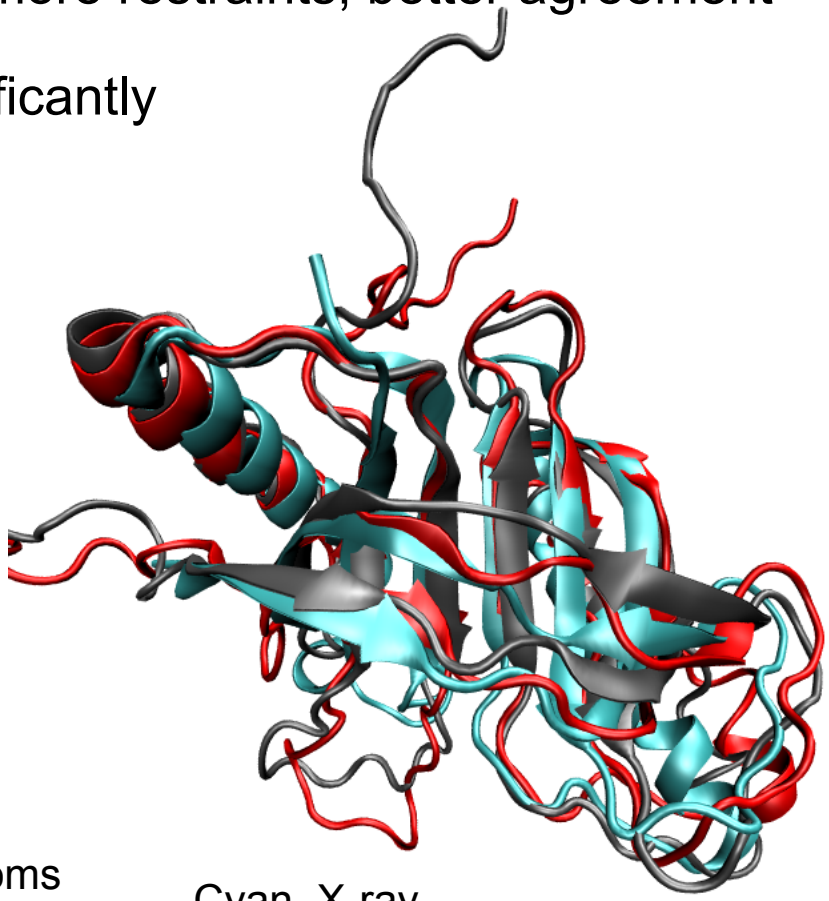




# Refinement with RDCs can Improve Quality
# CtR107 with and without RDCs

- Using RDCs to improve NOE based structures
  - RDCs can be used during initial stages ('anneal') and refinement stages ('refine')
  - RDCs improve initial ('anneal') RMSD (more restraints, better agreement among calculated structures)
  - RDCs added to refinement stages significantly improves RMSDs

| Refinement detail | Average RMSD to X-ray (best of 10) | RMSD of the ensemble |
|---|---|---|
| Anneal, no RDC | 4.4 | 4.2 |
| Anneal, with RDC | 3.4 | 2.7 |
| Refine, no RDC | 2.5 (2.0) | 1.4 |
| Refine, with RDC | 2.0 (1.6) | 1.4 |



Alignment performed by superimposing backbone atoms of residues 19 to 26, 30 to 39, 59 to 61, 69 to 71, 88 to 90, 97 to 102, 111 to 122, 132 to 134 and 149 to 152

Cyan, X-ray
Red, best refined with RDC
Gray, best refined without RDC