

SPIN RELAXATION IN PROTEINS

BCMB/CHEM 8190

Utility of Spin-Relaxation Studies in Proteins

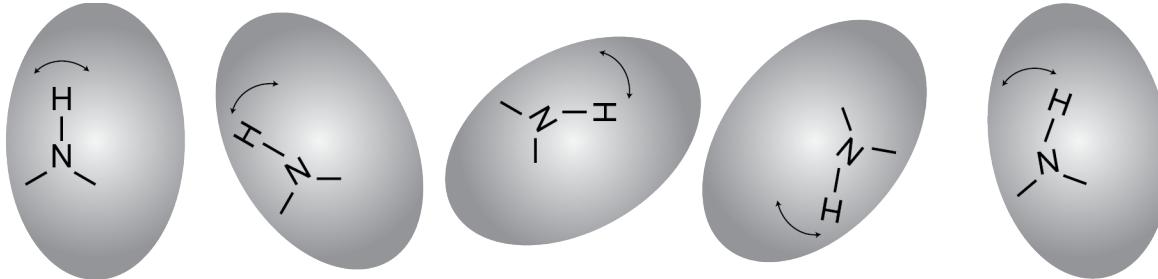
- Spin relaxation (dipole-dipole, CSA) are useful for structural studies of proteins
 - the nuclear Overhauser effect (NOE) is a consequence of spin relaxation and is the principle means for measuring distances used in structural studies
 - spin relaxation can be used to study overall molecular tumbling (τ_c)
- Spin relaxation methods are important for other types of NMR studies of proteins
 - spin relaxation measurements can reveal the magnitudes and timescales of internal motions ('internal dynamics') in proteins
 - important resolution enhancement methods (TROSY) have their origins in spin relaxation

Model-Free Approach to Internal Motions in Proteins

- The "Model-Free" approach (Lipari and Szabo) is used widely to study internal motions in proteins
 - this is one of the most common means used to analyze internal motions in proteins
- Lipari & Szabo, JACS 104, 4546 (1982)
- Palmer AG. Ann. Rev. Biophys. Biomol. Struc., 30, 129-155 (2001)
- Palmer AG, Kroenke CD, Loria JP, Meth. Enzymol. 339, 204-238 (2001)

Model-Free Approach to Internal Motions in Proteins

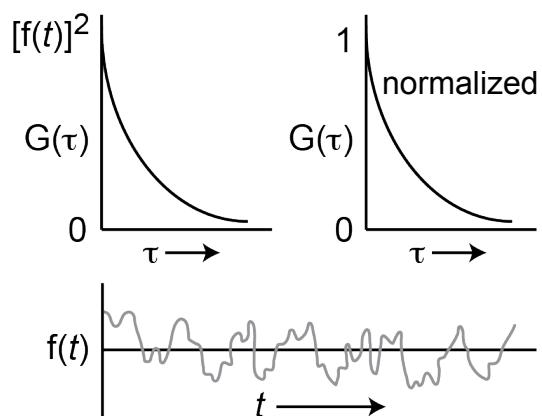
- The "Model-Free" approach (Lipari and Szabo) is used widely to study internal motions in proteins
 - this is one of the most common means used to analyze internal motions in proteins
 - the approach allows the timescale or rate of motions to be quantified, as well as the degree of spatial restriction of the motion (compared to isotropic motion)
 - no physical model describing what the motion looks like is necessary
- Goal is a framework to analyze fast internal motions (for instance, bond vector reorientations) in the context of overall tumbling of the molecule



- The analysis gives two parameters describing the internal motion
 - an internal ("effective") correlation time, τ_e , that describes the rate of the motion
 - an generalized "order parameter", S^2 , that describes the spatial restriction of the motion

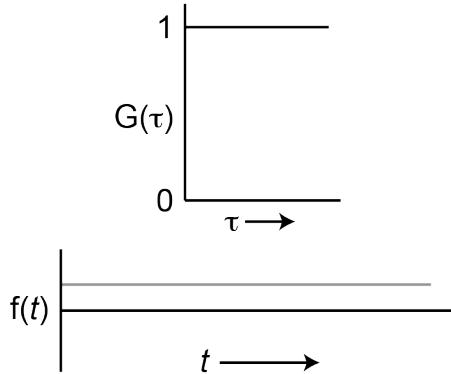
Model-Free Approach to Internal Motions in Proteins

- Recall the correlation function for a random process



- $G(\tau) = \overline{f(t+\tau) \times f(t)}$ i.e. time average of $f(t)$ and $f(t+\tau)$
- $f(t)$ averages to zero, and as $\tau \rightarrow \infty$, $G(\tau) \rightarrow 0$, i.e. $G(\infty)=0$
- the correlation function is an exponential decay, defined by a correlation time, τ_c
$$G(\tau) = G(0)\exp(-|\tau|/\tau_c)$$
- if we normalize such that $G(\tau)=[f(t)]^2=1$, then
$$G(\tau) = \exp(-|\tau|/\tau_c)$$

- This (above) is what is expected for a bond vector (N-H) in a protein whose motion is totally *unrestricted* (i.e. random, isotropic)
- Now, consider a bond vector (N-H) in a protein whose motion is totally *restricted* by tight packing in the protein
 - for the moment we also assume the protein is immobile (i.e. not tumbling)



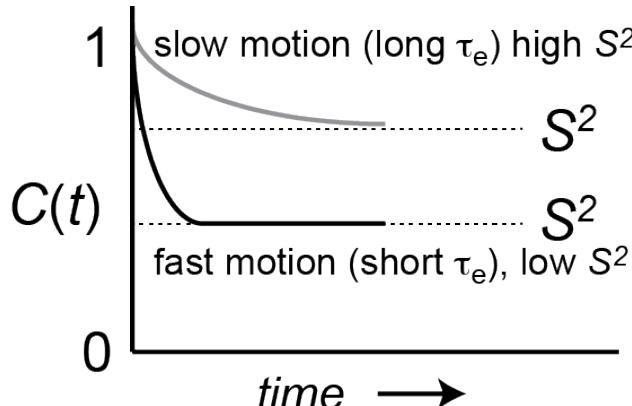
- here $G(\tau) = \overline{f(t+\tau) \times f(t)}$ is independent of τ
- here $G(\infty)=1$

Model-Free Approach to Internal Motions in Proteins

- The Model-Free approach defines a correlation function for internal motions (and subsequently combines this with overall tumbling of the molecule)
- The correlation function for internal motion has to account for the range of restriction (from unrestricted to totally restricted):

$$C_e(t) = S^2 + (1 - S^2)\exp(-t / \tau_e)$$

- although there is no 'standard' notation, we'll use a notation that is often used:
 $C_e(t)$ is the correlation time for internal motion, S^2 is a "generalized order parameter", τ_e is the correlation time for internal motion ("effective" correlation time)
- S^2 characterizes the degree of spatial restriction: $C_e(\infty)=S^2$
 - as $S^2 \rightarrow 1$ (totally restricted motion), $C_e(t) \rightarrow 1$, so $C_e(\infty)=1=S^2$
 - as $S^2 \rightarrow 0$ (totally unrestricted motion), $C_e(t) \rightarrow 0$, so $C_e(\infty)=0=S^2$



Model-Free Approach to Internal Motions in Proteins

- In the Model-Free approach the correlation function for random molecular tumbling is what we have seen previously

$$C_M(t) = \exp(-t / \tau_m)$$

- here, C_M is the correlation function for overall molecular tumbling, and τ_m is the molecular rotational correlation time

- For the simplest case, the Model-Free approach assumes that internal motion and global tumbling are independent (uncorrelated), so the overall correlation function is just the product of C_M and C_e :

$$C(t) = C_M(t)C_e(t) = \exp(-t / \tau_m) [S^2 + (1 - S^2)\exp(-t / \tau_e)] = S^2 \exp(-t / \tau_m) + (1 - S^2)\exp(-t / \tau)$$

$$C(t) = S^2 \exp(-t / \tau_m) + (1 - S^2)\exp(-t / \tau)$$

$$\text{where } \tau^{-1} = \tau_m^{-1} + \tau_e^{-1} \quad \tau = \frac{\tau_m \tau_e}{\tau_m + \tau_e} \quad (\text{as } \tau_e \text{ decreases, } \tau \rightarrow \tau_e)$$

- The Fourier Transform of $C(t)$ gives the spectral density

$$J(\omega) = \frac{2}{5} \left(\frac{S^2 \tau_m}{1 + \omega^2 \tau_m^2} + \frac{(1 - S^2) \tau}{1 + \omega^2 \tau^2} \right)$$

Separate timescales can be introduced for internal and overall motion

$$J(\omega) = \frac{2}{5} \left(\frac{S^2 \tau_m}{1 + \omega^2 \tau_m^2} + \frac{(1 - S^2) \tau}{1 + \omega^2 \tau^2} \right)$$

$$\text{where } \tau^{-1} = \tau_m^{-1} + \tau_e^{-1} \quad \tau = \frac{\tau_m \tau_e}{\tau_m + \tau_e} \quad (\text{as } \tau_e \text{ decreases, } \tau \rightarrow \tau_e)$$

- In the spectral density, the two terms describe separate timescale motions, that are scaled (differently) by the order parameter
 - the first term is the global tumbling term: S^2 (which varies from 0 to 1) scales (down) that term, so that efficiency due to global tumbling is less efficient
 - so, internal motion decreases the efficiency of relaxation due to global tumbling
- the second term is the internal motion term: this term is scaled (down) by $1 - S^2$ so that efficiency of relaxation due to internal motion is also decreased
- also, when τ is very small (τ_e is small), the second term becomes irrelevant (so, relaxation due to global tumbling dominates)

$$J(\omega) = \frac{2}{5} \left(\frac{S^2 \tau_m}{1 + \omega^2 \tau_m^2} + \frac{(1 - S^2) \tau}{1 + \omega^2 \tau^2} \right) \approx \frac{2}{5} \left(\frac{S^2 \tau_m}{1 + \omega^2 \tau_m^2} \right) \text{ when } \tau_e \text{ is very small}$$

Determining S^2 and τ_m for ^{15}N Relaxation

- The generalized order parameter, S^2 , and the global tumbling time, τ_m , are important for understanding protein dynamics
 - S^2 gives site-resolved measures of flexibility and motional restriction
 - τ_m gives global tumbling timescale
 - these are useful for a comprehensive understanding of protein dynamics
- For proteins, measurements of ^{15}N T_1 , T_2 (and sometimes the heteronuclear NOE) are used to determine S^2 and τ_m

$$R_{1S} = 1/T_1 = ((\mu_0 h^2 \gamma_I^2 \gamma_S^2 r_{IS}^{-6}) / (64\pi^3)) (J(\omega_I - \omega_S) + 3J(\omega_s) + 6J(\omega_I + \omega_S))$$

$$R_{2S} = 1/T_2 = ((\mu_0 h^2 \gamma_I^2 \gamma_S^2 r_{IS}^{-6}) / (128\pi^3)) (4J(0) + J(\omega_I - \omega_S) + 3J(\omega_s) + 6J(\omega_I) + 6J(\omega_I + \omega_S))$$

- here "I" is ^1H , "S" is ^{15}N , and r_{IS} is the N-H bond length (treated as a constant)
- we can group all the constants and rewrite as follows

$$R_{1S} = 1/T_1 = (D)(J(\omega_I - \omega_S) + 3J(\omega_s) + 6J(\omega_I + \omega_S))$$

$$R_{2S} = 1/T_2 = (D/2)(4J(0) + J(\omega_I - \omega_S) + 3J(\omega_s) + 6J(\omega_I) + 6J(\omega_I + \omega_S))$$

Determining S^2 and τ_m for ^{15}N Relaxation

- For large molecules and small τ_e (good assumptions), S^2 and τ_m can easily be determined from T_1 and T_2

$$J(\omega) = \frac{2}{5} \left(\frac{S^2 \tau_m}{1 + \omega^2 \tau_m^2} \right) R_{1S} = 1/T_1 = (D)(J(\omega_I - \omega_S) + 3J(\omega_s) + 6J(\omega_I + \omega_S))$$
$$R_{2S} = 1/T_2 = (D/2)(4J(0) + J(\omega_I - \omega_S) + 3J(\omega_s) + 6J(\omega_I) + 6J(\omega_I + \omega_S))$$

- for T_2 , $J(0)$ dominates, so

$$R_{2S} = 1/T_2 \cong (D/2)(4J(0))$$

- $J(\omega)$ gets very small for high frequencies (ω^2 in denominator), so, $J(\omega_I)$ (i.e. $J(\omega_H)$) terms in T_1 expression become insignificant because of the high H frequency, so the $3J(\omega_S)$ (i.e. $3J(\omega_N)$) term dominates

$$R_{1S} = 1/T_1 \cong (D)(3J(\omega_s))$$

$$\frac{R_2}{R_1} = \frac{T_1}{T_2} \cong \frac{(D/2)(4J(0))}{(D)(3J(\omega_s))} \cong \frac{2}{3} \frac{J(0)}{J(\omega_s)} \cong \frac{2}{3} \left[\frac{2}{5} S^2 \tau_m \right] \left/ \left[\frac{2}{5} \left(\frac{S^2 \tau_m}{1 + \omega_s^2 \tau_m^2} \right) \right] \right. \cong \frac{2}{3} (1 + \omega_s^2 \tau_m^2)$$

- So, τ_m can be calculated from R_2/R_1 , then S^2 can be calculated, knowing τ_m , from R_1 or R_2
- S^2 tends to be relatively high (~ 0.8) for structured regions of low flexibility, and lower for more mobile regions (loops, terminii)

Example from binding of phosphopeptides to SH2 domain *Biochemistry*, 33, 5987 (1994)

- SH2 domains are present in many proteins, and function to bind to phosphorylated peptides

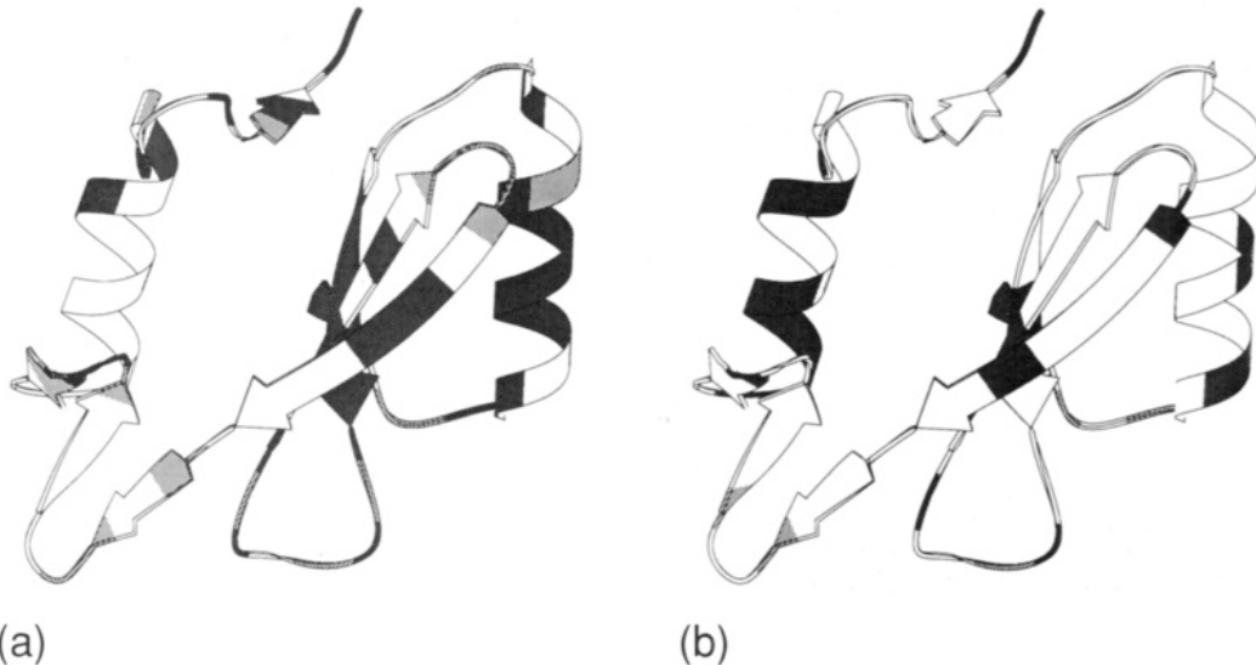
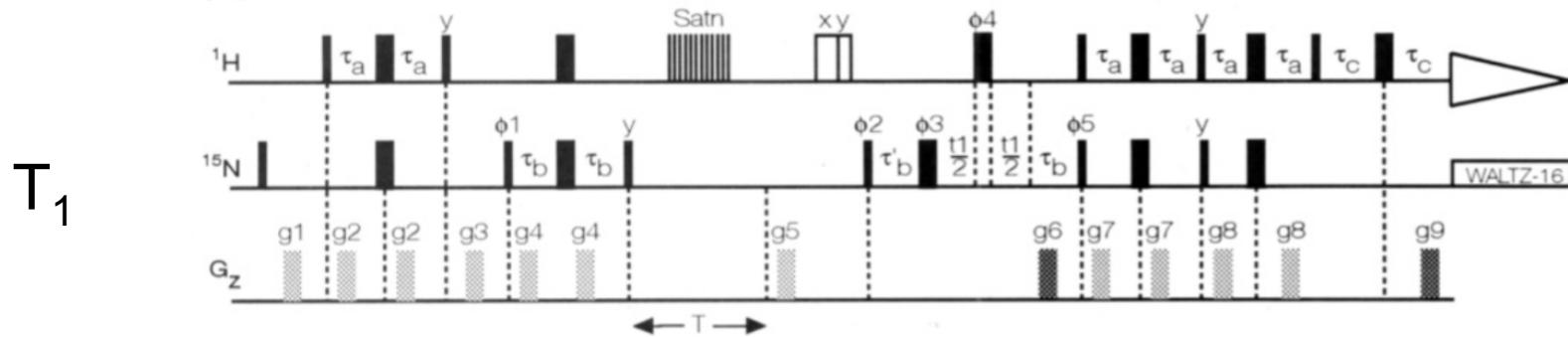


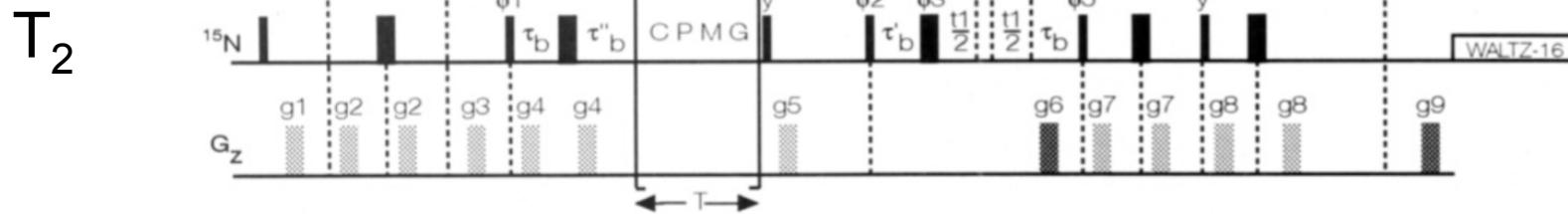
FIGURE 9: Schematic representations of the spatial distributions of residues requiring an R_{ex} term or two-time-scale spectral density functions to fit the measured relaxation data. Residues which required an R_{ex} term are highlighted in dark grey, and residues which required the two-time-scale spectral density function are shown in lighter grey. Residues which required neither of these spectral density functions or for which spectral density functions were not determined are white. The distribution for the uncomplexed form of PLC γ 1C is shown in (a); the complexed form is shown in (b). The distributions are imposed on a preliminary structure of PLC γ 1C in complex with pY1021 (Pascal et al., 1994). The phosphopeptide is not shown in the diagram. The figure was produced using SETOR (Evans, 1993).

$^{15}\text{N-T}_1$, $^{15}\text{N-T}_2$ and $^1\text{H},^{15}\text{N-NOE}$ Experiments

(a)

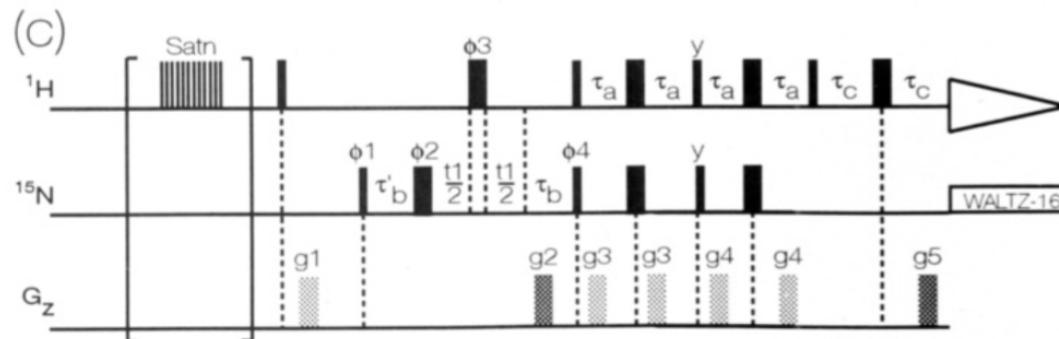


(b)



(c)

NOE



Typical ^{15}N -T₁, ^{15}N -T₂ Results

- Decay curves are fit to the appropriate exponential equations to extract T₁ and T₂

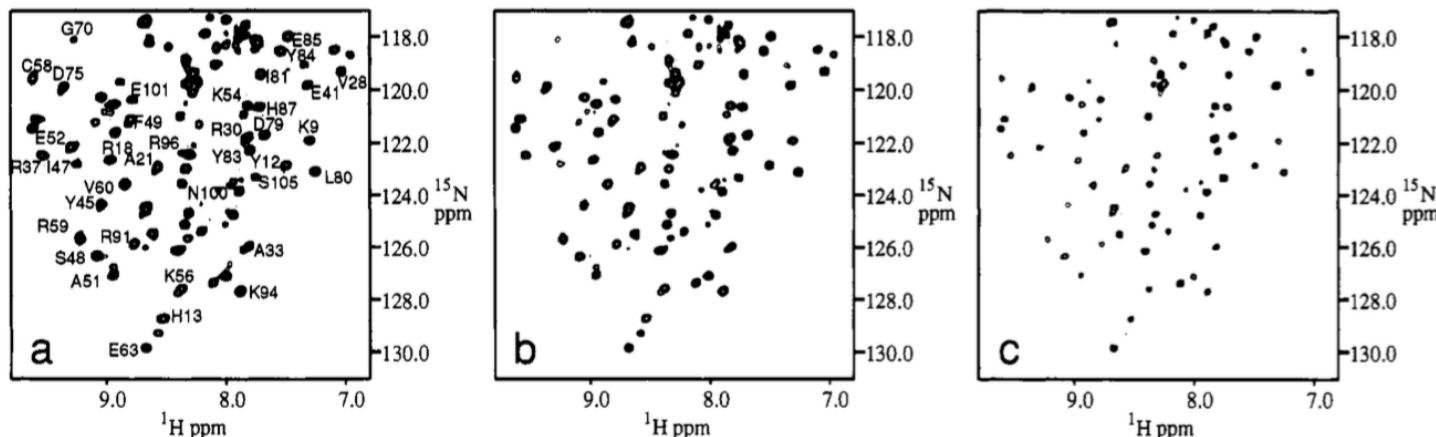


FIGURE 4: Region of the ^1H - ^{15}N shift correlation spectrum recorded with the pulse sequence used to determine T_1 as shown in Figure 3a and three relaxation delay values (T) of (a) 5, (b) 246, and (c) 757 ms.

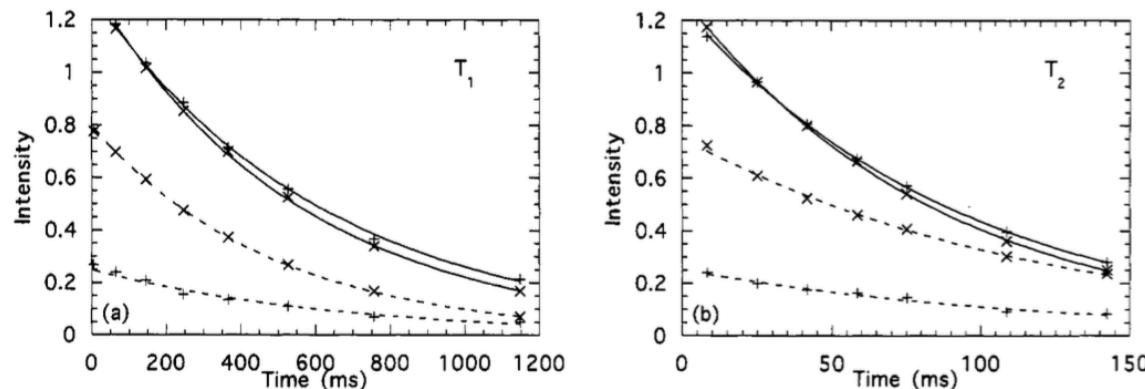


FIGURE 5: Examples of (a) T_1 and (b) T_2 decay curves for Ser 44 (+) and Leu 80 (x) in the uncomplexed (solid lines) and pY1021 peptide-complexed (dashed lines) forms. The curves indicate best fits to single-exponential decays. Error bars, if shown, would be smaller than the size of the characters used to indicate the data points.

Changes in Order Parameters on Complexation

- A typical result is that most residues of globular, folded proteins have relatively high (0.8) order parameters
 - loops, terminii typically show lower values for S^2
- When the ligand (phosphopeptide) binds, notice some increased flexibility (decreased S^2).
 - seems counterintuitive, but, not uncommonly, increased entropy (disorder, motion) accompanies binding to compensate for entropy loss of binding
- R_{ex} , or longer timescale motion (μs , ms), terms can be added to the expression for T_2 to account for anomalously fast T_2 relaxation

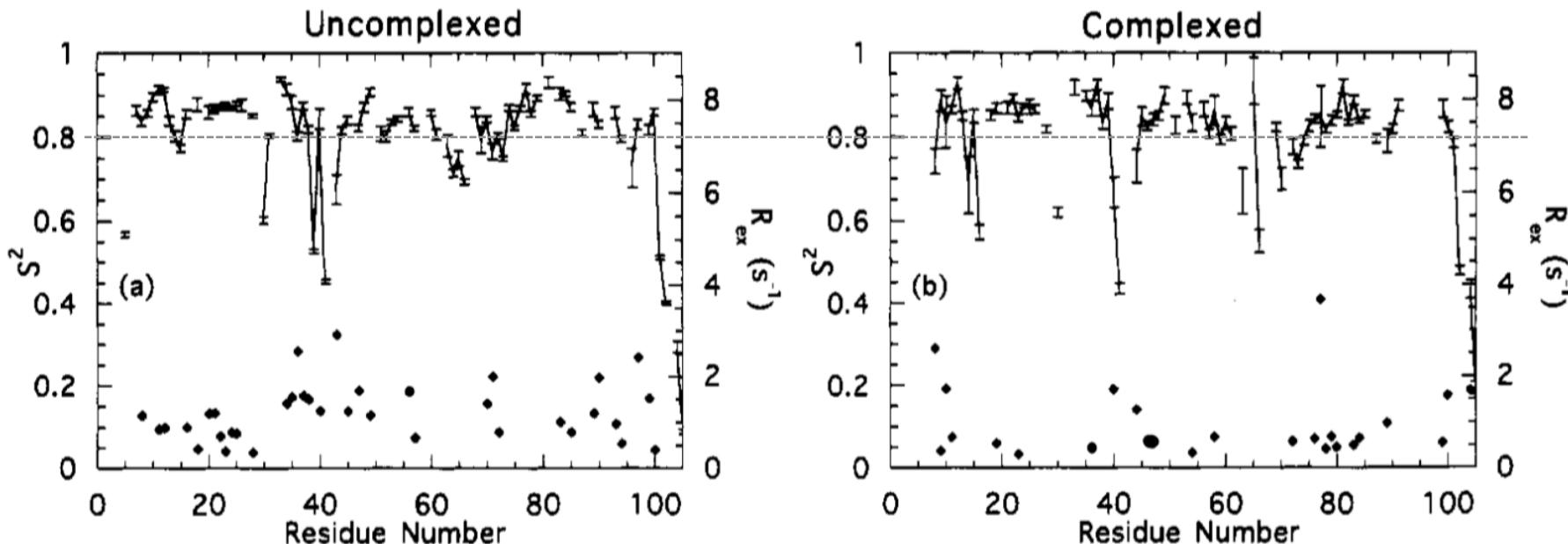


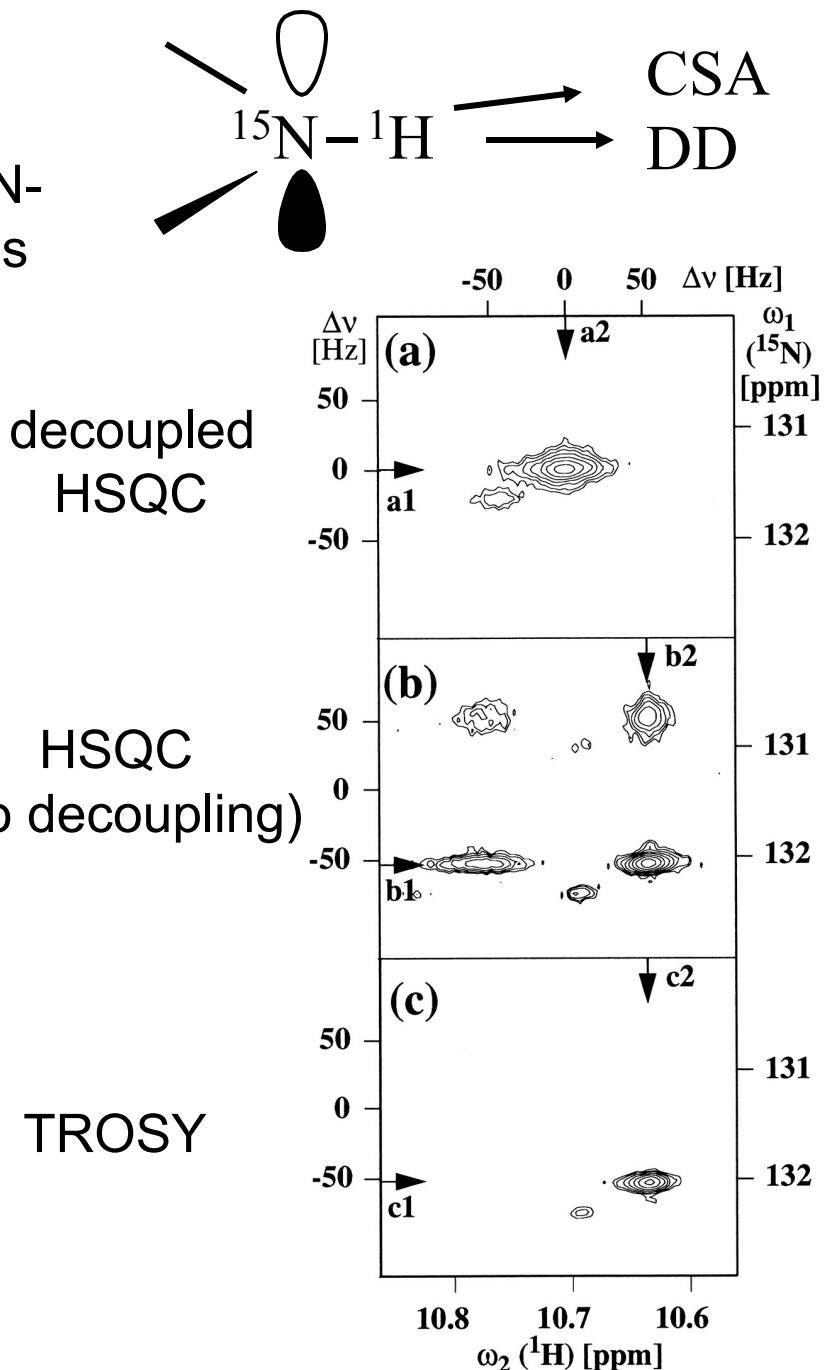
FIGURE 7: Plot of the order parameter S^2 (solid line) and R_{ex} (solid diamonds) as a function of residue number for (a) the uncomplexed form of SH2 and (b) the complexed form of SH2.

Internal Dynamics can Improve Resolution – Cross-Correlation Effects

- TROSY: Transverse Relaxation Optimized Spectroscopy
 - as proteins get large, T_2 decreases considerably, and leads to very broad lines (significantly decreased S/N, resolution)
 - TROSY can miraculously decrease linewidths
- TROSY - Pervushin, Riek, Wider & Wuthrich, PNAS 94, 12366 (1997)
- TROSY na CRINEPT - Riek, Pervushin & Wuthrich TIBS, 25, 462 (2000)
- Ca-N torsion angles -Reif, Hennig & Griesinger Science, 276, 1230-1233 (1997)

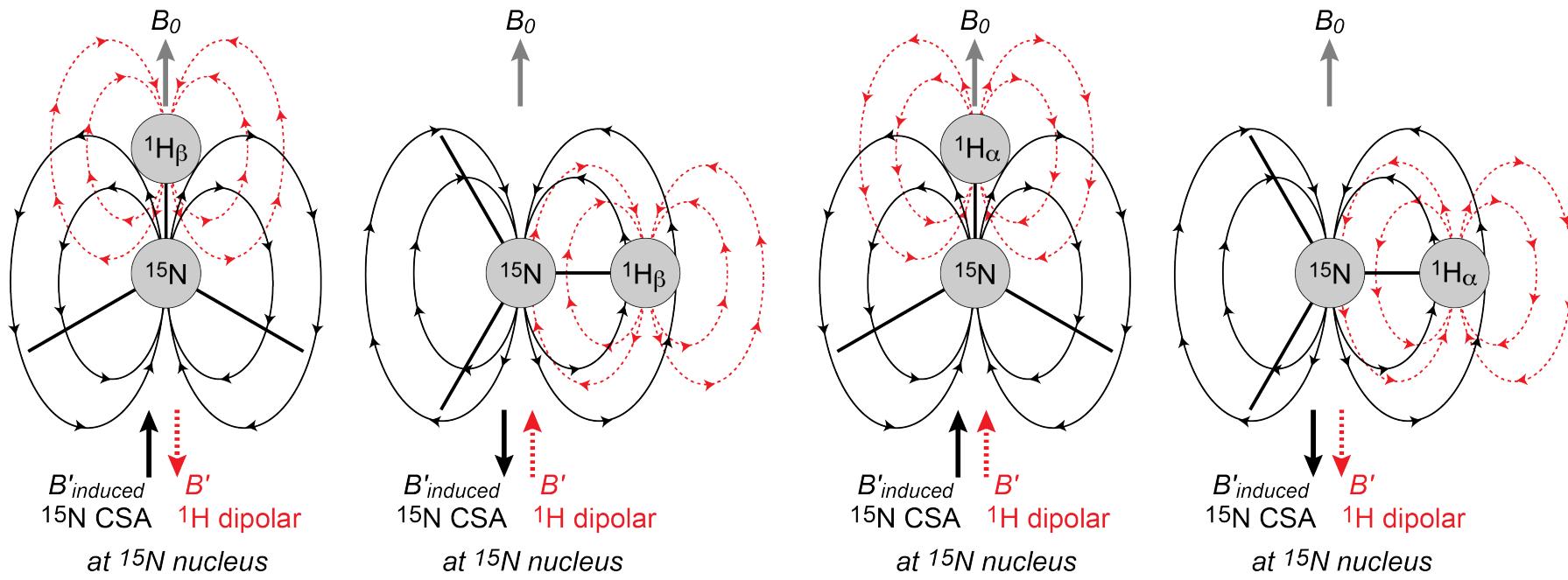
TROSY, Example

- TROSY most commonly applied to ^1H , ^{15}N -HSQC type experiments on large proteins
- In a decoupled ^1H , ^{15}N HSQC spectrum, each peak is an average of the four multiplet components
- The S/N and line widths of the individual multiplet components are very different: each has different contributions from CSA and dipole-dipole coupling to T_2
- TROSY selects for one of the components; for this component, the CSA and dipole-dipole contributions nearly cancel one another (highest S/N)



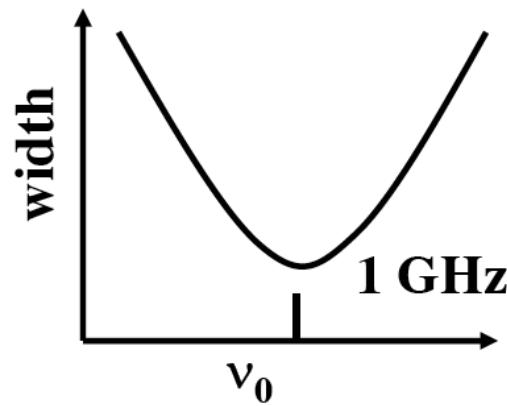
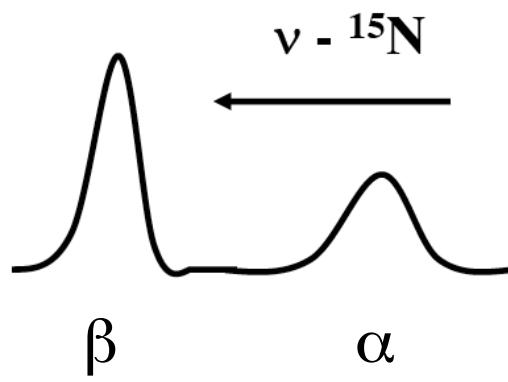
^{15}N CSA and ^1H - ^{15}N Dipole Interactions Interfere

- The TROSY effect results from a decrease in CSA and dipole-dipole contributions to relaxation as the molecule tumbles, due to a decrease in the local/induced magnetic fields by interference between CSA and dipole-dipole contributions to the local fields
 - as the molecule below tumbles, the induced field (CSA) and local dipolar field (from ^1H) at the ^{15}N nucleus fluctuate, leading to relaxation
 - however, for ^1H in the β state, the induced field and local field from the attached ^1H nucleus are opposite in sign, and this interference decreases the total local field at the ^{15}N nucleus, and the ^{15}N peak with ^1H in the β state is narrowed
 - for the ^{15}N peak with ^1H in the α state, the effects add and the peak is broadened



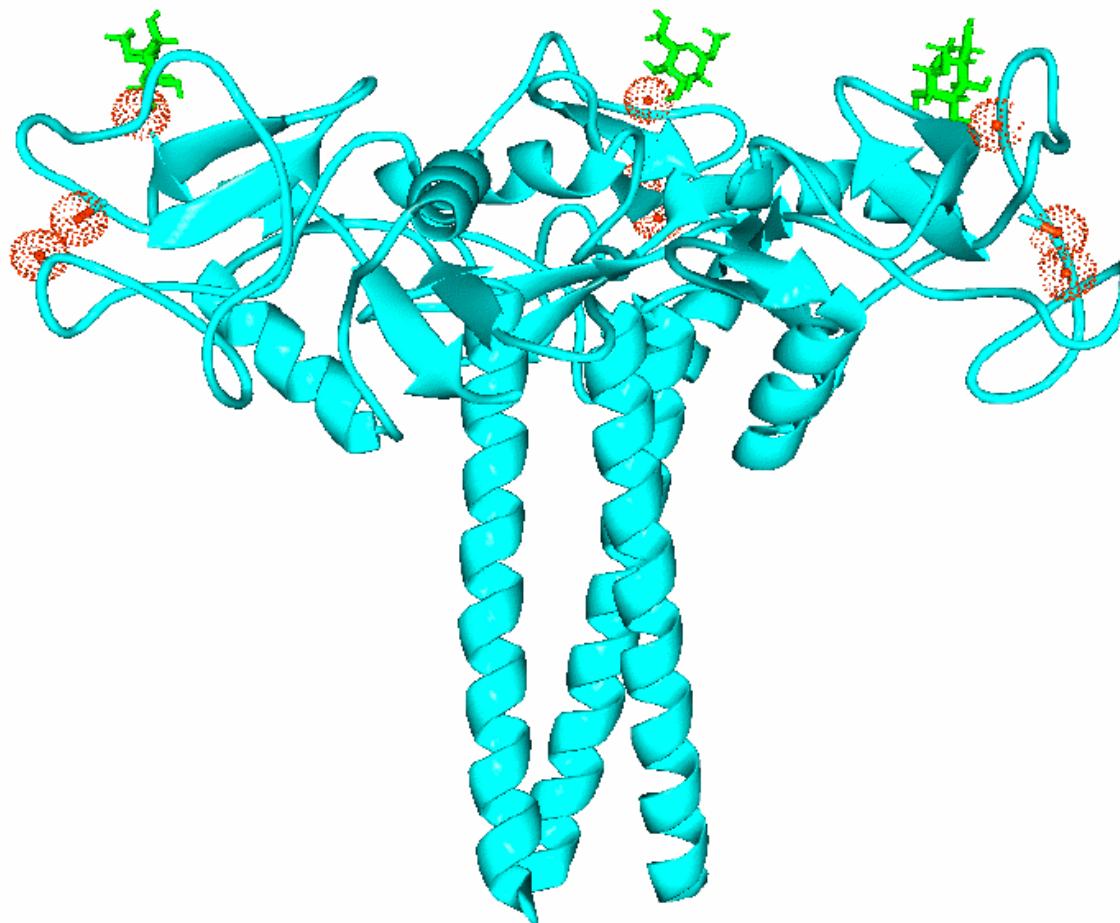
^{15}N CSA and ^1H - ^{15}N Dipole Interactions Interfere

- For the ^{15}N signal with attached ^1H , one component of the doublet (^1H in the β state) is narrowed, the other (^1H in the α state) is broadened
 - the effect is field dependent (remember CSA, $B_{\text{eff}} = (1-\sigma)B_0$)
 - the maximum effect (minimum line width of narrowed peak) is approximately 1.2 GHz (depends on the magnitude and orientation of the chemical shift tensor, etcetera)
 - effect not so noticeable at lower field (i.e. 500 MHz)
 - the CSA and DD components never cancel completely for N-H (CSA tensor is not quite aligned with B_0 , is about 20° off)
 - are also lots of other hydrogen nuclei in proteins near enough to any given ^{15}N nucleus to also contribute to relaxation (so, lines are never infinitely narrow)



Example: HSQC and TROSY spectra α -Methyl Mannose Bound to Mannose Binding Protein

- A three-fold symmetric trimer (total mass ~ 60 kDa)
- Symmetry means fewer peaks, but still very high mass

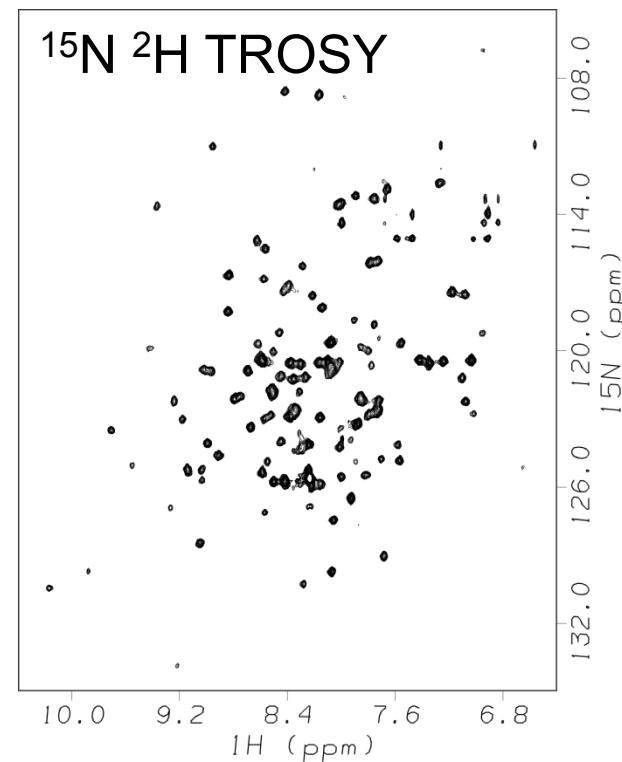
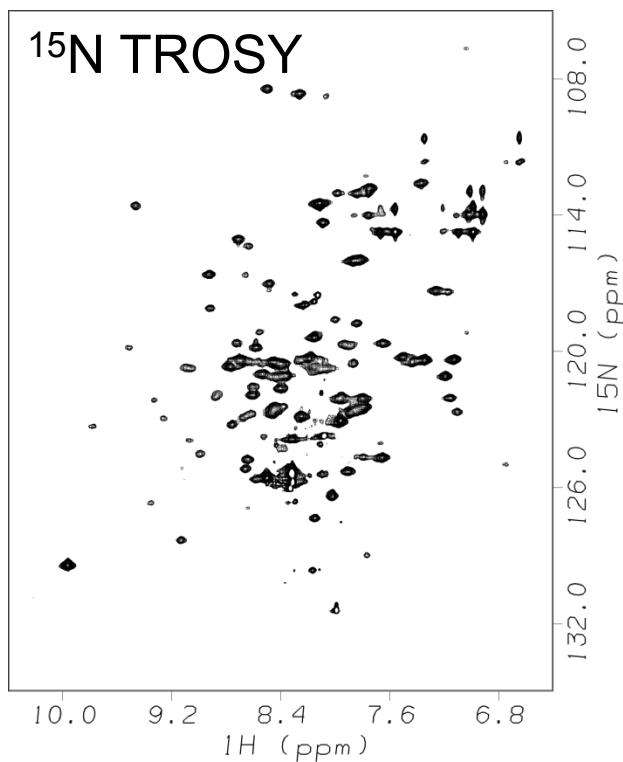
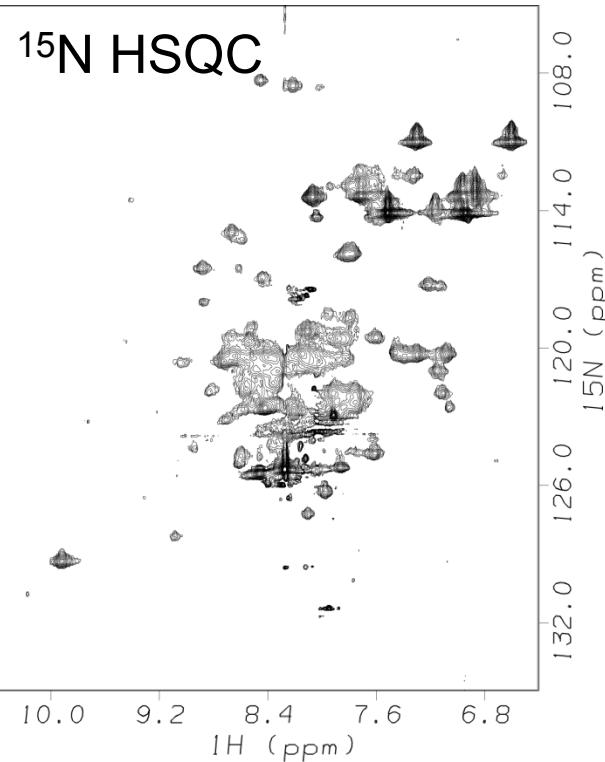


Deuteration and TROSY Greatly Improve Resolution

- Spectra of mannose binding protein illustrate dramatic resolution improvements due to TROSY, and also to deuteration

differential line
broadening due to
cross-correlation

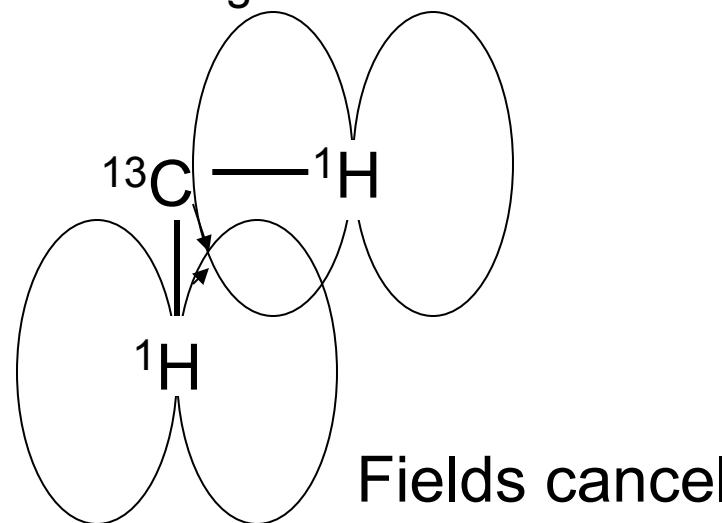
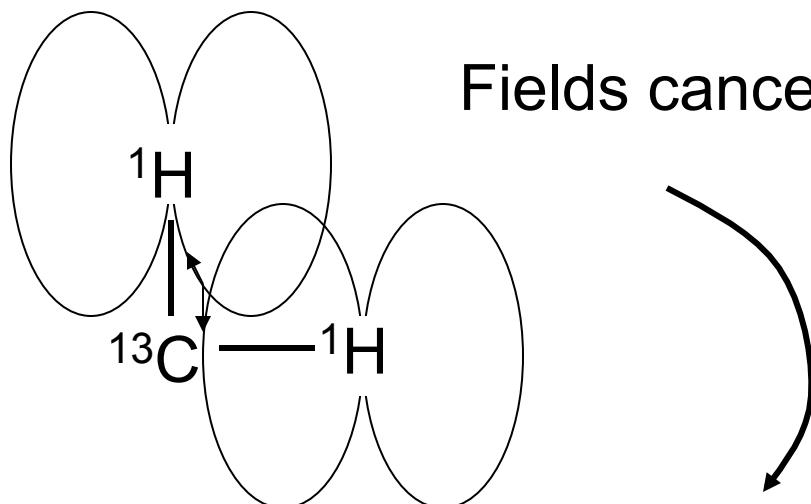
- TROSY has big effect
- deuterium is much less efficient at promoting relaxation of ^{15}N via dipole-dipole interactions, so replacing all non-exchangeable hydrogens (i.e. C-H) with deuterium also interferes with relaxation processes and sharpens lines



Other Cross-Correlated Relaxation Phenomena

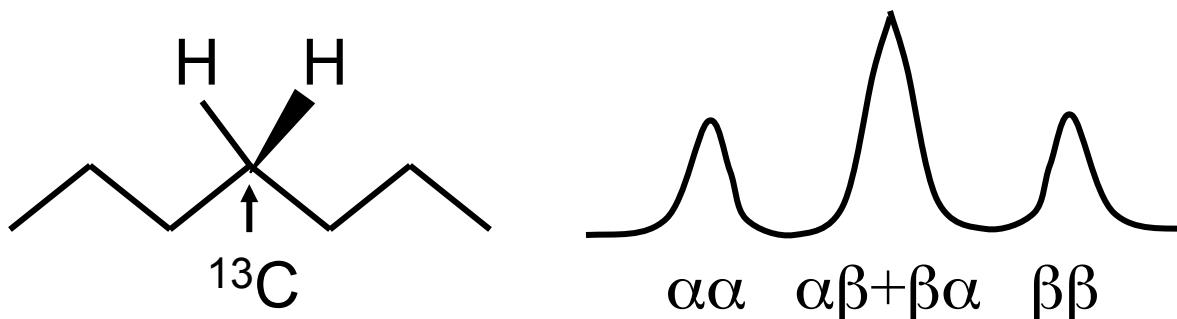
A general approach

- These types of cross-correlation interference effects are not limited to interactions between CSA and dipole-dipole relaxation (and certainly not limited to N-H groups)
- For instance, there may be instances where interference between two dipole-dipole mechanisms can reduce relaxation efficiency
 - motions must be correlated: if uncorrelated, no effect
 - will depend on geometry and motion
- Example: two hydrogens bonded to ^{13}C in a methylene group
 - each promotes relaxation of bonded ^{13}C via dipole-dipole mechanism, and these can interfere under certain conditions
 - below: dipole fields from two methylene hydrogens canceling at the ^{13}C nucleus

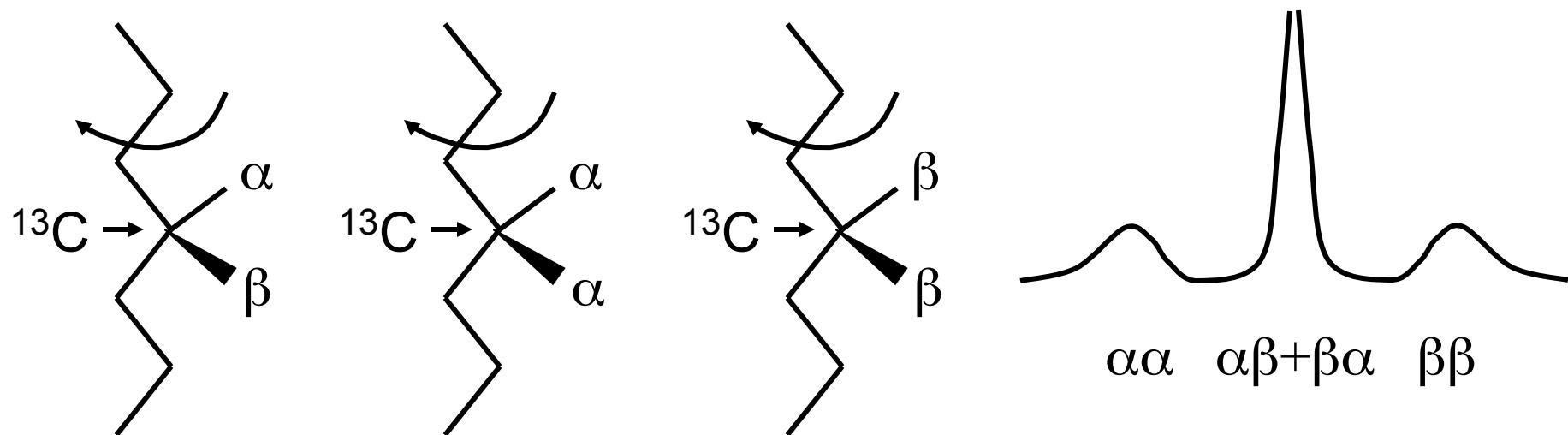


Example: Acyl Chain Rotation in Lipid Bilayers

- Without decoupling, the ^{13}C of a methylene group is a normal triplet
 - triplet shows typical 1:2:1 intensities for random tumbling of the molecule



- for *axially symmetric rotation* of the lipid in a bilayer the dipole-dipole fields interfere
- when one of the bonded ^1H nuclei is α and the other β , the dipolar fields from the ^1H nuclei oppose each other at the ^{13}C nucleus, so relaxation is less efficient and the line (center line of the triplet) narrows
- for the outer lines (both α or both β), the fields add and these lines broaden



Example: Acyl Chain Rotation in Lipid Bilayers

- Actual experimental result shown here
 - the lipids are in micelles, with mass ~1,000,000 Da
 - very fast axial rotation of the lipid sidechains in the micelle, and very sharp center lines of the triplet (if you decouple, they disappear)

J. Am. Chem. Soc., Vol. 105, No. 2, 1983 171

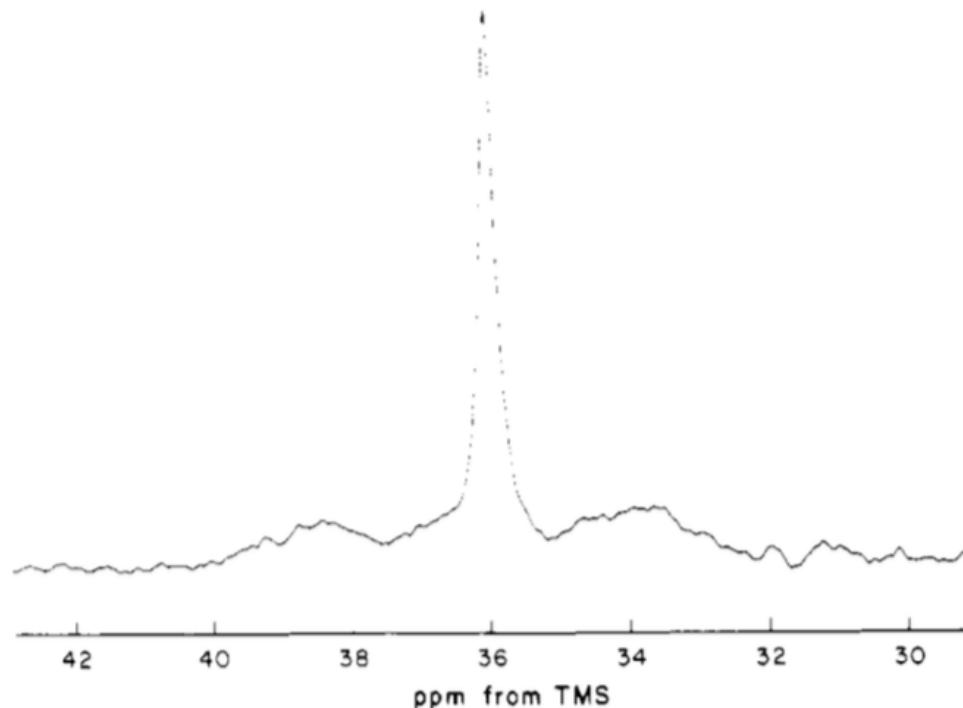
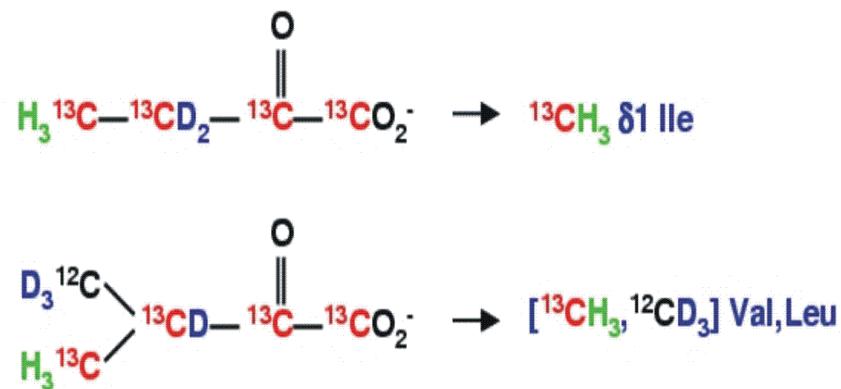
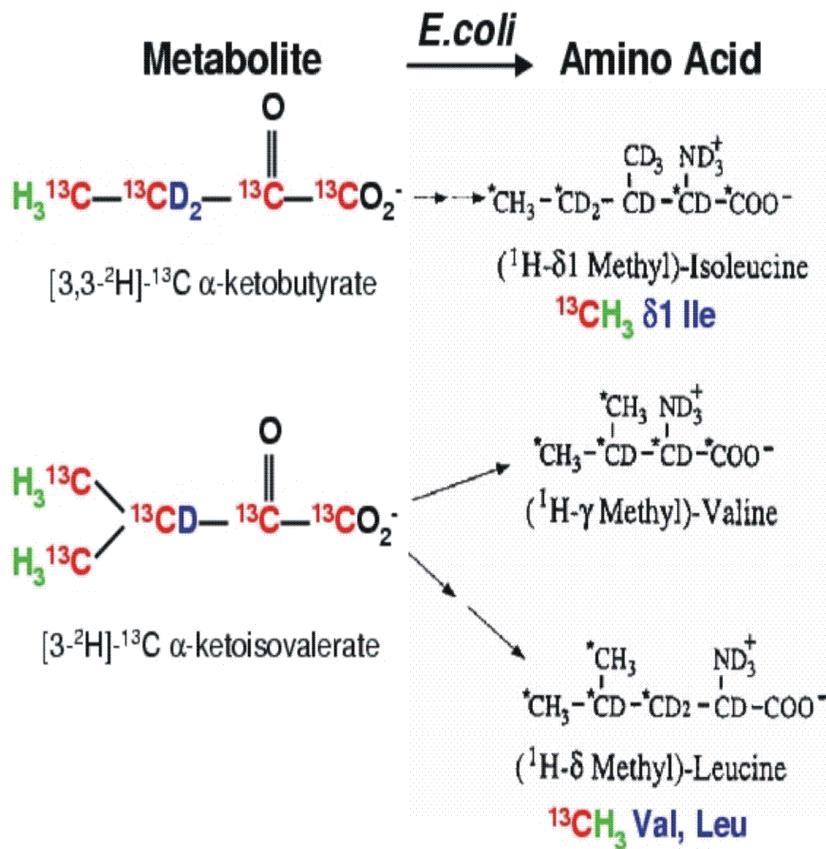


Figure 2. The 50.3-MHz ^{13}C difference spectrum of sonicated DMPC vesicles containing unenriched and 90% enriched 2- ^{13}C -MA; 7200 transients requiring 6 h were collected for each spectrum.

Fusion and Prestegard (1983) JACS

Selective Labeling of Methyl Groups Provides Sensitivity and Resolution

- Selective methyl group labeling first introduced for structure studies
 - protonated methyl groups in deuterated background decreases signal overlap and improves sensitivity by decreasing dipole-dipole relaxation from neighboring ^1H
 - particularly relevant for large proteins
 - subsequent new relaxation experiments also benefitted from this selective labeling

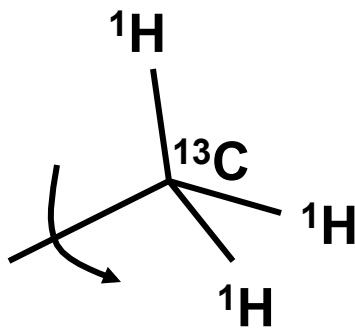


Gardner & Kay (1997) JACS **119** 7599
 Goto et al. (1999) J Biomol NMR **13** 369
 Tugarinov & Kay (2003) JACS **125**
 13868

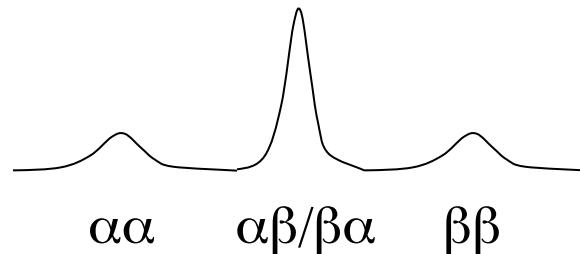
Methyl-TROSY

V. Tugarinov, R. Sprangers, L.E. Kay (2004) JACS **126**, 4921-4925 (2004)

- "Methyl-TROSY" is another example of cross-correlation effects
- Idea is getting methyl group to behave like $-\text{CH}_2-$, where there is perfect interference, in this case, for rotation about the C-C axis
- For a methyl group, for each of the bonded ^1H nuclei, in an HMQC experiment, ^{13}C evolution evolves as zero- and two-quantum coherence with the ^1H nucleus: this is coupled to the other two ^1H nuclei, so "looks" like a $-\text{CH}_2-$ group, and behaves that way (triplet)
- The dipolar contributions from the other two hydrogens cancel



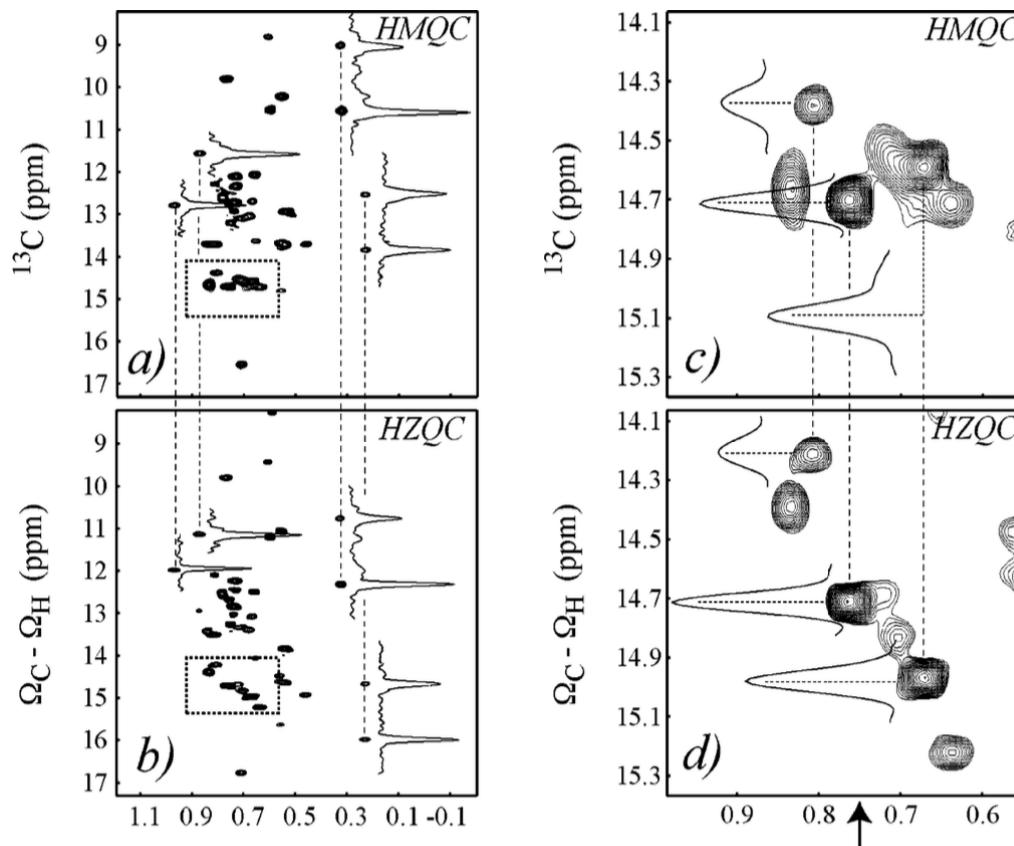
Proton coupled ZQ (H-C) spectrum



Comparison of HMQC and HZQC Data

80 kDa malate synthase G – Kay et al, 2004

- Simple HMQC will work (center line sharp, outer lines disappear for large proteins)
- Variations ("HZQC", zero-quantum part only)



Other Contributions to T_2 can Complicate Analysis (R_{ex})

Extracting and Exploiting R_{ex} is also Useful

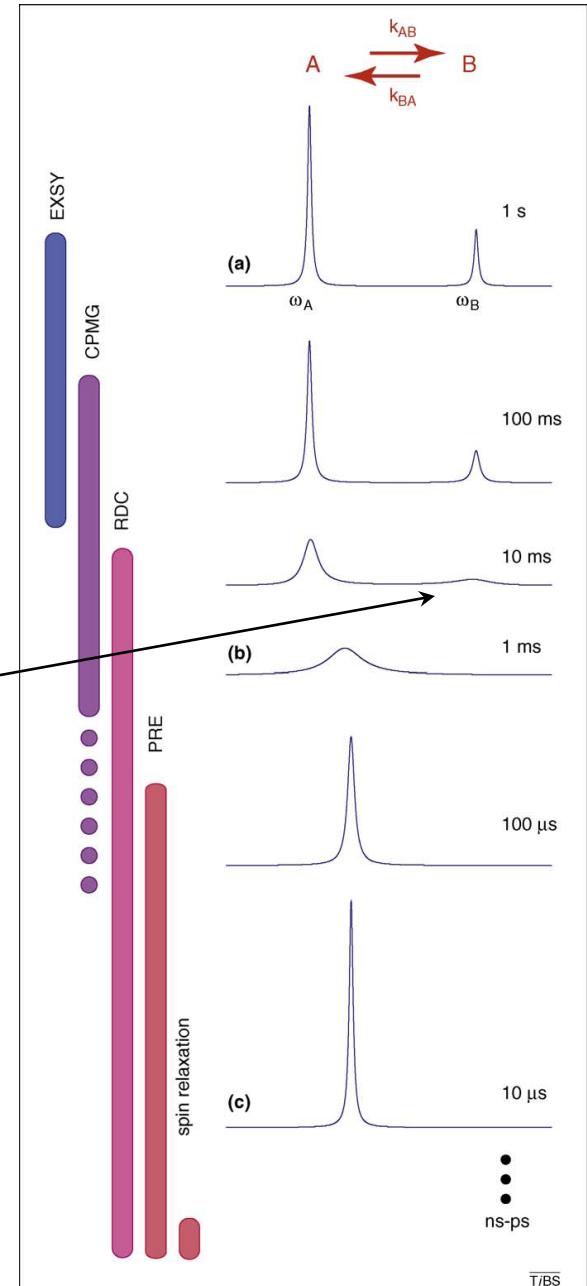
- Functionally important motional timescales in proteins are μ s-ms
- These manifest as "exchange contributions" to T_2 (shorten T_2)
 - enzyme catalysis associated with motions on these timescales
 - conformational exchange associated with functionally relevant binding states
- Experiments to measure these directly typically referred to as "relaxation dispersion" experiments
 - Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy, Vallurupalli P, Hansen DF, Kay LE, PNAS, 105, 11766-11771 (2008)
 - Characterization of enzyme motions by solution NMR relaxation dispersion, Loria JP, Berlow RB, Watt ED, Acc. Chem. Res., 41, 214-221 (2008)
 - Observing biological dynamics at atomic resolution using NMR, Mittermaier AK, Kay LE, Trends Biochem. Sci., 34, 601-611 (2009)

NMR senses dynamics on many time scales

- Chemical exchange (conformational exchange) is interconversion of states
 - slow interconversion of states A and B relative to their chemical shift difference gives two signals
 - for fast exchange, a single signal is observed, at a weighted average chemical shift
 - in between, lines broaden
- States with low population, with particular exchange rates, can disappear
 - under appropriate conditions, measuring R_{ex} gives access to populations of states and frequency difference (i.e. can "see" invisible states)

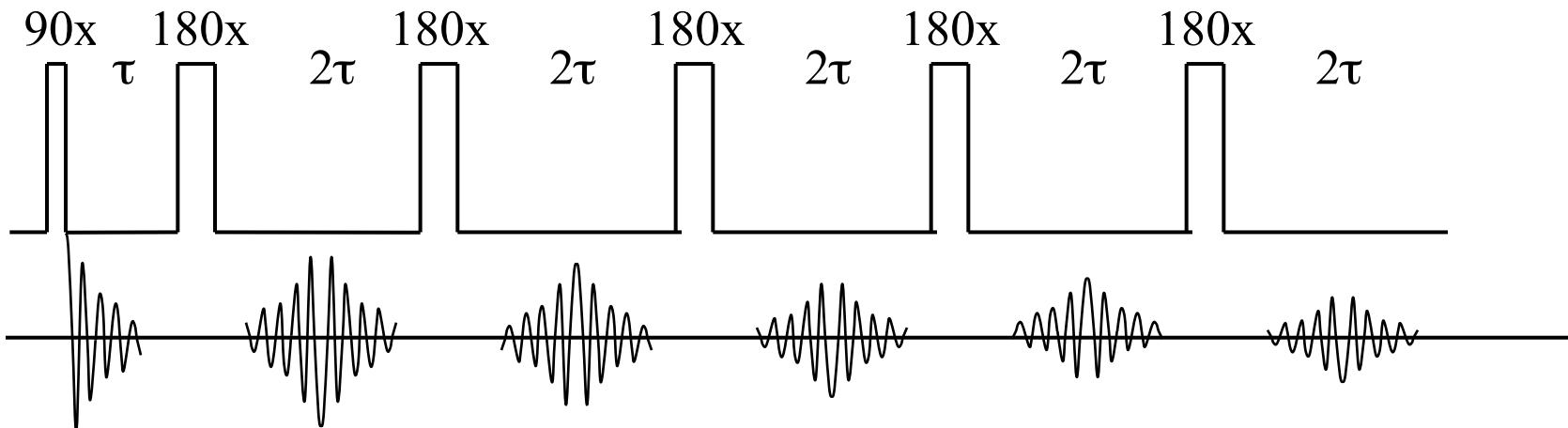
$$R_{ex} = \tau_{ex} p_A p_B \Delta\omega^2$$

$$\tau_{ex}^{-1} = \tau_A^{-1} + \tau_B^{-1}$$



Carr-Purcell Meiboom-Gill Sequence Can Remove Effects of Exchange

- At typical τ values, sequence removes effect of inhomogeneity
 - decay of echos is T_2 exponential decay, and includes exchange (R_{ex})
- If τ is very short, slow R_{ex} processes can't occur
 - thus, result is T_2 that does *NOT* include R_{ex}



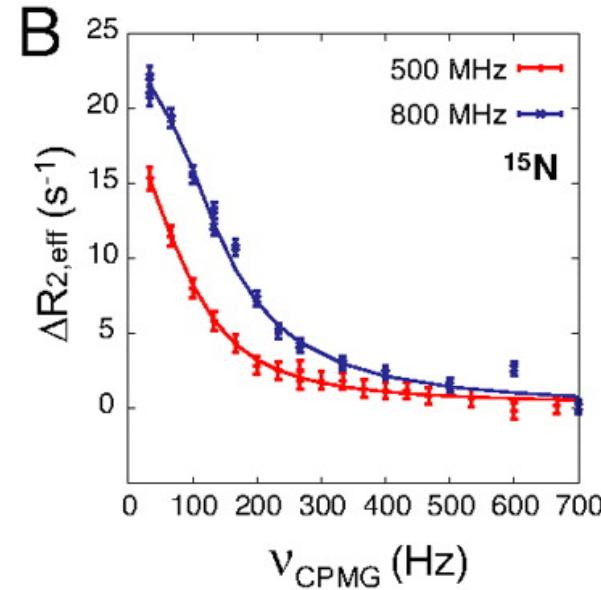
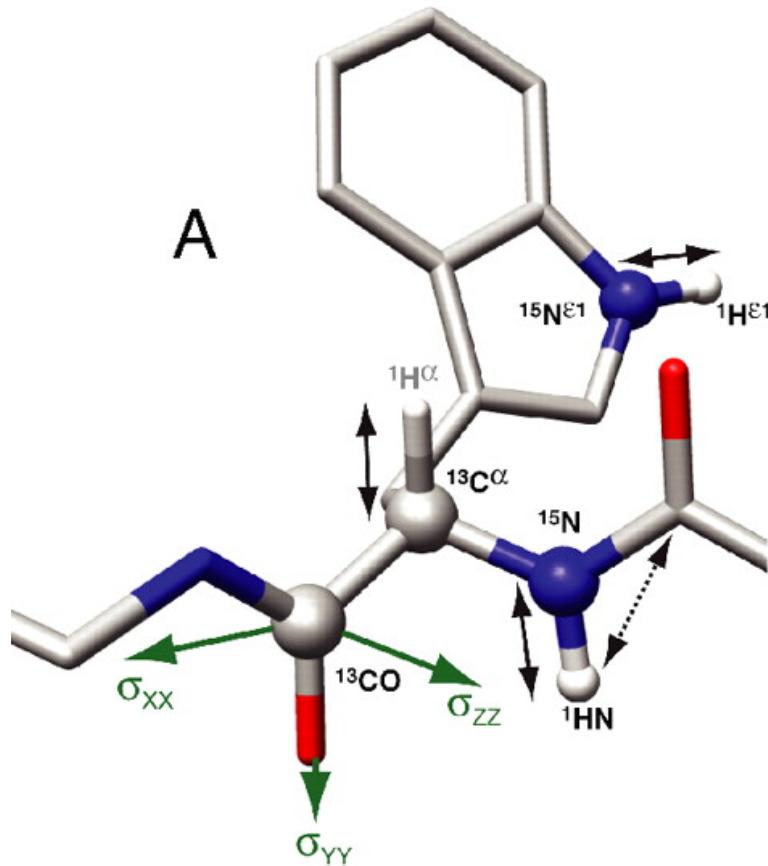
- So, relaxation dispersion experiment is performed as function of τ
 - no change with τ signifies no R_{ex} , change with τ signifies significant R_{ex}

$$R_2(1/\tau) = R_2^0 + \varphi_{\text{ex}}/k_{\text{ex}}[1 - 2\tanh(k_{\text{ex}}\tau/2)\gamma(k_{\text{ex}}\tau)]$$
$$\varphi_{\text{ex}}/k_{\text{ex}} = p_A p_B \Delta\omega^2$$

- simple experiment as function of τ permits detection of R_{ex} , but in order to extract populations and frequency difference, need to do experiment at >1 field strength

Field Dependent Measurement Separates $\Delta\omega$ and $P_{A,B}$ information (Kay, PNAS, 2008)

- Typical results showing significant R_{ex} at two field strengths



Detection of 5-10% minor species of peptide bound to SH3 domain

- Example: detection of very low (5%) minor state populations
 - have to be very careful in interpretation of results (proteins exist in many states, all the time)

