

Sequential Resonance Assignment Strategies for Proteins

BCMB/CHEM 8190

Chemical Shift Assignments

- In order to determine a structure by traditional, NOE-based ^1H - ^1H distance-based methods, the chemical shifts of the individual ^1H nuclei must be known
- Older methods were based on homonuclear (^1H) experiments only, whereas newer methods based on heteronuclear / triple resonance experiments
- Triple resonance techniques generally require ^1H , ^{13}C and ^{15}N chemical shift assignments in a protein
- Once chemical shifts are known, these can be used to assign signals in NOE spectra for distance measurements
- Chemical shifts so obtained are useful for predicting secondary structure of proteins including main chain ϕ and ψ angles

^1H Methods: The Sequential Assignment Procedure

- Due to Wüthrich and coworkers (early/mid 1980's)

“1. Identification of amino acid side chain spin systems. 2. Identification of neighboring residues in the amino acid sequence. 3. Suitable combinations of the results from 1 and 2 which provides individual resonance assignments in the primary structure of the protein.”

K. Wüthrich 1983 *Biopolymers* **22**, 131-138

- 1). Use J correlation spectroscopy (COSY, TOCSY) to identify individual spin systems (amino acids)
- 2). Use NOE spectroscopy (NOESY) to link amino acids

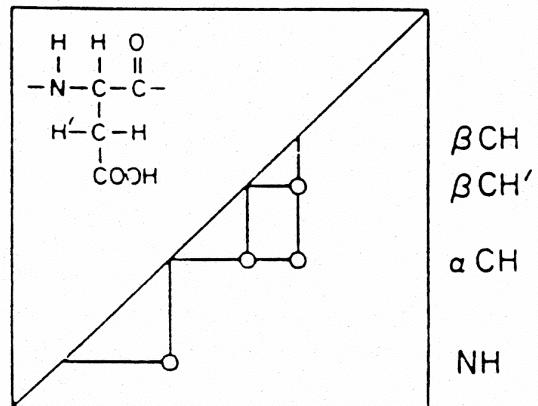
Book: “NMR of Proteins and Nucleic Acids”, Kurt Wüthrich, 1986 (Wiley)

^1H Methods: The Sequential Assignment Procedure

1). Identify amino acid spin systems using COSY and TOCSY

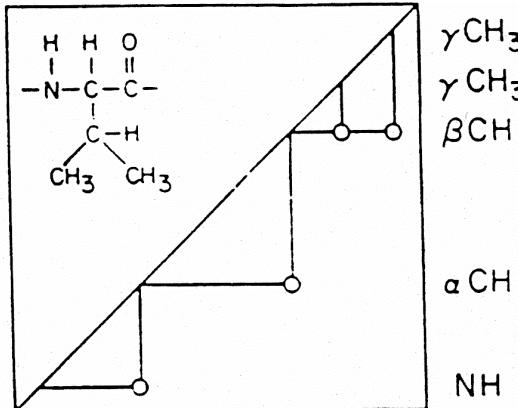
- Spin systems are classified based on the spins that can be correlated in COSY spectra, and then TOCSY spectra
 - remember, there is no way to connect amino acids in a protein by COSY or TOCSY
 - a COSY or TOCSY spectrum of a protein is comprised of the individual spectra of the amino acids in the protein
- The individual spin systems are identified as to amino acid type
 - how? Based on “amino acid specific patterns” of COSY and TOCSY cross peaks

Asp



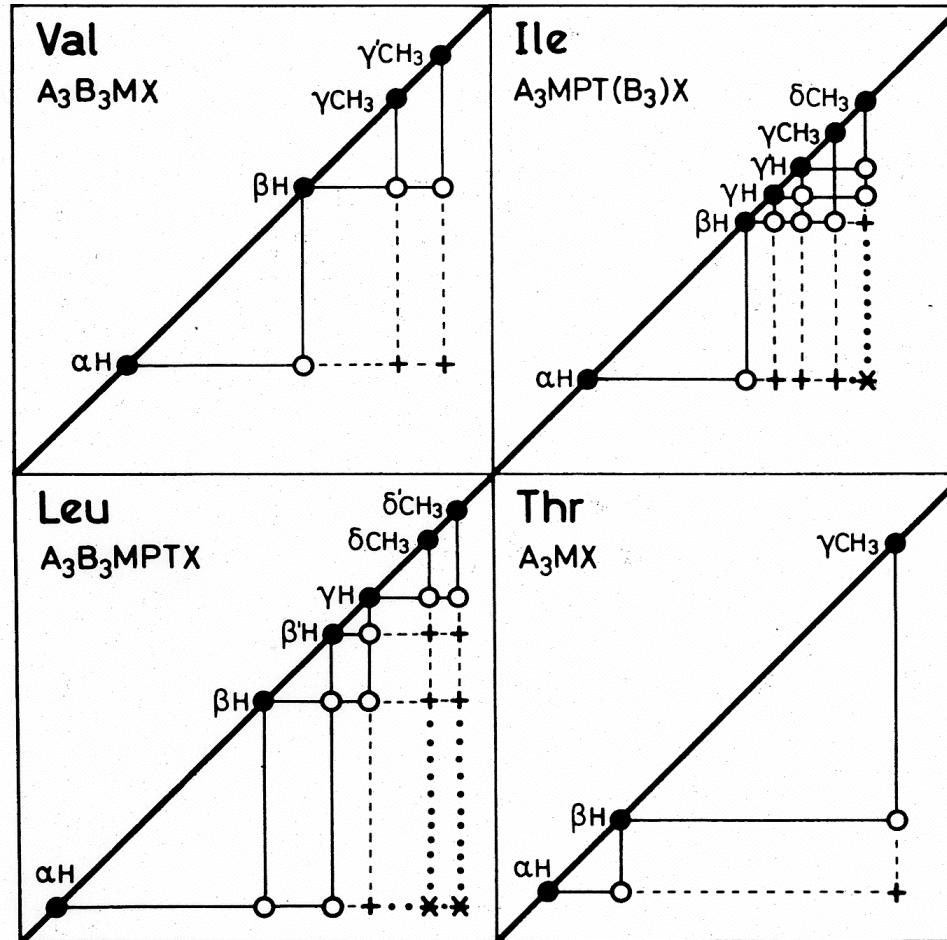
βCH
 $\beta\text{CH}'$
 αCH
NH

Val

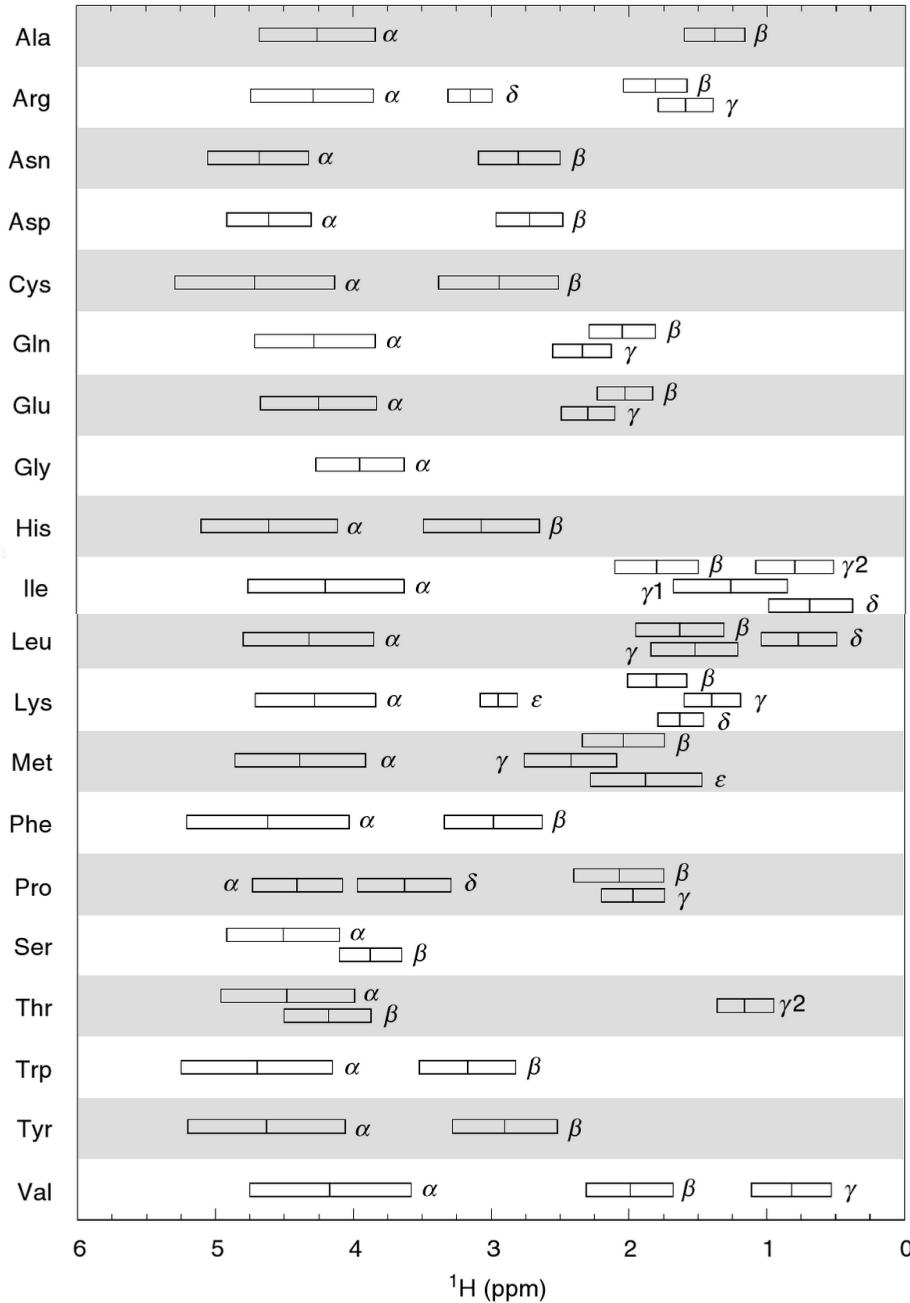


γCH_3
 γCH_3
 βCH
 αCH
NH

More Examples of COSY Patterns



^1H Chemical Shift Information

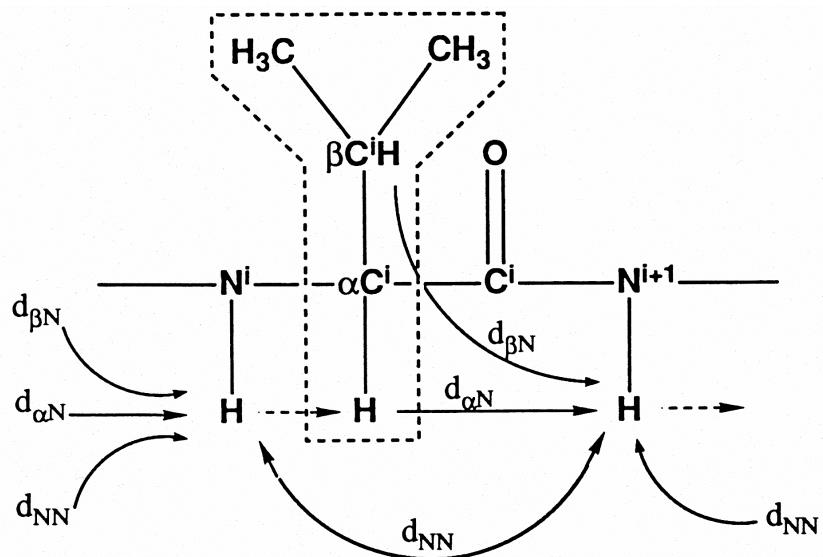


- ^1H chemical shift information is also very useful in assisting to identify spin systems
 - shown are distributions of ^1H chemical shifts for amino acids in proteins

^1H Methods: The Sequential Assignment Procedure

2). Link identified amino acids using NOESY data

- Critical observation: it is virtually impossible to align the main chain atoms of two adjacent amino acid residues in a protein so that there is not at least one pair of inter-residue distances between main chain hydrogens that is significantly less than the NOE detection limit ($\sim 5\text{\AA}$)



dashed lines: nonlabile protons correlated by J correlation experiments

arrows: sequential interresidue NOEs

Sequential Inter-residue NOEs

- Most short inter-proton ($^1\text{H}^{\text{N}}$, $^1\text{H}^{\text{a}}$, $^1\text{H}^{\text{b}}$) distances in proteins are between directly adjacent amino acid residues
 - intense NOEs indicate adjacent amino acids
- Less intense i , $i+2$ and i , $i+3$ NOEs (“nonsequential”) also observed and useful, particularly in well-defined secondary structures

TABLE 8.1. Statistics of Short ^1H – ^1H Distances in Protein Crystal Structures.^a

Distance (Å)	$j-i = 1$ (%)
$d_{\alpha\text{N}}(i,j) \leq 2.4$	98
≤ 3.0	88
≤ 3.6	72
$d_{\text{NN}}(i,j) \leq 2.4$	94
≤ 3.0	88
≤ 3.6	76
$d_{\beta\text{N}}(i,j) \leq 2.4$	79
≤ 3.0	76
≤ 3.6	66
$d_{\alpha\text{N}}(i,j) \leq 3.6, d_{\text{NN}} \leq 3.0$	99
$d_{\alpha\text{N}}(i,j) \leq 3.6, d_{\beta\text{N}} \leq 3.4$	95
$d_{\text{NN}}(i,j) \leq 3.0, d_{\beta\text{N}} \leq 3.0$	90

TABLE 7.1. Short (≤ 4.5 Å) Sequential and Medium-Range ^1H – ^1H Distances in Polypeptide Secondary Structures

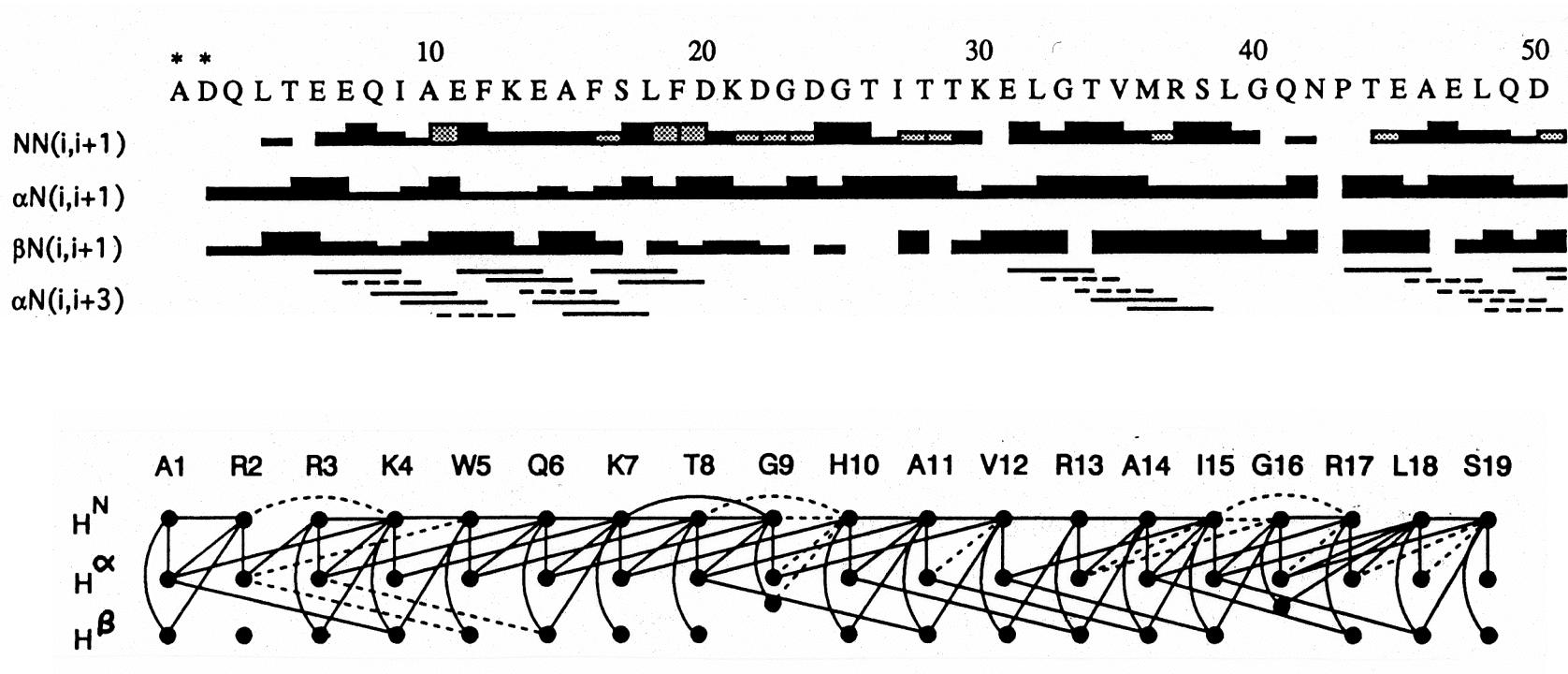
Distance	α -helix	β	β_{P}	turn I ^a	turn II ^a
$d_{\alpha\text{N}}$	3.5	3.4	2.2	2.2	3.4 2.2
$d_{\alpha\text{N}}(i,i+2)$	4.4	3.8		3.2	3.2
$d_{\alpha\text{N}}(i,i+3)$	3.4	3.3		3.6	3.3
$d_{\alpha\text{N}}(i,i+4)$	4.2			3.1–4.2	3.8–4.7
d_{NN}	2.8	2.6	4.3	4.2	2.6 4.5
$d_{\text{NN}}(i,i+2)$	4.2	4.1		2.4	2.4
$d_{\beta\text{N}}^{\text{b}}$	2.5–4.1	2.9–4.4	3.2–4.5	3.7–4.7	2.9–4.4 3.6–4.6
$d_{\alpha\beta}(i,i+3)^{\text{b}}$	2.5–4.4	3.1–5.1		3.6–4.6	3.6–4.6

^a For the turns, the first of two numbers applies to the distance between residues 2 and 3, the second to that between residues 3 and 4 (Fig. 7.12). The range indicated for $d_{\alpha\text{N}}(i,i+3)$ corresponds to the distances adopted if ψ_1 is varied between -180 and 180° .

^b The ranges given correspond to the distances adopted by a β -methine proton if χ^1 is varied between -180 and 180° .

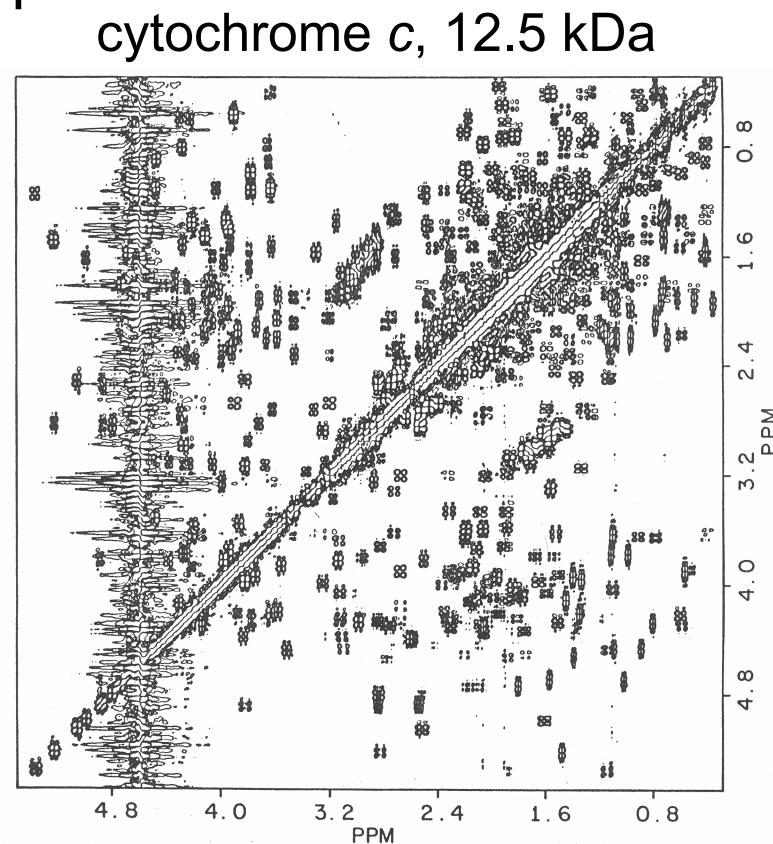
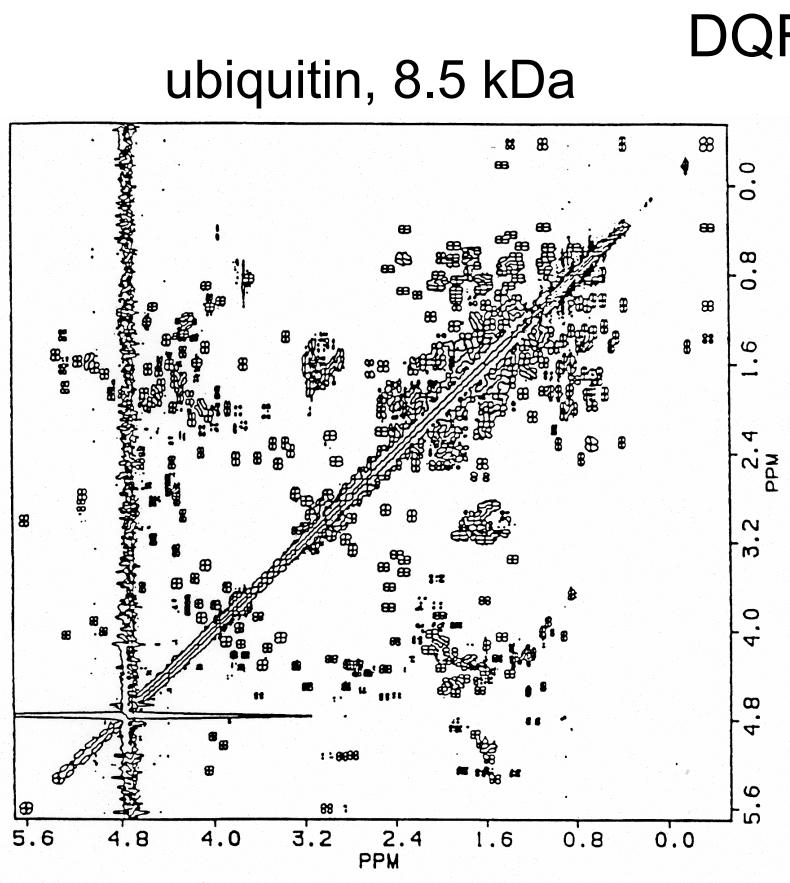
Sequential Inter-residue NOEs

- Examples of NOE connectivities
- Graphical means to summarize NOEs and demonstrate self-consistency of assignments



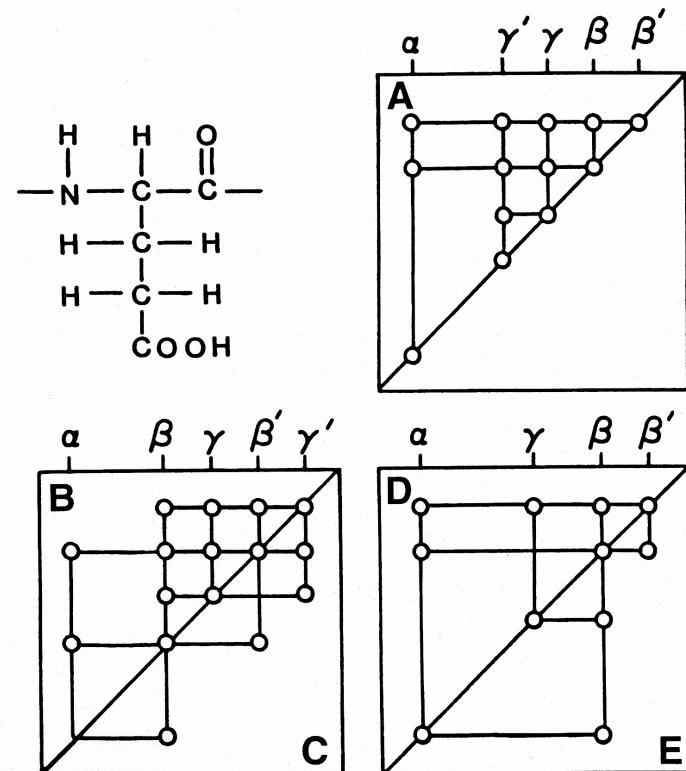
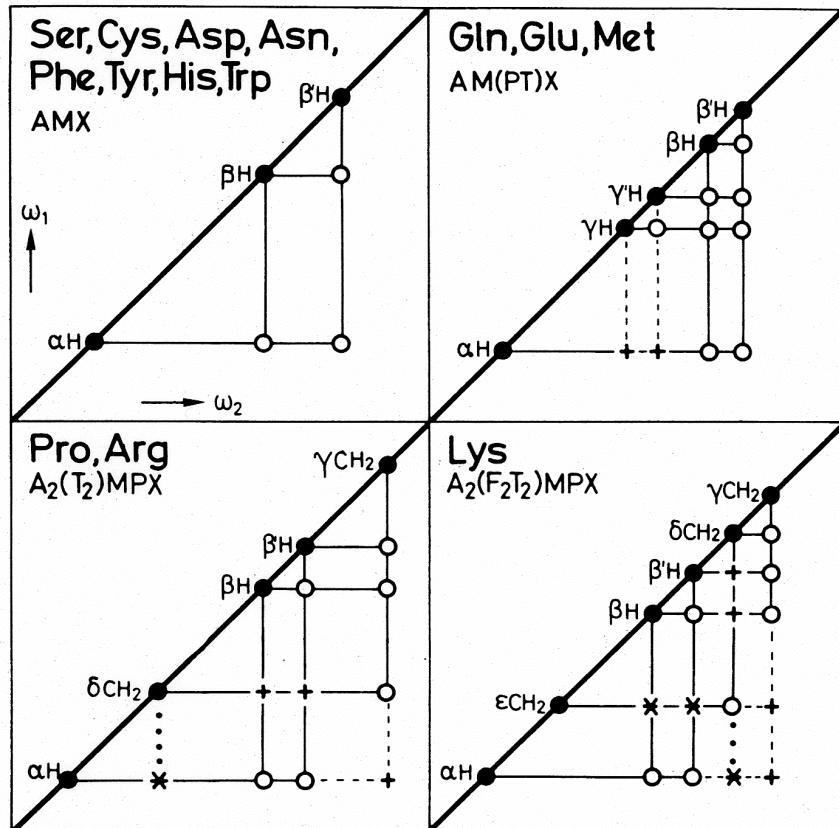
Problems with Sequential Assignment Procedure

- Spectral degeneracy can preclude complete COSY/TOCSY pattern identification



Problems with Sequential Assignment Procedure

- Degeneracy of COSY patterns
 - normally, only 8-12 species can be distinguished from among the 20 amino acids prior to NOE-based sequence alignment
- Plasticity of COSY patterns
 - in proteins, local environment effects on *J*-coupling and relaxation/dynamics lead to variability in COSY patterns



Main Chain Directed (MCD) Approach

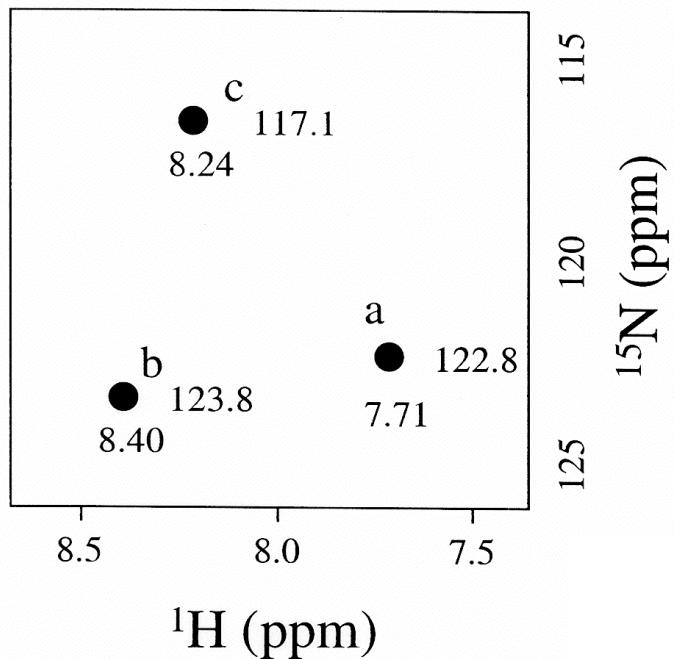
- Due to Wand/Englander and coworkers (mid/late 1980's)
- 1). use J correlated and spectra to first identify $^1\text{H}^{\text{N}}\text{-}^1\text{H}^{\text{a}}\text{-}^1\text{H}^{\text{b}}$ ("NAB") units
 - the $^1\text{H}^{\text{N}}$ region of COSY spectra is usually less crowded
 - individual amino acid type identification not attempted at this point
- 2). next align the NAB units sequentially using NOESY spectra
 - pattern recognition routines employed to search for well-established patterns of NOEs
 - amino acid type identification is then substantially restricted
- No initial reliance on identity of amino acid to establish connectivity
- Amenable to automation (pattern matching algorithms)

Triple Resonance Approach

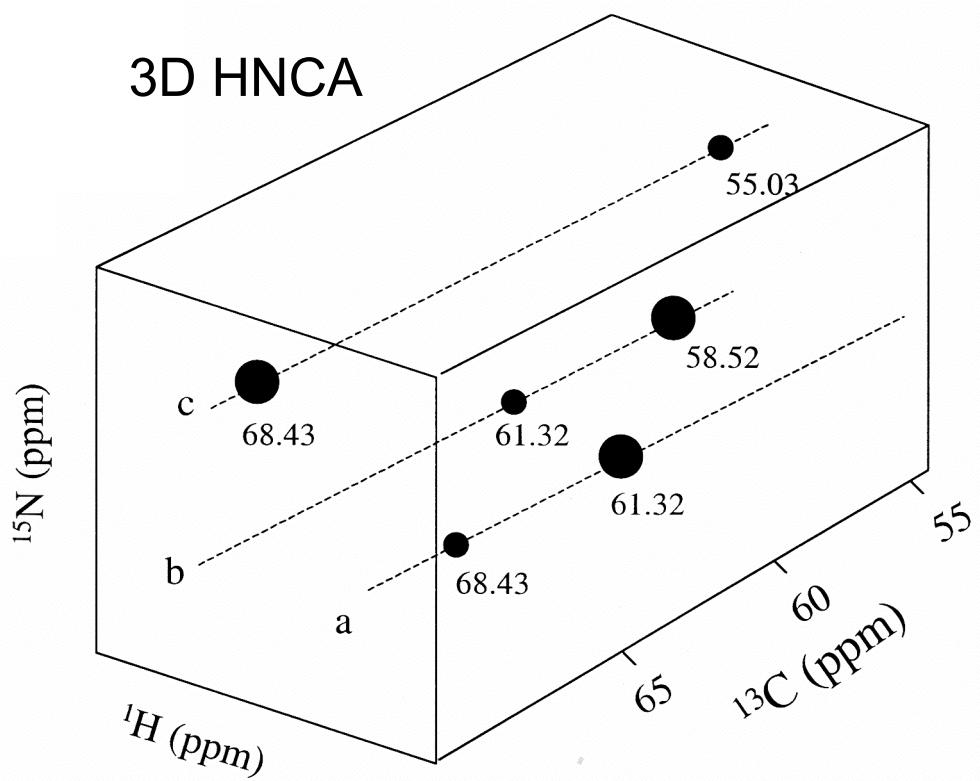
- The triple resonance approach has largely replaced ^1H -only methods in cases where the protein can be isotopically labeled
 - is the only method currently available for large proteins
- Experiments provide selective chemical shift correlation of main chain (plus $^1\text{H}^\beta$ and $^{13}\text{C}^\beta$) nuclei in adjacent amino acids
 - these correlations permit links between individual amino acids to be established, and thus assignment of the main chain
 - the chemical shifts of side chain nuclei are then correlated with assigned main chain nuclei to complete side chain assignments
- In the ideal case, no other information is required
 - in the ideal case, and in theory, amino acid type need not be established initially
 - in practice, many other types of information, including chemical shift/amino acid type information, NOE distance information, etc., play important roles

Triple Resonance Approach: A Simple Example

2D HNCA projection

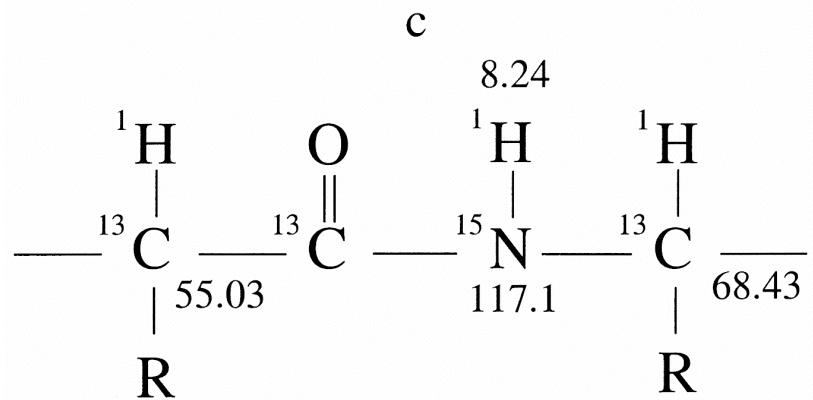
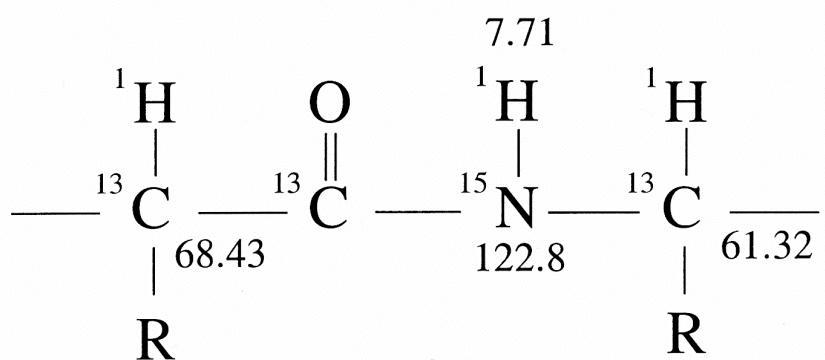
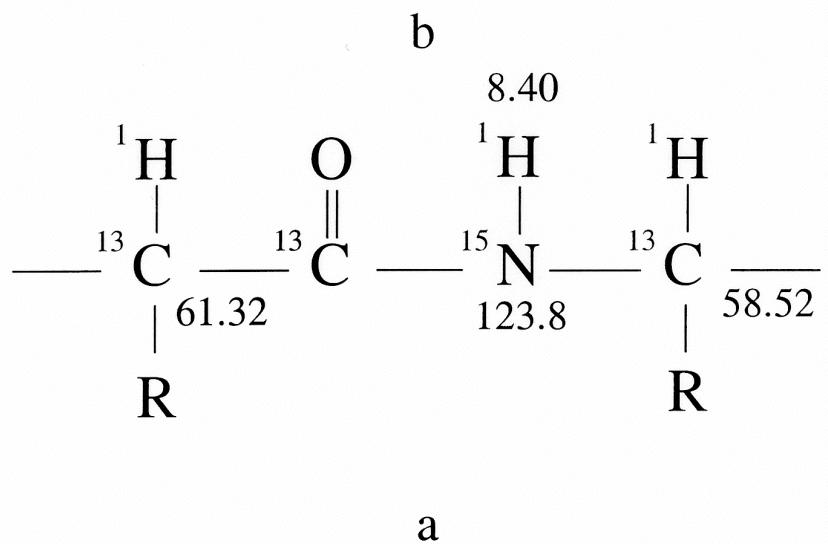


3D HNCA



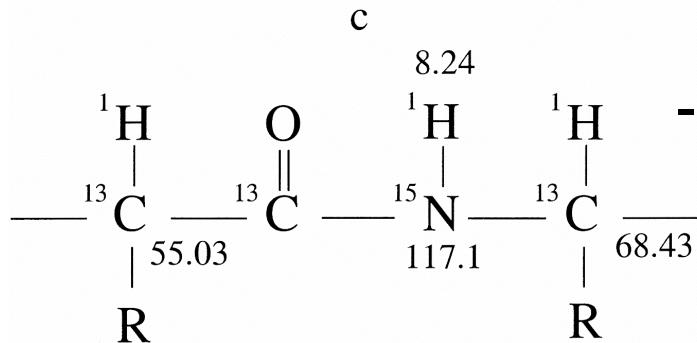
	c	a	b
¹ H ^N _i	8.24	7.71	8.40
¹⁵ N _i	117.1	122.8	123.8
¹³ C ^α _i	68.43	61.32	58.52
¹³ C ^α _{i-1}	55.03	68.43	61.32

Triple Resonance Approach: A Simple Example

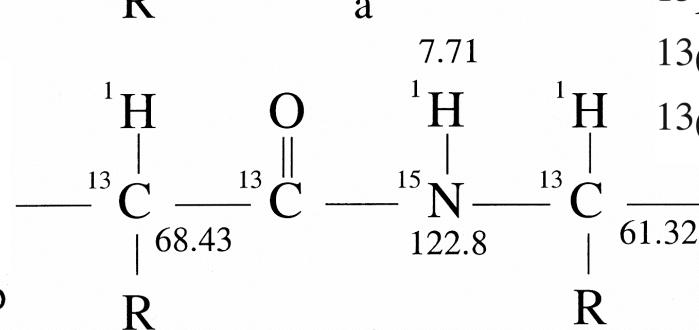


	c	a	b
${}^1\text{H}{}^{\text{N}}_i$	8.24	7.71	8.40
${}^{15}\text{N}_i$	117.1	122.8	123.8
${}^{13}\text{C}^{\alpha}_i$	68.43	61.32	58.52
${}^{13}\text{C}^{\alpha}_{i-1}$	55.03	68.43	61.32

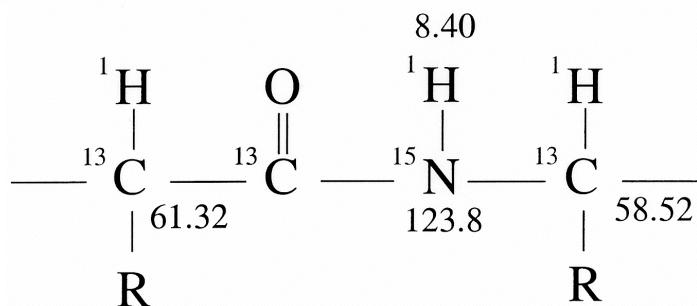
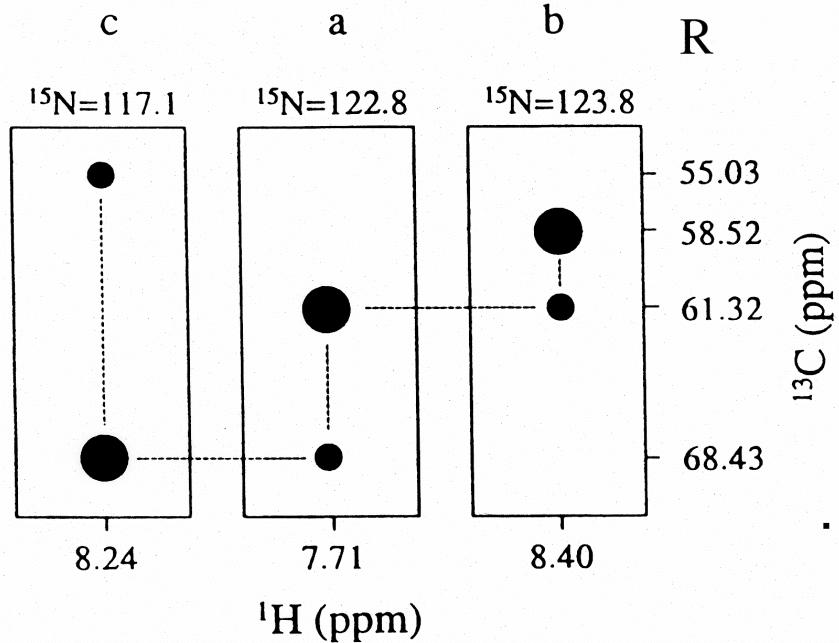
Triple Resonance Approach: A Simple Example



- link the correlated shifts numerically....

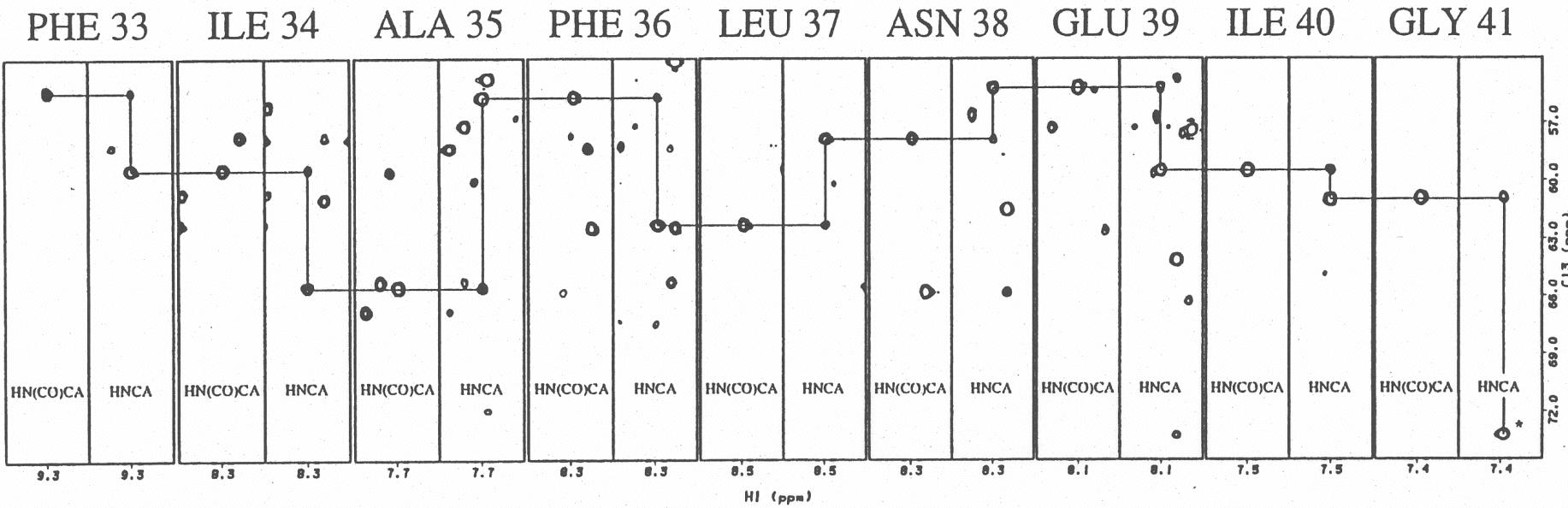


	c	a	b
$^1\text{H}^{\text{N}}_i$	8.24	7.71	8.40
$^{15}\text{N}_i$	117.1	122.8	123.8
$^{13}\text{C}^{\alpha}_i$	68.43	61.32	58.52
$^{13}\text{C}^{\alpha}_{i-1}$	55.03	68.43	61.32



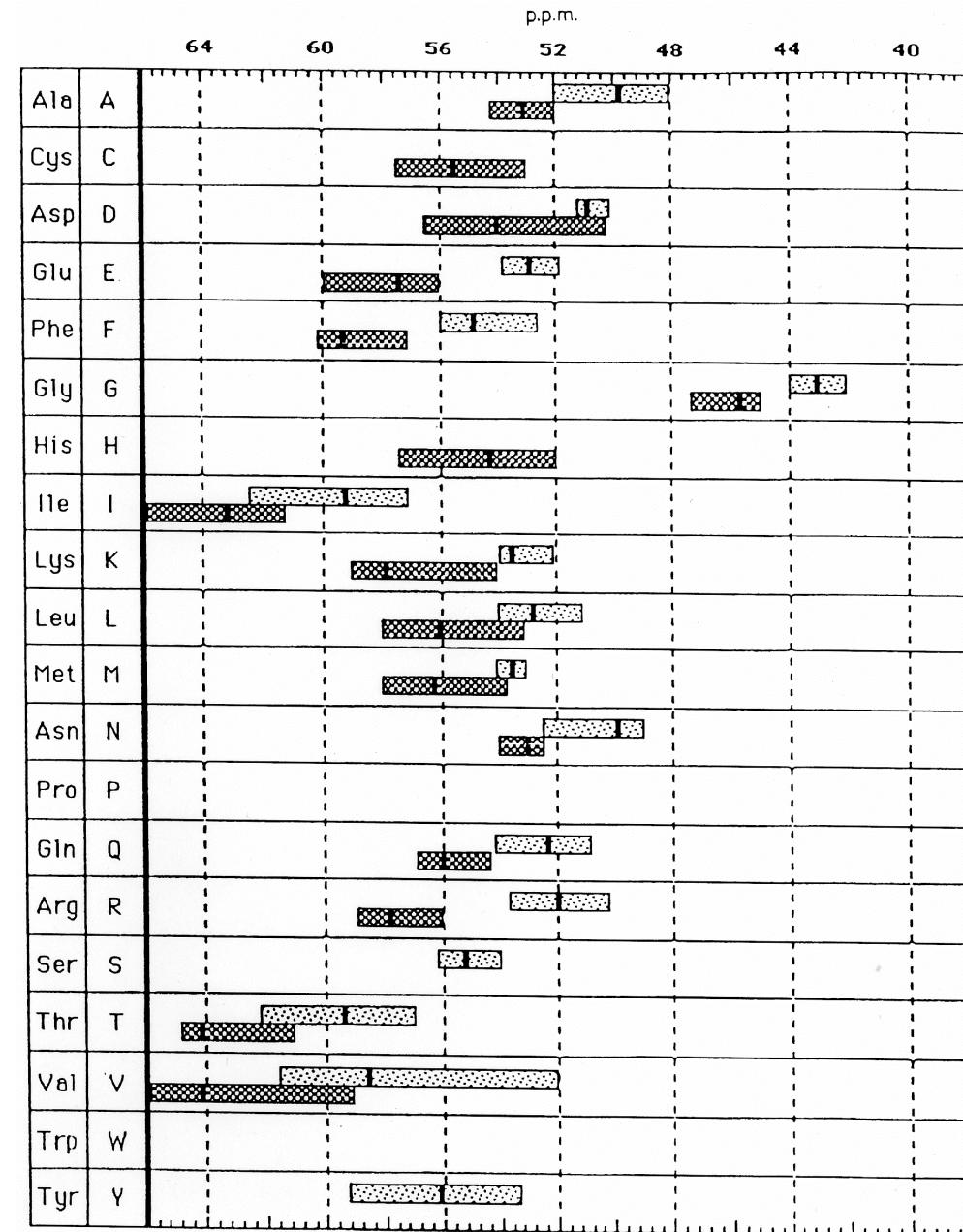
.... or visually

Triple Resonance Approach: HNCA HN(CO)CA Example



- Problems:
 - $^{13}\text{C}^\alpha$ chemical shift degeneracy in proteins
 - $^{13}\text{C}^\alpha$ linewidths/resolution
 - these preclude complete linkage via $^{13}\text{C}^\alpha$ alone
 - the same is true for $^{13}\text{C}^\beta$, $^{13}\text{C}'$

$^{13}\text{C}^\alpha$ Chemical Shifts

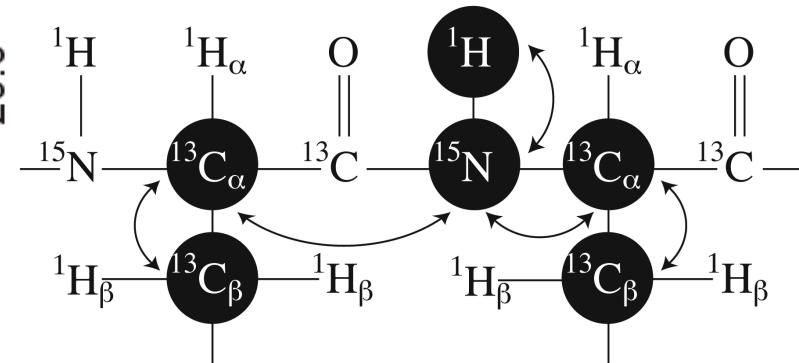
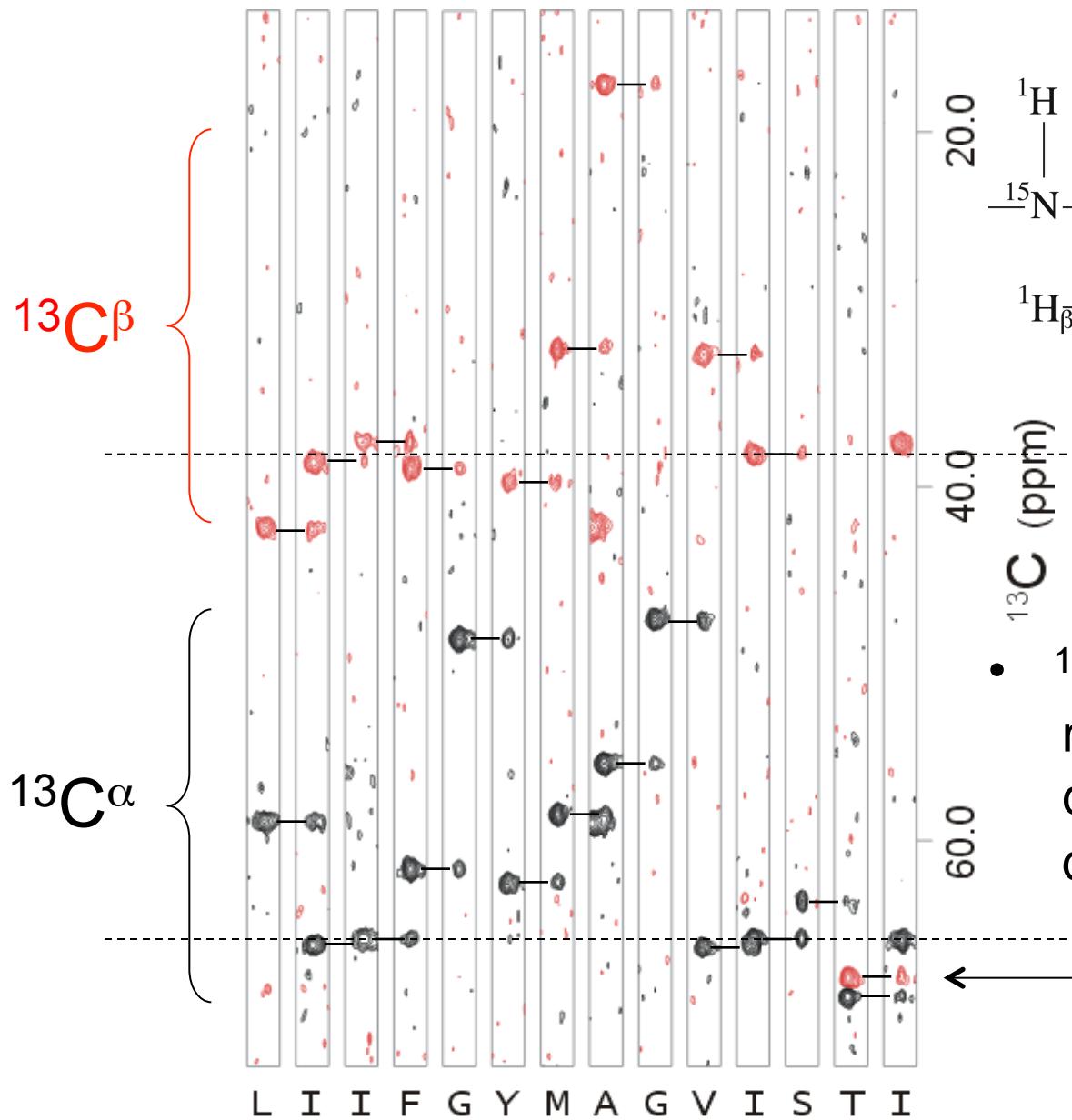


- $^{13}\text{C}^\alpha$ chemical shift degeneracy limits the utility of experiments that link amino acids via $^{13}\text{C}^\alpha$
- $^{13}\text{C}^\alpha$ chemical shifts can be useful for identifying some amino acids (Gly, Ile/Val)
- $^{13}\text{C}^\alpha$ chemical shifts are also good indicators of 2° structure (providing the amino acid type is known)

$^{13}\text{C}^\alpha$ chemical shifts in proteins

- **dark bars, alpha helices**
- light bars, beta sheets

HNCACB / CBCA(CO)NH, connectivity via $^{13}\text{C}^\beta$



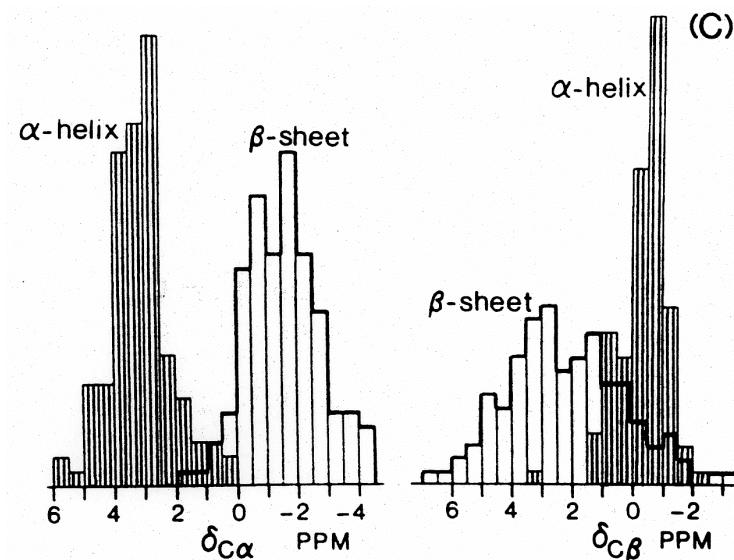
- $^{13}\text{C}^\beta$ correlations permit resolution of ambiguities in connectivities due to $^{13}\text{C}^\alpha$ degeneracy

$^{13}\text{C}^\beta$ Chemical Shifts

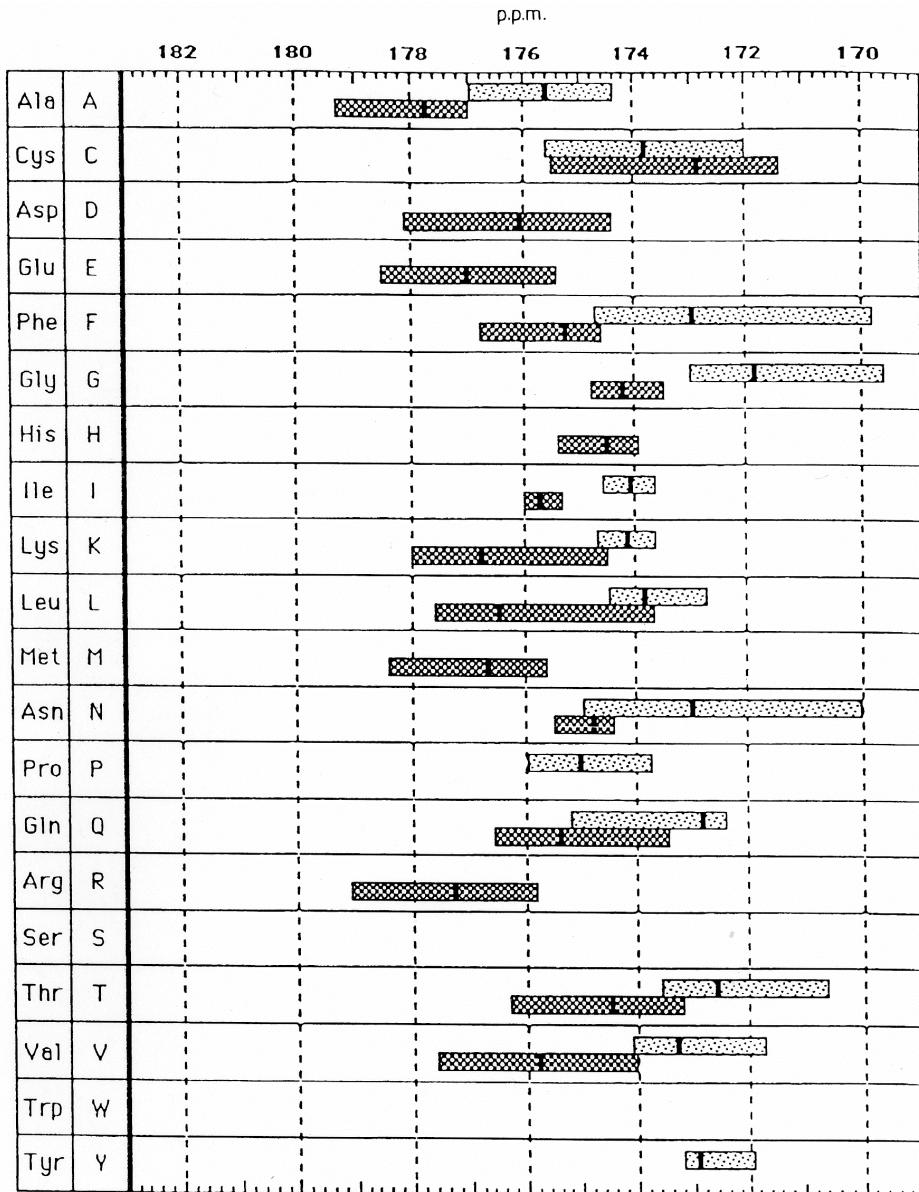
^{13}C chemical shifts in proteins



- $^{13}\text{C}^\beta$ chemical shift degeneracy, on average, is not as severe as $^{13}\text{C}^\alpha$ degeneracy
- $^{13}\text{C}^\beta$ chemical shifts are also useful for amino acid identification (Ser/Thr, Ala)
- $^{13}\text{C}^\alpha$ chemical shifts are also good indicators of 2° structure (providing the amino acid type is known)



HNCO / HN(CA)CO, Connectivity via $^{13}\text{C}'$ (Carbonyl)

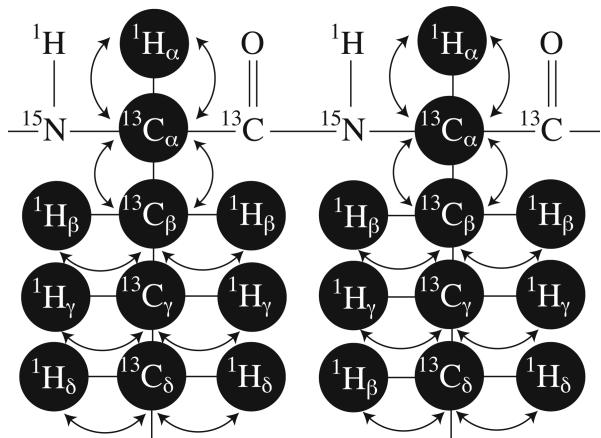
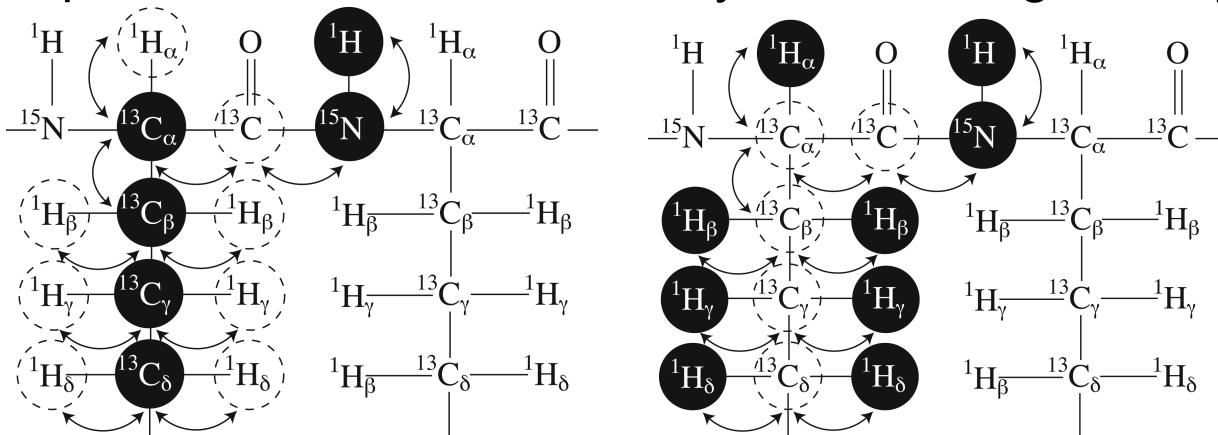


- $^{13}\text{C}'$ correlations permit resolution of ambiguities in connectivities due to $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ degeneracy
- $^{13}\text{C}'$ chemical shifts somewhat degenerate, but good indicators of 2° structure

$^{13}\text{C}'$ chemical shifts in proteins
-dark bars, alpha helices
-light bars, beta sheets

Sidechain Assignments

- Side chain ^1H and ^{13}C resonances are assigned based on correlations with main chain $^1\text{H}^\text{N}$, ^{15}N , $^1\text{H}^\alpha$, and $^1\text{H}^\beta$ and $^{13}\text{C}^\beta$ chemical shifts
 - $^1\text{H}^\text{N}$, ^{15}N resolved TOCSY experiments are often used, but don't provide for direct correlation of side chain ^{13}C and ^1H nuclei with one another
 - these experiments are also limited by $^1\text{H}^\text{N}$, ^{15}N degeneracy



- HCCH TOCSY and HCCH COSY experiments provide direct side chain ^1H , ^{13}C connectivities
- NOESY data are also used to assist in assignment of side chain resonances

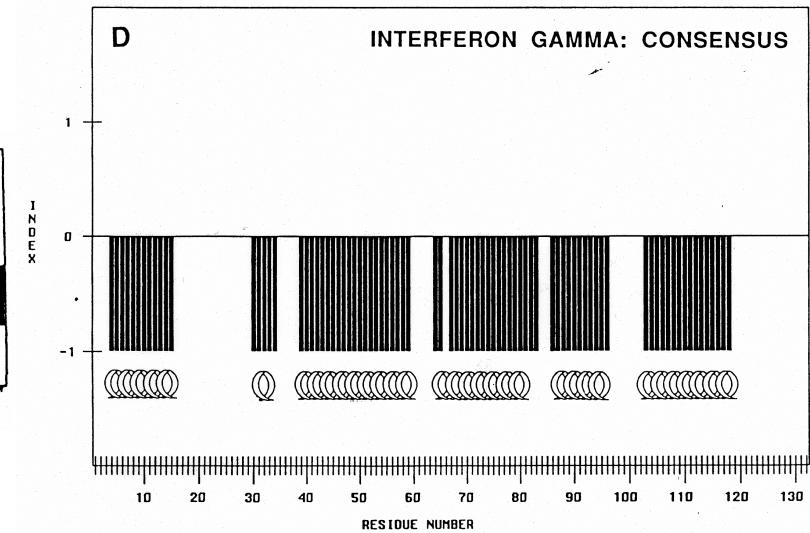
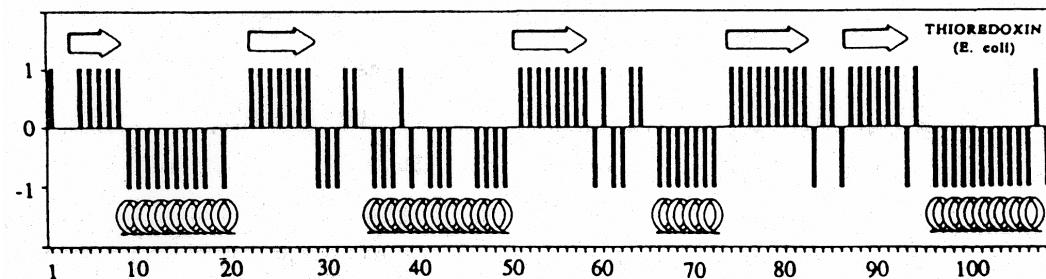
Other (Main Chain) Assignment Procedures

- Automated/semi-automated assignment procedures
 - “Autoassign” (1), “PACES” (2), “GARANT” (3), etc.
 - performance (usually) highly dependent on data quality, completeness, and peak picking
 - usually require, as input, correlated spins from many triple resonance experiments ($^1\text{H}^N_i$, $^{15}\text{N}_i$, $^{13}\text{C}^\alpha_i$, $^{13}\text{C}^\alpha_{i-1}$, $^{13}\text{C}^\beta_i$, $^{13}\text{C}^\beta_{i-1}$, $^{13}\text{C}'_i$, $^{13}\text{C}'_{i-1}$)
 - other types of useful information permitted (amino acid type if known, etc.)
 - usually require, or work best with, some interactive user intervention
- “no assignment” procedures
 - residual dipolar coupling method for simultaneous structure determination and resonance assignment (4, Prestegard group)
 - measured dipolar couplings and chemical shifts limit connectivity ambiguity via $^{13}\text{C}^\alpha$ permitting

1. Zimmerman et al. (1997) *J. Mol. Biol.* **269**, 592-610; Moseley and Montelione (1999) *Curr. Opin. Struct. Biol.* **9**, 635-642 (http://www.nesg.org/web_tools.html)
2. Coggins and Zhou (2003) *J. Biomol. NMR* **26**, 93-111
3. Bartels et al. (1996) *J. Biomol. NMR* **7**, 207-213; Malmodin et al. (2003) *J. Biomol. NMR* **27**, 69-79. (<http://www.bpc.uni-frankfurt.de/guentert/wiki/index.php/GARANT>)
4. Tian et al. (2001) *J. Am. Chem. Soc.* **123**, 11791-11796.

Some Uses of Chemical Shifts

- Predicting secondary structure
 - because $^1\text{H}^\alpha$, $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, and $^{13}\text{C}'$ chemical shifts depend on secondary structure, these shifts can be used to predict secondary structure
 - Chemical Shift Index (“CSI” (1))



- Predicting main chain dihedral angles (ϕ and ψ)
 - “TALOS” (2): uses chemical shift and sequence information and database matching to predict reliable values for ϕ and ψ

1. Wishart et al. (1992) *Biochemistry* **31**, 1647-1651; Wishart and Sykes (1994) *J. Biomol. NMR* **4**, 171-180
2. Cornilescu et al. (1999) *J. Biomol. NMR* **13**, 289-302; Shen et al. (2009), *J. Biomol. NMR* **44**, 213-223.
also <https://spin.niddk.nih.gov/bax/software/TALOS/>; <https://spin.niddk.nih.gov/bax/nmrserver/talos/>