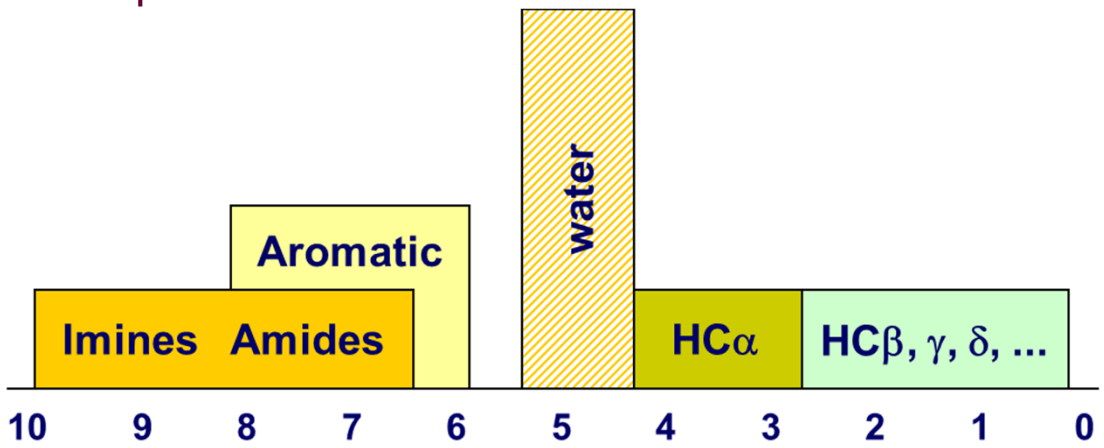
The very basics of NMR of proteins

• The first thing we need to know is were do the peaks of an amino acid residue show up in the ^{1}H spectrum:



• Since they are all very close, after we go pass 3 or 4 amino acids we need to do 2D spectroscopy to spread out the signals enough to resolve them.

• As we said before, there are no connections between different **AAs**: we cannot tell which one is which. One of the requirements in NMR structure determination is knowledge of the primary structure of the peptide chain.

• Now, in order to determine the structure we need to assign an amino acid in the chain to signals in the spectrum. This is the first step in the NMR study.

Spin system assignments.

• To do this we rely on the 1D (if the molecule is small enough), COSY, and TOCSY spectra. We have seen how a whole spin system is easily identified in a TOCSY.

• In peptides, there will be an isolated line for each amino acid starting from the NH that will go all the way down to the side chain protons.

• The only exceptions are **Phe**, **Tyr**, **Trp**, and **His** (and some others I don't remember) in which part of the side chain is separated by a quaternary or carbonyl carbon.

• We can either assign all the spin systems to a particular amino acid (good), or do only part of them due to spectral overlap (bad). If this happens, we may have to go to higher dimensions or fully labeled protein (next class…).

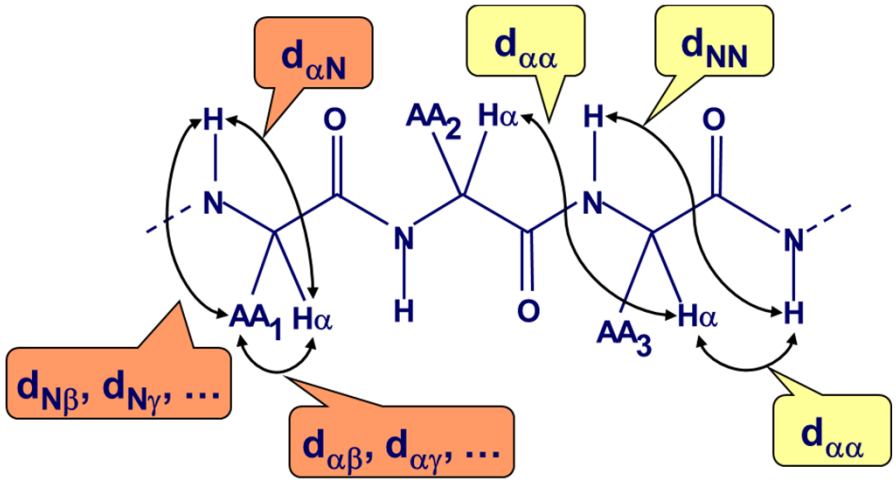
• In any case, once all possible spins systems are identified, we have to tie them together and identify the relative position of the signals in the primary structure.

• There are two ways of doing this. One is the ***sequential assignment approach***, and the other one the ***main-chain directed approach***.

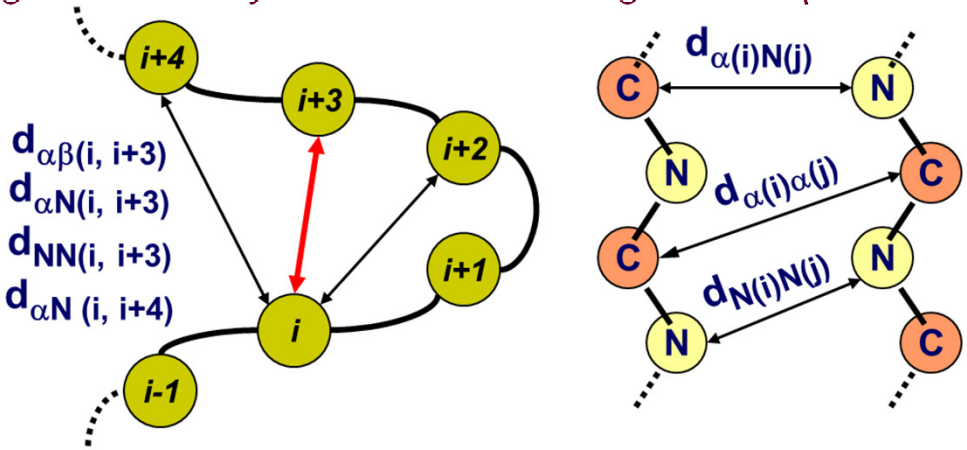
• Both rely on the fact that there will be characteristic NOE cross-peaks for protons of residue ***i*** to ***(i + 1)*** and ***(i - 1)***.

Characteristic NOE patterns.

• The easiest to identify are ***interesidue*** and ***sequential*** NOE, cross-peaks, which are NOEs among protons of the same residue and from a residue to protons of the ***(i + 1)*** and ***(i - 1)*** residues:



• Apart from those, regular secondary structure will have regular NOE patterns. For ***α-helices*** and ***β-sheets*** we have:



Sequential assignment

• In the sequential assignment approach, we try to tie spin systems by using sequential NOE connectivities (those from a residue to residues ***i + 1*** or ***i - 1***).

• The idea is to pick an amino acid whose signals are well resolved in the TOCSY, and then look in the NOESY for sequential NOE correlations from its protons to protons in other spin systems.

• These are usually the $\mathbf{d_{NN}}$, $\mathbf{d_{\alpha N}}$, and $\mathbf{d_{\beta N}}$ correlations. At this point we also look for the $\mathbf{d_{\beta\delta}}$ to establish the identity of aromatic amino acids, **Asn**, **Arg**, **Gln**, etc…

• After we found those, we go back to the TOCSY to identify to which amino acid those correlations belong. These protons will be in either the ***i + 1*** or ***i - 1*** residues.

• We do it until we run out of amino acids (when we get to the end of the peptide chain) or until we bump into a lot of overlapping signals.

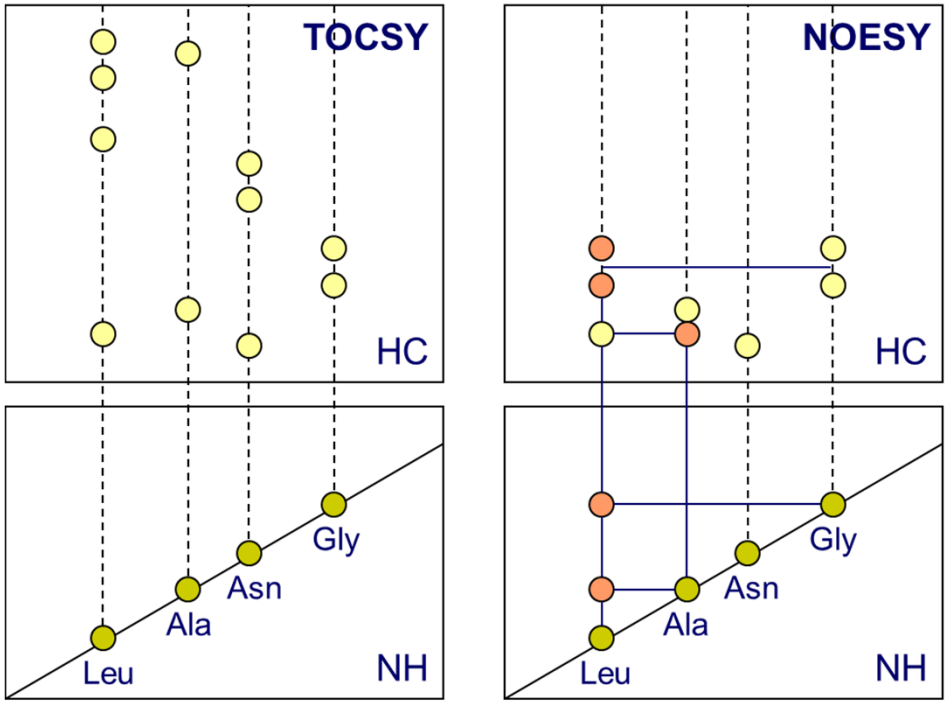
• Since we may have different starting points (and directions), the method has a built-in way of proofing itself automatically.

• Yes, hundreds of folks have some sort of a computerized algorithm that should do this. Their reliability varies, and there is a lot of user intervention involved…

Sequential assignment (continued)

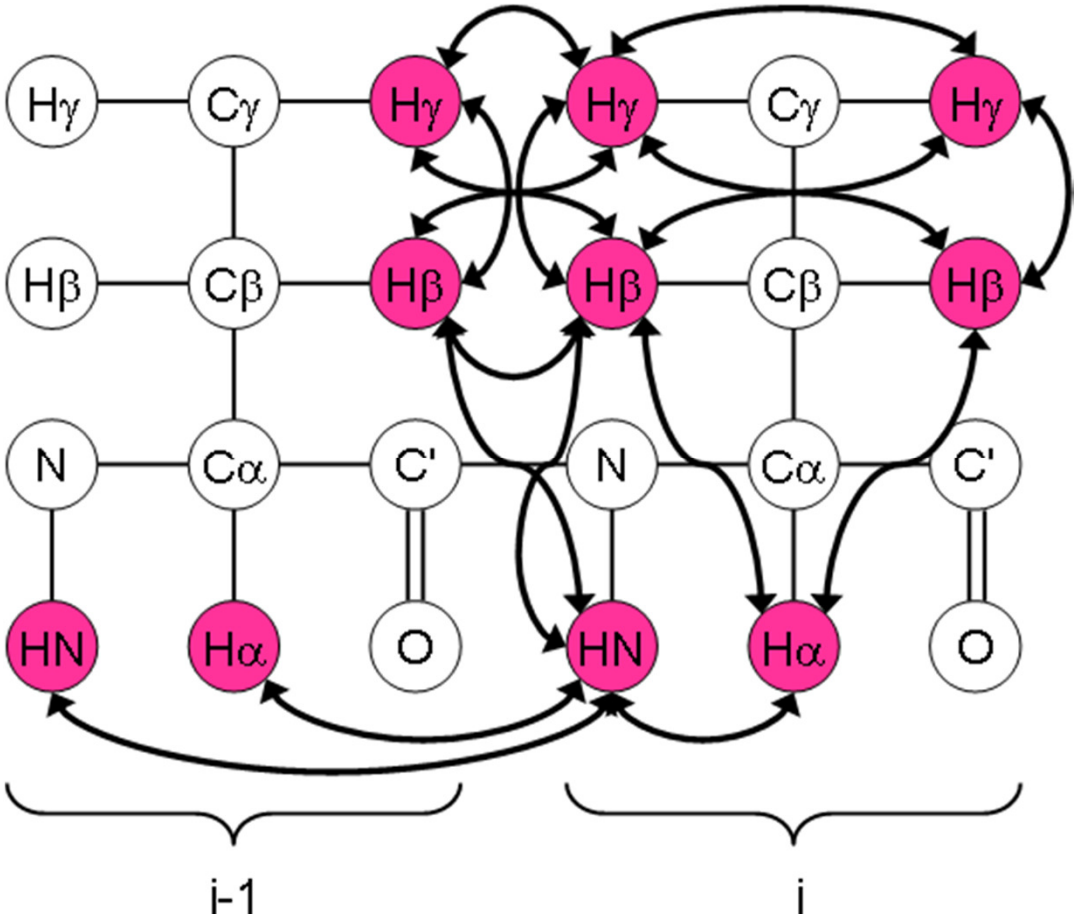
• We can see this with a simple diagram (sorry, could not find much good data among my stuff…).

• Say we are looking at four lines in a TOCSY spectrum that correspond to **Ala**, **Asn** , **Gly** and **Leu**. We also know that we have **Ala-Leu-Gly** in the peptide, but no other combination:

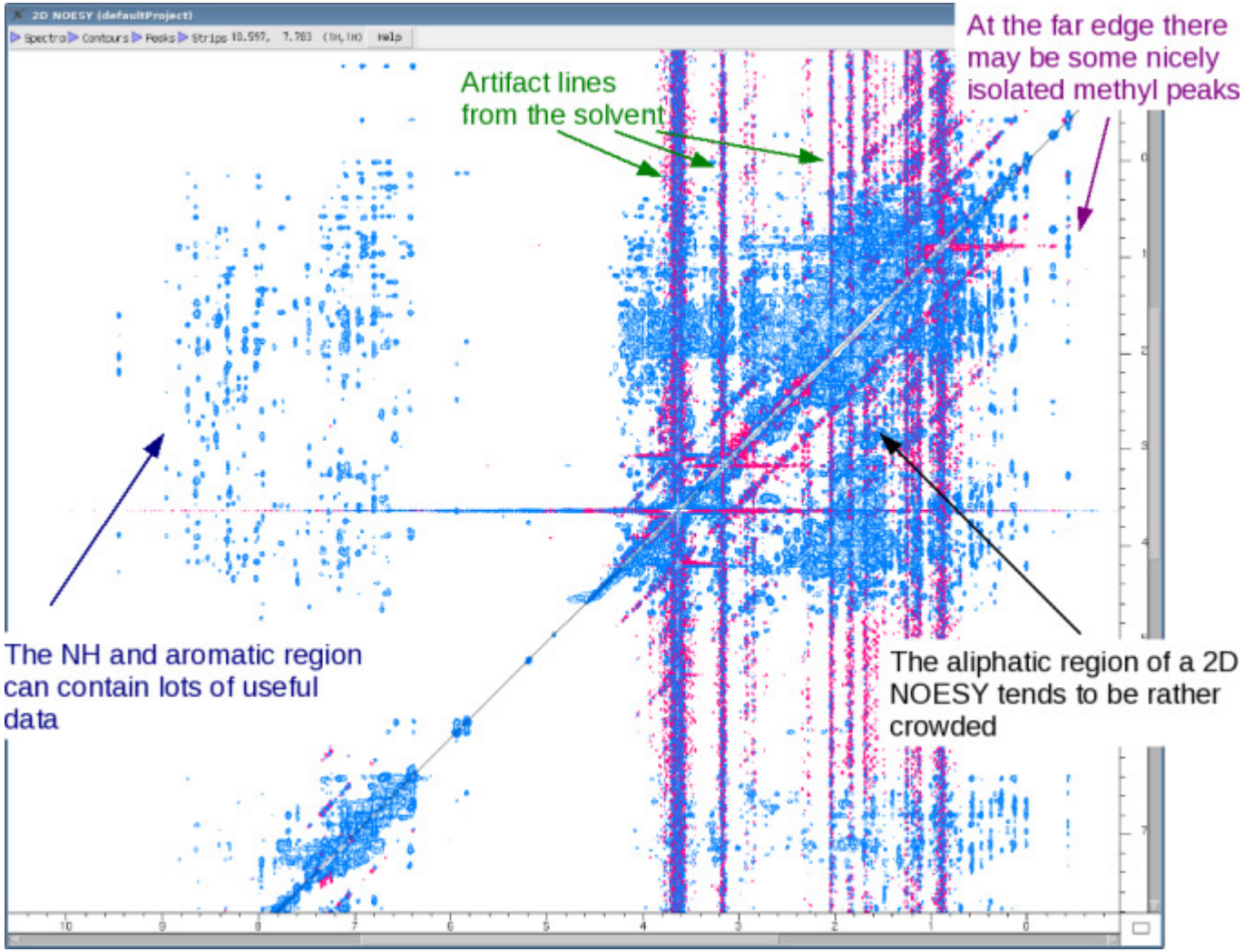


• In the TOCSY we see all the spins. The NOESY will have both intraresidue correlations ( ), as well as interesidue correlations ( ), which allows us to find which residue is next to which in the peptide chain.

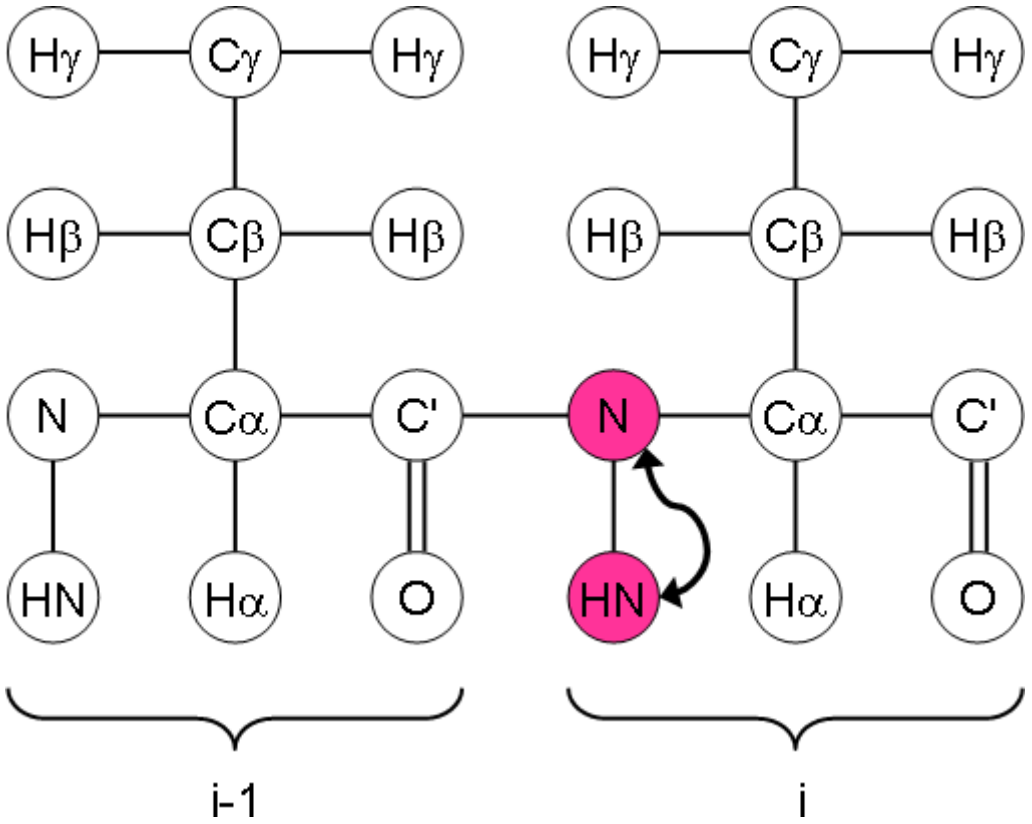
H-H NOESY



This spectrum can be used both to help assignment (especially of aromatic residues) and to get structural restraints. It can get very crowded for large proteins, but can nonetheless provide some useful constraints for methyl and aromatic hydrogens. The quality can be improved by using D_2O as the solvent, though this means that the exchangable NH protons will not be visible.

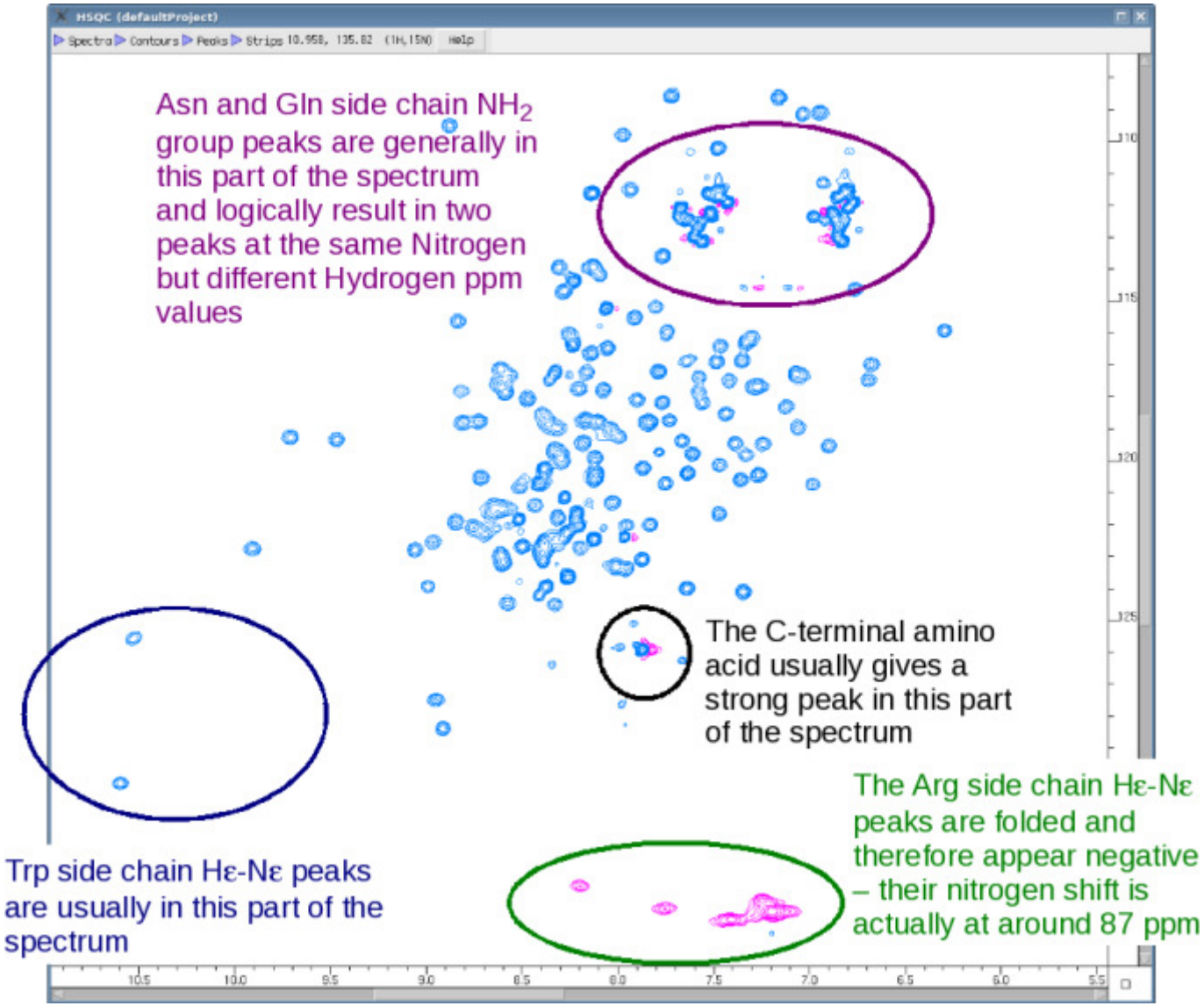


1H-15N HSQC



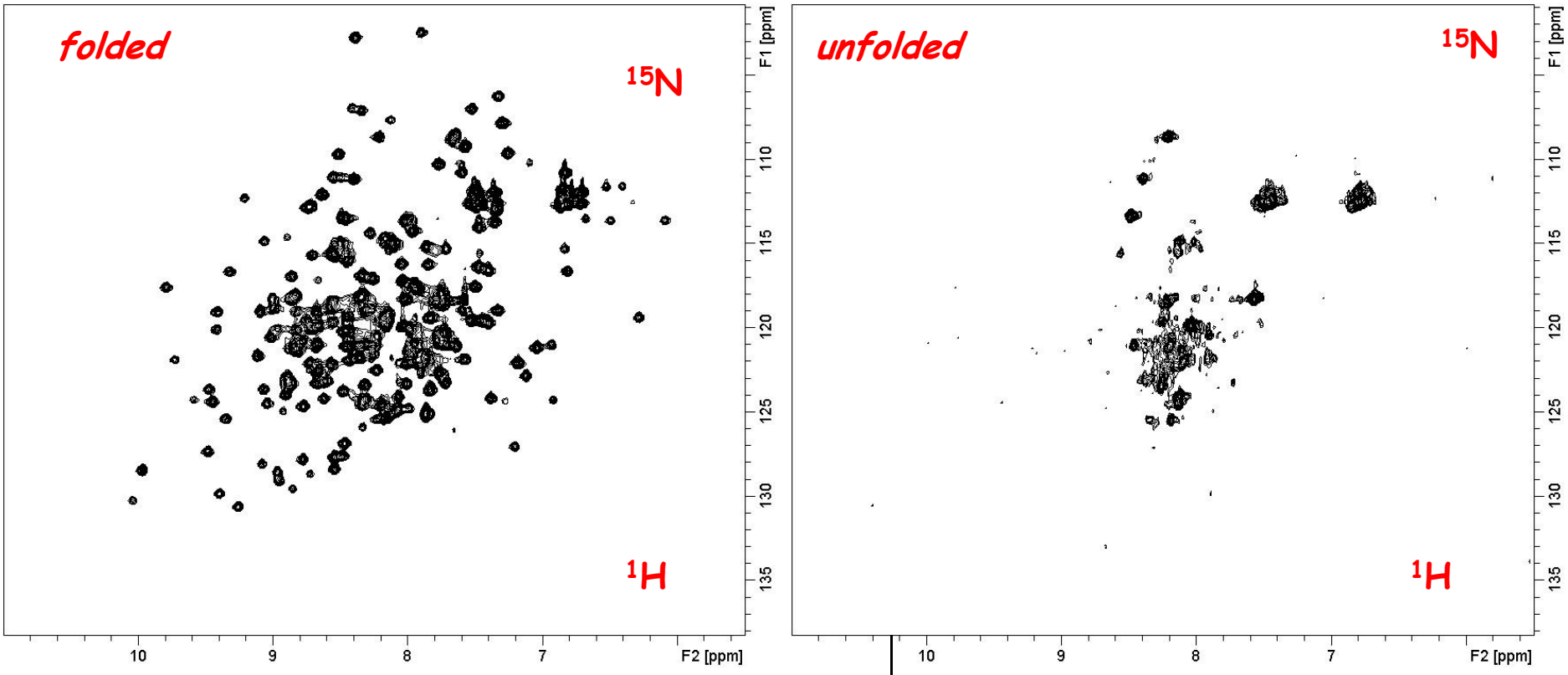
Magnetization is transferred from hydrogen to attached 15N nuclei via the J-coupling. The chemical shift is evolved on the nitrogen and the magnetisation is then transferred back to the hydrogen for detection.

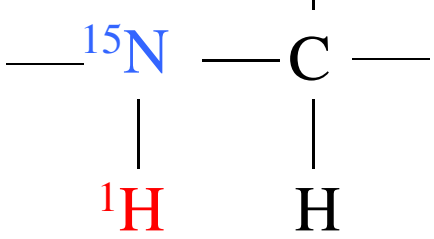
This is the most standard experiment and shows all H-N correlations. Mainly these are the backbone amide groups, but Trp side-chain Nε-Hε groups and Asn/Gln side-chain Nδ-Hδ2/Nε-Hε2 groups are also visible.
The Arg Nε-Hε peaks are in principle also visible, but because the Nε chemical shift is outside the region usually recorded, the peaks are folded/aliased (this essentially means that they appear as negative peaks and the Nε chemical shift has to be specially calculated).



Protein state verification

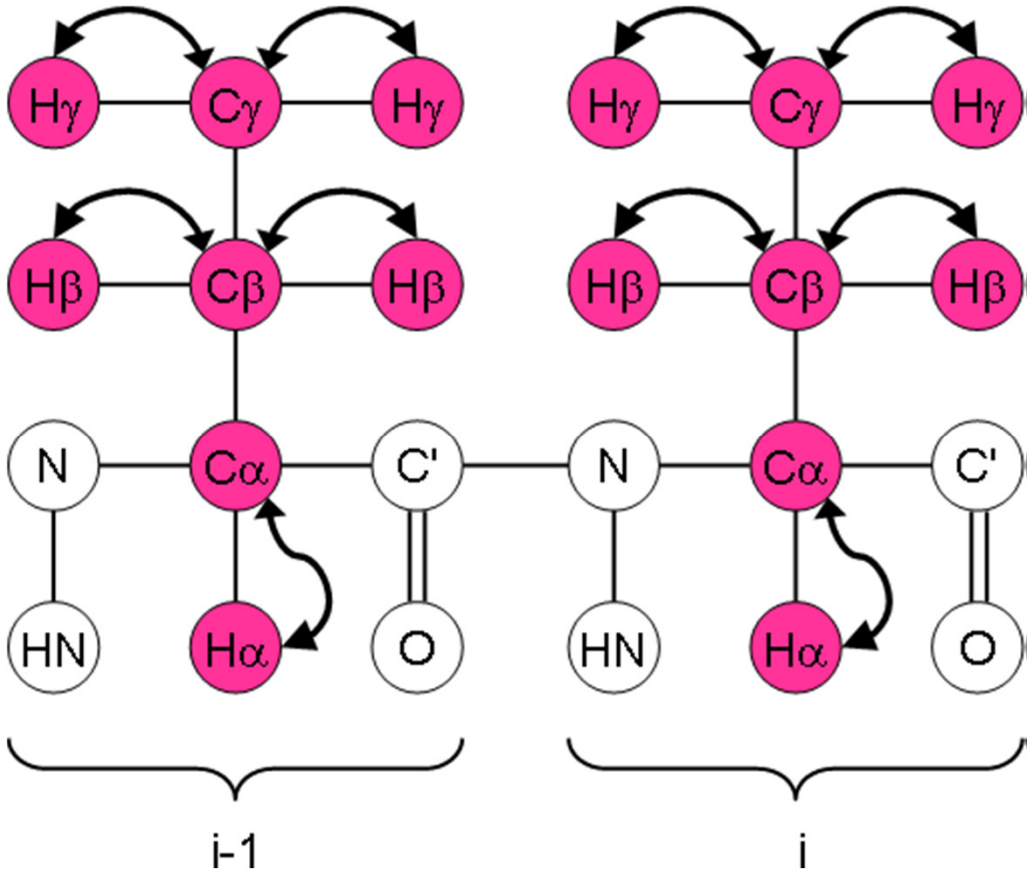
^{1}H-^{15}N HSQC



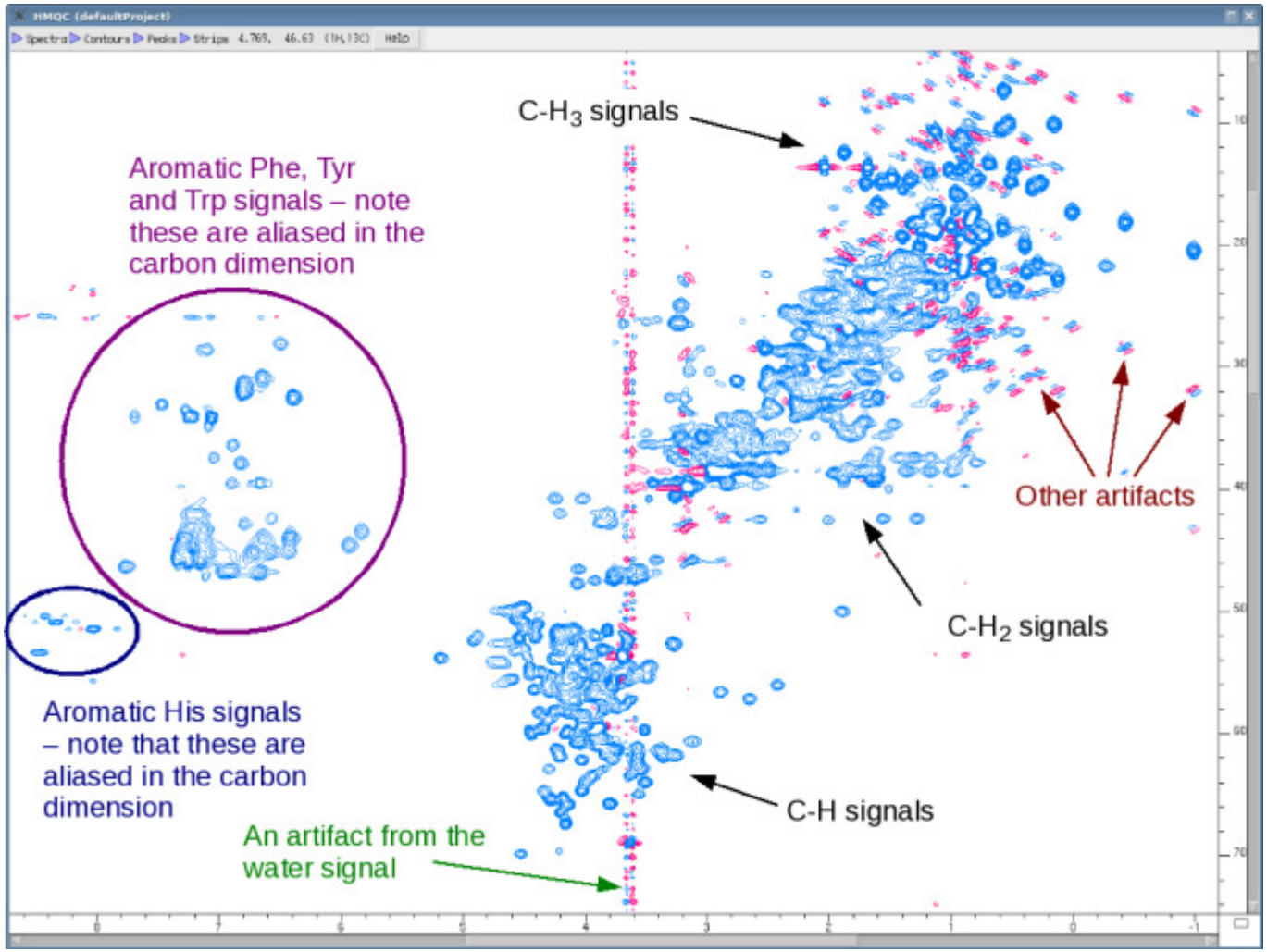


13C-HMQC

The magnetisation is transferred from 1H to 13C and then back again for detection and all 1H-13C moieties, regardless of chemical type, are observed.

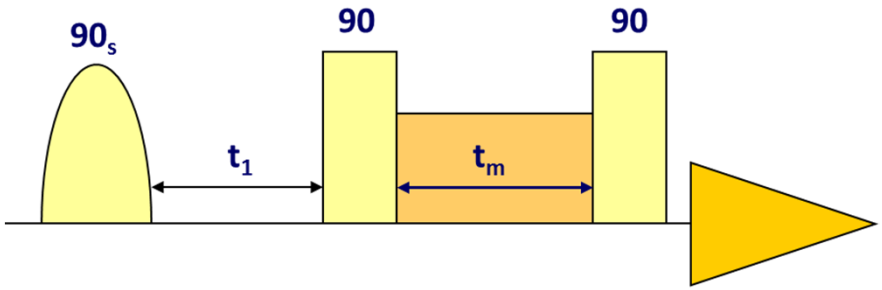


This is the carbon equivalent of the 1H-15N-HSQC. All H-C correlations are shown. It can be useful as a basis for picking a 3D 13C NOESY spectrum. It can also be a useful reference for the aromatic peaks.

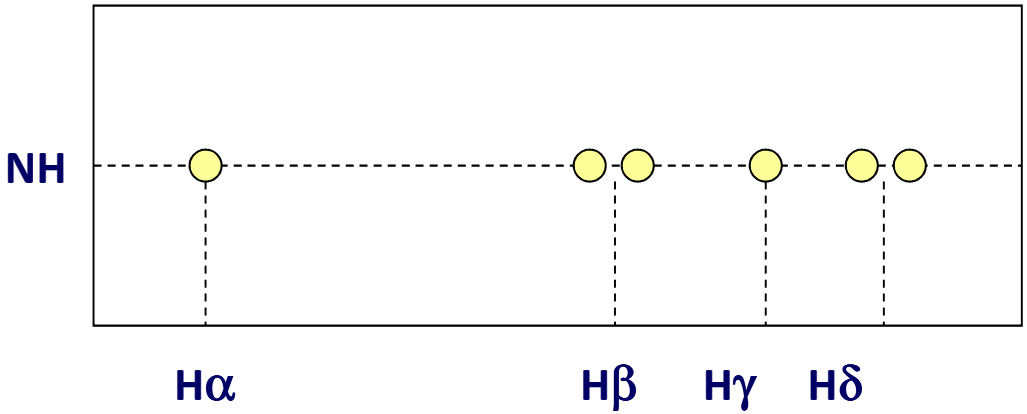


3D NMR spectroscopy (…)

• Say that we could somehow selectively tickle only certain amide protons in the sample (we'll see more on selective pulses today, but this is only an example). Only protons attached to this amide proton will give us cross-peaks:



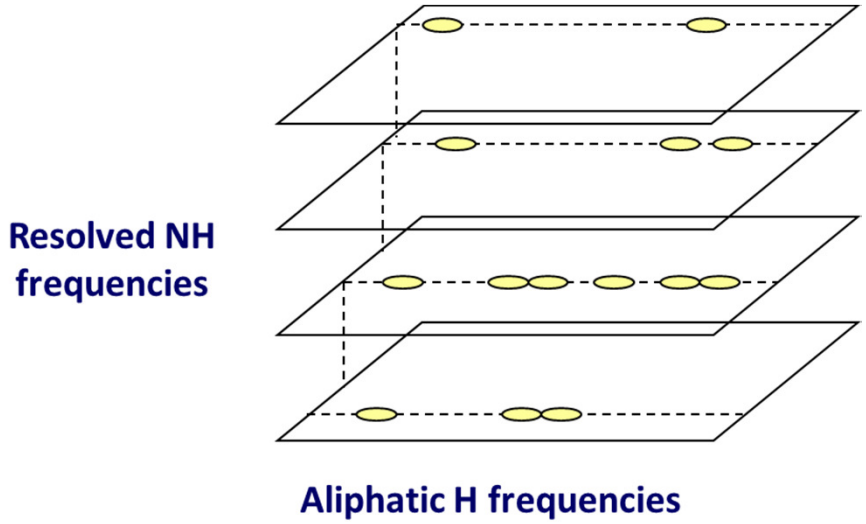
• So, we do this selective amide excitation followed by a 2D TOCSY experiment. Our 2D plot will only have the line that corresponds to the amide proton we selected. For a **Leu**:



• If we changed the frequency of the selective pulse to another NH we would get another spin system and so forth.

3D NMR spectroscopy (…)

• Now, we could put all the 2D experiments stacked like if they were posters in a rack, and each slice would have the connectivities of a particular spin system:



• This would be a ‘pseudo’ 3D experiment. The problem here is the way we do the selection of the NHs. Usually, we isolate each NH (or whatever we want to isolate) by doing a 2D experiment that resolves it.

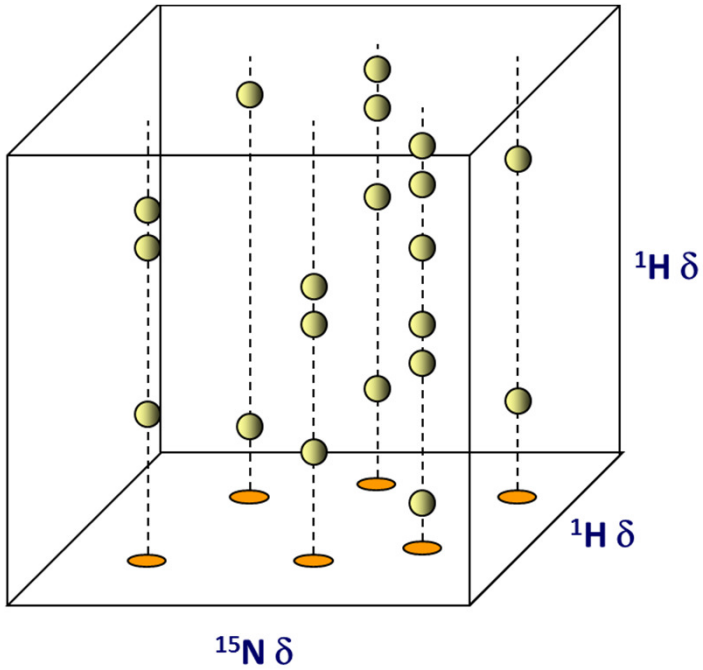
• From what we’ve seen, an NH correlation would be good for this purpose, because most of the cross-peaks are well resolved.

• Additionally, in the way we do 3Ds we don’t usually collect all the 2Ds into a 3D, but get the 3D which has all the cross- peaks and then analyze the slices at different frequencies.

3D NMR spectroscopy (…)

• Furthermore, peaks cross-peaks appearing in the cube arise due to a transfer of polarization between the nuclei that we look at in the 3 dimensions.

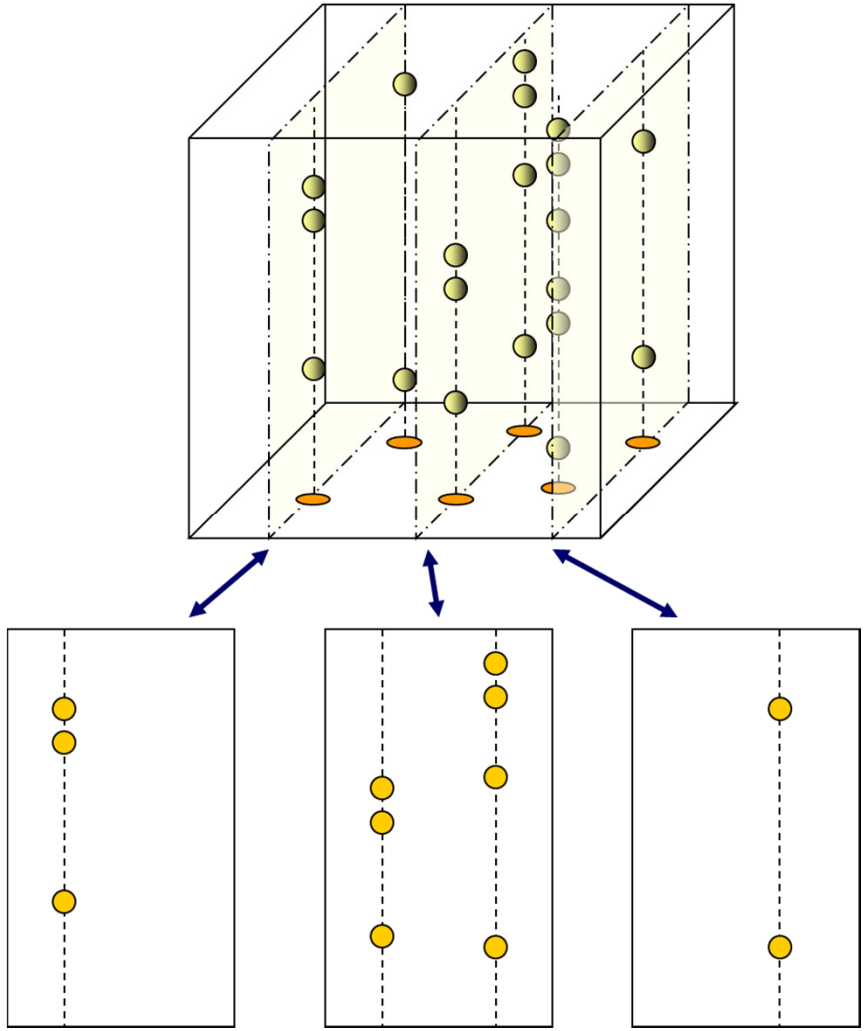
• A 3D using a ^{15}N-^{1}H correlation and TOCSY combination will look like this (hope you like it - it took me forever…):



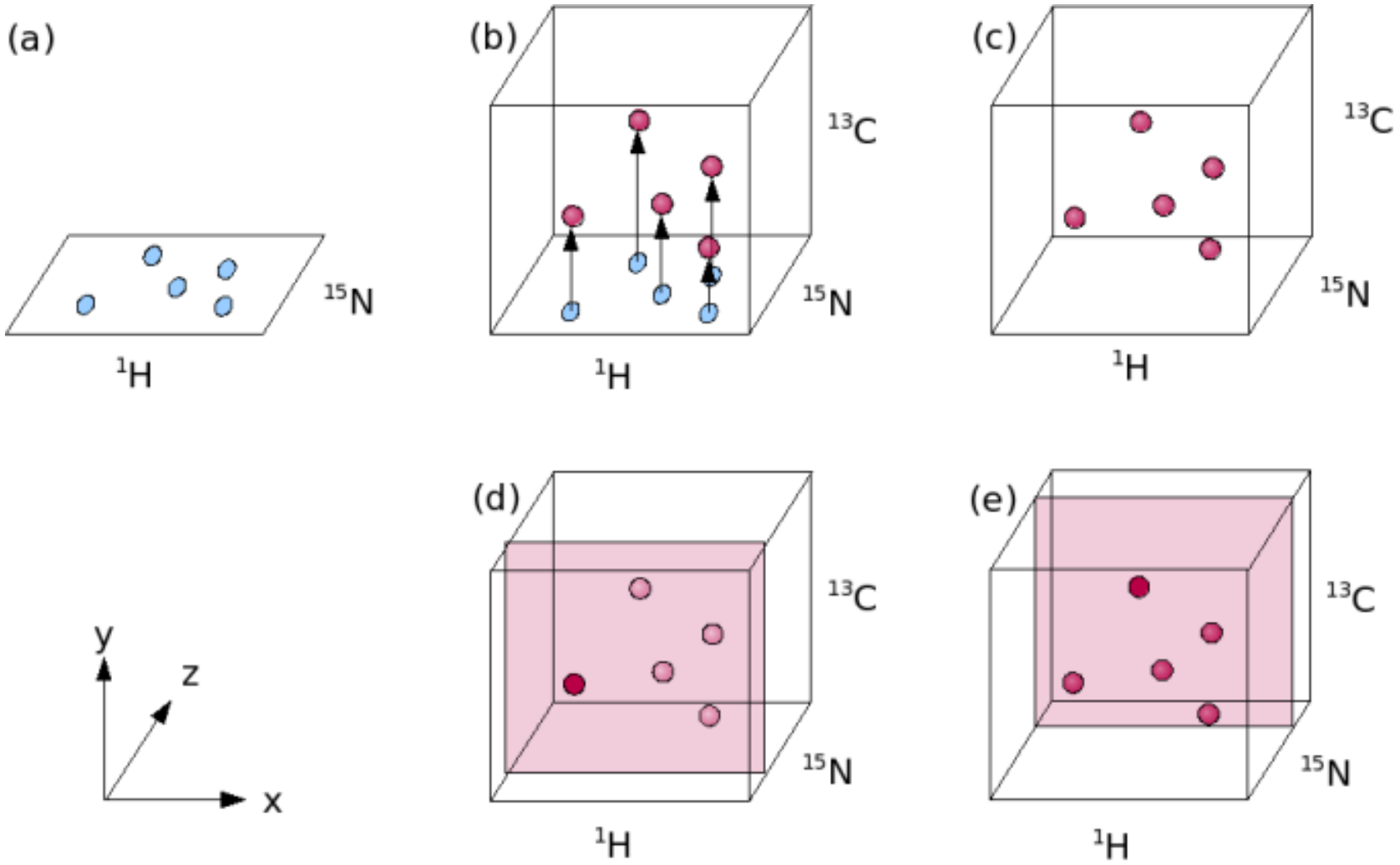
• We have each of the individual (hopefully) TOCSY cross- peaks in a single line, which starts in the ^{15}N-^{1}H cross- peak. The amide protons are separated by the ^{15}N δs.
• Looking at the cube is kind of hard when you have 200 amides. We usually take slices at different 15N frequencies.

3D NMR spectroscopy (…)

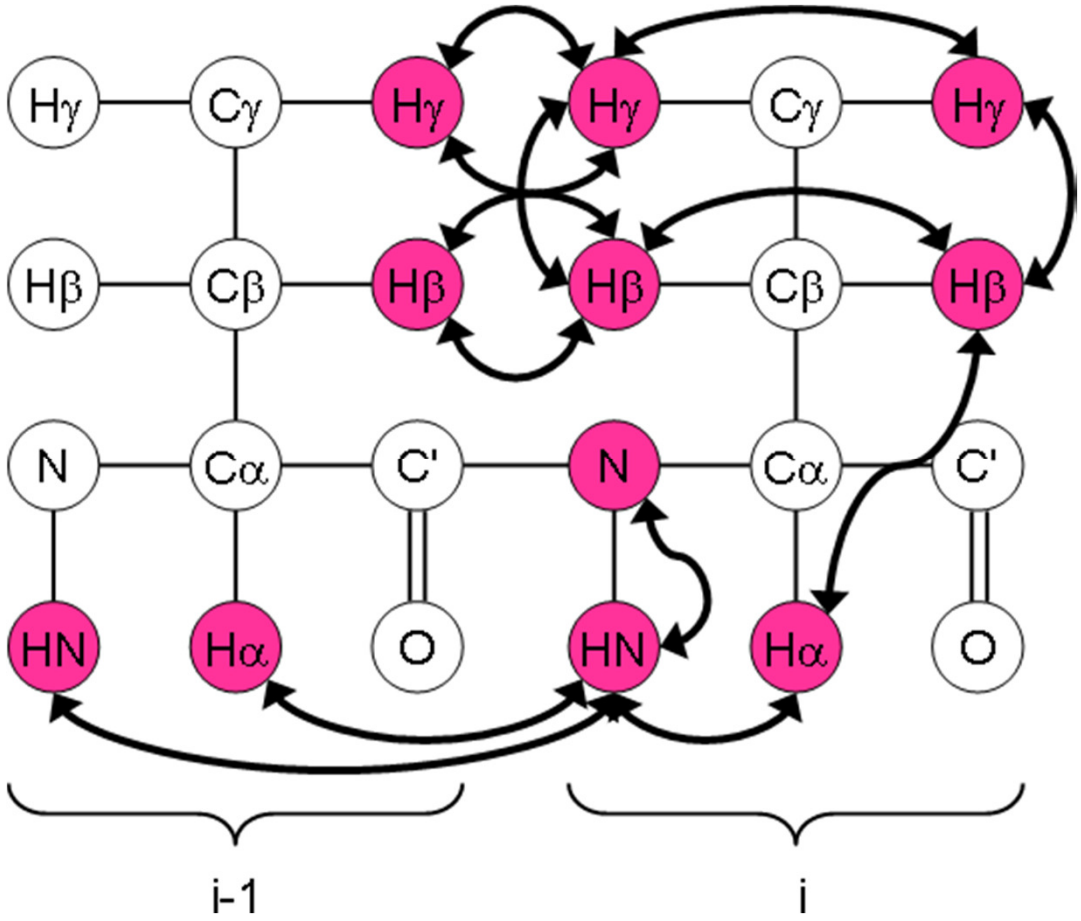
- Depending on the slice (plane) we chose, we'll have TOCSY spectra corresponding to different NHs:



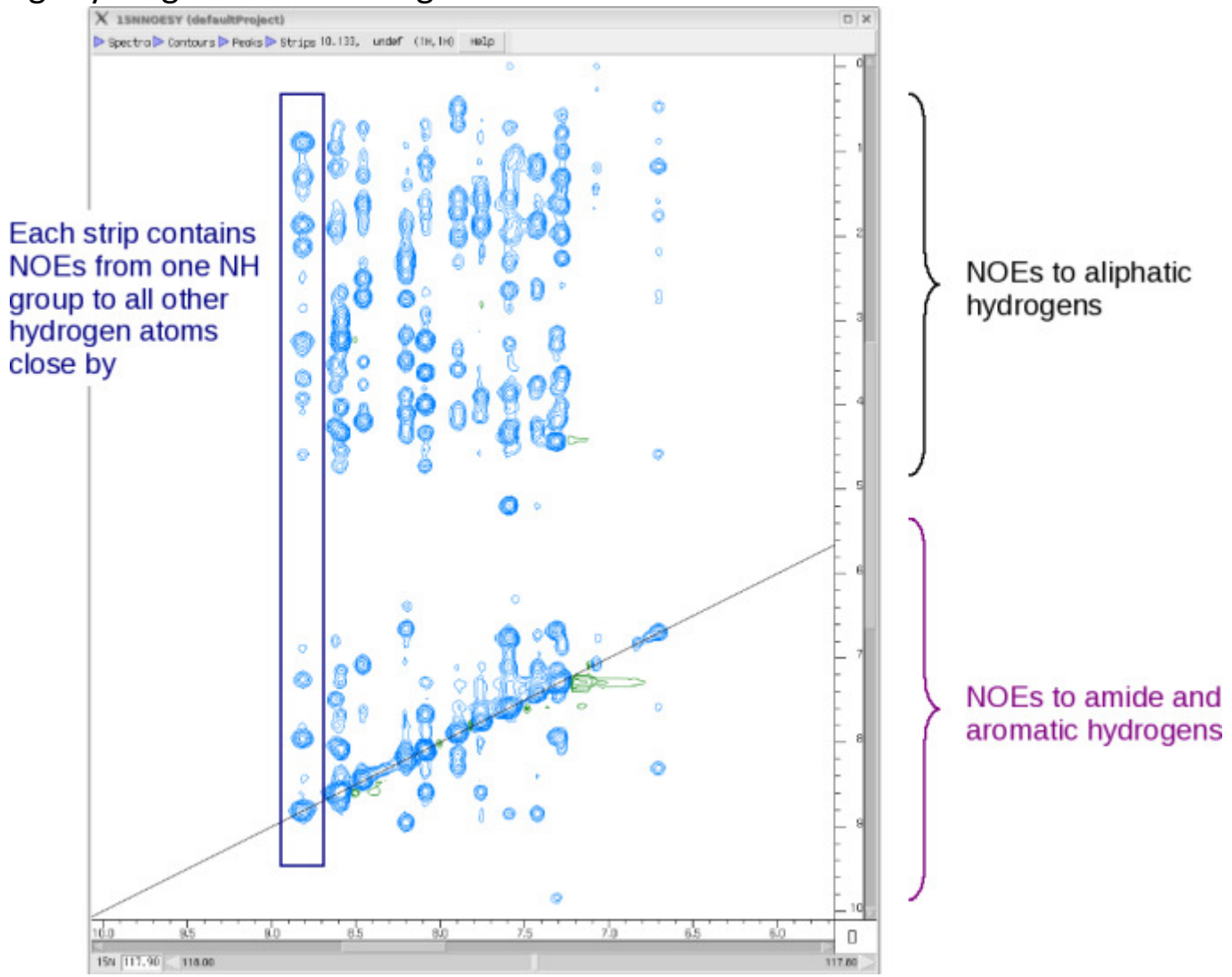
Visualizing 3D spectra



15N-NOESY-HSQC

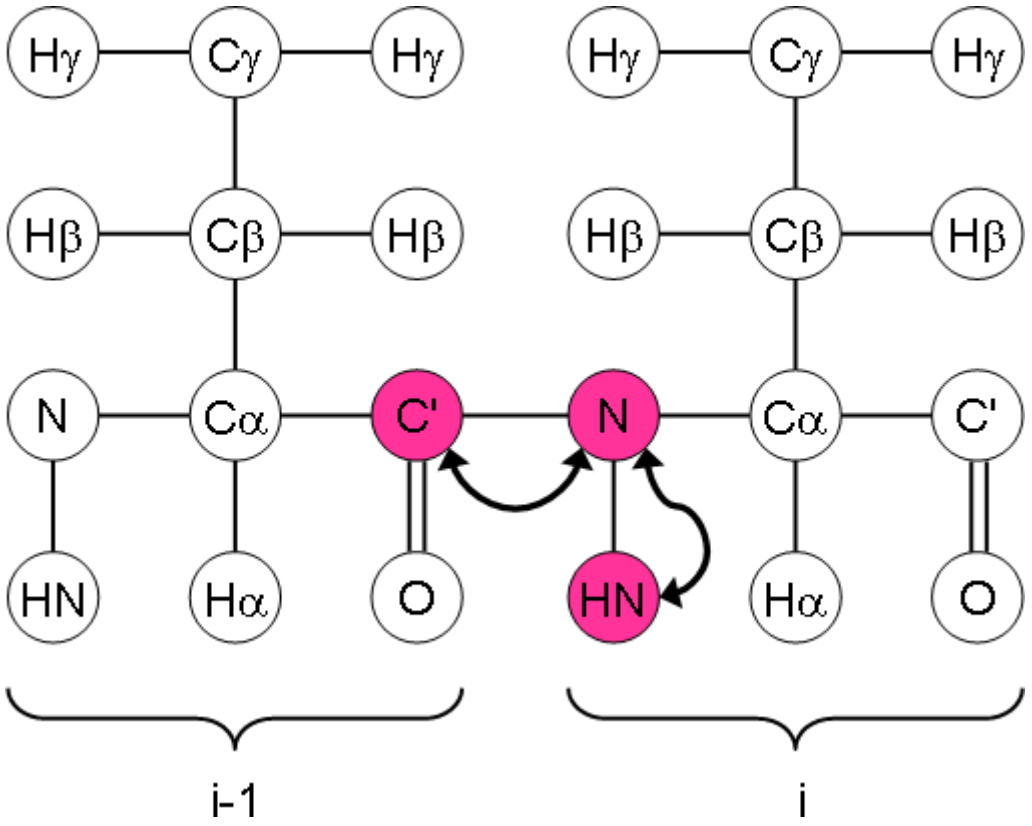


This spectrum can be used to obtain restraints for structure calculations. In this case the NOESY mixing time should probably be around 80ms. It can also be used to help assignment, and for small to medium-sized proteins, assignment using this and 15N-TOCSY-HSQC only is possible. In this case it may be useful to use a slightly longer NOESY mixing time.

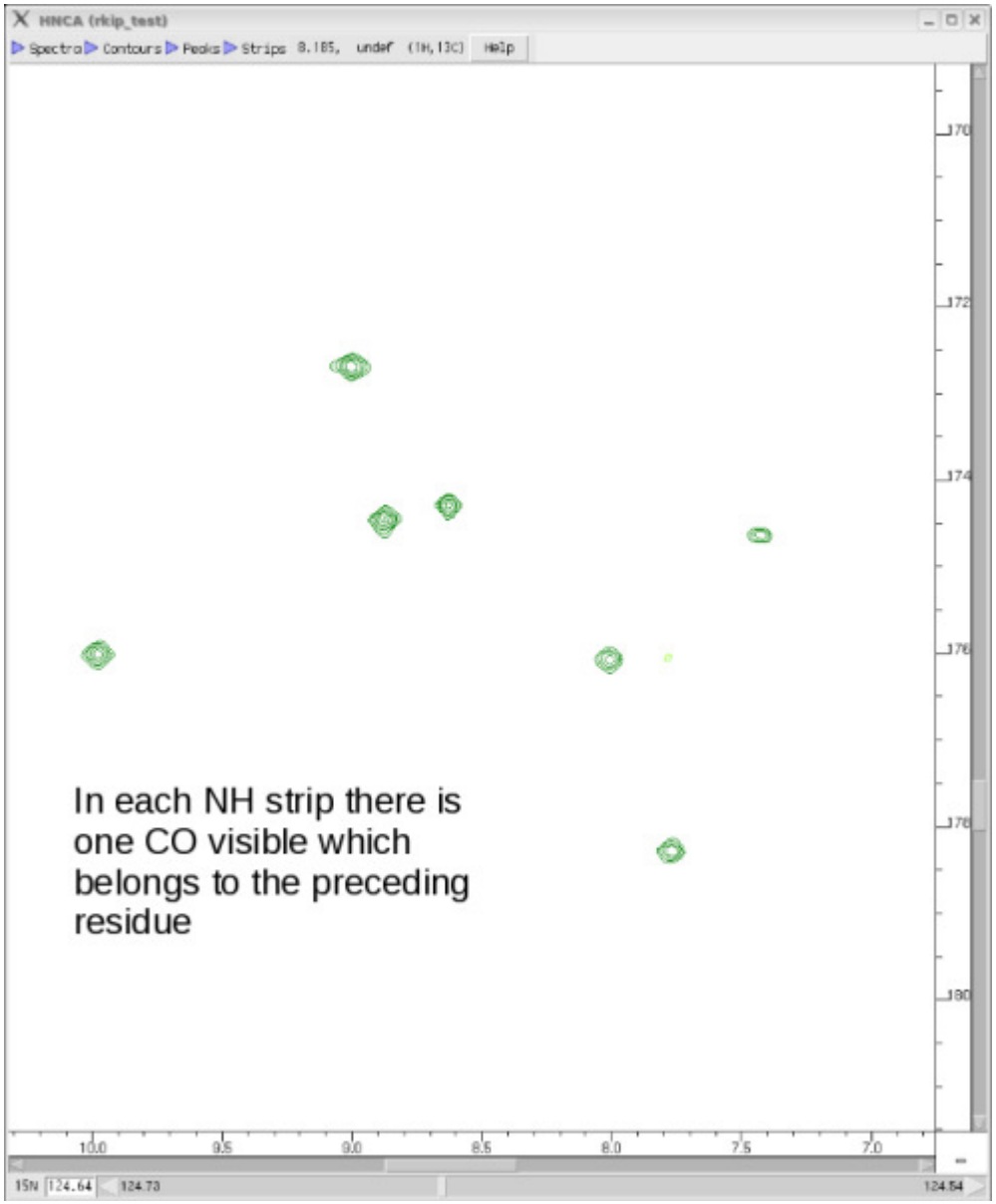


HNCO

Magnetisation is passed from 1H to 15N and then selectively to the carbonyl 13C via the 15NH-13CO J-coupling. Magnetisation is then passed back via 15N to 1H for detection. The chemical shift is evolved on all three nuclei resulting in a three-dimensional spectrum.

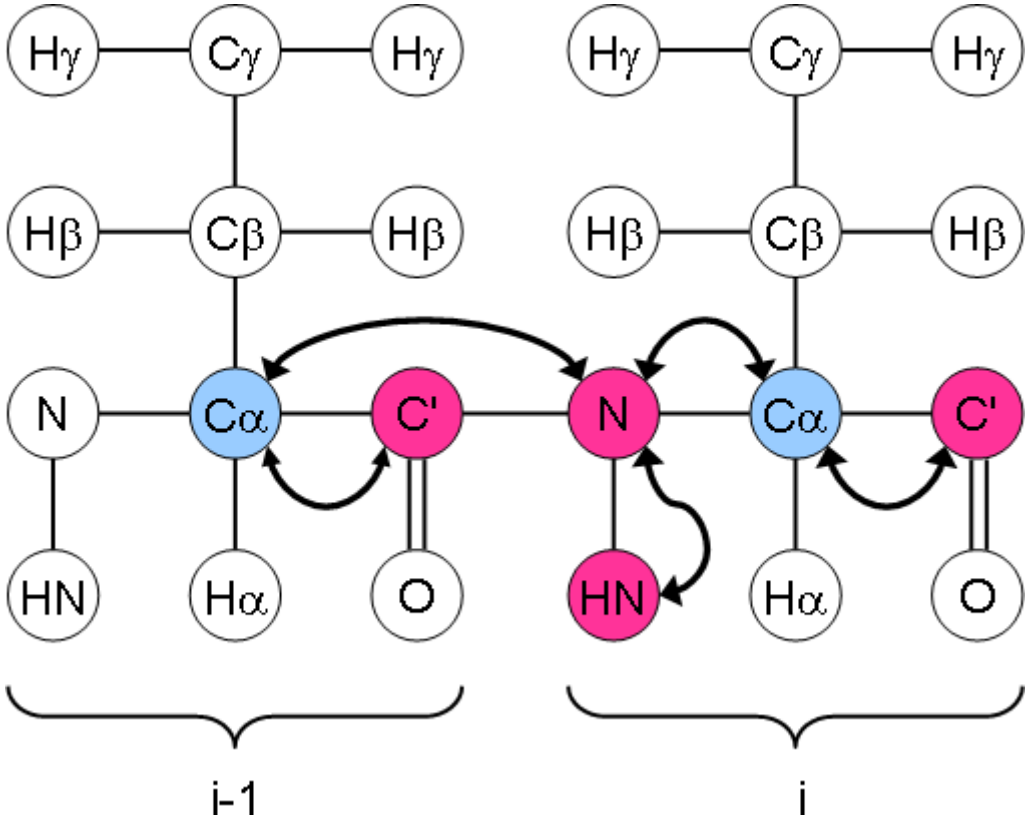


This is the most sensitive triple-resonance experiment. In addition to the backbone CO-N-HN correlations, Asn and Gln side-chain correlations are also visible.

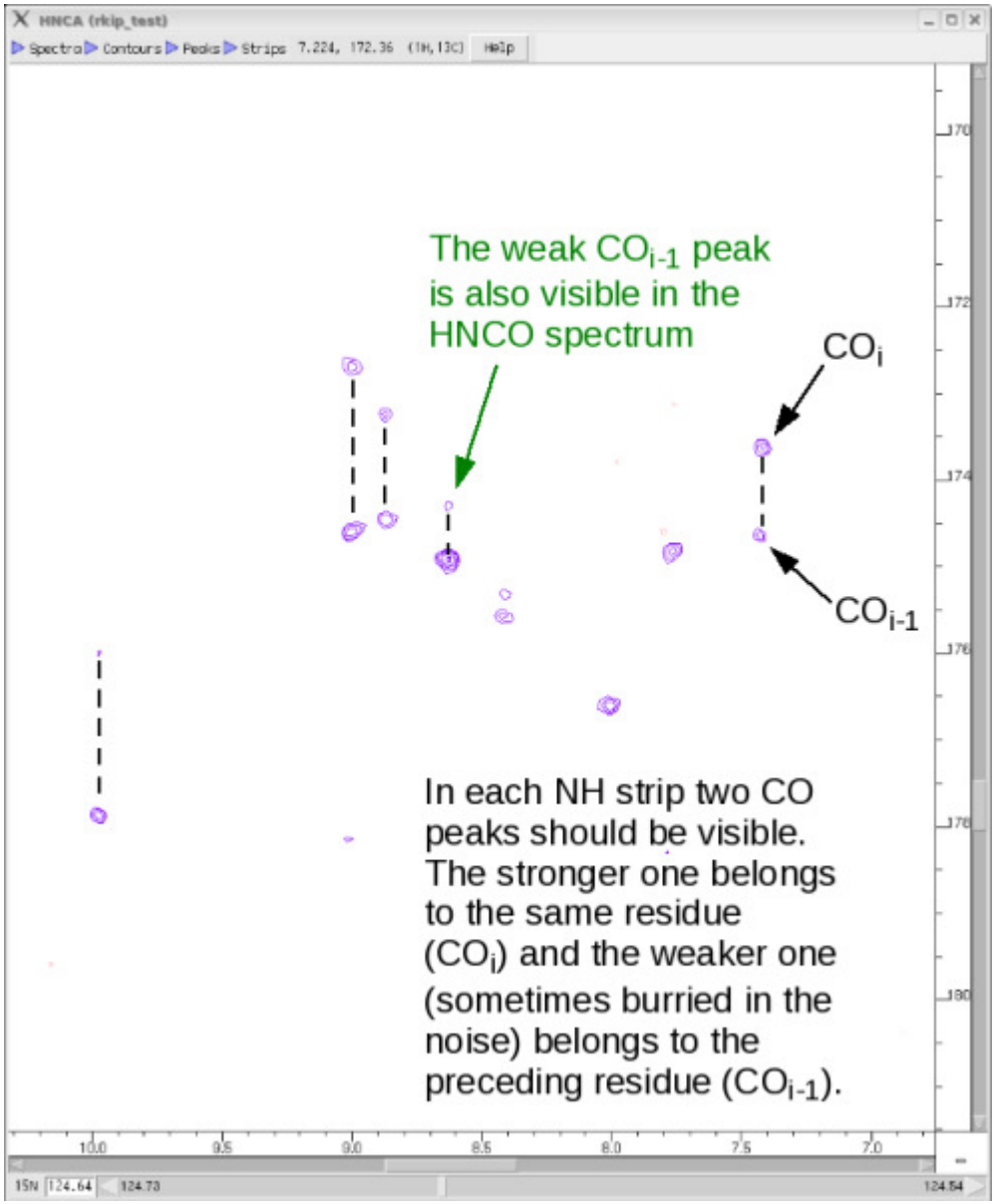


HN(CA)CO

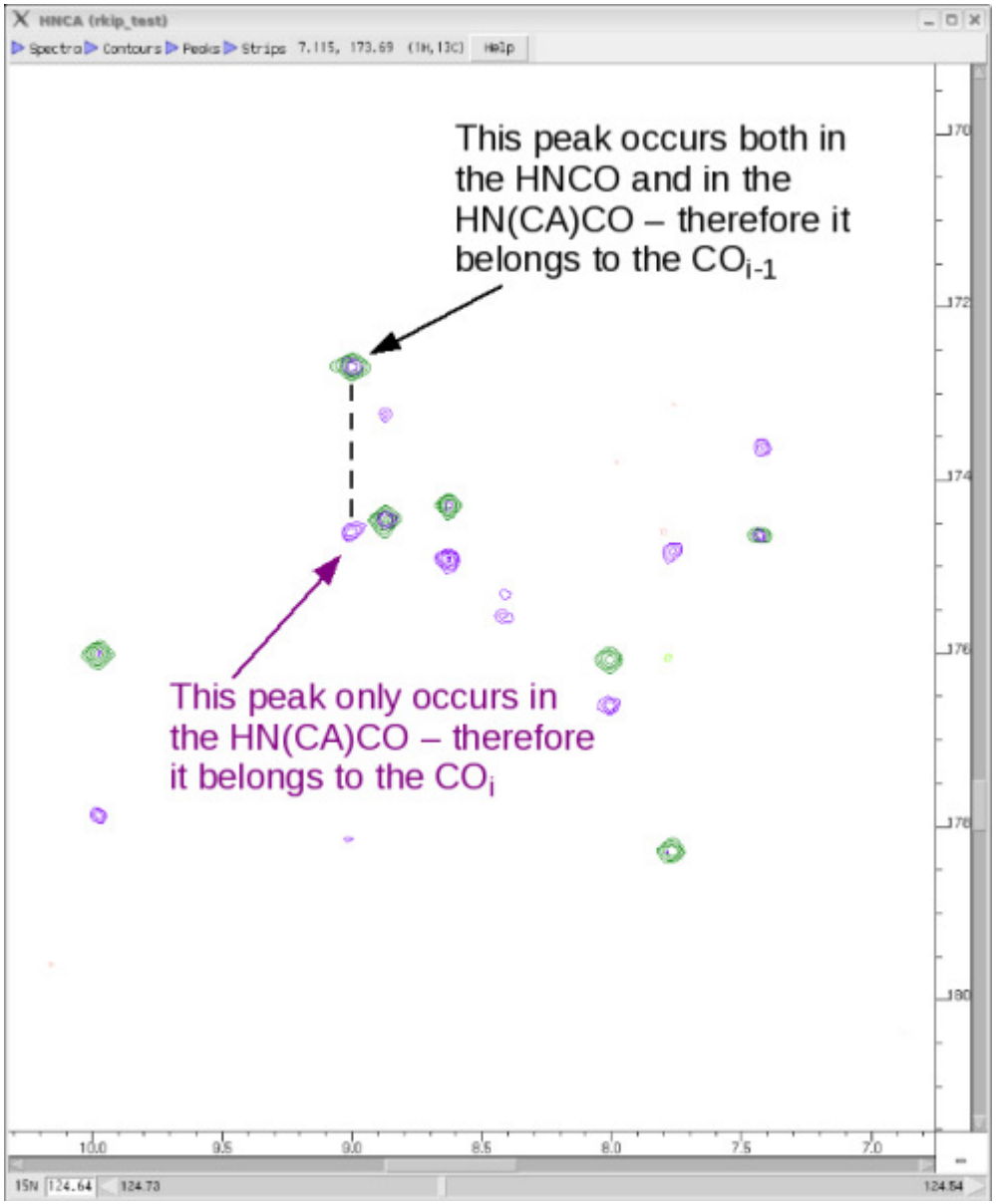
The Magnetisation is transferred from 1H to 15N and then via the N-Cα J-coupling to the 13Cα. From there it is transferred to the 13CO via the 13Cα-13CO J-coupling. For detection the magnetisation is transferred back the same way: from 13CO to 13Cα, 15N and finally 1H. The chemical shift is only evolved on 1H, 15N and 13CO and not on the 13Cα. The result is a three-dimensional spectrum. Because the amide nitrogen is coupled both to the Cα of its own residue and that of the preceding residue, both these transfers occur and transfer to both 13CO nuclei occurs. Thus for each NH group, two carbonyl groups are observed in the spectrum. But because the coupling between Ni and Cαi is stronger than that between Ni and Cαi-1, the Hi-Ni-COi peak generally ends up being more intense than the Hi-Ni-COi-1 peak.



This experiment can be useful for backbone assignment when used in conjunction with the HNCA, HN(CO)CA and HNCO if the CBCANNH and CBCA(CO)NNH spectra are of bad quality.

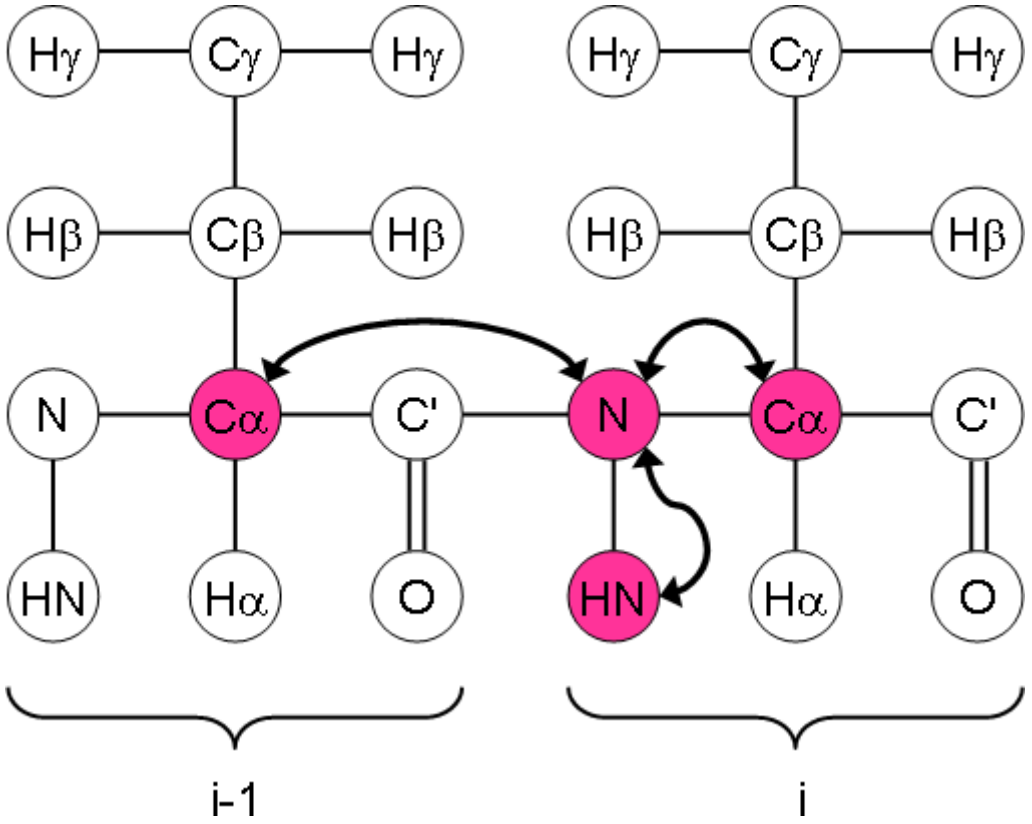


An overlay of the HNCO and HN(CA)CO spectra makes it very easy to distinguish between CO_i and CO_{i-1} for each NH group.

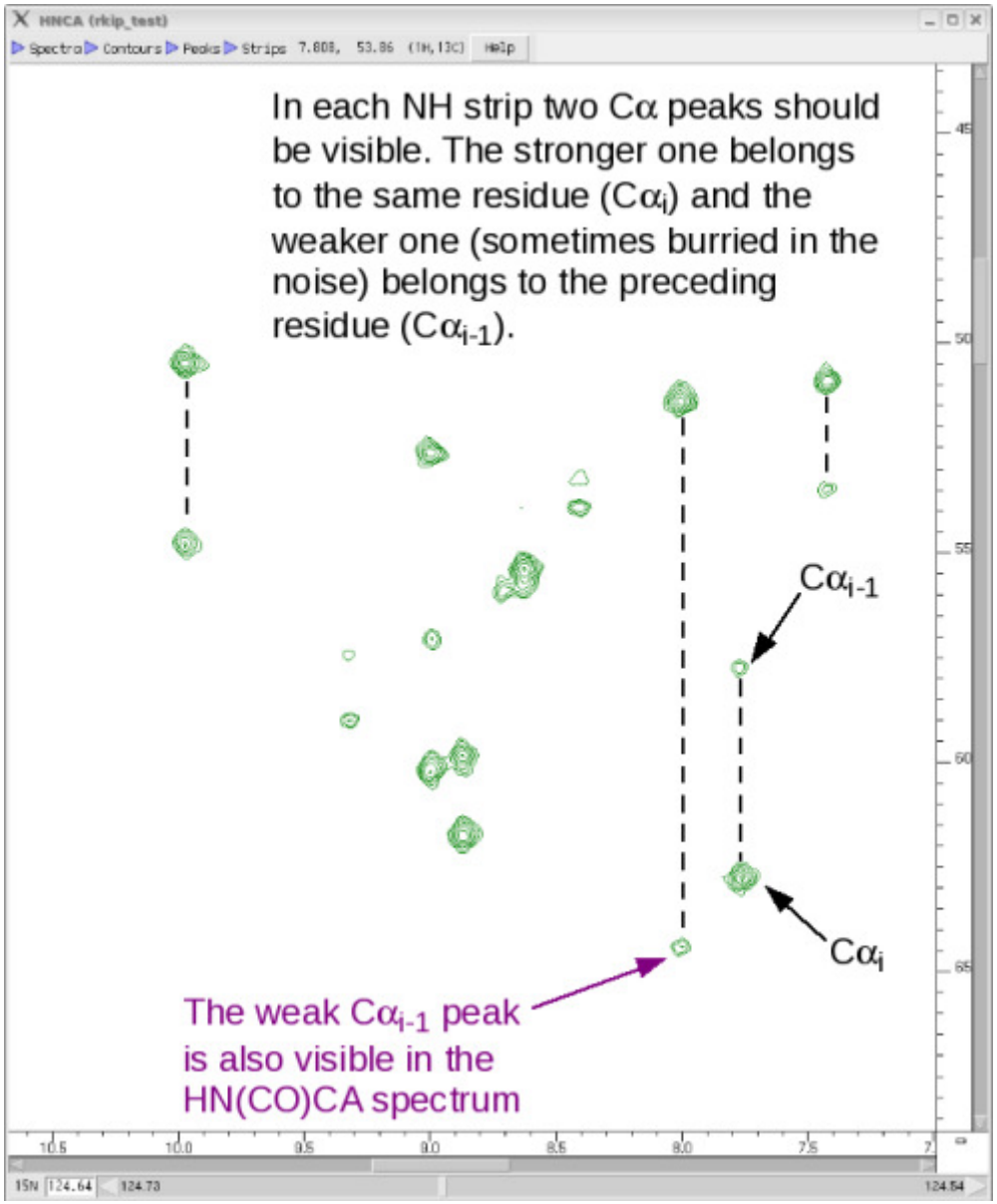


HNCA

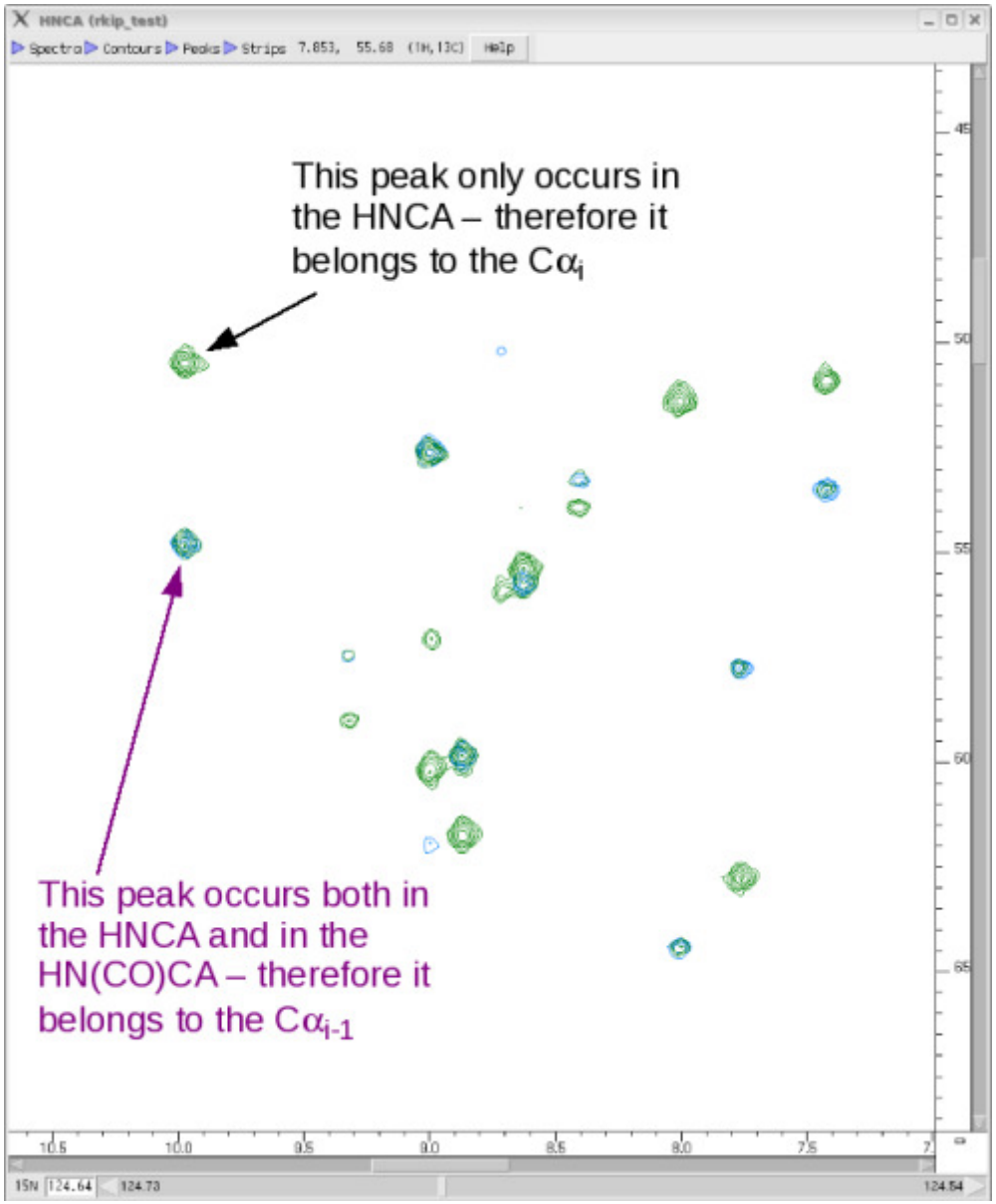
Here the magnetisation is passed from ^{1}H to ^{15}N and then via the N-Cα J-coupling to the $^{13}C\alpha$ and then back again to ^{15}N and ^{1}H hydrogen for detection. The chemical shift is evolved for $^{1}H^{N}$ as well as the $^{15}N^{H}$ and $^{13}C\alpha$, resulting in a 3-dimensional spectrum. Since the amide nitrogen is coupled both to the Cα of its own residue and that of the preceding residue, both these transfers occur and peaks for both Cαs are visible in the spectrum. However, the coupling to the directly bonded Cα is stronger and thus these peaks will appear with greater intensity in the spectra.



This experiment can be useful for backbone assignment when used in conjunction with the HN(CO)CA, HNCO and HN(CA)CO if the CBCANNH and CBCA(CO)NNH spectra are of bad quality.

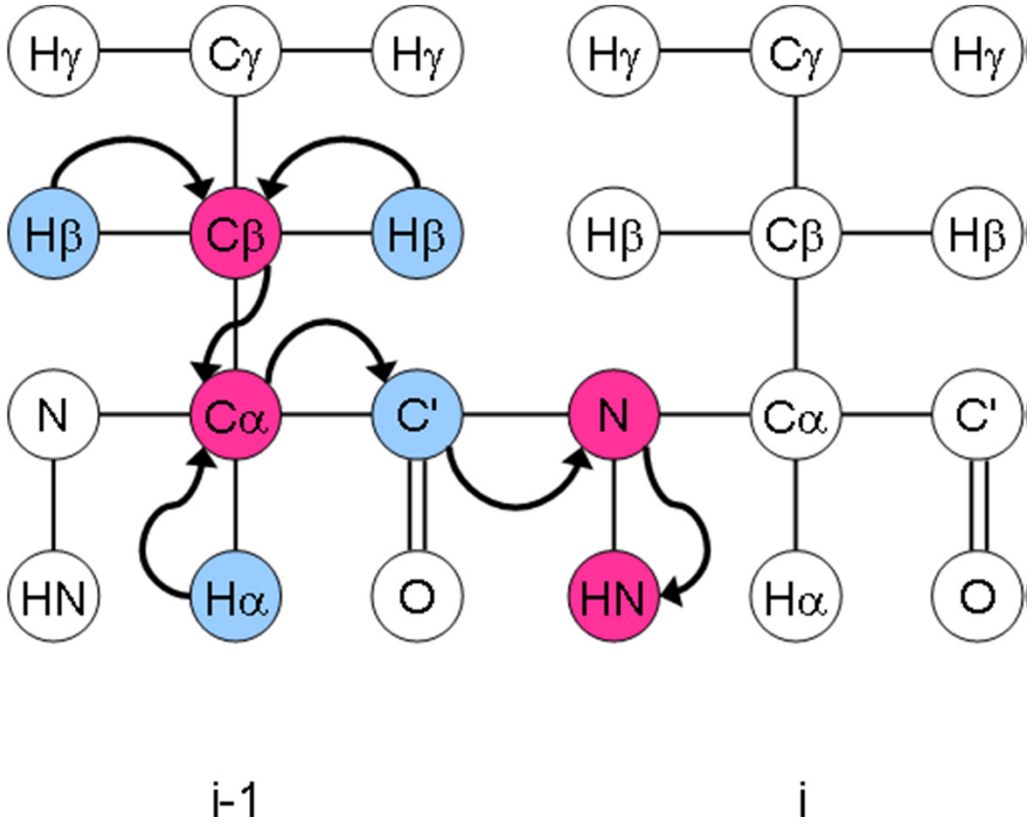


By overlaying the HN(CO)CA spectrum with the HNCA, it becomes even easier to identify and distinguish between all $C\alpha_i$ and $C\alpha_{i-1}$ peaks.

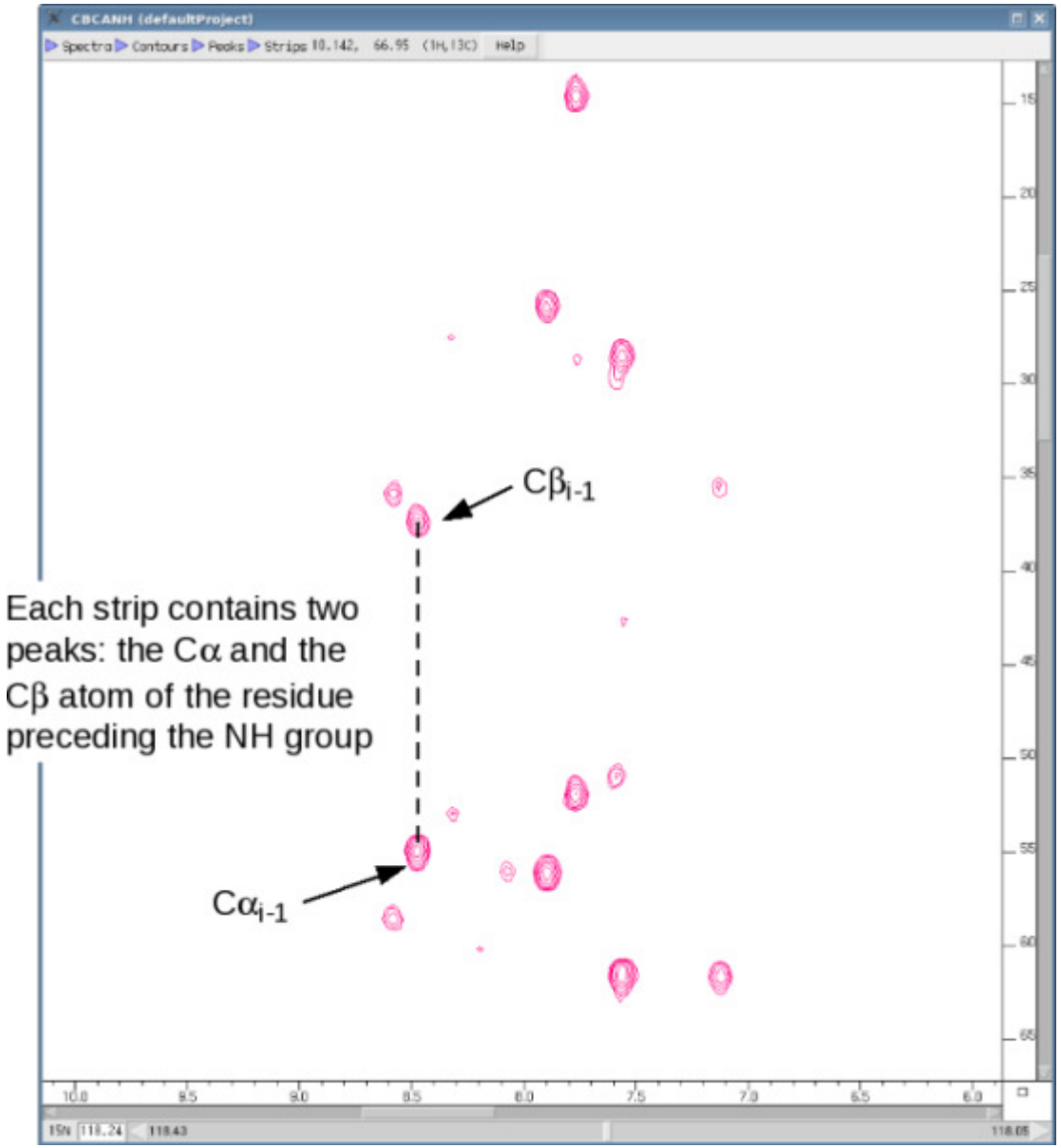


CBCA(CO)NH / HN(CO)CACB

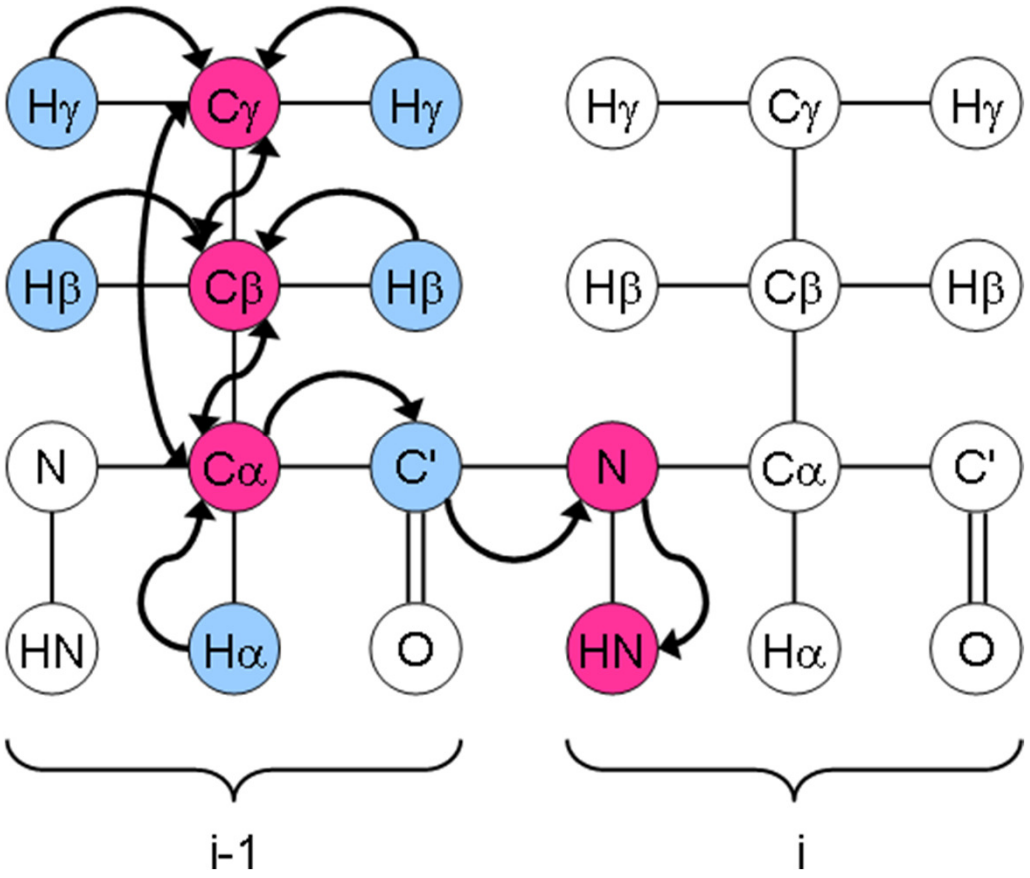
Magnetisation is transferred from $^{1}H\alpha$ and $^{1}H\beta$ to $^{13}C\alpha$ and $^{13}C\beta$, respectively, and then from $^{13}C\beta$ to $^{13}C\alpha$. From here it is transferred first to^{13}CO, then to $^{15}N^{H}$ and then to $^{1}H^{N}$ for detection. The chemical shift is evolved simultaneously on $^{13}C\alpha$ and $^{13}C\beta$, so these appear in one dimension. The chemical shifts evolved in the other two dimensions are $^{15}N^{H}$ and $^{1}H^{N}$. The chemical shift is not evolved on ^{13}CO.



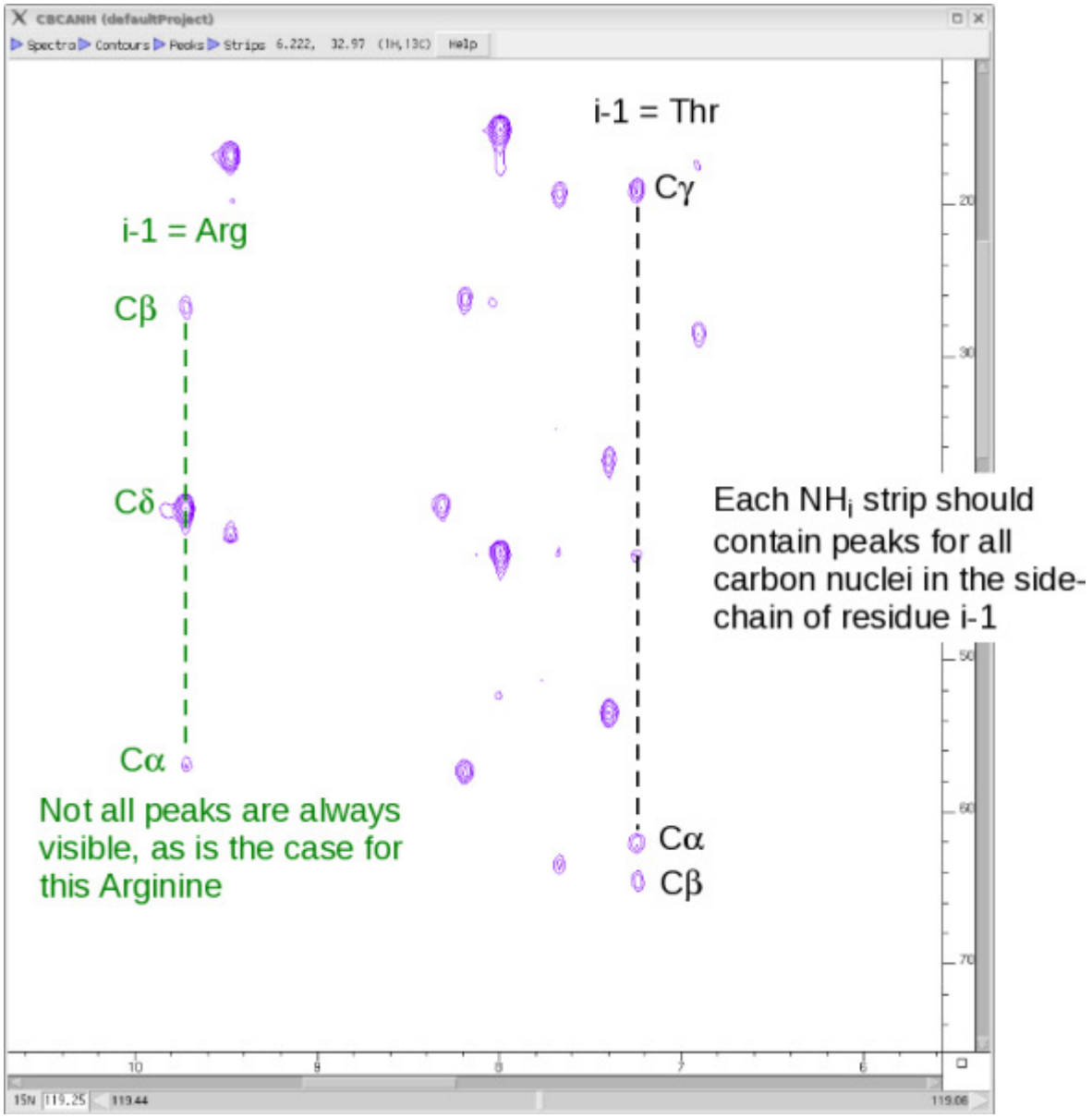
CBCA(CO)NH / HN(CO)CACB



CC(CO)NH

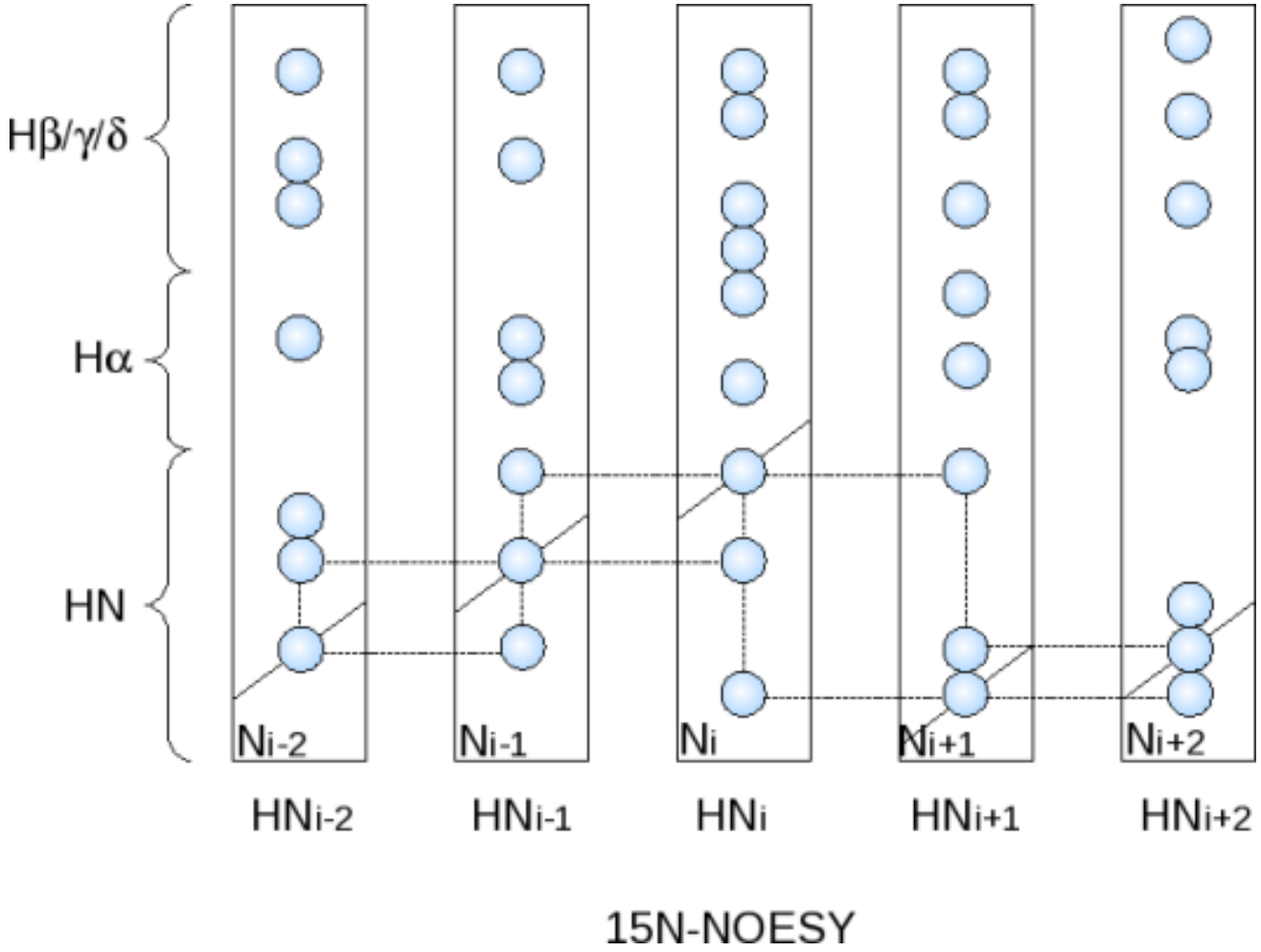


CC(CO)NH

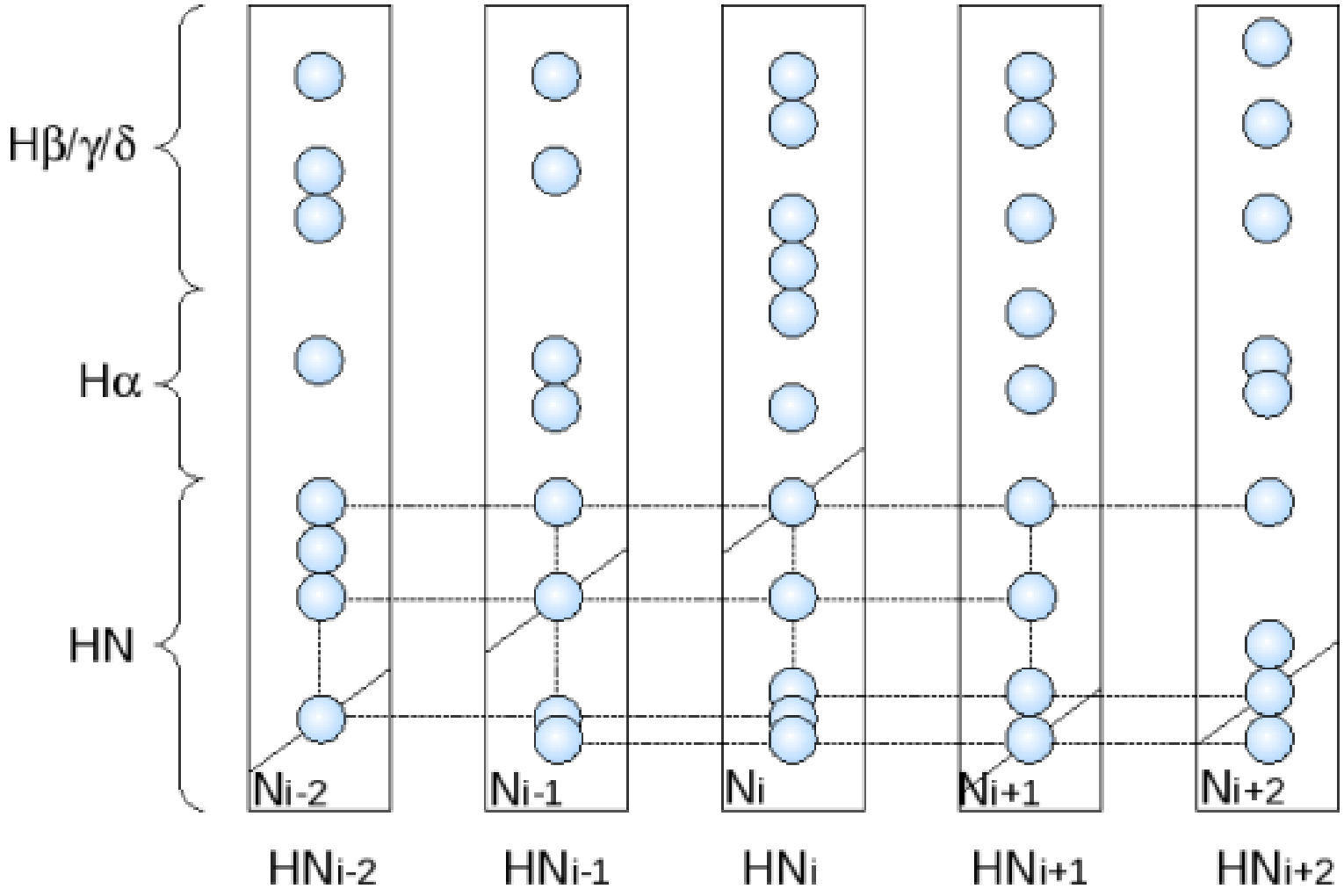

CBCANH (defaultProject)
Spectra Contours Peaks Strips 6.222, 32.97 (1H,13C) Help
i-1 = Thr
Cγ
i-1 = Arg
Cβ
Cδ
Cα
Not all peaks are always visible, as is the case for this Arginine
Each NHi strip should contain peaks for all carbon nuclei in the side-chain of residue i-1
Cα
Cβ
15N 119.25 119.44 119.06

Double Resonance Backbone Assignment

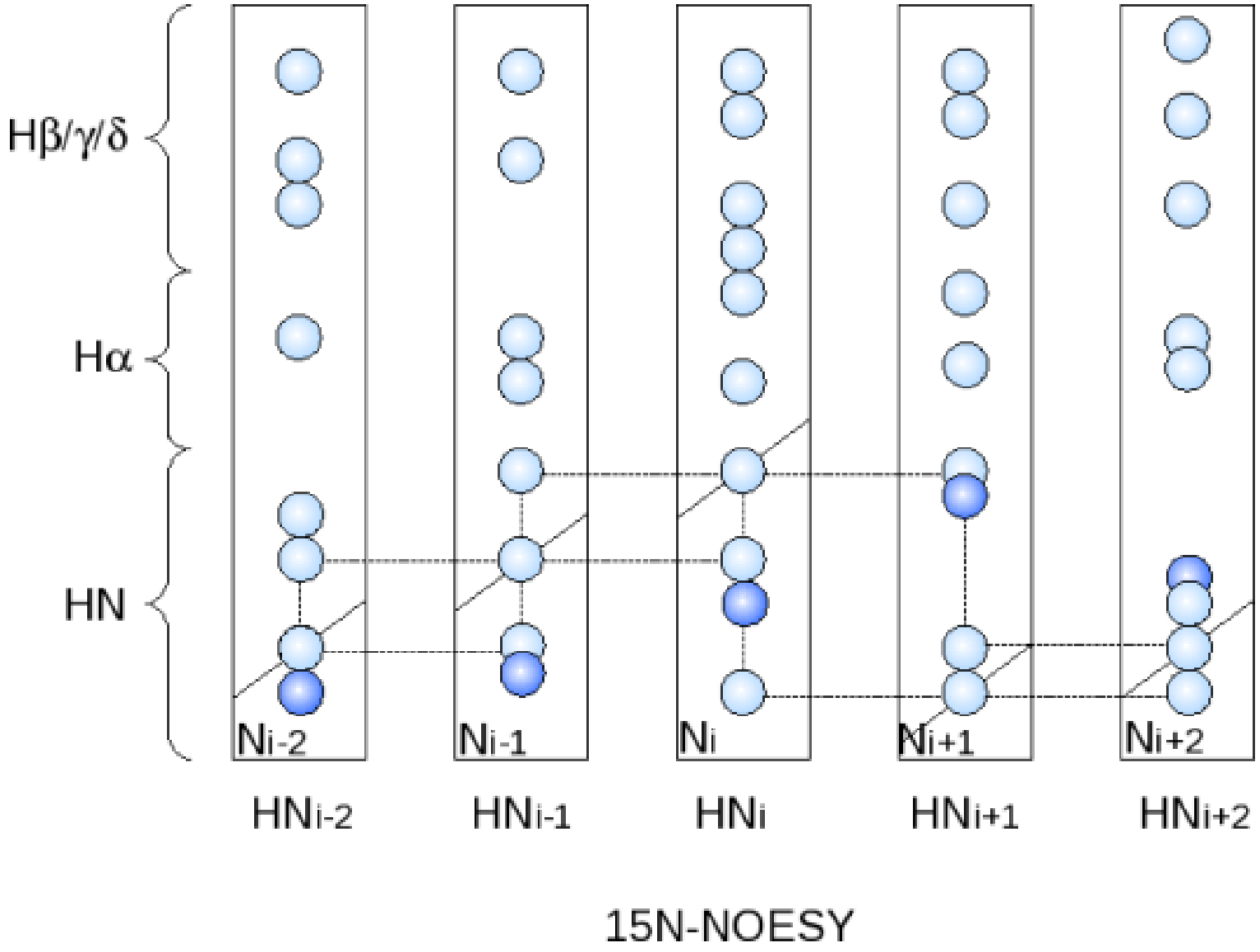
The 15N-NOESY-HSQC will show for each NH group all 1H resonances which are within about 5-7Å of the NH hydrogen. Assignment is done on the assumption that the two neighbouring NH groups are always visible.



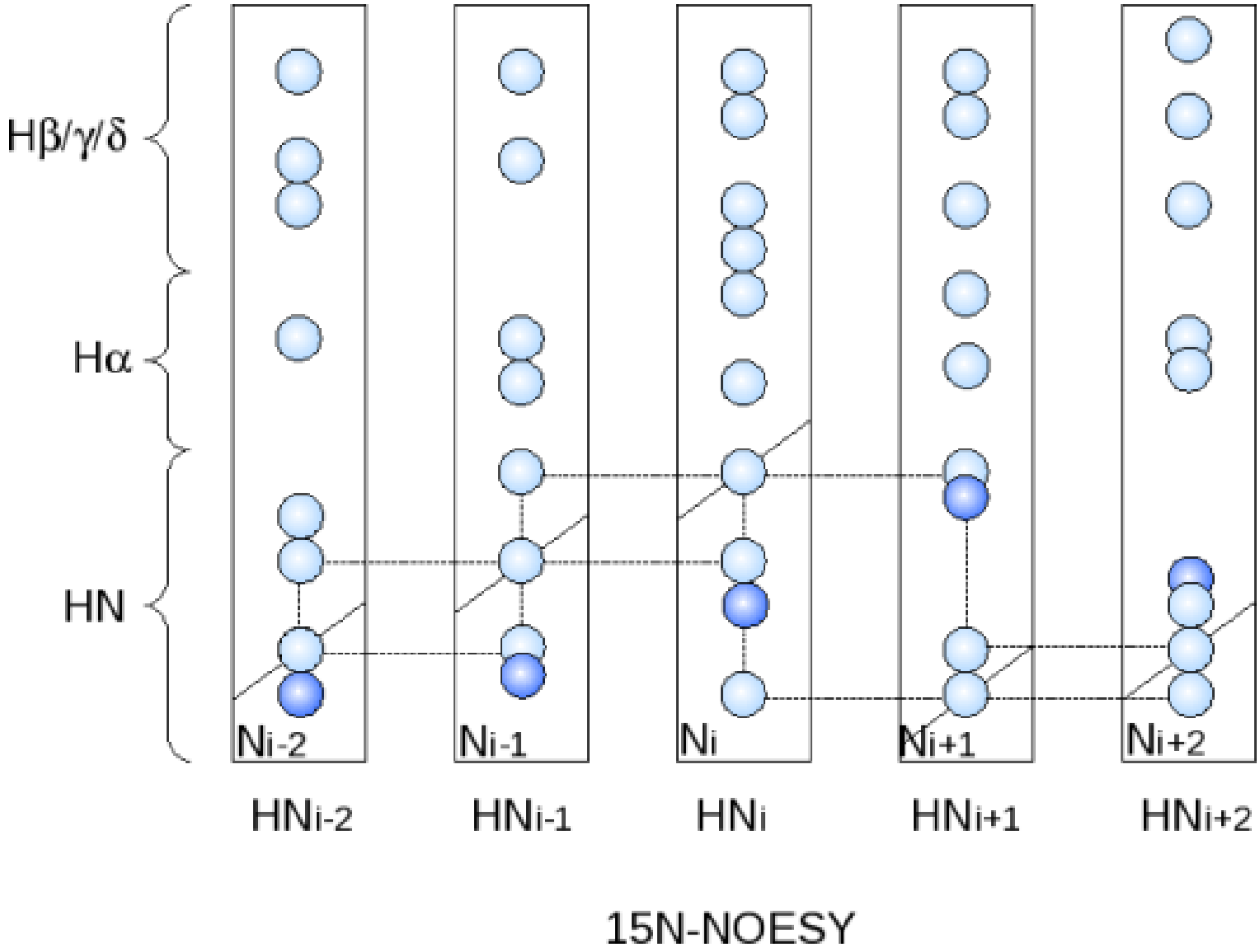
Helical sections are generally easier to assign, as NOEs from NH(i) are visible not only to NH(i±1), but also to NH(i±2) and sometimes NH(i±3).



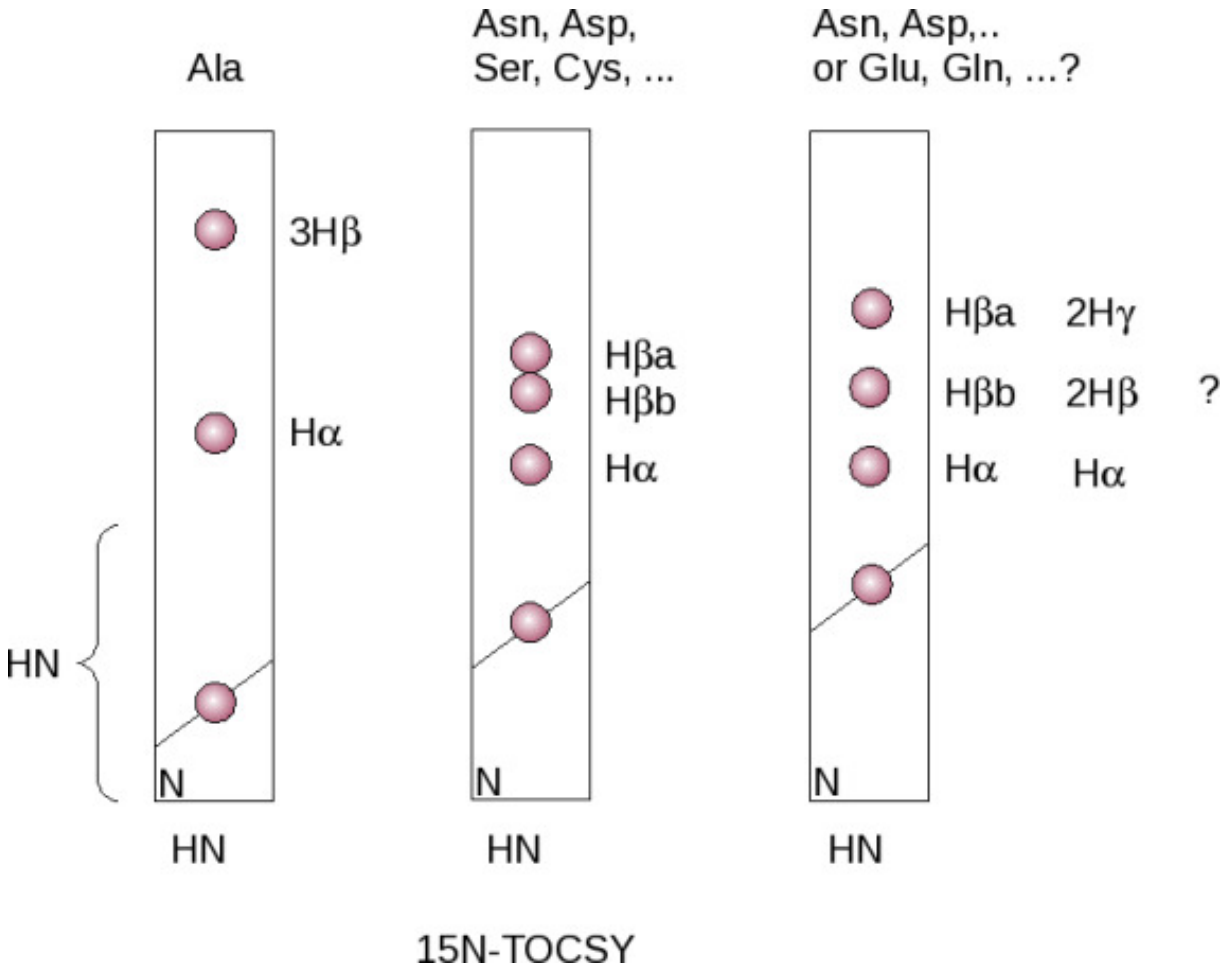
β-sheet structures include short NH-NH distances between the strands. This means that in addition to the NH(i±1) NOEs, a strong cross-strand NOE is also observed.



The side-chain NOEs from the 15N-NOESY-HSQC can also be useful during the assignment process, as NH(i)-Hα(i-1) are generally very strong, in particular in β-sheet sections.

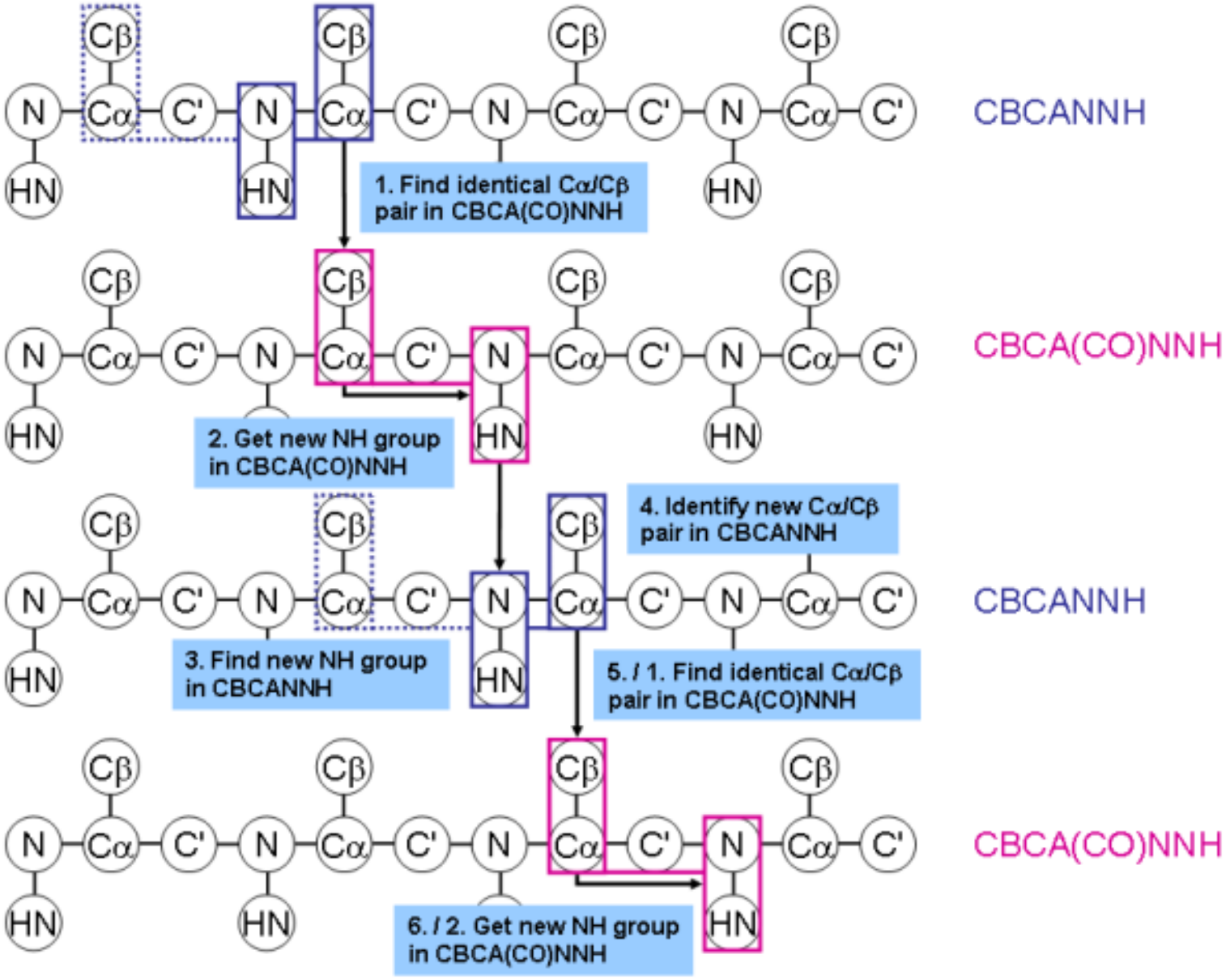


Double Resonance Side-chain Assignment

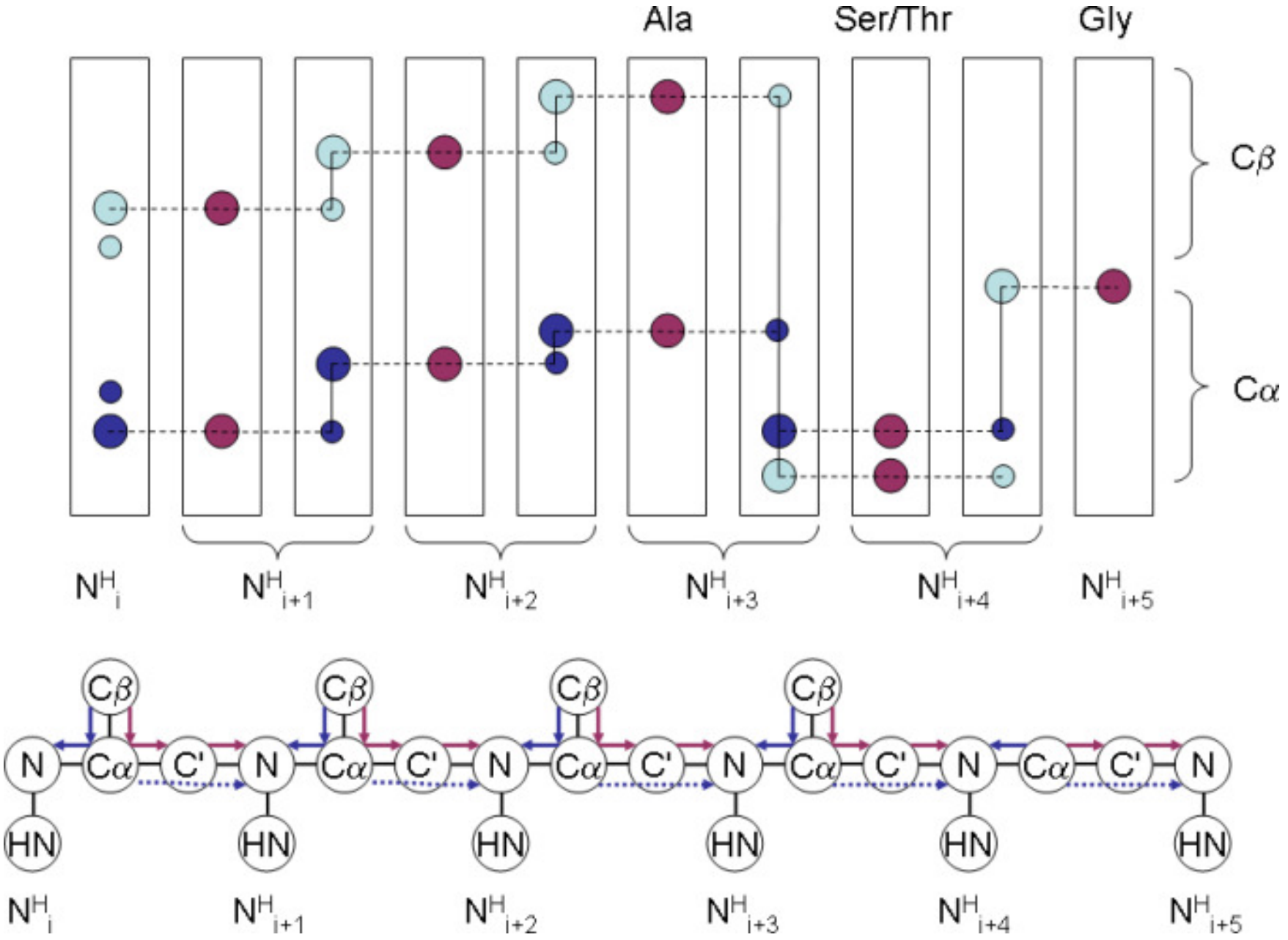


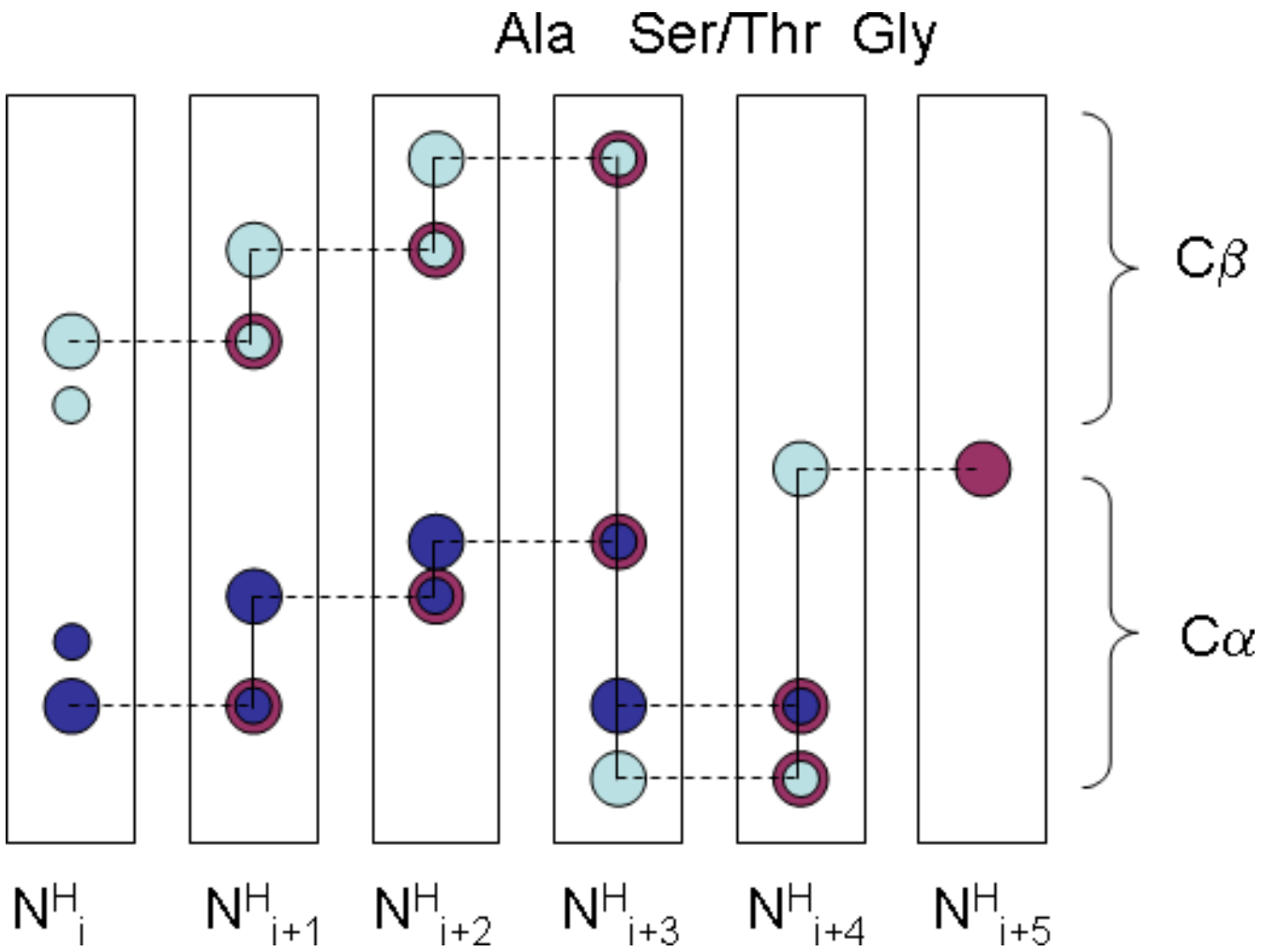
In many cases not all side-chain resonances may be visible, or for certain amino acid types it may not be possible to assign the resonances to particular side-chain groups. In particular it may in some cases be difficult to distinguish between Hβ and Hγ atoms and it may not always be possible to work out whether mythylene groups are degenerate or not.

Triple Resonance Backbone Assignment



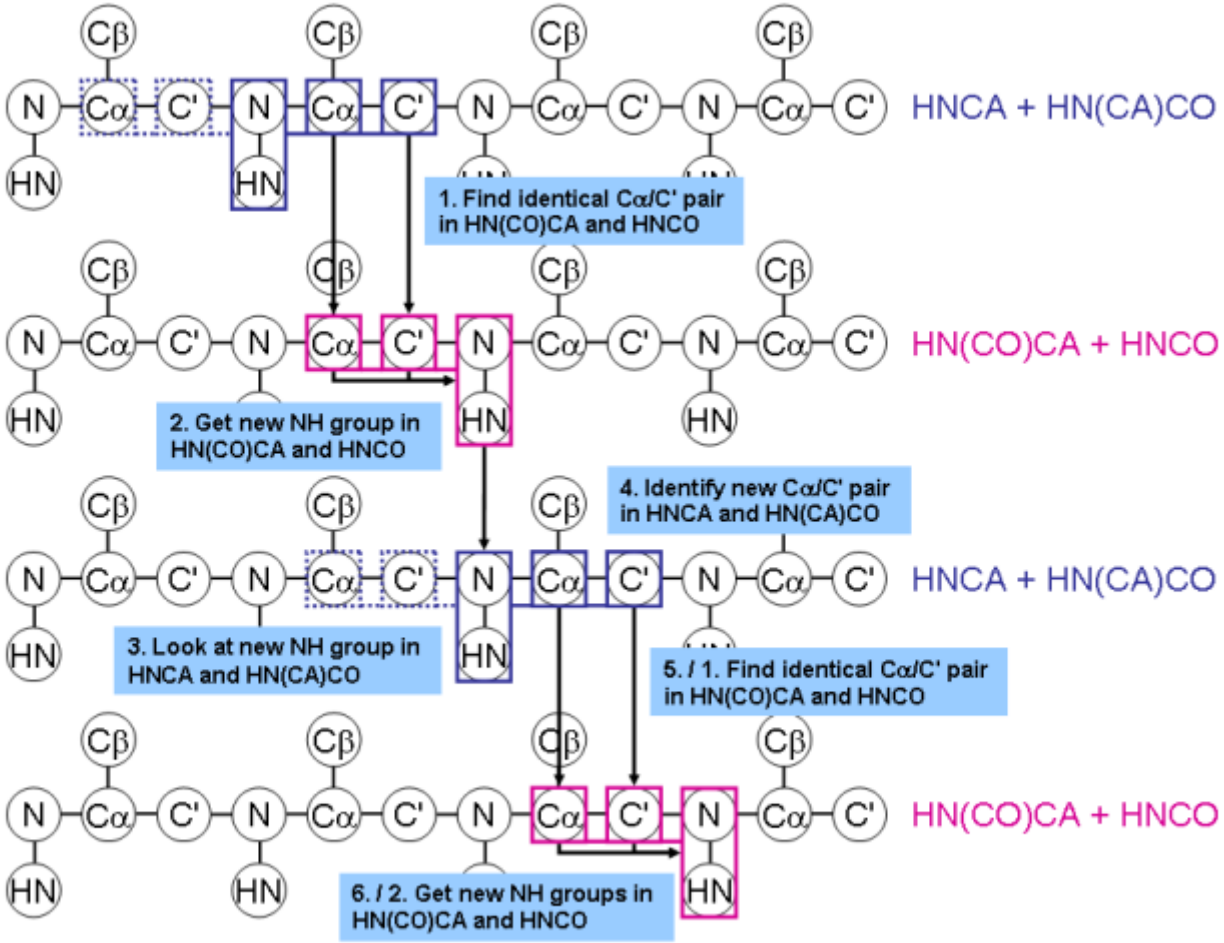
CBCANH and CBCA(CO)NH spectra



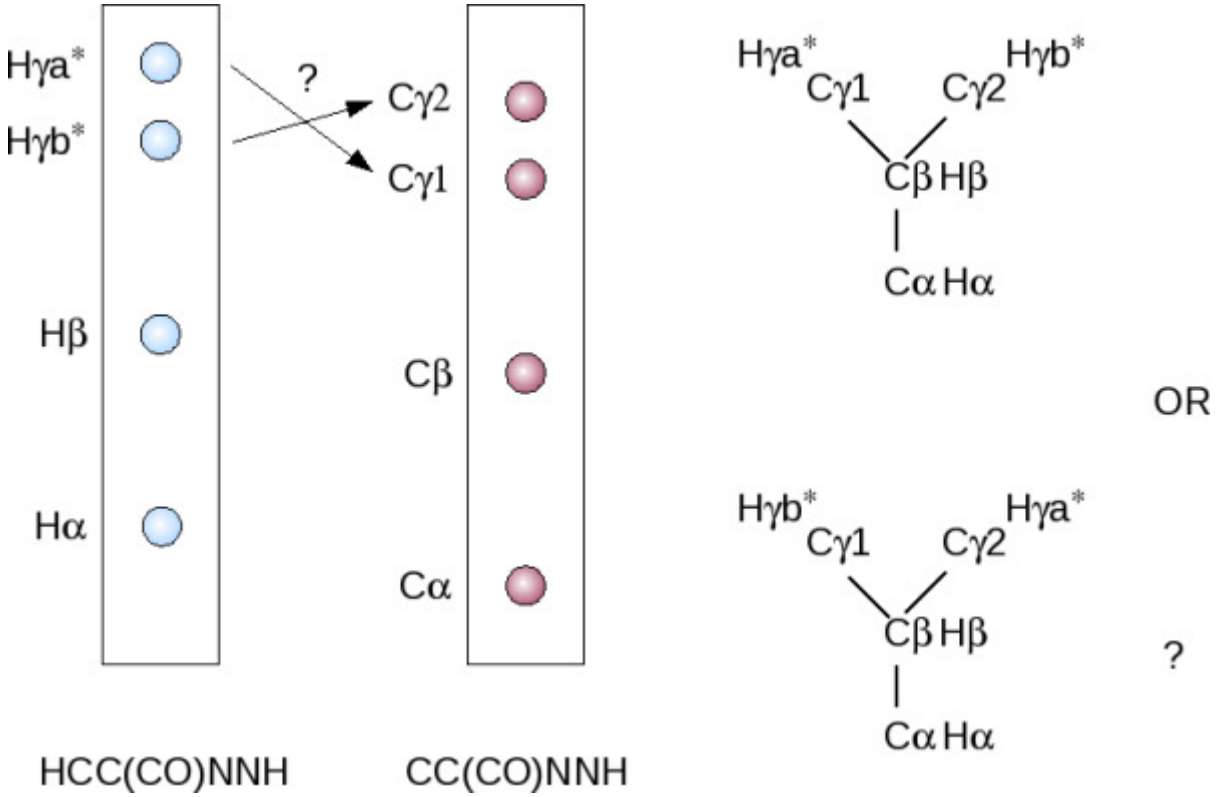


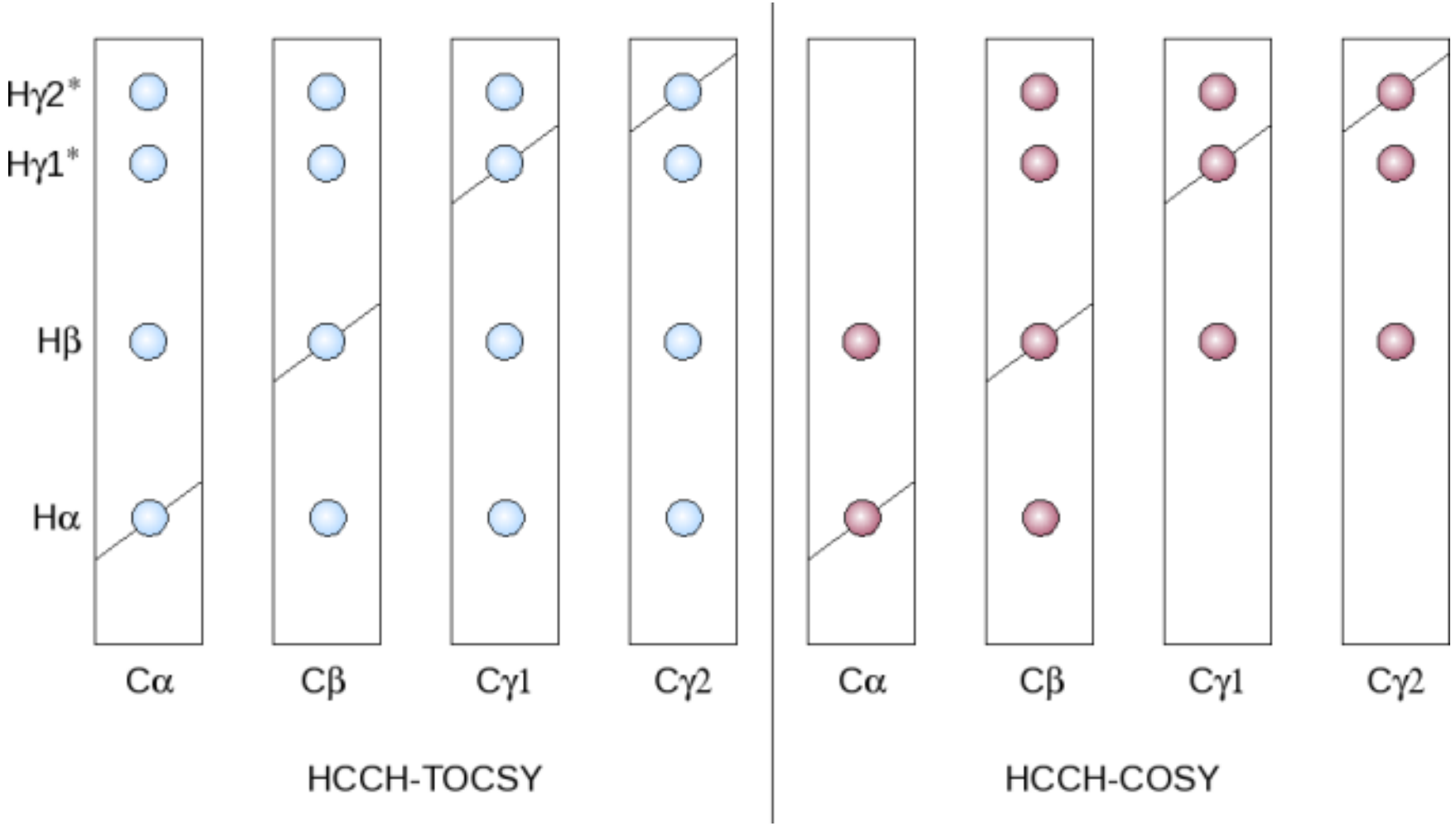
The Cα and Cβ chemical shifts adopt values characteristic of the amino acid type. Some of these, such as Alanine, Serine, Threonine and Glycine are very easy to spot as their Cβ chemical shifts are very different to those of the other amino acids (and in the case of Glycine there is no Cβ). Valine, Isoleucine and Proline are also likely to stand out by the fact that they have lower than normal Cα chemical shifts. Once a chain of NH groups with their corresponding Cα and Cβ chemical shifts has been built, then the identification of some of the amino acid types makes it possible to match this string to the sequence. E.g. a string of shifts may have been found that corresponds to xxxSxxAx – if this sequence only appears once in the sequence of the protein in question, then sequence-specific assignment can be made.

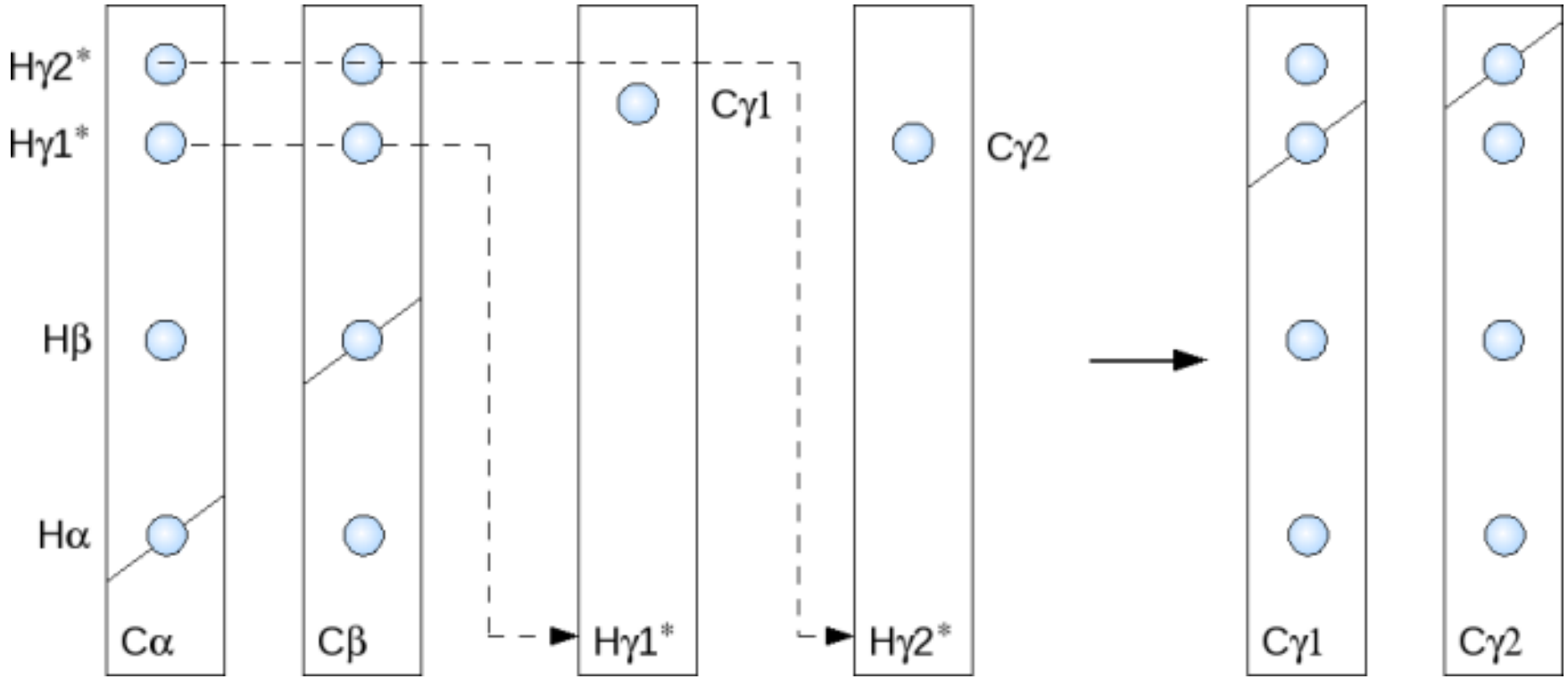
In some cases, in particular if your protein is fairly large (>200 residues, say), you may find that the quality of the CBCANNH and CBCA(CO)NNH spectra are not very good. The Cβ resonances may, for example, not be visible above the noise level. In this case it is possible to use the Cα and C' chemical shifts rather than the Cα and Cβ chemical shifts, as those which you use to walk from one residue to the next. The HNCA and HN(CO)CA experiments give you the same information as the CBCANNH and CBCA(CO)NNH spectra, except without the Cβ resonances. To complement this, you can then record the HNCO and HN(CA)CO experiments. These link each NH(i) group with the C'(i-1) (HNCO) or with C'(i) and C'(i-1) (HN(CA)CO).



Triple Resonance Side-chain Assignment




Hγ2*
Hγ1*
Hβ
Hα
Cα
Cβ
Cγ1
Cγ2
HCCH-TOCSY
Cα
Cβ
Cγ1
Cγ2
HCCH-COSY


Hγ2*
Hγ1*
Hβ
Hα
Cα
Cβ
Hγ1*
Cγ1
Hγ2*
Cγ2
Cγ1
Cγ2