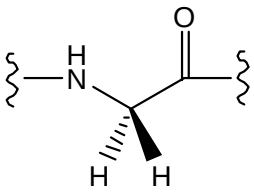
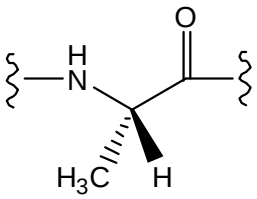
Problem Set 6 Solutions
TOCSY, NOE Correlation Spectroscopy

1. a. Our cyclic peptide contains only glycine and alanine, and these two simple amino acids should show characteristic sets of multiplets in the 1D spectrum:



glycine



alanine

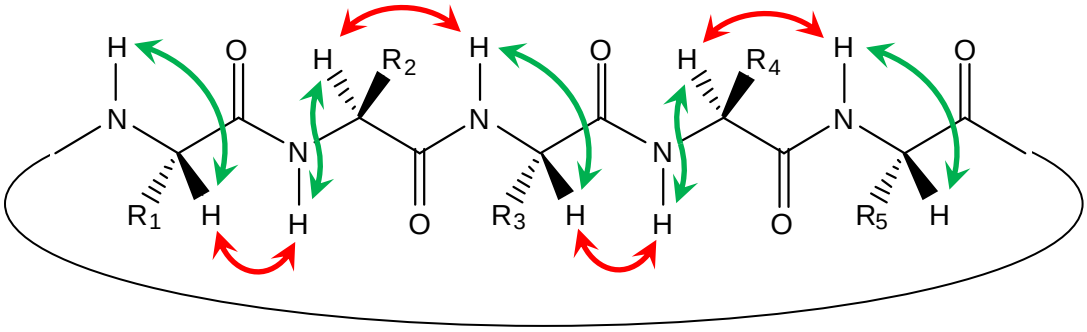
	Gly	Ala
NH	dd/t	d
$C_\alpha H$	dd (×2)	dq
$-CH_3$	-	d

Looking at the ^{1}H NMR spectrum, In the NH region (δ = 7-9 ppm), there are three doublets and two more complex multiplets. In the methyl region (δ = 1-2 ppm) there are three doublets. (Two of them overlap.) Clearly, or molecule had three alanines and two glycines.

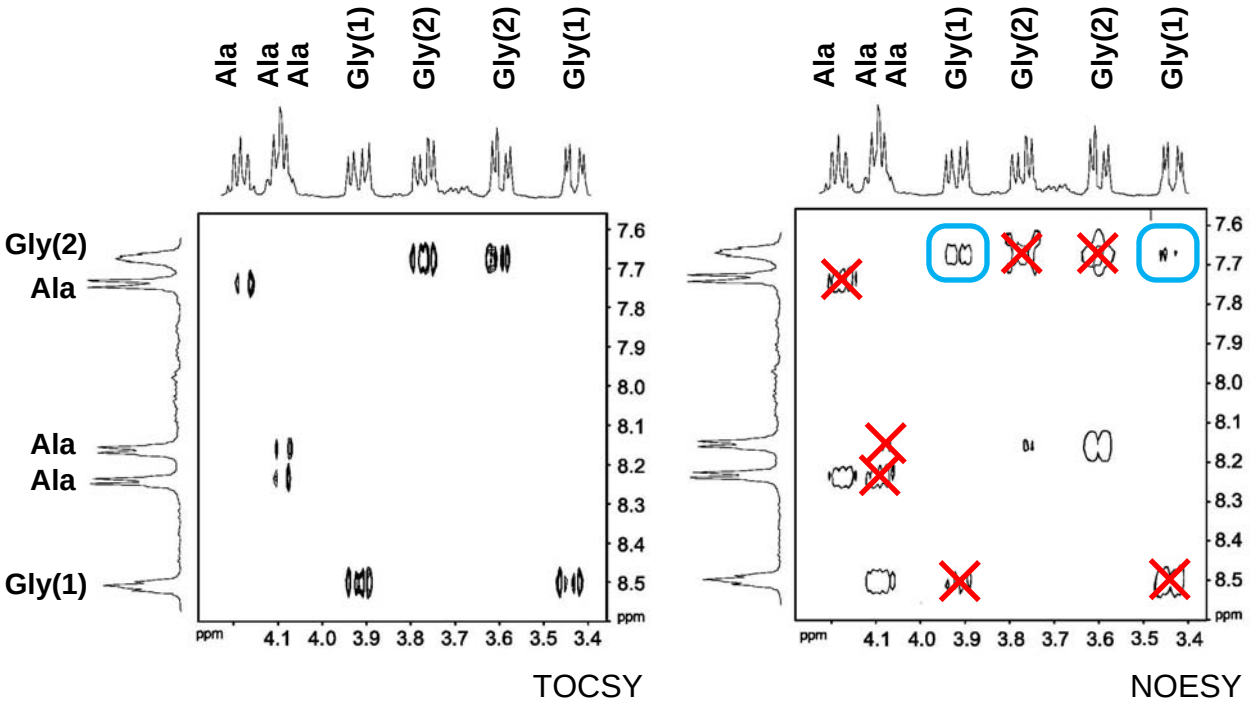
b. This means there are only two possible cyclic pentapeptide structures:

Ala - **Gly** - Ala - **Gly** - Ala *or* Ala - **Gly** - **Gly** - Ala - Ala

Basically, we're asking, are the two glycine residues adjacent, or not? We should be able to see all coupled connections in **TOCSY**, and all adjacent connections in **NOESY**:



We can identify most of these relationships in the NOESY, but most importantly, we can see a unique Gly-Gly NOE peak in the NOESY spectrum:



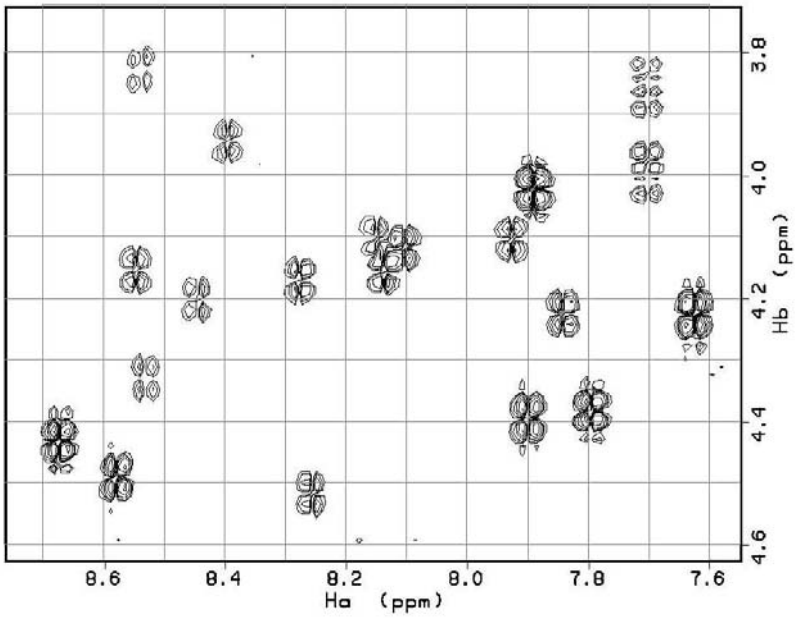
On the NOESY spectrum, red X's show uninformative crosspeaks that may correspond to TOCSY correlations rather than true NOESY correlations. Of the peaks that are unique to the NOESY spectrum, the circled ones are between a Gly and a neighboring Gly. As a result, we know that the structure is

Ala - **Gly** - **Gly** - Ala - Ala

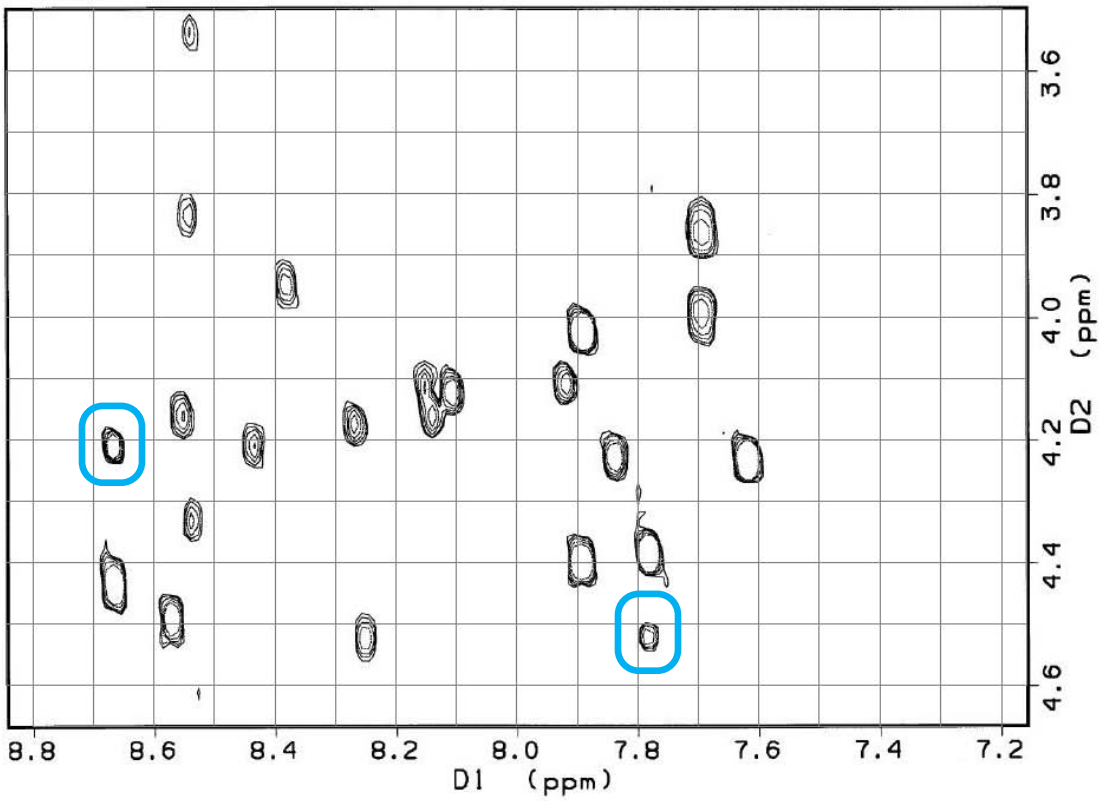
c. Guess I'm not really sure what I was asking here, but, yeah.

2. a. The COSY and TOCSY spectra are different sizes, but I tried to place them on their respective pages so that you could overlay the pages to find the peaks you needed. They are:

COSY

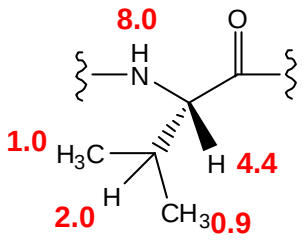


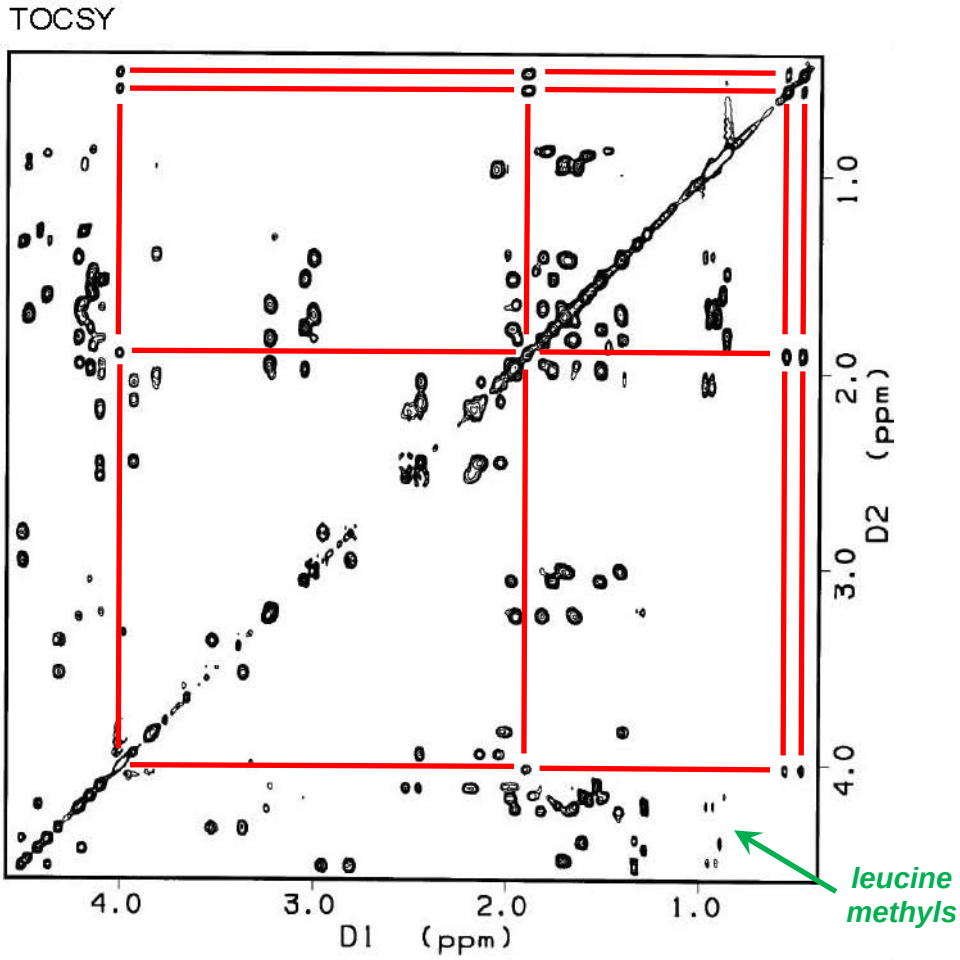
TOCSY



These circled resonances would have to correspond to amino acids that have, in addition to their C_α-proton in the δ = 3.6-4.6 ppm region, an additional one in that same region. I think the only possibilities are serine (C_β-H at 3.6-3.7 ppm) or threonine (C_β-H at ~4.4 ppm). Our peptide doesn't have any serine in it, but it does have two threonine residues, so we can tentatively assign those TOCSY crosspeaks to these two **Thr** residues.

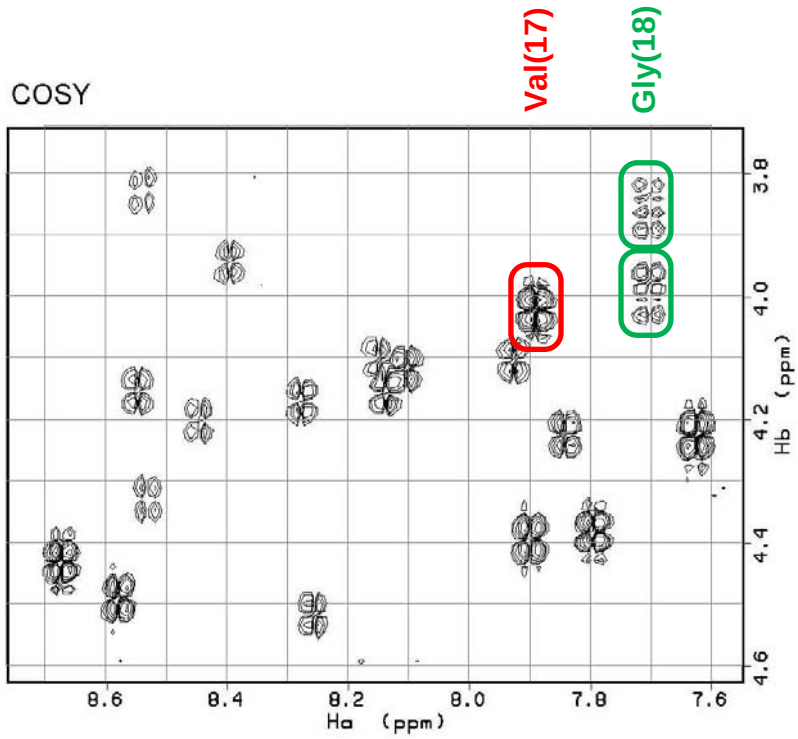
b. The amino acid chart says that valine should have the chemical shifts shown at right. If we can find this pattern of correlated chemical shifts in the TOCSY, we've found our amino acid. I actually thought the easiest resonances to find were the two inequivalent methyl groups. There are actually six amino acids in our peptide that have two inequivalent, upfield methyl groups: five leucines, where each methyl group is TOCSY-correlated to lots and lots of other inequivalent protons; and the one valine. Sure enough, in the TOCSY, we can see a number of paired (leucine) methyl groups that have lots and lots of correlations in the 1.5-2 ppm region, but just one (valine) pair that has only the correlations we want:



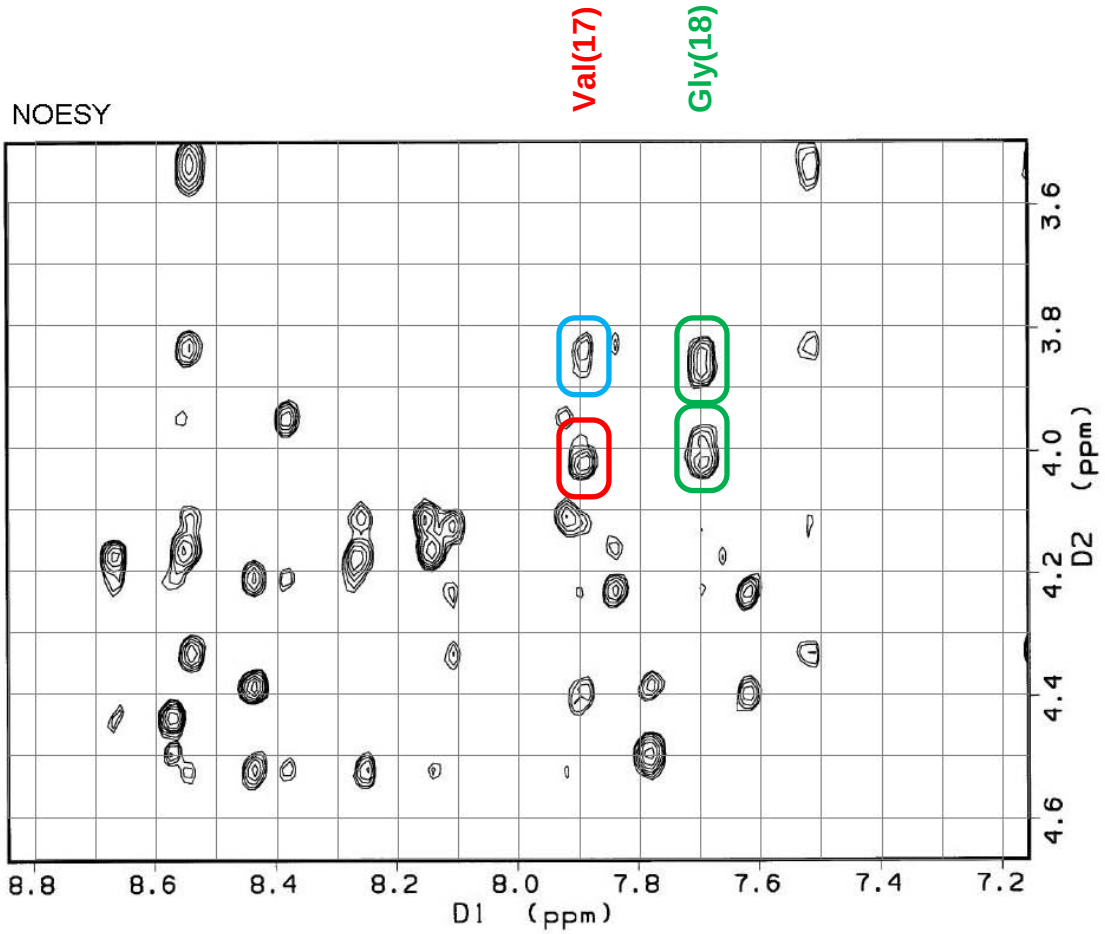


The TOCSY shows us the actual chemical shifts of those valine(17) protons; it shows that the methyl groups are upfield of the chart's prediction, and so is the $C_\alpha H$ (at almost exactly 4.00 ppm).

c. In the COSY closeup, we actually see two residues that have $\delta(C_\alpha H) = 4.00$ ppm. One has two adjacent $C_\alpha H$ protons; this must be a glycine residue, the only amino acid that has two coupled protons in this region. This matched pair (circled in green below) is characteristic of glycine. The other 4.00 ppm crosspeak must be valine(17).



These same crosspeaks also appear in the TOCSY spectrum. In the NOESY spectrum, there is one extra crosspeak that matches Val(17) and Gly(18):



Unfortunately, there is overlap between the TOCSY peaks for Val(17) and Gly(18) and what would be NOESY peaks between these two. But there is one unique NOESY peak (circled in blue above) that correlates these two amino acids. Oddly--I'm not sure why this is--the peak correlates the valine N-H with the glycine C_α-H. Because peptide and protein sequences are read N-to-C, this means that the NOESY correlation goes the long way: