

Simulation of Magnetization Pathways and Peak Inference in the new NMR Resonance Assignment Program Package CARA

Rochus Keller, Pascal Bettendorff, Fred Damberger and Prof. Kurt Wüthrich
Institute for Molecular Biology and Biophysics, ETH Zurich

The analysis of NMR spectra is both a critical step and today the most time consuming task in the process of protein structure determination. All relevant signals of a spectrum must be identified and the chemical shifts uniquely assigned to the atoms of the protein by logical inference. All NMR software packages known to the author base their information management on peak- and atomlists. Each spectrum is associated with a dedicated peaklist consisting of individual peaks. The synchronization of peaks among different peaklists is one of the major causes of the complexity related to resonance assignment. We present a novel approach implemented in CARA which infers the position of expected cross-peaks dynamically from a spectrum-independent signal repository, eliminating the need to synchronize peaklists.

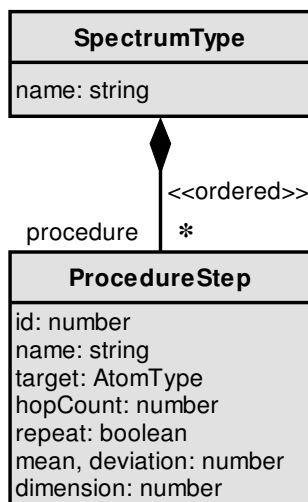


Fig. A: Class Diagram of a Spectrum Type

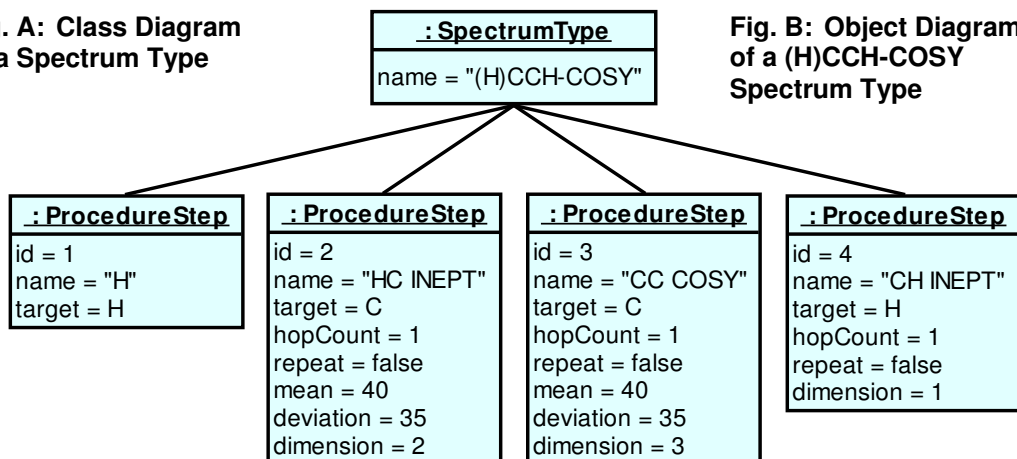


Fig. B: Object Diagram of a (H)CCH-COSY Spectrum Type

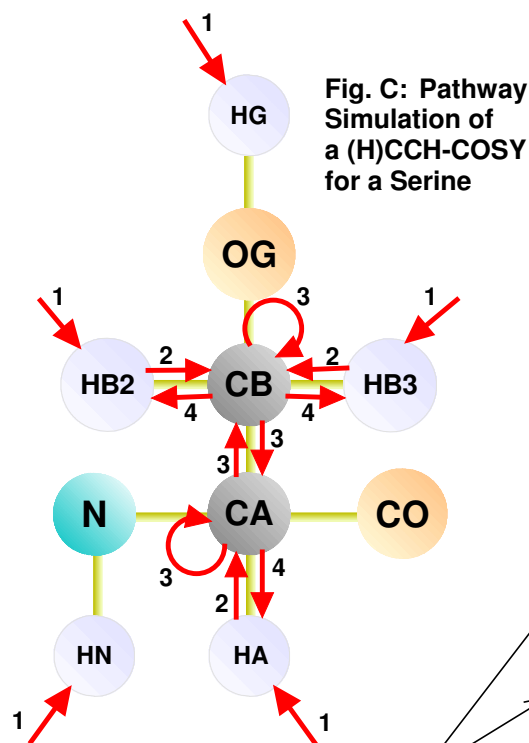


Fig. C: Pathway Simulation of a (H)CCH-COSY for a Serine

1. The domain of the first step corresponds to all *Atoms* of the Serine. Since *target* equals to *H*, the codomain of the relation are all *Atoms* with *type = H*.
2. The codomain of the first step becomes the candidate domain of the second. In the second step we are looking for all *Atoms* within one bond distance, whose *type* equals to *C* and *mean* is in the range of 40 ± 35 PPM. HG and HN don't meet these conditions and their path therefore break off.
3. The conditions of step three are identical to step two. CA and CB form the candidate domain for this step. Reflection is allowed, so CA and CB can relate to themselves. However the relation of CA and CO is not possible due to the restriction of the PPM range (i.e. the mean of CO is around 170 PPM).
4. This step tries to relate CA and CB to all *Atoms* with *type = H*. Due to the bond distance restriction (*hopCount* of one) this is only possible with HA, HB2 and HB3. If the step had defined a *hopCount* of two, then HN and HG would belong to the codomain.

Fig. D: Resulting State Tree of the Experiment

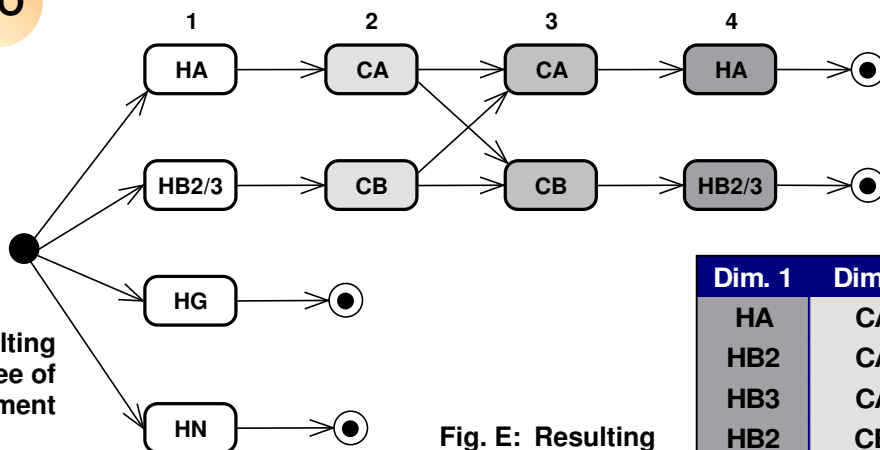


Fig. E: Resulting Table of Expected Cross-Peaks after Denormalization

Dim. 1	Dim. 2	Dim. 3
HA	CA	CA
HB2	CA	CB
HB3	CA	CB
HB2	CB	CB
HB3	CB	CB
HA	CB	CA