Efficient Amino Acid Type Identification in NMR Spectra of Proteins via β- and γ-Carbon Edited Experiments

J. Santoro, D. Pantoja-Uceda

Instituto de Química Física Rocasolano, CSIC, 2806 Madrid, Spain

For the sequential assignment of the protein backbone, a number of powerful multidimensional triple-resonance experiments are available. With these experiments, stretches of sequentially neighboring amino acids are obtained which, in a subsequent step, must be mapped on the known amino acid sequence. For this purpose, the amino acid type must be determined for some at least of the assigned peaks. Some amino acids, G, A, and the pair S/T, can be identified from their characteristic $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts. However, for the rest of amino acids chemical shifts are not unique, and categorization cannot be made on this basis. Thus, methods that provide a better amino acid classification will have a very positive impact in the assignment of protein backbone resonances, making automatic methods that include this kind of information more efficient and robust.

The 20 protein amino acid side chains form eight topology classes with respect to the number of hydrogen atoms attached to Cβ carbons and to the number and type of carbons at the γ position. Therefore, the Cβ carbons are ideal as starting point in pulse sequences for amino-acid-type identification. However, distinguishing the eight amino acid groups is difficult by NMR, since methods usually differentiate only between odd and even multiplicities. Here we present a sign-encoding scheme to differentiate among six amino-acid-type groups from a single 2D (CBCACO)NH experiment. The experiment yields several ^{1}H-^{15}N correlation subspectra, that reveal the nature of the amino acid preceding the NH moiety. Each ^{1}H-^{15}N correlation peak is unambiguously assigned to a group according to the subspectrum in which it is detected and to its sign. A large group, that consists of the seven amino acids with a single aliphatic γ carbon, is obtained. We have shown that this large group can be subdivided into two subgroups of similar size by using a sequential γ-carbon edited pulse sequence. The classification can be further extended by using an intraresidual β-carbon edited experiment. This experiment produces the same six amino-acid-type groups obtained with the sequential experiment. Unfortunately, in this case, no subdivision of the large group can be obtained. Besides, the classification of the intraresidual amino acid is possible only if its type does not coincide with the one of the sequential amino acid. Therefore, only approximately 75% of the intraresidual amino acids can be unambiguously classified.

Results of the application to two proteins, ubiquitin and PUB1, will be presented. For ubiquitin (76 residues, 1.7 mM) acquisition times of ~20 min for each 2D experiment proved to be sufficient to yield high-quality spectra**.** For PUB1 (101 residues, 0.2 mM) the three carbon-edited experiments were obtained in 15 hours and provided the assignment of all NH's.

This work has been carried out with financial aid of grant BFU2005-01855.