A generalized approach to automated NMR peak list editing: application to reduced dimensionality triple resonance spectra

Hunter N.B. Moseley, Nadeem Riaz, James M. Aramini, Thomas Szyperski, Gaetano T. Montelione

Abstract

We present an algorithm and program called Pattern Picker that performs editing of raw peak lists derived from multidimensional NMR experiments with characteristic peak patterns. Pattern Picker detects groups of correlated peaks within peak lists from reduced dimensionality triple resonance (RD-TR) NMR spectra, with high fidelity and high yield. With typical quality RD-TR NMR data sets, Pattern Picker performs almost as well as human analysis, and is very robust in discriminating real peak sets from noise and other artifacts in unedited peak lists. The program uses a depth-first search algorithm with short-circuiting to efficiently explore a search tree representing every possible combination of peaks forming a group. The Pattern Picker program is particularly valuable for creating an automated peak picking/editing process. The Pattern Picker algorithm can be applied to a broad range of experiments with distinct peak patterns including RD, G-matrix Fourier transformation (GFT) NMR spectra, and experiments to measure scalar and residual dipolar coupling, thus promoting the use of experiments that are typically harder for a human to analyze. Since the complexity of peak patterns becomes a benefit rather than a drawback, Pattern Picker opens new opportunities in NMR experiment design.

Keywords: Automated NMR data analysis; Depth-first search; Pattern Picker; Peak list editing; Reduced dimensionality

1. Introduction

Advances in sample preparation, hardware for data collection pulse sequence development and experiment design, and software for automated analysis provide a significant reduction in the time necessary to generate biomolecular NMR structures [11-13,21-23,35]. International efforts in structural genomics centering on “high throughput” analysis of protein resonance assignments and solution NMR protein structure determination are spearheading advances that reduce the time, effort, and expense to generate protein resonance assignments and structures [4,14,21,26]. Spurred in part by the demand and opportunities of structural genomics, many steps in the process of NMR data collection and analysis have been streamlined and greatly improved. Still, the first spectral analysis step in the NMR protein structure determination process, peak picking and peak list edit-