A Graph-based Peak Walking Algorithm for the Chemical Shift Mapping of Peaks in Serially Titrated ^{15}N-HSQC NMR Spectra

Richard Jang[1*], Xin Gao[2], Ming Li[1]

[1] David R. Cheriton School of Computer Science, University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1

[2] Division of Mathematics and Computer Sciences and Engineering,
King Abdullah University of Science and Technology,
Thuwal, Saudi Arabia

Email: {rjang, mli}@uwaterloo.ca, xin.gao@kaust.edu.sa

* Corresponding author

ABSTRACT

In NMR-based drug design and screening, chemical shift perturbation/mapping is a popular technique used to identify the atoms of a target protein that have changed when a potential drug molecule is introduced in increasing concentrations. The mapping of peaks in the reference spectrum of the unbound protein to peaks in perturbed spectra is often done semi-automatically, especially for large proteins, due to missing, missing but then reappearing, overlapped, noisy, and new peaks not associated with any peaks in the reference. Such errors can lead to ambiguous mappings. Nevertheless, automated methods are necessary for high-throughput drug screening. To model the above errors and multiple possible paths, we present a novel peak walking algorithm for fast-exchange systems, based on a constrained k-dimensional maximum weight matching problem formulation, where k is the number of spectra. With, on average, less than two paths connecting peaks in the final perturbed spectrum to peaks in the reference, we obtained an accuracy of 95% for the proteins UbcH5B titrated with Not4 (k=5), and hBclXL with BH3I-1 (k=11). This has implications for backbone resonance assignment methods that use contact information because it is known that the assignment accuracy increases if there is accurate spin system typing information with few possible residues per peak. The multiple path information provided by our algorithm can provide the typing if the resonance assignment of the reference spectrum is known, as is generally the case.