

## COMPLX - A COMPUTER ALGORITHM FOR THE DETECTION OF PROTEIN-LIGAND AND OTHER MACROMOLECULAR COMPLEXES IN MASS SPECTRA

Jason W.H. Wong and Kevin M. Downard

School of Molecular and Microbial Biosciences, University of Sydney, Sydney, Australia

Mass spectrometry is an important analytical technology for the study of proteins and other macromolecules. It has recently provided a unique window into the nature and stability of protein and macromolecular interactions. Despite some perceived and legitimate limitations of mass spectrometry for these investigations, including the ability to both preserve and detect protein and other complexes in the absence of solvent, the approach has offered some tantalizing glimpses of macromolecular complexes and assemblies that are difficult or even impossible to study by other methods.

We have devised and constructed a new algorithm specifically designed to detect protein complexes, or any other macromolecular complex, in mass spectral data. The program takes advantage of the appearance of multiply-charged ions common to electrospray mass spectra, and also detected in MALDI mass spectra for moderate to large molecular weight proteins [1]. This algorithm, known as COMPLX (COMposition of Protein-Ligand compleXes) [2], is capable of identifying complexes for any protein or macromolecule with a binding partner whose molecular weight ranges from 1 to 100,000 Daltons. Thus large protein-protein complexes can be detected as easily as the interactions between proteins and small molecule weight ligands.

Data is input as a mass-intensity list generated directly from the mass spectral data or typed manually into an input file. The program can operate using either positive or negative ion mass spectral data. The output consists of a list of  $m/z$  values for pairs of ions associated with one of the binding partners and the complex, together with their charge state and the calculated mass of the ligand. Ions possibly associated with the ligand are also identified in the mass spectral data. The output data is scored and ranked according to the standard deviation of the ligand mass, the number of continuous charge states associated with the set of ions pairs, and the presence of ions associated with the ligand in the mass spectrum.

Although developed primarily to advance the study of protein associations, the program can be implemented to identify any macromolecular complex for which mass spectral data have been generated. The program can be successfully implemented for complex mixtures. Thus it may assist mass spectrometry investigations to directly analyse protein and other macromolecular associations within cell lysates and other biological extracts to extend mass spectrometry's role in proteomics.

This presentation will describe the computational basis of the algorithm as well as the implementation of the program for both hypothetical and experimental mass spectral data of varying complexity. It will also demonstrate the web-interface to access and run the program across the internet.

[1] J.G. Kiselar, K.M. Downard (2000) *J. Am. Soc. Mass Spectrom.* 11(8): 746-750.

[2] J.W.H. Wong, K.M. Downard, *submitted for publication.*

---