COMPLEXITY AND SCORING FUNCTION OF MS/MS PEPTIDE DE NOVO SEQUENCING

Changjiang Xu and Bin Ma
Department of Computer Science, University of Western Ontario,
London, ON N6A 5B7, Canada

Tandem mass spectrometry (MS/MS) has become a standard way for identifying peptides and proteins. A scoring function plays an important role in the MS/MS data analysis. De novo sequencing is the computational step to derive a peptide sequence from an MS/MS spectrum, normally by constructing the peptide that maximizes the scoring function. A number of polynomial time algorithms have been developed based on scoring functions that consider only either the N-terminal or C-terminal fragment ions of the peptide. It remains unknown whether the consideration of the internal fragment ions will still be polynomial time solvable. In this paper, we prove that the internal fragment ions make the de novo sequencing problem NP-complete. We also propose a regression model based scoring method to incorporate correlations between the fragment ions. Our scoring function is combined with PEAKS de novo sequencing algorithm and tested on ion trap data. The experimental results show that the regression model based scoring method can remarkably improve the de novo sequencing accuracy.