Improving phosphorylation site assignment by statistical analysis on CID fragmentation

Sheng-Chieh Lu[1], Chia-Feng Tsai[2], Chih-Chiang Tsou[1], Han-Yin Yang[1], Lien-Chin Chen[1], Yu-Ju Chen[2], Ting-Yi Sung[1] and Wen-Lian Hsu[1]

1. Institute of Information Science, Academia Sinica, Taiwan ROC

2. Institute of Chemistry, Academia Sinica, Taiwan ROC

Introduction

Identification of a phosphoproteome using mass spectrometry consists of sequence identification and phosphorylation site assignment. Commonly used search engines, e.g., SEQUEST and MASCOT, can provide good phosphopeptide sequence identification, however, frequently suffer from the ambiguous site determination when dealing with phosphopeptides with multiple possible sites. Neutral loss fragment ions are commonly used to determine sites from multiple candidates for data generated from collision-induced dissociation (CID) experiments. However, the indistinguishability between fragment ions with neutral loss of phosphate and fragment ions with loss of water in MS/MS spectra may yield incorrect site assignment. In this study, we incorporate the frequencies of presence of known ion types and internal fragments to develop a scoring model to improve the phosphorylation site identification.

Method

Currently, we used a phospho-enriched sample undergoing CID fragmentation and electron transfer dissociation fragmentation and analyzed by MASCOT. Based on the search results confidently identified ($p<0.05$), i.e., in both experiments, the corresponding 1740 CID spectra were used for calculating the frequencies of different fragment ion types, including b and y ions, and b and y ions with intact phosphate and with neutral losses of ammonia, water, and phosphate, that could be matched in the MS/MS spectra. We used the result to determine the importance of respective ion types to design a scoring function for phospholylation site assignment. In addition to these commonly used ion types, we additionally incorporate internal fragments to help MS/MS sequencing and phosphorylation site assignment.

Result

The result showed that tyrosine-containing fragments seldom occurred with

neutral loss of phosphate, while serine-containing and threonine-containing fragments frequently occurred with loss of phosphate. Furthermore, more y ions were matched in the MS/MS spectra than b ions were, and this finding is consistent with previous studies. We also found that the presence of fragment ions with water loss and ammonia loss were less than that of loss of phosphate. Toward site assignment, we designed a scoring function on fragment ions with weights determined based on the above analysis. We applied our scoring method to a standard phosphoprotein experiment (bovine proteins, Q-TOF instrument) and used the phosphopeptides with site localization confirmed in the literature as benchmark. Our site assignment improved Mascot's results by 14%. Next, we included internal fragment ions in the scoring function. It showed that internal fragment ions helped increasing the confidence of site assignment results and reducing ambiguity of site assignment. This result reflected that the internal fragments could be a useful feature for determining phosphorylation sites and even peptide sequences, and worth further study.

Novel aspect

We incorporated the frequencies of fragment ions including internal fragments and developed a scoring model to improve phosphorylation site assignment.