



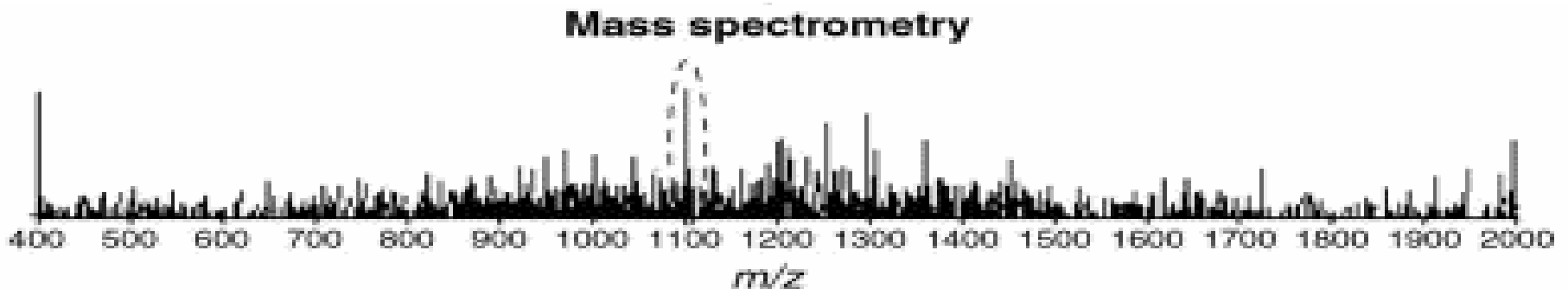
More Accurate Identification of Proteins Using Various Approaches

Nanjundaswamy A. and Cios K.

University of Colorado at Denver

Mass Spectrometry

- Mass Spectrometry is commonly used technique for protein identification
- MS measures mass-to-charge ratio, which gives the molecular weight and the pattern of peptides derived from the protein
- MS is therefore a general method for measuring all modifications that change the molecular weight of the peptides



Procedure - Identify Proteins

- Peptide Mass Mapping - identify proteins by comparing observed mass with the predicted masses of digested proteins stored in a database, through a search program like Mascot or MSFIT
- Unfortunately, many unmatched masses arise due to contaminated proteins, unexpected cleavage site, or common chemical modifications
- To explain those unmatched peptide fragments we developed software to detect both, the potential single amino acid substitutions in peptides, and the Post Translational Modifications (PTMs)

Amino Acid Substitution

- Often some of the amino acids are chemically modified, and this is one of the reasons that some of the mass sequences are not identified

- Procedure used in the software:

- each Amino Acid in the unmatched peptide sequence is substituted for each of the 20 Amino Acids, one after the other

```
MSFGSEHYLC SASSYRKVEFG DSSRLSARLS GPGGSGSFRS QSLSRSNVAS
TAACSSASSL GLGLAYRRLP ASDGLDLSQA AARTNEYKII RTNEKEQLQG
LNDREAVFIE KVHQLETQNR ALEAELAALR QRHAEPSRVG ELFQRELREL
RAQLEEASSA RAQALLERDG LAEEVQRLRA RCEEEESRGRE GAERALKAQQ
```

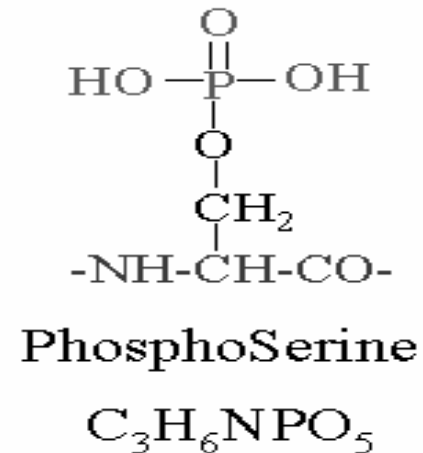
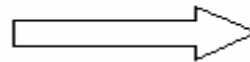
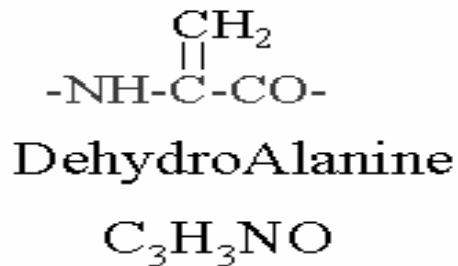
- then, the molecular weight is recalculated
- if a match is not found we continue the procedure for the rest of the unmatched sequence

Post Translational Modifications

- Proteins undergo a huge number of post translational modifications (PTMs)
- PTMs, or chemical modifications, are the result of covalent linkage of chemical group to amino acid in the protein
- PTMs of a protein can determine its activity state, localization, turnover, and interactions with other proteins
- All protein PTMs are associated with either an increase or a decrease in mass

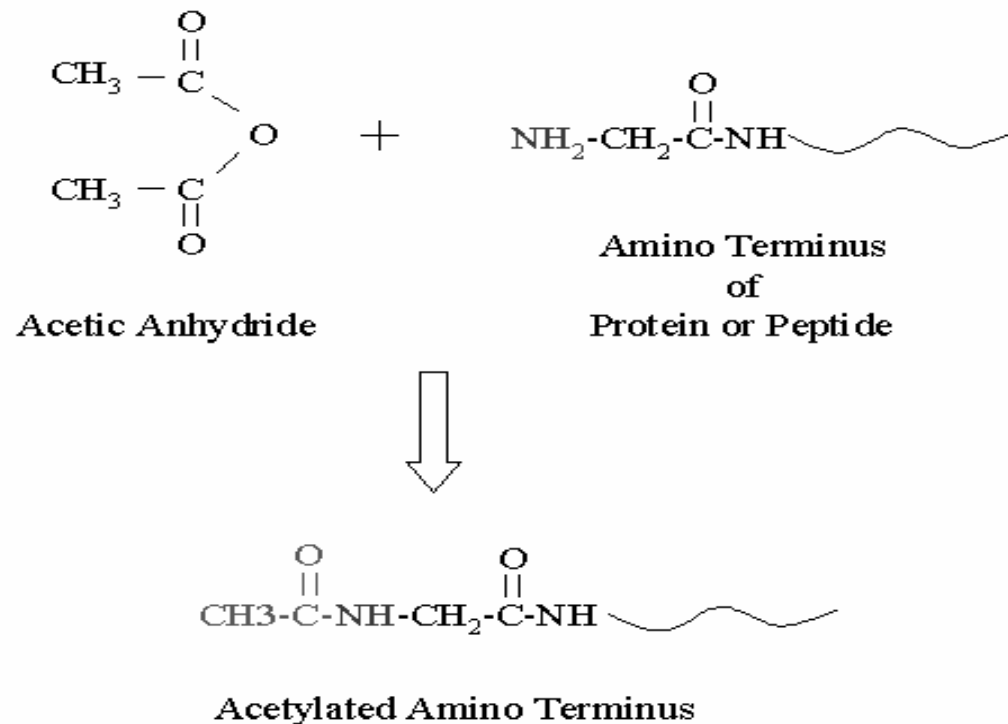
Phosphorylation

- Phosphorylation on Serine (S), Threonine (T) and Tyrosine (Y) residues is a very important modulator of protein function
- If in the unmatched sequence we find S, T or Y amino acid then we remove the phosphate group and compute the molecular weight of the sequence



Acetylation

- In Acetylation the proteins are modified at the N-Terminal
- If in an unmatched sequence we find N followed by any of the Amino Acids except for N,K,R,H,F,W,Y then we check whether that sequence has been modified by Acetylation





Results

True Protein	Number of Masses given	Masses matched by MASCOT	Name of the Protein identified	Masses matched after Amino Acid Substitution	Masses matched after Phosphorylation	Masses matched after Acetylation
heat shock 70kD protein 5 [Rattus norvegicus]	47	28	heat shock 70kD protein 5 [Rattus norvegicus]	9	3	1
hypothetical protein MGC25352 [Mus musculus]	53	31	hypothetical protein MGC25352 [Mus musculus]	7	5	1
Dihydropyrimidinase related protein-2 (DRP-2)	47	20	Dihydropyrimidinase related protein-2 (DRP-2)	11	5	0
stress-induced phosphoprotein 1; IEF SSP 3521; Hsp70/Hsp90 organizing protein	55	27	stress-induced phosphoprotein 1; IEF SSP 3521; Hsp70/Hsp90 organizing protein	19	2	1



Results

True Protein	Number of Masses given	Masses matched by MASCOT	Name of the Protein identified	Masses matched after Amino Acid Substitution	Masses matched after Phosphorylation	Masses matched after Acetylation
78 kDa glucose-regulated protein precursor (GRP 78)	49	28	78 kDa glucose-regulated protein precursor (GRP 78)	10	2	2
dnaK-type molecular chaperone HSPA5 precursor - human	47	25	dnaK-type molecular chaperone HSPA5 precursor - human	9	3	1
stress-induced phosphoprotein 1; IEF SSP 3521; Hsp70/Hsp90 organizing protein	55	27	stress-induced phosphoprotein 1; IEF SSP 3521; Hsp70/Hsp90 organizing protein	19	2	1



Conclusions and Future Work

- The software can be used to positively identify proteins that could not be identified by search engines like Mascot
- We plan to extend the software so it accounts for several more PTMs
- We also plan to modify the package so it works with other mass spectrometry data than MALDI-TOF