Journal of American Society of Mass Spectrometry

Large-Scale Examination of Factors Influencing Phosphopeptide Neutral Loss during Collision Induced Dissociation

Robert Brown[†], Scott S. Stuart[†], Stephane Houel, Natalie G. Ahn[†§] and William M. Old[†‡]

[†] Department of Molecular Cellular and Developmental Biology, [††]Department of Chemistry and Biochemistry, [§]Howard Hughes Medical Institute,

University of Colorado, Boulder, CO 80309

[‡]Corresponding author: William M. Old

Department of Molecular Cellular and Developmental Biology

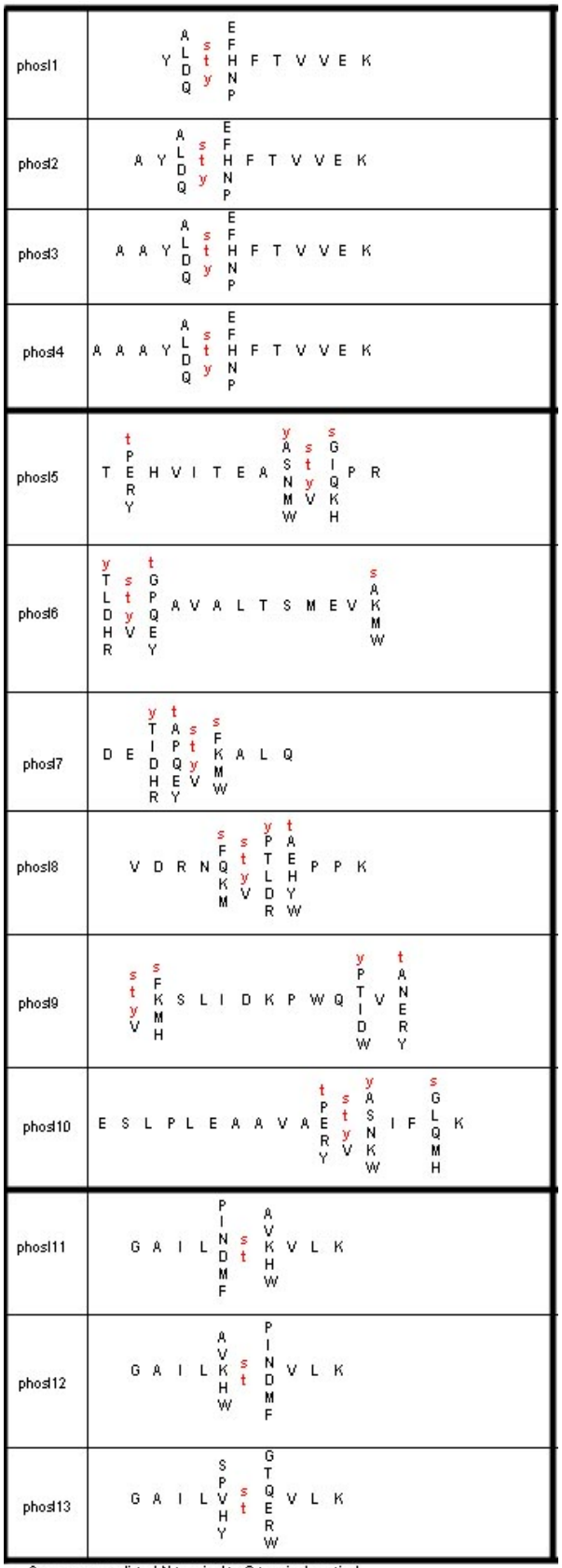
University of Colorado

Boulder, Colorado 80309

Phone: 303-492-5519

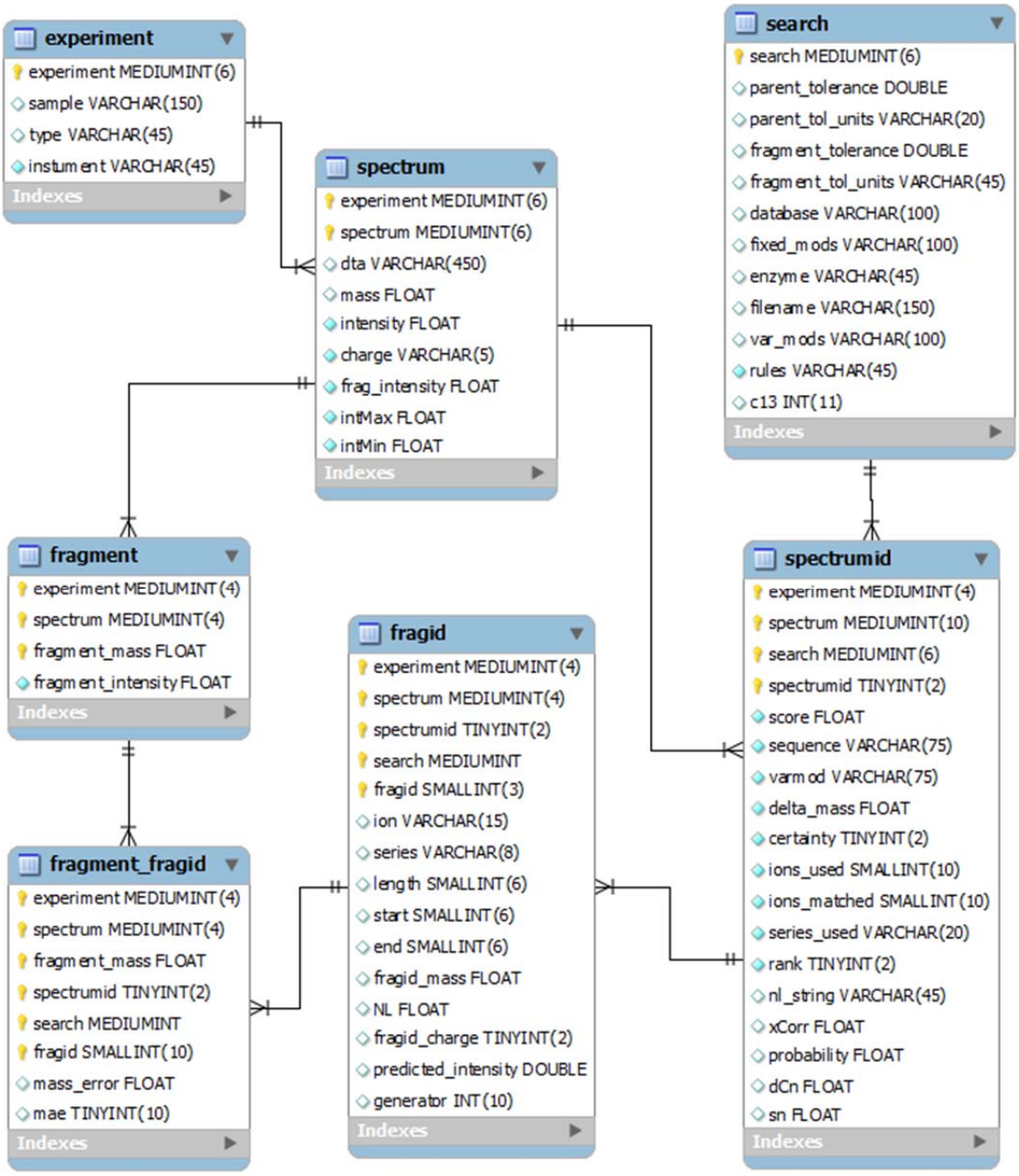
Fax: 303-492-2439

Email: william.old@colorado.edu



Sequences are listed N-terminal to C-terminal, vertical groups indicate a mixed pot at that site.
Lower case s, t, or y denote a phosphorylated residue

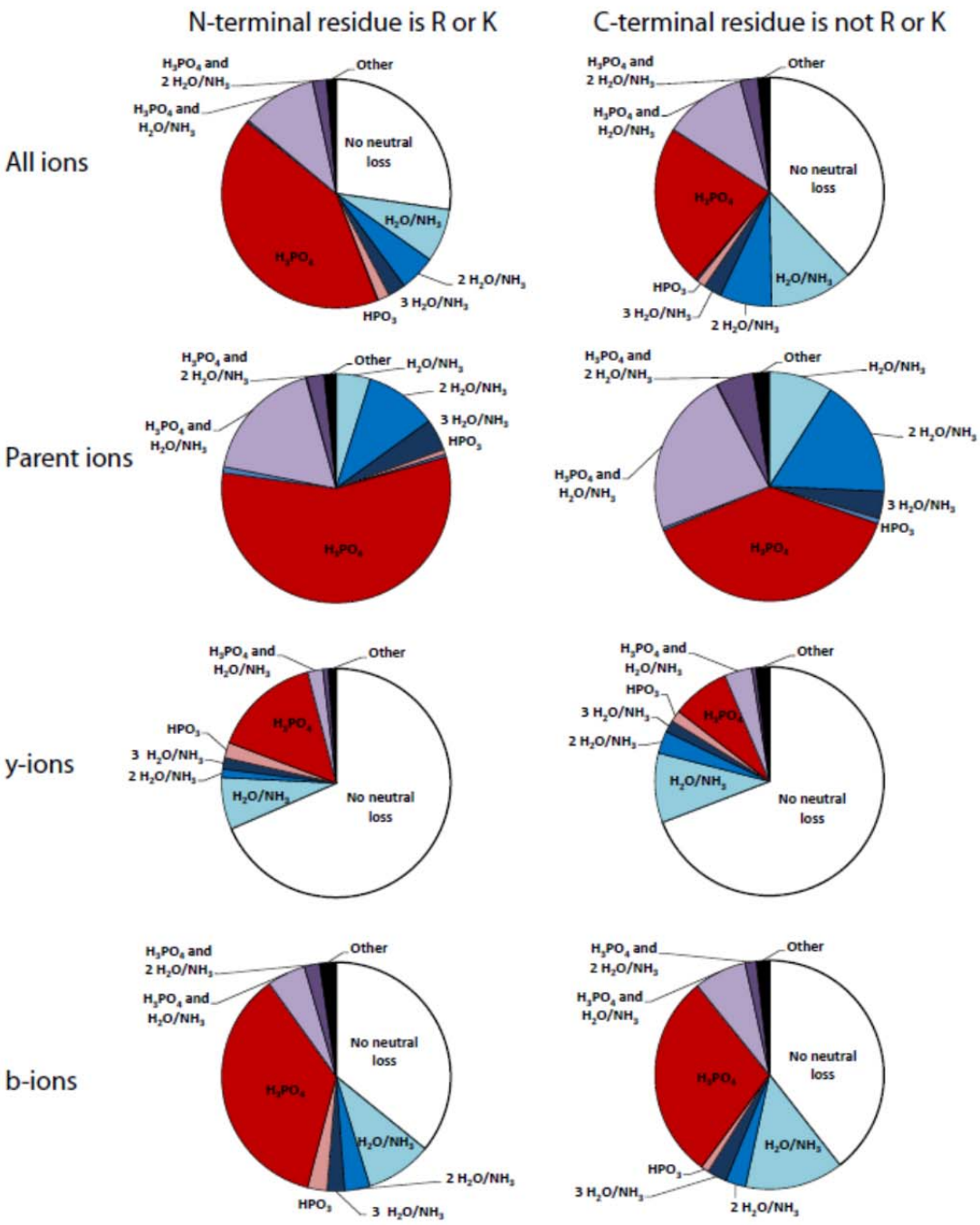
Supplemental Figure 1: The sequences of the combinatorial libraries used in this paper. Peptides are written N-terminal to C-terminal. Multiple amino acids in a column indicates that a mixture of the indicated amino acids was introduced at that site. Phosphorylated residues are indicated in red text and lower case.



Supplemental Figure 2:

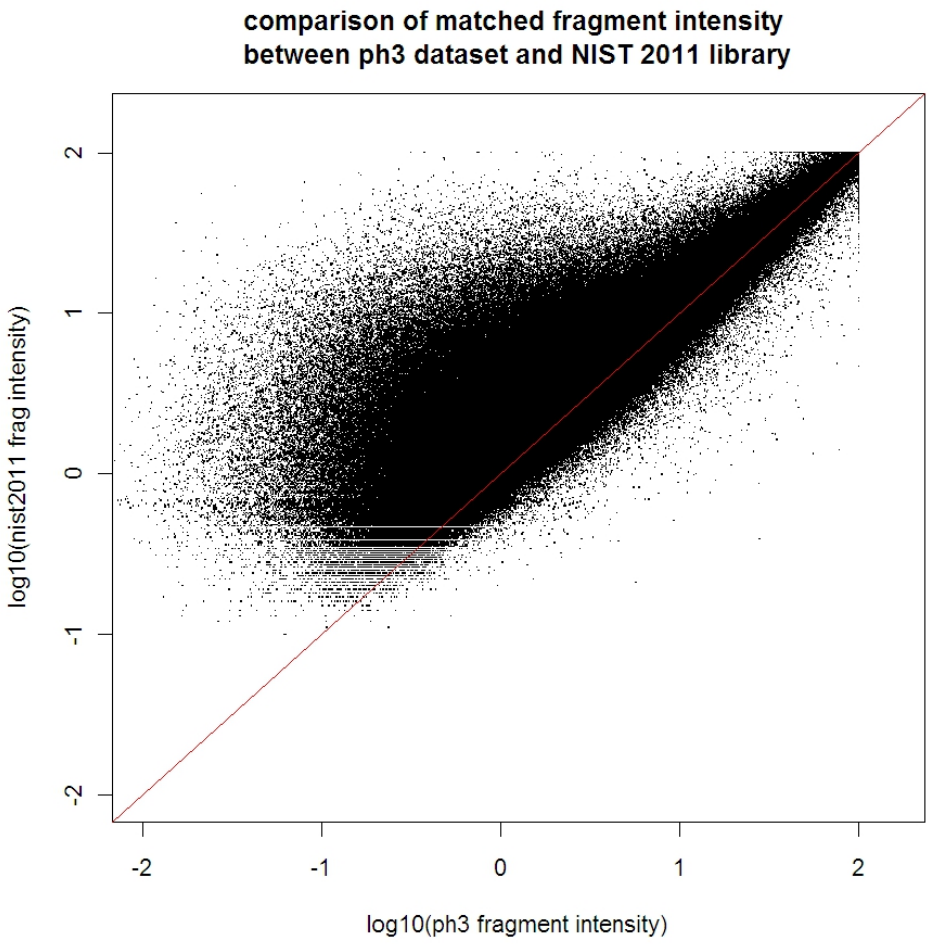
Schema of the relational database used for the storage of spectral information and annotations. The *experiment* table holds metadata about the sample used. *Spectrum* and *fragment* contain the information from MS1 and MS2 scans, respecitely. *Search* contains the parameters of the

Mascot search. *Spectrumid* holds the information about a particular identification, while *fragid* holds the annotations for individual fragments in the spectrum. *Fragment_fragid* exists to allow a many-to-many join of fragments and their annotations along with metadata for that annotation.

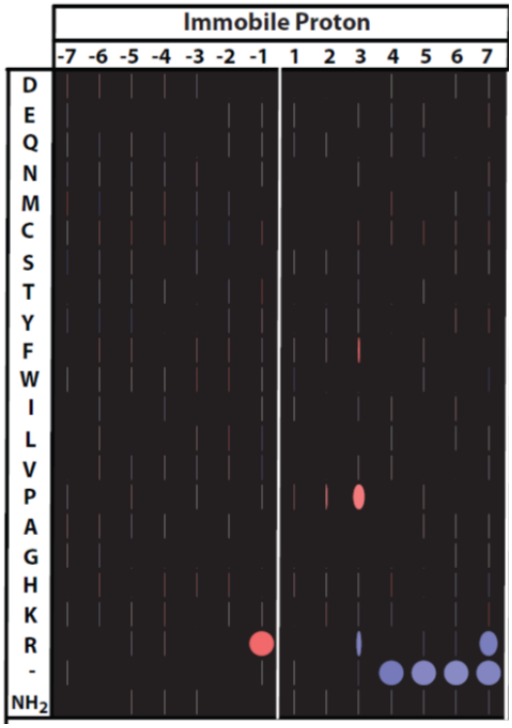


Supplemental Figure 3: The proportion of signal intensity representing each neutral loss state, as was shown in Figure 1 a-d, but in peptides that differ from typical tryptic peptides. The left column contains peptides with either R or K as the N-terminal amino acid, these spectra almost

all contain an additional R or K at the C-terminal amino acid. The right coumn represents spectra with non-tryptic C-termini.

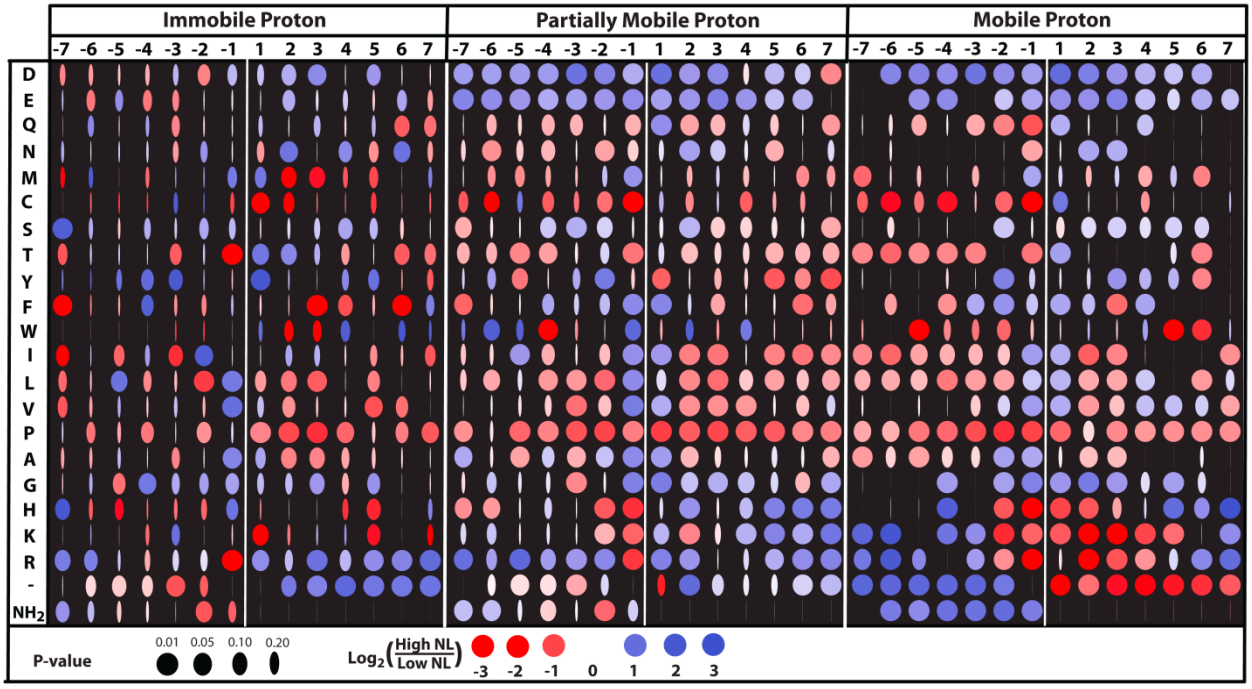


Supplementary Figure 4: Comparison of the observed intensity of all matched ions from the NIST 2011 library (y-axis) and locally obtained unprocessed spectra, ph3 (axis) For this analysis the intensity of the highest peak in the spectrum is defined as 100. This demonstrates that the ratio of intensities between the highest and lowest ions is significantly suppressed in the library spectra compared to spectra obtained locally.



Supplemental Figure 5:

The effects of flanking residue on immobile proton spectra. The data is presented exactly as in Figure 2. The low number of observed immobile cases leads to very few significant results.



Supplemental Figure 6: The effects of flanking sequence on the amount of neutral loss observed.

The same as Figure 2 except that no multiple hypothesis correction has been applied.

Position	IMMOBILE	MOBILE	PARTMOBILE	ALL
-7	0.0206	0.0326	0.0034	0.0243
-6	0.0093	0.0355	0.0055	0.0248
-5	0.0085	0.0316	0.0066	0.0236
-4	0.0144	0.0342	0.0046	0.0255
-3	0.0193	0.0387	0.0131	0.0384
-2	0.0165	0.0337	0.0174	0.0447
-1	0.0516	0.0334	0.0322	0.0153
1	0.0251	0.0233	0.0374	0.0267
2	0.0212	0.0291	0.0161	0.0146
3	0.0343	0.0313	0.0196	0.0162
4	0.0217	0.0277	0.0075	0.0109
5	0.0200	0.0274	0.0100	0.0131
6	0.0253	0.0248	0.0073	0.0119
7	0.0282	0.0164	0.0130	0.0143

Supplemental Table 1: The fraction of variance accounted for by knowledge of the amino acid at a given position for all peptides and stratified by charge mobility. With the fraction of variance explained defined as the sample variance divided by the weighted average of the variances of the stratified subclasses divided by the sample variance.