ABRF PRG 2008 Samples A&B
Results for ID# 12303

Each of the samples was solubilized in 100uL of 100mM ammonium bicarbonate buffer. Half of each sample was placed in –35C storage and the other half was reduced/alkylated to remove disulfide linkages. Reduced/alkylated samples were then digested with 250ng endoprotease Lys-C at 37C for 5hrs. Samples were frozen at –35C after digestion.

A 7% aliquot of each of the digested samples was analyzed via C18 RP-HPLC elution into a Thermo Fisher LTQ-FTMS mass spectrometer. MS/MS spectra of peptides were acquired using data-dependent precursor ion selection. Full MS data was acquired in FTMS mode, which enabled assignment of precursor charge state and accurate mass within 3ppm.

Data was searched using the OMSSA search algorithm against the human nr protein database. Variable modifications were: oxidized methionine and carbamidomethylated cysteine. Endoprotease Lys-C was specified and two missed cleavages were allowed. Results of the database search indicated the protein in each sample was some isoform of the receptor for advanced glycosylation end products (RAGE) protein. After inspecting theoretical Lys-C digests for several of the isoforms, and noting the presence of several highly charged species in the FTMS data, it was clear that ETD analysis would yield more peptide identifications than CAD alone.

A 10% aliquot of each of the digested Sample A and Sample B was analyzed via C18 RP-HPLC elution into a Thermo Fisher LTQ equipped for ETD MS/MS. Both ETD and CAD MS/MS of peptides were acquired in a data-dependent manner. Again, data was searched using the OMSSA search algorithm against the human nr protein database. Variable modifications were: oxidized methionine and carbamidomethylated cysteine. Endoprotease Lys-C was specified and two missed cleavages were allowed. Although the number of identified peptides doubled due to better quality MS/MS spectra from ETD of higher charged peptides, several abundant peptides remained unidentified. Two of these peptides were manually sequenced. One, EQTRRHPETGLF, was unique to Sample A. Another, SNYRVRVYQIPGKPEIVDSASE, was unique to Sample B. Longer versions of each of these peptides were present in the other sample, so they could not be declared as qualitative differences between the two samples. It appeared that each of these peptides resulted from non-specific cleavages, so ETD/CAD data were re-searched using the No Enzyme parameter. The new data search resulted in even more peptide identifications from the RAGE protein and several isoforms were eliminated as possible sources for Sample A due to identification of some peptides on the C-terminal end of the protein. Sample B, however, showed very little C-terminal peptides so only two isoforms could be eliminated as possible sources of the protein.

The N-terminus of each of samples A&B was determined by de novo sequencing of an ETD spectrum of an abundant peptide that had remained unidentified after database searching. The resulting peptide was GSHMAQNITARIGEPLVK, which contained the GSHM tag specified in the sample literature but indicated that the RAGE_HUMAN

protein had been expressed without the signal peptide (residues 1-22). The first amino acid from the RAGE protein in each of Samples A&B is residue 23 from each of the isoforms. Subsequent database searches were against a user-created database that included the biologically expressed N-terminus and 12 isoforms of the RAGE_HUMAN protein.

Several abundant peptides from the ETD/CAD analysis remained unidentified, so de novo sequencing was again utilized. Manual inspection of ETD spectra showed that residue Y150 was iodinated (in both Sample A and Sample B). Fragment ions were confirmed for the N-terminal and C-terminal ends of the VGTCVSEGSYPAGTLSWHLDGKPLVPNEK peptide but the stretch beginning with YPAG was modified by an additional 126 amu. Accurate mass information from the FTMS data showed the mass difference to be 125.9156, which corresponds to iodination and a loss of proton. A second MS/MS spectrum indicated an additional iodination on Y150. A new database search of all data was performed, with iodination of Y included as a variable modification. The results identified two other tyrosines as being iodinated: both tyrosines in the SNYRVRVYQIPGKPEIVDSASELTAGVPNK peptide. All iodinations were present in both samples A&B.

Amino acid sequence coverage of Sample A indicated that its identity was one of the following five isoforms of the RAGE protein: Q15109.1, AAA03574.1, AAX07274.1, AAX07273.1, AAX07275.1 No evidence of the intron 4 (VVEESRRSRKRPCEQE) or spice variant 1 (EGFDKVREAEDSPQHM) was found, which indicates that Q15109.1 or AAA03574.1 (differing by the M on the N-terminus; truncated in our samples) is the "true" Sample A but lack of MS evidence does not prove that the variant peptides are not present. All five isoforms are still possibilities for Sample A. Lack of C-terminal peptide information for Sample B indicates that it is most likely an isoform that is truncated on the C-terminus but, again, it cannot be proven conclusively by these data. The MS/MS data were queried to look for signature fragment ions specific for the C-terminal residues of each of the isoforms, but they were not observed.

Next steps for Samples A and B would be top-down analysis, using the 50% of original sample not subjected to proteolysis, to obtain intact-protein mass information. This would confirm the supposition that Sample B is a C-terminally truncated isoform and would indicate the presence/absence of intron or splice variants for Sample A.