

Figure 3. Theoretical analysis of N-terminal positional peptides from mouse

A set of 8,000 *Mus musculus* protein sequences were randomly extracted from the Swiss-Prot database and the trypsin (arginine selective) and endopeptidase Glu-C N-terminal peptides were extracted, after processing to remove the signal peptide and if appropriate, the N-terminal initiator methionine residue. The N-terminal peptides were then filtered to remove those smaller than 500 Da and larger than 5,000 Da to yield a realistic subset of analysable peptides. The utility of N-terminal peptide mass determination alone, recovered at different mass accuracies, was assessed for the sets of peptides generated by either endopeptidase, or for a combination study where the endopeptidases were used in a pair of parallel digestions. The identifiability of the protein set is expressed as a percentage, relative to the 8,000 mouse proteins initially extracted from the Swissprot database.

