Supplemental Information

Database searches

Peak lists from MS and MS/MS data were created using Mascot Distiller (MatrixScience, version 1.1) and used to interrogate the NCBInr 20060817 protein database (3879234 sequences; 1336925637 residues) restricted to mammalian entries (486696 sequences). We used Mascot Deamon (version 2.0) for batch searches and Mascot (MatrixScience, version 2.0) as the search engine. Specified mass accuracy for parent and fragment ions was 100 ppm and 150 ppm, respectively. For MS/MS data, proteins were considered identified when the addition of non-redundant peptide hits reached statistically significant values (i.e., Mascot scores were greater than 36) and with at least 2 matched peptides. Identifications based on single peptide matches were not included in the final report. We chose these parameters because they are restrictive and result in a low rate of false positive identifications. For peptide mass fingerprinting, threshold mass accuracy for searches was 20 ppm and a minimum coverage of 20% was used for protein identification. When the set of peptides matched more than one entry, we selected the entry with more descriptive protein name was used for tabulation. No attempt was made to distinguish between different members of proteins families and the entries reported should be taken as to indicate that the identification may be to any member of the protein family.