

TOP DOWN MASS SPECTROMETRY OF ARCHAEA
AND HUMAN TUMOR CELLS

Andrew J. Forbes, Ph.D.
Department of Chemistry
University of Illinois at Urbana-Champaign, 2004
Neil L. Kelleher, Adviser

Final Defense
Aug. 10, 2004

This work describes continued work in the realm of “top down” mass spectrometry and its application to proteomic scale experiments. The work centered around efficient characterization of intact proteins as seen in their native state as well as how these strategies could be applied to large proteins.

Large, overexpressed proteins of PchE (159 kDa) and PchF (199 kDa) were used as test cases for the characterization of proteins over >70 kDa. Digestion via Lys-C yielded 6 and 10 proteolytic peptides, respectively, and the resulting peptide mixtures were analyzed on a 4.7 T ESI-FTMS instrument. Three of the 10 peptides for PchF were detected, and all six PchE peptides were detected. Reconstruction of the PchE protein yielded an intact mass +3 Da from the predicted mass (19 ppm).

A general proteomic analysis platform was applied to the soluble fraction of the *Methanococcus jannaschii* proteome. Proteins were fractionated into mixtures of 2-10 proteins via Prep Cell/RPLC separations, and automated MS and tandem MS was performed on ~100 samples. Seventy-two proteins in all were identified and fully characterized through this method; very few post-translational modifications were unveiled, including none on any histones. Use of the separations platform in a targeted fashion rapidly isolated histones from the organism *Methanosarcina acetivorans*, and the subsequent MS data showed that they also were unmodified despite the annotation of histone modifying genes in its genome.

A final study was performed applying the top down platform to human HeLa tumor cells. Thirty-four proteins in all were identified, and >90% of these were fully characterized. Several proteins harbored post-translational modifications including N-terminal and internal acetylations, methylations, disulfide bonds, and protein truncations. Proteins were also characterized with known coding single nucleotide polymorphisms, alternative splice variants, and highly homologous gene family members.