Small HCPs in a 12 kDa Protein Drug Analyzed by GeLC-MS/MS

Rikke Raaben Lund1, Marie Grimstrup1, Maheen Sayeed2, Karin Abara Heidemann2, Ejvind Mørtz1
1Alphalyse A/S, Odense, Denmark, 2Rockland Immunochemicals Inc., Limerick, PA, USA

Introduction
The content of low molecular weight host cell proteins (HCPs) in purified protein drugs is often difficult to evaluate, due to their low immunogenicity and poor ability to be visualized in gel-based total protein stains. The proteome of commonly used expression organisms, such as E. coli and Chinese Hamster cells, contains 30-40% proteins with a molecular weight below 20kDa, and these are easily missed in both gel separations, Western blots and ELISA quantitation of the total HCP-content. To provide unbiased analysis of small as well as larger HCPs, we introduce the use of a mass spectrometry-based orthogonal method, well known from proteomics, called GeLC-MS/MS. Here, we analyze an in process protein drug, a 12 kDa protein produced in E. coli, as well as the corresponding null cell lysate, using 1D-PAGE and nano-flow LC-MS/MS (GeLC-MS/MS) to achieve high coverage of small HCPs.

1D-PAGE
Proteins were separated by 1D-PAGE and stained for isolation and visualization of the high concentration protein drug in separate gel fractions. Most proteins, 78% in the null cell lysate and 82% in the process sample were only identified in a single gel fraction (light green and blue bars). Further, the fractionation leads to a high number of protein identifications in the fractions without the protein drug (fig 2).

Protein separation by 1D-PAGE prior to LC-MS/MS leads to identification of a high percentage of small HCPs: 46% of the HCPs identified in the null cell lysate and 37% in the in process sample were smaller than 20kDa (fig. 3, table 1 and 2). Further, the HCPs that were identified covered the entire pl range and 99% of the molecular weight range of the entire E. coli proteome (fig. 3). Examples of small HCPs is given in table 3 and 4.

The sensitivity of the method was estimated by parallel analysis of the purified protein drug with two spiked-in standards, these standards were both identified at 50ppm.

Nano Flow LC-MS/MS
The proteins contained in the gel fractions were digested enzymatically into peptides using trypsin. Separation of the complex peptide mixture was achieved by nanoflow HPLC using a CSH (charged surface hybrid, Waters) column. This material has a high loading capacity and excellent peak shape in formic acid mobile phases. This enables highly sensitivity and accurate MS detection of low level HCPs using a qTOF mass spectrometer (Bruker Corp. fig. 4 and 5).

GeLC-MS/MS advantages
• 1D-PAGE requires very little sample preparation and has a wide mass and pl range, leading to unbiased sample analysis
• Protein digest combined with nano flow LC-MS/MS provides sensitive and accurate measurement of multiple peptides from each protein as well as sensitive and accurate measurement of peptide fragment masses
• Post translational modifications and processing
• Identification of the protein names and database accession numbers enables an in-depth analysis of each individual HCP present in the drug sample. Important features for a drug risk assessment include:
  • Immunological properties and presence of human B- and T-cell epitopes
  • Homology to human proteins with important biological function
  • Homology to the drug protein
  • Enzymatic activity to modify or cleave the drug product constituents
  • Hormone or hormone-like activity

Known information can be found at http://www.uniprot.org including:
• Molecular function, biological process, ligand binding and cellular component
• Post translational modifications and processing
• Expression, interaction and structure
• Protein family and domains
• Sequence variations
• Publications

Prediction of protein features can be investigated at http://www.cbs.dtu.dk/services including (fig. 5):
• Immunological features
• Post translational modifications
• Protein structure and function

Conclusion
The HCPs that were identified covered 99% of the entire E. coli proteome in terms of molecular weight and pl. This shows that the developed GeLC-MS/MS method has no inherent limitations with respect to pl or molecular weight for HCP identification. The obtained protein identity enables an in-depth analysis of each individual HCP and a more detailed risk assessment.