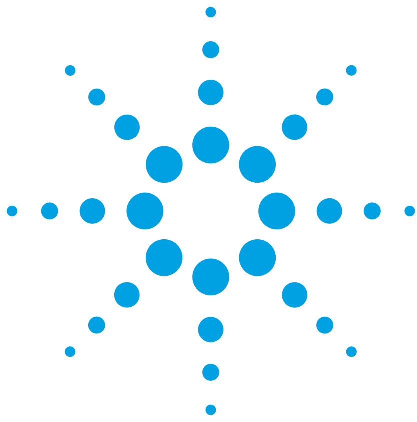

Characterization of bacteriophage derived anti-staphylococcal protein (P128) from production to purification using Agilent HPLC-Chip Q-TOF LC/MS system

Application Note

Authors

Ravindra Gudihal
Agilent Technologies India Pvt. Ltd.
Bangalore India

Jiya Asrani, Madhavi H.N,
Uma Maheshwari, P.M.Sundaram
GangaGen Biotechnologies Pvt. Ltd,
Bangalore, India

Abstract

The analysis of in-process samples of recombinant protein production using an Agilent HPLC-Chip system coupled to an Agilent Accurate-Mass Q-TOF LC/MS system is described. The superior mass accuracy of the HPLC-Chip LC/MS platform, combined with the powerful data processing capabilities of Agilent MassHunter and BioConfirm software, enabled identification of impurities during the production process. This type of analysis aids in the assessment of impurities found in protein pharmaceutical at the early phases of product development.

Agilent Technologies

Introduction

P128 is a recombinant protein expressed in *E.coli*. This protein has potent bactericidal activity against the methycillin resistant *Staphylococcus aureus* (MRSA) and is currently under early development for use in humans. The staphylococcal phageK ORF56, a tail associated muralytic enzyme was earlier shown to exhibit anti-staphylococcal activity. The active portion was found to reside in the C-terminal 16kDa region by deletion analysis of ORF56. To improve the specific activity of ORF56 and to develop it into a potent anti-staphylococcal molecule for therapy, the muralytic activity of 16kDa ORF56 was combined with the bacterial cell wall binding domain. The chimeric protein termed P128 displayed significantly higher activity than ORF56 and was found to be highly potent against drug resistant *Staphylococcus*. The P128 was also found to be active in body fluids indicating its potential use as both topical and systemic antibacterial. P128 is a staphylococcus peptidoglycan hydrolase with an unknown site of action and assumed to cleave the peptidoglycan cross bridges at Gly-Gly linkages.

LC/MS technology is a powerful and sensitive technique for the characterization and identification of proteins unlike the traditional SDS PAGE analysis which gives a less specific assessment of the samples being investigated. SDS-PAGE analyses lack the resolution to separate similar MW proteins, complex protein mixtures and may miss the small molecule impurities in the in-process samples. Detection of impurities also depends on the method of staining used for the SDS-PAGE, for example, some additives cannot be detected.

Nanoflow LC/MS enables rapid and more sensitive protein identification while minimizing sample and solvent consumption. Agilent's microfluidic based HPLC-Chip integrates sample preparation, chromatographic separation, and nanoelectrospray formation for efficient, high-sensitivity nanospray LC/MS. Agilent Accurate-Mass LC/MS TOF and Q-TOF systems deliver exceptional mass resolution, mass accuracy, sensitivity, and data processing capabilities for optimal MS characterization of proteins.

The HPLC-Chip provides improved run-to-run reproducibility which is important in the QC and in-process sample analysis. A portfolio of HPLC-Chips such as C_8, $C_{18,}$ PGC-chips (graphitized carbon chips) and phospho-chips gives additional advantage as a plug-and-play system to do multiple analyses of the biologics on the same instrument resulting in time saving in the overall characterization.

A key aspect to the success of analysis by LC/MS is the processing software. Protein electrospray spectra yield many multiply charged ions which need to be deconvoluted into zero charge mass spectra in order to determine the molecular weight. In a potentially complex mixture, separate deconvolution of each component can be very time consuming. A large molecule feature extraction procedure in the MassHunter software allows for this step to be automated and fast.

Chemical modifications and degradation that can potentially occur during manufacturing, formulation, and storage of proteins, necessitate reliable and sensitive methods for the characterization of purity and structural integrity. If sub-optimal characteristics of biologics are discovered after the initiation of clinical trials then serious and costly delays can occur in moving the drug to market.

Materials

In-process and purified protein samples were obtained from GangaGen Biotechnologies Pvt. Ltd.

Recombinant protein purification

The recombinant protein was purified as shown in figure 1. Briefly, the cells containing the protein of interest were lysed by sonication in lysis buffer containing Triton X-100, the host cell DNA was precipitated using polyethyleneimine(PEI) and then the lysate was subjected to different concentration of ammonium salt precipitation/fractionation starting from 20% to 50%. The 50% ammonium sulphate precipated protein pellet was dialysed and then subjected to series of ion exchange chromatography steps to yield the final purified protein P128.

Instrumentation

The Agilent 1200 Series HPLC-Chip/MS Interface was coupled with the Agilent 6520 Accurate-Mass Q-TOF LC/MS System for LC/MS analyses.

LC parameters

HPLC-Chip: 5 µm, ZORBAX 300SB-C8 (300Å), 40 nL enrichment column, and a 75 mm x 43 mm analytical column.

Flow rate: 3 µL/min from capillary pump to the enrichment column and 600 nL/min from nano LC pump to the analytical column.

Solvents: 0.1% formic acid in water (A); 90% acetonitrile in water with 0.1% formic acid (B).

Sample Loading: With an Agilent 1200 Series Capillary Pump at 3% B.

Amount of sample injected onto the chip: 200 ng of the in-process samples.

Sample analysis: Gradient with an Agilent 1200 Series Nanoflow LC pump as shown below.

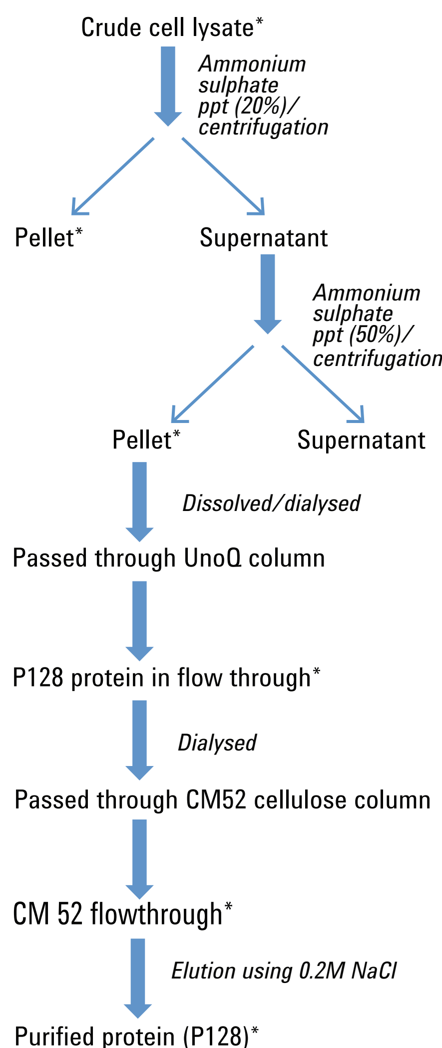

Figure 1. Purification scheme of P128 protein from crude cell lysate (used for LC/MS analysis).*

Time (min)	B (%)
Initial	3
40.00	60
45.00	95
46.00	95
47	3

Stop time: 50 min

MS parameters

Spectra were recorded in positive ion mode.

Vcap: 1900 V and drying gas flow of 5 L/min at 350°C was used.

Fragmentor voltage: 300 V

Data were acquired at 2 GHz (extended dynamic); MS only mode, range 300–3,200 *m/z.* An internal mass calibration was used during the LC/MS runs. This internal reference mass system allowed accurate, and automated mass calibration correction during the LC/MS runs.

Data analysis

The data obtained from LC/MS and MS/MS were analyzed using features contained in the following software packages: Agilent MassHunter Qualitative Analysis, Agilent MassHunter BioConfirm, and Agilent DA reprocessor. The raw data (chromatograms) were processed using the Large Molecular Feature Extractor (LMFE) algorithm within Agilent MassHunter Qualitative Analysis software. The purified protein was also deconvoluted using Maximum Entropy algorithm and matched to the target protein sequence using BioConfirm.

Results and Discussion

In-process samples at different stages of P128 purification (labeled * in Figure 1) were analyzed on the HPLC-Chip Q-TOF LC/MS system. The total ion chromatograms of these samples are shown in Figure 2. The broad peak eluting around 27-30 min corresponds to polyethyleneimine (PEI) or Triton X-100 which was used in the crude cell lysate. The SDS-PAGE analysis did not reveal such details, showing the importance and specificity of LC/MS analysis (data not shown). As further steps of purification were performed, most of the unwanted proteins as well as other small molecules were removed yielding the enriched purified protein.

LMFE was used to determine the number of components at each step in the purification[1]. The LMFE algorithm is specifically designed for the analysis of complex mixtures of intact proteins or large oligonucleotides. This feature extraction finds all the peaks in an LC/MS run and then groups them according to similar retention times and elution profiles into co-elution groups.

The co-elution group incorporates the different charge state of the same protein grouped together by algebraic charge state deconvolution. LMFE produces a list of compounds within the MassHunter Qualitative Analysis software with links to a compound spectrum containing the different charge states found for a given protein and the extracted compound chromatograms for each compound.

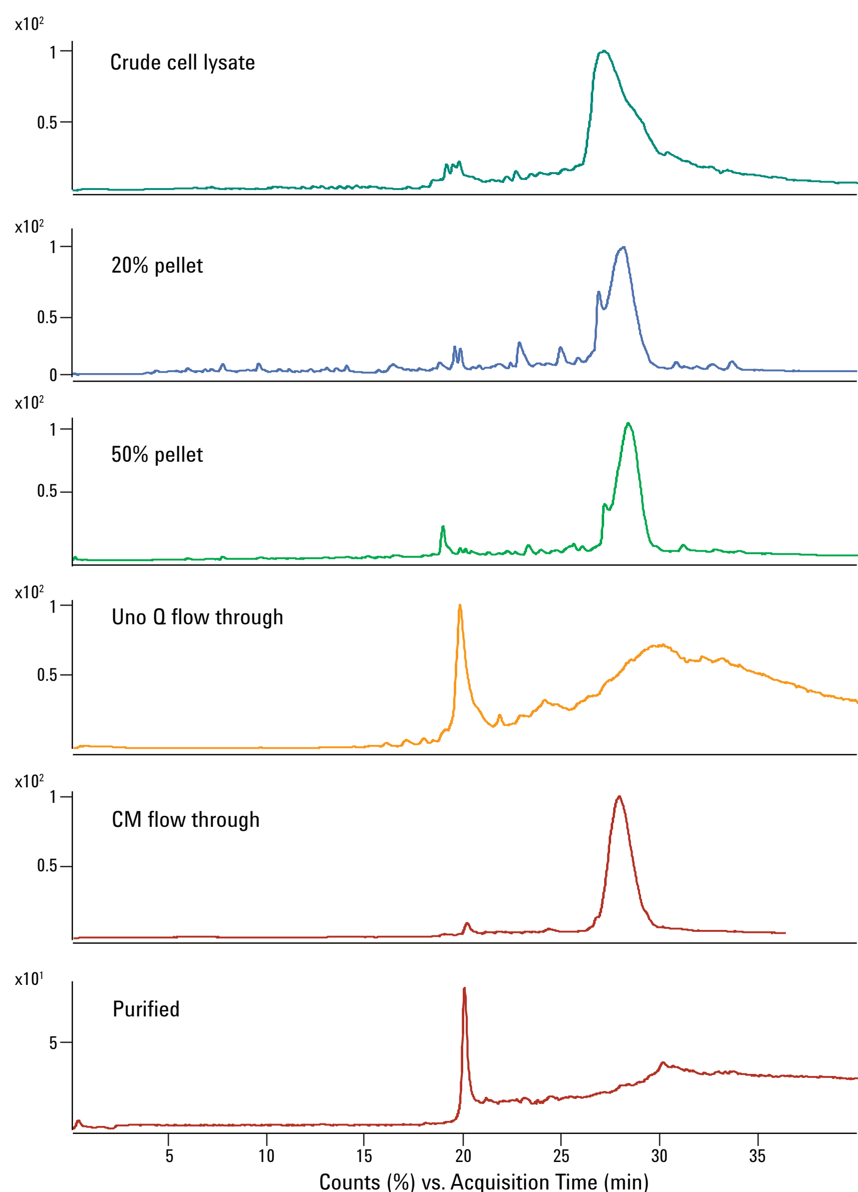

Figure 2. Total ion chromatogram for the different in-process samples.

Classic maximum entropy deconvolution took 45 minutes to integrate and extract peak spectra, deconvolute each spectrum to yield 112 proteins in the crude cell lysate while LMFE took 10 minutes to yield 256 proteins. The LMFE algorithm was approximately three times faster than maximum entropy deconvolution and identified two times the number of proteins in this example. Data processing with the LMFE algorithm proceeded with exceptional speed and yielded substantially more information for complex intact protein mixtures than compared to maximum entropy deconvolution, generally used for analysis of single or simpler protein mixtures[2].

The results obtained from using LMFE are tabulated in Table 1.

PEI and Triton X-100 treated and untreated crude cell lysates were analyzed. The peak eluting at 27-30 minutes could either be due to either PEI or Triton X-100. Therefore in order to compare the origin of peaks at 27-30 minutes, PEI and Triton X-100 were run independently. The spectrum obtained from Triton X-100 matched with the peak at 27-30 min (the error in mass measurement was 4.9 ppm) suggesting that the samples contained Triton X-100 during the early steps of purification, figure 3.

In-process sample	Number of compounds*
Crude cell lysate	256
20% pellet	199
50% pellet	113
Uno Q flowthrough	151
CM flowthrough	60
Purified	6

Table 1. Number of compounds at different stages of purification identified using LMFE
**Average of 4 replicates*

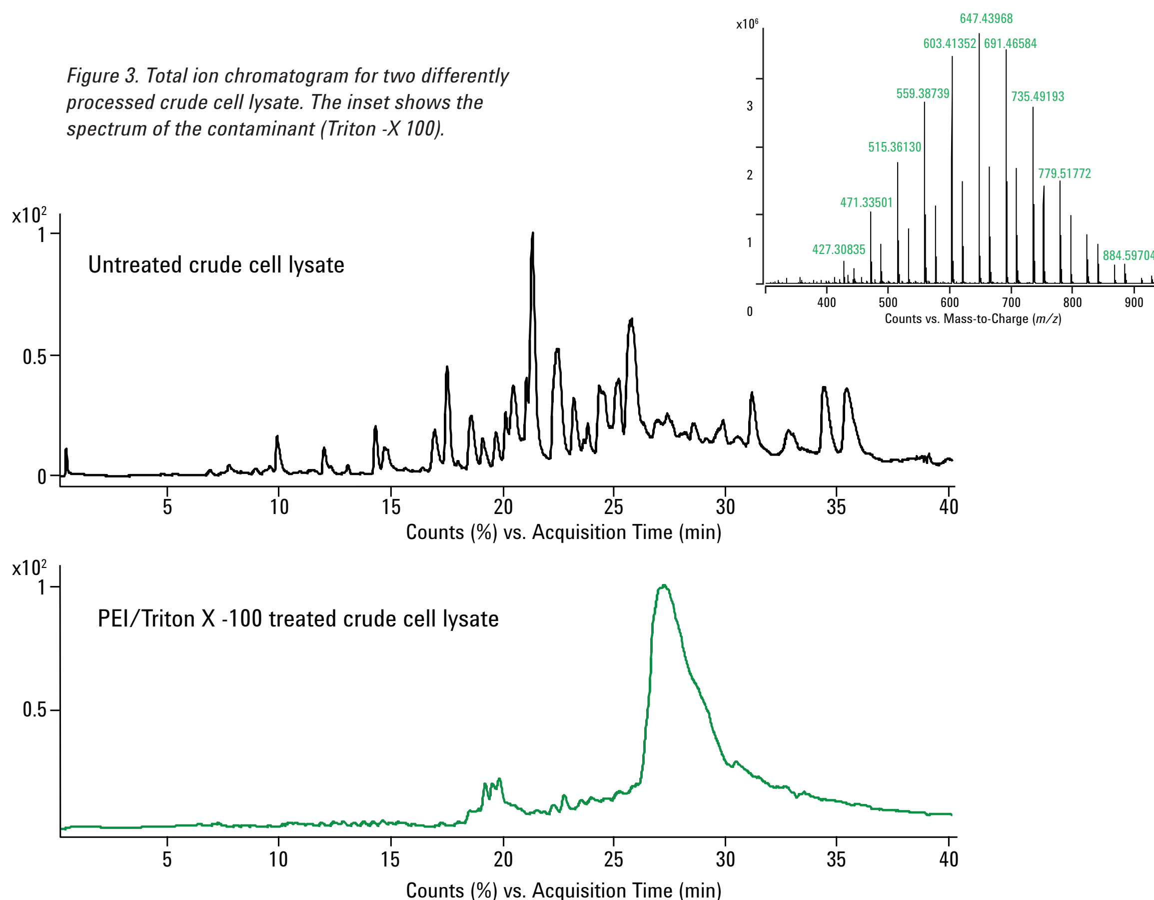

Figure 3. Total ion chromatogram for two differently processed crude cell lysate. The inset shows the spectrum of the contaminant (Triton -X 100).

The measured intact molecular weight of the purified P128 was found to be 26489.2 Da, with an error in mass measurement of 4ppm from the theoretical mass of 26489.13 Da.

Upon inspection of the mass spectrum of the purified protein, very minor additional peaks corresponding to oxidation/hydroxylation were found, figure 4.

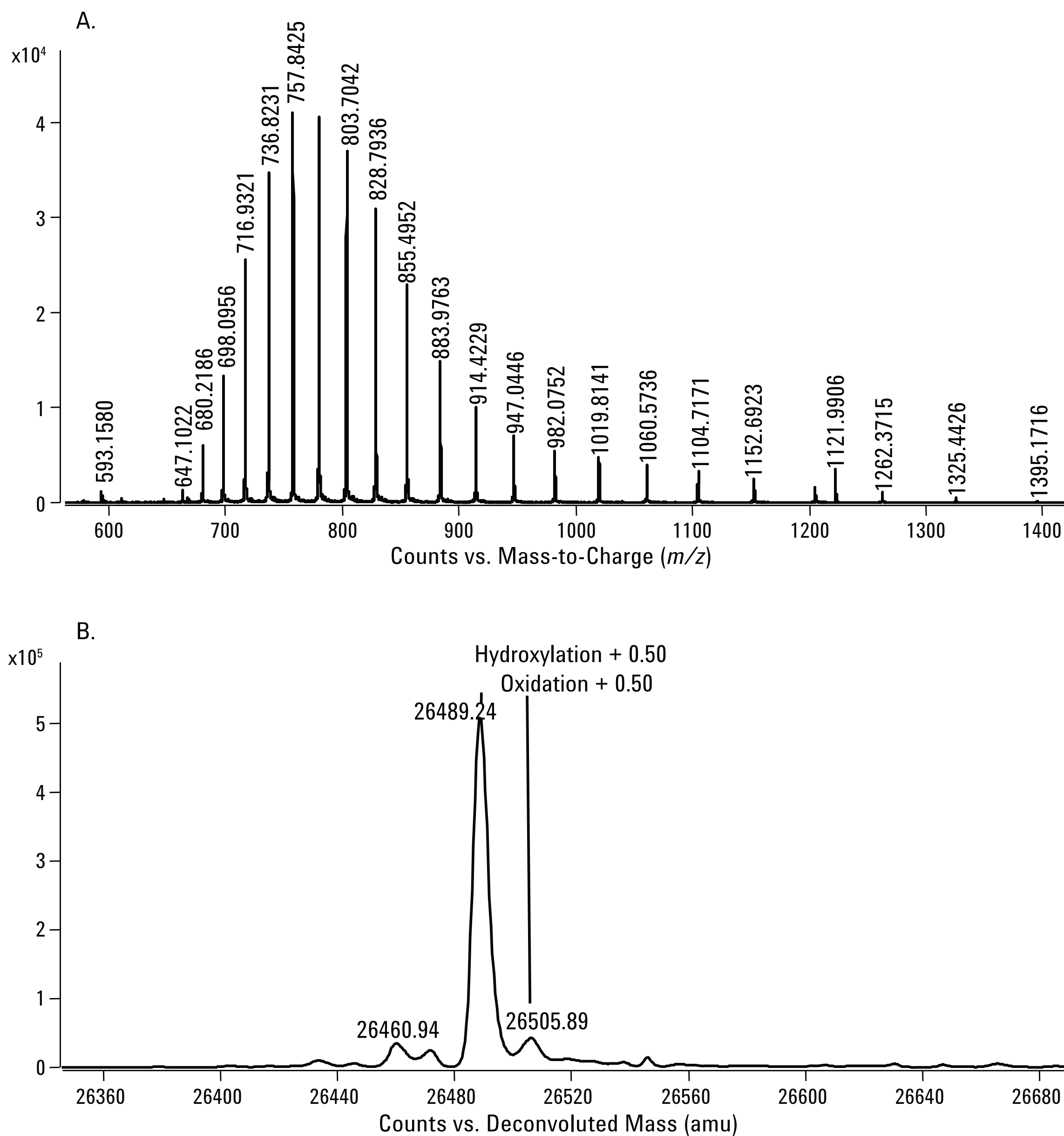

Figure 4. Mass spectrum (A) and Deconvoluted Spectrum (B) of P128

Conclusions

- Micro-fluidic HPLC-Chip Q-TOF LC/MS analysis provides a very convenient and detailed way of studying in-process samples as well as for characterization of purified proteins. Additives such as the polyethyleneimine/Triton X-100 added during purification were detected and identified which were not seen by SDS-PAGE analysis.
- Large Molecule Feature Extractor (LMFE) algorithm enables faster analysis and more information on the complex nature of samples than classic maximum entropy deconvolution. Hence, by providing a quick analysis of the protein biologics, LMFE can save valuable time during the production process.

References

1. Faster, More Accurate Characterization of Proteins and Peptides with Agilent MassHunter BioConfirm Software (5990-5096EN).
2. Characterization of Bacteriophage Derived Anti-staphylococcal Protein (P128) from Production to Purification Using Microfluidic based LC System Coupled to an Advanced QTOF-MS. Gudihal Ravindra *et al*., (poster at ASMS 2010)

www.agilent.com

This information is subject to change without notice.

© Agilent Technologies, Inc., 2011
Published in USA, May 16, 2011
Publication Number 5990-7952EN

Agilent Technologies