2014

INNOMOL

Innovation Pipeline

Proteomics

INNOMOL WORKSHOP

1. RATIONALE AND OBJECTIVES

1.1. RATIONALE

The proteomics is scientific discipline of characterizing and analysing the proteome. The term proteome is described as entire protein complement expressed by the genome in the lifetime of a given cell, tissue or organism, including isoforms, polymorphisms and modifications, protein-protein interactions and the structural description of proteins. Proteomics encompasses the identification, characterisation and quantification of the proteome.

The dominant proteome analysis workflow utilizes few important steps: protein extraction from the cells, protein/peptide separation, site-specific protein enzymatic digestion and mass spectrometry (MS)/bioinformatics analysis.

The extraction of proteins from the cells is a critical step for establishing a stable proteome and this is important experimental step in proteomics studies. The extraction of soluble proteins is simply performed by lysing the cells and collecting the supernatant. Different cell lysis methods are easily accessible in literature data.

Protein separation by two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) provides high-resolution separation of complex protein mixtures. 2-D PAGE can potentially separate several thousand proteins in a single experiment (on average, between 5000 and 15000 per cell type). During 2-D PAGE the proteins are firstly separated according to their pI (first dimension), followed by their separation according to their molecular weight (second dimension). Alternatively, a mixture of proteins or peptides can be separated by high performance liquid chromatography (HPLC). HPLC separation is based on the differences in interactions of compounds (analytes) between two phases, mobile (liquid) and stationary, as the compounds travel through column under a high pressure. In reverse-phase HPLC mode separation is performed according to hydrophobicity of analytes (proteins/peptides). The HPLC systems are capable to detect proteins of very low abundance and allow the analysis of complex protein/peptide mixtures in a single run. When HPLC separation method ends with mass spectrometry/bioinformatics analysis, the method is known as Shot-gun proteomics.

On the other hand, protein spots separated by 2-D PAGE are excised from the gels, properly rinsed, and in-gel digested, while in HPLC separation the proteins are usually digested in-solution and then separated. Both methods (2-D PAGE and HPLC) in conjunction with MS analysis and followed by peptide mass database search represent very powerful tools for protein identification and quantification.

MS is one of the most sensitive methods for the structural characterisation and protein identification. MS technique measures the mass to charge ratios (*m/z*) of gas-phase ions with extremely high sensitivity and resolution. The most common ionisation techniques are electrospray ionisation (ESI) and matrix-assisted laser desorption ionisation (MALDI), which allow the transfer of large, polar, thermally labile biomolecules into the gaseous phase for mass

analysis (e.g. peptides). Further, peptides can be fragmented in the mass spectrometer by tandem mass spectrometry (MS/MS), and the resulting MS/MS spectra are usually used to retrieve the corresponding peptide sequence from the database. Protein identification is accomplished by matching the list of observed *m/z* in MS and MS/MS spectra with a calculated list of all the expected peptide masses for each entry in the protein database.

Proteomics can be viewed as an array of biological or clinical assays and have a great impact on biology and medicine. Proteomics technology is applied to various standard biochemical and molecular biology studies: biomarker discovery, protein/protein interactions, biotypization, cell functioning, proteins or “small molecules” mechanism of action and so on.

1.2. OBJECTIVES

The objectives of this Innomol FP7-REGPOT-2012-2013 - Proteomics workshop course include:

- understanding of the fundamentals of mass spectrometry, with a focus on different peptides/proteins separation techniques (multidimensional nano- and capillary-liquid chromatography/1-D and 2-D SDS-PAGE),
- cell lysate handling (sample homogenization) and protein/peptide enrichment strategies,
- peptide and protein analysis by MALDI-TOF/TOF and ESI-Q-TOF,
- bioinformatics data-analysis, followed by database search, *de novo* sequencing and high-throughput data processing.

2. MATERIALS AND METHODS

2.1. MATERIALS

2.1.1 LABORATORY EQUIPMENT

2.0 mL PP tubes, Eppendorf, Germany

1.5 mL microcentrifuge tubes, Eppendorf, Germany

Tube rack for 1.5 mL-2.0 mL test tubesTube holder cool pack for 1.5 or 2.0 mL tubes

Vortexer, IKA, USA

Tissue homogenizer (TissueRuptor and TissueRuptor disposable probes), Quiagen, Germany

ZipTip C4, Millipore, USA

Analytical balance Kern, ALS 220-4N, Germany

Thermomixer, Eppendorf AG 22331, Germany

Concentrator 5301 (SpeedVac), Eppendorf, Germany

Ultrasonic bath Sonorex, Bandelin, Germany

Laboratory centrifuge Centric 400, Tehtnica, Slovenia

Centrifuge 5415 R, Eppendorf, Germany

Power supply – Power PAC 3000, BIO-RAD, USA

ReadyStripTM IPG Strip 17 cm, pH 3-10NL, BIO-RAD, USA

Electrode Wicks, Bio-Rad, USA

PROTEAN IEF Cell system, Bio-Rad, USA

PROTEAN II xi 2-D cell, Bio-Rad, USA

LC Column C_{18}, 300 µm x 150 mm I.D., 3.5 µm particle size, Symmetry 300TM, Waters, USA

LC Trap column, 180 µm x 10 mm ID, 3.5 µm particle size, Symmetry C_{18}, Waters, USA

LC Analytical column, 75 µm x 100 mm ID, 1.7 µm particle size, BEH dC_{18}, Waters, USA

Spin filters 0.22 µm, cellulose acetate membranes, Agilent, USA

Extraction manifold SPE system, Waters, USA

SPE 1 mL Diol cartridge, 100 mg Sorbent per Cartridge, Sep-Pak® Diol 1 cc Vac cartridge, Waters, USAF

Household microwave oven, HeatWave compact 800W Electrolux

2.1.2 SCIENTIFIC INSTRUMENTS

BioPhotometer, Eppendorf, Germany

VersaDoc™ 3000 imaging system, Bio-Rad, SAD

Liquid chromatography system: Capillary LC, Waters, USA

Tempo™ LC MALDI Spotting System, Applied Biosystems, MDS Sciex, USA

Mass spectrometer: 4800 MALDI TOF/TOF™ Analyzer, Applied Biosystems, MDS Sciex, USA

Liquid chromatography system: nanoACQUITY UPLC, Waters, USA

Mass spectrometer: SYNAPT G2-Si Mass Spectrometer, Waters, USA

2.1.3 CHEMICALS

Acetonitrile (ACN), Ultra gradient HPLC grade, JT Baker, USA

Ultrapure water, TOC < 5 ppb, resistivity < 18.2 MΩ cm

Ammonium carbonate, Riedel-de Haen, Germany

Ammonium sulfate, 99.5 %, Kemika, Croatia

Ammonium formate, for HPLC, ≥ 99.0% Sigma-Aldrich, USA

Trypsin, Merck, Germany

Trifluoracetic acid (TFA), Merck, Germany

Glacial acetic acid, min. 99.5 %, GRAM-MOL, d.o.o., Croatia

Methanol, MERCK, Germany, gradient grade for LC

Coomassie Brilliant Blue G250, Bio-Rad, USA

Trizma® base, ≥ 99.9 %, Sigma, USA

Hydrochloric acid, (37%; ϕ), p.a., Kemika, Croatia

Glycine, for electrophoresis, Sigma, USA

30% Acrylamide/Bis solution, 29:1 (Acrylamide: N,N´-methylene-bis-acrylamide), Bio-Rad, USA

Sodium dodecyl sulfate, 95 % based on total alkyl sulfate, Sigma, USA

Ammonium persulfate (APS), for electrophoresis ≥98 %, Sigma, USA

N,N,N`,N`-tetramethyl-ethylendiamine (TEMED), for electrophoresis ≥ 98 %, Sigma, USA

Urea, USP testing spec., Sigma, USA

Thiourea, ReagentPlus ≥99 %, Sigma, USA

CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate), Sigma, USA

Iodoacetamide, Sigma, USA

Dithiothreitol (DTT), Sigma, USA

Bromophenol Blue, Sigma, USA

Glycerol, redistilled, Kemika, Croatia

Agarose, Sigma, USA

Phosphoric acid (H_3PO_4), (85 %; ϕ), p.a., Kemika, Croatia

α-cyano-4-hydroxycinnamic acid (CHCA), Fluka, Switzerland

Mineral oil, biotechnology grade, Bio-Rad, USA

DNase I, Thermoscientific, USA

RNase A, Roche, Germany

Protease Inhibitor cocktail tablets, Complete Mini, EDTA-free, Roche, Germany

Triton X-100, Sigma-Aldrich, USA

Mag-Trypsin, trypsin immobilized on magnetic beads, Clontech, USA

Endoproteinase GluC (V8), sequencing grade, Roche, USA

CAF-/CAF+ reagent (5-formylbenzene-1,3-disulfonic acid disodium salt hydrate), p.a. synthetic product, Ruđer Bošković Institute

Sodium cyanoborohydride, for synthesis, Merck, Germany

Potasium dihydrogen phosphate, p.a., Kemika, Croatia

2.2. SOLUTION PREPARATION

2.2.1 CELL LYSIS

Cell lysis buffer 1 (25 mM NH_4HCO_3 + 0.1% Triton-X-100)

Dissolve 198 mg of NH_4HCO_3 in a total volume of 100 mL 0.1% Triton-X-100 (v/v).

Cell lysis buffer 2 (25 mM $(NH_4)_2CO_3$ + 0.1% Triton-X-100)

Dissolve 240.2 mg of $(NH_4)_2CO_3$ in a total volume of 100 mL 0.1% Triton-X-100 (v/v).

2.2.2. DETERMINATION OF PROTEIN CONCENTRATION

Bradford reagent

Dissolve 10 mg of Commassie Brilliant Blue G-250 in a total volume of 100 mL solution containing 85% water, 10% phosphoric acid and 5% ethanol (v/v/v).

2.2.3 2-D SDS-PAGE

1.5 M Tris-HCl buffer

Dissolve 18.2 g Trizma base in a total volume of 100 mL water. Adjust pH to pH 8.8 using 6 M HCl.

30% Acrylamide / bisacrylamide

Dissolve 29.2 g of acrylamide and 0.8 g of N,N´-methylen-bis acrylamide in a total volume of 100 mL water.

10% (w/v) Ammonium persulfate (APS)

Dissolve 100 mg of APS in a total volume of 1 mL water.

10% (w/v) Sodium dodecyl sulfate (SDS)

Dissolve 10 g of SDS in a total volume of 100 mL water.

M1 Buffer

Dissolve components in water according to Table 1.

Table 1. M1 stock buffer

Component	Quantity
Urea	4.2042 g
Thiourea	1.522 g

CHAPS	0.4 g
Water	Up to 10 mL

Aliquot stock solution and store it at -20 ^{o}C. Immediately prior to use, mix 857.2 µL of M1 buffer stock solution with 142.8 µL of 7× protease inhibitor cocktail and 10 mg of DTT to get ready to use M1 buffer.

Equilibration buffer

Prepare the buffer by combining all reagents according to Table 2.

Table 2. Equilibration buffer

Component	**Quantity**
1.5 M Tris-HCl, pH 8.8	10 mL
6 M Urea	72.07 g
Glycerol 99.5%	60.3 mL
2% SDS	4 g
Bromophenol Blue	0.1 µg
Water	Up to 200 mL

12% Acrylamid/Bis solution

Prepare monomer solution by combining all reagents according to Table 3.

Table 3. 12 % Acrylamid/Bis solution

Component	**Volume**
30% Acrylamid/Bis	60 mL
Water	50.2 mL
1.5 M Tris-HCl, pH 8.8	37.5 mL
10% SDS	1.5 mL
10% APS	750 µL
TEMED	75 µL

10× Running buffer, pH 8.3

Dissolve components in water according to Table 4.

Table 4. 10× Running buffer

Component	**Quantity**
Trizma base	16 g

Glycine	72.1 g
SDS	5 g
H_2O	Up to 500 mL

0.5% Agarose

Put the components in water according to Table 5. and dissolve in microwave oven.

Table 5. 0.5% Agarose

Component	Quantity
0.5% Agarose	0.5 g
Bromophenol blue 0.0002%	200 µg
Running buffer (5x)	20 mL
Water	80 mL

Coomassie blue staining solution

Prepare stock solution according to Table 6.

Table 6. Coomassie blue staining stock solution.

Component	Quantity
Coomassie Brilliant Blue G250	2 g
Phosphoric acid	47 mL
Ammonium sulfate	200 g
Water	Up to 2 L

Prepare working solution by mixing stock solution and methanol in 5:1 ratio (v/v).

Gel storage solution

Prepare solution containing 2% glacial acetic acid in water (v/v).

Destaining solution

Prepare solution containing 10% glacial acetic acid, 40% methanol and 50% water (v/v/v).

2.2.4 PROTEIN DIGESTION, PEPTIDE EXTRACTION AND PURIFICATION

2.2.4.1 In-solution digestion

Tryptic buffer, 25 mM NH_4HCO_3, pH 7.8

Dissolve 198 mg of NH_4HCO_3 in a total volume of 100 mL water. Adjust pH to 7.8 using 10 % glacial acetic acid.

Trypsin stock solution (1 mg/mL)

Dissolve lyophilized trypsin (50 µg) sequencing grade in 50 µL of water.

V8 protease buffer - 25 mM $(NH_4)_2CO_3$, pH 7.8

Dissolve 240.2 mg of $(NH_4)_2CO_3$ in a total volume of 100 mL water. Adjust pH to 7.8 using 10% glacial acetic acid.

V8 protease stock solution (1 mg/mL)

Dissolve lyophilized V8 protease (50 µg) sequencing grade in 50 µL of water.

2.2.4.2 Mag-Trypsin protein digestion (trypsin immobilized on magnetic beads)

Washing buffer (0.1M $NaHCO_3$, pH 8.3)

Dissolve 840 mg of $NaHCO_3$ in a total volume of 100 mL water.

Digestion buffer (0.1M $NaHCO_3$, pH 8.3, 1 M Urea)

Dissolve 840 mg of $NaHCO_3$ and 600.6 mg of urea in a total volume of 100 mL water.

2.2.4.3 In-gel digestion and extraction

50 mM NH_4HCO_3 buffer, pH 7.8

Dissolve 396 mg of NH_4HCO_3 in a total volume of 100 mL water.

ACN:50 mM NH_4HCO_3 (50:50 v/v)

Mix 5 mL of ACN with 5 mL of of 50 mM NH_4HCO_3 pH 7.8.

Extraction solution (ACN: aqueous solution of 5% TFA (50:50 v/v)

Mix 5 mL of ACN with 5 mL of aqueous solution of 5% TFA (v/v).

2.2.4.4 Protein purification

Solutions for Zip-Tip C4 and SPE

ACN: aqueous solution of 0.1% TFA (80:20 v/v)

ACN: aqueous solution of 0.1% TFA (50:50 v/v)

Aqueous solution of 0.1% TFA (v/v)

2.2.5 PEPTIDE DERIVATIZATION

Derivatization buffer (10 mM KH_2PO_4, pH 5)

Dissolve 34 mg of KH_2PO_4 in a total volume of 25 mL water. Adjust pH with H_3PO_4 to pH 5.

Derivatization solution

Dissolve 8 mg of CAF-/CAF+ reagent (5-formylbenzene-1,3-disulfonic acid disodium salt hydrate) and 12 mg of $NaBH_3CN$ in 1 mL of derivatization buffer.

2.2.6 LIQUID CHROMATOGRAPHY

2.2.6.1 Capillary liquid chromatography (CapLC)

Mobile phase A

Prepare aqueous solution of 0.1% TFA.

Mobile phase B

ACN:aqueous solution of 0.1% TFA (80:20 v/v).

2.2.6.2 2-dimensional ultra-performance liquid chromatography (2-D UPLC)

Mobile phase A1

Prepare 20 mM ammonium formate, pH 10 by dissolving 126 mg of ammonium formate in a total volume of 100 mL water. Adjust pH with NaOH to pH 10.

Mobile phase B1

ACN.

Mobile phase A2

Prepare aqueous solution of 0.1% formic acid (v/v).

Mobile phase B2

Prepare ACN containing 0.1% formic acid (v/v).

2.2.7 MASS SPECTROMETRY

CHCA matrix

Prepare 5 mg/mL of CHCA matrix by dissolving 5 mg of CHCA in 1 mL of ACN:water (50:50 v/v).

2.3. PROTOCOLS

2.3.1 PROTEIN SAMPLE PREPARATION

Note: This protocol contains 2-DE and LC steps denoted with small letters in addition (optional$_{2\text{-}DE}$ or optional$_{LC}$).

2.3.1.1 Removing cell suspension growth medium

- Centrifuge cell suspension at 2000 rcf for 15 min and aspirate the supernatant.
- Wash the pellet three times in 1.5 mL of 25 mM NH_4HCO_3.

2.3.2.2 Cell lysis

- Resuspend the pellet in 400 µL of lysis buffer 1 or lysis buffer 2. Keep the tubes incubated on cool pack.
- Proceed to cell rupture using TissueRuptor and disposable probe (Figure 1) with the following settings:

    - Power: Position II
    - Grinding cycle: 45 seconds ON / 45 seconds OFF
    - Total grinding times: 5-6 cycles

Figure 1. TissueRuptor

2.3.2.3 Blocking endoprotease activity (*optional*$_{LC}$)

- Put the sample in boiling water for 3 minutes to inhibit endoprotease activity.

2.3.2.4 Obtaining solution of proteins

- Centrifuge at 4000 *g* at 4°C for 15 min and transfer the supernatant to a clean 1.5 mL tube.

2.3.2.5 Nucleic acid removal (*optional*$_{2\text{-}DE}$)

- Add 1 µL of RNase A and 10 µL of DNase I in the sample.
- Vortex and incubate in thermomixer (25 °C, 500 rpm) for 20 min.

2.3.2.6 Determination of protein concentration

- Make the triplicate serial dilution of the sample using water as diluent.
- Mix 10 µL of each dilution with 100 µL of Bradford reagent.
- Make blank using the tube which contains no protein.
- Adjust the spectrophotometer to a wavelength of 595 nm, make blank and read absorbance at 595 nm wavelength.

2.3.2.7 Protein precipitation (*optional*$_{2\text{-DE}}$)

- Mix 4 volumes acetone to 1 volume protein solution and incubate overnight at -20˚C.
- Centrifuge at 4,000 rpm at 4°C for 30 min and discard the supernatant.
- Resuspend precipitated proteins in M1 buffer and incubate at 35˚C until the proteins are completely dissolved.

2.3.2. PROTEIN SEPARATION BY TWO-DIMENSIONAL SODIUM DODECYL SULFATE POLYACRYLAMIDE GEL ELECTROPHORESIS (2-D SDS-PAGE), IN-GEL DIGESTION AND PEPTIDE PURIFICATION

Note: Isolate proteins from the cells as was described above, but include Nucleic Acid Removing step (Step V) and Protein Precipitation Step (Step VII).

2.3.2.1 2-D SDS-PAGE

Note: Protein Separation by 2-D SDS-PAGE will be done before practicum starts.

Isoelectric Focusing (IEF)

- Assemble PROTEAN IEF Cell system (Bio-Rad) for isoelectric focusing (Figure 2) according to manufacturer's instructions.
  (*see* http://www.bio-rad.com/webroot/web/pdf/lsr/literature/4006164B.pdf).

Figure 2. PROTEAN IEF cell system.

- Select immobilized pH gradient (IPG) strip with the optimal pH range.
- Resuspend precipitated proteins in 200 µL of M1 buffer and determine protein concentration.
- Dilute the volume of protein sample that contains 700 µg of total proteins in M1 buffer up to 350 µL.
- Pipette the protein solution into the loading wells of the focusing tray and cover it with IPG strip.
- Apply mineral oil overlay IPG strip and close the lid of IEF tray.
- Place the tray on the Peltier platform, select the method (Table 7) and start the run (the program include a pause step after rehydration).
- After rehydration, pre-wet electrode wicks in Milli-Q water and insert them directly on top of both electrode wires in the focusing tray just prior to isoelectric focusing and continue the run.

Table 7. Programs for rehydration and isoelectric focusing.

Phase	**Voltage [V]**	**Duration**
Rehydration	50	840 min
Focusing		
S 1	200	75 min
S 2	500	15 min
S 3	500	60 min
S 4	1000	15 min
S 5	1000	60 min
S 6	7000	240 min
S 7	7000	90 000 Vh
S 8	500	1440 min

IPG Strip Equilibration

- Put the strip in equilibration buffer (Table 2) containing 10 mg/mL DTT and soak for 15 minutes.
- Discard the buffer and put the strip in the equilibration buffer containing 25 mg/mL iodoacetamide and soak for 15 minutes.

SDS-PAGE

- Assemble PROTEAN II xi 2-D cell (Bio-Rad) (Figure 3) according to manufacturer's instructions. (see http://www.natur.cuni.cz/biologie/servisni-laboratore/genomicka-a-proteomicka-laborator/soubory/protean-manual).

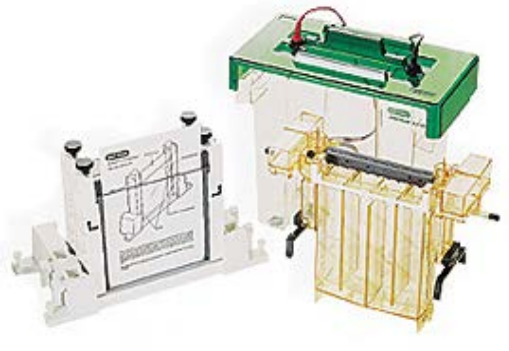

Figure 3. PROTEAN IEF Cell system.

- Prepare the 12% Acrylamid/Bis solution (Table 3) by combining all reagents except APS and TEMED and mix it well.
- Add APS and TEMED to the solution, and pour it between glass plates using a glass pipette. Immediately overlay with *n*-butanol.
- Allow the gel to polymerize for 45 minutes to 1 hour, and then completely rinse off the *n*-butanol from the gel with distilled water.
- Remove gel sandwich from casting stand and install it in the electrophoresis apparatus.
- Remove the remaining water from the top of the gel by inserting a small piece of filter paper between the glass slides.
- Immerse the IPG strip in 1x running buffer (Table 4) and place it onto the top of the gel (between the glass plates).
- Overlay with 0.5% molten agarose (Table 5). Fill the electrophoresis apparatus with 1× running buffer (Table 6), and run electrophoresis according to program in Table 8.

Table 8. Program for 2-D SDS-PAGE

Step	Current /	Duration
1	1.5	1 h
2	20	600 VHs
3	30	5 h

- After completion of electrophoresis, transfer the gel to Commassie blue staining solution (Table 6) for 16 h and destain it in water afterwards.
- Store the gel in the gel store solution.

Processing 2-D SDS-PAGE gel

- Scan the gel with VersaDoc™ 3000 imaging system (Figure 4) and analyse it with PDQuest 2-D Analysis Software. Example of 2-DE SDS-PAGE gel is shown in Figure 5.

Figure 4. VersaDoc™ 3000 imaging system.

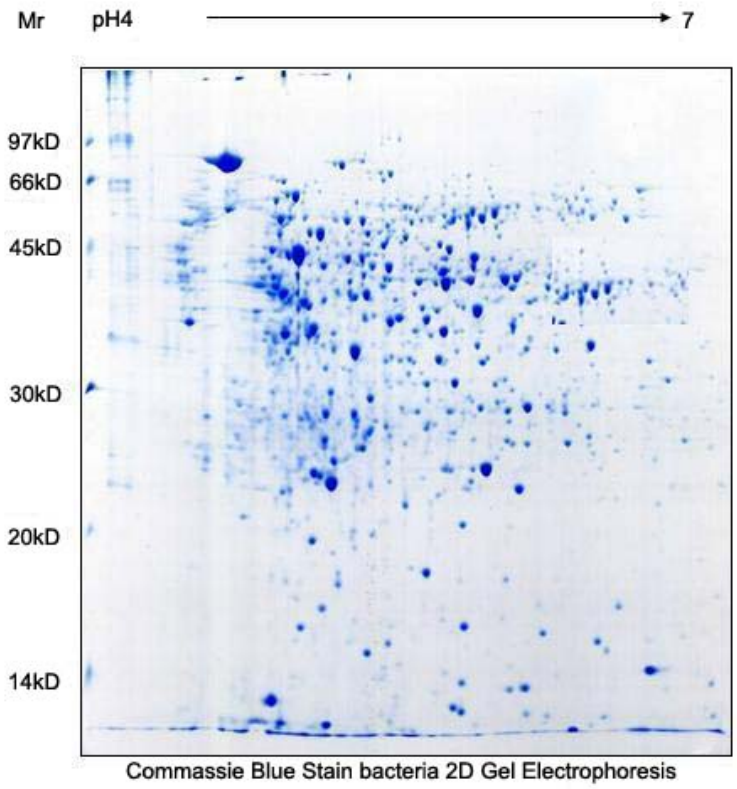

Figure 5. Example of Coomassie blue stained 2-D gel.

2.3.2.2 In-gel digestion

Gel Excision and Rinsing

- Excise the protein spot of the interest from the gel, place it in 1 mL of destaining solution and incubate in thermomixer (40°C, 700 rpm) for 3 hours. After incubation, discard the rest of the destaining solution.
- Add 0.5 mL of buffer 50 mM NH_4HCO_3, pH 7.8 to gel slice and incubate in thermomixer (25°C, 700 rpm) for 5 minutes.
- Discard the rest of the buffer (repeat this cycle twice).
- Add fresh 0.5 mL of 50 mM NH_4HCO_3, pH 7.8 buffer to gel slice and incubate in thermomixer (25°C, 700 rpm) for additional 30 minutes. Discard the rest of the buffer.
- Add 0.5 mL of solution 50 mM NH_4HCO_3, pH 7.8 / ACN (50:50 v/v) to gel slice and incubate in thermomixer (25°C, 700 rpm) for 30 minutes. After incubation, discard the rest of the solution.
- Add 100 µL of ACN to gel slice and incubate in thermomixer (25°C, 700 rpm) for 5 minutes. After incubation, discard the rest of ACN.
- Dry the gel slice in SpeedVac.

In-gel Digestion

- Add 10 µL of trypsin (final concentration 20 µg/mL) to dried gel slice and incubate in thermomixer at 37°C and 650 rpm for 18 hours.

Peptide Extraction

- After in-gel digestion, remove the enzyme buffer, place it in a clean Eppendorf tube and dry in SpeedVac.
- Add 10 µL of extraction solution to remaining gel slices and incubate for 30 minutes in ultrasonic bath. In addition, incubate gel samples in thermomixer (25°C and 500 rpm) for 15 min. Combine extracts with accompanying dried peptides and dry in SpeedVac.

2.3.2.3 Peptide Purification

- Resuspend dried peptides in 10 µL of aqueous solution of 0.1% TFA.
- Condition each ZipTip (Figure 6.):
  - 3 × ACN: aqueous solution of 0.1% TFA (80:20 v/v)
  - 3 × ACN: aqueous solution of 0.1% TFA (50:50 v/v)
  - 3 × aqueous solution of 0.1% TFA (v/v)
- Bind the peptide sample:
  - Bind peptides to Zip Tip by aspirating and dispensing 10×.
- Rinse sample 5× with aqueous solution of 0.1% TFA.

- Elute samples with 10 µL ACN: aqueous solution of 0.1% TFA (80:20 v/v). Dry peptide samples in SpeedVac.
- Resuspend dried peptide sample in 5 µL of CHCA matrix, deposit on MALDI plate and analyse by using MALDI-TOF/TOF mass spectrometer.

Figure 6. ZipTip pipette tips

2.3.3. PROTEIN IN-SOLUTION DIGESTION, PEPTIDE DERIVATIZATION AND PURIFICATION

Note: Isolate proteins from the cells as was described above but include Blocking endoprotease activity step (step III.).

2.3.3.1 Protein in-solution digestion

Use one of three suggested digestion procedures:

Mag-Trypsin (trypsin immobilized on magnetic beads)

- Mix protein solution with digestion buffer in 1:1 ratio.
- Heat the protein-buffer mixture at 80°C for 20 min.
- Prepare Mag-Trypsin suspension as follows:
  A. Mix the Mag-Trypsin suspension and transfer 1 mL into a 1.5 mL tube. Place the tube on a magnetic separator and remove the storage buffer.
  B. Remove the tube containing the beads from the separator and add 1 mL of Milli-Q water.
  C. Mix and place the tube containing the beads on the separator. Remove water.
  D. Repeat Steps B and C.
  E. Remove the beads from separator and add 1 mL of wash buffer.
  F. Repeat Steps B and C with wash buffer instead of water.
- Add the protein sample to the Mag-Trypsin suspension and incubate with mixing at 37°C for 2 h.
- Once digestion is complete, place the tube on magnetic separator and collect the reaction mixture.

Standard in-solution tryptic digestion

- Add trypsin in ratio 50:1 (w/w) to protein sample.
- Incubate in thermomixer (37 ºC, 700 rpm) for 18 h.

Standard in-solution V8 digestion

- Add V8 protease in ratio 50:1 (w/w) to protein sample.
- Incubate in thermomixer (25 ºC, 700 rpm) for 18 h.

2.3.3.2 Peptide derivatization

- Mix digested protein sample with derivatization solution in ratio 1:4.
- Place the tube in the holder (e.g. styrofoam) and perform the derivatization in a household microwave oven at 90W for 10 minutes.

2.3.3.3 Peptide purification

Solid Phase Extraction (SPE) – optional

- Purify the peptides after trypsin digestion and/or derivatization on Diol cartridges using vacuum manifold SPE system.
- Condition the Diol cartridges in three steps:
  1) Aspirate and dispense to waste three times with 0.5 mL of ACN:aqueous solution of 0.1% ammonium formate (80:20 v/v).
  2) Aspirate and dispense to waste three times with 0.5 mL of ACN:aqueous solution of 0.1% ammonium formate (50:50 v/v).
  3) Aspirate and dispense to waste three times with 0.5 mL of an aqueous solution of 0.1% ammonium formate (v/v).
- Load the peptide sample solution onto the cartridge and wash five times with 0.5 mL aqueous solution of 0.1% ammonium formate (v/v).
- Elute the peptides from the column to a clean 1.5 mL tube with 0.35 mL of ACN:aqueous solution of 0.2% ammonium formate (80:20 v/v).
- Evaporate the eluting solution to dryness in vacuum centrifuge.

2.3.3.4 Sample filtration – optional

- Filter the sample before loading LC column by passing it through cellulose acetate spin filter 0.22 µm using low-speed centrifugation (e.g at 4000 rpm for 30 s).

2.3.4. PEPTIDE SEPARATION BY LIQUID CHROMATOGRAPHY

2.3.4.1 Peptide Separation by Capillary Liquid Chromatography (CapLC)

The CapLC system equipped with a Photodiode Array (PDA) detector coupled to TempoTM LC MALDI spotter is used for peptide separation and collection directly onto the MALDI plate.

- Perform the chromatographic separation on a silica based LC column C18, 300 µm x 150 mm I.D., and 3.5 µm particle size at 30°C. Set the flow rate to 2 µL/min and injection volume to 8 µL. Vial temperature is maintained at 5°C in the autosampler tray. Eluted peptides are detected by UV absorbance at 280 nm. Program the 50 min gradient elution to increase the percentage of solvent B from 5% to 80% over 35 min and then to condition the column back to the initial conditions until completion of the run. Set complete gradient conditions following the Table 9.

Table 9. Complete gradient conditions.

Time (min)	A (%)	B (%)
0	95	5
7	95	5
35	20	80
40	95	5
50	95	5

- Set the spotter make-up flow to 2 µL/min (5 mg/mL CHCA matrix). Example of typical CapLC chromatogram is shown in Figure 7.

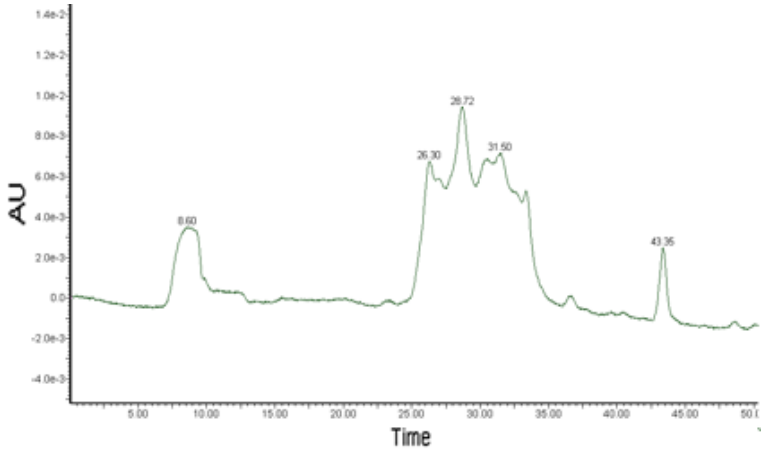

Figure 7. Example of typical CapLC chromatogram.

2.3.4.2 Peptide separation by Ultra Performance Liquid Chromatography (UPLC) 2-D method

The nanoACQUITY UPLC (Figure 8) system equipped with a TUV detector, Sample Manager and two Binary Solvent Managers is used for 2-D peptide separation. The nanoACQUITY UPLC System with 2-D technology runs with two reversed-phase separations at high and low pH, optionally strong cation exchange (SCX) followed by reversed-phase separations.

The nanoACQUITY UPLC is directly coupled to the mass spectrometer SYNAPT G2-Si.

Figure 8. *nano*ACQUITY UPLC.

- For first dimension (reversed phase or ion exchange), perform the chromatographic separation on a silica based LC column C18, 180 µm x 10 mm ID, 3.5 µm particle size or SCX based column 50 µm x 23.5 mm ID, 3.5 µm and set appropriate flow rate (e.g. 2-15 µL/min). Program the 20 min gradient elution to increase the percentage of solvent B from 5% to 11%.
- For second dimension (peptide separation), perform the chromatographic separation on a silica based LC column C18, 75 µm x 100 mm ID, 1.7 µm particle size and set the flow rate to 250 nL/min. Program the 30 min gradient elution to increase the percentage of solvent B from 5% to 80%.
- Inject the sample on the first-dimension column (peptide trap column, RP column 1) at pH 10 to wash unwanted solutes to waste. The first-dimension binary solvent manager (BSM2) elutes analytes (peptides) from RP column 1 to a mixing tee. The second-dimension binary solvent manager (BSM1) takes sample (peptides) from the first dimension and adjusts the pH from 10 to 2 while decreasing the organic content to trap the sample on the trapping column. The second-dimension BSM1 takes over the separation at pH 2 for high-resolution MS analyses. Eluted peptides are detected by UV absorbance at 280 nm. Example of chromatograms is shown in Figure 9.

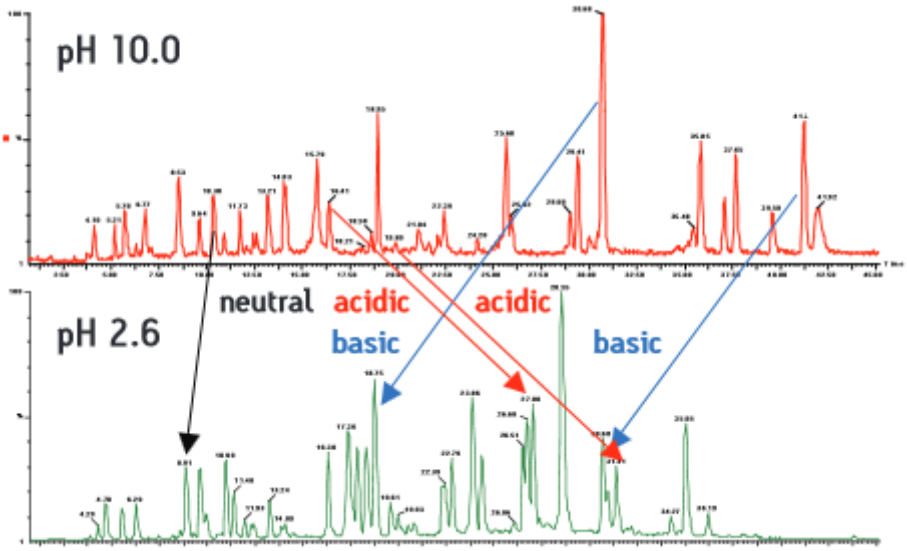

Figure 9. Comparison of peptide elution at pH 10.0 (top) and pH 2.6 (bottom) showing good UPLC chromatographic resolution and change in elution order.

2.3.5. MASS SPECTROMETRY

2.3.5.1 MALDI-TOF/TOF mass spectrometry

Use ions derived from the trypsin auto-digestion for external calibration of mass spectrometer in positive ion mode.

Use ions derived from the trypsin auto-digestion with CAF-/CAF+ derivatization for external calibration in negative ion mode according to Table 10.

Table 10. Calibration standard.

***m/z* (Da)**			
MS		**MSMS**	
Positive	**Negative**	**Positive**	**Negative**
1045.564 2211.104 3161.496	1090.4554 2459.0459	110.072 551.258 1091.549 1695.798	548.100 1027.313 1368.508 1796.674

MS acquisition

- MS acquisition is performed on a MALDI TOF/TOF 4800 Plus analyzer (Figure 10) equipped with a 200 Hz, 355 nm neodymium-doped yttrium aluminum garnet Nd:YAG laser. Ions are analyzed in reflectron positive or negative (for derivatized peptides) ion mode. The instrument parameters are set using the 4000 Series Explorer software version 3.5.3 (Applied Biosystems, USA). Mass spectra are obtained by averaging 1800 laser shots covering a mass range of *m/z* 1000 to 4000.
- MS/MS acquisition is achieved by 1kV collision energy in positive ion mode. Derivatized peptides are analyzed in negative than in positive ion mode without usage of collision gas (*de novo* sequencing). The same precursor ions generated by negative or positive ion MS were analyzed in negative and/or positive MS/MS.

Figure 10. 4800 Plus MALDI TOF/TOF Analyser.

Analysis parameters

- *MS analysis parameters*

  Mode: Detection of positive/negative ions
  Mirror: Reflectron
  Number of laser shots per spectrum: 1000
  Mass range: 900 – 4000 Da
  Ions retention time: 500 ns

- *MSMS analysis parameters*

  Mode: Detection of positive/negative ions
  Mirror: Reflectron
  Number of laser shots per spectrum: 1600
  Ions retention time: 80 ns

Example of MS and MS/MS spectra is shown in Figures 11 and Figure 12.

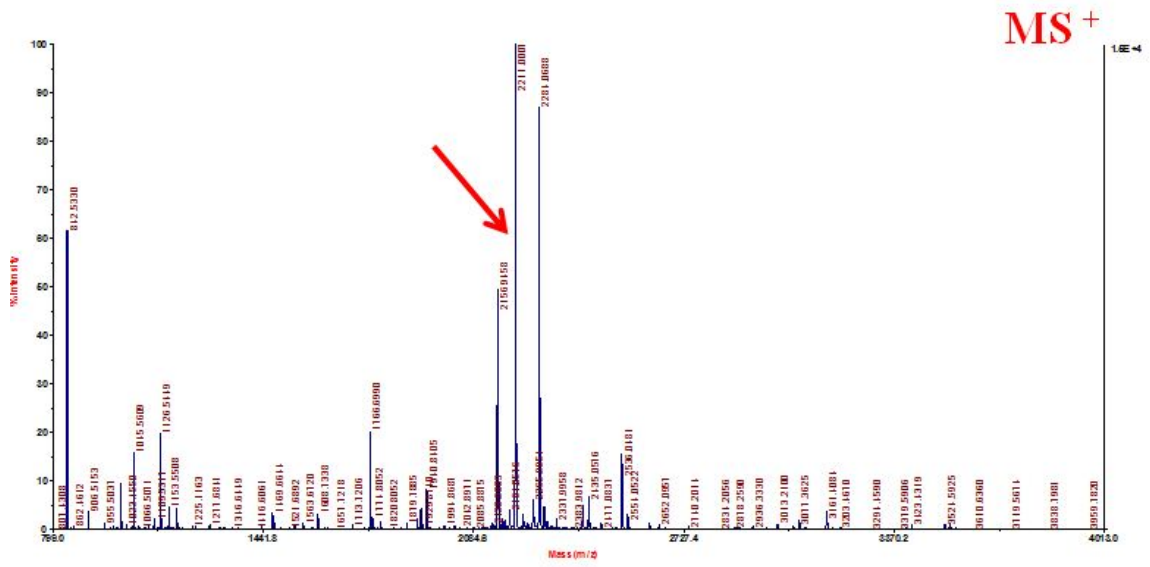

Figure 11. MALDI-TOF/TOF MS positive spectrum of the peptides in the *m/z* range between 799 and 4013. Peptide peak *m/z* 2211.0007 is marked with red arrow.

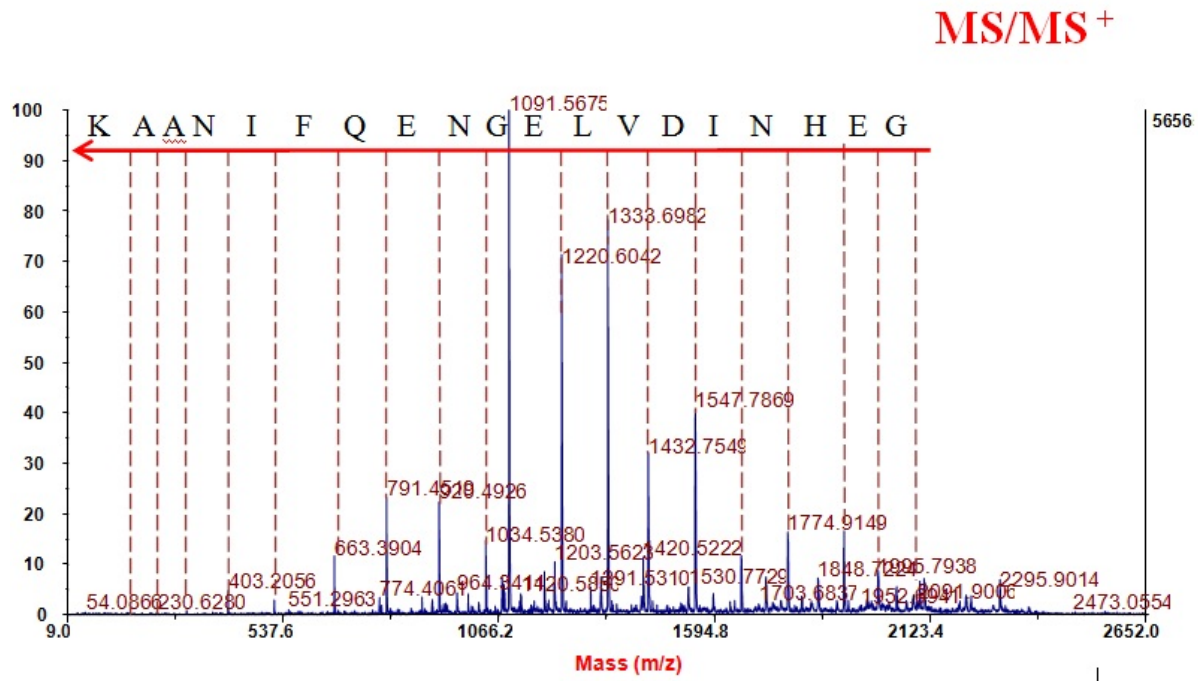

Figure 12. MALDI-TOF/TOF MS/MS positive spectrum of precursor ion at *m/z* 2211.0007 assigned to amino acid sequence LGEHNIDVLEGNEQFINAAK.

To check the amino acid sequence LGEHNIDVLEGNEQFINAAK, the peptide can be fragmented *in silico* using "Protein Prospector" program (Figure 13.) which can be accessed on the web page http://prospector.ucsf.edu/prospector/mshome.htm.

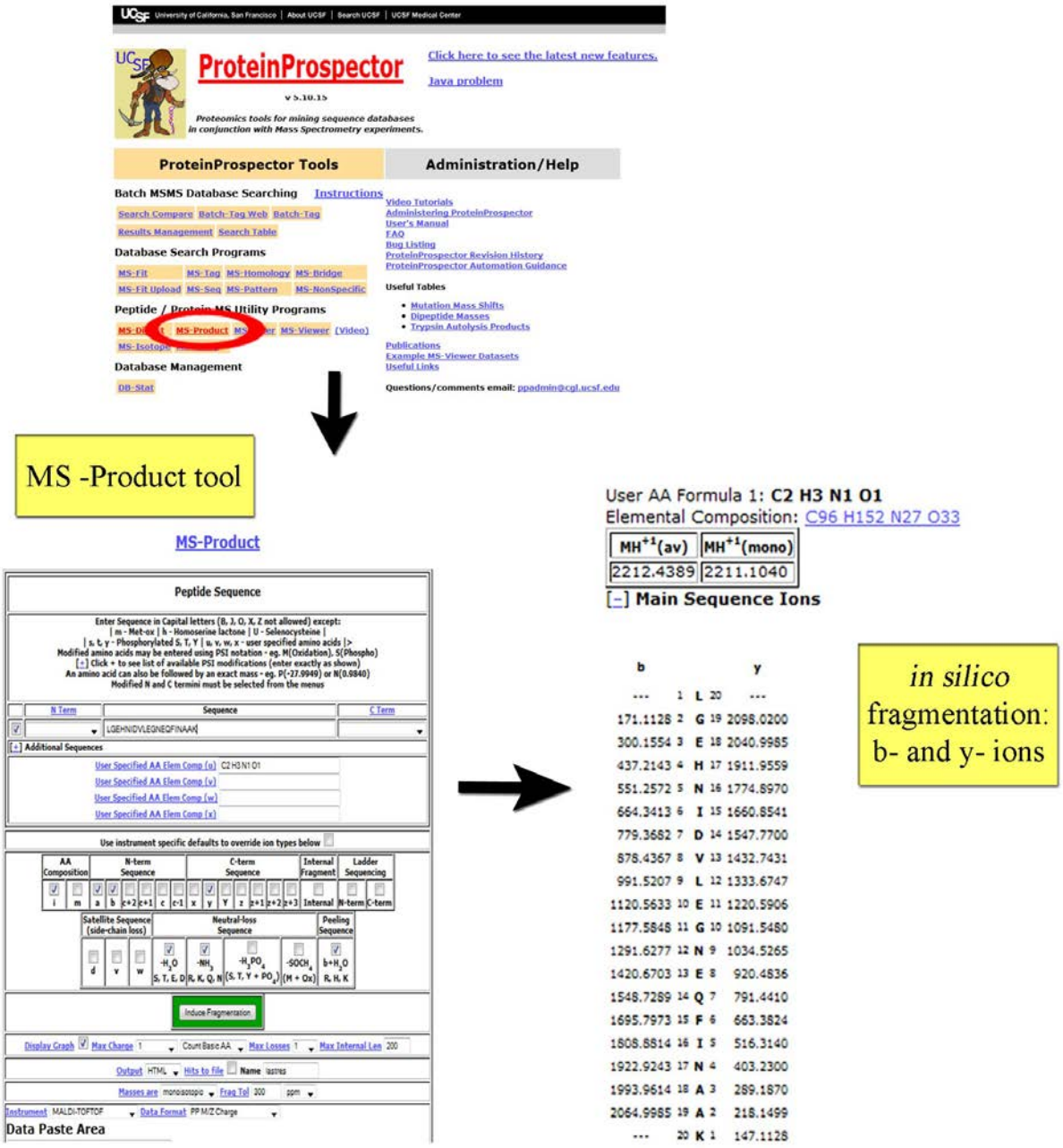

Figure 13. Example of *in silico* fragmented peptide LGEHNIDVLEGNEQFINAAK using "Protein Prospector" program.

2.3.5.2 ESI-qTOF mass spectrometry

After peptide separation by nanoACQUITY system, peptides are delivered to the ESI-qTOF mass spectrometer via NanoLockSpray interface using the auxiliary pump of nanoACQUITY system.

- MS acquisition is performed with ESI-qTOF SYNAPT G2-Si mass spectrometer (Figure 14). Ions are analyzed in positive and negative electrospray (ESI+ and ESI-) modes. The instrument parameters are set using the Mass Lynx software version 4.1. SCN902 (Waters, USA). Mass spectra are obtained in a mass range of *m/z* 1000 to 4000.
- Peptide fragmentation is performed in the trap collision cell with argon as the collision gas in ESI+ and ESI- modes. The voltage on the collision cell was adjusted with mass and charge to give a distribution of fragment ions across the mass scale (CE values: 80-120).

  The ion source conditions:
  - Temperature: 80 °C
  - Nitrogen flow: 0.7 Bar
  - Infusion needle potential: 1.2 kV
  - Cone voltage: 30
  - RF-1 voltage 150 V

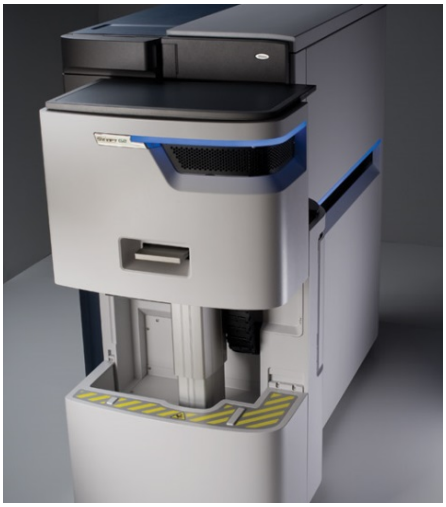

Figure 14. SYNAPT G2-Si mass spectrometer.

2.3.6. SPECTRA PROCESSING AND *DE NOVO* SEQUENCING

2.3.6.1 Protein Pilot (Mascot search engine)

For each MS and MS/MS spectrum, software named Protein Pilot is used to determine which peptide sequence in a database gives the best match. Each entry in the database is digested *in silico*, using the known specificity of the enzyme, and the masses of the intact peptides are calculated. If the calculated mass of a peptide matches that of an observed peptide, the masses of the expected fragment ions are calculated and compared with the experimental values.

Because the data in each MS/MS spectrum correspond to an isolated peptide, it makes no difference whether the original sample was a single protein or a mixture. Individual peptide sequences are identified, then the set of peptide sequences is used to infer which proteins may have been present. Unless a peptide is unique to one particular protein, there may be some ambiguity as to which protein it originated from. Mascot method is shown in Figure 15.

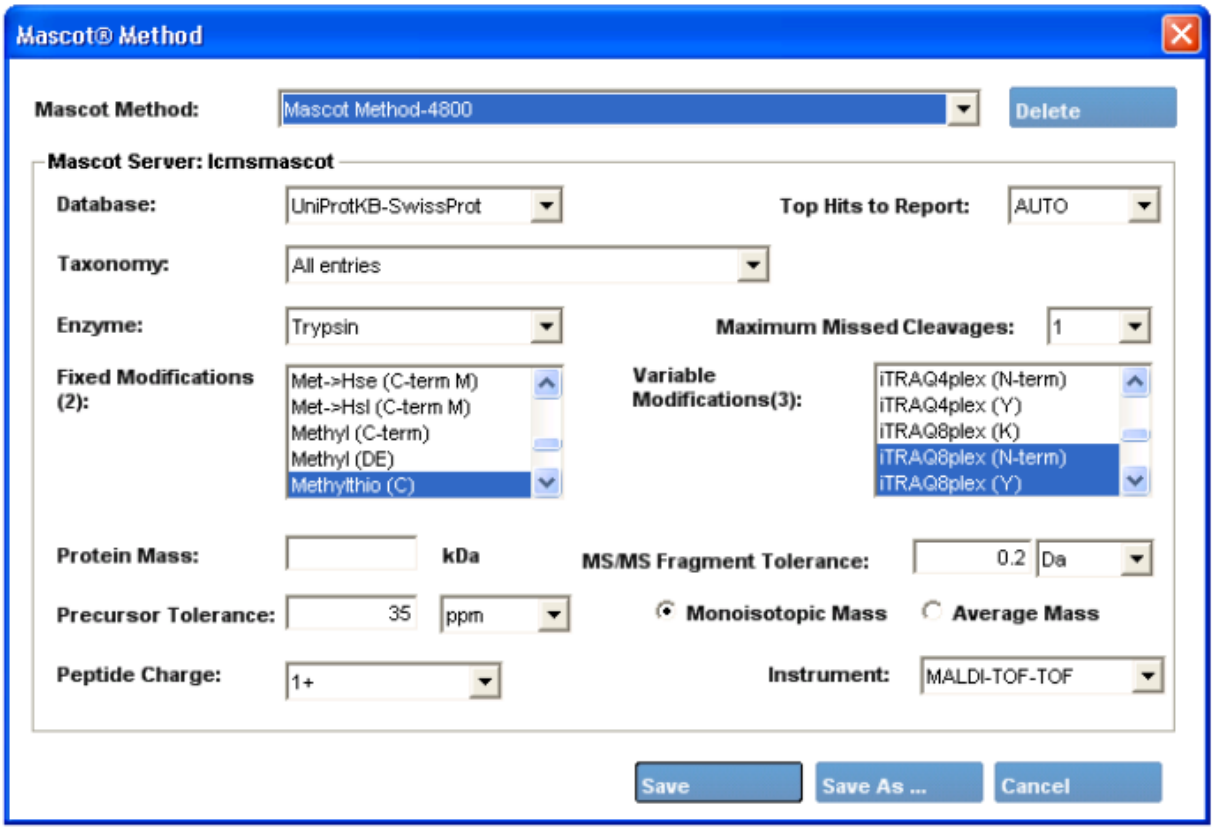

Figure 15. Search parameters of MS/MSMS spectra using Mascot method.

Search parameters:

a. Database: National Center for Biotechnology Information (NCBI)

   Choose protein database that will be used for peptide matching. The National Center for Biotechnology Information in the USA (http://www.ncbi.nlm.nih.gov/) houses a series of databases that contain sequences in the ubiquitous and non-redundant FASTA format.

b. Taxonomy: enter particular organism or taxonomic rank.

   If a database contains taxonomy information, use this to restrict the search to entries for a particular organism or taxonomic rank.

c. Mass tolerance

The search engine supports separate mass tolerances for precursors and fragments since many instruments may have very different accuracies for MS and MS/MS. Used instruments have the *m/z* accuracy of up to 0.3 both in MS and MS/MS ion mode.

d. Enzyme: enter the enzyme you used for protein digestion.

Setting the number of allowed missed cleavage sites to zero simulates a limit digest. If you are confident that the digest was perfect, with no partial fragments present, this will give maximum discrimination. But, if experience shows that the digest mixtures usually include some partials (peptides with missed cleavage sites) you should choose a setting of 1, or maybe 2 missed cleavage sites.

e. Post-Translational Modifications: phosphorylation, oxidation, deamination...

In database searching, modifications are handled in two ways. First, there are the quantitative modifications, usually called fixed or static. An example would be the efficient alkylation of cysteine. Since all cysteines are modified, this is effectively just a change in the mass of cysteine. It carries no penalty in terms of search speed or specificity.

In contrast, most post-translational modifications do not apply to all instances of a residue. For example, phosphorylation might affect just one serine in a peptide containing many serines. Non-quantitative modifications, usually called variable or differential, are expensive in the sense that they increase the time taken for a search and reduce its specificity.

2.3.6.2 ProteinReader (software developed by "Ruđer Bošković" Institute and Faculty of Food Technology and Biotechnology, University of Zagreb)

Protein Reader is a program that identifies an organism based on MS and MS/MS spectra. By searching the internal database (consisted of tryptic peptides of proteins taken from nrNCBI database) program identifies the organism on the basis of identified proteins that are specific for a particular organism. The named software has the ability to perform database match or *de novo* sequencing analysis.

a. Database Match

ProteinReader matches peaks from MS/MS^- and MS/MS^+ spectra along with MS^- data against NCBInr database to identify peptides/proteins. Processed and assembled data represents identified proteins and assigned species. Identifying the same peptide from MS/MS^- and MS/MS^+ spectra of the same precursor mass is referred as Sure Shot and it significantly increases reliability of matched result. Based on the number of matched peptides (primarily on quantitative number of Sure Shots), ProteinReader is able to

unambiguously determinate the species that is present in the NCBInr database (more than 3000 species). Sure Shot does not rely on probabilistic scoring algorithms used in similar database match search engines (e.g. MASCOT, PEAKS). It is a simple amino acid overlap reading between MS/MS$^-$ and MS/MS$^+$ spectra (from N-to C-terminus and back from C- to N-terminus, respectively).

Use parameters shown in Figure 16 to perform Database Match search.

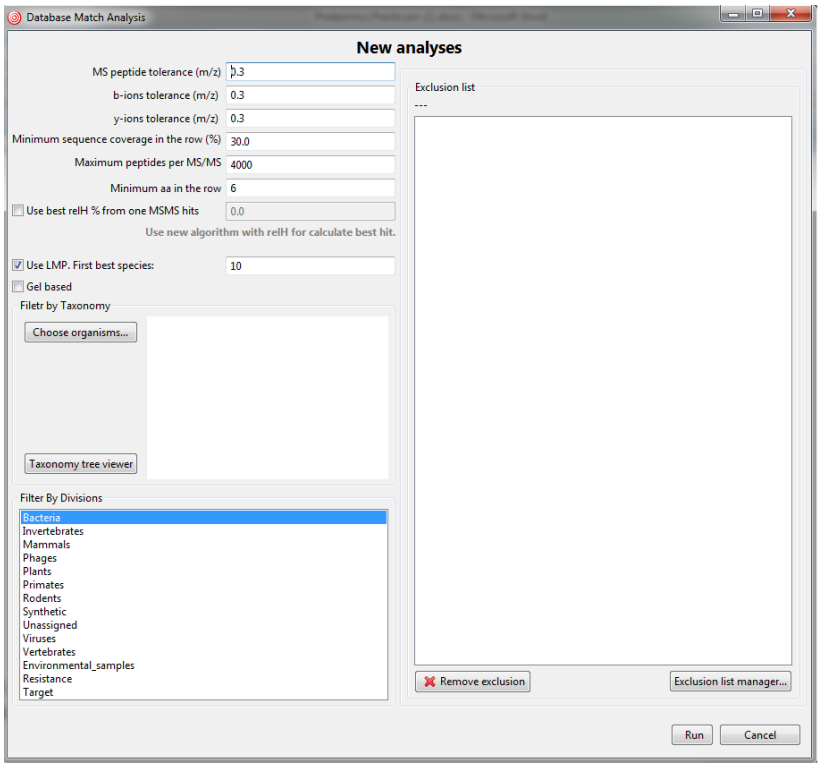

Figure 16. Search parameters of MS and MS/MS spectra using ProteinReader Database Match.

b. *De novo* Sequencing

Once ProteinReader Database Match identifies the species, *de novo* algorithm reads the amino acid sequence (≥ 5 amino acids in a series) from MS/MS$^-$ and MS/MS$^+$ spectra and makes alignment with protein sequences of the selected species in NCBInr database. Since *de novo* reading algorithm does not follow database match probabilistic logic it can be considered as independent and orthogonal peptide/protein identifier.

Use parameters shown in Figure 17 to perform *de novo* sequencing.

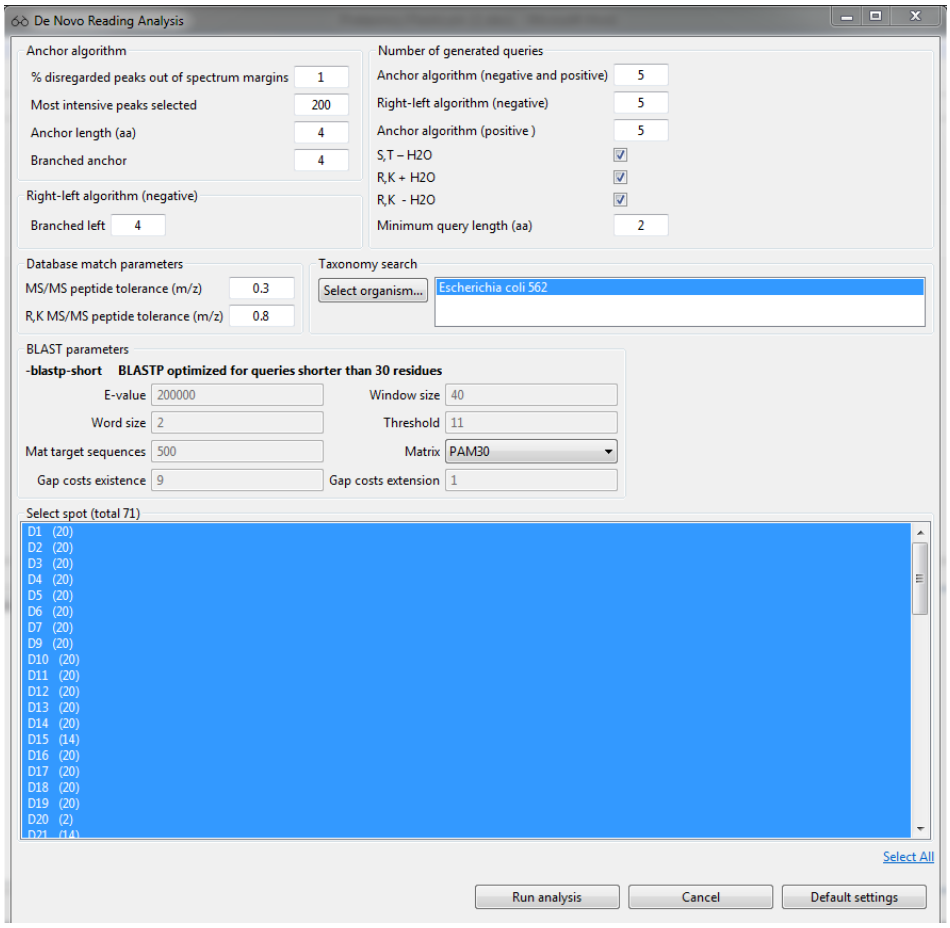

Figure 17. Search parameters of MS/MS spectra using ProteinReader *De novo* Sequencing.

3. RESULTS

3.1 MASCOT SEARCH ENGINE RESULTS

Result of Mascot search engine is shown at Figure 18.

Figure 18. Mascot search engine result.

3.2. PROTEINREADER RESULTS

Final results of analysis are reported for one species as intersection of results obtained from Database Match (Figure 19) and *De novo* Sequencing (Figure 20) as final Report (Figure 21). Proteins that are not confirmed by Sure Shot or Database Match and *De novo* Sequencing overlapping are not taken into consideration.

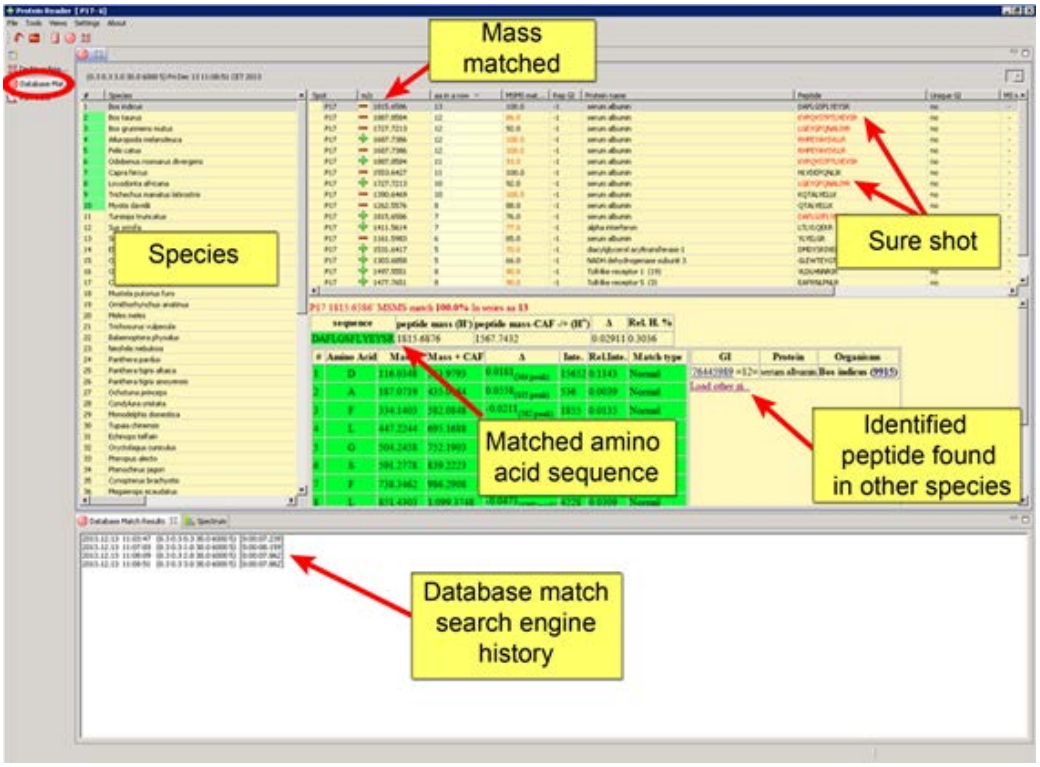

Figure 18. ProteinReader Database Match result.

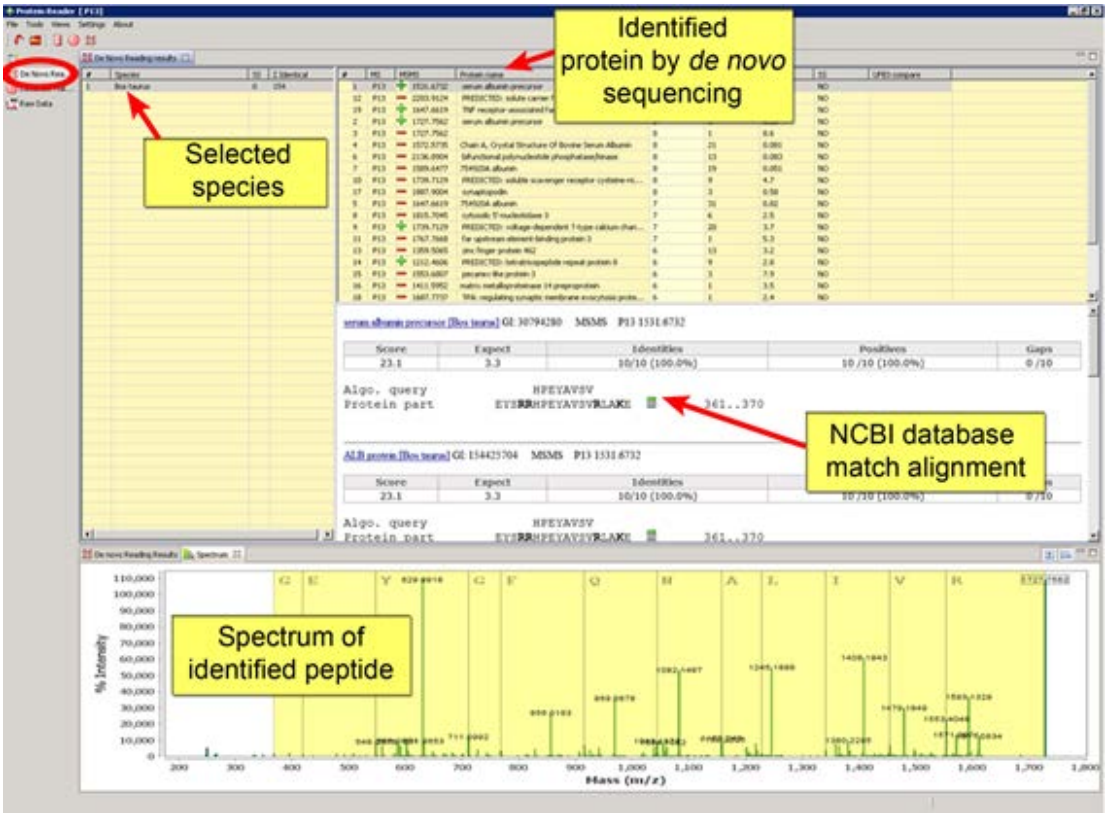

Figure 19. ProteinReader *De novo* Sequencing result.

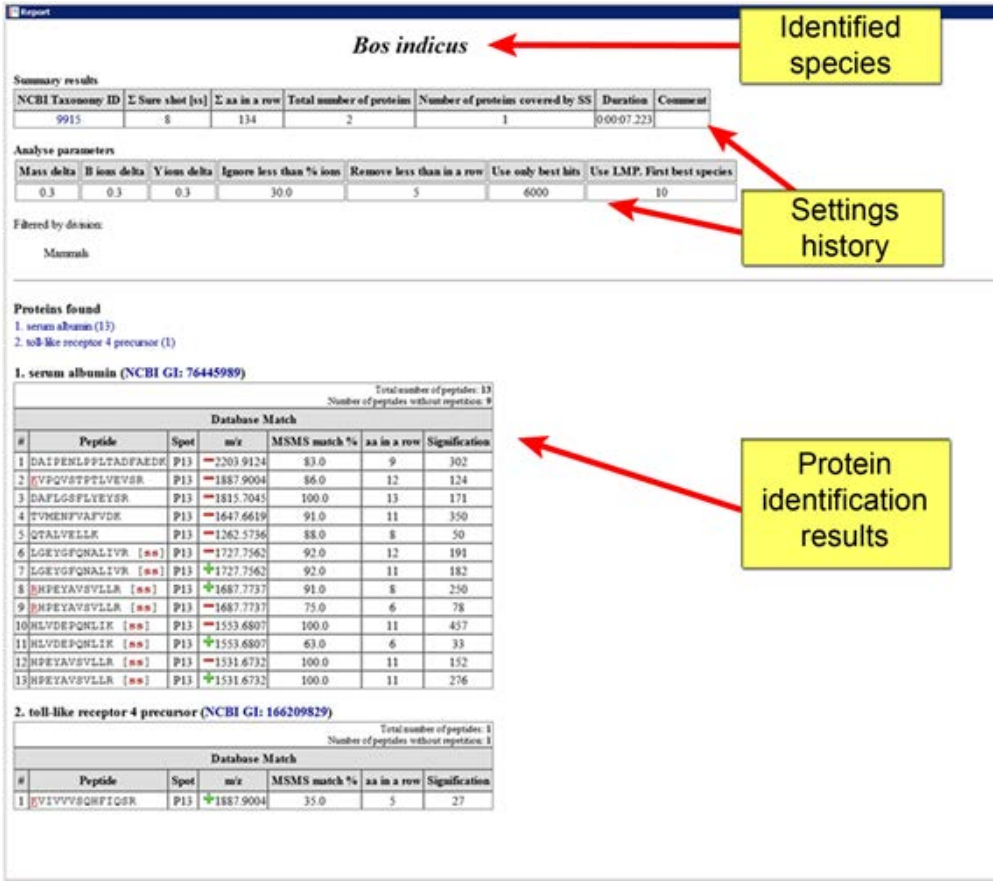

Figure 20. ProteinReader final Report.