**Title**
Top-Down, Bottom-Up, and Side-to-Side Proteomics with Virtual 2-D Gels



**Author**
Ogorzalek Loo, Rachel R

**Publication Date**
2004-11-23

Peer reviewed

# Top-Down, Bottom-Up, and Side-to-Side Proteomics with Virtual 2-D Gels

Rachel R. Ogorzalek Loo,[1,3*] Richard Hayes,[4] Yanan Yang,[2] Frank Hung,[2] Prasanna Ramachandran,[2] Nuri Kim,[2] Robert Gunsalus,[3,4] and Joseph A. Loo[1-3]

[1]Department of Biological Chemistry

[2]Department of Chemistry & Biochemistry

[3]Molecular Biology Institute

[4]Department of Microbiology, Immunology, and Molecular Genetics

University of California-Los Angeles

*Corresponding author:

rloo@mednet.ucla.edu
406A Boyer Hall
Molecular Biology Institute
Los Angeles, CA 90095-1570
(310) 206-1484 Phone
(310) 206-7286 FAX

**ABSTRACT**

Intact protein masses can be measured directly from immobilized pH gradient (IPG) isoelectric focusing (IEF) gels loaded with mammalian and prokaryotic samples, as demonstrated here with murine macrophage and *Methanosarcina acetivorans* cell lysates. Mass accuracy and resolution is improved by employing instruments which decouple the desorption event from mass measurement; e.g., quadrupole time-of-flight instruments. MALDI in-source dissociation (ISD) is discussed as a means to pursue top-down sequencing for protein identification. Methods have been developed to enzymatically digest all proteins in an IEF gel simultaneously, leaving the polyacrylamide gel attached to its polyester support. By retaining all gel pieces and their placement relative to one another, sample handling and tracking are minimized, and comparison to 2-D gel images is facilitated. MALDI-MS and MS/MS can then be performed directly from dried, matrix-treated IPG strips following whole gel trypsin digestion, bottom-up methodology. Side-to-side proteomics, highlighting the link between virtual and classical 2-D gel electrophoresis, is introduced to describe a method whereby intact masses are measured from one side (the IEF gel), while proteins are identified based on analyses performed from the other side (the SDS-PAGE gel).

**Keywords:** archaea, proteomics, virtual 2-D gel electrophoresis, immobilized *pH* gradient gels, *Methanosarcina acetivorans*

**Abbreviations:** 2-D PAGE, two-dimensional polyacrylamide gel electrophoresis; CHAPS, 3-(3-cholamidopropyl)-dimethylammonio-1-propanesulfonate; DTT, dithiothreitol; HDL, high density lipoprotein; ICAT, Isotope coded affinity tagging; IEF, isoelectric focusing; IPG, immobilized *pH* gradient; ISD, in-source dissociation; MudPIT, multidimensional protein identification technology; PSD, post-source decay; QqTOF, quadrupole time-of-flight; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; TOF, time-of-flight

## 1. Introduction

Intact mass measurements, in concert with protein identifications, can alert one to 1) co- or post-translationally modified proteins, 2) alternative splice variants, 3) RNA editing events, or 4) ribosomal frameshifts [1], alternate promoters and/or alternate initiation sites. Despite their potential to provide new insights, intact mass measurements have received scant attention in proteomics discussions. This lack may be due, in part, to their incompatibility with the new proteomics paradigm, in which complex samples are enzymatically digested and the products separated by multidimensional chromatography and analyzed by data-dependent tandem mass spectrometry. However, this "incompatibility" really reflects the complementary nature of such analyses. While tryptic peptide quantification can reveal a protein's abundance averaged over all forms of protein (assuming the peptide is common to *all* forms), quantitative 2-D PAGE (two-dimensional polyacrylamide gel electrophoresis) analysis reveals the abundance of *individual* protein components; e.g., different proteolytic products or modified forms.

Virtual 2-D Gel Electrophoresis [2-8] was developed to mass measure *intact*, polyacrylamide gel-isolated proteins for proteomic applications. In the virtual 2-D gel method, MS is performed directly from dried isoelectric focusing (IEF) gels, substituting MALDI-MS for the SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) 2nd dimension of classical 2-D PAGE. Useful features of the virtual 2-D gel approach for obtaining intact masses from complex samples include 1) analysis of *small* quantities of complex samples, 2) detection of low abundance proteins, (as revealed by comparison to silver-stained 2-D gels) [5] 3) superb resolution and good inter-laboratory reproducibility from immobilized *pH* gradient (IPG) gels (facilitating comparisons to 2-D gels prepared more than 15 years ago) [5], and 4) synergistic links to other analyses performed *via* IEF and 2-D gels (e.g., Western and far-Western blotting,

functional group-specific staining, phosphorimaging, peptide mapping, and carbohydrate analysis). Enormous value accrues from linking to the standardized separation/visualization/storage/dispensing *protein array* that 2-D gels comprise.

Without the link to protein identity, an intact mass has limited value. Hence, tying the intact masses recovered from virtual 2-D gels to unambiguous protein identities and further to descriptors is important. We consider 3 ways to forge this link; two are well-described in popular parlance as: "top-down" and "bottom-up" [9,10]. We call the third, "side-to-side", because it highlights the link to classical 2-D gel electrophoresis: namely that isoelectric focusing constitutes one side/dimension/axis/gel in classical 2-D analysis, while SDS-PAGE constitutes the second side/dimension/axis/gel. In side-to-side analysis, one measures the intact mass from one side (the IEF gel), while securing the protein identification from an analysis performed on the other side (the SDS-PAGE gel).

Top-down proteomics strategies provide both accurate intact mass and sequence data, bridging the gap between protein identification and characterization. As executed with MALDI-generated ions, top-down methods rely on in-source dissociation (ISD) performed on time-of-flight (TOF) analyzers [5,11-15] or tandem mass spectrometry on a TOF-TOF [16]. Very small proteins or biopolymers may succumb to collisional dissociation (e.g., on MALDI-source-equipped quadrupole time-of-flight instruments) or to MALDI post-source decay (PSD) on reflector time-of-flight analyzers. For example, MALDI-PSD has identified and characterized tetrahydrofolylpolyglutamic acid polymers in *E. coli* cell lysates [17].

Bottom-up strategies are most potent delivering sequence information from low abundance protein components of complex mixtures for which only limited sample is available (e.g., tissue), or from extraordinarily hard-to-handle proteins; i.e., integral membrane proteins. Important

approaches to bottom-up analysis include "sort, then break", as achieved by 2-D gel electrophoretic separation, in-gel tryptic digestion, and LC-MS/MS and "break, then sort" achieved by first cleaving proteins within a solubilized lysate followed by multidimensional chromatography interfaced to data-dependent tandem mass spectrometry (i.e., multidimensional protein identification technology (MudPIT) [18,19]. Isotope coded affinity tagging (ICAT) [20] also enzymatically digests proteins prior to separation, fitting the "break, then sort" mold. The bottom-up variation examined here, *whole-gel* tryptic digestion followed by MALDI-MS/MS performed directly from the dried gel, can be considered another "sort, then break" method.

Acquisition of intact masses has been combined with bottom-up analysis by VerBerkmoes, et al. [21] who mass-measured intact proteins in *Shewanella oneidensis* HPLC fractions and identified those proteins by enzymatically digesting the fractions). Because top-down is usually bestowed on combinations where the intact protein mass measurement and identification analyses are *directly* linked, i.e., dissociation of an intact protein ion *isolated within* the mass spectrometer, side-to-side will be employed to describe approaches such as those of the Oak Ridge laboratory [21], in which the intact mass measurement is linked to the identification analysis *indirectly*. Our side-to-side analysis measures intact masses from dried IEF gels, while obtaining additional information from the corresponding 2-D gel spot (e.g., by trypsin digestion, Edman degradation, or antibody binding). Alternatively, the corresponding region on an IEF gel could be digested and analyzed.

Bottom-up and side-to-side analyses are demonstrated here for proteins from a whole-cell lysate of the archaeon *Methanosarcina acetivorans*. Previous 2-D electrophoresis/N-terminal sequencing results have been reported for the related organism, *Methnosarcina thermophila* [22], and top-down sequencing data has been presented from *M. acetivorans* protein fractions [23].

The current studies employed IEF gels loaded with 35 and 85 μg of total protein, quantities that, when loaded onto 2nd dimension polyacrylamide gels, require staining by silver or Sypro Ruby to visualize protein spots and that demonstrate compatibility with the limited sample sizes available in many proteomics projects.

Strategies combining assets from top-down and bottom-up methods are needed if proteomic-scale mass spectrometry is to move beyond identification to characterization, if characterization is to include more than the one hundred most abundant proteins in the cell and if it is to be performed not only on cultured cells but also on tissue. Virtual 2-D gel electrophoresis has potential to fulfill these needs.

## 2. Experimental

*Methanosarcina acetivorans* C2A (methanogenic archae) were cultivated as single cells anaerobically in a 100mM methanol, 200mM NaCl-based medium at 37°C [24]. Optionally, protease inhibitor cocktail (Sigma P8340) was added immediately before harvest. Pelleted cells were stored at -80°C until lysis., for which pellets were resuspended in 300 μL of 0.3% SDS (w/v), 10 mM Tris, *pH* 7.5, and 0.2 M dithiothreitol (DTT) and incubated in a boiling water bath for 2 minutes, after which they were cooled on ice. To the cooled solution was added 30 μL of DNAse/RNAse cocktail [25] and the mixture was incubated on ice for 10 min. Following incubation, 1200 μL of "lysis buffer," 138 mg urea, and 50 mg thiourea were added and the mixture was vortexed extensively. Lysis buffer contained 7 M urea, 2 M thiourea, 50 mM DTT, 1.54% Pharmalyte™ 3-10 carrier ampholytes (v/v), and 2.55% CHAPS (w/v). After this procedure, the archaeal pellet was completely solubilized.

Immobilized *pH* gradient gels (*pH* 4-7, 18-cm lengths), purchased from Amersham Biosciences, were focused on an Amersham Multiphor II equipped with an EPS 3501XL power

supply for 64 kilovolt-hours. Sample (35 or 85 μg) was loaded by in-gel rehydration or by cathodal cup-loading. Following isoelectric focusing, polyacrylamide gels were stored at -80°C, until processed for direct mass spectrometry, whole-gel trypsin digestion, or the SDS-PAGE second dimension.

Murine macrophage cells (RAW 264.7) were cultured and lysed as described previously [26]. Isoelectric focusing was performed as described above.

*SDS-PAGE 2nd Dimension, Protein Visualization, and Trypsin Digestion of Excised Spots*

IPG strips were equilibrated for 10 minutes in 37.5 mM Tris-HCl, *pH* 8.8, 20% glycerol, 2% SDS, 6 M urea and 2% DTT, followed by 10 minutes in 37.5 mM Tris-HCl, 20% glycerol, 2% SDS, 6 M urea and 2.5% iodoacetamide. Following equilibration, strips were applied to precast Tris-HCl gels (Protean II Ready Gel®, 12%, Bio-Rad) within a Protean Plus Dodeca or Protean XL Cell (Bio-Rad) and sealed with molten agarose. Following electrophoresis, gels were stained with Sypro-Ruby (Molecular Probes, Eugene, OR) and fluorescence-imaged with ultraviolet excitation on a Molecular Imager FX Pro Plus™ (Bio-Rad). Spots were excised by a spot-excision robot (ProteomeWorks™, Bio-Rad) and automatically washed, trypsin-digested and extracted [27] by a robotic liquid-handling workstation (Mass*PREP*™, Micromass-Waters, Beverly, MA). The resulting peptides were spotted onto MALDI targets with α-cyano-4-hydroxycinnamic acid matrix either automatically (Mass*PREP*™) or manually.

### *2.1 Virtual 2-D Gel Electrophoresis*

Focused, unstained, IPG gels were prepared for mass spectrometry as described previously [5,8]. Once dry, the 3 mm wide matrix-treated gel strips (plastic backing still attached) were cut to 3.5-4 cm lengths and mounted onto the mass spectrometer's modified sample stage with double stick adhesive tape [2-5,8].

Prior to mass analysis, 0.2-0.3 μL of cytochrome *c*/sinapinic acid solution was dispensed along one edge of the dried IPG gel at multiple *pH* positions for external and internal mass calibration. Sinapinic acid was saturated in 33% $CH_3CN$/67% $H_2O$ (v/v) acidified to 0.1% with trifluoroacetic acid (TFA).

Positions of protein bands visualized by MALDI-MS or by silver staining were converted to isoelectric points by assuming a linear *pH* 4-6.6 gradient across the gel's full length (approximately 180 mm), based on an Amersham-supplied plot of *pH* versus distance for *pH* 4-7 Immobiline DryStrip™ gels at 20°C and 8 M urea.

### *2.2 Whole Gel Tryptic Digests*

In-gel trypsin digests utilized unstained IPG gels. After electrophoresis, focused gels were washed in 10% acetic acid/30% methanol for 10 minutes. Wet gels were scored (sliced) every 2 mm without detaching the polyacrylamide from its plastic support. After scoring, the plastic strip with attached gel slices was washed for 10 minutes in 30% (v/v) methanol/70% (v/v) 140 mM $NH_4HCO_3$/0.75% (w/v) D-sorbitol and allowed to dry overnight at room temperature. Two microliters of trypsin (10 ng/μL) were spotted on each gel segment and the gel strip was incubated in a warm, moist environment for at least two hours. After incubation, gels were immersed for 30 seconds in a saturated solution of α-cyano-4-hydroxycinnamic acid in 1:1 $CH_3CN$:0.1% TFA with 0.75% D-sorbitol and then allowed to dry at room temperature. Dried, matrix-deposited strips were mounted on the MALDI sample stage and analyzed as described in Section 2.1.

### *2.3 Mass Spectrometry*

MALDI mass spectra of intact proteins embedded in dried polyacrylamide gels were acquired on an Applied Biosystems Voyager DE-STR time-of-flight mass spectrometer

(Framingham, MA) operated with 337 nm irradiation, delayed extraction, and positive ion detection. The instrument was operated primarily in linear mode, but reflector mode was used to improve the mass accuracy of ions below 10 kDa, especially in regions where calibrant ions were not visible. Mass spectra were also acquired with a MALDI quadrupole time-of-flight (QqTOF) mass spectrometer (QSTAR®-XL, Applied Biosystems, Toronto, Canada). The latter instrument, also operated in positive ion mode with 337 nm irradiation, provided well-calibrated MALDI-MS spectra of intact proteins without added calibration standards.

MALDI peptide fingerprint mass spectra of extracts of trypsin-digested excised 2-D gel spots were obtained on reflector time-of-flight mass spectrometers (either M@LDI-R (Micromass-Waters) or Voyager DE-STR). Sequence information was obtained from excised protein spots with nano-liquid chromatography interfaced to the QqTOF mass spectrometer, employing electrospray ionization (ESI). Mass spectra of tryptic peptides desorbed from dried IEF gels were also acquired with the Voyager DE-STR, while corresponding tandem mass spectra were provided by the QSTAR®-XL.

Peptide fingerprint mass spectra and tandem mass spectra (MALDI-MS/MS and LC-MS/MS) were searched against proteins predicted by the *Methanosarcina acetivorans* genome sequence [28], using the Mascot search program (Matrix Science, London, United Kingdom) (www.matrixscience.com) [29].

### *2.4 Reagents*

Sinapinic acid, α-cyano-4-hydroxycinnamic acid, cytochrome *c*, Tris base, Tris hydrochloride, $NH_4HCO_3$, and puriss-grade thiourea were purchased from Sigma/Aldrich/Fluka (Milwaukee, WI), while TFA (sequencing grade) and CHAPS (3-(3-cholamidopropyl)-dimethylammonio-1-propanesulfonate) were obtained from Pierce. HPLC-grade water,

acetonitrile, methanol, and glacial acetic acid were purchased from EM Science (Darmstadt, Germany). Urea (electrophoresis grade), SDS, agarose, and Pharmalyte™ 3-10 ampholytes were provided by Amersham Biosciences (Piscataway, NJ), while DTT was obtained from Bio-Rad (Hercules, CA). Sequencing grade modified trypsin (porcine) and D-sorbitol were supplied by Promega (Madison, WI), and Acros Organics (Morris Planes, NJ), respectively.

## 3. Results and Discussion

### *3.1 Intact Mass Measurements*

Fig. 1 illustrates spectra obtained at discrete *pH* positions along the IEF gel. Its 180 mm long, linear *pH* gradient spans 2.6 *pH* units (Amersham 4-7 DryStrip™ format), corresponding to 0.014 *pH* units/mm. Spectra acquired 1 mm apart share ions, but also reveal differences (Figs. 1a and 1b). Low abundance, well focused proteins are observed over a 1 mm distance in mass analysis from IPG gels. This retention of spatial resolution by the immobilized *pH* gradient matrix following washing and matrix incubation [5], quite remarkable when compared to other methods imaging proteins from polyacrylamide gels and membranes, and when compared to related efforts from carrier ampholyte-based IEF gels [2,3], is precedented [30]. It is also consistent with common 2-D electrophoresis practice, in which IPG gels are incubated in multiple solutions to reduce, alkylate, and equilibrate between 1$^{st}$ and 2$^{nd}$ dimensions, without serious concerns regarding degraded spatial resolution.

Useful spectra for masses below about 5 kDa are obtained at selected locations, though they are often plagued by a high background sloping down in intensity toward higher *m/z*. The source of this background appears variable. The detergent CHAPS yields cluster ion backgrounds below 5000 *m/z* with a characteristic 615 Da spacing. It can be reduced by washing the IPG gel more extensively and/or by lowering the CHAPS concentration of the IEF

rehydration buffer. We have found cluster ion backgrounds to be most noticeable at CHAPS concentrations of 2-4% (w/v). Alternatively, the lower *m/z* range can be examined in reflector mode, i.e., instrumental conditions under which the clusters are less stable, to reduce the detergent background. Initially, we attributed the sloping, low mass background to incomplete removal of carrier ampholytes (a general additive to IPG gel rehydration solutions), but their impact may be less than previously thought [5,31]. Rather, non-proteinacious cell components, important constituents of cell lysates, particularly in the absence of pre-fractionation or protein precipitation steps, merit consideration. MALDI mass analysis from IEF gels loaded with *E. coli* whole-cell lysates reveal extremely intense, polymeric ions below 3 kDa in size, widely distributed across the *pI* span. That their spacing matches the profile ascribed by Reilly and colleagues [17] to tetrahydrofolylpolyglutamic acid, as observed from MALDI profiles of whole bacterial cells, leads us to reconsider sources of background ions. Glyco- and phospholipids [32,33] may also impart low *m/z* background, particularly at acidic *pH*, and MALDI matrix clusters are also likely to contribute background ions.

Despite sometimes unresolved low *m/z* background, excellent peptide spectra can be obtained at selected isoelectric points, likely reflecting the position of abundant peptides. Figure 2, an expansion of Fig. 1b, displays more than 50 ions focused at *pH* 4.34. The complexity, reproduced on multiple IPG gels, is unique to this isoelectric point, and may reflect the presence of a co-migrating protease, or autolyzed protein, unexplained at present. The ions do not arise from in-source dissociation, because their presence is independent of laser power and does not track with the abundance of particular higher molecular weight species. Although 2-D gels have been criticized abundantly for their failure to reveal low molecular weight proteins, the limitation is imposed only by the second dimension SDS-PAGE separation. Clearly, substituting MALDI-

MS for the SDS-PAGE size separation allows simultaneous access to small proteins and those already amenable to 2-D gel electrophoresis. Proteins under 3 kDa and up to 70 kDa have been detected from virtual 2-D gels of *Methanosarcina acetivorans*.

Large, heavily glycosylated proteins are important constituents of eukaryotic and, especially, mammalian proteomes. MALDI's ability to deliver ions in low charge states provides rare access to their intact masses and enables evaluations of heterogeneity. Capabilities for virtual 2-D gel analysis of glycoproteins were demonstrated in our study of human high density lipoprotein (HDL) particles [8]. Virtual 2-D gels' facility with glycoprotein analysis merges nicely with other polyacrylamide gel-based capabilities, including glycoprotein-specific staining, lectin blotting, carbohydrate composition analysis, and separation of charged isoforms. It is also an antidote to the inferior size measurements arising from anomalous migration of heavily glycosylated proteins on SDS-PAGE gels. These capabilities should facilitate studies of prokaryotic glycosylation, an established modification in both bacteria and archaea, primarily to cell surface-layer proteins. [34,35]

Previous studies mass analyzing intact proteins directly from dried immobilized *pH* gradient gels [3,5,8], searched databases employing a mass accuracy of ± 0.1-0.2% for proteins below 50 kDa in size. Generally, 0.1% accuracy or better is achievable for reasonably abundant proteins below 20 kDa in size. Those IEF-MALDI studies also established that proteins separated on *pH* 4-7 IPG strips are not alkylated by acrylamide, despite the reducing conditions employed during electrophoresis [4,5]. The *pH* range spanned by these gels is *not* conducive to acrylamide addition onto cysteine, in contrast to the *pH*>8 conditions encountered in classic tris-glycine SDS-PAGE separations [8]. Indeed, cysteine stability in lower *pH* polyacrylamide gels

was exploited to develop in-gel chemistries for cysteine modification and cleavage to identify and characterize proteins separated by gel electrophoresis [36].

Previous work comparing MALDI-MS under continuous extraction and delayed extraction (time-lag focusing) conditions [3,6], established that surface charging of the gel (an insulator) is not the major factor reducing mass accuracy when time-lag focusing is employed. Rather, mass accuracy on linear time-of-flight instruments is governed by the evenness of the surface from which ions are desorbed. Hence, accuracy can be improved by employing instruments which decouple the desorption event from mass measurement; e.g., quadrupole time-of-flight instruments. MALDI QqTOF mass spectra of murine RAW264.7 macrophage cell proteins desorbed from dried IEF gels are displayed in Fig. 3. Isoelectric points for the top and bottom spectra differ by approximately 0.075 *pH* units, and the resolving power of the QqTOF is sufficient to separate the isotopes at *m/z* 6000-7000. Alternatively, mass accuracy can be enhanced on standard time-of-flight instrumentation by analyzing proteins under 6 kDa in reflector mode.

### *3.2 Top-Down Proteomics*

Of means available to link protein identities to intact masses, top-down sequencing methods [9,10] are the most elegant and conceptually simple. ESI-equipped analyzers comprise the preferred platform, because highly charged protein ions are more amenable to dissociation than singly- or doubly-charged ions; charge contributes extensively to the center of mass collision energy.

On MALDI-equipped platforms, in-source dissociation provides limited access to top-down sequencing capability [5,11,14,15,37-42]. ISD has been demonstrated with ions desorbed directly from an IPG strip onto which *Escherichia coli* whole cell lysates had been focused [5].

From the 84.6-kDa protein, cobalamin-independent methionine synthase (*metE*), a c-ion series spanning 43 residues was obtained. Higher mass accuracy, as delivered by reflector ISD, is useful for confident sequence searching. Limitations include the larger quantity of protein required for ISD analyses than for intact mass measurements, potential difficulty linking ISD products to an intact precursor ion when complex mixtures are analyzed, and that the technique is far from universal, being both protein and matrix deposition/sample preparation-dependent..

Alternatively, MALDI-MS/MS of singly-charged, intact proteins from 5 to 12-kDa (0.2-10 picomoles), has been performed on a TOF/TOF™ mass spectrometer, inducing metastable fragmentation at aspartyl and prolyl residues predominantly, with, in some cases, additional sequence ions sufficient for protein identification [16]. In that study, collisional activation of intact thioredoxin cleaved all 11 Asp-Xxx peptide bonds. Interestingly, the same thioredoxin results were obtained earlier by MALDI-PSD on a reflector time-of-flight analyzer [43], providing an early example of top-down sequencing with MALDI-generated ions.

### *3.3 Bottom-Up Proteomics*

Generally, methods that identify proteins by analyzing cleaved peptides have been referred to as bottom-up, including LC-MS/MS approaches securing identifications by trypsin-digesting proteins embedded in 1-D gel bands or 2-D gel spots [44], as well as non-gel-based approaches that digest or cleave complex protein mixtures prior to separation and analysis (e.g., MudPIT [18,19] and ICAT [20]). Here, we apply a bottom-up protein identification and sequencing strategy relying on gel-wide enzymatic digestion.

Throughput, ease of automation, and the unrelenting need for higher sensitivity compel exploration of "whole gel" approaches, in which all proteins are cleaved simultaneously and the product peptides desorbed and analyzed directly from the gel. In principle, such approaches

reduce both effort and loss due to sample processing and transfer, because the sample is removed from the gel only by the laser desorption/ionization event. Moreover, a single gel is enzymatically treated, rather than 1000 isolated spots. The "molecular scanner" embodies a similar philosophy with 2-D electroblots [45,46]; gel-embedded proteins are simultaneously proteolyzed and electroblotted to a polyvinylidene difluoride membrane. Matrix is applied, and the cleavage products are analyzed directly from the membrane by MALDI-MS. Similarly, a previous study [5] combined the infrequent cleavage reagent cyanogen bromide (CNBr), cleaving C-terminal to methionine residues, with IPG gels. The larger peptides produced by chemical cleavage (relative to trypsin) are desirable, because fewer peptides are required for unambiguous identification. However, side-reactions (e.g., with tryptophan), cleavage at acid-labile bonds (Asp-Pro) and incomplete cleavage complicated data interpretation. Moreover, diffusion was observed for small peptide products when the cleavage was performed on wet gels and the degraded spatial resolution, in concert with ion suppression, compromised the method's utility. For mixtures encountered in the analysis of whole cell lysates from IPGs, it may be that the proteins' matched isoelectric points minimize ion suppression, while those of the corresponding peptide products vary, thus limiting the number of peptides detected.

A novel procedure has been developed to perform trypsin digestions on entire IPG gels [7]. Separated proteins are reduced and alkylated in-gel after focusing, and trypsin solutions are applied by airbrushing. Mass analysis is performed directly from dried gels. This promising approach does not allow digest conditions to be optimized independently of spatial (*pI*) resolution, however, because the small cleavage products diffuse rapidly. [30] To limit diffusion, we propose scoring focused, unfixed, unstained IPG gels with a sharp template, knife, or razor blade, creating ca. 2-mm gel sections, yet leaving the polyester backing to which the

slices are bound, intact. In this manner, individual handling and processing of tens of gel sections is avoided because each slice remains attached to the polyester strip. (See Fig. 4.) Ideally, diffusion would be limited to the volume of the sliced gel sections, enabling trypsin, matrix, and other reagents to be dispensed by spotting (or dipping in solution), rather than airbrushing.

Key requirements for this approach are (1) that the gel be washed carefully before digestion to remove carrier ampholytes, urea, detergent, and other IEF additives. Washing prior to digestion is preferred, because proteins should diffuse out of the porous IEF gels more slowly than peptides. (2) The polyacrylamide slice must remain bound to its polyester support throughout processing and analysis. Our preliminary data suggest that with added plasticizers such as sorbitol, the needed stability can be attained.

Feasibility for gel-wide digestion and analysis by MS and MS/MS is demonstrated in Fig. 5. Figure 5a illustrates a reflector time-of-flight mass spectrum obtained at *pI* 4.33 from an IPG strip prepared as described in Section 2.2. Several tryptic peptides are observed, along with a CHAPS dimer ion at *m/z* 1232, arising from residual detergent employed in electrophoresis. Intact proteins corresponding to this isoelectric point are displayed in Fig. 3a. In-gel digestion products were examined by tandem mass spectrometry, by loading the same gel onto a quadrupole time-of-flight mass spectrometer equipped with a MALDI source. The QqTOF mass spectrometer also simplifies calibration, by abrogating questions about the effect of surface evenness on mass accuracy. Dissociation of the *m/z* 2514 ion yields the MS/MS spectrum shown in Fig. 5b with sufficient sequence-related ions to identify the protein as methanol-5-hydroxybenzimidazolylcobamide co-methyltransferase, isozyme 2 (MA4391). Analogues of this

protein, important to methanogenesis, have been identified previously by Edman sequencing 2-D electroblots of acetate- or methanol-grown *Methanosarcina thermophila* [22].

The MS/MS data was searched with Mascot against a database of proteins predicted from the completed *Methanosarcina acetivorans* genome [28]. Note that intact mass and isoelectric point were *not* employed as search constraints. The protein's predicted *pI* and average molecular weight are 4.30 and 28100 Da, respectively. That the identified peptide matches the predicted C-terminus and that Fig. 1c reveals a 28140 Da protein, suggest that the protein MA4391 is likely to be full-length, at least at this isoelectric point. Because the mass accuracy of this measurement is expected to be within 0.1%, the discrepancy in mass suggests potential modification; additional work is needed to rule out an oxidation or sodium adduct. Stained, large format 2-D gels reveal a pair of partially resolved spots sized less than 30 kDa at this isoelectric point, consistent with multiple forms of this protein. Moreover, our virtual 2-D gel spectra display some heterogeneity for this 28 kDa protein, again implying multiple forms. Spectra acquired at positions of increasing pH show increasing intensity for a shoulder approximately 160 Da higher in mass.

### *3.4 Side-to-Side Proteomics*

We are using the term side-to-side to describe MS methods that link the intact mass measurement to the identification analysis *indirectly*. In contrast, top-down methods *directly* link intact mass measurements to identifications, while bottom-up methods need not be linked to an intact mass. Analyses ascribed to side-to-side include combinations which measure intact masses from dried IEF gels and obtain additional information from the corresponding 2-D gel spot. Here, that additional information is obtained by combined spot excision, in-gel digestion, elution, and peptide mass fingerprinting or tandem MS peptide sequence analysis. If isolated

analyses are to be linked, side-to-side methods require excellent correspondence between migration along the isoelectric point dimension, and a fair correspondence between migration along the size dimension (MALDI-MS and SDS-PAGE). Fig. 6 compares a virtual 2-D gel image to a Sypro-Ruby stained gel image obtained with an equivalent sample load over the same *pI* and molecular weight range.

A Sypro-Ruby stained protein spot, estimated by gel migration to be 24 kDa, *pI* 4.4, was excised, reduced, alkylated, and digested with trypsin. LC-ESI-MS/MS of the digest products revealed 5 peptide ions matching peptidylprolyl isomerase (MA3136). The genome sequence-predicted, full-length protein is calculated to have an isoelectric point of 4.39 and an average mass of 17756 Da. However, intact mass measurements performed from the IEF gel at *pI* 4.40 (Fig. 7) display an ion of average mass 17624 Da. Based on the protein's N-terminus (MTEET) and known propensities for initiator Met excision [47-49], Met-removal would be unremarkable, yielding a predicted mass of 17625 Da, in good agreement with the MALDI measurement. Moreover, the N-terminal tryptic peptide (TEETIKNPDK) was recovered from the in-gel digested gel spot, confirming Met excision. Thus, the complete side-to-side analysis revealed that the *pI* 4.4, 24 kDa 2-D gel spot corresponds to MA3136 lacking the N-terminal methionine and that no other modifications are present. The 36% discrepancy between 2-D gel and MALDI-MS size determinations (24 kDa vs. 17.624 kDa) is large, reflecting anomalous migration. In fact, most FKBP-type peptidyl-prolyl cis-trans isomerases are known to run anomalously slowly by SDS/PAGE [50].

Similarly, the 28 kDa, *pI* 4.3 protein identified by whole-gel trypsin digestion (Section 3.3) was also examined by standard in-gel digestion of an excised 2-D gel spot. MALDI-MS of the

resultant peptides matched the whole-gel digest assignment, again attributing the protein to MA4391.

## 4. Conclusion

Virtual 2-D gel electrophoresis has been applied to proteins from *Methanosarcina acetivorans* and murine macrophage cell cultures, revealing proteins from < 3 to 70 kDa. A parallel in-gel digestion method has been developed to enzymatically cleave proteins embedded in isoelectric focusing gel slices with minimal sample handling and little reduction in spatial resolution. The cleavage products from proteins digested in IEF gels can be analyzed by MS and MS/MS directly from the dried gels to yield unambiguous protein identifications (bottom-up). Identifications may also be obtained by MALDI-ISD in favorable cases (top-down) or by correspondence to a protein identified from a conventional 2-D gel (side-to-side). MALDI-MS protein detection from virtual 2-D gels compares favorably to fluorescently imaged Sypro Ruby-stained 2-D gels.

**Acknowledgments**

We thank Ping Yip for providing very useful software to display virtual 2-D gel images and Verne Schumaker, Lang Yam, Michael Collins, Andre Nel, and Ning Li, for providing excellent samples and collaborations. The authors acknowledge support from the U.S. Department of Energy for funding the UCLA-DOE Laboratory of Genomics & Proteomics (DE-FC03-87 ER60615), the David Geffen School of Medicine at UCLA, and the UCLA Molecular Biology Institute. Establishment and equipment in the UCLA Mass Spectrometry and Proteomics Technology Center were supplied by a generous gift from the W. M. Keck Foundation.

## References


[1] S.L. Alam, J.F. Atkins, R.F. Gesteland, Proc. National. Acad. Sci. USA 96 (1999) 14177-14179.

[2] R.R. Ogorzalek Loo, C. Mitchell, T. Stevenson, J.A. Loo, P.C. Andrews, in D.R. Marshak (Editor), Techniques in Protein Chemistry VII, Academic Press, San Diego, CA, 1996, p. 305-313.

[3] R.R. Ogorzalek Loo, C. Mitchell, T.I. Stevenson, S.A. Martin, W. Hines, P. Juhasz, D. Patterson, J. Peltier, J.A. Loo, P.C. Andrews, Electrophoresis 18 (1997) 382-390.

[4] J.A. Loo, J. Brown, G. Critchley, C. Mitchell, P.C. Andrews, R.R. Ogorzalek Loo, Electrophoresis 20 (1999) 743-748.

[5] R.R. Ogorzalek Loo, J.D. Cavalcoli, R.A. VanBogelen, C. Mitchell, J.A. Loo, B. Moldover, P.C. Andrews, Anal. Chem. 73 (2001) 4063-4070.

[6] A.K. Walker, G. Rymar, P.C. Andrews, Electrophoresis 22 (2001) 933-945.

[7] S. Iyer, J. Olivares, Rapid Commun. Mass Spectrom. 17 (2003) 2323-2326.

[8] R.R. Ogorzalek Loo, L. Yam, J.A. Loo, V.N. Schumaker, Electrophoresis 25 (2004) 2384-2391.

[9] N.L. Kelleher, H.Y. Lin, G.A. Valaskovic, D.J. Aaserud, E.K. Fridriksson, F.W. McLafferty, J. Am. Chem. Soc. 121 (1999) 806-812.

[10] F.W. McLafferty, International J. Mass Spectrom. 212 (2001) 81-87.

[11] D.C. Reiber, T.A. Grover, R.S. Brown, Anal. Chem. 70 (1998) 673-683.

[12] V. Schnaible, S. Wefing, A. Resemann, D. Suckau, A. Buecker, S. Wolf-Kuemmeth, D. Hoffmann, Anal. Chem. 74 (2002) 4980-4988.

[13] D. Suckau, D.S. Cornett, Analusis 26 (1998) M18-M21.

[14] D. Suckau, A. Resemann, Anal. Chem. 75 (2003) 5817-5824.

[15] M. Takayama, J. Mass Spectrom. Soc. Japan 50 (2002) 304-310.

[16] M. Lin, J.M. Campbell, D.R. Mueller, U. Wirth, Rapid Commun. Mass Spectrom. 17 (2003) 1809-1814.

[17] R.J. Arnold, J.P. Reilly, Anal. Biochem. 281 (2000) 45-54.

[18] D.A. Wolters, M.P. Washburn, J.R. Yates, III, Anal. Chem. 73 (2001) 5683-5690.

[19] M.P. Washburn, D. Wolters, J.R. Yates, Nature Biotech. 19 (2001) 242-247.

[20] S.P. Gygi, B. Reist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nature Biotech. 17 (1999) 994-999.

[21] N.C. VerBerkmoes, J.L. Bundy, L. Hauser, K.G. Asano, J. Razumovskaya, F. Larimer, R.L. Hettich, J.L. Stephenson, Jr., J. Proteome Res. 1 (2002) 239-252.

[22] Y.-H.R. Ding, S.-P. Zhang, J.-F. Tomb, J.G. Ferry, FEMS Microbiol. Lett. 215 (2002) 127-132.

[23] S.M. Patrie, M.J. Roth, D.E. Robinson, Y. Du, J.T. Ferguson, D.E. Whipple, M. Rother, W.W. Metcalf, N.L. Kelleher, in Proceedings of the 52nd ASMS Conference on Mass Spectrometry and Allied Topics, Nashville, TN, 2004.

[24] K.R. Sowers, R.P. Gunsalus, J. Bacteriol. 170 (1988) 998-1002.

[25] R.A. VanBogelen, F.C. Neidhardt, in A. Link (Editor), Methods in Molecular Biology: 2-D Proteome Analysis Protocols, Humana Press, Totowa, NJ, 1998, p. 21-29.

[26] G.G. Xiao, M. Wang, N. Li, J.A. Loo, A.E. Nel, J. Biol. Chem. 278 (2003) 50781-50790.

[27] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850-858.

[28] J.E. Galagan, C. Nusbaum, A. Roy, M.G. Endrizzi, P. Macdonald, W. FitzHugh, S. Calvo, R. Engels, S. Smirnov, D. Atnoor, A. Brown, N. Allen, J. Naylor, N. Stange-Thomann, K. DeArellano, R. Johnson, L. Linton, P. McEwan, K. McKernan, J. Talamas, A. Tirrell, W. Ye, A. Zimmer, R.D. Barber, I. Cann, D.E. Graham, D.A. Grahame, A.M. Guss, R. Hedderich, C. Ingram-Smith, H.C. Kuettner, J.A. Krzycki, J.A. Leigh, W. Li, J. Liu, B. Mukhopadhyay, J.N. Reeve, K. Smith, T.A. Springer, L.A. Umayam, O. White, R.H. White, E.C. de Macario, J.G. Ferry, K.F. Jarrell, H. Jing, A.J.L. Macario, I. Paulsen, M. Pritchett, K.R. Sowers, R.V. Swanson, S.H. Zinder, E. Lander, W.W. Metcalf, B. Birren, Genome Res. 12 (2002) 532-542.
[29] D.N. Perkins, D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (1999) 3551-3567.
[30] P. Sinha, M. Seidel, P.G. Righetti, I. Bause-Neidrig, E. Kottgen, J. Biochem. Biophys. Methods 18 (1989) 195-208.
[31] R.R. Ogorzalek Loo, P.C. Andrews, in Proceedings of the 46th ASMS Conference on Mass Spectrometry and Allied Topics, Orlando, FL, 1998, p. 1151.
[32] H.F. Sturt, R.E. Summons, K. Smith, M. Elvert, K.-U. Hinrichs, Rapid Commun. Mass Spectrom. 18 (2004) 617-628.
[33] G.D. Sprott, S. Sad, L.P. Fleming, C.J. Dicaire, G.B. Patel, L. Krishnan, Archaea 1 (2003) 151-164.
[34] Z. Konrad, J. Eichler, Biochem. J. 366 (2002) 959-964.
[35] S. Moens, J. Vanderleyden, Arch. Microbiol. 168 (1997) 169-175.
[36] M. Thevis, R.R. Ogorzalek Loo, J.A. Loo, J. Proteome Res. 2 (2003) 161-172.
[37] R.S. Brown, J.J. Lennon, Anal. Chem. 67 (1995) 3990-3999.
[38] J.J. Lennon, K.A. Walsh, Prot. Sci. 6 (1997) 2446-2453.
[39] V. Katta, D.T. Chow, M.F. Rohde, Anal. Chem. 70 (1998) 4410-4416.
[40] L.A. Marzilli, T.R. Golden, R.J. Cotter, A.S. Woods, J. Am. Soc. Mass Spectrom. 11 (2000) 1000-1008.
[41] M. Takayama, A. Tsugita, Electrophoresis 21 (2000) 1670-1677.
[42] M. Takayama, J. Am. Soc. Mass Spectrom. 12 (2001) 420-427.
[43] W. Yu, J.E. Vath, M.C. Huberty, S.A. Martin, Anal. Chem. 65 (1993) 3015-3023.
[44] J. Eng, A.L. McCormack, J.R. Yates, III, J. Am. Soc. Mass Spectrom. 5 (1994) 976-989.
[45] P.-A. Binz, M. Mueller, D. Walther, W.V. Bienvenut, R. Gras, C. Hoogland, G. Bouchet, E. Gasteiger, R. Fabbretti, S. Gay, P. Palagi, M.R. Wilkins, V. Rouge, L. Tonella, S. Paesano, G. Rossellat, A. Karmime, A. Bairoch, J.-C. Sanchez, R.D. Appel, D.F. Hochstrasser, Anal. Chem. 71 (1999) 4981-4988.
[46] M. Muller, R. Gras, R.D. Appel, W.V. Bienvenut, D.F. Hochstrasser, J. Am. Soc. Mass Spectrom. 13 (2002) 221-231.
[47] P.-H. Hirel, J.-M. Schmitter, P. Dessen, G. Fyat, S. Blanquet, Biochem. 86 (1989) 8247-8251.
[48] H.S. Dalboge, S. Bayne, J. Pedersen, FEBS Lett. 266 (1990) 1-3.
[49] R.R. Ogorzalek Loo, P. Du, J.A. Loo, T. Holler, J. Am. Soc. Mass Spectrom. 13 (2002) 804-812.
[50] J.E. Kay, Biochem. J. 314 (1996) 361-385.

**Figure Captions**

**Fig. 1.** MALDI mass spectra obtained at discrete isoelectric points on a *pH* 4-7 immobilized *pH* gradient isoelectric focusing gel. The gel was load with whole-cell lysate from the archaeon *Methanosarcina acetivorans* containing 35 μg of total protein. (a) *pH* 4.33 (b) *pH* 4.34 (c) *pH* 4.69.

**Fig. 2.** Expanded view of the low mass region of Fig. 1b. MALDI-MS directly from a *pH* 4.34 position on an IPG gel.

**Fig. 3.** MALDI mass spectra of murine RAW264.7 macrophage cell proteins acquired from dried IEF gels with a QqTOF mass spectrometer. Separation between the top and bottom spectra is 5 mm on the IPG strip, or 0.075 *pH* units. The resolving power of the QqTOF is sufficient to separate the isotopes at *m/z* 6000-7000.

**Fig. 4.** (top) Image of silver-stained IEF gel loaded with *Methanosarcina acetivorans* lysate. (bottom) Schematic of sliced polyacrylamide gel bound to polyester backing.

**Fig. 5.** Bottom-up analysis from IEF gels. (a) Reflector time-of-flight mass spectrum obtained at *pI* 4.33 position on an IPG gel after a whole-gel trypsin digest. The *m/z* 1232 ion corresponds to a dimer of the detergent CHAPS. (b) MALDI-MS/MS of the *m/z* 2514 ion yields sequence-related ions sufficient to identify the protein as methanol-5-hydroxybenzimidazolylcobamide co-methyltransferase, isozyme 2 (MA4391).

**Fig. 6.** Virtual 2-D gel image (left) and Sypro-Ruby stained gel (right) obtained with an equivalent load of *Methanosarcina acetivorans* lysate. The images span the same *pI* and molecular weight range.

**Fig. 7.** Side-to-side analysis. Mass Spectrum obtained from the *pI* 4.40 position on an IPG gel. This intact mass data, in combination with information from a corresponding excised, digested 2-D gel spot, established that this protein, an FKBP-type peptidyl-prolyl cis-trans isomerase (MA3136), has its initiator methionine excised, but is otherwise unmodified and full-length.

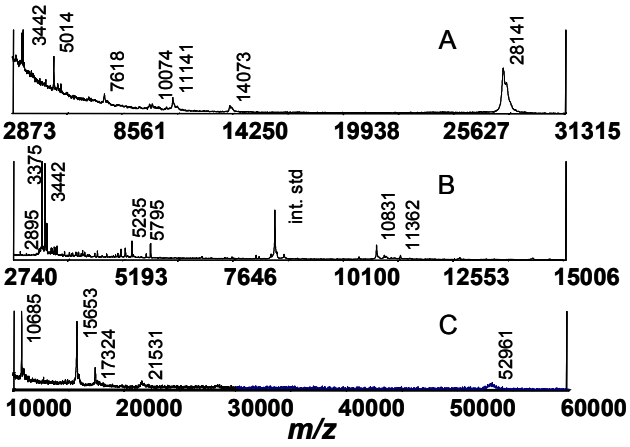
3442
5014
7618
10074
11141
14073
28141
A
2873
8561
14250
19938
25627
31315
3375
3442
2895
5235
5795
int. std
10831
11362
B
2740
5193
7646
10100
12553
15006
10685
15653
17324
21531
52961
C
10000
20000
30000
40000
50000
60000
m/z

**Fig. 1**

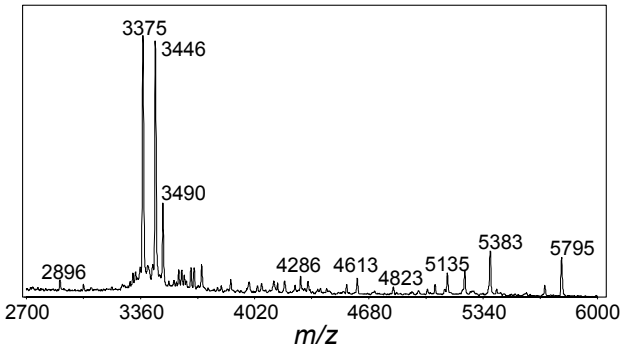
3375
3446
3490
2896
4286
4613
4823
5135
5383
5795
2700
3360
4020
4680
5340
6000
m/z

**Fig. 2.**

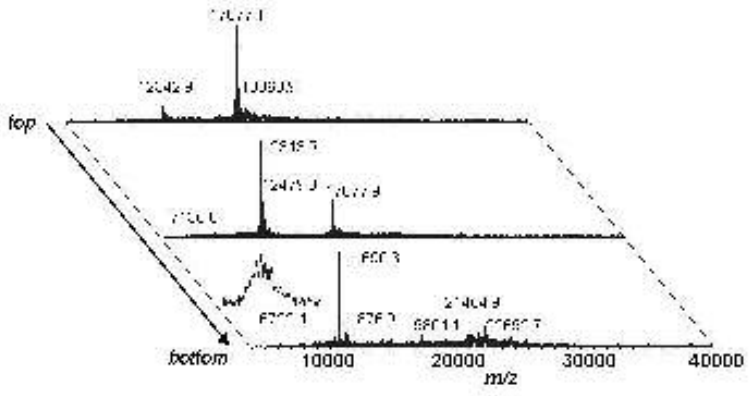


**Fig. 3**

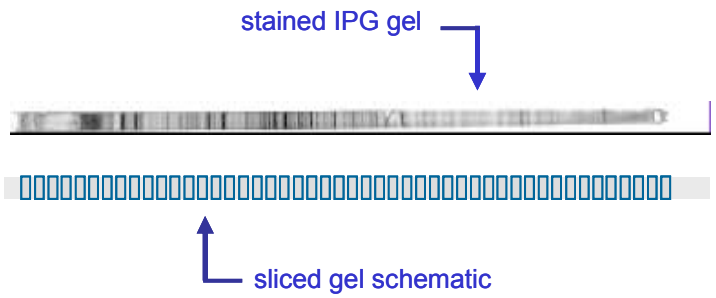


**Fig. 4**

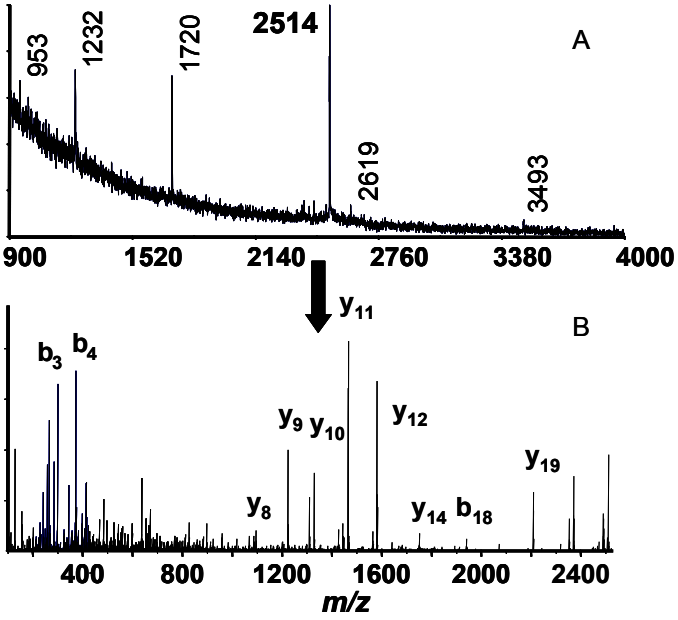
953
1232
1720
2514
2619
3493
A
900
1520
2140
2760
3380
4000
y11
B
b3
b4
y9
y10
y12
y19
y8
y14
b18
400
800
1200
1600
2000
2400
m/z

**Fig. 5**

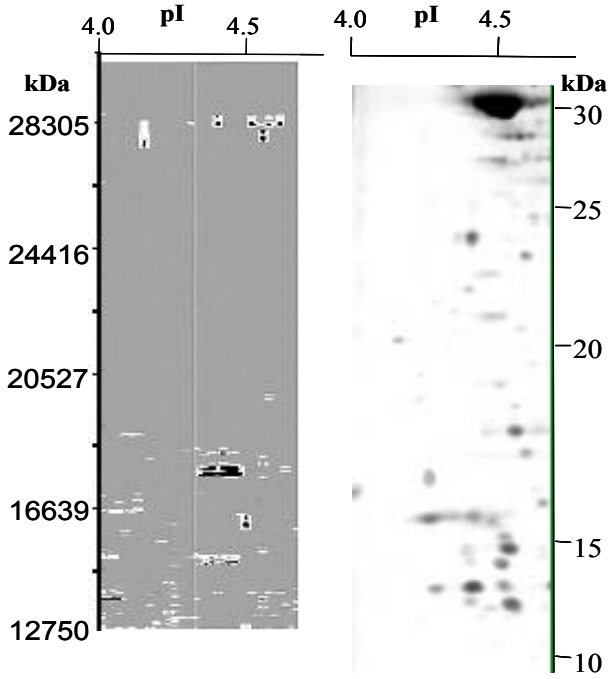
4.0
pI
4.5
4.0
pI
4.5
kDa
28305
24416
20527
16639
12750
kDa
30
25
20
15
10

**Fig. 6**

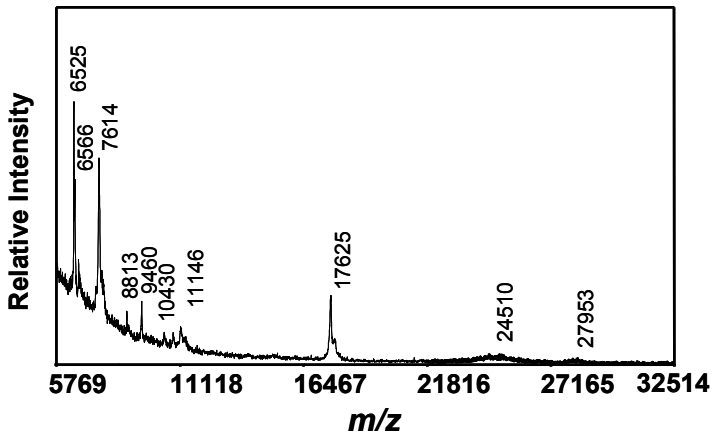
Relative Intensity
6525
6566
7614
8813
9460
10430
11146
17625
24510
27953
5769
11118
16467
21816
27165
32514
m/z

**Fig. 7**