Developments and applications of electrophoresis and small molecule laser desorption ionization mass spectrometry

by

Hui Zhang

A dissertation submitted to the graduate faculty

in partially fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Analytical Chemistry

Program of Study Committee:
Edward S. Yeung, Major Professor
Robert S. Houk
Victor S.-Y. Lin
Surya K. Mallapragada
Basil Nikolau

Iowa State University

Ames, Iowa

2007

Copyright © Hui Zhang, 2007. All rights reserved.

Graduate College

Iowa State Univeristy

This is to certify that the doctoral dissertation of

Hui Zhang

has met the dissertation requirement of Iowa State University

_____

Major Professor

_____

For the Major Program

To My Families

TABLE OF CONTENTS

IS-T 2519

ABSTRACT

Ultra-sensitive native fluorescence detection of proteins with miniaturized one- and two-dimensional polyacrylamide gel electrophoresis was achieved with laser side-entry excitation, which provides both high excitation power and low background level. The detection limit for R-phycoerythrin protein spots in 1-D SDS-PAGE was as low as 15 fg, which corresponds to 40 thousand molecules only. The average detection limit of six standard native proteins was 5 pg per band and the dynamic range spanned more than 3 orders of magnitude. Approximately 150 protein spots from 30 ng of total *Escherichia coli* extraction were detected on a 0.8 cm × 1 cm gel in two-dimensional separation.

Estrogen-DNA adducts as 4-OHE$_1$(E$_2$)-1-N3Ade and 4- OHE1(E2)-2-NacCys were hypothesized as early risk assessment of prostate and breast cancers. Capillary electrophoresis, luminescence/absorption spectroscopy and LC-MS were used to characterize and detect these adducts. Monoclonal antibodies against each individual adduct were developed and used to enrich such compounds from urine samples of prostate and breast cancer patients as well as healthy people. Adduct 4-OHE$_1$-1-N3Ade was detected at much higher level in urine from subjects with prostate cancer patients compared to healthy males. The same adduct and 4-OHE1-2-NacCys were also detected at a much higher level in urine from a woman with breast carcinoma than samples from healthy controls. These two DNA adducts may serve as novel biomarkers for early diagnostic of cancers.

The adsorption properties of R-phycoerythrin (RPE), on the fused-silica surface were studied using capillary electrophoresis (CE) and single molecule spectroscopy. The band shapes and migration times were measured in CE. Adsorption and desorption events were recorded at the single-molecule level by imaging of the evanescent-field layer using total internal reflection. The adsorbed RPE molecules on the fused-silica prism surface were counted with confidence based on ImageJ software. The capacity factor and desorption rate were estimated from the counting results. The mobility-based adsorption isotherms were constructed from both computer simulations and experiments to determine the capacity factor.

Colloidal graphite was investigated as an alternative matrix for laser desorption/ionization and imaging mass spectrometry (MS). Conventional matrix assisted desorption/ionization (MALDI) MS has strong interference peaks in low mass region (<500 Da) and this greatly limits the application of MALDI MS for analysis of small molecules. Graphite assisted laser desorption/ionization (GALDI) provides good sensitivity in the detection of small molecules such as organic acids, flavonoids and oligosaccharides. The detection limit of fatty acids and flavonoids in negative ion mode are in low femtomoles range. Another advantage of graphite over conventional MALDI matrix is that graphite sprayed surfaces are very homogeneous, which is required for IMS purposes. Fruits were chosen to evaluate the practical utility of GALDI since many types of small molecules are present in them. Distribution of these small molecules in fruit slices was investigated by using IMS and IMS/MS.

The possibility of discriminating isomeric disaccharides with different linkage types and different monosaccharide residues (glucose, galactose, and mannose at the non-reducing end) was investigated with tandem mass spectrometry (MS) and linear discriminant analysis (LDA). Acidic fullerene matrix is very suitable for the analysis of disaccharides as it has very clean background in the low mass region. Disaccharides with different linkage types give different tandem mass profiles from various cross-ring fragmentation pathways. Disaccharide isomers with the same linkage type but different monosaccharide residues bear the same fragmentation profiles, but the relative ratios of the fragment ion intensities were distinctively different. Statistical tool as linear discriminant analysis was used to analyze the tandem mass spectra. Disaccharide isomers with both different linkages and different monosaccharide residues were successfully discriminated.

CHAPTER 1. GENERAL INTRODUCTION

Dissertation Organization

This dissertation begins with a general introduction of the history and recent progress in electrophoresis and laser desorption ionization (LDI) mass spectrometry (MS) of small molecules with a list of cited references. The following chapters are arranged in such a way that published papers and manuscript to be submitted are each presented as separated chapters. Among them chapter two, three, and four are projects based on electrophoresis techniques; chapter five and six are projects related to LDI MS. References for each paper or manuscript are attached to the end of each chapter. The last chapter (chapter seven) presents the general conclusions and the dissertation finishes up with acknowledgements.

Electrophoresis

Electrophoresis is a very widely used technique for separating both large biomolecules (proteins and DNAs) and small organic molecules. Electrophoresis refers to the movement of an electrically charged particle under the influence of an electric field. The earliest study for the mobility of ions under electrical field and the theories of the mechanisms of electrophoresis can be traced back to the 19th century.[1] Tiselius developed the moving boundary electrophoresis (later known as zone electrophoresis) experiments for separation of proteins in free solution, which demonstrated the practical aspects of electrophoresis for the first time and later won him the 1948 Nobel prize.[2] Severe band broadening due to the electrical heating (Joule heating) was observed in early electrophoresis. Circulating cooling water (kept at 4°C) was used in Tiselius's apparatus but it was not sufficient to eliminate the heating effect and band broadening. Different supporting media including paper,[3] starch,[4] cellulose membrane,[5] agarose,[6] and polyacrylamide (PA)[7] were introduced to limit the diffusion of analyte bands as well as providing sieving matrix as extra separation power based on the analytes' size. Most of the paper or starch based matrixes have

been abandoned in favor of agarose and PAA gels because of the improved separation and higher loading capacities of later matrixes.

The movement of particles during electrophoresis is due to the two forces acting on them: the electrostatic force (F_e) that accelerates the charged particle, and the frictional force (F_f) that counteracts the electrostatic force and decelerates the charged particle (Figure 1):

$$F_e = Q \cdot E = Z \cdot e_0 \cdot E$$

$$F_f = -f \cdot v$$

in which Q is the net effective charge of the ion (Coulombs), E is the electric field strength ($V \cdot cm^{-1}$), Z is the number of charges on the particle, e_0 is the elementary charge (1.602×10^{-19} Coulombs), v is the ion velocity ($cm \cdot s^{-1}$) and f is the frictional coefficient ($g \cdot s^{-1}$).

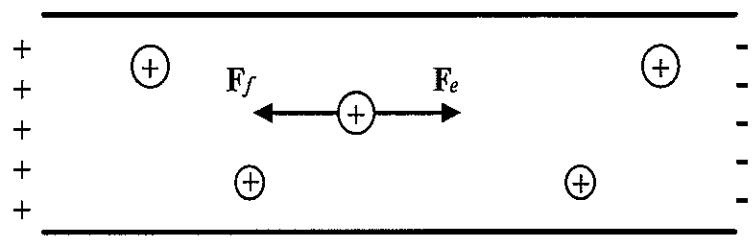


Figure 1. Forces acting on charged particles during electrophoresis

For an undeformable spherical particle moving through free solution and friction-dominant-by-viscosity scenario, the friction coefficient is given by:

$$f = 6 \cdot \pi \cdot \eta \cdot r$$

in which r is the radius of the particle(cm) and η is the viscosity of the matrix (Pa*s).

During electrophoresis, a balance between these two counteractive forces is obtained, at which point the forces are equal. Therefore,

$$F_e = -F_f$$

$$Q \cdot E = 6 \cdot \pi \cdot \eta \cdot r \cdot v$$

The electrophoretic mobility μ_{ep} ($cm^2 \cdot s^{-1} \cdot V^{-1}$) is defined as the velocity per unit electric field as:

$$\mu_{ep} = \frac{v}{E} = \frac{Q}{f} = \frac{Q}{6 \cdot \pi \cdot \eta \cdot r}$$

This equation implies that the electrophoretic mobility of each species in a free solution depends only on its charge and size. With different charge properties and shapes, different kinds of particles or ions will experience different driving and friction forces and migrate at different velocities under the applied electric field. This enables the separation of different species by electrophoresis in free solution.

In gel matrixes, many molecules, especially large biopolymers such as DNA or proteins, can not be considered as hard spherical particles because they may deform and squeeze through the gel network. The migration of such biopolymers can be explained by reptation model in functional dependencies rather than exact equations. The electrophoretic mobility μ_{ep} is still defined as Q/f, in which Q is the net charge for DNAs or proteins and f is the translational frictional coefficient. According to the structure, the total charge Q of DNAs or proteins is proportional to the size N (at proper pH), while the f is proportional to the N^2.[8] This overall will give the relation that

$$\mu_{ep} = \frac{Q}{f} \sim \frac{N}{N^2} = \frac{1}{N}$$

This clearly states that for the biopolymers undergoing electrophoresis in reptile motion, the electrophoretic mobility is inversely proportional to the molecular size. In other words, the larger DNA fragments or proteins will migrate slower than the smaller ones in gel electrophoresis.

Agarose electrophoresis has been the standard method for separation and purification of DNA and RNA fragments.[9] Among all other electrophoresis techniques, polyacrylamide gel electrophoresis (PAGE) and capillary electrophoresis (CE) with various modes are the two most widely used methods and each of them will be introduced separately in the following parts.

Polyacrylamide Gel Electrophoresis (PAGE)

PAGE is an indispensable tool for separation and purification of proteins[10] and DNAs.[11] Polyacrylamide gel can be formed by co-polymerization of acrylamide monomers with a cross-linking reagent—usually N,N'-methylenebisacrylamide (Bis), as shown in Figure 2. The pore size of the gel can be controlled in a reproducible manner by adjusting the total arylamide concentration T and the degree of cross-linking C. Generally, higher percentage gels (higher T), with smaller pores, are needed to separate smaller molecules and *vice versa*. The precise control of the pore size enables the superior resolving power for separation of biomolecules with 0.1% difference in size (i.e., one base difference for a 1kb DNA fragment).

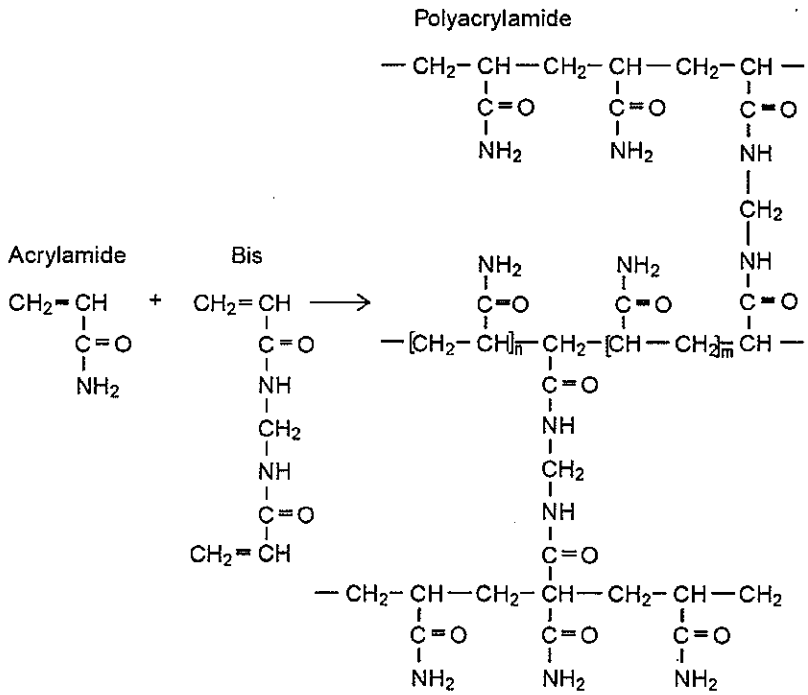


Figure 2. Polyacrylamide gel matrix

Anionic detergent sodium dodecyl sulfate (SDS) is always included for the PAGE of proteins for two major reasons: first, it can disrupt the hydrogen bonds in native proteins to

unfold proteins to linear shape; second, the charge of SDS can mask the charges of proteins and give the SDS-protein complex a constant net charge per mass unit according to the SDS-protein binding ratio as 1.4g SDS per g of protein. Overall a linear relationship between the logarithm of the molecular weight and the relative distance of migration can be achieved, and this enables the determination of molecular weight for unknown proteins.

The proteins can also be separated electrophretically on the basis of their relative abundance of acidic and basic residues. The isoelectric point (pI) is defined as the pH at which the protein has its net charge as zero, so at such pH the protein would have no electrophoretic mobility. Different proteins can be separated according to their different pIs by isoelectric focusing (IEF). The combination of IEF and SDS-PAGE provided one of the most frequently used methods for protein studies as two-dimensional (2-D) SDS-PAGE. As the two separation mechanism are orthogonal, very high resolution can be achieved as to resolve thousands of protein spots on one single slab gel.[10]

After separation with 2-D SDS-PAGE, proteins are usually stained with dyes in order to be detected. Coomassie brilliant blue dyes are the most popular protein staining materials due to their simplicity and reliability. The sensitivity is limited by the high background to 0.1 μg protein per spot.[12] Silver staining is much more sensitive than coomassie brilliant blue staining, with nanogram levels of detection ability.[13] Many new dyes and fluorescence scanners are now commercially available to facilitate sensitive detection based on fluorescence staining. Some fluorescence dyes such as SYPRO Ruby are capable of detecting as low as 1 ng protein per band with three orders of magnitude of dynamic range.[13] Radioactive labeling is the most sensitive method, but very long exposure times are needed for detection of low amounts of proteins.[10] Direct detection of protein spots without staining was possible based on both ultra-violet (UV) absorption[14] as all proteins absorbs strongly at 280nm, and native fluorescence[15] as most proteins contain the amino acid residues tryptophan, tyrosine, or phenylalanine which will fluoresce under UV irradiation.[16] Mass spectrometry has also been used to detect proteins separated by a thin polyacrylamide gel.[17]

Capillary Electrophoresis

It was as early as 1967 when Hjerten[18] predicted that the Joule heating can be dissipated more efficiently if electrophoresis was carried out in narrow diameter tubes. Theoretically peak broadening will be diminished and better separation resolution would be possible. The electrophoresis with very high resolution (with excess of 400,000 theoretical plates) was first demonstrated by Jorgenson and Lukacs[19] in thin glass capillaries with inner diameter of 75μm. Since then CE has been increasingly used for both academic researches and industrial applications. CE has many advantages over the conventional electrophoresis methods, such as extraordinary separation efficiency and resolution, high separation speeds, very low sample and buffer consumption, versatile and sensitive detection options, simplicity of instrumentation and ease of automation. More and more attentions were attracted to this research area in the past two decades. As a proof, there have been steady growth of peer reviewed papers using 'Capillary Electrophoresis' for the past two decades and the total number is over 26,000 by the time this thesis is written.[20]

The major driving force in CE is the electroosmotic flow (EOF) other than the electrostatic force applied directly on the charged particles. At pH higher than 3 silanol groups which cover the inner surface of fused-silica capillary will be negatively charged. Cationic ions in the buffer solution will migrate to the negatively charged capillary wall and a double-layer will be formed at the wall/solution interface. When electrical potential is applied across the capillary, the solvated cations will migrate towards the cathode carrying water molecules (from solvation). The cohesive nature of water due to hydrogen bonding will cause the whole buffer solution to be dragged to the cathode and this phenomenon is called EOF. The EOF can be strong enough to counter the electrostatic attraction of anions to the anode, so that all molecules including cations, neutral ones, and anions can migrate toward cathode.

Capillary zone electrophoresis (CZE) is the basic mode of CE as the separation matrix is free buffer solution only. Analyte ions are separated into discrete zones based on their different apparent mobilities. The separation can be finely tuned by changing the buffer condition (buffer composition, ionic strength, pH value, etc.) and running condition (electric

field, column dimensions, etc.). CZE has been successfully applied for the separation of peptides[21] and proteins,[22] organic acids[23], amino acids and peptides,[19, 24] oligosaccharides,[25, 26] and even inorganic ions.[27, 28]

CZE cannot separate neutral molecules as they will all migrate at the same mobility as EOF drives, nor could it separate those analytes if they have very similar electrophoretic mobility. To overcome these problems, Terabe introduced a new mode of CE as micellar electrokinetic chromatography (MEKC).[29] By adding ionic surfactant to the running buffer, the neutral molecules can form charged complexes with surfactant which will facilitate the migration under electrical field. More importantly, the micelles act as pseudo-stationary phase so that extra separation power is available based on the differential partitioning of analytes between micelle phase and aqueous phase. The movement of MEKC is driven by electric field as in CZE but this method is in fact a hybrid of electrophoresis and chromatography. MEKC renders both super separation efficiency and high separation speed for mixture of both ionic and neutral compounds which are not always available for CZE or conventional chromatography techniques only. The most important optimization of MEKC separation is altering the composition and concentration of surfactant(s) or other host molecules. Commonly used ones include SDS,[29] crown ethers,[30] bile salts,[31] cyclodextrins,[32] etc. In chapter three, crown ether $C_{16}E_8$ was used as surfactant additive for separation of DNA-estrogen adducts.

Simulation for CE

Computer based simulation for CE was possible because the behavior of analyte and buffer ions inside capillary under well controlled physical and chemical environment can be described in detail by three kind of equations: mass conservation, charge conservation or electroneutrality of solution, and various reaction equilibria.[33-35] In each simulation, the appropriate initial and boundary conditions need to be specified, as well as other CE conditions. According to each specific scenario, approximations are usually applied in order to simplify the calculation to a doable manner by current computers within a reasonable time.

By solving sets of nonlinear partial differential equations for mass and charge conservation and algebraic equilibrium equations, the separation of analytes can be simulated with many detailed information such as peak shape and distortion.[36] Simulation can also be used to optimize separation conditions such as which buffer system or injection process to choose, and to explain unusual phenomena such as system peaks.[37, 38] Many CE simulations are one-dimensional (1-D) as they focus on the longitudinal distribution of analytes only but totally ignore the wall adsorption by assuming the constant radial distribution. The assumption is not valid when analytes absorb to the capillary wall. Pseudo two-dimensional (2-D) model was proposed by Ermakov et al. in which the wall adsorption was considered for a very thin adsorbing layer and radial distribution in the rest part was still considered as constant.[39] A true 2-D model was developed by Gas et al. to simulate the peak shapes and variances due to the adsorption.[40] In chapter 4, a real 2-D model was used to predict the adsorption of R-phycoerythrin (RPE) protein on the fused silica capillary wall. Adsorption isotherms were constructed from simulation and were validated by both CE experiments and direct observation of the RPE molecules on fused silica surface by single molecule spectroscopy.

Mass Spectrometry

Mass spectrometry (MS) has been an indispensable tool in chemistry, biochemistry, pharmacy, medicine, environment, and geology.[41] The basic principle of this technique is to generate ions from analytes (ion sources), then to separate the ions usually by the different mass to charge ratio (*m/z*) of ions (mass analyzer), and to accurately detect the number of ions at each *m/z* (detector). MS can be used to detect both small molecules such as inorganic elements or small organic compounds, and large molecules such as proteins and DNAs. It is commonly used to identify unknown compounds by the mass of the molecules as well as the isotopic composition of elements, to elucidate the structure of a compound by observing its fragmentation, and to study the fundamentals of gas phase ion chemistry as well as other physical, chemical, and biological properties of molecules.

One reason for the big success of MS is the versatile options of instrumentation for different applications. For example, different mass analyzers including magnetic sectors,[42]

time of flight (TOF),[43] quadrupole,[44] ion traps (IT)[45, 46] and Fourier transform ion cyclotron resonance (FT-ICR)[47] were utilized as each of them has their own merits. During the past century, various ion sources have been developed: from historic but still widely used ones as electron ionization (EI),[48] field ionization (FI)[49] and chemical ionization (CI)[50] (before 1950s'); to the newer ones as fast atom bombardment ionization (FAB),[51] electrospray ionization (ESI)[52], laser desorption ionization (LDI)[53] and matrix assisted laser desorption ionization (MALDI),[54](1960-1990s'); and to the most recent ambient ionization methods including desorption electrospray ionization (DESI),[55] direct analysis in real time (DART)[56] and atmospheric solids analysis probe (ASAP),[57] (2004-2005). Multi Nobel prizes have been awarded for MS related researches include the 1989 physics award for the development of ion trap, and the 2002 chemistry award for the impact to the biological macromolecules study with two soft ionization methods as ESI and soft LDI. In fact, ESI and MALDI are the two most widely used ionization methods for biomolecules nowadays.

MALDI MS and Small Molecule MALDI

Since the introduction in late 1980s, MALDI has grown exponentially as an indispensable tool used in chemistry and bio-related sciences.[54] The big success of MALDI comes from its soft ionization characteristic, in which large molecules could be preserved; and the simplified spectra as mostly singly charged species are generated by MALDI.[58] Besides proteins[59] and synthetic polymers,[60] MALDI has also been very successful in detecting lipids,[61] carbohydrates and glycans,[62, 63] and nucleic acids and DNAs.[64] In order to be detected by MALDI, analytes are mixed with matrix by so-called cocrystallization process. Matrix molecules are usually small organic acids and have strong absorption at the laser wavelength, such as α-cyano-4-hydroxycinnamic acid (CHCA)[65] or 2,5-dihydroxybenzoic acid (DHB).[66] Upon the laser irradiation, matrix molecules will strongly absorb the laser energy and assist the desorption and ionization of analytes by energy and charge transfer. The real mechanisms which lead to the final ionization of analytes are poorly understood as yet, though models have been proposed. One model assumes the analyte molecules as incorporated in the matrix crystals are neutral.[54] After the matrix molecules are photoionzed

the charge will be transferred to analyte molecules. Recent "lucky survivor" model[62] assumes the analytes are incorporated into matrixes as charged species in the beginning, but only those '"lucky" ones will remain charged and be detected while others will be re-neutralized within the desorbed clusters of matrix and analyte mixture.

It was recognized very early that the analytes should be diluted to tens of thousand times by matrix molecules in order to get good signals.[67] This excess amount of matrix molecules will be ionized and give very strong background peaks upon laser irradiation. As the conventional MALDI matrixes are usually small organic acids with molecular weight less than 500Da, the small *m/z* region of MALDI mass spectra (<500Da) are always omitted for detection of large analytes in order to prevent detector saturation and space charge effect. The presence of such high matrix peaks would also severely suppress the analyte peaks which are in the same region. For these reasons the application of MALDI to small molecules has lagged far behind the application to large ones. To overcome this problem, various approaches have been introduced such as to use organic matrix with a larger molecular weight (e.g., porphyrin with molecular weight of 974.6);[68] to use different inorganic matrixes including metal powders and metal oxides to eliminate matrix peaks;[69-72] and to skip matrix as in desorption ionization on silicon (DIOS).[73, 74] Several carbon based matrixes (colloidal graphite, fullerol, and acidic fullerene) were successfully developed in our lab for detection of small molecules. Detailed discussion can be found in chapter five and six.

Imaging Mass Spectrometry

Imaging MS as it is called, allows one to visualize the 2-D relative concentration profiles of molecules on the surface of biological samples by creating images with a set of mass spectra acquired stepwise with defined lateral assignment for each spectrum. One big advantage of imaging MS over conventional molecular imaging methods is that no labeling is needed and many kinds of molecules can be detected simultaneously.[75]

Various MS techniques have been used for imaging purposes, include secondary ion mass spectrometry (SIMS),[76-79] MALDI,[80-84] and DESI.[85] Decades ago SIMS was used as an

imaging tool for detecting inorganic impurities on silicon surfaces in semiconductor industries.[77] Nowadays SIMS can be used to detect both inorganic elements[76] and small organic molecules.[76, 78] It also provides the best spatial resolution (nm level) among all the three imaging MS methods. The diverse choices of lasers and matrixes make MALDI MS suitable for fast, simultaneous and high-throughput imaging of both small molecules as lipids and peptides and large molecules as proteins from various biological samples.[80-84] The spatial resolution of MALDI IMS is in between that of SIMS and DESI IMS, usually ranging from 80-200 μm.[86] DESI requires the least sample preparation and allows true *in situ* measurement with the simplest instrumentation, but it is destructive to samples and the resolution is worse than other two methods.[85]

The general process of MALDI imaging can be shown as in Figure 3. First, a thin sample slice (normally cut by cryosectioning) is placed on a MALDI plate, and the plate is

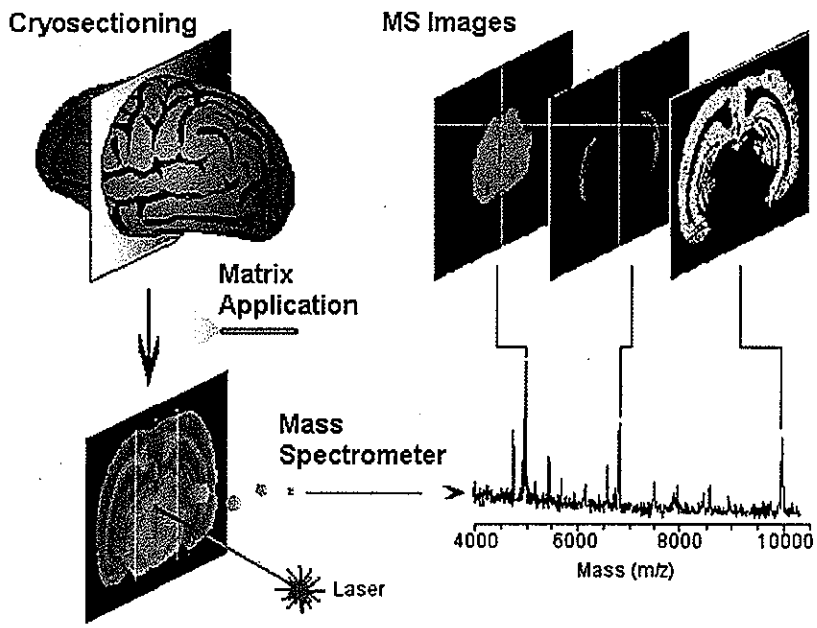


Figure 3. General procedure of MALDI-imaging mass spectrometry[75]

then sprayed uniformly with appropriate matrix. Mass spectra are collected at each of the rastering points of laser beam for a pre-selected region. Special designed imaging software is used to process the collected mass spectra and to generate 2-D optical images.

Due to the limitation of conventional MALDI as discussed previously, it is a big challenge to do MALDI imaging MS for small molecules. With graphite matrix developed in our lab, imaging MS of small organic acids, flavonoids and oligosaccharides were demonstrated with fruit samples in Chapter five.

Structural Identification of Carbohydrates

Carbohydrates are very important biomolecules in nature as they serve crucial structural and functional roles in almost all living organisms.[87] Besides functioning as energy storage for many biological pathways and structural components for cells, carbohydrates and their derivatives also play major roles in molecular recognition and signaling.[87] The very diverse biological activities of carbohydrates depend on their detailed structures, which include the composition of monosaccharide residues, positions of inter-residue linkages and branches, and the anomeric configuration. In many cases it is a must to find the exact structure of carbohydrates, as isomers with very similar structures may function very differently. After the proteins can be routinely sequenced and human genome is essentially complete, more and more ambitious scientists participated in carbohydrates sequencing.[88-90] Many traditional biochemistry methods for carbohydrate sequencing based on the use of different enzymes with various specificities that oligosaccharides can be cleaved at one type of linkage or position but not the others. Those methods required a lot of sample preparation and separation thus were very time consuming.[91] NMR was frequently used to determine the structure of oligosaccharides but pure individual component was needed.[92] Thanks to the invention of two soft ionization methods as MALDI[59] and ESI,[52] carbohydrates can be easily analyzed with MS techniques. Tandem MS can provide detailed structural information of carbohydrates from the fragmentation profiles.[93-98] In the past decades, tandem MS was successfully used to identify the derivatization type and position, to elucidate the branching

sites, and to distinguish different linkage positions.[93-95] The anomeric configuration (α or β) was also distinguished for four pairs of disaccharide isomers recently.[99] However, one fundamental question for carbohydrate sequencing, as to identify the different monosaccharide residues as the building blocks of oligosaccharides, has not been done. Chapter six was devoted to address this problem as differentiating the different monosaccharide residues of disaccharides by using tandem MS and discriminant data analysis.

References

(1) Kohlrausch, F. *Ann. Phys. (Leipzig)* **1897**, *62*, 209.

(2) Tiselius, A. *Trans. Faraday. Soc.* **1937**, *33*, 524-531.

(3) Brattsten, I.; Nilsson, A. *Arkiv for Kemi* **1951**, *3*, 337-345.

(4) Smithies, O. *Biochem. J.* **1955**, *61*, 629-641.

(5) Kohn, J. *Nature* **1957**, *180*, 986-987.

(6) Hjerten, S. *Biochim. Biophys. Acta* **1961**, *53*, 514-517.

(7) Raymond, S.; Weintraub, L. *Science* **1959**, *130*, 711-711.

(8) Lerman, L. S.; Frisch, H. L. *Biopolymers* **1982**, *21*, 995-997.

(9) Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Moleular cloning: a laboratory manual* **1982**, *Cold Spring Laboratory.*

(10) Ofarrell, P. H. *J. Biol. Chem.* **1975**, *250*, 4007-4021.

(11) Sanger, F.; Coulson, A. R. *J. Mol. Biol.* **1975**, *94*, 441-448.

(12) Rabilloud, T. T. *Proteome research: Two-Dimensional Gel Electrophoresis and Identification Methods*; Berlin ; New York : Springer, 2000.

(13) Berggren, K.; Chernokalskaya, E.; Steinberg, T. H.; Kemper, C.; Lopez, M. F.; Diwu, Z.; Haugland, R. P.; Patton, W. F. *Electrophoresis* **2000**, *21*, 2509-2521.

(14) Yamamoto, H.; Nakatani, M.; Shinya, K.; Kim, B. H.; Kakuno, T. *Anal. Biochem.* **1990**, *191*, 58-64.

(15) Sluszny, C.; Yeung, E. S. *Anal. Chem.* **2004**, *76*, 1359-1365.

(16) Roegener, J.; Lutter, P.; Reinhardt, R.; Bluggel, M.; Meyer, H. E.; Anselmetti, D. *Anal. Chem.* **2003**, *75*, 157-159.

(17) Loo, R. R. O.; Stevenson, T. I.; Mitchell, C.; Loo, J. A.; Andrews, P. C. *Anal. Chem.* **1996**, *68*, 1910-1917.

(18) Hjertén, S. *Chromatogr. Rev.* **1967**, *9*, 122-219.

(19) Jorgenson, J. W.; Lukacs, K. D. *Anal. Chem.* **1981**, *53*, 1298-1302.

(20) *ISI Web of Science* http://portal.isiknowledge.com/portal.cgi, Accessed on Oct. 2007.

(21) Jorgenson, J. W.; Lukacs, K. D. *Clin. Chem.* **1981**, *27*, 1551-1553.

(22) Gordon, M. J.; Lee, K. J.; Arias, A. A.; Zare, R. N. *Anal. Chem.* **1991**, *63*, 69-72.

(23) Bondoux, G.; Jandik, P.; Jones, W. R. *J. Chromatogr.* **1992**, *602*, 79-88.

(24) Zacharis, C. K.; Tempels, F. W. A.; Theodoridis, G. A.; Voulgaropoulos, A. N.; Underberg, W. J. M.; Somsen, G. W.; de Jong, G. J. *J. Chromatogr. A* **2006**, *1132*, 297-303.

(25) Liu, J. P.; Shirota, O.; Novotny, M. *Anal. Chem.* **1991**, *63*, 413-417.

(26) Chen, D. H.; Peterson, M. D.; Brumley, R. L.; Giddings, M. C.; Buxton, E. C.; Westphall, M.; Smith, L.; Smith, L. M. *Anal. Chem.* **1995**, *67*, 3405-3411.

(27) Weston, A.; Brown, P. R.; Jandik, P.; Heckenberg, A. L.; Jones, W. R. *J. Chromatogr.* **1992**, *608*, 395-402.

(28) Weston, A.; Brown, P. R.; Heckenberg, A. L.; Jandik, P.; Jones, W. R. *J. Chromatogr.* **1992**, *602*, 249-256.

(29) Terabe, S.; Otsuka, K.; Ichikawa, K.; Tsuchiya, A.; Ando, T. *Anal. Chem.* **1984**, *56*, 111-113.

(30) Markushin, Y.; Gaikwad, N.; Zhang, H.; Kapke, P.; Rogan, E. G.; Cavalieri, E. L.; Trock, B. J.; Pavlovich, C.; Jankowiak, R. *Prostate* **2006**, *66*, 1565-1571.

(31) Okerberg, E. S.; Elshihabi, S.; Carmichael, P. T.; Woody, K. A.; Barckholtz, T. A.; Burke, J. A.; Bushey, M. M. *J. Microcolumn Sep.* **2000**, *12*, 391-397.

(32) Valle, B. C.; Billiot, F. H.; Shamsi, S. A.; Zhu, X. F.; Powe, A. M.; Warner, I. M. *Electrophoresis* **2004**, *25*, 743-752.

(33) Okhonin, V.; Krylova, S. M.; Krylov, S. N. *Anal. Chem.* **2004**, *76*, 1507-1512.

(34) Andreev, V. P.; Lisin, E. E. *Chromatographia* **1993**, *37*, 202-210.

(35) Bier, M.; Palusinski, O. A.; Mosher, R. A.; Saville, D. A. *Science* **1983**, *219*, 1281-1287.

(36) Fang, N.; Zhang, H.; Li, J.; Li, H. W.; Yeung, E. S. *Anal. Chem.* **2007**, *79*, 6047-6054.

(37) Hruska, V.; Stedry, M.; Vcelakova, K.; Lokajova, J.; Tesarova, E.; Jaros, M.; Gas, B. *Electrophoresis* **2006**, *27*, 4610-4617.

(38) Stedry, M.; Jaros, M.; Gas, B. *J. Chromatogr. A* **2002**, *960*, 187-198.

(39) Ermakov, S. V.; Zhukov, M. Y.; Capelli, L.; Righetti, P. G. *J. Chromatogr. A* **1995**, *699*, 297-313.

(40) Gas, B.; Stedry, M.; Rizzi, A.; Kenndler, E. *Electrophoresis* **1995**, *16*, 958-967.

(41) Gross, J. H. *Mass Spectrometry: A Textbook* **2004**, *Berlin ; New York : Springer.*

(42) Ligon, W. V. *Science* **1979**, *205*, 151-159.

(43) Stephens, W. E. *Physical Review* **1946**, *69*, 691-691.

(44) Paul, W.; Steinwedel, H. *Z. Naturforsch.* **1953**, *8A*, 448-450.

(45) Paul, W. *Angew. Chem.* **1990**, *102*, 780-789.

(46) Stafford, G. *J. Am. Soc. Mass. Spectrom.* **2002**, *13*, 589-596.

(47) Comisarow, M. B.; Marshall, A. G. *J. Mass Spectrom.* **1996**, *31*, 581-585.

(48) Field, F. H.; Franklin, J. L. *Electron Impact Phenomena and the Properties of Gaseous Ions* **1957**, *1st ed.*, Academic Press: New York.

(49) Gomer, R.; Inghram, M. G. *J. Am. Chem. Soc.* **1955**, *77*, 500-500.

(50) Munson, M. S. B.; Field, F. H. *J. Am. Chem. Soc.* **1966**, *88*, 2621.

(51) Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. N. *J. Chem. Soc., Chem. Commun.* **1981**, 325-327.

(52) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.

(53) Fenner, N. C.; Daly, N. R. *Rev. Sci. Instrum.* **1966**, *37*, 1068.

(54) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 53-68.

(55) Takats, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G. *Science* **2004**, *306*, 471-473.

(56) Cody, R. B.; Laramee, J. A.; Durst, H. D. *Anal Chem* **2005**, *77*, 2297-2302.

(57) McEwen, C. N.; McKay, R. G.; Larsen, B. S. *Anal Chem* **2005**, *77*, 7826-7831.

(58) Hillenkamp, F.; Peter-Katalinić, J. *MALDI MS : a practical guide to instrumentation, methods and applications* **2007**, *Wiley-VCH.*

(59) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

(60) Bahr, U.; Deppe, A.; Karas, M.; Hillenkamp, F.; Giessmann, U. *Anal. Chem.* **1992**, *64*, 2866-2869.

(61) Cha, S.; Yeung, E. S. *Anal. Chem.* **2007**, *79*, 2373-2385.

(62) Mechref, Y.; Novotny, M. V.; Krishnan, C. *Anal. Chem.* **2003**, *75*, 4895-4903.

(63) Harvey, D. J. *Mass Spectrom. Rev.* **2006**, *25*, 595-662.

(64) Doktycz, M. J.; Hurst, G. B.; Habibigoudarzi, S.; McLuckey, S. A.; Tang, K.; Chen, C. H.; Uziel, M.; Jacobson, K. B.; Woychik, R. P.; Buchanan, M. V. *Anal. Biochem.* **1995**, *230*, 205-214.

(65) Beavis, R. C.; Chaudhary, T.; Chait, B. T. *Organic Mass Spectrom.* **1992**, *27*, 156-158.

(66) Strupat, K.; Karas, M.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1991**, *111*, 89-102.

(67) Beeson, M. D.; Murray, K. K.; Russell, D. H. *Anal. Chem.* **1995**, *67*, 1981-1986.

(68) Ayorinde, F. O.; Hambright, P.; Porter, T. N.; Keith, Q. L. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2474-2479.

(69) Sluszny, C.; Yeung, E. S.; Nikolau, B. J. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 107-115.

(70) Kinumi, T.; Saisu, T.; Takayama, M.; Niwa, H. *J. Mass Spectrom.* **2000**, *35*, 417-422.

(71) Chen, C. T.; Chen, Y. C. *Anal. Chem.* **2005**, *77*, 5912-5919.

(72) McLean, J. A.; Stumpo, K. A.; Russell, D. H. *J. Am. Chem. Soc.* **2005**, *127*, 5304-5305.

(73) Wei, J.; Buriak, J. M.; Siuzdak, G. *Nature* **1999**, *399*, 243-246.

(74) Wen, X. J.; Dagan, S.; Wysocki, V. H. *Anal. Chem.* **2007**, *79*, 434-444.

(75) Stoeckli, M.; Chaurand, P.; Hallahan, D. E.; Caprioli, R. M. *Nat. Med.* **2001**, *7*, 493-496.

(76) Arlinghaus, H. F.; Kriegeskotte, C.; Fartmann, M.; Wittig, A.; Sauerwein, W.; Lipinsky, D. *Appl. Surf. Sci.* **2006**, *252*, 6941-6948.

(77) Castaing, R.; Slodzian, G. *J Microsc* **1962**, *1*, 395-410.

(78) Ostrowski, S. G.; Van Bell, C. T.; Winograd, N.; Ewing, A. G. *Science* **2004**, *305*, 71-73.

(79) Altelaar, A. F. M.; Klinkert, I.; Jalink, K.; de Lange, R. P. J.; Adan, R. A. H.; Heeren, R. M. A.; Piersma, S. R. *Anal. Chem.* **2006**, *78*, 734-742.

(80) Caprioli, R. M.; Farmer, T. B.; Gile, J. *Anal. Chem.* **1997**, *69*, 4751-4760.

(81) Chaurand, P.; Caprioli, R. M. *Electrophoresis* **2002**, *23*, 3125-3135.

(82) Chaurand, P.; Cornett, D. S.; Caprioli, R. M. *Curr. Opin. Biotechnol.* **2006**, *17*, 431-436.

(83) Garrett, T. J.; Prieto-Conaway, M. C.; Kovtoun, V.; Bui, H.; Izgarian, N.; Stafford, G.; Yost, R. A. *Int. J. Mass Spectrom.* **2007**, *260*, 166-176.

(84) Garrett, T. J.; Yost, R. A. *Anal. Chem.* **2006**, *78*, 2465-2469.

(85) Wiseman, J. M.; Ifa, D. R.; Song, Q. Y.; Cooks, R. G. *Angew. Chem. Int. Ed.* **2006**, *45*, 7188-7192.

(86) Zhang, H.; Cha, S.; Yeung, E. S. *Anal. Chem.* **2007**, *79*, 6575-6584.

(87) Berg, J. M.; Tymoczko, J. L.; Lubert, S. *Biochemistry* **2001**, Fifth Edition, International Edition.

(88) Ashline, D.; Singh, S.; Hanneman, A.; Reinhold, V. *Anal. Chem.* **2005**, *77*, 6250-6262.

(89) Harvey, D. J. *Mass Spectrom. Rev.* **1999**, *18*, 349-450.

(90) Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161-227.

(91) Edge, C. J.; Rademacher, T. W.; Wormald, M. R.; Parekh, R. B.; Butters, T. D.; Wing, D. R.; Dwek, R. A. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6338-6342.

(92) Townsend, R. R.; Hilliker, E.; Li, Y. T.; Laine, R. A.; Bell, W. R.; Lee, Y. C. *J. Biol. Chem.* **1982**, *257*, 9704-9710.

(93) Hofmeister, G. E.; Zhou, Z.; Leary, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5964-5970.

(94) Konig, S.; Leary, J. A. *J. Am. Soc. Mass. Spectrom.* **1998**, *9*, 1125-1134.

(95) Sheeley, D. M.; Reinhold, V. N. *Anal. Chem.* **1998**, *70*, 3053-3059.

(96) Lewandrowski, U.; Resemann, A.; Sickmann, A. *Anal. Chem.* **2005**, *77*, 3274-3283.

(97) Yamagaki, T.; Nakanishi, H. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1069-1074.

(98) Ren, S. F.; Guo, Y. L. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 255-260.

(99) Simoes, J.; Domingues, P.; Reis, A.; Nunes, F. M.; Coimbra, M. A.; Domingues, M. R. M. *Anal. Chem.* **2007**, *79*, 5896-5905.

CHAPTER 2. ULTRA-SENSITIVE NATIVE FLUORESCENCE DETECTION OF PROTEINS WITH MINIATURIZED POLYACRYLAMIDE GEL ELECTROPHORESIS BY LASER SIDE-ENTRY EXCITATION

A paper published in Electrophoresis*

Hui Zhang and Edward S. Yeung

Abstract

Direct detection of separated proteins inside polyacrylamide gels has many advantages compared with staining methods. Ultra-sensitive native fluorescence detection of proteins with miniaturized one- and two-dimensional polyacrylamide gel electrophoresis was achieved with laser side-entry excitation. The detection limit for R-phycoerythrin protein spots in 1-D SDS-PAGE with 532 nm excitation was as low as 15 fg, which corresponds to only 40 thousand molecules. The average detection limit of six standard native proteins was 5 pg per band with 275 nm excitation. The dynamic range spanned more than 3 orders of magnitude. By using the same detection setup, approximately 150 protein spots from 30 ng total Escherichia coli extraction were detected on a 0.8 cm × 1 cm gel in two-dimensional separation. The significant improvement in sensitivity for laser side-entry excitation comes from higher excitation power and lower background level compared with other excitation modes.

---

* Reprint with permission from Electrophoresis **2006,** 27(18), 3609-3618.

Copyright © 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Introduction

Gel electrophoresis is one of the most frequently used techniques for the separation of complex biopolymer mixtures such as DNAs and proteins.[1-3] For decades sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) has been the core analytical tool for the separation of protein molecules, estimation of molecular mass and assessment of sample purity.[4] Since two-dimensional (2-D) electrophoresis was introduced by O'Farrell,[5] the method has not changed significantly. First proteins are separated according to their isoelectric points (pI) in the first dimension by isoelectric focusing (IEF). Disulfide bonds in the proteins are then cleaved by thiol reducing agents, the protein subunits are dissociated and the polypeptide chains are saturated with SDS. The resulting complexes with similar charge-to-mass ratios are then separated by their molecular sizes in the second dimension.[6]

Sensitive detection of the separated proteins is always a challenge in SDS-PAGE. Coomassie Brilliant Blue dyes are the most popular protein staining materials due to simplicity and reliability. The sensitivity is limited by the high background to 0.1 μg per protein spot.[6] Silver staining is much more sensitive than Coomassie Brilliant Blue staining, with nanogram levels of detection ability.[7] However, it suffers from some inherent drawbacks such as low dynamic range, tedious staining/destaining steps, and the use of toxic chemicals.[6] Many new dyes and fluorescence scanners are now commercially available to facilitate sensitive detection based on fluorescence staining. Some fluorescence dyes such as SYPRO Ruby are capable of detecting as low as 1 ng protein per band with three orders of magnitude of dynamic range.[7] In all cases, high background is an intrinsic shortcoming and a subsequent destaining step is required. Radioactive labeling (with ^{14}C, ^{3}H, etc.) is the most sensitive method, but very long exposure times are needed for detection of low amounts of proteins.[5] Moreover, samples are limited to those which could be incorporated with radio-active nuclei,[6] and environmental hazards are always an issue.

In proteomics, gels are often analyzed by mass spectrometry after separation.[2] Unfortunately, the coupling of gel separation and mass spectrometry suffers from the presence of dyes or fluorescence labels required to first visualize the proteins.[8, 9] Direct detection methods are more desirable than staining methods as intact proteins are preserved.

Direct detection of protein spots from a thin polyacrylamide gel by mass spectrometry was demonstrated at the 10-ng range.[10] However, mass spectrometry in proteomics is mainly used for analyzing peptide fragments digested from individual protein spots, instead of characterizing all proteins separated in a gel.

Among the 20 common amino acids, tryptophan, phenylalanine, and tyrosine have significant UV absorption between 250 nm and 300 nm. Proteins that contain tryptophan or tyrosine groups fluoresce between 300 nm and 400 nm when excited into these absorption bands.[11] Database search performed by Roegener et al. showed that more than 99% of the 1 million proteins with molecular mass larger than 10 kD have at least one amino acid residue of tryptophan or tyrosine.[12] This makes direct detection methods based on both UV absorption and native fluorescence of proteins possible.[11-16] For example, Yamamoto et al. detected microgram-level protein bands by measuring the 280 nm UV absorbance of proteins in polyacrylamide (PAA) gels.[14] Roegener et al. expanded a frequency-tripled Ti:Sapphire laser (operated at 280 nm) and excited PAA gels from the top for direct fluorescence detection of proteins.[12] Detection limits of 5-10 ng for various protein bands were obtained with 35 mW/cm^2 excitation power. Sluszny et al. coupled micro-electrophoresis with native fluorescence detection with top excitation by a UV lamp.[16] The detection limit achieved in that project was as low as 40 pg per protein spot.

Side-entry excitation has been suggested as a more efficient way for DNA detection in slab gels than top excitation.[17] In this work, the sensitivity of protein detection by native fluorescence was improved to the picogram level with 275 nm laser side-entry excitation. Furthermore, femtogram-level detection limit was achieved for R-phycoerythrin (RPE), B-phycoerythrin (BPE) and Alexa Fluor-532 labeled proteins by 532 nm Nd:YAG laser side-entry excitation.

Materials and Methods

Chemicals

IEF and SDS-PAGE: Dry IEF PhastGel, DTT, Urea, and CHAPS were purchased from Amersham Biosciences (Piscataway, NJ). Acrylamide, bisacrylamide, bromophenol

blue, 1.5 M Tris-HCl solution (pH 8.8) and 0.5 M Tris-HCl solution (pH 6.8) were purchased from Bio-Rad (Hercules, CA). SDS, Tris, glycine, TEMED, β-mercaptoethanol (5% v/v) ammonium persulfate and glycerol were obtained from Sigma (St. Louis, MO). Agarose M was purchased from Amersham. The sample buffer (Laemmli buffer) contained 62.5 mM Tris-HCl at a pH of 6.8, 25% glycerol, 2% SDS and 0.01% bromophenol blue. The gel running buffer contained 25 mM Tris, pH 8.3, 192 mM glycine, and 0.1% SDS. All water was obtained from a MilliQ water purification system (Millipore, Billerica, MA).

Protein samples: Alexa Fluor-532 protein labeling kit, RPE and BPE were purchased from Invitrogen-Molecular Probes (Carlsbad, CA). Protein mixtures containing 2 μg/μl each of six native proteins lysozyme (MW 14,400), trypsin inhibitor (MW 21,500), carbonic anhydrase (MW 31,000), ovalbumin (MW 45,000), serum albumin (MW 66,200) and phosphorylase b (MW 97,000) were purchased from Bio-Rad. Standard proteins of bovine serum albumin (BSA, MW 66,000) and conalbumin (MW 77,000) were purchased from Sigma. Proteins were dissolved to a final concentration of 2 mg/ml and stored at –20 °C until use. *Escherichia coli* (E. coli) protein sample used for 2-D separations was purchased from Bio-Rad.

Gel Electrophoresis

Gel Cassette: Quartz gel cassettes for both 1-D and 2-D electrophoresis (Fig. 1) were manufactured by NSG Precision Cells (Farmingdale, NY). The cassettes were 15 mm long rectangular tubes with 15 mm × 1 mm inner cross section. All surfaces were flat and polished. For 1-D SDS-PAGE experiments, six 0.5 mm holes were drilled in one quartz gel cassette by CiDRA Corporation (Wallingford, CT). The six holes were 1 mm away from the edge of top plate and the separation between holes was 2.2 mm.

Gel Composition: For both 1-D and 2-D electrophoresis, the stacking gel was composed of 4% T and 3.3% C at a pH of 6.8 and the resolving gel was 12% T and 3.3% C at a pH of 8.8, except for the experiments with 532 nm laser excitation in which the resolving gel concentration was 8%T and 3.3% C at a pH of 8.8.

Gels for 1-D SDS PAGE were cast as follows. First, parafilm was stretched to cover the bottom of the quartz gel cassette in order to prevent leaking. Monomer solution of the resolving gel was poured 3 mm-to-top into the sealed gel cassette. Isopropanol was added on top of the resolving gel solution to make the gel surface flat. The gel was polymerized for 30 min in the dark. Then six fused-silica capillaries (1.5 cm long, 360 μm outer diameter and 75 μm inner diameter, Polymicro Technologies, Phoenix, AZ) were inserted through each of the six drilled holes to about 0.2 mm away from the bottom plate of gel cassette. An aluminium holder with six holes that were designed to match those in the quartz gel cassette was used to hold the capillaries and to keep them straight. The stacking gel was then poured into the gel cassette and polymerised for 30 min. The capillaries were then pulled out and injection wells were made with dimensions of 0.8 mm in depth and 0.4 mm in diameter. Prior to injection polymerized gels were photobleached 10 min by a 254 nm mercury lamp (Spectroline, Westbury, NY). Gels for 2-D separations were cast the same way as those for 1-D separation except that no injection wells were made.

Sample Loading and Electrophoresis of 1-D SDS-PAGE: BSA and conalbumin protein were labeled with Alexa Fluor-532 fluorescence dye according to manufacturer's instruction. The two labeled proteins were then mixed with RPE and BPE and diluted to the desired concentration with H_2O. Such a mixture was ready for sample loading in experiments with 532 nm laser excitation. Protein mixtures used in experiments with 275 nm laser excitation were prepared following Laemmli's protocol: pre-mixed protein standard solution from Bio-Rad was diluted to the designated concentration with H_2O. Then, protein mixtures were mixed with sample buffer (final concentration: 50%) and beta-mercaptoethanol (final concentration: 5%) and heated to 100 °C for 10 min. Samples were then cooled down and were ready for injection. The micro-injector was made by connecting a short fused-silica capillary (2 cm long, 150 μm outer diameter and 50 μm inner diameter, Polymicro) with a 0.5 μl syringe (Hamilton Co., Reno, NV), as shown in Fig. 2. A protein solution of 0.08 μl was injected into each injection well made as described previously. Gels were then put inside the running chamber. Melted agarose solution (0.5%, in H_2O) was added in between the gel cassette and running chamber to prevent current leakage. Electrophoresis was carried out horizontally at a constant voltage of 200 volts for 5 min. The whole quartz gel cassette with

the separated proteins inside was taken out from the running chamber, rinsed with water and dried with air. The gel was imaged inside the quartz gel cassette.

2-D SDS-PAGE Separations: Isoelectric focusing was performed according to the manufacturer's instructions. Dry sample proteins extracted from E. coli were dissolved in 8 M urea, 2 mM DTT and 2% CHAPS and stored at –70 °C until use. Dry IEF gels were cut into 12 mm × 1 mm strips. The IEF strip was soaked 1 h with rehydration solution contained 6 M urea, 2% CHAPS, 0.3% (w/v) DTT and 0.002% bromophenol blue. Various amounts of E. coli protein, ranging from 30 ng-2 μg were added to the rehydration buffer solution. After the rehydration step high voltage was applied across the 12 mm gel with the following protocol: 40 V 10 min, 80 V 5 min, 150 V 5 min, 200 V 30 min and 300 V 20 min. Subsequently, IEF strips were equilibrated for 15 min in a solution containing 50 mM Tris-HCl (pH 8.8), 6 M urea, 30% glycerol, 2% SDS, 1% DTT and 0.002% bromophenol blue. The second-dimension separation (i.e., SDS-PAGE) was performed similarly to the 1-D experiments after the IEF strips were delivered on top of the stacking gels. The gels were then rinsed and imaged.

Fluorescence Detection

The detection set-up is presented in Fig. 3. The gel inside the quartz cassette was placed horizontally on a holder. Either the 275 nm laser line isolated from an argon-ion laser (Model 2045, Spectra-Physics, Mountain View, CA) or the 532 nm Nd:YAG laser (Uniphase, San Jose, CA) was used to excite the gels from the side. The laser beam was expanded horizontally by the first cylindrical lens (40.0 mm × 25.4 mm, f = 3 cm, CVI Laser, Albuquerque, NM) and then focused vertically by the second cylindrical lens (25.4 mm in diameter, f = 15 cm, CVI Laser) into a thin sheet. The dimensions of the laser beam at the gel were approximately 15 mm wide and 300 μm thick. The laser passed through the gel perpendicular to the side of the gel cassette to ensure maximum transmittance of light through the gel. The thin laser beam was prevented from hitting the top or the bottom plate of quartz gel cassette.

UV Fluorescence Detection: For experiments with UV laser, the 275 nm laser line was isolated from other laser lines and plasma emission by a prism. All UV mirrors and UV cylindrical lenses were purchased from CVI laser. The average irradiation power measured at the side entrance of the gel was 60-120 mW/cm^2. In order to reduce background light, the set-up was contained in a closed box. The irradiated gel was imaged by a 105 mm UV camera lens (f = 4.5, Nikon, Japan) onto a back illuminated, 16 bit CCD camera (TE/CCD-512-TKB, Princeton Instruments, Princeton, NJ), which had an approximate quantum efficiency of 40% for UV light. A long-pass color glass filter (WG-305, Schott Glass, Duryea, PA) and a band-pass filter (330WB80, Omega Filters, Brattleboro, VT) were used to select the fluorescence of native proteins. The exposure of the CCD was set at 10-60 s according to the fluorescence intensity emitted from the gels.

Visible Fluorescence Detection: Experiments with 532 nm laser side-entry excitation were based on a similar detection setup as shown in Fig. 3 except all the optics were designed for visible light. All the mirrors were Nd:YAG High Energy Laser Mirrors (CVI). Two cylindrical lenses had the same specifications as those for the UV experiments except that no UV coating was applied. The average irradiation power measured at the side entrance of the gel was 50-200 mW/cm^2. The irradiated gel was imaged by a 28 mm Nikkor camera lens (AF28/1.4D). A 532 holographic notch filter (Kaiser Optical Systems, Ann Arbor, MI) and a band pass filter (580WB80, Omega Filters) were used to select the protein fluorescence. The CCD exposure was set at 1-60 s according to the fluorescence intensity emitted from the gels. Laser top-entry excitation was also carried out by expanding the laser beam with a symmetric convex lens (5 mm diameter, f = 4.5mm, Melles Griot, Rochester, NY). The enlarged laser beam was then introduced onto the gel cassette from the top at an excitation power of about 15 mW/cm^2. The same filters and camera lens as in the experiments with 532 nm side-entry excitation were used for imaging.

Data Processing

Image analysis was performed with the Winview/32 (Princeton Instruments) and Pdquest software (Bio-Rad). Quantification was based on the integrated fluorescence

intensity of each protein spot after background subtraction and normalization to the excitation power.

Safety concerns

Care should be taken when handling acrylamide monomers as they are toxic. Appropriate protective eyewear is required for working with both lasers used in this research.

Results and Discussion

General Concern and Optimization of System

Injection: Compared with conventional gels, miniaturized gels can provide higher speeds of separation, higher sensitivity and smaller quantities of reagents consumed [16]. However sample loading for miniaturized gels is quite challenging. In normal SDS-PAGE experiments, injection wells for sample introduction are in the order of several mm and several tens of μL sample are injected into each well. The miniaturized gels here are only 15 mm × 15 mm × 1 mm. Various approaches have been suggested for sample delivery into ultra-thin or miniaturized slab gels. For example, Stein et al. [18] fabricated 0.1 μl injection wells with a sharp comb made by spark erosion of a razor blade. Zheng et al. [19] applied a specially designed ceramic comb to form rectangularly shaped wells, which enabled the use of a regular micropipette for sample delivery. Alternatively, Liu and Sweedler [20] and Hietpas et al. [21] used a capillary to deliver samples electrokinetically into ultra-thin gels. Our group developed a method to cast small injection wells by immersing a capillary tip inside the gel during polymerization. Then, 0.1 μl of sample solution was delivered through a capillary splitter assembly.[16] Here we simplify the injection process by replacing the capillary splitter with a home-made micro-injector. Precise liquid deliveries in the range of 20-100 nL were achieved.

Optimization of Laser Beam: Various excitation sources could be utilized for the detection of biopolymers separated in gels, including lamps, lasers, light emitting diodes, etc.

Among them lasers bear many advantages due to high excitation power, low divergence and monochromatic characteristics. Lasers have been applied to excite both native proteins and labeled proteins in gels.[12, 22] In both cases lasers were introduced from the top and did not provide comparable sensitivity as classic staining methods. Also, the gel image had to be assembled from several sequential images because only partial irradiation of the whole gel was achieved due to difficulties in focusing and concerns about excitation power loss. For the miniaturized gels, the laser beam could be expanded easily to excite the whole gel without significant power decrease. Introducing the laser beam by side entry has been applied to DNA separations,[17] but no analogous protein detection has been reported before.

Ideally the laser beam should be able to excite all the proteins inside the entire gel area efficiently and it should not hit the top or bottom plates of the gel cassette to give high background. So the laser beam was expanded horizontally by the first cylindrical lens to match the gel width, as shown in Fig. 3. Meanwhile the distance between the gel and the first cylindrical lens should not be too far in order to prevent over-expansion and excitation power loss. The expanded laser beam had to be focused vertically to fit the 1 mm thick gel. Two parameters are involved, namely a thin beam of $2w_0$ (Gaussian beam radius) at the focal plane and a large depth of focus (DOF), which represents the distance in the direction of propagation where the laser beam thickness is $<\sqrt{2}$ times of the beam waist. Compromise was made since both could not be achieved simultaneously. With the optics used in our experimental setup, the laser beam was maintained to less than 0.3 mm in thickness for at least 2 cm long distance.

Absorption by Gel Matrix and Fluorescence Background: One big issue for side-entry excitation is the attenuation of the laser beam as it traveled through the gel matrix. Considering the low amount of protein loading (pg to ng levels), the laser excitation power was not affected much as a result of absorption by proteins. Both monomers of acrylamide and PAA gels are totally transparent of 532 nm. Hence proteins loaded in different lanes experience the same excitation power in experiments with 532 nm laser excitation. For experiments at 275 nm, polyacrylamide became non-transparent for UV light. Acrylamide monomer also has strong absorption below 300 nm due to the C=C double bonds. During

polymerization, absorbance in the 260-280 nm region decreases gradually as the polymer chain increases. However, even the 12% polyacrylamide gel polymerized from the best acrylamide monomer we could obtain commercially still had absorption of about 0.45 for 1 cm beam path at 275 nm. Normalization with respect to the laser power loss is thus necessary, as described in detail later.

Similarly, PAA gel matrix has negligible fluorescence background under 532 nm excitation, but the gel matrix fluoresces considerably under 275 nm excitation. It was observed that brief photobleaching under a 254 nm mercury lamp could decrease the fluorescence background of the gel matrix to a moderate level. The residual fluorescence background is mostly in the 300-360 nm region, which falls in the same region as the native fluorescence of proteins. So there are no suitable spatial filters to selectively block such background light. Fortunately, the excitation power provided by side-entry configuration was so high that as low as 20 pg per protein band could still be detected over fairly high background levels, as presented *vide infra*.

Visible Fluorescence Detection in 1-D SDS-PAGE with 532 nm Laser Side Entry Excitation

1-D SDS-PAGE experiments with visible fluorescence were carried out to demonstrate the superior sensitivity provided by laser side-entry excitation. Phycobiliproteins like R-phycoerythrin (RPE) and B-phycoerythrin (BPE) found in cyanobacteria and several groups of eucaryotic algae [23] have strong visible fluorescence in the 550-650 nm region when excited by 450-600 nm visible light. Two native proteins (BSA and conalbumin) were labeled with Alexa Fluor-532 dye to give visible fluorescence at similar wavelengths. Labeled proteins were mixed with RPE and BPE, and then diluted to the desired concentration and injected into each lane. As shown in Fig. 4A, all four protein spots were easily visible in the 0.8 pg lane. Signal-to-noise (S/N) ratio of 200 fg spots of labeled BSA, labeled conalbumin, BPE and RPE were 15, 3, 20 and 40, respectively. The calculated detection limit for each of the four proteins were 40 fg, 200 fg, 30 fg and 15 fg at S/N = 3,

which correspond to only 100, 500, 80, 40 thousand of protein molecules each. It is worth mentioning that the concentration of the two labeled proteins was calculated by assuming 100% labeling and recovery efficiency from the labeling reaction, which in practice could be much lower. For comparison, laser top-entry excitation was also carried out after beam expansion by a convex lens. Protein spots below the 1 pg level were not visible, as shown in Fig. 4B. With the same detection setup as employed by Sluszny et al.,[16] 530 nm-550 nm light was also selected from a 500 W Hg (Xe) lamp (Oriel, Stamford, CT) and utilized to excite the gels from the top. The sensitivity achieved was even lower than laser top excitation, presumably due to the lower excitation power.

The excitation power density in laser side-entry excitation was much higher than top entry excitation, as the laser was focused more tightly. A careful comparison between the two gel images in Fig. 4 also indicates that after normalization with respect to the excitation power, the side-entry configuration provides lower background levels than the top-entry configuration. A major part of background light in the gel images came from reflection and scattering from the gel surface. By focusing the laser beam into a thin sheet and applying side-entry excitation, such reflection and scattering were decreased effectively. The noise level, which was dominated by the fluctuation of background light in this case, was decreased as well. This combination led to lower detection limits for side-entry excitation.

Calibration for Heterogeneous Excitation Power

PAA matrix has negligible absorption at 532 nm so that excitation power is homogeneous across the gel. However, it has considerable absorption at 275 nm. As a result, excitation power for different locations of the gel decreases exponentially as the laser travels through it, as is clearly shown in Fig 5A. By fitting the background level to an exponential decay model, absorption by the PAA gel matrix could be accounted for. This result agreed well with measurements made with a spectrometer. Control experiments were carried out by loading trypsin inhibitor protein into each lane as internal standards. The integrated fluorescence signal of each protein spot was plotted as shaded columns in the histogram of

Fig. 5B. Clearly, the fluorescence intensity of protein spots had the same decreasing trend as the background and could be fitted to the same exponential decay model. By assigning different correction factors to different pixels in the raw gel images, excitation power at different locations of the gel could thus be calibrated. The process was further simplified by assigning one correction factor for each protein spot. Such calibration for the control experiments proved to be valid and effective, as the blank columns show in Fig. 5B. All the 1-D SDS-PAGE gels with 275 nm laser side-entry excitation were calibrated by this means.

Native Fluorescence Detection of 1-D SDS-PAGE with 275 nm Laser Side-Entry Excitation

The one-dimensional SDS-PAGE separation of the six standard proteins was presented in Fig. 6. In order to prevent saturation of the CCD, gels with higher protein loading were imaged with shorter exposure times than those with lower loading. Fig. 6 emphasizes the superior detection capabilities of native fluorescence detection with laser side-entry excitation compared to those of standard staining methods or other fluorescence detection methods. All six protein spots (20 pg each) could be clearly visualized in Fig. 6B. For example, the S/N of the 20 pg carbonic anhydrase spot was about 18, indicating a detection limit of less than 4 pg (S/N = 3). All six proteins gave average detection limits of about 5 pg, which is about 4 and 2 orders of magnitude better than those of the coomassie stain and silver staining, respectively.

The detection limit for proteins with UV fluorescence is considerably higher than those with visible fluorescence. The major difference is that the very low background level in visible fluorescence detection is not achievable with UV laser excitation. The PAA gel matrix has high fluorescence background under 275 nm excitation. In addition, the filters are much less effective in resolving protein fluorescence from background compared to those used in experiments with 532 nm laser excitation.

Fig. 7 presents the calibration plots for the 1-D SDS-PAGE separations. The high-range data points of 0.8 ng-16 ng correspond to Fig. 6A, and the low-range data points of 20

pg-0.8 ng correspond to Fig. 6B. The fluorescence signal of each protein spot was calculated by subtracting the adjacent background level from the total signal. Calibration was applied to correct for the attenuation of excitation laser power, as discussed previously. The 0.8 ng lane is present in both high-range gels and low-range gels, so proteins in this lane were used as internal standards to account for the different exposure times. Fig. 7 shows that the system exhibits a large and linear dynamic range, from the pg region to an upper limit of 16 ng. This feature is also better than the corresponding dynamic range obtained by silver staining and coomassie dyes.[6, 24]

Native Fluorescence Detection of 2-D SDS-PAGE with 275 nm Laser Side-Entry Excitation

Previous work of our group demonstrated the feasibility of miniature gels to separate complex protein mixtures.[16] 2 cm × 1.5 cm miniaturized gels provided similar resolution as commercial mini-gel (7 cm × 7 cm) but with better sensitivity than conventional staining methods. About 115 protein spots were detected with 0.1 μg protein loading when the gel was irradiated by UV lamp via top-entry excitation.[16] Fig. 8 represents the 2-D separation of E. coli protein extracts with 1 μg, 0.25 μg, and 30 ng loading. Images in the 2-D experiments were not calibrated for laser excitation power as in 1-D experiments, but there is no reason that such calibration could not be done in the same way for quantitative analysis. Within the 10 mm × 8 mm gel area, the number of protein spots detected with the Pdquest software were 280, 280, 250, 210 and 150 for 1 μg, 0.25 μg, 0.125 μg (data not shown), 65 ng (data not shown) and 30 ng of protein loading, respectively. Protein loading was calculated by multiplying the protein concentration in the rehydration solution by the volume of solution taken up by the dry IEF gel, which was about 6 μl for 12 mm × 1 mm gel strips. These results underscore the high sensitivity of native protein fluorescence detection with laser side-entry excitation.

Conclusions

Native fluorescence detection of separated protein spots in gel electrophoresis by side-entry excitation has many advantages over staining methods, such as high speed, high sensitivity, elimination of tedious staining/destaining or labeling steps, and compatibility with subsequent mass spectrometric analysis. The method was applied to both one and two-dimensional separations of native proteins and Alexa Fluor-532 labeled proteins. In 1-D SDS-PAGE, a simple micro-injector was designed for precise delivery of less than 0.1 μl of sample solution. Reproducible results were obtained for various protein loadings. The detection limit for 6 native proteins and two Alexa Fluor-532 labeled proteins were as low as 15 fg. E. coli protein extracts were separated by two-dimensional electrophoresis. On a 10 mm × 8 mm area, as many as 280 and 150 protein spots were detected for 1 μg and 30 ng protein loading, respectively.

When compared with other excitation modes, including lamp top-entry excitation and laser top-entry excitation, laser side-entry excitation provided higher excitation power and lower background levels. As a result, significant improvements in sensitivity were achieved. Although manual operation of the system is presented here, automation of the 2-D separation was presented recently by Xu et al.[25] Throughput will be improved as well as sensitivity, if such automated separation is coupled to this fluorescence detection method with laser side-entry excitation.

Acknowledgements

We thank Dr. Chanan Sluszny, Aoshuang Xu and Guoxin Lu for helpful discussions. E.S.Y. thanks the Robert Allen Wright Endowment for Excellence for support. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. W-7405-Eng-82. This work was supported by the Director of Science, Office of Basic Energy Sciences, Division of Chemical Sciences.

References

(1) Laemmli, U. K. *Nature* **1970**, *227*, 680-685.

(2) Rabilloud, T. *Proteomics* **2002**, *2*, 3-10.

(3) Guttman, A.; Ronai, Z. *Electrophoresis* **2000**, *21*, 3952-3964.

(4) Ganesh, G.; Kumar, T. K. S.; Pandian, S. T. K.; Yu, C. *J. Biochem. Bioph. Methods* **2000**, *46*, 31-38.

(5) Ofarrell, P. H. *J. Biol. Chem.* **1975**, *250*, 4007-4021.

(6) Rabilloud, T. *Proteome Research: Two Dimensional Gel Electrophoresis and Identification Methods*; Springer-Verlag: Berlin, 2000.

(7) Berggren, K.; Chernokalskaya, E.; Steinberg, T. H.; Kemper, C.; Lopez, M. F.; Diwu, Z.; Haugland, R. P.; Patton, W. F. *Electrophoresis* **2000**, *21*, 2509-2521.

(8) Rabilloud, T.; Strub, J. M.; Luche, S.; Van Dorsselaer, A.; Lunardi, J. *Proteomics* **2001**, *1*, 699-704.

(9) Richert, S.; Luche, S.; Chevallet, M.; Van Dorsselaer, A.; Leize-Wagner, E.; Rabilloud, T. *Proteomics* **2004**, *4*, 909-916.

(10) Loo, R. R. O.; Stevenson, T. I.; Mitchell, C.; Loo, J. A.; Andrews, P. C. *Anal. Chem.* **1996**, *68*, 1910-1917.

(11) Koutny, L. B.; Yeung, E. S. *Anal. Chem.* **1993**, *65*, 183-187.

(12) Roegener, J.; Lutter, P.; Reinhardt, R.; Bluggel, M.; Meyer, H. E.; Anselmetti, D. *Anal. Chem.* **2003**, *75*, 157-159.

(13) Hogan, B. L.; Yeung, E. S. *Appl. Spectrosc.* **1989**, *43*, 349-350.

(14) Yamamoto, H.; Nakatani, M.; Shinya, K.; Kim, B. H.; Kakuno, T. *Anal. Biochem.* **1990**, *191*, 58-64.

(15) Kazmin, D.; Edwards, R. A.; Turner, R. J.; Larson, E.; Starkey, J. *Anal. Biochem.* **2002**, *301*, 91-96.

(16) Sluszny, C.; Yeung, E. S. *Anal. Chem.* **2004**, *76*, 1359-1365.

(17) Chen, D. H.; Peterson, M. D.; Brumley, R. L.; Giddings, M. C.; Buxton, E. C.; Westphall, M.; Smith, L.; Smith, L. M. *Anal. Chem.* **1995**, *67*, 3405-3411.

(18) Stein, A.; Hill, S. A.; Cheng, Z. Q.; Bina, M. *Nucleic Acids Res.* **1998**, *26*, 452-455.

(19) Zheng, J. J.; Odake, T.; Kitamori, T.; Sawada, T. *Anal. Chem.* **1999**, *71*, 5003-5008.

(20) Liu, Y. M.; Sweedler, J. V. *Anal. Chem.* **1996**, *68*, 3928-3933.

(21) Hietpas, P. B.; Bullard, K. M.; Gutman, D. A.; Ewing, A. G. *Anal. Chem.* **1997**, *69*, 2292-2298.

(22) Guttman, A.; Csapo, Z.; Robbins, D. *Proteomics* **2002**, *2*, 469-474.

(23) Apt, K. E.; Collier, J. L.; Grossman, A. R. *J. Mol. Biol.* **1995**, *248*, 79-96.

(24) Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248-254.

(25) Xu, A. H.; Sluszny, C.; Yeung, E. S. *J. Chromatogr. A* **2005**, *1087*, 177-182.

Figure Captions

Figure 1. Gel running chamber. The gel running chamber has dimensions of 50 mm × 65 mm. Volume of the buffer containers was 2 mL each. Inset: quartz gel cassette for 1-D SDS-PAGE. Dimensions of separation gel were 15 mm × 15 mm × 1 mm. Six 0.5 mm diameter injection holes were drilled on the top plate of the cassette with 2.2 mm separation. Gel cassettes for 2-D separation were the same except no injection holes were drilled.

Figure 2. Micro-injector made by connecting a 0.5 μl syringe with a short fused-silica capillary. A 3-D translational stage was employed for precise positioning.

Figure 3. Experimental setup for fluorescence detection of proteins in miniaturized gels with laser side-entry excitation.

Figure 4. Visible fluorescence detection in 1-D SDS-PAGE with a 532 nm laser. Alexa Fluor-532 labeled BSA and conalbumin were mixed with RPE and BPE at the same concentration. Different amounts of each protein were loaded into each lane as indicated on top of the images. Images were taken of the same gel by: A, 532 nm laser side-entry excitation; the laser entered from right side with excitation power of 100 mW/cm^2; and B, 532 nm laser top-entry excitation with excitation power of 15 mW/cm^2. Both images were taken with 10 s exposure, and both images were inverted.

Figure 5. A, background level of a blank gel excited at 275 nm with the laser entering from the right side. B, fluorescence intensity of five trypsin inhibitor protein spots with 275 nm laser side-entry excitation. Shaded columns represent the data from the raw image, and blank columns represent the normalized data. Protein loading for each lane was 1.0 ng except lane 5 which was 0.8 ng. Bottom: raw gel image of five trypsin inhibitor protein spots.

Figure 6. Native fluorescence detection in one-dimensional miniaturized gels with 275 nm laser side-entry excitation. Samples were divided into two groups: A, high range, 16 ng-0.8 ng per protein band and B, low range, 800 pg-20 pg per protein band. The amounts of the protein bands were indicated on top of both images. For both gels, the laser entered from the right side with excitation power of 90 mW/cm^2. The high range gel image was taken with

10 s exposure and the low range gel was taken with 30 s exposure. Both images were inverted and the contrasts of images were adjusted separately for proper visualization.

Figure 7. Calibration plots for 1-D SDS-PAGE separations with UV laser side-entry excitation. Each point represents the average of three replicates.

Figure 8. Native fluorescence detection of proteins in 2-D gels by 275 nm laser side-entry excitation. A, 1 μg E. coli protein extract; the image was taken with 60 mW/cm^2 excitation power and 30 s exposure; B, 0.25 μg E. coli protein extract; the image was taken with 90 mW/cm^2 excitation power and 30 s exposure; and C, 30 ng E. coli protein extract; the image was taken with 120 mW/cm^2 excitation power and 30 s exposure. All the images were inverted.

Figure 1

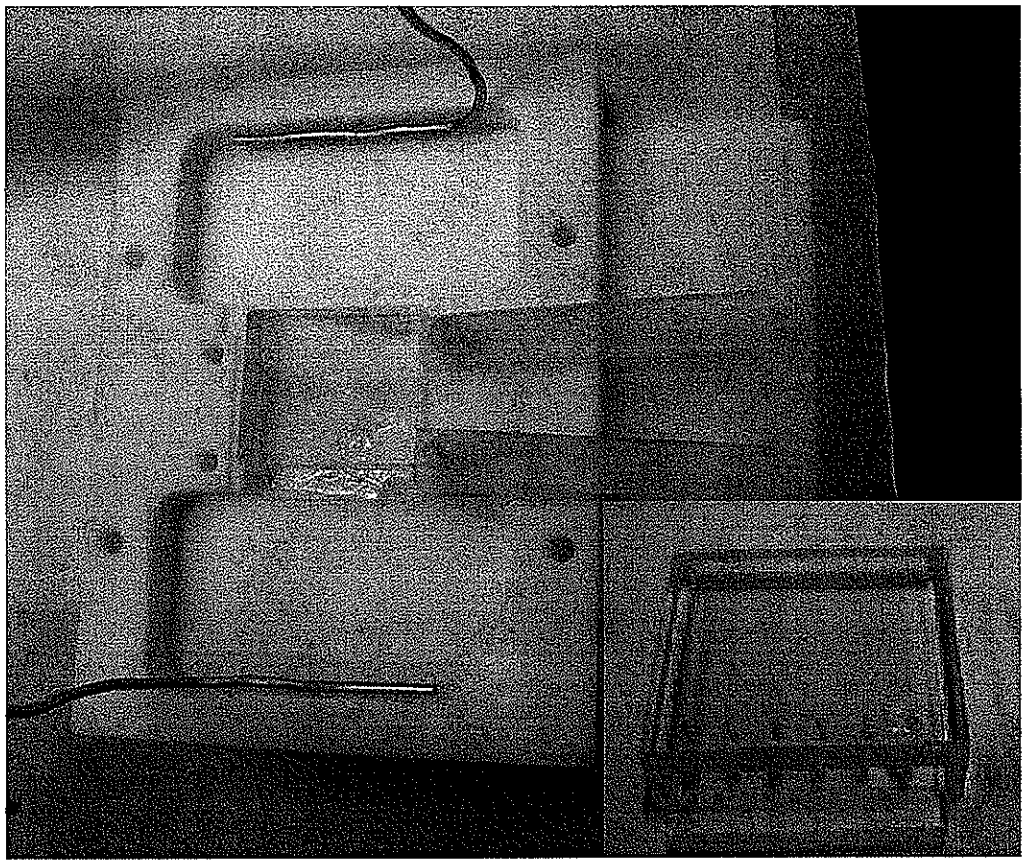


Figure 2

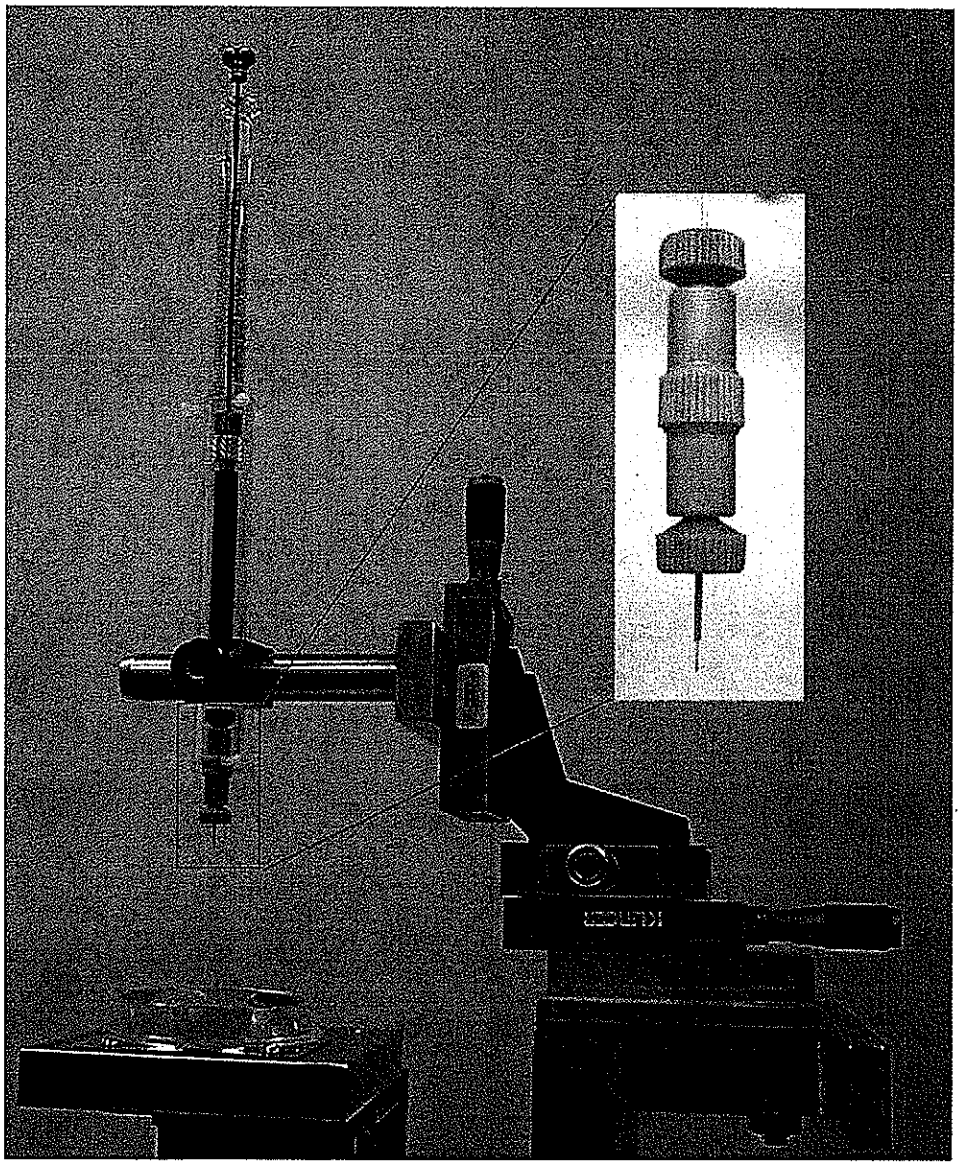


Figure 3

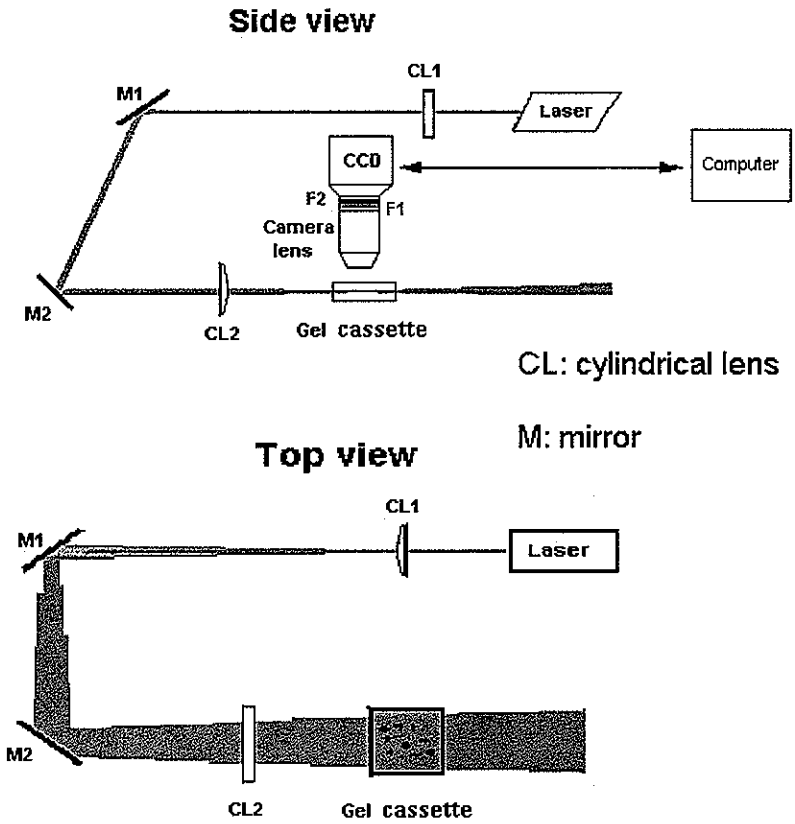

Side view
CL1
M1
Laser
CCD
Computer
F2
F1
Camera lens
M2
CL2
Gel cassette
CL: cylindrical lens
M: mirror
Top view
CL1
M1
Laser
M2
CL2
Gel cassette

Figure 4

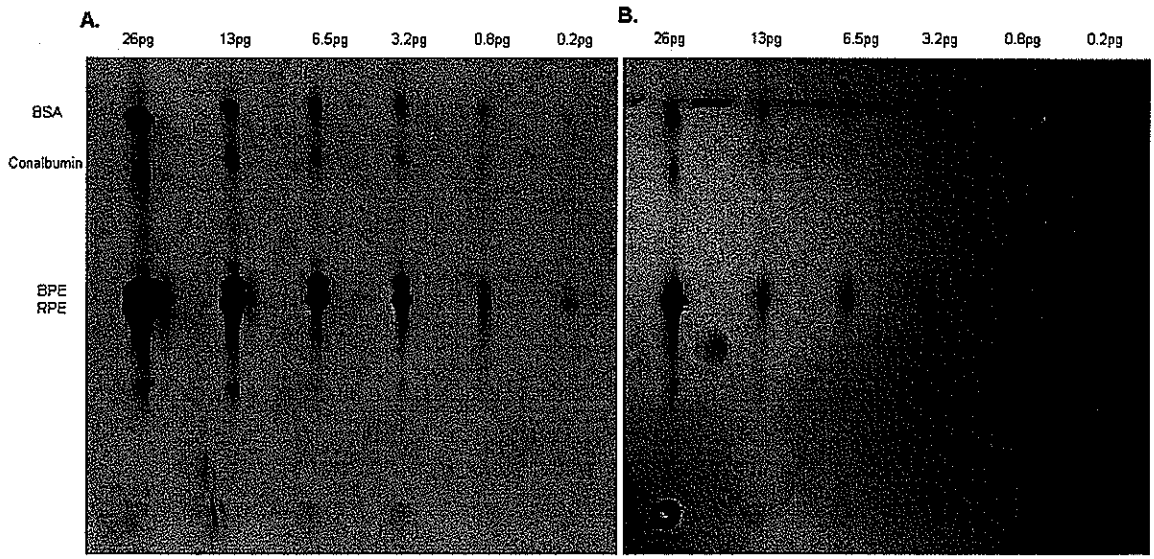
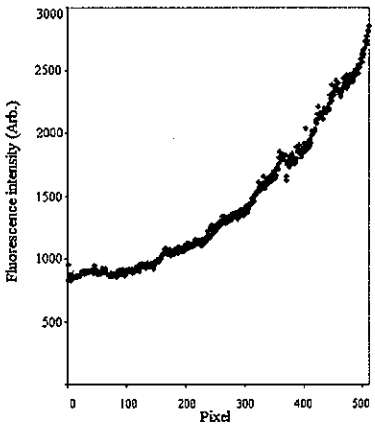

A.
26pg
13pg
6.5pg
3.2pg
0.8pg
0.2pg
BSA
Conalbumin
BPE
RPE
B.
26pg
13pg
6.5pg
3.2pg
0.8pg
0.2pg

Figure 5

A. Decay of excitation power: background level in one blank gel



B. Fluorescence intensity of 5 trypsin inhibitor protein spots

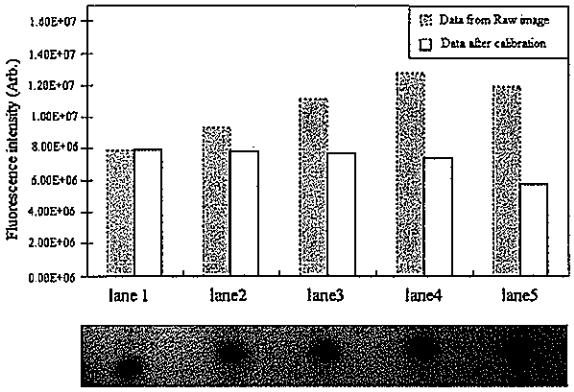


Figure 6

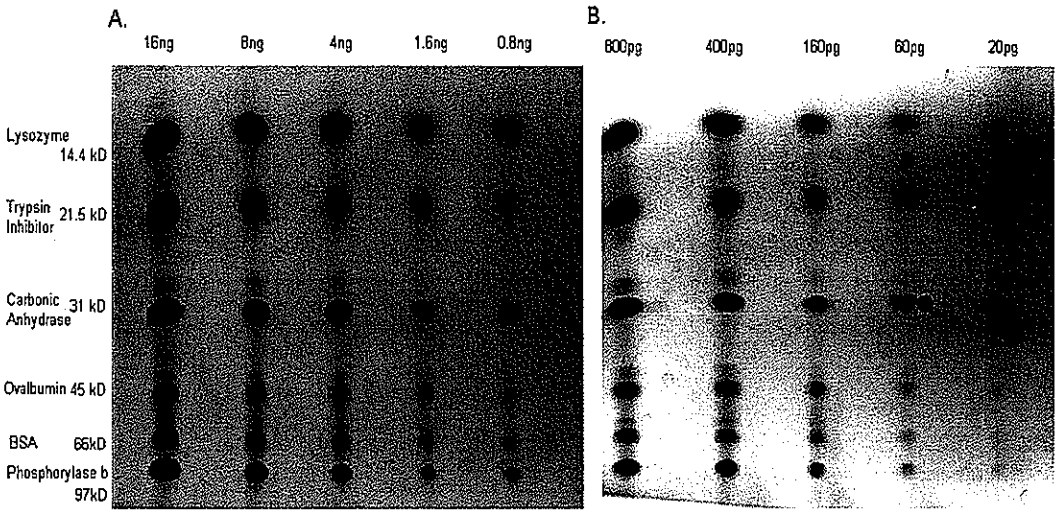

A.
16ng
8ng
4ng
1.6ng
0.8ng
B.
800pg
400pg
160pg
60pg
20pg
Lysozyme 14.4 kD
Trypsin Inhibitor 21.5 kD
Carbonic Anhydrase 31 kD
Ovalbumin 45 kD
BSA 66kD
Phosphorylase b 97kD

Figure 7

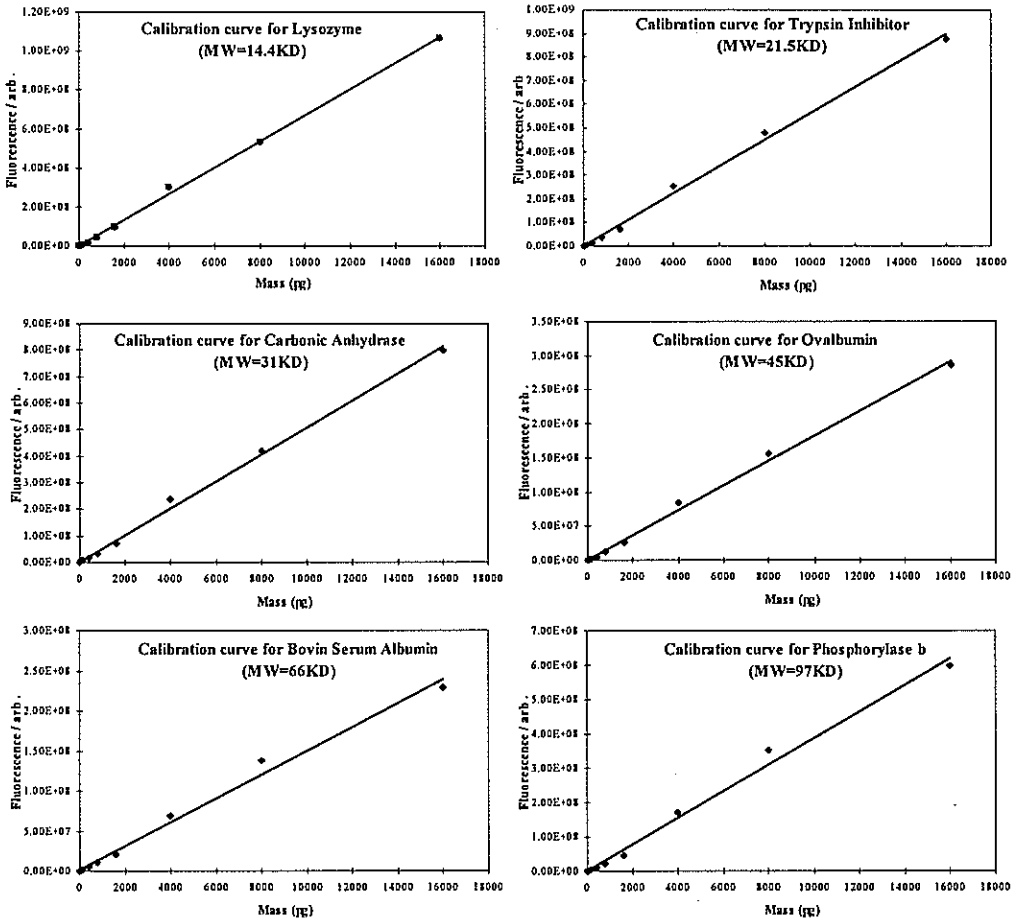
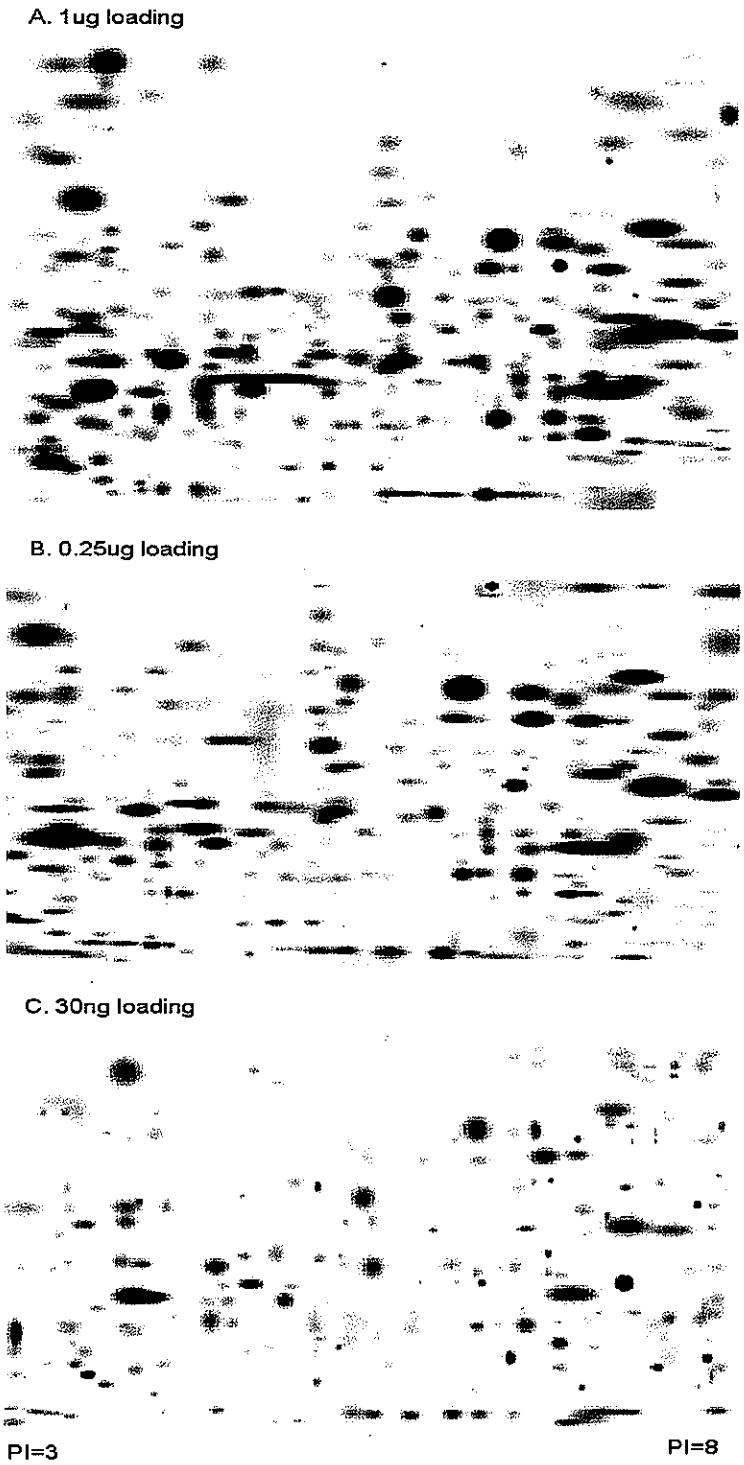

Calibration curve for Lysozyme
(MW=14.4KD)
Calibration curve for Trypsin Inhibitor
(MW=21.5KD)
Calibration curve for Carbonic Anhydrase
(MW=31KD)
Calibration curve for Ovalbumin
(MW=45KD)
Calibration curve for Bovin Serum Albumin
(MW=66KD)
Calibration curve for Phosphorylase b
(MW=97KD)
Fluorescence / arb.
Mass (pg)

Figure 8



CHAPTER 3. DETECTION OF NOVEL BIOMARKERS FOR EARLY RISK ASSESSMENT OF PROSTATE AND BREAST CANCERS

Papers published in Chemical Research in Toxicology and The Prostate*

Y. Markushin, H. Zhang, P. Kapke, N. Gaikwad, E.G. Rogan, E.L. Cavalieri, E.S. Yeung, and R. Jankowiak

Abstract

Estrogens, including the natural hormones estrone (E1) and estradiol (E2) are thought to be involved in tumor induction. Catechol estrogen quinones (CEQ) derived from 4-hydroxyestrone (4-OHE1) and 4-hydroxyestradiol (4-OHE2) react with DNA and form depurinating –N7Gua and –N3Ade adducts. This damage leads to mutations that can initiate breast and prostate cancer. To determine whether this type of DNA damage occurs in humans, urine samples from prostate and breast cancer patients and healthy subjects were analyzed in a blind study. A primary objective was to determine whether any of the cancer patients had formed 4-OHE1(E2)-1-N3Ade or 4-OHE1(E2)-2-NacCys, which were suspected to be the major adduct formed by CEQ. Monoclonal antibodies (MAb) to 4-OHE1(E2)-2-NacCys and 4-OHE1(E2)-1-N3Ade were developed and characterized. The interested adducts were extracted from the urine samples using affinity columns equipped with the specifically developed monoclonal antibodies (MAb). A CE spiking procedure with synthesized DNA adduct-standards, absorption/ luminescence spectroscopies, and mass spectrometry were used to identify the biomarkers of interest. The depurinating DNA adduct 4-OHE_1-1-N3Ade was detected at much higher level in urine samples from subjects with

---

* Modified from Chemical Research in Toxicology **2005**, *18*, 1520-1527. Copyright © 2007 American Chemical Society & The Prostate **2006**, *66*, 1565-1571. Copyright © 2007 Wiley-Liss, Inc.

prostate cancer and urological condition compared to healthy males. The same adduct and 4-OHE1-2-NacCys were also identified in urine samples from a woman with breast carcinoma; and no adduct was detected in urine from a control woman. Since the level of the two CEQ-derived DNA adducts is significantly elevated in the urine of cancer patients, we suggest that their presence could be used as an indicator of breast and prostate cancer risk.

Introduction

The natural estrogens, estrone (E1) and estradiol (E2), are metabolized at the 2- or 4-position with the formation of catechol estrogens, which, in turn, are metabolically oxidized into catechol estrogen quinones (CEQ). The latter have been implicated in the etiology of human breast cancer by various research studies.[1-5] The reaction of the CEQ, in particular CE-3,4-Q, with DNA forms the depurinating 4-OHE$_1$(E$_2$)-1-N3Ade and 4-OHE$_1$(E$_2$)-1-N7Gua adducts.[1, 5, 6] These two adducts constitute >99% of the total adducts formed. The CE-2,3-Q form only small amounts of the depurinating 2-OHE$_1$(E$_2$)-6-N3Ade adducts.[7] These adducts generate apurinic sites that may lead to cancer-initiating mutations,[8, 9] which transform cells,[10, 11] thereby initiating cancer.

Exposure to estrogens is a well-established risk factor for breast cancer.[12] The possible role of estrogens in prostate cancer is thus far less well-established than in breast cancer. Limited evidence exists of an association between estrogens and risk of prostate cancer.[13] One important piece of evidence is the higher level of circulating estrogens observed in African-American men, who have a two-fold higher risk of prostate cancer, compared to European-American men.[14] More direct evidence in support of the role of estrogens in prostate carcinogenesis comes from experiments using Noble rats treated with testosterone plus E$_2$.[15, 16] This combined treatment induces ductal adenocarcinoma of the prostate in 100% of the rats,[15] whereas treatment with only testosterone causes prostate cancer in only 40% of the rats. Treatment with 5α-dehydrotestosterone, which unlike testosterone cannot be converted to E$_2$, results in only a 4% incidence of prostate cancer.

These carcinomas have been suggested to arise from estrogen-induced initiation and testosterone-produced promotion of the prostate tissue.[13]

In our earlier work, CE-3,4-Q-derived DNA adducts were identified in tissue extracts from breast cancer patients.[17] In this case, samples were analyzed by CE interfaced with room temperature absorption and low-temperature (laser-excited) phosphorescence spectroscopies. The level of the 4-OHE1-1-N3Ade in the breast tissue extracted from a patient with breast carcinoma (8.40±0.05 pmol/g of tissue) was larger by a factor of about 30 than the level in the breast tissue sample from a woman without breast cancer (0.25±0.05 pmol/g of tissue).[17] Although more breast tissue samples from women with and without breast cancer need to be studied, these results suggested that the -N3Ade adducts could serve as biomarkers to predict the risk of breast cancer.

To determine whether this type of DNA damage occurs in men, urine samples from subjects with prostate cancer, benign tumors, benign prostate hyperplasia, and a urological condition, as well as healthy males, were analyzed in a blind study. A primary objective was determining whether any of the subjects had formed 4-OHE1(E2)-1-N3Ade and 4-OHE1(E2)-2-NAcCys (Fig. 1), as the major adducts formed by CE- 3,4-Q. We showed that CE-3,4-Q-derived DNA adducts are present in human urine samples, and that their identification can be accomplished by a combination of techniques.

Materials and Methods

Caution: Catechol estrogen quinones are hazardous chemicals and should be handled carefully and in accordance with NIH guidelines.

Chemicals and CEQ-DNA Adduct Standards: 4-OHE_1 and 4-OHE_2 were synthesized according to Dwivedy et al.[18] The 4-OHE_1- and 4-OHE_2-derived DNA adduct standards were synthesized as previously described.[6, 19] Structural analysis of the above standards was accomplished via NMR and mass spectrometry (MS).[6, 19] Ultra-pure grade

glycerol was obtained from Spectrum Chemical (Gardena, CA). The purity of standards for CEQ-derived DNA adduct standards, originally separated by HPLC, was verified in our laboratory by capillary electrophoresis (CE) and low-temperature luminescence spectroscopy. CEQ-derived DNA adducts, which are heat- and oxygen-senstive, were kept for longer-term storage at –80°C under inert atmosphere (N_2 or Ar). Samples were dissolved in methanol:buffer (80:20), with the following buffer content: 0.1 M ammonium acetate and 1mg/L ascorbic acid in nanopure water, pH 4.5. Tris-[hydroxymethyl]-aminomethane was purchased from Fisher Scientific (Fairlawn, NJ). Phosphoric acid and polyoxyethylene 8 cetyl ether ($C_{16}E_8$) were obtained from Sigma-Aldrich (St. Louis, MO).

Monoclonal Antibodies (MAb): Ovalbumin (OA) and keyhole limpet hemocyanin (KLH) were purchased from Pierce Biotechnology, Inc., Rockford, IL. Delbecco's Modified Eagle medium and horse serum were purchased from Mediatech, Inc., Herndon, VA, and Valley Biomedical, Inc., Winchester, VA, respectively. *N*-(9-Fluorenyl) methoxycarbonyl multiple antigenic peptides (Fmoc MAP) resin was purchased from Applied Biosystems, Foster City, CA. Well-established methods[20] were used to generate an immune response in the mice. The 4-$OHE_1(E_2)$-2-NAcCys-16α, β-MCC linker was conjugated to KLH and used in an immunization protocol with 25 μg of antigen/mouse/injection using Freund's incomplete adjuvant. Serum titers were established using 4-$OHE_1(E_2)$-2-NAcCys conjugated to OA. Mouse spleen cells were fused with an equal number of SP2/O cells (40 million of each) and plated in 16 × 96 well microtiter plates. When hybridoma wells started to turn yellow, the plates were screened using an ELISA assay. Five hundred nanograms of OA-4-$OHE_1(E_2)$-2-NAcCys-16α, β-MCC in binding buffer (100 mM $NaHCO_3$, pH 9.3) was used to coat each well of a Nunc maxisorb plate. An affinity column was made to purify MAb by immobilizing the 4-$OHE_1(E_2)$-2-NAcCys-16 α, β-MCC on a MAP resin core used to commonly synthesize peptides. The hapten was immobilized on the MAP resin bead column using the same chemistry used to attach it to the carrier proteins.[21] This column was used to purify MAb by passage of 3 mL of the supernatant fluid from the selected hybridoma over the column. The column was washed with 50 mL of PBS, and antibody was eluted with 100

mM acetic acid, pH 2.5. The eluted antibody was isotyped using a kit specific for mouse antibody, confirming that the antibody was of mouse origin and not from the horse serum used in growing the cells. This purified MAb was immobilized on an agarose bead column (Aminolink kit, Pierce Inc.) and used to detect 4-OHE_1-2-NAcCys in PBS buffer that was spiked with various concentrations of the conjugate.

The 4-OHE1-1-N3Ade-16α,β-MCC linker was conjugated to KLH and used in an immunization protocol with 25 μg of antigen/mouse/injection using Freund's incomplete adjuvant. Serum titers were established using 4-OHE1-1-N3Ade conjugated to OA. Mouse spleen cells were fused with an equal number of SP2/O cells (40 million of each) and plated in 16 x 96-well microtiter plates. When hybridoma wells started to turn yellow, the plates were screened using an ELISA assay. Five hundred ng of OA-4-OHE_1-1-N3Ade-16α,β-MCC in binding buffer (100 mM $NaHCO_3$, pH 9.3) was used to coat each well of a Nunc maxisorb plate. MAb were produced by immunizing mice with 4-OHE_1-1-N3Ade attached to an appropriate linker that was conjugated to keyhole limpet hemocyanin (KLH). Hybridoma cell lines were screened using 4-OHE_1-1-N3Ade conjugated to ovalbumin (OA). Since there is no immunological cross reactivity between KLH and OA, positive hybridoma cell lines secreting antibody against 4-$OHE_1(E_2)$-1-N3Ade could be rapidly identified using OA–4-OHE_1-1-N3Ade. Affinity columns were developed and used to purify MAb against 4-OHE_1-1-N3Ade. MAbs were concentrated by centrifugal filters (Millipore Corporation, Bedford, MA) using Amicon Ultra 100,000 molecular weight cut. This allowed separating solutes from low MW compounds. The purified MAb was immobilized on an agarose bead column. We have demonstrated that this developed MAb binds with high affinity to 4-OHE_1-1-N3Ade adducts ($K_a = 10^8\ M^{-1}$, data to be published elsewhere), highly discriminating against a large spectrum of closely related CEQ-derived metabolites, and has been used here for detecting 4-$OHE_1(E_2)$-DNA adducts in urine samples. Columns were used to capture and preconcentrate the hapten of interest from the urine samples.

Capillary Electrophoresis: The analysis of urine extracts was done using a P/ACE MDQ capillary electrophoresis (CE) system (Beckmancoulter, CA) with a photodiode array

(PDA) detector for simultaneous detection of electropherograms and UV absorption spectra of separated analytes. Bare fused-silica capillary (Polymicro Technologies, Phoenix, AZ) with 21cm effective length and 31.2 cm total length (75 μm I.D. and 360 μm O.D.) was used. The running buffer was 0.5% $C_{16}E_8$ in 0.25 mM Tris-phosphate (pH 3.5). Before injection, the solvent in the sample was evaporated by vacuum pump; the sample residue was then diluted with the same volume of 75 μM H_3PO_4 solution. The same extract was also separated with an ISCO (Lincoln, NE) model 3140 Electropherograph System, and re-analyzed by low-temperature luminescence spectroscopy. The CEQ-derived DNA adduct standards, and the extracts from the immunoaffinity columns were analyzed with field amplified sample stacking (FASS) conditions. FASS[22, 23] was used for analyte preconcentration. To achieve reproducible and accurate stacking results, a water plug was injected into the capillary before the sample (at 0.2 psi for 12 seconds) followed by the electrokinetic injection of urine extract sample at +10 kV for 30 seconds. The applied electric field for separation was 480 V/cm, and the running temperature was 25°C. The absorption detection was set at the PDA mode to obtain the electropherograms under different UV wavelengths and the absorption spectra of the separated analytes. After each run, the capillary was rinsed with 0.1 M NaOH for 2 min, and running buffer for 5 min. Electropherograms were obtained in the absorbance mode. CE-separated DNA adducts were identified based on the characteristic migration times and corresponding absorption spectra. Various detection wavelengths for the CE electropherograms were utilized (e.g. 214, 260, and 276 nm).

Luminescence and Absorption Spectroscopy: Luminescence spectra were obtained using an excitation wavelength of 257 nm of a Lexel 95-SHG-257 CW laser. Emission was dispersed by a Model 218 0.3-m monochromator (McPherson, Acton, MA), equipped with a 300 G/mm grating, providing a resolution of ~1 nm. Spectra were detected with an intensified CCD camera (Princeton Instruments, Trenton, NJ) using gated and non-gated modes of detection. A fast shutter, operated by a Uniblitz driver control (model SD-12 2B), was synchronized with the CCD camera (ICCD-1024 MLDG-E1) and used for time-resolved phosphorescence measurements. Using this setup, time-resolved phosphorescence spectra could be measured in 0.5 sec intervals with a gate width of 0.5 sec. To ensure good glass

formation in off-line spectroscopic measurements, glycerol (50% by volume) was added to the samples in buffer just prior to cooling to 77 K in a liquid nitrogen optical cryostat with suprasil optical windows. Samples (ca. 20 μL) were contained in suprasil tubes (2-mm i.d.). All spectra were background corrected.

HPLC and Mass Spectrometry: The sample was diluted 1:10 in 50% methanol in water and analyzed twice by Ultra performance liquid chromatography. The parent-daughter transition used in analysis was *m/z*=420.1 ->*m/z*=296.0. Instruments: Waters Acquity Binary solvent manager and Sample Manager Micromass Quattro Micro mass spectrometer Gradient: 80% (H_2O, 0.1% Formic Acid), 20% (Acetonitrile, 0.1% Formic Acid) to 79% (H_2O, 0.1% Formic Acid), 21% (Acetonitrile, 0.1% Formic Acid) in 4 minutes, 79% (H_2O, 0.1% Formic Acid), 21% (Acetonitrile, 0.1% Formic Acid) to 45% (H_2O, 0.1% Formic Acid), 55% (Acetonitrile, 0.1% Formic Acid) in 6 minutes. Standard curve and quantitation was done using QuanLynx v4.0.

Results and Discussion

MAb Raised Against 4-OHE1(E2)-2-NAcCys Conjugates. Mice were immunized with KLH-[4-$OHE_1(E_2)$-2-NAcCys]. KLH was used as the carrier to immunize the mice because of its extremely large size (molecular mass over 5 million Da). The small hapten molecule (i.e., 4-$OHE_1(E_2)$-2-NAcCys, which is not immunogenic alone) attached to a large protein carrier (KLH) made the complex suitable for immunogenic reaction. The mice were tested for an immune response to 4-OHE_1-2-NAcCys using OA-[4-$OHE_1(E_2)$-2-NAcCys] and OA alone. Mice demonstrated an elevated antibody titer with OA-[4-$OHE_1(E_2)$-2-NAcCys] as compared to OA alone (data not shown). The mouse with the highest titer was IP boosted with KLH-4-$OHE_1(E_2)$-2-NAcCys and used for hybridoma production. After fusing the immunized mouse spleen cells to the SP2/O cells to make the hybridomas and plating in 16 × 96 well plates, positive hybridoma wells were identified by ELISA using OA-4-$OHE_1(E_2)$-2-NAcCys as an antigen to immobilize captured MAb. Most of the wells had optical density (OD) values of less than 0.1. However, the wells that had hybridomas

secreting antibody to the 4-$OHE_1(E_2)$-2-NAcCys hapten were quite apparent (data not shown). That is, the wells where the antibody was produced had much higher OD values, typically in the range of 0.5-1.0, clearly indicating that antibody(assigned as 2E9) was produced against the 4-$OHE_1(E_2)$-2-NAcCys conjugates.

Detection of 4-OHE1-2-NAcCys and Related Analytes Released from the 2E9 MAb-Based Column. To further evaluate the binding affinities of the 2E9 MAb developed for the detection of 4-$OHE_1(E_2)$-2-NAcCys conjugates, CE has been used to analyze a water-based buffer sample spiked with five analytes of interest: 1, 4-OHE_1-1-N3Ade; 2, 4-OHE_1; 3, 4-OHE_2; 4, 4-OHE_1-2-NAcCys; and 5, NAcCys. The concentration of analytes 1, 2, 3, and 5 used for the CE separation was about 10^{-6} M, while the concentration of the key analyte of interest was smaller by a factor of 100, i.e., 10^{-8} M. The corresponding room temperature CE absorbance-based electropherogram (λ_{obs} = 214 nm) is shown in Figure 2 (curve a). The solid arrow in Figure 2 indicates the position of analyte 4 as confirmed by standard spiking procedure with a higher concentration of 4-OHE_1-2-NAcCys (data not shown). As expected, the peak corresponding to this analyte is hardly discernible in curve a of Figure 2. A specially prepared 2E9 MAb-based affinity column was used to capture and concentrate the 4-OHE_1-2-NAcCys out of the above solution. The recovery rate for 4-OHE_1-2-NAcCys (based on the integrated phosphorescence intensity measurements obtained for various aliquots eluted from the affinity column) was about 80% (data not shown). The resulting electropherogram, shown as curve b in Figure 2, shows that mostly one analyte has been preconcentrated by the affinity column. This peak (near 5 min) corresponds to the 4-OHE_1-2-NAcCys conjugate, as confirmed by standard spiking procedures and phosphorescence spectroscopy (not shown for brevity). The very small peaks near 3.5 and 10 min most likely correspond to peaks 1-3 and 5, respectively; however, only the identity of peak 5 was confirmed by the spiking procedure. Because the sample corresponding to electropherogram b (Figure 2) was preconcentrated by 2 orders of magnitude, we conclude that the binding efficiency for analytes 1, 2, 3, and 5 is negligibly small. Comparison of the integrated peak intensities in the electropherograms a and b reveals that the column preferentially captures the 4-OHE_1-2-NAcCys conjugate.

CE with FASS and Absorbance/Phosphorescence Detection in Spiked Urine Samples. Figure 3 shows four absorbance-based CE electropherograms to further demonstrate the selectivity of the MAb raised against one of the analytes of interest, i.e., 4-$OHE_1(E_2)$-2-NAcCys. Spectrum a corresponds to a CE electropherogram obtained for a mixture of 4-OHE_1-1-N3Ade (peak 1; $c = 5 \times 10^{-5}$ M), 4-OHE_1-2-NAcCys (peak 2; $c = 10^{-4}$ M), 4-OHE_1 (peak 3; $c = 5 \times 10^{-5}$ M), and 4-OHE_1-1-N7Gua (peak 4; $c = 10^{-5}$ M) in a buffer solution. Curve b is the electropherogram of PBS buffer (2 mL) spiked with the mixture of the above four analytes diluted by a factor of 100 and run through the 2E9 MAb-based affinity column. Only peak 2 is observed, with an ~80% efficiency of recovery. Two orders of magnitude higher concentrations of 4-OHE_1-1-N3Ade, 4-OHE_1, and 4-OHE_1-1-N7Gua in comparison with that of 4-OHE_1-2-NAcCys provided similar recovery of the latter compound (data not shown). Spectrum c shows another CE electropherogram obtained after a 10-fold buffer-diluted human urine sample was spiked with 4-OHE_1-2-NAcCys ($c = 10^{-6}$ M) and subsequently run through the affinity column. A remarkably simple CE electropherogram was obtained with the major peak (#2) corresponding to 4-OHE_1-2-NAcCys. The identification of this peak was confirmed by the standard spiking procedure. Again, an excellent recovery of ~80% was obtained. Finally, curve d in Figure 3 was obtained for the 4-OHE_1-2-NAcCys standard ($c = 10^{-4}$ M) and is shown for comparison. These data clearly demonstrate that very efficient recovery of analytes of interest can be obtained, which, in combination with the various separation and identification methods described in this article, should provide the means for analyzing human samples.

Detection of 4-OHE1(E2)-1-N3Ade Adducts in Urine from Subjects with Prostate Cancer and Urological Condition and Healthy Individuals. Immunoaffinity columns were used to purify the antibody by passing 3 mL of supernatant fluid from the selected hybridoma over the column. Then the columns were washed with 50 mL of PBS and antibody eluted with 100 mM acetic acid, pH 2.5. The eluted antibody was isotyped using a kit specific for mouse IgG antibody, confirming that the antibody was of mouse origin and

not from the horse serum used in growing the cells. This purified MAb (assigned as 15G8) was immobilized on an agarose bead column (Aminolink kit, Pierce Inc.) and used to detect 4-OHE$_1$-1-N3Ade in a PBS buffer that was spiked with various concentrations of –N3Ade adduct. Characterizing of the 15G8 antibody and affinity column with that antibody were carried in a similar manner as discussed previously for antibody 2E9 (data not shown).

Urine samples (20 mL each) from sixteen subjects were analyzed in blind studies using different detection methods as described above, as listed in Table 1. Among them six were newly diagnosed prostate cancer patients; two had benign urological conditions (erectile dysfunction or benign prostatic hyperplasia); two were undergoing prestate biopsy (later proved to be both benign tumors, or negative) and five were from health people as controls. All specimens were initially analyzed using affinity column purification, i.e. the adducts of interest were extracted from the urine samples using home-built columns equipped with 15G8 MAbs. Eluted extracts from immunoaffinity columns were studied by laser-exited low-temperature phosphorescence spectroscopy and liquid chromatography (LC) interfaced with mass spectrometry (LC/MS/MS). In addition, urine samples after lyophilization and methanol extraction were pre-concentrated and analytes therein were separated by capillary electrophoresis (CE) with field amplified sample stacking (FASS) and detected by absorbance-based electropherograms. A spiking procedure with synthesized DNA adduct-standards and absorption/ luminescence spectroscopies were used to identify the biomarkers of interest.

The bars in Figure 4 (row #1) correspond to the integrated (normalized) area of the absorbance based CE electropherogram peaks assigned to 4-OHE$_1$(E$_2$)-1-N3Ade. Only the samples from: i) the subjects with prostate cancer (samples #3-6, 8, 10, and 11), ii) a subject with a benign prostate tumor (sample #1), iii) a subject with a benign prostatic hyperplasia (BPH) (sample #9), iv) a subject with urological condition (sample #2), and v) a patient tentatively diagnosed with a prostate cancer (sample #7), contained 4-OHE$_1$(E$_2$)-1-N3Ade adduct. The adduct level (normalized to creatinine concentration) varied from sample to sample with concentration levels of about 15-240 pmole per mg of creatinine. The identity of samples #1-11 in the row #1 was confirmed by low-temperature (77K) luminescence

spectroscopy. The bars in the row #2 of Figure 4 correspond to the integrated (normalized) area of the low temperature phosphorescence spectra obtained for urine samples #1-11 eluted from the immunoaffinity columns. An example of the phosphorescence spectra obtained for samples #1, 4 and 6 are shown in the right inset of Figure 4; the red spectrum overlapping with the phosphorescence spectrum of sample #6 is that of the standard adduct. In fact, all spectra measured for samples #1-11 revealed emission identical to the phosphorescence spectrum of the 4-OHE$_1$(E_2)-1-N3Ade adduct standard (data not shown), thus proving again that the analyte eluted from the 15G8-MAb based column corresponds to the 4-OHE$_1$(E_2)-1-N3Ade. The amount of this adduct in samples #1-11 using low temperature phosphorescence-based calibration curves was about 10-150 pmole per mg of creatinine, depending on the sample. With the detection limit of about 10^{-9} M[17] no 4-OHE$_1$-1-N3Ade adducts were observed in samples #12-16 in agreement with the CE/FASS results. The observed emission intensity was near the background level.

Finally, ultra-performance liquid chromatography (LC) interfaced with mass spectrometry (MS), LC/MS/MS, was used for further validation of the above findings. That is, all samples eluted from the immunoaffinity columns were also analyzed by the LC/MS/MS. The findings are summarized in the row #3 of Figure 4. Also in this case only the samples #1-11 revealed the presence of the 4-OHE$_1$-1-N3Ade adducts. Although similar adduct distribution is observed in all samples using three different methodologies, the relative adduct concentration observed in elutions from immunoaffinity columns was somewhat smaller than that observed by CE/FASS with absorbance detection. The latter is not surprising as an efficiency recovery of a typical column is ~70-80%.[17, 22] An example of the HPLC chromatogram obtained for sample #11 is shown in the left inset of Figure 4; the main peak near 2 min in the chromatogram corresponds to the 4-OHE$_1$-1-N3Ade adduct indicating that the extract eluent from the immunoaffinity column was relatively pure. The spectrum corresponds to the fragmented daughters, m/z 135.9 and 296.0, which were obtained from MS of the adduct parent ion, m/z 420.1. Note that 4-OHE$_1$-1-N3Ade (and not 4-OHE$_2$-1-N3Ade) is being excreted into the urine of all subjects with prostate cancer (samples #3-6,8,10, and 11). This suggests that this adduct may be a biomarker for risk of developing prostate cancer. Interestingly, the same adduct was also observed in a subject

with a benign prostate tumor (sample #1), and in two subjects with benign prostatic hyperplasia (BPH) (sample #9) and an urological condition (sample #2), respectively. Sample #7 was obtained from a patient suspected to have a prostate cancer whose biopsy test, at the present time, was negative for cancer. However, LC/MS/MS did not reveal the presence of any 4-OHE_1-1-N3Ade adduct in samples #12-16 obtained from healthy individuals in perfect agreement with low temperature phosphorescence and CE/FASS studies. Several other healthy individuals tested by LTP (data not shown) also did not reveal any presence of the 4-OHE_1-1-N3Ade adduct. This might suggest that the subject who provided sample #7 is at high risk of developing prostate cancer and that BPH might be related to CEQ-induced DNA damage.

Detection of 4-OHE1-2-NAcCys Conjugates and 4-OHE1-1-N3Ade Adducts in Urine of Breast Cancer Patients. An example of an absorbance based electropherogram obtained at 214 nm for a 20 mL urine sample from a breast cancer patient run through the 2E9 MAb based column and subsequently eluted for CE/FASS analysis is shown in curve (b) of Figure 5A. As expected, only one major peak (#1) is observed, which corresponds to 4-OHE_1-2-NAcCys as proven by spectrum (a) obtained for the 4-OHE_1-2-NAcCys standard. Identification of this analyte was also confirmed by the standard spiking procedure and room T absorption spectra of peak #1 (data not shown). A similar procedure was used to identify the presence of 4-OHE_1-1-N3Ade in urine of the same patient; namely, spectra c, d, and e of frame B correspond to a urine sample from a breast cancer patient, 4-OHE_1-2-NAcCys conjugate standard, and urine from a healthy individual, respectively. The migration time of the main peak in curve c is identical to that of peak #2 obtained with 4-OHE_1-2-NAcCys standard (curve d), clearly suggesting that this conjugate is excreted into urine of the breast cancer patient. Note that this conjugate is not observed in the urine sample from the healthy individual (see curve e). However, quantitation of the above analytes was impossible, as both MAb-columns were saturated, suggesting relatively high concentrations (research in progress).

The above finding is also supported by the data shown in Figure 5C; here, spectrum f is the electropherogram obtained with CE/FASS for the same methanol-extracted and pre-

concentrated (by evaporation) urine sample. Curves g and h were obtained for the same urine extract after spiking with standards of 4-OHE_1-2-NAcCys (curve g) and 4-OHE_1-1-N3Ade (curve h), respectively. Comparison of curves f, g, and h indicates that peaks 1 and 2 correspond to the 4-OHE_1-2-NAcCys and 4-OHE_1-1-N3Ade, respectively. Quantitative data revealed that the concentration of 4-OHE_1-2-NAcCys and 4-OHE_1-1-N3Ade in this urine sample is about 10^{-6} and 2.5×10^{-7}M, respectively. We hasten to add that so far none of the above two analytes has been identified in urine samples from healthy women, suggesting that both 4-OHE_1-2-NAcCys and 4-OHE_1-1-N3Ade could constitute excellent biomarkers of breast cancer risk.

Conclusions

Monocolonal antibody 2E9 to 4-$OHE_1(E_2)$-2-NAcCys and monocolonal antibody 15G8 to 4-OHE1-1-N3Ade were synthesized and characterized. They were used for detection and quantitation of 4-OHE1(E2) conjugates and related analytes in human urine. The binding specificity study of the MAbs revealed a high degree of discrimination between 4-OHE_1-2-NAcCys/4-OHE_1-1-N3Ade and 4-OHE_1/4-OHE_2/NAcCys. The purified MAbs were individually immobilized on agarose bead columns, which were used to capture and preconcentrate the hapten of interest out of urine samples. A number of structurally related standards were used to estimate the selectivity and specificity of the chosen MAb. CE with FASS in the absorbance mode and off-line spectral characterization of samples released from the affinity column were used for identification and quantitation of 4-OHE_1-2-NAcCys, 4-OHE1-1-N3Ade and related analytes in water-based buffer and/or human urine matrixes from both healthy group and cancer patients.

4-OHE_1-1-N3Ade has been detected at much higher level in urine from subjects with prostate cancer and urological condition compared to healthy males. The same adduct and 4-OHE1-2-NacCys were also identified in urine from a woman with breast carcinoma. We believe that this is the first-ever indication that CEQ-derived DNA adducts are present in the urine samples from subjects with prostate or breast cancer. The significant elevation of

depurinating adducts in the urine of the prostate/breast cancer patients compared to the controls clearly indicates that exposure to catechol estrogen quinones is associated with increased risk of breast and prostate cancer. Therefore, we propose that the presence of depurinating adducts in human urine samples could be used as a risk factor for estrogen induced-cancer. We hasten to add that these biomarkers could serve as a model to investigate the hypothesis that metabolically activated endogenous estrogens are involved in initiating both prostate and breast cancers.

Acknowledgment

This work was supported by a grant from the National Cancer Institute (Program Project Grant 2PO1 CA49210-12) and in part by the NIH COBRE grant.

Table 1. Selected Characteristics of the Subject Population

Patient ID # in Figure 2	Diagnosis	Age	PSA (Prostate Specific Antigen)	Smoking Status
3	Prostate Ca	68	4.2	Non-smoker
4	Prostate Ca	70	4.8	Smoker
5	Prostate Ca	47	5.6	Non-smoker
6	Prostate Ca	63	6.7	Non-smoker
8	Prostate Ca	61	3.6	N/A
10	Prostate Ca	67	5.6	Non-smoker
11	Prostate Ca	53	N/A	Non-smoker
1	Benign biopsy	67	7.3	N/A
2	Erectile dysfunction	59	N/A	N/A
7	Benign biopsy	76	1.4	Non-smoker
9	BPH (post-TUNA)[a]	80	1.2	Smoker
12	Health control	21	[b]	Non-smoker
13	Health control	23	[b]	Non-smoker
14	Health control	17	[b]	Non-smoker
15	Health control	21	[b]	Non-smoker
16	Health control	35	[b]	Non-smoker

a. TUNA: trans-urethral needle ablation (a minimally invasive treatment for BPH)

b. Not measured

References

(1) Cavalieri, E. L.; Stack, D. E.; Devanesan, P. D.; Todorovic, R.; Dwivedy, I.; Higginbotham, S.; Johansson, S. L.; Patil, K. D.; Gross, M. L.; Gooden, J. K.; Ramanathan, R.; Cerny, R. L.; Rogan, E. G. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10937-10942.

(2) Cavalieri, E., Frenkel, K., Liehr, J.G., Rogan, E. and Roy, D. *JNCI Monograph: Estrogens as Endogenous Carcinogens in the Breast and Prostate (E.Cavalieri and E. Rogan, Eds.), Oxford Press,* **2000**, 75-93.

(3) Liehr, J. G. *Endocrine Reviews* **2000**, *21*, 4054-4058.

(4) Rogan, E. G.; Badawi, A. F.; Devanesan, P. D.; Meza, J. L.; Edney, J. A.; West, W. W.; Higginbotham, S. M.; Cavalieri, E. L. *Carcinogenesis* **2003**, *24*, 697-702.

(5) Cavalieri, E.; Rogan, E.; Chakravarti, D. *Methods in Enzymology, Quinones and Quinone Enzymes, Part B Editors: Sides, H. and Packer, L.* **2004**, 293.

(6) Li, K. M.; Todorovic, R.; Devanesan, P.; Higginbotham, S.; Kofeler, H.; Ramanathan, R.; Gross, M. L.; Rogan, E. G.; Cavalieri, E. L. *Carcinogenesis* **2004**, *25*, 289-297.

(7) Zahid, M.; Kohli, E.; Saeed, M.; Rogan, E.; Cavalieri, E. *Chem. Res. Toxicol.* **2006**, *19*, 164-172.

(8) Chakravarti, D.; Mailander, P. C.; Li, K. M.; Higginbotham, S.; Zhang, H. L.; Gross, M. L.; Meza, J. L.; Cavalieri, E. L.; Rogan, E. G. *Oncogene* **2001**, *20*, 7945-7953.

(9) Zhao, Z.; Kosinska, W.; Khmelnitsky, M.; Cavalieri, E. L.; Rogan, E. G.; Chakravarti, D.; Sacks, P.; Guttenplan, J. B. *Chem. Res. Toxicol.* **2006**, *19*, 475-479.

(10) Lareef, M. H.; Garber, J.; Russo, P. A.; Russo, I. H.; Heulings, R.; Russo, J. *Int. J. Oncol.* **2005**, *26*, 423-429.

(11) Russo, J.; Lareef, A. H.; Balogh, G.; Guo, S.; Russo, I. H. *J. Steroid Biochem. Mol. Biol.* **2003**, *87*, 1-25.

(12) Dorgan, J. F.; Longcope, C.; Franz, C.; et al., e. *J. Natl. Cancer Inst.* **2002**, *94*, 606-616.

(13) Bosland, M. C. *J. Natl. Cancer Inst. Monogr.* **2000**, *27*, 39-66.

(14) Ross, R.; Bernstein, L.; Judd, H.; Hanisch, R.; Pike, M.; Henderson, B. *J. Natl. Cancer Inst.* **1986**, *76*, 45-48.

(15) Leav, I.; Merk, F. B.; Kwan, P. W. L.; Ho, S. M. *Prostate* **1989**, *15*, 23-40.

(16) Bosland, M. C.; Ford, H.; Horton, L. *Carcinogenesis* **1995**, *16*, 1311-1317.

(17) Markushin, Y.; Zhong, W.; Cavalieri, E. L.; Rogan, E. G.; Small, G. J.; Yeung, E. S.; Jankowiak, R. *Chem. Res. Toxicol.* **2003**, *16*, 1107-1117.

(18) Dwivedy, I.; Devanesan, P.; Cremonesi, P.; Rogan, E.; Cavalieri, E. *Chem. Res. Toxicol.* **1992**, *5*, 828-833.

(19) Stack, D. E.; Byun, J.; Gross, M. L.; Rogan, E. G.; Cavalieri, E. L. *Chem. Res. Toxicol.* **1996**, *9*, 851-859.

(20) Harlow E , L. D. *Antibodies: A Laboratory Manual, editors. Plainview, NY: Cold Spring Habor; Chapter 6: Monoclonal Antibodies.* **1988**, 139-243.

(21) Jue, R.; Lambert, J. M.; Pierce, L. R.; Traut, R. R. *Biochemistry* **1978**, *17*, 5399-5405.

(22) Jankowiak, R.; Markushin, Y.; Cavalieri, E. L.; Small, G. J. *Chem. Res. Toxicol.* **2003**, *16*, 304-311.

(23) Quirino, J. P.; Terabe, S. *J. Chromatogr. A* **2000**, *902*, 119-135.

Figure Captions

Figure 1. Chemical structures of 4-$OHE_1(E_2)$, 4-$OHE_1(E_2)$-2-NAcCys, and 4-$OHE_1(E_2)$-1-N3Ade.

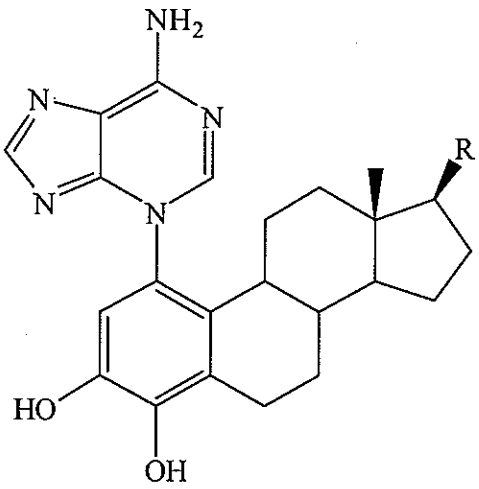
Figure 2. Curve a is the CE electropherogram (observation wavelength at 214 nm); peaks 1, 2, 3, and 5 correspond to 4-OHE_1-1-N3Ade, 4-OHE_1, 4-OHE_2, and NAcCys, respectively (concentration, $c = 10^{-6}$ M). Peak 4 (near 5 min migration time labeled by a solid arrow) corresponds to the 4-OHE_1-2-NAcCys at a significantly lower concentration (i.e., 10^{-8} M). Curve b is the CE electropherogram obtained for the same mixture passed through the 2E9 MAB-based affinity column and preconcentrated by a factor of 100. The major peak 4 corresponds to the captured and highly concentrated 4-OHE_1-2-NAcCys conjugate.

Figure 3. Curve a: CE electropherogram of a mixture of four analytes in a buffer solution; peaks 1, 2, 3, and 4 correspond to 4-OHE_1-1-N3Ade, 4-OHE_1-2-NAcCys, 4-OHE_1, and 4-OHE_1-1-N7Gua, respectively. Curve b: electropherogram of a PBS buffer sample spiked with analytes 1-4 listed above and run through the affinity column [only 4-OHE_1-2-NAcCys (peak 2) is recovered]. Curve c: CE electropherogram obtained after a diluted human urine sample was spiked with 4-OHE_1-2-NAcCys is run through the affinity column. Peak 2 reveals an excellent recovery of 4-OHE_1-2-NAcCys. Curve d: electropherogram of 4-OHE_1-2-NAcCys standard (see text).

Figure 4. Identification of the 4-OHE_1-1-N3Ade adduct in human samples from men with prostate cancer or urological conditions and healthy men as control. Right inset: The spectra labeled 1, 4 and 6 refer to individual samples 1, 4 and 6, respectively; the red spectrum is that of the standard. Left inset: Identification of the parent compound and m/z 135.9 and 296 are the fragmentation daughters selected for the unequivocal identification of the adduct.

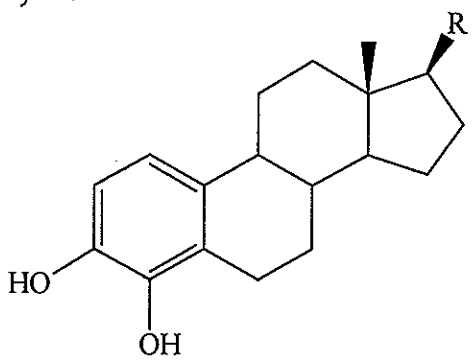
Figure 5. Identification of 4-OHE_1-2-NAcCys (peak 1) and 4-OHE_1-1-N3Ade (peak 2) in human urine from a woman with breast carcinoma (see text).

Figure1.

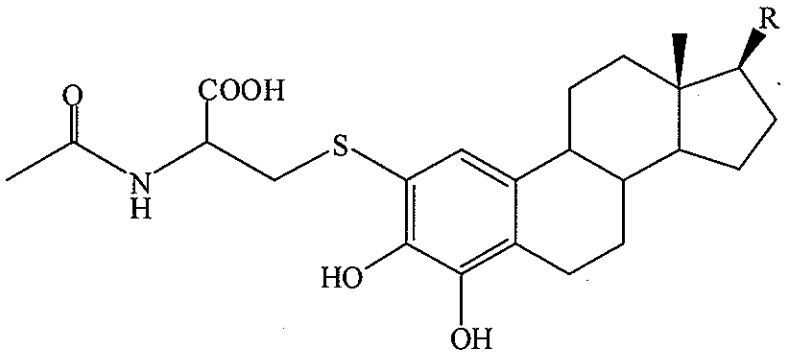


4-OHE1(E2)-1-N3Ade

E1: R, =O
E2: R, -OH



4-OHE1(E2)



4-OHE1(E2)-2-NAcCys

Figure 2.

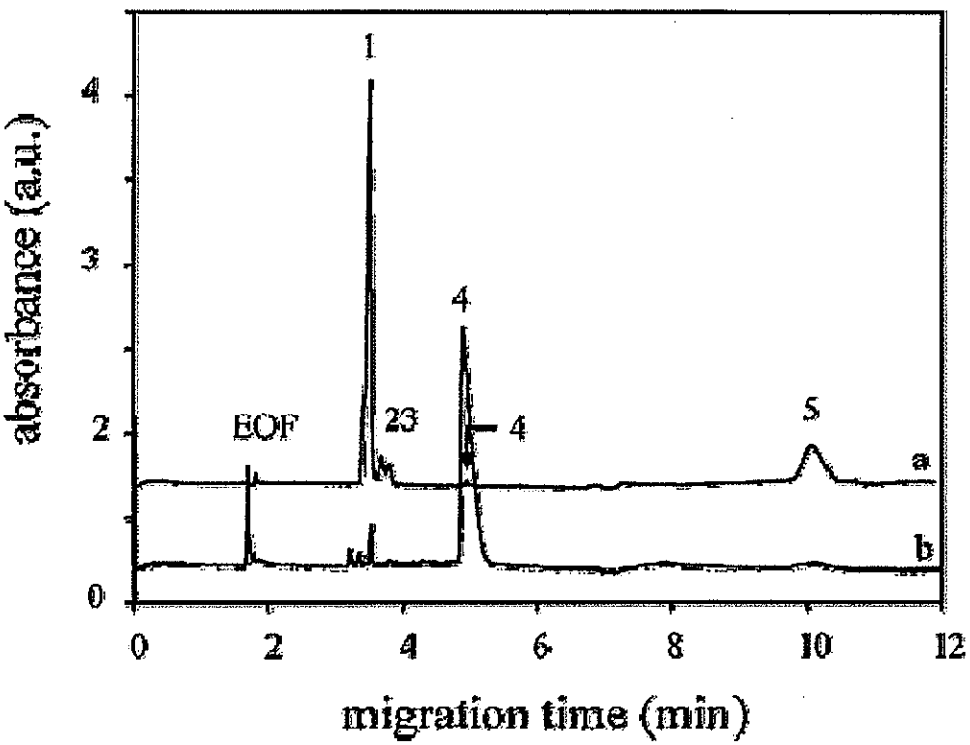

absorbance (a.u.)
migration time (min)
EOF
1
23
4
4
5
a
b
0
2
3
4
0
2
4
6
8
10
12

Figure 3.

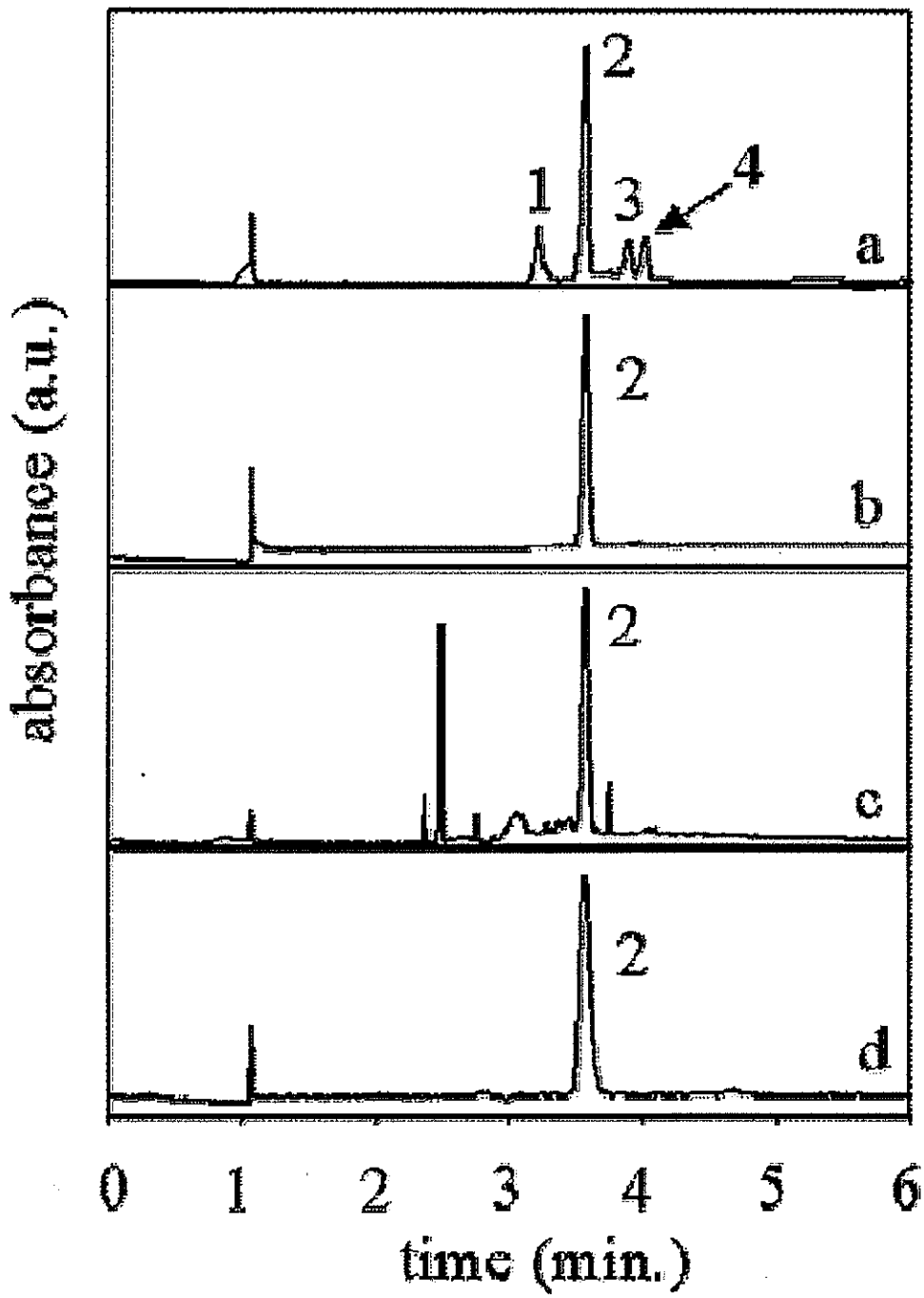


Figure 4.

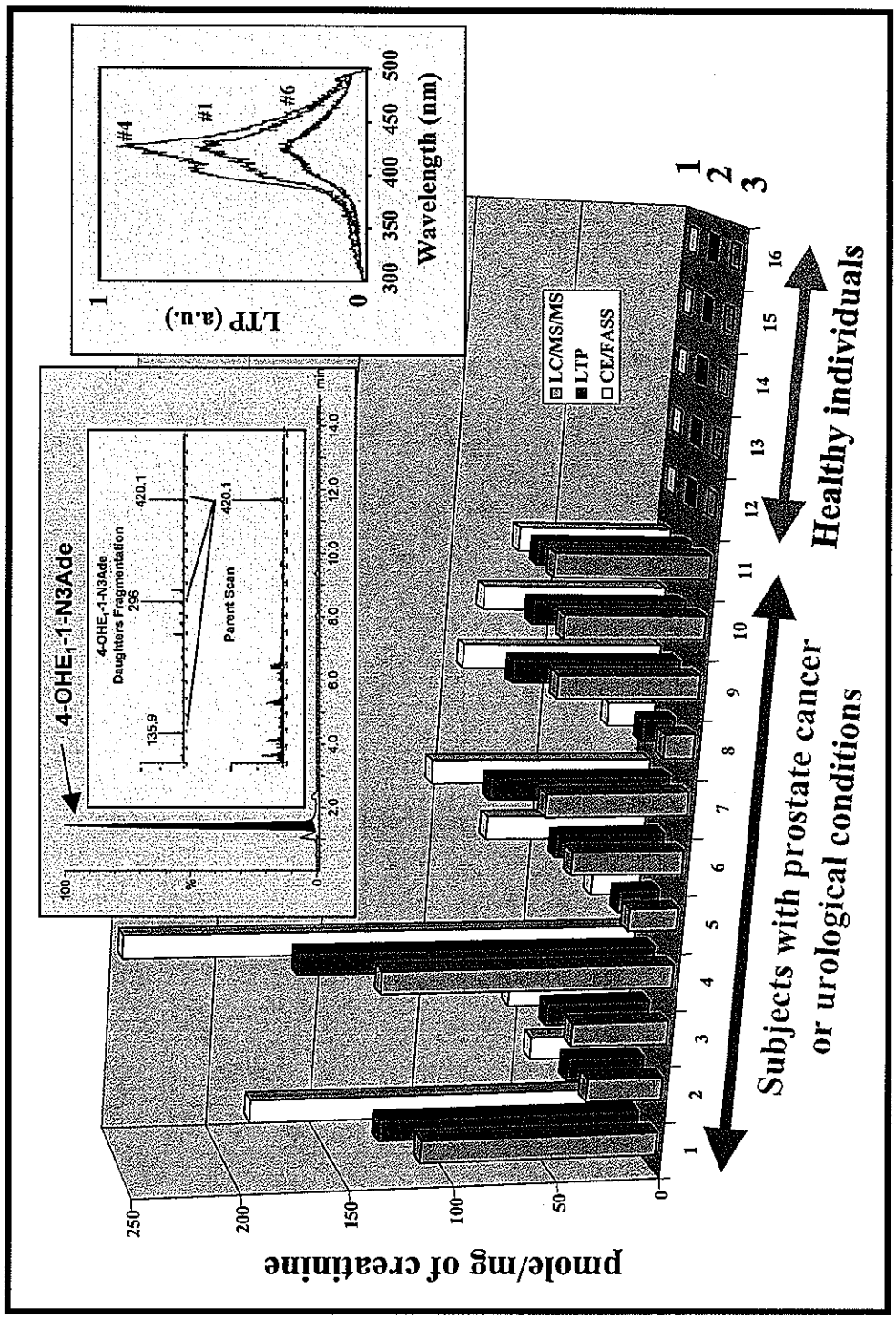
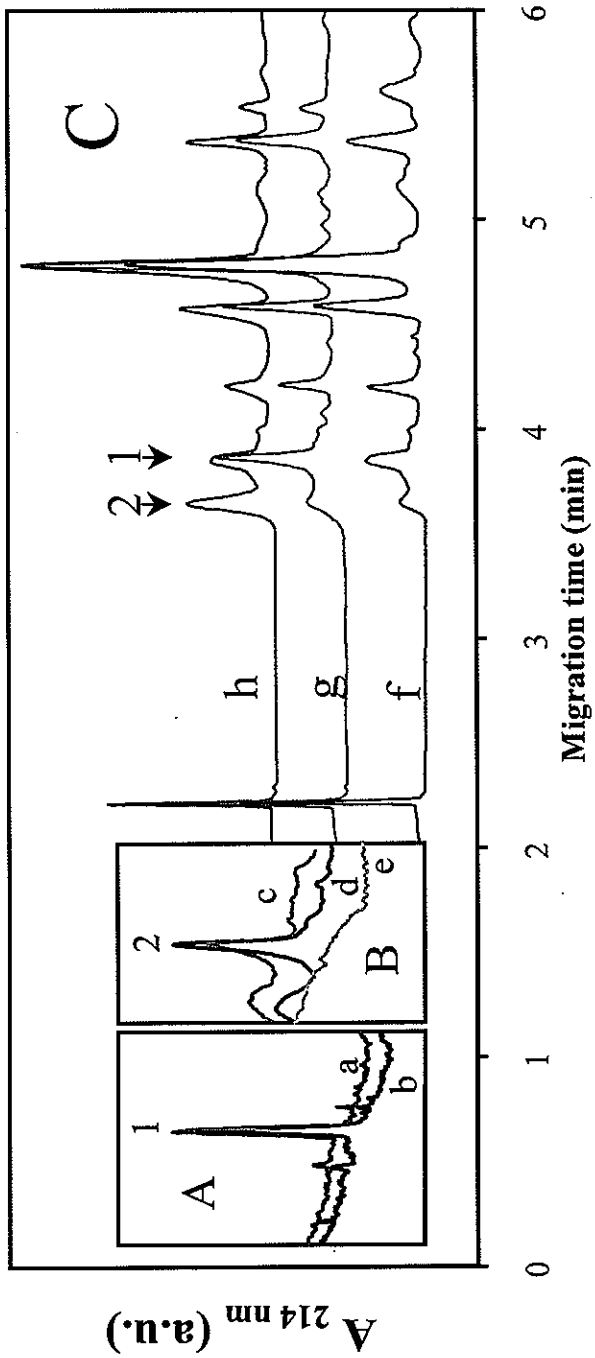

pmole/mg of creatinine
250
200
150
100
50
0
1
2
3
Subjects with prostate cancer
or urological conditions
Healthy individuals
LC/MS/MS
LTP
CE/FASS
LTP (a.u.)
Wavelength (nm)
#4
#1
#6
4-OHE-1-N3Ade
Daughters Fragmentation
Parent Scan
135.9
420.1
296
min

Figure 5.



CHAPTER 4. MOBILITY-BASED WALL ADSORPTION ISOTHERMS FOR COMPARING CAPILLARY ELECTROPHORESIS WITH SINGLE-MOLECULE OBSERVATIONS

A paper published in Analytical Chemistry*

Ning Fang, Hui Zhang, Jiangwei Li, Hung-Wing Li and Edward S. Yeung

Abstract

The adsorption properties of R-phycoerythrin (RPE), an autofluorescent protein, on the fused-silica surface were studied in capillary electrophoresis (CE) and in single-molecule experiments. The band shapes and migration times were measured in CE and adsorption and desorption events were recorded at the single-molecule level by imaging within the evanescent-field layer using total internal reflection fluorescence microscopy. The adsorbed RPE molecules on the surface of the fused-silica prism were counted with confidence based on ImageJ software. The capacity factor and desorption rate were estimated from the counting results. The mobility-based adsorption isotherms were constructed from both computer simulations and experiments to determine the capacity factor.

---

* Reprint with permission from Analytical Chemistry **2007,** 79(17), 6047-6054.

Copyright © 2007 American Chemical Society

Introduction

Understanding molecular dynamics at liquid-solid interfaces is important in many research areas, such as retention in liquid chromatography (LC) and gas chromatography (GC), peak distortion caused by wall adsorption in capillary electrophoresis (CE), biocatalysis at solid membranes, etc. Chromatographic separation is usually attributed to the differential rate of migration induced by solute distribution between the stationary and mobile phases. The sorption isotherm, a plot of the solute concentration in the stationary phase against that in the mobile phase at a constant temperature, can provide important knowledge of the underlying thermodynamic functions. Adsorption on the surface of the solid stationary phase is often the main mechanism, and the methods for experimental determination of the adsorption isotherms in LC have been reviewed.[1] In CE, it is common that the analytes are adsorbed to the capillary wall to induce non-uniform zeta-potential and differential transport velocities,[2] causing peak broadening or distortion. Theoretical studies and modeling of capillary wall adsorption have been carried out by several groups.[3-8] Graf *et al.* demonstrated qualitatively several adsorption properties of proteins on fused-silica- and polyacrylamide-coated capillaries.[9] However, there were few studies to measure the adsorption strength quantitatively in CE mainly due to two complicated issues: kinetics (non-equilibrium CE wall adsorption vs. generally-assumed instant equilibrium in LC/GC) and geometry (adsorption on the inner surface of cylindrical capillary column vs. even distribution of solutes in the stationary phase).

The effectiveness of a separation is a function of the component's capacity factor, as well as the column's efficiency and selectivity. Capacity factor of a component is of central importance for the description of chromatographic migration and is a measure of the degree to which it partitions into the stationary phase. In this study, a two-dimensional CE simulation model[7] is utilized to establish a quantitative description of adsorption in CE and to allow the use of mobility-based adsorption isotherms for determining adsorption parameters.

It is well known that both electrostatic and hydrophobic interactions govern protein adsorption at liquid-solid interfaces.[10, 11] Both rate theory and band broadening in chromatography and electrophoresis have been explained by using statistical theory.[12] Direct

monitoring of individual molecular motion and interactions at the liquid-solid interface would provide valuable information regarding these and other related phenomena. The total internal reflection (TIR) geometry provides excellent background rejection and confines excitation to the evanescent field layer (EFL), where the motions of individual molecules are recorded.[13-15]

R-phycoerythrin (RPE), as a model protein, is a 240-kDa autofluorescent protein consisting of 7 subunits and about 30 chromophores.[16-18] Phycobiliproteins such as RPE, β-phycoerythrin (BPE), and allophycocyanin (APC) are stable and highly water-soluble globular proteins derived from cyanobacteria and eukaryotic algae.[19] They are slightly acidic, with isoelectric points around pH 5, and are stable over the pH range 5-9.[20] The pI of RPE depends on the source of the protein. Kang *et al.* reported the pI of RPE from cyanobacteria and eukaryotic algae at pH 4.3-4.7.[21]

In the present work, we studied the adsorption properties of RPE both on the inner wall of cylindrical fused-silica capillary columns under the influence of an electric field and on the flat fused-silica prism surface by single-molecule imaging. Using CE, the capacity factor of RPE at pH 5.0 was determined by the mobility-based adsorption isotherm. Using total internal reflection fluorescence microscopy, motions of single RPE molecules were imaged at the water/fused silica interface, and a reliable counting method was developed for the adsorbed RPE molecules. We demonstrate the potential application of single-molecule imaging as a means to measure the capacity factor and desorption rate of a molecule on a chromatographic surface.

Experimental Section

Buffer Solutions. The various buffer systems used were as follows: sodium acetate/ acetic acid (20 mM, pH 4.0, 5.0, 5.5, 6.0), Gly-His/HCl (10 mM, pH 7.0), and Gly-Gly/NaOH (10 mM, pH 8.2). A.C.S. grade or higher glacial acetic acid and sodium acetate were purchased from Fisher Scientific (Fair Lawn, NJ). Gly-His (G1627) and Gly-Gly

(G3915) were purchased from Sigma (St. Louis, MO). Ultra-pure 18-MΩ water was used in all solutions. All solutions were filtered through 0.22 μm filters.

Sample Preparation. The RPE stock solution (4 mg mL^{-1} or 16.7 μM) was purchased from Molecular Probes (Eugene, OR). RPE samples were prepared in different ways for CE and single-molecule detection experiments.

Coumarin 535 (Exciton, Dayton, OH) was chosen as the EOF marker for CE experiments. Coumarin 535 was first dissolved in about 200 μL of methanol and diluted to 10 mL with the pH 5.0 buffer solution. The coumarin solution was then centrifuged at 12000 rpm for 5 min, and finally the clear solution was taken out and used to dilute the RPE stock solution to the following concentrations: 0.25, 0.5, 1, 2, 4, 20, and 40 μg L^{-1}.

For single-molecule imaging, the RPE stock solution was first diluted to a concentration of 16.7 nM with the pH 8.2 buffer solution. This RPE solution was then diluted to 16.7 or 1.67 pM with appropriate buffer solutions right before imaging.

CE Apparatus and Procedure. The experiments were carried out on a Beckman Coulter MDQ System (Fullerton, CA) with a laser-induced fluorescence (LIF) detector. A Beckman Laser Module 488 was used as the excitation source, and a 600 nm (FWHM 70 nm) band-pass filter (CVI Laser, Albuquerque, NM) was placed in the filter housing for RPE detection. 50 cm long (40 cm to detector) × 50 or 75 μm inner diameter, fused-silica capillary columns (Polymicro Technologies, Phoenix, AZ) were used. A temperature of 25 °C was maintained in all CE experiments.

Prior to use, the capillary columns were treated with 1 M NaOH, methanol, and purified water for 15 min each. Then, the capillary was flushed with pH 5.0 buffer solution for 30 min and was conditioned overnight. Between CE runs, the column was rinsed with 10 mM NaOH for 5 min, water for 2 min, and the buffer solution for at lease 10 min. In each CE run, an RPE sample solution was injected at 0.5 psi for 5 s, and then a 10 kV potential was applied to the capillary. Each RPE sample was analyzed four times.

CE Simulation Conditions. The newly-developed 2D CE simulation program[7] was used to demonstrate the effects of wall adsorption. In order to run the simulation within a

reasonable computing time, the capillary conditions were chosen as: capillary length, $L_t = 5$ cm; length to detector, $L_d = 4$ cm; inner radius, $r_c = 25$ μm; applied voltage, 1000 V; length of the rectangular injection plug, $L_A = 1.8$ mm; initial concentration of A, from 0.005 to 100 mol m^{-3}; diffusion coefficient, $D_A = 10^{-9} m^2 s^{-1}$; electrophoretic mobility, $\mu_{ep,As} = 2 \times 10^{-8} m^2 V^{-1} s^{-1}$; maximum concentration of available binding sites on the wall, $w_0 = 4 \times 10^{-6}$ mol m^{-2}; adsorption and desorption rate constants, $k_a = 100 m^3 mol^{-1} s^{-1}$, $k_d = 100 s^{-1}$; electroosmotic flow, $\mu_{eo} = 0$. In addition, the three key parameters for numerical calculation, longitudinal space increment Δz, radial space increment Δr, and time increment Δt, were set to 1.0×10^{-5} m, 1.25×10^{-5} m, and 0.001 s, respectively.

Single-Molecule Detection System. The evanescent-wave excitation geometry was similar to that described previously.[12] Each 5-μL of sample solution was sandwiched between a No. 1 (22×22 mm^2) Corning glass coverslip and the hypotenuse face of the right-angle UV-grade fused-silica prism (Part #01 PQB 002; Melles Griot, Irvine, CA). A laser beam was focused and directed through the prism to the sample interface. The angle of incidence was about 66°. Therefore, the laser beam was totally internally reflected at the prism/solution interface and an evanescent field of ~100 nm thick was created.

A 532-nm solid-state continuous wave laser (45 mW, μ-Green model 4611, Uniphase, San Jose, CA) was used as the excitation source. A Uniblitz mechanical shutter (model LS2Z2, Vincent Associates, Rochester, NY) and a driver (model T132, Vincent Associates) were synchronized to the Pentamax 512-EFT/1EIA intensified charge-coupled device (ICCD) camera (Princeton Instruments, Princeton, NJ). The shutter only lets the laser beam pass through when the ICCD was on to reduce photobleaching. Two 532-nm, long-pass edge filters (Semrock, Rochester, NY) were placed between the objective lens and the ICCD to cut off stray light from the excitation beam. A microscope with a 100× oil immersion objective lens (Zeiss Plan-Neofluar 100×/1.3 oil) was employed. The objective was coupled to the coverslip with immersion oil (type FF, n = 1.48, Cargille, Cedar Grove, NJ). The ICCD exposure frequency was 5 Hz (0.2 s/frame), and the exposure time for each frame was 10 ms. A sequence of frames were acquired for each sample via WinView/32 (Princeton Instruments).

Image Analysis. He *et al.* demonstrated a method to count free DNA molecules using ImageJ.[22] In the present work, we need to not only identify and count RPE molecules, but also study the dynamics of adsorption/desorption processes by determining whether a molecule is exactly at the surface or just close to the surface within the evanescent field layer and how long it is adsorbed. In general, a molecule is considered adsorbed if it stays at the same position for two or more consecutive frames (images). However, the actual counting process based on this simple idea is challenging because of inhomogeneous illumination of the laser beam, fluctuations in the fluorescence intensity, photobleaching, aggregation, and Brownian motion of RPE molecules. We herein propose a method based exclusively on the functions and plug-ins within ImageJ, a public-domain image processing program (version 1.38, National Institutes of Health, USA), for counting adsorbed RPE molecules.

Single-molecule videos were recorded by repeating the following steps: moving to a new observation region, focusing the objective to the surface, and taking a set of 10-20 images at a frequency of 5 Hz before the RPE molecules started to be photobleached. The videos (in Winview/32 SPE file format) required pre-treatment to become uniform and to reduce the background. The video files were first loaded into ImageJ using a SPE format plug-in. Then the background subtraction function, which implements the rolling-ball algorithm, was applied on the image stacks to bring the backgrounds of all frames to the same level. Finally, the smoothing function was used to improve the signal-to-noise ratio. The mean and standard deviation (σ) of the reduced background calculated from the images of blank samples were approximately 20 ± 18. Similar noise levels were found in the frames taken with the low-concentration (1.67 pM) RPE samples. However, the noise level for the samples at 16.7 pM increased by ~50% due to Brownian motion of a larger number of RPE molecules in the solution.

To decide whether RPE molecules were adsorbed on the surface or were free in the solution, a threshold was applied to convert the entire stack of images into black (value 0, background) and white (value 255, bright RPE molecules) images. A reasonable choice for the threshold was the mean of the background plus 4σ. The number of white spots in each frame did not necessarily give the right count of RPE molecules because of Brownian motion

and random noise. Noise could be greatly reduced, however, by removing all the white particles that were less than 4 pixels in size. The size requirement of 4 pixels was justified by carefully inspecting the images. Less-than-4-pixel bright spots were most likely caused by the ICCD camera's shot noise or by the RPE molecules that came close enough to the surface but were not actually adsorbed or were only adsorbed for a fraction of the exposure time (10 ms).

The next step was to sum up multiple ($n \geq 2$) consecutive well-focused, black and white frames taken at the exact same position by using the stacks/Z project function in ImageJ. Each pixel in the resulting image had only one of the n+1 discrete intensity values (0, 1×255, 2×255, …, n×255), corresponding to the number of the original frames in which the same pixel was bright. The threshold was then set to a value greater than (n-1)×255 to identify the bright particles in all original frames. The particle analysis function was used to count the number of such particles that satisfied a minimum size requirement. This counting result gave the number of adsorbed molecules in the first frame that stayed at the same positions for at least n consecutive frames, or a time period calculated as n times the 10-ms exposure time plus (n-1) times the 190-ms frame interval. Next, we moved one frame forward in the same set of images or to a new set (taken at a different position) to continue counting.

The minimum size requirement in particle analysis was of great importance to accurate counting. When $n = 2$, the chance of having two randomly moving molecules at the same position was relatively high, so the minimum size was set to 4 to reduce-false positive counts. When $n \geq 3$, the chance of false-positive counts was reduced substantially, but the chance of false negative counts due to fluctuations in fluorescence intensity was increased; therefore, the minimum size was set to 1 in that case. The results based on these practical rules matched quite well with manual counting.

Results and Discussion

In the present work, quantitative analysis of adsorption at liquid/solid interfaces was carried out in a three-stage approach: (1) discovery of the relationship between analyte mobility and concentration in the presence of wall adsorption using a 2-dimensional computer simulation of CE; (2) calculation of the adsorption capacity factor from the experimental mobility-based adsorption isotherm; (3) direct observation and counting of adsorbed molecules at the liquid/solid interface. The results from these three stages will be discussed sequentially.

STAGE ONE: Discovery with Computer Simulation of CE

In a simple CE experiment, an uncoated fused silica-capillary (total length L_t, length to detector L_d, inner radius r_c) is filled with a buffer solution. A sample solution containing one analyte (A) is injected to form a narrow plug of a length L_A. An electric field is then applied across the capillary, and the analyte plug migrates toward the detector. The velocity (v) is determined by the electric field strength (E), the electroosmotic mobility (μ_{eo}), and the electrophoretic mobility (μ_{ep}): $v = (\mu_{ep} + \mu_{eo})E$. During electrophoretic migration, the analyte also interacts with the capillary wall, as described by eq. 1.

$$As + W = Aw \quad (1)$$

in which As is the free analyte (in solution), Aw is the analyte adsorbed to the wall, and W is the binding site on the wall. The adsorption and desorption processes are considered to take place between the wall and the layer of analyte solution closest to the wall. The width of this layer (l_s) is determined by many factors, including the diffusion coefficient of the analyte (D_A), the adsorption and desorption rates (k_a and k_d), the analyte concentration, and the number of available binding sites on the wall.

This type of CE experiments with consideration of wall adsorption can be simulated with the newly-developed 2D simulation model of CE.[7] The mathematical model and

computer implementation have been discussed extensively in the previous paper. To describe the wall interaction, a nonlinear Langmuir second-order kinetic law is used:

$$\frac{da_w}{dt} = k_a a_s w - k_d a_w = k_a a_s (w_0 - a_w) - k_d a_w \qquad (2)$$

in which a_w is the concentration of the analyte adsorbed to the wall, a_s is the analyte concentration in the layer closest to the wall, w is the concentration of wall binding sites available to the analyte, and w_0 is the total concentration of binding sites. Note that a_s is in the unit of mol m^{-3}, while a_w and w_0 are in the unit of mol m^{-2}. Due to the cylindrical shape of the capillary, the change in a_s has to be corrected according to the radius of the capillary and the thickness of the layer:[7]

$$-\frac{da_s}{dt} = \left[k_a a_s (w_0 - a_w) - k_d a_w\right] \frac{2r_c}{l_s(2r_c - l_s)} \qquad (3)$$

When the number of analyte molecules is small compared with the number of total binding sites, the wall interaction can be treated approximately as a linear case with eqs. 4 and 5.

$$\frac{da_w}{dt} = k_a w_0 a_s - k_d a_w \qquad (4)$$

$$-\frac{da_s}{dt} = \left(k_a w_0 a_s - k_d a_w\right) \frac{2r_c}{l_s(2r_c - l_s)} \qquad (5)$$

Average Mobility. Due to the adsorption/desorption processes, analyte migration slows down as the analyte spends a certain fraction of time on the wall. The average apparent mobility (μ^A) can be calculated by the following equation:

$$\mu^A = f_{As}\mu_{As} + f_{Aw}\mu_{Aw} \qquad (6)$$

in which f_{As} and f_{Aw} are the fractions of the different forms ($f_{As} + f_{Aw} = 1$), μ_{As} is the apparent mobility of A in the solution ($\mu_{As} = \mu_{ep,As} + \mu_{eo}$), and μ_{Aw} is the apparent mobility on the capillary wall. An adsorbed molecule does not move on the wall. It can only migrate in the solution after being released from the wall; therefore, $\mu_{Aw} = \mu_{ep,Aw} + \mu_{eo} = 0$, or $\mu_{ep,Aw} = -\mu_{eo}$. Substituting the electrophoretic mobilities for the apparent mobilities in eq. 6 gives eq. 7:

$$\mu_{ep}^{A} = f_{As}\mu_{ep,As} + f_{Aw}\mu_{ep,Aw} = f_{As}\mu_{ep,As} - f_{Aw}\mu_{eo} \qquad (7)$$

Equations 6 and 7 are similar to the nonlinear regression equation for determining binding constants in affinity capillary electrophoresis (ACE).[23, 24] However, they cannot be used in a similar manner to determine wall adsorption parameters because the average mobility and the fractions of free and adsorbed analyte can be affected by the experimental conditions and the characteristics of the analyte: kinetic constants, diffusion coefficient, inner radius of capillary, conditions of the capillary wall (concentration of binding sites, charge distribution, etc.), electrophoretic mobilities, electroosmotic flow, and length and initial concentration of the analyte plug. It is difficult to study all these factors in actual CE experiments. The present work demonstrates that computer simulation is a convenient tool in assessing the effects of these factors.

Computer-Simulated Electropherograms. The electropherograms simulated with the conditions listed in the experimental section are shown in Figure 1. When the analyte is not adsorbed to the capillary wall at all, the rectangular injection plug results in a perfect Gaussian peak due solely to longitudinal diffusion. When wall adsorption occurs, the elution of the analyte peak is delayed, and peak distortion is observed. At the low end of the concentration range, the amount of the analyte in the injection plug is small compared with the number of available binding sites on the wall. Therefore, nearly all analyte molecules in the plug undergo a similar degree of interaction with the wall. The capillary wall reaches its full impact on the analyte migration and the migration time increases to a maximum value, which in turn means minimum average analyte mobility. The resulting peaks are near-Gaussian-shaped, but broader than the ones without wall adsorption. At the high end of the concentration range, the analyte can easily saturate the available binding sites on the wall. The analytes at the front of the plug can interact with the wall as dictated by the kinetic rate constants, but the analytes at the back cannot be adsorbed to the already-saturated wall. Thus, the latter group travels almost exclusively in the bulk solution and at a faster speed. As a consequence, the peak migration times are closer to those without wall adsorption, and long tails on the analyte peaks are observed.

Another interesting observation in Figure 1 is that the trailing ends of all peaks meet at the same position. This is predictable because the analyte molecules at the front edges of all injection plugs in the very beginning are affected by adsorption equally and most efficiently to produce identical migration times.

Because the fraction of time that an analyte molecule spends on the wall is related to its position in the analyte plug and the initial analyte concentration, f_{As} and f_{Aw} in eqs. 6 and 7 are not uniform for the entire analyte plug. Therefore, the regression methods for ACE cannot be used directly for determining wall adsorption parameters.

Computer-Simulated Adsorption Isotherms. The electrophoretic mobilities of the analyte peaks are calculated from the peaks in Figure 1 and then plotted against the initial analyte concentrations in Figure 2A. At the high end of the concentration range, the electrophoretic mobility rises to a maximum value. At the low end, the electrophoretic mobility decreases nearly linearly towards a minimum value as the analyte concentration reduces to 0. This set of data points can be fitted with an exponential-rise-to-maximum equation, which has the following general form:

$$y = y_0 + a_1(1-e^{-b_1 x}) + a_2(1-e^{-b_2 x}) + \ldots + a_N(1-e^{-b_N x}) \quad (8)$$

in which a_1, b_1, a_2, b_2, …, a_N, b_N, and y_0 are constants. The R-squared value (R^2) is used to evaluate how well the equation fits the data set. This theoretical nonlinear adsorption isotherm is generated with well-defined experimental conditions and kinetic parameters; therefore, a highly significant fit ($R^2 \rightarrow 1$) is expected from regression. Note that in eq. 8, y_0 is the only linear parameter, and the remaining parameters (a_1, b_1, …) are nonlinear (exponential). The number of exponential terms can be adjusted to reflect the complexity of the problem.

Equation 8 was found to provide the best fit among all the equations that have been tried. As shown in Figure 2A, the three-parameter (dot-dash), five-parameter (long-dash), and seven-parameter (solid) forms are tested, and the seven-parameter form gives a near-perfect fit (R^2=1.0000). The maximum mobility (2.000×10^{-8} $m^2V^{-1}s^{-1}$) is equal to the expected value: $\mu_{ep,As}$. The minimum mobility (1.516×10^{-8} $m^2V^{-1}s^{-1}$) indicates the full impact

of wall adsorption. These mathematical outcomes can be related to the system's experimental conditions. The one linear factor is $\mu_{ep,As}$ and related conditions (L_t, L_d, E). There are also a number of nonlinear factors: kinetics (k_a, k_d, and w_0), wall capacity (w_0, L_d, r_c, and L_A), radial diffusion (D_A and r_c), etc. One parameter may belong to more than one group. The effects of some nonlinear factors (L_A, L_d, D_A, and r_c) cannot be easily described by mathematical equations directly. Computer simulation makes it possible to reveal the outcomes of the complex system.

Buffer conditions (composition, ionic strength, pH, etc.) may also change the adsorption isotherm in actual CE experiments, but they are not part of the discussion in this paper for the following reasons: CE buffers can often be optimized to minimize peak distortion caused by undesirable conditions. Once a well-buffered CE system is established, the electric field strength may be assumed to be constant throughout the capillary and the buffer conditions can be incorporated into the linear factor ($\mu_{ep,As}$). If unnecessary complexities are brought to the CE system with complicated buffer conditions, the large number of contributing factors cannot be easily determined.

For lower analyte concentrations, a linear adsorption isotherm (Figure 2B, short-dash) can be made using simple linear regression. The linear and nonlinear adsorption isotherms are analogous to the ones used in conventional chromatography, in which the analyte concentration in the mobile phase is plotted against the analyte concentration in the stationary phase.[25] The linear adsorption isotherm is also regulated by the linear kinetic equations (eqs. 4 and 5).

Wall Adsorption Parameters. The fraction of free analyte can be calculated by eq. 9, which is derived from eqs. 6 and 7.

$$f_{As} = \frac{n_{As}}{n_{As} + n_{Aw}} = \frac{\mu^A}{\mu_{As}} = \frac{\mu^A}{\mu_{ep,As} + \mu_{eo}}, \text{ or } f_{As} = \frac{\mu_{ep}^A}{\mu_{ep,As}} \text{ when } \mu_{eo} = 0 \quad (9)$$

in which n_{As} and n_{Aw} are the mole number (or mass) of the analyte in the solution and on the wall, respectively, and $\mu_{ep,As}$ is given by the maximum mobility in Figure 2A. The

capacity factor k' (also known as retention factor or partition ratio) is defined in eq. 10. Using eqs. 9 and 10, k' at each analyte concentration can be calculated and plotted in Figure 2C.

$$k' = \frac{n_{Aw}}{n_{As}} = \frac{1 - f_{As}}{f_{As}} \qquad (10)$$

Note that f_{As} and k' are the average values for the entire analyte peak. Only when the analyte concentration approaches 0 that these values become uniform for the entire peak. From the y-intercept in Figure 2D, k' at zero analyte concentration (k_0') is 0.319.

The radius-dependent partition ratio can be converted to the concentration ratio:

$$\frac{a_w}{a_s} = k'\beta = k'\frac{r_c}{2} \qquad (11)$$

in which β (defined as half of the radius r_c) is the phase volume ratio. Plugging in the values of k_0' and r_c, $a_w/a_s = 7.98\times10^{-6}$ m for the simulated CE system as the concentration approaches zero. The concentration ratio may also be derived from the partition coefficient (K) multiplied by w, if adsorption and desorption kinetics is much faster than electrophoretic migration and an equilibrium can be assumed at any moment.

CE Experimental Considerations. Computer simulation and mathematical derivation provide the framework for understanding wall adsorption in CE. However, more factors need to be considered in actual experiments.

First, subtle information on the adsorption isotherm is lost due to actual experimental errors. Therefore, seven- or five-parameter regression equations are no longer required. Nonetheless, the maximum and minimum mobilities can still be used to calculate the adsorption parameters.

Second, the adsorption isotherm is often obtained for an analyte with weak wall adsorption. It may take a long time (or forever) for an analyte with strong wall adsorption to migrate out of the capillary, which makes mobility measurements difficult.

Third, analyte sample solutions over the full concentration range of the adsorption isotherm are required. If high concentrations are not feasible due to solubility or cost, nonlinear regression can still be carried out to get the maximum mobility with compromised

accuracy. In addition, the analyte must be detectable at the low end of the concentration range.

Fourth, the theoretical adsorption isotherm is demonstrated assuming $\mu_{eo} = 0$; however, the adsorption-induced non-uniform zeta-potential can change the EOF constantly during an actual CE run.[8] This error is monitored and partially compensated for by using a non-adsorbing neutral EOF marker. In addition, changes in the bulk flow velocity may also affect adsorption/desorption kinetics.

STAGE TWO: Experimental Adsorption Isotherms in CE

Adsorption of RPE on the fused-silica capillary wall (rc = 25 or 37.5 μm) in 20-mM acetic acid/sodium acetate buffer (pH 5.0) was chosen as the model system. Electropherograms obtained with different RPE concentrations (0.25 ~ 40 mg L^{-1}) are displayed in Figure 3, in which the horizontal axis is a mobility scale as opposed to the standard time scale. The low-concentration RPE peaks were narrower and migrated slower than the higher-concentration ones. Distortions at the right-hand slopes of the peaks were caused by heterogeneity of the protein sample rather than adsorption. Similar shapes were observed by Kang and Yeung even at pH 8.2 when no adsorption would occur.[21] The experimental adsorption isotherms were fitted nonlinearly over the full concentration range (Figure 4A) and linearly in the low concentration portion (Figure 4B) to give the maximum and minimum effective mobilities. The correlation between the partition ratio and RPE concentration was further constructed in Figure 4C (nonlinear) and 4D (linear) using eqs. 11 and 12. Because the actual EOF mobility was not zero and could change from run to run, each individual measurement was plotted instead of the average value of four consecutive measurements. The value of $k_0^{'}$ was found to be 0.069±0.005, which means that ~6.5% of the total RPE molecules, calculated as 0.069/(1+0.069), was adsorbed on the capillary wall at very low concentration. This value is much smaller than the $k_0^{'}$ obtained for the exaggerated conditions used in the simulated experiments so peak tailing at high concentrations in Figure 3 is not as noticeable as that in Figure 1. The concentration ratio can be calculated using eq.

11 to be $(8.6\pm0.6)\times10^{-7}$ m. The experiments were repeated with a 75-μm-i.d. capillary, and the same calculation procedure was performed to give $k_0^{'} = 0.040\pm0.005$ and aw/as = $(7.5\pm0.8)\times10^{-7}$ m. Due to the larger volume of the 75-μm-i.d. capillary, this $k_0^{'}$ is smaller than that for the 50-μm-i.d. capillary.

STAGE THREE: Capacity Factor from Single Molecule Imaging

The capacity factor can be simply represented by the ratio of the number of molecules that adsorbed at the interfacial region to the number of molecules that merely diffuse without adsorption within the experimental region. The number of adsorbed molecules can be counted manually or by image-processing software. But the number of diffusing molecules can only be estimated from sample concentrations, volumes, and the specifications of the microscope and ICCD camera. In our experiments, the 5-μL RPE solution dispensed on the surface formed a layer 10.3 μm thick between the prism and the coverslip (an area of 484 mm^2). Assuming a uniform distribution of RPE molecules, the estimated molecule numbers within the observation region (an area of 13271 μm^2) were 1378 for 16.7 pM and 138 for 1.67 pM. Native fluorescent images of RPE at the fused-silica surface were recorded at the pH range of 4.0-8.2. Our observations were similar to those reported by Kang and Yeung.[21] At pH 5.0, a significant fraction of RPE molecules were adsorbed and remained on the surface for a relatively long time. That made the counting results more reliable than at other pH conditions.

Experimental Counting Results. The simplest way of counting the adsorbed RPE molecules regardless of the time these molecules stay adsorbed is to project two consecutive frames together and count the number of bright particles as described in the experimental section. However, this method would definitely underestimate the adsorbed molecule number because the count is actually the number of molecules that stay adsorbed for at least two frames (or 200 ms) and not the number actually adsorbed in the first frame. A better way of counting is demonstrated in Figure 5A, in which the adsorbed molecule numbers are plotted against the numbers of projected frames ($n \geq 2$). The data points (black dots) are fitted with

an exponential decay function, which intercepts with the vertical line, $y = 1$, to give the actual count.

Over 30 sets of images taken at different surface positions and/or with different samples were processed for each concentration, and the average counts and standard deviations were 220±40 for 16.7 pM and 21±5 for 1.67 pM. Using the counting results and the estimated total number of RPE molecules in the observation region, the capacity factor was calculated to be $k_0^{'}$=0.25, with a ~20% relative standard deviation. This capacity factor can be compared with those determined in the capillary. A plot of $k_0^{'}$ vs. $1/r_c$ (r_c in μm) gives a straight line with a slope of 1.22, an intercept of 0.014 and a correlation coefficient of 0.998. These values are thus in good agreement with one another.

There may be additional differences in $k_0^{'}$ due to adsorption and desorption kinetics. Desorption in the capillary is greatly enhanced by the bulk flow driven by EOF and the electrophoretic migration of RPE molecules in contrast to the equilibrated distribution on the fused-silica prism surface. In addition, the surface properties are different: fused-silica capillary tubing produced by pulling a larger semi-fluid cylinder has smoother internal surface than the polished prism surface (20-10 scratch and dig).

The accuracy of this counting method could be affected by photobleaching of RPE. The effect can be estimated by counting the numbers of adsorbed molecules in every pair of consecutive frames in a set of 20+ images taken at the same position. The results showed that the counts decreased by 3 ~ 4 per 10-ms exposure to the 45-mW laser for the 16.7 pM samples and 0.5 ~ 1 per exposure for the 1.67 pM samples. These counts were much smaller than the deviations caused by uneven surface conditions and by molecule-counting statistics. Therefore, photobleaching was ignored in the calculations.

The desorption rate can also be estimated from the changes in counts over time. The white circles in Figure 5A correspond to the differences between every two adjacent molecule counts. These differences were then multiplied by the ratio of the total adsorbed molecule number to the number of adsorbed molecules remaining since the first frame. In this way, the white circles were converted to the black triangles in Figure 5B, showing nearly

constant desorption rates (~40 per frame or 200 ms). This behavior was expected because equilibrium has been established on the surface.

Conclusions

In the present work, three powerful tools – computer simulation, CE experiments, and single-molecule imaging – were united to advance the understanding of adsorption/desorption phenomena at solid/liquid interfaces. Computer simulation has proven to be an invaluable tool in nearly every field of science. Here simulated and experimental results together provided convincing support for the proposed mobility-based adsorption isotherm. This constitutes the first report of quantitative analysis of the adsorption capacity factor of target molecules on the bare fused-silica or coated capillary wall. On the other hand, single-molecule imaging experiments revealed the activities of individual molecules at the fused-silica prism surface, and these microscopic observations were complementary to the ensemble averages obtained from actual and simulated CE experiments. The adsorbed molecules were counted with confidence by extrapolation of the data points constructed from the adsorption times of individual molecules. The CE results and single-molecule detection results demonstrated good agreement, although further studies of the correlation between these results were somewhat limited by the relatively large random errors inherent to the single-molecule approach and by the differences of surface conditions. The ability to measure the adsorption strength with CE or with single-molecule detection is an important step forward in understanding the basic separation mechanisms of CE and LC.

Acknowledgments

E.S.Y. thanks the Robert Allen Wright Endowment for Excellence for support. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. DE-AC02-07CH11358. This work was supported by the Director of Science, Office of Basic Energy Sciences, Division of Chemical Sciences.

References

(1) Seidel-Morgenstern, A. *J. Chromatogr. A* **2004**, *1037*, 255-272.

(2) Towns, J. K.; Regnier, F. E. *Anal. Chem.* **1992**, *64*, 2473-2478.

(3) Schure, M. R.; Lenhoff, A. M. *Anal. Chem.* **1993**, *65*, 3024-3037.

(4) Ermakov, S. V.; Zhukov, M. Y.; Capelli, L.; Righetti, P. G. *J. Chromatogr. A* **1995**, *699*, 297-313.

(5) Gaš, B.; Stedry, M.; Rizzi, A.; Kenndler, E. *Electrophoresis* **1995**, *16*, 958-967.

(6) Ghosal, S. *Anal. Chem.* **2002**, *74*, 771-775.

(7) Fang, N.; Li, J.; Yeung, E. S. *Anal. Chem.* **2007**, *79*, 5343-5350.

(8) Ghosal, S. *J. Fluid Mech.* **2003**, *491*, 285-300.

(9) Graf, M.; Garcia, R. G.; Watzig, H. *Electrophoresis* **2005**, *26*, 2409-2417.

(10) Enderlein, J. *Biophys. J*. **2000**, *78*, 2151-2158.

(11) Hill, E. K.; de Mello, A. J. *Analyst* **2000**, *125*, 1033-1036.

(12) Kang, S. H.; Shortreed, M. R.; Yeung, E. S. *Anal. Chem.* **2001**, *73*, 1091-1099.

(13) Funatsu, T.; Harada, Y.; Tokunaga, M.; Saito, K.; Yanagida, T. *Nature* **1995**, *374*, 555-559.

(14) Xu, X. H.; Yeung, E. S. *Science* **1997**, *275*, 1106-1109.

(15) Houseal, T. W.; Bustamante, C.; Stump, R. F.; Maestre, M. F. *Biophys. J*. **1989**, *56*, 507-516.

(16) Glazer, A. N. *Annu. Rev. Microbiol.* **1982**, *36*, 173-198.

(17) Glazer, A. N. *Methods Enzymol.* **1988**, *167*, 291-303.

(18) Goulian, M.; Simon, S. M. *Biophys. J*. **2000**, *79*, 2188-2198.

(19) Haugland, R. P. *The Handbook: A Guide to Fluorescent Probes and Labeling Technologies*, 10th ed.; Invitrogen Corp., 2005.

(20) Glazer, A. N.; Stryer, L. *Trends Biochem. Sci.* **1984**, *9*, 423-427.

(21) Kang, S. H.; Yeung, E. S. *Anal. Chem.* **2002**, *74*, 6334-6339.

(22) He, Y.; Li, H. W.; Yeung, E. S. *J. Phys. Chem. B* **2005**, *109*, 8820-8832.

(23) Bowser, M. T.; Chen, D. D. Y. *J. Phys. Chem. A* **1998**, *102*, 8063-8071.

(24) Bowser, M. T.; Chen, D. D. Y. *J. Phys. Chem. A* **1999**, *103*, 197-202.

(25) Miller, J. M. *Chromatography: Concepts and Contrasts, 2nd ed.*, Wiley-Interscience: Hoboken, New Jersey, 2005.

Figure Captions

Figure 1. Normalized simulated peak profiles for various injected analyte concentrations. The concentration (in mol m^{-3}) in the initial injection plug is displayed on top of each curve. Other conditions are listed in the text.

Figure 2. (A) Simulated adsorption isotherms over the full concentration range using three nonlinear regression equations: the 3-parameter form (dot-dash), $y = 1.5286\times10^{-8} + 0.4598(1-e^{-0.5082x})$; the 5-parameter form (long-dash), $y = 1.5168\times10^{-8} + 0.1299(1-e^{-2.1743x}) + 0.3479(1-e^{-0.3181x})$; and the 7-parameter form (solid) as shown above the plot. (B) Linear adsorption isotherm for low concentrations (short-dash) is compared with the 7-parameter regression curve (solid). (C) Plot of partition ratio vs. analyte concentration over the full range. Three forms of regression equations were tested, and the one that gave a highly significant fit (solid) is shown. (D) Linear (short-dash) regression curve vs. the nonlinear (solid) one. All four panels use concentration as the x-axis.

Figure 3. Electropherograms of RPE adsorption experiments. The electropherograms are presented with a transformed mobility scale (based on the EOF peak). The left panel displays seven normalized RPE peaks. The right panel shows the location of the coumarin 535 peaks and their recorded fluorescence intensities (relative fluorescence units). The distortion of the coumarin peak becomes more severe as the RPE concentration gets higher mainly because of the increasingly altered EOF.

Figure 4. (A) Experimental adsorption isotherm. The error bars indicate the standard deviations of four consecutive runs at each RPE concentration. The maximum mobility obtained from the nonlinear regression is -0.0002177 $cm^2\ V^{-1}\ s^{-1}$ with a ~0.5% RSD. (B) Linear regression of the points in the low end of the RPE concentration range. The minimum mobility given by the y-intercept is -0.0002428 $cm^2\ V^{-1}\ s^{-1}$ with a ~0.5% RSD. (C) Nonlinear regression of the correlation between the partition ratio and RPE concentration. (D) Linear regression for the low concentration portion of (C). All four panels use RPE concentration as the x-axis.

Figure 5. (A) Plot of the average adsorbed molecule number vs. the number of projected frames. The white circles that show the differences between consecutive counts are positioned in the middle of the two corresponding black circles. (B) Original and converted differences between two consecutive counts.

Figure 1.

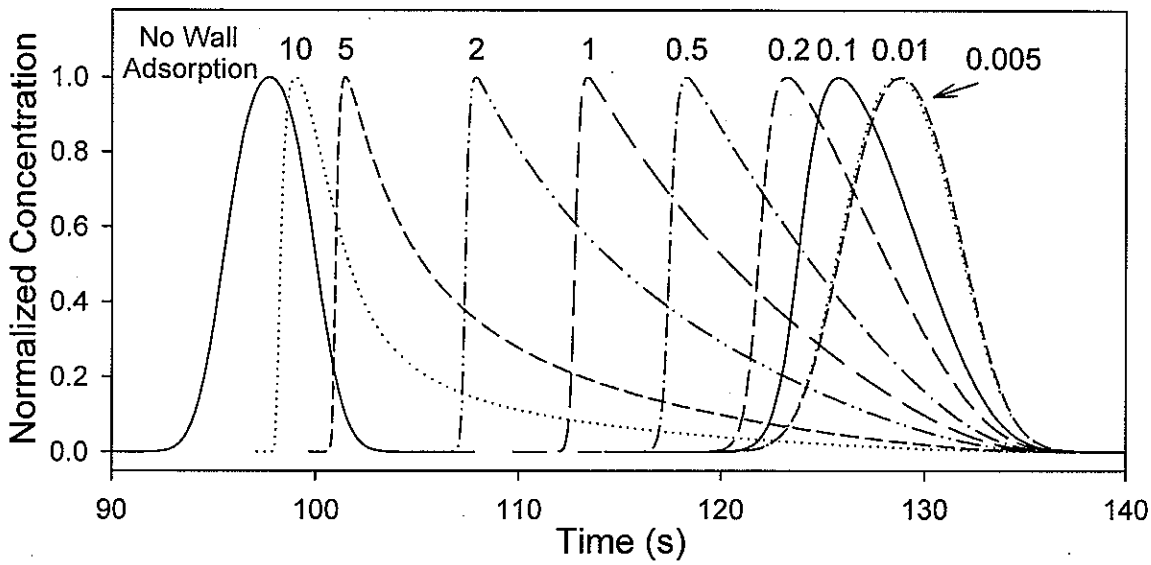

No Wall Adsorption
10
5
2
1
0.5
0.2
0.1
0.01
0.005
Normalized Concentration
1.0
0.8
0.6
0.4
0.2
0.0
90
100
110
120
130
140
Time (s)

Figure 2.

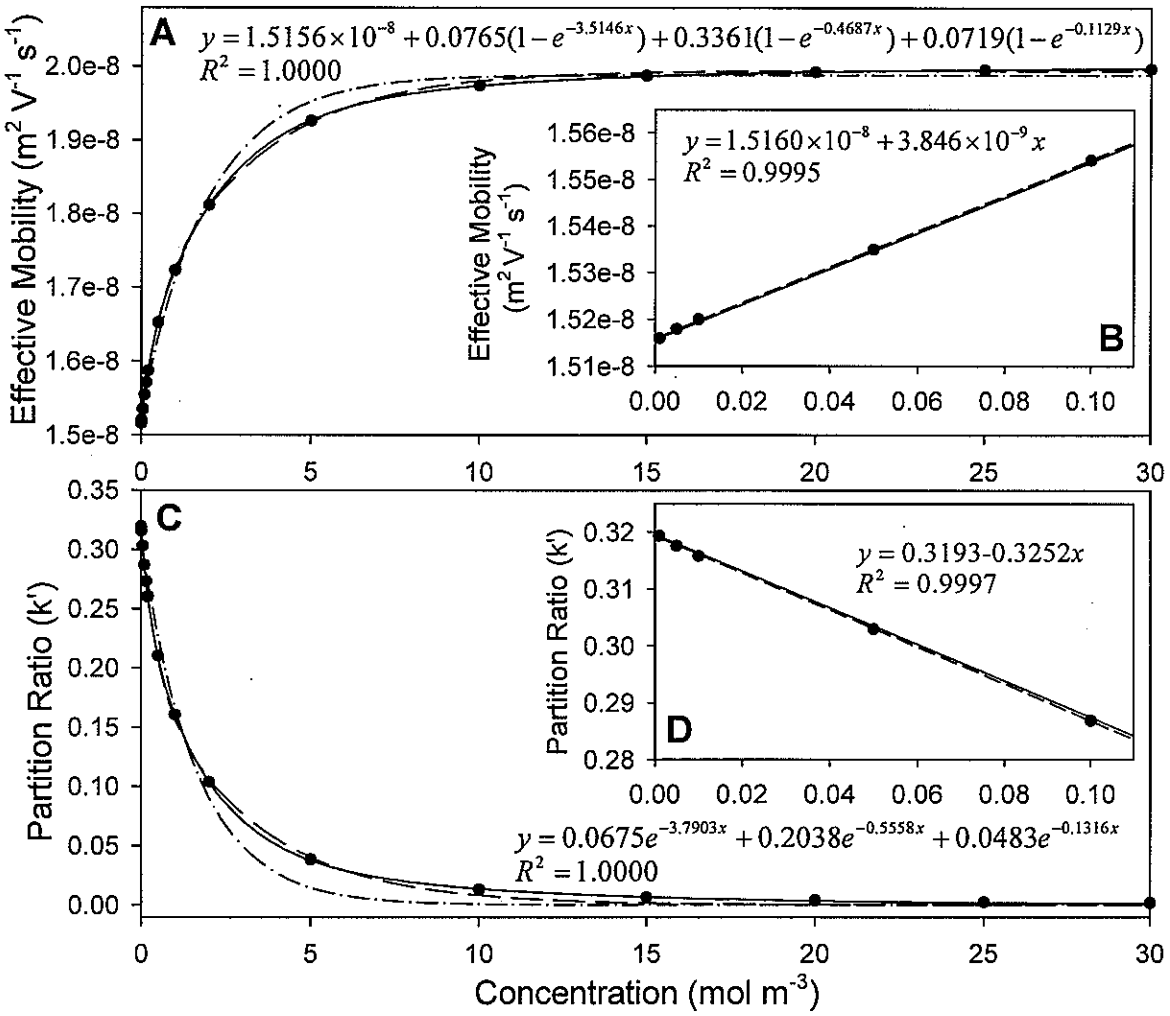

A
$y = 1.5156\times10^{-8} + 0.0765(1-e^{-3.5146x}) + 0.3361(1-e^{-0.4687x}) + 0.0719(1-e^{-0.1129x})$
$R^2 = 1.0000$
Effective Mobility ($m^2 V^{-1} s^{-1}$)
2.0e-8
1.9e-8
1.8e-8
1.7e-8
1.6e-8
1.5e-8
Effective Mobility ($m^2 V^{-1} s^{-1}$)
$y = 1.5160\times10^{-8} + 3.846\times10^{-9} x$
$R^2 = 0.9995$
1.56e-8
1.55e-8
1.54e-8
1.53e-8
1.52e-8
1.51e-8
B
0.00 0.02 0.04 0.06 0.08 0.10
0 5 10 15 20 25 30
C
Partition Ratio (k')
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Partition Ratio (k')
$y = 0.3193-0.3252x$
$R^2 = 0.9997$
0.32
0.31
0.30
0.29
0.28
D
0.00 0.02 0.04 0.06 0.08 0.10
$y = 0.0675e^{-3.7903x} + 0.2038e^{-0.5558x} + 0.0483e^{-0.1316x}$
$R^2 = 1.0000$
0 5 10 15 20 25 30
Concentration (mol m^{-3})

Figure 3.

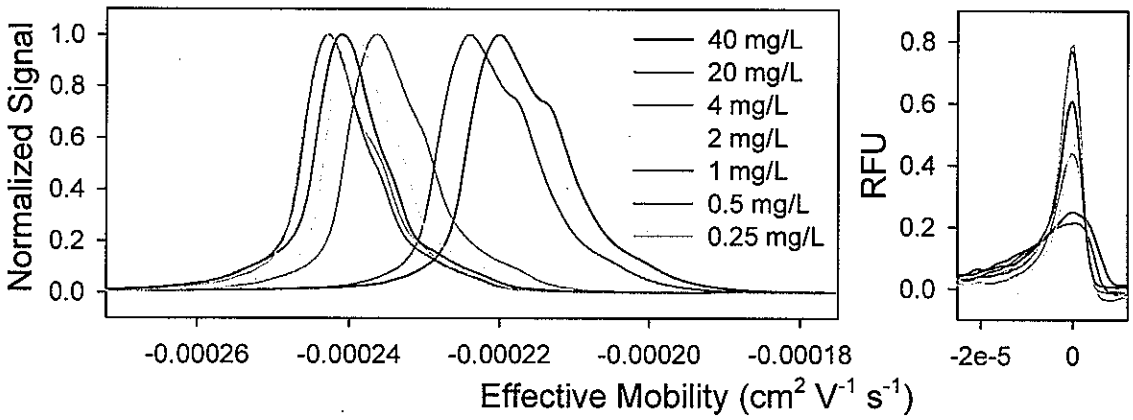

Normalized Signal
1.0
0.8
0.6
0.4
0.2
0.0
40 mg/L
20 mg/L
4 mg/L
2 mg/L
1 mg/L
0.5 mg/L
0.25 mg/L
-0.00026
-0.00024
-0.00022
-0.00020
-0.00018
Effective Mobility (cm^2 V^{-1} s^{-1})
RFU
0.8
0.6
0.4
0.2
0.0
-2e-5
0

Figure 4.

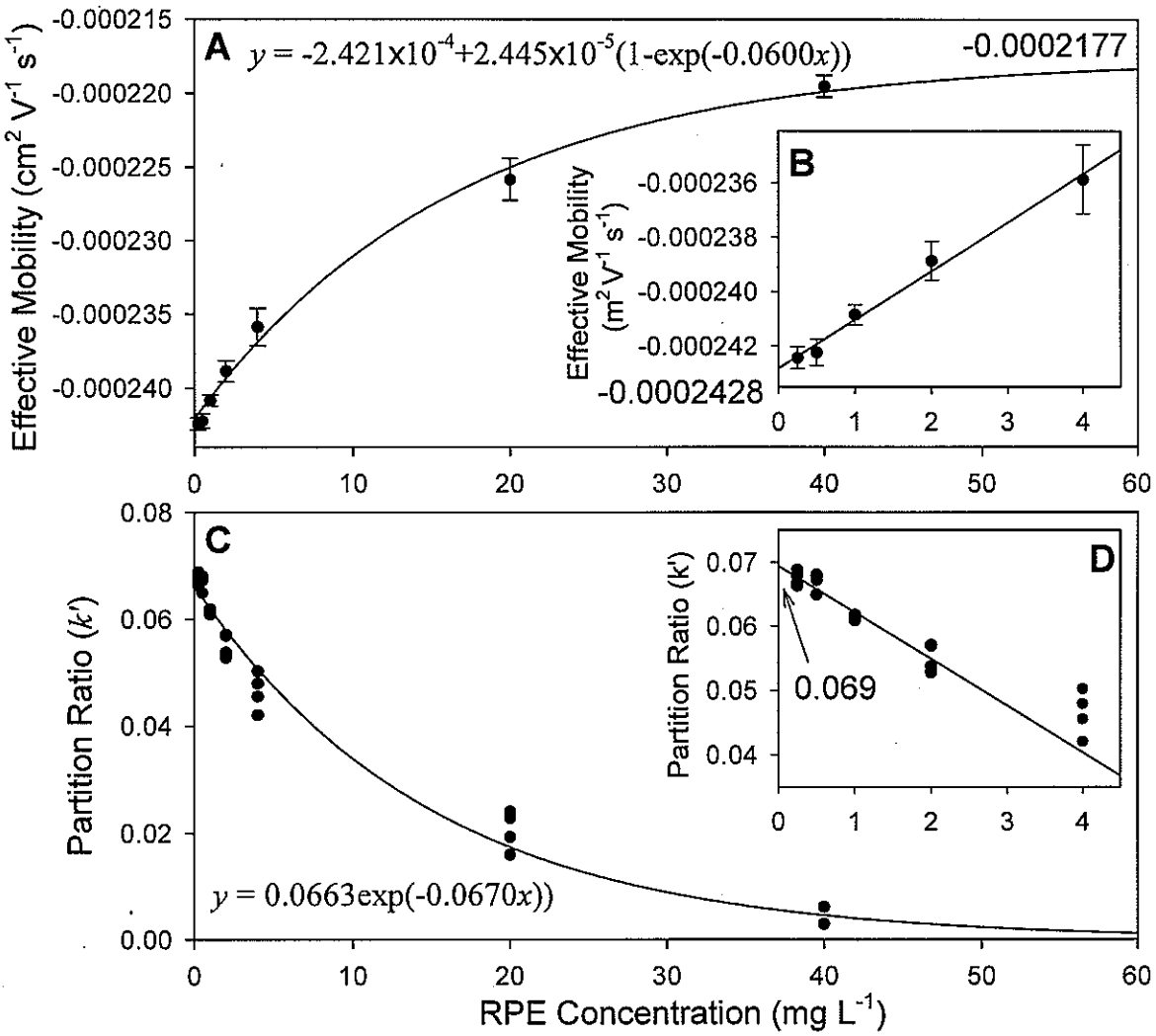
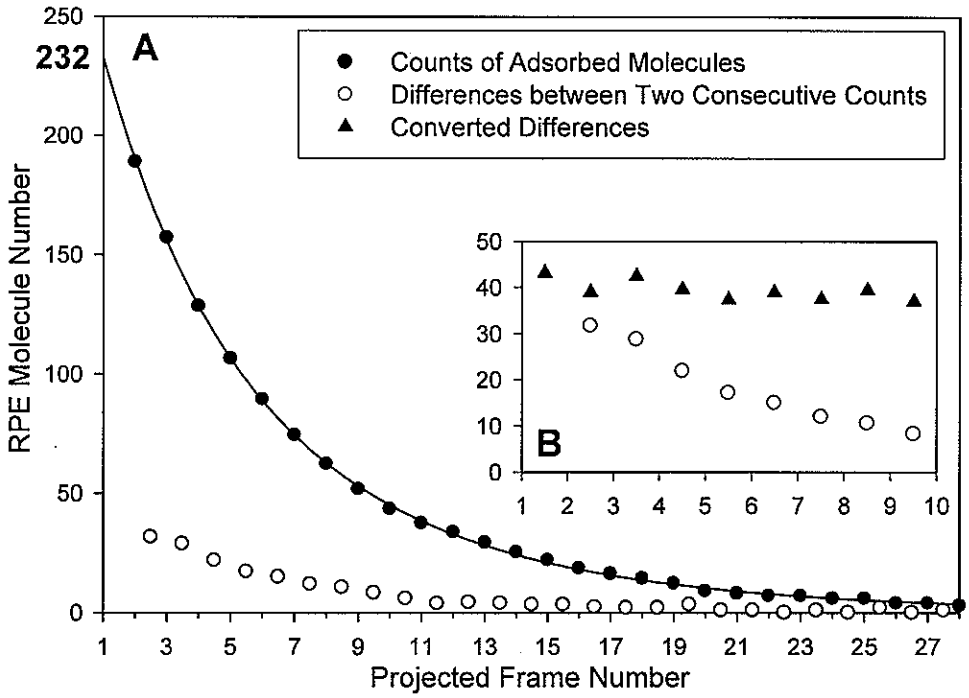

A
$y = -2.421x10^{-4}+2.445x10^{-5}(1-\exp(-0.0600x))$
-0.0002177
Effective Mobility ($cm^2 V^{-1} s^{-1}$)
B
Effective Mobility ($m^2 V^{-1} s^{-1}$)
-0.0002428
C
$y = 0.0663\exp(-0.0670x))$
Partition Ratio (k')
D
0.069
RPE Concentration (mg L^{-1})

Figure 5.


A
232
Counts of Adsorbed Molecules
Differences between Two Consecutive Counts
Converted Differences
RPE Molecule Number
Projected Frame Number
B
250
200
150
100
50
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10

CHAPTER 5. COLLOIDAL GRAPHITE ASSISTED LASER DESORPTION/IONIZATION (GALDI) MS AND MS^n OF SMALL MOLECULES: DIRECT PROFILING AND MS IMAGING OF SMALL METABOLITES FROM FRUITES

A paper published in Analytical Chemistry*

Hui Zhang, Sangwon Cha, Edward S. Yeung

Abstract

Due to a high background in the low mass region, conventional MALDI is not as useful for detecting small molecules (molecular weights < 500 Da) as it is for large ones. Also, spatial inhomogeneity that is inherent to crystalline matrixes can degrade resolution in imaging mass spectrometry (IMS). In this study, colloidal graphite was investigated as an alternative matrix for laser desorption/ionization (GALDI) in IMS. We demonstrate its advantages over conventional MALDI in the detection of small molecules such as organic acids, flavonoids and oligosaccharides. GALDI provides good sensitivity for such small molecules. The detection limit of fatty acids and flavonoids in negative ion mode are in low femtomoles range. Molecules were detected directly and identified by comparing the MS and MS/MS spectra with those of standards. Various fruits were chosen to evaluate the practical utility of GALDI since many types of small molecules are present in them. Distribution of these small molecules in the fruit was investigated by using IMS and IMS/MS.

---

* Reprint with permission from Analytical Chemistry **2007,** 79(17), 6575-6584.
Copyright © 2007 American Chemical Society

Introduction

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) has been extensively used for the analysis of large molecules such as proteins[1] and synthetic polymers.[2] The key feature of MALDI MS are its soft ionization characteristic and simplified spectra as mostly singly charged species are generated. Compared to electrospray ionization (ESI) MS, MALDI bears other advantages such as better tolerance to interference from salts and buffers and simpler sample preparation. MALDI has also proven to be very useful for the analysis of medium-size molecules (500-10 kDa) such as peptides,[3] oligonucleotides,[4] and oligo-saccharides.[5] However, the analysis of small molecules (< 500 Da) by conventional MALDI MS is far less successful than that of larger molecules because the analyte ions are strongly interfered with or are suppressed by the matrix-related ions that are predominant at the low *m/z* range.

Different approaches have been employed in MALDI MS to minimize the background in the low mass range. Reports includes derivatization of the analyte molecules to a higher molecular weight[6] or using a matrix with higher molecular weight such as porphyrin (MW 974.6).[7] Extra sample preparation was then needed, thereby limiting the classes of analytes that can be detected. It has been observed that matrix ions can be suppressed dramatically and sometimes complete suppression can be achieved under well controlled conditions.[8, 9] For example, surfactant additives such as cetyltrimethylammonium bromide (CTAB) have been reported to substantially suppress the background from α-cyano-4-hydroxycinnamic acid (CHCA).[10] Laser intensity and the relative molar ratio of matrix to analyte are the major parameters to adjust. However a suitable molar ratio is not always achievable especially for native biological samples.

Many inorganic materials have been tested as matrixes for surface-assisted laser desorption/ionization (SALDI), including different metal powders and metal oxide nanoparticles such as Ag, Au, Co, Al, Mn, Mo, Zn, Sn, W, Fe_3O_4, SnO_2, TiO_2, WO_3, ZnO, etc.[11-15] Generally, those SALDI MS can provide a cleaner background than conventional MALDI MS as no interference peaks from fragment ions of the organic matrixes were present. Another matrix-free approach for laser desorption/ionization on porous silicon

(DIOS) was extensively studied since 1999.[16, 17] Porous silicon surfaces were etched from crystalline silicon chips with hydrofluoric acid and functionalized as the laser desorption/ionization matrix as well as trapping agents for analyte molecules. Small molecules including pharmaceuticals, nucleic acids, carbohydrates, and steroids were successfully detected.[18-20] In a more recent work, commercially available silicon nanoparticles were utilized as an LDI matrix and the silicon powder preparation was optimized for the analysis of small molecules.[21] Different kinds of carbon materials, including graphite particles,[22] graphite plates,[23, 24] graphite suspension in different solvents,[25-27] graphite trapped in silicone polymer,[28] activated carbon powders,[29] functionalized carbon nanotubes[30-33] and fullerenes,[34] and more recently pencil lead,[35-37] have been suggested as alternative matrixes for LDI MS. Many kinds of analyte molecules over a wide mass range (100-6000 Da) have been detected, such as peptides,[20, 26-29, 31, 33-36] phospholipids,[25] oligosaccharides,[30-33, 35] fatty acids,[24, 36] synthetic polymers[7, 23, 26, 31, 32, 35, 37] and other various organic compounds.[7, 15, 22-29, 31, 33-37] A more detailed description of graphite-LDI can be found in our previous paper.[38] Recent reviews about small molecules MALDI MS[39] and matrix-free LDI MS can be found elsewhere.[40]

Imaging mass spectrometry (IMS) has proven to be a powerful technology for direct profiling and imaging of elements and biomolecules in tissue sections. Secondary ion mass spectrometry (SIMS),[41, 42] MALDI[43-47] and desorption electrospray ionization (DESI)[48] have been applied as desorption/ionization techniques for the IMS of molecules such as metal elements, peptides, proteins, lipids and other metabolites. SIMS has the best spatial resolution among the three and DESI requires the least sample preparation and allows true *in situ* measurement with the simplest instrumentation.[48] The spatial resolution of MALDI IMS is in between that of SIMS and DESI IMS, usually ranging from 80-200 μm in diameter. The diverse choices of lasers and matrixes make MALDI MS suitable for fast, simultaneous and high-throughput analyses of metabolites from tissue samples. MALDI equipped with UV laser has been successfully demonstrated for the imaging of peptides, proteins and lipids.[43-47] Due to the high background problem as discussed earlier there is limited application of UV-MALDI for imaging of small molecules (< 500 Da). Infrared (IR) MALDI was introduced recently as a technique for imaging small metabolites from fruit samples.[49] Water is used as

the natural matrix for IR MALDI, but it is inevitable that the sample may dry out during the process of IR irradiation. Different locations will thus give different sensitivities due to inhomogeneous water content. So far, molecules that can be detected by IR MALDI are quite limited, either because of the low desorption/ionization efficiency or low detection sensitivity. Furthermore, the spatial resolution of IR-IMS is inherently worse than that of UV-IMS.

Previously we demonstrated that colloidal graphite was a good LDI matrix for the analysis of molecules in 500-1000 Da range, such as different lipid species.[38] This matrix contains fine particles and is spatially homogeneous, making it suitable for quantitative imaging. The colloidal property also allows it to be easily sprayed to form a layer on top of tissue samples and thus simplifies imaging experiments. In this study, we investigated the applicability of colloidal graphite as an alternative LDI matrix for the analysis of even smaller metabolite molecules. Fruits contain many kinds of small molecules such as long-chain fatty acids, small oligosaccharides, and flavonoids, so they serve as good systems to test the performance. GALDI MS and tandem MS were used to identify the ionized species, while IMS and IMS/MS were utilized to map the distribution of those molecules in fruit slices.

Experimental Section

Standards such as long-chain fatty acids, oligosaccharides and flavonoids were purchased from Sigma-Aldrich (St. Louis, MO). Dihydroxybenzoic acid (DHB) from Bruker Daltonics (Billerica, MA) and CHCA solution from Agilent Technologies (Palo Alto, CA) were used as standard MALDI matrixes. 2-Propanol-based colloidal graphite aerosol spray (Aerodag G) was obtained from Acheson Colloids (Port Huron, MI). Pure water was obtained from a MilliQ water purification system (Billerica, MA). All other chemicals were purchased from Fisher Scientific (Fairlawn, NJ).

Apple and strawberry fruits were purchased from a local grocery store. Apple skin was peeled off by a sharp razor blade and attached to the stainless steel plate by double-sided tape. Apple juice was collected from crushed flesh onto a glass slide, dried and used directly.

A cryostat from International Equipment Co. (Needham Heights, MA) was used for cryosectioning. Fruit chunks were dipped in liquid N_2 before they were cryosectioned into about 15-μm thick specimens and stored at –20 °C before mass spectrometric analysis. No optimum cutting temperature (OCT) compounds were used to embed the fruit samples as the interference to mass spectra from OCT compounds is known.[50] Sectioned fruit slices were directly transferred and mounted onto the stainless-steel plate. Before applying colloidal graphite solution, the slices were dried under moderate vacuum (~50 Torr) at room temperature for half an hour.

Long-chain fatty acid standards were prepared by dissolving stearic acid (C18, MW 284.48), pentacosanoic acid (C25, MW 382.66), hexacosanoic acid (C26, MW 396.69), octacosanoic acid(C28, MW 424.74), and melissic acid (C30, MW 452.80) in chloroform to a final concentration of 200 pmole/μl each. For flavonoid standards, quercetin (MW 302.24), kaempferol (MW 286.23), phloretin (MW 274.27), and apigenin (MW 270.24) were dissolved individually in DMSO to give concentration of 5mg/ml of each; then the four standard solution were mixed and further diluted to a final concentration of 200 pmole/μl each in water/acetonitrile/trifluoroacetic acid (49.95/49.95/0.1). Oligosaccharide standards were prepared by dissolving ribose (MW 150.13), glucose (MW 180.16), sucrose (MW 342.30), N-acetyl-D-lactosamine (LacNAc) (MW 383.35), maltotriose (MW 504.44), and maltotetraose (MW 666.58) in water/acetonitrile/trifluoroacetic acid (49.95/49.95/0.1) to a final concentration of 100 ng/μl each. 20 mg/mL DHB solution in 70% methanol and 30% water (containing a 0.1% trifluoroacetic acid) was prepared. Commercial Agilent CHCA solution at 6 mg/ml in 36/56/8 methanol/acetonitrile/water was purchased and used directly. Four times dilution of colloidal graphite solution with 2-propanol was used for GALDI MS and IMS.

For all mass spectrometric analysis and IMS, an LTQ linear ion trap mass spectrometer equipped with vMALDI source (Thermo Electron, Mountain View, CA) was used. The N_2 laser (337 nm) is guided to the source by a fiber-optic cable and has a maximum output of 280 μJ/pulse (before entering the optical fiber cable). The measured

laser spot size is ~100 μm in diameter on the sample plate surface. A more detailed description of the LTQ with vMALDI source has been reported elsewhere.[46]

For conventional MALDI MS, 1 μL of DHB or CHCA matrix solution was applied onto the stainless-steel sample plate and let to dry in air, followed by 1 μL of sample solution on top of the matrix crystals. For GALDI MS of standards, 0.5 μL of diluted colloidal graphite solution was applied onto the stainless-steel sample plate by a micropipette and let to dry in air. Then 1 μL of standard solution was applied on top of the graphite spot. To obtain mass spectra from apple juice with GALDI, 1 μl of fresh apple juice was applied onto a dried (0.5 μl) graphite spot. For apple peels and fruit slices, diluted colloidal graphite solution was applied by a double-action airbrush (Aztek A470 with a 0.30 mm nozzle from Testor, Rockford, IL). The whole fruit slice was covered with colloidal graphite homogeneously by spraying with 20 psi air pressure and 15 cm away from the sample plate for 30 s. Peak identification was made by comparing both mass and tandem mass spectra with those of standards.

Optical images of fruit slices were taken inside the vMALDI source before IMS. Serial optical images were taken every 1 mm movement of the sample stage in either *x*- or *y*-direction. Each segment of the images has a size of 140 pixels by 170 pixels. These segments of optical images were reconstructed as one optical image for one fruit slice. To collect mass spectra, the same sample plate was rastered with 100 μm steps. For each raster point, a mass spectrum was recorded for desorbed ions and integrated over 3-5 laser shots. In the cases of IMS/MS, target precursor ions (*m/z* 191 or *m/z* 301) were first selected based on the mass spectral profiles of strawberry. Then the first-generation product-ion spectra of the selected precursor ion were collected from all rastering points on the strawberry slice. More laser shots were required for MS^2 experiments and so 9 laser shots were averaged for each raster point.

Custom software "vMALDI data browser" (Version 1.0) was used to extract mass spectra from specific locations and generate chemically selective images. This software was provided by the instrument vendor (Thermo Electron, Mountain View, CA). The mass window for generating images was 0.5 Da. Intensities of the selected ion were normalized by

dividing by the total ion current of each mass spectrum. Then, chemically selective images were plotted as 3-D maps with the 3rd dimension being the normalized intensity.

Results and Discussion

MALDI and GALDI MS of Standard Mixtures of Fatty Acids, Flavonoids and Oligosaccharides

Different classes of compounds were selected to compare the performance of conventional MALDI and GALDI. Long-chain fatty acids (C18-C30) were selected as one group of standards because of their important roles in many metabolic pathways of living organisms.[51] Fatty acids can be detected by GC-MS with proper derivatization such as esteritication to reduce the polarity and increase volatility.[52] HPLC-ESI MS can detect native fatty acids in the negative-ion mode through deprotonization.[53] However, the basic pH condition which is required to produce such negative ions is not compatible with most reverse-phase C18 columns that typically require acidic mobile phases. To overcome this problem, judicious derivatization was needed.[54] MALDI MS can eliminate the separation step but conventional MALDI matrixes do not work well for detection due to a high background and ion suppression in that mass region. This is underscored by the fact that none of the selected standard fatty acids were detected with DHB or CHCA in either positive-ion mode or negative-ion mode (data not shown). With GALDI, all five fatty acids (100 ng each) were detected as deprotonated peaks ($[M - H]^-$), as shown in Figure 1. It is noteworthy that the background in the negative-ion mode of GALDI is very clean up to at least *m/z* 1000 without any pretreatment. In fact, there are only a few low number carbon cluster ions, such as C_{12}^--C_{14}^-; and the intensity of such peaks are much lower than those for the analytes. GALDI is very sensitive to detect those fatty acids in negative ion mode. Fig. 2 shows the spectrum of the fatty acid mixture with sample loading of 100fmole of each on a 3mm-in-diameter spot. Such fatty acids can also be detected as potassium adduct ions ($[M + K]^+$) in positive ion mode (data not shown), though the detection sensitivity (detection limit: 50pmole/spot) is not as good as in negative ion mode .

Another group of standards tested were natural phenolic molecules, the flavonoids. It was estimated that 2% of all carbon photosynthesized by plants is converted into flavonoids or related compounds.[55] They have been reported to have antioxidant, antiatherosclerotic and anti-neurodegenerative properties, and are also known to be beneficial for the prevention of chronic diseases like cancer and heart diseases.[56, 57] In the positive-ion mode, quercetin, kaempferol and apigenin were detected as $[M + H]^+$ in both MALDI and GALDI experiments (Fig. 3a-c). $[M + Na]^+$ and $[M + 2Na-H]^+$ ions for these three flavonoids can also be detected with GALDI. Phloretin was not detected with any of the three matrixes. This suggests that the center ring in the other three flavonoids plays an important role for protonation. However, it was possible to detect phloretin with GALDI in the negative-ion mode (Fig. 3f). Several peaks in the region of *m/z* 150-200 in the spectra were identified as in-source-fragments for the flavonoids by using tandem MS. For example, *m/z* 167 was identified as a fragment of phloretin while *m/z* 151 and *m/z* 179 were fragments of quercetin. Unlike GALDI, which has a clean background, with DHB and CHCA, matrix peaks are predominant and none of the four flavonoid standards were detected in the negative-ion mode (Fig. 3d-e).With GALDI, negative ion mode provides better sensitivity for detection of those flavonoid standards than positive ion mode does, and the detection limit are 50 fmole and 200 fmole/3mm-in-diameter spot, respectively.

The third group of standards selected were oligosaccharides. Due to the lack of acidic or basic groups, they are difficult to be ionized with conventional MALDI. DHB can be used to detect large oligosaccharides (> 1000 Da)[58] but smaller oligosaccharides are strongly interfered with by matrix ions. We tested oligosaccharides between 150-700 Da with DHB (Fig. 4a), CHCA (Fig. 4b), and graphite matrixes (Fig. 5). With CHCA, only larger molecules like LacNAc, maltotriose and maltotetraose can be detected as $[M + Na]^+$ but not the smaller ones such as ribose, glucose and sucrose. DHB works slightly better than CHCA in that the small oligosaccharides such as glucose and sucrose can be detected. However, the intensity of the small saccharide peaks was very low. For example $[\text{ribose} + Na]^+$ and $[\text{glucose} + Na]^+$ peaks are strongly suppressed as they are very close to matrix peaks $[\text{DHB} + Na]^+$ and $[\text{DHB} + 2Na - H]^+$ respectively. LacNAc has the highest detection sensitivity among all the oligosaccharides. This may suggest that the N-acetyl group has more affinity

for sodium ions than the other moieties of oligosaccharides. In both the CHCA and DHB MALDI mass spectra (Fig. 4a-b), the matrix peaks were predominant and oligosaccharides peaks were strongly interfered with. With GALDI, all six oligosaccharides were detected as $[M + Na]^+$ with decent signal-to-noise ratios, as shown in Fig. 5. $[M + Na -18]^+$ peaks were also observed for all five oligosaccharides as water loss (Fig. 5). Neutral loss (NL) of 60 are known to result from the major cross-ring cleavage fragments of oligosaccharides with 1,4 linkages.[59] Here, *m/z* 629 and *m/z* 467 were observed as such fragments of maltotetraose and maltotriose respectively. The peak at *m/z* 305 corresponds to the major fragment of LacNAc. All of the above fragments were verified by tandem MS (data not shown). All cross-ring cleavage fragments were marked with asterisks in Fig. 5. That the positive-ion mode of GALDI gave a noisier background compared to the negative-ion mode did not prevent the successful detection of small oligosaccharides. Small oligosaccharides such as hexose (glucose or fructose) and sucrose can also be detected as $[M - H]^-$ under the negative-ion mode peaks, *vide infra*. The detection sensitivity of oligosaccharides with GALDI under positive ion mode is better than with GALDI under negative ion mode. For instance, the detection limit of sucrose is 20pmole and 100pmole/3mm-in-diameter spot, respectively. This can be understood as oligosaccharides have neither carboxyl group as fatty acids, or aromatic rings as flavonoids which can stabilize the deprotonated phenol group.

Direct Detection of Small Metabolite Molecules from Fruit Samples

Because of the sensitivity and background issues, GALDI in the positive-ion mode was used to detect oligosaccharides and the negative-ion mode was used to detect other metabolite molecules from fruit samples. Unlike experiments with the standards, colloidal graphite solution was applied on top of the sample, otherwise the graphite particles may not be accessed by laser irradiation. For imaging purposes, homogeneous coverage over the sample area is a must. The application methodology has been optimized for IMS of mouse brain tissues in our previous work[38] and was used in this study without modification.

A list of small molecules detected directly from apple and strawberry samples were summarized in Table I, with concentration data wherever available.[60-65] All identifications were made by comparing the MS spectra with those of standards. MS/MS data were also compared with those of standards except those at too low a concentration to give meaningful MS/MS data, as marked with asterisks. Tandem MS data is indispensable to identify isobaric ions, such as *m/z* 191, which were observed from both apple and strawberry. Tandem mass data of *m/z* 191 from apple and strawberry flesh are shown in Fig. 6. Notably, citric acid gave a predominant product ion at *m/z* 111 while quinic acid has specific product ions at *m/z* 85, 93 and 127, as reported previously.[66] By tandem MS it was confirmed that the peaks at *m/z* 191 were from quinic acid in apple while such peak came from citric acid in strawberry.

Fig. 7 shows typical mass spectra taken from different parts of the apple. Fruits in supermarkets are always coated with a thin layer of wax for preservation and for better appearance. Fatty acids are one major class of components of wax and many of those up to C28 fatty acids were detected from apple peel, as shown in Fig. 7a. Other compounds such as sugars or flavonoids were absent from the mass spectrum. The reason may be that such compounds were covered by the wax layer and were not accessible. Fatty acids may be naturally present on the apple peels as well; however, those cannot be discriminated from the species in the artificial wax layer.

As shown in Fig. 7b, fresh apple juice gave malic acid, quinic acid, palmitic acid, and linolenic acid in the negative-ion mode. Hexose (glucose or fructose) and sucrose were detected as deprotonated ions as *m/z* 179 and *m/z* 341, respectively. Quercetin is one of the major flavonoids contained in apple and it was also detected. Fig. 7c shows the spectrum in the positive-ion mode. Sodium adduct ions of hexose and sucrose were detected, as well as potassium adduct ions. According to the USDA nutrient database, fruits usually contain much higher amounts of potassium than sodium (90 mg vs. 0 mg/100 g for apple and 292 mg vs. 37 mg/100 g for strawberry).[64]

Fig. 7d is the representative spectrum taken from the apple core (endocarp). Here, organic acids such as malic acid and quinic acid were still observed, but the relative intensities were not as high as those from juice. On the other hand, higher amounts of

flavonoids such as phloretin and quercetin accumulated in the endocarp region. Apple peel also contains a lot of flavonoids and quercetin was the predominant one found on the inside of apple peel (data not shown).

Fig. 8a and 8b are representative mass spectra taken from strawberry. As reported previously with IR-MALDI,[49] strawberry contains a lot of citric acid. In our study, citric acid (*m/z* 191) was always the most intense peak in the negative-ion mode. GALDI can also detect those compounds which were too low in concentration to be detected by IR-MALDI, such as ascorbic acid and ellagic acid. Quercetin and kaempferol was also detected in the red part of strawberry (more details in IMS data below). Other deprotonated ions include palmitic acid, oleic acid, apigenin, hexose, and sucrose, as marked in Fig. 8a. Similar to the previous report,[49] citric acid and sucrose were barely detected in the seed region, as shown in Fig. 8b. Instead, C16 and unsaturated C18 fatty acids were the major components.

IMS of Metabolites from Apple and Strawberry

IMS in the negative-ion mode was performed to show the detailed distributions of different metabolite molecules on apple and strawberry slices, as Fig. 9 and Fig. 10 respectively. Small molecules such as malic acid, quinic acid, and sucrose distributed relatively evenly over the apple flesh part, as shown in Fig. 9b, c, and e. Long-chain fatty acid such as linoleic acid was detected all over the slice but more accumulated along the core line and bundles, as shown in Fig. 9d. Flavonoids also accumulated more in the core region but not in the flesh, as shown in Fig. 9f-i. Unlike linoleic acid, they seem to be only enriched in the bundles (both sepal and pedal), but not along the core line. Another interesting feature is that flavonoids were found in the *ventral carpellary* bundle (center of the apple slice) except for quercetin.

The images shown in Fig. 10 were scanned from a pie-shaped strawberry slice, with the red skin on the right-hand side and one seed in the middle-right part. Citric acid, apigenin, hexose and sucrose were distributed all over the flesh, while fatty acids such as linolenic and linoleic acids accumulated on the seed. The peak at *m/z* 301.2 was detected all over the

strawberry with a relatively constant concentration in the flesh, but showed higher local intensity on the outside (red part) and the seed areas. The candidates were quercetin (MW 302.24) and ellagic acid (MW 302.20) and the two species could be distinguished in tandem MS. Fig. 11a and 6b show the MS/MS spectra of ellagic acid and quercetin standards. *m/z* 151 and *m/z* 179 were specific fragments of quercetin due to the cleavage of the center ring, as reported previously;[56, 67] while ellagic acid is more rigid and only small fragments such as NL 28 and NL 44 were observed. MS/MS product ion spectra from strawberry flesh were similar to those from ellagic acid (Fig. 11c), while those at *m/z* 301.2 on the edge gave quercetin-like fragments (Fig. 11d). The chemically selective ion image at *m/z* 179 with precursor ions at *m/z* 301.2 is shown in Fig. 11e. The ambiguity in Fig. 10i is thus resolved.

Similarly, both citric acid (MW 192.13) and quinic acid (MW 192.17) were detected as *m/z* 191 with GALDI in the negative-ion mode. IMS/MS (data not shown) with precursor ions at *m/z* 191.2 showed that product ions at *m/z* 111.3 were detected all over the strawberry slice, while no specific fragments of quinic acid (*m/z* 85 or *m/z* 93) were observed. This suggests that all *m/z* 191 ions from strawberry were from citric acid (Fig. 10c).

Acknowledgements

E.S.Y. thanks the Robert Allen Wright Endowment for Excellence for support. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. DE-AC02-07CH11358. This work was supported by the Director of Science, Office of Basic Energy Sciences, Divisions of Chemical Sciences and Biosciences.

Table I. List of compounds detected directly from apple and strawberry and their average concentrations from the literature. All species have been checked by *m/z* against standards, and all but those with marked with asterisks have been checked by tandem MS.

	Apple			Strawberry		
	m/z	Peaks assigned	Expected Concentrations g/Kg	m/z	Peaks assigned	Expected Concentrations g/Kg
Organic acids	133 $[M-H]^-$	Malic Acid	$4.9 * 10^{-3}$ [60]	175 $[M-H]^-$	Ascorbic Acid*	0.59
	191 $[M-H]^-$	Quinic Acid	$0.76 * 10^{-3}$ [60]	191 $[M-H]^-$	Citric Acid	6.9-12.6 [62]
	255 $[M-H]^-$	Palmitic Acid	0.24	255 $[M-H]^-$	Palmitic Acid	0.12
	277 $[M-H]^-$	Linolenic Acid	0.09	277 $[M-H]^-$	Linolenic Acid	0.65
	279 $[M-H]^-$	Linoleic Acid	0.43	279 $[M-H]^-$	Linoleic Acid	0.90
	281 $[M-H]^-$	Oleic Acid	0.07	281 $[M-H]^-$	Oleic Acid	0.42
Phenolics	273 $[M-H]^-$	Phloretin	$4.7 * 10^{-3}$ [60]	269 $[M-H]^-$	Apigenin	0.00-0.01
	289 $[M-H]^-$	Epicatechin	$47.1 * 10^{-3}$	285 $[M-H]^-$	Kaempferol	$4.6 * 10^{-3}$
	301 $[M-H]^-$	Quercetin	$45.7 * 10^{-3}$	301 $[M-H]^-$	Ellagic Acid	0.06-0.5 [63]
	435 $[M-H]^-$	Phloridzin*	$55.9 * 10^{-3}$ [61]	301 $[M-H]^-$	Quercetin	$11.4 * 10^{-3}$
	447 $[M-H]^-$	Quercetin Glucosides*	0.13 [61]	431 $[M-H]^-$	Apigenin Glucosides*	
Oligo-saccharides	179 $[M-H]^-$ 203 $[M+Na]^+$ 219 $[M+K]^+$	Glucose/ Fructose*	83.3	179 $[M-H]^-$ 203 $[M+Na]^+$ 219 $[M+K]^+$	Glucose/ Fructose*	44.3
	341 $[M-H]^-$ 365 $[M+Na]^+$ 381 $[M+K]^+$	Sucrose	20.7	341 $[M-H]^-$ 365 $[M+Na]^+$ 381 $[M+K]^+$	Sucrose	4.7

Except noted, all other concentration data come from (1) See the USDA National Nutrient Database at http://www.nal.usda.gov/fnic/foodcomp/search/, and (2) See the USDA National Nutrient Database at **http://www.nal.usda.gov/fnic/foodcomp/Data/Flav/flav.pdf**

References

(1) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

(2) Bahr, U.; Deppe, A.; Karas, M.; Hillenkamp, F.; Giessmann, U. *Anal. Chem.* **1992**, *64*, 2866-2869.

(3) Kaufmann, R.; Kirsch, D.; Spengler, B. *Int. J. Mass Spectrom. Ion Processes* **1994**, *131*, 355-385.

(4) Lecchi, P.; Pannell, L. K. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 972-975.

(5) Finke, B.; Stahl, B.; Pfenninger, A.; Karas, M.; Daniel, H.; Sawatzki, G. *Anal. Chem.* **1999**, *71*, 3755-3762.

(6) Tholey, A.; Wittmann, C.; Kang, M. J.; Bungert, D.; Hollemeyer, K.; Heinzle, E. *J. Mass Spectrom.* **2002**, *37*, 963-973.

(7) Ayorinde, F. O.; Hambright, P.; Porter, T. N.; Keith, Q. L. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2474-2479.

(8) Knochenmuss, R.; Dubois, F.; Dale, M. J.; Zenobi, R. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 871-877.

(9) Knochenmuss, R.; Karbach, V.; Wiesli, U.; Breuker, K.; Zenobi, R. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 529-534.

(10) Guo, Z.; Zhang, Q. C.; Zou, H. F.; Guo, B. C.; Ni, J. Y. *Anal. Chem.* **2002**, *74*, 1637-1641.

(11) Sluszny, C.; Yeung, E. S.; Nikolau, B. J. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 107-115.

(12) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.

(13) McLean, J. A.; Stumpo, K. A.; Russell, D. H. *J. Am. Chem. Soc.* **2005**, *127*, 5304-5305.

(14) Kinumi, T.; Saisu, T.; Takayama, M.; Niwa, H. *J. Mass Spectrom.* **2000**, *35*, 417-422.

(15) Chen, C. T.; Chen, Y. C. *Anal. Chem.* **2005**, *77*, 5912-5919.

(16) Wei, J.; Buriak, J. M.; Siuzdak, G. *Nature* **1999**, *399*, 243-246.

(17) Shen, Z. X.; Thomas, J. J.; Averbuj, C.; Broo, K. M.; Engelhard, M.; Crowell, J. E.; Finn, M. G.; Siuzdak, G. *Anal. Chem.* **2001**, *73*, 612-619.

(18) Pihlainen, K.; Grigoras, K.; Franssila, S.; Ketola, R.; Kotiaho, T.; Kostiainen, R. *J. Mass Spectrom.* **2005**, *40*, 539-545.

(19) Compton, B. J.; Siuzdak, G. *Spectroscopy-an International Journal* **2003**, *17*, 699-713.

(20) Li, Q.; Ricardo, A.; Benner, S. A.; Winefordner, J. D.; Powell, D. H. *Anal. Chem.* **2005**, *77*, 4503-4508.

(21) Wen, X. J.; Dagan, S.; Wysocki, V. H. *Anal. Chem.* **2007**, *79*, 434-444.

(22) Zumbuhl, S.; Knochenmuss, R.; Wulfert, S.; Dubois, F.; Dale, M. J.; Zenobi, R. *Anal. Chem.* **1998**, *70*, 707-715.

(23) Kim, H. J.; Lee, J. K.; Park, S. J.; Ro, H. W.; Yoo, D. Y.; Yoon, D. Y. *Anal. Chem.* **2000**, *72*, 5673-5678.

(24) Park, K. H.; Kim, H. J. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1494-1499.

(25) Peng, S.; Edler, M.; Ahlmann, N.; Hoffmann, T.; Franzke, J. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2789-2793.

(26) Dale, M. J.; Knochenmuss, R.; Zenobi, R. *Anal. Chem.* **1996**, *68*, 3321-3329.

(27) Sunner, J.; Dratz, E.; Chen, Y. C. *Anal. Chem.* **1995**, *67*, 4335-4342.

(28) Li, X. P.; Wilm, M.; Franz, T. *Proteomics* **2005**, *5*, 1460-1471.

(29) Chen, Y. C.; Shiea, J.; Sunner, J. *J. Chromatogr. A* **1998**, *826*, 77-86.

(30) Pan, C. S.; Xu, S. Y.; Hu, L. G.; Su, X. Y.; Ou, J. J.; Zou, H. F.; Guo, Z.; Zhang, Y.; Guo, B. C. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 883-892.

(31) Ren, S. F.; Guo, Y. L. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 255-260.

(32) Ren, S. F.; Zhang, L.; Cheng, Z. H.; Guo, Y. L. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 333-339.

(33) Xu, S. Y.; Li, Y. F.; Zou, H. F.; Qiu, J. S.; Guo, Z.; Guo, B. C. *Anal. Chem.* **2003**, *75*, 6191-6195.

(34) Ugarov, M. V.; Egan, T.; Khabashesku, D. V.; Schultz, J. A.; Peng, H. Q.; Khabashesku, V. N.; Furutani, H.; Prather, K. S.; Wang, H. W. J.; Jackson, S. N.; Woods, A. S. *Anal. Chem.* **2004**, *76*, 6734-6742.

(35) Black, C.; Poile, C.; Langley, J.; Herniman, J. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1053-1060.

(36) Langley, G. J.; Herniman, J. M.; Townell, M. S. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 180-190.

(37) Berger-Nicoletti, E.; Wurm, F.; Kilbinger, A. F. M.; Frey, H. *Macromolecules* **2007**, *40*, 746-751.

(38) Cha, S.; Yeung, E. S. *Anal. Chem.* **2007**, *79*, 2373-2385.

(39) Cohen, L. H.; Gusev, A. I. *Anal. Bioanal. Chem.* **2002**, *373*, 571-586.

(40) Peterson, D. S. *Mass Spectrom. Rev.* **2007**, *26*, 19-34.

(41) Altelaar, A. F. M.; Klinkert, I.; Jalink, K.; de Lange, R. P. J.; Adan, R. A. H.; Heeren, R. M. A.; Piersma, S. R. *Anal. Chem.* **2006**, *78*, 734-742.

(42) Ostrowski, S. G.; Van Bell, C. T.; Winograd, N.; Ewing, A. G. *Science* **2004**, *305*, 71-73.

(43) Caprioli, R. M.; Farmer, T. B.; Gile, J. *Anal. Chem.* **1997**, *69*, 4751-4760.

(44) Chaurand, P.; Caprioli, R. M. *Electrophoresis* **2002**, *23*, 3125-3135.

(45) Chaurand, P.; Cornett, D. S.; Caprioli, R. M. *Curr. Opin. Biotechnol.* **2006**, *17*, 431-436.

(46) Garrett, T. J.; Prieto-Conaway, M. C.; Kovtoun, V.; Bui, H.; Izgarian, N.; Stafford, G.; Yost, R. A. *Int. J. Mass Spectrom.* **2007**, *260*, 166-176.

(47) Garrett, T. J.; Yost, R. A. *Anal. Chem.* **2006**, *78*, 2465-2469.

(48) Wiseman, J. M.; Ifa, D. R.; Song, Q. Y.; Cooks, R. G. *Angew. Chem. Int. Ed.* **2006**, *45*, 7188-7192.

(49) Li, Y.; Shrestha, B.; Vertes, A. *Anal. Chem.* **2007**, *79*, 523-532.

(50) Schwartz, S. A.; Reyzer, M. L.; Caprioli, R. M. *J. Mass Spectrom.* **2003**, *38*, 699-708.

(51) Claudio, G. A., P.; Simopoulos, E. T. *Effects of fatty acids and lipids in health and disease; Basel: New York* **1994**.

(52) Halket, J. M.; Zaikin, V. G. *Eur. J. Mass Spec.* **2004**, *10*, 1-19.

(53) Nagy, K.; Jakab, A.; Fekete, J.; Vekey, K. *Anal. Chem.* **2004**, *76*, 1935-1941.

(54) Yang, W. C.; Adamec, J.; Regnier, F. E. *Anal. Chem.* **2007**, *ASAP*.

(55) Robards, K.; Antolovich, M. *Analyst* **1997**, *122*, R11-R34.

(56) Seeram, N. P.; Lee, R.; Scheuller, H. S.; Heber, D. *Food Chem.* **2006**, *97*, 1-11.

(57) Törrönen, R. a. M., K. *Acta Hort. (ISHS)* **2002**, 797-803.

(58) Harvey, D. J.; Martin, R. L.; Jackson, K. A.; Sutton, C. W. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2997-3007.

(59) Asam, M. R.; Glish, G. L. *J. Am. Soc. Mass. Spectrom.* **1997**, *8*, 987-995.

(60) Gokmen, V.; Artik, N.; Acar, J.; Kahraman, N.; Poyrazoglu, E. *Eur. Food Res.Technol.* **2001**, *213*, 194-199.

(61) Lee, K. W.; Kim, Y. J.; Kim, D. O.; Lee, H. J.; Lee, C. Y. *J. Agric. Food. Chem.* **2003**, *51*, 6516-6520.

(62) Skupien, K.; Oszmianski, J. *Eur. Food Res.Technol.* **2004**, *219*, 66-70.

(63) Williner, M. R.; Pirovani, M. E.; Guemes, D. R. *Journal of the Science of Food and Agriculture* **2003**, *83*, 842-845.

(64) *See the USDA National Nutrient Database at* http://www.nal.usda.gov/fnic/foodcomp/search/ Accessed on Mar. **2007**.

(65) *See the USDA National Nutrient Database at* http://www.nal.usda.gov/fnic/foodcomp/Data/Flav/flav.pdf Accessed on Mar. **2007**.

(66) Ng, L. K.; Lafontaine, P.; Vanier, M. *J. Agric. Food. Chem.* **2004**, *52*, 7251-7257.

(67) Chen, M. L.; Song, F. R.; Guo, M. Q.; Liu, Z. Q.; Liu, S. Y. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 264-271.

Figure Captions

Figure 1. Mass spectrum of fatty acid standards (C18-C30 fatty acids) with GALDI in the negative-ion mode. Sample loading: 200 pmole of each.

Figure 2. Mass spectrum and detection limit of fatty acid standards (C18-C30 fatty acids) with GALDI in the negative-ion mode. Sample loading: 100 fmole of each.

Figure 3. Comparison of MALDI and GALDI spectra of flavonoid standards: quercetin (MW 302.24), kaempferol (MW 286.23), phloretin (MW 274.27) and apigenin (MW 270.24); Sample loading: 200 pmole of each. Apigenin, kaempferol and quercetin were observed as $[M + H]^+$ with DHB (a) and CHCA (b) in the positive ion mode. With GALDI in the positive-ion mode (c), apigenin, kaempferol and quercetin are detected as $[M + H]^+$ (*m/z* 271, *m/z* 286, and *m/z* 303, respectively), $[M+Na]^+$ (*m/z* 293, *m/z* 309, and *m/z* 325, respectively), and $[M + 2Na - H]^+$ (*m/z* 315, *m/z* 331, and *m/z* 347, respectively). Phloretin was not observed either. None of the four flavonoids can be detected with DHB (d) or CHCA (e) in the negative ion mode; However, all the four flavonoids are detected as $[M - H]^-$ ions with GALDI in the negative-ion mode (f). Fragment from phloretin (*m/z* 167) and fragments from quercetin (*m/z* 151 and *m/z* 179) are observed in the lower mass range as well.

Figure 4. Mass spectrum of oligosaccharide standards with DHB (a) and CHCA (b) in the positive-ion mode. Sample loading: 100 ng of each.

Figure 5. Mass spectrum of oligosaccharide standards with GALDI in the positive-ion mode. Sample loading: 100 ng of each. Peaks with •: water loss fragments; peaks with *: major ring-cleavage fragments.

Figure 6. Product ion spectra of *m/z* 191 from (a) quinic acid standard; (b) apple flesh; (c) citric acid standard; and (d) strawberry flesh. All spectra were collected in the negative-ion mode.

Figure 7. Representative GALDI-MS spectra from different parts of apple. (a) fatty acids composition on the outside of apple peel, negative-ion mode; (b) fresh apple juice,

negative-ion mode; (c) hexose and sucrose from fresh apple juice, positive-ion mode; and (d) apple core, negative-ion mode. See Table I for peak identification in (b) and (d).

Figure 8. Representative GALDI-MS spectra (negative-ion mode) from strawberry (a) flesh; and (b) seed.

Figure 9. Chemically selective images of major ionic species identified from apple endocarp region with GALDI in the negative-ion mode. (a) optical image taken with reversed color; (b) malic acid; (c) quinic acid; (d) linoleic acid; (e) sucrose; (f) phloretin; (g). epicatechin; (h) quercetin; and (i) phloridzin. All peak intensities were normalized by dividing by the total ion current (TIC) of each spectrum.

Figure 10. Chemically selective images of the major ionic species identified from strawberry with GALDI in the negative-ion mode. (a) optical image taken with reversed color. One seed was present as the darkest dot on the right-hand side; (b) ascorbic acid; (c) citric acid; (d) linoleic acid; (e) hexose; (f) sucrose; (g) apigenin; (h) kaempferol; and (i) *m/z* 301-301.5 ellagic acid + quercetin. All peak intensities were normalized by dividing by the total ion current of each spectrum.

Figure 11. Product ion spectra of *m/z* 301 from (a) ellagic acid standard; (b) strawberry flesh; (c) quercetin standard; (d) edge of strawberry; and (e) chemically selective image for product ion at *m/z* 179 (precursor ion *m/z* 301) of strawberry slice with GALDI. All spectra were collected in the negative-ion mode. The proposed fragment pathways are shown as inserts.[67]

Figure 1.

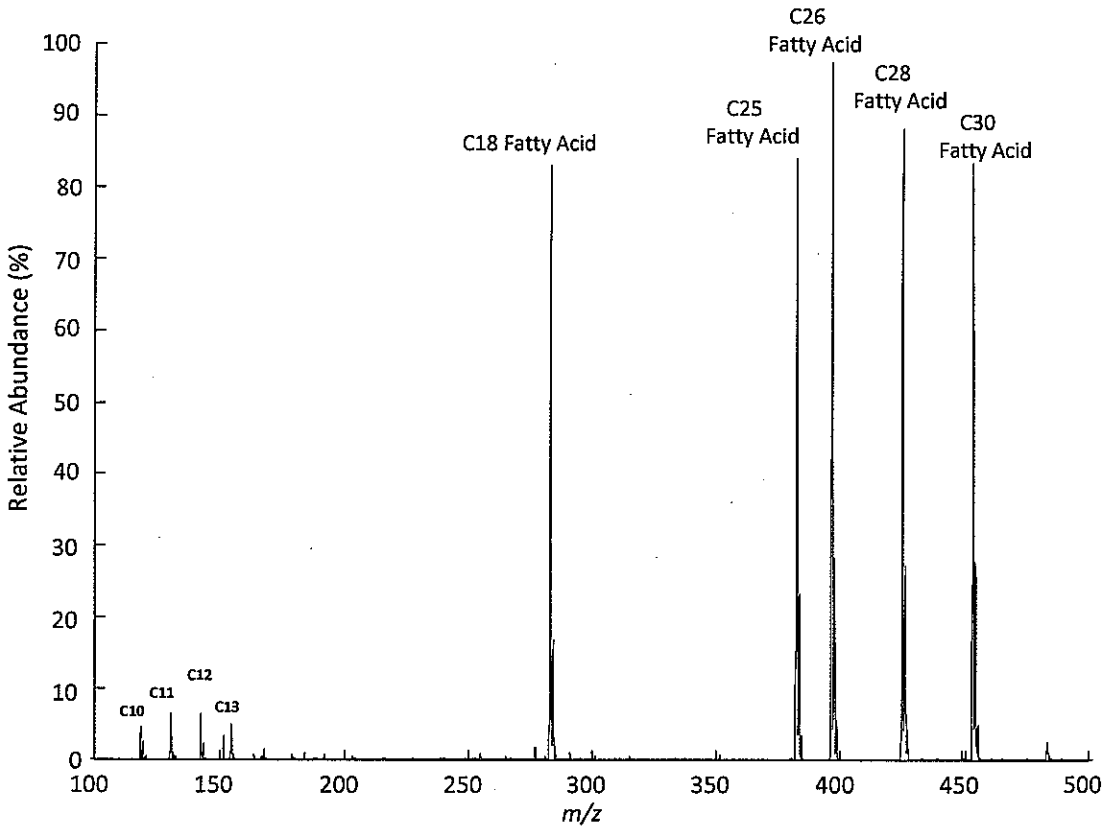

C26
Fatty Acid
C28
Fatty Acid
C30
Fatty Acid
C25
Fatty Acid
C18 Fatty Acid
C10
C11
C12
C13
Relative Abundance (%)
m/z
100
90
80
70
60
50
40
30
20
10
0
100
150
200
250
300
350
400
450
500

Figure 2.

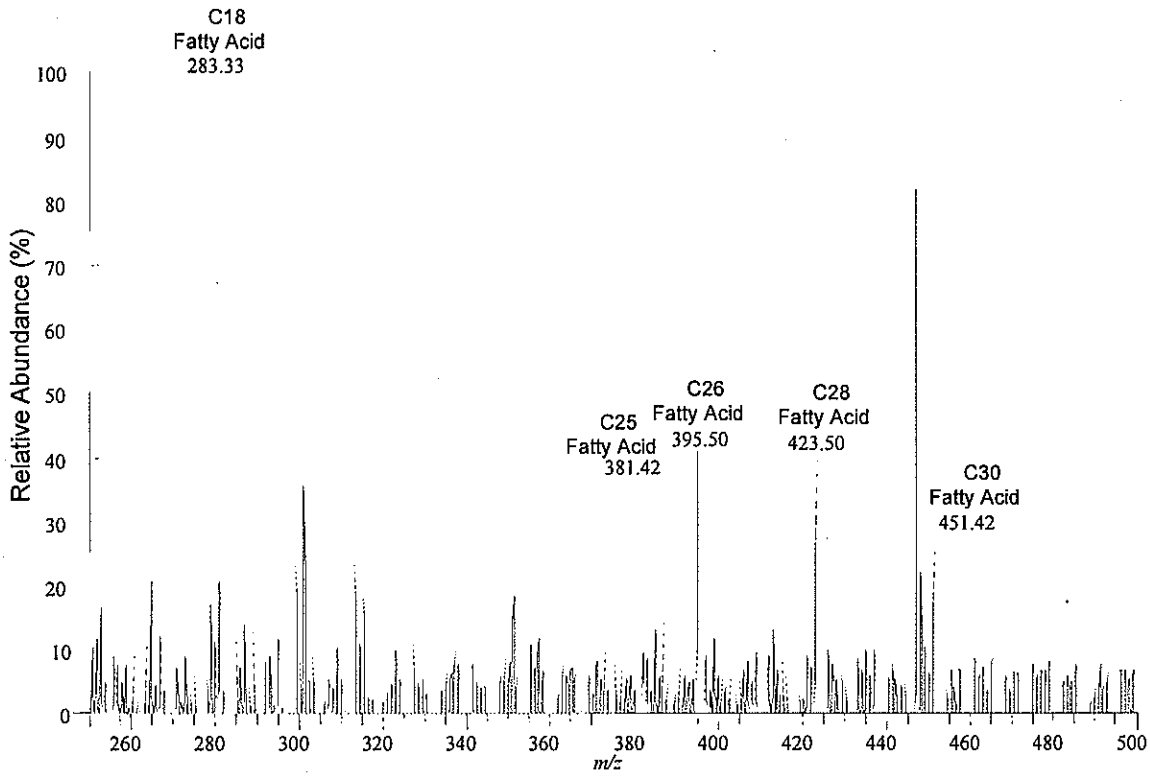

C18
Fatty Acid
283.33
C25
Fatty Acid
381.42
C26
Fatty Acid
395.50
C28
Fatty Acid
423.50
C30
Fatty Acid
451.42
Relative Abundance (%)
m/z
100
90
80
70
60
50
40
30
20
10
0
260
280
300
320
340
360
380
400
420
440
460
480
500

Figure 3.

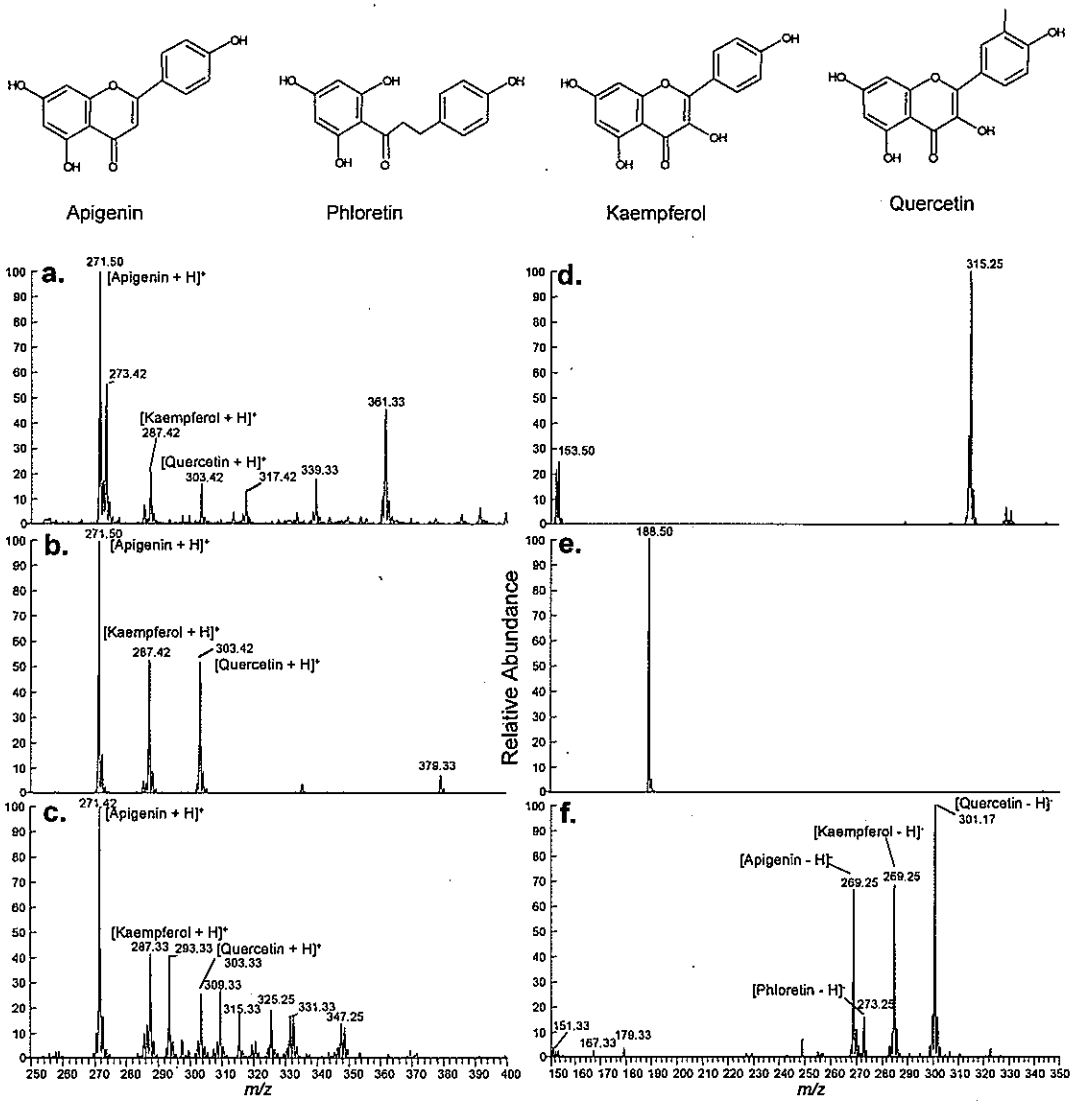

Apigenin
Phloretin
Kaempferol
Quercetin
a.
[Apigenin + H]+
271.50
273.42
[Kaempferol + H]+
287.42
[Quercetin + H]+
303.42
317.42
339.33
361.33
b.
[Apigenin + H]+
271.50
[Kaempferol + H]+
267.42
303.42
[Quercetin + H]+
379.33
c.
[Apigenin + H]+
271.42
[Kaempferol + H]+
287.33
293.33
[Quercetin + H]+
303.33
309.33
315.33
325.25
331.33
347.25
d.
315.25
153.50
e.
188.50
f.
[Quercetin - H]-
301.17
[Kaempferol - H]-
[Apigenin - H]-
269.25
269.25
[Phloretin - H]-
273.25
151.33
167.33
179.33
Relative Abundance
m/z
m/z

Figure 4.

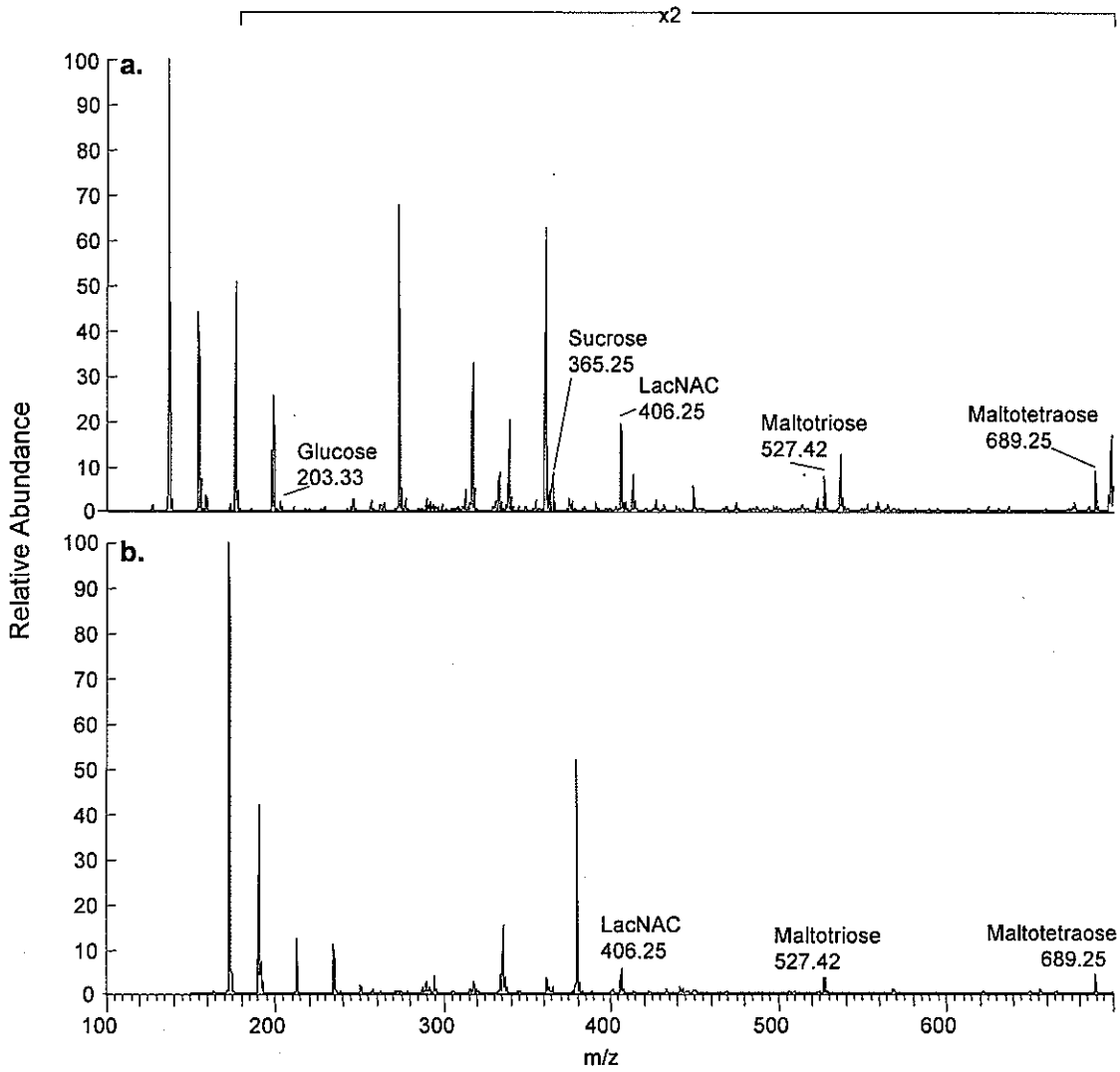

x2
a.
b.
Relative Abundance
Glucose
203.33
Sucrose
365.25
LacNAC
406.25
Maltotriose
527.42
Maltotetraose
689.25
LacNAC
406.25
Maltotriose
527.42
Maltotetraose
689.25
m/z

Figure 5.

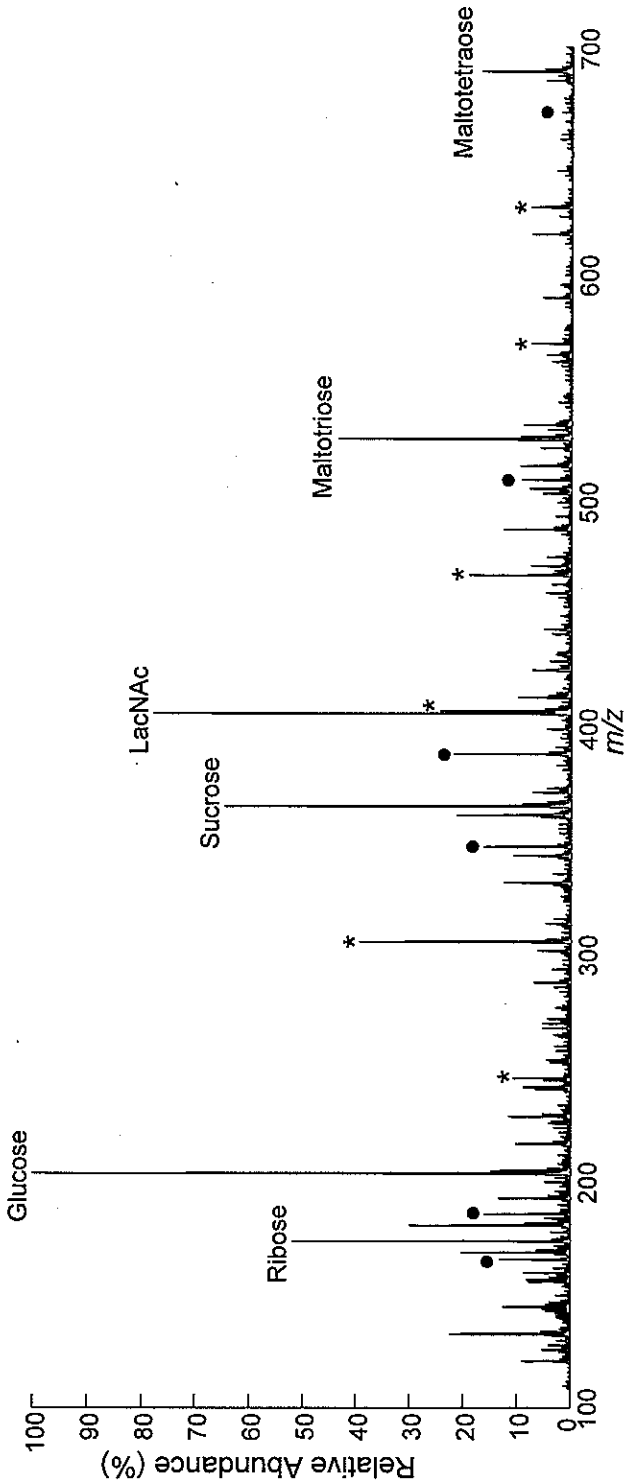

Relative Abundance (%)
m/z
Glucose
Ribose
Sucrose
LacNAc
Maltotriose
Maltotetraose
*

Figure 6.

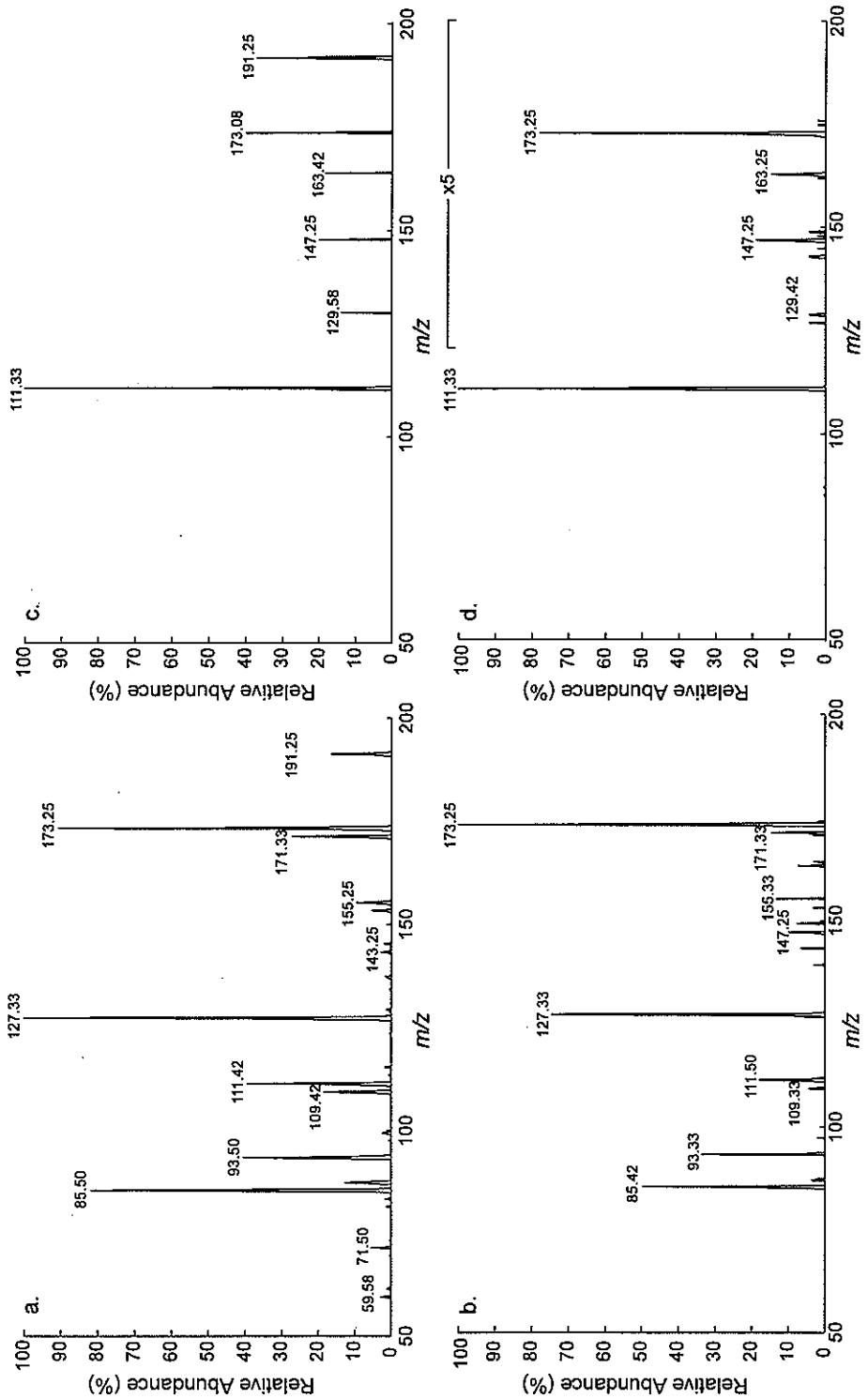

a.
b.
c.
d.
Relative Abundance (%)
m/z
x5
59.58
71.50
85.50
93.50
109.42
111.42
127.33
143.25
155.25
171.33
173.25
191.25
85.42
93.33
109.33
111.50
127.33
147.25
155.33
171.33
173.25
111.33
129.58
147.25
163.42
173.08
191.25
111.33
129.42
147.25
163.25
173.25

Figure 7.

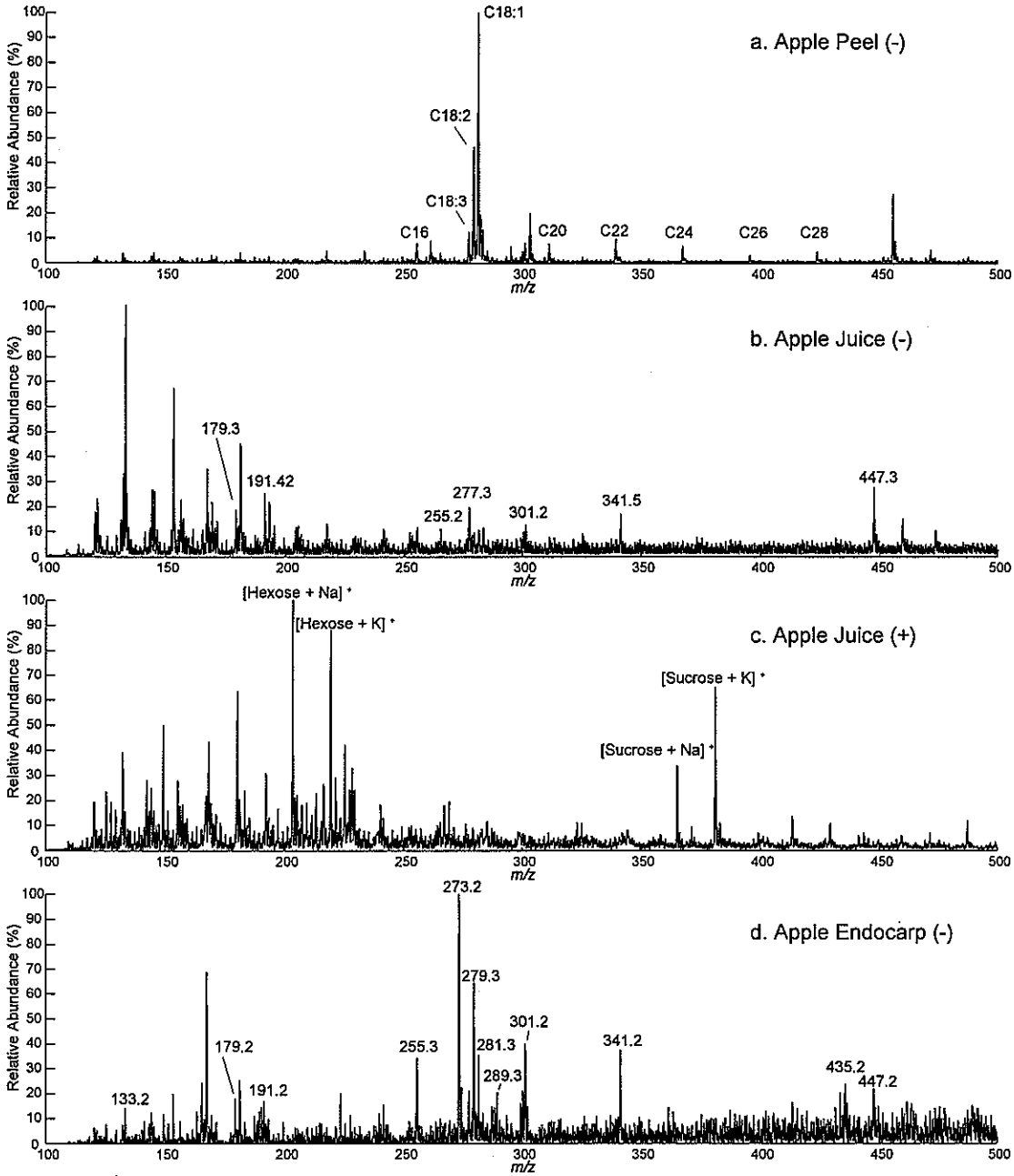

a. Apple Peel (-)
C18:1
C18:2
C18:3
C16
C20
C22
C24
C26
C28
b. Apple Juice (-)
179.3
191.42
277.3
255.2
301.2
341.5
447.3
c. Apple Juice (+)
[Hexose + Na] +
[Hexose + K] +
[Sucrose + K] +
[Sucrose + Na] +
d. Apple Endocarp (-)
273.2
279.3
301.2
179.2
255.3
281.3
341.2
289.3
435.2
447.2
191.2
133.2
Relative Abundance (%)
m/z

Figure 8.

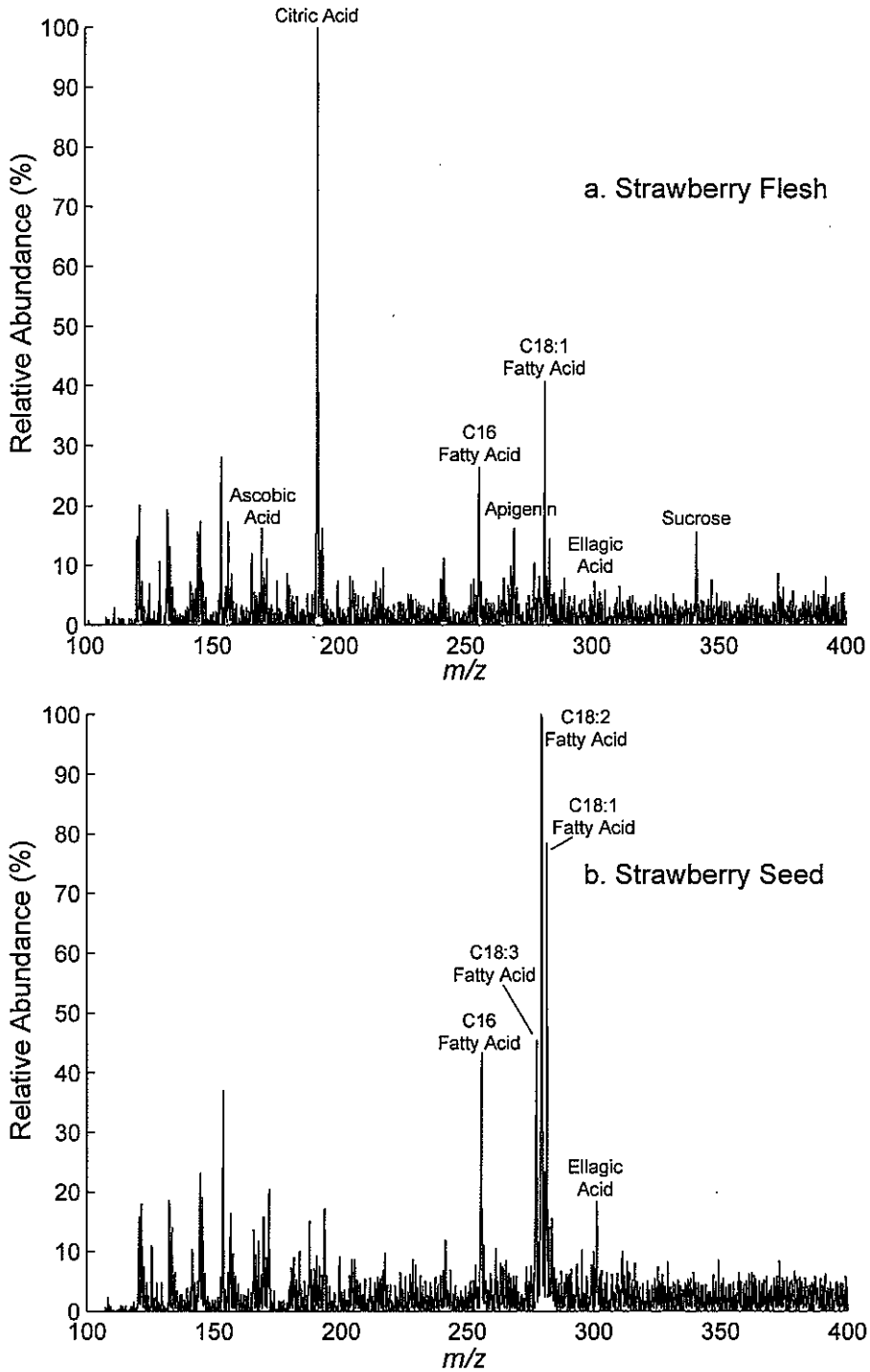

Citric Acid
a. Strawberry Flesh
C18:1
Fatty Acid
C16
Fatty Acid
Ascobic
Acid
Apigenin
Sucrose
Ellagic
Acid
Relative Abundance (%)
m/z
C18:2
Fatty Acid
C18:1
Fatty Acid
b. Strawberry Seed
C18:3
Fatty Acid
C16
Fatty Acid
Ellagic
Acid
Relative Abundance (%)
m/z

Figure 9.

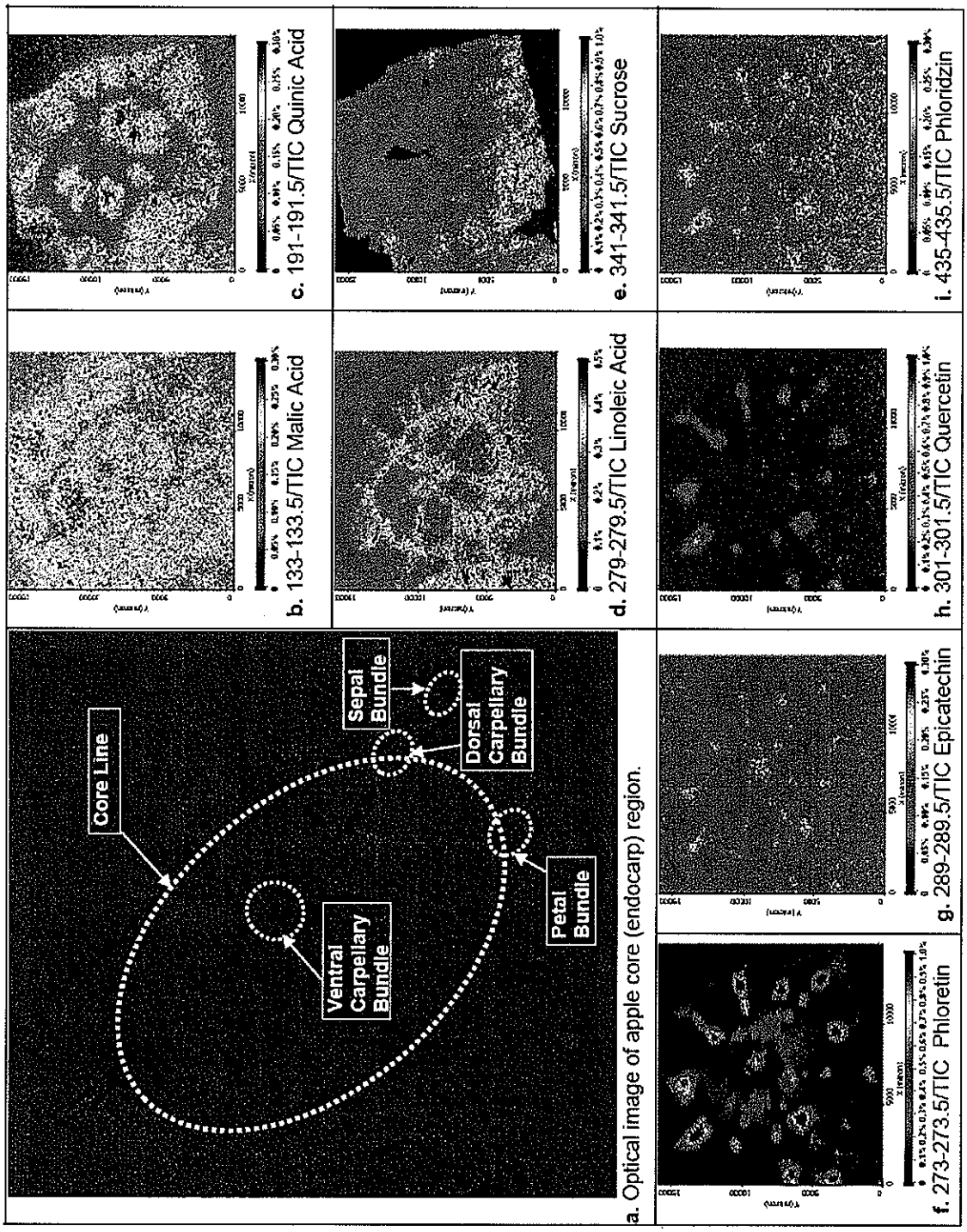

Core Line
Sepal Bundle
Dorsal Carpellary Bundle
Ventral Carpellary Bundle
Petal Bundle
a. Optical image of apple core (endocarp) region.
b. 133-133.5/TIC Malic Acid
c. 191-191.5/TIC Quinic Acid
d. 279-279.5/TIC Linoleic Acid
e. 341-341.5/TIC Sucrose
f. 273-273.5/TIC Phloretin
g. 289-289.5/TIC Epicatechin
h. 301-301.5/TIC Quercetin
i. 435-435.5/TIC Phloridzin

Figure 10.

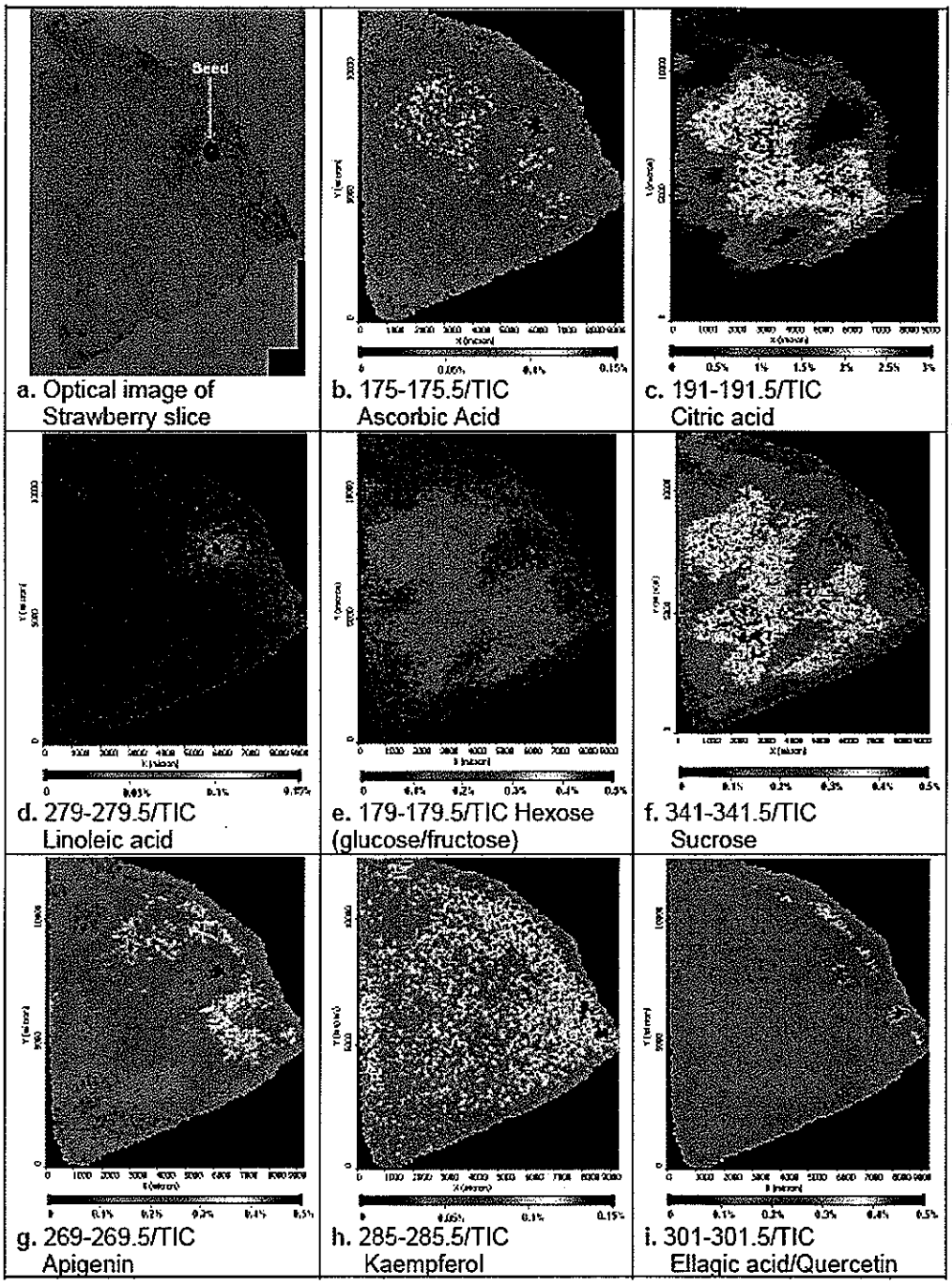
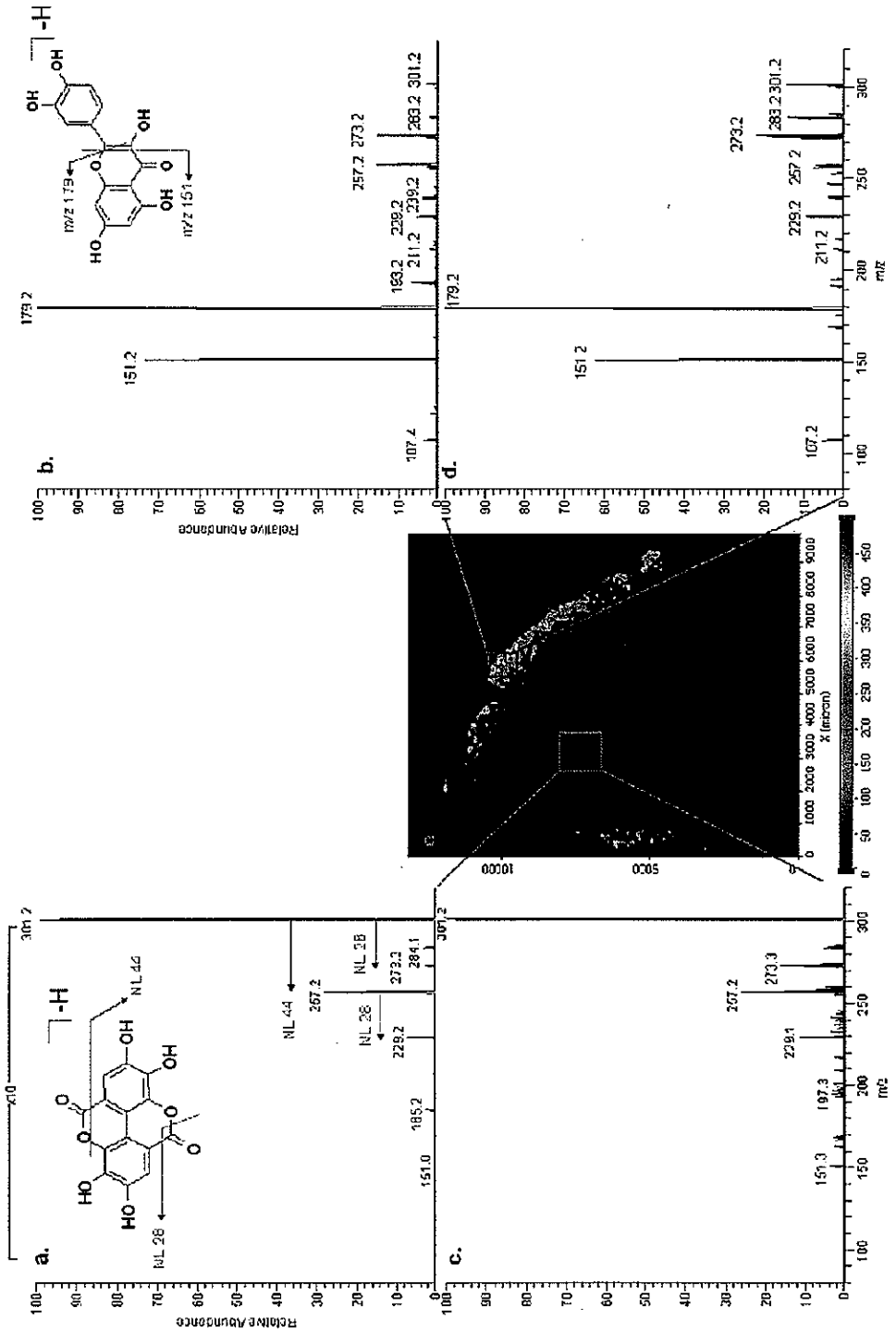

Seed
a. Optical image of Strawberry slice
b. 175-175.5/TIC Ascorbic Acid
c. 191-191.5/TIC Citric acid
d. 279-279.5/TIC Linoleic acid
e. 179-179.5/TIC Hexose (glucose/fructose)
f. 341-341.5/TIC Sucrose
g. 269-269.5/TIC Apigenin
h. 285-285.5/TIC Kaempferol
i. 301-301.5/TIC Ellagic acid/Quercetin

Figure 11.


a.
b.
c.
d.
Relative Abundance
m/z
NL 44
NL 28
NL 28
-H
m/z 179
m/z 151
X (micron)
Y (micron)

CHAPTER 6. LINKAGE POSITION AND MONOSACCHARIDE RESIDUE DISCRIMINATION OF DISACCHARIDES BY TANDEM MS AND LINEAR DISCRIMINANT ANALYSIS

A paper presented as a poster at 234th ACS meeting and prepared for submission to Rapid Communications in Mass Spectroscopy

Hui Zhang, Ning Fang, Steve M. Brokman, Nicola L. Pohl, and Edward S. Yeung

Abstract

The discrimination of isomeric disaccharides with different linkage types and different monosaccharide residues: glucose (Glc), galactose (Gal), and mannose (Man) at the non-reducing end was investigated with tandem mass spectrometry (MS) and linear discriminant analysis. Conventional matrix assisted laser desorption/ionization (MALDI) MS has strong interference peaks from matrix ions in low mass region (<500 Da). This greatly limits the application of MALDI MS for analysis of small molecules such as disaccharides. Acidic fullerene matrix gives very clean background in the low mass region thus it serves as a good candidate for small molecules analysis. Disaccharides with different linkage types give different tandem mass profiles from various cross-ring fragmentation pathways. Disaccharides with the same linkage type but three different kinds of monosaccharide residues bear the same fragmentation profiles. However, the relative ratios of the fragment ion intensities were distinctively different among these three kinds of disaccharide isomers. Statistical tools including linear discriminant analysis were used to analyze the tandem mass spectra. Disaccharide isomers with both different linkages and different monosaccharide residues were successfully classified.

Introduction

Carbohydrates are one of the most widely distributed molecules in nature and they serve crucial structural and functional roles in almost all living organisms.[1] To name a few: they are the energy storage and fuels for many biological pathways as well as metabolic intermediates; the structural framework for RNAs/DNAs are formed with small carbohydrates as ribose and deoxyribose while celluloses are the major constituents of plant cell walls; carbohydrate conjugates with proteins or lipids are important mediators of biological recognitions and cellular interactions. It is not surprising that ambitious scientists have moved on to sequence carbohydrates after the other two major biopolymers as proteins and DNAs were successfully sequenced.[2-5]

The very diversely biological activities of carbohydrates depend on their detailed structures, namely, the composition of monosaccharides, the positions of interresidue linkages and branches, and the anomeric configurations. A successful sequence of carbohydrates should cover all of the above aspects. Many traditional biochemistry methods for carbohydrate sequencing based on the use of different enzymes with various specificities that oligosaccharides can be cleaved at one type of linkage or position but not the others. Those methods require a lot of sample preparation and separation thus are very time consuming.[6] Nuclear magnetic resonance (NMR) spectroscopy was frequently used to determine the structure of oligosaccharides but pure individual component was needed.[7] Mass spectrometry (MS) provides higher sensitivity and requires simpler sample preparation[8] compared with conventional methods. MS coupled with soft ionization such as MALDI and electrospray ionization (ESI) have been widely used in the study of carbohydrates in the past decades.[9-13] Tandem MS is an indispensable tool for the structural analysis of carbohydrates as the fragments can provide detailed information about the structure. Collision induced dissociation (CID)[10, 14, 15] and post source decay (PSD) [16, 17] are the two major methods to break oligosaccharides into fragments. Wavelength-tunable infrared multiple-photon dissociation tandem MS also enabled the differentiation of linkage and anomeric configuration of disaccharides.[18] It is well known that neutral oligosaccharides are difficult to be ionized due to the lack of acidic or basic functional groups. Oligosaccharides have

stronger affinity to alkali metal, or alkaline earth metal ions than to protons, so it is common to include some additives such as lithium or sodium salts for detection of oligosaccharides.[15, 18] Ionization efficiency can also be improved by permethylation of the hydroxyl groups.[19, 20] This method also proves very effective in determining the branching feature of oligosaccharides by simply counting how many methyl groups were attached after fragmentation.[19, 20] In the past decades, tandem MS was successfully used to identify the derivatization type and position, to elucidate the branching sites, and to distinguish different linkage positions.[10, 14, 20] Most recently, the anomeric configuration (α or β) was distinguished for four pairs of disaccharide isomers.[21] Monosaccharide isomers include glucose, galactose and mannose have been differentiated using tandem MS,[22] as well as ion mobility MS.[23] However, the work of identifying the different monosaccharide residues as building blocks of oligosaccharides, which is a fundamental question for carbohydrate sequencing, is scare. To the best of our acknowledge, there was only one related report in which branched β-cyclodextrins (CD) with glucose, or galactose, or mannose residue showed different relative intensities of *y* type fragments[24] with PSD MALDI MS.[16] Thus this work was devoted to address this problem as differentiating the different monosaccharide residues of disaccharides by using MALDI-tandem MS and discrimination/classification analyses. We believe this study will eventually benefit the real carbohydrate sequencing.

One problem with conventional MALDI is that it is not suitable for the analysis of small molecules such as disaccharides (< 500 Da), because the analyte ions are strongly interfered with or suppressed by the matrix-related ions that are predominant at the low *m/z* range. Different approaches have been employed in MALDI MS to minimize the background in the low mass range. Reports include derivatization of the analyte molecules to a higher molecular weight[25] or using a matrix with higher molecular weight such as porphyrin (MW 974.6).[26] Extra sample preparation was then needed, thereby limiting the classes of analytes that can be detected. Different inorganic materials including different metal powders and metal oxide nanoparticles[27-29] have been tested as matrixes for surface-assisted laser desorption/ionization (SALDI). Those SALDI MS can provide a cleaner background than conventional MALDI MS as no interference peaks from the organic matrixes were present, but the application was very limited. Another matrix-free approach as laser

desorption/ionization on porous silicon (DIOS) was extensively studied since 1999.[30] Porous silicon surfaces were etched from crystalline silicon chips with hydrofluoric acid and functionalized as the laser desorption/ionization matrix as well as trapping agents for analyte molecules. Small molecules including pharmaceuticals, nucleic acids, carbohydrates, and steroids were successfully detected.[31] Recently, colloidal graphite has been reported as a successful matrix for detection of small molecules include oligosaccharides, fatty acids, phenolic compounds, and phospholipids.[32, 33] However, there were still some background peaks of this colloidal graphite matrix in the positive ion mode. In this work, acidic fullerene was used as a novel matrix for detection of disaccharides, as it provides very clean background in lower mass range (<500Da), as well as decent detection sensitivity for small oligosaccharides.

Discriminant and classification analyses are statistical techniques for description of group separation (discrimination) and prediction or allocation of observations to groups (classification). The basic idea is to establish group separation of original samples with regard to the discriminant functions, which are linear combinations of variables that best separate groups; and then to use the classification functions, which are derived from the discriminant functions, to predict the group membership of unknown samples.[34] These methods have been used widely in many applications from classification of petroleum pollutants based on infrared spectra[35] to differentiation of healthy tissues from cancer tissues based on the multiple metal elemental concentrations determined by X-ray emission.[36] In this work, discriminant and classification analyses were successfully used to discriminate disaccharide isomers with different linkage types and different monosaccharide residues.

Experimental Section

Chemicals: The disaccharides trehalose, sophorose, 2α-mannobiose, 3α-galactobiose, 3α-mannobiose, maltose, cellobiose, lactose, isomaltose, melibiose (as listed in Table 1) as well as monosaccharides ribose and glucose were purchased from Sigma-Aldrich (St. Louis, MO). Dihydroxybenzoic acid (DHB) from Bruker Daltonics (Billerica, MA) was used as standard MALDI matrix. Acidic fullerene ($C_{60}CHCOOH$) was obtained from Sigma-Aldrich

(St. Louis, MO). Pure water was obtained from a MilliQ water purification system (Billerica, MA). Organic solvents and all other chemicals were purchased from Fisher Scientific (Fairlawn, NJ).

Synthesis of the Propyl-linked Disaccharides: Nine propyl-linked disaccharides were synthesized as listed in Table 2. The syntheses of the 1,2-linked disaccharides DS198/DS204/DS223, and the 1,3-linked disaccharides DS90/DS101/DS95 began with the known glycosyl acceptors Allyl 3,4,6-tri-*O*-benzyl-β-D-glucopyranoside(1)[37] and Allyl-4,6-*O*-benzylidene-2-*O*-pivaloyl-α-D-glucopyranoside(2),[38] respectively. Trichloracetimidate donors: 2,3,4,6-Tetra-*O*-acetyl-α-D-glucopyranosyl trichloroacetimidate(3),[37] 2,3,4,6-Tetra-*O*-acetyl-α-D-galactopyranosyl trichloroacetimidate(4),[39] and 2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyl trichloroacetimidate(5)[37] were synthesized according to the literatures. Coupling of (1) or (2) with each of the three trichloracetimidate donors gave intermediate disaccharides with desired 1-2 or 1-3 linkages and all free hydroxyl groups protected. The final 1-2 or 1-3 linked propyl-disaccharides can be achieved after debenzylidenation, deacetylation, and catalytic hydrogenation steps.

The syntheses of the 1,4-linked disaccharides DS232 (Glc-Glc-propyl) and DS247 (Gal-Glc-propyl) started with the known trichloroacetimidates from cellobiose and lactose, respectively. The deprotected propyl glycosides DS232 and DS247 can be achieved after coupling the above intermediate disaccharides with n-propanol (to give the corresponding propyl-glycosides) and deacetylating. 1-4 linked disaccharide DS249 (Man-Glc-propyl) started with the known acceptor allyl 2,3,6-tri-*O*-benzyl- α-D-glucopyranoside which can be prepared from D-glucose in four steps.[40] Glycosylation of this acceptor with donor 5 provided disaccharide with protecting groups. The final 1-4 linked disaccharide DS249 can be achieved after debenzylidenation, deacetylation, and catalytic hydrogenation.

Mass Spectrometry: For all mass spectrometric analysis, an LTQ linear ion trap mass spectrometer equipped with vMALDI ion source (Thermo Electron, Mountain View, CA) was used. The N_2 laser (337 nm) is guided to the source by a fiber-optic cable and the laser power was set at 8-15 (arbitrary unit) for all MS experiments. 20 mg/mL DHB solution in 70% methanol and 30% water (containing a 0.1% trifluoroacetic acid) was prepared.

Acidic fullerene was dissolved in 50% iso-proponal and 50% water to give a saturated solution (about 0.1mg/ml).

All disaccharides were dissolved in water to a concentration of 100 ng/μl before loading onto MALDI plate. Saccharide standards mixture was prepared by dissolving ribose, glucose, and lactose in water/acetonitrile/trifluoroacetic acid (49.95/49.95/0.1) to a final concentration of 500 pmol/μl each.

Two-layer method was used for sample loading of all MALDI experiments: 2 μL of DHB, or acidic fullerene matrix solution was applied onto the stainless-steel sample plate and let to dry in air; then 1 μL of sample solution was applied on top of the dried matrix and let to dry in air before introduction into the mass spectrometer.

Tandem MS spectra of disaccharides were collected by applying CID on sodiated disaccharide ions as mass selected precursors using standard isolation and excitation procedures (activation q value of 0.25, and activation time of 30 ms). The collision energies used were 30, 40, and 50 (arbitrary units). Data acquisition was carried out with Xcalibur data system. For each of the disaccharide sample, nine spectra were taken with three replicas for each of the three collision energies. Each tandem MS spectrum was averaged with 100 laser shots.

Discriminant and Classification Analyses: All statistical analyses were carried out in SPSS (Version 15.0 for Windows, SPSS Inc., Chicago, IL). The relative abundance of the selected product ions of disaccharides were tested for significance regarding different linkage types or different monosaccharide residues using a one factorial multivariate analysis of variance (MANOVA). Then, to determine which peaks that discriminate between groups defined by linkage types or monosaccharide residues, a step-wise discriminant analysis was carried out to calculate the canonical discriminant functions. The first two canonical discrimination functions normally can represent over 98% of the separation of different disaccharides. The classification functions were then obtained to predict the group which unknown samples belong to.

Results and Discussion

Selection of Matrix: DHB vs. Acidic Fullerene

Conventional MALDI matrixes such as DHB and CHCA are usually small organic acids and they give strong background in low mass region (< 500Da) under laser irradiation, which strongly interferes with the detection of small molecules such as mono- and di-saccharides. Figure 1 shows the detection of the standards mixture of ribose, glucose and lactose with DHB and with acidic fullerene as matrix. With DHB, the three saccharide peaks were very low due to the strong interference and suppression by the strong matrix peaks nearby, as shown in the red circles in Fig.1a. For example, the $[\text{ribose} + \text{Na}]^+$ peak was very close to the matrix peak of $[\text{DHB} + \text{Na}]^+$, while the $[\text{glucose} + \text{Na}]^+$ peak was very close to the matrix peak of $[\text{DHB} + 2\text{Na-H}]^+$. With the acidic fullerene matrix, all the three small saccharides were well detected as sodium adducts. The background of acidic fullerene is quite clean in low mass range because the matrix molecule has larger molecular weight and the fullerene skeleton is hard to break into small fragments. The peak at *m/z* 305 was a fragment of lactose, which corresponds to a very common fragment of oligosaccharides with neutral loss of $C_2H_4O_2$. Another big advantage of acidic fullerene over DHB is that more homogeneous sample preparation can be achieved with the acidic fullerene matrix as no co-crystallization process is involved. There are no more 'sweet spots' of samples but the analytes can be detected consistently over the whole sample loading area. For the detection and homogeneity reasons, acidic fullerene was selected as the matrix for this study.

Linkage Type Discrimination of Disaccharides

Ten disaccharides as listed in Table 1 are isomers of $C_{12}H_{22}O_{11}$ with 5 different linkage types: 1-1, 1-2, 1-3, 1-4, and 1-6. Fragmentation of the sodium adducts of these 10 disaccharides were studied. Figure 2 shows five MS/MS spectra of five disaccharides (trehalose, 2α-mannobiose, 3α-galactobiose, maltose, and isomaltose, each one represents one linkage type.) The ion at *m/z* 365 is the precursor ion of sodiated disaccharide, and the ions at *m/z* 203 and 185 are the *y* and *b* ions from glycosidic cleavage.[24] Water loss was

observed as the ion at *m/z* 347 for all of the disaccharides except trehalose (1-1 linkage). Fragments with neutral losses of CH_2O, $C_2H_4O_2$, $C_3H_6O_3$, and $C_4H_8O_4$, as common ions of disaccharides with cross-ring cleavages as reported previously,[10, 14] were observed at *m/z* 335, 305, 275, and 245, respectively. The tandem MS spectra of other five disaccharides, sophorose, 3α-mannobiose, cellobiose, lactose, and melibiose, are not shown, as they all have similar pattern as the disaccharides with the same linkage type as shown in Fig. 2. The spectra clearly show that disaccharides with different linkage types give distinct fragmentation profiles, such as 1-1 linked disaccharide gives strong y ion (*m/z* 203) but very few fragments from cross-ring cleavages; while 1-2 linked disaccharide gives strong neutral loss of $C_4H_8O_4$ (*m/z* 245). This kind of observation has been used to identify the linkage type of oligosaccharides by manually confirming the presence or absence of these fragments.[10]

The tandem MS spectra of the 10 disaccharides were tested by MONOVA against 5 different linkage types, and significant differences were found among the different linkage types. Subsequently linear discriminant analysis (LDA) was performed to calculate discriminant functions for differentiation of those linkage isomers of disaccharides. Four canonical discriminant functions (CDF) were generated in which the first two explained 98.8% of the total variance with a very good correlation value (1.00). The first two CDFs were plotted in Figure 3 for all the MS/MS spectra of 10 disaccharides (nine spectra included for each of the disaccharide: three replicas for each of the three collision energies). Clearly, the ten disaccharides were grouped into five classes, with each class represents one kind of linkage type. Based on this, 100% of the original data input can be correctly classified into the right groups. The distance between the centroid of each class reflects the difference between different groups. For example, disaccharides with 1-4 linkage are close to those with 1-6 linkage, while disaccharides with both linkage types are very different from those with 1-2 linkage. This conclusion can be also appreciated from the similarities of fragment patterns of the disaccharides as shown in Fig. 2.

Residue Discrimination of Disaccharides with Three Different Linkage Types

Due to the unavailability of commercial disaccharide isomers with different monosaccharide residues, nine disaccharides (as listed in Table 2) were synthesized according to the procedures described in the experimental section. These nine disaccharides were designed to contain the same glucose residue (with propyl- group attached to the C1 position) on the reducing end, but three different monosaccharide residues (glucose or galactose or mannose) on the non-reducing end. On the other hand, the nine disaccharides were synthesized with three different linkage types: 1-2, 1-3, and 1-4 (see Table 2). MS/MS spectra of all nine synthesized disaccharides were collected as shown in Figure 4. As discussed earlier, disaccharides with the same linkage type gave a similar fragment pattern. It can be observed as the similar presence/absence of specific fragments for the three spectra in each column (A1-A3, B1-B3, and C1-C3) of Fig. 4. If we compare the three spectra in each column, one trend can be observed for all three different linkage types that the intensity ratio of each smaller fragment ion to the ion at *m/z* 347 was always the highest for disaccharides with galactose residue, and the lowest for disaccharides with mannose residue.

Relative fragment ion intensities of each MS/MS spectra were input for residue discriminant analysis. The first two canonical discriminant functions were plotted as Figure 5. Even though the statistics show that all the nine disaccharides could be correctly classified into three groups as shown in different colors, clearly the discrimination was not as good as in Fig. 3, especially for the data points in the highlighted rectangular in the middle region of Fig. 5. Such data points represented the MS/MS spectra inputs of disaccharides with 1-3 linkage only. Compared with the MS/MS spectra of 1-2 or 1-4 linked disaccharides, relative fragments intensities to the base peak were much lower for 1-3 linked disaccharides (as shown in Fig. 4 that the lower mass region was ten times magnified). It is believed that such low relative intensities cause high error when the CDFs were calculated for all the nine disaccharides.

Two-Step Residue Discrimination of Disaccharides

To improve the discrimination of disaccharide isomers with different residues, a two-step discrimination was performed. In the first step, the disaccharides with 1-3 linkage were separated from those with 1-2 and 1-4 linkages by discriminant analysis. This step was straightforward and the resulting plot (not shown) was similar to Fig. 3. In the second step, the disaccharides with each individual linkage type were tested for residue discrimination by discriminant analysis. Good discrimination of residues were achieved for disaccharides with all three linkage types, as shown in Fig. 6a for the disaccharides with 1-3 linkage and Fig. 6b for those with 1-2 and 1-4 linkages. Compared with direct residue discrimination as shown in Fig. 5, this two-step discrimination provides much better separation between three different groups, which can be judged by the distance between centroid of each group relative to the spread of data points within each group. One interesting feature about this two-step discrimination method is that the same CDFs can be applied for the disaccharides with both 1-2 linage and 1-4 linkage types without ruining the discrimination power, as shown in Fig. 6b.

The classification functions can be used to determine to which group each case most likely belongs to. There are as many classification functions as there are groups. Each function allows us to compute classification scores of the case for each groups, and the case would be assigned to the group for which it has the highest classification score. This analysis is useful to test the classification of unkown samples, unfortunately there are no more propyl-labeled disaccharides available. For this reason all the data used for discriminant analysis as in Fig. 6 are *cross* validated for classification, in which each case was classified by the functions derived from all cases but the one need to be validated. In another word, that case was regarded as an unknown but all other cases were used to calculate the classification functions. For all the cases, the cross validation was 100% correct.

Discrimination of Bi- and Tri-nary Mixtures of 1-3 Linked Propyl-disaccharides

In principle, the CDFs are calculated as the linear combination of the original variable input, in this case, the relative ion intensities of each fragment.[34] This linear characteristic allows LDA to distinguish the composition of a mixture if all individual components are available. To show this capability of LDA, three 1-3 linked propyl-disaccharides (DS 90, DS 101, and DS 95 as listed in Table 2) were selected and mixed to give binary or trinary mixtures. MS/MS spectra of those mixtures as well as those from three individual standards were collected and the relative intensities of each fragment ions were inputted for LDA. The discrimination of those mixtures from the standards was plotted as Figure 7. Clearly seven groups can be observed and each one represents one species (individual standard or mixture). If we draw a triangle with the group centroids of the three standards, all three binary mixtures were along the three edges, while the data points of trinary mixture were inside the triangle. The location of mixture data points reflects the prediction of the origins of the mixture samples regarding to the three individual standards. Based on the location in the CDF plot, the composition of the mixtures can be qualitatively determined.

Conclusions

With acidic fullerene as a matrix, small oligosaccharides were well detected. By analyzing tandem MS spectra with linear discriminant analysis, disaccharide isomers with five different linkage types and three different monosaccharide residues (on the non-reducing end) were successfully differentiated. Residues discrimination can be done for all the disaccharides with three linkage types simultaneously, but it was not as good as the linkage types discrimination, presumably due to the too low relative intensities of the fragment ions for the 1-2 linked disaccharides. In order to improve the discrimination of residues, disaccharides were first tested to differentiate the linkage type, and were later tested for the residue discrimination of disaccharides with individual linkage type. This two-step discrimination method improved the performance of residue differentiation effectively. Interestingly, residues of disaccharides with 1-2 linkage and 1-4 linkage can be differentiated with the same set of discriminant functions. The discrimination was not sensitive to the

collisional energies(CE) and tandem MS data from three different CEs (30-50, arbitrary units) were incorporated together for this study. The robustness of this method was also reflected by the fact that different residues of disaccharide isomers were successfully discriminated, no matter what anomeric configurations they are (as shown in Table 2).

Due to the limitation of the availability of disaccharide samples, the discrimination of monosaccharide residues were performed for nine synthesized disaccharides with 1-2, 1-3, and 1-4 linkages only. It is worth to mention that the nine synthesized disaccharides were attached with a propyl- group to facilitate the synthesis process. However this propyl group complicated the interpretation of tandem MS spectra of such disaccharides. For example, both neutral loss of water (H_2O, 18Da) and that of the propyl group (C_3H_6, 42Da) were observed, while it is also very common to lose $C_2H_4O_2$ (60Da) during fragmentation of oligosaccharides. It requires a very high resolution (R~10,000) to resolve the fragment ion with neutral loss of $C_2H_4O_2$ from that with losing both H_2O and the propyl group. Unfortunately, such high resolution is not always available for many of the mass instruments include the one used in this study. Nevertheless this current study proved the feasibility that linkage types and monosaccharide residues of disaccharide isomers could be differentiated by LDA of the tandem MS data, and this method can also be applied for analysis of the mixtures, as long as all the standard compounds are available. We strongly believe that this result will benefit the structure determination of oligosaccharides and eventually the real carbohydrate sequencing.

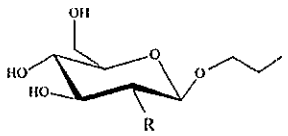
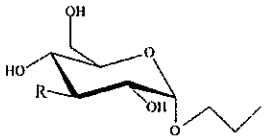
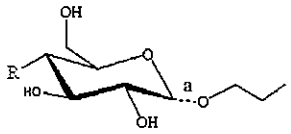
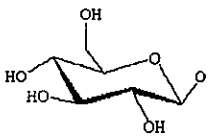
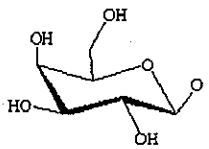
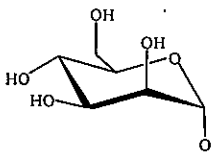
Acknowledgements

E.S.Y. thanks the Robert Allen Wright Endowment for Excellence for support. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. DE-AC02-07CH11358. This work was supported by the Director of Science, Office of Basic Energy Sciences, Divisions of Chemical Sciences and Biosciences.

Table 1. List of ten commercially available disaccharides purchased from Sigma and used for linkage type determination. The MS/MS spectra showed in Figure 2 were collected from disaccharides of the top row of each linkage type.

	1-1 linkage	1-2 linkage	1-3 linkage	1-4 linkage	1-6 linkage
$C_{12}H_{22}O_{11}$	Trehalose	2α-Mannobiose	3α-Galactobiose	Maltose	Isomaltose
		Sophorose	3α-Mannobiose	Cellobiose	Melibiose
				Lactose	

Table 2. List of nine synthesized disaccharides used for monosaccharide residues determination. The MS/MS spectra of each disaccharide were shown in Figure 4.

			
	1-2 linkage β-Glc	1-3 linkage α -Glc	1-4 linkage α / β-Glc[a]
R = β-Glc 	**DS 198**	**DS 90**	**DS 232**
R = β-Gal 	**DS 204**	**DS 101**	**DS 247**
R = α-Man 	**DS 223**	**DS 95**	**DS 249**

a. β for DS232 and DS247, α for DS249

References

(1) Berg, J. M.; Tymoczko, J. L.; Lubert, S. *Biochemistry* **2001**, Fifth Edition, International Edition.

(2) Harvey, D. J. *Mass Spectrom. Rev.* **1999**, *18*, 349-450.

(3) Harvey, D. J. *Mass Spectrom. Rev.* **2006**, *25*, 595-662.

(4) Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161-227.

(5) Ashline, D.; Singh, S.; Hanneman, A.; Reinhold, V. *Anal. Chem.* **2005**, *77*, 6250-6262.

(6) Edge, C. J.; Rademacher, T. W.; Wormald, M. R.; Parekh, R. B.; Butters, T. D.; Wing, D. R.; Dwek, R. A. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6338-6342.

(7) Townsend, R. R.; Hilliker, E.; Li, Y. T.; Laine, R. A.; Bell, W. R.; Lee, Y. C. *J. Biol. Chem.* **1982**, *257*, 9704-9710.

(8) Perez, S.; Mulloy, B. *Curr. Opin. Struct. Biol.* **2005**, *15*, 517-524.

(9) Asam, M. R.; Glish, G. L. *J. Am. Soc. Mass. Spectrom.* **1997**, *8*, 987-995.

(10) Konig, S.; Leary, J. A. *J. Am. Soc. Mass. Spectrom.* **1998**, *9*, 1125-1134.

(11) Zeng, C. H.; Biemann, K. *J. Mass Spectrom.* **1999**, *34*, 311-329.

(12) Mechref, Y.; Novotny, M. V.; Krishnan, C. *Anal. Chem.* **2003**, *75*, 4895-4903.

(13) Harvey, D. J.; Martin, R. L.; Jackson, K. A.; Sutton, C. W. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2997-3007.

(14) Hofmeister, G. E.; Zhou, Z.; Leary, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5964-5970.

(15) Lewandrowski, U.; Resemann, A.; Sickmann, A. *Anal. Chem.* **2005**, *77*, 3274-3283.

(16) Yamagaki, T.; Nakanishi, H. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1069-1074.

(17) Ren, S. F.; Zhang, L.; Cheng, Z. H.; Guo, Y. L. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 333-339.

(18) Polfer, N. C.; Valle, J. J.; Moore, D. T.; Oomens, J.; Eyler, J. R.; Bendiak, B. *Anal. Chem.* **2006**, *78*, 670-679.

(19) Weiskopf, A. S.; Vouros, P.; Harvey, D. J. *Anal. Chem.* **1998**, *70*, 4441-4447.

(20) Sheeley, D. M.; Reinhold, V. N. *Anal. Chem.* **1998**, *70*, 3053-3059.

(21) Simoes, J.; Domingues, P.; Reis, A.; Nunes, F. M.; Coimbra, M. A.; Domingues, M. R. M. *Anal. Chem.* **2007**, *79*, 5896-5905.

(22) Gaucher, S. P.; Leary, J. A. *Anal. Chem.* **1998**, *70*, 3009-3014.

(23) Leavell, M. D.; Gaucher, S. P.; Leary, J. A.; Taraszka, J. A.; Clemmer, D. E. *J. Am. Soc. Mass. Spectrom.* **2002**, *13*, 284-293.

(24) Domon, B.; Costello, C. E. *Glycoconjugate J.* **1988**, *5*, 397-409.

(25) Tholey, A.; Wittmann, C.; Kang, M. J.; Bungert, D.; Hollemeyer, K.; Heinzle, E. *J. Mass Spectrom.* **2002**, *37*, 963-973.

(26) Ayorinde, F. O.; Hambright, P.; Porter, T. N.; Keith, Q. L. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2474-2479.

(27) Kinumi, T.; Saisu, T.; Takayama, M.; Niwa, H. *J. Mass Spectrom.* **2000**, *35*, 417-422.

(28) McLean, J. A.; Stumpo, K. A.; Russell, D. H. *J. Am. Chem. Soc.* **2005**, *127*, 5304-5305.

(29) Sluszny, C.; Yeung, E. S.; Nikolau, B. J. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 107-115.

(30) Wei, J.; Buriak, J. M.; Siuzdak, G. *Nature* **1999**, *399*, 243-246.

(31) Pihlainen, K.; Grigoras, K.; Franssila, S.; Ketola, R.; Kotiaho, T.; Kostiainen, R. *J. Mass Spectrom.* **2005**, *40*, 539-545.

(32) Zhang, H.; Cha, S.; Yeung, E. S. *Anal. Chem.* **2007**, *79*, 6575-6584.

(33) Cha, S.; Yeung, E. S. *Anal. Chem.* **2007**, *79*, 2373-2385.

(34) Dunteman, G. H. *Introduction To Multivariate Analysis* **1984**, Thousand Oaks, CA: Sage Publications.

(35) Mattson, J. S.; Mattson, C. S.; Spencer, M. J.; Spencer, F. W. *Anal. Chem.* **1977**, *49*, 500-502.

(36) Banas, K.; Jasinski, A.; Banas, A. M.; Gajda, M.; Dyduch, G.; Pawlicki, B.; Kwiatek, W. M. *Anal. Chem.* **2007**, *79*, 6670-6674.

(37) Koketsu, M.; Kuwahara, M.; Sakurai, H.; Ishihara, H. *Synth. Commun.* **2004**, *34*, 239-245.

(38) Yang, Z. Y.; Lin, W. B.; Yu, B. *Carbohydr. Res.* **2000**, *329*, 879-884.

(39) Schmidt, R. R.; Kinzy, W. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 21-123.

(40) Kanakamma, P. P.; Mani, N. S.; Maitra, U.; Nair, V. *J. Chem. Soc., Perkin Trans. 1* **1995**, 2339-2344.

Figure Captions

Figure 1. Mass spectra of saccharide mixture: ribose, glucose, and lactose with DHB (a) and acidic fullerene matrix (b). The peak at *m/z* 305 in (b) is a fragment of lactose. Sample loading: 500 pmole of each. All three standards were detected as $[M+Na]^+$ ions.

Figure 2. Representative MS/MS spectra of disaccharides with 1,-1, 1-2, 1-3, 1-4, and 1-6 linkage types. Every spectrum was averaged for 100 laser shots.

Figure 3. Linear discriminant analysis of linkage types: all ten disaccharides were grouped into five classes that each class represents one kind of linkage type. Nine spectra were inputted for each disaccharide: three replicas for each of three different collision energies. Every spectrum was averaged for 100 laser shots.

Figure 4. Representative MS/MS spectra of nine propyl-disaccharides with three different monosaccharide residues (glucose, galactose, and mannose) and three different linkage types (1-2, 1-3, and 1-4). Every spectrum was averaged for 100 laser shots.

Figure 5. Discrimination of three monosaccharide residues for all the nine synthesized disaccharides with three kinds of linkage types. Three groups can be recognized but the classification is not good especially for those center data points, which correspond to the disaccharides with 1-3 linkage.

Figure 6. Discrimination of monosaccharide residue for disaccharides with 1-3 linkage only (a) and disaccharides with 1-2 and 1-4 linkages (b). The same discriminant function can be used for both disaccharides with 1-2 and 1-4 linkages.

Figure 7. Discrimination of binary and trinary mixture of three 1-3 linked disaccharides. The location of the mixture cluster reflects the origin of the original components.

Figure 1.

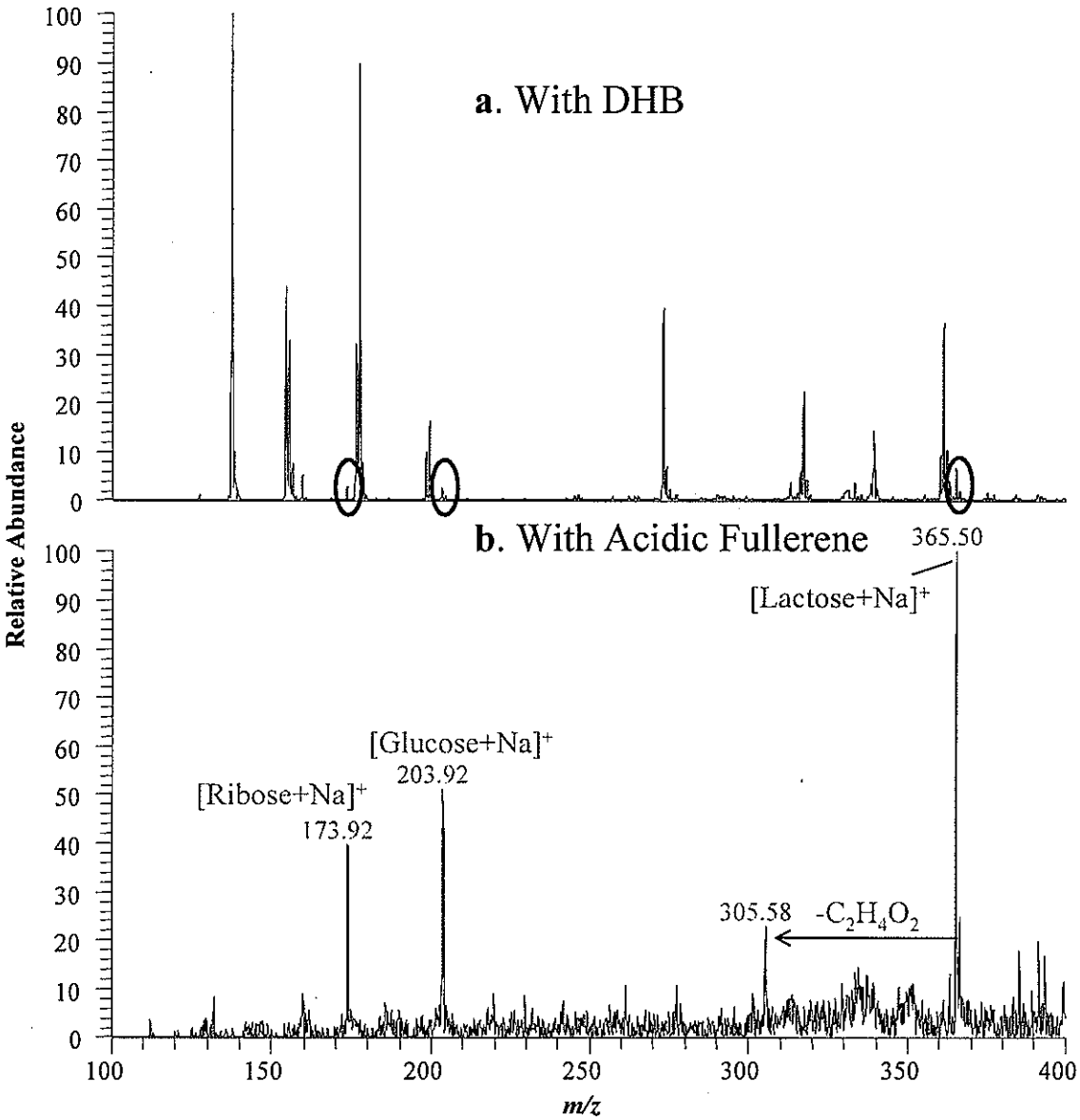

a. With DHB
b. With Acidic Fullerene
Relative Abundance
m/z
365.50
[Lactose+Na]+
[Glucose+Na]+
203.92
[Ribose+Na]+
173.92
305.58
-C2H4O2

Figure 2.

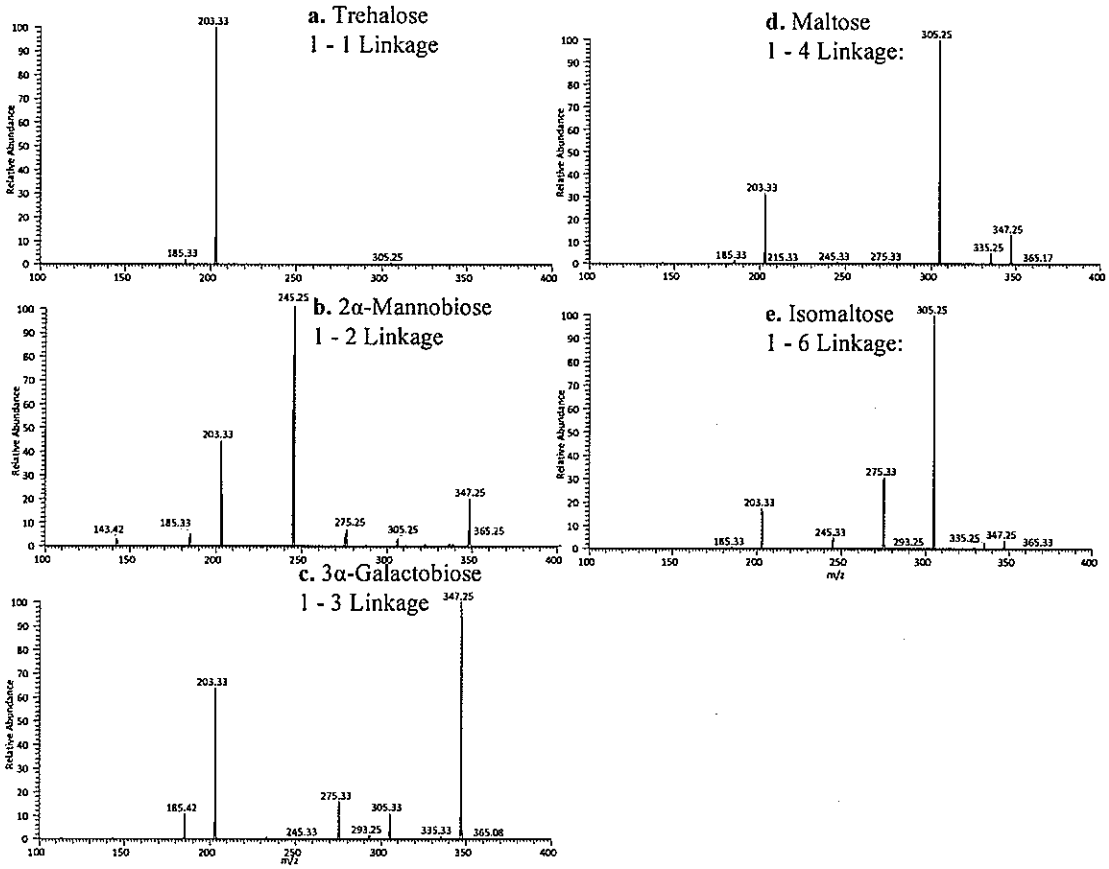

a. Trehalose
1 - 1 Linkage
b. 2α-Mannobiose
1 - 2 Linkage
c. 3α-Galactobiose
1 - 3 Linkage
d. Maltose
1 - 4 Linkage:
e. Isomaltose
1 - 6 Linkage:
Relative Abundance
m/z

Figure 3.

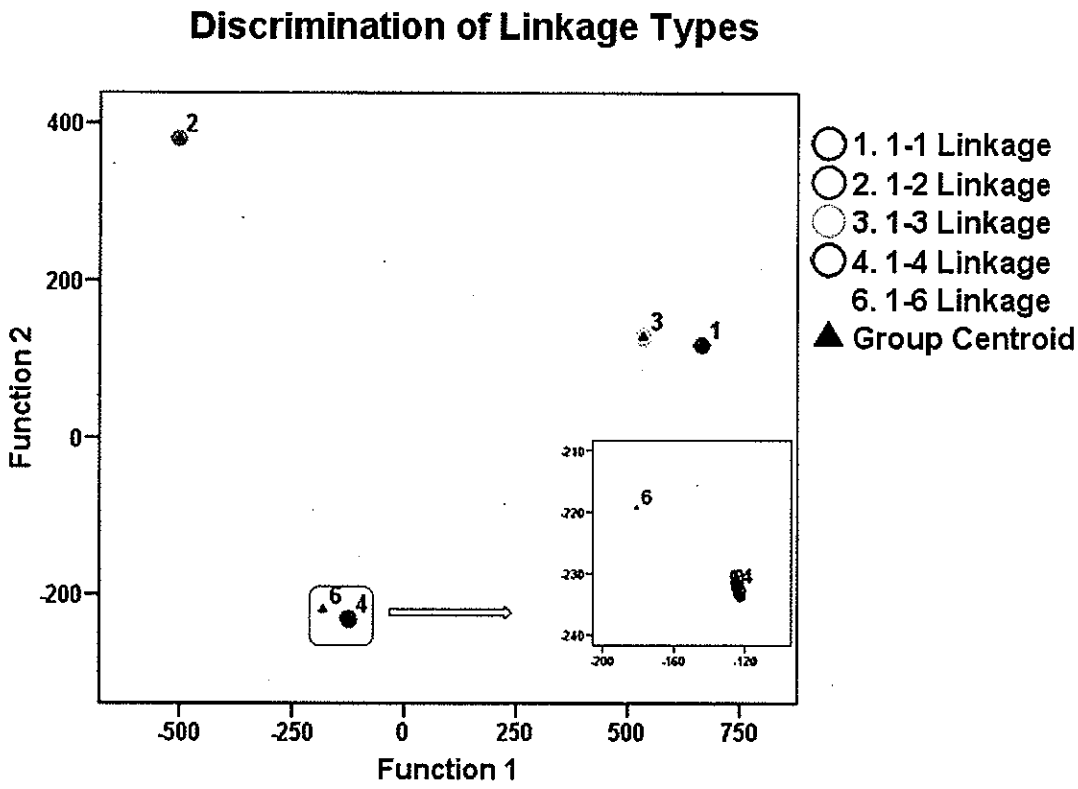

Discrimination of Linkage Types
Function 2
Function 1
400
200
0
-200
-500
-250
0
250
500
750
1. 1-1 Linkage
2. 1-2 Linkage
3. 1-3 Linkage
4. 1-4 Linkage
6. 1-6 Linkage
Group Centroid

Figure 4.

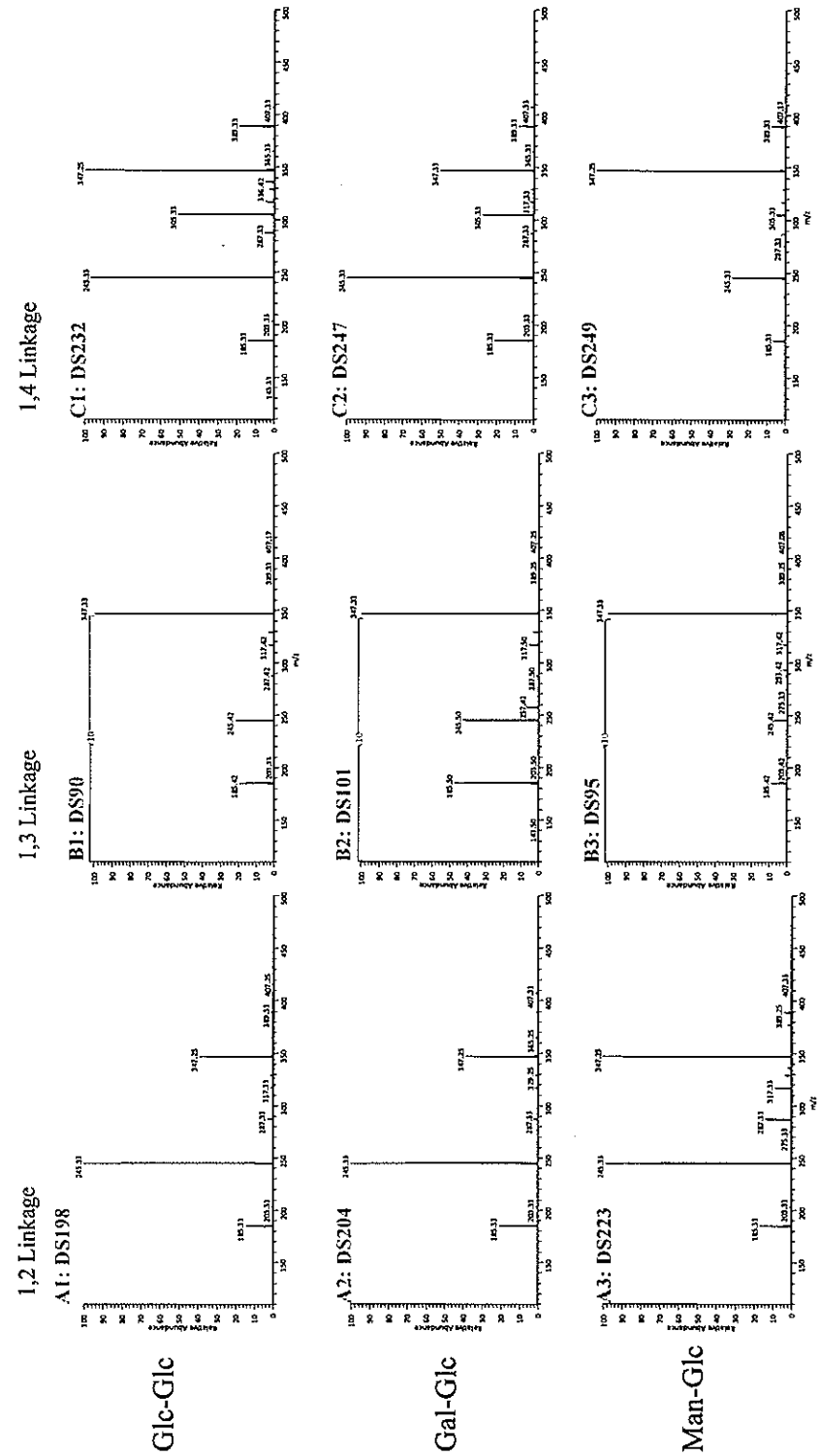

1,2 Linkage
1,3 Linkage
1,4 Linkage
A1: DS198
B1: DS90
C1: DS232
A2: DS204
B2: DS101
C2: DS247
A3: DS223
B3: DS95
C3: DS249
Glc-Glc
Gal-Glc
Man-Glc

Figure 5.

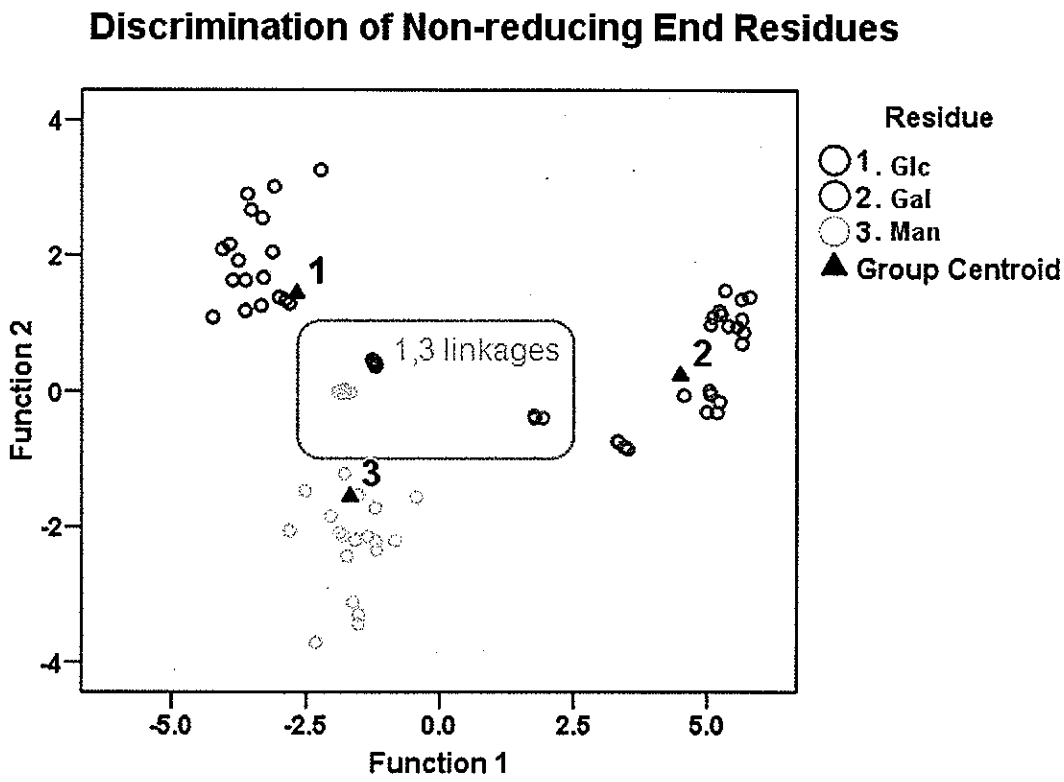

Discrimination of Non-reducing End Residues
Residue
1. Glc
2. Gal
3. Man
Group Centroid
1,3 linkages
1
2
3
Function 2
Function 1
4
2
0
-2
-4
-5.0
-2.5
0.0
2.5
5.0

Figure 6.

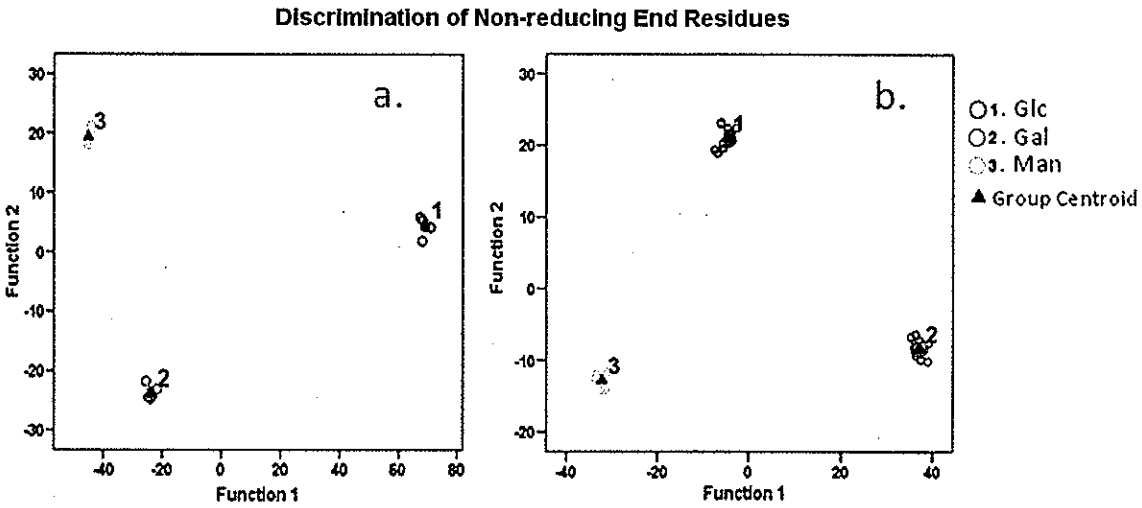
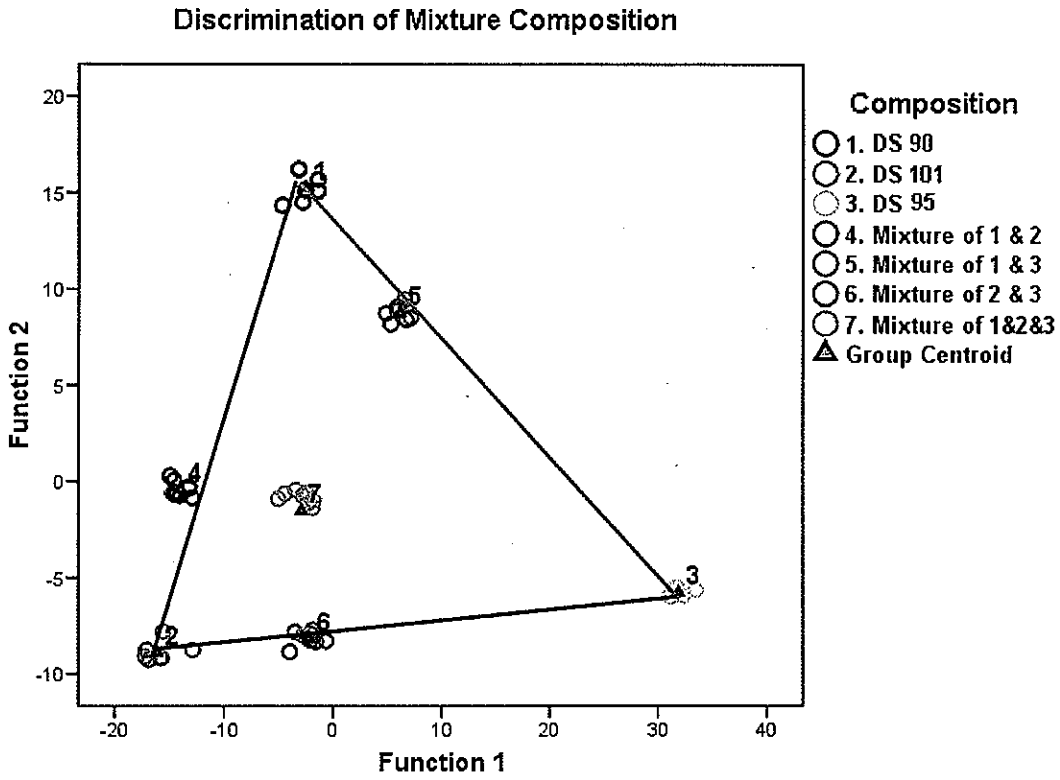

Discrimination of Non-reducing End Residues
a.
b.
Function 2
Function 1
1. Glc
2. Gal
3. Man
Group Centroid

Figure 7.


Discrimination of Mixture Composition
Function 2
Function 1
20
15
10
5
0
-5
-10
-20
-10
0
10
20
30
40
Composition
1. DS 90
2. DS 101
3. DS 95
4. Mixture of 1 & 2
5. Mixture of 1 & 3
6. Mixture of 2 & 3
7. Mixture of 1&2&3
Group Centroid

CHAPTER 7. GENERAL CONCLUSIONS

Analytical chemistry is the science of chemical measurements which include detection, separation, identification and quantification. It is our job as analytical chemists to develop new methods as well as to keep improving the existing analytical chemistry techniques to meet the increasing demands from all kinds of chemical and biological researches. This is well reflected in the projects discussed in this dissertation.

In this dissertation, the detection sensitivity for proteins separated by SDS-PAGE was improved hundreds of times to pictogram level, by using side entry laser induced native fluorescence detection. Novel biomarkers for prostate and breast cancers were separated and identified using CE and other spectroscopic methods. Adsorption properties of protein on fused silica surface was studied and cross validated by computational simulation, CE, and single molecule spectroscopy. Carbon based matrixes include colloidal graphite and fullerene derivatives were developed for detection, profiling and imaging of small molecules with MALDI MS. Identification of disaccharide isomers with different non-reducing end monosaccharide residues were demonstrated for the first time with tandem MS and discriminant analysis.

We hope that all these discoveries and developments in this dissertation will find a large number of applications in analytical and bioanalytical studies.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude and appreciation to my major advisor, Professor Edward S. Yeung, who generously guided, supported, and encouraged me through all these years in Iowa State. Besides a very inspiring and patient mentor, he is such a devoted scientist and decent person who I can always look up to. I was so fortunate to have such a wonderful supervisor as Dr. Yeung.

I would also like to acknowledge the support of Dr. Houk, Dr. Lin, Dr. Mallapragada, Dr. Nikolau and Dr. Badman for kindly serving on my committee and for all the guidance.

I would like to express my appreciation to Dr. Markushin, Dr. Jankowiak, Dr. Phol and her graduate student, Steve Brokman for their cooperation and help in this work.

My colleagues as the Yeung group members supported my reach and I thank them all for their help and friendship: Dr. Frank Li, Dr. Dragon Isailovic, Dr. Jiyoung Lee, Dr. Hungwing Li, Guoxin Lu, Slavica Isailovic, Dr. Bob Hsin, Becky Li, Melissa Zhang, Wenjun Xie, Wei Sun and Dr. Wei Wei. Specially thanks to Sangwon Cha, Dr. Ning Fang, Dr. Wenwan Zhong, Aoshuang Xu, and Dr. Chanan Sluszny, for their direct assistance to my own research. I really enjoy the time working with you all.

Thank you Mom and Dad for everything you have done for me. I could never acknowledge all my families enough for their endless love and support. Your short visit to Ames was definitely one of the brightest and happiest times for me among these years.

Especially, I would like to give my special thanks to my girlfriend Becky, for her tremendous support during my graduate study. You make my life in Ames more colorful and enjoyable.

This work was performed at Ames Laboratory under Contract No. DE-AC02-07CH11358 with the U.S. Department of Energy. The United States government has assigned the DOE Report number IS-T2519 to this thesis.