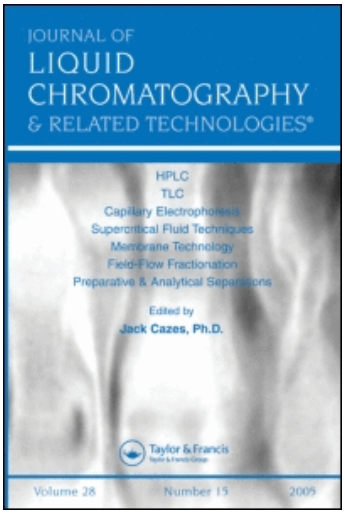


## Journal of Liquid Chromatography & Related Technologies

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

# Profiling of Urinary Proteins by Nano-High Performance Liquid Chromatography/Tandem Mass Spectrometry

Soon Sun Hong[a]; Sung Won Kwon[b]
[a] College of Pharmacy, Seoul National University, Seoul, Korea [b] Protein Chemistry Technology Center, UT Southwestern Medical Center, Dallas, Texas, USA

**To cite this Article** Hong, Soon Sun and Kwon, Sung Won(2005) 'Profiling of Urinary Proteins by Nano-High Performance Liquid Chromatography/Tandem Mass Spectrometry', Journal of Liquid Chromatography & Related Technologies, 28: 6, 805 — 822
**To link to this Article: DOI:** 10.1081/JLC-200051465
**URL:** http://dx.doi.org/10.1081/JLC-200051465

*Journal of Liquid Chromatography & Related Technologies*®, 28: 805–822, 2005

ISSN 1082-6076 print/1520-572X online
DOI: 10.1081/JLC-200051465

# Profiling of Urinary Proteins by Nano-High Performance Liquid Chromatography/ Tandem Mass Spectrometry

**Soon Sun Hong**
College of Pharmacy, Seoul National University, Seoul, Korea

**Sung Won Kwon**
Protein Chemistry Technology Center, UT Southwestern Medical Center, Dallas, Texas, USA



**Abstract:** There has been an increasing interest in proteomics which lies beyond the genomics era. Urine is one of the popular tools in proteomic field and has been the common focus to find proteins of interest, such as biomarkers, because urine is easy to collect and represents the body's status directly as does blood. To enhance our understanding of the urinary proteome, the water soluble urinary proteins were prepared in centrifuge filter tubes; one-dimensional gel electrophoresis (1DE) coupled nano-high performance liquid chromatography/tandem mass spectrometry (nano-LC/MS/MS) was used to profile the urinary proteins. From the male human urine profiled in this study, a total of 589 peptides, corresponding to 195 unique proteins, were identified by nano-LC/MS/MS. Each identified peptide's tandem mass spectrum was reconfirmed by manual verification. Some peptides originated from oncogene family proteins, which may be potential biomarker candidates related to certain disease states. The protein profiling method using 1DE coupled nano-LC/MS/MS showed sensitive, accurate, and unbiased protein identification. This technique will provide a powerful application tool for disease diagnosis and monitoring in the future clinic.

**Keywords:** Urinary proteins, Nano-HPLC, LC/MS/MS, Protein identification

Address correspondence to Sung Won Kwon, Protein Chemistry Technology Center, The University of Texas Southwestern Medical Center, Dallas, 5323 Harry Hines Blvd., Dallas, Texas 75390, USA; E-mail. sungwon.kwon@utsouthwestern.edu

## INTRODUCTION

Since the last decade, there has been a growing interest in the proteome of human urine that can be used for diagnosis and clinical monitoring.[1–20] Other components in human urine, such as xenobiotics, minerals, hormones, and ions also have important physiological functions, and have been studied for clinical kidney function.[21–23] They have been specifically examined, especially, to study metabolic diseases and toxicology using modern technologies.

The urine proteome is of great potential in clinical kidney research because relatively high amounts of proteins from the kidney can be acquired directly from urine without any barrier. In practice, most of the studies have been focused on proteinuria, using electrophoresis or diagnostic imaging, due to their easy detection.[2] However, it is not easy to investigate the detailed protein components from normal human subjects because only trace amounts of proteins are formed in urine. Furthermore, urinary proteins are not quantitatively enough to represent the body's state because of their filtration from blood before being transported into urine. The correlations between urinary proteins and related diseases are also difficult, due to some complicating factors such as race, sex, age, heredity, overlapped diseases, food, climate, etc. In spite of these pitfalls, the urinary proteome has infinite potential for diagnosis and disease monitoring because of its non-invasive, simple, and stress-free sampling from patients, while tumor and tissue biopsies are not obtained as easily.



Nowadays, various applications of the urinary proteome have been employed. Pregnancy test kits detecting human chorionic gonadotropin (HCG) have become a universal tool for pregnancy diagnosis. Besides, urine has been widely used to detect protein markers for the diagnosis of Bence Jones proteinuria, cancers of the urogenital tract, nephrolithiasis, cadmium toxicity, rheumatoid arthritis, duchenne muscular dystrophy, pregnancy-induced hypertension, and other renal diseases.[6]

One of the features of the last few years is that many urinary proteins have been reported using mass spectrometry techniques and its related techniques. Spahr et al. identified 124 gene products (proteins and translations of expressed sequence tags) in normal male urine by 1DLC/MS/MS (Q-TOF).[14] Pang et al. identified 90 proteins in healthy and diseased urine by 2DLC/MS/MS.[15] Thongboonkerd et al. identified 67 protein forms of unique proteins in normal urine by 2DE/MALDI/TOF.[18] Moon et al. identified 13 proteins by LC/IM/CID/TOF.[19] But, these findings reported fewer numbers of identified proteins. However, clinical application requires exhaustive protein profiling, which can be achieved by acquiring a more sensitive and simple technique. Alternatively, the variations of urinary proteins have been monitored by ELISA.[4]

Based on the need for identifying an increased number of proteins, the profiling of human urinary proteins has been carried out using simple

1DE/nano-1DLC/MS/MS. One hundred ninety-five unique proteins were identified from male human urine, which offers an alternative for the present protein profiling methods. It is also predicted that profiling of urinary proteins by nano-LC/MS/MS not only improves our knowledge of renal physiology, but also finds potential biomarker candidates associated with specific disease stages.

## EXPERIMENTAL

### Materials

The following reagents were purchased from commercial companies and used without further treatment. HPLC grade-acetonitrile, water, and methanol were from EM Science (Gibbstown, NJ, USA); trifluoroacetic acid (TFA) was from Fluka (Buchs, Switzerland); acetic acid was from Aldrich (St. Louis, MO, USA); ammonium bicarbonate, bromophenol blue, sodium dodecyl sulfate (SDS), monobasic sodium phosphate, and dibasic sodium phosphate were from Fisher (Fair Lawn, NJ, USA); ammonium persulfate, acrylamide, bis acrylamide, Tris-HCl, Tris base, glycine, glycerol, 2-mercaptoethanol, and bovine serum albumin (BSA) were from Sigma (St. Louis, MO, USA); TEMED was from Invitrogen (Carlsbad, CA, USA); protein standards and Bradford protein assay kit were from Bio-Rad (Hercules, CA, USA); centrifugal filter was from Millipore (Billerica, MA, USA); trypsin was from Promega (Madison, WI, USA).



### Protein Separation and Digestion

The male human urine sample was used in this study. The urine sample was centrifuged at 10,000 g for 30 min at 4°C and the supernatant was filtered with a 0.22 μm pore size cellulose acetate membrane filter (Corning Inc., Corning, NY, USA) to remove cell debris and insoluble compounds. The soluble urinary proteins were recovered by centrifugation, using a 5 kDa regenerated cellulose membrane centrifugal filter tube. Protein concentration was determined as 526 ng/μL by simple colorimetric assay, based on a Bradford dye-binding procedure.[24] Twenty seven μg was mixed in 50 μL 1 X SDS sample buffer with a vortex and boiled for 5 min. After attaining room temperature, followed by centrifugation at 10,000 g for 1 min, the supernatant was loaded into a home-made 12% (w/v) polyacrylamide gel (80 mm × 70 mm × 1.5 mm) with protein standards. The gel was stained by colloidal Coomassie Blue. Proteins bands were excised from the gel slab. Each protein band was cut into one cubic millimeter pieces, and destained with 25 mM $NH_4HCO_3$ in 50% methanol/50% water (v/v) two times, for

10 min. They were washed with 10% acetic acid/50% methanol/40 % water (v/v/v) two times for an hour, and swollen in water twice for 20 min. The gel pieces were dehydrated with acetonitrile and dried in the SpeedVac (Thermo Savant, Holbrook, NY, USA). The gel pieces were, again, rehydrated with modified porcine trypsin at the concentration of 10 ng/μL in 100 mM $NH_4HCO_3$ and then subjected to trypsin proteolytic digestion at 37°C overnight. Tryptic peptides were sequentially extracted with 50% acetonitrile/45% water/5% TFA, (v/v/v) and 75% acetonitrile/24.9% water/0.1% TFA (v/v/v) solutions. The peptide extracts were combined and dried in the SpeedVac. The peptide samples were cleaned with ZipTip $C_{18}$ (Millipore, Billerica, MA, USA) prior to nano-LC/MS/MS analysis.

## Nano-HPLC/Mass Spectrometry for Protein Identification



Nano-LC/MS/MS analysis was performed with an LCQ DECA XP ion-trap mass spectrometer (ThermoFinnigan, SanJose, CA, USA) equipped with a nano electrospray ionization (ESI) source. The electrospray source was coupled online with an Agilent 1100 series capillary HPLC system (Agilent, Palo Alto, CA, USA). Two μL of the peptide solution in buffer A (2% acetonitrile/97.9% water/0.1% acetic acid (v/v/v)) was manually loaded into a capillary HPLC column (70 mm length X 75 μm ID, 5 μm particle size, 300 Å pore diameter), packed in-house with Luna C18 resin (Phenomenex, St. Torrance, CA, USA). The peptides were eluted from the column with a gradient of 15% to 80% buffer B (90% acetonitrile/9.9% water/0.1% acetic acid (v/v/v)) in buffer A for 30 min. The eluted peptides were electrosprayed directly into the LCQ mass spectrometer. The MS/MS spectra were acquired in a data-dependent mode that determined the masses of the parent ions and fragments of three strongest ions. All MS/MS spectra were searched against the NCBI non-redundant human protein sequence database, using the Knexus program (Genomic Solutions, Ann Arbor, MI, USA) for protein identification. The redundant proteins were removed and all the identified spectra were manually analyzed to confirm the true protein identification.

## RESULTS

SDS-PAGE analysis of human urine resolved the proteins, as shown in Figure 1. According to the densities of proteins in the gel, 32 gel slices were excised. Dark bands were excised into small pieces of 1–3 mm size and light bands were excised into larger pieces of $<3$ mm size. Excised gel slices were subjected to in-gel digestion and the extracted peptides were analyzed by nano-LC/MS/MS for protein identification.

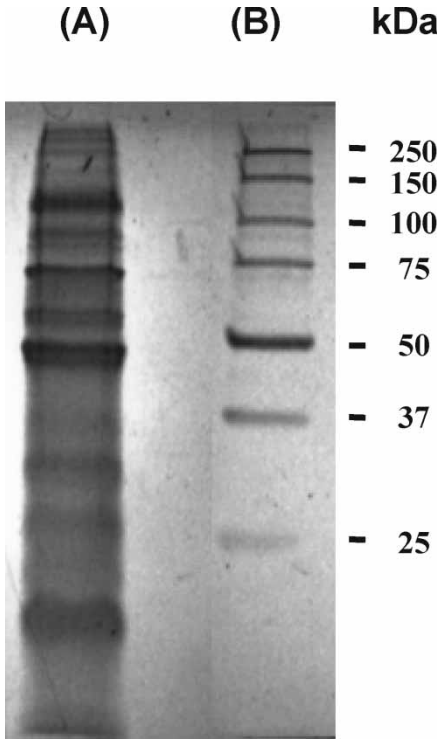


***Figure 1.*** SDS-PAGE separation of urinary proteins. The proteins were resolved in 12% (w/v) acrylamide and stained with colloidal coomassie blue. (a), urinary proteins isolated by centrifugal filtration were run and stained with Colloidal Coomassie Blue; (b), molecular weight of each molecular marker was shown in the right lane.



Figure 2 shows the representative nano-LC/MS/MS analysis of one of the peptides of a protein. Three strongest parent ions of one full MS spectrum (400–1500 m/z range) were selected for fragmentation. Each MS/MS spectrum was searched against the NCBI non-redundant protein sequence database using the Knexus program, where the enzyme was specified as trypsin with 1 missing cleavage. Mass error for parent ion mass was set as $+/-4$ Da and for fragment ion as $+/-0.6$ Da. Manual analysis was applied to validate protein identification results. The following criteria were used for manual verification. All the major isotope-resolved peaks should match the theoretically calculated fragment ions list; y, b, and a ions, as well as their water loss or amine loss peaks, are considered. The isotope-resolved peaks were emphasized because a single peak could come from an electronic spark and are less likely to be relevant to peptide

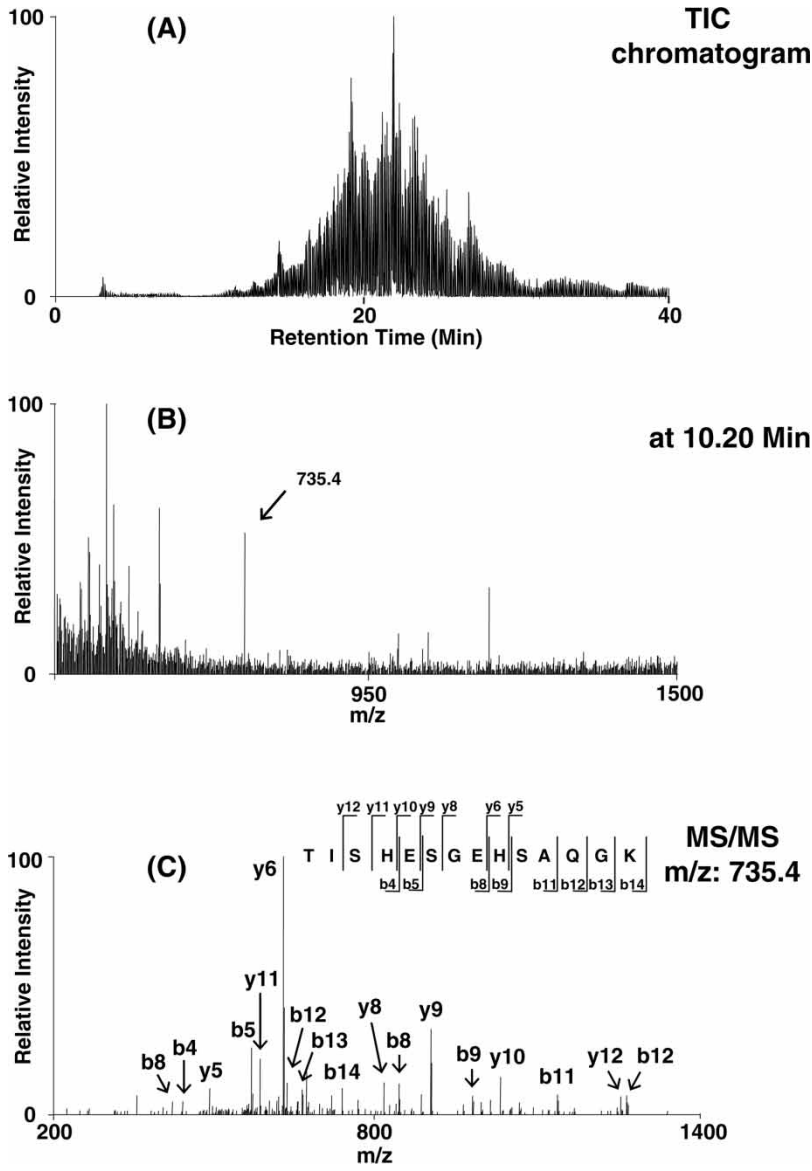


***Figure 2.*** An example of nano-HPLC/tandem MS analysis results. (A), total ion current (TIC) chromatogram of a nano-HPLC/tandem mass spectrometry; (B), Full MS spectrum at the retention time of 10.28 min (400–1500 m/z); (C), MS/MS spectrum of 735.4 m/z, which identified the peptide TISHESGEHSAQGK, unique to gastric cancer antigen Ga34 from *Homo sapiens.*

fragments. Specific amino acids were also considered for manual verification, since proline residues, acidic residues, and large hydrophobic residues showed relatively more intense peaks. The solid validation procedures and well established confirmation criteria described above will allow a large diffusion of such kind of methods for clinical research purposes.

Five hundred eighty nine peptides, corresponding to 195 unique proteins, were unambiguously identified in this study. All the redundant proteins were removed by confirming the unique peptides, and keratin proteins were also excluded to remove false-positive protein identification. Overlapped peptides were also removed by listing identified peptides. Table 1 shows the representative unique peptide of each protein. Unique peptides were selected by searching amino acid sequences of all identified proteins.

The reason that the urinary proteome of this study could not cover the other papers' profiling list may be due to the intra- and inter-individual variations of the protein expression in urine and also different sample preparation methods.



## DISCUSSION

For the purpose of unbiased broad urinary protein profiling, two dimensional separation techniques, SDS PAGE by molecular weight, and nano-HPLC by polarity, were applied to maximize the number of protein identification. Other methods such as shotgun (LC/MS/MS), 2DE/LC/MS/MS, and LC/MALDI/TOF can also be considered for profiling, but they have their own defects in achieving the maximum number of proteins. The disadvantage of LC/MS/MS analysis of an unfractionated tryptic digest of urinary proteins (shotgun method) is that the dominating human albumin peptides prevent other low abundance peptides from being selected by a data-dependent mode, which requires an additional albumin removal step or multidimensional LC techniques. The former has a danger of losing proteins, while the latter has a difficulty of optimal performance due to its complexity. In the case of 2DE, proteins out of the pI range are missed and it is also difficult to separate high MW proteins. Though a large number of samples could be generated from 2DE, protein identification by this method is a time consuming process, and it is also difficult to maintain good reproducibility. For the technique using MALDI/TOF, the number of identified proteins are confined because the peptide complex in a sample shows an ion suppression effect. Due to the low number of peptides generated by enzyme digestion, low molecular weight proteins are also difficult to identify.

The 1.5 mm home-made polyacrylamide gel was used so as to have higher loading capacity compared to the normally used 1 mm gel. By this simple method, a greater number of proteins was identified, even without following

***Table 1.*** Identified proteins in human urine by nano-HPLC/MSMS

| Identified protein | Mass (kDa) | gi number | Unique peptide | # of peptides |
|---|---|---|---|---|
| kininogen light chain | 4.2 | 224060 | IGEIKEETTSHLR | 1 |
| Ubiquitin | 8.4 | 229532 | TITLEVEPSDTIENVK | 3 |
| Ig kappa light chain variable region | 10.8 | 4323920 | ASGVPDKFSGSGSGTDFTLK | 1 |
| Ig kappa chain | 10.9 | 106717 | SSQSLLHR | 2 |
| Cystic fibrosis antigen | 10.9 | 225541 | ALNSIIDVYHK | 2 |
| Recombinant Human Cystatin A | 11.0 | 1311047 | TQVVAGTNYYIK | 2 |
| H4 histone family, member A | 11.3 | 4504301 | DAVTYTEHAK | 1 |
| Ig kappa light chain variable region | 11.4 | 4378334 | SPSSLSASVGDR | 1 |
| IgM | 11.8 | 510844 | LTQSPGTLSLSPGER | 3 |
| Ig light chain variable region | 12.0 | 12003250 | EIVLTQSPATLSLSPGER | 1 |
| Ig alpha heavy chain variable region | 12.3 | 15777241 | VTVSSASPTSPK | 1 |
| Ig kappa light chain variable region A2 | 12.8 | 5731233 | TVAAPSV | 1 |
| RBP | 13.2 | 296672 | VKENFDK | 1 |
| S100 calcium-binding protein A9 | 13.2 | 4506773 | KDLQNFLK | 3 |
| Heart fatty acid binding protein | 13.3 | 458862 | LVHLQK | 2 |
| Ig heavy chain | 13.7 | 435983 | LVESGGGLVKPGGSLR | 1 |
| beta-2-Microglobulin | 13.7 | 4757826 | VEHSDLSFSK | 2 |
| beta-Subunit signal transducing proteins GS/GI | 13.8 | 3387975 | AGVLAGHDNR | 1 |
| Ig G heavy chain | 13.9 | 1552322 | LLESGGGLVQPGGSLR | 1 |
| Ly-6 neurotoxin-like protein 1 isoform a | 14.0 | 29337281 | VLSNTEDLPLVTK | 1 |
| CD59 antigen p18–20 | 14.1 | 10835165 | AGLQVYNK | 1 |
| Eosinophil-Derived Neurotoxin | 15.6 | 14488647 | DPPQYPVVPVHLDR | 2 |
| Superoxide Dismutase Mutant | 15.7 | 5822065 | GDGPVQGIINFEQK | 2 |
| Prealbumin | 15.8 | 386998 | TSESGELHGLTTEEEFVEGIYK | 1 |
| Cystatin M precursor | 16.5 | 4503113 | AQSQLVAGIK | 2 |

| | | | | |
|---|---|---|---|---|
| Prolactin-induced protein | 16.5 | 4505821 | FYTIEILKVE | 2 |
| Chromosome 6 open reading frame 82 | 16.8 | 7706244 | TQSSLVPALTDFVR | 1 |
| Hypothetical protein XP_211608 | 17.1 | 27500930 | TTATRPSQGGDAPAGAK | 1 |
| Gm2 Activator Structure | 17.5 | 14278639 | VDLVLEK | 1 |
| Macrophage colony-stimulating factor | 17.6 | 234089 | TFYETPLQLLEK | 1 |
| Complement component C4A | 18.0 | 553210 | VLSLAQEQVGGSPEK | 1 |
| Ig J chain | 18.1 | 21489959 | SSEDPNEDIVER | 2 |
| Similar to common salivary protein 1 | 18.8 | 21687060 | LGALGGNTQEVTLQPGEYITK | 3 |
| Neutrophil gelatinase associated lipocalin monomer | 20.5 | 7245433 | SYPGLTSYLVR | 2 |
| Neutrophil elastase precursor | 20.7 | 182051 | VVLGAHNLSR | 1 |
| Prostatic binding protein | 21.0 | 4505621 | LYTLVLTDPDAPSR | 1 |
| Prostaglandin D2 synthase 21kDa | 21.0 | 32171249 | APEAQVSVQPNFQQDK | 4 |
| Nuclear mitotic apparatus protein-retinoic acid receptor alpha fusion protein | 21.4 | 3510458 | FYSTR | 1 |
| Similar to vitelline membrane outer layer protein I | 21.5 | 27500341 | GLGDDTALNDAR | 2 |
| Syndecan 4 | 21.6 | 4506861 | AGSGSQVPTEPK | 3 |
| Sorcin | 21.6 | 4507207 | SGTVDPQELQK | 1 |
| Peptidoglycan recognition protein | 21.7 | 4827036 | VPTPQAIR | 2 |
| Tetranectin | 22.5 | 4507557 | GGTLSTPQTGSENDALYEYLR | 2 |
| Lysosomal proteinase cathepsin B | 23.0 | 181178 | HYGYNSYSVSNSEK | 1 |
| Gastric cancer antigen Ga34 | 23.1 | 21724166 | TISHESGEHSAQGK | 1 |
| alpha-1-Acid glycoprotein 1 precursor | 23.3 | 1197209 | YVGGQEHFAHLLILR | 1 |
| Small GTP binding protein Rab7 | 23.4 | 1174149 | DPENFPFVVLGNK | 1 |
| alpha-1-Acid glycoprotein 2 precursor | 23.4 | 29170378 | EHVAHLLFLR | 3 |
| Dermatopontin precursor | 24.0 | 1082328 | GATTTFSAVER | 1 |

(*continued*)

***Table 1.*** Continued

| Identified protein | Mass (kDa) | gi number | Unique peptide | # of peptides |
|---|---|---|---|---|
| Twisted gastrulation | 25.0 | 10190664 | STVEELHEPIPSLFR | 1 |
| HGFL(S) protein | 25.3 | 28195103 | EDQTSPAPGLR | 2 |
| Dendritic cell nectin-like protein 1 short isoform | 25.6 | 25777796 | KGDQELHGEPTR | 1 |
| Ligand-Free Heterodimeric Human Glutathione S-transferase M2-3 | 26.4 | 5822511 | LKPQYLEELPGQLK | 1 |
| Triosephosphate isomerase 1 | 26.6 | 4507645 | TATPQQAQEVHEK | 1 |
| Endothelial protein C receptor precursor | 26.6 | 13626550 | HISAENTK | 3 |
| Secreted and transmembrane 1 precusor | 27.1 | 17389346 | AHGQESAIFNEVAPGYFSR | 1 |
| Apolipoprotein D | 28.0 | 619383 | IPTTFENGR | 4 |
| Aquaporin 1 | 28.1 | 19387211 | YPVGNNQTAVQDNVK | 1 |
| Ig kappa light chain VLJ region | 28.2 | 21669315 | FSGSGSGTDFTLR | 4 |
| Ig lambda light chain VLJ region | 28.2 | 21669581 | FSGSSSGAER | 1 |
| Ig kappa light chain VLJ region | 28.4 | 21669437 | LLIYDASNR | 1 |
| Glutaminyl-peptide cyclotransferase | 28.5 | 525241 | DHSLEGR | 5 |
| Ig lambda light chain VLJ region | 28.8 | 21669611 | RPSGVPDR | 2 |
| Insulin-like growth factor binding protein 7 | 29.1 | 4504619 | GHYGVQR | 3 |
| Ig kappa light chain VLJ region | 29.4 | 21669367 | SSQSVLYSSNNK | 1 |
| Dimethylarginine dimethylaminohydrolase 2 | 29.6 | 7524354 | TVVAGSSDAAQK | 1 |
| Leukocyte-associated Ig-like receptor 1 isoform b precursor | 29.8 | 11231177 | IDSVSEGNAGPYR | 1 |
| Ficolin 2 isoform b precursor | 30.2 | 8051586 | VDLVDFEDNYQFAK | 1 |
| Proapo-A-I protein | 30.7 | 178777 | AELQEGAR | 8 |
| Deoxyribonuclease I precursor | 31.4 | 1197173 | DSHLTAVGK | 4 |
| Zinc-Alpha-2-Glycoprotein | 31.6 | 4699583 | AGEVQEPELR | 8 |
| Syndecan | 32.5 | 338634 | EGEAVVLPEVEPGLTAR | 1 |

| | | | | |
|---|---|---|---|---|
| Leucine-rich alpha-2-glycoprotein | 34.3 | 72059 | DLLLPQPDLR | 2 |
| Procathepsin L | 35.8 | 5822035 | VFQEPLFYEAPR | 1 |
| Structure Of Gamma-Glutamyl Hydrolase | 35.9 | 20664327 | LDLTEKDYEILFK | 2 |
| Triggering receptor expressed on myeloid cells 4 | 36.0 | 21687218 | KGGILFSR | 1 |
| Truncated cathepsin H | 36.3 | 16506815 | GNFVSPVK | 1 |
| Glycoprotein VI-1 | 36.9 | 25392151 | YGFDQFALYK | 1 |
| Aldo-keto reductase family 7 | 37.1 | 19343681 | FYAFNPLAGGLLTGK | 1 |
| Hypothetical protein | 37.2 | 31873256 | ITVNYPPYISK | 1 |
| Ig alpha-1 heavy chain constant region | 37.5 | 184749 | QEPSQGTTTFAVTSILR | 4 |
| beta-2-Glycoprotein I precursor | 38.3 | 543826 | ATVVYQGER | 1 |
| alpha-1-Microglobulin/bikunin precursor | 39.0 | 4502067 | AVLPQEEEGSGGGQLVTEVTK | 3 |
| alpha-2-HS-glycoprotein | 39.3 | 4502005 | HTLNQIDEVK | 1 |
| Aldolase B | 39.4 | 178357 | LDQGGAPLAGTNK | 1 |
| CD44 antigen | 39.4 | 21429241 | FAGVFHVEK | 2 |
| KIAA0174 protein | 39.9 | 12653323 | TNQIGTVNDR | 1 |
| Monocyte antigen CD14 precursor | 40.0 | 3983127 | LTVGAAQVPAQLLVGALR | 5 |
| Lectin | 40.2 | 5803023 | DNFHGLAIFLDTYPNDETTER | 4 |
| Ciliary neurotrophic factor receptor alpha precursor | 40.6 | 69714 | HSPQEAPHVQYER | 1 |
| Actin, beta | 41.0 | 14250401 | SYELPDGQVITIGNER | 2 |
| Biglycan preproprotein | 41.6 | 4502403 | IQAIELEDLLR | 1 |
| Pepsinogen 5 | 41.9 | 23943854 | HNLNPAR | 1 |
| Serine (or cysteine) proteinase inhibitor | 42.7 | 13489087 | LGVQDLFNSSK | 1 |
| TSLC1-like 2 | 42.7 | 21686977 | FQLEEFSPR | 2 |
| Chitobiase | 43.7 | 4758092 | APYYNYK | 2 |
| Alpha1-Antitrypsin | 44.2 | 1942629 | TDTSHHDQDHPTFNK | 9 |

(*continued*)

***Table 1.*** Continued

| Identified protein | Mass (kDa) | gi number | Unique peptide | # of peptides |
|---|---|---|---|---|
| Prostatic acid phosphatase precursor | 44.5 | 14250150 | FQELESETLK | 8 |
| Protein Z | 44.7 | 4506121 | DFAEHLLIPR | 1 |
| Lysosomal membrane glycoprotein-2 | 44.8 | 307110 | IPLNDLFR | 1 |
| Golgi phosphoprotein 2 | 45.3 | 29550838 | IQSSHNFQLESVNK | 1 |
| Protein C inhibitor | 45.7 | 1144561 | TLYLADTFPTNFR | 10 |
| Thyroxine-binding globulin precursor | 46.3 | 1351236 | NALALFVLPK | 1 |
| SERPINA3 protein | 47.6 | 21961493 | AVLDVFEEGTEASAATAVK | 6 |
| Procollagen C-endopeptidase enhancer | 47.9 | 4505643 | YDALEVFAGSGTSGQR | 1 |
| GW112 protein | 48.1 | 29126831 | DYSPQHPNK | 2 |
| Kallistatin precursor | 48.5 | 1708609 | LGFTDLFSK | 1 |
| Plasminogen activator | 48.5 | 4505863 | KPSSPPEELK | 1 |
| Aspartyl protease 3 | 48.5 | 6561816 | FAIQYGTGR | 2 |
| Ecay-accelerating factor | 48.7 | 87316 | REPSLSPK | 2 |
| Limitrin | 49.1 | 14150145 | AYGPLFLR | 1 |
| G protein-coupled receptor | 49.3 | 13112057 | VPSEGAYDIILPR | 1 |
| Recombinant IgG2 heavy chain | 49.5 | 14030849 | VVSVLTVVHQDWLNGK | 6 |
| Ig G1 | 49.7 | 230581 | TTPPVLDSDGSFFLYSK | 1 |
| Physiological Dimer Hpp Precursor | 51.4 | 2098347 | APDQDEIQR | 1 |
| Hemopexin precursor | 51.5 | 386789 | GECQAEGVLFFQGDR | 2 |
| alpha-1-B-glycoprotein | 51.9 | 69990 | LETPDFQLFK | 3 |
| Arylsulfatase A | 51.9 | 2981949 | DPGENYNLLGGVAGATPEVLQALK | 1 |
| Clusterin | 52.2 | 32891795 | LFDSDPITVTVPVEVSR | 6 |
| Complement component 1 | 53.4 | 7706083 | LGNHPVHR | 2 |
| Annexin A11 | 54.3 | 4557317 | GTITDAPGFDPLR | 3 |
| Vitronectin precursor | 54.3 | 18201911 | FEDGVLDPDYPR | 1 |

| | | | | |
|---|---|---|---|---|
| Plasma protease (C1) inhibitor precursor | 55.1 | 179619 | TNLESILSYPK | 9 |
| PTPNS1 protein | 55.2 | 20070655 | AKPSAPVVSGPAAR | 1 |
| Podocalyxin-like precursor | 55.4 | 33598950 | LGDQGPPEEAEDR | 3 |
| Prolylcarboxypeptidase | 55.7 | 4826940 | YYGESLPFGDNSFK | 1 |
| The R337q variant of human pancreatic alpha-mylase | 55.8 | 18655893 | LTGLLDLALEK | 10 |
| LIR-D1 precursor | 55.9 | 22266726 | VVAGEVQVQR | 2 |
| alpha-Amylase | 57.6 | 178585 | LSGLLDLALGK | 1 |
| Amylase | 57.7 | 10280622 | LVGLLDLALEK | 1 |
| Carboxypeptidase N 83 kDa chain | 58.6 | 115877 | LSNNALSGLPQGVFGK | 1 |
| Hypothetical protein XP_293219 | 59.2 | 30159017 | VTSGSTFNENIR | 1 |
| Histidine-rich glycoprotein precursor | 59.5 | 4504489 | DGYLFQLLR | 1 |
| beta-Galactosidase-related protein precursor | 60.5 | 114947 | SLYPLTFIQVK | 2 |
| KIAA0653 protein | 60.8 | 3327120 | GLYDVVSVLR | 1 |
| Nectin 3 | 61.0 | 11386199 | VTNSLGQR | 1 |
| Biotinidase precursor | 61.1 | 4557373 | SHLIIAQVAK | 2 |
| Similar to RIKEN cDNA 2610528G05 gene | 64.1 | 15489339 | SLTLGIEPVSPTSLR | 9 |
| Unnamed protein product | 64.7 | 16554039 | LTVLSQPK | 5 |
| Galectin 3 binding protein | 65.3 | 5031863 | YSSDYFQAPSDYR | 11 |
| I factor (complement) | 65.7 | 4504579 | AQLGDLPWQVAIK | 1 |
| Serum albumin in a complex with myristic acid and tri-iodobenzoic acid | 66.0 | 4389275 | KLVAASQAALG | 24 |
| Lutheran blood group | 67.3 | 31543106 | AGAAGTAEATAR | 1 |
| Peptidoglycan recognition protein L precursor | 67.9 | 21361845 | AGLLRPDYALLGHR | 1 |
| Radixin | 68.5 | 4506467 | IGFPWSEIR | 1 |
| Unnamed protein product | 69.2 | 28590 | KVPEVSTPTLVEVSR | 3 |

(*continued*)

***Table 1.*** Continued

| Identified protein | Mass (kDa) | gi number | Unique peptide | # of peptides |
|---|---|---|---|---|
| Uromodulin precursor | 69.7 | 137116 | TLDEYWR | 14 |
| Coagulation factor II precursor | 69.9 | 30802115 | DKLAACLEGNCAEGLGTNYR | 1 |
| Kininogen | 71.9 | 386852 | KYFIDFVAR | 11 |
| MASP-2 protein | 75.6 | 5459324 | LASPGFPGEYANDQER | 5 |
| Transferrin | 77.0 | 4557871 | EGYYGYTGAFR | 11 |
| Cadherin 13 preproprotein | 78.2 | 4502719 | DIQGSLQDIFK | 2 |
| Lactoferrin | 78.3 | 186833 | NLLFNDNTECLAR | 16 |
| Bile-salt-activated lipase precursor | 78.3 | 231629 | LGLLGDSVDIFK | 2 |
| N-acetylglucosaminidase | 82.1 | 4505327 | QLAGLVANYYTPR | 5 |
| alpha-Glucosidase | 82.8 | 31609 | STFAGHGR | 4 |
| Polymeric immunoglobulin receptor | 83.2 | 31377806 | IIEGEPNLK | 5 |
| Gelsolin | 85.6 | 4504165 | DSQEEEKTEALTSAK | 2 |
| Membrane metallo-endopeptidase | 85.7 | 4505203 | AIAQLNSK | 1 |
| Bone-derived growth factor | 86.3 | 1203965 | NNEEYLALIFEK | 1 |
| OB-cadherin-1 | 87.9 | 1377894 | FFTINPEDGFIK | 2 |
| mu-Protocadherin | 88.1 | 11245187 | GINQPIIYSIFR | 1 |
| Dipeptidyl peptidase iv | 88.2 | 35336 | TYTLTDYLK | 7 |
| E-cadherin | 91.1 | 1617084 | NTGVISVVTTGLDR | 4 |
| Desmocollin 2 isoform Dsc2b preproprotein | 93.7 | 13435366 | GPGVDQEPR | 3 |
| Fibrinogen | 94.9 | 4503689 | DEAGSEADHEGTHSTK | 2 |
| HP95 | 95.9 | 13375569 | LANQAADYFGDAFK | 3 |
| PROM1 protein | 96.2 | 15082356 | VLPIEQSLSTLYQSVK | 1 |
| Discoidin domain receptor family | 96.7 | 5453814 | QVLDGNSNPYDIFLK | 1 |
| Tyrosine kinase receptor | 97.3 | 292870 | APLQGTLLGYR | 1 |
| Ceruloplasmin | 97.6 | 180249 | GAYPLSIEPIGVR | 1 |

| | | | | |
|---|---|---|---|---|
| N-cadherin | 99.7 | 253483 | FLEAGIYEVPIIITDSGNPPK | 3 |
| Inter-alpha-trypsin inhibitor heavy chain-related protein precursor | 103.3 | 1082547 | NPLVWVHASPEHVVVTR | 5 |
| Collagen | 108.4 | 30851190 | VPSYQALLR | 5 |
| Glutamyl aminopeptidase | 109.2 | 1082404 | ASLIDDAFALAR | 1 |
| Membrane alanine aminopeptidase precursor | 109.4 | 4502095 | FSTEYELQQLEQFK | 14 |
| Contactin 1 isoform 2 precursor | 111.8 | 28373119 | GPPGPPGGLR | 1 |
| Solute carrier family 12 member 3 | 113.0 | 1717801 | VFVGGQINR | 1 |
| Thyrotropin-releasing hormone degrading ectoenzyme | 116.9 | 7019561 | LIEFYEDYFK | 4 |
| Sodium potassium chloride cotransporter 2 | 121.2 | 4557849 | ITDAELEAVK | 3 |
| Mucin 1 precursor | 122.0 | 547937 | EGTINVHDVETQFNQYK | 1 |
| KIAA1336 protein | 132.2 | 27465513 | FYLSK | 1 |
| Attractin | 133.4 | 3676347 | SVNNVVVR | 3 |
| Urogastrone | 133.8 | 4503491 | ITAVSLDVLDK | 16 |
| Type XV collagen | 141.7 | 3893879 | FTGSLQQLTVHPDPR | 4 |
| Maltase-glucoamylase | 209.6 | 4758712 | NPFGIEIR | 9 |
| Agrin precursor | 212.7 | 2988422 | TFVEYLNAVTESEK | 1 |
| PTPsigma | 213.1 | 1407625 | VLAFTSVGDGPLSDPIQVK | 1 |
| Hypothetical protein | 268.6 | 31874109 | LGVRPSQGGEAPR | 4 |
| Filamin | 280.5 | 1203969 | DAGEGLLAVQITDPEGKPK | 4 |
| alpha 3 Type VI collagen isoform 5 precursor | 321.0 | 17149811 | VAVVQYSDR | 1 |
| Cubilin | 398.7 | 4557503 | FHADYAR | 24 |
| Heparan sulfate proteoglycan perlecan | 466.3 | 11602963 | EVSEAVVDTLESEYLK | 10 |
| Lipoprotein, Lp(a) | 501.0 | 5031885 | GTYSTTVTGR | 1 |
| Glycoprotein receptor gp330 | 521.6 | 32816595 | DDQPFLITVR | 23 |

the albumin removal step. An albumin removal kit was not used in this study to avoid loss of some trace proteins, which may have an affinity for this kit. While precipitation and ultracentrifugation have been widely used for protein isolation, an acetone-precipitated sample had a problem of narrow protein molecular weight range and an ultracentrifuged sample lacked adhesion proteins such as cadherin. However, the centrifugal filter sample used in this study showed a broad molecular weight range of 4.2 to 521.6 kDa, including cadherin proteins.[18]

In Table 1, the identified proteins are displayed according to their molecular weights. Some previous reports displayed proteins by the number of identified peptides, followed by relative quantification.[14,19] However, this seems to be illogical because it is based on the number of peptides only. Molecular weight, preparation with molecular weight cutoff membrane, solubility, enzyme resistant proteins, and proteins having repeating sequences were not considered in those studies. The higher molecular weight proteins yield higher peptides. Some proteins will be lost in the process of filtration with membranes and will be precipitated with the change of temperature and buffer. As an example of proteins having repeatable sequences, lipoprotein, Lp(a) protein identified in this study, has 31 sequences of GTYSTTVTGR.



Some of the proteins identified in this study include oncogene proteins; Neutrophil Gelatinase Associated Lipocalin Monomer, small GTP binding protein Rab7, and gastric cancer antigen Ga34. Neutrophil Gelatinase Associated Lipocalin has been shown to be associated with early urinary biomarkers for ischemic renal injury, small GTP binding protein Rab7 with vesicle transport, and gastric cancer antigen Ga34 with tumor cells.[25–27] Mental or physical stress may potentially promote cancer cell growth under certain circumstances. To monitor these proteins among healthy and diseased people, further study should be directed towards fresh sample collection with a regulated procedure and a sufficient number of samples. In practice, urine samples have been collected over 24 hours by patients. Thus, different collecting and storing methods would generate a large deviation in protein components.

There are still many unresolved issues concerning urine proteomics, even if the urinary protein library, a human protein database, is completed in the future. Answering ample questions such as low molecular polypeptide signature ions, protein patterns, etc. could be a next proteomic challenge. Therefore, developing liquid chromatography and mass spectrometry techniques may be very useful for elucidating the above questions.

## ACKNOWLEDGMENT

I thank Dr. Shama P. Mirza and Teresa Macnee for various suggestions.

## REFERENCES


1. Greive, K.A.; Balazs, N.D.H.; Comper, W.D. Protein fragments in urine have been considerably underestimated by various protein assays. Clin. Chem. **2001**, *47*, 1717–1719.
2. Wittke, S.; Fliser, D.; Haubitz, M.; Bartel, S.; Krebs, R.; Hausadel, F.; Hillmann, M.; Golovko, I.; Koester, P.; Haller, H.; Kaiser, T.; Mischak, H.; Weissinger, E.M. Determination of peptides and proteins in human urine with capillary electrophoresis-mass spectrometry, a suitable tool for the establishment of new diagnostic markers. J. Chromatogr. A **2003**, *1013*, 173–181.
3. Hampel, D.J.; Sansome, C.; Sha, M.; Brodsky, S.; Lawson, W.E.; Goligorsky, M.S. Toward proteomics in uroscopy: urinary protein profiles after radiocontrast medium administration. J. Am. Soc. Nephrol. **2001**, *12*, 1026–1035.
4. Kiernan, U.A.; Tubbs, K.A.; Nedelkov, D.; Niederkofler, E.E.; McConnell, E.; Nelson, R.W. Comparative urine protein phenotyping using mass spectrometric immunoassay. J. Proteome Res. **2003**, *2*, 191–197.
5. Davis, M.T.; Spahr, C.S.; McGinley, M.D.; Robinson, J.H.; Bures, E.J.; Beierle, J.; Mort, J.; Yu, W.; Luethy, R.; Patterson, S.D. Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry II. Limitations of complex mixture analyses. Proteomics **2001**, *1*, 108–117.
6. Marshall, T.; Williams, K.M. High resolution two-dimensional electrophoresis of human urinary proteins. Anal. Chim. Acta **1998**, *372*, 147–160.
7. Tateda, N.; Matsuhisa, K.; Hasebe, K.; Miura, T. Sensitive and specific determination of histidine in human serum, urine, and stradum corneum by a flow injection method based on fluorescence derivatization with o-phthalaldehyde. J. Liq. Chromatogr. & Relat. Technol. **2001**, *24*, 3181–3196.
8. Adelberg, A.M.; Miller, J.; Doerzbacher, M.; Lambers, D.S. Correlation of quantitative protein measurements in 8-, 12-, and 24-hour urine samples for the diagnosis of preeclampsia. Am. J. Obstet. Gynecol. **2001**, *185*, 804–807.
9. Bouchelouche, K.; Kristensen, B.; Nordling, J.; Horn, T.; Bouchelouche, P. Increased urinary leukotriene E4 and eosinophil protein X excretion in patients with interstitial cystitis. J. Urol. **2001**, *166*, 2121–2125.
10. Rodriguez-Thompson, D.; Lieberman, E.S. Use of a random urinary protein-to-creatinine ratio for the diagnosis of significant proteinuria during pregnancy. Am. J. Obstet. Gynecol. **2001**, *185*, 808–811.
11. Pannacciulli, N.; Cantatore, F.P.; Minenna, A.; Bellacicco, M.; Giorgino, R.; De Pergola, G. Urinary albumin excretion is independently associated with C-reactive protein levels in overweight and obese nondiabetic premenopausal women. J. Intern. Med. **2001**, *250*, 502–507.
12. Cutillas, P.R.; Norden, A.G.; Cramer, R.; Burlingame, A.L.; Unwin, R.J. Detection and analysis of urinary peptides by on-line liquid chromatography and mass spectrometry: application to patients with renal Fanconi syndrome. Clin. Sci. **2003**, *104*, 483–490.
13. Hortin, G.L.; Meilinger, B.; Drake, S.K. Size-selective extraction of peptides from urine for mass spectrometric analysis. Clin. Chem. **2004**, *50*, 1092–1095.
14. Spahr, C.S.; Davis, M.T.; McGinley, M.D.; Robinson, J.H.; Bures, E.J.; Beierle, J.; Mort, J.; Courchesne, P.L.; Chen, K.; Wahl, R.C.; Yu, W.; Luethy, R.; Patterson, S.D. Towards defining the urinary proteome using liquid

chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. Proteomics **2001**, *1*, 93–107.
15. Pang, J.X.; Ginanni, N.; Dongre, A.R.; Hefta, S.A.; Opitek, G.J. Biomarker discovery in urine by proteomics. J. Proteome Res. **2002**, *1*, 161–169.
16. Schaub, S.; Wilkins, J.; Weiler, T.; Sangster, K.; Rush, D.; Nickerson, P. Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. Kidney Int. **2004**, *65*, 323–332.
17. Bakoush, O.; Grubb, A.; Rippe, B.; Tencer, J. Urine excretion of protein HC in proteinuric glomerular diseases correlates to urine IgG but not to albuminuria. Kidney Int. **2001**, *60*, 1904–1909.
18. Thongboonkerd, V.; McLeish, K.R.; Arthur, J.M.; Klein, J.B. Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation. Kidney Int. **2002**, *62*, 1461–1469.
19. Moon, M.H.; Myung, S.; Plasencia, M.; Hilderbrand, A.E.; Clemmer, D.E. Nanoflow LC/ion mobility/CID/TOF for proteomics: analysis of a human urinary proteome. J. Proteome Res. **2003**, *2*, 589–597.
20. Vlahou, A.; Giannopoulos, A.; Gregory, B.W.; Manousakas, T.; Kondylis, F.I.; Wilson, L.L.; Schellhammer, P.F.; Wright, G.L.; Semmes, O.J. Protein profiling in urine for the diagnosis of bladder cancer. Clin. Chem. **2004**, *50*, 1438–1441.
21. Van Eenoo, P.; Delbeke, F.T.; De Jong, F.H.; De Backer, P. Endogenous origin of norandrosterone in female urine: indirect evidence for the production of 19-norsteroids as by-products in the conversion from androgen to estrogen. J. Steroid Biochem. Mol. Biol. **2001**, *78*, 351–357.
22. Desroches, M.C.; Mathurin, J.C.; Richard, Y.; Delahaut, P.; de Ceaurriz, J. Urinary 19-norandrosterone purification by immunoaffinity chromatography: application to gas chromatography/combustion/isotope ratio mass spectrometric analysis. Rapid Commun. Mass Spectrom. **2002**, *16*, 370–374.
23. Tsikas, D. Mass spectrometry-validated HPLC method for urinary nitrate. Clin. Chem. **2004**, *50*, 1259–1261.
24. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **1976**, *72*, 248–254.
25. Mishra, J.; Ma, Q.; Prada, A.; Mitsnefes, M.; Zahedi, K.; Yang, J.; Barasch, J.; Devarajan, P. Identification of neutrophil gelatinase-associated lipocalin as a novel early urinary biomarker for ischemic renal injury. J. Am. Soc. Nephrol. **2003**, *14*, 2534–2543.
26. Novick, P.; Zerial, M. The diversity of Rab proteins in vesicle transport. Curr. Opin. Cell Biol. **1997**, *9*, 496–504.
27. Line, A.; Stengrevics, A.; Slucka, Z.; Li, G.; Jankevics, E.; Rees, R.C. Serological identification and expression analysis of gastric cancer-associated genes. Br. J. Cancer **2002**, *86*, 1824–1830.

Received October 20, 2004
Accepted November 25, 2004
Manuscript 6523