Discovery Research of Novel Biomarkers in Osteoarthritis using Glycoproteomics

2014

Dissertation

Takeshi Ishihara

Transdisciplinary Life Science Course,
Graduate school of Life Science,
Hokkaido University
&
Innovative Drug Discovery Research Laboratories,
SHIONOGI & CO., LTD

Contents

Acknowledgements

I would like to express my sincere gratitude to my thesis adviser, professor Shin-Ichiro Nishimura, of Hokkaido University for his suggestions, supports, discussions and encouragement. The work described herein would not have been successful without him.

I am grateful to professor Norimasa Iwasaki and assistant professor Maho Amano of Hokkaido University, professor Naoshi Fukui of Sagamihara hospital for their suggestions, supports, discussions, and encouragement.

I am grateful to Dr. Hiroshi Takemoto, Dr. Yoshito Numata, Dr. Hiroyuki Okamoto, Dr. Atsushi Morita, Dr. Hiroko Togame, Dr. Yoshikazu Tanaka and Mr. Koji Takahashi of Shionogi & Co., Ltd. for valuable discussions, critical reading and suggestions on the dissertation manuscript. They drew my attention to unexpected discoveries and encouraged me to conquer prejudice with experimentation. I also thank to my colleagues of Shionogi & Co., Ltd, for their supports and encouragement.

I would like to thank all of my friends and my family for their invaluable mental supports and encouragement. I could not have completed this thesis without them.

Dedication

I would like to dedicate this dissertation to my parents, Suehiro Ishihara and Michiko Ishihara.

Abbreviation

2D	two-dimentional
2-DE	two-dimensional gel electrophoresis
aoWR	N^{α}-((aminooxy)acetyl)tryptophanylarginine methyl ester
BSA	bovine serum albumin
CD	cluster of differentiation
CHAPS	3-((3-cholamidopropyl)dimethylammonium)-1-propanesulfonate
ConA	concanavalin A
CV	coefficient of variation
dHex	deoxyhexose
DIG	digoxigenin
DMEM	Dulbecco's modified Eagle's medium
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
ESI	electrospray ionization
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
Hex	hexose
HexNAc	N-acetylhexosamine
HPLC	high performance liquid chromatography
iTRAQ	Isobaric tags for relative and absolute quantitation
JSW	joint space width
LC	liquid chromatography
MALDI	matrix-assisted laser desorption/ionization
MARS	multiple affinity removal system

MS	mass spectrometry
MSC	Mesenchymal stem cell
MSC	mesenchymal stem cell
NeuAc	N-acetylneuraminic acid
NOR	normal volumteer
OA	Osteoarthritis
PC	positive control
PNGase	peptide-N-Glycosidase
PTM	posttranslational modification
qPCR	quantitative polymerase chain reaction
RP	reverse-phase
SCX	strong cation exchange
SDS	sodium dodecyl sulfate
siRNA	small interfering RNA
SSEA	stage-specific embryonic antigen
TOF	time-of-flight
WGA	wheat germ agglutinin

Chapter 1

General Introduction

1-1. Osteoarthritis and Regenerative medicine

Osteoarthritis (OA) is the most prevalent form of arthritis, and is characterized by a gradual loss of cartilage matrix that often extends over years and decades[1-3]. Age is the most powerful risk factor for OA. Thus, with increasing longevity, OA has now become a leading cause of disability for older adults in developed countries. Currently available pharmacological therapies are mainly used for pain relief, and include acetaminophen, non-steroidal anti-inflammatory drugs (NSAIDs), cyclooxgenase-2 inhibitors, glucocorticoids, and opioids. Other palliative drugs affecting various targets such as catabolic enzymes (ex. matrix metalloproteinases, etc.) or cytokine-activated signalling cascades (ex. IL-1, TNF-α, etc.) are currently in development, but a curative treatment to induce regeneration and repair of cartilage is still lacking.

Although OA can affect any synovial joint, OA in knee joints has the highest prevalence rate and greatest association with disability[4,5]. Knee OA has now become a large economic and medical burden on society[6,7]. Therefore, there is increasing demand to establish effective therapies for this disease. However, no effective treatments have yet been established for OA that can inhibit or even delay disease progression. This may be partly due to the lack of reliable biomarkers to predict disease progression. In knee joints, progression of OA is most often determined by changes in the joint space measured on plain radiographs obtained in weight-bearing positions[8,9]. However, as progression of OA occurs gradually over several years, observation over an extended period of time is required to determine whether a given therapy can effectively inhibit progression of the disease. This necessity for a prolonged trial period significantly hinders the development of effective treatments for the disease. Thus, the establishment of reliable prognostic markers would greatly facilitate the process of drug

development by reducing the time required for clinical trials. Identification of proteins that are closely associated with disease progression may also be of benefit in research regarding the etiology of OA, which is still unclear.

Currently, regenerative medicine using mesenchymal stem cells (MSCs) has been approached to cure OA disease[2,10]. MSCs in various adult tissues, such as bone marrow, adipose tissue, and synovial fluid, have multipotency and self-renewal capabilities, including induction to undergo chondrogenic differentiation[11-15]. Especially, MSCs from the synovial membrane have the ability to differentiate into chondrocytes[13], indicating that they potentially have a physiological role in the repair and degeneration of the articular cartilage. Therefore, it is expected that medications induce proliferation and differentiation of MSCs *in vitro* and *in vivo* will contribute to repair and regeneration of damaged cartilage. Advent of new sensitive markers that contribute to development of novel therapeutic reagents to induce differentiation of MSCs are strongly required for improving quality of life of patients suffering OA.

The purpose of this thesis is discovery of novel biomarkers to progress innovative drug discovery to cure osteoarthritis.

1-2. Glycoproteomics for discovery of chondrogenic differentiation markers

Protein glycosylation is one of the most important posttranslational modifications in cell surface proteins and extracellular matrix proteins, which is considered to have a variety of biological functions, including enhancement of protein stability, controlling cell-to-cell communication and adhesion[16]. Recently, it has been well documented that glycan modifications of proteins greatly contribute to the pathogenesis of many diseases[17,18] and cell surface markers(ex. SSEA-1, Tra1-60, etc.). One of the characteristics of cartilage is that chondrocytes exist in the extracellular matrix. It has been well documented that glycan modifications the characteristics of cartilage is that chondrocytes exist in the extracellular matrix. It is also well known that glycoproteins are abundant on the cell surface and in cartilage extracellular matrix. Recently, we reported that some high-mannose type *N*-glycan levels are decreased significantly both in human OA cartilage and in degraded mouse cartilage[19], strongly suggesting an association between *N*-glycans of cell surface glycoproteins and the pathogenesis of OA. Pabst *et al.*[20] also demonstrated that the levels of glycophenotype of primary human chondrocytes were altered by inflammatory cytokines. These clearly mean that proteins with such characteristic *N*-glycoforms might become novel chondrogenic cell markers. However, glycoproteins bearing high-mannose type *N*-glycans in cartilage and their functional roles in the process of cartilage degradation have not yet been uncovered. Glycoproteomics is a new potential strategy among focused proteomics approaches for the characterization of plasma membrane proteins[21]. This approach has facilitated comprehensive identification of cell surface glycoproteins with low abundance and insolubility by hydrazide chemistry[22,23] and lectin affinity chromatography[24]. However, determination of the total glycan structures from whole cellular glycoproteins had long remained to be a challenging and extremely difficult task due to the lack of general platforms for high throughput glycomics[25]. We have developed a standardized protocol of glycomics based on the simple chemical enrichment method, namely glycoblotting method[26] that allows for rapid and large-scale enrichment analysis of human serum glycans[27,28].

We also recently demonstrated the versatility of this glycoblotting protocol for the monitoring and characterization of the processes of dynamic cellular differentiation of mouse embryonic carcinoma cells and mouse embryonic stem cells into cardiomyocytes or neural cells[29]. In *Chapter 2 and 3*, we discovered novel chondrogenic differentiation markers through high throughput glycoproteomics uncovering proteins displaying characteristic glycan structures, notably a combination of glycoblotting-based quantitative cellular *N*-glycomics and glycoforms-focused reverse proteomics/genomics.

1-3. Glycoproteomics for discovery of plasma OA biomarkers

To date, several proteins are known to associate with the progression of knee OA. Cartilage oligomeric matrix protein (COMP) in serum[30,31] and the C-telopeptide of type II collagen (CTX-II) in urine[32-34] are the examples. Hyaluronic acid and C-reactive protein (CRP) in serum may also be associated with the progression of the disease[35-38]. However, none of these markers are reliable enough to be used clinically.

Proteomic analysis coupled with mass spectrometry (MS) is a powerful tool for exploring biomarkers, and this approach has also been used in the discovery of OA biomarkers[39]. However, it is very difficult to find new biomarkers from plasma samples by these proteomics approaches because the “classical proteins” account for 99% of total plasma proteins with a high dynamic range of 10 orders of magnitude[40]. In addition, the plasma proteins are known to have many variations of posttranslational modifications, such as phosphorylation, glycosylation, glycation, etc. Several effective procedures have been reported for improving these issues by removing high-abundance proteins from the plasma by affinity chromatography, e.g., multiple affinity removal system (MARS) [41], and concentrating the proteins focusing on a specific modification of interest coupled with proteomics analysis [42-44].

Glycosylation, one of the most common forms of PTM of proteins, plays important roles in the maintenance of protein stability, binding of ligands to receptors, and adhesion of cells to the matrix[45,46]. Some protein biomarkers in cancer are known to have glycosylation modifications, such as alpha-fetoprotein in hepatocellular carcinoma, CA125 in ovarian cancer, Her2/neu in breast cancer, and prostate-specific antigen in prostate cancer[47-49]. The carbohydrate moieties of plasma proteins are changed in the case of cancer and rheumatoid diseases, and these changes are considered to be useful markers[50,51]. In addition, the differences in carbohydrate structures on total serum proteins demonstrated the possibility of the efficient detection of patients with hepatocellular carcinoma[28]. These findings suggest that changes in the glycosylation modification profiles (expression levels and carbohydrate modification patterns) of plasma

glycoproteins are closely linked to disease states. Thus, enrichment of glycosylated proteins should be an effective procedure for discovering biomarkers in blood.

In *Chapter 4*, we demonstrated that glycoproteomic analysis is a powerful tool for finding novel biomarkers in plasma. This method may be used for the exploration for OA biomarkers, because matrix components of articular cartilage, such as proteoglycans, are known to be highly glycosylated[52]. In *Chapter 5*, the glycoproteomic approach using 2D-LC-MALDI was used to discover factors the concentration of which changes with the progression of knee OA.

1-4. References

1. Wieland HA, Michaelis M, Kirschbaum BJ, Rudolphi KA (2005) Osteoarthritis - an untreatable disease? *Nat Rev Drug Discov* **4:** 331-344.
2. Clouet J, Vinatier C, Merceron C, Pot-vaucel M, Maugars Y, Weiss P, Grimandi G, Guicheux J (2009) From osteoarthritis treatments to future regenerative therapies for cartilage. *Drug Discov Today* **14:** 913-925
3. Hunter DJ (2011) Pharmacologic therapy for osteoarthritis-the era of disease modification. *Nat Rev Rheumatol* **7:** 13-22
4. Murphy L, Schwartz TA, Helmick CG, et al. (2008) Lifetime risk of symptomatic knee osteoarthritis. *Arthritis Rheum* **59**:1207-1213
5. Jinks C, Jordan K, Croft P. (2007) Osteoarthritis as a public health problem: the impact of developing knee pain on physical function in adults living in the community: (KNEST 3). *Rheumatology (Oxford)* **46**:877-881
6. Kotlarz H, Gunnarsson CL, Fang H, Rizzo JA. (2009) Insurer and out-of-pocket costs of osteoarthritis in the US: Evidence from national survey data. *Arthritis Rheum* **60**:3546-3553
7. Kim S. (2008) Changes in surgical loads and economic burden of hip and knee replacements in the US: 1997-2004. *Arthritis Rheum* **59**:481-488
8. Fukui N, Yamane S, Ishida S, et al. (2010) Relationship between radiographic changes and the symptoms or physical examination findings in the subjects with symptomatic medial knee osteoarthritis: a three-year prospective study. *BMC Musculoskeletal Disorders* **11**:269-278
9. Emrani PS, Katz JN, Kessler CL, et al. (2008) Joint space narrowing and Kellgren-Lawrence progression in knee osteoarthritis: an analytic literature synthesis. *Osteoarthritis Cartilage* **16**:873-882
10. Rachel AO. (2012) Cell sources for the regeneration of articular cartilage: the past, the horizon and the future. *Int. J. Exp. Path* **93**: 389-400

11. De Bari C, Dell'Accio F, Tylzanowski P, Luyten FP (2001) Multipotent mesenchymal stem cells from adult human synovial membrane. *Arthritis Rheum* **44:** 1928-1942
12. Jones EA, English A, Henshaw K, Kinsey SE, Markham AF, Emery P, McGonagle D (2004) Enumeration and phenotypic characterization of synovial fluid multipotential mesenchymal progenitor cells in inflammatory and degenerative arthritis. *Arthritis Rheum* **50:** 817-827
13. Sakaguchi Y, Sekiya I, Yagishita K, Muneta T (2005) Comparison of human stem cells derived from various mesenchymal tissues: superiority of synovium as a cell source. *Arthritis Rheum* **52:** 2521-2529
14. Nöth U, Steinert AF, Tuan RS (2008) Technology insight: adult mesenchymal stem cells for osteoarthritis therapy. *Nat Clin Pract Rheumatol* **4:** 371-380
15. Kolf CM, Cho E, Tuan RS (2007) Mesenchymal stromal cells. Biology of adult mesenchymal stem cells: regulation of niche, self-renewal and differentiation. *Arthritis Res Ther* **9:** 204
16. Gabius HJ, André S, Jiménez-Barbero J, Romero A, Solís D (2011) From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem Sci.* **36**: 298-313
17. Amano M, Eriksson H, Manning JC, Detjen KM, André S, Nishimura S, Lehtiö J, Gabius HJ (2012) Tumour suppressor p16(INK4a) - anoikis-favouring decrease in N/O-glycan/cell surface sialylation by down-regulation of enzymes in sialic acid biosynthesis in tandem in a pancreatic carcinoma model. *FEBS J.* **279**: 4062-4080
18. André S, Sanchez-Ruderisch H, Nakagawa H, Buchholz M, Kopitz J, Forberich P, Kemmner W, Böck C, Deguchi K, Detjen KM, Wiedenmann B, von Knebel Doeberitz M, Gress TM, Nishimura S, Rosewicz S, Gabius HJ (2007) Tumor suppressor p16INK4a--modulator of glycomic profile and galectin-1 expression to increase susceptibility to carbohydrate-dependent induction of anoikis in pancreatic carcinoma cells. *FEBS J.* **274**: 3233-3256
19. Urita A, Matsuhashi T, Onodera T, Nakagawa H, Hato M, Amano M, Seito N, Minami A,

Nishimura S-I, Iwasaki N (2011) Alterations of high-mannose type N-glycosylation in human and mouse osteoarthritis cartilage. *Arthritis Rheum* **63**: 3428-3438

20. Pabst M, Wu SQ, Grass J, Kolb A, Chiari C, Viernstein H, Unger FM, Altmann F, Toegel S (2010) IL-1beta and TNF-alpha alter the glycophenotype of primary human chondrocytes in vitro. *Carbohydr Res.* **345**: 1389-1393
21. Cordwell SJ, Thingholm TE (2010) Technologies for plasma membrane proteomics. *Proteomics* **10:** 611-627
22. Zhang H, Li XJ, Martin DB, Aebersold R (2003) Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* **21:** 660–666
23. Chen R, Jiang XN, Sun DG, Han GH, Wang FJ, Ye ML, Wang LM, Zou HF (2009) Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J Proteome Res* **8:** 651-661
24. Kullolli M, Hancock WS, Hincapie M (2008) Preparation of a high-performance multi-lectin affinity chromatography (HP-M-LAC) adsorbent for the analysis of human plasma glycoproteins. *J Sep Sci* **31:** 2733–2739
25. Nishimura S-I (2011) Toward automated glycan analysis. *Adv Carbohydr Chem Biochem* **65**: 219-271.
26. Nishimura S-I, Niikura K, Kurogochi M, Matsushita T, Fumoto M, Hinou H, Kamitani R, Nakagawa H, Deguchi K, Miura N, Monde K, Kondo H (2004) High-throughput protein glycomics: combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew Chem Int Ed* **44:** 91-96
27. Kita Y, Miura Y, Furukawa J, Nakano M, Shinohara Y, Ohno M, Takimoto A, Nishimura S-I (2007) Quantitative glycomics of human whole serum glycoproteins based on the standardized protocol for liberating *N*-glycans. *Mol Cell Proteomics* **6:** 1437-1445
28. Miura Y, Hato M, Shinohara Y, Kuramoto H, Furukawa JI, Kurogochi M, Shimaoka H,

M, Nakanishi K, Ozaki M, Todo S, Nishimura S-I (2008) BlotGlycoABCTM, an integrated glycoblotting technique for rapid and large scale clinical glycomics. *Mol Cell Proteomics* **7:** 370-377

29. Amano M, Yamaguchi M, Takegawa Y, Yamashita T, Terashima M, Furukawa J, Miura Y, Shinohara Y, Iwasaki N, Minami A, Nishimura S-I (2010) Threshold in stage-specific embryonic glycotypes uncovered by a full portrait of dynamic *N*-glycan expression during cell differentiation. *Mol Cell Proteomics* **9:** 523-537
30. Sharif M, Kirwan JR, Elson CJ, Granell R, Clarke S (2004) Suggestion of nonlinear or phasic progression of knee osteoarthritis based on measurements of serum cartilage oligomeric matrix protein levels over five years. *Arthritis Rheum* **50**:2479-2488.
31. Vilim V, Olejarova M, Machacek S, Gatterova J, Kraus VB, Pavelka K (2002) Serum levels of cartilage oligomeric matrix protein (COMP) correlate with radiographic progression of knee osteoarthritis. *Osteoarthritis Cartilage* **10**:707-713.
32. Bruyere O, Collette J, Kothari M, Zaim S, White D, Genant H, Peterfy C, Burlet N, Ethgen D, Montague T, et al (2006) Osteoarthritis, magnetic resonance imaging, and biochemical markers: a one year prospective study. *Ann Rheum Dis* **65**:1050-1054.
33. Dam EB, Byrjalsen I, Karsdal MA, Qvist P, Christiansen C (2009) Increased urinary excretion of C-telopeptides of type II collagen (CTX-II) predicts cartilage loss over 21 months by MRI. *Osteoarthritis Cartilage* **17**:384-389.
34. Garnero P, Sornay-Rendu E, Arlot M, Christiansen C, Delmas PD (2004) Association between spine disc degeneration and type II collagen degradation in postmenopausal women: the OFELY study. *Arthritis Rheum* **50**:3137-3144.
35. Bruyere O, Collette JH, Ethgen O, Rovati LC, Giacovelli G, Henrotin YE, Seidel L, Reginster JY (2003) Biochemical markers of bone and cartilage remodeling in prediction of longterm progression of knee osteoarthritis. *J Rheumatol* **30**:1043-1050.
36. Pavelka K, Forejtova S, Olejarova M, Gatterova J, Senolt L, Spacek P, Braun M, Hulejova

M, Stovickova J, Pavelkova A (2004) Hyaluronic acid levels may have predictive value for the progression of knee osteoarthritis. *Osteoarthritis Cartilage* **12**:277-283.

37. Spector TD, Hart DJ, Nandra D, Doyle DV, Mackillop N, Gallimore JR, Pepys MB (1997) Low-level increases in serum C-reactive protein are present in early osteoarthritis of the knee and predict progressive disease. *Arthritis Rheum* **40**:723-727.
38. Sharif M, Shepstone L, Elson CJ, Dieppe PA, Kirwan JR (2000) Increased serum C reactive protein may reflect events that precede radiographic progression in osteoarthritis of the knee. *Ann Rheum Dis* **59**:71-74.
39. Ruiz-Romero C, Blanco FJ (2010) Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage* **18:** 500-509
40. Anderson NL, Anderson NG. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1**: 845-867
41. Echan LA, Tang HY, Ali-Khan N, Lee K, Speicher DW. (2005) Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* 5: 3292-3303
42. Schiess R, Wollscheid B, Aebersold R. (2009) Targeted proteomic strategy for clinical biomarker discovery. *Mol. Oncol.* **3**: 33-44.
43. Zhang H, Li XJ, Martin DB, Aebersold R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **6**: 660-666.
44. Lee HJ, Na K, Kwon MS, Kim H, Kim KS, Paik YK. (2009) Quantitative analysis of phosphopeptides in search of the disease biomarker from the hepatocellular carcinoma specimen. Proteomics **9**: 3395-3408.
45. Shental-Bechor D, Levy Y (2009) Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. *Curr Opin Struct Biol* **19**: 524-533
46. Mitra N, Sinha S, Ramya TN, Surolia A (2006) N-linked oligosaccharides as outfitters for

glycoprotein folding, form and function. *Trends Biochem Sci*. **31**: 156-163

47. Bast RC Jr, Klug TL, St John E, Jenison E, Niloff JM, Lazarus H, Berkowitz RS, Leavitt T, Griffiths CT, Parker L, Zurawski VR Jr, Knapp RC. (1983) A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. *N. Engl. J. Med.* **309**: 883-887.
48. Kath R, Höffken K, Otte C, Metz K, Scheulen ME, Hülskamp F, Seeber S. (1993) The neu-oncogene product in serum and tissue of patients with breast carcinoma. *Ann. Oncol.* **4**: 585-590.
49. Kuriyama M, Wang MC, Papsidero LD, Killian CS, Shimano T, Valenzuela L, Nishiura T, Murphy GP, Chu TM. (1980) Quantitation of prostate-specific antigen in serum by a sensitive enzyme immunoassay. *Cancer Res.* **40**: 4658-4662
50. Miyoshi E, Noda K, Yamaguchi Y, Inoue S, Ikeda Y, Wang W, Ko JH, Uozumi N, Li W, Taniguchi N. (1999) The alpha1-6-fucosyltransferase gene and its biological significance. *Biochim. Biophys. Acta.* **1473**: 9-20
51. Parekh RB, Dwek RA, Sutton BJ, Fernandes DL, Leung A, Stanworth D, Rademacher TW, Mizuochi T, Taniguchi T, Matsuta K, et al. (1985) Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* **316**: 452-457
52. Pratta MA, Tortorella MD, Arner EC (2000) Age-related changes in aggrecan glycosylation affect cleavage by aggrecanase. *J. Biol. Chem.* **275**: 39096-39102.

Chapter 2

Glycomics in chondrogenic differentiation

2-1. Abstract

Osteoarthritis (OA) is one of the most common chronic diseases among adults, especially the elderly, which is characterized by destruction of the articular cartilage. Despite affecting more than 100 million individuals all over the world, therapy is currently limited to treating pain, which is a principal symptom of OA. New approaches to the treatment of OA that induce regeneration and repair of cartilage are strongly needed. In this chapter, to discover potent markers for chondrogenic differentiation, glycoblotting-based comprehensive glycomic approach using high-throughput MALDI-TOF mass spectrometry were performed. As a result, 45 kinds of glycoforms were detected and quantified reproducibly and expression levels of high-mannose type *N*-glycans were up-regulated significantly at the late stage of differentiation of the mouse chondroprogenitor cells. In conclusion, we clarified that high-mannose type *N*-glycans increased specifically during chondrocyte maturation by quantitative monitoring of *N*-glycans during cell differentiation,.

2-2. Introduction

Protein glycosylation is one of the most important posttranslational modifications in cell surface proteins and extracellular matrix proteins, which is considered to have a variety of biological functions, including enhancement of protein stability, controlling cell-to-cell communication and adhesion[9]. Recently, it has been well documented that glycan modifications of proteins greatly contribute to the pathogenesis of many diseases[1,2]. One of the characteristics of cartilage is that chondrocytes exist in the extracellular matrix. It is also well known that glycoproteins are abundant on the cell surface and in cartilage extracellular matrix. Recently, we reported that some high-mannose type *N*-glycan levels are decreased significantly both in human OA cartilage and in degraded mouse cartilage[3], strongly suggesting an association between *N*-glycans of cell surface glycoproteins and the pathogenesis of OA. Pabst *et al.*[4] also demonstrated that the levels of glycophenotype of primary human chondrocytes were altered by inflammatory cytokines. These means clearly that proteins with such characteristic *N*-glycoforms might become novel chondrogenic cell markers. However, glycoproteins bearing high-mannose type *N*-glycans in cartilage and their functional roles in the process of cartilage degradation have not yet been uncovered.

We have developed a standardized protocol of glycomics based on the simple chemical enrichment method, namely glycoblotting method[5] that allows for rapid and large-scale enrichment analysis of human serum glycans[6,7]. We also recently demonstrated the versatility of this glycoblotting protocol for the monitoring and characterization of the processes of dynamic cellular differentiation of mouse embryonic carcinoma cells and mouse embryonic stem cells into cardiomyocytes or neural cells[8]. Here, we report *N*-glycans changes in chondrogenic differentiation by using the glycoblotting method.

2-3. Results

Entire N-glycan profiling during chondrogenic differentiation of ATDC5 cells.

We performed glycoblotting for comprehensive analysis of *N*-glycan alterations during chondrogenic differentiation of ATDC5 cells, a mouse chondroprogenitor cell line to serve as a common model for studying chondrogenic differentiation after insulin induction[9]. As shown in Figure 2-1a, ATDC5 cells were grown to confluence and formed into cartilage nodules through cellular condensation in the presence of insulin (10 μg/mL). To confirm the maturation of cartilage nodules, Alcian blue staining, a representative method to detect cartilage matrix, was performed. As shown in Figure 2-1b, Alcian blue-positive cells were observed on Day 7 and increased in number after Day 14. Moreover, the levels of expression of the major cartilage matrix proteins, type II collagen and aggrecan, were confirmed by quantitative polymerase chain reaction (qPCR) (Figure 2-1c). Both were increased after Day 3 and peaked at Day 7-11. As described above, we confirmed that ATDC5 cells could be differentiated into mature chondrocytes after Day 7 of induction with insulin.

Whole *N*-glycans obtained from ATDC5 cells with or without induction into chondrocytes were profiled by glycoblotting-based high-throughput MALDI-TOF mass spectrometry (Figure 2-2). As summarized in Table 2-I, 45 kinds of glycoforms were detected and quantified reproducibly. High-mannose type *N*-glycans were the major components with the peak expression twice on Day 7 compared with undifferentiated state (Figure 2-2b, Table 2-II). The expression pattern of high-mannose type *N*-glycan was similar to the results of analysis of chondrogenic differentiation markers (Figure 2-1b-c). These findings clearly suggested that expression levels of high-mannose type *N*-glycans were significantly accompanied by chondrogenic differentiation. Figures 2-2c-g show the expression level of each high-mannose type *N*-glycan. The amount of $Hex_5(HexNAc)_2$ increased by fivefold throughout chondrogenic differentiation and retained at the elevated level on mature chondrocytes (Figure 2-2c). Those of $Hex_6(HexNAc)_2$, $Hex_7(HexNAc)_2$, and $Hex_8(HexNAc)_2$ increased and peaked on Day 7 (Figure

2-2d-f). On the other hand, $Hex_9(HexNAc)_2$ was downregulated during maturation into chondrocytes. These results indicated that conversion of $Hex_9(HexNAc)_2$ into $Hex_5(HexNAc)_2$ was promoted and that $Hex_5(HexNAc)_2$ accumulated eventually in mature chondrocytes while other high-mannose type *N*-glycanss decreased. We next confirmed the gene expression level of glycohydrolases and glycosyltransferases related to the synthesis of a series of *N*-glycans (Table 2-III, Figure 2-3). The levels of mRNA for *Ganab* and *Ugcgl 1*, which are enzymes involved in converting into $Hex_8(HexNAc)_2$, were upregulated and peaked on Day 3. However, those of *Man1a* and *Man1c*, which convert $Hex_8(HexNAc)_2$ into $Hex_5(HexNAc)_2$, were gradually downregulated. *Mgat1*, which is a key enzyme responsible for switching from high-mannose type *N*-glycans to other hybrid/complex type *N*-glycans, was also downregulated. These results of qPCR analysis showed that both trimming of mannose residues of high-mannose type *N*-glycans and conversion from these glycans into other *N*-glycan types were suppressed during chondrogenic differentiation following activation of biosynthesis of the high-mannose type *N*-glycans at the early stages of differentiation. This finding was strongly correlated with those of quantitative *N*-glycan analysis (Figure 2-2).

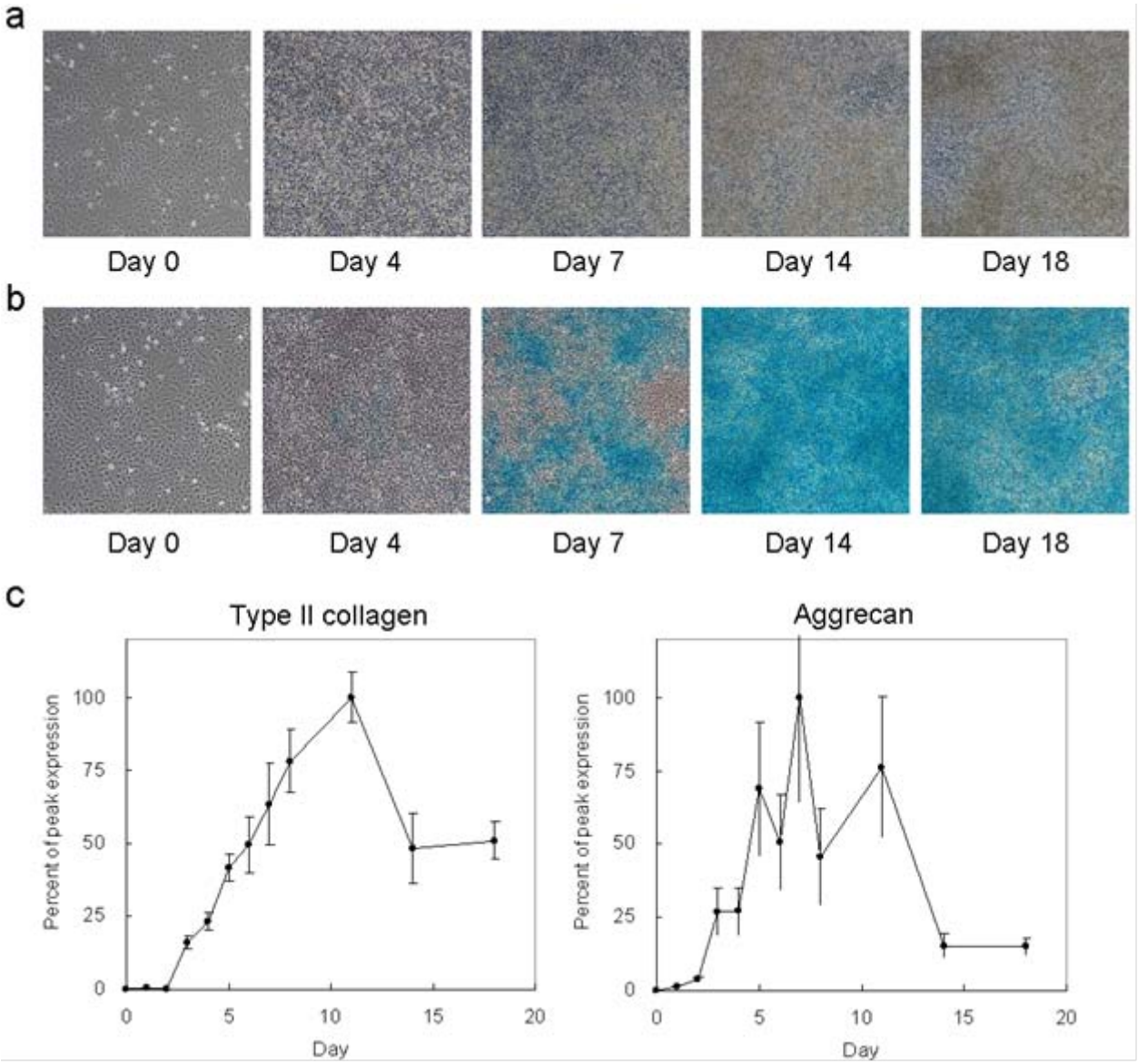
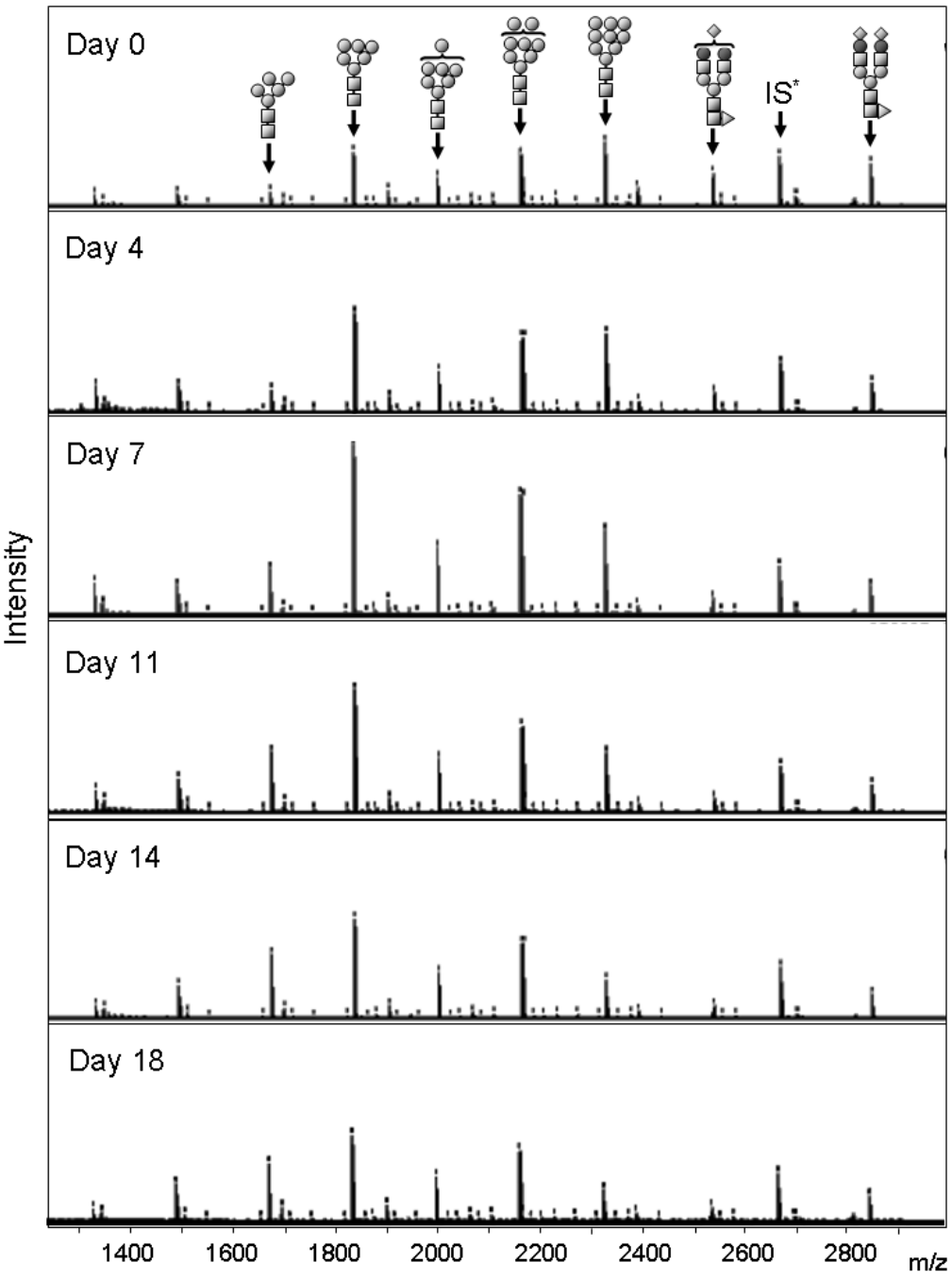


Figure 2-1. Chondrogenic differentiation of ATDC5 cells. (a) ATDC5 cells formed cartilage nodules through cellular proliferation and condensation after induction with insulin. (b) Cultured ATDC5 cells were stained with Alcian blue. The Alcian blue-positive area was observed on Day 7 and increased after Day 14. (c) Type II collagen and aggrecan expression were analyzed by qPCR in biological triplicate. Levels of both transcripts increased after the cells reached confluence and peaked at Day 7-11.

a


Day 0
IS*
Day 4
Day 7
Day 11
Day 14
Day 18
Intensity
1400
1600
1800
2000
2200
2400
2600
2800
m/z

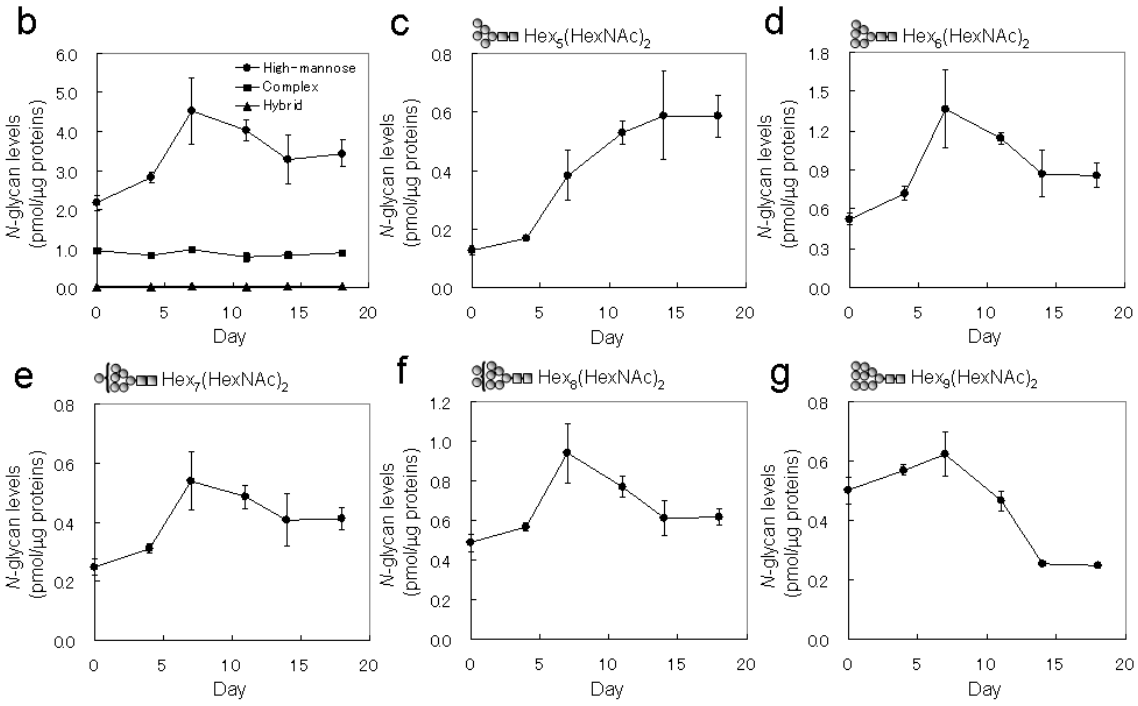


Figure 2-2. *N*-glycan profiling by glycoblotting during chondrogenic differentiation of ATDC5 cells. (a) All samples derived from Day 0 to Day 18 after induction with insulin were measured by MALDI-TOF-MS in biological triplicate. Forty-five glycans shown in Table 2-I were profiled. (b) Relative expression levels of high-mannose, hybrid, and complex *N*-glycan subtypes are shown. High mannose type *N*-glycans are major components and the expression levels increased throughout chondrogenic differentiation. (c – g) Relative expression levels of each high-mannose type *N*-glycan are shown. The signal intensity was normalized relative to an internal standard.

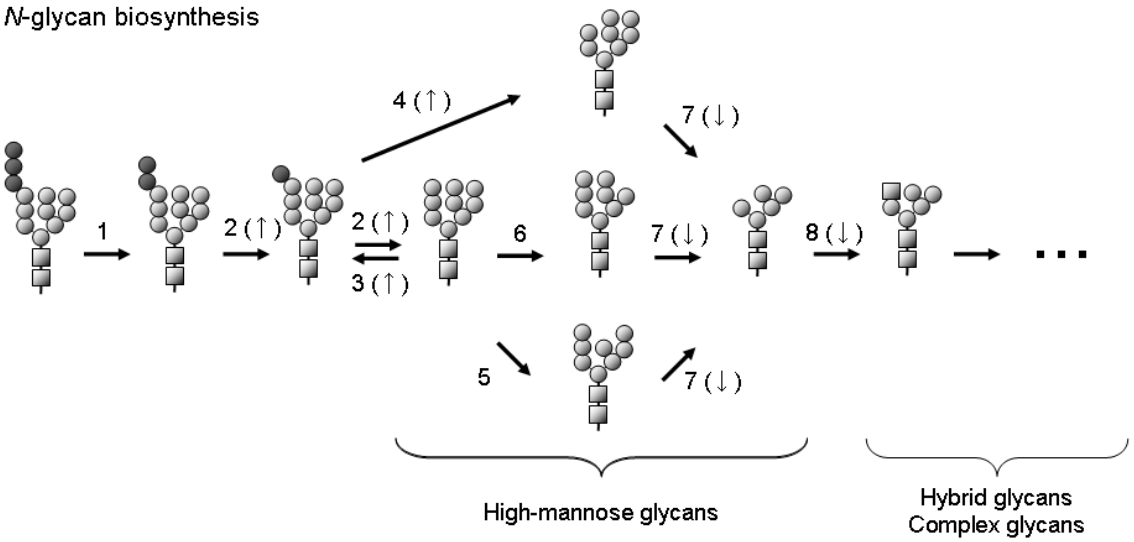


Figure 2-3. Biosynthesis of high-mannose type *N*-glycans according to alterations of the glycohydrolases and glycosyltransferases. The levels of mRNA for the glycosyltransferases during chondrogenic differentiation by microarray analysis were shown in Table 2-III. Up- and downregulated glycosyltransferases are indicated by upward and downward pointing arrows, respectively. The numbers in figure correspond to those of each glycosyltransferase shown in Table 2-III.

Table 2-I: Glycoforms detected during chondrogenic differentiation of ATDC5 cells and averages and standard deviation of the mass peaks (n=3).

Hex, hexose; dHex, deoxyhexose; HexNAc, *N*-acetylhexosamine; NeuAc, *N*-acetylneuraminic acid.

			Day 0		Day 4		Day 7		Day 11		Day 14		Day 18	
Peak no.	m/z	Composition	Ave.	SD	Ave.	SD	Ave.	SD	Ave.	SD	Ave.	SD	Ave.	SD
1	1324.55	Hex2(HexNAc)2dHex1	0.11716	0.02061	0.20729	0.02235	0.29314	0.05589	0.19458	0.01472	0.11807	0.02509	0.14051	0.02271
2	1340.55	Hex3(HexNAc)2	0.03498	0.00561	0.05888	0.00178	0.08700	0.02082	0.09654	0.01195	0.08286	0.02674	0.10396	0.01903
3	1486.61	Hex3(HexNAc)2dHex1	0.12553	0.01804	0.21283	0.01143	0.26404	0.06566	0.29328	0.03224	0.29992	0.04942	0.39738	0.04356
4	1502.6	Hex4(HexNAc)2	0.01337	0.00327	0.01994	0.00344	0.04060	0.00679	0.05580	0.01046	0.05269	0.01676	0.06153	0.00568
5	1543.63	Hex3(HexNAc)3	0.00642	0.00113	0.01049	0.00166	0.01123	0.00370	0.01031	0.00197	0.01053	0.00132	0.01422	0.00250
6	1648.66	Hex4(HexNAc)2dHex1	0.00156	0.00023	0.00301	0.00095	0.00724	0.00154	0.00898	0.00148	0.01132	0.00153	0.01790	0.00222
7	1664.65	Hex5(HexNAc)2	0.13170	0.01560	0.17374	0.00792	0.39575	0.08601	0.54641	0.03993	0.60708	0.15564	0.60508	0.07286
8	1689.69	Hex3(HexNAc)3dHex1	0.03789	0.00434	0.05941	0.00320	0.07845	0.01000	0.07924	0.01726	0.09919	0.02116	0.11715	0.00817
9	1705.68	Hex4(HexNAc)3	0.00608	0.00118	0.00579	0.00017	0.00502	0.00068	0.00387	0.00056	0.00429	0.00115	0.00592	0.00037
10	1746.71	Hex3(HexNAc)4	0.00949	0.00110	0.00721	0.00044	0.00849	0.00130	0.00677	0.00120	0.00959	0.00228	0.01196	0.00114
11	1810.71	Hex5(HexNAc)2dHex1	0.00155	0.00035	0.00312	0.00037	0.00680	0.00116	0.00663	0.00083	0.00947	0.00218	0.01529	0.00143
12	1826.71	Hex6(HexNAc)2	0.54122	0.04942	0.74513	0.05619	1.40941	0.30964	1.17767	0.04494	0.89772	0.18278	0.88515	0.09720
13	1851.74	Hex4(HexNAc)3dHex1	0.01651	0.00232	0.01483	0.00045	0.01092	0.00036	0.01067	0.00142	0.01146	0.00136	0.01526	0.00020
14	1867.73	Hex5(HexNAc)3	0.00390	0.00098	0.00379	0.00119	0.00691	0.00108	0.00502	0.00116	0.00521	0.00110	0.00718	0.00031
15	1892.76	Hex3(HexNAc)4dHex1	0.11877	0.01099	0.11005	0.00835	0.13590	0.01612	0.10542	0.01359	0.12070	0.02487	0.13720	0.01092
16	1908.76	Hex4(HexNAc)4	0.00424	0.00108	0.00361	0.00061	0.00548	0.00074	0.00410	0.00032	0.00493	0.00070	0.00619	0.00043
17	1949.79	Hex3(HexNAc)5	0.00396	0.00068	0.00837	0.00066	0.01053	0.00169	0.00597	0.00086	0.00527	0.00078	0.00637	0.00123
18	1988.76	Hex7(HexNAc)2	0.25650	0.02627	0.32023	0.01774	0.55759	0.10041	0.50136	0.04154	0.41968	0.09126	0.42650	0.03923
19	2010.79	Hex4(HexNAc)3(NeuAc)1	0.00554	0.00112	0.00485	0.00059	0.00668	0.00009	0.00632	0.00140	0.00674	0.00164	0.00867	0.00014
20	2029.79	Hex6(HexNAc)3	0.00522	0.00050	0.00635	0.00053	0.01114	0.00095	0.00852	0.00199	0.00996	0.00132	0.01189	0.00050
21	2054.82	Hex4(HexNAc)4dHex1	0.03545	0.00204	0.03683	0.00192	0.04615	0.00303	0.03797	0.00647	0.04398	0.00806	0.04798	0.00482
22	2070.81	Hex5(HexNAc)4	0.01647	0.00234	0.00731	0.00043	0.01014	0.00027	0.00677	0.00150	0.00676	0.00038	0.00931	0.00043
23	2095.84	Hex3(HexNAc)5dHex1	0.00471	0.00013	0.00787	0.00137	0.01216	0.00124	0.00959	0.00155	0.01190	0.00114	0.01632	0.00265

Table 2-I: (continued)

			Day 0		Day 4		Day 7		Day 11		Day 14		Day 18	
Peak no.	m/z	Composition	Ave.	SD	Ave.	SD	Ave.	SD	Ave.	SD	Ave.	SD	Ave.	SD
24	2150.81	Hex8(HexNAc)2	0.50339	0.04494	0.58532	0.02284	0.97029	0.15495	0.79510	0.05405	0.63223	0.09095	0.63611	0.04195
25	2156.85	Hex4(HexNAc)3dHex1(NeuAc)1	0.01853	0.00213	0.02146	0.00306	0.03169	0.00122	0.02917	0.00545	0.03271	0.00430	0.03927	0.00328
26	2172.84	Hex5(HexNAc)3(NeuAc)1	0.00823	0.00204	0.01102	0.00124	0.01324	0.00152	0.00929	0.00138	0.01079	0.00207	0.01257	0.00040
27	2191.84	Hex7(HexNAc)3	0.01314	0.00047	0.01486	0.00146	0.02645	0.00295	0.01922	0.00354	0.01948	0.00090	0.02096	0.00429
28	2216.87	Hex5(HexNAc)4dHex1	0.05153	0.00373	0.02922	0.00234	0.04014	0.00114	0.02955	0.00532	0.03563	0.00324	0.03753	0.00197
29	2257.9	Hex4(HexNAc)5dHex1	0.00255	0.00092	0.00742	0.00097	0.02035	0.00194	0.01769	0.00220	0.02068	0.00190	0.02357	0.00472
30	2298.92	Hex3(HexNAc)6dHex1	0.00268	0.00081	0.00423	0.00161	0.01227	0.00286	0.01064	0.00137	0.01482	0.00271	0.02297	0.00361
31	2312.86	Hex9(HexNAc)2	0.51816	0.04682	0.58736	0.01775	0.64384	0.07614	0.48181	0.03534	0.26269	0.00854	0.25648	0.01176
32	2318.9	Hex5(HexNAc)3dHex1(NeuAc)1	0.00519	0.00026	0.01156	0.00149	0.01562	0.00051	0.01231	0.00157	0.01174	0.00117	0.01608	0.00128
33	2334.9	Hex6(HexNAc)3(NeuAc)1	0.01099	0.00208	0.02028	0.00205	0.02561	0.00113	0.01829	0.00341	0.02148	0.00326	0.02357	0.00076
34	2359.93	Hex4(HexNAc)4dHex1(NeuAc)1	0.00915	0.00076	0.01045	0.00166	0.01427	0.00080	0.01015	0.00223	0.01209	0.00197	0.01504	0.00217
35	2375.92	Hex5(HexNAc)4(NeuAc)1	0.11367	0.01270	0.07740	0.00313	0.07456	0.00294	0.05365	0.01167	0.05108	0.00424	0.05571	0.00193
36	2419.95	Hex5(HexNAc)5dHex1	0.00361	0.00072	0.00611	0.00130	0.01184	0.00127	0.01055	0.00125	0.01252	0.00126	0.01504	0.00230
37	2521.98	Hex5(HexNAc)4dHex1(NeuAc)1	0.19440	0.01187	0.12353	0.01296	0.12846	0.00593	0.09058	0.01689	0.08615	0.00539	0.07884	0.00719
38	2537.98	Hex6(HexNAc)4dHex1	0.02220	0.00290	0.01420	0.00146	0.01856	0.00056	0.01313	0.00377	0.01364	0.00205	0.01825	0.00153
39	2563.01	Hex4(HexNAc)5dHex1(NeuAc)1	0.00058	0.00009	0.00224	0.00033	0.00855	0.00038	0.00837	0.00154	0.01280	0.00128	0.01546	0.00392
40	2681.03	Hex5(HexNAc)4(NeuAc)2	0.04511	0.00420	0.03093	0.00496	0.03183	0.00182	0.02390	0.00586	0.02410	0.00383	0.02213	0.00255
41	2684.03	Hex6(HexNAc)4dHex1(NeuAc)1	0.05043	0.00478	0.03709	0.00626	0.04254	0.00270	0.03244	0.00531	0.03387	0.00222	0.03181	0.00217
42	2827.09	Hex5(HexNAc)4dHex1(NeuAc)2	0.19185	0.00654	0.17875	0.01998	0.19547	0.01037	0.16846	0.02752	0.15831	0.01778	0.13280	0.01812
43	3192.22	Hex6(HexNAc)5dHex1(NeuAc)2	0.00807	0.00117	0.00963	0.00705	0.00420	0.00168	0.00499	0.00355	0.00381	0.00089	0.00307	0.00062
44	3351.28	Hex6(HexNAc)5(NeuAc)3	0.00100	0.00018	0.00103	0.00036	0.00105	0.00046	0.00092	0.00043	0.00098	0.00063	0.00081	0.00013
45	3497.34	Hex6(HexNAc)5dHex1(NeuAc)3	0.00597	0.00197	0.00643	0.00233	0.00630	0.00101	0.00355	0.00128	0.00441	0.00133	0.00299	0.00092

Table 2-II: Averages and standard deviation of three major subclasses in *N*-glycans

		Day 0	Day 4	Day 7	Day 11	Day 14	Day 18
High-mannose	Ave.	2.175	2.826	4.530	4.028	3.288	3.435
	SD	0.201	0.127	0.845	0.271	0.628	0.336
Complex	Ave.	0.964	0.825	0.969	0.782	0.837	0.895
	SD	0.048	0.062	0.013	0.109	0.093	0.064
Hybrid	Ave.	0.019	0.029	0.042	0.031	0.036	0.041
	SD	0.003	0.002	0.002	0.006	0.005	0.001

Table 2-III: Gene expression levels of glycosyltransferases during chondrogenic differentiation by microarray analysis

		Fold change				
No.	Gene Symbol	Day 0	Day 3	Day 5	Day 7	Annotation
1	Gcs1	1	1.08	1.10	0.88	-
2	Ganab	1	2.04	1.86	1.52	↑ (upregulated)
	Prkcsh	1	1.28	1.36	1.15	-
3	Ugcgl 1	1	3.06	2.13	1.39	↑ (upregulated)
	Ugcgl 2	1	0.89	1.02	0.82	-
4	Manea	1	1.32	1.41	2.23	↑ (upregulated)
5	Man1b1	1	0.93	0.85	0.78	-
6	Man2c1	1	1.36	1.35	1.28	-
7	Man1a	1	0.46	0.59	0.57	↓ (downregulated)
	Man1a2	1	0.69	0.47	0.29	↓ (downregulated)
	Man1c1	1	0.97	0.75	0.44	↓ (downregulated)
8	Mgat1	1	0.78	0.67	0.64	↓ (downregulated)

2-4. Discussion

Comprehensive *N*-glycan profiling by a glycoblotting method was first performed on chondrogenic differentiation of mouse ATDC5 cells. By quantitative monitoring of the 45 *N*-glycans during cell differentiation, we clarified that high-mannose type *N*-glycans increased specifically during chondrocyte maturation (Figure 2-2, Table 2-I and Table 2-II). The increases in this type of *N*-glycans were corroborated by gene expression analysis of glycohydrolases and glycosyltransferases (Figure 2-3, Table 2-III). On the other hand, considering that ATDC5 cells differentiate into osteocytes after chondrogenic differentiation with insulin[10], it seems that the alterations of high-mannose type *N*-glycans observed here may be influenced by osteogenic differentiation. However, it was also reported that an elevation of alkaline phosphatase activity and expression of Type X collagen mRNA induced during osteogenic differentiation of ATDC5 cells are first observed during Days 16-20 after induction of insulin[10]. In the present study, the alterations of high-mannose type *N*-glycans were observed during Day 7-14, indicating that there must be little effect of osteogenic differentiation on the alteration of high-mannose type *N*-glycans. Moreover, Heiskanen *et al.*[11] demonstrated that the levels of high-mannose type *N*-glycans on osteoblasts differentiated from human MSCs were lower than those of human MSCs. In addition, we revealed previously that no significant change in the expression level of high-mannose type *N*-glycans was observed during mouse cardiac and neural differentiation while increase in other glycotypes such as complex type *N*-glycans carrying fucose or bisect-type GlcNAc residues were detected[22]. Therefore, it is concluded that the increases in high-mannose *N*-glycans observed here may have specificity in the chondrogenic differentiation. Interestingly, it was uncovered that the level of a key high-mannose type *N*-glycan, $Hex_5(HexNAc)_2$, decreased significantly in mouse and human OA cartilage[3]. As anticipated, the gene expression of *Mgat1*, which converts specifically this glycoform into other hybrid and/or complex type *N*-glycans, was significantly upregulated in degraded mouse cartilage, and mouse chondrocytes with suppressed *Mgat1* expression exhibited drastically reduced expression levels

of matrix metalloprotease 13 and aggrecanase 2 mRNA following stimulation with interleukin-1α[12]. These observations suggest clearly that the alteration of high-mannose type *N*-glycans is attributable to gene expression of *Mgat1*, which regulates the expression of key proteases, matrix metalloprotease 13 and aggrecanase 2, in a cartilage degradation process. Futher, Pabst et al. also demonstrated that the levels of high-mannose type *N*-glycans of primary human chondrocytes decrease under inflammatory cytokines, IL-1β and TNF-α[4]. Moreover, it was also demonstrated that mouse neural cells differentiated from embryonic stem cells induce upregulation of various bisect-type *N*-glycans during cell differentiation[8]. Taking into account all these results, it was strongly suggested that alteration and accumulation of high-mannose type *N*-glycans might have essential roles in the homeostatic and functional maintenance of chondrocytes and an association with the pathogenesis of OA.

2-5. Experimental section

Cell culture and chondrogenic differentiation

ATDC5 cells were obtained from RIKEN Cell Bank (Ibaraki, Japan) and cultured in a 1:1 mixture of Dulbecco's modified Eagle's medium (DMEM) and Ham's F-12 medium (Sigma, St. Louis, MO) containing 5% fetal bovine serum (FBS) (Life Technologies, Carlsbad, CA), 2 mM L-glutamine (Life Technologies), 100 units/mL penicillin (Life Technologies), and 100 μg/mL streptomycin (Life Technologies). The cells were seeded into 6-well tissue culture plates at a density of 1.0×10^5 cells and cultured in the above medium supplemented with 10 μg/mL transferrin (Roche Diagnostics Co., Basel, Switzerland). For induction of chondrogenesis, 10 μg/mL bovine insulin (Sigma) was reacted to the sub-confluent cells. Cells were maintained at 37°C in a humidified atmosphere of 5% CO_2 in air. The medium was replaced every other day.

Alcian blue staining

Accumulation of glycosaminoglycans associated with chondrocyte differentiation was assessed by staining with Alcian blue. Differentiating ATDC5 cells were rinsed with PBS (Life Technologies) and fixed with 95% methanol (Wako Pure Chemical, Osaka, Japan) for 20 min. They were then stained with 0.1% Alcian blue 8GX (Merck Chemicals, Darmstadt, Germany) in 0.1 M HCl (Nacalai Tesque, Kyoto, Japan) overnight.

Real-time qPCR

Total RNA was prepared from ATDC5 cells and human MSCs cultured in differentiation medium using TRIzol reagent (Life Technologies) according to the manufacturer's instructions. After reverse transcription using SuperScript III First-Strand Synthesis SuperMix (Life Technologies), real-time qPCR was performed with the 7500 Real-Time PCR System using Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA).

Glycoblotting-based quantitative cellular N-glycomics

Release of total *N*-glycans was carried out directly using whole-cell lysates as follows. After inducing differentiation, ATDC5 cells were scraped in PBS containing 10 mM EDTA (Wako Pure Chemical) at various time points. The scraped cells were washed three times with PBS containing 10 mM EDTA and lysed by incubation with 1% Triton X-100 for 1 h on ice. The lysates were centrifuged at 15000 rpm for 10 min at 4°C, and the obtained supernatant was added to cold acetone (1:4) to precipitate proteins. The precipitates were collected by centrifugation at 12000 rpm for 15 min at 4°C followed by serial washing with acetonitrile. The resulting precipitates were dissolved in 50 μL of 100 mM ammonium bicarbonate containing 0.04% of 1-propanesulfonic acid, 2-hydroxyl-3-myristamido followed by incubation at 60°C for 10 min. The solubilized proteinaceous materials were reduced by 10 mM DTT at 60°C for 30 min followed by alkylation with 20 mM iodoacetamide in the dark at room temperature for 30 min. The mixture was then treated with 800 units of trypsin (Sigma-Aldrich, Cat. No. T0303, St. Louis, MO) at 37°C overnight followed by heat inactivation of the enzyme at 90°C for 10 min. After cooling to room temperature, *N*-glycans of glycopeptides were released from trypsin-digested samples by incubation with 2.5 units of peptide-*N*-glycosidase F (Roche Applied Science, Basel, Switzerland) at 37°C overnight. The sample mixture was then dried by vacuum centrifugation and stored at –20°C until use.

Glycoblotting of *N*-glycans was performed according to the procedure described previously[6]. BlotGlyco H beads (50 μL) (10 mg/mL suspension; Sumitomo Bakelite Co., Ltd.) were aliquoted onto the wells of a MultiScreen Solvinert filter plate (Millipore, Billerica, MA). Peptide-*N*-glycosidase F-digested samples were added with an internal standard (A2 amide glycan, $Hex_5(HexNAc)_4(NeuAcAmide)_2$, prepared from egg yolks) to the wells followed by addition of 360 μL of 2% acetic acid in acetonitrile. The plates was incubated at 80°C for 100 min to specifically capture total glycans in sample mixtures onto beads via stable hydrazone bonds. The plate was washed with 200 μL of 2 M guanidine HCl in ammonium bicarbonate

followed by washing with the same volume of water and 1% triethylamine in methanol (MeOH). Each washing step was performed twice. Unreacted hydrazide functional groups on beads were capped by incubation with 10% acetic anhydride in MeOH for 30 min at room temperature. Then, the solvent was removed under vacuum, and the beads were serially washed with 2 × 200 μL of 10 mM HCl, MeOH, and dioxane, respectively. On-bead methyl esterification of carboxyl groups in sialic acids[12] was carried out by incubation with 100 mM 3-methyl-1-*p*-tolyltriazene in dioxane at 60°C for 80 min. Then, the beads were serially washed with 200 μL of dioxane, water, MeOH, and water. The glycans blotted on beads were subjected to *trans*-iminization reaction with N^{α}-((aminooxy)acetyl)tryptophanylarginine methyl ester (aoWR, aminooxy-functionalized peptide reagent) for 70 min at 80°C. WR-tagged glycans were eluted by adding 100 μL of water and then purified on a Mass PREP™ hydrophilic interaction chromatography (HILIC) μElution Plate (Waters, Manchester, UK) according to the manufacturer's protocol. Each purified aoWR-derivatized sample was dried and diluted with 5 μL of 2,5-dihydroxybenzoic acid (DHB; 10 mg/mL in 30% acetonitrile; Bruker Daltonics, Bremen, Germany), and an aliquot (1 μL) was deposited on the stainless steel target plate. The analytes were then subjected to MALDI-TOF-MS analysis using an Ultraflex time-of-flight mass spectrometer II (Bruker Daltonics) in reflector, positive ion mode typically summing 1000 shots. The detected *N*-glycan peaks in MALDI-TOF-MS spectra were picked using the software FlexAnalysis version 2.0 (Bruker Daltonics). The intensities of the isotopic peaks of each glycan were normalized relative to 1.5 pmol of internal standard (A2 amide glycan) in each status. The glycan structures were estimated by input of peak masses into the GlycoMod Tool (Swiss Institute of Bioinformatics) and GlycoSuite (Proteome Systems). All glycomic analyses were performed three times.

2-6. References

1. Amano M, Eriksson H, Manning JC, Detjen KM, André S, Nishimura S, Lehtiö J, Gabius HJ(2012) Tumour suppressor p16(INK4a) - anoikis-favouring decrease in N/O-glycan/cell surface sialylation by down-regulation of enzymes in sialic acid biosynthesis in tandem in a pancreatic carcinoma model. *FEBS J.* **279**: 4062-4080
2. André S, Sanchez-Ruderisch H, Nakagawa H, Buchholz M, Kopitz J, Forberich P, Kemmner W, Böck C, Deguchi K, Detjen KM, Wiedenmann B, von Knebel Doeberitz M, Gress TM, Nishimura S, Rosewicz S, Gabius HJ(2007) Tumor suppressor p16INK4a--modulator of glycomic profile and galectin-1 expression to increase susceptibility to carbohydrate-dependent induction of anoikis in pancreatic carcinoma cells. *FEBS J.* **274**: 3233-3256
3. Urita A, Matsuhashi T, Onodera T, Nakagawa H, Hato M, Amano M, Seito N, Minami A, Nishimura S-I, Iwasaki N (2011) Alterations of high-mannose type N-glycosylation in human and mouse osteoarthritis cartilage. *Arthritis Rheum* **63**: 3428-3438
4. Pabst M, Wu SQ, Grass J, Kolb A, Chiari C, Viernstein H, Unger FM, Altmann F, Toegel S(2010) IL-1beta and TNF-alpha alter the glycophenotype of primary human chondrocytes in vitro. *Carbohydr Res.* **345**: 1389-1393
5. Nishimura S-I, Niikura K, Kurogochi M, Matsushita T, Fumoto M, Hinou H, Kamitani R, Nakagawa H, Deguchi K, Miura N, Monde K, Kondo H (2004) High-throughput protein glycomics: combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew Chem Int Ed* **44:** 91-96
6. Kita Y, Miura Y, Furukawa J, Nakano M, Shinohara Y, Ohno M, Takimoto A, Nishimura S-I (2007) Quantitative glycomics of human whole serum glycoproteins based on the standardized protocol for liberating *N*-glycans. *Mol Cell Proteomics* **6:** 1437-1445
7. Miura Y, Hato M, Shinohara Y, Kuramoto H, Furukawa JI, Kurogochi M, Shimaoka H, Tada M, Nakanishi K, Ozaki M, Todo S, Nishimura S-I (2008) BlotGlycoABCTM, an integrated

glycoblotting technique for rapid and large scale clinical glycomics. *Mol Cell Proteomics* **7:** 370-377

8. Amano M, Yamaguchi M, Takegawa Y, Yamashita T, Terashima M, Furukawa J, Miura Y, Shinohara Y, Iwasaki N, Minami A, Nishimura S-I (2010) Threshold in stage-specific embryonic glycotypes uncovered by a full portrait of dynamic *N*-glycan expression during cell differentiation. *Mol Cell Proteomics* **9:** 523-537
9. Atsumi T, Miwa Y, Kimata K, Ikawa Y (1990) A chondrogenic cell line derived from a differentiating culture of AT805 teratocarcinoma cells. *Cell Differ Dev* **30:** 109-116
10. Shukunami C, Ishizeki K, Atsumi T, Ohta Y, Suzuki F, Hiraki Y (1997) Cellular hypertrophy and calcification of embryonal carcinoma-derived chondrogenic cell line ATDC5 in vitro. *J. Bone Miner. Res.* **12:** 1174-1188
11. Heiskanen A, Hirvonen T, Salo H, Impola U, Olonen A, Laitinen A, Tiitinen S, Natunen S, Aitio O, Miller-Podraza H, Wuhrer M, Deelder AM, Natunen J, Laine J, Lehenkari P, Saarinen J, Satomaa T, Valmu L (2009) Glycomics of bone marrow-derived mesenchymal stem cells can be used to evaluate their cellular differentiation stage. *Glycoconj J* **26:** 367-384
12. Miura Y, Shinohara Y, Furukawa JI, Nagahori N, Nishimurara S-I (2007) Rapid and simple solid-phase esterification of sialic acid residues for quantitative glycomics by mass spectrometry. *Chem Eur J* **13**: 4797-4804

Chapter 3

Discovery of novel differentiation markers in chondrogenesis by glycoform-focused reverse proteomics and genomics

3-1. Abstract

In *Chapter 2,* we clarified that high-mannose type *N*-glycans increased specifically during chondrogenic differentiation. In this chapter, we discovered novel differentiation markers in chondrogenesis by glycoform-focused reverse proteomics and genomics. At first, among 246 glycoproteins carrying this glycotype identified by ConA affinity chromatography and LC/MS, it was demonstrated that 52% are classified as cell surface glycoproteins. Second, gene expression levels indicated that mRNAs for 15 glycoproteins increased distinctly in the earlier stages during differentiation compared with Type II collagen. Finally, the feasibility of mouse chondrocyte markers in human chondrogenesis model was demonstrated by testing gene expression levels of these 15 glycoproteins during differentiation in human mesenchymal stem cells. In Conclusion, the results showed clearly an evidence of up-regulation of 5 genes, ectonucleotide pyrophosphatase/phosphodiesterase family member 1, collagen alpha-1(III) chain, collagen alpha-1(XI) chain, aquaporin-1, and netrin receptor UNC5B, in the early stages of differentiation.

3-2. Introduction

One of the characteristics of cartilage is that chondrocytes exist in the extracellular matrix. It is also well known that glycoproteins are abundant on the cell surface and in cartilage extracellular matrix. In *chapter 2*, we described that alteration and accumulation of high-mannose type *N*-glycans might have essential roles in the homeostatic and functional maintenance of chondrocytes and an association with the pathogenesis of OA. Furthermore, we previously reported that some high-mannose type *N*-glycan levels are decreased significantly both in human OA cartilage and in degraded mouse cartilage[1], strongly suggesting an association between *N*-glycans of cell surface glycoproteins and the pathogenesis of OA. However, glycoproteins bearing high-mannose type *N*-glycans in cartilage and their functional roles in the process of chondrogenesis have not yet been uncovered.

Glycoproteomics is a new potential strategy among focused proteomics approaches for the characterization of plasma membrane proteins[2]. This approach has facilitated comprehensive identification of cell surface glycoproteins with low abundance and insolubility by hydrazide chemistry[3,4] and lectin affinity chromatography[5]. However, determination of the total glycan structures from whole cellular glycoproteins had long remained to be a challenging and extremely difficult task due to the lack of general platforms for high throughput glycomics[6].

Mesenchymal stem cells (MSCs) in various adult tissues, such as bone marrow, adipose tissue, and synovial fluid, have multipotency and self-renewal capabilities, including induction to undergo chondrogenic differentiation[7-11]. Especially, cells from the synovial membrane have the ability to differentiate into chondrocytes[9], indicating that they potentially have a physiological role in the repair and degeneration of the articular cartilage. Therefore, it is expected that medications induce proliferation and differentiation of MSCs *in vitro* and *in vivo* will contribute to repair and regeneration of damaged cartilage. Advent of new sensitive markers that contribute to development of novel therapeutic reagents to induce differentiation of MSCs are strongly required for improving quality of life of patients suffering OA.

In this chapter, we discovered novel chondrogenic differentiation markers through high throughput glycoproteomics uncovering proteins displaying characteristic glycan structures.

3-3. Results

Identification of high-mannose type N-glycopeptides

Driven by the observed unique *N*-glycan profile in chondrogenic differentiation of ATDC5 cells, proteins that carry high-mannose type oligosaccharides were investigated. Concanavalin A (ConA) was used as an affinity reagent to selectively recover the glycopeptides with high-mannose type oligosaccharides[12,13]. Following the enrichment of glycopeptides of interest in a crude tryptic digest of the chondrogenic differentiated cells, the ConA-bound fraction was further incubated with peptide-*N*-glycosidase F and examined by LC/MS to identify the peptides. The *N*-glycosylation sites were assessed by conversion of asparagine into aspartic acid residues at the position of glycan attachment caused by the peptide-*N*-glycosidase F-catalyzed deglycosylation reaction. The consensus amino acid sequence of *N*-glycosylation (Asn-X-Ser/Thr) was also confirmed. 246 *N*-glycoproteins and 434 *N*-glycopeptides (including 460 *N*-glycosylation sites) were identified (Table 3-II). Further, we analyzed the subcellular localization of the identified proteins using information from the UniProt database (http://www.uniprot.org/). Of the total of 246 identified proteins, 74 proteins (30%) were localized at the plasma membrane and 53 proteins (22%) were localized to the extracellular matrix or were secreted (Figure 3-1). In this study, glycoforms-focused reverse proteomics of the chondrogenic differentiated cells revealed many proteins carrying high-mannose type *N*-glycans. Interestingly, 52% of the identified glycoproteins were cell surface proteins, which were easily detectable by flow cytometry or immunochemistry, and may include some differentiation markers.

Gene expression analysis of the identified glycoproteins

Quantitative and qualitative alterations of high-mannose type *N*-glycans are significantly correlated to chondrogenic differentiation. To investigate whether expression of the identified high-mannose glycoproteins is regulated at the mRNA level during differentiation, we

performed gene expression analysis of the identified proteins. A total of 246 candidates were preliminary screened by microarray analysis (Table 3-III), and genes up-regulated at the earlier stages of chondrogenic differentiation compared with Type II collagen were selected. Then, we focused on cell surface proteins by using information from the UniProt database. Finally, qPCR were performed to validate the expression of the selected genes. The results identified 15 genes showing peak expression at the earlier stages of chondrogenic differentiation compared with Type II collagen (Figure 3-2, Table 3-I).

Gene expression analysis of 15 differentiation marker candidates during human chondrogenic differentiation

To evaluate the applicability of 15 marker candidates for human chondrogenic differentiation, gene expression analysis was performed in a human MSC model. MSCs have the ability to differentiate into cells of the chondrogenic lineage. Human MSCs were cultured and differentiated into chondrocytes. To confirm the chondrogenic differentiation of the cells, Alcian blue staining was performed. As shown in Figure 3-3a, Alcian blue-positive cells on extracellular matrices were observed after Day 7 of induction. Moreover, the expression of the major cartilage matrix proteins type II collagen was confirmed to be induction-dependent as well as in mouse ATDC5 by qPCR. In the human chondrogenesis model, five of the 15 genes—*Aqp1*, *Col3a1*, *Col11a1*, *Enpp1*, and *Unc5b*—showed distinct up-regulation of expression at the mRNA level in the earlier stage of differentiation compared with Type II collagen (Figure 3-3b). These results clearly indicated that these 5 genes (*Aqp1*, *Col3a1*, *Col11a1*, *Enpp1*, and *Unc5b*) are highly sensitive markers during human chondrocyte differentiation.

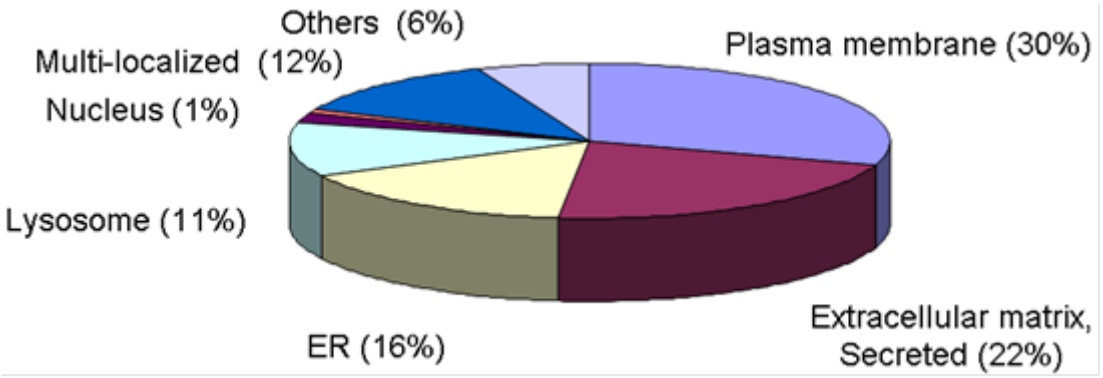


Figure 3-1. Subcellular locations of the 246 glycoproteins identified. The identified *N*-glycoproteins were classified based on localization in the cell using information from the UniProt database (http://www.uniprot.org/).

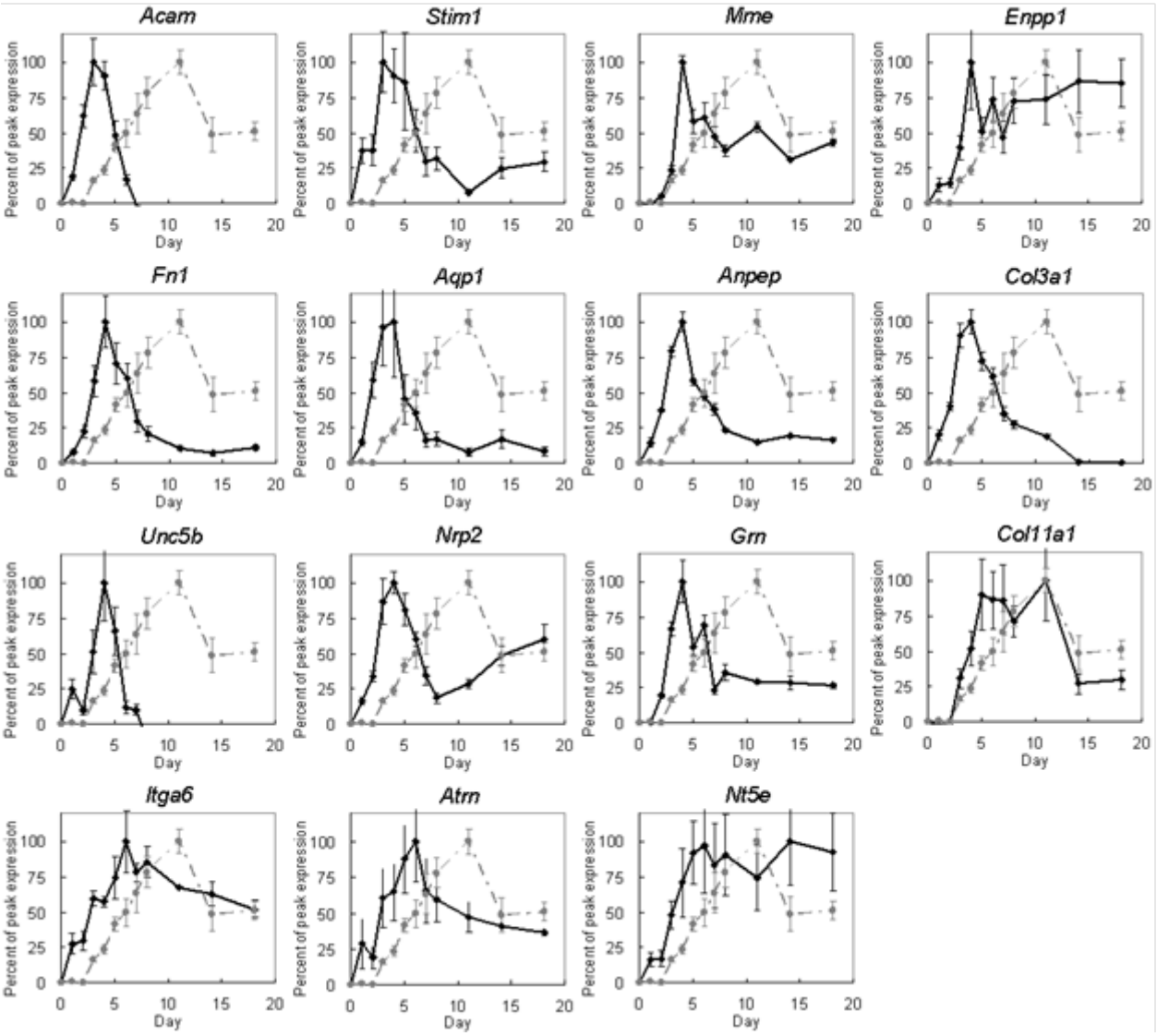


Figure 3-2. Gene expression profiles of 15 differentiation marker candidates with increasing transcript levels at the early stages of chondrogenic differentiation in ATDC5 cells. The solid lines indicate gene expression profiles of each of the 15 differentiation marker candidates and the dotted line indicates the gene expression profile of Type II collagen. The profiles were performed in biological triplicate.

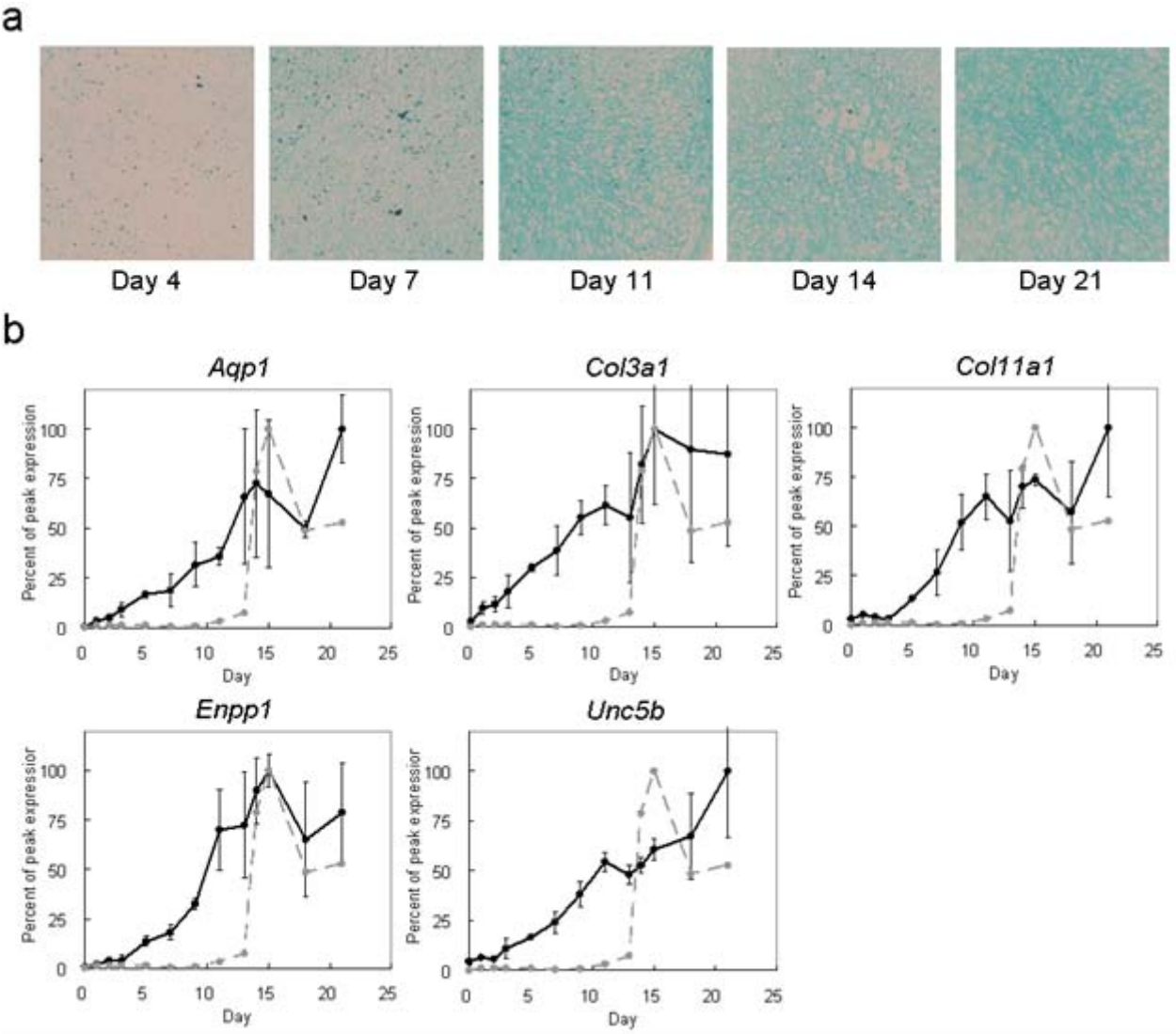


Figure 3-3. Chondrogenic differentiation of human MSCs and gene expression profiles of five differentiation marker candidates. (a) Human MSCs differentiated after induction with TGF-β3 were stained with Alcian blue. The Alcian blue-positive area was observed on Day 7 and increased after Day 11. (b) Gene expression levels of five differentiation marker candidates are shown according to the time after induction of differentiation in comparison with Type II collagen. The profiles were performed in biological triplicate. The solid lines indicate gene expression profiles of each of the five differentiation marker candidates and the dotted line indicates the gene expression profile of Type II collagen.

Table 3-I: Fifteen cell surface *N*-glycoproteins showing increased expression during chondrogenic differentiation

Gene Symbol	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]
Nt5e	5NTD_MOUSE	5^{-}-nucleotidase	LDNYSTQELGR	59	plasma membrane
Acam	ACAM_MOUSE	Adipocyte adhesion molecule	HVYNNLTEEQK	87	plasma membrane
Anpep	AMPN_MOUSE	Aminopeptidase N	FTCNQTTDVIIIHSK	57	plasma membrane
			KLNYTLK	53	
			LNYTLK	36	
			SGQEDHYWLDVEKNQSAK	68	
Aqp1	AQP1_MOUSE	Aquaporin-1	NQTLVQDNVK	45	plasma membrane
Atrn	ATRN_MOUSE	Attractin	IDSTGNVTNELR	107	plasma membrane
Col3a1	CO3A1_MOUSE	Collagen alpha-1(III) chain	LLSSRASQNITYHCK	57	Extracellular matrix, Secreted
Col11a1	COBA1_MOUSE	Collagen alpha-1(XI) chain	VYCNFTAGGETCIYPDK	109	Extracellular matrix,
			NTSEDILYGNK	45	Secreted
Enpp1	ENPP1_MOUSE	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1	VYNGSVPFEER	48	plasma membrane
Fn1	FINC_MOUSE	Fibronectin	DQCIVDDITYNVNDTFHK	114	Extracellular matrix,
			HEEGHMLNCTCFGQGR	67	Secreted
			LDAPTNLQFVNETDR	121	
			WTPLNSSTIIGYR	81	
			NYTDCTSEGR	71	
Grn	GRN_MOUSE	Granulins Grn	NYTTDLLTK	45	Secreted
Itga6	ITA6_MOUSE	Integrin alpha-6	LWNSTFLEEYSK	77	plasma membrane
Mme	NEP_MOUSE	Neprilysin	SCINESAIDSR	98	plasma membrane
			EIANATTKPEDR	72	
Nrp2	NRP2_MOUSE	Neuropilin-2	NFTSPNGTIESPGFPEK	85	plasma membrane
Stim1	STIM1_MOUSE	Stromal interaction molecule 1	LAVTNTTMTGTVLK	117	plasma membrane
Unc5b	UNC5B_MOUSE	Netrin receptor UNC5B	LSDTANYTCVAK	55	plasma membrane

[a] $P < 0.05$ by Mascot Search.

[b] Subcellular localizations of the identified proteins using information from the UniProt database (http://www.uniprot.org)

Table 3-II: The 246 proteins identified in chondrogenic differentiated ATDC5 cells

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
1	4F2_MOUSE	4F2 cell-surface antigen heavy chain	LMNAPLYLAEWQNITK	105	plasma membrane	CD98
2	5NTD_MOUSE	5~-nucleotidase	LDNYSTQELGR	59	plasma membrane	CD73
3	ACAM_MOUSE	Adipocyte adhesion molecule	HVYNNLTEEQK	87	plasma membrane	
4	ADA10_MOUSE	Disintegrin and metalloproteinase domain-containing protein 10	INTTSDEKDPTNPFR	65	plasma membrane	CD156c
			NISQVLEK	58		
5	ADAM9_MOUSE	Disintegrin and metalloproteinase domain-containing protein 9	NFSSCSAEDFEK	86	plasma membrane	
6	AEBP1_MOUSE	Adipocyte enhancer-binding protein 1	GVVTDEQGIPIANATISVSGINHGVK	34	Secreted, Nucleus, Cytoplasm	
7	AER61_MOUSE	Uncharacterized glycosyltransferase AER61	LNITQEGPK	53	Secreted	
8	AGAL_MOUSE	Alpha-galactosidase A	LGIYADVGNK	59	Lysosome	
			YMALALNR	56		
9	AMPN_MOUSE	Aminopeptidase N	FTCNOTTDVIIIHSK	57	plasma membrane	CD13
			KLNYTLK	53		
			SGQEDHYWLDVEKNQSAK	68		
10	AMRP_MOUSE	Alpha-2-macroglobulin receptor-associated protein	VIDLWDLAQSANFTEK	120	Cytoplasm, Endoplasmic reticulum	
11	ANO6_MOUSE	Anoctamin-6	LNITCESSK	67	plasma membrane	
12	ANPRC_MOUSE	Atrial natriuretic peptide clearance receptor	IIQQTWNR	64	plasma membrane	
13	ANTR1_MOUSE	Anthrax toxin receptor 1	DFNETQLAR	63	plasma membrane	
14	APMAP_MOUSE	Adipocyte plasma membrane-associated protein	AGPNGTLFVVDAYK	95	plasma membrane	
15	AQP1_MOUSE	Aquaporin-1	NOTLVQDNVK	45	plasma membrane	
16	ARSB_MOUSE	Arylsulfatase B	IYAGMVSLMDEAVGNVTK	125	Lysosome	
17	ASAH1_MOUSE	Acid ceramidase	NLSFATIYDVLSTKPVLNK	62	Lysosome	
			SVLENTTSYEEAK	98		
18	ASM3B_MOUSE	Acid sphingomyelinase-like phosphodiesterase 3b	WLGDVLSNASR	90	Secreted	
19	AT1B3_MOUSE	Sodium/potassium-transporting ATPase subunit beta-3	EENATIATYPEFGVLDLK	67	plasma membrane	CD298
			NLTSCPDGAPFIQHGPDYR	44		
20	ATRN_MOUSE	Attractin	IDSTGNVTNELR	107	plasma membrane	
21	BASI_MOUSE	Basigin	SQLTISNLDVNVDPGTYVCNATNAQGTTR	96	plasma membrane	CD147
			TQLTCSLNSSGVDIVGHR	132		
			TSDTGEEEAITNSTEANGK	147		
22	BGLR_MOUSE	Beta-glucuronidase	YGIVVIDECPGVGIVLPQSFGNESLR	149	Endoplasmic reticulum, Lysosome	
23	BTD_MOUSE	Biotinidase	FNDTEVLQR	79	Secreted	
24	CA2D1_MOUSE	Voltage-dependent calcium channel subunit alpha-2/delta-1	IDVNSWIENFTK	98	plasma membrane	
25	CADH2_MOUSE	Cadherin-2	SNISILR	47	plasma membrane	CD325
26	CALR_MOUSE	Calreticulin	SGTIFDNFLITNDEAYAEEFGNETWGVTK	77	Endoplasmic reticulum	
27	CALU_MOUSE	Calumenin	NATYGYVLDDPDPDDGFNYK	126	Endoplasmic reticulum, Secreted	
28	CATD_MOUSE	Cathepsin D	YYHGELSYLNVTR	82	Lysosome	
29	CATF_MOUSE	Cathepsin F	VYINDSVELSR	84	Lysosome	
30	CATL1_MOUSE	Cathepsin L1	AEFAVANDTGFVDIPQQEK	135	Lysosome	
31	CATZ_MOUSE	Cathepsin Z	ECHTIQNYTLWR	68	Lysosome	
32	CBPD_MOUSE	Carboxypeptidase D	FANEYPNITR	54	plasma membrane	
			NNSNNFDLNR	64		
33	CBPE_MOUSE	Carboxypeptidase E	DLQGNPIANATISVDGIDHDVTSAK	55	Secreted	
			GNETIVNLIHSTR	89		
34	CCD80_MOUSE	Coiled-coil domain-containing protein 80	LLGVGEEVGGVLELFPINGSSIVER	35	Extracellular matrix, Secreted	
35	CD109_MOUSE	CD109 antigen	QLNGSVIAK	38	plasma membrane	CD109
36	CD276_MOUSE	CD276 antigen	TALFPDLLVQGNASLR	92	plasma membrane	CD276
			VVLGANGTYSCLVR	108		
37	CD44_MOUSE	CD44 antigen	TEAADLCQAFNSTLPTMDQMK	77	plasma membrane	CD44

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
38	CD47_MOUSE	Leukocyte surface antigen CD47	DAMVGNYTCEVTELSR	119	plasma membrane	CD47
39	CERU_MOUSE	Ceruloplasmin	EYEGAVYPDNTTDFQR	65	Secreted	
40	CLDN1_MOUSE	Claudin domain-containing protein 1	YNGSLGLWR	59	unknown (membrane)	
41	CLP1L_MOUSE	Cleft lip and palate transmembrane protein 1-like protein	DLMVINR	38	plasma membrane	
			TVNVSVPK	70		
42	CLPT1_MOUSE	Cleft lip and palate transmembrane protein 1 homolog	DYYPINESLASLPLR	97	plasma membrane	
43	CNPY3_MOUSE	Protein canopy homolog 3	VVMDIPYELWNETSAEVADLK	129	Secreted	
44	CO1A1_MOUSE	Collagen alpha-1(I) chain	LMSTEASQNITYHCK	105	Extracellular matrix, Secreted	
45	CO1A2_MOUSE	Collagen alpha-2(I) chain	ASQNITYHCK	55	Extracellular matrix, Secreted	
46	CO2A1_MOUSE	Collagen alpha-1(II) chain	LLSTEGSQNITYHCK	114	Extracellular matrix, Secreted	
47	CO3A1_MOUSE	Collagen alpha-1(III) chain	LLSSRASQNITYHCK	57	Extracellular matrix, Secreted	
48	CO5A1_MOUSE	Collagen alpha-1(V) chain	KNVTLILDCK	90	Extracellular matrix, Secreted	
			VYCNFTAGGSTCVFPDK	107		
49	CO5A2_MOUSE	Collagen alpha-2(V) chain	EASQNLTYICR	77	Extracellular matrix, Secreted	
50	CO6A1_MOUSE	Collagen alpha-1(VI) chain	GEDGPPGNGTEGFPGFPGYPGNR	92	Extracellular matrix, Secreted	
			NITAQICIDK	76		
51	CO6A2_MOUSE	Collagen alpha-2(VI) chain	GTFTDCALANMTQQIR	129	Extracellular matrix, Secreted	
			NMTLFSDLVAEK	99		
52	COBA1_MOUSE	Collagen alpha-1(XI) chain	VYCNFTAGGETCIYPDK	109	Extracellular matrix, Secreted	
			NTSEDILYGNK	45		
53	COCA1_MOUSE	Collagen alpha-1(XII) chain	EAGNITTDGYEILGK	96	Extracellular matrix, Secreted	
			MLEAYNLTEK	78		
			NLQVYNATSNSLTVK	143		
54	COL12_MOUSE	Collectin-12	ETLQNNSFLITTVNK	114	plasma membrane	
			NFTILQGPPGPR	65		
			VQSLQTLAANNSALAK	61		
55	CREG1_MOUSE	Protein CREG1	VVTPEEYFNVTLQ	71	Secreted	
56	CREL2_MOUSE	Cysteine-rich with EGF-like domain protein 2	NETHSICSACDESCK	99	Secreted, Endoplasmic reticulum	
57	CRRY_MOUSE	Complement regulatory protein Crry	INYTCNQGYR	75	plasma membrane	
58	CRTAP_MOUSE	Cartilage-associated protein	NCSAATPAPAPAGPASHAELR	92	Extracellular matrix, Secreted	
			WGLSDEHFQPRPEAVQFFNVTTLQK	50		
59	CS063_MOUSE	UPF0510 protein C19orf63 homolog	GSEVEDEDLELFNTSVQLRPPSTAPGPETAAFIER	127	plasma membrane	
60	CTL2_MOUSE	Choline transporter-like protein 2	KNITDLVEGAK	83	plasma membrane	
			NESLQCPTAR	54		
61	CYTC_MOUSE	Cystatin-C	SQTNLTDCPFHDQPHLMR	64	Secreted	
62	DAG1_MOUSE	Dystroglycan	NCSSITLQNITR	100	Secreted	
63	DAPLE_MOUSE	Protein Daple	DPATLSRDGNTSGR	35	unknown	
64	DPP2_MOUSE	Dipeptidyl-peptidase 2	ALAGLVYNSSGTEPCYDIYR	87	Lysosome	
65	ECE1_MOUSE	Endothelin-converting enzyme 1	LGGWNITGPWAK	31	plasma membrane	
66	ELOV4_MOUSE	Elongation of very long chain fatty acids protein 4	GLLDSEPGSVLNAMSTAFNDTVEFYR	50	Endoplasmic reticulum (membrane)	
67	EMB_MOUSE	Embigin	YIINGSHANETR	78	plasma membrane	
			SCNISVTEK	52		
			YSNLSLK	36		
68	EMIL1_MOUSE	EMILIN-1	ESNSTSLTQAALLEK	96	Extracellular matrix, Secreted	
			FNSTLGPSEEQEK	68		
			LEGLLANVSR	80		
			LGALNNSLLLLEDR	117		
			LNLTAAQLSQLEGLLQAR	131		
69	ENPL_MOUSE	Endoplasmin	EEEAIQLDGLNASQIR	124	Endoplasmic reticulum	

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
			ELISNASDALDK	84		
			HNNDTQHIWESDSNEFSVIADPR	129		
			LGVIEDHSNR	52		
70	ENPP1_MOUSE	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1	VYNGSVPFEER	48	plasma membrane	
71	ENPP4_MOUSE	Ectonucleotide pyrophosphatase/phosphodiesterase family member 4	INVTEVYDK	65	plasma membrane	
72	ENPP5_MOUSE	Ectonucleotide pyrophosphatase/phosphodiesterase family member 5	VQPIVAVADEGWYILQNK	89	Plasma membrane, Secreted	
73	EPDR1_MOUSE	Mammalian ependymin-related protein 1	NYTVVMSTR	74	Secreted	
74	EPHA2_MOUSE	Ephrin type-A receptor 2	TASVSINQTEPPK	49	plasma membrane	
75	EPHB4_MOUSE	Ephrin type-B receptor 4	SVVHHLNGSTLR	46	plasma membrane	
76	ERLN1_MOUSE	Erlin-1	NYTADYDK	33	Endoplasmic reticulum (membrane)	
77	ERO1A_MOUSE	ERO1-like protein alpha	WGHNVTEFQQR	87	Endoplasmic reticulum (membrane)	
78	FGFR3_MOUSE	Fibroblast growth factor receptor 3	GNYTCVVENK	72	plasma membrane	CD333
79	FINC_MOUSE	Fibronectin	DQCIVDDITYNVNDTFHK	114	Extracellular matrix, Secreted	
			HEEGHMLNCTCFGQGR	67		
			LDAPTNLQFVNETDR	121		
			WTPLNSSTIIGYR	81		
			NYTDCTSEGR	71		
80	FKB10_MOUSE	FK506-binding protein 10	ITVPPHLAYGENGTGDK	41	Endoplasmic reticulum	
			NHTYNTYVGQGYIIPGMDQGLQGACIGER	74		
			TLSRPPENCNETSK	76		
			YHYNCSLLDGTR	80		
			YHYNGSLMDGTLFDSSYSR	136		
			YHYNGTFEDGKK	66		
81	FKBP7_MOUSE	FK506-binding protein 7	IEVLHRPENCSK	76	Endoplasmic reticulum	
82	FKBP9_MOUSE	FK506-binding protein 9	NHTFDTYIGQGYVIPGMDEGLLGVCIGER	69	Endoplasmic reticulum	
			YHYNASLLDGTLLDSTWNLGK	104		
			YHYNGTFLDGTLFDSSHNR	143		
			YHYNGTLLDGTLFDSSYSR	141		
83	FMOD_MOUSE	Fibromodulin	LYLDHNNLTR	46	Extracellular matrix, Secreted	
84	FPRP_MOUSE	Prostaglandin F2 receptor negative regulator	DLDLSCNITTDR	87	plasma membrane, Golgi apparatus, etc.	CD315
			LENWTDASR	35		
85	FSTL1_MOUSE	Follistatin-related protein 1	GSNYSEILDK	76	Secreted	
			FVEQNETAINITTYADQENNK	53		
86	FUCO_MOUSE	Tissue alpha-L-fucosidase	WGQNCSCHHGGYYNCQDK	47	Lysosome	
87	FUCO2_MOUSE	Plasma alpha-L-fucosidase	SQNDTVTPDVWYTSKPEK	49	Secreted	
88	FUT11_MOUSE	Alpha-(1,3)-fucosyltransferase 11	VEEIPENDSWK	48	Golgi apparatus(membrane)	
89	GGH_MOUSE	Gamma-glutamyl hydrolase	LPLNFTEGAR	55	Secreted, Lysosome	
90	GL8D1_MOUSE	Glycosyltransferase 8 domain-containing protein 1	ASTCSFNPGVFVANLTEWK	119	unknown (membrane)	
91	GLCE_MOUSE	D-glucuronyl C5-epimerase	NLTEKPPHIEVYETAEDR	34	Golgi apparatus (membrane)	
92	GLCM_MOUSE	Glucosylceramidase	DLGPALANSSHDVK	95	Lysosome (membrane)	
			MELSVGAIQANR	84		
			VYTYADTPNDFQLSNFSLPEEDTK	89		
93	GLT11_MOUSE	Polypeptide N-acetylgalactosaminyltransferase 11	SFGNISER	41	Golgi apparatus (membrane)	
94	GNS_MOUSE	N-acetylglucosamine-6-sulfatase	TPMTNSSIR	56	Lysosome	
			YPHNHHVVNNTLEGNCSSK	71		
			HGENYSVDYLTDVLANLSLDFLDYK	55		
			MLVSNIDLGPTILDLAGYDLNK	78		
			YYNYTLSINGK	52		
95	GOLM1_MOUSE	Golgi membrane protein 1	AVLVNNITTGEK	41	Golgi apparatus(membrane)	

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
96	GREM1_MOUSE	Gremlin-1	AQHNDSEQTQSPPQPGSR	39	Secreted	
97	GRN_MOUSE	Granulins	NYTTDLLTK	45	Secreted	
98	GSLG1_MOUSE	Golgi apparatus protein 1	GNITEYQCHQYITK	54	plasma membrane, Golgi apparatus (membrane)	
			LNLTTDPK	58		
99	GT251_MOUSE	Glycosyltransferase 25 family member 1	AMNTSQVEAMGIQMLPGYR	76	Endoplasmic reticulum	
			TALWVATDHNTDNTSAILR	143		
100	GT252_MOUSE	Glycosyltransferase 25 family member 2	ALNTSQLK	46	Endoplasmic reticulum	
101	HA17_MOUSE	H-2 class I histocompatibility antigen, Q7 alpha chain	TAQSYYNOSK	56	plasma membrane	
102	HA1B_MOUSE	H-2 class I histocompatibility antigen, K-B alpha chain	TLLGYYNOSK	46	plasma membrane	
103	HEXA_MOUSE	Beta-hexosaminidase subunit alpha	SAEGTFFINK	90	Lysosome	
			LWSSNLTTNIDFAFK	80		
104	HEXB_MOUSE	Beta-hexosaminidase subunit beta	LLYISAEDFSIDHSPNSTAGPSCSLLQEAFR	90	Lysosome	
			TQVFGPVDPTVNTTYAFFNTFFK	80		
105	HPLN1_MOUSE	Hyaluronan and proteoglycan link protein 1	GGNVTLPCK	76	Extracellular matrix, Secreted	
106	HS2ST_MOUSE	Heparan sulfate 2-O-sulfotransferase 1	YHVLHINTTK	36	Golgi apparatus (membrane)	
107	HSP13_MOUSE	Heat shock 70 kDa protein 13	NSTIQAANLAGLK	90	Endoplasmic reticulum (membrane)	
108	HYOU1_MOUSE	Hypoxia up-regulated protein 1	AEPPLNASAGDQEEK	104	Endoplasmic reticulum	
			ENGTDAVQEEEESPAEGSK	108		
			LSALDNLLNHSSIFLK	103		
			VFGSQNLTTVK	86		
			VINDTWAWK	69		
			FQISPQLQFSPEEVLGMVLNYSR	35		
109	IGF1R_MOUSE	Insulin-like growth factor 1 receptor	NTTVADTYNITDPEEFETEYPFFESR	35	plasma membrane	CD221
110	IGSF8_MOUSE	Immunoglobulin superfamily member 8	IGPGEPLELLCNVSGALPPPGR	126	plasma membrane	CD316
			GETASLLCNISVR	47		
111	IKIP_MOUSE	Inhibitor of nuclear factor kappa-B kinase-interacting protein	FQNITDFWK	68	Endoplasmic reticulum (membrane)	
112	IL6RB_MOUSE	Interleukin-6 receptor subunit beta	ETYLETNYTLK	52	plasma membrane	CD130
113	ITA5_MOUSE	Integrin alpha-5	NALNLTFHAQNLGEGGAYEAELR	88	plasma membrane	CD49e
			TEKDPQNDPVGTCYLSTENFTR	113		
			HPGNFSSLSCDYFAVNOSR	100		
114	ITA6_MOUSE	Integrin alpha-6	LWNSTFLEEYSK	77	plasma membrane	CD49f
115	ITA7_MOUSE	Integrin alpha-7	LWNSTFLEEYMAVK	63	plasma membrane	
116	ITAV_MOUSE	Integrin alpha-V	ANTTQPGIVEGGQVLK	91	plasma membrane	CD51
			TAADATGLQPILNQFTPANVSR	123		
117	ITB1_MOUSE	Integrin beta-1	DTCAQECSHFNLTK	87	plasma membrane	CD29
			NPCTSEQNCTSPFSYK	133		
118	ITB5_MOUSE	Integrin beta-5	NFTALIPGTTVEILHGDSK	69	plasma membrane	
			SNLTVLR	55		
119	ITPR1_MOUSE	Inositol 1,4,5-trisphosphate receptor type 1	VETGENCTSPAPK	96	Endoplasmic reticulum(membrane)	
120	JIP1_MOUSE	C-jun-amino-terminal kinase-interacting protein 1	SQDTLNNNSLGKK	35	Cytoplasm, Nucleus	
121	K0090_MOUSE	Uncharacterized protein KIAA0090	FINYNQTVSR	75	plasma membrane	
122	K1199_MOUSE	Protein KIAA1199 homolog	GIQFYDGPINIQNCTFR	44	unknown	
123	KDEL1_MOUSE	KDEL motif-containing protein 1	EMNCSETISQIQK	91	Endoplasmic reticulum	
			SNLSDLLEK	38		
124	KTEL1_MOUSE	KTEL motif-containing protein 1	FLSYNVTR	44	Endoplasmic reticulum	
125	LAMA4_MOUSE	Laminin subunit alpha-4	NASGIYAEIDGAK	67	Extracellular matrix, Secreted	
126	LAMC1_MOUSE	Laminin subunit gamma-1	LLNNLTSIK	85	Extracellular matrix, Secreted	
			NISQDLEK	39		
			VNSSLHSQISR	61		

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules	
127	LAMP1_MOUSE	Lysosome-associated membrane glycoprotein 1	AFNISPNDTSSGSCGINLVTLK	130	plasma membrane, Endosome, Lysosome		CD107a
			DATIQAYLSSGNFSK	90			
			GYLLTLNFTKNTTR	35			
			LNMTLPDALVPTFSISNHSLK	66			
			NGSSCGKENVSDPSLTITFGR	67			
			NVTVVLR	45			
			VYMKNVTVVLR	39			
			YSVQHMYFTYNLSDTEHFPNAISK	48			
128	LAMP2_MOUSE	Lysosome-associated membrane glycoprotein 2	CNSVLTYNLTPVVQK	101	plasma membrane, Endosome, Lysosome		CD107b
			NLSFWDAPLGSSYMCNK	113			
			VPFIFNINPATTNFTGSCQPQSAQLR	100			
			VQPFNVTK	36			
129	LASS2_MOUSE	LAG1 longevity assurance homolog 2	LWLPVNLTWADLEDK	106	Endoplasmic reticulum, Nucleus(membrane)		
130	LASS6_MOUSE	LAG1 longevity assurance homolog 6	FWLPHNVTWADLK	32	Endoplasmic reticulum, Nucleus(membrane)		
131	LCAP_MOUSE	Leucyl-cystinyl aminopeptidase	DIILHSTGHNISR	70	plasma membrane		
			DSLNSSHPISSSVQSSEQIEEMFDSLSYFK	115			
			EETLLYDNATSSVADR	109			
			KDSLNSSHPISSSVQSSEQIEEMFDSLSYFK	110			
			NLSQDVNGTLVSVYAVPEK	68			
			SDVINLTEQVQWVK	103			
			NETHTAPITEALFQTNLIYNLLEK	32			
132	LG3BP_MOUSE	Galectin-3-binding protein	ALGYENATQALGR	86	Extracellular matrix, Secreted		
			APIPTALDTNSSK	71			
			GLNLTEDTYKPR	47			
133	LGMN_MOUSE	Legumain	SHTNTSHVMQYGNK	55	Lysosome		
134	LICH_MOUSE	Lysosomal acid lipase/cholesteryl ester hydrolase	YDLPASINYILNK	105	Lysosome		
135	LIFR_MOUSE	Leukemia inhibitory factor receptor	IEGLTNETYR	56	Plasma membrane, Secreted	CD118	
136	LOXL3_MOUSE	Lysyl oxidase homolog 3	NITAEDCSHSQDAGVR	34	Secreted		
137	LPP2_MOUSE	Lipid phosphate phosphohydrolase 2	GSPANVTEAR	70	unknown(membrane)		
138	LRP1_MOUSE	Prolow-density lipoprotein receptor-related protein 1	DNATDSVPLR	71	Plasma membrane, Cytoplasm, Nucleus	CD91	
			DNTTCYEFK	54			
			FGTCSQLCNNTK	100			
			FNSTEYQVVTR	100			
			GVTHLNISGLK	77			
			IETILLNGTDR	100			
			LNGTDPIVAADSK	119			
			LNLDGSNYTLLK	85			
			LYWISSGNHTINR	77			
			MHLNGSNVQVLHR	48			
			TCVSNCTASQFVCK	124			
			WTGHNVTVVQR	55			
			GPCSHLCLINYNR	36			
			LTSCATNASMCGDEAR	97			
			QTGDVTCNCTDGR	66			
139	LRP4_MOUSE	Low-density lipoprotein receptor-related protein 4	VLINNSLDEPR	42	plasma membrane		
140	LY75_MOUSE	Lymphocyte antigen 75	EAPFGTNCNLTITSR	102	plasma membrane	CD205	
			ENYNITMR	39			
			LANISGEEQK	84			
			SPLTGTWNFTSCSER	108			

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
141	LYAG_MOUSE	Lysosomal alpha-glucosidase	EVTVLGVATAPTQVLSNGIPVSNFTYSPDNK	71	Lysosome	
			GVFITNETGQPLIGK	107		
			LENLSSTESGYTATLTR	148		
			QVVENMTR	44		
142	LYPA3_MOUSE	1-O-acylceramide synthase	VSLQELPGSEHIEMLANATTLAYLK	68	Secreted, Lysosome	
			VFVYTPTTNYTLR	42		
143	MA2B1_MOUSE	Lysosomal alpha-mannosidase	ANLTWTVK	53	Lysosome	
			ELNISICPVSQTSER	94		
			TALVQEVHQNFSAWCSQVIR	43		
			LVNAQQANGSLVHVLYSTPTCYLWELNK	48		
144	MATN4_MOUSE	Matrilin-4	SLDVGLNATR	77	Secreted	
145	MFGM_MOUSE	Lactadherin	NNTIPDSQMSASSSYK	61	plasma membrane	
146	MINP1_MOUSE	Multiple inositol polyphosphate phosphatase 1	NETALYHVEAFK	71	plasma membrane, Endoplasmic reticulum	
147	MPRD_MOUSE	Cation-dependent mannose-6-phosphate receptor	EASNHSSGAGLVQINK	107	Lysosome (membrane)	
			INETHIFNGSNWIMLIYK	97		
			LRPLFNK	31		
148	MPRI_MOUSE	Cation-independent mannose-6-phosphate receptor	AACAVRPQEVTMVNGTLTNPVTGK	108	Lysosome (membrane)	CD222
			AGINASYSEK	80		
			DAGVGFPEYQEEDNSTYNFR	145		
			HGNLYDLKPLGLNDTIVSVGEYTYYLR	80		
			ISTNITLVCKPGDLESAPVLR	104		
			MSVINFECNK	66		
			NGSSIIDLSPLIHR	60		
			SLLEFNTTMGCQPSDSQHR	140		
			TQGCAVTDEQLLYSFNLTSLSTSTFK	55		
			MNYTGGDTCHK	51		
149	MRC2_MOUSE	C-type mannose receptor 2	GTDPSLSPSPAATPPAPGAELSYLNHTFR	112	plasma membrane	CD280
			KPNATVEPIQPDR	97		
			WNDSPCNQSLPSICK	115		
150	NAGAB_MOUSE	Alpha-N-acetylgalactosaminidase	VNYTEVSR	68	Lysosome	
151	NCAM1_MOUSE	Neural cell adhesion molecule 1	DGQLLPSSNYSNIK	44	plasma membrane	CD56
152	NCLN_MOUSE	Nicalin	VIYNLTEK	48	Endoplasmic reticulum (membrane)	
153	NEP_MOUSE	Neprilysin	SCINESAIDSR	98	plasma membrane	CD10
			EIANATTKPEDR	72		
154	NEUR1_MOUSE	Sialidase-1	VNLTLR	43	plasma membrane, Cytoplasm, Lysosome	
			WSFSNGTSWQK	56		
155	NICA_MOUSE	Nicastrin	ANNSWFQSILK	78	plasma membrane	
			NISGVVLADHSGSFHNR	78		
156	NPC1_MOUSE	Niemann-Pick C1 protein	LIASNITETMR	93	plasma membrane, Lysosome, Endosome	
157	NRP2_MOUSE	Neuropilin-2	NFTSPNGTIESPGFPEK	85	plasma membrane	
158	OSTM1_MOUSE	Osteopetrosis-associated transmembrane protein 1	NIGNTSEGPR	38	plasma membrane	
159	P2RX4_MOUSE	P2X purinoceptor 4	AAENFTLLVK	92	plasma membrane	
			GVAVTNTSQLGFR	94		
			NILPNITTSYLK	62		
160	P3H1_MOUSE	Prolyl 3-hydroxylase 1	MYYNVTEK	50	Extracellular matrix, Secreted, Endoplasmic reticulum	
161	P3H2_MOUSE	Prolyl 3-hydroxylase 2	EGGPLLYENITFVYNSEQLNGTQR	114	Endoplasmic reticulum, Golgi apparatus	
162	P3H3_MOUSE	Prolyl 3-hydroxylase 3	DALLMEGVTLTQDAQQLNGSER	92	Endoplasmic reticulum	
163	P4HA1_MOUSE	Prolyl 4-hydroxylase subunit alpha-1	DMSDGFISNLTIQR	133	Endoplasmic reticulum	

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
164	P4HA2_MOUSE	Prolyl 4-hydroxylase subunit alpha-2	LNTDWPALGDLVLQDASAGFVANLSVQR	118	Endoplasmic reticulum	
			SLSNOTDAGLATQENLYERPTDYLPER	141		
165	PCYOX_MOUSE	Prenylcysteine oxidase	LLNOTLR	58	Lysosome	
			MSNITFR	51		
166	PGBM_MOUSE	Basement membrane-specific heparan sulfate proteoglycan core protein	ALVNFTR	36	Extracellular matrix, Secreted	
			SLTQGSLIVGNLAPVNGTSQGK	95		
			LTVPSSQNSSFR	30		
167	PGCA_MOUSE	Aggrecan core protein	AFCAAQNATLASTGQLYAAWSQGLDK	125	Extracellular matrix, Secreted	
			DTNETYDVYCFAEEMEGEVFYATSPEK	56		
			TVYLHANOTGYPDPSSR	98		
			TVYLYPNOTGLPDPLSK	99		
168	PGCP_MOUSE	Plasma glutamate carboxypeptidase	EVMNLLQPLNVTK	89	Secreted	
169	PGH2_MOUSE	Prostaglandin G/H synthase 2	SWEAFSNLSYYTR	101	Endoplasmic reticulum, Microsome	
			TGFYGENCTTPEFLTR	119		
170	PGS1_MOUSE	Biglycan	LLQVVYLHSNNITK	60	Extracellular matrix, Secreted	
			MIENGSLSFLPTLR	94		
171	PIGS_MOUSE	GPI transamidase component PIG-S	IYNASELPVR	81	Endoplasmic reticulum(membrane)	
172	PLBL2_MOUSE	Putative phospholipase B-like 2	LEDGFHPDAVAWANLTNAIR	113	Lysosome	
			SDLNPANGSYPFQALHQR	124		
173	PLOD1_MOUSE	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	EQINISLDHR	70	Endoplasmic reticulum(membrane)	
			YIHENYTK	51		
174	PLOD2_MOUSE	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	EAFNITLDHK	67	Endoplasmic reticulum(membrane)	
			LISTANYNTSHLNNDFWQIFENPVDWK	102		
			YFNYTVK	40		
			YNCSIESPR	72		
175	PLOD3_MOUSE	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	EQYIHENYSR	62	Endoplasmic reticulum(membrane)	
			FLQSAEFFNYTVR	114		
176	PLTP_MOUSE	Phospholipid transfer protein	FQIYSNOSALESLALIPLQAPLK	42	Secreted	
			GHFYYNISDVR	70		
			VSNVSCEASVSK	107		
177	PLXA1_MOUSE	Plexin-A1	LSGNLTLLR	36	plasma membrane	
178	PO210_MOUSE	Nuclear pore membrane glycoprotein 210	GATNNTCIIR	53	Endoplasmic reticulum, Nucleus	
			GLMVGNGSVLGVVQAVDAETGK	112		
			VNFTLEASEGCYR	44		
179	POSTN_MOUSE	Periostin	GVNETLLVNELK	109	Extracellular matrix, Secreted	
180	PPAL_MOUSE	Lysosomal acid phosphatase	NLTLMATTSQFPK	64	Lysosome (membrane)	
			YEQLQNETR	76		
181	PPBT_MOUSE	Alkaline phosphatase, tissue-nonspecific isozyme	CNTTQGNEVTSILR	110	plasma membrane	
			HSHYVWNR	35		
			NNLTDPSLSEMVEVALR	95		
			NRTDVEYELDEK	79		
182	PPGB_MOUSE	Lysosomal protective protein	LDPPCTNTTAPSNYLNNPYVR	90	Lysosome	
			MYVTNDTEVAENNYEALK	178		
183	PPT1_MOUSE	Palmitoyl-protein thioesterase 1	FFNDSIVDPVDSEWFGFYR	154	Lysosome	
			NYSIFLADINQER	105		
184	PPT2_MOUSE	Lysosomal thioesterase PPT2	DHPNATAWR	55	Lysosome	
185	PSL2_MOUSE	Signal peptide peptidase-like 2A	IAQEGGAAALLIANNSVLIPSSR	53	unknown (membrane)	
186	PTK7_MOUSE	Tyrosine-protein kinase-like 7	STNASFNIK	56	plasma membrane	
187	PTPRK_MOUSE	Receptor-type tyrosine-protein phosphatase kappa	GPLANPIWNVTGFTGR	65	plasma membrane	

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
188	PTTG_MOUSE	Pituitary tumor-transforming gene 1 protein-interacting protein	NVSCLWCNENK	79	Plasma membrane, Cytoplasm, Nucleus	
			VGCSEYTNR	45		
189	PXDN_MOUSE	Peroxidasin homolog	ILCDNSDNITR	43	Secreted	
190	RCN3_MOUSE	Reticulocalbin-3	NATYGHYEPGEEFHDVEDAETYK	92	Endoplasmic reticulum	
191	RNT2_MOUSE	Ribonuclease T2	AEDCNOSWHFNLDEIK	49	Secreted	
192	RPN1_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	DEIGNVSTSHLLILDDSVEMEIRPR	142	Endoplasmic reticulum (membrane)	
193	RPN2_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	SNLDPSNVDSLFYAAQSSQVLSGCEISVSNETK	116	Endoplasmic reticulum (membrane)	
194	S12A4_MOUSE	Solute carrier family 12 member 4	SIFDPPVFPVCMLGNR	43	plasma membrane	
195	S12A9_MOUSE	Solute carrier family 12 member 9	HGHFTGFNGSTLR	34	plasma membrane	
196	S29A1_MOUSE	Equilibrative nucleoside transporter 1	LDVSQNVSSDTDQSCESTK	122	plasma membrane	
197	SAP_MOUSE	Sulfated glycoprotein 1	DNATQEEILHYLEK	103	Secreted	
			FSELIVNNATEELLVK	123		
			NSTKEEILAALEK	98		
			TNSSFIQGFVDHVK	107		
198	SC65_MOUSE	Synaptonemal complex protein SC65	EEAMLYHNOTSELR	82	Nucleus	
199	SCRB1_MOUSE	Scavenger receptor class B member 1	ESGIQNVSTCR	67	plasma membrane	
			FTAPDTLFANGSVYPPNEGFCPCR	91		
200	SCRB2_MOUSE	Lysosome membrane protein 2	NGTNDGEYVFLTGEDNYLNFSK	126	Lysosome (membrane)	
			NOSVGDPNVDLIR	102		
			TMVFPVMYLNESVLIDK	133		
			TSLDWWTTDTCNMINGTDGDSFHPLISK	123		
			VPAEILANTSENAGFCIPEGNCMDSGVLNISICK	43		
			ANIQFGENGTTISAVTNK	97		
			NMVLQNGTK	43		
201	SE1L1_MOUSE	Protein sel-1 homolog 1	EATIVGENETYPR	93	Endoplasmic reticulum (membrane)	
202	SERPH_MOUSE	Serpin H1	SLSNSTAR	33	Endoplasmic reticulum	
203	SGCE_MOUSE	Epsilon-sarcoglycan	LNAINITSALDR	56	Plasma membrane, Cytoplasm	
204	SHPS1_MOUSE	Tyrosine-protein phosphatase non-receptor type substrate 1	GIANLSNFIR	38	plasma membrane	CD172a
			NVSYNISSTVR	56		
205	SIAE_MOUSE	Sialate O-acetylesterase	NSSDYGFPEIR	80	Cytoplasm, Lysosome	
			NLTFQGPLPK	57		
206	SIL1_MOUSE	Nucleotide exchange factor SIL1	FNSSSSSLEEK	86	Endoplasmic reticulum	
			LINKFNSSSSSLEEK	44		
207	SORT_MOUSE	Sortilin	DITNLINNTFIR	72	plasma membrane, Endoplasmic reticulum, etc.	
208	SSRA_MOUSE	Translocon-associated protein subunit alpha	DLNGNVFQDAVFNQTVTVIER	90	Endoplasmic reticulum(membrane)	
			YPQDYQFYIQNFTALPLNTVVPPQR	65		
209	SSRB_MOUSE	Translocon-associated protein subunit beta	AGYFNFTSATITYLAQEDGPVVIGSTSAPGQGGILAQR	116	Endoplasmic reticulum (membrane)	
			IAPASNVSHTVVLRPLK	111		
210	STIM1_MOUSE	Stromal interaction molecule 1	LAVTNTTMTGTVLK	117	plasma membrane, Endoplasmic reticulum	
211	STT3A_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	TILVDNNTWNNTHISR	135	Endoplasmic reticulum (membrane)	
			VMSWWDYGYQITAMANR	102		
212	STT3B_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	AMSSNETAAYK	83	Endoplasmic reticulum (membrane)	
			TTLVDNNTWNNSHIALVGK	111		
			VMSWWDYGYQIAGMANR	93		
213	SUMF1_MOUSE	Sulfatase-modifying factor 1	FVNSTGYLTEAEK	107	Endoplasmic reticulum	
214	T106B_MOUSE	Transmembrane protein 106B	LNNITNIGPLDMK	91	unknown (membrane)	
215	TENA_MOUSE	Tenascin	ASTEEVPSLENLTVTEAGWDGLR	136	Extracellular matrix, Secreted	
			EPEIGNLNVSDVTPK	102		
			GNFSAEGCGCVCEPGWK	106		

Table 3-II: (continued)

	Swiss-Prot Number	Protein name	Sequence	Mascot score[a]	Subcellular location[b]	CD molecules
			GPNCSEPDCPGNCNLR	81		
			LLQTAEHNISGAER	98		
			LNYSLPTGQSMEVQLPK	82		
			NFFIQVLEADTTQTVQNLTVPGGLR	38		
216	TENN_MOUSE	Tenascin-N	ADLQTFNESAYAVYDFFQVASSK	65	Extracellular matrix, Secreted	
			LHNLTTGTTTR	53		
217	TFR1_MOUSE	Transferrin receptor protein 1	NITAFNETLFR	85	plasma membrane	CD71a
218	TGFB1_MOUSE	Transforming growth factor beta-1	LASPPSQGEVPPGPLPEAVLALYNSTR	41	Secreted	
219	TIMP1_MOUSE	Metalloproteinase inhibitor 1	FMGSPEINETTLYQR	102	Secreted	
220	TM2D1_MOUSE	TM2 domain-containing protein 1	LSITNETFR	57	plasma membrane	
221	TM2D3_MOUSE	TM2 domain-containing protein 3	NFVINMTCR	58	unknown (membrane)	
			YFANCTVR	55		
222	TM9S1_MOUSE	Transmembrane 9 superfamily member 1	IIFANVSVR	41	unknown (membrane)	
223	TM9S3_MOUSE	Transmembrane 9 superfamily member 3	IVDVNLTSEGK	64	unknown (membrane)	
224	TMED9_MOUSE	Transmembrane emp24 domain-containing protein 9	FTFTSHTPGEHQICLHSNSTK	88	Endoplasmic reticulum (membrane)	
225	TMEM2_MOUSE	Transmembrane protein 2	GDPSIISVNGTDFTFR	56	unknown (membrane)	
			FDTHEYHNESR	46		
226	TMTC4_MOUSE	Transmembrane and TPR repeat-containing protein 4	NLADQGNQTAAIK	96	unknown (membrane)	
227	TOIP2_MOUSE	Torsin-1A-interacting protein 2	HLNASNPSEPATIIFTAAR	73	Endoplasmic reticulum (membrane)	
228	TOR1B_MOUSE	Torsin-1B	FTECCHEERPLNTSALK	52	Endoplasmic reticulum	
			GNVSACGSSVFIFDEMDK	93		
229	TOR2A_MOUSE	Torsin-2A	SWVQGNLTACGR	67	Secreted	
230	TPP1_MOUSE	Tripeptidyl-peptidase 1	DVGSGTTNNSQACAQFLEQYFHNSDLTEFMR	135	Lysosome	
			SSSHLPPSSYFNASGR	43		
231	TSN3_MOUSE	Tetraspanin-3	NQSVPLSCCR	55	plasma membrane	
232	TSN31_MOUSE	Tetraspanin-31	NTQADVINASWSVLSNSTR	153	plasma membrane	
233	TSP1_MOUSE	Thrombospondin-1	GCSSSTNVLLTLDNNVVNGSSPAIR	127	Secreted	
			VSCPIMPCSNATVPDGECCPR	67		
			VVNSTTGPGEHLR	86		
234	TSP2_MOUSE	Thrombospondin-2	VVNSTTGTGEHLR	101	Secreted	
			NMSACVQEGR	77		
235	TTYH3_MOUSE	Protein tweety homolog 3	VWDTAAALNR	34	plasma membrane	
236	TWSG1_MOUSE	Twisted gastrulation protein homolog 1	NYSDTPPTSK	58	Secreted	
237	TXD15_MOUSE	Thioredoxin domain-containing protein 15	IFIFNQTGIEAK	78	plasma membrane	
238	U513_MOUSE	UPF0513 transmembrane protein	NVTTVLSDPNPIWLVGR	87	unknown (membrane)	
			YNVSVINGTSPFAQDYDLTHIVAAYQER	77		
239	UGGG1_MOUSE	UDP-glucose:glycoprotein glucosyltransferase 1	GTEVNATVIGESDPIDEVQGFLFGK	113	Endoplasmic reticulum	
240	UN84A_MOUSE	Protein unc-84 homolog A	TLSPTGNISSAPK	80	Cytoplasm, Nucleus (membrane)	
241	UN84B_MOUSE	Protein unc-84 homolog B	ALSPNSTISSAPK	70	Nucleus (membrane)	
242	UNC5B_MOUSE	Netrin receptor UNC5B	LSDTANYTCVAK	55	plasma membrane	
243	VA0E1_MOUSE	V-type proton ATPase subunit e 1	NETIWYLK	60	unknown(membrane)	
244	VAS1_MOUSE	V-type proton ATPase subunit S1	LNASLPALLLIR	87	unknown (membrane)	
245	VASN_MOUSE	Vasorin	LHEISNETFR	51	plasma membrane	
246	WNT5A_MOUSE	Protein Wnt-5a	WNCSTVDNTSVFGR	52	Extracellular matrix, Secreted	

[a] $P < 0.05$ by Mascot Search.

[b] Subcellular localizations of the identified proteins using information from the UniProt database (http://www.uniprot.org/).

Table 3-III: Gene expression profiles of the identified 246 glycoproteins during cell differentiation by microarray analysis

	Uniprot ID	Protein Name	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
1	4F2_MOUSE	4F2 cell-surface antigen heavy chain	1.00	1.08	0.90	1.34	1.31	1.38	0.76	0.84
2	5NTD_MOUSE	5~-nucleotidase	1.00	1.08	1.23	2.17	2.90	5.01	6.30	6.33
3	ACAM_MOUSE	Adipocyte adhesion molecule	1.00	1.13	2.03	2.96	2.82	2.39	1.85	1.54
4	ADA10_MOUSE	Disintegrin and metalloproteinase domain-containing protein 10	1.00	0.92	0.99	0.98	1.01	0.86	0.92	0.84
5	ADAM9_MOUSE	Disintegrin and metalloproteinase domain-containing protein 9	1.00	0.90	0.78	0.72	0.71	0.68	0.49	0.60
6	AEBP1_MOUSE	Adipocyte enhancer-binding protein 1	1.00	0.95	0.98	0.97	0.97	0.76	0.74	0.61
7	AER61_MOUSE	Uncharacterized glycosyltransferase AER61	1.00	1.11	0.93	0.84	0.83	1.07	1.51	1.30
8	AGAL_MOUSE	Alpha-galactosidase A	1.00	1.06	1.22	1.92	1.98	1.27	1.43	1.10
9	AMPN_MOUSE	Aminopeptidase N	1.00	1.18	2.10	3.04	3.83	2.97	3.02	2.77
10	AMRP_MOUSE	Alpha-2-macroglobulin receptor-associated protein	1.00	0.85	0.96	1.08	1.17	0.84	0.86	0.65
11	ANO6_MOUSE	Anoctamin-6	1.00	0.65	0.76	0.82	1.41	1.49	0.95	0.49
12	ANPRC_MOUSE	Atrial natriuretic peptide clearance receptor	1.00	0.26	0.25	0.19	0.22	0.15	0.40	0.16
13	ANTR1_MOUSE	Anthrax toxin receptor 1	1.00	0.66	1.00	1.36	1.53	1.26	0.90	0.81
14	APMAP_MOUSE	Adipocyte plasma membrane-associated protein	1.00	0.52	0.68	0.58	0.57	0.55	0.62	0.84
15	AQP1_MOUSE	Aquaporin-1	1.00	1.00	1.92	2.32	2.32	1.80	1.58	1.56
16	ARSB_MOUSE	Arylsulfatase B	1.00	1.17	1.50	1.29	1.28	1.18	1.38	1.80
17	ASAH1_MOUSE	Acid ceramidase	1.00	0.86	1.08	1.04	1.31	1.30	1.42	1.48
18	ASM3B_MOUSE	Acid sphingomyelinase-like phosphodiesterase 3b	1.00	0.73	0.74	0.87	0.88	0.99	0.73	0.81
19	AT1B3_MOUSE	Sodium/potassium-transporting ATPase subunit beta-3	1.00	0.96	0.79	0.57	0.89	0.97	1.60	0.90
20	ATRN_MOUSE	Attractin	1.00	1.21	2.22	1.30	1.28	2.04	1.55	1.71
21	BASI_MOUSE	Basigin	1.00	0.94	1.05	1.57	2.17	2.68	4.50	4.43
22	BGLR_MOUSE	Beta-glucuronidase	1.00	0.83	0.99	1.15	1.27	1.07	1.09	0.87
23	BTD_MOUSE	Biotinidase	1.00	0.70	0.92	0.65	0.80	0.53	0.91	0.65
24	CA2D1_MOUSE	Voltage-dependent calcium channel subunit alpha-2/delta-1	1.00	1.18	1.09	1.26	1.03	1.47	1.09	1.47
25	CADH2_MOUSE	Cadherin-2	1.00	1.40	0.66	0.42	0.40	1.41	1.42	0.91
26	CALR_MOUSE	Calreticulin	1.00	1.12	0.97	1.03	1.12	1.19	1.99	1.77
27	CALU_MOUSE	Calumenin	1.00	0.89	0.86	0.89	0.95	1.08	1.57	1.49
28	CATD_MOUSE	Cathepsin D	1.00	0.97	1.07	1.21	1.68	1.41	2.53	2.24
29	CATF_MOUSE	Cathepsin F	1.00	0.56	0.80	1.56	2.75	2.61	4.07	3.60
30	CATL1_MOUSE	Cathepsin L1	1.00	0.88	0.97	1.17	1.78	1.57	2.17	1.92
31	CATZ_MOUSE	Cathepsin Z	1.00	0.83	1.06	1.36	1.99	1.85	2.30	1.97
32	CBPD_MOUSE	Carboxypeptidase D	1.00	1.12	0.91	0.91	0.96	0.96	1.43	1.02
33	CBPE_MOUSE	Carboxypeptidase E	1.00	0.81	1.01	1.91	2.68	3.30	5.38	5.10
34	CCD80_MOUSE	Coiled-coil domain-containing protein 80	1.00	0.63	0.59	0.63	0.81	0.65	1.38	1.51
35	CD109_MOUSE	CD109 antigen	1.00	1.13	1.00	1.36	1.45	1.82	3.26	2.52
36	CD276_MOUSE	CD276 antigen	1.00	1.08	1.40	1.42	2.17	1.40	1.45	1.25
37	CD44_MOUSE	CD44 antigen	1.00	1.09	0.41	0.36	0.88	1.42	2.33	1.10
38	CD47_MOUSE	Leukocyte surface antigen CD47	1.00	1.12	1.39	1.05	0.78	0.93	0.96	1.05
39	CERU_MOUSE	Ceruloplasmin	1.00	0.69	1.03	1.30	1.65	1.73	1.95	2.14
40	CLDN1_MOUSE	Claudin domain-containing protein 1	1.00	1.03	0.98	0.92	0.94	1.14	1.40	1.29
41	CLP1L_MOUSE	Cleft lip and palate transmembrane protein 1-like protein	1.00	1.04	0.93	0.96	0.92	0.98	0.82	1.02
42	CLPT1_MOUSE	Cleft lip and palate transmembrane protein 1 homolog	1.00	0.97	0.95	0.98	0.95	0.90	0.74	0.71
43	CNPY3_MOUSE	Protein canopy homolog 3	1.00	1.05	1.13	1.32	1.21	1.21	1.33	1.15
44	CO1A1_MOUSE	Collagen alpha-1(I) chain	1.00	1.08	1.26	1.71	1.93	2.09	2.23	1.94
45	CO1A2_MOUSE	Collagen alpha-2(I) chain	1.00	0.96	1.07	1.20	1.28	1.23	1.36	1.24
46	CO2A1_MOUSE	Collagen alpha-1(II) chain	1.00	0.88	0.93	5.61	8.40	15.35	19.77	23.72
47	CO3A1_MOUSE	Collagen alpha-1(III) chain	1.00	1.44	2.82	3.95	4.35	3.77	3.23	2.57
48	CO5A1_MOUSE	Collagen alpha-1(V) chain	1.00	0.85	0.78	0.96	1.05	1.11	1.43	1.34

Table 3-III: (continued)

	Uniprot ID	Protein Name	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
49	CO5A2_MOUSE	Collagen alpha-2(V) chain	1.00	0.69	0.60	0.83	1.07	1.16	1.75	1.54
50	CO6A1_MOUSE	Collagen alpha-1(VI) chain	1.00	1.17	1.51	1.50	1.43	1.47	1.62	1.91
51	CO6A2_MOUSE	Collagen alpha-2(VI) chain	1.00	1.30	1.69	1.82	1.59	1.76	1.65	2.17
52	COBA1_MOUSE	Collagen alpha-1(XI) chain	1.00	0.71	0.41	2.71	4.40	8.17	6.73	7.67
53	COCA1_MOUSE	Collagen alpha-1(XII) chain	1.00	0.63	0.58	1.20	2.47	2.04	2.00	1.34
54	COL12_MOUSE	Collectin-12	1.00	0.94	1.22	1.15	1.61	1.16	1.81	1.44
55	CREG1_MOUSE	Protein CREG1	1.00	0.71	1.03	1.35	1.69	1.72	2.33	2.36
56	CREL2_MOUSE	Cysteine-rich with EGF-like domain protein 2	1.00	1.10	1.16	1.23	1.19	1.32	1.37	1.49
57	CRRY_MOUSE	Complement regulatory protein Crry	1.00	1.20	1.60	1.32	4.58	1.30	1.47	2.15
58	CRTAP_MOUSE	Cartilage-associated protein	1.00	0.75	0.85	1.01	1.18	1.21	1.51	1.50
59	CS063_MOUSE	UPF0510 protein C19orf63 homolog	1.00	0.94	1.11	1.14	1.17	1.20	1.16	1.15
60	CTL2_MOUSE	Choline transporter-like protein 2	1.00	0.77	0.70	0.90	1.10	1.41	1.82	1.76
61	CYTC_MOUSE	Cystatin-C	1.00	0.76	0.92	1.08	1.75	1.63	3.43	3.19
62	DAG1_MOUSE	Dystroglycan	1.00	0.91	1.11	1.21	1.53	1.31	1.56	1.46
63	DAPLE_MOUSE	Protein Daple	1.00	0.88	2.70	6.43	7.74	15.07	12.59	13.58
64	DPP2_MOUSE	Dipeptidyl-peptidase 2	1.00	0.93	1.30	1.61	2.00	1.66	1.82	1.60
65	ECE1_MOUSE	Endothelin-converting enzyme 1	1.00	0.44	0.43	1.81	1.18	0.61	1.18	2.18
66	ELOV4_MOUSE	Elongation of very long chain fatty acids protein 4	1.00	0.80	0.97	1.24	1.28	1.47	1.12	1.43
67	EMB_MOUSE	Embigin	1.00	0.76	0.91	1.28	1.35	2.04	2.63	2.65
68	EMIL1_MOUSE	EMILIN-1	1.00	0.81	0.94	0.94	1.01	0.86	1.55	1.25
69	ENPL_MOUSE	Endoplasmin	1.00	1.07	1.00	1.21	1.21	1.36	1.82	1.64
70	ENPP1_MOUSE	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1	1.00	2.85	3.33	6.41	13.59	8.71	15.22	12.12
71	ENPP4_MOUSE	Ectonucleotide pyrophosphatase/phosphodiesterase family member 4	1.00	0.87	1.04	1.20	1.10	1.09	1.15	1.07
72	ENPP5_MOUSE	Ectonucleotide pyrophosphatase/phosphodiesterase family member 5	1.00	0.62	1.36	2.33	3.83	3.68	6.20	5.51
73	EPDR1_MOUSE	Mammalian ependymin-related protein 1	1.00	1.20	1.58	1.32	1.35	1.28	1.45	1.93
74	EPHA2_MOUSE	Ephrin type-A receptor 2	1.00	0.59	0.29	0.25	0.37	0.30	0.37	0.36
75	EPHB4_MOUSE	Ephrin type-B receptor 4	1.00	1.12	1.05	1.35	1.46	1.37	1.63	1.16
76	ERLN1_MOUSE	Erlin-1	1.00	1.06	0.87	0.80	0.71	0.81	0.61	0.65
77	ERO1A_MOUSE	ERO1-like protein alpha	1.00	1.02	1.16	2.58	2.40	6.45	3.39	8.62
78	FGFR3_MOUSE	Fibroblast growth factor receptor 3	1.00	0.55	0.51	1.70	1.90	6.42	14.79	27.25
79	FINC_MOUSE	Fibronectin	1.00	1.16	1.84	2.93	4.68	3.84	3.54	2.53
80	FKB10_MOUSE	FK506-binding protein 10	1.00	0.75	0.84	1.06	1.14	1.26	1.56	1.45
81	FKBP7_MOUSE	FK506-binding protein 7	1.00	0.76	1.27	1.64	1.98	1.88	3.03	2.78
82	FKBP9_MOUSE	FK506-binding protein 9	1.00	0.73	0.92	1.19	1.48	1.70	2.66	2.43
83	FMOD_MOUSE	Fibromodulin	1.00	0.96	1.19	5.96	10.42	18.58	21.92	39.19
84	FPRP_MOUSE	Prostaglandin F2 receptor negative regulator	1.00	0.83	0.96	0.95	0.89	0.90	1.40	0.94
85	FSTL1_MOUSE	Follistatin-related protein 1	1.00	0.63	0.74	0.65	0.69	0.62	0.67	0.63
86	FUCO_MOUSE	Tissue alpha-L-fucosidase	1.00	0.81	1.03	1.27	1.67	1.55	2.17	1.78
87	FUCO2_MOUSE	Plasma alpha-L-fucosidase	1.00	0.89	1.14	1.25	1.39	1.24	1.38	1.29
88	FUT11_MOUSE	Alpha-(1,3)-fucosyltransferase 11	1.00	0.89	1.04	1.02	1.11	0.99	1.10	0.93
89	GGH_MOUSE	Gamma-glutamyl hydrolase	1.00	0.83	1.16	1.52	1.82	1.70	1.66	2.14
90	GL8D1_MOUSE	Glycosyltransferase 8 domain-containing protein 1	1.00	0.82	1.02	1.25	1.50	1.50	2.03	1.72
91	GLCE_MOUSE	D-glucuronyl C5-epimerase	1.00	0.90	0.76	0.92	0.89	1.16	0.77	0.84
92	GLCM_MOUSE	Glucosylceramidase	1.00	0.86	1.07	1.35	1.81	1.47	2.16	1.66
93	GLT11_MOUSE	Polypeptide N-acetylgalactosaminyltransferase 11	1.00	0.94	0.70	1.20	1.27	1.39	1.14	1.03
94	GNS_MOUSE	N-acetylglucosamine-6-sulfatase	1.00	1.14	1.33	1.42	2.05	1.94	2.72	1.61
95	GOLM1_MOUSE	Golgi membrane protein 1	1.00	1.24	1.03	0.92	0.71	0.52	0.46	0.34
96	GREM1_MOUSE	Gremlin-1	1.00	0.81	0.89	1.01	0.63	0.81	0.75	0.98
97	GRN_MOUSE	Granulins	1.00	1.46	1.98	2.27	2.42	2.12	1.81	1.76

Table 3-III: (continued)

	Uniprot ID	Protein Name	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
98	GSLG1_MOUSE	Golgi apparatus protein 1	1.00	0.86	0.98	1.24	1.32	1.34	1.28	1.17
99	GT251_MOUSE	Glycosyltransferase 25 family member 1	1.00	0.97	0.92	1.04	1.12	1.24	1.67	1.55
100	GT252_MOUSE	Glycosyltransferase 25 family member 2	1.00	0.64	2.44	22.46	29.72	72.80	172.47	236.01
101	HA17_MOUSE	H-2 class I histocompatibility antigen, Q7 alpha chain	1.00	0.73	0.86	0.89	1.26	1.30	2.02	1.71
102	HA1B_MOUSE	H-2 class I histocompatibility antigen, K-B alpha chain	1.00	0.72	1.04	1.11	1.37	1.39	2.64	1.99
103	HEXA_MOUSE	Beta-hexosaminidase subunit alpha	1.00	0.77	1.15	1.70	2.44	2.26	2.95	2.36
104	HEXB_MOUSE	Beta-hexosaminidase subunit beta	1.00	0.72	0.98	1.46	2.10	1.87	2.77	2.34
105	HPLN1_MOUSE	Hyaluronan and proteoglycan link protein 1	1.00	1.46	2.69	19.37	46.85	150.24	325.32	571.33
106	HS2ST_MOUSE	Heparan sulfate 2-O-sulfotransferase 1	1.00	1.22	1.19	1.15	1.01	0.83	0.55	0.55
107	HSP13_MOUSE	Heat shock 70 kDa protein 13	1.00	3.90	1.59	2.63	3.25	3.33	4.65	6.70
108	HYOU1_MOUSE	Hypoxia up-regulated protein 1	1.00	1.76	1.46	1.24	1.19	2.26	4.58	1.85
109	IGF1R_MOUSE	Insulin-like growth factor 1 receptor	1.00	0.69	0.44	0.54	0.81	0.99	1.58	0.91
110	IGSF8_MOUSE	Immunoglobulin superfamily member 8	1.00	1.17	1.00	1.33	1.84	2.34	3.41	2.63
111	IKIP_MOUSE	Inhibitor of nuclear factor kappa-B kinase-interacting protein	1.00	0.86	1.15	1.36	1.66	1.64	1.66	1.69
112	IL6RB_MOUSE	Interleukin-6 receptor subunit beta	1.00	0.92	0.73	0.66	0.72	0.69	0.78	0.66
113	ITA5_MOUSE	Integrin alpha-5	1.00	0.92	1.02	1.02	1.14	1.09	1.42	1.12
114	ITA6_MOUSE	Integrin alpha-6	1.00	2.29	2.41	3.59	3.44	5.72	9.80	6.52
115	ITA7_MOUSE	Integrin alpha-7	1.00	0.84	0.81	1.98	5.77	4.59	17.27	6.02
116	ITAV_MOUSE	Integrin alpha-V	1.00	1.90	1.30	1.40	3.55	2.32	1.63	1.60
117	ITB1_MOUSE	Integrin beta-1	1.00	0.77	1.00	1.39	1.53	1.59	2.70	1.18
118	ITB5_MOUSE	Integrin beta-5	1.00	0.80	0.75	0.76	0.93	0.74	1.05	0.77
119	ITPR1_MOUSE	Inositol 1,4,5-trisphosphate receptor type 1	1.00	0.93	0.99	1.24	1.65	1.41	1.46	0.88
120	JIP1_MOUSE	C-jun-amino-terminal kinase-interacting protein 1	1.00	1.05	1.84	1.46	2.19	1.50	2.96	1.39
121	K0090_MOUSE	Uncharacterized protein KIAA0090	1.00	1.07	0.99	1.21	1.36	1.41	1.61	1.63
122	K1199_MOUSE	Protein KIAA1199 homolog	1.00	0.38	0.32	0.34	0.71	0.21	0.20	0.12
123	KDEL1_MOUSE	KDEL motif-containing protein 1	1.00	1.02	1.05	1.27	1.10	1.33	0.94	1.05
124	KTEL1_MOUSE	KTEL motif-containing protein 1	1.00	1.32	0.93	0.78	2.81	0.75	1.16	1.14
125	LAMA4_MOUSE	Laminin subunit alpha-4	1.00	0.54	0.86	1.00	2.84	4.93	5.96	5.96
126	LAMC1_MOUSE	Laminin subunit gamma-1	1.00	0.87	0.93	1.13	1.27	1.28	1.26	0.93
127	LAMP1_MOUSE	Lysosome-associated membrane glycoprotein 1	1.00	0.92	1.02	1.13	1.36	1.10	1.37	1.29
128	LAMP2_MOUSE	Lysosome-associated membrane glycoprotein 2	1.00	0.95	1.19	1.41	1.89	1.81	2.17	1.94
129	LASS2_MOUSE	LAG1 longevity assurance homolog 2	1.00	0.97	0.86	0.92	0.90	0.91	0.76	0.80
130	LASS6_MOUSE	LAG1 longevity assurance homolog 6	1.00	0.98	0.81	0.96	0.95	0.95	0.60	0.63
131	LCAP_MOUSE	Leucyl-cystinyl aminopeptidase	1.00	0.79	0.81	0.65	0.53	0.51	0.47	0.38
132	LG3BP_MOUSE	Galectin-3-binding protein	1.00	0.56	1.08	1.29	2.00	1.97	1.85	1.44
133	LGMN_MOUSE	Legumain	1.00	0.86	1.02	1.13	1.56	1.62	3.14	2.55
134	LICH_MOUSE	Lysosomal acid lipase/cholesteryl ester hydrolase	1.00	0.56	1.14	0.41	0.93	1.16	1.77	0.98
135	LIFR_MOUSE	Leukemia inhibitory factor receptor	1.00	0.71	1.01	1.45	1.61	3.17	6.11	9.23
136	LOXL3_MOUSE	Lysyl oxidase homolog 3	1.00	0.93	1.12	1.50	2.10	1.65	1.45	1.64
137	LPP2_MOUSE	Lipid phosphate phosphohydrolase 2	1.00	1.66	1.35	1.17	0.95	1.05	1.13	1.30
138	LRP1_MOUSE	Prolow-density lipoprotein receptor-related protein 1	1.00	0.84	0.77	1.02	1.69	1.97	3.36	2.01
139	LRP4_MOUSE	Low-density lipoprotein receptor-related protein 4	1.00	1.17	0.63	0.92	5.06	2.30	4.97	2.66
140	LY75_MOUSE	Lymphocyte antigen 75	1.00	0.86	0.65	0.58	0.54	0.55	0.52	0.39
141	LYAG_MOUSE	Lysosomal alpha-glucosidase	1.00	0.82	1.05	1.63	2.38	2.15	2.80	2.14
142	LYPA3_MOUSE	1-O-acylceramide synthase	1.00	1.10	1.05	1.08	1.07	1.19	1.46	1.20
143	MA2B1_MOUSE	Lysosomal alpha-mannosidase	1.00	0.81	0.98	1.37	1.92	1.61	1.87	1.49
144	MATN4_MOUSE	Matrilin-4	1.00	0.92	1.52	10.08	32.75	74.59	299.22	265.58
145	MFGM_MOUSE	Lactadherin	1.00	0.86	0.89	1.13	1.27	1.15	1.46	1.26
146	MINP1_MOUSE	Multiple inositol polyphosphate phosphatase 1	1.00	1.00	0.86	0.56	0.58	0.64	0.31	0.40

Table 3-III: (continued)

	Uniprot ID	Protein Name	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
147	MPRD_MOUSE	Cation-dependent mannose-6-phosphate receptor	1.00	1.02	0.95	0.91	0.75	0.78	0.64	0.79
148	MPRI_MOUSE	Cation-independent mannose-6-phosphate receptor	1.00	0.74	1.20	1.97	2.99	2.38	3.58	2.93
149	MRC2_MOUSE	C-type mannose receptor 2	1.00	0.87	1.15	1.64	2.47	2.03	2.77	1.90
150	NAGAB_MOUSE	Alpha-N-acetylgalactosaminidase	1.00	0.67	1.02	1.54	2.43	2.10	3.77	2.48
151	NCAM1_MOUSE	Neural cell adhesion molecule 1	1.00	0.85	0.51	0.60	0.95	1.03	1.48	0.73
152	NCLN_MOUSE	Nicalin	1.00	1.15	0.85	0.78	0.69	0.64	0.36	0.31
153	NEP_MOUSE	Neprilysin	1.00	0.51	1.08	2.76	10.45	7.85	11.18	9.29
154	NEUR1_MOUSE	Sialidase-1	1.00	0.73	1.28	1.60	2.11	1.72	2.01	2.26
155	NICA_MOUSE	Nicastrin	1.00	0.78	0.88	1.22	1.22	1.04	0.94	0.93
156	NPC1_MOUSE	Niemann-Pick C1 protein	1.00	1.22	1.60	1.37	1.33	1.28	1.47	1.93
157	NRP2_MOUSE	Neuropilin-2	1.00	1.20	1.98	7.12	4.95	10.24	1.46	4.53
158	OSTM1_MOUSE	Osteopetrosis-associated transmembrane protein 1	1.00	0.88	1.01	0.82	0.84	0.71	0.71	0.58
159	P2RX4_MOUSE	P2X purinoceptor 4	1.00	0.77	0.94	1.06	1.42	1.07	1.77	1.54
160	P3H1_MOUSE	Prolyl 3-hydroxylase 1	1.00	0.86	1.04	1.55	1.76	2.35	3.73	2.95
161	P3H2_MOUSE	Prolyl 3-hydroxylase 2	1.00	0.62	0.64	1.24	1.37	3.59	5.22	9.14
162	P3H3_MOUSE	Prolyl 3-hydroxylase 3	1.00	0.75	1.02	1.68	2.46	3.04	3.84	3.92
163	P4HA1_MOUSE	Prolyl 4-hydroxylase subunit alpha-1	1.00	0.94	1.56	3.33	4.18	6.16	8.36	8.29
164	P4HA2_MOUSE	Prolyl 4-hydroxylase subunit alpha-2	1.00	3.68	1.04	5.06	7.94	17.32	12.44	16.27
165	PCYOX_MOUSE	Prenylcysteine oxidase	1.00	0.94	0.97	1.10	1.14	0.99	1.16	0.97
166	PGBM_MOUSE	Basement membrane-specific heparan sulfate proteoglycan core protein	1.00	0.79	0.59	0.87	1.00	1.40	1.63	1.43
167	PGCA_MOUSE	Aggrecan core protein	1.00	1.44	3.43	17.66	18.17	46.72	38.49	79.98
168	PGCP_MOUSE	Plasma glutamate carboxypeptidase	1.00	0.65	0.96	1.39	2.02	2.05	2.80	2.68
169	PGH2_MOUSE	Prostaglandin G/H synthase 2	1.00	0.85	0.66	0.96	0.40	1.25	0.72	2.27
170	PGS1_MOUSE	Biglycan	1.00	1.51	1.11	1.32	1.43	1.43	1.86	0.90
171	PIGS_MOUSE	GPI transamidase component PIG-S	1.00	0.80	0.96	1.19	1.38	1.35	1.55	1.21
172	PLBL2_MOUSE	Putative phospholipase B-like 2	1.00	0.88	1.07	1.23	1.57	1.45	1.75	1.46
173	PLOD1_MOUSE	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	1.00	0.93	0.86	1.20	1.35	1.37	2.11	1.81
174	PLOD2_MOUSE	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	1.00	0.75	1.02	1.63	1.36	2.35	1.98	3.25
175	PLOD3_MOUSE	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	1.00	0.99	0.97	1.28	1.33	1.55	1.80	1.71
176	PLTP_MOUSE	Phospholipid transfer protein	1.00	0.69	0.81	1.37	2.66	3.00	4.38	3.79
177	PLXA1_MOUSE	Plexin-A1	1.00	0.97	0.62	0.53	0.60	0.48	0.52	0.35
178	PO210_MOUSE	Nuclear pore membrane glycoprotein 210	1.00	0.58	1.24	3.39	4.57	7.80	10.52	6.47
179	POSTN_MOUSE	Periostin	1.00	0.76	0.53	0.69	0.75	0.64	0.73	0.66
180	PPAL_MOUSE	Lysosomal acid phosphatase	1.00	0.78	0.86	1.37	1.85	1.63	1.60	1.39
181	PPBT_MOUSE	Alkaline phosphatase, tissue-nonspecific isozyme	1.00	0.74	0.64	0.92	1.03	1.51	1.48	1.76
182	PPGB_MOUSE	Lysosomal protective protein	1.00	0.74	1.44	1.81	2.60	2.16	2.43	2.50
183	PPT1_MOUSE	Palmitoyl-protein thioesterase 1	1.00	0.74	1.00	1.28	1.55	1.34	1.79	1.86
184	PPT2_MOUSE	Lysosomal thioesterase PPT2	1.00	0.97	0.91	0.95	0.91	0.97	1.03	1.04
185	PSL2_MOUSE	Signal peptide peptidase-like 2A	1.00	0.99	1.02	1.41	1.78	1.95	2.53	2.51
186	PTK7_MOUSE	Tyrosine-protein kinase-like 7	1.00	0.76	0.79	0.95	1.12	1.01	0.91	0.85
187	PTPRK_MOUSE	Receptor-type tyrosine-protein phosphatase kappa	1.00	0.58	0.51	0.42	0.40	0.21	0.43	0.27
188	PTTG_MOUSE	Pituitary tumor-transforming gene 1 protein-interacting protein	1.00	0.86	0.86	1.07	1.22	1.14	1.32	1.25
189	PXDN_MOUSE	Peroxidasin homolog	1.00	1.21	1.60	1.34	1.36	1.31	1.48	2.07
190	RCN3_MOUSE	Reticulocalbin-3	1.00	0.79	0.99	1.39	1.70	1.65	2.65	2.27
191	RNT2_MOUSE	Ribonuclease T2	1.00	0.89	1.11	1.30	1.46	1.28	1.27	1.20
192	RPN1_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	1.00	0.76	1.74	1.00	1.44	1.52	1.74	1.72
193	RPN2_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	1.00	1.23	2.09	1.41	1.57	1.29	1.50	1.91
194	S12A4_MOUSE	Solute carrier family 12 member 4	1.00	0.67	0.78	0.95	1.15	1.21	1.60	1.38
195	S12A9_MOUSE	Solute carrier family 12 member 9	1.00	0.88	0.88	1.00	1.10	0.96	0.80	0.68

Table 3-III: (continued)

	Uniprot ID	Protein Name	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
196	S29A1_MOUSE	Equilibrative nucleoside transporter 1	1.00	1.23	1.00	1.19	1.27	1.19	1.33	1.15
197	SAP_MOUSE	Sulfated glycoprotein 1	1.00	0.95	1.26	1.55	2.11	1.76	2.28	1.91
198	SC65_MOUSE	Synaptonemal complex protein SC65	1.00	0.79	0.94	1.17	1.37	1.42	1.71	1.69
199	SCRB1_MOUSE	Scavenger receptor class B member 1	1.00	0.89	1.03	0.98	1.09	1.49	2.02	2.08
200	SCRB2_MOUSE	Lysosome membrane protein 2	1.00	0.79	0.92	1.20	1.41	1.47	1.96	1.25
201	SE1L1_MOUSE	Protein sel-1 homolog 1	1.00	1.06	0.97	1.09	1.27	1.63	2.57	2.79
202	SERPH_MOUSE	Serpin H1	1.00	0.96	1.08	1.20	1.38	1.40	1.54	1.45
203	SGCE_MOUSE	Epsilon-sarcoglycan	1.00	0.82	1.04	1.29	1.38	1.32	1.34	1.21
204	SHPS1_MOUSE	Tyrosine-protein phosphatase non-receptor type substrate 1	1.00	1.09	0.90	0.83	0.80	1.14	1.60	1.69
205	SIAE_MOUSE	Sialate O-acetylesterase	1.00	1.00	1.05	1.45	1.79	2.01	3.15	2.52
206	SIL1_MOUSE	Nucleotide exchange factor SIL1	1.00	0.78	0.85	1.29	1.67	1.94	3.71	3.03
207	SORT_MOUSE	Sortilin	1.00	0.57	1.12	2.02	5.28	4.38	11.59	9.25
208	SSRA_MOUSE	Translocon-associated protein subunit alpha	1.00	0.97	0.94	1.17	1.38	1.40	2.89	1.82
209	SSRB_MOUSE	Translocon-associated protein subunit beta	1.00	0.88	0.97	1.07	1.27	1.40	1.79	1.70
210	STIM1_MOUSE	Stromal interaction molecule 1	1.00	1.84	1.58	3.48	3.05	3.67	5.52	2.75
211	STT3A_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	1.00	0.89	0.67	0.71	0.65	0.73	1.02	0.88
212	STT3B_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	1.00	1.46	2.00	2.71	2.97	2.16	2.00	1.95
213	SUMF1_MOUSE	Sulfatase-modifying factor 1	1.00	0.80	0.86	0.93	1.02	0.95	0.90	0.92
214	T106B_MOUSE	Transmembrane protein 106B	1.00	0.91	1.00	1.23	1.35	1.24	1.40	1.54
215	TENA_MOUSE	Tenascin	1.00	0.70	0.50	0.53	0.67	0.83	0.60	0.56
216	TENN_MOUSE	Tenascin-N	1.00	0.91	0.89	0.75	0.45	0.56	0.54	0.56
217	TFR1_MOUSE	Transferrin receptor protein 1	1.00	1.10	0.90	0.79	0.59	0.47	0.12	0.29
218	TGFB1_MOUSE	Transforming growth factor beta-1	1.00	0.80	0.93	1.30	1.43	1.45	2.86	2.11
219	TIMP1_MOUSE	Metalloproteinase inhibitor 1	1.00	1.52	0.93	0.76	0.76	0.90	1.14	1.36
220	TM2D1_MOUSE	TM2 domain-containing protein 1	1.00	1.02	1.19	1.19	1.19	1.27	1.45	1.67
221	TM2D3_MOUSE	TM2 domain-containing protein 3	1.00	1.18	1.20	1.19	1.09	1.14	1.08	1.19
222	TM9S1_MOUSE	Transmembrane 9 superfamily member 1	1.00	0.90	0.90	0.88	0.90	0.93	0.91	0.86
223	TM9S3_MOUSE	Transmembrane 9 superfamily member 3	1.00	1.01	1.01	0.95	0.92	0.92	0.72	0.79
224	TMED9_MOUSE	Transmembrane emp24 domain-containing protein 9	1.00	1.06	1.13	1.16	1.19	1.27	1.40	1.37
225	TMEM2_MOUSE	Transmembrane protein 2	1.00	0.72	0.84	0.90	0.94	0.52	0.60	0.37
226	TMTC4_MOUSE	Transmembrane and TPR repeat-containing protein 4	1.00	0.85	0.83	0.86	0.76	0.79	0.91	0.89
227	TOIP2_MOUSE	Torsin-1A-interacting protein 2	1.00	1.09	1.25	1.15	1.77	2.65	3.29	5.49
228	TOR1B_MOUSE	Torsin-1B	1.00	1.01	0.96	0.90	0.94	0.86	0.81	0.64
229	TOR2A_MOUSE	Torsin-2A	1.00	1.13	1.07	1.27	1.32	1.35	1.04	1.11
230	TPP1_MOUSE	Tripeptidyl-peptidase 1	1.00	0.73	0.97	1.16	1.74	1.41	1.44	1.10
231	TSN3_MOUSE	Tetraspanin-3	1.00	1.01	1.33	1.61	1.72	1.71	2.12	2.18
232	TSN31_MOUSE	Tetraspanin-31	1.00	0.94	1.14	1.20	1.33	1.26	0.99	1.06
233	TSP1_MOUSE	Thrombospondin-1	1.00	0.33	0.19	0.20	0.23	0.28	0.61	0.48
234	TSP2_MOUSE	Thrombospondin-2	1.00	0.76	0.78	0.66	0.87	0.77	2.05	1.00
235	TTYH3_MOUSE	Protein tweety homolog 3	1.00	0.78	1.40	1.36	1.24	0.84	0.92	0.51
236	TWSG1_MOUSE	Twisted gastrulation protein homolog 1	1.00	1.39	1.10	0.92	0.51	0.53	0.63	0.52
237	TXD15_MOUSE	Thioredoxin domain-containing protein 15	1.00	1.30	1.22	1.36	1.75	1.58	1.82	1.58
238	U513_MOUSE	UPF0513 transmembrane protein	1.00	0.87	0.96	1.19	1.31	1.29	1.40	1.40
239	UGGG1_MOUSE	UDP-glucose:glycoprotein glucosyltransferase 1	1.00	0.93	0.70	3.06	2.34	2.13	0.89	1.39
240	UN84A_MOUSE	Protein unc-84 homolog A	1.00	0.93	1.06	1.38	1.43	1.80	1.76	1.59
241	UN84B_MOUSE	Protein unc-84 homolog B	1.00	0.89	0.89	1.06	0.91	0.92	0.94	0.89
242	UNC5B_MOUSE	Netrin receptor UNC5B	1.00	1.52	1.39	2.16	2.28	2.18	0.64	0.85
243	VA0E1_MOUSE	V-type proton ATPase subunit e 1	1.00	1.02	1.37	1.25	1.20	1.22	1.22	1.61
244	VAS1_MOUSE	V-type proton ATPase subunit S1	1.00	0.96	1.07	1.36	1.44	1.45	1.88	1.58

Table 3-III: (continued)

	Uniprot ID	Protein Name	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
245	VASN_MOUSE	Vasorin	1.00	0.69	0.78	0.84	0.84	0.68	0.66	0.66
246	WNT5A_MOUSE	Protein Wnt-5a	1.00	0.72	0.88	1.15	1.56	2.23	3.19	4.43

3-4. Discussion

By performing proteomics focusing on the glycoproteins bearing high-mannose type *N*-glycans during chondrogenic differentiation, we identified 246 glycoproteins that included cartilage-specific extracellular matrix proteins (e.g., aggrecan core protein and collagen α-1(II) chain). It was expected that most of the identified glycoproteins, in which 52% were plasma membrane residences (Figure 3-1), should have important roles in maintenance of chondrocytes because decrease of high-mannose type *N*-glycans were observed both in mouse and human OA cartilage[1]. Recently, proteomics based on differential gel electrophoresis (DIGE) revealed alteration of 39 proteins during chondrogenic differentiation of human umbilical cord stromal MSCs[14]. Although five of these 39 proteins (12.8%) were cell surface proteins, it seems likely that they are not candidates of chondrocyte specific markers during cell differentiation. Merit of our strategy to discover cells specific marker is evident because most *N*-glycans are distributing on cell surfaces as multiple branching structures involved in the glycoproteins and have essential functions in cellular adhesion. In addition, rapid structural alterations in carbohydrate moieties of the glycoproteins appear to be beneficial for rapid cell proliferation and differentiation that needs to exchange cell surface receptors/ligands at each stage during differentiation[15]. On the other hand, it is clear that proteins may alter only the expression levels without any drastic changes in the primary amino acid sequence directly regulated by genetic information. Approach focusing on the glycoforms altered specifically during cell differentiation apparently facilitated the process to enrich potential glycoprotein markers from a large set of tryptic peptides of whole cellular proteins. In the present study, although ConA may become a suited tool for the enrichment of glycopeptides bearing high-mannose type *N*-glycans, it is noteworthy that other options such as reverse glycoblotting method focusing sialic acid residues[16,17] would expand glycoform-focused reverse proteomics and genomics approach.

The gene expression levels of the identified glycoproteins on mouse ATDC5 cells revealed that 15 cell surface glycoproteins showed an enhanced expression during chondrogenic

differentiation with a peak at the earlier stages of differentiation compared with Type II collagen (Figure 3-2). These genes may have several functional roles in development and are expected to be useful as markers indicating the initial stage of differentiation. Five of these 15 genes—*Col3a1*, *Col11a1*, *Enpp1*, *Fn1*, and *Itga6*—have already been reported to be upregulated in chondrogenic differentiation of ATDC5 cells or human MSCs[18-23]. Although the significance of the other 12 genes to chondrogenic differentiation are still unclear, we uncovered that 5 genes—*Aqp1*, *Col3a1*, *Col11a1*, *Enpp1*, and *Unc5b*—are significantly enhanced expression levels at the earlier stages of chondrogenic differentiation of human MSCs compared with Type II collagen (Figure 3-3). As the expression profiles of these 5 genes were quite similar in both human and mouse cases, we may suggest that they have essential roles in the maintenance of chondrocytes and potential for use as chondrocyte specific differentiation markers.

Two proteins, COL3A1 and COL11A1, are well-known cartilage collagens. COL3A1 is expressed in the superficial and deepest area of articular cartilage along with type I collagen. The deepest zone contains ellipsoidal chondrocytes that synthesize lubricin and types I, II, and III collagen fibers. COL11A1 exists in transitional and deep area of articular cartilage[24]. ENPP1 is an enzymatic generator of pyrophosphate in differentiated chondrocytes. The regulation of extracellular pyrophosphate levels is necessary for appropriate mineral homeostasis in cartilage. In aging and OA cartilage, a dysregulated increase in pyrophosphate elaboration in chondrocytes promotes calcium pyrophosphate dihydrate crystal deposition[25]. AQP1 is a well-characterized membrane glycoprotein that belongs to the aquaporin family and is selectively permeated by water driven by osmotic gradients. Cell adhesion to type II collagen-coated plates was significantly reduced by AQP1 deletion. AQP1-mediated plasma membrane water permeability plays an important role in chondrocyte migration and adhesion[26]. UNC5B, netrin receptor, is highly expressed in the brain and has functions in axon growth of neurons and angiogenesis. It is also expressed at lower levels in developing cartilage. The level of UNC5B transcript increases in mouse embryos from day 10.5 in the cartilaginous primordia

of many bones and cartilage[27] while its function during development remains to be unclear. All differentiation marker candidates found in this study are present in cartilage. Considering that some of these molecules may be related closely to the pathogenesis of OA, they are expected to become not only differentiation markers but biomarkers and therapeutic targets.

3-5. Experimental section

Cell culture and chondrogenic differentiation

ATDC5 cells were obtained from RIKEN Cell Bank (Ibaraki, Japan) and cultured in a 1:1 mixture of Dulbecco's modified Eagle's medium (DMEM) and Ham's F-12 medium (Sigma, St. Louis, MO) containing 5% fetal bovine serum (FBS) (Life Technologies, Carlsbad, CA), 2 mM L-glutamine (Life Technologies), 100 units/mL penicillin (Life Technologies), and 100 μg/mL streptomycin (Life Technologies). The cells were seeded into 6-well tissue culture plates at a density of 1.0×10^5 cells and cultured in the above medium supplemented with 10 μg/mL transferrin (Roche Diagnostics Co., Basel, Switzerland). For induction of chondrogenesis, 10 μg/mL bovine insulin (Sigma) was reacted to the sub-confluent cells. Cells were maintained at 37°C in a humidified atmosphere of 5% CO_2 in air. The medium was replaced every other day.

Human MSCs were obtained from Lonza Walkersville, Inc. (Walkersville, MD). Cell culture and chondrogenic differentiation of human MSCs were performed according to the manufacturer's protocols. Human MSCs were expanded in monolayer culture by passaging twice in MSC basal medium (MSCBM) with MSC growth supplement (MCGS), L-glutamine, and GA-1000. For induction of chondrogenic differentiation, human MSCs were suspended in chondrogenic basal medium with ITS supplement (containing insulin, transferrin, and selenite), dexamethasone, ascorbate, sodium pyruvate, proline, GA-1000, L-glutamine, and 10 ng/mL TGF-β3. The seeding cell number for human MSCs was 2.5×10^5 cells/15 mL polypropylene culture tube (Sumitomo Bakelite Co., Ltd., Tokyo, Japan) and the cells were maintained at 37°C in a humidified atmosphere of 5% CO_2 in air. The medium was replaced every other day.

Alcian blue staining

Differentiating human MSCs were rinsed with PBS (Life Technologies) and fixed with 4% formaldehyde (Wako Pure Chemical). The formaldehyde-fixed cells were embedded in paraffin and sectioned. The paraffin-embedded sections were rinsed with *p*-xylene (Wako Pure

Chemical), ethanol (Wako Pure Chemical), and 3% acetic acid (Wako Pure Chemical) and stained with 10 mg/mL Alcian blue 8GX in 3% acetic acid for 30 min.

Enrichment and identification of high-mannose type N-glycopeptides

Enrichment of high-mannose type *N*-glycopeptides was carried out directly using whole-cell lysates as follows. After inducing differentiation, ATDC5 cells were scraped in PBS containing 10 mM EDTA at various time points. The scraped cells were washed three times with PBS containing 10 mM EDTA and lysed by incubation with 1% Triton X-100 for 1 h on ice. The lysates were centrifuged at 15000 rpm for 10 min at 4°C, and the obtained supernatant was added to cold acetone (1:4) to precipitate proteins. The precipitates were dissolved in 500 mM Tris-HCl (pH 8.5) containing 1 mM EDTA, 8 M urea, and 0.1% SDS. The solubilized proteinaceous materials were reduced by 40 mM DTT at 37°C for 90 min followed by alkylation with 60 mM iodoacetamide by incubation in the dark at room temperature for 30 min. The mixture was then treated with 50 μg of trypsin (Promega, Madison, WI) at 37°C overnight. The digested proteins were applied to a concanavalin A (ConA)-agarose (Seikagaku Co., Tokyo, Japan) column equilibrated with a solution of 1.5 M NaCl, 10 mM $MgCl_2$, 10 mM $CaCl_2$, and 100 mM Tris-HCl buffer (pH 7.5). After the column was washed with equilibrated buffer, the high-mannose type *N*-glycopeptides were eluted with buffer containing 10 mM methyl α-D-glucopyranoside followed by elution with buffer containing 10 mM methyl α-D-mannopyranoside. The eluate was then dried by vacuum centrifugation, dissolved in 100 mM NH_4HCO_3, and incubated with 2.5 units of peptide-*N*-glycosidase F at 37°C overnight. The sample was then dried by vacuum centrifugation and stored at –20°C until use. The high-mannose type *N*-glycopeptides were analyzed using a NanoFrontier (L) system (Hitachi High-Technologies, Tokyo, Japan), consisting of a capillary HPLC system based on a nanoGR generator and electrospray ionization linear ion trap time-of-flight mass spectrometry (ESI-LIT-TOFMS). Linear ion trap-time of flight (LIT-TOF) and collision-induced dissociation

(CID) modes were used for MS detection and peptide fragmentation. The reverse-phase nano-capillary columns (NTCC-360/75-3) were prepared by Nikkyo Technos (Tokyo, Japan). The separated peptides were then ionized with a capillary voltage of 1600 V. The ionized peptides were detected at a detector potential of 2000 V. The mobile phases consisted of 2% acetonitrile (0.2% formic acid) in water (A) and 98% acetonitrile (0.2% formic acid) in water (B). After loading peptides onto the column, the mobile phase was held at 95% A and 5% B for 6 min. Exponential gradient elution was performed by increasing the mobile phase composition from 5% to 50% B over 55 min at a flow rate of 100 nL/min. Raw data were converted into the Mascot format using the Data Processing software (Hitachi High-Technologies), and analyzed subsequently using MASCOT ver. 2.2 (Matrix Science K.K., Tokyo, Japan) against the Swiss-Prot protein database(mouse, version 57.4). The following dynamic modifications were considered: carbamidomethylation of cysteine, oxidation of methionine, and PNGase F-catalyzed conversion of asparagine to aspartic acid at the site of carbohydrate attachment.

Real-time qPCR

Total RNA was prepared from ATDC5 cells and human MSCs cultured in differentiation medium using TRIzol reagent (Life Technologies) according to the manufacturer's instructions. After reverse transcription using SuperScript III First-Strand Synthesis SuperMix (Life Technologies), real-time qPCR was performed with the 7500 Real-Time PCR System using Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA).

Microarray analysis

Total RNA was prepared as above. Using the total RNA as templates, cyanine 3-labeled cRNA targets were prepared and purified with Quick Amp Labeling kit according to the "One-Color Microarray-Based Gene Expression Analysis" protocol provided by the manufacturer (Agilent Technologies). The resultant cRNA targets were hybridized to Whole

Mouse Genome Oligo Microarray (Agilent Technologies, 41,174 probes). After hybridization, the microarray slide was washed, dried, and scanned on an Agilent DNA microarray scanner (Agilent Technologies), and hybridization images were quantified using Feature Extraction Software 9.1 (Agilent Technologies). Data analyses were conducted with GeneSpring GX 7.3.1 (Agilent Technologies).

3-6. References

1. Urita A, Matsuhashi T, Onodera T, Nakagawa H, Hato M, Amano M, Seito N, Minami A, Nishimura S-I, Iwasaki N (2011) Alterations of high-mannose type N-glycosylation in human and mouse osteoarthritis cartilage. *Arthritis Rheum* **63**: 3428-3438
2. Cordwell SJ, Thingholm TE (2010) Technologies for plasma membrane proteomics. *Proteomics* **10:** 611-627
3. Zhang H, Li XJ, Martin DB, Aebersold R (2003) Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* **21:** 660–666
4. Chen R, Jiang XN, Sun DG, Han GH, Wang FJ, Ye ML, Wang LM, Zou HF (2009) Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J Proteome Res* **8:** 651-661
5. Kullolli M, Hancock WS, Hincapie M (2008) Preparation of a high-performance multi-lectin affinity chromatography (HP-M-LAC) adsorbent for the analysis of human plasma glycoproteins. *J Sep Sci* **31:** 2733–2739
6. Nishimura S-I (2011) Toward automated glycan analysis. *Adv Carbohydr Chem Biochem* **65**: 219-271.
7. De Bari C, Dell'Accio F, Tylzanowski P, Luyten FP (2001) Multipotent mesenchymal stem cells from adult human synovial membrane. *Arthritis Rheum* **44:** 1928-1942
8. Jones EA, English A, Henshaw K, Kinsey SE, Markham AF, Emery P, McGonagle D (2004) Enumeration and phenotypic characterization of synovial fluid multipotential mesenchymal progenitor cells in inflammatory and degenerative arthritis. *Arthritis Rheum* **50:** 817-827
9. Sakaguchi Y, Sekiya I, Yagishita K, Muneta T (2005) Comparison of human stem cells derived from various mesenchymal tissues: superiority of synovium as a cell source. *Arthritis Rheum* **52:** 2521-2529
10. Nöth U, Steinert AF, Tuan RS (2008) Technology insight: adult mesenchymal stem cells for

osteoarthritis therapy. *Nat Clin Pract Rheumatol* **4:** 371-380

11. Kolf CM, Cho E, Tuan RS (2007) Mesenchymal stromal cells. Biology of adult mesenchymal stem cells: regulation of niche, self-renewal and differentiation. *Arthritis Res Ther* **9:** 204
12. Kobata A, Endo T (1992) Immobilized lectin columns: useful tools for the fractionation and structural analysis of oligosaccharides. *J Chromatogr* **597:** 111–122
13. Uematsu R, Furukawa J, Nakagawa H, Shinohara Y, Deguchi K, Monde K, Nishimura S-I (2005) High throughput quantitative glycomics and glycoform-focused proteomics of murine dermis and epidermis. *Mol Cell Proteomics* **4:** 1977-1989
14. De la Fuente A, Mateos J, Lesende-Rodriguez I, Calamia V, Fuentes-Boquete I, de Toro FJ, Arufe MC, Blanco FJ (2012) Proteome analysis during chondrocyte differentiation in a new chondrogenesis model using human umbilical cord stroma mesenchymal stem cells. *Mol Cell Proteomics* in press
15. Lau KS, Partridge EA, Grigorian A, Silvescu CI, Reinhold VN, Demetriou M, Dennis JW (2007) Complex *N*-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* **129**: 123-134.
16. Kurogochi M, Amano M, Fumoto M, Takimoto A, Kondo H, Nishimura S-I (2007) Reverse glycoblotting allows for rapid enrichment glycoproteomics of biopharmaceuticals and disease-related biomarkers. *Angew Chem Int Ed* **46**: 8808-8813.
17. Kurogochi M, Matsushista T, Amano M, Furukawa J, Shinohara Y, Aoshima M, Nishimura S-I (2010) Sialic acid-focused quantitative mouse serum glycoproteomics by multiple reaction monitoring assay. *Mol Cell Proteomics* **9**: 2354-2368.
18. Wahl M, Shukunami C, Heinzmann U, Hamajima K, Hiraki YJ, Imai KJ (2004) Transcriptome analysis of early chondrogenesis in ATDC5 cells induced by bone morphogenetic protein 4. *Genomics* **83:** 45-58
19. Chen L, Fink T, Zhang XY, Ebbesen P, Zachar V (2005) Quantitative transcriptional

profiling of ATDC5 mouse progenitor cells during chondrogenesis. *Differentiation* **73:** 350-363

20. Sekiya I, Vuoristo JT, Larson BL, Prockop DJ (2002) *In vitro* cartilage formation by human adult stem cells from bone marrow stroma defines the sequence of cellular and molecular events during chondrogenesis. *Proc Natl Acad Sci USA* **99:** 4397-4402
21. Xu J, Wang W, Ludeman M, Cheng KV, Hayami T, Lotz JC, Kapila S (2008) Chondrogenic differentiation of human mesenchymal stem cells in three-dimensional alginate gels. *Tissue Eng Part A* **14:** 667-680
22. Ng F, Boucher S, Koh S, Sastry KSR, Chase L, Lakshmipathy U, Choong C, Yang Z, Vemuri MC, Rao MS, Tanavde V (2008) PDGF, TGF-beta, and FGF signaling is important for differentiation and growth of mesenchymal stem cells (MSCs): transcriptional profiling can identify markers and signaling pathways important in differentiation of MSCs into adipogenic, chondrogenic, and osteogenic lineages. *Blood* **112:** 295-307
23. Lee HJ, Choi BH, Min BH, Park SR (2009) Changes in surface markers of human mesenchymal stem cells during the chondrogenic differentiation and dedifferentiation processes *in vitro*. *Arthritis Rheum* **60:** 2325-2332
24. Clouet J, Vinatier C, Merceron C, Pot-vaucel M, Maugars Y, Weiss P, Grimandi G, Guicheux J (2009) From osteoarthritis treatments to future regenerative therapies for cartilage. *Drug Discov Today* **14:** 913-925
25. Johnson K, Vaingankar S, Chen Y, Moffa A, Goldring MB, Sano K, Jin-Hua P, Sali A, Goding J, Terkeltaub R (1999) Differential mechanisms of inorganic pyrophosphate production by plasma cell membrane glycoprotein-1 and B10 in chondrocytes. *Arthritis Rheum* **42:** 1986-1997
26. Liang HT, Feng XC, Ma TH (2008) Water channel activity of plasma membrane affects chondrocyte migration and adhesion. *Clin Exp Pharmacol Physiol* **35**: 7-10
27. Przyborski SA, Knowles BB, Ackerman SL (1998) Embryonic phenotype of Unc5h3

mutant mice suggests chemorepulsion during the formation of the rostral cerebellar boundary *Development* **125:** 41-50

Chapter 4

Development of quantitative plasma *N*-glycoproteomics using label-free 2D-LC-MALDI MS for the biomarker discovery

4-1. Abstract

There has been rapid progress in the development of clinical proteomics methodologies with improvements in mass spectrometric technologies and bioinformatics, leading to many new methodologies for biomarker discovery from human plasma. However, it is not easy to find new biomarkers because of the wide dynamic range of plasma proteins and the need for their quantification. In this chapter, we report a new methodology for relative quantitative proteomics analysis combining large-scale glycoproteomics with label-free 2-D LC-MALDI MS. In this method, enrichment of glycopeptides using hydrazide resin enables focusing on plasma proteins with lower abundance corresponding to the tissue leakage region. On quantitative analysis, signal intensities by 2-D LC-MALDI MS were normalized using a peptide internal control, and the values linked to LC data were treated with DeView™ software. Our proteomics method revealed that the quantitative dynamic ranged from 10^2 to 10^6 pg/mL of plasma proteins with good reproducibility, and the limit of detection was of the order of a few ng/mL of proteins in biological samples. To evaluate the applicability of our method for biomarker discovery, we performed a feasibility study using plasma samples from patients with hepatocellular carcinoma.

4-2. Introduction

Proteomics approaches using human plasma samples are now being used to discover biomarkers against a number of clinical diseases[1-4]. However, it is very difficult to find new biomarkers from plasma samples by these proteomics approaches because the “classical proteins” account for 99% of total plasma proteins with a high dynamic range of 10 orders of magnitude[5]. In addition, the plasma proteins are known to have many variations of posttranslational modifications, such as phosphorylation, glycosylation, glycation, *etc*. Several effective procedures have been reported for improving these issues by removing high-abundance proteins from the plasma by affinity chromatography, *e.g.*, multiple affinity removal system (MARS) [6], and concentrating the proteins focusing on a specific modification of interest coupled with proteomics analysis [7-9].

Proteomics analysis is usually carried out by separating proteins or peptides prior to their quantification and identification by MS. Separation is mainly performed by gel-based techniques, such as 2-DE[10] and LC-based techniques, such as ionic exchange and reverse phase column chromatography[11]. LC-based separation is easier and exhibits high-throughput performance because samples globally digested with trypsin can be used and automated on-line LC systems can be constructed. In the case of plasma protein separation, gel-based technology has some disadvantages with respect to spot resolution because of the heterogeneity of posttranslational modifications of plasma proteins.

In some types of MS quantification, LC-separated peptides are quantified by labeling using stable isotopes, such as iTRAQ reagents[12]. Although this technology is very useful and effective, the labeling reagents are expensive and the labeling steps generate additional complexity, which reduces quantitative accuracy. Theoretically, a label-free method would be suitable for comprehensive analysis because sample preparation without labeling steps reflects the non-biased diversity of the original protein levels. There have been many previous reports regarding discovery of biomarkers by label-free methods using LC-ESI MS systems with

continuous data acquisition and analysis[13]. On the other hand, there have been only a few reports describing label-free quantification by LC-MALDI MS[14-16]. LC-MALDI MS methods are not used widely due to the discontinuous separation on LC (or LCs) needed to drop onto the MALDI target plate followed by integration and rearrangement of all of the intermittent MS data by analytical informatics. Thus, LC-MALDI MS methods have some issues for systematic and automatic high-throughput data analysis in contrast to LC-ESI MS methods. However, the peptides/proteins identified using MALDI MS are different from those determined by ESI MS[17-18] because of the differences in ionization type. Therefore, it may be possible to find unknown novel biomarkers when using LC-MALDI MS for biomarker discovery.

In this chapter, we developed a new methodology for large-scale *N*-glycoproteomics analysis combined with a quantitative label-free 2-D LC-MALDI MS system. Using this method, we analyzed plasma samples from patients with hepatocellular carcinoma as a feasibility study to evaluate its applicability for the discovery of new biomarkers in human plasma.

4-3. Results

Enrichment of N-glycopeptides from human plasma proteins

Glycoproteomics analysis using hydrazide resin[8,19,20] and lectin affinity capture[21,22] have been reported as applicable platforms for plasma proteomics to identify relatively low-abundance plasma proteins designated as tissue leakage proteins. To select the most useful method for enrichment of *N*-glycopeptides in our glycoproteomics analysis, we compared the capacity and efficiency of *N*-glycopeptide enrichment among these methods. Prior to this evaluation, some conditions of the respective methods were optimized. The *N*-glycopeptides, including the peptide samples eluted by these methods, were examined for identification of peptides by LC-ESI MS, and the *N*-glycosylation sites were assessed by conversion of asparagine into aspartic acid residues at the position of glycan attachment caused by the PNGase F-catalyzed deglycosylation reaction. The consensus amino acid sequence of *N*-glycosylation (Asn-X-Ser/Thr) was also confirmed. Figure 4-1 shows the number of peptides identified and the efficiency of *N*-glycopeptide identification among the three lectin affinity capture (ConA, WGA, and both in combination) and hydrazide resin methods.

Multilectin affinity capture methods are widely used for enrichment of *N*-glycopeptides because of their high efficiency. However, multilectin affinity capture was not effective in the present study, even with the capture of large numbers of *N*-glycopeptides. Thus, the hydrazide resin method led to the identification of the largest number of *N*-glycopeptides (94 *N*-glycopeptides). Furthermore, the number of non-glycopeptides identified (21 peptides) was lower in comparison to the lectin-based methods. Therefore, we adopted the hydrazide resin method for *N*-glycopeptide enrichment in our glycoproteomics analysis.

2-D LC-MALDI MS analysis focusing on N-glycopeptides

Figure 4-2 and 4-3 shows a schematic outline of our newly developed method of 2-D LC-MALDI MS coupled with the enrichment of *N*-glycopeptides. The present method was used to analyze samples of 100 μL of human plasma, which provided about 30 μg of *N*-glycopeptides. From this, 10 μg of *N*-glycopeptides was separated by two-dimensional LC using strong cation exchange chromatography followed by reverse-phase chromatography. The set of two chromatography modes was selected based on references as reported previously[19]. We also used a long silica-based monolithic column (400 mm) with a low flow rate of 3 μL/min for higher resolution on glycopeptide separation. This 2-D LC system finally generated 1520 fractions as spots on Prespotted AnchorChip MALDI target plates. These spots containing 10 fmol of angiotensin II as an internal standard were analyzed using MALDI MS. In this procedure, we were always able to obtain more than 3000 ion peaks derived from *N*-glycopeptides on the MS spectra.

To confirm the performance of the developed 2-D LC-MALDI MS system coupled with the enrichment of *N*-glycopeptides on human plasma samples, the peaks detected by the MALDI MS were identified by MS/MS. A total of 3031 MS peaks were observed and a subsequent Mascot search indicated 252 peptides as *N*-glycoproteins. Table 4-I shows a list of the identified *N*-glycoproteins other than the classical plasma proteins, *e.g.*, coagulation and complement factors, immunoglobulins, and apolipoproteins. Among the listed plasma proteins, some proteins that were present at $< 10^5$ pg/mL order by ELISA, such as cation-independent mannose-6-phosphate receptor (CD222; 2.2×10^3 pg/mL as the limit of detection in the present study), metalloproteinase inhibitor 1

(1.8×10^5 pg/mL), and hepatocyte growth factor-like protein (6.6×10^4 pg/mL), were included. These observations suggested that our method could detect and identify some plasma proteins at lower concentrations of the order of a few ng/mL, which were designated as tissue leakage proteins.

2-D LC-MALDI MS for plasma proteomics differential analysis

The present study was performed to develop a new method for label-free quantitative proteome analysis using 2-D LC-MALDI MS for low-abundance *N*-glycoproteins to enable the discovery of biomarkers. To perform quantitative differential analysis from the MS data of 2-D LC-MALDI MS described above, all MS data with the information of LC fractions and retention times were analyzed with DeView™ software (MCBI Inc.). The DeView™ software imported all ion signals obtained from MALDI MS after normalization using the intensity of angiotensin II as an internal standard in each spectrum, followed by data matching across all spectra by peak clustering analysis. Background and noise peaks were extracted and removed by DeView™ software. For good reproducibility of 2-D LC separation, the two separation columns were always checked in each analysis to reduce the shift of LC retention time within eight fractions among samples.

The quantitative properties of MALDI MS were confirmed using BSA tryptic digest peptides (Fig. 4-4a). Each signal of tryptic BSA digest was normalized using the internal standard (10 fmol of [Sar^1, Thr^8]-angiotensin II). The dynamic range depended on the sample amount ranging from 200 amol to 1 pmol. Reproducibility was confirmed by triplicate analysis of human plasma samples obtained from healthy volunteers. Figure 3-4b shows a scattergram of all signals detected and identified by 2-D

LC-MALDI MS with *N*-glycopeptide enrichment analysis. The distributions of detected signals showed a dynamic range of more than four orders of magnitude. In addition, the dispersion was unrelated to peak intensity, which contributed to the lower intensity peaks with reproducibility equal to the higher signals. The correlation coefficients were 0.873 and 0.849 for two samples of the triplicate data, and the average coefficient of variation (CV) of all signals was 34%. The reproducibility of our method was equivalent to those described in previous reports[16,23,24].

Feasibility study of biomarker discovery in hepatocellular carcinoma

To evaluate the applicability of the newly developed method for biomarker discovery from biological fluids, we performed a feasibility study using plasma samples from patients with hepatocellular carcinoma (HCC). Table 4-II shows the information for six patients with HCC and three healthy volunteers (NOR). These nine plasma samples were analyzed according to the method described above. A total of 5644 cluster peaks were detected in these samples after analysis using DeView™ software. To extract the peaks that were increased in HCC patients compared to NOR, the peaks detected in all HCC plasma samples were extracted (1826 peaks), and those showing threefold increases in intensity were selected (18 peaks). The peaks that were decreased in HCC patients were also extracted, and 2006 were confirmed. The peaks showing a threefold decreases in intensity were selected (27 peaks). After extraction by quantitative analysis, these peaks were identified based on MS/MS spectra data obtained by MALDI MS analysis. The representative proteins except those present in high abundance and those that were ambiguous are listed in Table 4-III. Alpha-1-antichymotrypsin and ceruloplasmin have been reported previously to be

present at elevated levels in blood samples of patients with HCC[25-29]. To quantitatively validate this molecule, we performed western blotting analysis using anti-alpha-1-antichymotrypin antibody. There was a good correlation between the normalized intensities of our method and the signals obtained by western blotting in all plasma samples from HCC and NOR (Fig. 4-5).

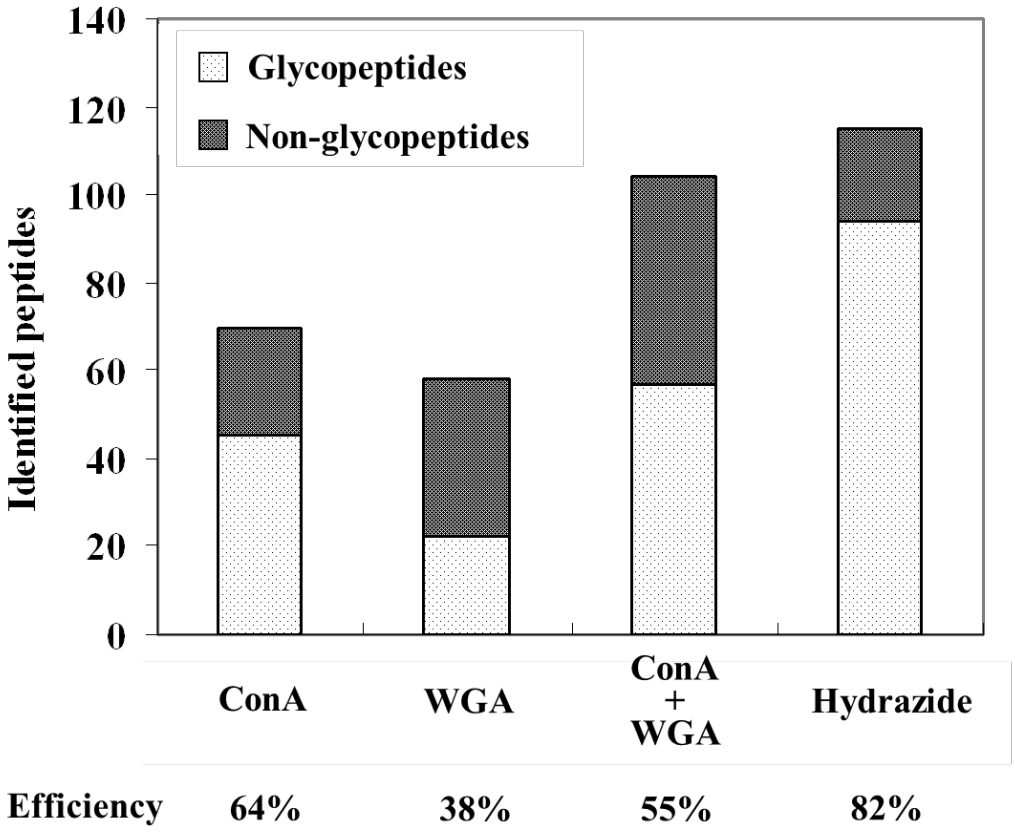


Figure 4-1. The number of identified peptides and enrichment efficiency of *N*-glycopeptides among three lectin affinity capture and hydrazide resin methods. The stippled bars indicate the number of identified *N*-glycopeptides and the shaded bars indicate the number of identified non-glycopeptides by LC-ESI MS. The percentage efficiency was calculated as the ratio of the number of identified *N*-glycopeptides to that of all identified peptides.

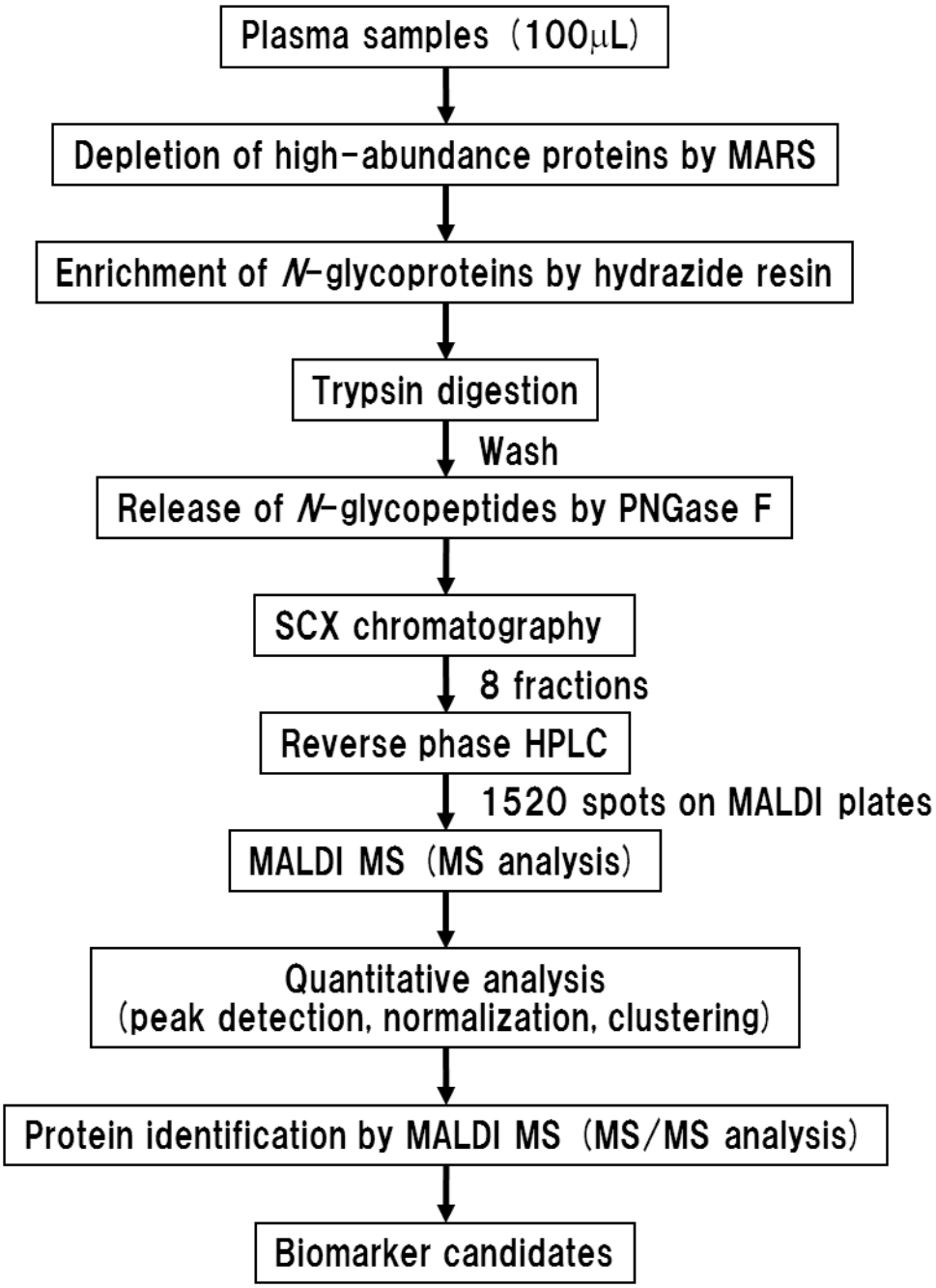
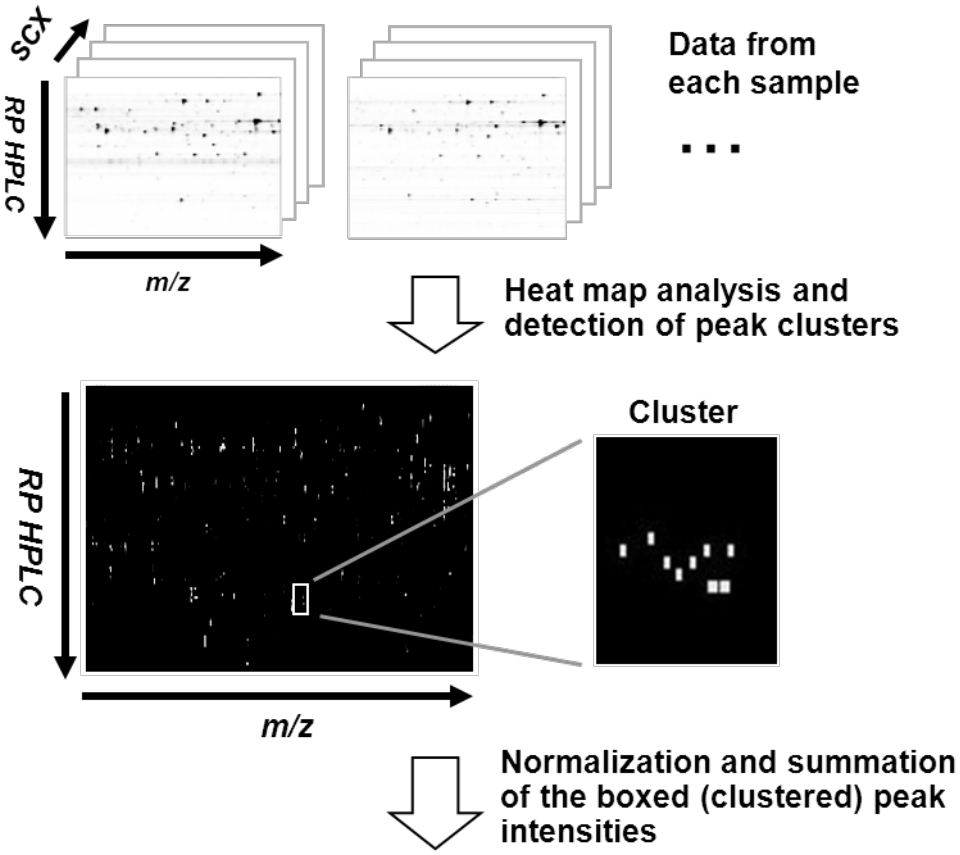


Figure 4-2. Schematic outline of our procedure: label-free 2-D LC-MALDI MS focused on plasma *N*-glycopeptides.



Cluster ID	Sample	Average Mass (*m/z*)	Normalized intensity
2214	A	1612.00	3,524
	B	1611.07	38,209
	C	1612.02	89,394
	:	:	:

Figure 4-3. Schematic illustration of the analytical flow for cluster analysis using DeView™ software. The peaks appearing in 2-D LC-MALDI MS data were picked up and grouped by cluster analysis using the parameters described in the *Experimental section*. The signal intensities were normalized relative to those of the internal standard for each mass spectrum and summed for each sample in each cluster. The obtained peak list was analyzed statistically for differential expression analysis.

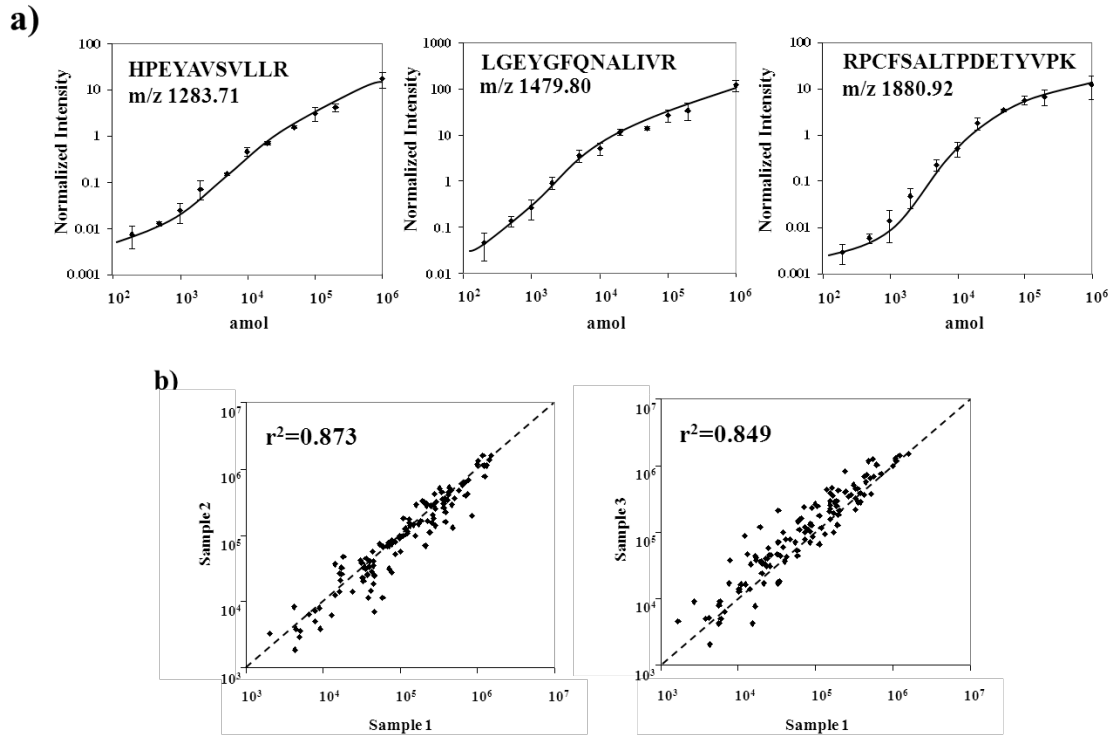


Figure 4-4. The quantitative properties of 2-D LC-MALDI MS. (a) Dynamic range of MS intensities from MALDI MS analysis. Data indicate the ion intensities obtained from BSA tryptic digest (ranging from 200 amol to 1 pmol). The signal intensities were normalized relative to 10 fmol of [Sar^{1}, Thr^{8}]-angiotensin II as an internal standard. (b) Scattergrams of all signals of the identified *N*-glycoproteins from a human plasma sample by 2-D LC-MALDI MS coupled with *N*-glycopeptide enrichment analysis. The plotted normalized ion intensities were obtained from triplicate analyses. The correlation coefficient (r-value) was calculated by Pearson's product-moment correlation coefficient analysis.

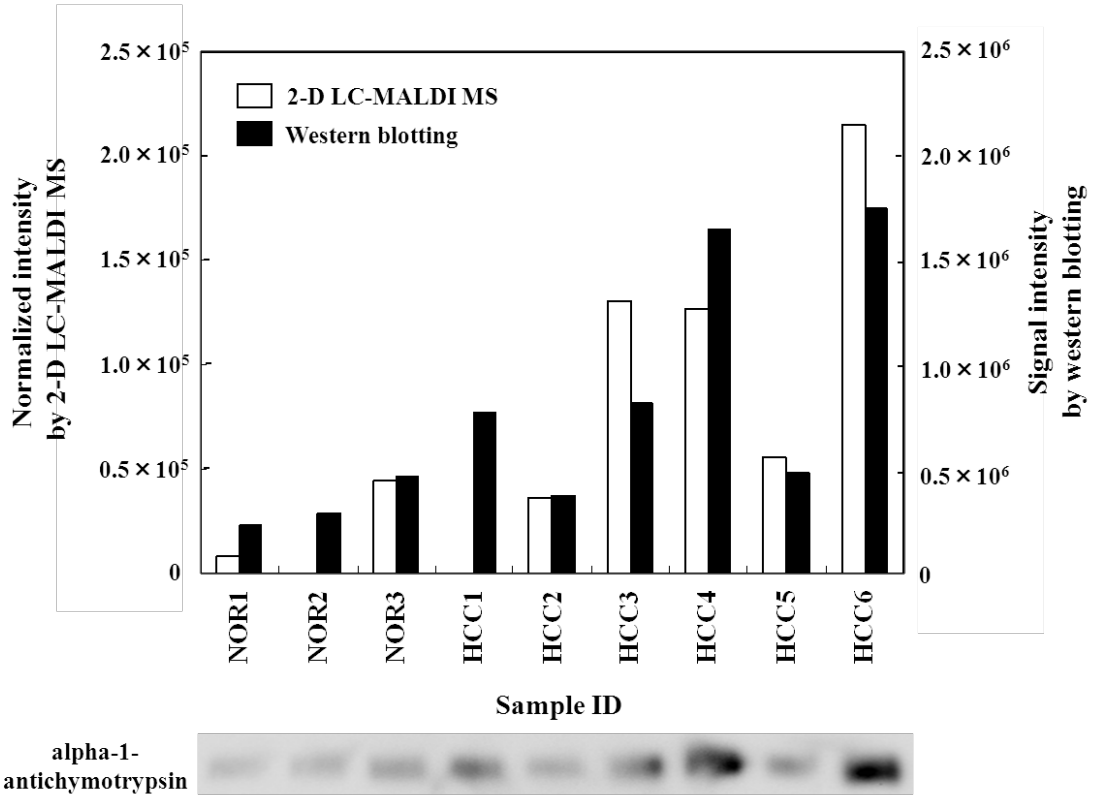


Figure 4-5. Comparison of the detected intensities of alpha-antichymotrypsin between the 2-D LC-MALDI MS method and western blotting analysis. The upper graph shows the normalized intensities by the present method (open bars) and the band intensities by western blotting (closed bars) obtained from representative plasma samples. The intensity by western blotting was calculated from the image of each band shown in the lower picture. Three normal volunteers and eight patients with HCC indicated in Table 4-II were examined.

Table 4-I. N-glycoproteins identified by our 2-D LC-MALDI MS focusing on N-glycopeptide

Protein ID	Protein name	Sequence	Mascot Score[b]	Concentration (pg/mL)[c]
Adhesion molecule[a]				
IPI00018305	Insulin-like growth factor-binding protein 3	AYLLPAPPAPGNASESEEDR	123	*2.5 × 10^6
IPI00020996	Insulin-like growth factor-binding protein complex acid labile chain	FVQAICEGDDCQPPAYTYNNITCASPPEVVGLDLR	64	
IPI00166729	alpha-2-glycoprotein 1, zinc	DIVEYYNDSNGSHVLQGR	117	
IPI00299059	Isoform 2 of Neural cell adhesion molecule L1-like protein (CD171)	DGEAFEINGTEDGR	75	
IPI00009030	Isoform LAMP-2A of Lysosome-associated membrane glycoprotein 2	VASVININPNTTHSTGSCR	78	
IPI00220737	Isoform N-CAM 120 of Neural cell adhesion molecule 1, 120 kDa isoform	IYNTPSASYLEVTPDSENDFGNYNCTAVNR	43	
IPI00218795	L-selectin (CD62L)	DNYTDLVAIQNK	59	*3.6 × 10^6
IPI00003813	Nectin-like protein 2	VSLTNVSISDEGR	64	
Receptor and membrane protein				
IPI00289819	Cation-independent mannose-6-phosphate receptor (CD222)	DAGVGFPEYQEEDNSTYNFR	75	*2.2 × 10^3
IPI00029260	Monocyte differentiation antigen CD14	LRNVSWATGR	34	**1.2 - 3.1 × 10^6
IPI00290328	Receptor-type tyrosine-protein phosphatase eta (CD148)	YNATVYSQAANGTEGQPQAIEFR	34	
IPI00163207	Isoform 1 of N-acetylmuramoyl-L-alanine amidase	GFGVAIVGNYTAALPTEAALR	119	
IPI00395488	Vasorin	LHEITNETFR	66	
IPI00791343	261 kDa protein (Multiple epidermal growth factor-like domains 8)	VNSTELFHVDR	57	
Extracellular matrix protein				
IPI00003351	Extracellular matrix protein 1	HIPGLIHNMTAR	39	
IPI00411462	fibronectin 1 isoform 7 preproprotein	HEEGHMLNCTCFGQGR	78	
IPI00023673	Galectin-3-binding protein	ALGFENATQALGR	122	
IPI00022418	Isoform 1 of Fibronectin	LDAPTNLQFVNETDSTVLVR	174	
IPI00020986	Lumican	LHINHNNLTESVGPLPK	126	
IPI00032292	Metalloproteinase inhibitor 1	FVGTPEVNQTTLYQR	63	*1.8 × 10^5
IPI00296099	Thrombospondin-1	VVNSTTGPGEHLR	104	*8.7 × 10^6
IPI00298971	Vitronectin	NNATVHEQVGGPSLTSDLQAQSK	173	
Protease and peptidase				
IPI00029193	Hepatocyte growth factor activator	CFLGNGTGYR	66	**4.0 × 10^5
IPI00654888	Kallikrein B, plasma (Fletcher factor) 1	IYPGVDFGGEELNVTFVK	81	
IPI00216882	mannan-binding lectin serine protease 1 isoform 3	FGYILHTDNR.T	41	
IPI00006114	Pigment epithelium-derived factor	VTQNLTLIEESLTSEFIHDIDR	111	**5.0 × 10^6
IPI00293057	Isoform 2 of Carboxypeptidase B2	QVHFFVNASDVDNVK	107	
IPI00221224	Aminopeptidase N (CD13)	NATLVNEADKLR	64	
IPI00064667	Carnosine dipeptidase 1	LVPHMNVSAVEK	37	
Enzyme				
IPI00218732	Serum paraoxonase/arylesterase 1	VTQVYAENGTVLQGSTVASVYK	121	
IPI00218413	biotinidase	NPVGLIGAENATGETDPSHSK	172	
IPI00013179	Prostaglandin-H2 D-isomerase	SVVAPATDGGLNLTSTFLR	62	
IPI00029061	Selenoprotein P	EGYSNISYIVVNHQGISSR	60	
IPI00299503	Phosphatidylinositol-glycan-specific phospholipase D 1	NLTTSLTESVDR	83	
IPI00017601	Ceruloplasmin	ENLTAPGSDSAVFFEQGTTR	158	**1.7-4.8 × 10^8
Growth factor, hormone, peptide etc…				
IPI00292218	Hepatocyte growth factor-like protein	GTANTTTAGVPCQR	85	*6.6 × 10^4
IPI00027235	Isoform 1 of Attractin	NHSCSEGQISIFR	100	
IPI00218460	Isoform 3 of Attractin	IDSTGNVTNELR	119	
IPI00219131	Isoform 1 of ICOS ligand (CD275)	TVVTYHIPQNSSLENVDSR	52	
IPI00032220	Angiotensinogen	HLVIHNESTCEQLAK	105	*3.7 × 10^7

[a] The identified proteins on the list were classified by molecular function as described in the Plasma Protein Database (http://www.plasmaproteomedatabase.org).

[b] $P < 0.05$ by Mascot Search.

[c] The concentrations in plasma are indicated as the values measured by ELISA (*, see Materials and Methods 2.6 or the previously reported values (**, [ref.12,19] or from product references of Quantikine Immunoassay Kits (R&D Systems, Minneapolis, MN)

Table 4-II. Information of the patients with HCC and normal volunteers

Category	Sample ID	Age at Excision	Sex	AJCC/UICC Stage Group
Hepatocellular carcinoma patients				
	HCC1	79	M	I
	HCC2	49	M	II
	HCC3	75	M	III
	HCC4	51	F	IV
	HCC5	56	M	IV
	HCC6	55	F	IV
Normal volunteers				
	NOR1	45	F	-
	NOR2	77	F	-
	NOR3	60	M	-

Table 4-III. Differentially expressed proteins in the plasma from patients with hepatocellular carcinoma

Cluster ID	Protein name	Sequence	Mascot score[a]	Ratio (HCC/NOR)[b]
1828	Afamin	YAEDKFNETTEK	83	33.1
3363	Ceruloplasmin	ELHHLQEQNVSNAFLDK	83	6.8
2214	Alpha-1-antichymotrypsin	LSLGAHNTTLTEILK	91	5.3
1651	Multimerin-1	KPTVNLTTVLIGR	115	4.6
5018	Prothrombin (Fragment)	SRYPHKPEINSTTHPGADLQENFCR	126	-4.5
5049	Isoform 1 of Attractin	GCSCFSDWQGPGCSVPVPANQSFWTR	93	-4.0
2773	Apolipoprotein B-100	VNQNLVYESGSLNFSK	130	-3.7
1477	Isoform 1 of Complement factor H	SHNMTTTLNYR	52	-3.7

[a] $P < 0.05$ by Mascot Search

[b] The ratios were calculated from the average of the normalized intensity generated from 2-D LC-MALDI MS.

4-4. Discussion

At present, proteomics analysis for biomarker discovery is mainly performed using two techniques that are based on two-dimensional gel electrophoresis (2-DE) and mass spectrometry coupled with multi-dimensional LC separation. The 2-DE technique is widely used for comprehensive analysis of samples from tissues and cells in culture but is not applicable for serum and plasma samples. One possible reason for this is the heterogeneous posttranslational modifications of some high-abundance proteins in the blood, which can mask low-abundance proteins of interest. To discover meaningful biomarkers by 2-DE, Pieper *et al*. performed extensive fractionation of human serum samples and then analyzed 66 2-DE gels per sample[30]. This procedure, although effective for identifying proteins present at low levels (ng/mL order) in the blood, is laborious and is not a practical means of finding biomarkers in multiple samples.

Two strategies are generally used in the LC-MS technique—one based on ESI MS and another based on MALDI MS. On discovering biomarkers in blood, the method using the ESI MS system has mainly been applied. However, the principles of ionization differ between ESI MS and MALDI MS, and thus the proteomics approach using MALDI MS may find new biomarkers that cannot be detected using ESI MS systems. In the present study, we developed a new label-free method based on MALDI MS for discovery of novel biomarkers in human plasma. Our method consists of two steps, *i.e.*, enrichment of *N*-glycopeptides and quantitative 2-D LC-MALDI MS analysis. Glycoproteomics analysis using hydrazide resin[8,19,20] or lectin affinity capture[21,22] coupled with LC-ESI MS has also been reported to identify tissue leakage proteins and to be useful as a platform for plasma proteomics. Here, we first compared two conventional methods for enrichment of *N*-glycopeptides, and found that the technique using hydrazide resin was the more effective of the two methods (Figure 4-1). The hydrazide resin permits more thorough washing because of the covalent hydrazone bonds compared to the lectin affinity capture method, which is limited with respect to binding specificity and treatment. Further optimization of the hydrazide resin methodology (washing and elution conditions, *etc*.)

made it possible to detect and identify plasma proteins corresponding to the tissue leakage region (ng/mL orders) as shown in Table 4-I.

Two approaches are available for differential analysis using LC-MS—a stable isotope labeling procedure and a label-free procedure. Patel *et al*. compared the two approaches using bacterial proteins[31], and reported that they were able to identify and quantify many different proteins. They demonstrated that the label-free procedure is advantageous in terms of sample volume, sample preparation, and handling time. Label-free quantification has a number of advantages, but has not been widely used, especially in combination with MALDI MS. One reason is the lack of commercial software for LC-MALDI MS[14,32], and it is difficult to analyze the complicated data from multidimensional liquid chromatography and large numbers of mass spectra. To overcome this problem, we used the new software DeView™ for data informatics for label-free quantification. This software, specifically developed for LC-MALDI MS analysis, calculates normalized peak intensity and matches the number of peaks across different spectra for easy operation. We further optimized the parameters of DeView™ before utilizing it for comprehensive analysis, and the experimental conditions in the sample preparation steps were optimized to achieve sufficient reproducibility for quantitative comparison. The results indicated that the quantitative data from all steps from sample preparation to informatics led to suitable performance (Figure 4-4). The correlation coefficients were 0.873 and 0.849 calculated from between two analysis of the triplicate, and the mean of CV values from all signals was 34%. This reproducibility was equivalent to that of label-free quantitative proteomics using the LC-ESI MS system reported by Yamada *et al*. [33]. Further, the mean values from the previously reported LC-MALDI MS methods were 25% – 35%[16,23,24]; therefore, our developed LC-MALDI MS method including *N*-glycoprotein enrichment steps had sufficient performance for comprehensive analysis to discover plasma biomarkers.

Using the method described here, we performed a feasibility study using plasma samples with HCC to evaluate its applicability for discovery of new biomarkers from biological fluids.

HCC is one of the leading causes of cancer death worldwide, and early diagnosis is considered important for success of clinical treatment. In the present study, many differentially expressed *N*-glycoproteins were found in the plasma samples from HCC patients, and we identified eight plasma biomarker candidates, as shown in Table 4-III. Among these candidates, elevated levels of alpha-1-antichymotrypsin and ceruloplasmin in HCC patients have been reported previously. Western blotting analysis was performed to verify the quantitative results of proteomics analysis. The results shown in Figure 4-5 indicated that the levels of alpha-1-antichymotrypsin protein were elevated in the plasma of HCC patients, and almost corresponded with the quantitative results of proteomics analysis. However, there were some differences between the results of the two quantitative analyses in a few plasma samples (*i.e.*, HCC1 and HCC3). The quantitative results obtained from the present *N*-glycoproteomics analysis included both qualitative (changes in extent of glycosylation modification) and quantitative (changes in protein expression levels) variations. Changes in extent of glycosylation modification are well-known in cancer diseases, and some of these modifications are associated with malignant transformation, cell proliferation, invasion, and metastases[34]. Serum alpha-fetoprotein (AFP) is widely used as a tumor marker for HCC with an FDA-approved diagnostic test[35,36]. Recently, it was suggested that it would be better to use AFP-L3, an isoform of AFP bound to a lectin of *Lens culinaris* agglutinin. Changes in *N*-glycosylation can be associated with the progression of oncogenic transformation. In addition, Comunale *et al.* performed proteomics analysis of fucosylated glycoproteins in plasma from patients with HCC using LC-ESI MS[26]. The biomarker candidates found in the present study must be verified for changes in protein expression level and/or extent of glycosylation modification by other types of quantitative analyses (*e.g.*, western blotting, ELISA) using larger samples. However, the present method showed suitable characteristics as the first step in the discovery of effective biomarkers. Indeed, multimerin-1 was shown to be of interest in the present study but has not been reported as a biomarker for HCC previously. Multimetrin-1 has also not been reported as a plasma protein but is found in vascular tissues of the placenta, lung,

and liver. The findings of *N*-glycoprotein profiling analysis suggested that the level of the multimerin-2 isoform may be increased in lung adenocarcinoma[37]. This suggests that multimerin-1 may also be increased in the plasma of HCC patients and may be linked to the progression of oncogenic transformation.

In conclusion, this chapter described improvement of quantitative proteomics analysis combining large-scale glycoproteomics with label-free 2-D LC-MALDI MS and a feasibility study to evaluate its usefulness for biomarker discovery in human plasma.

4-5. Experimental section

Chemicals

Dithiothreitol, iodoacetamide, and formic acid were purchased from Wako Pure Chemical Industries (Osaka, Japan). RapiGest SF was from Waters (Milford, MA) and sequencing-grade trypsin was from Promega (Madison, WI). PNGase F was purchased from New England Biolabs Inc. (Ipswich, MA). Ammonium carbonate and [Sar^1, Thr^8]-angiotensin II were purchased from Sigma-Aldrich (St. Louis, MO). Alpha-1-antichymotrypsin monoclonal antibody was from LifeSpan Biosciences (Seattle, WA). Blocking reagent and horseradish peroxidase-conjugated anti-mouse IgG antibody were from GE Healthcare (Uppsala, Sweden). HPLC-grade acetonitrile was purchased from Merck Chemicals (Darmstadt, Germany). Ammonium acetate was from Kanto Chemical (Tokyo, Japan). Trifluoroacetic acid was obtained from Nacalai Tesque (Kyoto, Japan). Other chemicals were purchased from Sigma-Aldrich.

Human plasma samples

Human EDTA-plasma samples from healthy volunteers were used for development and evaluation of the enrichment of *N*-glycopeptides and 2-D LC-MALDI MS. To perform a feasibility study, human EDTA-plasma samples from six patients with hepatocellular carcinoma (HCC) and three healthy volunteers were purchased from Asterand Plc. (Royston, UK). All plasma samples (100 μL) were passed through MARS-human 7 LC columns (4.6 × 100 mm; Agilent Technologies, Santa Clara, CA) using ÄKTAexplorer 10S system (GE Healthcare) to deplete the seven high-abundance proteins in plasma: albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, and fibronectin.

Enrichment of N-glycopeptides

An Affi-Gel Hz Immunoaffinity Kit (Bio-Rad Laboratories, Hercules, CA) was used for

enrichment of *N*-glycopeptides. The MARS-depleted plasma samples were diluted to 500 μL with the coupling buffer and oxidized with 10 mM sodium periodate at room temperature for 1 h in the dark. Excess oxidizer was removed using an equilibrated NAP-5 column (GE Healthcare) with the coupling buffer. The oxidized proteins were collected and allowed to react with 100 μL of hydrazide resin slurry at room temperature overnight. The resin was washed three times each with methanol and 8 M urea in 100 mM NH_4HCO_3. The reduction reaction was performed in 50 mM dithiothreitol at 37°C for 90 min and the alkylation reaction in 25 mM iodoacetamide at 37°C for 30 min, followed by exposure to 100 mM NH_4HCO_3 containing 8 M urea in the dark. After washing the resin three times with 1 M urea in 100 mM NH_4HCO_3, the reduced and alkylated *N*-glycoproteins on the resin were digested with 10 μg of trypsin in 100 mM NH_4HCO_3 containing 1 M urea and 0.1% RapiGest SF at 37°C overnight. The resin was washed three times with 2 M sodium chloride, methanol, and 50 mM NH_4HCO_3, the *N*-glycopeptides bound to the hydrazide resin were incubated with 500 units of PNGase F in 50 mM NH_4HCO_3 at 37°C overnight. The absorbance of the eluted fractions was measured at 280 nm, and they were then dried and stored at –20°C until analysis.

Enrichment of *N*-glycopeptides by a lectin affinity capture method was performed using a ConA Glycoprotein Isolation Kit (Pierce, Rockford, IL) and WGA Glycoprotein Isolation Kit (Pierce) according to the manufacturer’s protocols. The *N*-glycopeptides were also dried and stored at –20°C until use.

Protein identification by LC-ESI MS

The prepared samples of *N*-glycopeptides were analyzed using an in-house optimized nano-flow HPLC system driven by a DiNa nano HPLC pump (KYA TECH Corp., Tokyo, Japan) coupled online with a Q-Tof Ultima API quadrupole/time-of-flight mass spectrometer (Micromass-Waters, Manchester, UK) through a nano-LC probe (ESI). The reverse-phase monolithic capillary column (50 μm i.d. × 300 mm) was prepared by Kyoto Monotech (Kyoto,

Japan). The mobile phases consisted of 2% acetonitrile (0.2% formic acid) in water (A) and 98% acetonitrile (0.2% formic acid) in water (B). After loading peptides onto the column, the mobile phase was held at 90% A and 10% B for 5 min. Exponential gradient elution was performed by increasing the mobile phase composition from 10% to 60% B over 50 min at a flow rate of 120 nL/min. To identify the eluting peptides, the quadrupole/time-of flight mass spectrometer was operated in data-dependent MS/MS mode (m/z 350 – 1300), in which a full MS scan was followed by four MS/MS scans. The acquired product ion spectra were identified using the Mascot search engine (Matrix Science, Tokyo, Japan) against the human International Protein Index (IPI) database. The following dynamic modifications were considered: carbamidomethylation of cysteine, oxidation of methionine, and PNGase F-catalyzed conversion of asparagine to aspartic acid at the site of carbohydrate attachment.

2-D LC-MALDI MS

The prepared samples of *N*-glycopeptides were separated using a dual gradient capillary LC system (Ultimate™ 3000; Dionex, San Francisco, CA). The LC system was controlled by Hyster software (Bruker Daltonics, Bremen, Germany). For separation in the 1st dimension, a PolySULFOETHYL A column (1.0 mm i.d. × 150 mm; PolyLC Inc., Columbia, MD) was used for strong cation exchange (SCX) chromatography. The separations were performed at a flow rate of 5 μL/min and the mobile phases consisted of 0.1% formic acid in 2% acetonitrile (A), 400 mM ammonium acetate in solvent A (B), and 1 M ammonium acetate in solvent A (C). After 10 μL (containing 10 μg peptides) of sample was loaded onto the column, the peptides were separated using a step-wise gradient of 0, 6, 8, 10, 12, 16, 20, and 800 mM ammonium acetate. Each fraction of the eluate from the SCX column was loaded onto a reverse-phase capillary trap column (200 μm i.d. × 100 mm; Kyoto Monotech). The peptides trapped in this silica-based monolithic column were desalted and separated by chromatography in the second dimension on a reverse-phase capillary column; this was a silica-based monolithic column (200

μm i.d. × 400 mm; Kyoto Monotech). The separations were performed at a flow rate of 3 μL/min and the mobile phases consisted of 0.1% TFA and 2% acetonitrile (A) and 0.1% TFA and 90% acetonitrile (B). The peptides were separated using a linear gradient from 2% to 46% acetonitrile over 95 min. The eluate was fractionated on 384 Prespotted AnchorChip MALDI target plates (Bruker Daltonics) at 30-s intervals. For each sample, 190 fractions were spotted for the first-dimension SCX column chromatography, and hence the total spot number was 1520 per plasma sample. Next, 10 fmol of [Sar^1, Thr^8]-angiotensin II was applied to all fractions of the 384 Prespotted AnchorChip as an internal standard before fraction collection. The target plates were washed with ice-cold 0.1% TFA containing 10 mM ammonium carbonate. The dried target plates were stored under nitrogen gas until MALDI MS analysis.

The glycopeptides on Prespotted AnchorChip target plates were analyzed using an Ultraflex II MALDI TOF/TOF mass spectrometer (Bruker Daltonics). MS-mode acquisitions consisted of 1000 laser shots accumulated from 20 sample positions and the laser strength was constant for all samples. The data from the mass spectra were input into DeView™ (MCBI Inc., Ibaraki, Japan). The peaks on all mass spectra were detected using the following parameters: mass range, 800 – 5000; maximum number of iterations, 50; correlation threshold, 0.95; minimum S/N ratio, 1000. The signal intensities were normalized relative to those of the internal standard for each mass spectrum. The peaks were grouped by cluster analysis using the following parameters: mass range, 800 – 5000; intensity cut-off, 1000; WindowSizeMpz, 0.8; WindowSizeFraction, 8. The obtained data were analyzed statistically for differential expression analysis. Figure 4-3 shows a schematic outline of the procedure for differential expression analysis based on MS data using DeView™ software with representative results obtained in the present study.

Enzyme-linked immunosorbent assay (ELISA)

The concentrations of *N*-glycoproteins in human plasma samples were measured with commercially available ELISA kits as follows: insulin-like growth factor-binding 3,

thrombospondin-1, hepatocyte growth factor-like protein, cation-independent mannose-6-phosphate receptor (R&D Systems, Minneapolis, MN); metalloproteinase inhibitor 1 (Invitrogen, Carlsbad, CA); L-selectin (Biovender Laboratory Medicine, Modrice, Czech Republic); angiotensinogen (Immuno-Biological Laboratories, Gunma, Japan). The assays were performed according to the respective manufacturer's protocol, and the results were determined in duplicate using EDTA-plasma samples from healthy volunteers as controls.

Western blotting

MARS-depleted plasma proteins were used. After separation by SDS-PAGE on 10% – 20% polyacrylamide gradient gels (DRC Co., Ltd., Tokyo, Japan) , the proteins was transferred onto PVDF membranes, which were blocked, and incubated with alpha-1-antichymotrypsin monoclonal antibody (1:1000 dilution) and horseradish peroxidase-conjugated anti-mouse IgG antibody (1:5000 dilution) according to standard procedures. Immunodetection was performed using ECL-plus reagent (GE Healthcare). Signals were visualized using an image reader (LAS-4000; Fuji Film, Tokyo, Japan).

4-6. References

1. Service RF. (2008) Proteomics. Proteomics ponders prime time. *Science* **321**: 1758-61.
2. Beretta L. (2009) HUPO Highlights. *Proteomics* **21**: 4830-3.
3. Archakov A, Bergeron JJ, Khlunov A, Lisitsa A, Paik YK. (2009) The Moscow HUPO Human Proteome Project workshop. *Mol. Cell. Proteomics* **9**:2199-2200.
4. Anderson NL, Anderson NG, Pearson TW, Borchers CH, Paulovich AG, Patterson SD, Gillette M, Aebersold R, Carr SA. (2009) A human proteome detection and quantitation project. *Mol. Cell. Proteomics* **5**: 883-6.
5. Anderson NL, Anderson NG. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1**: 845-67.
6. Echan LA, Tang HY, Ali-Khan N, Lee K, Speicher DW. (2005) Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* **5**: 3292-303.
7. Schiess R, Wollscheid B, Aebersold R. (2009) Targeted proteomic strategy for clinical biomarker discovery. *Mol. Oncol.* **3**: 33-44.
8. Zhang H, Li XJ, Martin DB, Aebersold R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **6**: 660-666.
9. Lee HJ, Na K, Kwon MS, Kim H, Kim KS, Paik YK. (2009) Quantitative analysis of phosphopeptides in search of the disease biomarker from the hepatocellular carcinoma specimen. *Proteomics* **9**: 3395-3408.
10. Kim MR, Kim CW. (2007) Human blood plasma preparation for two-dimensional gel electrophoresis. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **849**: 203-210.
11. Qian WJ, Jacobs JM, Liu T, Camp DG 2nd, Smith RD. (2006) Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol. Cell. Proteomics* **10**: 1727-1744.

12. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics.* **3**: 1154-1169.
13. Mueller LN, Brusniak MY, Mani DR, Aebersold R. (2008) An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J. Proteome Res.* **7**: 51-61.
14. Mueller DR, Voshol H, Waldt A, Wiedmann B, Van Oostrum J. (2007) LC-MALDI MS and MS/MS--an efficient tool in proteome analysis. *Subcell. Biochem.* **43**: 355-380.
15. Zhen Y, Xu N, Richardson B, Becklin R, Savage JR, Blake K, Peltier JM. (2004) Development of an LC-MALDI method for the analysis of protein complexes. *J. Am. Soc. Mass Spectrom.* **15**: 803-822.
16. Neubert H, Bonnert TP, Rumpel K, Hunt BT, Henle ES, James IT. (2008) Label-free detection of differential protein expression by LC/MALDI mass spectrometry. *J. Proteome Res.* **7**: 2270-2279.
17. Bodnar WM, Blackburn RK, Krise JM, Moseley MA. (2003) Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J. Am. Soc. Mass. Spectrom.* **14**: 971-979.
18. Mollé D, Jardin J, Piot M, Pasco M, Léonil J, Gagnaire V. (2009) Comparison of electrospray and matrix-assisted laser desorption ionization on the same hybrid quadrupole time-of-flight tandem mass spectrometer: application to bidimensional liquid chromatography of proteins from bovine milk fraction. *J. Chromatogr. A.* **1216**:2424-2432.
19. Liu T, Qian WJ, Gritsenko MA, Camp DG 2nd, Monroe ME, Moore RJ, Smith RD. (2005) Human plasma *N*-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J. Proteome Res.* **4**: 2070-2080.
20. Chen R, Jiang X, Sun D, Han G, Wang F, Ye M, Wang L, Zou H. (2009) Glycoproteomics

analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J. Proteome Res.* **8**: 651-661.

21. Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, Kasai K, Takahashi N, Isobe T. (2003) Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat Biotechnol.* **21**: 667-672.
22. Madera M, Mechref Y, Klouckova I, Novotny MV. (2006) Semiautomated high-sensitivity profiling of human blood serum glycoproteins through lectin preconcentration and multidimensional chromatography/tandem mass spectrometry. *J. Proteome Res.* 5: 2348-2363.
23. Tammen H, Schulte I, Hess R, Menzel C, Kellmann M, Mohring T, Schulz-Knappe P. (2005) Peptidomic analysis of human blood specimens: comparison between plasma specimens and serum by differential peptide display. *Proteomics* **5**: 3414-3422.
24. Hattan SJ, Parker KC. (2006) Methodology utilizing MS signal intensity and LC retention time for quantitative analysis and precursor ion selection in proteomic LC-MALDI analyses. *Anal. Chem.* **78**: 7986-7996.
25. Pousset D, Piller V, Bureaud N, Piller F. (2001) High levels of ceruloplasmin in the serum of transgenic mice developing hepatocellular carcinoma. *Eur. J. Biochem.* **268**: 1491-1499.
26. Comunale MA, Lowman M, Long RE, Krakover J, Philip R, Seeholzer S, Evans AA, Hann HW, Block TM, Mehta AS. (2006) Proteomic analysis of serum associated fucosylated glycoproteins in the development of primary hepatocellular carcinoma. *J. Proteome Res.* **5**: 308-315.
27. Ordóñez NG, Manning JT Jr. (1984) Comparison of alpha-1-antitrypsin and alpha-1-antichymotrypsin in hepatocellular carcinoma: an immunoperoxidase study. *Am. J. Gastroenterol.* **79**: 959-963.
28. Xu KC, Wei Q, Meng XY. (1989) Serum alpha-1-antitrypsin and alpha-1-antichymotrypsin in the diagnosis of primary hepatocellular carcinoma. *Chin. Med. J.***102**: 834-838.

29. Matsuzaki S, Iwamura K, Itakura M, Kamiguchi H, Katsunuma T. (1981) A clinical evaluation of serum alpha-1-antichymotrypsin levels in liver disease and cancers. *Gastroenterol. Jpn.* **16**: 582-591.

30. Pieper R, Gatlin CL, Makusky AJ, Russo PS, Schatz CR, Miller SS, Su Q, McGrath AM, Estock MA, Parmar PP, Zhao M, Huang ST, Zhou J, Wang F, Esquer-Blasco R, Anderson NL, Taylor J, Steiner S. (2003) The human serum proteome: display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. *Proteomics* **3**:1345-1364.

31. Patel VJ, Thalassinos K, Slade SE, Connolly JB, Crombie A, Murrell JC, Scrivens JH. (2009) A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J. Proteome Res.* **8**: 3752-3759.

32. Vandenbogaert M, Li-Thiao-Té S, Kaltenbach HM, Zhang R, Aittokallio T, Schwikowski B. (2008) Alignment of LC-MS images, with applications to biomarker discovery and protein identification. *Proteomics* **8**: 650-672.

33. Ono M, Shitashige M, Honda K, Isobe T, Kuwabara H, Matsuzuki H, Hirohashi S, Yamada T. (2006) Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectrometry. *Mol. Cell. Proteomics* **5**: 1338-1347.

34. Dennis JW, Laferté S, Waghorne C, Breitman ML, Kerbel RS. (1987) Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. *Science*. **236**: 582-585.

35. Wright LM, Kreikemeier JT, Fimmel CJ. (2007) A concise review of serum markers for hepatocellular cancer. *Cancer Detect. Prev*. **31**: 35-44.

36. Gomaa AI, Khan SA, Leen EL, Waked I, (2009) Taylor-Robinson SD. Diagnosis of hepatocellular carcinoma. *World J. Gastroenterol.* **15**: 1301-1314.

37. Soltermann A, Ossola R, Kilgus-Hawelski S, von Eckardstein A, Suter T, Aebersold R, Moch H. (2008) *N*-glycoprotein profiling of lung adenocarcinoma pleural effusions by

shotgun proteomics. *Cancer.* **114**:124-133.

Chapter 5

Potential plasma biomarkers for progression of knee osteoarthritis

5-1. Abstract

Although osteoarthritis (OA) is the most prevalent joint disease in industrialized countries, effective treatments for the disease have not yet been established. This is partly ascribed to the lack of reliable biomarkers for OA, particularly those which predict disease progression. In this chapter, we attempted to find factors the amounts of which significantly change in association with the progression of knee OA. A total of 68 subjects with primary knee OA were enrolled in the study. These subjects were followed up over an 18-month period, and plasma and serum samples were obtained during that period at 0, 6, 12 and 18 months, together with knee radiographs. From the changes on radiographs, 3 subjects were found to undergo disease progression during the study period (progressors). A total of 6 plasma and serum samples obtained from these subjects were subjected to the analysis, together with another 6 samples from other 3 subjects who were confirmed to be in a stable condition at the time of sample acquisition (non-progressors).

These samples were subjected to *N*-glycoproteomic 2D-LC-MALDI analysis (described in *Chapter 4*), in which more than 3000 MS peaks were identified. Among them, 4 peaks were chosen to represent potential prognostic biomarkers that significantly increased during disease progression. MS/MS analysis revealed that these peaks represented clusterin, hemopexin, alpha-1 acid glycoprotein-2, and macrophage stimulating protein, respectively. The expression of these molecules was confirmed in synovial tissues collected from 10 OA knees by *in situ* hybridization. For clusterin and hemopexin, the expression in the OA synovium was also confirmed by Western blotting analysis. In addition, the protein levels of hemopexin and clusterin in serum were measured by ELISA, which revealed that the amount of either protein

did not differ significantly between progressors and non-progressors.

In conclusion, 4 potential biomarkers the concentrations of which may associate with the progression of knee OA were identified through *N*-glycoproteomic analysis. For clusterin and hemopexin, the change in the glycosylation level with the progression of the disease was suggested. This is the first report regarding the use of glycoproteomic technology in exploring potential biomarkers for knee OA.

5-2. Introduction

Knee OA is a highly heterogeneous disease in terms of progression. Previous studies have indicated that the course of progression of knee OA differs considerably among patients and knee joints; some OA knees may undergo rapid progression, while others remain almost unchanged for years or decades[1-3]. To establish the optimum treatment protocol and to facilitate drug development, it is important to identify patients at risk of disease progression. Investigators have attempted to establish reliable biomarkers that can predict the progression of the disease [4]. However, despite these efforts, no surrogate protein biomarkers have been used clinically, primarily due to the lack of predictability. Therefore, it is necessary to establish reliable biomarkers for OA.

Proteomic analysis coupled with mass spectrometry (MS) is a powerful tool for exploring biomarkers, and this approach has also been used in the discovery of OA biomarkers[5]. Attempts to discover biomarkers in plasma are often hampered by the paucity of target proteins, because plasma contains a large variety of proteins at various concentrations ranging over ten orders of magnitude[6]. Various techniques have been developed to overcome this problem. In *Chapter 4*, we demonstrated that glycoproteomic analysis is a powerful tool for finding novel biomarkers in plasma. This method may be used for the exploration for OA biomarkers, because matrix components of articular cartilage, such as proteoglycans, are known to be highly glycosylated[7].

In this chapter, a glycoproteomic approach using 2D-LC-MALDI was used to discover factors the concentration of which changes with the progression of knee OA. Recently, we conducted a follow-up study of knee OA patients, and confirmed that the course of radiographic progression of knee OA differed considerably among patients[8]. Using plasma obtained from these subjects, we compared the samples from the subjects who were in the middle of disease progression with those from the subjects who were in a stable condition of the disease. Four proteins were found to fulfill our criteria, which we expect to serve as prognostic factors for the disease.

5-3. Results

Exploration of biomarker candidates by glycoproteomic analysis

As described earlier, the purpose of this study was to find factors the amounts of which increase or decrease in association with disease progression, expecting that such factors would serve prognostic biomarkers for knee OA. Thus, the factors we attempted to find were termed biomarker candidates hereafter. To find biomarker candidates, *N*-Glycoproteomic 2D-LC-MALDI analysis was carried out using approximately 30 μL of plasma samples. Prior to this study, we evaluated the performance of our proteomic analysis with human plasma samples in *Chapter 4*, which demonstrated excellent sensitivity (down to 10^3 pg/mL) and sufficient reproducibility (CV < 34%) of the system.

Glycoproteomic analysis yielded 3456 and 3357 MS peaks in set #1 and set #2 samples, respectively, using DeView software. We chose the peaks for potential biomarkers by the following three selection criteria (Figure 5-1). First, for each MS peak, linear regression analysis was performed between the rate of narrowing of joint space width (JSW; this indicates the rate of disease progression) and the intensity of the peak, and the peak was taken as a candidate when the correlation coefficient (r) was above 0.7. By these criteria, 435 and 350 peaks were chosen for set #1 and set #2 samples, respectively. Next, the MS peaks were chosen with the condition that the average of the intensity for the progressors was over two-fold greater than that for the non-progressors, and that the difference in intensity was statistically significant between the progressors and the non-progressors ($P < 0.05$ by Mann–Whitney U test). Through this selection, 176 and 92 MS peaks were chosen in set #1 and set #2 samples, respectively. Next, we chose peaks found in common in both set #1 and set #2 samples, and four MS peaks were finally selected for further analysis.

Identification of biomarker candidates

The glycoproteins represented by the four selected MS peaks were subjected to

MALDI-TOF/TOF MS analysis for identification (Table 5-I). Plasma samples collected from the progressors were used for the analysis, after confirming that the peaks had sufficient intensities for MS/MS analysis. The obtained MS/MS data were analyzed using Mascot software, and the reliability of identification was confirmed by the significant high values of the scores. The data were also analyzed by the MS-Tag software (ProteinProspector v 5.10.1; http://prospector.ucsf.edu/prospector/mshome.htm), and identical results were obtained. In the search using the Mascot software, the possibility was considered a possibility of conversion of asparagine to aspartic acid by deglycosylation with PNGase F, which causes a mass increase of 0.98 Da. To confirm the identification of the protein, the analysis was repeated at least twice for each MS peak. After identification, the presence of the consensus Asn-X-Thr/Ser motif of *N*-linked glycopeptides was confirmed with each determined sequence (Table 5-I).

Among the four identified proteins, a peptide fragment of $G^{235}HGHRN^{*}GTGHGN^{*}STHHGPEYMR^{256}$ (*m/z* 2,414.1 with two glycosylated asparagine residues and an oxidized methionine residue) was found to have originated from hemopexin. Two *N*-glycosylation motifs were found within this fragment. For this fragment, the peak intensities of progressors were 19.3-fold and 10.0-fold greater than those of non-progressors in the set #1 and set #2 samples, respectively. This difference was greatest among the four biomarker candidates found in this study, suggesting that hemopexin may be a promising biomarker for the disease.

A peptide fragment of $L^{26}VPVPITN^{*}ATLDRITGK^{42}$ (*m/z* 1,809.1 with one glycosylated asparagine residue) was identified to be a peptide generated from alpha-1 acid glycoprotein-2 (AGP-2). The N-terminal amino acid sequence of the mature AGP-2 molecule was Q^{19}IPLCANLVPVPITNATLDRITGK... according to the UniProt/Swiss-Prot database (version 117). Thus, the identified AGP-2 fragment lacked the first seven amino acids at the N-terminus. Although the samples underwent trypsin digestion prior to the analysis, this treatment was unlikely to cause the loss of the seven amino acids, because trypsin cleaves proteins only at the

C-terminus of arginine (R) or lysine (K). Therefore, the identified AGP-2 fragment was considered to be generated *in vivo* by the processing of an endogenous protease. This fragment was identified with a sufficiently high score by Mascot search, and was observed consistently in all plasma samples analyzed in this work. Moreover, the samples from the progressors and non-progressors were completely discriminated by the MS peak intensities for this fragment. Thus, this cleaved form of AGP-2 could be a novel but reliable prognostic factor for the disease.

To evaluate the usefulness as a biomarker, MS peak intensities of the plasma samples were compared between the progressors and non-progressors. As anticipated, the peak intensities with the progressors were significantly greater than those with the non-progressors for all four potential biomarkers (Figure 5-2). The statistical significance of the difference was the least with clusterin ($P = 0.033$), followed by hemopexin and macrophage stimulating protein (MSP), and the greatest with AGP-2 ($P = 0.001$). For hemopexin (Figure 5-2C) and AGP-2 (Figure 5-2D), progressors and non-progressors were completely separated by the MS peak intensities.

Serum levels of potential biomarkers

Among the four proteins identified, the concentrations of clusterin and hemopexin in the serum samples were determined by ELISA. The results of this assay indicated that the concentration of neither protein differed significantly between the progressors and non-progressors (Table 5-II).

Expression of biomarker candidates in OA synovial tissue

We next investigated the expression of the four biomarker candidates within OA synovial tissue. For this, synovial tissues were obtained from 10 end-stage OA knees at the time of prosthetic surgery. First, mRNA expression for these molecules was investigated by *in situ* hybridization, which confirmed abundant expression of all 4 genes coding the candidate proteins in the synovial tissues from OA knees (Figure 5-3A).

For clusterin and hemopexin, their presence in the synovial tissues was examined by Western blotting analysis using specific antibodies for respective proteins. The results of this analysis showed that these proteins were in fact present at substantial levels in the synovial tissues obtained from OA-affected knee joints (Figure 5-3B).

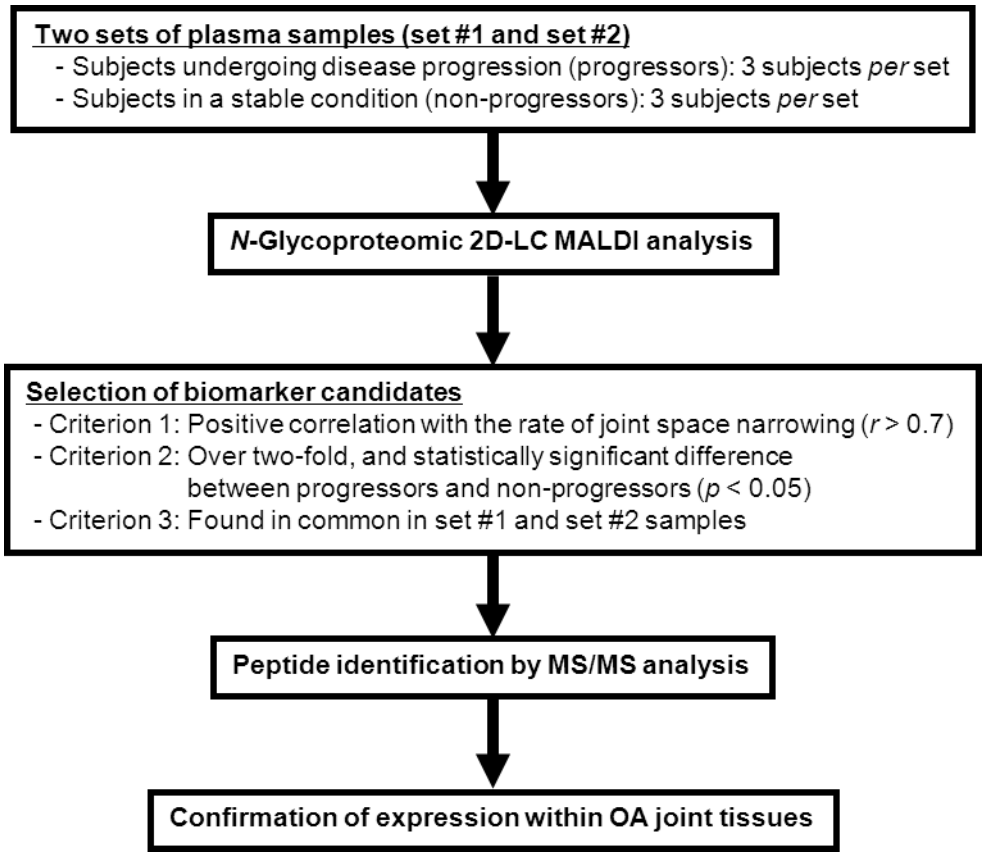


Figure 5-1. Schematic flowchart of analysis strategy.

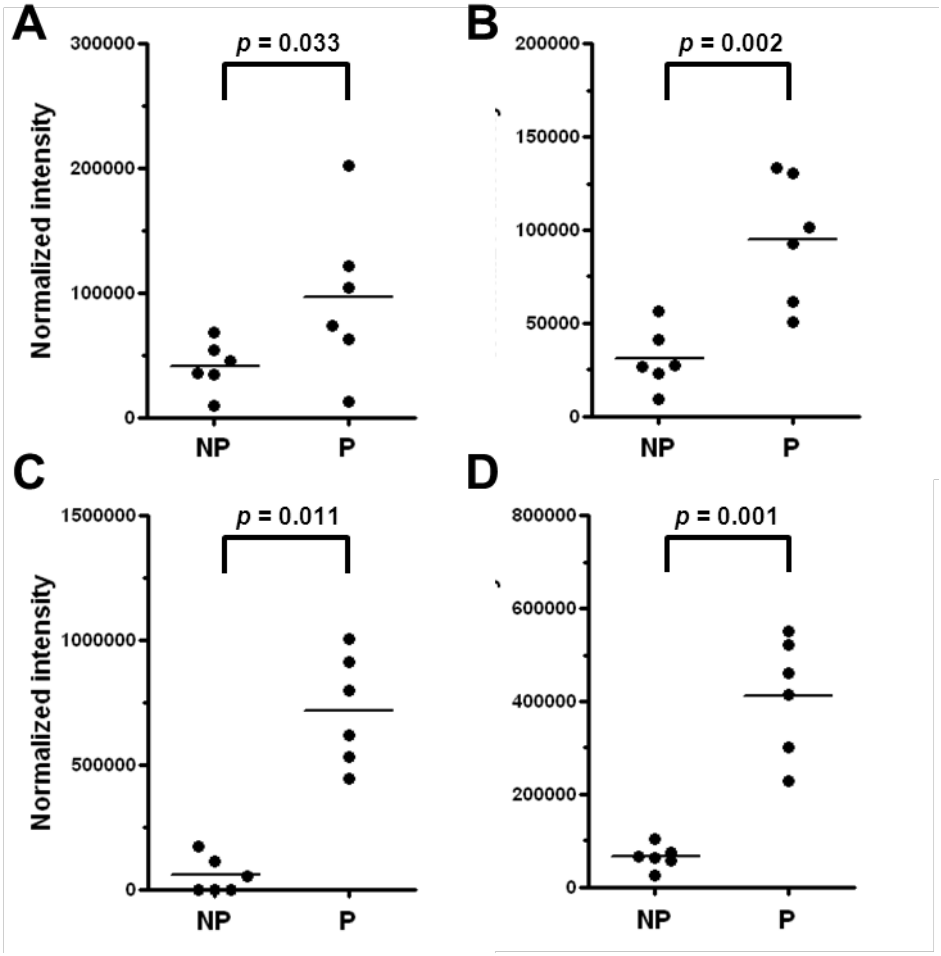


Figure 5-2. Comparison of MS peak intensities of potential biomarkers between progressors and non-progressors. Results are shown for clusterin (A), MSP (B), hemopexin (C), and AGP-2 (D). P and NP represent progressors and non-progressors, respectively. *P*-values were determined by Mann–Whitney U test.

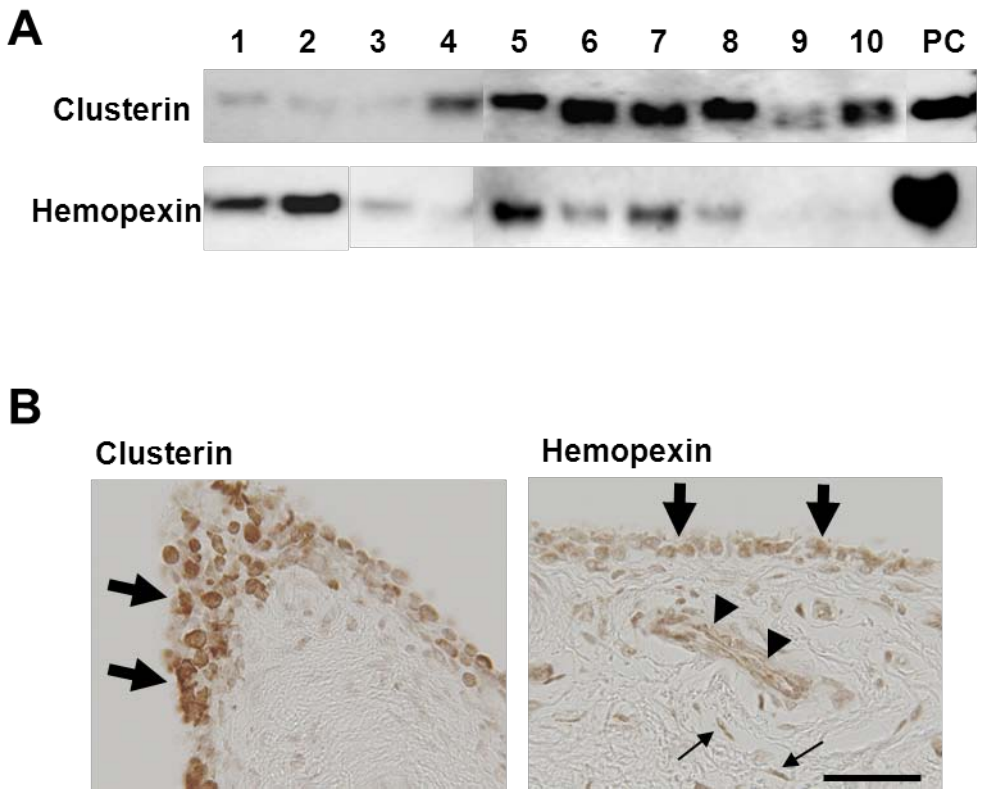


Figure 5-3. Expression of candidate biomarkers in synovial tissues collected from OA knee joints. (A) Results of *in situ* hybridization for clusterin, hemopexin, MSP, and AGP-2 mRNA in OA synovial tissue. Clusterin, hemopexin, and MSP expression were observed primarily in synovial lining cells (thick arrows), while the expression of AGP-2 was mostly observed in endothelial cells (arrowheads). Expression of hemopexin was also observed in endothelial cells (arrowheads) and fibroblast-like cells (thin arrows). Bars, 50 μm. (B) Results of Western blotting analysis of synovial tissue lysates. Clusterin and hemopexin were confirmed in all samples. Numbers above the lanes indicate numbers of synovium samples. PC: positive control.

Table 5-I. Biomarker candidates identified by proteomic analysis

Peak cluster (sample set)	Correlation coefficient (*r*)*	Fold change**	*P* value***	Identified peptide	Protein name	Accession No.	Mascot score	Error, Da (ppm)
2932 (#1)	0.799	6.4	0.012	LVPVPIT**N*AT**LDRITGK	α-1 acid glycoprotein-2	NP_000599	98	0.131
2424 (#2)	0.769	2.9	0.007					(72.4)
4512 (#1)	0.894	2.3	0.199	ML**N*TS**SLLEQLNEQFNW	Clusterin	NP_976084	66	0.052
3819 (#2)	0.812	3.5	0.009	VSR				(21.6)
4878 (#1)	0.916	3.1	0.028	GTG**N*DT**VLNVALLNVIS	Macrophage stimulating protein	NP_066278	43	0.221
4126 (#2)	0.720	2.1	0.061	NQECNIK				(85.3)
4483 (#1)	0.731	19.3	0.005	GHGHR**N*GT**GHG**N*ST**H	Hemopexin	NP_000604	96	0.118
3797 (#2)	0.916	10.0	0.007	HGPEYMR				(49.0)

N-Glycosylation motifs (N-X-S/T) are indicated in bold, and glycosylated asparagine residues are shown with asterisks.

*Correlation coefficient between the peak intensity and the rate of JSW narrowing. **Calculated from the mean intensity of progressors (n=3) to that of non-progressors from respective sample set (#1, #2). ***Calculated by Student's *t*-test.

Table 5-II. Concentrations of clusterin and hemopexin in serum samples

Protein	Concentration (μg/mL)		Statistical significance (*P*)*
	Progressors (*n* = 6)	Non-progressors (*n* = 6)	
Clusterin	239 ± 18.5	243 ± 18.4	N.S. (0.353)
Hemopexin	1241 ± 133	1259 ± 36.3	N.S. (0.373)

*Statistical analyses were performed by Mann-Whitney U test. N.S., not significant.

Table 5-III. Demographic and clinical characteristics of patients who provided plasma and serum samples for analysis

Age	Sex	Height (cm)	Body weight (kg)	BMI	K/L grade*	
					Right	Left
Progressors						
71	Female	150	58	25.8	Grade 2	Grade 3
76	Female	152	65	28.1	Grade 2	Grade 2
67	Female	157	57	23.1	Grade 1	Grade 2
Non-progressors						
75	Male	165	70	25.7	Grade 1	Grade 2
71	Female	151	70	30.7	Grade 2	Grade 3
72	Female	145	57	27.1	Grade 2	Grade 2

*Kellgren–Lawrence grade[9] on enrollment.

5-4. Discussion

To date, several proteins are known to associate with the progression of knee OA. Cartilage oligomeric matrix protein (COMP) in serum[2,10] and the C-telopeptide of type II collagen (CTX-II) in urine[11-13] are the examples. Hyaluronic acid and C-reactive protein (CRP) in serum may also be associated with the progression of the disease[14-17]. However, none of these markers are reliable enough to be used clinically. This lack of established biomarkers for OA may be partly ascribed to the difficulty in specifying the phase of disease progression. In most previous studies, blood or urine samples were obtained only once, at the beginning of the study period. Such samples may not be necessarily obtained during the progression of the disease, because the progression may have started after the collection of the initial samples. Considering this possibility, in the present study, we obtained knee radiographs repeatedly during the study period, and specified the samples acquired in the middle of disease progression. Due to this study design, we consider the four molecules identified here would be highly promising as prognostic biomarkers for the disease.

Another advantage of this study is the application of a glycoproteomic approach to the discovery of biomarkers. To our knowledge, this is the first study to use this technology for the exploration of biomarkers for OA. In this approach, glycoproteins in plasma were specifically concentrated and fractionated by 2D-LC, and the trypsin-digested glycoproteins were subjected to comprehensive analysis by 2D-LC-MALDI. In understanding the result, it should be noted that the changes in peak intensities reflect not only the change in the amount of a protein but also the change in the level of glycosylation on that protein. Considering this issue, for hemopexin and clusterin, we determined the amounts of the proteins as well. These measurements revealed that the concentration of neither protein differed significantly between the progressors and the non-progressors. Considering this result together with the differences in the amounts of glycosylated proteins, we currently assume that for these proteins, the level of glycosylation may increase during the progression of the disease. In order to validate this

assumption, we have performed western blotting analysis, but failed to demonstrate the anticipated changes in glycosylation levels between the progressors and non-progressors, possibly due to technical limitations of the method. Recent studies have shown that the changes in the patterns of glycosylation on hemopexin and clusterin may be useful for the diagnosis for hepatocellular carcinoma[18,19]. We consider similar scenarios might be true with these two proteins for knee OA patients.

For the remaining AGP-2 and MSP, we are currently unable to determine the protein concentrations, and it is therefore still not clear whether the observed differences in peak intensities between the progressors and non-progressors were related to the difference in the amounts of proteins and/or in the levels of glycosylation. Further studies are necessary to clarify why the peak intensities for these proteins differed between the two groups of subjects.

Among the four identified biomarkers, clusterin is a heterodimeric glycoprotein which is known to be expressed in a variety of biological events, such as inflammation, tissue injury, cold stress, and oxidative stress[20,21]. In these conditions, clusterin is expressed by various types of cells, and exhibits protective effects on the cells. In OA, both chondrocytes and synoviosytes are known to express clusterin[22-26]. As mentioned above, we assume that the level of glycosylation on clusterin would elevate in those cells in association with the progression of the disease.

Hemopexin is one of the acute-phase response proteins produced primarily by hepatocytes[27]. This protein is abundant in plasma, and plays significant roles in inflammation and tissue repair, likely by suppressing the adhesion of polymorphonuclear leukocytes[28]. During tissue repair, hemopexin suppresses the release of proinflammatory cytokines from macrophages[29]. In analogy to clusterin, we observed that the level of glycosylation on hemopexin, but not the amount of total protein, was increased in patients undergoing disease progression. We assume that the increased glycosylation on this protein would reduce the anti-inflammatory effects of hemopexin, which could facilitate disease progression. This possibility may be worth examined by future studies.

AGP-2, also known as orosomucoid-2, is another acute-phase response protein that exists abundantly in plasma. AGP-2 is secreted by hepatocytes and released into the circulation in highly *N*-glycosylated forms[30]. The concentration of AGP-2 in plasma is known to elevate with tissue injury, inflammation, and certain types of cancer[31]. In interpreting the present findings, it should be noted that the peptide identified by MS/MS was not the entire AGP-2 molecule but a processed fragment lacking an N-terminal region. Although speculative, it seems plausible that the amount of this fragment increased in the progressors by the increase in protein processing, rather than by the rise in the glycosylation level. This suggests the possibility of a specific enzyme being induced in OA knees during the progression of the disease, which could yield this AGP-2 fragment. Such processing may be involved in the pathology of OA through the modulation of the biological activity of AGP-2.

Similar to hemopexin and AGP-2, MSP is synthesized primarily by hepatocytes as an inactive precursor and released into the circulation [42]. In the blood, most MSP is present in this biologically inactive form. To become biological active, the precursor should be processed into MSP-α and MSP-β, and these two proteins should be dimerized by disulfide bonds. This active MSP (α-β heterodimer) is known to have biphasic effects on macrophages. While the protein enhances spreading, migration, and phagocytosis of macrophages, it may also inhibit macrophage production of inflammatory mediators such as TNF-α[33]. Thus, increased levels of MSP in OA patients may account for the unique pathology of OA synovium, which is characterized by macrophage infiltration without overt signs of inflammation. In this study, we found that the peak intensity for MSP-β was increased in association with disease progression. Since we were unable to determine whether this increase was caused by the increase in the amount of MSP-β protein or by the enhancement of glycosylation on this protein, the significance of the change in peak intensity is not yet clear. Further studies are necessary to determine its significance in the pathology of OA.

Thus, in this chapter, we described that the concentrations of four molecules are increased

in plasma of knee OA patients in association with the progression of the disease, suggesting that these molecules have a potential to be prognostic markers for the disease. Hopefully, future studies demonstrate reliability of these factors as biomarkers.

5-5. Experimental section

Study subjects and collection of plasma samples

This study was approved by the institutional review boards of the participating institutions, and informed consent was obtained in writing from each subject. To obtain plasma samples from the subjects with and without progression of knee OA (progressors and non-progressors, respectively), we followed up knee OA patients over an 18-month period, which was conducted as a part of a 3-year follow-up study of knee OA patients[8]. The subjects had been diagnosed to have primary knee OA with radiographic manifestations. Subjects with any serious health problems other than knee OA were excluded from the study.

In the 18-month study period, plasma and serum samples were obtained every 6 months, together with knee radiographs. To prepare the plasma samples, peripheral blood was drawn into EDTA-containing tubes, and the plasma was separated by centrifugation. For serum samples, blood was obtained in collection tubes, and the serum was obtained by centrifugation after 60 min of incubation at room temperature. The plasma and serum samples were stored at –80°C until use.

The progression of knee OA was determined by the changes in JSW between the femur and tibia in the involved compartment of the knee, which was measured on postero-anterior radiographs obtained in a weight-bearing position[34,35]. In this study, progressors and non-progressors were determined using the following conditions. Progressors were patient in whom the JSW decreased continuously for 18 months or three consecutive 6-month periods (four consecutive follow-up visits) in at least one knee. In non-progressors, the JSW did not show any detectable changes for three consecutive 6-month periods in either knee. For each progressor, the annual rate of JSW narrowing was determined for every 6-month period. If a progressor showed disease progression bilaterally, the higher rate was chosen for the subject.

Selection of plasma samples for 2D-LC-MALDI analysis

Among all subjects enrolled in the study, 3 were determined to be progressors, and another 12 were categorized as non-progressors according to our criteria. The rate of JSW narrowing for these 3 progressors was 0.71 – 4.4 mm/year. Among the 12 non-progressors, 3 patients were chosen for the study considering equivalence to the progressors in age, BMI, and radiographic severity of OA[9] (Table 5-III). The samples from these 3 non-progressors were analyzed together with those from the 3 progressors.

During the 18-month follow-up period, these 6 patients (3 progressors and 3 non-progressors) visited the clinic 4 times as requested, and thus 4 plasma samples were obtained from each subject at 6-month intervals. Among these 4 samples, those obtained at the second (set #1) and third visits (set #2) were used for the analysis. These samples were chosen because at those visits, it could be determined exactly whether a subject was undergoing disease progression or was in a stable (non-progressing) condition, because the changes in JSW were known for both the 6-month period before the visit and the next 6-month period after the visit. If JSW reduced successively in these two 6-month periods, the subject was surely in the middle of disease progression at that visit.

2D-LC-MALDI analysis focusing on N-glycoproteins

In this study, all plasma samples obtained from the patients were analyzed independently, and no samples were pooled or mixed for the analysis. The method for the analysis of *N*-glycoproteins in plasma samples was described in the *Chapter 4*.

ELISA

Serum levels of clusterin and hemopexin were determined by human clusterin Quantikine ELISA (R&D Systems, Minneapolis, MN) and human hemopexin ELISA (Kamiya Biomedical Co., Seattle, WA), respectively, in accordance with the manufacturers' instructions. The assays were performed in duplicate.

Acquisition of synovial tissues

To evaluate the expression of possible prognostic factors in the synovium, synovial tissues were obtained from 10 end-stage OA knees at the time of prosthetic surgery. Immediately upon harvest, the synovial tissue was divided equally into two parts each weighing 200 – 400 mg wet weight. One portion was snap-frozen in liquid nitrogen and stored at –80°C for Western blotting analysis. The other was processed for *in situ* hybridization.

Western blotting analysis

Tissue lysates of OA synovium were prepared as follows. A sample of synovial tissue was pulverized while frozen in a tissue homogenizer (SK-200; Tokken, Chiba, Japan), and sample extraction buffer (8 M urea, 4% CHAPS, 65 mM dithioerythritol) containing protease inhibitors (Complete Protease Inhibitor Cocktail; Roche Diagnostics, Basel, Switzerland) was added at a rate of 10 mL per 1 g wet weight of the tissue. After gentle rotation at 4°C overnight, the sample was centrifuged at 15000 × *g* for 15 min, and the tissue lysate was obtained as a supernatant. After determining protein concentration, the lysate was stored at –80°C until use.

The lysate containing 1 μg of protein was electrophoresed in each lane of a 4% – 12% SDS-polyacrylamide gel. The separated proteins were transferred onto a PVDF membrane, which was blocked with 5% skimmed milk in TBS for 1 h at 25°C. The membrane was then incubated with anti-human clusterin monoclonal antibody (1:1000, R&D Systems) or anti-human hemopexin polyclonal antibody (1:2500; Lifespan Biosciences, Seattle, WA) in TBST (0.05% Tween 20 in TBS) containing 5% skimmed milk overnight at 4°C. Subsequently, the membrane was washed 3 times with TBST for 5 min each time and then incubated with HRP-conjugated anti-mouse IgG polyclonal antibody (1:2000; GE Healthcare, Uppsala, Sweden) in TBST containing 0.5% skimmed milk for 1 h at 25°C. After washing 3 times with TBST for 5 min each time, the protein of interest was detected with ECL-plus (GE Healthcare)

on a LAS-3000 image analysis system (Fuji Film, Tokyo, Japan).

In situ hybridization

For *in situ* hybridization, synovial tissues were cut into small pieces (3-mm cubes) and fixed in 4% paraformaldehyde fixative at room temperature overnight. The tissues were then embedded in paraffin.

DIG-labeled cRNA probes for human clusterin and hemopexin were synthesized from PCR-generated cDNA fragment. The clusterin cDNA fragment was amplified using the primer pair, 5'-TTGGAGGCATGATGAAGACT-3' (forward) and 5'-TTCAGGAACTCCTCAAGCTG-3' (reverse). The hemopexin cDNA fragment was amplified using the primer pair, 5'-CCATTGCTCATCAGTGGCCC-3' (forward) and 5'-GTGAGTGCAGCCCAGGAGAC-3' (reverse). The MSP cDNA fragment was amplified using the primer pair, 5'-CGCACAAGCCGCAGTTCACG-3' (forward) and 5'-CCATCTTGGCTACTGGGACC-3' (reverse). The AGP-2 cDNA fragment was amplified using the primer pair, 5'-TTGCTGGGCTCCAAGTGACC-3' (forward) and 5'-CATCAAGGTCTTGGTGTCCC-3' (reverse). Each fragment was purified and subcloned into pCRII-TOPO vector (Invitrogen, Carlsbad, CA). The subcloned samples were linearized with the restriction enzyme *Hin*dIII as the templates of probes, and the labeled cRNA probes were synthesized with T7 RNA polymerase in the presence of digoxigenin (DIG)-labeled UTP (Roche Diagnostics) for antisense probe, or they were linearized with the restriction enzyme *Eco*RV, and the labeled cRNA probes were synthesized with Sp6 RNA polymerase for the sense probe. DIG incorporation of each probe was assessed by dot blotting. Paraffin sections (10 μm thick) were dewaxed in xylene and rehydrated through a graded alcohol series. The sections were treated with 3 μg/mL proteinase K (Roche Diagnostics) for 10 min at 37°C and acetylated with 0.25% acetic anhydride in 0.1 mol/L triethanolamine (pH 8.0) for 10 min. The sections were prehybridized at 52°C for 2 h with hybridization buffer containing 50% deionized

formamide, 1× Denhardt's solution, 15% dextran sulfate, 600 mM NaCl, 0.1% SDS, 5 mM EDTA, 0.25 mg/mL yeast tRNA, and 10 mM Tris-HCl (pH 8.0). Next, the slides were incubated in hybridization buffer containing 500 ng/mL DIG-labeled antisense or sense cRNA probe for 16 h at 55°C. After hybridization, the sections were rinsed with 5× sodium chloride-sodium citrate buffer (SSC) for 5 min at 50°C. The sections were washed twice with 2× SSC for 30 min at 50°C and thoroughly washed twice with 0.2× SSC for 30 min at 50°C. The sections were immersed in 5% normal rabbit serum for 30 min. The sections were then incubated with 0.4 μg/mL sheep anti-DIG antibody (Roche Diagnostics) for 60 min. The sections were incubated with biotinylated rabbit anti-sheep IgG (Vector Laboratories, Burlingame, CA), followed by treatment of avidin-biotin complex solution (Vector Laboratories). 3,3-Diaminobenzidine (DAB) was used as a chromogen. The sections were dehydrated with alcohol, cleared with xylene, coverslipped, and mounted with Entellan-new (Merck, Darmstadt, Germany). Photomicrographs of bright field images were obtained with an Olympus BH-2® microscope (Olympus, Tokyo, Japan).

5-6. References

1. Spector TD, Dacre JE, Harris PA, Huskisson EC (1992) Radiological progression of osteoarthritis: an 11 year follow up study of the knee. *Ann Rheum Dis* **51**:1107-1110.
2. Sharif M, Kirwan JR, Elson CJ, Granell R, Clarke S (2004) Suggestion of nonlinear or phasic progression of knee osteoarthritis based on measurements of serum cartilage oligomeric matrix protein levels over five years. *Arthritis Rheum* **50**:2479-2488.
3. Dieppe PA, Lohmander LS (2005) Pathogenesis and management of pain in osteoarthritis. *Lancet* **365**:965-973.
4. Elsaid KA, Chichester CO: Review (2006) Collagen markers in early arthritic diseases. *Clin Chim Acta* **365**:68-77.
5. Ruiz-Romero C, Blanco FJ (2010) Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage* **18**:500-509.
6. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* **1**:845-867.
7. Pratta MA, Tortorella MD, Arner EC (2000) Age-related changes in aggrecan glycosylation affect cleavage by aggrecanase. *J Biol Chem* **275**:39096-39102.
8. Fukui N, Yamane S, Ishida S, Tanaka K, Masuda R, Tanaka N, Katsuragawa Y, Fukui S (2010) Relationship between radiographic changes and symptoms or physical examination findings in subjects with symptomatic medial knee osteoarthritis: a three-year prospective study. *BMC Musculoskelet Disord* **11**:269.
9. Kellgren JH, Lawrence JS (1957) Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* **16:**494-502
10. Vilim V, Olejarova M, Machacek S, Gatterova J, Kraus VB, Pavelka K (2002) Serum levels of cartilage oligomeric matrix protein (COMP) correlate with radiographic progression of knee osteoarthritis. *Osteoarthritis Cartilage* **10**:707-713.
11. Bruyere O, Collette J, Kothari M, Zaim S, White D, Genant H, Peterfy C, Burlet N, Ethgen

D, Montague T, et al (2006) Osteoarthritis, magnetic resonance imaging, and biochemical markers: a one year prospective study. *Ann Rheum Dis* **65**:1050-1054.

12. Dam EB, Byrjalsen I, Karsdal MA, Qvist P, Christiansen C (2009) Increased urinary excretion of C-telopeptides of type II collagen (CTX-II) predicts cartilage loss over 21 months by MRI. *Osteoarthritis Cartilage* **17**:384-389.
13. Garnero P, Sornay-Rendu E, Arlot M, Christiansen C, Delmas PD (2004) Association between spine disc degeneration and type II collagen degradation in postmenopausal women: the OFELY study. *Arthritis Rheum* **50**:3137-3144.
14. Bruyere O, Collette JH, Ethgen O, Rovati LC, Giacovelli G, Henrotin YE, Seidel L, Reginster JY (2003) Biochemical markers of bone and cartilage remodeling in prediction of longterm progression of knee osteoarthritis. *J Rheumatol* **30**:1043-1050.
15. Pavelka K, Forejtova S, Olejarova M, Gatterova J, Senolt L, Spacek P, Braun M, Hulejova M, Stovickova J, Pavelkova A (2004) Hyaluronic acid levels may have predictive value for the progression of knee osteoarthritis. *Osteoarthritis Cartilage* **12**:277-283.
16. Spector TD, Hart DJ, Nandra D, Doyle DV, Mackillop N, Gallimore JR, Pepys MB (1997) Low-level increases in serum C-reactive protein are present in early osteoarthritis of the knee and predict progressive disease. *Arthritis Rheum* **40**:723-727.
17. Sharif M, Shepstone L, Elson CJ, Dieppe PA, Kirwan JR (2000) Increased serum C reactive protein may reflect events that precede radiographic progression in osteoarthritis of the knee. *Ann Rheum Dis* **59**:71-74.
18. Debruyne EN, Vanderschaeghe D, Van Vlierberghe H, Vanhecke A, Callewaert N, Delanghe JR (2010) Diagnostic value of the hemopexin N-glycan profile in hepatocellular carcinoma patients. *Clin Chem* **56**:823-831.
19. Comunale MA, Wang M, Rodemich-Betesh L, Hafner J, Lamontagne A, Klein A, Marrero J, Di Bisceglie AM, Gish R, Block T, Mehta A (2011) Novel changes in glycosylation of serum Apo-J in patients with hepatocellular carcinoma. *Cancer Epidemiol Biomarkers Prev*

20:1222-1229.

20. Viard I, Wehrli P, Jornot L, Bullani R, Vechietti JL, Schifferli JA, Tschopp J, French LE (1999) Clusterin gene expression mediates resistance to apoptotic cell death induced by heat shock and oxidative stress. *J Invest Dermatol* **112**:290-296.
21. Jones SE, Jomary C (2002) Clusterin. *Int J Biochem Cell Biol* **34**:427-431.
22. Connor JR, Kumar S, Sathe G, Mooney J, O'Brien SP, Mui P, Murdock PR, Gowen M, Lark MW (2001) Clusterin expression in adult human normal and osteoarthritic articular cartilage. *Osteoarthritis Cartilage* **9**:727-737.
23. Kumar S, Connor JR, Dodds RA, Halsey W, Van Horn M, Mao J, Sathe G, Mui P, Agarwal P, Badger AM, et al (2001) Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries. *Osteoarthritis Cartilage* **9**:641-653.
24. Aronow BJ, Lund SD, Brown TL, Harmony JA, Witte DP (1993) Apolipoprotein J expression at fluid-tissue interfaces: potential role in barrier cytoprotection. *Proc Natl Acad Sci U S A* **90**:725-729.
25. Devauchelle V, Marion S, Cagnard N, Mistou S, Falgarone G, Breban M, Letourneur F, Pitaval A, Alibert O, Lucchesi C, et al (2004) DNA microarray allows molecular profiling of rheumatoid arthritis and identification of pathophysiological targets. *Genes Immun* **5**:597-608.
26. Devauchelle V, Essabbani A, De Pinieux G, Germain S, Tourneur L, Mistou S, Margottin-Goguet F, Anract P, Migaud H, Le Nen D, et al (2006) Characterization and functional consequences of underexpression of clusterin in rheumatoid arthritis. *J Immunol* **177**:6471-6479.
27. Tolosano E, Altruda F (2002) Hemopexin: structure, function, and regulation. *DNA Cell Biol* **21**:297-306.
28. Suzuki K, Kobayashi N, Doi T, Hijikata T, Machida I, Namiki H (2003) Inhibition of

Mg2+-dependent adhesion of polymorphonuclear leukocytes by serum hemopexin: differences in divalent-cation dependency of cell adhesion in the presence and absence of serum. *Cell Struct Funct* **28**:243-253.

29. Liang X, Lin T, Sun G, Beasley-Topliffe L, Cavaillon JM, Warren HS (2009) Hemopexin down-regulates LPS-induced proinflammatory cytokines from macrophages. *J Leukoc Biol* **86**:229-235.
30. Fournier T, Medjoubi NN, Porquet D (2000) Alpha-1-acid glycoprotein. *Biochim Biophys Acta* **1482**:157-171.
31. Budai L, Ozohanics O, Ludanyi K, Drahos L, Kremmer T, Krenyacz J, Vekey K (2009) Investigation of genetic variants of alpha-1 acid glycoprotein by ultra-performance liquid chromatography-mass spectrometry. *Anal Bioanal Chem* **393**:991-998.
32. Bezerra JA, Witte DP, Aronow BJ, Degen SJ (1993) Hepatocyte-specific expression of the mouse hepatocyte growth factor-like protein. *Hepatology* **18**:394-399.
33. Nikolaidis NM, Gray JK, Gurusamy D, Fox W, Stuart WD, Huber N, Waltz SE (2010) Ron receptor tyrosine kinase negatively regulates TNFalpha production in alveolar macrophages by inhibiting NF-kappaB activity and Adam17 production. *Shock* **33**:197-204.
34. Peterfy C, Li J, Zaim S, Duryea J, Lynch J, Miaux Y, Yu W, Genant HK (2003) Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test-retest reproducibility. *Skeletal Radiol* **32**:128-132.
35. Kothari M, Guermazi A, von Ingersleben G, Miaux Y, Sieffert M, Block JE, Stevens R, Peterfy CG (2004) Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis. *Eur Radiol* **14**:1568-1573.

Chapter 6
Concluding remarks

Osteoarthritis (OA) is the most prevalent form of arthritis, and is characterized by a gradual loss of cartilage matrix that often extends over years and decades. Age is the most powerful risk factor for OA. Thus, with increasing longevity, OA has now become a leading cause of disability for older adults in developed countries. Currently available pharmacological therapies are mainly used for pain relief. Other palliative drugs affecting various targets are currently in development. However, a curative treatment to induce regeneration and repair of cartilage is still lacking. Moreover, no effective treatments have yet been established for OA that can inhibit or even delay disease progression. These may be partly due to the lack of reliable biomarkers for regeneration of cartilage and prediction of disease progression.

In *Chapter 2 and 3,* we discovered novel chondrogenic differentiation markers through high throughput glycoproteomics uncovering proteins displaying characteristic glycan structures, notably a combination of glycoblotting-based quantitative cellular *N*-glycomics and glycoforms-focused reverse proteomics/genomics[1]. The 5 chondrogenic differentiation markers were ectonucleotide pyrophosphatase/phosphodiesterase family member 1, collagen alpha-1(III) chain, collagen alpha-1(XI) chain, aquaporin-1, and netrin receptor UNC5B. All candidates were cell surface glycoproteins and showed increased expression levels at the earlier stages of differentiation compared with Type II collagen, a well-known chondrogenic differentiation marker. These markers can be detected specifically with much higher sensitivity than Type II collagen in differentiation by immunochemistry or fluorescence-activated cell sorting (FACS). Moreover, these molecules may have key functional roles in chondrogenesis and are also expected to be potential as pharmaceutical targets in OA, though further extensive investigation is needed to assess the practical availability of differentiation markers and pharmaceutical targets.

In *Chapter 4 and 5*, we discovered novel potential plasma biomarkers of knee OA by quantitative plasma *N*-glycoproteomics using label-free 2-D LC-MALDI MS[2,3]. To discover new plasma biomarkers, we developed a new methodology for large-scale *N*-glycoproteomics analysis combined with a quantitative label-free 2-D LC-MALDI MS system[2]. The methodology had high sensitivity and quantification. Compared with progressors and non-progressors in OA patients by using the method, we found the 4 potential plasma biomarkers, clusterin, hemopexin, alpha-1 acid glycoprotein-2 and macrophage stimulating protein[3]. The concentrations of four molecules are increased in plasma of knee OA patients in association with the progression of the disease. We suggested that these molecules have a potential to be prognostic markers for the disease.

In many previous reports, proteomic approaches in cartilage biology and OA research have uncovered a lot of cartilage- and OA-related protein markers. The increase or decrease in the levels of the identified proteins could have an important role in to the pathology of OA but most of those proteins were already known to be present in the extracellular matrix of cartilage. In this thesis, we discovered novel biomarkers in osteoarthritis using glycoproteomics. The significance of protein glycosylations is one of the most important research areas in cartilage biology and related clinical/pharmaceutical fields. We also demonstrated the importance of combined glycoform-focused reverse proteomics and genomics to discover new biomarkers, notably the integration of “omics” approach. I strongly hope that these novel biomarkers contribute to drug discovery of osteoarthritis in the near future.

1. **Ishihara T**, Kakiya K, Takahashi K, Miwa H, Rokushima M, Yoshinaga T, Tanaka Y, Togame H, Takemoto H, Amano M, Iwasaki N, Minami A, and Nishimura S. (2014) Discovery of novel differentiation markers in the early stage of chondrogenesis by glycoform-focused reverse proteomics and genomics. *BBA-General Subjects* **1840**, 645-655.
2. **Ishihara T**, Fukuda I, Morita A, Takinami Y, Okamoto H, Nishimura S, and Numata Y. (2011) Development of quantitative plasma N-glycoproteomics using label-free 2-D LC-MALDI MS and its applicability for biomarker discovery in hepatocellular carcinoma. *J. Proteomics* **74**, 2159-2168.
3. Fukuda I, **Ishihara T**, Ohmachi S, Sakikawa I, Morita A, Ikeda M, Yamane S, Toyosaki-Maeda T, Takinami Y, Okamoto H, Numata Y, and Fukui N. (2012) Potential plasma biomarkers for progression of knee osteoarthritis using glycoproteomic analysis coupled with a 2D-LC-MALDI system. *Proteome Science* **10**, 36