Role of HBV replication in host cell metabolism : a proteomics and metabolomics analysis

Laleh Sadrolodabaee

2014

Laleh Sadrolodabaee. (2014). Role of HBV replication in host cell metabolism : a proteomics and metabolomics analysis. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/61086>

<https://doi.org/10.32657/10356/61086>

**ROLE OF HBV REPLICATION IN HOST
CELL METABOLISM:
A PROTEOMICS AND METABOLOMICS
ANALYSIS**

LALEH SADROLODABAE

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

2014

Acknowledgements

First and foremost, I wish to express my sincere gratitude to my supervisor Professor Chen Wei Ning William, who guided this project and helped me in the best way whenever I was in need. I am deeply grateful of his help in the completion of this thesis.

I am also indebted to Agency for Science, Technology and Research (A*STAR) for providing the scholarship as well as School of Chemical and Biomedical Engineering / Nanyang Technological University (NTU) for providing the opportunity for my graduate studies and also financial support.

I would also like to extend my gratitude to my lab colleagues and alumni:

Dr. Feng Huixing, Dr. Zhang Jianhua, Dr. Bai Jing, Dr. Wang Mingxuan Dr. Niu Dandan, Ms. Tang Xiaoling, Ms. Shi Jiahua, Ms. Tan Yi Ling Jane, Ms. Chen Liwei, Ms. Li Xiang, Ms. Lee Jie Lin Jaslyn, Mr. Low Tiong Keng William and Mr. Tan Kee Yang for their valuable discussions and cooperation with my experiments.

Warmest thanks to my dear friends Monireh Soroush and Amir Morshedi for their kind help and constructive suggestions.

Words are unable to express my appreciation to my beloved family especially my parents for their love and kind support. I could not have made it this far in my life without them.

Last but not least, I would like to give my special thanks to my husband, Arash, whose patient love, understanding and encouragements enabled me to complete this thesis.

Table of Contents

Acknowledgements	I
List of Figures.....	VI
List of Tables	VIII
Abbreviations	IX
Publications	XII
Abstract.....	XIII
Chapter 1 : Introduction	1
1.1 Hepatocellular Carcinoma.....	1
1.2 Hepatitis B Virus and Epidemiology	3
1.3 Structure and Genomic Organization of HBV	5
1.4 HBV Replication.....	8
1.5 Hepatitis B Virus Genotypes.....	10
1.6 Hepatitis B Virus X Protein (HBx)	12
1.6.1 HBx and HBV Replication.....	13
1.6.2 Proline-Serine Rich Region in HBx.....	14
1.6.3 HBx and Cellular Signaling Pathways.....	15
1.6.4 HBx and Apoptosis	18
1.6.5 HBx and HCC	18
1.7 HCC and Liver Metabolism	19
1.7.1 Glycolysis and Krebs Cycle	20
1.7.2 Warburg Effect.....	22
1.8 Introduction to "Omics"	24
1.9 Proteomics.....	26
1.9.1 Liquid Chromatography-Mass Spectrometry (LC-MS).....	27
1.9.2 Tandem Mass Spectrometry.....	32
1.9.3 iTRAQ Coupled 2-D LC-MS/MS.....	35
1.9.4 Applying Proteomics to Biomarker Discovery	37
1.10 Metabolomics.....	39
1.10.1 MS-Based Metabolomics	40
1.10.2 Gas Chromatography-Mass Spectrometry (GC-MS).....	42
1.10.3 Applying Metabolomics to Biomarker Discovery	45

1.11	Objectives of Research Project	47
Chapter 2	: Materials and Methods	50
2.1	Cloning.....	50
2.1.1	Mammalian Expression Vector Used.....	50
2.1.2	Cloning of HBx from Different Genotypes.....	51
2.1.3	Gel Extraction and Purification of DNA Fragments	53
2.1.4	Restriction Enzyme Digestion and Ligation	54
2.1.5	Preparation of Competent Cells by CaCl ₂ Method.....	55
2.1.6	Plasmid Transformation	55
2.1.7	Glycerol Stock.....	56
2.1.8	Plasmid Preparation by QIAGEN Mini Kit	56
2.1.9	Plasmid Preparation by QIAGEN Midi Kit	57
2.1.10	Plasmid DNA Quantification by UV Spectrophotometer	58
2.2	Mammalian Cell Culture and Transfection.....	59
2.2.1	Mammalian Cell Culture and Cell Passage.....	59
2.2.2	Freezing, Storage and Recovery of Cells	60
2.2.3	Cell Counting	61
2.2.4	Cell Transfection.....	62
2.3	Proteomics.....	63
2.3.1	Total Protein Extraction, Protein Digestion and Labeling with iTRAQ Reagents 63	
2.3.2	Online 2D Nano LC-MS/MS Analysis	69
2.3.3	Protein Identification and Data Analysis	73
2.4	Western Blot Analysis	74
2.5	Cellular Reactive Oxygen Species Detection.....	78
2.6	Metabolomics.....	80
2.6.1	Cell Lysis, Metabolite Extraction and Derivatization.....	80
2.6.2	GC-MS Analysis	81
2.6.3	Metabolites Detection and Quantification.....	81
2.6.4	Statistical Analysis	82
Chapter 3	: Proteomic Analysis of HBx Transfected Rat Primary Hepatocytes and HepG2 Cells: Role of HBx Protein in Cellular Metabolism	83
3.1	Introduction	83
3.2	Experimental procedures.....	85
3.3	Results and discussion	87

3.3.1	Cloning of HBx and transfection efficiency	87
3.3.2	The expression of HBx in transfected cells.....	89
3.3.3	Differentially expressed proteins identified by iTRAQ-coupled 2D LC- MS/MS 89	
3.3.4	Differentially expressed metabolic proteins.....	98
3.3.5	Validation of differentially expressed proteins by Western blot.....	100
3.3.6	ROS levels are increased in HepG2 cells transfected by HBx.....	103
3.3.7	HBx and reprogramming energy metabolism	105
3.4	Conclusions.....	108
Chapter 4 : Metabolomic Analysis of HBx Transfected HepG2 Cells: Insights on Cellular Metabolism and Biomarker Discovery		109
4.1	Introduction.....	109
4.2	Experimental procedures.....	111
4.3	Results and discussion	112
4.3.1	Cloning of HBx and transfection efficiency	112
4.3.2	Validation of HBx genotypes expression in HepG2 cells.....	113
4.3.3	GC-MS-based metabolomic profiling of HepG2 cells transfected by HBx....	114
4.3.4	Up-regulation of glucose-6-phosphate, lactic acid and succinic acid by HBx	118
4.3.5	Up-regulation of amino acids by HBx	119
4.3.6	Down-regulation of palmitic acid and stearic acid by HBx	122
4.4	Conclusions.....	124
Chapter 5 : Protein Profile of Mutant HBx Transfected Cells: Investigation of SH3 Binding Domain in Virus-Host Interactions		125
5.1	Introduction.....	125
5.2	Experimental of procedures	127
5.2.1	Mutation and Cloning of HBx Genotypes (converting proline to alanine).....	127
5.2.2	PCR Amplification.....	127
5.2.3	Cell Lysis, Protein Digestion, and Labeling with iTRAQ Reagents.....	130
5.3	Results and discussion	131
5.3.1	Cloning of mutant HBx and transfection efficiency	131
5.3.2	The expression of mutant HBx in transfected cells.....	132
5.3.3	Differentially expressed proteins identified by iTRAQ-coupled 2D LC-MS/MS 133	
5.3.4	Comparison of the protein profile of wild type and mutant HBx	135
5.3.5	Validation of differentially expressed proteins by western blot.....	137
5.4	Conclusions.....	140

Chapter 6 : Conclusions and Future Direction	141
References	149
Appendix	161

List of Figures

Figure 1-1 Risk factors for HCC development [10].	2
Figure 1-2 World prevalence of chronic HBV infection and HCC development [14].	4
Figure 1-3 Geographic distribution of chronic HBV infection [16].	5
Figure 1-4 Organization of the HBV genome [21].	7
Figure 1-5 HBV replication cycle [25].	9
Figure 1-6 Prevalence of HBV genotypes [30].	11
Figure 1-7 The schematic shows HBx structural domains [35].	12
Figure 1-8 Multiple biological activates of HBx [25].	17
Figure 1-9 Different steps of glycolysis and overview of the TCA cycle [66].	21
Figure 1-10 Energy production in normal and cancer cells [67].	22
Figure 1-11 Schematic represents that the dissimilarity between oxidative phosphorylation, anaerobic glycolysis and aerobic glycolysis (Warburg effect).	23
Figure 1-12 Schematic diagram of the “Omics” cascade.	26
Figure 1-13 Simplified schematic of two ionization sources: MALDI and ESI.	30
Figure 1-14 Structure of Q-TOF mass analyzer [94].	33
Figure 1-15 On-line LC-MS/MS Analysis Procedure of Complex Peptide Mixture.	34
Figure 1-16 Working Mechanism of iTRAQ Reagents.	37
Figure 1-17 Classes of metabolites and the MS instruments (GC-MS, LC-MS) [121].	42
Figure 1-18 Different part of GC-MS [118].	43
Figure 1-19 A schematic GC-MS-based metabolomics workflow.	45
Figure 2-1 Schematic map of mammalian expression vector pXJ40.	51
Figure 2-2 Reference line obtained for protein quantization.	64
Figure 2-3 Layout for the experiment design (using iTRAQ reagents).	68
Figure 2-4 Position 1 in 2-D Nano LC setup	71
Figure 2-5 Position 2 in 2-D Nano LC setup.	72
Figure 2-6 Layout Setup for the Western Blot Transfer Semi-dry System (Bio-Rad).	76
Figure 2-7 Mechanism of detection the ROS in live cells.	79
Figure 3-1 Cloning of different genotypes of HBx in pXJ40.	87
Figure 3-2 GFP is used for tracing the efficiency of transfection in RPHs and HepG2 cells.	88
Figure 3-3 Western-blot showing the expression of pXJ40-HBx A/B/C in RPHs.	89
Figure 3-4 Differentially expressed proteins identified in RPHs and HepG2 cell.	91
Figure 3-5 The differential expression level of several metabolic enzymes in HBx transfected HepG2 cells (a) and RPHs cells (b).	99
Figure 3-6 Western blot confirmations of the differentially expressed proteins in RPHs (a) and HepG2 cells (b) transfected with HBx A/B/C.	101

Figure 3-7 Fold changes of protein level by western blot were plotted using Image J software in RPHs (a) and HepG2 cells (b) transfected with HBx A/B/C.	102
Figure 3-8 Detecting the ROS levels in HBx transfected HepG2 cells by ROS assay kit.	104
Figure 3-9 HBx can shift the metabolic pattern of cells from oxidative phosphorylation to glycolysis.	107
Figure 4-1 Cloning of different genotypes of HBx in pXJ40 vector and transfection efficiency.	113
Figure 4-2 Western-blot showing the expression of pXJ40-HBx A/B/C/D.	114
Figure 4-3 Overlay of representative GC-MS chromatography of HepG2 cells transfected with HBx.	115
Figure 4-4 Distribution of all metabolites identified by GC-MS.	115
Figure 4-5 Differential expression levels of intracellular metabolites in HBx transfected HepG2 cells.	117
Figure 4-6 HBx induces aerobic glycolysis and glutaminolysis, leading to hepatocellular carcinogenesis.	121
Figure 4-7 Schematic represents role of selected differentially expressed metabolites in hepatocellular carcinogenesis.	123
Figure 5-1 Alignments of HBx amino acids in genotypes A/B/C/D.	126
Figure 5-2 Targeted Mutation in HBx Genotypes.	127
Figure 5-3 Gel electrophoresis of PCR products.	129
Figure 5-4 Cloning of different genotypes of mutant HBx in pXJ40 vector and transfection efficiency.	132
Figure 5-5 Western-blot analysis of mutant HBx genotype A, B and C in transfected cells.	133
Figure 5-6 Difference in protein profile of wild type and P-A mutant HBx.	136
Figure 5-7 Western blot confirmations of the differentially expressed proteins in HepG2 cells transfected with mutant HBx A/B/C.	138
Figure 5-8 Fold changes of protein level by western blot were analyzed using Image J software.	139
Figure 6-1 Proposed protocol workflow of iTRAQ-coupled LC-MS/MS and GC-MS platform for Identification the differentially expressed cellular proteins and metabolites in HBx transfected HepG2 cells.	142
Figure 6-2 Differentially expressed cellular proteins and metabolites indicate that HBx may shift the metabolic pattern of cells form oxidative phosphorylation to glycolysis which leads to hepatocellular carcinogenesis.	145

List of Tables

Table 1-1 The open reading frames of HBV genome and the functions of the proteins they encode [21].....	6
Table 1-2 Geographic distribution of HBV genotypes [31].....	11
Table 2-1 Primers used for different genotypes of HBx	52
Table 2-2 PCR Program	53
Table 2-3 OD and concentration of the prepared plasmids.....	59
Table 2-4 (a) Different concentrations of separating gels. (b) Components of stacking gel for SDS-PAGE.....	74
Table 3-1 Antibodies and their dilution ratios.	86
Table 3-2 List of differentially expressed proteins in HBx transfected HepG2 cells. .	94
Table 3-3 List of differentially expressed proteins in HBx transfected RPHs.....	96
Table 4-1 List of differentially expressed intracellular metabolites in HepG2 cells transfected by HBx.	116
Table 5-1 F1 Primers and theirs characteristics	128
Table 5-2 F2 Primers and theirs characteristics	128
Table 5-3 Primers and theirs characteristics	130
Table 5-4 Differentially expressed proteins in HepG2 cells transfected by mutant HBx.....	134
Table 5-5 Antibodies and their dilution ratios	137

Abbreviations

HCC	Hepatocellular carcinoma
TACE	Transarterial chemoembolization
NASH	Nonalcoholic steatohepatitis
NAFLD	Nonalcoholic fatty liver disease
WHO	World health organization
HBV	Hepatitis B virus
HBsAg	Hepatitis B virus surface antigen
HBcAg	Hepatitis B virus core antigen
HBx	Hepatitis B virus X protein
RC DNA	Relaxed circular DNA
cccDNA	Covalently close circular DNA
pgRNA	Pregenomic RNA
kDa	Kilo dalton
Kb	Kilo basepair
HepG2	Human hepatocellular liver carcinoma cell
NLS	Nuclear localization signal
NES	Nuclear export signal
SH3	Src homology-3
HIV	Human immunodeficiency virus
PSR	Proline-serine rich
TFIIB	Transcription factor IIB
TFIIH	Transcription factor II Human
TBP	TATA-binding protein
RBP5	Retinol-Binding Protein 5
CREB	cAMP response element binding
AFT	Activating transcription factors
MAPK	Mitogen-activated protein kinase
JAK/STAT	Janus Kinase/Signal transducer and activators of transcription
DDB1	DNA damage-binding protein 1
ATP	Adenosine triphosphate

NADH	Nicotinamide adenine dinucleotide
TCA	Tricarboxylic acid
FAD	Flavin adenine dinucleotide
LC	Liquid chromatography
MS	Mass spectrometry
HPLC	High performance liquid chromatography
iTRAQ	Isobaric tag for relative and absolute quantitation
LC-MS/MS	Liquid chromatography -tandem mass spectrometry
2D	Two-dimensional
ESI	Electrospray ionization
MALDI	Matrix assisted laser desorption/ionization
TOF	Time of flight
Q-TOF	Quadrupole-time of flight
FTICR	Fourier transform ion cyclotron resonance
CID	Collision induced dissociation
ICPL	Isotope coded protein labels
ICAT	Isotope-coded affinity tag
SILAC	Stable isotope labeling by amino acids in cell culture
GC-MS	Gas chromatography - mass spectrometry
EI	Electron Impact Ionization
CI	Chemical Ionization
PCR	Polymerase chain reaction
MCS	Multiple cloning sites
SDS-PAGE	Sodium dodecyl sulfate –polyacrylamide gel electrophoresis
RPHs	Rat primary hepatocytes
FBS	Fetal bovine serum
MEM	Minimal essential medium
GFP	Green fluorescent protein
MMTs	Methyl-methanethiosulfonate
HBSS	Hank's Balanced Salt Solution

PBS	Phosphate buffered saline
TBHP	Tert-butyl hydroperoxide
H2DCFDA	Dichlorodihydrofluorescein diacetate
ROS	Reactive oxygen species
DMSO	Dimethyl sulfoxide
SDS	Sodium Dodecyl Sulfate
SCX	Strong cation exchange
MSTFA	N-methyl-N-(trimethylsilyl)trifluoroacetamide
TMCS	Trimethylchlorosilane
AART	Automatic Adjustment of Retention Time
NIST	National Institute of Standards and Technology
SI	Similarity index
PCA	Principle Component Analysis
LDH	Lactate dehydrogenase
HK	Hexokinase
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
PGK	Phosphoglycerate kinase
PK	Pyruvate kinase
SDH	Succinate dehydrogenase
FH	Fumarate hydratase
HIF	Hypoxia-inducible factor
Ac-CoA	Acetyl coenzyme A
MC-CoA	Methacrylyl-coenzyme A
GLS	Glutaminase

Publications

Sadrolodabaee L, Low TK, Feng H, Chen W (2013). Role of HBV replication in host cell metabolism: a proteomics analysis. *Current Proteomics*, 10: 29-37.

Feng H, Zhang J, Tan JYL, **Sadrolodabaee L**, Chen W (2012). Proteomics-related biomarkers for HBV-associated hepatocellular carcinoma: current status and future prospects. *Future Virology*, 7(2): 161-171.

Bai J, **Sadrolodabaee L**, Chen W (2012). Recent Pharmacoproteomics Studies of Warfarin in the Asia-Pacific: A New Strategy for Personalized Medicine? *Current Pharmacogenomics and Personalized Medicinenomics*, 10(4): 322-331.

Saminathan R*, Bai J*, **Sadrolodabaee L***, Karthik GM, Singh O, Subramaniyan K, Ching C, Chowbay B, Chen W (2010). *VKORC1* Pharmacogenetics and Pharmacoproteomics in Patients On Warfarin Anticoagulant Therapy: Transthyretin Precursor as a Potential Biomarker. *PloS ONE*, 5(12):e15064 (*, joint first author).

Bai J, **Sadrolodabaee L**, Ching C, Chowbay B, Chen W (2010). A comparative proteomic analysis of HepG2 cells incubated by S(-) and R(+) enantiomers of anti-coagulating drug warfarin. *Proteomics*, 10(7):1463-1473.

Abstract

Hepatocellular carcinoma (HCC) with more than 700,000 deaths every year is the most prevalent type of liver cancer and a global concern. It has a poor general prognosis and low survival rate, hence improving HCC therapy is highly important. HCC is mainly developed as the result of chronic hepatitis B virus (HBV) infection. It is estimated that about 53% of HCC patients are afflicted with chronic HBV. The smallest protein of HBV, HBx, has vital role in hepatocellular carcinogenesis.

In this study, we aim to investigate the association between HBV replication and the host cell metabolism. Rat primary hepatocytes (RPHs) and human hepatocellular liver carcinoma (HepG2) cells were transfected with different genotypes of HBx. Total proteins and metabolites were extracted and analyzed using LC-MS/MS and GC-MS.

LC-MS/MS based proteomics and GC-MS based metabolomics were applied in order to establish a comprehensive cellular proteomic as well as metabolomic profiling of HBx transfected cells.

The iTRAQ-coupled LC-MS/MS results indicated that a number of glycolytic enzymes including lactate dehydrogenase (LDH), Hexokinase (HK), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK) and pyruvate kinase (PK) were significantly up-regulated while two TCA cycle enzymes, succinate dehydrogenase (SDH) and fumarate hydratase (FH) were down-regulated in RPHs and HepG2 cells transfected by HBx compared to the pXJ40 control group.

These findings suggested that HBV replication could alter host cell metabolism by reducing TCA cycle and increasing the rate of glycolysis to provide important metabolic requirements for nucleotides, amino acids and lipids synthesis.

Our proteomics approach may provide candidate biomarkers which improves insights into the future of HBV-associated HCC diagnosis.

In addition, iTRAQ-coupled LC-MS/MS was applied in order to investigate the role of SH3 binding domain in virus-host interactions. For this purpose, four proline residues at proline rich region of HBx were converted to alanine. Comparing the protein profile of wild type HBx and P-A mutant HBx, may provide valuable information on the role of SH3 binding domain in virus-host interactions.

Our findings demonstrated that a series of differentially expressed cellular cytoskeletal and signaling proteins including vimentin and annexin A2 may interfere with HBx via its proline rich domain. These differentially expressed cellular proteins play a significant role in cytoskeleton organization and cell migration which are involved in hepatocellular carcinogenesis.

As a complement to our proteomics approach, metabolomic analysis was applied to explore the relevance of our proteomics data. GC-MS based metabolomics was carried out to analyze the intracellular metabolites in HepG2 cells transfected by different genotypes of HBx.

In our metabolomic profiling, a total of 71 metabolites were identified out of which 10 showed significant changes compared to the control group. Lactic acid, succinic acid, glucose 6-phosphate, valine, glycine, proline, alanine and glutamate were significantly up-regulated while palmitic acid and stearic acid were down-regulated in HBx

transfected HepG2 cells. These metabolites are involved in several key metabolic pathways such as glycolysis, glutaminolysis, TCA cycle, amino acids and fatty acids metabolism.

Our comprehensive metabolomic profiling of HBx transfected cells may indicate differentially expressed metabolites for developing potential biomarkers which could improve antiviral strategies, thus better managing HBV-related HCC patients.

Taken together, the combined proteomics and metabolomics information on the status of HBx transfected cells provides a comprehensive understanding on the role of HBV replication in host cell metabolism and may improve therapeutic strategies for HBV-associated HCC.

Chapter 1 : Introduction

1.1 Hepatocellular Carcinoma

Hepatocellular carcinoma (HCC) is accounting for 70–85% of all primary liver cancers. HCC is ranked fifth regarding its prevalence worldwide [1, 2]. On the other hand, HCC has the third highest mortality rate among all cancers. Every year, more than 630,000 new cases are reported to be HCC-diagnosed and around 700,000 deaths take place due to HCC [3, 4]. It has a poor general prognosis and low survival rates of 3-5%; hence improving HCC therapy is highly important [5].

Many treatments such as chemotherapy, liver transplantation, surgical resection, transarterial chemoembolization (TACE) and hormonal therapy have been carried out to improve the prognosis of HCC [4]. However, there are still difficulties in both diagnosis and treatment of the disease. Such problems can be solved by finding suitable biomarkers for early detection and effective HCC therapy [6].

Fatality rate of HCC is mainly due to late diagnosis. The symptoms only surface when the tumor is relatively large and causes substantial pain due to the stretching of the liver capsule. Upon realization, HCC could have already entered its late stage which considerably reduces the patient's chance for recovery. Hence, early detection of HCC is crucial for improving the chances of recovery [7].

Hepatocellular carcinoma has a variety of risk factors according to its geographic location. These include chronic viral hepatitis (HBV and HCV), toxins (e.g., alcohol,

tobacco, aflatoxin), hereditary metabolic liver diseases, autoimmune hepatitis, obesity, diabetes mellitus, Nonalcoholic steatohepatitis (NASH) and nonalcoholic fatty liver disease (NAFLD) as shown in figure 1.1 [8].

Any of these mentioned factors can be present in an individual who has a metabolic liver disease and thus influence the final outcome of the disease and possibly elevate the risk for hepatic cancer. Viral hepatitis is involved in up to 80% of all HCC and HBV infection is associated with two thirds of all cases [1]. In spite of the remarkable evidence for control and prevention of HBV infection in childhood by vaccination and antiviral therapies, hepatocellular occurrence is rather increasing in the western world than decreasing [9].

Figure 1-1 Risk factors for HCC development [10].

The risk factors for HCC are hepatitis virus (HBV, HCV), aflatoxin, alcohol, NASH, autoimmune and inherited factors.

1.2 Hepatitis B Virus and Epidemiology

Hepatitis B virus (HBV) is still a global health concern though being discovered for more than 30 years. Based on the statistic data released from world health organization (WHO), it has been estimated that around two billion people have been contaminated with HBV and there are more than 400 million chronic carriers of this virus in the whole world, which leads to 620,000 annual deaths [11].

HBV liver infections usually lead to liver cirrhosis, liver failure and primary liver cancer. Chronic HBV infection highly contributes to the increased risk of HCC. It has been reported that about 53% of HCC patients are afflicted with chronic HBV [12]. The individuals infected with HBV are 100 times more prone to develop HCC comparing to non-infected ones. As shown in figure 1.2, number of HCC cases is highly correlated with the rates of HBV infection in a specific area. Therefore, HBV infection has correlation with HCC incidence [13].

Figure 1-2 World prevalence of chronic HBV infection and HCC development [14].

As the schematic shows, HBV infection is correlated with HCC incidence. (Permission from ref.14 was obtained from publisher to use this figure)

The incidence of HBV infection alters geographically as shown in figure 1.3. Hepatitis B virus surface antigen (HBsAg) is used as a marker for chronic HBV infection [15].

HBsAg prevalence more than 8% is considered as high, 2-7% is intermediate and less than 2% is low. In some areas such as Far East, Africa, the Middle East and some parts of South America the prevalence of HBV is high (8% to 15%), while in Northern Europe, Canada, Australia and some parts of South Africa, the prevalence is less than 2% [16].

Figure 1-3 Geographic distribution of chronic HBV infection [16].

The incidence of HBV infection alters geographically. HBsAg is used as a marker for chronic HBV infection. (Permission from ref.16 was obtained from publisher to use this figure)

1.3 Structure and Genomic Organization of HBV

HBV is a member of the family of hepadnaviridae, with circular genome composed of partially double stranded DNA. The DNA genome of HBV is 3.2kb, surrounded by outer lipid envelope and inner nucleocapsid [17].

There are four genes encoded by the viral genome: C, P, S and X. Gene C encodes core protein, known as hepatitis B virus core antigen (HBcAg). The P region encodes polypeptides responsible for DNA polymerase activity. It has also an important role in reverse transcription activity. The core and polymerase are essential for the viral DNA

replication. Gene S is the gene that codes for HBsAg. The envelope proteins come from the surface gene (figure 1.4) [18].

The X gene encodes hepatitis B x protein (HBx) with 154 amino acids. HBx is multifunctional non-structural regulatory protein. Antigen x of HBV (HBxAg) induces viral gene expression and replication so it plays significant role in the development of HCC [19, 20]. A summary of the various proteins expressed by the HBV genome is listed in table 1.1.

Table 1-1 The open reading frames of HBV genome and the functions of the proteins they encode [21].

Open Reading Frame	Proteins and Function
C	C (core) proteins, together with the surface proteins, form a double layer shell which encapsulates the genome
P	Encodes for viral polymerase and the synthesis of both strands of DNA through a reverse transcriptase enzyme (RT)
S	Surface antigens. Encodes for S protein (HBsAg), M protein and L protein
X	Protein X (HBx) is a nonstructural, multifunctional protein

Three different particles exist in the blood circulation of patients carrying HBV. The Dane particle, the most important of these particles, has a size of 42-47 nm. Smaller filamentous and spherous particles make up the other two. All these particles harbor the surface antigen (HBsAg). On the other hand the envelope proteins encoded by

ORF are namely the Large (L), the middle (M), and the small (S) proteins. The most significant surface protein is S known as the surface antigen of Hepatitis B. The two other surface proteins are; L which is involved in cell surface adhesion and M the function of which is not yet clear [18].

Figure 1-4 Organization of the HBV genome [21].

There are four viral genes including C, S, P and X in HBV genome. Core protein encoded by gene C, gene P encodes for viral polymerase, envelope proteins by S gene and X gene encodes HBx which is the smallest protein of HBV and plays an important role in virus and host cell interaction. (Permission from ref. 21 was obtained from publisher to use this figure)

1.4 HBV Replication

HBV genome contains partially double-stranded, relaxed circular DNA (RC DNA) with the length of 3.2kb. For viral replication, RC DNA should be converted to covalently close circular DNA (cccDNA) [22].

cccDNA is used as template for viral RNA transcription. HBV surface receptors are employed to enter the cell via endocytosis. Upon infection the RC DNA is transferred into the host cell nucleus for replication, where it is converted to cccDNA [20]. Subsequently the viral mRNA is transcribed from the CCC DNA by RNA polymerase II. There are four open reading frames, producing four series of viral mRNAs: 0.7kb, 2.1kb, 2.4kb, and 3.5kb. The longest RNA (3.5kb) is a pregenomic RNA (pgRNA). pgRNA is used as mRNA for synthesis of core protein and reverse transcriptase [23].

The shortest RNA (0.7kb) encodes the X protein which plays a critical role in the virus and host cell interaction. These RNAs are transferred to the cytoplasm for expression into viral proteins. Viral genome replication occurs by reverse transcription of the pgRNA. Completed viral particles either exit the cell via exocytosis or return to the nucleus to produce more copies (figure 1.5) [24].

Figure 1-5 HBV replication cycle [25].

Prior to infection HBV genome is RC DNA. Upon infection RC DNA is converted to cccDNA which acts as template for pregenomic RNA intermediate. pgRNA is used as mRNA for synthesis of core protein and polymerase, while X protein is translated from smaller RNA species. (Permission from ref.25 was obtained from publisher to use this figure)

1.5 Hepatitis B Virus Genotypes

HBV has eight different genotypes categorized from A to H on the basis of their nucleotide sequence. Distinct genotypes have greater than 8% nucleotide sequence difference between them [26]. The prevalence of HBV genotypes are shown in figure 1.6. Among all HBV genotypes, A, B and C are the most coincident genotypes. Genotypes B and C more prevalent in Asia whereas genotypes A and D more dominant in Europe, Africa and Middle East. Genotype A has been shown to have major influence in the prevalence of HCC in sub-Saharan Africa [27].

Study shows that HBs-Ag positive Africans infected with genotype A have 4.5 fold increased risk of HCC in comparison with other HBV genotypes. Several studies have also associated an increased risk of HCC with genotypes B and C. Some studies showed that, increased risk of HCC is associated with genotype C in younger individuals (<50 years old) whereas genotype B has major influence in older individuals (>60 years old). In addition, one of these studies shows that the overall risk of HCC is not significantly different between HBV genotypes B and C while genotype C increases the risk of HCC at a younger age [28].

Each HBV genotypes has its related clinical circumstances and prognosis. In individuals with chronic HBV infections, genotype C is the main cause of cirrhosis while genotype B is mostly associated with HCC [29]. Genotypes D and F seem to cause a more severe form of the disease with a less chance of recovery than genotype A. The geographic distribution of HBV genotypes are shown in table 1.2 [30] .

Figure 1-6 Prevalence of HBV genotypes [30].

Figure shows that genotypes A, B and C are the most dominant genotypes.

Table 1-2 Geographic distribution of HBV genotypes [31].

Genotype	Geographic distribution
A	Africa, India, Northern Europe, United States
B	Asia, United States
C	Asia, United States
D	India, Middle East, Southern Europe, United States
E	West and South Africa
F	Central and South America
G	Europe, United States
H	Central and South America, United States

1.6 Hepatitis B Virus X Protein (HBx)

HBx is the smallest protein of HBV with 154 amino acids (17.5 kDa) which is encoded by the x region. It is a multifunctional protein and has important role in hepatocarcinogenesis [32].

HBx is involved in many cellular processes such as transcription, signaling pathways, cell proliferation, cell cycle, protein degradation and apoptosis. HBx interacts with the cell through a series of protein-protein interactions since it is not able to make direct binding to DNA. It also regulates transcription of virus and cellular genes. HBx has a highly conserved sequence across the different species of mammalian hepadnaviruses [33]. There is no X-ray crystallography and nuclear magnetic resonance from HBx structure because it is insoluble. Therefore, little is known about the 3-Dimensional HBx structure. The function of HBx is complicated and unclear, so it is called X [34].

Figure 1.7 shows the basic structure of the HBx protein.

Figure 1-7 The schematic shows HBx structural domains [35].

HBx has 154 amino acids (17.5 kDa). It has two important domains: regulatory domain and trans-activation domain. They are linked by a proline-serine rich (PSR) domain. BD: basic domain, CPBD: cellular protein binding domain.

Some studies have shown that HBx is mainly located in the cytoplasm; however other studies have revealed that it is mainly located in the nucleus. Some studies indicated a correlation between the localization of HBx and the quantity of its expression. In low expression levels, it is mostly found in the nucleus, however in higher expression amounts, it is mainly located in the cytoplasm [36]. In recent studies, it has been shown that HBx can be detected in mitochondria. Its presence in mitochondria may induce loss of mitochondrial membrane properties, triggering mitochondrial dependant cell death [37].

1.6.1 HBx and HBV Replication

HBV replication is stimulated by HBx at the level of DNA replication; however this is not seen at pgRNA transcription. Previous studies have demonstrated the critical influence of HBx on HBV replication [38]. Furthermore, it has been indicated that the efficiency of wild type HBV replicon is 5-10 times higher than the X-ORF (-) mutant in HepG2 cells [20, 39]. Many transcription factors like p53, Smad4 and leucine zipper family interact with HBx, suggesting that HBx may regulate transport across nucleus and cytoplasm [40].

Recent reports have demonstrated that there is a linkage between nuclear and cytoplasmic HBx with stimulation of viral replication. It was demonstrated that HBx linked either to NLS (nuclear localization signal) or NES (nuclear export signal) is able to induce viral replication at similar levels. In overall these findings suggest that HBx is significantly associated with HBV replication [41].

1.6.2 Proline-Serine Rich Region in HBx

Protein-protein interaction is very important in understanding cellular signal transduction processes. Src homology-3 domains (SH3) and WW domain play a critical role in signal transduction. In eukaryotic cells, SH3 domain is one of the most important domains found in many signaling proteins such as Src, Fyn, Lyn, Hck, Abl, Grb2, kinases, phosphatases, phospholipases and adaptor proteins [42].

SH3 domain is small protein domain which contains 55-70 amino acids, enabling it to specifically interact with proline-rich domain. SH3 domains have been implicated in cancer development. SH3 containing proteins can interact with signalling, cytoskeletal and structural proteins [43].

It was reported that SH3 domains preferred to interact with the proline rich motifs which possess “PXXP” motif. In “PXXP” motif P indicates Proline whereas X indicates any other amino acids [44].

It has been shown that the Nef protein of HIV can interact with host SH3 domains proteins through its proline-rich motifs [45].

HBx has two main domains: regulatory domain in the amino-terminal of the HBx and transactivating domain in the carboxy-terminal. They are linked by a proline-serine rich (PSR) domain. HBx possess a PSR domain in amino acids 30-46, so it can be activated by SH3 domains in cellular proteins as shown in figure 1.8. SH3 domains are critical links between the virus and the host cells [35].

Another proline rich binding motif is WW domain. It has about 40 amino acids which fold as triple stranded β -sheets. They can interact with proline/serine rich ligands. HBx has also been found to interact with pin1, through the WW domain [46].

Our laboratory has recently reported that HBx contains a proline/serine motif which appears to interact with a number of SH3-containing cellular proteins [44].

1.6.3 HBx and Cellular Signaling Pathways

HBx interacts with the cell through a series of protein-protein interactions since it is not able to make a direct binding to the DNA. It regulates transcription of virus and cellular genes. HBx enhances transcription by interacting with nuclear transcriptional factors, including TFIIB, TFIIF, TBP and RBP5. It also directly interacts with CREB/ATF family which have roles in proliferation and liver metabolism [47, 48].

Moreover, HBx can bind to CBP/p300 coactivators and activates various cellular transcription factors. It has been shown that HBx can interact with different signaling pathways such as mitogen-activated protein kinase (MAPK), Janus Kinase/Signal transducer and activators of transcription (JAK/STAT) [49].

HBx can up-regulate the Ras-Raf-MAPK signal cascade and so deregulating the cell cycle. Up-regulation of Ras-Raf-MAPK signal pathway activates NF- κ B transcription factor that increases the gene transcription. HBx also interacts with DDB1 and activates viral replication and transcription. The interactions between HBx and signalling pathways lead to alteration in normal pathways, which causes hepatocellular carcinogenesis [50].

It has been shown that HBx alters the expression of some tumor suppressors and oncogenes such as p53 and c-Myc. HBx plays a role in apoptosis via its binding to tumor suppressor protein p53. It can bind to p53 and block p53-induced apoptosis [51, 52].

HBx also has a role in protein degradation through the interaction with proteasome subunits. To sum up, the HBx protein is a multifunction viral protein and plays a critical role in hepatocellular carcinogenesis caused by HBV infection [53].

Studying the HBx impact on metabolites variation may give us high potential of finding effective biomarkers related with HCC development, which contributes novel clues on future improvement of HBV-associated HCC diagnosis.

In addition, it may also give us opportunities to study the significant activity of HBx and understand its role in hepatocarcinogenesis. Multiple biological activities of HBx are summarized in figure 1.8 [25].

Figure 1-8 Multiple biological activates of HBx [25].

HBx binds various cellular transcription factors including TFIIB, TFIIF, TBP and RBP5. It can regulate the cellular signaling pathways such as MAPKs and JAK/STAT. (Permission from ref.25 was obtained from publisher to use this figure)

1.6.4 HBx and Apoptosis

One of the essential cellular processes for homeostasis in multicellular organisms is apoptosis or programmed cell death. It has been reported that HBx binds to tumor suppressor protein, p53, and inhibits its function. Therefore, the p53 concentration in the nucleus is decreased, which leads to reduction in cell apoptosis and deregulation of cell cycle control [54, 55].

It has been shown that HBx decreases cell apoptosis through several signal pathways. For example, the expression of some proteins such as survivin which is involved in apoptosis inhibition is overexpressed by HBx [56].

Overexpression of this protein by HBx leads to decrease in cell apoptosis. In addition, HBx increases cells proliferation by its effects on the cell cycle. HBx-mediated JAK/STAT pathways activation is associated with enhanced S phase of cell cycle and increased cells proliferation [57, 58]. Thus, down regulation of apoptosis and acceleration of cells proliferation will be the main reasons for the progression of hepatocarcinogenesis.

1.6.5 HBx and HCC

The main cause of HCC worldwide is HBV. HCC is a major liver malignancy and is ranked fifth regarding its prevalence worldwide. HBx plays many diverse functions in different cellular pathways. Its role in HBV associated HCC has been studied extensively due to its importance [33, 59]. HBx has been associated with HCC development with possible linkage to individual genotypes [33].

It has been found that, by deleting the X gene from HBV genome viral infection may be suppressed. Infection with mammalian HBV containing x protein is associated with HCC development whereas infection with duck HBV doesn't lead to HCC due to the deficiency of x protein in duck HBV genome [60]. In addition, during the process of HBV replication, viral DNA integrated into the genome of hepatocytes of chronic HBV carriers and the X gene is mostly conserved. Therefore, HBx is closely related to the HBx-mediated hepatocarcinogenesis [6, 53, 61].

1.7 HCC and Liver Metabolism

Previous studies suggested that metabolic abnormalities in the liver are related with the development of HBV-related HCC [62, 63]. The main organ which carries out most of the metabolic and detoxification processes is the liver. Metabolic abnormalities develop in the liver as a result of exogenous factors or pathological agents like HCC. Irregular amounts of metabolites may show changes in the biochemical pathways in the whole body due to pathological state [64]. As a result, monitoring alterations in biochemical pathways can bring about new perspectives into how diseases function and may reveal diagnostic biomarkers, making it very useful in therapy and prognosis [65].

Cellular respiration is one of the significant metabolic reactions in cells of organisms for converting energy and nutrients into adenosine triphosphate (ATP), releasing debris. Respiration includes different catabolic reactions that involve the redox reaction. Respiration is one of the principal methods which a cell can use to access useful energy as fuel for cellular reformations.

1.7.1 Glycolysis and Krebs Cycle

In the cytosol of all living organisms a metabolic process called glycolysis takes place. In this process pyruvic acid (pyruvate) is formed by breaking up glucose molecules. This in turn yields two ATP and two high energy carrying nicotinamide adenine dinucleotide (NADH) molecules as shown in figure 1.9. Glycolysis can either happen in the presence or absence of oxygen. When oxygen is available, glycolysis is the primary stage of cellular respiration [65].

The overall reaction of glycolysis is the following:

Citric acid cycle -also called 'Krebs' or tricarboxylic acid (TCA) cycle- consists of a sequence of enzyme-mediated chemical reactions and is one of the most significant pathways in all living cells, especially cells that use oxygen as a part of their respiration. The citric acid cycle happens in the matrix of the mitochondrion in eukaryotic cells. This pathway initiates when acetyl-CoA is formed by the three carbon sugars produced previously in glycolysis. After several intermediate steps, different high energy containing compounds along with two ATP molecules are generated. These energy carrying compounds known as nicotinamide adenine dinucleotide (NAD) and flavin adenine dinucleotide (FAD) are reduced along the pathway as shown in figure 1.9. TCA cycle dysfunction may lead to human diseases and tumor formation [66]. Glycolysis and TCA cycle are deregulated in cancer cells. Most cancer cells derive a substantial amount of energy from aerobic glycolysis, converting most incoming glucose to lactate rather than metabolizing it in the mitochondria through

oxidative phosphorylation. Energy production of normal and cancer cells are shown in figure 1.10 [65].

Figure 1-9 Different steps of glycolysis and overview of the TCA cycle [66].

Left: Different steps of glycolysis in cytosol. Right: An overview of the TCA cycle (Krebs cycle) in the matrix of the mitochondrion. As the figure shows pyruvate produced from glycolysis is converted into acetyl CoA, which enters TCA cycle in mitochondria.

Figure 1-10 Energy production in normal and cancer cells [67].

(a) Normal cell energy production. In normal cells pyruvate produced from glycolysis is converted into acetyl CoA, which enters TCA cycle in mitochondria. **(b)** Cancer cell energy production. Most cancer cells shift the energy metabolism from oxidative phosphorylation to glycolysis. (Permission from ref.67 was obtained from publisher to use this figure)

1.7.2 Warburg Effect

According to Warburg effect, in most cancer cells, the rate of glycolysis and lactate generation in the existence of oxygen has been increased [68]. In majority of cancers, pyruvate produced from glycolysis is altered to lactate by lactate dehydrogenase rather than being metabolized in mitochondria through oxidative phosphorylation [67, 69]. Aerobic glycolysis is less efficient in ATP generation than oxidative phosphorylation as shown in figure 1.11.

Figure 1-11 Schematic representation of the dissimilarity between oxidative phosphorylation, anaerobic glycolysis and aerobic glycolysis (Warburg effect).

In the condition of sufficient oxygen, the differentiated tissue oxidized the pyruvate generated by glycolysis to CO₂ in the mitochondria during oxidative phosphorylation. When the level of oxygen is not enough, they converted pyruvate into lactate away from oxidative phosphorylation (anaerobic glycolysis). In proliferative tissue even in the presence of oxygen pyruvate generated by glycolysis is converted into lactate by lactate dehydrogenase rather than being metabolized in mitochondria through oxidative phosphorylation (aerobic glycolysis) [67]. (Permission from ref.67 was obtained from publisher to use this figure)

As noted above, Warburg suggested that cancer cells are increasing the rate of glycolysis and lactate generation in the existence of enough oxygen. This aerobic glycolysis produces only 2 ATPs per glucose molecule, while from oxidative phosphorylation up to 36 ATPs can be generated per glucose molecule [70]. So, why cancer cells select a less efficient metabolic pathway?

The study suggests two possible explanations for switching of cancer cells to aerobic glycolysis [67]:

1) As cancer cells have access to constant source of glucose and other supplements in circulating blood, insufficient amount of ATP is not a concern for these cells.

2) For lipids, amino acids and nucleotides synthesis, cancer cells have crucial metabolic requirements such as acetyl CoA, glycolytic and ribose intermediates that extend beyond ATP.

For instance, palmitic acid synthesis consumes 7 molecules of ATP, 8 molecules of acetyl CoA and 14 molecules of NADPH. In addition to ATP, more equivalents of carbon and NADPH are required for synthesis of amino acids and nucleotides.

1.8 Introduction to "Omics"

Technologies that study the characteristics of molecules in the organisms' cells such as genes, proteins and small metabolites are known by their "omics" suffix e.g. "genomics." In general, Omics technologies identify the roles of various molecules within a cell and investigate their relationships and actions [71]. Omics technologies include:

Genomics: Genome is defined as the information of a complete set of genes within a single cell of an organism. Genomics is the science of discovery, analysing and noting the entire genome of a particular organism. Hence, Genomics, in general, is the study of genetic structure of organisms [72].

Transcriptomics: is the science of cataloguing the complete set of RNA transcripts produced by the genome. The RNA transcripts set include mRNAs, non-coding RNAs and small RNAs [73].

Proteomics: studies the function of all expressed proteins. Large-scale data sets have been generated for interactions between proteins as well as protein activity patterns, protein profiles and organelle composition in cancer patients. Particularly, proteomics aims to find the physiological roles of proteins in an organism by identifying when and where the proteins are expressed in a cell [74].

Metabolomics: Among the ‘omics’ sciences, Metabolomics is one of the newest ones. The complete set of low molecular weight compounds in a sample is called metabolome [75]. Metabolome directly affects the phenotype of the cell. The reason is that these compounds are the substrates and by-products of enzymatic reaction within a cell. Hence, the main objective in metabolomic studies is to determine the profile of these compounds in a sample at a specific time and under specific environmental state [76].

While genomics and proteomics studies provide abundance of information on genotype, they provide little information on phenotype. Metabolites, compounds with low molecular weight, are closely associated with phenotype [77]. Metabolomics, the study of metabolites, fits at the central dogma of biology and sits within the hierarchy of proteomics, transcriptomics and genomics as shown in figure 1.12 [78].

In contrast to the large genome, transcriptome and proteome, the metabolome is small and tightly conserved across the organisms. It is estimated that the human body has more than one million proteins, 25,000 genes and 100,000 transcripts while only has 6,500 (Human Metabolome Database (HMDB)) metabolites [79]. The unique genome of every organism leads to unique transcriptome and proteome. Although the genome

of every species defines a unique collection of metabolites, the fundamental biochemical molecules are highly similar [80].

Figure 1-12 Schematic diagram of the “Omics” cascade.

The figure shows the Omics technologies which include, genomics, transcriptomics, proteomics and metabolomics as describe in page 25. Omics technologies identify the roles of various molecules within a cell and investigate their relationships and actions.

1.9 Proteomics

Proteomics, study of proteome, is involved in both qualitative and quantitative assessment of proteins in large scale. Instead of studying one particular gene or protein in an organism, a comprehensive understanding of genomics, proteomics and metabolomics makes the systematic analysis of all cellular components become possible [81].

Proteomics with lots of applications in practical aspects has become promising methods these last years. Proteomics studies can be considered as useful tools in

medicine for identification of new drugs in order to treat diseases. Finding a suitable biomarker with the help of proteomics technology is another application, which is outstanding in biomedical studies [82]. Biomarkers are biological molecules in the body fluids like plasma, cell lines and tissues. They are specific substances used as indicators of a biological state. Biomarkers have many applications in medicine including: monitoring the treatment, efficacy of disease progression, disease diagnosis and predicting response to treatment [83]. Therefore, in the field of biomarker discovery, mass spectrometry (MS)-based proteomics can be named as a useful tool. Indeed, protein biomarkers can cause a breakthrough in disease diagnosis which might be as a result of proteomics. It is of high importance in detecting a disease and further measuring its progression which in fact means the effect of a treatment [84].

With proteomics guide, we will face new avenues to clarify unknown mechanisms for diseases. It will also help us identify new markers for diagnosis approaches as well as therapeutic aims.

Typically, there are two steps in proteomics studies [85]:

- 1) Separation of proteins done by two-dimensional electrophoresis or liquid chromatography
- 2) Identification of separated proteins, which in turn includes characterizing post-translational modifications, done by mass spectrometry.

1.9.1 Liquid Chromatography-Mass Spectrometry (LC-MS)

As mentioned earlier, protein identification commonly can be fulfilled by mass spectrometry technique. Success in protein identification depends on many factors

such as mass spectrometer accuracy or database availability and modifications which might take place after translation as well as other affecting factors [86]. During the last 20 years, our information and understanding about protein chemistry has been increased a lot and we owe all of these progresses to advances in instrumentation of MS and related techniques. MS has been successful in finding its position in researches and experiment and has been selected to analyze samples containing mixture of proteins. Use of liquid chromatography-mass spectrometry (LC-MS) would be the suitable choice in detecting low molecular weight proteins [87]. In two-dimensional gels, low-molecular weight proteins together with heavy proteins are not detected efficiently [88]. Recently, mass spectrometry (MS) has technologically improved a lot which opened the way for its usage in the field of proteins, peptides, DNAs, carbohydrates as well as many other biological molecules such as drugs [89].

Mass spectrometer plays crucial role in proteomics and has the ability to determine the mass of molecule e.g. protein by measuring m/z in which m is mass of the molecule while z indicates its charge. As we know, inducing molecules to take or lose charge will lead to form an ion, negative or positive respectively. MS technique has the ability to measure m/z ratio in highly accurate way and in fact this measuring provides spectra with high resolution [90]. Developing of mass spectrometry gave rise to tandem mass spectrometry (MS/MS) allowing obtaining of information from de novo protein sequence. In fact this progress (tandem mass spectrometry) caused enhancement of MS application in proteomics studies [91]. The complexes of two mass analyzers make this machine more sensitive and accurate. Tandem MS allows a single charged peptide to be selected, separated, sequenced and detected [92].

Another tremendous progress in mass spectrometry was joining of high performance liquid chromatography (HPLC) to this technique which led to extraordinary and powerful technique (LC-MS) with lots of applications in which both sensitivity and accuracy are achieved [86]. Generally there exist 4 major parts in LC-MS system including:

1) Chromatography column(s): allowing pre-separation of peptide mixtures based on one or more physico-chemical properties (e.g. hydrophobicity, charges) [93].

2) Ionization source: converting eluted peptides into ions in gas phase. There are two main types of ionization sources including electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI) as shown in figure 1.13. [94]. In ESI, a high voltage causes the plume of droplets to form and the liquid in the droplets to evaporate while leaves the peptides charged. A soft ionization technique named 'MALDI' has two step processes. In the first step, analyte molecules are incorporated into a suitable matrix. The second step involves protonation or deprotonation of analyte molecules in the hot plume [95, 96].

3) Mass analyzer(s): separate ions based on m/z ratio. There are mainly four types of mass analyzer used in proteomics research including 1) quadrupole, 2) reflector time-of-flight (TOF), 3) quadrupole ion trap and 4) Fourier transform ion cyclotron resonance (FTICR) analyzers [97]. Among these analyzers, quadrupole-time of flight (Q-TOF) is the mass analyzer that was applied in our two-dimensional 2-D LC-MS/MS [94].

4) Detector: enabling detection of the relative abundance of ions at discrete m/z . Detector has the ability to provide electrical signals from ion energy. These signals then can be transferred to a computer [98].

Protein identification by LC-MS was performed by bioinformatic software (e.g. ProteinPilotTM from Applied Biosystems) and databases, including SWISS-PROT, TrEMBL PIR, NCBI Inr, ESTdb.

Figure 1-13 Simplified schematic of two ionization sources: MALDI and ESI.

(Detailed description in page 29) [94]. (Permission from ref.94 was obtained from publisher to use this figure)

In liquid chromatography, peptides are separated based on their physiochemical characterization. Liquid chromatography has four columns and two pumps. The most commonly used columns for separation of mixed peptides is reverse-phase liquid chromatography. It separates and elutes the peptides according to their hydrophobicity. A second dimensional chromatography usually is strong cation-exchange (SCX) liquid chromatography, which separates the peptides by the nature of charges. Moreover, it contains 2 identical columns called enrichment columns by which trapping of peptides is done (ZORBAX 300SBC18 enrichment column 1 and ZORBAX 300SB-C18 enrichment column 2). High performance liquid chromatography (HPLC) which has lots of application in analytical chemistry and biochemistry is categorized as a type of column chromatography and has been used for separating and identifying compounds as well as determining their quantity. Two pumps with different characteristics have been applied in HPLC: one is isocratic pump associated with buffer C (95% acetonitrile, 0.1% formic acid). The other pump called nano pump and consists of two buffers: buffer A (5% acetonitrile, 0.1% formic acid) and buffer B (0.1% formic acid). These pumps produce solvent gradient for better separation in complex mixtures of compounds [99].

In the analysis of certain classes of proteins, LC–MS can be considered as a complementary technology for two-dimensional gels/MALDI-TOF-MS. In comparison between 2-D electrophoresis with mass spectrometry, we can see that it has not progressed to the same amount as mass spectrometry. Limitations of two-dimensional gels in detection of different proteins with different characteristics of pH, size and hydrophobicity have not yet been solved. Also it has possibility of overlapping proteins and limited dynamic range [85].

1.9.2 Tandem Mass Spectrometry

The combination of two mass analyzers in tandem mass spectrometry enhances the efficiency of proteomics analysis. Q-TOF mass analyzer has high sensitivity and is able to analyze the quantity and quality of complex samples with high accuracy. Q-TOF mass analyzer is illustrated in figure 1.14. Q-TOF is the most popular mass analyzer and consists of three elements, quadrupole ion selection (Q1), quadrupole collision cell (Q2) and TOF mass analyzer [98].

Precursor or parent ions are chosen and transmitted in MS/MS mode and Q1 in mass filter mode. Ions are fragmented by collision induced dissociation (CID) with neutral gas molecules such as, argon or nitrogen, in Q2 collision cell after they were accelerated to reach the energy range of 20 to 200 eV. Prior to transmission into TOF for further analysis, the remaining parent ions and fragmented ions continue to collide. These collisions result in the loss of kinetic energy. Hence, the ions are focused and their kinetic energy is controlled by collisional damping. Subsequently, TOF analyzer separates the resulting fragments prior to reaching MS/MS spectra detector.

Firstly, in MS, the mass of the intact peptide ion (precursor ion) is measured. In the next step, product ions are obtained by fragmentation of specific precursor ions through CID. MS/MS spectra are obtained based on the masses of the product ions. These spectra are then matched with protein databases using software. Such process is shown in figure 1.15 [100].

Prior to the LC-MS/MS analysis proteins (e.g. A, B, C) are digested into peptides. A survey scan detects peptides in MS. Subsequently, sequence analysis by MS/MS is

performed on the co-eluted peptides (e.g. A1, B2, C4) where peptide ions with different charges states are randomly selected and sequenced. Software detects similar pattern in MS/MS spectra of the peptide sequencing. Several MS/MS scans follow every survey scan [100].

A single scan is performed in MS/MS sequence analysis. As the position of the fragmented peptide could be determined, software can find matching in the spectra database.

Software and databases used in LC-MS protein identification are, ProteinPilot™ from Applied Biosystems and SWISS-PROT, TrEMBL PIR, NCBIInr, ESTdb respectively.

Figure 1-14 Structure of Q-TOF mass analyzer [94]

Q-TOF consists of three elements, quadrupole ion selection (Q1), quadrupole collision cell (Q2) and TOF mass analyzer (Detailed description in page 32)

Figure 1-15 On-line LC-MS/MS Analysis Procedure of Complex Peptide Mixture [100].

(a) A protein mixture is proteolyzed to a highly complex peptide mixture. The peptide mixture is separated by reversed-phase LC. (b) Timeline of automated peptide sequencing by LC/MS/MS analysis. (Detailed description in pages 32-33)

(Permission from ref.100 was obtained from publisher to use this figure)

1.9.3 iTRAQ Coupled 2-D LC-MS/MS

Specific mass tag recognizable by the mass spectrometer while providing the quantification has been generated using differential stable isotope labeling. These tags include isotope coded protein labels (ICPL), isotope-coded affinity tag (ICAT), stable isotope labeling by amino acids in cell culture (SILAC) and isobaric tagging and relative quantification (iTRAQ). In ICAT method, cysteine residues are derivetized with a reagent containing eight zero or eight deuterium atoms and are analyzed by MS analysis. One of the advantages of ICAT is that this labeling method reduces the complexity of samples due to scarcity of cysteine amino acid. However, ICAT has some limitation such as it does not perform well for the proteins with no cysteine residues. Furthermore, its performance on the analysis of post-translational modified proteins is limited [92, 101].

iTRAQ is another method which is widely used in proteomics and biomarkers researches. iTRAQ, abbreviation for isobaric tags for relative and absolute quantification, has been used to find biomarkers for various diseases including cancer [102]. A whole molecule consists of two components, isobaric tag and reactive group. Isobaric tag component is composed of two parts, balance group and reporter group. The mass of the reporter group and balance group are 114-117 and 31-28 respectively, which yields a total mass of the isobaric tag of 145 as shown in figure 1.16 [103]. Currently, commercially existing iTRAQ reagent kits contain four or eight label per kit and hence support labeling of up to four or eight samples per experiment respectively. Trypsin is used to digest the proteins and the resulting fragmented peptides are labeled with different isobaric tags. A mass spectrometer is then used to analyze the peptide

samples. Reporter group was detached from the balance group during MS/MS fragmentation while the balance group remained attached to the peptide. Released reporter groups provide a well-defined MS signal. The differential amount of the reporter groups are in proportion of the labeled peptides from each sample. This is because, the common physicochemical parameters excepting the mass shared by all reporter groups of isobaric labels. The peak area in the spectrum is proportional to the relative abundance of the peptide which in turn is proportional to the original protein in the tandem MS analysis with a labeled reagent. Peptide abundance in different samples can be obtained through the relative concentration of different isobaric reporter ions by mixing four differentially labeled samples and analyzing on MS/MS [104, 105].

Figure 1-16 Working Mechanism of iTRAQ Reagents.

(a) Diagram showing the components of the multiplexed isobaric tagging chemistry. (b) When reacted with a peptide, the tag forms an amide linkage to any peptide amine (N-terminal or α -amino group of lysine). (c) Illustration of the isotopic tagging used to arrive at four isobaric combinations with four different reporter group masses.[103]. (Permission from ref.103 was obtained from publisher to use this figure)

1.9.4 Applying Proteomics to Biomarker Discovery

Biomarker discovery have attracted attention of many researchers mainly due to the development of recent technologies such as genomics and proteomics approaches. Using biomarkers leads to finding more efficient methods in disease prevention and diagnosis as well as development of new therapies and more individualized disease prognosis and treatments. Among the class of all known biomarkers, protein

biomarkers have been one of the most useful biomarkers over the former century [82, 106, 107].

Protein identification and quantitation have been greatly enhanced by the development of novel separation technologies combined with current mass spectrometry techniques. Such proteomics techniques support identification of thousands of proteins from complex compounds [86].

Although existing platforms have been successfully applied in protein identification, there are some research challenges that have to be addressed. As an example, detecting and measuring protein markers in body fluids such as plasma is very important in medical practice. A large dynamic range of protein relative abundance is observed in body fluids due to the greatly complex mixture of proteins content. Hence, current proteomics technologies may leave some important proteins and potential markers undetected [108].

Moreover, very few of the existing clinically used biomarkers are discovered by means of proteomics techniques. This is mainly because of the costly and time consuming process of validating or invalidating biomarkers detected through proteomics studies [109]. This can be greatly enhanced by fast and cost-effective validation approaches through mass spectrometry.

The main challenges in discovery and development of protein biomarkers can be categorized into two distinct proteomics problems [110]:

1. Proteins with low abundance can be detected, identified and quantified from complex biological samples.

2. For validation of the identified protein markers through proteomics approaches, a high throughput and low cost mass spectrometry should be developed.

In summary, although the process of biomarker discovery and development can be potentially revolutionized by proteomics technologies, some technical challenges need to be addressed before such technologies can be realized.

1.10 Metabolomics

The metabolome has been simply defined as the complete set of small molecule metabolites present in the cells, involving in general metabolic reactions needed for normal function of a cell such as growth and maintenance. Small molecules which are present in an organism are considered the basic units of macromolecules formation. A small molecule needs to have two criteria to be considered as a part of metabolome. One is that the molecular weight should be less than 1000 Dalton, and the other one is that they should have been involved in a metabolism or been generated in a metabolic pathway. These metabolites can be signaling molecules and hormones as well as other molecules like metabolic intermediates and secondary metabolites [111, 112]. The metabolome is dynamic meaning that it is different from one second to another like proteome or transcriptome. So far, global profiling of metabolites in an organism is not achievable. However, it is doable to measure large scale cellular metabolites both quantitatively and qualitatively [113]. This can be achieved by applying different technologies. Metabolomics, scientific study of metabolome, is of high significance in identification of a biomarker for disease diagnosis purpose. The reason for this claim is that the metabolite biomarkers are usually at the end point of biological processes or effects [114].

At present, metabolic profiling and metabolic fingerprinting are the two approaches to study the metabolome. In the first approach (metabolic profiling), physiological interactions are identified in a specific biological pathway showing the enzymatic activity. Comparing can be done quantitatively between information produced from normal vs. patient samples. In metabolic fingerprinting, unlike the first approach, comparing is mainly of different patterns which are caused to happen by the environment changes [115]. There are several methods for metabolomics analysis such as Nuclear magnetic resonance (NMR) and mass spectrometry (MS). They are the most widely used analytical tools for metabolomics analysis. MS-based technologies have higher sensitivities comparing to NMR-based ones, meaning they are able to identify low-abundance molecules. Hence, they can have more coverage especially for those metabolites which have lower concentration range leading to more steady results. MS is the most commonly used technology in metabolomics and it also identifies metabolites in a more effective way, both quantitatively and qualitatively [116].

1.10.1 MS-Based Metabolomics

One of the main tools in metabolomics research is mass spectrometry. Metabolites are separated by using chromatographic separation methods including, gas chromatography (GC), capillary electrophoresis (CE) and high-performance liquid chromatography (HPLC) in conjunction with mass spectrometry. Effective methods of separation are crucial in metabolomics studies due to the complex nature of metabolite extraction process [117]. Metabolites can be separated and detected based on their specific retention time especially isomers. On the other hand, mass spectrometry is unable to identify isobaric compounds. Hence, separation techniques such as Gas and

liquid chromatography are generally applied to separate metabolites. Volatile compounds can be effectively analyzed by combination of mass spectrometry and gas chromatography (GC-MS) [118]. Hence, chemical derivatization must be applied to metabolite extracts in case they are not volatile. GC-MS analysis is more effective on some volatile metabolite classes such as esters, lipids and plant terpenes which could not be analyzed by LC-MS analysis because of its poor ionization (Figure 1.17). Amino acids, alcohols, fatty acids and steroids are also well suited for GC-MS analysis. Characteristics of GC-MS such as higher sensitivity, peak resolution and repeatability make it advantageous for metabolite profiling. In addition, GC-MS electron impact spectral library is available for metabolite identification. High performance liquid chromatography (HPLC) is more effective in separating non-volatile metabolites and hence no chemical derivatization is needed. This leads to a wider compound analysis for LC-MS than GS-MS. Polyamines, amino acids and organic acid are usually analyzed by LC-MS. LC-MS analysis has been increasingly used in metabolomics studies recently [119, 120].

Figure 1-17 Classes of metabolites and the MS instruments (GC-MS, LC-MS) [121].

The figure classifies the metabolites based on their polarity. On the left, more polar metabolites are listed which are effectively analyzed by LC-MS. On the right, less polar metabolites are shown. GC-MS is a more effective tool in analyzing these metabolites. Metabolites shown in the middle can be analyzed with both LC-MS and GC-MS.

1.10.2 Gas Chromatography-Mass Spectrometry (GC-MS)

In order to better identify different substances within a sample, Gas Chromatography (GC) features has been combined with the MS ones to generate Gas Chromatography-Mass Spectrometry (GC-MS) which was the most broadly used technique for analysis of small molecules. The use of gas chromatography (GC) as an analytical tool extends for half a century. In these methods, carrier gas elutes analytes from the surface of GC column in a temperature gradient based on their volatility [122]. The detector responds to the carbon content of the analyte proportionally and thus quantifies the analyte. On the other hand, mass spectrometry is commonly used in combination to GC to obtain more information on the analyte such as mass which is useful in identifying unknown analytes. Relatively hydrophobic molecules are best separated by GC. However, polar,

water soluble analytes must be derivatized before GC analysis [123]. GC-MS was first used for metabolite analysis in 1970s. On the other hand, in metabolomics studies, the aim is to detect and identify as many metabolites as possible. GC-MS is still in use and quite helpful in many metabolic diseases detection and with the relative comprehensive databases is widely applied in metabolite studies as well as disease diagnosis. With the help of GC-MS, key small molecules like amino acids, fatty acids as well as organic acids have been identified in blood or other biofluids such as urine [118]. The ionization mechanism between GC-MS and LC-MS is different. Electron Impact Ionization (EI) and Chemical Ionization (CI) are two methods that are used for GC-MS sample ionization. In EI samples collide with an electron beam to create ions; while in CI a reagent gas is used to create ions. In MS, mass analyzers can be single quadrupole, TOF, FT-ICR and Orbitrap. Different part of GC-MS is described in figure 1.18 [118].

Figure 1-18 Different part of GC-MS [118].

The solvent containing metabolites is injected into the GC and vaporized. The vaporized chemicals are carried through a column in the oven by an inert gas where high volatile chemicals pass more quickly than chemicals with low volatility. In MS, the chemicals are bombarded with electrons and turn into ions. The resulting ions are then filtered based on their mass using an electromagnetic field. In the end, a detector quantifies the number of ions with the same mass. These data is transferred to computer and is plotted as mass spectrum.

In GC-MS based metabolite profiling the existing steps are as follows [121]:

- 1) Metabolite Extraction: comprehensiveness should be highly considered in extraction of metabolites from biological sample. Simultaneously, it must prevent metabolite degradation or modification.
- 2) Metabolite Derivatization: In order to prepare metabolites volatile and suitable to gas chromatography, derivatization is necessary. Trimethylsilylation is the most common one among many accessible derivatization methods.
- 3) Separation by GC: In this step, high standard gas-flow condition and temperature programming as well as standardized column material should be performed in order to get reproducible results.
- 4) Ionization: The other step is ionization of compounds while eluting from GC which is necessary for mass spectrometry. Electron impact (EI) ionization is extensively applied. Fragmentation patterns and molecular ions resulting from this method could be repeated and thus makes mass spectra assessment easier.
- 5) Detection of molecular ions: Different mass-detection devices can be applied for detection purposes. These devices include single quadrupole detectors (QUAD), ion trap technology (TRAP) and time-of-flight detectors (TOF).
- 6) Data Evaluation: To assess the data, retention times and mass spectral fragmentation patterns must be compared with existing databases for a possible match. Hence, GC-MS results are given to software which yields mass spectra extracted from

chromatogram, co-eluting metabolites corrections, GC retention time indices and suitable mass fragments for quantification.

A detailed GC-MS-based metabolomics workflow is described in figure 1.19.

Figure 1-19 A schematic GC-MS-based metabolomics workflow.

GC-MS based metabolite profiling has three main steps: sample preparation, GC-MS analysis and metabolites identification as described in page 44 [121].

1.10.3 Applying Metabolomics to Biomarker Discovery

Using Metabolomics for biomarker discovery primarily deals with characterizing differences in metabolites level during abnormal physiological states including altered and stressed states. Such studies demand carefully designed samples collection protocols as well as accurate statistical analysis. Every physiological state has a specific set of metabolite levels that could be used as the candidate biomarker of the state. Each biomarker metabolite is assigned to a specific biochemical assay

corresponding to the candidate biomarker analytes upon validation. Such assays are valuable indicators for clinical diagnostics, similar to the use of glucose levels for diabetes diagnosis [76, 124].

In addition to finding specific biomarkers, metabolomics has a crucial role in developing drugs. Metabolomics analysis, provide invaluable information based on the effectiveness of the drug, specific biomarkers and toxicity analysis of the new compounds and support assessments at pre-clinical and clinical levels. Although metabolomics is a relatively new member of the "Omics" family, it has already started to contribute to the disease diagnosis. In this regard, metabolomics has been applied to cancer diagnosis and treatment. Analysis of the metabolic pattern of cancerous tissue or bodily fluids have confirmed changes in molecular pathology in comparison with healthy tissue and revealed the effects of specific drugs. Several studies have shown the relation between cancer and deviations in main metabolic pathways such as glycolysis, pentose phosphate pathway, tricarboxylic acid (TCA) cycle and biosynthetic pathways of fatty acids and amino acids. In addition to cancer diagnosis, metabolomics studies have found applications in finding new biomarkers for cardiovascular disease [125-128]

Nevertheless, metabolomics is still at the beginning of its development in comparison to other members of 'omics' family and hence there are challenges which need to be solved. Applying the current technologies has limited analysis abilities due to the abundance of chemical characteristics of the metabolites as well as vast dynamic range of metabolite levels. Furthermore, there is a need for new algorithms and software in metabolomics studies and analysis.

1.11 Objectives of Research Project

Hepatocellular carcinoma (HCC) is ranked fifth regarding its prevalence worldwide. On the other hand, it has the third highest mortality rate among all cancers. HCC has poor general prognosis and low survival rate; hence, improving HCC therapy is of high importance. It is estimated that about 53% of HCC patients are afflicted by HBV. The smallest protein of HBV, HBx, has vital role in hepatocellular carcinogenesis [129, 130].

Although many treatments like chemotherapy, liver transplantation, surgical resection, transarterial chemoembolization (TACE) and hormonal therapy have been carried out to improve the prognosis of HCC, there is still much needed to be investigated about HCC as there are still problems in both diagnosis and treatment of the disease. This can be solved by finding suitable biomarkers for early detection and effective HCC therapy [131].

Biomarkers have many applications in medicine including monitoring the treatment of a disease, efficacy of disease progression and disease diagnosis as well as predicting response to a treatment. In this regard, 2-D LC-MS/MS coupled with iTRAQ reagents and GC-MS are effective tools for disease diagnosis and biomarker screening [6, 33, 132].

The main objective of this project is to identify and characterize cellular proteins and metabolites in HepG2 cells transfected by different genotypes of HBx in order to better understand the role of HBV replication in host cell metabolism. We applied omics technologies including LC-MS based proteomics and GC-MS based metabolomics to

detect novel biomarkers for the diagnosis, prognosis and the treatment of HBV-related HCC patients.

LC-MS/MS based proteomics and GC-MS based metabolomics analysis allow the detection and identification of large scale cellular proteins and metabolites, which provides a platform to screen for particular biomarkers as well as establishes cell based HBV replication system which simulates early stages of HCC development. These newly identified proteins and metabolites are involved in several key metabolic pathways such as glycolysis, glutaminolysis, TCA cycle, fatty acids and amino acids metabolism.

The combined proteomic and metabolomic profiling may provide novel information on the fundamental mechanism of HBV replication in host cell metabolism to identify differentially expressed proteins and metabolites, potential as biomarker, for the early diagnosis of HBV-associated HCC. By finding these proteins and metabolites potential biomarkers, it would perhaps facilitate and improve strategy against HBV- associated HCC with better prognosis and effective therapy.

As discussed in chapter 6, a proline rich domain has recently been identified in HBx. The proline rich region of HBx plays significant role in interaction with cellular proteins containing SH3 domains. Therefore, the proline rich domain is very important in virus-host cell interplay.

In this project, we aimed to create point mutations in four proline residues in the proline rich region of HBx (amino acids 28-46) due to the high variability of this region among different genotypes of HBx. We investigated the effect of proline rich

domain on interaction between HBx and cellular proteins containing SH3 domain in order to provide valuable information on how HBx might interfere with host proteins.

Overall, we aim to establish comprehensive cellular proteomic and metabolomic profiling to improve antiviral strategy and to provide new insights for efficient prognosis and diagnosis of HBV-related HCC patients.

Chapter 2 : Materials and Methods

The following section describes the materials and methodologies used in this project including molecular cloning techniques, proteomics and metabolomics analysis in order to establish the comprehensive cellular proteomic and metabolomic profiling of cells transfected by HBx.

2.1 Cloning

2.1.1 Mammalian Expression Vector Used

The pXj40 vector was used in this study in order to clone the construct of interest. Compared with other mammalian vectors such as pcDNA3.1, pXJ40 vector provides a high expression efficacy in mammalian cells. As shown in figure 2.1, in the multiple cloning sites (MCS) of pXJ40 vector, hemagglutinin (HA) tag is located within endogenous restriction sites EcoRI and BamHI. This HA tag is used to detect the expression of target protein. The gene of interest could be inserted into the multiple cloning sites after T7 promoter. The expression of cloned gene is driven by HCMV promoter. In our experiments, different genotypes of HBx were cloned individually into pXJ40 vector using the HindIII and PstI restriction enzymes.

Figure 2-1 Schematic map of mammalian expression vector pXJ40.

Gene of interest could be inserted into the multiple cloning sites after T7 promoter. The expression of cloned gene is driven by HCMV promoter. Ampicillin resistance gene is included for selection of clones.

2.1.2 Cloning of HBx from Different Genotypes

HBx genes of HBV genotypes A, B, C and D were amplified and cloned individually into mammalian expression vector pXJ40. HindIII and PstI were used as restriction enzymes. The coding sequences of the four genes were amplified by polymerase chain reaction (PCR). PCRs were accomplished using the oligonucleotide primers to amplify the virus gene as described in table 2.1. The plasmid is usually diluted at 1:50. PCR was carried out using the provided reagents including: Pfu DNA polymerase (2.5U/μL, Promega), 10X Pfu polymerase reaction buffer (100mM KCl, 200mM Tris-HCl pH 8.8, 100mM (NH₄)₂SO₄), 1% Triton X-100, 20mM MgSO₄ and 1mg/ml nuclease-

free BSA), 50 ng of DNA template, 125 ng each of forward and reverse oligonucleotides primers, dNTP mix (0.2 mM each, Fermentas) were added and PCR was carried out. The cycling conditions are as described in table 2.2.

The confirmation of sequences cloned were obtained through DNA sequencing and compared to HBx sequences archived in NCBI.

Table 2-1 Primers used for different genotypes of HBx

Gene	Direction	Sequence (5'→3')
HBx-A	Forward	CCCAAGCTTGCTGCTAGGCTGTACT
	Reverse	AACTGCAGTTAGGCAGAGGTGAAA
HBx-B	Forward	CCCAAGCTTGCTGCTAGGCTGTGCT
	Reverse	AACTGCAGTTAGGCAGAGGTGAAA
HBx-C	Forward	CCCAAGCTTGCTGCTAGGGTGTGCT
	Reverse	AACTGCAGTTAGGCAGAGGTGAAA
HBx-D	Forward	CCCAAGCTTGCTGCTAGGCTGTGCT
	Reverse	AACTGCAGTTAGGCAGAGGTGAAA

Table 2-2 PCR Program

	Step	Temperature (°C)	Time
	Initial Denaturation	94	5 min
25 cycles	Denaturation	94	50 Sec
	Annealing	60	50 Sec
	Elongation	72	45 Sec
	Final Extension	72	10 min

2.1.3 Gel Extraction and Purification of DNA Fragments

After performing PCR, the products were separated by DNA electrophoresis using 1.5% agarose gel electrophoresis. Ethidium bromide (Bio-Rad) was used for visualization of the DNA fragments under UV. Desired DNA fragments were cut out from the gel followed by purification using QIAquick gel extraction kit.

First we applied QG buffer from the QIAquick kit. QG buffer should be added in a ratio of 3: 1 (volume: weight) to resolve the gel and release the DNA fragment. Then the tubes were incubated in 50°C for 10 minutes with shaking every 3 minutes to dissolve the agarose thoroughly. When the gel was melted completely, 1 gel volume of isopropanol was added and mixed to the QIAquick spin column followed by centrifugation at 13,000 rpm for 2 minutes for binding DNA to the silica membrane. After that, 500 µl buffer QG was added and centrifuge for 1 minute at 13,000 rpm to ensure removal of residual agarose and ethidium bromide (EB). In the next step, 750 µl buffer PE was added as washing buffer and incubated for 5 minutes in room

temperature and centrifuged for 1 minute. Then the QIAquick spin column was transferred to 1.5 ml tube and 35 μ l elution buffer (EB) (10mM Tris-Cl, pH 8.5) or 20 μ l distilled water was added to the centre of column. Incubation was applied for 5 minutes in room temperature followed by centrifugation for 2 minutes at 13,000 rpm.

2.1.4 Restriction Enzyme Digestion and Ligation

Digestion was performed with restriction enzymes (REs) purchased from NEB (New England Biolabs). For each digestion reaction, the restriction enzymes were added in the concentration of 1 unit. 100 $\times$ BSA was added for some enzymes in order to optimize the digestion reaction. For cutting the pXJ40 vector as well as inserts, Hind III and Pst I restriction enzymes were used.

In summary, 0.5-1 μ g DNA (for inserts, 28 μ l purified PCR products) together with 1 μ l of each restriction enzymes (Hind III and Pst I), 4 μ l of NEB buffer (10X), 0.4 μ l BSA and 5.6 μ l H₂O to a final volume of 40 μ l. Then the mixture was incubated at 37°C for 2 hours. 1 μ l Calf intestinal alkaline phosphatase (CIAP) was added to the vector digestion and incubated at 50°C for 30 minutes. This step was applied in order to reduce the self-ligation of vector DNA. Digested DNA products were then re-checked by agarose gel electrophoresis.

After restriction digestion, ligation was applied in order to ligate the DNA insert to the vector. T4 DNA ligase and 10 $\times$ T4 DNA ligase buffer (NEB) were used for ligation purpose. The concentration of vector: insert ratio used in ligation was “1:3” or “1:4”. Ligation mixture included: insert DNA, vector DNA, T4 DNA ligase (10U/ μ l), 10 $\times$

T4 DNA ligase buffer (NEB). The optimal temperature for ligation was 16°C for overnight.

2.1.5 Preparation of Competent Cells by CaCl₂ Method

In our experiments, *Escherichia coli* (DH5 α) strain was used as competent cells. After preparing the LB/Agar plates, 0.2 μ L of DH5 α cells were streaked on agar plates and incubated for overnight at 37°C. After that a single colony of DH5 α cells was selected and added to 5 mL LB broth (without selective antibiotics) followed by incubation at 37°C and shaking at 240rpm for overnight. In the next day, 1 mL of the cultured bacteria was inoculated in to 100 mL LB broth and grown to OD^{600nm} of 0.3 to 0.4. Subsequently, the culture is aliquoted in to 50 mL pre-chilled tubes and incubated on ice for 5 to 10 minutes. Then the tubes were centrifuged for 5 minutes at 5000rpm at 4°C. The cell pellet was re-suspended with 30 ml cold MgCl₂ (0.1M) followed by centrifuging for 5 minutes at 5000 rpm at 4°C. The supernatant was removed and the cell pellet was re-suspended in 20 ml cold CaCl₂ (0.1M) thoroughly and put on ice for 30 minutes. After that, the tubes were centrifuged for 5 minutes at 5000rpm at 4°C followed by re-suspension of cell pallet in 1.5 ml cold CaCl₂ (0.1M) + 15% glycerol. Finally, the prepared competent cells were aliquoted and stored in -80°C for long term usage.

2.1.6 Plasmid Transformation

In order to propagate the plasmids prepared by ligation, 10 μ L of ligation mixture (50ng plasmid DNA) was added to 100 μ L of competent DH5 α bacteria (as described above), mixed then incubated for 30 minutes on ice. Heat shock (45 sec at 42°C followed by 5

minutes cooling down on ice quickly) was carried out on the competent cells and plasmid mixture. After that 800 µl of Luria-Bertani (LB) broth was added into mixture and incubated at 37°C with shaking at 200 rpm for 45 minutes. 100 µl of the transformed cells were spread onto prepared LB agar plate supplemented with ampicillin (100µg/ml) and incubated at 37°C overnight in an upside down position. Single colonies were picked subsequently and inoculated in 5mL fresh LB containing ampicillin in 50mL tubes. The cultures were incubated at 37°C with shaking at 200 rpm for overnight. Preparation of the vectors from transformed *E.coli* cell pellets were carried out via QIAGEN miniprep kit.

2.1.7 Glycerol Stock

After transformation, an individual colony was isolated from LB agar plate and inoculated into 5 ml of LB broth with the selective antibiotics followed by incubation at 37°C (240rpm) for overnight. Subsequently, 200 µL of glycerol was added to 80 µL bacterial culture followed by vortexing to mix thoroughly. The glycerol stock can be store at -80°C for long term.

2.1.8 Plasmid Preparation by QIAGEN Mini Kit

Randomly picked colonies were inoculated into 5ml LB broth containing ampicilline and then incubated at 37°C overnight while shaking at 200 rpm. Plasmid preparation was performed on the overnight culture using QIAGEN miniprerp kit. 5 ml culture was centrifuged at 5,000 rpm for 10 minutes. After that, supernatant was removed and cells pellet was re-suspended in 250 µL Buffer P1 (50 mM Tris-Cl pH 8.0, 10 mM EDTA, RNase A (100µg/mL)). Then, 250 µL lysis buffer P2 (200 mM NaOH, 1% SDS (w/v))

was added and mixed completely by reversing the tube 4-6 times. Concurrently, 350 μ L buffer N3 was added as a neutralization buffer followed by centrifugation at 13,000 rpm for 10 minutes. Supernatant containing plasmid DNA was then transferred into the QIAprep purification column with centrifugation at 13,000 rpm for 30-60 sec. For washing the adsorbed plasmid DNA, 750 μ L buffer PE was added followed by centrifugation for 30-60 sec. For removing the residual completely an additional 1 min centrifugation was performed. In order to elute the DNA 50 μ L buffer EB (10mM Tris-Cl, pH 8.5) or distilled water was added to the centre of QIAprep spin column followed by incubation for 1 minute at room temperature and centrifugation at 13,000 rpm.

2.1.9 Plasmid Preparation by QIAGEN Midi Kit

QIAGEN Plasmid Midi Kit was applied to remove endotoxin and other contaminants in order to purify the plasmid DNA of mammalian cell after transfection.

Transformed bacteria were incubated at 37°C with shaking at 240 rpm for overnight. In the next day, the bacteria were centrifuged at 5,000 rpm for 10 minutes. The supernatant was removed and the cell pellet was resuspended by adding 4 ml buffer P1 ((50 mM Tris-Cl pH 8.0, 10 mM EDTA, supplemented with 100 μ g / ml RNase A). After that, 4 ml buffer P2 (200 mM NaOH, 1 % SDS (w / v)) was added as a lysis buffer. They were mixed thoroughly and incubated at room temperature for 5 minutes. Concurrently, 4 ml buffer P3 (3M potassium acetate, pH 5.5) was added as neutralization buffer followed by reversing the tube 3-4 times for precipitation of proteins, genomic DNA and cell debris. The mixture was incubated on ice for 15 minutes and centrifuged at 15,000 rpm for 30 minutes at 4°C. The supernatant

containing plasmid DNA was centrifuged again at 15,000 rpm for 15 minutes at 4°C. Subsequently, 4 ml of buffer QBT (750 mM NaCl, 50 mM MOPS, pH 7.0, 15 % isopropanol (v / v), 0.15 % Triton X-100 (v / v)) was added and gravity flow was applied to drain the QIAGEN-tip completely. Supernatant was added to QIAGEN-tip immediately and allowed to flow through in order to absorb plasmid on QIAGEN-tip membrane. For washing the adsorbed plasmid DNA, 10 ml buffer QC (1 M NaCl, 50 mM MOPS pH 7.0, 15 % isopropanol (v / v)) was added to QIAGEN-tip for two times. For eluting the plasmid DNA, buffer QF (1.25 M NaCl, 50 mM Tris-Cl pH 8.5, 15 % isopropanol (v / v)) was added as an elution buffer. After that, 3.5 ml isopropanol was added to the eluted DNA for precipitation. The centrifuge was applied at 15,000 rpm for 30 minutes at 4°C. Supernatant was removed carefully and the DNA pellet was washed with 2 ml 70% ethanol followed by centrifugation at 15,000 rpm for 10 minutes. After that, the pellet was collected and dried for 10 minutes. Finally, DNA was dissolved in 100 µl buffer TE (10 mM Tris-Cl pH 8.0, 1 mM EDTA)

2.1.10 Plasmid DNA Quantification by UV Spectrophotometer

Plasmid DNA purified by QIAGEN mini kit or midi kit was quantified by UV spectrophotometer. Reading was taken at wavelengths of 260 nm called OD^{260nm} for DNA plasmid. The plasmids were diluted to proper concentration (100 times) with distilled water. Pure water was used for zero calibration. Considering that 1 OD^{260nm} for double-stranded DNA = 50 ng/ µl, DNA concentration was calculated by multiplication of OD^{260nm} x 50 ng/µl and a further multiplication by the dilution factor. Formula shown as follows:

$$\text{OD}^{260\text{nm}} \times 50 \text{ ng } / \mu\text{l} \times 100 (\text{dilution factor}) = \text{Concentration ng } / \mu\text{l}.$$

The resultant concentrations are shown in table 2.3.

Table 2-3 OD and concentration of the prepared plasmids

	OD (260nm)	Concentration ($\mu\text{g}/\mu\text{l}$)
HBx-A	0.093	0.465
HBx-B	0.124	0.62
HBx-C	0.097	0.48
HBx-D	0.127	0.63

2.2 Mammalian Cell Culture and Transfection

2.2.1 Mammalian Cell Culture and Cell Passage

Human hepatocellular carcinoma cell line (HepG2) and rat primary hepatocytes (RPHs) were used in our experiments. HepG2 cells were cultured in Minimal Essential Medium (MEM), supplemented with 10% fetal bovine serum (FBS) (Gibco, USA), 1% anti-mycotic solution (penicillin, streptomycin and amphotericin B) (Sigma-Aldrich, USA) at 37°C, 5% CO₂. RPHs were cultured in Williams E Medium (WEM) (Invitrogen, USA), supplemented with 5% FBS, 1 μM Dexamethasone (10Mm) and 18 mL cocktail-A (Invitrogen, USA) containing: 5mL penicillin/streptomycin (10,000 $\mu\text{g}/\text{mL}$), 500 μl human recombinant insulin (4 mg/mL), 5 mL GlutaMAX (200Mm) and 7.5 mL HEPES, PH 7.4 (1M).

The cells were passaged when reached 80-90% confluency. For cell passaging, the cells were washed by phosphate buffered saline (PBS) (Gibco, USA) and treated with trypsin (Gibco, USA) for several minutes. The cells were incubated at 37°C for 3 to 5 minutes. After that the cells were examined under the microscope to make sure that they are rounded up and detached from the surface. When about 90% of the cells becoming round and detaching from each other culture medium was added. The cells are then pipetted gently to a centrifuge tube. The suspended cells were centrifuged at 1,700 rpm for 5 minutes at room temperature. The supernatant is removed and the cells pellets are re-suspended in fresh medium. A cell count is determined using the Neubauer hemacytometer (Baxter Scientific, Germany). About 2×10^6 cells are seeded into new petri dishes.

2.2.2 Freezing, Storage and Recovery of Cells

When cells reached to 80-90 % confluency, the culture medium was removed and cells were washed with PBS twice followed by trypsinization for 3 minutes at 37°C. The suspended cells were centrifuged at 1,700 rpm for 5 minutes. 1 ml of frozen buffer (40% FBS + 10% dimethyl sulfoxide (DMSO) + 50% MEM) was added to cell pellet. The cell pellet was re-suspended in frozen buffer and transferred to plastic cryogenic vial (IWAKI, Japan) at 4°C for 1 hour. After that, the cryogenic tubes were placed at -20°C for two hours and finally stored at -80°C for overnight. For long term storage the cryogenic tubes should be placed at liquid nitrogen.

For de-freezing the cells the temperature should be increased very quickly. The tubes were transferred from liquid nitrogen to hot water bath at 37°C quickly. After thawing, 10 ml of fresh culture medium was added. Centrifugation was applied at 1000rpm for

5 minutes in order to remove DMSO. Supernatant containing DMSO was removed and cell pellet was re-suspended in fresh pre-warmed MEM (10 % FBS, 1% anti-mycotic) followed by incubation at 37°C and 5% CO₂. In order to remove residual DMSO completely, cells should be passaged for a minimum of two cycles.

2.2.3 Cell Counting

Neubauer hemacytometer (Baxter Scientific, Germany) was applied for cell counting. Cells were treated with trypsin followed by centrifugation. After that, cell pellet was re-suspended with MEM. About 10 µl of the diluted cell suspension was loaded into the hemacytometer chambers. The number of cells in each of the four corner and central squares was counted using a hand-held counter under an inverted microscope (100X, Olympus). Total cell number was determined using the following equation:

$$Cell\ density\left(\frac{cells}{ml}\right) = \frac{total\ number\ of\ cells}{number\ of\ 1mm^2\ squares\ counted} \times 10^4 \times dilution\ factor$$

For estimating the cell viability, trypan blue was used for cell staining. After cell staining, dead cells were observed as dark blue spots under the microscope because their membranes were permeable to trypan blue. Therefore, cell viability is defined as the ratio of live cells number to dead cells number in the suspension.

2.2.4 Cell Transfection

Transfections of HepG2 cells were performed using Nucleofection solution V (Amaxa GmbH, USA). This method of transfection provides high transfection efficiency. Effectene Transfection Reagent Kit (Qiagen) was applied for transfection of RPHs. Two days prior to transfection, 2×10^6 cells were seeded in 100mm cell culture dishes. The medium was removed 1 day before transfection and replaced with fresh medium. Cells were ready for transfection when they reached 70 to 80% confluency. The medium is first removed and washed with PBS. Trypsin was added and cells were incubated at 37 °C for 5 min. When about 90% of the cells has detached from the surface, trypsin is neutralised with culture medium. The suspension is then centrifuged at $300 \times g$ for 5 minutes at room temperature. The supernatant is then removed completely.

Briefly, HepG2 cells were re-suspended in 100 µl of solution V. Plasmids were then added to the mixture. The cells were transfected with 5 µg of various constructed plasmids. pmaxGFP Vector is added to a sample so that the efficiency of the transfection can be ascertained. The cell/DNA suspension is then transferred into cuvettes using sterile pipettes.

Transfection was performed by using Program T-28 (HepG2 high efficiency). 500 µl of pre-warmed medium was immediately added to the transfected cells and the mixture was carefully transferred into 100 mm cell culture dishes. The transfected cells were incubated at 37 °C/ 5% CO₂ for 4-6h. The medium was removed after the incubation and the fresh pre-warmed medium was added to the cells. Transfected cells were maintained at 37°C and 5% CO₂ for 24h to allow introduction of gene expression.

Green fluorescent protein (GFP) was used as a control to measure the efficiency of transfection under the fluorescence microscope.

2.3 Proteomics

2.3.1 Total Protein Extraction, Protein Digestion and Labeling with iTRAQ Reagents

The cell pellets were harvested 24h after transfection by tripsinization and then lysed in 200µl of lysis buffer (8M urea, 4 % (w/v) CHAPS and 0.05% SDS (w/v))). The mixtures were then sonicated for several seconds and then put on ice for about 20-30 minutes with regular vortexing. After that the mixtures were centrifuged at 15000 x g for 30 minutes at 4°C. Supernatant containing the proteins was then collected for quantification. Protein concentration was determined using Bradford Protein Assay or 2-D Quant Kit as described below.

Total protein concentration was determined by applying Quick Start Bradford Protein Assay (Bio-Rad, USA). In this method, Coomassie Brilliant Blue was bound to proteins and a stable un-protonated blue compound was formed which can be detected at 595nm using a spectrophotometer. 1 mg/ml Bovine serum albumin (BSA) was used as standard control. Different volumes of BSA (0, 5, 10, 15, 20, 25 and 30 µl) were added to 1 ml disposable plastic PLASTIBRAND® cuvettes (Brand, Germany). 1mL of the dye and the standard proteins or the protein to be quantified was added to cuvettes and mixed well by pipetting. The mixture was incubated at room temperature for 5 minutes. A standard curve (shown in figure 2.2) was obtained using various quantity of BSA. OD was obtained at 595nm. Concentration was determined against the BSA standard curve.

Figure 2-2 Reference line obtained for protein quantization.

Protein concentration was measured by using the 2-D Quant Kit (GE Healthcare).

This analysis is generally on the basis of specific attachment of copper ions to proteins. Proteins which are precipitated are isolated and diluted in a copper containing solution and the measurement of the unbound copper is carried out with a colorimetric element. The colour density and the protein concentration related inversely. The analysis shows a linear reaction between 0 and 50 μg .

2-D quant Kit includes the following components:

Precipitant: renders proteins insoluble

Co-precipitant: Reagents and materials components of this reagent co-precipitate with proteins.

Copper solution: This solution is used for re-suspension of precipitated proteins.

Colour reagent A: Reagent A is mixed with reagent B to indicate unbound copper amounts.

Colour reagent B: Reagent B is mixed with reagent A to indicate unbound copper amounts.

Standard Bovine serum albumin: For preparation of standard curve.

We follow the protocol for determining protein concentration.

1) Preparing working solution (100 part color reagent A + 1 part color reagent B) it should be stored at 4 °C.

2) Standard procedure: we use 6 standard tubes and prepare a standard curve according to the table below.

Tube number	1	2	3	4	5	6
standard solution volume	0 µL	5 µL	10 µL	15 µL	20 µL	25 µL
Protein quantity	0 µg	10µg	20µg	30µg	40µg	50µg

3) Add the precipitant to an amount of 500 µL, vortex them and then incubate the tubes at room temperature for 2-3 min.

4) 500 µL of co-precipitant is added to the next tube and vortex them.

5) Centrifuge the tubes at $10\,000 \times g$ for 5 min.

6) At this stage the pellet can be seen. Remove the supernatants.

- 7) Re-centrifuge the tubes at $10\,000 \times g$ for 2 min. use micropipette to remove the remaining supernatant. The visible liquid should not remain in the tubes.
- 8) 100 μ L of copper + distilled H₂O is added followed by vortex.
- 9) For each tube add 1ml of working solution and subsequently mix by inverting the tube.
- 10) Incubation for 15-20 min at room temperature.
- 11) Using the spectrophotometer, measure the absorbance at 480 nm for each sample.

Recovered proteins from 70% confluent 10cm plates are measured as follows:

HepG2 cells transfected with HBx-A: 0.83 mg

HepG2 cells transfected with HBx-B: 0.78 mg

HepG2 cells transfected with HBx C: 0.86 mg

HepG2 cells transfected with empty vector pXJ40 (control): 0.71 mg

After quantification, a total of 100 μ g of each protein sample was precipitated by the addition of five volume of cold acetone. Samples were put at -20 °C at least for 2h. Then tubes were centrifuged at 15,000 rpm for 10 min and the precipitated proteins were collected. Before labelling with iTRAQ reagents, the proteins were dissolved in 20 μ l dissolution buffer (0.5 M triethylammonium bicarbonate, PH8.5 and 1.5 μ l 2% SDS) followed by denaturation and then reduction by 2 μ l 50 mM reducing buffer (TCEP-HCL, tris (2-carboxy-ethyl) phosphine) and incubated at 60 °C for 1h. After

that 1 μ l cysteine blocking reagent (200 mM methylmethanethiosulfonate (MMTs)) was added to alkylate cysteines as described in the iTRAQ protocol (Applied Biosystems, USA). In the next step, 20 μ l of 0.25 μ g / μ l sequence grade modified trypsin (Promega, USA) was added to digest the samples followed by incubation at 37 °C overnight (12 to 16h).

Finally samples were labelled with iTRAQ reagent as follows:

HepG2 transfected with empty vector pXJ40 is labelled with iTRAQ 114.

HepG2 transfected with HBx genotype A (pXJ40-HBx A) is labelled with iTRAQ 115.

HepG2 transfected with HBx genotype B (pXJ40-HBx B) is labelled with iTRAQ 116.

HepG2 transfected with HBx genotype C (pXJ40-HBx C) is labelled with iTRAQ 117.

The labelled samples were then mixed and analyzed by LC-MS/MS as shown in figure 2.3.

Figure 2-3 Layout for the experiment design (using iTRAQ reagents).

Different genotypes of HBx were amplified and cloned individually into mammalian expression vector pXJ40. RPHs and HepG2 cells were then transfected by different genotypes of HBx. Total proteins were extracted, precipitated, digested and labeled by iTRAQ reagents. All samples were analyzed in three independent experiments by using LC-MS/MS.

2.3.2 Online 2D Nano LC-MS/MS Analysis

Mass spectrometry is a feature to detect proteins as biomarkers. Quantitative protein analysis in this approach is performed by labelling the peptides via using isotopes. The analysis was performed on Agilent 1200 nanoflow LC system (Agilent Technologies, USA) interfaced with a QSTAR XL mass spectrometer (Applied Biosystems/MDS Sciex, USA) [133].

Total 10 μ L of the protein mixture was injected in to the PolySulfoethyl A strong cation exchange column, SCX, (0.32 x 50 mm, 5 μ m) of the HPLC system. HPLC is combined of four columns including: SCX column, reversed-phase column and two identical enrichment columns. SCX column and reversed-phase column are used for first and second dimensions respectively while, two other enrichment columns are applied for trapping the peptides. Therefore, two-dimensional separation was performed by these four columns. In LC system, in order to elute the retained peptide from SCX column in the first dimension, buffer D containing various molar concentrations of KCL solution (10, 20, 30, 40, 50, 60, 80, 100, 300 and 500 mM) was injected in sequence [134, 135].

During the first run the 10-port valve is in position 1 (figure 2.4). The flow-through containing peptides that were not binding to the SCX column was fixed on the ZORBAX 300SBC18 enriching column1 (0.3 x 5 mm, 5 μ m). Subsequently a wash step is performed for 100 min with buffer A (5% acetonitrile, 0.1% formic acid) at the rate of 0.005 mL/min for getting rid of any excess reagent. Meantime, ZORBAX 300SB-C18 enrichment column 2 was changed into the path of nanopump. There was

no trace of peptides on the column 2; therefore, it was anticipated that no peptide will be observed after running for 100 min [136].

10-port valve was set into position 2 in the second run (figure 2.5). The retained peptides were eluted to column 2 by injecting 5 μ L of 10mM KCL solution into SCX column. Column 2 was later washed with loading buffer according to the protocol for removing extra reagent. The nanopump was connected to Column 1 containing the unbound peptides from the first run. Then elution of buffers was carried out by buffer B (0.1% formic acid) and buffer C (a nanoflow gradient of 5-80% acetonitrile + 0.1% formic acid) at the rate of 500 nL/min. Increasing concentration of acetonitrile was used for elution of samples which were concentrated. Analytical Zorbax 300SB C-18 reversed-phase column (75 μ m x 50 mm, 3.5 μ m) was subjected to additional separation [136].

In the next run, solvent output of the Nano-pump was fed to the enrichment column 1 while eluted peptides from SCX were trapped by enrichment column 2. After finishing the separations, peptides were moved to the ionization source, electro spray ionization (ESI). ESI converting eluted peptides into ions in gas phase and sends these charged molecules to the mass analyzer. The mass analyzer was combined of quadrupole and time of flight (Q-TOF) that separate peptides based on mass to charge ratio (m/z). Totally, 13 runs were performed to complete one experiment.

Totally, the quality of achieved profiling studies is determined by considering all of the factors like overall sensitivity, detection and fragmentation efficiency and dynamic range, as well as mass resolution and accuracy.

Figure 2-4 Position 1 in 2-D Nano LC setup

The 10-port valve was in position 1: bypass of SCX column, washing of enrichment column1, and analysis of peptides from enrichment column 2. (See page 69 for detailed explanations) [136]. (Permission from ref.136 was obtained from publisher to use this figure)

Figure 2-5 Position 2 in 2-D Nano LC setup.

The 10-port valve was in position 2: bypass of SCX column, washing of enrichment column2, and analysis of peptides from enrichment column 1. (See page 70 for detailed explanations) [136]. (Permission from ref.136 was obtained from publisher to use this figure)

2.3.3 Protein Identification and Data Analysis

Data analysis in LC-MS/MS is performed with ProteinPilot Software 2.0 (Applied Biosystems, Software Revision 50861, USA). This software works based on some principal parameters such as fixed modification of methylmethanethiosulfate labeled cysteine, variable iTRAQ modification of tyrosine, fixed iTRAQ modification of N-terminal and side reaction of serine and threonine residues with iTRAQ reagent. Proteins will be quantified based on ProtScore more than 2.0 that give the confidence value of 99%. Protein identification is based on single discarded unique peptides. Relative concentration of peptides in the samples was analyzed by dividing the monitored peak areas at 115.1, 116.1, and 117.1 m/z by what was monitored at 114.1 m/z. Three criteria were required for proteins for further analysis: 1) high confidence > 95% should be identified for two or more unique peptides. 2) Score of unused protein to be more than 2 with a 99% confidence. 3) Protein Quant must have a P value <0.05. Subsequently, the ratios of the mentioned calculated peak were corrected for their overlap in isotopic contributions, and used for the estimation of relative amounts of a peculiar protein. For proteins which matched in more than two qualified peptide, three intermediate peak area ratios (specified as 115/114, 116/114, and 117/114) were calculated using the peak area ratios of the peptides originating from the same protein. Those peptides which do not possess modifications of N-terminus free amine or modification of iTRAQ related to free amine in lysine were not taken into account in the calculation of protein ratios. To include the effect of protein loading difference, as suggested by Applied Biosystems, the overall ratios of all proteins in the sample was used to normalize the protein ratios [92, 137].

2.4 Western Blot Analysis

Sodium dodecyl sulphate-poly acrylamide gel electrophoresis (SDS-PAGE) was applied for separation of the proteins based on their molecular weight and electrical charges. SDS-PAGE consists of two gels: separation gel (top) and staking gel (bottom). The concentration of staking gel is lower than separation gel, so proteins were first stacked in staking gel. In our experiment, the different concentration of separating gel (6%, 8%, 12% and 15%) were used for better separation efficiency based on the size of our target proteins as shown in table 2.4.

Table 2-4 (a) Different concentrations of separating gels. (b) Components of stacking gel for SDS-PAGE.

(a) Different concentration of separating gels for SDS-PAGE

% Gel	Components	Volume of components (ml)			
		5 ml	10 ml	15 ml	20 ml
8%	H2O	2.3	4.6	6.9	9.3
	30% acryl-bisacrylamide mix	1.3	2.7	4.0	5.3
		1.3	2.5	3.8	5.0
	1.5M Tris (pH8.8)	0.05	0.1	0.15	0.2
	10% SDS	0.05	0.1	0.15	0.2
	10% ammonium persulfate	0.003	0.005	0.009	0.012
10%	TEMED				
	H2O	1.9	4.0	5.9	7.9
	30% acryl-bisacrylamide mix	1.7	3.3	5.0	6.7
		1.3	2.5	3.8	5.2
	1.5M Tris (pH8.8)	0.05	0.1	0.15	0.2
	10% SDS	0.05	0.1	0.15	0.2
12%	10% ammonium persulfate	0.002	0.004	0.005	0.008
	TEMED				
	H2O	1.6	3.3	4.9	6.6
	30% acryl-bisacrylamide mix	2.0	4.0	6.0	8.0
		1.3	2.5	3.8	5.0
	1.5M Tris (pH8.8)	0.05	0.1	0.15	0.2
15%	10% SDS	0.05	0.1	0.15	0.2
	10% ammonium persulfate	0.002	0.004	0.006	0.008
	TEMED				
	H2O	1.1	2.3	3.4	4.6
	30% acryl-bisacrylamide mix	2.5	5.0	7.5	10.0
		1.3	2.5	3.8	5.0
	1.5M Tris (pH8.8)	0.05	0.1	0.15	0.2
	10% SDS	0.05	0.1	0.15	0.2
	10% ammonium persulfate	0.002	0.004	0.006	0.008
	TEMED				
	H2O	1.1	2.3	3.4	4.6
	30% acryl-bisacrylamide mix	2.5	5.0	7.5	10.0

(b) Components of 5% stacking gels for SDS-PAGE

Components	Volume of components (ml)							
	1 ml	2 ml	3 ml	4 ml	5 ml	6 ml	8 ml	10 ml
H ₂ O	0.68	1.4	2.1	2.7	3.4	4.1	5.5	6.8
30% acryl-bisacrylamide mix	0.17	0.33	0.5	0.67	0.83	1.0	1.3	1.7
1.0M Tris (pH6.8)	0.13	0.25	0.38	0.5	0.63	0.75	1.0	1.25
10% SDS	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.1
10% ammonium persulfate	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.1
TEMED	0.001	0.002	0.003	0.004	0.005	0.006	0.008	0.01

After gel preparation, different types of combs (0.8mm, 1.0mm and 1.5mm) were used according to the amount of our proteins. The casted gel was placed in a chamber filled with 1 X running buffer (0.025M Tris, 0.192 M glycine, pH 8.3, 0.1 % SDS). After that, 2 X SDS loading buffer was added to protein samples and incubated at 98 °C for 10 minutes. For protein ladder molecular weight rainbow markers RPN 756 (GE Healthcare, USA) was used. Around 20 µL of the samples was loaded into each well. For stacking gel 80V for 30 minutes was applied and after that 150 V was performed for separating gel.

After electrophoresis, the gel was carefully removed and proteins were transferred onto the 0.45 µm nitrocellulose membrane (GE Hybond) or 0.2 µm PVDF membrane (Bio-Rad) via Bio-Rad's semi-dry transfer system as shown in figure 2.6 for western blot detection.

Figure 2-6 Layout Setup for the Western Blot Transfer Semi-dry System (Bio-Rad) [33].

The proteins on the gels were then transferred onto the transfer buffer pre-wetted nitrocellulose membrane (NC) and transfer buffer washed activated polyvinylidene difluoride (PVDF) by semidry transfer system (Bio-rad, USA). (Detailed description in page 75)

Nitrocellulose membranes were immersed in transfer buffer (25 mM Tris, 192 mM glycine, 10 % (v/v) methanol) 15 minutes before transfer. The setting for the transfer is at constant voltage (21 V) for about 50 minutes (depends on the size of proteins).

The transferred membranes were then subjected to blocking in 5 % non-fat milk solved in 1 × PBS for 1 hour at room temperature. After blocking, specific primary antibodies, dissolved in 3 % non-fat milk / 1 X PBS solution, was reacted with the proteins on the membrane at 4 °C overnight. Unbounded antibodies were then removed by 3 washes of 10 minutes with utilizing cold 1× PBS/0.005 % Tween-20. Secondary antibody conjugated with Horseradish Peroxidase (HRP), was dissolved in 5 % nonfat milk / 1 X PBST solution and incubated for 1 hour at room temperature. This was followed by washing three times with cold 0.005 % Tween-20 / 1X PBS solution. After that, enhanced chemiluminescence (ECL) was applied on the membrane, using

SuperSignal West Pico Chemiluminescent Substrate kit (Thermo). This kit contains two solutions: Stable Peroxide Buffer and Luminol/Enhancer Buffer.

300 µL of each Stable Peroxide buffer and Luminol/Enhancer buffer were added to the membrane followed by incubation at room temperature for 3 minutes. After that the membrane was exposed to Hyperfilm X-ray films (GE healthcare) for different time durations at dark room. In the last step, the membranes were transferred to developing reagent (Kodak) for 5 minutes and fixing reagent (Kodak) for 2 minutes.

The antibodies used in this experiment are from Santa Cruz, Sigma, Abcam and Pierce were as follows:

-
- **Primary antibody mouse anti-β-actin (AC-74, Sigma)**
 - **Primary antibody mouse anti-HA (F-7: sc-7392, Santa Cruz)**
 - **Primary antibody mouse anti-LDH (H-10: sc-133123, Santa Cruz)**
 - **Primary antibody mouse anti-GAPDH (0411: sc-47724, Santa Cruz)**
 - **Primary antibody rabbit anti-PKM (H-59: sc-292640, Santa Cruz)**
 - **Primary antibody mouse anti-PGK1/2 (A-2: sc-166432, Santa Cruz)**
 - **Primary antibody rabbit anti-HXK II (H-95: sc-28889, Santa Cruz)**
 - **Primary antibody mouse Anti-SDH (ab14715, Abcam)**
 - **Primary antibody rabbit anti-FH (ab171948, Abcam)**
 - **Primary antibody mouse anti-Vimentin (RV202: sc-32322, Santa Cruz)**
 - **Primary antibody mouse anti-Annexin A2 (C-10: sc-28385, Santa Cruz)**
 - **Secondary antibody, HRP-conjugated goat anti-mouse IgG (#31430, Pierce)**
 - **Secondary antibody, HRP-conjugated goat anti-rabbit IgG (sc-2004, Santa Cruz)**
-

2.5 Cellular Reactive Oxygen Species Detection

The Reactive oxygen species (ROS) level of HepG2 cells transfected with three different genotypes of HBx was measured by Image-IT™ LIVE Green Reactive Oxygen Species Detection Kit (I36007, Invitrogen).

Fluorescence microscopy is used for detecting the oxidatively stressed and non-stressed cells in ROS detection kit protocol. In aerobic organisms, the production of ROS happens at a regulated rate. When the cells are under oxidative stress, the production of ROS is severely raised. As a result of oxidative stress, the membrane lipids, proteins, and nucleic acids were altered. Variations in these biomolecules lead to different pathological diseases such as atherosclerosis, carcinogenesis and neurodegenerative disorders.

In order to detect the ROS level in live cells, 5-(and-6)-carboxy-2', 7'-dichlorodihydrofluorescein diacetate (carboxy-H2DCFDA) was provided by Image-iT™ LIVE Green Reactive Oxygen Species (ROS) Detection Kit. This reagent is used as a fluorogenic marker for ROS in live cells. ROS detection kit also provided tert-butyl hydroperoxide (TBHP) as a positive control. TBHP acts as a common inducer of ROS production. The mechanism of this reaction is shown in figure 2.7.

Image-iT™ LIVE Green Reactive Oxygen Species (ROS) Detection Kit consists of three components: Component A: 5-(and-6)-carboxy-2', 7'-dichlorodihydrofluorescein diacetate (carboxy-H2DCFDA), Component B: tert-butyl hydroperoxide (TBHP) solution and Component C: Dimethylsulfoxide (DMSO).

Figure 2-7 Mechanism of detection of ROS in live cells.

In live cells the non-fluorescent (carboxy-H₂DCFDA) is cleaved by cellular esterases. When ROS exists, the oxidized carboxy-H₂DCFDA radiates green fluorescence which is detectable by fluorescence microscopy. More green fluorescent cells indicate higher intracellular ROS level [138].

Briefly, 2×10^4 cells were seeded into each well plate. Upon 70-80% confluency, the cells were washed with warm HBSS buffer and 25 μ M carboxy-H₂DCFDA working solution was added to cover the cells for 30 minutes at 37 °C in the dark. After that, cells were gently washed three times with warm HBSS. TBHP was provided in the kit as inducer of ROS production. 100 μ M TBHP was added to the cells and incubated at 37 °C/ 5% CO₂ for 90 minutes. Subsequently, fluorescence was measured at 495 nm excitation and 529 nm emission using a fluorescence microscope (Olympus).

2.6 Metabolomics

2.6.1 Cell Lysis, Metabolite Extraction and Derivatization

The cell pellets were harvested 24h after transfection by trypsinization. 70% Methanol solution was used to extract metabolites; 1 ml of 70% methanol was added to each sample. Metabolite loss during sample preparation was compensated by adding 10 µl internal standard solvent Ribitol (2mg/ml). After that samples were sonicated for about 2 minutes on ice bath. Subsequently, the samples were transferred to 1.5ml centrifugation tubes and centrifuged at 4°C at 15300rpm for 10 minutes. Supernatant containing the metabolite was then collected. The supernatant was evaporated for overnight to complete dryness.

Metabolites were derivetized before GC-MS analysis. Derivatization was performed in two stages. The first step is methoxymation which is the conversion of enol forms of aldehydes and ketones in the metabolites to oximes or alkyloximes. To perform this, samples were dissolved in 50µl of methoxylamine hydrochloride (20mg/mL in pyridine, Sigma Aldrich) followed by room temperature incubation for 1 hour. In the next step, MSTFA (N-methyl-N-(trimethylsilyl) was used in silylation which reduces the metabolites polarity and makes them more volatile. In this process, 100µl MSTFA with 1% TMCS (Trimethylchlorosilane, Sigma Aldrich) was added to each sample. Finally, incubation was performed at 70°C for 30 minutes and then samples were analyzed by GC-MS [139, 140].

2.6.2 GC-MS Analysis

The analysis of the samples was performed by a Shimadzu QP2010Plus GC-MS system (Shimadzu, Kyoto, Japan), with a 30 m $\times$ 0.25 mm i.d.DB5 capillary column (122-5033 Agilent Technologies). The GC-MS system was set to splitless mode to inject 1 μ l of derivetized samples into the system. Cutoff time for the solvent was set at 3.5 min. Carrier gas consists of Helium with a flow of 1.1 mL/min. The temperature at which samples were injected was set to 280°C. Also, the temperature of the ion source was set to 200°C. The oven temperature was set to 100°C for 4 min before it was increased to 320°C at 4°C /min rate where it was maintained for 1.56 min. Electron impact (EI) at 70eV was used as ionization source for mass spectrometry. The data was scanned in 35 m/z to 600 m/z range. Acetone was injected as blank between each sample to avoid cross-contamination of the samples. Automatic Adjustment of Retention Time (AART) correction was performed through a standard alkane series mixture (C-10 to C-40) injection into the GC-MS system before analysis [123].

2.6.3 Metabolites Detection and Quantification

Shimadzu GC-MS solution was used to obtain total ion chromatogram and deconvolution. Chromatograms were processed to reduce noise and correct baseline. Peak deconvolution was performed at peak width of 2 s and smoothing equal to 3. Peaks with half-height width of more than 2 s were used in library search. Two libraries were used to search for candidate compounds: National Institute of Standards and Technology (NIST) Mass Spectra Library and Shimadzu's GC-MS Metabolite Mass Spectra Database.

A similarity index (SI) above 70% was the criterion for naming the compound. In case more than one compound would fall under the same category retention indices in GC–MS metabolite mass spectra database was used to identify candidate compounds more specifically. Principle component analysis (PCA) software SIMCA-P (version 12.0; Umetrics, Umea, Sweden) was employed to find linear combinations between the variables on large amount of metabolites where individual samples have largest difference. Relative chromatograph peak areas were normalized and compared for relative quantification of selected compounds [123, 126].

2.6.4 Statistical Analysis

The significance of intracellular proteins and metabolites level changes from HepG2 cells transfected with different genotypes of HBx was calculated using the Student's t-test from at least 3 different experiments. P-values <0.05 were considered to be statistically significant for all comparisons.

Chapter 3 : Proteomic Analysis of HBx Transfected Rat Primary Hepatocytes and HepG2 Cells: Role of HBx Protein in Cellular Metabolism

(Part of this section was from our paper published in Current Proteomics, 2013, 10, 29-37. Permission was obtained from publisher to use this published article in this dissertation)

3.1 Introduction

Hepatocellular carcinoma (HCC) is a cancer of the hepatocyte, one of the main functional cell types in the liver. HCC is considered as the fifth most prevalent cancer worldwide and, at the same time, has the third rank among all cancer-related deaths. Chronic hepatitis B virus (HBV) infection is a major cause of HCC [61].

Hepatitis B x protein (HBx) is the smallest protein of HBV with 154 amino acids which is encoded by the X region. HBx is a multifunctional protein and has important role in hepatocarcinogenesis. Its role in HBV-associated HCC has been studied extensively due to its importance. HBx has also been related to HCC development with possible linkage to individual genotypes [131, 141].

Many treatments such as chemotherapy, liver transplantation, surgical resection, TACE and hormonal therapy have been carried out to improve the prognosis of HCC. However, there are still difficulties in both diagnosis and treatment of the disease. Such problems can be solved by finding suitable biomarkers for early detection and effective HCC therapy [3]. Biomarkers which are specific biological molecules in the body

fluids like plasma can be used as indicators of a biological state. They have many applications in medicine including monitoring the treatment of a disease, efficacy of disease progression and disease diagnosis as well as predicting response to a treatment [108]. Proteomic studies have had enormous impact on biomarker discovery which is an outstanding field in biomedical sciences. Recent developments of mass spectrometry (MS) and chromatography provide useful tools in proteomic analysis, especially the combination of liquid chromatography (LC) or gas chromatography (GC) and mass spectrometry [83]. Liquid or gas chromatography acts as separation tools while mass spectrometry is used for protein identification. In most cases, LC coupled with tandem MS is needed for a more accurate detection which is also referred to as LC-MS/MS. Tandem MS allows a single charged peptide to be selected, separated, sequenced and detected. Isobaric tag for relative and absolute quantitation (iTRAQ) method is a useful approach in quantitative proteomics [92]. The reagents are composed of an amino reactive NHS group coupled to a balancer and reporter group. The most commonly used method is iTRAQ 4-plex. In this method, up to four samples can be processed in a single experiment, where four different reporter groups (MW: 114Da, 115Da, 116Da, 117Da) will connect to a balancer, contributing to the total molecular weight of 145 Da [101].

The protein samples are digested by trypsin and fragmented peptides are tagged with different isobaric tags. Peptide samples are combined and analyzed in a mass spectrometer. During the MS/MS fragmentation, the bond between reporter group and balance group is cleaved and only the balance group remains bonded to the peptide. Meanwhile, the released reporter groups yield a defined MS signal [105].

In the context of HBV-associated liver diseases, several studies have applied proteomic analysis to identify the effects of HBx on cytoskeletal proteins, calcium binding proteins, cell growth proteins, proteasome pathway proteins, extracellular matrix proteins as well as signaling pathways proteins [6, 33, 129, 130].

Our study is mainly focused on significant metabolic enzymes and the role of HBV replication in host cell metabolism in order to better understand the role of HBV replication in HCC development. In this study, total proteins are extracted from rat primary hepatocytes (RPHs) and HepG2 cells transfected with different genotypes of HBx and analyzed by iTRAQ-coupled 2-D LC-MS/MS.

Our proteomics approach may provide valuable information on the fundamental mechanism of HBV replication in host cell metabolism to identify differentially expressed metabolic enzymes, potential as biomarkers, to improve HBV-associated HCC diagnosis.

3.2 Experimental procedures

In this study, different genotypes of HBx were amplified and cloned individually into mammalian expression vector pXJ40 by HindIII and PstI restriction enzymes. DNA sequencing was applied to confirm the sequence of the clones (described in section 2.1). After that, RPHs and HepG2 cells were transfected by different genotypes of HBx (described in section 2.2). Total proteins were extracted and quantified using 2-D Quant Kit (described in section 2.3.1). Proteins were separated, precipitated, digested and labeled by iTRAQ reagents (described in section 2.3.2). All samples were analyzed in

three independent analyses by using LC-MS/MS. (described in section 2.3.3). The differentially expressed proteins were confirmed by western blot analysis (described in section 2.4). The significance of intracellular proteins level changes from RPHs and HepG2 cells transfected by different genotypes of HBx was calculated using the Student's t-test from at least 3 different experiments. P-values <0.05 were considered to be statistically significant. In addition, ROS detection kit was applied to show the metabolic changes under the effects of HBx (described in section 2.5).

The antibodies were used in this experiment and the dilution ratios are shown in table 3.1.

Table 3-1 Antibodies and their dilution ratios.

Antibody	Dilution ratios
Primary antibody mouse anti- β -actin (AC-74, Sigma)	1:10,000
Primary antibody mouse anti-HA (F-7: sc-7392, Santa Cruz)	1:1000
Primary antibody mouse anti-LDH (H-10: sc-133123, Santa Cruz)	1:500
Primary antibody mouse anti-GAPDH (0411: sc-47724, Santa Cruz)	1:500
Primary antibody rabbit anti-PKM (H-59: sc-292640, Santa Cruz)	1:500
Primary antibody mouse anti-PGK1/2 (A-2: sc-166432, Santa Cruz)	1:500
Primary antibody rabbit anti-HXK II (H-95: sc-28889, Santa Cruz)	1:500
Primary antibody mouse anti-SDH (ab14715, Abcam)	1:1000
Primary antibody rabbit anti-FH (ab171948, Abcam)	1:1000
Secondary antibody, HRP-conjugated goat anti-mouse IgG (#31430, Pierce)	1:5000
Secondary antibody, HRP-conjugated goat anti-rabbit IgG (sc-2004, Santa Cruz)	1:10,000

3.3 Results and discussion

3.3.1 Cloning of HBx and transfection efficiency

Fragments of HBx A/B/C were amplified by PCR and ligated into pXJ40 vector by HindIII and PstI restriction enzyme sites. The results shown in figure 3.1 indicated that all three genotypes of HBx have been successfully cloned into pXJ40 vector. Transfections of RPHs and HepG2 cells were performed using Nucleofection solution V (Amaxa, USA), which provides high transfection efficiency. Figure 3.2 shows the results of RPHs and HepG2 cells 24h after transfection with Green fluorescent protein (GFP).

Figure 3-1 Cloning of different genotypes of HBx in pXJ40.

Electrophoresis of agarose gel is shown. DNA samples included HBx A/B/C were cloned in pXJ40 vector. Lane M is the 1kb DNA marker (Gene Ruler, #SM0311). Lanes 1, 2, 3 represent HBx -A, HBx -B and HBx -C in pXJ40 respectively.

Figure 3-2 GFP is used for tracing the efficiency of transfection in RPHs and HepG2 cells.

(a) and (c) are phase contrast images of RPHs and HepG2 cells 24 hours after transfection. (b) and (d) are corresponding images of the green fluorescent microscope using FITC filter (Olympus fluorescent microscopy).

3.3.2 The expression of HBx in transfected cells

The expression of different genotypes of HBx were confirmed by western blot using anti-HA antibody (figure 3.3). β -actin was used as internal control. The result showed that pXJ40-HBx of three genotypes (A, B and C) were effectively expressed in RPHs and HepG2 cells.

Figure 3-3 Western-blot showing the expression of pXJ40-HBx A/B/C in RPHs.

Lane1: empty vector pXJ40; Lane2: pXJ40-HBx A; Lane3: pXJ40-HBx B; Lane4: pXJ40-HBx C. The Tag antibodies used here were anti-HA (Santa Cruz). β - actin is the internal reference.

3.3.3 Differentially expressed proteins identified by iTRAQ-coupled 2D LC-MS/MS

In order to establish the global protein profile in RPHs and HepG2 cells transfected with genotype specific HBx or with empty vector pXJ40, the iTRAQ was used to label as follows:

RPHs / HepG2 cells transfected with empty vector pXJ40 is labelled with iTRAQ 114, RPHs / HepG2 cells transfected with HBx genotype A (pXJ40-HBx A) is labelled with iTRAQ 115, RPHs / HepG2 cells transfected with HBx genotype B (pXJ40-HBx B) is

labelled with iTRAQ 116 and RPHs / HepG2 cells transfected with HBx genotype C (pXJ40-HBx C) is labelled with iTRAQ 117.

A total of 137 and 152 unique proteins were identified respectively in RPHs and HepG2 cells transfected with HBx in three independent experiments.

Data analysis in LC-MS/MS was performed with ProteinPilot Software 2.0 (Applied Biosystems, Software Revision 50861, USA). This software works based on some principal parameters such as fixed modification of methylmethanethiosulfate labeled cysteine, variable iTRAQ modification of tyrosine, fixed iTRAQ modification of N-terminal and side reaction of serine and threonine residues with iTRAQ reagent. Proteins were quantified based on ProtScore more than 2.0 which gives a confidence value of 99%. Protein identification is based on single discarded unique peptides [92].

The relative amount of a peptide in each sample was calculated by dividing the peak areas observed at 115.1, 116.1, and 117.1 m/z by the observation at 114.1 m/z. Three criteria were required for proteins to be considered for further analysis: 1) high confidence > 95% should be identified for two or more unique peptides. 2) Unused protein score should be more than 2 (99% confidence). 3) P value in the Protein Quant had to be $P < 0.05$.

Among 152 identified proteins in HepG2 cells, 13 proteins satisfied these criteria and the results were reproducible when repeated in triplicates as shown in table 3.2.

From 137 detected proteins in RPHs, 12 proteins display changes in their expression levels as shown in table 3.3.

These differentially expressed proteins were clustered into 5 categories based on their cellular functions. They included metabolism, cytoskeleton, signal transduction, stress and defense and biosynthesis proteins as shown in figure 3.4.

Figure 3-4 Differentially expressed proteins identified in RPHs and HepG2 cell.

These proteins categorized into metabolism, cytoskeleton, signal transduction, stress and defense and biosynthesis proteins.

Among the 19 differentially expressed proteins identified in RPHs and HepG2 cells transfected by HBx, 6 proteins are common in RPHs and HepG2 cells.

These 6 proteins are lactate dehydrogenase (LDH), phosphoglycerate kinase (PGK), pyruvate kinase (PK), succinate dehydrogenase (SDH), vimentin and annexin A2.

All these proteins displayed similar pattern in up and down regulation of their levels in RPHs and HepG2 cells transfected by HBx when compared to the pXJ40 control group.

In the group of metabolic proteins, a number of glycolytic enzymes including lactate dehydrogenase (LDH), hexokinase (HK), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK) and pyruvate kinase (PK) were significantly up-regulated while two TCA cycle enzymes, succinate dehydrogenase (SDH) and fumarate hydratase (FH) were down-regulated in RPHs and HepG2 cells transfected by HBx.

The up-regulation of glycolytic enzymes and down-regulation of TCA cycle enzymes may suggest that HBx can shift the metabolic pathways of cells from oxidative phosphorylation to glycolysis. This metabolic alternation will induce a condition in cells known as oxidative stress. Under conditions of oxidative stress, ROS production is dramatically increased which, in turn, leads to hepatocellular carcinogenesis.

In the group of cytoskeleton proteins, vimentin was up-regulated by HBx in both RPHs and HepG2 cells. Vimentin is one of the major components of intermediate filament. It has been found to be over-expressed in various cancers such as prostate cancer, breast cancer, malignant melanoma, and lung cancer [142, 143].

Up-regulation of cytoskeletal proteins such as vimentin, filamin and tubulin may have influence on cytoskeleton organization and can affect the cell movement tendency. On

the other hand, down-regulation of lamin A/C has been found in several cancers such as gastric carcinoma and colon cancer. Down-regulation of lamin A/C by HBx may associate with reduction in focal adhesion and increase in cell motility [33, 144].

Previous studies have shown that HBx is involved in the cytoskeletal reorganization in response to HBV replication. Reorganization of cytoskeleton may have an effect on cell adhesion and migration which leads to HCC development [44]. Therefore up-regulation of vimentin by HBx may reveal a strong cell movement tendency.

In the group of signalling proteins, annexin A2 was up-regulated by HBx in RPHs as well as HepG2 cells. Annexin A2 is calcium-dependent, phospholipid-binding protein. It has a role in many biological processes such as angiogenesis, proliferation, apoptosis, cell migration, tumor invasion and progression [145]. It is up-regulated in various cancer types such as gastric cancer [146], renal carcinoma [147], colorectal cancer [148], breast cancer [149], pancreatic cancer [150] and hepatocellular carcinoma [151]. In the context of HCC, annexin A2 can be up-regulated in higher metastatic HCC cell lines and over-expressed in malignant hepatocytes in HCC [151].

Studies have shown that annexin A2 also up-regulated in HBV transfected HepRG cells [152]. This is consistent with our results on the up-regulation of annexin A2. Therefore, up-regulation of annexin A2 by HBx may be associated with HCC development.

Table 3-2 List of differentially expressed proteins in HBx transfected HepG2 cells.

Protein name	Avg. A:N ($\pm$ SD)	<i>p</i> (A:N)	Avg. B:N ($\pm$ SD)	<i>p</i> (B:N)	Avg. C:N ($\pm$ SD)	<i>p</i> (C:N)	Function
Metabolism							
Lactate dehydrogenase (P00338)	1.40 $\pm$ 0.08	0.013	1.49 $\pm$ 0.07	0.001	1.68 $\pm$ 0.11	0.003	Catalysis the inter-conversion of pyruvate to lactate
Glyceraldehyde-3- phosphate dehydrogenase (P04406)	1.63 $\pm$ 0.16	0.009	1.51 $\pm$ 0.06	0.004	1.42 $\pm$ 0.08	0.001	Glycolytic enzymes. Catalyzes step 6 of glycolysis
Phosphoglycerate kinase (P00558)	1.39 $\pm$ 0.09	0.002	1.28 $\pm$ 0.13	0.005	1.32 $\pm$ 0.08	0.003	Glycolytic enzymes. Catalyzes step 7 of glycolysis
Pyruvate kinase M1/M2 (P14618)	1.47 $\pm$ 0.09	0.006	1.56 $\pm$ 0.04	0.009	1.37 $\pm$ 0.12	0.005	Glycolytic enzymes. Catalyzes step 10 of glycolysis
Succinate dehydrogenase (P21912)	0.61 $\pm$ 0.04	0.012	0.72 $\pm$ 0.08	0.002	0.77 $\pm$ 0.10	0.01	TCA cycle enzyme, converts succinate to fumarate
Cytoskeleton							
Vimentin (P08670)	1.24 $\pm$ 0.09	0.003	1.17 $\pm$ 0.01	0.005	1.20 $\pm$ 0.03	0.003	Class-III intermediate filaments
Lamin-A/C (P02545)	0.82 $\pm$ 0.04	0.01	0.84 $\pm$ 0.08	0.007	0.65 $\pm$ 0.01	0.02	Components of the nuclear lamina
Filamin-B (O75369)	1.07 $\pm$ 0.10	0.04	1.15 $\pm$ 0.01	0.001	1.12 $\pm$ 0.01	0.02	Link between cell membrane and cytoskeletal network, regulates skeletal development

Signal transduction	Avg. A:N ($\pm$ SD)	<i>p</i> (A:N)	Avg. B:N ($\pm$ SD)	<i>p</i> (B:N)	Avg. C:N ($\pm$ SD)	<i>p</i> (C:N)	Function
Annexin A2(P07355)	1.25 $\pm$ 0.02	0.009	1.39 $\pm$ 0.04	0.004	1.23 $\pm$ 0.07	0.001	Calcium-regulated membrane-binding protein, plays role in cell motility and cell matrix interactions
Protein S100-A4(P26447)	1.19 $\pm$ 0.07	0.01	1.22 $\pm$ 0.05	0.003	1.05 $\pm$ 0.12	0.008	contribute to cellular calcium signaling, regulation of cell cycle and differentiation
Stress and defense							
Heat shock protein HSP 90(P08238)	1.15 $\pm$ 0.05	0.001	1.20 $\pm$ 0.10	0.002	1.14 $\pm$ 0.11	0.04	Molecular chaperone; Implicated in protein folding and degradation
Biosynthesis							
60s acidic ribosomal protein P1 (P05386)	1.21 $\pm$ 0.03	0.02	1.18 $\pm$ 0.05	0.007	1.22 $\pm$ 0.01	0.005	plays an important role in the elongation step of protein synthesis
Protein disulfide-isomerase A3 (P30101)	0.80 $\pm$ 0.02	0.008	0.92 $\pm$ 0.10	0.01	0.96 $\pm$ 0.03	0.009	Ribonucleoprotein, Implicated in wide range of cellular processes

Data were reported as average fold changes $\pm$ SD (standard deviation) from at least three independent experiments. A: N is the ratio of different protein expression levels in HepG2 cells transfected by HBx-A to the control; B:N is HBx-B to the control; C:N is HBx-C to the control. Control is protein extracts from HepG2 cells transfected by empty vector pXJ40. P-values <0.05 were considered to be statistically significant.

Table 3-3 List of differentially expressed proteins in HBx transfected RPHs.

Protein name	Avg. A:N ($\pm$ SD)	p (A:N)	Avg. B:N ($\pm$ SD)	p (B:N)	Avg. C:N ($\pm$ SD)	p (C:N)	Function
Metabolism							
Lactate dehydrogenase (P04642)	1.62 $\pm$ 0.06	0.002	1.57 $\pm$ 0.05	0.005	1.55 $\pm$ 0.05	0.008	Catalysis the inter-conversion of pyruvate to lactate
Hexokinase 2 (P27881)	1.40 $\pm$ 0.02	0.04	1.49 $\pm$ 0.01	0.001	1.31 $\pm$ 0.10	0.003	Glycolytic enzymes. Catalyzes first step of glycolysis
Phosphoglycerate kinase (P16617)	1.29 $\pm$ 0.04	0.006	1.37 $\pm$ 0.07	0.004	1.34 $\pm$ 0.08	0.002	Glycolytic enzymes. Catalyzes step 7 of glycolysis
Pyruvate kinase M1/M2 (P11980)	1.56 $\pm$ 0.05	0.005	1.38 $\pm$ 0.11	0.007	1.25 $\pm$ 0.04	0.006	Glycolytic enzymes. Catalyzes step 10 of glycolysis
Succinate dehydrogenase (Q920L2)	0.58 $\pm$ 0.03	0.01	0.66 $\pm$ 0.09	0.002	0.62 $\pm$ 0.01	0.02	TCA cycle enzyme, converts succinate to fumarate
Fumarate hydratase (P14408)	0.83 $\pm$ 0.01	0.03	0.91 $\pm$ 0.08	0.008	0.78 $\pm$ 0.06	0.01	TCA cycle enzyme, converts fumarate to malate
Cytoskeleton							
Tubulin alpha-1B chain (Q6P9V9)	1.21 $\pm$ 0.08	0.04	1.15 $\pm$ 0.01	0.01	1.13 $\pm$ 0.11	0.001	Tubulin is a major constituent of microtubules
Vimentin (P31000)	1.19 $\pm$ 0.01	0.005	1.22 $\pm$ 0.12	0.01	1.28 $\pm$ 0.02	0.007	Class-III intermediate filaments

Signal transduction	Avg. A:N ($\pm$ SD)	<i>p</i> (A:N)	Avg. B:N ($\pm$ SD)	<i>p</i> (B:N)	Avg. C:N ($\pm$ SD)	<i>p</i> (C:N)	Function
Annexin A2 (Q07936)	1.29 $\pm$ 0.11	0.001	1.32 $\pm$ 0.03	0.008	1.17 $\pm$ 0.09	0.009	Calcium-regulated membrane-binding protein, plays role in cell motility and cell matrix interactions
Stress and defense							
60 kDa heat shock protein, mitochondrial (P63039)	1.25 $\pm$ 0.05	0.007	1.28 $\pm$ 0.13	0.01	1.1 $\pm$ 0.01	0.003	Involved in Import and folding of proteins
Biosynthesis							
40S ribosomal protein SA (P38983)	1.08 $\pm$ 0.03	0.006	1.16 $\pm$ 0.05	0.02	1.13 $\pm$ 0.06	0.005	catalyze protein folding, formation and breakage of disulfide bonds between cysteine residues
T-complex protein 1 subunit epsilon(Q68FQ0)	0.90 $\pm$ 0.04	0.002	0.88 $\pm$ 0.02	0.03	0.96 $\pm$ 0.04	0.01	Molecular chaperone; proteins folding upon ATP hydrolysis

Data were reported as average fold changes $\pm$ SD (standard deviation) from at least three independent experiments. A: N is the ratio of different protein expression levels in RPHs transfected by HBx-A to the control; B:N is HBx-B to the control; C:N is HBx-C to the control. Control is protein extracts from RPHs transfected by empty vector pXJ40. P-values <0.05 were considered to be statistically significant.

3.3.4 Differentially expressed metabolic proteins

Our LC-MS/MS results demonstrated that expression level of several metabolic enzymes involved in glycolysis and TCA cycle were altered in RPHs and HepG2 cells transfected by HBx.

Glycolytic enzymes including LDH, HK, GAPDH, PGK and PK were significantly up-regulated while the TCA cycle enzymes SDH and FH were down-regulated in both RPHs and HepG2 cells transfected with different genotypes of HBx compared to the pXJ40 control group as shown in figure 3.5.

LDH catalyzes the inter-conversion of pyruvate to lactate and is finally excreted from the cell. In the process, NADH gives up its electrons and is converted into NAD^+ . This regeneration of NAD^+ is required to maintain the reaction of glycolysis. Up-regulation of LDH has been found in several types of cancers, therefore measuring the LDH levels can be helpful in monitoring treatment of cancer [65, 153].

HK, GAPDH, PGK and PK have roles in the first step (glucose converted to glucose-6-phosphate), step 6 (glyceraldehyde phosphate converted to 1, 3-bisphosphoglycerate), step 7 (1, 3-bisphosphoglycerate converted to 3- phosphoglycerate) and step 10 (phosphoenolpyruvic acid converted to pyruvate) of glycolysis respectively. These glycolytic enzymes play an important role in promoting tumor progression and metastasis in some cancers such as gastric, lung and prostate cancers [154-156].

Several studies found that there is a link between mitochondrial dysfunction and tumorigenesis. As several mitochondrial proteins including FH and SDH act as tumor

suppressor therefore, down-regulation of these TCA cycle enzymes can lead to oncogenesis [157, 158].

Figure 3-5 The differential expression level of several metabolic enzymes in HBx transfected HepG2 cells (a) and RPHs cells (b).

Three independent experiments were carried out by LC-MS/MS. The four columns of each group represent pXJ40-HepG2 (control), pXJ40-HBx A, pXJ40-HBx B and pXJ40-HBx C respectively. Asterisk indicates $p < 0.05$.

3.3.5 Validation of differentially expressed proteins by Western blot

RPHs and HepG2 cells transfected with different genotype of HBx were harvested 24 hours after transfection and proteins were extracted by adding lysis buffer and centrifugation. The changes of proteins levels obtained from LC-MS/MS analysis were also confirmed by Western blot analysis.

Figure 3.6 shows the expression level of GAPDH, LDH, PK, PGK, HK, SDH and FH in RPHs and HepG2 cells.

The expressions of glycolytic enzymes were up-regulated significantly while FH and SDH were down-regulated in RPHs and HepG2 cells transfected by HBx A/B/C in comparison with HepG2 cells transfected by empty vector pXJ40. Western bands were analyzed and quantified by ImageJ software as shown in figure 3.7.

The Western blot analysis of these differentially expressed proteins was consistent with the results of LC-MS/MS.

(a)

(b)

Figure 3-6 Western blot confirmations of the differentially expressed proteins in RPHs (a) and HepG2 cells (b) transfected with HBx A/B/C.

Beta-actin used as internal control. GAPDH, LDH PK, PGK and ACC were up-regulated whereas SDH was down-regulated compared with negative control (empty vector pXJ40).

Figure 3-7 Fold changes of protein level by western blot were plotted using Image J software in RPHs (a) and HepG2 cells (b) transfected with HBx A/B/C.

The four columns of each group represent pXJ40-HepG2 (control), pXJ40-HBx A, pXJ40-HBx B and pXJ40-HBx C respectively. Beta-actin used as internal control.

3.3.6 ROS levels are increased in HepG2 cells transfected by HBx

Molecular oxygen generates reactive oxygen species (ROS). This ROS is mainly originated from the mitochondria and metabolism of glucose. Transcription of different genes involved in glycolysis and ROS is induced during hypoxia in response to hypoxia inducible factor-1 (HIF-1). Elevated levels of ROS can damage genome stability, tumor suppressor gene functions and regulation of cell growth.

ROS levels of HBx transfected HepG2 cells were detected using ROS detection kit. In the presence of ROS, the oxidized carboxy-H2DCFDA radiates green fluorescence which is detectable by fluorescence microscopy.

More green fluorescent cells indicate higher intracellular ROS level. As shown in figure 3.8, ROS levels in HepG2 cells transfected by HBx A/B/C were significantly higher than HepG2 cells transfected by empty vector. Therefore, HBx induces oxidative stress and ROS production which can contribute to hepatocellular carcinogenesis associated with HBV infection.

Figure 3-8 Detecting the ROS levels in HBx transfected HepG2 cells by ROS assay kit.

(A) Left panel is fluorescent view and right panel is light view of HepG2 cells. **(B)** The histogram plotted using ImageJ software indicates fluorescence intensity. a) HepG2 cells transfected by the empty vector pXj40 as control; b) HepG2 cells transfected by HBx-A; c) HepG2 cells transfected by HBx-B; d) HepG2 cells transfected by HBx-C.

3.3.7 HBx and reprogramming energy metabolism

According to Warburg effect, in most cancer cells, the rate of glycolysis and lactate production in the presence of oxygen has been increased [159]. In majority of cancers, pyruvate generated by glycolysis is converted into lactate by lactate dehydrogenase rather than being metabolized in mitochondria through oxidative phosphorylation [69, 70].

Our results indicated that up-regulation of some glycolytic enzymes including LDH, HK, GAPDH, PGK and PK enhanced glycolysis. On the other hand, down-regulation of SDH and FH lead to reduction of the TCA cycle. Therefore, HBx transfected RPHs and HepG2 cells may switch their metabolism from oxidative phosphorylation to glycolysis.

Virus infection can be a cause of cancer in some cases. It is interesting that some viruses such as HBV are altering gene expression of tumor suppressors and oncogenes by disrupting mitochondrial function and energy metabolism [160]. HBx has been shown to bind to tumor suppressor protein p53 and hence suppress its function. HBx also blocks hypoxia-inducible factor (HIF) ubiquitination and causes increasing HIF stability and activity [161]. HIF acts as oncogene in many types of cancer. Overexpression of HIF causes a shift from oxidative phosphorylation to glycolysis through up-regulation of glucose transporters and multiple enzymes of the glycolytic pathway. In addition, HIF stimulates ROS production [160, 162].

Cooperation between MYC, a known oncogene, and HIF suppresses pyruvate dehydrogenase and induces lactate dehydrogenase; thus, pyruvate produced by

glycolysis is converted to lactate instead of entering to TCA cycle as shown in figure 3.9 [68].

Our results showed that HBx can shift the metabolic pathways of cells from oxidative phosphorylation to glycolysis. This metabolic alternation will induce a condition in cells known as oxidative stress. Under conditions of oxidative stress, ROS production is dramatically increased which, in turn, leads to hepatocellular carcinogenesis [163].

Figure 3-9 HBx can shift the metabolic pattern of cells from oxidative phosphorylation to glycolysis.

HBx stimulates HIF activity and leads to increasing lactate dehydrogenase. Thus, pyruvate produced from glycolysis which doesn't enter in to Krebs cycle, is converted to lactate by lactate dehydrogenase. Abbreviations, LDH: lactate dehydrogenase; PDH: pyruvate dehydrogenase; HIF: hypoxia-inducible factor; Ac-CoA: Acetyl-CoA; SDH: succinate dehydrogenase; FH: fumarate hydratase. Induction / repression are illustrated by -----> / ----| respectively. Up-regulation/ down-regulation are illustrated by ↑ / ↓ respectively.

3.4 Conclusions

In summary, LC-MS/MS results showed that a series of glycolytic enzymes including LDH, HK, GAPDH, PGK and PK were significantly up-regulated while SDH and FH were down-regulated in RPHs and HepG2 cells transfected with different genotypes of HBx in comparison to control group.

Up-regulation of glycolytic enzymes suggests that HBV replication could alter host cell metabolism by accelerating the rate of glycolysis to provide important metabolic requirements for nucleotides, amino acids and lipids synthesis. Our findings indicate that even in the presence of oxygen, HBx can reprogram the energy metabolism of cells from oxidative phosphorylation to glycolysis by altering expression of tumor suppressor genes and oncogenes such as p53 and HIF. HBx has the ability to bind to p53 and suppress its function. In addition, it stimulates HIF stability and activity. As a result, p53 knockdown and HIF overexpression can induce a shift from oxidative phosphorylation to lactic fermentation. This metabolic alternation will induce oxidative stress and ROS production which, in turn, leads to hepatocellular carcinogenesis. Our proteomics approach provides information on the principles of HBV replication in the metabolism of host cell, elucidating the role of HBV replication in HCC development.

Taken together, our cell-based system demonstrated that a number of glycolytic enzymes were expressed differentially in HBx transfected cells compared to the control group. These proteins may have a potential to be used as HCC markers. Such a potential could be validated by testing their presence in HCC cells from clinical samples in future.

Chapter 4 : Metabolomic Analysis of HBx Transfected HepG2 Cells: Insights on Cellular Metabolism and Biomarker Discovery

4.1 Introduction

Although HBV has been discovered for more than 30 years, it remains a global health concern. HBV liver infections usually lead to liver cirrhosis, liver failure and primary liver cancer. HCC is mainly developed as the result of chronic HBV infection [6].

HCC is considered as the fifth most prevalent cancer worldwide and, at the same time, has the third rank among all cancer-related death. So far, many treatments have been tried to improve the prognosis of HCC. Yet, there are still problems seen in both diagnosis and treatment of the disease and still much needed to be investigated about HCC. This can be potentially solved by finding suitable and effective biomarkers to trace the disease and hence to improve HCC therapy [3, 130].

The aim of the current study is to identify intracellular metabolites potential biomarkers by using a metabolomics approach in order to further improve HBV-associated HCC diagnosis.

Metabolomics is assigned as the study of metabolite profiles of different processes occurring in a biological system, which includes cells, tissues and organisms.

The metabolome has been simply defined as complete set of small molecule such as amino acids, carbohydrates, organic acids, nucleotides, fatty acids, ketones and

aldehydes which are involved in general metabolic reactions needed for normal function of a cell such as growth and maintenance. Metabolites which are present in an organism are considered the basic units of macromolecules formation. These metabolites can be signalling molecules and hormones as well as other molecules like metabolic intermediates and secondary metabolites [111]. Metabolomics approach to cancer diagnosis has drawn attention of the researchers during the past years. The reason is that metabolomics study yields better understanding of the cancer pathogenesis as well as finding tumor biomarkers which leads to early detection of the disease and thus improved treatment [164].

Metabolomics is of high significance in identification of a biomarker for disease diagnosis purposes. The reason for this claim is that the metabolite biomarkers are usually at the end point of biological processes or effects. Several studies suggested that metabolic abnormalities are related with the development of HBV-associated HCC [62, 64]. Metabolic abnormalities develop in the liver as a result of exogenous factors or pathological agents like HCC. Irregular amounts of metabolites may show changes in the biochemical pathways in the whole body due to pathological state. As a result, monitoring alterations in biochemical pathways can bring about new perspectives in how diseases function and reveal diagnostic biomarkers, making it very useful in therapy and prognosis [127, 128].

Gas chromatography and mass spectrometry are effectively combined to identify and quantify metabolites. GC-MS is a powerful technique in metabolomics due to its high GC resolution which gives rise to a good separation and also due to its sensitive MS

which results in a reliable detection; hence, it can provide confident metabolite information both qualitatively and quantitatively [118].

Today, GS-MS is quite helpful in many metabolic diseases detection and with the relative comprehensive databases, it has been applied in metabolite studies as well as disease diagnosis [114].

In this study, GC-MS based metabolomics was applied to reveal the effective metabolites potential as biomarkers in HepG2 cells transfected with different genotypes of HBx (A, B, C and D) and those transfected by empty vector as a control. GC-MS results determined that HBx can alter the levels of several metabolites incorporated in glycolysis, TCA cycle and fatty acids or amino acids metabolism. These findings suggested that HBV replication could alter host cell metabolism by increasing the rate of glycolysis and inhibiting the TCA cycle. By finding these metabolites potential as biomarkers, it would be much easier to improve the strategy against HBV associate-HCC with better prognosis and effective therapy.

4.2 Experimental procedures

In this study, four genotypes of HBx (A, B, C and D) were amplified and cloned individually into mammalian expression vector pXJ40 by HindIII and PstI restriction enzymes. DNA sequencing was applied to confirm the sequence of the clones (described in section 2.1). After that HepG2 cells were transfected by different genotypes of HBx (described in section 2.2). Western blot analysis was applied for validation of HBx genotypes expression in HepG2 cell line (described in section 2.4). In the next step, cells were lysed and metabolites were extracted and derivetized before

GC-MS analysis. All samples were analyzed in three independent analyses using GC-MS. Large amounts of metabolites were detected and quantified with PCA by software SIMCA-P (described in section 2.6).

The significance of intracellular metabolites level changes from HepG2 cells transfected by different genotypes of HBx was calculated using the Student's t-test from at least 3 different experiments. P-values <0.05 were considered to be statistically significant.

4.3 Results and discussion

4.3.1 Cloning of HBx and transfection efficiency

Fragments of HBx A/B/C/D were amplified by PCR and ligated into pXJ40 vector cut by HindIII and PstI restriction enzymes. The results indicated that all four genotypes of HBx have been successfully cloned into pXJ40 vector (figure 4.1(a)). Transfections of HepG2 cells were performed using Nucleofection solution V (Amaxa, USA) which provides high transfection efficiency. HepG2 cells demonstrated efficient transfection with green fluorescent protein 24 hours after transfection (figure 4.1(b)).

Figure 4-1 Cloning of different genotypes of HBx in pXJ40 vector and transfection efficiency.

(a) Agarose gel electrophoresis of DNA samples including HBx A/B/C/D cloned into the pXJ40 vector. Lane M is the 1kb DNA marker (Gene Ruler, #SM0311). Lanes 1, 2, 3, 4 represented HBx -A, HBx -B, HBx -C and HBx -D in pXJ40 respectively. **(b)** HepG2 cells after transfection with GFP using Amaxa Nucleofector. GFP was used for tracing the efficiency of transfection. Left is phase contrast image of HepG2 cells 24 hours after transfection. Right is the corresponding image of the green fluorescent cells by GFP in the same field (Olympus fluorescent microscopy).

4.3.2 Validation of HBx genotypes expression in HepG2 cells

The expression of different genotypes of HBx were confirmed by western blot using anti-HA antibody. Figure 4.2 shows that pXJ40-HBx of four genotypes (A, B, C and D) were effectively expressed in HepG2 cells.

Figure 4-2 Western-blot showing the expression of pXJ40-HBx A/B/C/D.

Lane1: empty vector pXJ40; Lane2: pXJ40-HBx A; Lane3: pXJ40-HBx B; Lane4: pXJ40-HBx C; Lane5: pXJ40-HBx D. The Tag antibodies used here were anti-HA (Santa Cruz). β -actin was used as internal control.

4.3.3 GC-MS-based metabolomic profiling of HepG2 cells transfected by HBx

In this study, GC-MS based metabolomics was performed in order to establish a comprehensive cellular metabolomic profiling of HepG2 cells transfected with different genotypes of HBx. Figure 4.3 shows an overlay of representative GC-MS chromatography of HepG2 cells transfected by HBx. The overlay of GC-MS chromatography data included a total of 71 intracellular metabolites identified from three independent experiments. These identified metabolites included amino acids, hydroxy acids, carboxylic acids, fatty acids, carbohydrates, alcohols, ketones and aldehydes, nucleosides and nucleoside conjugates as well as other organic acids. The distribution of these metabolites is shown in figure 4.4.

Out of 71 metabolites, 10 showed significant changes in HepG2 cells transfected with different genotypes of HBx (table 4.1). Lactic acid, succinic acid, glucose-6-phosphate, glutamate, valine, glycine, alanine and proline are significantly up-regulated while palmitic acid and stearic acid are down-regulated in HepG2 cells transfected with different genotypes of HBx as shown in figure 4.5.

Figure 4-3 Overlay of representative GC-MS chromatography of HepG2 cells transfected with HBx.

Figure 4-4 Distribution of all metabolites identified by GC-MS.

Table 4-1 List of differentially expressed intracellular metabolites in HepG2 cells transfected by HBx.

Metabolites	RT (min)	Avg. A:N($\pm$ SD)	<i>p</i> (A:N)	Avg. B:N($\pm$ SD)	<i>p</i> (B:N)	Avg. C:N($\pm$ SD)	<i>p</i> (C:N)	Avg. D:N($\pm$ SD)	<i>p</i> (D:N)	Chemical class
L-Lactic acid	9.90	1.45 $\pm$ 0.07	0.027	1.73 $\pm$ 0.06	0.007	1.62 $\pm$ 0.19	0.001	1.57 $\pm$ 0.09	0.01	Hydroxy acid
Succinic acid	18.84	1.51 $\pm$ 0.12	0.007	1.53 $\pm$ 0.05	0.004	1.48 $\pm$ 0.08	0.004	1.35 $\pm$ 0.06	0.005	Carboxylic Acid
Glutamic acid	28.66	1.72 $\pm$ 0.03	0.005	1.49 $\pm$ 0.10	0.002	1.60 $\pm$ 0.05	0.032	1.28 $\pm$ 0.13	0.001	Amino acid
L-Valine	15.56	1.27 $\pm$ 0.12	0.009	1.31 $\pm$ 0.08	0.012	1.47 $\pm$ 0.08	0.003	1.23 $\pm$ 0.07	0.009	Amino acid
L-Alanine	11.54	1.43 $\pm$ 0.06	0.005	1.22 $\pm$ 0.11	0.01	1.29 $\pm$ 0.01	0.02	1.19 $\pm$ 0.05	0.01	Amino acid
Glycine	14.25	1.33 $\pm$ 0.11	0.004	1.42 $\pm$ 0.09	0.001	1.36 $\pm$ 0.07	0.005	1.30 $\pm$ 0.10	0.004	Amino acid
L-Proline	18.46	1.40 $\pm$ 0.13	0.03	1.35 $\pm$ 0.01	0.005	1.17 $\pm$ 0.10	0.002	1.22 $\pm$ 0.02	0.004	Amino acid
Palmitic acid	39.78	0.85 $\pm$ 0.08	0.001	0.51 $\pm$ 0.06	0.003	0.59 $\pm$ 0.09	0.045	0.92 $\pm$ 0.07	0.03	Fatty acid
Stearic acid	42.63	0.47 $\pm$ 0.09	0.008	0.61 $\pm$ 0.04	0.006	0.73 $\pm$ 0.07	0.005	0.72 $\pm$ 0.05	0.011	Fatty acid
Glucose-6-phosphate	43.95	1.59 $\pm$ 0.04	0.006	1.41 $\pm$ 0.08	0.04	1.52 $\pm$ 0.05	0.008	1.39 $\pm$ 0.14	0.001	Carbohydrates

A: N is the ratio of different metabolite levels in HepG2 cells transfected by HBx-A to the control; B:N is HBx-B to the control; C:N is HBx-C to the control; D:N is HBx-D to control. Control is metabolites extracts from HepG2 cells transfected by empty vector pXJ40. The standard deviation (SD) was generated from three independent experiments. The p-values of metabolite changes were $p < 0.05$.

Figure 4-5 Differential expression levels of intracellular metabolites in HBx transfected HepG2 cells.

Three independent experiments were carried out. Columns from each group represent pXJ40-HepG2 (control), pXJ40-HBx A, pXJ40-HBx B, pXJ40-HBx C and pXJ40-HBx D respectively. *, $p < 0.05$; **, $p < 0.01$.

4.3.4 Up-regulation of glucose-6-phosphate, lactic acid and succinic acid by HBx

Our results suggest that the level of lactic acid, an end product of glycolysis, and succinic acid were significantly up-regulated in HBx transfected HepG2 cells.

Glucose is the main carbon source for most tumor cells. The rate of glucose uptake is increased considerably in tumor cells in order to provide the required energy to support rapid tumor progression [154]. According to Warburg effect, most cancer cells are accelerating the rate of glycolysis and lactate production in the presence of oxygen [70, 165]. Glycolysis could have been up-regulated by HBx in an attempt to harness extra energy to meet its metabolic requirement for replication. This could be achieved through interaction with other cellular proteins such as hypoxia-inducible factor (HIF) and c-Myc (a known oncogene in many human cancers). Overexpression of HIF and c-Myc by HBx causes a shift from oxidative phosphorylation to glycolysis through up-regulation of glucose transporters and multiple enzymes of the glycolytic pathway [160]. Thus, pyruvate produced by glycolysis is converted to lactate instead of entering to TCA cycle (figure 4.6) [68]. Therefore, up-regulation of glucose-6-phosphate and lactic acid indicates the increasing rate of glycolysis by HBx.

On the other hand, accumulation of succinic acid, the intermediate products of the TCA cycle, reveals that HBx increased the rate of glycolysis but inhibited the TCA cycle (This is consistent with our proteomics results in chapter 3).

Our results have shown that HBx can shift the metabolic pathways of HepG2 cells from oxidative phosphorylation to glycolysis to provide important metabolic requirements for nucleotides, amino acids and lipids synthesis. Although replication

might appear to be an energy intensive event, it also requires other metabolic requirements such as ribose, amino acids and acetyl CoA that extend beyond ATP [67]. Hence, switching to glycolysis might prove to be a neat solution to this problem. HBV appears to have adopted this strategy via HBx, giving itself a selective advantage.

Enhanced glycolysis induces lactic acid accumulation which in turn favours tumour metastasis [166]. Lactic acid is a known substrate for tumorous cells requiring energy for growth, therefore measuring the level of lactic acid may prove to be useful for a novel metabolic classification. This could profoundly enhance methods used for diagnosis and prognosis of HBV-related HCC.

4.3.5 Up-regulation of amino acids by HBx

We observed that glycine, valine, alanine, proline and glutamate were up-regulated in HepG2 cells transfected with HBx compared to the control.

Glycine is the smallest amino acid which plays a key role in rapid cancer cell proliferation. It is synthesized from glycolytic intermediate (3-phosphoglycerate). Several studies have demonstrated that there is a link between glycine metabolism and tumorigenesis [167, 168]. Therefore, up-regulation of glycine by HBx may lead to hepatocellular carcinogenesis.

Valine and alanine are synthesized from pyruvic acid. The amino acid alanine is another important end-product of glucose utilization which is up-regulated in most cancer cells [169]. Valine has been a candidate for study in previous research due to specific catabolism properties in HCC that produce the toxic and mutagenic compound

Methacrylyl-coenzyme A (MC-CoA) [125]. Our data indicated that up-regulation of valine by HBx leads to accumulation of MC-CoA in the cells that causes cell mutation which support carcinogenesis.

Glutamate acts as nitrogen donor for the synthesis of many nonessential amino acids. It can be used to support the production of NADPH or converted to the metabolic intermediates pyruvate and α -ketoglutarate. Metastatic properties of cancer cells like migration, proliferation and invasion are particularly influenced by glutamate [170]. Since the TCA cycle generates crucial metabolites essential for synthesis of lipids, non-essential amino acids and nucleotides, different mechanism are necessary for an actively dividing cell to refill this cycle. Many tumor cells depend on glutamate anaplerosis for replenishment of the TCA cycle [171, 172]. Glutamate is converted to α -ketoglutarate upon entrance into the TCA cycle. The studies demonstrated that HIF/c-Myc enhanced expression of the mitochondrial enzyme glutaminase (GLS) which catalyzes the conversion of glutamine to glutamate (figure 4.6) [173, 174]. On the other hand biosynthesis of proline is significantly increased by Myc [175].

Our study suggests that HBx induces both aerobic glycolysis and glutaminolysis to provide ATP and important metabolic intermediates for supporting various biosynthetic processes critical for viral replication.

Figure 4-6 HBx induces aerobic glycolysis and glutaminolysis, leading to hepatocellular carcinogenesis.

Overexpression of HIF and c-Myc by HBx causes a shift from oxidative phosphorylation to glycolysis. Thus; pyruvate produced by glycolysis is converted to lactate instead of entering to TCA cycle. Therefore, up-regulation of glucose-6-phosphate and lactic acid indicates the increasing rate of glycolysis by HBx. On the other hand, accumulation of succinic acid, the intermediate products of the TCA cycle, reveals that HBx increased the rate of glycolysis but inhibited the TCA cycle. Moreover, up-regulation of glutamate indicates that the glutaminolysis is increased by HBx. Overexpression of HIF/c-Myc enhanced expression of GLS which catalyzes the conversion of glutamine to glutamate. LDH: lactate dehydrogenase; PDH: pyruvate dehydrogenase; HIF: hypoxia-inducible factor; GLS: glutaminase; Succ: succinate; α-KG: alpha-ketoglutarate; G-6-P: glucose-6-phosphate; 3-PG: 3-phosphoglycerate. Induction is illustrated by (---->) and repression is illustrated by (—|).

4.3.6 Down-regulation of palmitic acid and stearic acid by HBx

Our findings indicate that the levels of palmitic acid and stearic acid are down-regulated in HepG2 cells transfected with different genotypes of HBx. Palmitic acid and stearic acid are common saturated fatty acids in human body. They have a role in regulation of membrane rigidity and inducing apoptosis [176, 177]. Down-regulation of these fatty acids leads to decrease in membrane rigidity as well as apoptosis which, in turn, induce carcinogenesis. Several studies demonstrated that the level of palmitic acid and stearic acid are down-regulated in hepatocellular carcinoma [127, 178, 179].

Reduction in the level of palmitic acid and stearic acid may be due to enhancement of fatty acid β -oxidation. HBx may elevate the oxidation of fatty acids in order to support cell survival and energy production [180].

Altered lipid metabolism by HBx could lead to pathogenesis of HCC. Therefore, measuring the level of palmitic acid and stearic acid may lead to new therapeutic strategies for managing HBV-related HCC patients.

The role of these selected metabolites in hepatocellular carcinogenesis is summarized in figure 4.7.

Figure 4-7 Schematic representation of the role of selected differentially expressed metabolites in hepatocellular carcinogenesis.

Metabolites in red/blue display up-regulation/down-regulation in HBx transfected HepG2 cells. Lactic acid, succinic acid, glucose-6-phosphate, glutamate, valine, glycine, alanine and proline were significantly up-regulated while palmitic acid and stearic acid were down-regulated in HBx transfected HepG2 cells. These metabolites are involved in several key metabolic pathways such as glycolysis, glutaminolysis, TCA cycle, amino acids and fatty acids metabolism.

4.4 Conclusions

In this study, we applied GC-MS metabolomic profiling of HepG2 cells transfected with different genotypes of HBx to detect novel biomarkers for the diagnosis, prognosis and treatment of HCC.

Our metabolomics analysis provides insights into the mechanism of HBV replication and metabolic changes in host cell to better understand the role of HBV replication in HCC development. 71 metabolites from various categories were totally identified. In our metabolic analysis the changes in the level of 10 intracellular metabolites were observed in HBx transfected HepG2 cells by comparing them with control group.

These metabolites include lactic acid, succinic acid, glucose-6-phosphate, glutamate, valine, glycine, alanine, proline, palmitic acid and stearic acid which are involved in several key metabolic pathways such as glycolysis, glutaminolysis, TCA cycle, amino acids and fatty acids metabolism. Studying the HBx impact on metabolites variation may give us high potential for finding effective biomarkers related with HCC development, which would contribute to new prognosis strategies for managing HBV-related HCC patients as well as contribute to the clinical therapy of HCC. In addition, it may also give us opportunities to study the significant activity of HBx and understand its role in hepatocarcinogenesis. Our findings suggested that these metabolic alternations are closely related with the development of hepatocellular carcinogenesis associated with HBV infection. Taken together, our metabolomic profiling approach may provide differentially expressed metabolites potential as biomarkers in order to improve therapeutic strategies for better managing HBV-related HCC patients.

Chapter 5 : Protein Profile of Mutant HBx Transfected Cells: Investigation of SH3 Binding Domain in Virus-Host Interactions

5.1 Introduction

Protein-protein interaction is an important aspect in understanding cellular signal transduction processes. SH3 domain, important in protein interactions, plays a critical role in signal transduction. SH3 domain is small conserved sequence domain which contains 55-70 amino acids [42, 43]. It can interact specifically with proline-rich motifs known as SH3-binding domains. HBx contains a proline/serine motif which appears to interact with a number of SH3-binding cellular proteins [181, 182].

In this study, HBx genotypes A/B/C were cloned into mammalian expression vector pXJ40, and their expressions were confirmed by Western blot analysis. Figure 5.1 shows the alignment of HBx A/B/C/D. As shown in this figure, the proline/serine rich region of HBx is comparatively different among the genotypes.

We aimed to generate point mutations in four proline residues of the proline rich region of HBx (amino acids 28-46) due to the high variability of this region among different genotypes of HBx. Four proline residues in the proline rich region were mutated to alanine in order to investigate the protein profile of mutant HBx by applying iTRAQ-coupled 2D LC-MS/MS-based quantitative proteomics. Comparing the protein profile of wild type HBx and P-A mutant HBx, may provide valuable information on the role of SH3 binding domain in virus-host interactions.

Figure 5-1 Alignments of HBx amino acids in genotypes A/B/C/D.

Most of the HBx sequences were conserved among different genotypes. Considerable variation between 28-46aa was observed.

5.2 Experimental of procedures

5.2.1 Mutation and Cloning of HBx Genotypes (converting proline to alanine)

The plasmid constructed used in this research was mutated HBx for each genotype with proline being converted to alanine. Mutations in these genotypes (as shown in figure 5.2) were created by polymerase chain reaction (PCR) method followed by cloning of the generated PCR products into pXJ40 vector. The cloned fragments were confirmed by sequencing.

Genotype A	RPVGAESRGR	P	LAG	P	L	G	T	L	S	S	P	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G		
Genotype B	RPVGAESRGR	P	L	P	G	P	L	G	A	L	P	P	A	S	P	P	I	V	P	T	D	H	G	A	H	L	S	L	R	G
Genotype C	RPVGAESRGR	P	V	S	G	P	F	G	T	L	P	S	P	S	S	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G
Genotype D	RPVGAESRGR	P	F	S	G	P	L	G	T	L	S	S	P	S	P	S	A	V	S	S	D	H	G	A	H	L	S	L	R	G

- Prolines in red mutated to alanines

Figure 5-2 Targeted Mutation in HBx Genotypes.

Proline in wild type genotypes has been changed to alanine [22].

5.2.2 PCR Amplification

In order to generate the mutated constructs, two fragments (F1 and F2) for each genotype were generated. After amplifying the two fragments, F1 and F2 were combined by PCR to make the complete mutated HBx genotype. PCRs were carried out with Pfu DNA polymerase using the primers described in table 5.1 and 5.2, for F1 and F2 respectively. Wild type HBx was used as template.

Table 5-1 F1 Primers and their characteristics

Gene	Direction	Sequence (5'→3')
HBx-A	Forward	GAATCCCGCGGACGACCCCTCGCGGGGCGCTTGGGACTCTCTCG
	Reverse	CGAGAGAGTCCCAAGCGGCCCCGCGAGGGGTCGTCCGCGGGATTC
HBx-B	Forward	GAATCCCGCGGACGACCCCTCCCGGGGCGCTTGGGGCTCTACCG
	Reverse	CGGTAGAGCCCCAAGCGGCCCCGGGAGGGGTCGTCCGCGGGATTC
HBx-C	Forward	GAATCCCGCGGACGACCCGTCTCGGGGCGCTTGGGACTCTA
	Reverse	TAGAGTCCCAAACGGCCCCGAGACGGGTCGGCGGGATTC
HBx-D	Forward	GAATCCCGCGGACGAGCCTTCTCGGGGCGCTTGGGACTCTCTCG
	Reverse	CGAGAGAGTCCCAAGCGCCCCCGAGAAGGCTCGTCCGCGGGATTC

Table 5-2 F2 Primers and their characteristics

Gene	Direction	Sequence (5'→3')
HBx-A	Forward	GGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCAGCC
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC
HBx-B	Forward	GGGGCTCTACCGCCCCTTCTCCGCCTATTGTACCGACC
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC
HBx-C	Forward	TTT GGGACTCTACCGTCCCCTTCTTCGTCTGCCGTTCCGGCC
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC
HBx-D	Forward	GGGACTCTCTCGTCCGCTTCTGCGTCTGCCGTTTCGTCC
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC

The gel electrophoresis of PCR products under UV light, after running on 1.5% agarose gel, is illustrated in figure 5.3 (a).

After performing PCR, the products were separated by using 1.5% agarose gel electrophoresis. Desired DNA fragments were visualized by ethidium bromide (Bio-Rad) under UV light and then were cut out from the gel followed by purification using QIAquick gel extraction kit. Subsequently, two fragments, i.e., F1 and F2, were attached together by PCR to generate the mutant HBx construct (figure 5.3 (b)). Primers and their characteristics are shown in table 5.3

Figure 5-3 Gel electrophoresis of PCR products.

(a) Gel picture of F1 and F2 after running over 1.5% agarose gel. **(b)** HBx mutated full length (F1 + F2) PCR product.

Table 5-3 Primers and theirs characteristics

Gene	Direction	Sequence (5'→3')
HBx-A	Forward	GAATCCCGCGGACGACCCCTCGCGGGGCCGCTTGGGACTCTCTCG
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC
HBx-B	Forward	GAATCCCGCGGACGACCCCTCCCGGGGCCGCTTGGGGCTCTACCG
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC
HBx-C	Forward	GAATCCCGCGGACGACCCGTCTCGGGGCCGTTTGGGACTCTA
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC
HBx-D	Forward	GAATCCCGCGGACGAGCCTTCTCGGGGGCGCTTGGGACTCTCTCG
	Reverse	GGACGAAACGGCAGACGCAGAAGCGGACGAGAGAGTCCC

5.2.3 Cell Lysis, Protein Digestion, and Labeling with iTRAQ Reagents

DNA digestions were performed with restriction enzymes (REs) purchased from NEB (New England Biolabs) in a volume of 50 µl adjusted by sterile deionized water. In order to cut the pXJ40 vector as well as inserts, Hind III and Pst I were used. After restriction digestion, the purified DNA fragments were used for ligation. T4 DNA ligase was used for ligation purpose. In order to propagate the plasmids prepared by ligation, 10 µl of ligation mixture (10-100ng) was added to just-thawed 20-25 µl of competent DH5α bacteria (previously prepared in the lab), mixed and incubated for 30 min on ice. Heat shock (45 sec at 42°C followed by 2-3 min cooling down on ice) was carried out on the bacteria-vector mixture. Preparation of the vectors from transformed *E.coli* cell pellets were carried out via QIAGEN miniprep kit. Sequencing was performed by National Neuroscience Institute (NNI) Core Facility, Tan Tock Seng

Hospital (TTSH). DNA from clones verified by sequencing was purified using Qiagen endo-free maxi kit and used in subsequent molecular biology or cell culture work. After that plasmid DNAs purified by miniprep or maxiprep were quantified by UV spectrophotometer. In the next step, HepG2 cells were transfected with different genotypes of mutant HBx using Nucleofection solution V (Amata GmBH). Two days after transfection, the transfected cells were harvested by trypsinization. The cells were then lysed in 200 µl lysis buffer on ice for 20 min with regular vortexing. Centrifugation at 15,000g for 10min at 4°C was performed. Supernatant containing the proteins were then collected and the proteins were quantified using the 2-D Quant Kit. Finally samples were labelled with iTRAQ reagent then mixed and analyzed by LC-MS/MS.

5.3 Results and discussion

5.3.1 Cloning of mutant HBx and transfection efficiency

Fragments of HBx A/B/C were amplified by PCR and ligated into pXJ40 vector by HindIII and PstI restriction enzyme sites. The results shown in figure 5.4 (a) indicated that all three genotypes of mutant HBx were successfully cloned into pXJ40 vector. Transfections of HepG2 cells were performed using Nucleofection solution V (Amata, USA), which provides high transfection efficiency. Figure 5.4 (b) shows the results of HepG2 cells 24h after transfection with Green fluorescent protein (GFP).

Figure 5-4 Cloning of different genotypes of mutant HBx in pXJ40 vector and transfection efficiency.

(a) Cloning of different genotypes of mutant HBx in pXJ40. Electrophoresis of agarose gel is shown. DNA samples included HBx A/B/C cloned in pXJ40 vector. Lane M is the 1kb DNA marker (Gene Ruler, #SM0311). Lanes 1, 2, 3 represented mHBx –A, mHBx –B and mHBx –C in pXJ40 respectively. **(b)** HepG2 cells after transfection with GFP using Amaxa Nucleofector. GFP was used for tracing the efficiency of transfection. Left is phase contrast image of HepG2 cells 24 hours after transfection. Right is the corresponding image of the green fluorescent cells by GFP in the same field (Olympus fluorescent microscopy).

5.3.2 The expression of mutant HBx in transfected cells

The expression of different genotypes of mutant HBx were confirmed by western blot using anti-HA antibody (figure 5.5). β -actin was used as internal control. The result showed that pXJ40-mHBx of three genotypes (A, B and C) were effectively expressed in HepG2 cells.

Figure 5-5 Western-blot analysis of mutant HBx genotype A, B and C in transfected cells.

Lane1 was the empty vector pXJ40; Lane 2: mHBx A - pXJ40; Lane3: mHBx B - pXJ40; Lane4: mHBx C - pXJ40. The Tag antibodies used here were anti-HA (Santa Cruz). β - actin was the internal control.

5.3.3 Differentially expressed proteins identified by iTRAQ-coupled 2D LC-MS/MS

In order to establish the protein profile in the HepG2 cells transfected with P-A mutant HBx compared with empty vector pXJ40, the iTRAQ was used to label as follows:

HepG2 transfected with empty vector pXJ40 is labelled with iTRAQ 114.

HepG2 transfected with mutant HBx -A (pXJ40-mHBx A) is labelled with iTRAQ 115.

HepG2 transfected with mutant HBx -B (pXJ40-mHBx B) is labelled with iTRAQ 116.

HepG2 transfected with mutant HBx -C (pXJ40-mHBx C) is labelled with iTRAQ 117.

A total of 146 proteins were identified in three independent experiments. Identification and quantification of proteins were based on ProtScore more than 2.0 that gives the

confidence value of 99%. The proteins identified with ProtScore less than 2.0 were not considered for further analysis. The high confidence > 95% was identified for two or more unique peptides and p value in the Protein Quant had to be $P < 0.05$.

Our LC-MS results demonstrated that cytoskeletal proteins including vimentin and annexin A2 were significantly down-regulated in HepG2 cells transfected with different genotypes of mutant HBx compared to the pXJ40 control group as shown in table 5.4.

Table 5-4 Differentially expressed proteins in HepG2 cells transfected by mutant HBx.

Protein name	Avg. A:N	<i>p</i> (A:N)	Avg. B:N	<i>p</i> (B:N)	Avg. C:N	<i>p</i> (C:N)
Vimentin	0.81 ± 0.02	0.001	0.78 ± 0.01	0.008	0.65 ± 0.01	0.004
Annexin A2	0.77 ± 0.05	0.003	0.69 ± 0.06	0.005	0.73 ± 0.04	0.001

Data were reported as means ±SD (standard deviation) from at least three independent experiments. A: N is the ratio of different protein expression levels in HepG2 cells transfected by mutant HBx-A to the control; B:N is mutant HBx-B to the control; C:N is mutant HBx-C to the control. Control is protein extracts from HepG2 cells transfected by empty vector pXJ40. *p*-values <0.05 were considered to be statistically significant.

5.3.4 Comparison of the protein profile of wild type and mutant HBx

As shown in figure 5.6, the difference in protein profile of wild type and P-A mutant HBx was mainly observed in cytoskeleton and signalling proteins.

Our results demonstrate that vimentin and annexin A2 are differentially expressed between wild type and mutant HBx. It may suggest that the influence of mutant HBx on vimentin/ annexin A2 is weaker compared to wild type HBx.

Annexin A2 and vimentin were up-regulated by wild type HBx in our previous work (table 3.2, pages 93-94). Studies have shown that HBx is involved in the cytoskeletal reorganization in response to HBV replication [44, 61]. HBx causes up-regulation in the level of cytoskeletal proteins which in turn may indicate an increase in cell migration. Therefore, cytoskeletal reorganization by HBx may effect on cell adhesion and migration which leads to HCC development. Overexpression of vimentin and annexin A2 by HBx may cause cytoskeletal reorganization which can alter the cell migration and adhesion [33, 44].

Conversely, in P-A mutant HBx the expression of annexin A2 and vimentin were down-regulated significantly. It has been suggested that the proline rich region of HBx may be responsible for its interaction with annexin A2 and vimentin. It should be noted that the influence of HBx on expression levels of these genes may be indirected.

The down-regulation of annexin A2 and vimentin by mutant HBx may induce a decrease in cell migration and enhancement of cell adhesion.

Figure 5-6 Difference in protein profile of wild type and P-A mutant HBx.

(a) The differential expression level of proteins in HepG2 cells transfected by wild type HBx.
 (b) The differential expression level of proteins in HepG2 cells transfected by mutant HBx.
 Three independent experiments were carried out. Asterisk indicates $p < 0.05$. Annexin A2 and vimentin were up-regulated by wild type HBx, while down-regulated by mutant HBx.

5.3.5 Validation of differentially expressed proteins by western blot

HepG2 cells transfected with different genotypes of wild type and mutant HBx were harvested 24 hours after transfection and proteins were extracted by adding lysis buffer and centrifugation. The changes of proteins levels obtained from LC-MS/MS analysis were also confirmed by Western blot analysis. Figure 5.7 shows the expression level of annexin A2 and vimentin between wild type and mutant HBx. In P-A mutant HBx the expression level of annexin A2 and vimentin were significantly down-regulated compared with wild type HBx.

Western bands were analyzed and quantified by ImageJ software as shown in figure 5.8. The Western blot analysis of these differentially expressed proteins is consistent with the results of LC-MS/MS.

The antibodies used in this experiment and the dilution ratios are shown in table 5.5.

Table 5-5 Antibodies and their dilution ratios

Antibody	Dilution ratios
Primary antibody mouse anti- β -actin (AC-74, Sigma)	1:10,000
Primary antibody mouse anti-HA (F-7: sc-7392, Santa Cruz)	1:1000
Primary antibody mouse anti-Vimentin (RV202: sc-32322, Santa Cruz)	1:500
Primary antibody mouse anti-Annexin A2 (C-10: sc-28385, Santa Cruz)	1:1000
Secondary antibody, HRP-conjugated goat anti-mouse IgG (#31430, Pierce)	1:5000

Figure 5-7 Western blot confirmations of the differentially expressed proteins in HepG2 cells transfected with mutant HBx A/B/C.

(a) the expression level of annexin A2 and vimentin in empty vector pXj40, HBx A (wild type), HBx A_m (mutant). **(b)** The expression level of annexin A2 and vimentin in empty vector pXj40, HBx B (wild type), HBx B_m (mutant). **(c)** The expression level of annexin A2 and vimentin in empty vector pXj40, HBx C (wild type) and HBx C_m (mutant).

Figure 5-8 Fold changes of protein level by western blot were analyzed using Image J software.

(a) the expression level of annexin A2 and vimentin in empty vector pXj40, HBx A (wild type) and HBx Am (mutant). (b) The expression level of annexin A2 and vimentin in empty vector pXj40, HBx B (wild type) and HBx Bm (mutant). (c) The expression level of annexin A2 and vimentin in empty vector pXj40, HBx C (wild type) and HBx Cm (mutant).

5.4 Conclusions

SH3 domains have been found in many signalling and cytoskeleton proteins. Some viral proteins such as HBx possess proline-rich SH3 binding motifs. They can interact with host SH3 domains containing cellular proteins via their proline-rich domains. By converting the proline residues to alanine in mutant HBx the bindings is supposed to have been abolished.

Our proteomic profiling demonstrated that the expression level of annexin A2 and vimentin were significantly down-regulated in P-A mutant HBx compared to wild type HBx. Therefore, the proline-rich domain of HBx may be responsible for its influence on SH3 domain-containing cellular proteins such as annexin A2 and vimentin. Down-regulation of annexin A2 and vimentin by mutant HBx decrease cell migration and improves cell adhesion. Our results suggest that HBx may influence cytoskeletal reorganization and consequently cell adhesion and migration which leads to HCC development.

Altogether, LC-MS/MS proteomics analysis of mutant HBx transfected cells sheds new lights on the understanding of the role of SH3 binding domain in virus-host interactions as well as HCC development.

Chapter 6 : Conclusions and Future Direction

HCC is the fifth most prevalent cancer worldwide. On the other hand; it has the third highest mortality rate among all cancers. Chronic HBV infection is a major cause of HCC development. The morbidity rate of HCC is mainly due to late diagnosis; hence, early detection of HCC is crucial for improving the chances of recovery.

In this study, the association between HBV replication and the host cell metabolism was investigated in order to better understand the role of HBV replication in HCC development. The smallest protein of HBV, HBx, has vital role in hepatocellular carcinogenesis. It is important to know how HBx interacts with host cell proteins and causes deregulation of normal metabolic pathways leading to hepatocellular carcinogenesis.

We applied omics technologies such as proteomics and metabolomics to detect potential biomarkers for early diagnosis of HBV-related HCC patients. LC-MS/MS based proteomics and GC-MS based metabolomics are effective tools for screening biomarkers useful for both disease prognosis and diagnosis.

Our cell-based proteomics/metabolomics platform was explored to identify differentially expressed cellular proteins/metabolites which are involved in virus-host metabolic pathways as presented in figure 6.1.

Figure 6-1 Proposed protocol workflow of iTRAQ-coupled LC-MS/MS and GC-MS platform for Identification the differentially expressed cellular proteins and metabolites in HBx transfected HepG2 cells.

In this work, different genotypes of HBx (A, B, C, D) were amplified and cloned into mammalian expression vector pXJ40. Their successful expressions were confirmed in RPHs and HepG2 cells. Total proteins were extracted and analyzed by iTRAQ-coupled 2D LC-MS/MS.

Our proteomics results indicated that a number of glycolytic enzymes including HK, GAPDH, PK, PGK and LDH were significantly up-regulated while two TCA cycle enzymes, SDH and FH were down-regulated in both RPHs and HepG2 cells transfected by HBx compared to control group.

These findings suggest that HBx can shift the metabolic pathways of cells from oxidative phosphorylation to glycolysis in order to provide important metabolic requirements for nucleotides, amino acids and lipids synthesis as illustrated in figure 6.2.

Therefore, our cell based assay may provide valuable information on the principles of HBV replication in host cell metabolism to understand the early cellular response to HBV infection. Differentially expressed proteins identified could be further used to define how HBV and HBX can contribute to hepatocarcinogenesis. In addition, the rat primary hepatocytes will provide crucial validation of the potential biomarkers identified in this project.

Identification of cellular metabolites by comparative metabolomics may complement the proteomic analysis for improving HBV-related HCC therapy. Metabolic factors play an important role in the pathogenesis of HCC. GC-MS based metabolomics was

applied to establish a comprehensive cellular metabolomic profiling of HBx transfected HepG2 cells.

A total of 71 metabolites were identified out of which 10 showed significant changes in compared to control group. Lactic acid, succinic acid, glucose-6-phosphate, glutamate, valine, glycine, alanine and proline were significantly up-regulated while palmitic acid and stearic acid were down-regulated in HBx transfected HepG2 cells.

These findings suggested that HBx can alter the levels of several metabolites incorporated in glycolysis, TCA cycle and fatty acids or amino acids metabolism. Hence, HBV replication could alter host cell metabolism by increasing the rate of glycolysis and inhibiting the TCA cycle (figure 6.2).

Such metabolites could act as potential biomarkers for improvement of HBV-related HCC patients. Our comprehensive metabolomic profiling of HBx transfected cells may provide new lights in understanding of early events of HCC development associated with HBV infection.

Figure 6-2 Differentially expressed cellular proteins and metabolites indicate that HBx may shift the metabolic pattern of cells from oxidative phosphorylation to glycolysis which leads to hepatocellular carcinogenesis.

Abbreviations: LDH: lactate dehydrogenase; SDH: succinate dehydrogenase; FH: fumarate hydratase; HK: hexokinase; GAPDH: glyceraldehyde-3-phosphate dehydrogenase; PK: pyruvate kinase; PGK: phosphoglycerate kinase; G-6-P: glucose-6-phosphate. Up-regulation/down-regulation are illustrated by ↑ / ↓ respectively.

Our laboratory has recently reported that HBx possess proline rich region located in amino acids 28-46. This region has high variability among different genotypes of HBx and it can interact with cellular proteins containing SH3 domains. Hence, the proline rich region of HBx plays critical role in virus-host cell interactions. In this project, in order to investigate the cellular proteins which interact with HBx through SH3 domains, four proline residues at proline rich region converted to alanine.

2D coupled LC-MS/MS proteomics analysis was applied to provide a comprehensive protein profile of P- A mutant HBx. Comparing the protein profile of wild type and mutant HBx demonstrated that a series of differentially expressed cellular cytoskeletal and signaling proteins including vimentin and annexin A2 may be affected by HBx via its proline rich domain. These differentially expressed cellular proteins play significant role in cytoskeleton organization and cell migration. Cytoskeletal reorganization by HBx may affect cell adhesion and migration which leads to hepatocellular carcinogenesis.

Altogether, proteomic profiling of mutant HBx transfected cells should shed new lights on understanding the role of SH3 binding domain in virus-host interactions as well as HCC development.

To gain a more comprehensive and further understanding of the mechanisms of HBV/HBx in HCC development, we propose the following experimental plan for the future.

Besides SH3 domain, WW domain is also involved in virus-host interactions. For example, Pin1 can interact with HBx through its WW domain. Hence, association of

HBx with WW domain containing proteins can be analyzed in future for illustrating overall network of virus-host interactions as well as provide new insights on cellular metabolism and biomarker discovery for future improvement of HCC therapy.

In addition to cell-based systems, our project can be further expanded into proteomics and metabolomics analysis for clinical samples.

Cell-based platform can be combined with clinical analysis for validation of candidate biomarkers in order to improve therapeutic strategies for better managing HBV-related HCC patients.

Blood plasma is considered a significant proteome in clinical studies. It can be easily extracted from blood samples of the patients and used for identification of biomarkers. In the same way, cancerous liver tissue as well as adjacent normal tissue could be used for biomarker identification in cancer metastasis, recurrence and poor prognosis.

Although using blood plasma to identify and analyse biomarkers is convenient due to the ease of sample collection, it suffers some limitation. A limiting factor of these methods is that, while biomarkers identified in clinical trials are valuable in clinical level they need to be analysed in vitro or validated by a large number of clinical samples. In addition, biomarkers identified through clinical samples are mostly useful in therapeutic monitoring as they are highly correlated with the late stages of HCC.

To diagnose the disease at the early stages, a cell-based system can be applied. This method is less expensive and yields more homogenous genetic background and hence results in replicable data.

In conclusion, our research demonstrated the feasibility of using proteomics and metabolomics cell based platform to understand the principal mechanism of HBV-HBx-host cell interplays.

Taken together, our study has established a comprehensive cell-based system by applying LC-MS/MS based quantitative proteomics as well as GC-MS based metabolomics in order to identify intracellular proteins / metabolites potential biomarkers for early diagnosis of HBV-related HCC patients.

A combined approach of proteomics/metabolomics and molecular biology should lead to an improved strategy against HBV-associated HCC with better prognosis and effective therapy.

It is hoped that this platform provides novel clues on the underlying mechanism of HBV replication and HBx involvement in hepatocellular carcinogenesis.

References

1. Schutte, K., J. Bornschein, and P. Malfertheiner, *Hepatocellular carcinoma--epidemiological trends and risk factors*. Dig. Dis., 2009. 27(2): p. 80-92.
2. Venook, A.P., et al., The incidence and epidemiology of hepatocellular carcinoma: a global and regional perspective. Oncologist, 2010. 4: p. 5-13.
3. Thomas, M.B. and A.X. Zhu, *Hepatocellular carcinoma: the need for progress*. J. Clin. Oncol., 2005. 23(13): p. 2892-9.
4. Zhu, A.X., Systemic therapy of advanced hepatocellular carcinoma: how hopeful should we be? Oncologist, 2006. 11(7): p. 790-800.
5. Motola-Kuba, D., et al., *Hepatocellular carcinoma. An overview*. Ann Hepatol, 2006. 5(1): p. 16-24.
6. Feng, H., et al., Proteomics-related biomarkers for HBV-associated hepatocellular carcinoma: current status and future prospects. Future Virology, 2012. 7(2): p. 161-171.
7. Coleman, W.B., *Mechanisms of human hepatocarcinogenesis*. Curr Mol Med, 2003. 3(6): p. 573-88.
8. Gomaa, A.I., et al., *Hepatocellular carcinoma: epidemiology, risk factors and pathogenesis*. World J Gastroenterol, 2008. 14(27): p. 4300-8.
9. Arbuthnot, P. and M. Kew, *Hepatitis B virus and hepatocellular carcinoma*. Int. J. Exp. Pathol., 2001. 82(2): p. 77-100.
10. Shariff, M.I., et al., Hepatocellular carcinoma: current trends in worldwide epidemiology, risk factors, diagnosis and therapeutics. Expert Rev Gastroenterol Hepatol, 2009. 3(4): p. 353-67.
11. Lemoine, M., S. Nayagam, and M. Thursz, Viral hepatitis in resource-limited countries and access to antiviral therapies: current and future challenges. Future Virol, 2013. 8(4): p. 371-380.
12. Perz, J.F., et al., The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. J Hepatol, 2006. 45(4): p. 529-38.
13. Elgouhari, H.M., T.I. Abu-Rajab Tamimi, and W.D. Carey, *Hepatitis B virus infection: understanding its epidemiology, course, and diagnosis*. Cleve Clin J Med, 2008. 75(12): p. 881-9.
14. Kew, M.C., Epidemiology of chronic hepatitis B virus infection, hepatocellular carcinoma, and hepatitis B virus-induced hepatocellular carcinoma. Pathologie-biologie, 2010. 58(4): p. 273-7.

15. Lavanchy, D., Worldwide epidemiology of HBV infection, disease burden, and vaccine prevention. *J Clin Virol*, 2005. 34(1): p. S1-3.
16. Hou, J., Z. Liu, and F. Gu, *Epidemiology and Prevention of Hepatitis B Virus Infection*. International journal of medical sciences, 2005. 2(1): p. 50-57.
17. Seeger, C. and W.S. Mason, *Hepatitis B virus biology*. Microbiol. Mol. Biol. Rev., 2000. 64(1): p. 51-68.
18. Yokosuka, O. and M. Arai, Molecular biology of hepatitis B virus: effect of nucleotide substitutions on the clinical features of chronic hepatitis B. *Med Mol Morphol*, 2006. 39(3): p. 113-20.
19. Zhang, X.D., H. Zhang, and L.H. Ye, *Effects of hepatitis B virus X protein on the development of liver cancer*. Journal of Laboratory and Clinical Medicine, 2006. 147(2): p. 58-66.
20. Tang, H., et al., Molecular functions and biological roles of hepatitis B virus x protein. *Cancer Science*, 2006. 97(10): p. 977-983.
21. Rizzetto, M. and A. Ciancio, *Chronic HBV-related liver disease*. Molecular aspects of medicine, 2008. 29(1-2): p. 72-84.
22. Chen, W.N. and C.J. Oon, Human hepatitis B virus mutants: significance of molecular changes. *FEBS Lett.*, 1999. 453(3): p. 237-42.
23. Block, T.M., H. Guo, and J.T. Guo, *Molecular virology of hepatitis B virus for clinicians*. Clinics in liver disease, 2007. 11(4): p. 685-706, vii.
24. Beck, J. and M. Nassal, *Hepatitis B virus replication*. World journal of gastroenterology : WJG, 2007. 13(1): p. 48-64.
25. Neuveut, C., Y. Wei, and M.A. Buendia, *Mechanisms of HBV-related hepatocarcinogenesis*. *J Hepatol*, 2010. 52(4): p. 594-604.
26. Bartholomeusz, A. and S. Schaefer, *Hepatitis B virus genotypes: comparison of genotyping methods*. *Rev Med Virol*, 2004. 14(1): p. 3-16.
27. Schaefer, S., *Hepatitis B virus taxonomy and hepatitis B virus genotypes*. *World J Gastroenterol*, 2007. 13(1): p. 14-21.
28. McMahon, B.J., The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B. *Hepatology international*, 2009. 3(2): p. 334-42.
29. Chen, W.N., C.J. Oon, and G.K. Lim, Frequent occurrence of hepatitis B virus surface antigen mutants in subtype adw in vaccinated Singapore infants. *Vaccine*, 2001. 20(5-6): p. 639-40.
30. Chu, C.J., et al., Hepatitis B virus genotypes in the United States: results of a nationwide study. *Gastroenterology*, 2003. 125(2): p. 444-51.

31. Khawaja, R.A. and A.A. Khawaja, *Hepatitis B virus genotypes: "clinical & therapeutic implications"*. JPMA. The Journal of the Pakistan Medical Association, 2009. 59(2): p. 101-4.
32. Azam, F. and A. Koulaouzidis, *Hepatitis B virus and hepatocarcinogenesis*. Annals of hepatology, 2008. 7(2): p. 125-9.
33. Feng, H., et al., Protein profile in HBx transfected cells: a comparative iTRAQ-coupled 2D LC-MS/MS analysis. J. Proteomics, 2010. 73(8): p. 1421-32.
34. Bouchard, M.J. and R.J. Schneider, *The enigmatic X gene of hepatitis B virus*. Journal of virology, 2004. 78(23): p. 12725-34.
35. Sirma, H., et al., Hepatitis B virus X mutants, present in hepatocellular carcinoma tissue abrogate both the antiproliferative and transactivation effects of HBx. Oncogene, 1999. 18(34): p. 4848-59.
36. Cha, M.Y., et al., Stimulation of hepatitis B virus genome replication by HBx is linked to both nuclear and cytoplasmic HBx expression. The Journal of general virology, 2009. 90(Pt 4): p. 978-86.
37. Henkler, F., et al., *Intracellular localization of the hepatitis B virus HBx protein*. The Journal of general virology, 2001. 82(Pt 4): p. 871-82.
38. Xu, Z., et al., Enhancement of hepatitis B virus replication by its X protein in transgenic mice. Journal of virology, 2002. 76(5): p. 2579-84.
39. Reifenberg, K., et al., Virus replication and virion export in X-deficient hepatitis B virus transgenic mice. J. Gen. Virol., 2002. 83(Pt 5): p. 991-6.
40. Chen, W.N., et al., Expression of integrated hepatitis B virus X variants in human hepatocellular carcinomas and its significance. Biochem. Biophys Res. Commun., 2000. 276(3): p. 885-92.
41. Clippinger, A.J. and M.J. Bouchard, Hepatitis B virus HBx protein localizes to mitochondria in primary rat hepatocytes and modulates mitochondrial membrane potential. Journal of virology, 2008. 82(14): p. 6798-811.
42. Larson, S.M. and A.R. Davidson, The identification of conserved interactions within the SH3 domain by alignment of sequences and structures. Protein Sci, 2000. 9(11): p. 2170-80.
43. Kay, B.K., M.P. Williamson, and M. Sudol, *The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains*. FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2000. 14(2): p. 231-41.
44. Tan, T.L., et al., Adhesion contact kinetics of HepG2 cells during Hepatitis B virus replication: Involvement of SH3-binding motif in HBX. Biochimica et biophysica acta, 2006. 1762(8): p. 755-66.

45. Saksela, K., G. Cheng, and D. Baltimore, Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4. *Embo J*, 1995. 14(3): p. 484-91.
46. Datta, S., et al., Pin1-HBx interaction: a step toward understanding the significance of hepatitis B virus genotypes in hepatocarcinogenesis. *Gastroenterology*, 2007. 133(2): p. 727-8; author reply 728-9.
47. Lupberger, J. and E. Hildt, *Hepatitis B virus-induced oncogenesis*. *World J Gastroenterol*, 2007. 13(1): p. 74-81.
48. Bar-Yishay, I., Y. Shaul, and A. Shlomai, *Hepatocyte metabolic signalling pathways and regulation of hepatitis B virus expression*. *Liver International*, 2011. 31(3): p. 282-290.
49. Waris, G., K.W. Huh, and A. Siddiqui, Mitochondrially associated hepatitis B virus X protein constitutively activates transcription factors STAT-3 and NF-kappa B via oxidative stress. *Mol Cell Biol*, 2001. 21(22): p. 7721-30.
50. Park, Y.K., et al., Could HBx protein expression affect signal pathway inhibition by gefitinib or selumetinib, a MEK inhibitor, in hepatocellular carcinoma cell lines? *J Korean Med Sci*, 2011. 26(2): p. 214-21.
51. Truant, R., et al., Direct interaction of the hepatitis B virus HBx protein with p53 leads to inhibition by HBx of p53 response element-directed transactivation. *Journal of virology*, 1995. 69(3): p. 1851-9.
52. Lee, S.G. and H.M. Rho, Transcriptional repression of the human p53 gene by hepatitis B viral X protein. *Oncogene*, 2000. 19(3): p. 468-71.
53. Tan, T.L. and W.N. Chen, A proteomics analysis of cellular proteins associated with HBV genotype-specific HBX: potential in identification of early diagnostic markers for HCC. *J. Clin. Virol.*, 2005. 33(4): p. 293-8.
54. Knoll, S., et al., Dissection of cell context-dependent interactions between HBx and p53 family members in regulation of apoptosis: a role for HBV-induced HCC. *Cell cycle*, 2011. 10(20): p. 3554-65.
55. Dewantoro, O., R.A. Gani, and N. Akbar, *Hepatocarcinogenesis in viral Hepatitis B infection: the role of HBx and p53*. *Acta Med Indones*, 2006. 38(3): p. 154-9.
56. Zhang, X., et al., Hepatitis B virus X protein upregulates survivin expression in hepatoma tissues. *J Med Virol*, 2005. 77(3): p. 374-81.
57. Lee, Y.H. and Y. Yun, *HBx protein of hepatitis B virus activates Jak1-STAT signaling*. *J Biol Chem*, 1998. 273(39): p. 25510-5.
58. Tang, R.X., et al., HBx activates FasL and mediates HepG2 cell apoptosis through MLK3-MKK7-JNKs signal module. *World J Gastroenterol*, 2012. 18(13): p. 1485-95.

59. Park, N.H., I.H. Song, and Y.H. Chung, *Molecular Pathogenesis of Hepatitis-B-virus-associated Hepatocellular Carcinoma*. Gut and liver, 2007. 1(2): p. 101-17.
60. Chang, S.F., et al., Duck hepatitis B virus expresses a regulatory HBx-like protein from a hidden open reading frame. Journal of virology, 2001. 75(1): p. 161-70.
61. Feng, H., et al., HBX-mediated migration of HBV-replicating HepG2 cells: insights on development of hepatocellular carcinoma. J Biomed Biotechnol, 2009. 2009: p. 930268.
62. Zhao, J., et al., Association between metabolic abnormalities and HBV related hepatocellular carcinoma in Chinese: a cross-sectional study. Nutr. J., 2011. 10: p. 49.
63. Paradis, V., et al., Hepatocellular carcinomas in patients with metabolic syndrome often develop without significant liver fibrosis: a pathological analysis. Hepatology, 2009. 49(3): p. 851-9.
64. Li, Z.F., et al., Gas chromatography/time-of-flight mass spectrometry-based metabolomics of hepatocarcinoma in rats with lung metastasis: elucidation of the metabolic characteristics of hepatocarcinoma at formation and metastasis. Rapid Commun. Mass Spectrom., 2010. 24(18): p. 2765-75.
65. Locasale, J.W. and L.C. Cantley, *Altered metabolism in cancer*. BMC biology, 2010. 8: p. 88.
66. Briere, J.J., et al., Tricarboxylic acid cycle dysfunction as a cause of human diseases and tumor formation. Am J Physiol Cell Physiol, 2006. 291(6): p. 7.
67. Vander Heiden, M.G., L.C. Cantley, and C.B. Thompson, *Understanding the Warburg effect: the metabolic requirements of cell proliferation*. Science, 2009. 324(5930): p. 1029-33.
68. Ferreira, L.M., A. Hebrant, and J.E. Dumont, *Metabolic reprogramming of the tumor*. Oncogene, 2012.
69. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. 144(5): p. 646-74.
70. Hsu, P.P. and D.M. Sabatini, *Cancer cell metabolism: Warburg and beyond*. Cell, 2008. 134(5): p. 703-7.
71. Gubb, E. and R. Matthiesen, *Introduction to omics*. Methods Mol Biol, 2010. 593: p. 1-23.
72. Calzone, K., D. Wattendorf, and B.K. Dunn, *The application of genetics and genomics to cancer prevention*. Semin Oncol, 2010. 37(4): p. 407-18.
73. Borgan, E., et al., Merging transcriptomics and metabolomics--advances in breast cancer profiling. BMC cancer, 2010. 10: p. 628.
74. Jain, K.K., Role of pharmacoproteomics in the development of personalized medicine. Pharmacogenomics, 2004. 5(3): p. 331-6.

75. Griffiths, W.J., et al., *Targeted metabolomics for biomarker discovery*. Angew Chem Int Ed Engl, 2010. 49(32): p. 5426-45.
76. Monteiro, M.S., et al., Metabolomics analysis for biomarker discovery: advances and challenges. Curr Med Chem, 2013. 20(2): p. 257-71.
77. Pesce, F., S. Pathan, and F.P. Schena, *From -omics to personalized medicine in nephrology: integration is the key*. Nephrol Dial Transplant, 2013. 28(1): p. 24-8.
78. Zhang, W., F. Li, and L. Nie, Integrating multiple 'omics' analysis for microbial biology: application and methodologies. Microbiology, 2010. 156(Pt 2): p. 287-301.
79. Wishart, D.S., et al., *HMDB: the Human Metabolome Database*. Nucleic Acids Res, 2007. 35(Database issue): p. D521-6.
80. Patti, G.J., O. Yanes, and G. Siuzdak, *Innovation: Metabolomics: the apogee of the omics trilogy*. Nat Rev Mol Cell Biol, 2012. 13(4): p. 263-9.
81. Phizicky, E., et al., *Protein analysis on a proteomic scale*. Nature, 2003. 422(6928): p. 208-15.
82. Hanash, S., *Disease proteomics*. Nature, 2003. 422(6928): p. 226-32.
83. Gramolini, A.O., S.M. Peterman, and T. Kislinger, *Mass spectrometry-based proteomics: a useful tool for biomarker discovery?* Clin. Pharmacol. Ther., 2008. 83(5): p. 758-60.
84. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics*. Nature, 2003. 422(6928): p. 198-207.
85. Garbis, S., G. Lubec, and M. Fountoulakis, *Limitations of current proteomics technologies*. J Chromatogr A, 2005. 1077(1): p. 1-18.
86. Molloy, M.P. and F.A. Witzmann, *Proteomics: technologies and applications*. Brief Funct Genomic Proteomic, 2002. 1(1): p. 23-39.
87. Tyers, M. and M. Mann, *From genomics to proteomics*. Nature, 2003. 422(6928): p. 193-7.
88. Gorg, A., W. Weiss, and M.J. Dunn, *Current two-dimensional electrophoresis technology for proteomics*. Proteomics, 2004. 4(12): p. 3665-85.
89. Mann, M., R.C. Hendrickson, and A. Pandey, *Analysis of proteins and proteomes by mass spectrometry*. Annu Rev Biochem, 2001. 70: p. 437-73.
90. Sui, J., et al., Comparative proteomic analysis of extracellular proteins reveals secretion of T-kininogen from vascular smooth muscle cells in response to incubation with s-enantiomer of propranolol. Mol Pharm, 2008. 5(5): p. 885-90.

91. Delahunty, C. and J.R. Yates, 3rd, *Protein identification using 2D-LC-MS/MS*. Methods, 2005. 35(3): p. 248-55.
92. Sui, J., et al., iTRAQ-coupled 2D LC-MS/MS analysis on protein profile in vascular smooth muscle cells incubated with S- and R-enantiomers of propranolol: possible role of metabolic enzymes involved in cellular anabolism and antioxidant activity. J Proteome Res, 2007. 6(5): p. 1643-51.
93. Cravatt, B.F., G.M. Simon, and J.R. Yates, 3rd, *The biological impact of mass-spectrometry-based proteomics*. Nature, 2007. 450(7172): p. 991-1000.
94. Koomen, J., D. Hawke, and R. Kobayashi, Developing an understanding of proteomics: an introduction to biological mass spectrometry. Cancer Invest, 2005. 23(1): p. 47-59.
95. Street, J.M. and J.W. Dear, *The application of mass-spectrometry-based protein biomarker discovery to theragnostics*. British Journal of Clinical Pharmacology, 2010. 69(4): p. 367-378.
96. Mirgorodskaya, O.A., et al., Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. Rapid Commun Mass Spectrom, 2000. 14(14): p. 1226-32.
97. McAlister, G.C., et al., Implementation of electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. Anal Chem, 2007. 79(10): p. 3525-34.
98. Frohlich, T. and G.J. Arnold, Proteome research based on modern liquid chromatography--tandem mass spectrometry: separation, identification and quantification. J Neural Transm, 2006. 113(8): p. 973-94.
99. Zhang, J., et al., Protein profile in neuroblastoma cells incubated with S- and R-enantiomers of ibuprofen by iTRAQ-coupled 2-D LC-MS/MS analysis: possible action of induced proteins on Alzheimer's disease. Proteomics, 2008. 8(8): p. 1595-607.
100. Peng, J. and S.P. Gygi, *Proteomics: the move to mixtures*. J Mass Spectrom, 2001. 36(10): p. 1083-91.
101. Griffiths, S.D., et al., The use of isobaric tag peptide labeling (iTRAQ) and mass spectrometry to examine rare, primitive hematopoietic cells from patients with chronic myeloid leukemia. Mol Biotechnol, 2007. 36(2): p. 81-9.
102. Glen, A., et al., iTRAQ-facilitated proteomic analysis of human prostate cancer cells identifies proteins associated with progression. J Proteome Res, 2008. 7(3): p. 897-907.
103. Ross, P.L., et al., Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics, 2004. 3(12): p. 1154-69.
104. Pierce, A., et al., Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. Mol Cell Proteomics, 2008. 7(5): p. 853-63.

105. Griffin, T.J., et al., iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. *J Proteome Res*, 2007. 6(11): p. 4200-9.
106. Qian, W.J., et al., Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol Cell Proteomics*, 2006. 5(10): p. 1727-44.
107. Takahashi, N., et al., *Proteomics: advanced technology for the analysis of cellular function*. *J Nutr*, 2003. 133(6 Suppl 1): p. 2090S-2096S.
108. Chambers, G., et al., *Proteomics: a new approach to the study of disease*. *J Pathol*, 2000. 192(3): p. 280-8.
109. Krejsa, C., M. Rogge, and W. Sadee, *Protein therapeutics: new applications for pharmacogenetics*. *Nat Rev Drug Discov*, 2006. 5(6): p. 507-21.
110. Rifai, N., M.A. Gillette, and S.A. Carr, Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol*, 2006. 24(8): p. 971-83.
111. Hollywood, K., D.R. Brison, and R. Goodacre, *Metabolomics: current technologies and future trends*. *Proteomics*, 2006. 6(17): p. 4716-23.
112. Wishart, D.S., *Metabolomics: the principles and potential applications to transplantation*. *Am. J. Transplant*, 2005. 5(12): p. 2814-20.
113. Bino, R.J., et al., *Potential of metabolomics as a functional genomics tool*. *Trends Plant Sci*, 2004. 9(9): p. 418-25.
114. Xue, R., et al., Gas chromatography/mass spectrometry screening of serum metabolomic biomarkers in hepatitis B virus infected cirrhosis patients. *Clin. Chem. Lab. Med.*, 2009. 47(3): p. 305-10.
115. Brown, M., et al., Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst*, 2009. 134(7): p. 1322-32.
116. Dunn, W.B. and D.I. Ellis, *Metabolomics: Current analytical platforms and methodologies*. *TrAC Trends in Analytical Chemistry*, 2005. 24(4): p. 285-294.
117. Dettmer, K., P.A. Aronov, and B.D. Hammock, *Mass spectrometry-based metabolomics*. *Mass Spectrom Rev*, 2007. 26(1): p. 51-78.
118. Want, E.J., B.F. Cravatt, and G. Siuzdak, *The expanding role of mass spectrometry in metabolite profiling and characterization*. *Chembiochem.*, 2005. 6(11): p. 1941-51.
119. Kanani, H., P.K. Chrysanthopoulos, and M.I. Klapa, *Standardizing GC-MS metabolomics*. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.*, 2008. 871(2): p. 191-201.
120. Jonsson, P., et al., A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS. *Anal Chem*, 2004. 76(6): p. 1738-45.

121. Koek, M.M., et al., Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics*, 2011. 7(3): p. 307-328.
122. Jonsson, P., et al., High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. *Anal Chem*, 2005. 77(17): p. 5635-42.
123. Bai, J., et al., Metabolic profiling of HepG2 cells incubated with S(-) and R(+) enantiomers of anti-coagulating drug warfarin. *Metabolomics* 2011. 7(3): p. 353-362.
124. Xia, J., et al., Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics*, 2013. 9(2): p. 280-299.
125. Xue, R., et al., A serum metabolomic investigation on hepatocellular carcinoma patients by chemical derivatization followed by gas chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom.*, 2008. 22(19): p. 3061-8.
126. Wang, M., et al., Metabolomic profiling of cellular responses to carvedilol enantiomers in vascular smooth muscle cells. *PloS one*, 2010. 5(11): p. e15441.
127. Chen, T., et al., Serum and urine metabolite profiling reveals potential biomarkers of human hepatocellular carcinoma. *Mol Cell Proteomics*, 2011. 10(7): p. M110 004945.
128. Wang, X., A. Zhang, and H. Sun, Power of metabolomics in diagnosis and biomarker discovery of hepatocellular carcinoma. *Hepatology*, 2013. 57(5): p. 2072-7.
129. Niu, D., H. Feng, and W.N. Chen, Proteomic analysis of HBV-associated HCC: insights on mechanisms of disease onset and biomarker discovery. *J. Proteomics*, 2010. 73(7): p. 1283-90.
130. Feng, H., M. Wang, and W.N. Chen, iTRAQ-coupled 2D LC-MS/MS analysis of secreted proteome of HBV-replicating HepG2 cells: potential in biomarkers for prognosis of HCC. *Curr. Microbiol.*, 2010. 61(4): p. 280-4.
131. Tan, T.L. and W.N. Chen, A proteomics analysis of cellular proteins associated with HBV genotype-specific HBX: potential in identification of early diagnostic markers for HCC. *J Clin Virol*, 2005. 33(4): p. 293-8.
132. Zhang, J., et al., Protein profile in hepatitis B virus replicating rat primary hepatocytes and HepG2 cells by iTRAQ-coupled 2-D LC-MS/MS analysis: Insights on liver angiogenesis. *Proteomics*, 2009. 9(10): p. 2836-45.
133. Wang, M., et al., Proteomic profiling of cellular responses to Carvedilol enantiomers in vascular smooth muscle cells by iTRAQ-coupled 2-D LC-MS/MS. *J Proteomics*, 2010. 73(8): p. 1601-11.
134. Bai, J., et al., A comparative proteomic analysis of HepG2 cells incubated by S(-) and R(+) enantiomers of anti-coagulating drug warfarin. *Proteomics*, 2010. 10(7): p. 1463-73.

135. Saminathan, R., et al., VKORC1 pharmacogenetics and pharmacoproteomics in patients on warfarin anticoagulant therapy: transthyretin precursor as a potential biomarker. PLoS One, 2010. 5(12): p. e15064.
136. Sui, J., et al., Comparative proteomics analysis of vascular smooth muscle cells incubated with S- and R-enantiomers of atenolol using iTRAQ-coupled two-dimensional LC-MS/MS. Mol Cell Proteomics, 2008. 7(6): p. 1007-18.
137. Xu, J., et al., Comparative proteomics profile of osteoblasts cultured on dissimilar hydroxyapatite biomaterials: an iTRAQ-coupled 2-D LC-MS/MS analysis. Proteomics, 2008. 8(20): p. 4249-58.
138. Niu, D., et al., HBx genotype D represses GSTP1 expression and increases the oxidative level and apoptosis in HepG2 cells. Molecular oncology, 2009. 3(1): p. 67-76.
139. Mal, M., et al., Development and validation of a gas chromatography/mass spectrometry method for the metabolic profiling of human colon tissue. Rapid Commun Mass Spectrom, 2009. 23(4): p. 487-94.
140. Halket, J.M. and V.G. Zaikin, *Derivatization in mass spectrometry--I. Silylation*. Eur J Mass Spectrom (Chichester, Eng), 2003. 9(1): p. 1-21.
141. Lu, Y.W., et al., Cellular apoptosis induced by replication of hepatitis B virus: possible link between viral genotype and clinical outcome. Virol J, 2007. 4: p. 117.
142. Satelli, A. and S. Li, Vimentin in cancer and its potential as a molecular target for cancer therapy. Cell Mol Life Sci, 2011. 68(18): p. 3033-46.
143. Wei, J., et al., Overexpression of vimentin contributes to prostate cancer invasion and metastasis via src regulation. Anticancer research, 2008. 28(1A): p. 327-34.
144. Lu, Q.Y., et al., Effects of green tea extract on lung cancer A549 cells: proteomic identification of proteins associated with cell migration. Proteomics, 2009. 9(3): p. 757-67.
145. Lokman, N.A., et al., *The role of annexin A2 in tumorigenesis and cancer progression*. Cancer Microenviron, 2011. 4(2): p. 199-208.
146. Emoto, K., et al., Annexin II overexpression is correlated with poor prognosis in human gastric carcinoma. Anticancer research, 2001. 21(2B): p. 1339-45.
147. Ohno, Y., et al., Annexin II represents metastatic potential in clear-cell renal cell carcinoma. Br J Cancer, 2009. 101(2): p. 287-94.
148. Duncan, R., et al., Characterisation and protein expression profiling of annexins in colorectal cancer. Br J Cancer, 2008. 98(2): p. 426-33.
149. Sharma, M.R., et al., Angiogenesis-associated protein annexin II in breast cancer: selective expression in invasive breast cancer and contribution to tumor invasion and progression. Exp Mol Pathol, 2006. 81(2): p. 146-56.

150. Esposito, I., et al., Tenascin C and annexin II expression in the process of pancreatic carcinogenesis. *J. Pathol.*, 2006. 208(5): p. 673-85.
151. Ji, N.Y., et al., Evaluation of annexin II as a potential serum marker for hepatocellular carcinoma using a developed sandwich ELISA method. *Int J Mol Med*, 2009. 24(6): p. 765-71.
152. Narayan, R., et al., Proteomic analysis of HepaRG cells: a novel cell line that supports hepatitis B virus infection. *J Proteome Res*, 2009. 8(1): p. 118-22.
153. Annibaldi, A. and C. Widmann, *Glucose metabolism in cancer cells*. *Curr Opin Clin Nutr Metab Care*, 2010. 13(4): p. 466-70.
154. Shaw, R.J., *Glucose metabolism and cancer*. *Curr Opin Cell Biol*, 2006. 18(6): p. 598-608.
155. Hamanaka, R.B. and N.S. Chandel, *Targeting glucose metabolism for cancer therapy*. *J Exp Med*, 2012. 209(2): p. 211-5.
156. Wolf, A., et al., Hexokinase 2 is a key mediator of aerobic glycolysis and promotes tumor growth in human glioblastoma multiforme. *J Exp Med*, 2011. 208(2): p. 313-26.
157. Chen, J.Q. and J. Russo, Dysregulation of glucose transport, glycolysis, TCA cycle and glutaminolysis by oncogenes and tumor suppressors in cancer cells. *Biochimica et biophysica acta*, 2012. 2: p. 370-84.
158. King, A., M.A. Selak, and E. Gottlieb, Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. *Oncogene*, 2006. 25(34): p. 4675-82.
159. Warburg, O., *On the origin of cancer cells*. *Science*, 1956. 123(3191): p. 309-14.
160. Seyfried, T.N. and L.M. Shelton, *Cancer as a metabolic disease*. *Nutr. Metab.*, 2010. 7: p. 7.
161. Lin, C.P., et al., *Targeting c-Myc as a novel approach for hepatocellular carcinoma*. *World J. Hepatol.*, 2010. 2(1): p. 16-20.
162. Ward, P.S. and C.B. Thompson, Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. *Cancer Cell*, 2012. 21(3): p. 297-308.
163. Trachootham, D., J. Alexandre, and P. Huang, *Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach?* *Nat Rev Drug Discov*, 2009. 8(7): p. 579-91.
164. Spratlin, J.L., N.J. Serkova, and S.G. Eckhardt, *Clinical applications of metabolomics in oncology: a review*. *Clin Cancer Res*, 2009. 15(2): p. 431-40.
165. Gatenby, R.A. and R.J. Gillies, *Why do cancers have high aerobic glycolysis?* *Nat. Rev. Cancer*, 2004. 4(11): p. 891-9.
166. Zhao, Y., E.B. Butler, and M. Tan, *Targeting cellular metabolism to improve cancer therapeutics*. *Cell Death Dis*, 2013. 4: p. e532.

167. Jain, M., et al., Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science*, 2012. 336(6084): p. 1040-4.
168. Zhang, W.C., et al., Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. *Cell*, 2012. 148(1-2): p. 259-72.
169. Tessem, M.B., et al., Evaluation of lactate and alanine as metabolic biomarkers of prostate cancer using ¹H HR-MAS spectroscopy of biopsy tissues. *Magn Reson Med*, 2008. 60(3): p. 510-6.
170. Newsholme, P., et al., Glutamine and glutamate--their central role in cell metabolism and function. *Cell Biochem. Funct.*, 2003. 21(1): p. 1-9.
171. DeBerardinis, R.J. and T. Cheng, Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene*, 2010. 29(3): p. 313-24.
172. Dang, C.V., Glutaminolysis: supplying carbon or nitrogen or both for cancer cells? *Cell Cycle*, 2010. 9(19): p. 3884-6.
173. Gao, P., et al., c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature*, 2009. 458(7239): p. 762-5.
174. Dang, C.V., Rethinking the Warburg effect with Myc micromanaging glutamine metabolism. *Cancer Res*, 2010. 70(3): p. 859-62.
175. Liu, W., et al., Reprogramming of proline and glutamine metabolism contributes to the proliferative and metabolic responses regulated by oncogenic transcription factor c-MYC. *Proc Natl Acad Sci U S A*, 2012. 109(23): p. 8983-8.
176. Artwohl, M., et al., Free fatty acids trigger apoptosis and inhibit cell cycle progression in human vascular endothelial cells. *FASEB J*, 2004. 18(1): p. 146-8.
177. Mu, Y.M., et al., Saturated FFAs, palmitic acid and stearic acid, induce apoptosis in human granulosa cells. *Endocrinology*, 2001. 142(8): p. 3590-7.
178. Muir, K., et al., Proteomic and lipidomic signatures of lipid metabolism in NASH-associated hepatocellular carcinoma. *Cancer Res*, 2013.
179. Beyoglu, D., et al., Tissue metabolomics of hepatocellular carcinoma: Tumor energy metabolism and the role of transcriptomic classification. *Hepatology*, 2013. 58(1): p. 229-38.
180. Santos, C.R. and A. Schulze, *Lipid metabolism in cancer*. *FEBS J*, 2012. 279(15): p. 2610-23.
181. Hwang, G.Y., et al., Dominant mutations of hepatitis B virus variants in hepatoma accumulate in B-cell and T-cell epitopes of the HBx antigen. *Virus Res*, 2003. 92(2): p. 157-64.
182. Pang, R., et al., Pin1 interacts with a specific serine-proline motif of hepatitis B virus X-protein to enhance hepatocarcinogenesis. *Gastroenterology*, 2007. 132(3): p. 1088-103.

Appendix

Appendix 1: List of differentially expressed proteins in HepG2 cells affected by HBx A/B/C identified by LC/MS/MS.

	Accession #	Protein Name
1	P10809	60 kDa heat shock protein, mitochondrial
2	P11021	78 kDa glucose-regulated protein
3	P63261	Actin, cytoplasmic 2
4	P08729	Keratin, type II cytoskeletal 7
5	P05787	Keratin, type II cytoskeletal 8
6	P02768	Serum albumin
7	P08238	Heat shock protein HSP 90
8	P06733	Alpha-enolase
9	P02765	Alpha-2-HS-glycoprotein
10	P14618	Pyruvate kinase isozymes M1/M2
11	P08670	Vimentin
12	P06576	ATP synthase subunit beta, mitochondrial
13	P02545	Lamin-A/C
14	P31327	Carbamoyl-phosphate synthase
15	P05783	Keratin, type I cytoskeletal 18
16	P07237	Protein disulfide-isomerase
17	P68363	Tubulin alpha-1B chain
18	P25705	ATP synthase subunit alpha, mitochondrial
19	P07437	Tubulin beta chain
20	P08758	Annexin A5
21	P00558	Phosphoglycerate kinase
22	P04695	Keratin, type I cytoskeletal 17
23	P15084	Protein disulfide-isomerase A6
24	P04406	Glyceraldehyde-3-phosphate dehydrogenase
25	P05386	60S acidic ribosomal protein P1
26	P05387	60S acidic ribosomal protein P2
27	P25788	Proteasome subunit alpha
28	P00338	L-Lactate dehydrogenase
29	P12429	Annexin A3
30	Q6FI13	Histone H2A type 2-A
31	Q99879	Histone H2B type 1-M
32	P26447	Protein S100-A4
33	P35579	Myosin-9

34	P30101	Protein disulfide-isomerase A3
35	Q5JR01	Elongation factor 1-alpha 1
36	P06703	Protein S100-A6
37	P31949	Protein S100-A11
38	P06748	Nucleophosmin
39	P07355	Annexin A2
40	P61604	10 kDa heat shock protein, mitochondrial
41	P04075	Fructose-bisphosphate aldolase A
42	P62805	Histone H4
43	P07108	Acyl-CoA-binding protein
44	P46782	40S ribosomal protein S5
45	P05388	60S acidic ribosomal protein P0
46	P23284	Peptidyl-prolyl cis-trans isomerase B
47	P26641	Elongation factor 1-gamma
48	P60709	Actin, cytoplasmic 1
49	P21333	Filamin-A
50	O75369	Filamin-B
51	P57053	Histone H2A type 2-F
52	Q99878	Histone H2A type 1-J
53	P04908	Histone H2A type 1-B
54	Q16777	Histone H2A type 2-C
55	P15531	Nucleoside diphosphate kinase A
56	P61769	Beta-2-microglobulin precursor
57	P04114	Apolipoprotein B-100 precursor
58	Q99497	Protein DJ-1
59	P18669	Phosphoglycerate mutase 1
60	P06744	Glucose-6-phosphate isomerase
61	P17931	Galectin-3
62	O43707	Alpha-actinin-4
63	P12814	Alpha-actinin-1
64	P63165	Small ubiquitin-related modifier 1
65	P09429	High mobility group protein B1
66	P0C0L4	Complement C4-A precursor
67	P02766	Transthyretin precursor
68	P08708	40S ribosomal protein S17
69	P62081	40S ribosomal protein S7
70	P78371	T-complex protein 1 subunit beta
71	P30041	Peroxiredoxin-6
72	P14625	Endoplasmin precursor
73	P40926	Malate dehydrogenase, mitochondrial
74	P01009	Alpha-1-antitrypsin precursor
75	P60866	40S ribosomal protein S20

76	P04080	Cystatin-B
77	P02763	Alpha-1-acid glycoprotein 1 precursor
78	P01023	Alpha-2-macroglobulin precursor
79	P0278	Serotransferrin precursor
80	P0C0L5	Complement C4-B precursor
81	P02679	Fibrinogen gamma chain precursor
82	P01857	Ig gamma-1 chain C region
83	P02671	Fibrinogen alpha chain precursor
84	P02647	Apolipoprotein A-I precursor
85	P02675	Fibrinogen beta chain precursor
86	P00738	Haptoglobin precursor
87	P02751	Fibronectin precursor
88	P02790	Hemopexin precursor
89	P01834	Ig kappa chain C region
90	P02774	Vitamin D-binding protein precursor
91	P49720	Proteasome subunit beta type 3
92	P01031	Complement C5 precursor
93	P68871	Hemoglobin subunit beta
94	P10643	Complement component C7 precursor
95	P02655	Apolipoprotein C-II precursor
96	P11142	Heat shock cognate 71 kDa protein
97	P21912	Succinate dehydrogenase
98	P49327	Fatty acid synthase
99	P29401	Transketolase
100	P61978	Heterogeneous nuclear ribonucleoprotein
101	P23921	Ribonucleoside-diphosphate reductase large subunit
102	P55072	Transitional endoplasmic reticulum ATPase
103	P15121	Aldose reductase
104	P32119	Peroxiredoxin-2
105	P53396	ATP-citrate synthase
106	P22392	Nucleoside diphosphate kinase B
107	P09651	Heterogeneous nuclear ribonucleoprotein A1
108	Q08211	ATP-dependent RNA helicase A
109	P13797	Plastin-3
110	P27797	Calreticulin
111	Q06830	Peroxiredoxin-1
112	P38646	Stress-70 protein, mitochondrial
113	P63241	Eukaryotic translation initiation factor 5A-1
114	P60842	Eukaryotic initiation factor 4A-I
115	P50395	Rab GDP dissociation inhibitor beta
116	P37802	Transgelin-2
117	Q14974	Importin subunit beta-1

118	P23528	Cofilin-1
119	O00299	Chloride intracellular channel protein 1
120	P00505	Aspartate aminotransferase, mitochondrial
121	Q12906	Interleukin enhancer-binding factor 3
122	Q16719	Kynureninase
123	P55072	Transitional endoplasmic reticulum ATPase
124	Q16576	Histone-binding protein RBBP7
125	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1
126	P62826	GTP-binding nuclear protein Ran
127	P42224	Signal transducer and activator of transcription 1- alpha/beta
128	P49588	Alanyl-tRNA synthetase, cytoplasmic
129	O60664	Mannose-6-phosphate receptor-binding protein 1
130	P02786	Transferrin receptor protein 1
131	Q13283	Ras GTPase-activating protein-binding protein 1
132	P05141	ADP/ATP translocase 2
133	P63244	Guanine nucleotide-binding protein subunit beta-2-like 1
134	Q15181	Inorganic pyrophosphatase
135	P55209	Nucleosome assembly protein 1-like 1
136	Q13501	Sequestosome-1
137	P23246	Splicing factor, proline- and glutamine-rich
138	P07954	Fumarate hydratase, mitochondrial
139	P35241	Radixin
140	P12268	Inosine-5'-monophosphate dehydrogenase 2
141	P22314	Ubiquitin-like modifier-activating enzyme 1
142	P55060	Exportin-2
143	P09960	Leukotriene A-4 hydrolase 2
144	P99999	Cytochrome c
145	P60891	Ribose-phosphate pyrophosphokinase 1
146	P32322	Pyrroline-5-carboxylate reductase 1, mitochondrial
147	Q9BWD1	Acetyl-CoA acetyltransferase, cytosolic
148	P26639	Threonyl-tRNA synthetase, cytoplasmic
149	P00491	Purine nucleoside phosphorylase
150	P29966	Myristoylated alanine-rich C-kinase substrate
151	P60660	Myosin light polypeptide 6
152	O76021	Ribosomal L1 domain-containing protein 1

Appendix 2: List of differentially expressed proteins in RPHs affected by HBx A/B/C identified by LC/MS/MS.

	Accession #	Protein Name
1	P63259	Actin, cytoplasmic 2
2	Q6IG12	Keratin, type II cytoskeletal 7
3	Q10758	Keratin, type II cytoskeletal 8
4	P02770	Serum albumin
5	P34058	Heat shock protein HSP 90-beta
6	P63039	60 kDa heat shock protein, mitochondrial
7	P06761	78 kDa glucose-regulated protein
8	Q6IFU8	Keratin, type I cytoskeletal 17
9	P14668	Annexin A5
10	P14669	Annexin A3
11	P16617	Phosphoglycerate kinase
12	P10719	ATP synthase subunit beta, mitochondrial
13	P0CC09	Histone H2A type 2-A
14	Q62812	Myosin-9
15	Q5BJY9	Keratin, type I cytoskeletal 18
16	P11598	Protein disulfide-isomerase A3
17	P04764	Alpha-enolase
18	P15999	ATP synthase subunit alpha, mitochondrial
19	P85108	Tubulin beta-2A chain
20	P24090	Alpha-2-HS-glycoprotein
21	P31000	Vimentin
22	Q6P9V9	Tubulin alpha-1B chain
23	P26772	10 kDa heat shock protein, mitochondrial
24	P19945	60S acidic ribosomal protein P0
25	P19944	60S acidic ribosomal protein P1
26	P02401	60S acidic ribosomal protein P2
27	P05065	Fructose-bisphosphate aldolase A
28	Q07936	Annexin A2
29	P32089	Tricarboxylate transport protein, mitochondrial
30	Q05982	Nucleoside diphosphate kinase A
31	Q00729	Histone H2B type 1-A
32	Q6P6V0	Glucose-6-phosphate isomerase
33	P07756	Carbamoyl-phosphate synthase
34	P62083	40S ribosomal protein S7
35	O88767	Protein DJ-1
36	P02262	Histone H2A type 1
37	P27881	Hexokinase 2

38	Q8K4T4	Filamin-A-interacting protein 1
39	P04642	L-Lactate dehydrogenase
40	Q63081	Protein disulfide-isomerase A6
41	P18422	Proteasome subunit alpha type-3
42	P07151	Beta-2-microglobulin precursor
43	P08699	Galectin-3
44	Q68FQ0	T-complex protein 1 subunit epsilon
45	P04797	Glyceraldehyde-3-phosphate dehydrogenase
46	P04644	40S ribosomal protein S17
47	Q9QXQ0	Alpha-actinin-4
48	Q9Z1P2	Alpha-actinin-1
49	P13084	Nucleophosmin
50	P11980	Pyruvate kinase isozymes M1/M2
51	P05964	Protein S100-A6
52	Q6B345	Protein S100-A11
53	Q64598	Histone H2A type 1-F
54	Q00715	Histone H2B type 1
55	Q920L2	Succinate dehydrogenase
56	P62804	Histone H4
57	Q7TMA5	Apolipoprotein B-100 precursor
58	P62630	Elongation factor 1-alpha 1
59	P02680	Fibrinogen gamma chain
60	P05942	Protein S100-A4
61	P11030	Acyl-CoA-binding protein
62	Q68FR6	Elongation factor 1-gamma
63	P60711	Actin, cytoplasmic 1
64	P04276	Vitamin D-binding protein precursor
65	P52925	High mobility group protein B2
66	O88426	Nucleoside diphosphate kinase 6
67	P02767	Transthyretin precursor
68	P24368	Peptidyl-prolyl cis-trans isomerase B
69	P04785	Protein disulfide-isomerase
70	P38983	40S ribosomal protein SA
71	Q62826	Heterogeneous nuclear ribonucleoprotein M
72	P28075	Proteasome subunit beta type-5
73	P01835	Ig kappa chain C region
74	P14408	Fumarate hydratase
75	P18418	Calreticulin
76	P62828	GTP-binding nuclear protein Ran
77	P08650	Complement C5 precursor
78	P50137	Transketolase
79	Q63598	Plastin-3

80	P0C170	Histone H2A type 1-E
81	P25113	Phosphoglycerate mutase 1
82	P02091	Hemoglobin subunit beta-1
83	P20059	Hemopexin precursor
84	P04639	Apolipoprotein A-I
85	P14480	Fibrinogen beta chain
86	P06866	Haptoglobin precursor
87	P04937	Fibronectin precursor
88	P06399	Fibrinogen alpha chain
89	P04636	Malate dehydrogenase, mitochondrial
90	Q5XHZ0	Heat shock protein 75 kDa, mitochondrial
91	Q66HD0	Endoplasmic
92	P02764	Alpha-1-acid glycoprotein
93	P20759	Ig gamma-1 chain C region
94	P16290	Phosphoglycerate mutase 2
95	P55797	Apolipoprotein C-IV
96	P63018	Heat shock cognate 71 kDa protein
97	P82995	Heat shock protein HSP 90-alpha
98	P29266	3-hydroxyisobutyrate dehydrogenase
99	P06238	Alpha-2-macroglobulin
100	O35244	Peroxiredoxin-6
101	P14841	Cystatin-C
102	P00884	Fructose-bisphosphate aldolase B
103	Q09073	ADP/ATP translocase 2
104	P18420	Proteasome subunit alpha type-1
105	P62260	14-3-3 protein epsilon
106	A7VJC2	Heterogeneous nuclear ribonucleoproteins A2/B1
107	P04256	Heterogeneous nuclear ribonucleoprotein A1
108	P63245	Guanine nucleotide-binding protein subunit beta-2-like 1
109	P12346	Serotransferrin
110	P60868	40S ribosomal protein S20
111	P30349	Leukotriene A-4 hydrolase
112	Q66HF8	Aldehyde dehydrogenase
113	P56571	ES1 protein homolog, mitochondrial
114	P20673	Argininosuccinate lyase
115	Q9Z2G8	Nucleosome assembly protein 1-like 1
116	P10633	Cytochrome P450 2D1
117	P62898	Cytochrome c
118	Q5XHY5	Threonine--tRNA ligase, cytoplasmic
119	P30904	Macrophage migration inhibitory factor
120	Q5U300	Ubiquitin-like modifier-activating enzyme 1
121	Q9JIL3	Interleukin enhancer-binding factor 3

122	P70712	Kynureninase
123	P46462	Transitional endoplasmic reticulum ATPase
124	B2GUY2	Transferrin receptor protein 2
125	P19804	Nucleoside diphosphate kinase B
126	P01041	Cystatin-B
127	P46462	Transitional endoplasmic reticulum ATPase
128	P62632	Elongation factor 1-alpha 2
129	P16638	ATP-citrate synthase
130	P17764	Acetyl-CoA acetyltransferase, mitochondrial
131	P07943	Aldose reductase
132	P11497	Acetyl-CoA carboxylase
133	P50399	Rab GDP dissociation inhibitor beta
134	Q64119	Myosin light polypeptide 6
135	P48721	Stress-70 protein, mitochondrial
136	G3V679	Transferrin receptor protein 1
137	P48500	Triosephosphate isomerase

Appendix 3: List of differentially expressed proteins in HepG2 cells affected by mutant HBx A/B/C identified by LC/MS/MS.

	Accession #	Protein Name
1	P02768	Serum albumin
2	P63261	Actin, cytoplasmic 2
3	P08729	Keratin, type II cytoskeletal 7
4	P11021	78 kDa glucose-regulated protein
5	P00558	Phosphoglycerate kinase
6	P08238	Heat shock protein HSP 90
7	P10809	60 kDa heat shock protein, mitochondrial
8	P02545	Lamin-A/C
9	Q99879	Histone H2B type 1-M
10	P07437	Tubulin beta chain
11	P68363	Tubulin alpha-1B chain
12	P06576	ATP synthase subunit beta, mitochondrial
13	P06733	Alpha-enolase
14	P31327	Carbamoyl-phosphate synthase
15	P05783	Keratin, type I cytoskeletal 18
16	P07237	Protein disulfide-isomerase
17	P08670	Vimentin
18	Q5JR01	Elongation factor 1-alpha 1
19	P14618	Pyruvate kinase isozymes M1/M2
20	P31949	Protein S100-A11
21	P05787	Keratin, type II cytoskeletal 8
22	P00338	L-Lactate dehydrogenase
23	P06748	Nucleophosmin
24	P07355	Annexin A2
25	P35579	Myosin-9
26	P05387	60S acidic ribosomal protein P2
27	P06703	Protein S100-A6
28	P04695	Keratin, type I cytoskeletal 17
29	P12429	Annexin A3
30	Q6FI13	Histone H2A type 2-A
31	P02765	Alpha-2-HS-glycoprotein
32	P26447	Protein S100-A4
33	P05386	60S acidic ribosomal protein P1
34	P30101	Protein disulfide-isomerase A3
35	P25705	ATP synthase subunit alpha, mitochondrial
36	P25788	Proteasome subunit alpha

37	P08758	Annexin A5
38	P15084	Protein disulfide-isomerase A6
39	P04406	Glyceraldehyde-3-phosphate dehydrogenase
40	P21333	Filamin-A
41	P18669	Phosphoglycerate mutase 1
42	P57053	Histone H2A type 2-F
43	P07108	Acyl-CoA-binding protein
44	P46782	40S ribosomal protein S5
45	O75369	Filamin-B
46	P23284	Peptidyl-prolyl cis-trans isomerase B
47	P06744	Glucose-6-phosphate isomerase
48	P60709	Actin, cytoplasmic 1
49	P61604	10 kDa heat shock protein, mitochondrial
50	P05388	60S acidic ribosomal protein P0
51	P04114	Apolipoprotein B-100 precursor
52	Q99878	Histone H2A type 1-J
53	Q99497	Protein DJ-1
54	Q16777	Histone H2A type 2-C
55	P15531	Nucleoside diphosphate kinase A
56	P61769	Beta-2-microglobulin precursor
57	P62805	Histone H4
58	P04908	Histone H2A type 1-B
59	P04075	Fructose-bisphosphate aldolase A
60	P26641	Elongation factor 1-gamma
61	P17931	Galectin-3
62	O43707	Alpha-actinin-4
63	P02655	Apolipoprotein C-II precursor
64	P63165	Small ubiquitin-related modifier 1
65	P09429	High mobility group protein B1
66	P0C0L4	Complement C4-A precursor
67	P02766	Transthyretin precursor
68	P08708	40S ribosomal protein S17
69	P61978	Heterogeneous nuclear ribonucleoprotein
70	P53396	ATP-citrate synthase
71	P30041	Peroxiredoxin-6
72	P14625	Endoplasmic precursor
73	P40926	Malate dehydrogenase, mitochondrial
74	P01009	Alpha-1-antitrypsin precursor
75	P68871	Hemoglobin subunit beta
76	P02774	Vitamin D-binding protein precursor
77	P02763	Alpha-1-acid glycoprotein 1 precursor
78	P01023	Alpha-2-macroglobulin precursor

79	P0278	Serotransferrin precursor
80	P49720	Proteasome subunit beta type 3
81	P02679	Fibrinogen gamma chain precursor
82	P01857	Ig gamma-1 chain C region
83	P02790	Hemopexin precursor
84	P02647	Apolipoprotein A-I precursor
85	P02675	Fibrinogen beta chain precursor
86	P00738	Haptoglobin precursor
87	P02751	Fibronectin precursor
88	P02671	Fibrinogen alpha chain precursor
89	P01834	Ig kappa chain C region
90	P04080	Cystatin-B
91	P0COL5	Complement C4-B precursor
92	P01031	Complement C5 precursor
93	P60866	40S ribosomal protein S20
94	P10643	Complement component C7 precursor
95	P12814	Alpha-actinin-1
96	P11142	Heat shock cognate 71 kDa protein
97	Q06830	Peroxiredoxin-1
98	P63244	Guanine nucleotide-binding protein subunit beta-2-like 1
99	P32119	Peroxiredoxin-2
100	P62081	40S ribosomal protein S7
101	P23921	Ribonucleoside-diphosphate reductase large subunit
102	P55072	Transitional endoplasmic reticulum ATPase
103	P15121	Aldose reductase
104	P29401	Transketolase
105	P78371	T-complex protein 1 subunit beta
106	O76021	Ribosomal L1 domain-containing protein 1
107	P42224	Signal transducer and activator of transcription 1-alpha/beta
108	P22314	Ubiquitin-like modifier-activating enzyme 1
109	P13797	Plastin-3
110	Q16576	Histone-binding protein RBBP7
111	P21912	Succinate dehydrogenase
112	P55060	Exportin-2
113	P63241	Eukaryotic translation initiation factor 5A-1
114	P60842	Eukaryotic initiation factor 4A-I
115	P50395	Rab GDP dissociation inhibitor beta
116	P37802	Transgelin-2
117	Q14974	Importin subunit beta-1
118	P35241	Radixin

119	O00299	Chloride intracellular channel protein 1
120	P00505	Aspartate aminotransferase, mitochondrial
121	Q12906	Interleukin enhancer-binding factor 3
122	Q16719	Kynureninase
123	P55072	Transitional endoplasmic reticulum ATPase
124	P27797	Calreticulin
125	P00491	Purine nucleoside phosphorylase
126	P12268	Inosine-5'-monophosphate dehydrogenase 2
127	P09651	Heterogeneous nuclear ribonucleoprotein A1
128	P60891	Ribose-phosphate pyrophosphokinase 1
129	O60664	Mannose-6-phosphate receptor-binding protein 1
130	Q13501	Sequestosome-1
131	Q13283	Ras GTPase-activating protein-binding protein 1
132	P05141	ADP/ATP translocase 2
133	P49327	Fatty acid synthase
134	P99999	Cytochrome c
135	P55209	Nucleosome assembly protein 1-like 1
136	P02786	Transferrin receptor protein 1
137	P23246	Splicing factor, proline- and glutamine-rich
138	P09960	Leukotriene A-4 hydrolase 2
139	P23528	Cofilin-1
140	P62826	GTP-binding nuclear protein Ran
141	Q08211	ATP-dependent RNA helicase A
142	P38646	Stress-70 protein, mitochondrial
143	P07954	Fumarate hydratase, mitochondrial
144	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1
145	P49588	Alanyl-tRNA synthetase, cytoplasmic
146	Q9BWD1	Acetyl-CoA acetyltransferase, cytosolic

Appendix 4: List of Metabolites detected, with the RT (retention time), Name and Class in HBx-transfected HepG2 cells.

No.	RT (min)	Identified Metabolites	Chemical class
1	7.48	Acetamide	Other organic acids
2	7.77	Boroic acid	Other organic acids
3	9.52	Benzamide	Ketones and Aldehydes
4	9.90	L-Lactic acid	Hydroxy acid
5	10.48	Glycolic acid	Hydroxy acid
6	11.54	L-Alanine	Amino acid
7	12.20	Hydroxylamine	Alcohols
8	12.80	Oxalic acid	Carboxylic Acid
9	14.25	Glycine	Amino acid
10	14.58	Malonic acid	Carboxylic Acid
11	15.56	L-Valine	Amino acid
12	16.15	Urea	Ketones and Aldehydes
13	16.56	Acetoacetic acid	Other organic acids
14	16.58	Benzoic acid	Other organic acids
15	17.42	L-Leucine	Amino acid
16	17.52	Phosphoric acid	Other organic acids
17	17.73	Propanoic acid	Carboxylic Acid
18	17.86	Glycerol	Alcohols
19	18.26	L-Isoleucine	Amino acid
20	18.46	L-Proline	Amino acid
21	18.84	Succinic acid	Carboxylic Acid
22	19.63	Glyceric acid	Hydroxy acid
23	19.85	Uracil	Nucleosides and Nucleoside conjugates
24	20.59	L-Serine	Amino acid
25	21.47	L-Threonine	Amino acid
26	22.1	Thymine	Nucleosides and Nucleoside conjugates
27	23.77	3-methyl-2-ketobutyric acid	Other organic acids
28	23.99	hydroxypyruvic	Hydroxy acid
29	24.10	Carbamic acid	Amino acid
30	24.73	Malic acid	Carboxylic Acid
31	25.69	L-aspartic acid	Amino acid
32	25.79	L-methionine	Amino acid
33	25.88	5-Oxoproline	Amino acid
34	28.24	Ribofuranoside	Nucleosides and Nucleoside conjugates
35	28.66	L-Glutamic acid	Amino acid
36	29.15	L-Phenylalanine	Amino acid
37	30.22	Asparagine	Amino acid
38	31.90	Aconitic acid	Carboxylic Acid
39	33.14	L-Glutamine	Amino acid
40	34.34	Glucuronic acid	Ketones and Aldehydes
41	34.44	Citrulline	Amino acid
42	34.87	Ribonic acid	Hydroxy acid

43	34.89	Myristic acid	Fatty acid
44	35.75	Adenine	Nucleosides and Nucleoside conjugates
45	36.72	D-glucose	Carbohydrates
46	36.90	D-mannose	Carbohydrates
47	37.08	L-Lysine	Amino acid
48	37.16	L-histidine	Amino acid
49	37.30	D-glucitol	Alcohols
50	37.61	L-tyrosine	Amino acid
51	38.01	D-Galactose	Carbohydrates
52	38.11	Fructose	Carbohydrates
53	39.78	Palmitic acid	Fatty acid
54	40.91	D-Ribose	Carbohydrates
55	41.10	Oleic acid	Fatty acid
56	41.36	Glucosamine	Carbohydrates
57	41.79	Myo-inositol	Alcohols
58	42.63	Stearic acid	Fatty acid
59	43.72	Vaccenic acid	Fatty acid
60	43.95	Glucose-6-phosphate	Carbohydrates
61	44.11	L-cysteine	Amino acid
62	45.20	Cadaverine	Other organic acids
63	47.12	Arachidonic acid	Fatty acid
64	48.50	Cholesterol	Steroids
65	48.91	Inositol	Alcohols
66	49.08	Uridine	Nucleosides and Nucleoside conjugates
67	51.54	Inosine	Nucleosides and Nucleoside conjugates
68	51.71	3-hydroxyadipic acid	Hydroxy acid
69	52.19	Maltose	Carbohydrates
70	53.28	Gluconic acid	Hydroxy acid
71	54.95	Adenosine	Nucleosides and Nucleoside conjugates