Investigation of SH3 binding domain in Hepatitis B virus X protein : role in virus-host interactions and cell migration

Feng, Huixing

2010

Feng, H. X. (2010). Investigation of SH3 binding domain in Hepatitis B virus X protein : role in virus-host interactions and cell migration. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/47460>

<https://doi.org/10.32657/10356/47460>

Nanyang Technological University

Downloaded on 26 Apr 2023 17:30:49 SGT

**INVESTIGATION OF SH3 BINDING DOMAIN IN
HEPATITIS B VIRUS X PROTEIN: ROLE IN VIRUS-HOST
INTERACTIONS AND CELL MIGRATION**

FENG HUIXING

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University
In partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2010

Acknowledgements

Upon the completion of my thesis, I would like to firstly give the greatest gratitude to my dear supervisor and mentor, Prof. William Chen, for his constant support, considerate instructions and warmest encouragement from the beginning to the end. He has helped me to walk through all the stages of the thesis writing. Without his illuminating and consistent instruction, this thesis could not have reached its present form. It is my great pleasure and honor to be his student while pursuing my Ph.D. degree.

Warmest thanks to my lab fellows and friends, Dr. Tan Tuan Lin, Dr. Lu Yiwei, Dr. Pan Hong, Dr. Sui Jianjun, Ms. Niu Dandan, Mr. Zhang Jianhua, Ms. Bai Jing, Mr. Wang Mingxuan, Ms. Bahareh Haji Rasouliha and Ms. Laleh Sadrolodabae for their generous help, discussion, coordination and many valuable suggestions on research and life. Working with them is a wonderful experience.

I would also like to express my heartfelt gratefulness to Prof. Vincent Chan for his technical support by sharing the cell migration track platform, which greatly complements my research.

Many thanks to all my school mates in SCBE, especially friends in PG rooms B2-06 and B1-36 for the experience and happiness you shared with me during the

past years' study.

Many thanks also to all my friends not only located in Singapore, but also anywhere. Thanks for your time in listening to me and giving me precious advice during the difficult time. No matter where you are, your support is always around me.

Special thanks to Mr. Li Xi for the support, understanding and helping me work out my problems all the time. My life would be much lonely in the absence of your company.

Thanks from my deepest heart to my beloved parents and sister for their great confidence and loving considerations in me all these years.

Last but not least, my sincere thanks to Nanyang Technological University for offering me the chance of PhD study and providing me graduate research scholarship.

Table of Contents

Acknowledgements.....I

Table of Contents III

List of Figures VI

List of Tables IX

Abbreviations X

List of Publications XIII

Summary XV

Chapter 1 : Introduction..... 1

1.1 Hepatitis B virus and epidemiology 1

1.2 Structure and genomic organization of HBV4

1.3 The replication of HBV.....7

1.4 Hepatitis B virus genotypes and disease prognosis..... 10

1.5 Hepatitis B virus X protein (HBx) 12

1.5.1 Expression of HBx 13

1.5.2 Subcellular localization of HBx 13

1.5.3 HBx and HBV replication 14

1.5.4 HBx and cellular signaling pathways..... 15

1.5.5 HBx and apoptosis 18

1.5.6 HBx and HCC 19

1.6 Proline-rich domain in HBx and SH3-mediated Signaling 19

1.7 HBV and cell migration 23

1.8 HBV and Rho GTPase-mediated Signaling 24

1.9 iTRAQ coupled 2-D LC-MS/MS based Proteome profiling.....26

1.10 Objectives of project 32

Chapter 2 : Materials and Methods..... 34

2.1 Cloning constructions 34

2.1.1 Viral DNA extraction from serum sample..... 34

2.1.2 Bacterial and mammalian expression vectors applied 35

2.1.3 Cloning construction of HBx from different genotypes..... 38

2.1.4 Cloning construction of SH3 binding domain from HBx and SH3 domains from
Vinexin-β..... 39

2.1.5 Preparation of competent cells by CaCl₂ method..... 39

2.1.6	Plasmid transformation	41
2.1.7	Glycerol stock	41
2.1.8	Plasmid amplification by QIAGEN QIAprep Miniprep	42
2.1.9	Plasmid amplification by Invitrogen Midiprep	43
2.1.10	Plasmid DNA quantification by UV spectrophotometer	44
2.1.11	Restriction enzyme digestion	44
2.1.12	Gel extraction	45
2.1.13	Ligation	46
2.1.14	Multiple sequence alignment analysis	47
2.1.15	Site-directed mutagenesis	47
2.2	Mammalian cell culture and transfection	48
2.2.1	Mammalian cell culture and cell passage	48
2.2.2	Freezing, storage and recovery of cells	49
2.2.3	Cell transfections	50
2.3	SDS-PAGE (Sodium dodecyl Sulfte - polyacrylamide gel electrophoresis)	51
2.4	Western Blot analysis	55
2.5	Total proteins extraction, protein digestion, and labeling with iTRAQ reagent for LC-MS/MS	57
2.6	Concentration of secreted proteins	60
2.7	Protein expression and purification	60
2.7.1	Bacterial protein expression	61
2.7.2	Protein purification	62
2.8	GST-pulldown assay	62
2.9	Panomics transSignal SH3 domian arrays	63
2.10	IMX analysis	64
2.11	RNA extraction, quantitation and RT-PCR	65
2.11.1	RNA extraction	65
2.11.2	RNA quantitation	66
2.11.3	RT-PCR	66
2.11.4	Real-time RT-PCR	67
2.12	Cell motility assay	69
2.12.1	Transwell assay	69
2.12.2	Surface coating and fluorescent beads based cell motility assay	70
2.13	Microscopy based real-time cell migration tracking assay	71

Chapter 3 : Genotype Specific SH3 Binding Array Assay ... 72

3.1	Cloning of HBx genotypes A/B/C	74
3.2	Genotyping of HBV genotype D	75
3.3	Cloning of HBx genotype D	77
3.4	Alignments of HBx of distinct genotypes (A/B/C/D)	78
3.5	Generation of mutations within proline/serine rich domain	80
3.6	Protein expression and purification	81
3.7	Identification of interacting SH3 domain-containing cellular proteins	84
3.8	Activation of HBV replication by SH3 domain of vinexin β	87

3.10	Discussion and Conclusions.....	92
Chapter 4 : Protein Profile in Cells Transfected with		
Genotype Specific HBx: a Comparative iTRAQ-coupled 2D		
LC-MS/MS Analysis.....		
4.1	HBx genotypes and their expression in transfected cells.....	96
4.2	iTRAQ analysis and differentially expressed proteins.....	98
4.3	Validation of differentially expressed proteins by real-time RT-PCR.....	102
4.4	Validation of differentially expressed proteins by Western blot.....	103
4.5	Cell migration assay.....	105
4.6	Discussion and Conclusions.....	108
Chapter 5 : Extracellular Protein Profile of HepG2 Cells		
Affected by HBV		
5.1	Secreted protein profile of HBV replicating HepG2 cells.....	117
5.2	Discussion and Conclusions.....	119
Chapter 6 : HBV Activated Rho GTPases Induced Cell		
Motility		
6.1	Rac1 activated high mobility of HepG2 cells infectd by HBV	124
6.2	RhoA activation by genotype specific HBx	133
6.3	Discussion and Conclusions.....	137
Chapter 7 : Conclusions.....		
Chapter 8 : Future Work.....		
References		
Appendix		

List of Figures

Chapter 1

Figure 1.1 : Global distribution of HBV infection and incidence of HCC [6].....	2
Figure 1.2: HBV virion (Dane particle) and genomic organization [16].....	5
Figure 1.3: Organization of the HBV genome.	7
Figure 1.4: HBV replication cycle [22].	9
Figure 1.5: Geographic distribution of HBV genotypes and subgenotypes [16, 40].....	12
Figure 1.6: HBx associated cytoplasmic signaling [71].	16
Figure 1.7: Common quantitative mass spectrometry.	28
Figure 1.8: The components of the multiplexed isobaric tagging chemistry [142].	30
Figure 1.9: Example of a typical MS/MS spectrum [142].	31

Chapter 2

Figure 2.1: Feature map of prokaryotic expression vector pGEX-5x-1.	36
Figure 2.2: Feature map of mammalian expression vector pcDNA3.1+.....	36
Figure 2.3: Feature map of pGEM-T Vector circle map.....	37
Figure 2.4: Feature map of mammalian expression vector pXJ40.	37
Figure 2.5: Semi-dry sandwich transfer system (BioRad).	55

Chapter 3

Figure 3.1: Cloning of HBx - A/B/C in pXJ40.....	75
Figure 3.2: Western-blot analysis of HBx genotype A, B and C in transfected cells.	75
Figure 3.3: PCR result of entire Surface antigen (1174bp) for genotyping.....	76
Figure 3.4: Restriction Map created in silico by VectorNTI after sequencing of PCR products from HBsAg.	77
Figure 3.5: Cloning of HBx of Genotype D in pXJ40.	78
Figure 3.6: Western blot result of HBx genotype D.....	78
Figure 3.7: Alignments of amino acids of HBx of genotypes A/B/C/D.....	79
Figure 3.8: Cloning strategy of GST fusion protein proline rich region of HBx on pGEX-5X-1.....	80

Figure 3.9: Expression of Proline/serine rich domain of HBx genotype D on pGEX-5X-1 vector with 6×His.....	82
Figure 3.10: Gradient concentration of standard BSA (2µg/µl) was used for protein quantification.	83
Figure 3.11: Purification of Proline/serine rich domain of HBx genotype D on pGEX-5X-1 vector with 6×His.	83
Figure 3.12: Western Blot Analysis of Expression of GST fusion PXXP and AXXA from genotype D in <i>E.coli</i>	84
Figure 3.13: Panomics SH3 domains arrays interaction with HBx probes containing SH3 binding domain from HBx genotypes A/B/C/D.....	86
Figure 3. 14: Schematic diagram of domain organization of vinexin β and protein sequences.	88
Figure 3.15: PCR products of SH3 binding domain of HBx genotype B and SH3 domain of vinexin β.	89
Figure 3.16: Enzyme digestion results of SH3 binding domain of HBx genotype B (BSH3), left panel, and SH3 domain of vinexin β (VSH3), right panel.....	89
Figure 3.17: Secreted HBsAg detection by IMX.	91
Figure 3.18: Real time RT-PCR analysis of HBcAg expression in HepG2 cells.	92

Chapter 4

Figure 4.1: Transfection of HepG2 cells using Amaxa Nucleofector and monitored by transient transfection of pEGFP vector.	98
Figure 4.2: A representative MS/MS spectrum showing a peptide, HVGMAVAGLLADAR, from PSMA3.	101
Figure 4.3: The differential expression level of several selected proteins.....	102
Figure 4.4: Real-time PCR analysis of mRNA levels of each gene.	103
Figure 4.5 Western blot confirmations of the differentially expressed proteins in HepG2 cells transfected with HBx A/B/C.....	104
Figure 4.6: Fluorescent beads based cell motility assay.....	106
Figure 4.7: Microscopy based real time cell migration assay.	107

Figure 4.8: The overlay of cell migration fluorescence images. 108

Chapter 5

Figure 5.1: A representative MS/MS spectrum showing a peptide,
EQLGPVTQEFWDNLEK, from Apolipoprotein A-I precursor..... 119

Chapter 6

Figure 6.1: Western blot analysis of Rac1, Cdc42 and RhoA expression 125

Figure 6.2: Expression of HBsAg and HBcAg by RT-PCR..... 126

Figure 6.3: Transwell cell motility assay..... 129

Figure 6.4: Graphical representation of transwell cell migration assay. 130

Figure 6.5: Microscopy based real time cell migration assay of HBx activated Rac1..... 132

Figure 6.6: Microscopy based real time cell migration assay of HBx activated Cdc42..... 133

Figure 6.7: Activation status of RhoA by HBx genotypes A/B/C/D. 136

Figure 6.8: Microscopy based real time cell migration assay of HBx activated RhoA..... 137

Chapter 7

Figure 7.1: Hypothetical HBV and HBx regulatory intracellular and secreted proteins..... 143

List of Tables

Table 1.1: Fundamental properties of genomes and differences between HBV genotypes
[39]..... 11

Table 1.2: Proline-rich sequences that bind to specific protein interaction motifs [99].22

Table 2.1: Components of stacking gel and different concentration separating gels.52

Table 3.1: Identified proteins from SH3 binding array.....87

Table 4.1: List of Differentially Expressed Cellular Proteins.....99

Table 5.1: List of Differentially Expressed Extracellular Proteins affected by HBV..... 118

Abbreviations

AHSG	α 2-HS glycoprotein
ATCC	American type culture collection
APS	Ammonium persulfate
Bp	Basepair
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic
ECL	Enhanced chemiluminescence
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetra-acetic acid
ER	Endogenous restriction (sites)
ERKs	Extracellular signal regulated kinases
FAs	Actin filaments
FAK	Focal adhesion kinase
FBS	Fetal bovine serum
GAPs	GTPase activating proteins
GC	Gas chromatography
GEFs	Guanine nucleotide exchange factors
GFP	Green fluorescent protein
GST	Glutathione-S-transferase
HBV	Hepatitis B virus
HBsAg	Hepatitis B virus surface antigen
HBeAg	Hepatitis B virus E antigen
HBcAg	Hepatitis B virus core antigen
HBx	Hepatitis B virus X protein
HCC	Hepatocellular carcinoma

HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HMGB1	High Mobility Group Box 1
hnRNPs	Heterogeneous nuclear ribonucleoproteins
IPTG	Isopropyl- β -D-thiogalactopyranoside
iTRAQ	Isobaric tag for relative and absolute quantitation
Kb	Kilo basepair
kD	Kilo dalton
LC-MS/MS	Liquid chromatography mass spectrometry
MACF1	Microtubule-actin cross-linking factor 1
MEM	Minimal Essential Medium Eagle
MMPs	Matrix metalloproteinases
mRNA	Messenger ribonucleic acid
MTs	Microtubules
MW	Molecular weight
NC	Nitrocellulose
NES	Nuclear export signal
NF- κ B	Nuclear factor κ B
NLS	Nuclear localization signal
OD	Optical density
ORF	Open reading frame
PAK	p21-activating kinase
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PI3K	Phosphatidylinositol 3-kinase
PSMA3	Proteasome subunit alpha type 3
PVDF	Polyvinylidene difluoride
PXXP	Proline-Xaa-Xaa-Proline
RBD	Rho binding domain

RFLP	Restriction fragment length polymerase
R.T.	Room temperature (~ 25 °C)
RT-PCR	Reverse transcriptional polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulfate –polyacrylamide gel electrophoresis
SH3	Src-homology 3
TAE	Tris, acetic acid and EDTA buffer
TEMED	N, N, N, N-tetramethyl-ethylenediamine
2 D	2- Dimensional

List of Publications

1. **Huixing Feng**, Mingxuan Wang and Wei Ning Chen. (2009) iTRAQ-coupled 2D LC-MS/MS analysis of secreted proteome of HBV-replicating HepG2 cells: potential in biomarkers for prognosis of HCC. *Curent Microbiology*, in press.
2. Xi Li, **Huixing Feng**, Wei Ning Chen and Chan Vincent. (2010) Thermal Induced Coupling of Cell Deadhesion and Migration under HBV replication. *Biomaterials*, 31(7):1894-903.
3. **Huixing Feng**, Xi Li, Jianjun Sui, Jianhua Zhang, Dandan Niu, Wei Ning Chen. (2009) Protein Profile in Cells Transfected with Genotype Specific HBx: a Comparative iTRAQ-coupled 2D LC-MS/MS Analysis. *Journal of Proteomics*, in press.
4. **Huixing Feng**, Jianhua Zhang, Xi Li, and Wei Ning Chen. (2009) HBX-Mediated Migration of HBV-Replicating HepG2 Cells: insights on development of hepatocellular carcinoma. *J. Biomed. Biotechnol*, in press.
5. Dandan Niu*, **Huixing Feng*** and Wei Ning Chen. (2009) Proteomics of hepatitis: insights in liver diseases and prognostics biomarkers. *Journal of Proteomics*, in press (* joint first authors)
6. **Huixing Feng**, Tuan Lin Tan, Dandan Niu and Wei Ning Chen. (2009) Hepatitis B Virus X protein interacts with cytoskeletal signaling proteins through SH3 binding. *Front. Biosci*, 1; 2:143-50.

7. Dandan Niu, JianJun Sui, Jianhua Zhang, **Huixing Feng** and Wei Ning Chen. (2009) iTRAQ-coupled 2D LC-MS/MS analysis of cellular protein profile with aberrant HBV-modulated DNA methylation: implications in hepatocarcinogenesis. *Proteomics*, 9(15):3856-68.
8. Dandan Niu, Jianhua Zhang, Yudan Ren, **Huixing Feng**, Wei Ning Chen. (2009) HBx genotype D represses GSTP1 expression and increases the oxidative level and apoptosis in HepG2 cells. *Molecular Oncology*. 3: 67-76.
9. Hong Pan, Dandan Niu, **Huixing Feng**, FP Ng, Ee Chee Ren and Wei Ning Chen. (2007) Identification of cellular transcription factors that bind to hepatitis B virus core promoter with naturally occurring deletions, and analysis of their impact on viral replication. *Biochim. Biophys. Acta – Mol. Basis Dis.* 1772:1075-84

Summary

Hepatitis B virus infection remains a worldwide health concern. A great proportion of the infected individuals carry the risk of developing liver cancer and liver cirrhosis. HBV genotypes have been linked with different clinical disease prognosis.

The x protein of HBV (HBx), exhibits a variety of functions in modulating cell transcription and signal transduction, and possibly causes abnormal cellular changes associated with liver diseases. Nevertheless, the underlying mechanism remains obscure. Our recent investigation has suggested a close interaction between HBx and cellular proteins through SH3 binding. In particular, we have identified a proline rich region located in HBx, and such a motif is also well known as SH3 binding domain. The widely existed SH3 domain was believed to be a pivotal link in the cellular signal transduction intracellularly and between virus and host cells. In our lab's established system, we have identified a number of SH3 domain containing proteins having possible interaction with HBx by using the commercial SH3 arrays. Some of the proteins identified in this assay were cytoskeleton related proteins. However, the consequences of such interactions remain to be elucidated. Here, by the application of 2D coupled LC-MS/MS proteomics analysis, we aimed to identify the HBx-induced protein profile in HepG2 cells. Our results indicated that proteins including cytoskeleton associated

proteins were differentially expressed in HepG2 cells transfected by HBx of various genotypes. These may in turn result in the decrease of focal adhesion and increase of cell migration in response to HBx. The differential protein profile identified in this study was also supported by our functional assay which indicated that cell migration was enhanced by HBx. Our preliminary study provided a new platform to establish a comprehensive cellular protein profile by LC-MS/MS proteomics analysis.

Our recent investigation and findings from this project have strongly implicated the Rho GTPases family of proteins, the key regulators of actin cytoskeletal organization. Deregulation of the Rho GTPases may lead to uncontrolled dynamic cytoskeletal reorganization and metastasis, one common feature of cancer developments. In our recent investigation, HBV viral replication has been shown to be up-regulated by HBV activated Rac1, a member of Rho GTPases family. Here, we showed that HBV replication enhanced cell migration by modulating cellular proteins via the PXXP motif in HBx. Significantly, genotype specific HBx showed differential activation of RhoA, another member of Rho GTPases. And the activation was supported by the microscopy based real-time migration assay.

Taken together, our work established a useful sequence-function platform to analyse HBV-host interactions. In particular, the changes in protein profile influenced by HBx expression, using 2D coupled iTRAQ LC-MS/MS based

quantitative proteomics, allowed the identification of cellular proteins which may well be the downstream effectors of HBV-host interactions via SH3 binding. A series of differentially expressed cellular cytoskeletal proteins identified in this work were highly associated with cell migration, and thus providing new mechanistic insights on our cell observation of changes in cellular adhesion/deadhesion and migration under the influence of HBV/HBx. It is to be hoped that our platform can be expanded into further functional analyses for the comprehensive understanding of the underlying mechanism of the involvement of HBx in HCC, as well as on HBV-associated HCC.

Chapter 1 : Introduction

1.1 Hepatitis B virus and epidemiology

More than thirty years after the discovery, Hepatitis B is still a global health concern with significant associated causalities. It is estimated that 2 billion people have been contracted this virus and there remains to be more than 350 million chronic carriers worldwide [1]. About 15% to 40% of the infected people will end in liver cirrhosis, liver failure, and primary liver cancer (hepatocellular carcinoma or HCC) [2]. Chronic hepatitis B virus (HBV) infection is causative of HCC. Annually, around 620,000 million people died from suffering HBV related cirrhosis or HCC. Chronic HBV carriers have more than 100-fold increased risk of the development of HCC [3].

The prevalence of HBV is geographically related. As shown in Fig. 1.1 left panel, in the areas of Far East, Africa, the Middle East and some parts of South America, the HBsAg rates is very high, varying from 8% to 15% [4]. In contrast, in Northern Europe, Canada, Australia and some parts of South Africa, the prevalence is less than 2% [5]. Moreover, the global distributions of HBV infection and HCC show close correlation. It is found that HBsAg prevalence in HCC incidence can reach up to 85% (Fig. 1.1).

Figure 1.1 : Global distribution of HBV infection and incidence of HCC [6].

1.2 HBV vaccination and clinical treatments

One reason for the setting of chronic infections is that HBV causes non cytopathic infections of the hepatocytes, which is the main cell type (up to 70%) of the liver. And the hepatocytes have a life span estimated 6-12 months or even longer. In addition, the virus was shed into the bloodstream by the hepatocytes, which then infect other hepatocytes until all the hepatocytes are infected. When the immune system is not strong enough to clear the virus, persistent infection is established under the two characteristics, longevity of the hepatocytes and stable cell-virus interaction. It was hypothesized that the transient and chronic of HBV infection is also due to the response of host's immune system. The immune system can induce strong hepatocytes destruction, which will lead to various tissue damages [7] [8]. The consequences of the response of immune system make little contribution to the clearance of the virus, but instead, cause destruction of the hepatocytes. Surprisingly, the elimination of the virus sometime occurs spontaneously many years after the virus production. What has happened between the virus and the host

cells? Why such immune response is turned on? Why sometimes the infection is persistent, while other time is terminate? To ravel these questions, a comprehensive understanding of the virus host cell interactions is needed. It is hoped that the development of the knowledge of the interactions between HBV and host cells' proteins and signal pathways, would lead to better antiviral strategy against the infection and subsequently reduced the development rate of HCC.

The vaccination of HBV has been made universal for than 20 years. It has been estimated to be 95% effective for the preventing of HBV chronic infection. The first nationwide vaccination of HBV infection has been launched in Taiwan in 1984 [9-11]. Unlike some other vaccinations, the protection by HBV vaccination is time limiting. For most of the people, boosters are needed every several years.

Although the vaccination is the best way to prevent the transmission of HBV infection, it has no effect on chronic HBV carriers. For these individuals, treatment must be applied to avoiding long term liver disease. Up to now, two therapies are widely used globally for the treatment of chronic HBV: conventional interferon alpha (IFN α) and lamivudine [6]. Some of the patients can have improved disease outcome and prognosis [12]. However, alpha interferon, ineffective in more than 50% carriers [13]. In addition, both of treatments display undesirable side effects. Overtime, the drug resistance due to the mutation of the virus will allow virus to proliferate even in the presence of the drug. Some other new drugs such as adefovir (has been approved in the USA and in Europe), entecavir [14] and

pegylated IFN α are to be launched. But the promising applications of the new drugs still need to be tested in the large scale, multicenter clinical trials. Long term potentials of these recent drugs are to be seen.

Due to the shortcomings of the drugs and viral resistance generated by the mutation in the virus in escaping with the treatment, HBV remains a worldwide healthy problem with considerable mortality. Thus, continuous effort in depicting the molecular mechanisms of HBV should provide new insights in improving antiviral strategy.

1.2 Structure and genomic organization of HBV

HBV belongs to the family of hepadnaviridae, which are composed of hepatotropic DNA viruses infecting selected avian or mammalian hosts such as woodchuck, squirrel, duck and human. Structurally, these viruses show the following typical characteristics: (1) they are partially double stranded DNA genome surrounded by an inner nucleocapsid or core protein and an outer lipoprotein envelope. (2) The polymerase which is coded by the viruses themselves has reverse transcription activity. (3) Viral envelope proteins, such as HBsAg of HBV, are overproduced. (4) Relatively hepatotropism [15].

There are three kinds of particles in the blood serum, the most important one is the

infectious Dane particle, which has a size of 42-47nm. The other two particles are much smaller, 20nm spheres and 20nm filamentous particle. All the three particles contain the hepatitis B surface antigen (HBsAg). Normally, HBV virion means the Dane particle. It is double-shelled particle. There are three kinds of envelope proteins, namely the large (L), the middle (M), and the small (S), and all are encoded by the surface ORF. The major protein of the surface is S protein, which is commonly known as hepatitis B surface antigen (HBsAg). The other two surface proteins, L is involved in the surface binding to the receptor of cell, while the function of M is not very clear.

Figure 1.2: HBV virion (Dane particle) and genomic organization [15].

HBV genome carries four open reading frames (ORF). They encode four major proteins: surface (S), polymerase (P), core (C), and X protein (X) respectively. There are three types of surface proteins, L (large), M (Middle), and S (small).

As shown in Fig. 1.2, inside the surface proteins of the Dane particle, there is

another shell which is consisted of the nucleocapsid with 25-27 nm in diameter. It is essentially composed of core protein. The double layers shells, contains the partially double-stranded DNA of the virus.

The DNA genome of HBV is 3.2kb and has all the typical characteristics of hepadnaviridae as mentioned above [16]. The HBV genome is relaxed circular and partially double- stranded DNA. An RNA intermediate is produced during the reverse transcription process of HBV replication, referred to as the pregenomic RNA which is around 3.5 kb. By the tiny size of the genome, the virus is capable of possessing promoter and enhancer elements essential for virus survival and the overlapping coding regions can provide all the information for the proteins transcription. It has four open reading frames that encode four proteins respectively, core, polymerase, surface, and HBx protein. The core and polymerase are essential for the viral DNA replication. The envelope proteins come from the surface gene. On the other hand, HBx is multifunctional non-structural regulatory protein.

As indicated in Fig. 1.3, the DR1 and DR2 are two Direct Repeats of 11 nucleotides sequences. These two short repeat sequences are essential for the beginning of the synthesis of two DNA strands, the minus (-) and plus (+) strands.

Figure 1.3: Organization of the HBV genome.

Four major viral genes (preC/C, preS1/preS2/S, P and X) are found in HBV genome. The capsid protein encoded by the core (C) gene, the viral polymerase by P gene, X protein by X gene. In addition, three envelope proteins are translated by a single ORF with three start codons in the same open reading frame. <http://www.hixonparvo.info/parvovirus.html>

1.3 The replication of HBV

Largely due to the lack of cell lines which support HBV infection, the early stages of the viral infection, such as receptor binding, uncoating of virus and nuclear transfer of the DNA genome can not be well analyzed.

The replication cycle of HBV starts with the synthesis of a ~3.2 kb relaxed circular DNA genome (partially double stranded) by using the reverse transcription of a pregenomic RNA intermediate. The synthesis of genome DNA is directed by RNA [17]. The process of HBV replication can be broadly separated into three phases (Fig. 1.4): the first event is the formation of a supercoiled plasmid like covalently closed circular (cccDNA) from relaxed circular DNA (rcDNA); second, the expression of the virus genes by using the pgRNAs as templates; the third, the assembly and reverse transcription [8].

During the initiation of the virus infection, the first event is the covalently closed circular DNA (cccDNA) formation from the relaxed circular DNA (rcDNA) and reverse transcriptase bound to the 5' end of minus (-) strand. Concurrently, a short oligonucleotide RNA is attached to the 5' end of the plus (+) strand. The formation of cccDNA indicates the successful virus infection, since the cccDNA will be used as the template for the production of the pregenome.

With the cccDNA, the following RNAs are formed: the longest pregenomic RNA (3.5 kb) and other subgenomic RNA with varying lengths of 2.4 kb, 2.1kb and 0.7 kb. The RNAs serves as the templates for the synthesis of all the HBV proteins, respectively. The pregenomic RNA serves as the RNA templates for the synthesis of viral polymerase and core protein (nucleocapsid subunit) encoded by different open reading frame (ORF). The shortest 0.7 kb RNA is responsible for the code of X protein. X protein is believed to play a pivotal role in the virus-host cell

interactions. The study of gene expression control showed that each gene of the HBV has one or more promoters which regulate the gene expression. In addition, the promoters are regulated by enhancers which are located at the upstream of the core promoter. The regulation also occurs at the translation level, since the pregenomic is more than the template of core protein but also the template of viral polymerase.

The next crucial step is the package of pregenomic RNA and reverse transcriptase into the newly formed viral core particles. The production of the polymerase coupled with the packaging of pgRNA into the nucleocapsids are closely associated [18-20].

Figure 1.4: HBV replication cycle [21].

HBV infects hepatocytes and releases supercoiled cccDNA. HBV DNA replicates by reverse transcription via a 3.5-kb-pair pregenomic RNA intermediate. HBcAg and polymerase are subsequently produced from the pregenomic RNA template, while HBsAg

and X protein are translated from smaller RNA species [21].

1.4 Hepatitis B virus genotypes and disease prognosis

There are eight genotypes of HBV. They have been designated from A to H based on nucleotide sequences differences which is greater than 8% over the entire genome between any two distinct genotypes [22-26]. The conventional determination by serological subtype is based on the immune reactivity of an antibody to a number of antigenic amino acids in the major surface antigen HBsAg. The eight genotypes are distributed in different geographical locations as shown in Fig. 1.5 [27]. Different genotypes have slightly variations in genome size as shown in Table 1.1. Genotype A is mainly discovered in North America and Northwest Europe, with less presence in East and South Africa, Hong Kong and Philippines [24, 28, 29]. Genotypes B and C are prevalent in Southeast Asia, and also appear in the Pacific islands [22, 30, 31]. Genotype D distributed most widely, spreading from Southern Europe and India to North, South and West Africa [24, 29, 32] and among intravenous drug takers all over the world [33]. Genotype E is the most similar genotype to Genotype D [24, 33, 34], and is essentially found in West Africa. The most divergent genotype F, can be found in South and Central America [35, 36]. For Genotype G, there is little information about its geographic distribution. Genotype G has only been found in the USA and France [25, 37, 38]. The last genotype up to now, genotype H, seems to be an American variant of genotype F [26, 36].

HBV genotypes are associated with different clinical outcomes and prognosis. In the case of chronic HBV infection, genotype C is more relevant to liver cirrhosis and genotype B is more relevant to HCC. It has also been reported that compared with genotype A, Genotype D and F are linked with more severe liver diseases and the prognosis is less promising.

Although the genotypic and serological classifications of HBV have been observed and summarized in recent years, the association between genotypes and stages of chronic HBV infection remained unclear. Molecular characterization of this association would help gain further information for aetiological, clinical, virological, and anthropological investigations. Findings from such an analysis may help us to a better understanding of the relationship between HBV infections and different prognosis.

Table 1.1: Fundamental properties of genomes and differences between HBV genotypes [39].

Genotype	Genome length in bp	ORF-differences
A	3221	Insertion of aa 153 and 154 in HBc
B	3215	
C	3215	
D	3182	Deletion of aa 1-11 in preS1
E	3212	Deletion of aa 11 in preS1
F	3215	
G	3248	Insertion of 12 aa in HBc
H		Deletion of aa 11 in preS1
I	3215

Subtypes	Genotype	Geographic distribution
ayw 1 adw 2	A	Northern Europe, United States
ayw 1 adw	B	Eastern Asia, Far East
ayr adr	C	Eastern Asia, Far East
adrq ayw2 ayw3	D	Worldwide; prevalent in Mediterranean area, ayw2 near and Middle East, South Asia
ayw4	E	Western sub-saharan areas
ayw4q- adw2	F G	United States, Africa Recently isolated in France and United States

Figure 1.5: Geographic distribution of HBV genotypes and subgenotypes [15, 39].

1.5 Hepatitis B virus X protein (HBx)

HBx, the smallest HBV protein with 154 amino acids (17.5 kDa), is a versatile protein that modulates the cell transcription, signaling pathways, cell cycle progression, apoptosis, and protein degradation through direct or indirect interaction with host cell's signal proteins or other factors [8]. HBx was not shown to bind to the DNA directly in the past research, and the interaction between HBx

and the cell is through a series of protein-protein interactions. HBx is able to activate the transcription of a variety of viral and cellular genes [40]. As HBx is found to be insoluble *in vitro*, the X-ray crystallization structure and nuclear magnetic resonance (NMR) information can not be obtained, and the three-dimensional structure is not available up to now. Sequence analysis showed that HBx is relatively highly conserved between different species of the mammalian hepadnaviruses.

1.5.1 Expression of HBx

HBx ORF is located downstream of Enhancer I and overlapped by the PreC ORF at its C- terminal, and Polymerase ORF at its N- terminal part. HBx is transcribed from a dependent RNA transcript, the 0.7 kb transcript and the transcription is controlled by Enh I and Promoter X. HBx is under autoregulation since activation of Enh I and X promote can be augmented by HBx.

1.5.2 Subcellular localization of HBx

Most studies revealed that HBx mainly located in the cytoplasm but also can be detected in the nucleus [41-47]. HBx was also showed mainly located at nucleus in some cell lines [48, 49]. Recently, the localization of HBx is believed to show an expression quantity dependent characteristic. When it is expressed at low level, it is predominantly found in the nucleus, whereas when expression level increases, it is mainly found in the cytoplasm [44, 50]. Recently, HBx is also reported to be

localized at mitochondria [44, 51-56]. The presence of HBx in mitochondria may lead to the loss of membrane potential which in turn will induce mitochondrial dependent cell death [53-55]. Such a location may also end in the increase of cytosolic calcium level [52]. The dynamic distribution of HBx may imply the multiple functions of HBx.

1.5.3 HBx and HBV replication

The exact function of HBx in HBV replication remains a long-debate issue. First, HBx has been shown not to affect viral replication in Huh7 hepatoma cell line and is not critical for the viral replication in transgenic mice, however it was also shown the ability of enhancing the replication of virus via the intermediate of other HBV gene expressions [57-59]. On the other hand, in the woodchuck model, HBx is required for persistent infection [60, 61]. Similar result has been recapitulated in human hepablastoma cell line HepG2 [62-64].

It is unclear whether the localization of HBx in the cells is related to the stimulation of viral genome replication. The cytosolic HBx promote a high level of viral replication by the activation of Src kinase [65]. This finding is supported by the experiment demonstrating that HBx fails to stimulate viral replication when it a nuclear localization signal (NLS) is included in HBx [65]. Work from other groups shows that the NLS-HBx can stimulate viral replication while the NES (nuclear export signal)-HBx can not. Another group shows that the increase in viral

replication by HBx is on the transcriptional level, mutation analysis is then confirmed that the two-thirds of the C-terminal region (amino acids 51 to 154), which contains the transactivation domain, is needed [66]. The most recent report has demonstrated that both nuclear and cytoplasmic HBx are linked with the stimulation of viral replication. Although the mechanism is not clear, it has been shown that HBx linked to either NLS or NES is able to stimulate viral replication to a comparable level [67]. Collectively, these studies suggest that HBx contributes significantly to HBV replication.

1.5.4 HBx and cellular signaling pathways

HBx has been shown to stimulate transcription of a series of viral and cellular transcription factors [68]. Fig. 1.6 shows the cellular signal pathways influenced by HBx. HBx does not bind DNA directly, but it plays a role of transcriptional activator by interacting with nuclear transcriptional components and factors, including TFIIB, TFIIH, TBP (TATA-binding protein). HBx also directly interacts with one of the transcription factors CREB (cAMP response element-binding protein), a member of the leucine zipper family. The addition of NLS to HBx has facilitated the study of interaction between HBx and nuclear transcriptional machinery. Results have demonstrated that NLX-HBx enhances the stimulation of transcription from HBV Enh I, but fails to activate transcription dependent upon AP-1 (activating protein-1), and NF- κ B (nuclear factor κ B) [43, 69].

Figure 1.6: HBx associated cytoplasmic signaling [70].

Arrows represent targets of HBx-mediated activities. Filled arrows represent established primary targets of HBx, whereas dotted arrows are their likely targets [70]. HBx binds to CREB directly and they bind to transcription factors such as TFIIB and TFIH. In mitochondria, HBx causes calcium release and in turn activates Pyk2/FAK, Src and Ras-Raf-MAPK-JNK signal pathway. The stimulation of proteasome by HBx has also been reported.

HBx is believed to be primarily localized in cytoplasm when its expression level is high. HBx interacts with a wide range of protein kinases and consequently stimulates cytoplasmic signal pathways. HBx has been reported to stimulate the extracellular signal regulated kinases (ERKs), p38 kinase and the stress-activated protein kinases/HN2-terminal-Jun kinases (SAPK/JNKs) [69, 71-73]. And the activations can be sustained to various extents in a cell-type dependent manner.

HBx is shown to up-regulate the RAS-RAF-MAPK signal cascade [74-78]. This

transduction activation is associated with de-regulation of cell cycle check point controls leading to early entry of cells into S phase, which may result in severe damages to the cells. The RAS-RAF-MAPK signal pathway activation by HBx has thus been associated with cell transformation. Up-regulation of RAS-RAF-MAPK signal pathway activates NF- κ B and AP-1 transcription factors [43] and increase the production of TBP inducing RNA polymerase II dependent gene transcription.

HBx up-regulates Src kinase which is an upstream activator of Ras GTPases. Subsequently Src is tyrosine phosphorylated and ultimately activates Ras. The activation of Src by HBx is indirect and no direct interaction has been observed between Src and HBx [65]. The activation of Src not only increases the activity of Ras, but also affects replication of HBV by stimulating the polymerase. It has been shown that HBx activates the nuclear factor of activated T cells (NF-AT) [79] which is regulated by cytosolic calcium signaling. It was later demonstrated by Schneider's group that HBx activates the focal adhesion kinase (FAK) and the cytosolic calcium-dependent proline-rich tyrosine kinase-2 (Pyk2), a Src kinase activator. Interestingly, the inhibition of Pyk2 or calcium signaling would block the replication activated by HBx [63]. The calcium signaling in HBx activation is importance and this has separately been confirmed also in the activation of STAT3 and MAPK p38 which is also directly related to Src kinase [73]. Furthermore, the activation of Src by HBx promotes the disruption of adherent junctions [80].

Be part of the Rel proteins family, transcription factor NF- κ B is ubiquitous in the

cell [81]. NF- κ B is usually localized in the cytoplasm and binds to its inhibitor such as I κ B α and I κ B β . And the migration to the nucleus of NF- κ B will be blocked by such bindings. External stimulations, such as viruses, IL-1 and TNF stimulate IKK- α and IKK- β which will let the inhibitors phosphorylated resulting in the degradation and the release of NF- κ B. The released NF- κ B then moves into the nucleus and exhibits transcriptional activities. NF- κ B activation pathway prevents the apoptosis of hepatocytes and is benefit for the liver generation. NF- κ B is one of the earliest identified HBx responsive proteins and the activation of NF- κ B by HBx is observed [82, 83]. Previous studies showed that all major forms of NF- κ B can interact with HBx. The expression of cytosolic HBx triggers the release of NF- κ B by phosphorylating the then degrading the inhibitors [74, 84]. HBx directly interacts with I κ B α and I κ B ϵ , causing entry of NF- κ B and HBx into nucleus [47]. The migration into nucleus and performing the transcriptional activities may in turn facilitate the viral replication.

1.5.5 HBx and apoptosis

Apoptosis, also called programmed cell death, is a crucial process to maintain the homeostasis of the environment in a wide range of organisms. Caspases which are cysteine proteases, form a cascade of effectors from activation to execution of apoptosis pathway. The relationship between HBx and apoptosis has been controversial. Interestingly, HBx has been reported to promote apoptosis, or prevent apoptosis, make the cells sensitive to proapoptotic stimuli and have no impact on apoptosis.

HBx was shown to inhibit the function of tumor suppressor protein p53, which is highly related with apoptosis. It has been reported that HBx binds to p53 and resulted in decrease of p53 in the nuclei which led to the inhibition of hepatocytes apoptosis. The down regulation of the apoptosis may well be the main reason for the progression of liver cancer [85]. In support of this, the over-expressed HBx has been found to alter mitochondrial membrane permeability of Ca^{2+} and block the caspase 3 activities. The protease is associated with the cleavage of the Ca^{2+} ATPase on plasma membrane. The ectopic expressed HBx thus disturbs the homeostasis of intracellular Ca^{2+} by blocking caspase 3 activities and finally inhibits apoptosis.

1.5.6 HBx and HCC

HBx displays numerous functions and is implicated in many cellular processes. Naturally, its role in HBV-associated HCC has been extensively investigated. HBx expression facilitates HCC development in HBx transgenic mice, but it is not clear whether HBx directly induces tumor formation or just makes the mice sensitive to other carcinogenic factors [86-90]. Significantly, chronic infections with mammalian HBV containing X protein lead to HCC development, while chronic infections with duck HBV (DHBV) which lacks X protein do not result in HCC.

1.6 Proline-rich domain in HBx and SH3-mediated Signaling

For the understanding of the cellular signal transduction, a key is to clarify protein-protein interaction. Such interactions can be identified and characterized by various techniques including cDNA library screening, yeast two hybrid, co-IP (immunoprecipitation), mutagenesis and truncation. Protein-protein interactions can occur through noncatalytic, conserved small domains as short as 10 amino acids in a single protein. Among the primary sequence of the interacting motifs, proline is a rather special one. Importantly, the well defined SH3 domain, WW domain and several other domains specifically recognize ligands which are rich in proline residues.

Src, SH3 (Src Homology 3) domain, as part of the Rous sarcoma oncogene product, plays an important role in signal transduction. SH3 domain is typically a small, conserved sequence with 60 amino acids. It interacts with proline-rich binding signature motifs known as SH3 binding domains. These domains usually are consisted by 6 to 12 amino acids, with a conserved Proline-X-X-Proline (PXXP) motif [91], where Xaa indicates any amino acid.

SH3 domains can be considered as an adaptor molecule, and function by attracting downstream proteins in a signaling pathway cascade. They facilitate the assembly of signaling complexes, such as epidermal growth factor receptor signaling [92], cellular localization of cytoplasmic proteins [93], and activation of PI3K in response to IgM crosslinking [94]. In the context of viral infections, SH3 domains are also believed to be pivotal links between the virus and the host cells. Further

more, SH3 domains have been implicated in cancer development. It has been reported that Nef, the HIV viral protein, interacts with the p21-activating kinase (PAK1) via its SH3 binding domain and viral replication is enhanced as a result. It is also believed that NS5A, non-structural protein of HCV, interacts through its C-terminal PXXP motif with the cellular linker protein Grb2 [31], resulting in a reduction of the mitogenic signaling pathway and inhibition of the activating protein 1 (AP-1). As for HBV, it has been shown that HBx activates SH3 domain containing signal protein c-Src and the activation although no direct protein-protein interaction motif has been identified [95]. As a result of the wide range of interacting targets, the SH3 domain holds the great promise in the finding of new therapeutic agents that are particularly specific to a certain signaling pathway. SH3 domain act could be inhibited by the dissociation of its binding activity [96, 97].

Another proline rich binding motif which has similar binding preference with SH3 domain is known as WW domain. WW domains are composed of around 40 amino acids which fold as triple stranded β -sheet and they recognize proline/serine rich ligands. SH3 and WW binding ligands share similar but not identical characteristics. The typical peptide ligand motifs are shown in the Table 1.2. It is apparent to find that generally the ligands of SH3 domain generally have a core PXXP sequence while the ligands of WW are more complicated with four classes displaying minor differences. Recently, HBx has been found to interact with Pin1, which contains a classic class IV WW domains [98]. Based on the sequencing

alignment, it is likely that HBx genotype A activates Pin1, through the WW domain.

Table 1.2: Proline-rich sequences that bind to specific protein interaction motifs [99].

<u>Proteins with SH3 domains</u>	Peptide ligand motif	Reference
Src	RPLPPLP	[100-103]
	PPVPPR	
Yes	RxLPxLP	[104]
Lyn	RxxRPLPPLPxP	[101]
Abl	PPPψPPPPψP	[101]
PI3K	RxxRPLPPLPP	[101]
PLCγ	PPVPPRP	[104]
Cortactin	+PPψPψKP	[104]
p53BP2	RPxψPψR	[104]
Grb2 A	PLPxLP	[104]
Crk A	PxLPx(K/R)	[105]
Amphiphysin I	PxRPx (R/H) (R/H)	[106]
Nck, SH3 B	PxPPxRxxSL	[107]
CAP, SH3 C	PxPPxRxSSL	[108]
<u>Proteins with WW domains</u>		
Class I WW's		
YAP65	PPxY	[109]
Nedd-4	PPxY	[110]
Dystrophin	PPxY	[111]
Class II WW's		
FBPs	PPPPPPL/R	[112]
FE65	PPPPPPL/R	[113]
Class III WW's		
FBPs	(PxxGMxPP)2*	[114]
Class IV WW's		
Essl/Pin1	(pS/pT)P	[115-117]
Nedd-4	(pS/pT)P	[116]
<u>Other</u>		
<u>Proteins</u>		
Mena (EVH1)	(D/E)FLPPPP	[118]
Homer (EVH1)	PPxxFr	[119]
CD2-BP (BHB)	(PPPGHR)2*	[120]
Profilin	G/L PPPPPP(PP)	[121]

Residues labeled ψ, +, and x correspond to aliphatic (A, I, L, V, and P), positively charged

(R, K, or H), and any residues, respectively. Lowercase letters refer to positions that are often (but not always) a certain residue. pS and pT refer to phosphoserine and phosphothreonine, respectively.

Our laboratory has recently reported HBx contains a proline/serine motif which appears to interact with a number of SH3-binding cellular proteins by SH3 Domain Arrays [122].

1.7 HBV and cell migration

HBV infection has been shown essential to the development of HCC. The infection of HBV is related with series of signal transduction and transcriptional factors. The HBV-mediated activation of diverse signaling pathways may lead not only to the activation of transcription factors, but also to cellular changes which may contribute to transformation and tumor development. One example is the acquisition of metastatic HCC which is frequently observed at the early stages of HCC and may cause liver failure and ultimately death [123, 124].

Deregulation of cell adhesion and uncontrolled high motility are hallmarks of tumor progression and cell metastasis. Metastasis is a major cause of death among cancers. A number of proteolytic enzymes participate in the processes of cell invasion and metastasis to degrade noticeably the extracellular matrix (ECM) [125]. Thus, the abnormal cytoskeletal reorganization could be a sign of tumor development and cell metastasis.

It has been found the infection of HBV shows to increase the invasiveness of HCC cells through the transcriptional activation of Matrix metalloproteinase-2 (MMP-2), MMP-3 and MMP-9 [126-128]. HBx is also found to induce cell migration in a CD44-dependent manner [129]. Such activation is mediated by PI-3K-Akt/ protein kinase B (PKB) and extracellular signal-regulated kinase 1/2 (ERK 1/2). In addition, altered Rho GTPase signaling is implicated in the initiation and the progression of HCC [130]. Rho downstream effector Rho kinase 2 (ROCK2) is frequently over-expressed in HCC and the knockout of ROCK2 influences the formation of stress fibers, filopodia, and lamellipodia [131].

1.8 HBV and Rho GTPase-mediated Signaling

Rho GTPases (guanine nucleotide triphosphatases) are a family of signaling proteins which function like molecular switches. They play important roles in controlling various signal transduction pathways; especially in regulating the cellular actin cytoskeleton. They cycle between two conformational states to realize the functions: GTP bound active form and GDP bound inactive form. These GTPases are able to hydrolyze GTP to GDP either by the intrinsic activity or by GTPase-activating proteins (GAPs). When they are bound to GTP, the Rho GTPases interact with their target proteins. To activate the downstream signaling pathway, the activation is turned off when Rho GTPases are bound with GDP. With this controlling strategy, the Rho GTPases participate in a series of cellular

functions such as cell polarity, cell migration, intracellular trafficking, cell cycle and transcription [132].

There are another two protein families named GTPase activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs) related to Rho GTPases. The function of GEFs is to activate Rho GTPases by catalyzing the exchange of the binding of GDP for GTP, while the function of GAPs is to accelerate the intrinsic GTPases activity, which means the GAPs are the inactivators of the Rho GTPases [133].

Viral infections can induce filopodia and lamellipodia formation. The main purpose is to facilitate the transport of progeny viral proteins along the polymerized cytoskeleton, before their assembly and budding out of the cells. However, the sustained formation of protruding cellular structure may lead to the onset of abnormal state and in some cases cancer development. Small Rho GTPases including Rac1, RhoA and Cdc42 regulate the formation of these cytoskeleton structures [134]. Rac1 induces lamellipodia formation and membrane ruffling while Cdc42 induces filopodia formation. RhoA induces stress fibers formation [18, 135-137].

Previous work has linked Rho GTPases with viral replication and cytoskeleton changes. It has been shown that by the activation of HIV envelope protein and co-receptor, Rac1 stimulates actin filament and facilitate membrane fusion with

host cells [138]. In the previous work of our lab, the expression of HBx in transfected cell activated the endogenous Rac1 and induced membrane ruffling. Significantly, the interaction between HBV and Rac1 led to cell morphological changes through the Rac1 signal transduction pathway [139].

1.9 iTRAQ coupled 2-D LC-MS/MS based Proteome profiling

In the past few years, systems approach to discover new biomarkers of cancers has arisen. The main purpose is to discover and validate biosignatures for the classification of cancers. Combined with clinical disease outcomes, these signatures can be associated with treatment outcome. With proper validation, these biomarkers can be used to guide the treatment of patients with the similar biomarker patterns. Biomarkers include genomic changes, polymorphisms and proteomic changes. In addition, biomarkers can be used for early detection of onset of cancer development. The latter refers to differential protein expression or modification in healthy or diseased tissues of patients.

Proteomics is the study of a large scale of proteins. The proteome of a cell is a dynamic entity which changes with the different physiological states of the cell. As opposed to the transcriptome in DNA microarray, the proteomics analysis aims to understand the end state production of specific proteins with response to the external stimuli.

In the context of HBV-associated liver diseases, proteomics has recently been applied for the understanding of the disease states such as cirrhosis and especially hepatocellular carcinoma [140].

In proteomics research, the quantification of differences between more than one related samples or physiological states in a same biological system is a critical but challenging technical task.

Traditionally, two-dimensional electrophoresis (2DE) is a broadly used method for the separation of proteins first by isoelectric point followed by molecular weight. The technique often forms the first screening protocol for the proteomics. It has been first introduced by O'Farrell 1975 and J. Klose in the same year. This technique allows resolution of 1000s of proteins and has since been refined to generate highly reproducible proteome maps by the advent of immobilized pH gradient (IPG) gel strips.

The 2-D based proteomics analysis has been useful in the early stage of biomarkers identification as well as drug discovery. New development of dyes, fluorophores have led to improvements in sensitivity and dynamic range associated with 2-D gel based proteomics. However, there are obvious shortcomings. First, the high resolution protein separation is required. Second, 2-D gel-based analysis is unable to reveal the identity of underlying protein. In addition, the reproducibility of protein spots on 2D gels remains a challenge over time. These problems can be

solved by the application of the increasingly popular LC-MS/MS techniques.

These LC-MS/MS techniques are based on mass-spectrometry quantification methods. Commonly, the cells or tissues targets to be analyzed are firstly purified and the purified proteins were broke into peptides by gentle trypsinization, and finally the peptides were applied to MS analysis as shown in Fig. 1.7.

Figure 1.7: Common quantitative mass spectrometry.

In recent years, stable isotope labeling chemistries combined with MS have been developed. Differential stable isotope labeling has been used to generate a specific mass tag which can be recognized by the mass spectrometer while providing the quantification. These tag 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) [141]. ICAT method involves derivatization of cysteine residues with a reagent containing eight zero or eight deuterium atoms, followed by MS analysis. One of the advantages of ICAT is that since cysteine is a rare amino acid, such a label method can reduce the complexity of samples. ICAT is not appropriate for the proteins that do not have cysteines. In addition, it has also limitation in the analysis of post-translational modified proteins.

Another widely used approach for proteomics and biomarkers researches is the iTRAQ method [141]. iTRAQ, or short for isobaric tags for relative and absolute quantification, has been used for the identification of cancer and other diseases biomarkers [139, 142-147]. The complete molecule is composed of two parts, isobaric tag and reactive group. In the part of isobaric tag, there are also two parts, balance group and reporter group. The total mass of the isobaric tag is 145, and the mass of reporter group is 114-117, while the mass of balance group is 31-28, which is used to maintain the total mass to be 145 (Fig. 1.8). The current commercial iTRAQ reagents contain four or eight labels in the kit which can provide the labeling of up to four samples or eight samples in one experiment respectively. The protein samples are digested into short peptides by trypsin and fragmented peptides are labeled with different isobaric tags. The labeled peptides then was mixed and analyzed in a tandem mass spectrometer. During the fragmentation by tandem MS, the balance group and reporter group is separated by cleavage the bond between them and only the balance group remained bond to the peptide. Meanwhile, the

released reporter groups can give a defined MS signal. Since all reporter groups of isobaric labels share the same physicochemical parameters excepting the mass [144], the differential abundance of the reporter groups are proportional to the separately labeled peptides from each sample. In the tandem MS analysis, with a reagent labeled, the peak area in the spectrum is proportional to the relative abundance of the peptide which is also proportional to the original protein. Four differentially labeled samples were mixed first and after application on MS/MS, the relative concentration between different isobaric reporter ions can be obtained to show the peptide abundance in different samples. Fig 1.9 shows a typical Mass spectrum of a single peptide prepared by labeling four separate digests with each of the four isobaric reagents, and combined the reaction in 1:1:1:1 ratio. In ii) of Fig. 1.9, the four isobaric labels show the same abundance of the peak and it is consistent with the labeling ratio of mixture.

Figure 1.8: The components of the multiplexed isobaric tagging chemistry [141].

Figure 1.9: Example of a typical MS/MS spectrum [141].

It shows the MS/MS spectrum of one peptide. After digestion, the peptide was labeled separately by four isobaric reagents at the ratio of 1:1:1:1. The labeled peptides were mixed together and analyzed by the MS/MS [141].

Mass spectrometry and gas chromatography (GC/MS) have been extensively used to identify and quantify relatively small molecules (with molecular mass less than 1,000 Da). Over the past few years, this technology has been increasingly used for the comprehensive analysis of cellular components. Development of new technologies have been reported to effectively make macromolecules volatilized and ionized, using various matrices and lasers (MALDI) or electrospray ionization (ESI) [148]. High resolution of spectra display is offered by measuring the accurate mass-to-charge ratio [139]. The tandem mass spectrometry facilitates data acquisition (protein sequence) and transformed the field of proteomics [139]. This has been further strengthened by the coupling of mass spectrometers with high performance liquid chromatography (HPLC) [139]. In MS/MS, precursor ions

are labeled in full-scan mode (all m/z values) [139]. This is followed by selective ion isolation and fragmentation for sequence identification [139]. Ions generated in solution from ESI make them suitable for coupling to LC. In addition, a reverse-phase HPLC setup allows the column to slow down the migration of peptides based on selective hydrophobic interaction affinities [139]. Peptides can then be eluted with a gradient of organic solvent, such as acetonitrile. They are then ionized before being injected into the mass spectrometer. In General, the quality of proteomics profiling needs to be sensitive, display wide coverage, possess efficient protein fragmentation, shows high mass resolution and reproducible [139]. These needs have been met by LC-MS, as femtomole or even lower detection limits can be reached even in the setting of complex cellular mixtures.

1.10 Objectives of project

The objectives of this project are to identify and characterize cellular proteins involved in cellular behavior in response to HBx expression or HBV replication. Our recent investigation has revealed a proline/serine rich domain in HBx which may play functional role in HBV/HBx-cell interactions in particular in cell morphology. Changes in cell morphology have wide impact on cellular behavior including cell motility. To this end, genotype specific non-structural protein HBx will be isolated from clinical samples, determined by Restriction Fragment Length Polymorphism (RFLP), and cloned into mammalian expression vector.

As an extension of our recent work, other genotypes of HBx will be used to screen for cellular proteins which interact with HBx via SH3 binding. We will further analyze the downstream events of the HBx-cell interactions by established global protein profile, using proteomics analysis. Such an analysis should generate expression profiles of proteome between untransfected cells and those transfected with HBx. Information from the differential protein profiles would shed new light on the underlying mechanisms of disease development in relation to a particular genotype. In addition, the secreted proteins from the HBV replicating cells will also be analyzed which may identify biomarkers for the development of HBV related liver diseases. This is particularly relevant as chronic HBV infection has been closely correlated with HCC development. The impact of the cellular changes in proteome in the context of HBx/HBV will be examined by functional assay such as cell motility assay.

Overall, we aim to establish a sequence-to-function platform for the comprehensive analysis of protein-protein interactions. Findings from such characterization should not only provide new insights on better antiviral strategy, but also identify biomarkers for efficient prognosis of related human diseases.

Chapter 2 : Materials and Methods

The following sections describe methodology used in this project. They include molecular biology and proteomics with a common goal of establishing protein profile in cells exposed to genotype specific HBx and replicative, as well as to toward a better understanding of molecule mechanism.

2.1 Cloning constructions

2.1.1 Viral DNA extraction from serum sample

Serum samples were obtained from HBV-infected patients, from National University of Singapore (NUH). Hepatitis B viral DNA from Genotypes D specifically, was extracted using QIAamp Blood Mini Kit (QIAGEN).

Genotype was confirmed by a modified RFLP method. First, the entire surface antigen gene, comprising the PreS1 and PreS2 region and the entire S gene proper (2848bp to 835bp), was amplified by primers P1 and P2. This was followed by a sequencing using primer P3 which is prior to the position 155, essentially the start codon of the S gene. The sequences of primers P1, P2, and P3 were shown as follows.

P1	5'-TCACCATATTCTTGGAACAA-3'
P2	5'-GTTTTGTTAGGGTTAAATG-3'
P3	5'-TCCTGCTGGTGGCTCCAGTTC-3'

2.1.2 Bacterial and mammalian expression vectors applied

Mammalian cloning vector pcDNA3.1 and pXJ40, *E.coli* expression vector pGEX-5X-1 pGEM-T were used here. Besides, during the cloning construction, the pGEM-T vector was applied for the quick ligation directly from PCR products. Among of these, pcDNA3.1, pGEX-5X-1 and pGEM-T were traditional commercial vector which were purchased from Invitrogen. pXJ40 is a nice gift from Dr. Cheng Gee Koh (School of Biological Sciences, NTU), and the relevant information can be found in the reference [137, 149]. The vector pXJ40 (Fig. 2.4) can provide a high expression efficacy compared with widely used mammalian vector such as pcDNA3.1. In the multicloning sites, there is a HA tag between the endogenous restriction sites EcoRI and BamHI, and such a tag can be conveniently used for the expression detection of the inserted target protein. In our work, genotypes specific HBx were cloned into pXJ40, and expression of HBx

Figure 2.1: Feature map of prokaryotic expression vector pGEX-5x-1.

Factor Xa is the Protease site which can be used to move out the purification tag GST.

Figure 2.2: Feature map of mammalian expression vector pcDNA3.1+.

MluI restriction site, prior to the pCMV promoter was used for cloning of the replicative HBV genome instead of the MCS.

Figure 2.3: Feature map of pGEM-T Vector circle map.

Figure 2.4: Feature map of mammalian expression vector pXJ40.

The expression vector pGEX-5X-1 (Invitrogen, USA) contains a purification tag GST, which improves the solubility of the protein and more importantly the

fusion protein can be purified by Glutathione-Sepharose column (GE healthcare, see Fig. 2.1). The factor Xa protease site provides cleavage site to separate GST tag away from the target protein.

The vector pcDNA3.1 (Fig. 2.2) is a widely used commercial mammalian expression. The information was shown in the Fig. 2.2 as follows. Our genotype specific replicative HBV genomes were constructed on this vector. The successful expression can be conformed by RT-PCR of core gene and surface gene of HBV. In addition, HBV specific ELISA (IMX) was applied to obtain the quantity of HBV particles secreted into medium.

2.1.3 Cloning construction of HBx from different genotypes

Clinical samples were provided by A/P. Lim Seng Gee from NUH. HBx gene of HBV genotypes A, B and C were cloned individually into pXJ40 using the HindIII and PstI sites in the Multiply Cloning Site (MCS). The oligonucleotide primers used to amplify the virus gene from template by PCR were as follows:

P1	HBx A	Forward: 5'- CCCAAGCTTGCTGCTAGGCTGTACT-3'
P2	HBx B	Forward: 5'- CCCAAGCTTGCTGCTAGGCTGTGCT-3'
P3	HBx C	Forward: 5'- CCCAAGCTTGCTGCTAGGGTGTGCT-3'
P4	HBx A/B/C/D	Reverse: 5'- AACTGCAGTTAGGCAGAGGTGAAA-3'

In particularly, P1 was the upstream primer for HBx-A. P2 was the upstream primer for HBx-B/D. P3 was the upstream primer for HBx-C. P4 was the downstream primer for all the four genotypes HBx. The individual inserts were sequenced and compared with archived HBx sequences in NCBI to ensure is the

cloning of the right genotype. This was important as the clinical samples may contain different genotypes.

2.1.4 Cloning construction of SH3 binding domain from HBx and SH3 domains from Vinexin-β

HBx genotype B SH3 binding domain and two SH3 domains of vinexin β were cloned individually into pXJ40 via the BamHI and XhoI sites in the Multiply Cloning Site (MCS). The oligonucleotide primers used were listed as follows:

pXJ40-BSH3	Forward: 5' - ACGGGATCCCGTCCCGTCGGCGCTGAA - 3'
	Reverse: 5' -CACCTCGAGTTATTATCCGCGTAAAGAGAGGTG-3'
pXJ40-VSH3	Forward: 5' -ACGGGATCCAAGAAGAGA AAGGCCGCCAGG-3'
	Reverse: 5' -CACCTCGAGTTACTTACTGCACGTAGCTGGCAGGGAA-3'

Plasmids having full length of HBx and vinexin β were used as templates for the amplification of SH3 binding domain. The target fragments were amplified by PCR, followed by direct ligation into pGEM-T vector, transformed into competent cells and colonies were selected. The fragments were ligated into pre-cut pXJ40 vector. The positive clonings were confirmed by enzyme digestion and analysed by electrophoresis of agarose gel. Positive clones were sent for DNA sequencing and the results were aligned to make sure no mutation was generated during PCR amplification.

2.1.5 Preparation of competent cells by CaCl₂ method

The preparation of LB broth, 2×YT broth and Agar plates were according to Molecular Clonings (Cold Spring Harbor Laboratory, Edition 3). The mixtures were autoclaved at 121°C 20min and cooled to 45~50°C to the addition of required

antibiotics. LB agar was then be used to prepare agar plates by pouring onto the Petri dish. The broth and plates cab be store at 4°C for several months.

Fresh agar plates were prepared without selective antibiotics. DH5 α or BL21 strains were streaked onto a fresh agar plate and cultured at 37°C for 12-16 hours. Single colony was selected from the agar plate and transferred into 5 mL LB broth without selective antibiotics and incubated at 37°C overnight with shaking at 220rpm. The cultured bacteria was re-inoculated the next day into fresh 100 mL LB broth at 1:100 ratio and allow to grow to a concentration of OD₆₀₀ 0.3 to 0.4. The culture in the exponential phase was then aliquoted into 50 mL pre-chilled centrifuge tubes and incubated on ice for 10 min followed by centrifugation for 5 min at 5,000 rpm at 4°C. And the cell pallet was re-suspended in 30 mL of 0.1 M ice-cold MgCl₂ by vortex and pipetting. The thoroughly suspended mixture was then centrifuged for 5 min at 5,000 rpm in 4°C. The supernatant buffer was then decanted carefully and the pellet was re-suspended in 20 mL of 0.1 M ice-cold CaCl₂ and incubated on ice for 30 min and then centrifuged for 5 min at 5,000 rpm at 4°C. And again the pellet was re-suspended in 1.5 mL of ice-cold 0.1 M CaCl₂/15% glycerol. The prepared competent cells were then aliquoted in to 1.5 mL autoclaved and fresh opened Eppendorf tubes and stored in -80°C freezer for long term storage. All the manipulations were carried out in Biological Safety Cabinet since the no antibiotics was applied.

2.1.6 Plasmid transformation

Prepared and stored DH5 α or BL21 competent cells prepared as described above were thawed on ice. 50ng of plasmid DNA was added to the competent cells (around 100 μ L) followed with soft flicking and incubated on ice for 30 min. The mixture of competent cells and plasmid was then subjected for incubation for 90 sec at 42°C, a process named heat shock. This was quickly followed by incubation on ice for 10 min allowing the cool down. 800 μ L LB broth was added to the mixture and the cells were recovered by incubation at 37°C shaking at 150rpm for 45 min. 100 μ L of the transformed cells were then plated onto prepared LB agar plate supplemented with selective antibiotics. The plate was then placed in 37°C incubator downside up until the all the liquid was absorbed by the LB broth and the plate was incubated overnight in an upside down position. The colonies, visible after an overnight incubation, were picked out from the plate and selection of positive transformation colony with insert was carried out as described below (2.1.8).

2.1.7 Glycerol stock

One single colony was selected from the LB broth agar plate after transformation, and was inoculated into a 5 mL fresh LB broth with proper antibiotics followed by shaking at 220 rpm overnight at 37°C. The cultured bacteria were then mixed thoroughly with sterile glycerol with the ratio of 15-30 % in the sterile EP tube and all relevant data (host strain, vector, selection criteria, date prepared) were recorded on the tube. The tube was stored at -20°C. For long term store, -80°C

freezer was applied. The glycerol stocked strain can be used for inoculation or recover by streaking.

2.1.8 Plasmid amplification by QIAGEN QIAprep Miniprep

Plasmids of interest were transformed into the DH5 α strain designed for gene recombination and plasmids amplification. Single colony was selected and inoculated into 5 mL fresh LB broth with proper antibiotics followed by incubation at 37°C shaking at 220 rpm overnight. The culture was centrifuged at 5,000 for 10 min to harvest the cell pellet. Supernatant was discarded and the bacteria cells pellet was re-suspended in 250 μ L ice-cod re-suspension Buffer P1 supplemented with provided RNase A (100 μ g/mL). The bacteria were completely re-suspended by vortexing and pipetting up and down until no cell clumps was visible. 250 μ L lysis buffer P2 was added and mixed thoroughly by gently inverting the tube 4-6 times. 350 μ L neutralization buffer N3 was added and mixed 4-6 times gently and thoroughly. The mixture was then centrifuged for 10 min at a speed around 13,000 rpm (~17,900g). Supernatant containing plasmids was then transferred into the QIAprep purification column without agitating the white pellet and subjected to centrifugation at 13,000 rpm for 30-60 sec. The absorbed plasmids on the silica membrane were than washed by 500 μ L buffer PB and 750 μ L buffer PE followed by centrifugation of 30-60 sec each time. An additional 1 min centrifugation was applied to thoroughly take away the residual washing buffer. The column with plasmids was placed into a new 1.5 mL EP tube and proper volume (30~ 50 μ L) of

distilled water or buffer Elution Buffer was added to the column. After incubation of 3 min, by 1 min centrifugation at 13,000 rpm, the plasmids were eluted from the membrane and dissolved in EB buffer or water. For high yield, second elution can be applied. Plasmid DNA collected by the elution was quantitated (see 2.1.10) and stored at -20°C until further use.

2.1.9 Plasmid amplification by Invitrogen Midiprep

Midiprep was applied to remove impurities such as endotoxin in DNA preparation which may harm the mammalian cell after transfection. Invitrogen Midiprep kit was used for this purpose. Briefly, 50ml of transformed *E.coli* bacterial cell culture was collected after overnight growth by 10 minutes centrifugation at 10,000g. The supernatant was discarded and pellet re-suspended in 3ml Cell Suspension Solution, followed by adding 3ml Cell Lysis Solution. The mixture was incubated for 3 min at R.T. 5ml Neutralization Solution was added to the lysed cells and mixed by gently inverting the tube 3-5 times. Another incubation of 2-3 min was added to allow a white flocculent precipitate to form. The lysate was then poured into PureYield Clearing Column (blue for lysate clearing) into a new 50ml disposable plastic tube and incubate for 2 min. The column was centrifuge 1,500rpm for 5 min. The PureYield Binding Column (white) was placed onto the vacuum manifold and the cleared lysate was poured into the white column. The lysate was forced pass through the binding column by applying maximum vacuum. The membrane bound plasmids were washed by 5ml Endotoxin Removal and 20ml

Column Wash Solution and the vacuum was applied to draw the solution through. The membrane was made by applying the vacuum for 30 seconds. The DNA on the membrane was eluted by adding 600µl Nuclease Free water. The purified plasmids was then quantitated (see 2.1.10) and stored.

2.1.10 Plasmid DNA quantification by UV spectrophotometer

Plasmids DNA purified by QIAprep Miniprep or Midiprep was quantified by UV spectrophotometry at OD_{260nm} and OD_{280nm}. The plasmid was diluted to proper concentration with distilled water at 125 ×. Pure water was used for zero calibration. DNA Yield was calculated by multiplication of by OD₂₆₀ 50 µg/mL and a further multiplication by the dilution factor. Formula shown as follows:

$$\text{OD}_{260\text{nm}} \times 50\text{ng} / \mu\text{l} \times 125 (\text{dilution factor}) = \text{Concentration ng} / \mu\text{l}.$$

The purity of the plasmid DNA was obtained by dividing OD_{260nm} with OD_{280nm}. A value between 1.7 and 1.9 indicated a good preparation of the plasmid DNA.

2.1.11 Restriction enzyme digestion

For cloning purpose, single restriction enzyme or double restriction enzyme digestion were used according to the initial cloning strategy. Typically, the vector plasmid DNA, PCR amplified product were digested by New England Biolab's (NEB) restriction enzymes. Enzymes were usually added in the concentration of 1 unit for each digestion reaction. 10 × reaction buffer was chosen to the appropriate enzymes to 1× in the final digestion mixture. For a number of enzymes, BSA was

needed to optimal the digestion and $100 \times$ BSA is provided by NEB. The mixture was incubated at 37°C for more than 1 hour. Longer time of incubation was applied whenever appropriate. The digestion products were separated on agarose gel by electrophoresis and the DNA fragments extracted and purified (see 2.1.12). For cloning purpose, calf intestinal alkaline phosphatase (CIP) was added to the vector digestion mixture at the end of usual restriction reaction and incubated for additional 30 min at 37°C . By catalyzing the release of 5' and 3' phosphate groups of the DNA digested fragments, the enzyme CIP was particularly useful to reduce self ligation of the vector DNA when single restriction sites were used.

2.1.12 Gel extraction

QIAquick gel extraction kit was used for the gel extraction and DNA fragments purification. After the restriction enzyme digestion, the products were separated by DNA electrophoresis, and the concentration of gel was selected on the size of target fragments. Generally, 1 % agarose gel was used. The target DNA fragment was cut out from the gel under the UV light and placed into pre-weighed EP tube. To resolve the gel and release the DNA fragment, the QG buffer from the QIAquick kit was added in a ratio of 3 : 1 (volume: weight) to the excised agarose gel containing the target fragment. The mixture was placed in the $50\text{-}60^{\circ}\text{C}$ water bath and incubated for 10 min with vortexing every 2 or 3 min to melt the agarose to facilitate the release of DNA. When the gel was thoroughly melted, $100\ \mu\text{l}$ isopropanol was added and mixed, followed by applying to the QIAquick spin

column. A centrifugation at 13,000 rpm for 1 min allowed the DNA binding to the silica membrane while the QE buffer it was removed. 500 μ l buffer QG was added following centrifugation for 1 min at 13,000 rpm to wash the bind DNA and remove the residual of agarose and ethidium bromide (EB). After removing the flow through, 750 μ l washing buffer PE was added to the column and incubated for 3 min, followed by centrifugation for 1 min. An additional centrifugation for 1 min was applied to remove any residual washing buffer. Then the column with binding DNA was transferred to a clean 1.5 mL EP tube and proper volume (according to the estimated quantity of target DNA and the concentration needed) of elution buffer (EB) (10mM Tris-Cl, pH 8.5) or distilled water was added to the center of the column membrane. At then end of 3 min incubation, the EB buffer containing the target DNA fragment was collected by centrifugation for 1 min at 13,000 rpm.

2.1.13 Ligation

After restriction digestion and gel extraction, the purified DNA fragments were used for ligation, to allow the insertion of DNA insert to the vector. T4 ligase from NEB was used for all ligations in this project. 10 $\times$ T4 ligase buffer was diluted into 1 $\times$ to for the reaction of ligation. The purified insertion DNA fragment and vector with the same sticky enzyme ends were added into the reaction mixture. 1 unit of T4 ligase was added to the mixture. While ligase displayed activity at a broad range of temperatures, and the optimal incubation temperature for T4 DNA ligase was 16°C.. For routine subcloning, the incubation time and temperature

were set at either 4°C overnight or at room temperature for 30 minutes. Following the ligation, transformation was applied prior to the selection of positive clonings. In contrast to the transformation of intact plasmids, the transformation efficiency for cloning was usually lower. Competent cells used in transformation were collected by centrifugation at 5,000 rpm for 1 min and 700 µl broth was removed and the cell pellet was softly resuspended in the left 200 µl broth. The suspended cells were plated onto LB agar plates with selective antibiotics. After overnight incubation, the colonies were selected, plasmid DNA isolated for confirmation of correct insert by enzyme digestion. Inserts of correct size were sent for sequencing with designed primers to ensure that there was no mutation in the insert.

2.1.14 Multiple sequence alignment analysis

Full length sequences for alignment were from the sequencing results of the individual clonings. The genotype was determined based on the sequences obtained from Genbank, NCBI (<http://www.ncbi.nlm.nih.gov>). The alignment was performed using the software BioEdit.

2.1.15 Site-directed mutagenesis

To generate mutant constructs, site-directed mutagenesis was used (QuickChange Site-Directed Mutagenesis kit, Stratagene). The primers were designed according to the sequence flanking the target mutation sites. Short primers of 17~18 nucleotides were generally used. Following the instructions of the manufacturer,

PCR was applied with the designed primers, the reagents, and the diluted (usually 1:50) template plasmid of interest. The applied reaction system was as follows: 5 µl of 10 × 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), 125 ng of the forward and reverse primers each, 50 ng of dsDNA template, 1 µl of dNTP mixture and RNase-free water was added to top up the reaction mixture to be 49 µl. 1 µL of pfu Turbo DNA polymerase (2.5U/µL) was then added to the reaction mixture. The PCR reaction was carried out in this cycling conditions: 95°C 30 seconds; 18 cycling of 95°C 30 seconds, 55°C 30 seconds, 68°C 5 minutes (may vary depending on the size of the target plasmid with a speed of 1 kb/min). The completed PCR reaction should contain the newly synthesized plasmids with site mutations. The parental strands were removed by adding DpnI (1 µl) and an incubation at 37°C for 2 hours. The products were transformed into *E.coli* cells provided in the kit as high efficiency competent cell DH5α. And the positive colonies were selected. The efficiency of site-directed mutagenesis was confirmed by sequencing the mutation site with a primer upstream.

2.2 Mammalian cell culture and transfection

2.2.1 Mammalian cell culture and cell passage

HepG2 (human hepatoma cell line) and HEK293T (human embryonic kidney cell line) cells were from ATCC and cultured and passaged in Dulbecco's Minimal Essential Medium (DMEM) plus 10% fetal bovine serum (FBS) (Gibco, USA), 1%

anti-mycotic (penicillin, streptomycin and amphotericin) (Invitrogen, USA) at 37°C 5% CO₂. HepG2.2.15 was cultured with DMEM (Gibco, USA), plus 10% FBS, 1% P+S (penicillin+streptomycin), 150 mg/ml of G418 (Geneticin) (Gibco, Invitrogen), 1 mM sodium pyruvate, 2 mM glutamine and 0.1 mM MEM NEAA under 37 °C 5% CO₂. The cells were passaged every 2~3 days until it reached approximately 80-90% confluency. For the cell passage, the cells were washed by PBS (Gibco, USA), and treated with trypsin (Gibco, USA) for several minutes (depending on the cell type, 3-5 minutes usually). The cells were observed under microscopy. The trypsinization was stopped by adding culture medium when the cells becoming round and detaching from each other and the cells were softly pipetted to disperse and detach from the surface and one to another. The suspended cells were centrifuged at less than 200g for 5 minutes before adding fresh medium and aliquoted into flask in a proper concentration, usually 3 to 5 times dilution.

2.2.2 Freezing, storage and recovery of cells

Cells were allowed to grow to 80-90% confluency in T75 tissue culture flask (Corning, US). The culture medium was removed and cells were washed with pH7.2 1 × PBS (Gibco, USA) twice. Cells were then trypsinized with 2 × Trypsine/EDTA in 37°C for several minutes,. When the cells were detached from the flask, the trypsinization was stopped by adding MEM with 10% FBS and cells were collected after centrifugation at 1,000rpm for 5min at room temperature (RT). The supernatant was discarded and the cell pellet was suspended in 3 mL cell

freezing medium with 10% DMSO, 20% FBS even distributed in MEM. 1 mL of the suspended cell medium was then aliquoted into 2 mL sterile plastic ampoule with screw cap. The ampoule was placed in -20°C for 2 hours and followed by -80°C freezer overnight to allow gradual temperature decrease. The frozen cells were then rapidly transferred into liquid nitrogen tank for long term storage.

When taken out from the liquid nitrogen, cells were thawed by quickly and gently swirling the stored ampoule in 37°C water bath in less than 2min. The cell suspension was removed from the ampoule to sterile 15 mL tube, by adding 9 mL fresh MEM with 10% FBS. The medium with DMSO was removed by centrifugation at 1,000 /min for 5min at RT. The removal of DMSO was to minimize detrimental effect on cell growth. Cell pellet was suspended in 10 mL pre-warmed fresh MEM with 10% FBS and 1% anti-mycotic in T75 cell flask. The cells were maintained in 37°C and 5% CO₂ cell culture incubator. Cells freshly recovered were allowed to passage for 2 cycles before experiments to ensure a normal cell growth.

2.2.3 Cell transfections

Transfections were performed using Effectene Transfection Reagent (QIAGEN, USA) and Nucleofection solution V (Amaza GmbH). For HepG2 cells, Nucleofection transfection was applied. 24 hours prior to transfection, 2.5×10^6 cells were seeded in 75cm² flasks. Upon adherence, the cells were transfected with

2-5 µg of various constructed plasmids. The following section was described in our previous published work [150]. Briefly, HepG2 cells were resuspended in 100 µl of provided solution V. Plasmids were added into the resuspended cells and mixed. Transfection was carried out applying HepG2 specific Program T-28 to obtain a high efficiency. 500µl prewarmed culture medium was added immediately and quickly and the mixture was carefully transferred onto culture dish or flask containing prewarmed medium. After transfection, the cells were incubated at 37 °C 5% CO₂. 4 hours after transfection, medium including solution V and non-adherence cells were removed and prewarmed fresh medium was added.

For 293T cells, Effectene Transfection was used. 24 hours prior to transfection, cells were seeded in 6-well plates. Upon adherence, the cells were transfected with 0.4ug plasmids processed as per instruction. Briefly, the constructed plasmids were mixed and incubated with 3.2 µL Enhancer followed by addition of 10 µl of Effectene Transfection Reagent. 2-4 hours after-transfection, the transfection medium was removed, and fresh medium was added. Transfected cells were maintained at 37 °C with 5% CO₂ for one day followed by cell harvesting.

2.3 SDS-PAGE (Sodium dodecyl Sulfate - polyacrylamide gel electrophoresis)

SDS-PAGE was a mature widely used protein separation technique based on the difference between the molecular weight of protein complex regarding their electrical charges [151]. It has been successfully used in biochemistry, forensic

science, genetics and molecular biology. The Laemmli discontinuous system in which the proteins were first stacked in the low concentration stacking gel before entering the separating gel to obtain a good resolution of the gel has been in use since 1970.

Different proteins with similar molecular weights may have different migration velocity due to differences in mass charge ratio. Addition of SDS, allows the unfolding of the protein, which will give a near uniform negative charge along the length of the polypeptide. This leads to an approximately uniform mass: charge ratio for most proteins and the difference among proteins caused by isoelectric points are eliminated. As a result the migration velocity is assumed to be directly related to only the size of the protein.

The following gel preparation recipes were applied.

Table 2.1: Components of stacking gel and different concentration separating gels.

(A) Separating gels for denaturing SDS-PAGE

% Gel	Components	Volume of components (ml) per gel mold volume of			
		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
	1.5M Tris (pH8.8)	1.3	2.5	3.8	5.0
	10% SDS	0.05	0.1	0.15	0.2
	10% ammonium persulfate	0.05	0.1	0.15	0.2
	TEMED	0.003	0.005	0.009	0.012
10%	H2O	1.9	4.0	5.9	7.9
	30% acryl-bisacrylamide mix	1.7	3.3	5.0	6.7
	1.5M Tris (pH8.8)	1.3	2.5	3.8	5.2
	10% SDS	0.05	0.1	0.15	0.2
	10% ammonium persulfate	0.05	0.1	0.15	0.2

TEMED		0.002	0.004	0.005	0.008
12%	H2O	1.6	3.3	4.9	6.6
	30% acryl-bisacrylamide mix	2.0	4.0	6.0	8.0
	1.5M Tris (pH8.8)	1.3	2.5	3.8	5.0
	10% SDS	0.05	0.1	0.15	0.2
	10% ammonium persulfate	0.05	0.1	0.15	0.2
	TEMED	0.002	0.004	0.006	0.008
15%	H2O	1.1	2.3	3.4	4.6
	30% acryl-bisacrylamide mix	2.5	5.0	7.5	10.0
	1.5M Tris (pH8.8)	1.3	2.5	3.8	5.0
	10% SDS	0.05	0.1	0.15	0.2
	10% ammonium persulfate	0.05	0.1	0.15	0.2
	TEMED	0.002	0.004	0.006	0.008

(B) 5% stacking gels for denaturing SDS-PAGE

Components	Volume of components (ml) per gel mold volume of							
	1 ml	2 ml	3 ml	4 ml	5 ml	6 ml	8 ml	10 ml
H2O	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

In our experiment, 6%, 8%, 12% and 15% separating gel were used for better separation efficiency based on the size of our target proteins. For good separation of small size proteins, low concentration separating gel was adopted while for the separation of large size proteins, 6% and 8% were applied. The components are shown in Table 2.1. Three types of gel thickness were used, 0.8mm, 1.0mm and 1.5mm and the choice of combs was made depending on the purpose and the concentration of the protein. Thick gel allowed more sample loading capacity when the length and width were consistent. First, the separating gel was prepared and applied into the glass plate sandwich, and water was added to cover the un-casted gel. When the gel was set, the water was removed and the stacking gel

can be prepared and aliquoted onto the separating gel. This was followed by the insertion of comb of proper thickness. The comb was removed after the gel was set and the formed wells filled with protein.

The electrophoresis was carried out by placing the casted gels with short glass plate facing inwards towards the inner buffer chamber. The inner buffer chamber was filled with fresh 1× Tris Glycin running buffer (0.025M Tris, 0.192 M glycine, pH 8.3, 0.1 % SDS). The prepared protein samples were mixed with 2× or 6× SDS loading buffer and boiled at 98 °C for 10min. Molecular weight rainbow markers (GE, USA) were loaded into the wells to indicate the size of the proteins in prepared samples. And the bands of rainbow markers can be transferred to membranes (NC or PVDF membrane) for Western blot detection. Around 20 µl of the samples was loaded into the well. The initial stacking runs were performed at 80V about 30min until the samples have entered the separating gel, indicated by dye contained in the loading buffer. For the separating gel, 150 V was applied until the dye reached the bottom. The gel was separated from the glass plates carefully. It was then stained with Coomassie Blue staining (50 % (v/v) methanol, 40 % (v/v) distilled water, 10 % (v/v) glacial acetic acid, 0.05 % (w/v) Brilliant Blue R-250) overnight. This was followed by destaining using destaining buffer which shared the same components with staining buffer except the Coomassie Blue. The destaining took several minutes to hours to show clear bands of the gel. The gels can also be used for Western blot detection.

2.4 Western Blot analysis

For Western blot analysis, the gels were not incubated in staining solution but immersed in transfer buffer (25 mM Tris, 192 mM glycine, 10 % (v/v) methanol). The proteins on the gels were then transferred onto the transfer buffer pre-wetted nitrocellulose membrane (NC) or methanol soaked and transfer buffer washed activated polyvinylidene difluoride (PVDF) (GE Hybond, USA) by semidry transfer system (Bio-rad, USA) (Fig. 2.5). The transfer was applied at constant voltage of 21 V for 50min. The membrane was subjected to Ponceau S staining (0.1 % (w/v) Ponceau S, 5 % (v/v) glacial acetic acid, 96 % distilled water) to show the transfer efficiency. The Ponceau S staining buffer was then washed off by distilled water.

Figure 2.5: Semi-dry sandwich transfer system (BioRad).

The transferred membranes were then subjected to blocking in 5 % non-fat milk solved in 1 × PBS for 1 hour at R.T. or overnight at 4 °C with rolling at a low

speed. After blocking, specific primary antibodies in 3 % non-fat milk dissolved in $1 \times$ PBS were incubated for 1 hour at R.T. or overnight at 4 °C. Non-specific interactions were then removed by 3 washes of 10 min with utilizing cold $1 \times$ PBS/0.005 % Tween-20. Secondary antibody conjugated with Horseradish Peroxidase (HRP), in 3 % non-fat milk dissolved in $1 \times$ PBS, were incubated for 1 hour at R.T. or 4 overnight. This was followed by 3 washes under the same condition with the washes after first antibody reaction. The detection of the reaction proteins to antibodies was performed using Enhanced Chemiluminescence (ECL), with SuperSignal West Pico Chemiluminescent reagent (Pierce). Equal volume of stable peroxide solution and luminol/enhancer solution were mixed and applied to the membrane processed above. The membrane was incubated with the reaction substrate for less than 3min; subsequently the excess substrate was decanted.

The detection of ECL was performed by using Pierce CL exposure X-ray film (blue color). The signals on the membrane were exposed to the X-ray film for varying durations, dependent on the concentration and sensitivity of the antibody and the intrinsic concentration of the target protein. The exposure time ranged from 10 seconds to overnight, usually several minutes. Developing and fixing buffer were purchased from Kodak and diluted 5 times in distilled water. Film was developed manually and varying exposure timings were adjusted to obtain properly intensity with clear background. After exposure, the film was immersed into the diluted developing buffer and agitated until the visualization of targeted

bands, followed by washing off the excess developer with water and then the film was put into diluted fixing buffer by agitating for around 2min. Developed X-ray film was then scanned using Bio-Rad's Densitometer and analyzed by using Bio-Rad's QualityOne software.

The antibodies in use were from Santa Cruz biotechnology, Sigma and Pierce. The catalog numbers were as follows:

Anti- β -actin (AC-74); anti-HA (sc-7392, Santa); anti-Annexin A2 (sc-28385); anti-S100A11 (sc-98427); anti-PSMA3 (sc-58414); secondary antibody, HRP-conjugated anti-mouse IgG (#31430, Pierce); secondary antibody, ImmunoPure Goat Anti-Rabbit IgG (H1L), peroxidaseconjugated antibody (31460, Pierce). The dilution ratios were different for each antibody, ranging from 1:100 to 1:5000.

2.5 Total proteins extraction, protein digestion, and labeling with iTRAQ reagent for LC-MS/MS

This section was largely similar to our published work and most descriptions were similar or identical [139]. Two days after-transfection, cells were harvested and lysed in 250~400 μ L (depending on the cell concentration) of 8 M urea, 4% CHAPS (w/v), and 0.05% SDS (w/v) on ice for 30 min with regular vortexing. The mixture was centrifuged at 15, 000g for 60 min at 4 °C. After the supernatant was removed, the concentration of protein was measured 2-D Quant Kit (GE Healthcare) or Bradford protein quantification assay. A standard curve was made

using BSA as a control. For the secreted proteins in cell culture medium, the medium was collected and concentrated, and followed by protein quantification. Specifically, a total of maximum 100 µg of each protein sample was precipitated by adding 4 volume of pre-cold acetone at -20 °C for from 2 h to overnight. Then the precipitated proteins were collected by 15,000rpm centrifugation for 10minutes. Before labeling with iTRAQ reagents, the proteins were dried and dissolved in 20 µL dissolution buffer (0.5M triethylammonium bicarbonate, pH8.5 and 1.5 µL 2% SDS), and denatured, and reduced by 2 µl 50mM reduction buffer (TCEP-HCl, tris (2-carboxy-ethyl) phosphine) followed by incubation for 1 hour at 60°C, and cysteines were alkylated by adding 1 µL cysteine blocking buffer (200mM methyl-methanthiosulfonate (MMTs)) as described in the iTRAQ kit provided protocol (Applied Biosystems, MA, USA). Sample being analyzed was then digested with 20 µL of 0.25µg/ µL sequence grade modified trypsin (Promega, USA) solution at 37 °C for 12~16 hrs and the iTRAQ tags were used to label them respectively as follows: HepG2 transfected with empty vector pXJ 40 = iTRAQ 114; HepG2 transfected with HBx genotype A = iTRAQ 115; HepG2 transfected with HBx genotype B = iTRAQ 116; HepG2 transfected with HBx genotype C = iTRAQ 117. The labeled samples were then combined for further analysis. To make sure that there was no interference between sample preparation and digestion/labeling, the bovine serum albumin (BSA) standard solution (Pierce) was used separately in the acetone precipitation, followed by enzymatic digestion with trypsin and labeling with the iTRAQ reagents as described above. These differentially labeled digests were mixed at a ratio of 1:2:1.5:2 and analyzed by

LC-MS/MS [139].

The equipments in use were Agilent 1200 nanoflow LC system (Agilent Technologies, US) interfaced with a QSTAR XL mass spectrometer (Applied Biosystems/MDS Sciex, US) as described [139]. A total 5 μ L of the combined peptide mixture was injected into the PolySulfethyl A SCX column (0.32×50 mm, 5μ m) of the HPLC system as described [139]. In the HPLC system, four columns were applied, which were SCX, reverse column and two other enrichment columns. By the four columns, two-dimensional separation was realized as described [139].

In the first dimension of separation in LC, a series of increasing concentration of KCl salt solution as 10, 20, 30, 40, 50, 60, 80, 100, 300, and 500 mM was used to elute the retained peptides in a stepwise manner from the SCX column by sequential injection as described [139]. Followed by the second dimension, with the 10-port valve being in position 1, the sequentially eluted peptides were released to C18 enrichment column 1 (0.3×5 mm, 5μ m) (ZORBAX 300SB) and then was washed isocratically with buffer A (5% acetonitrile, 0.1% formic acid) at 0.005 mL/min for 100 min as described [139]. The same time, the peptides bound on the C18 enrichment column 2 during the previous run were eluted with buffer B (0.1% formic acid) and buffer C (a nanoflow gradient of 5-80% acetonitrile + 0.1% formic acid) at a flow rate of 500 nL/min as described [139]. Then the samples were further separated by a C-18 reversed-phase column (75μ m $\times$ 50 mm,

3.5 μ m) (Zorbax 300SB) as described [139]. After finishing the separation, the 10-valve was switched to position 2, and sample 1 was switched into the solvent path of the nanopump and column 2 was used to trap the newly eluted peptides from SCX as described [139]. 13 runs were performed altogether for one injection as described [139]. For MS/MS analysis, survey scans were acquired from m/z 300-2000 with up to two precursors selected for MS/MS from m/z 100-2000 using dynamic exclusion, and the rolling collision energy was used to promote fragmentation as described [139].

2.6 Concentration of secreted proteins

For the identification of biomarkers, the changes in secreted proteins were investigated. 48 hours after-transfection, the culture medium was carefully collected, and subjected to centrifugation for 3,000g for 10 min to remove cell debris in the culture medium. The medium was then transferred into 5 kDa ultra-filter concentrators (Sartorius). With 30min centrifugation at 4°C in swing bucket rotor, the medium was 30 times concentrated. And the concentrated protein samples were stored for quantification and further use. For the LC-MS/MS analysis, the proteins were quantitated with Bradford as described above and 100 μ g proteins of each sample was precipitated and labeled with iTRAQ reagents, the same with the intracellular protein identification.

2.7 Protein expression and purification

2.7.1 Bacterial protein expression

Bacterial *E. coli* strain BL21 (DE3) was used. Individual constructed clonings on bacterial expression vectors, pGEX or pET vectors were transformed into the prepared competent bacterial cell BL21 prepared in 2.1.3. Isolated colonies were selected and grown in 5 mL of Luria Bertani (LB) with selective antibiotics (100ug/mL for ampicillin and 30ug/mL for kanamycin) overnight at 37°C 220rpm. The culture was scaled up by re-seeding the cultured cell medium in 1:100 ratios, which was 1 mL into 100 mL fresh LB culture medium with selective antibiotics. And it was cultured for approximately 3 hours to obtain the cell density of OD_{600nm} 1.0. 100mM filter-sterilized (0.2um filter) Isopropyl-β-D-thiogalactopyranoside (IPTG) (Fermentas, Lithuania) was added in the range of 0.1 mM to 1mM for the induction of protein expression. The cultures were then incubated in varying temperatures of 37°C, 30°C, 18°C, 4°C for varying hours (4, 6, 8 hours, overnight respectively) at 220rpm for protein expression. The *E. coli* cells were collected by a centrifugation at 5000 rpm for 30 min at 4°C. The LB medium supernatant was removed and the bacterial cell pellet was suspended with appropriate amount of PBS (100mL LB cultured cell, suspended in 10 to 20 mL PBS) supplemented with EDTA-free complete protease inhibitor (Roche, Switzerland) and Trinton X-100. The cells were broken by sonication disruption for several minutes depending on the volume and concentration of the cell suspension followed by 15,000 rpm centrifugation at 4°C for 1 hour to separate the insoluble and soluble proteins. The supernatant, containing the soluble proteins was obtained and quantified (Bio-rad's Quick Start Bradford Protein Assay, USA). The quantified proteins were stored at

-80°C for further application. The insoluble cell pellet was re-suspended in SDS-PAGE sample buffer and may be used for SDS-PAGE gel detection.

2.7.2 Protein purification

The fusion proteins in our experiments were glutathione-S-transferase (GST) tagged to facilitate affinity purification. Microspin GST purification modules from GE Healthcare were used. Briefly, quantified GST fusion protein was mixed with sepharose 4B beads for 30 min with head to end rotation at 4°C. The GST fusion protein would bind to the GST specific beads during rotation. The buffer containing impurities and un-combined GST fusion proteins were removed by centrifugation at 3000rpm for 45 seconds at R.T. and followed by three washes with 600 µL ice-cold PBS. Reduced glutathione buffer which will competing with the binding with beads was added and the mixture was rotated for another 30 min at 4°C. Under completion, the GST fusion proteins bound on the beads were now cleaved by glutathione and solved in the buffer and were collected by centrifugation at 3,000 rpm for 1 min.

2.8 GST-pulldown assay

As described above, the GST fusion protein was allowed to be immobilized on Glutathione Sepharose 4B beads by rotation for 30 min at 4°C. Excess of unbound proteins were removed by three washes with PBS. The cell lysate or expressed proteins which may interact with the bound proteins on the beads were added and

allowed to interact by rotation for 4 hours at 4°C. Non-interacting proteins were removed by centrifugation and more than three washes with PBS/0.005% Tween-20. The interacted proteins were then eluted using glutathione buffer (10mM glutathione in 50 mM Tris-Cl pH 8.0). The eluted interacting proteins were store at -80°C prior to further analysis.

2.9 Panomics transSignal SH3 domain arrays

TransSignal SH3 domain arrays (Panomics, USA) were applied for the identification of potential interacting proteins with our SH3 binding domain-containing protein. The array was made of the nitrocellulose (NC) membranes which contain expressed and spotted in duplicates SH3 domains from various known cellular proteins. GST fusion proteins with six histidine tag were generated for use. GST-PXXP-6×His were generated from HBx genotypes A, B, C, and D (prepared by Dr. Tan Tuan Lin) and the PXXP mutants were generated from these templates. All the eight constructs (4 wildtype and 4 mutants) were expressed in *E.coli* and the purified proteins were used for interaction analysis with Panomics membrane. Briefly, the amino acids sequence of HBx genotypes A, B, C, and D between amino acids 19 and 58 was cloned with an N-terminal GST (pGEX-5X-1) by additional adding of C-terminal 6X Histidine tag to generate the probe designated GST-PXXP. To explore the potential importance of the proline residues within this region for interaction with cellular proteins, site-directed mutagenesis (QuikChange Site-directed mutagenesis, Stratagene) was carried out as per

manufacturer's instructions and specifically, proline residues at position 29, 33, 42 and 43 were mutated to alanine residues to obtain the mutant GST-AXXA probe. For genotypes B and C, the wild type and mutant clonings had been constructed by lab colleague, and in this work, genotypes A and C were completed to show much more comprehensive understandings. More than 15 µg of 4B Sepharose purified GST fusion proteins were interacted with the membrane as the manufacturer's instructions. Non-interacting proteins were washed off with the provided wash buffer. And provided 400 × anti-histidine antibody was used to react with the membrane for the detection of the GST fusion protein. Subsequently, the membrane was washed by provided washing buffer and reacted with second antibody similar with Western blot procedure. Positive interactors of SH3 domain were identified by the use of SuperSignal West Pico chemiluminescent substrate (Pierce). Controls of the membrane were provided in the form of histidine peptides spotted in duplicates on the right and bottom of the membrane by the manufacturer. Importantly, true positive interaction shows up as duplicated spots.

2.10 IMX analysis

IMX system from Abbott Laboratories was used for qualitative and semi-quantitative analysis of HBV replication. In this system, the detection of HBsAg in culture medium was measured by the use of Microparticle Enzyme Immunoassay (MEIA) technology. This method has been widely used in clinical diagnosis for chronic HBV carriers.

The assay functions on the specific interaction of kit provided mouse monoclonal anti-HBs coated microparticles with HBsAg. The anti-HBs were biotinylated and when added into the antibody-antigen mixture, an antibody-antigen-antibody sandwich complex was formed. The complexes were then bound to the glass fiber matrix and followed by reaction with anti-biotin (Alkaline phosphatase conjugated). The unbound materials were washed away, and the substrate 4-Methylumbelliferyl phosphate was then added and the measurement was taken by the IMX MEIA optical assembly. In our experiments, cell transfected with plasmids, carrying HBV replicative genome was incubated for 48 hours at normal cell culture condition to allow HBV replication and HBsAg secretion. The culture medium containing the secreted HBsAg was collected by centrifugation at 2,000 rpm for 10 minutes to remove the cell debris. Five drops (around 200 μ l) of culture medium was added to the proper carousel. Separately, the positive and negative controls were also added to different carousels. Detection results were reflected numerically. The waste tank was cleaned after each single use.

2.11 RNA extraction, quantitation and RT-PCR

2.11.1 RNA extraction

Total RNAs were extracted from HepG2 cells transfected by various plasmids by using the RNeasy Mini kit (QIAGEN, USA) according to the manufacturer's instructions. Briefly, cells were pelleted by trypsinization and resuspended in 350

or 600 μl ($>5 \times 10^6$ cells) of Buffer RLT (supplied by the kit) followed by homogenization for 30 sec by vortexing. An equal volume of 70% analytical grade ethanol was then added, and the mixture was applied to the spin columns (within 700 μl) and centrifuged at 13,200rpm for 30 sec. The columns were then washed with 700 μl of Buffer RW1 and centrifuged for 30 sec. After two additional washes with 500 μl Buffer RPE, additional centrifugation was applied to remove residual buffer RPE. The total RNA bound on the membrane of the column was eluted with 50 μl RNase-free water or DEPC treated water. Total RNA was quantified and stored at -80°C for long term storage.

2.11.2 RNA quantitation

The concentration and quality of RNA were measured by the absorbance reader at OD_{260} and the absorbance ratio of OD_{260} and OD_{280} . Extracted total RNA was diluted 100 to 200 $\times$ to perform the quantification. Yield was calculated by multiplication of by OD_{260} 40 $\mu\text{g}/\text{mL}$ and a further multiplication by the dilution factor. Quality and Purity of the RNA were calculated by division of OD_{260} over OD_{280} .

2.11.3 RT-PCR

Reverse transcript-polymerase chain reaction (RT-PCR) was applied to detect the *HBcAg* and *HBsAg* gene expression from HBV genome. Total mRNA was extracted with RNeasy Mini Kit (QIAGEN) according to the manufacture's

protocol. Briefly, 1 µg of each type of total mRNA was used for amplification by using One Step RT-PCR Kit (QIAGEN). Primer sequences specific for *HBcAg*, *HBsAg* and *β-actin* gene in RT-PCR were shown as follows:

HBsAg	Forward: 5'-TCACCATATTCTTGGAACAA-3'	1247bp
	Reverse: 5'-GTTTTGTTAGGGTTTAAATG-3'	
HBcAg	Forward: 5'-ATCTCCTAGACACCGCCTCA-3'	146bp
	Reverse: 5'-TTCCAAATTATTACCCACCC-3'	
β- actin	Sense: 5'- CTTAGTTGCGTTACACCCTTTC - 3'	151bp
	Antisense: 5'- ACCTTCACCGTCCAGTTT - 3'	

The programs for *HBcAg* and *HBsAg* gene were as follows: reverse transcription (RT) at 50 °C for 1 h; initial denaturation at 95 °C for 5 min; 25 cycles of PCR (30 s of denaturation at 95 °C, 30 s of annealing at 60 °C and 1 min of extension at 72 °C); final extension at 72 °C for 10 min. At the end of RT-PCR, 5 µl of the amplification products were loaded in 1.0% agarose gel and separated by electrophoresis for 30 min at 120 V. DNA bands on the gel were stained with EB (Sigma, USA), visualized under UV light and photoed by ChemDot system (BioRad, USA).

2.11.4 Real-time RT-PCR

Technique described in the section was similar if not the same to our published work [152]. Real-time RT-PCR was used for accurate quantification of the mRNAs level in differential samples by iScript One-step RT-PCR kit (BioRad, USA) according to manufacturer’s instructions as described [152]. The real-time RT-PCR was carried out in a thermal cycler and analyzed by IQ5 multicolor Real-time PCR analysis system (BioRad, USA), with the following program: 50°C 10 minutes for

first strand cDNA reverse transcription; 95°C 5 minutes for the inactivation of reverse transcriptase; followed by 40 cycles of PCR (one cycle at 95°C 10 s, 60°C 30 s); and analysis of a melting curve of 95°C 60 s, 55°C 60 s, 55°C 10 s with 80 cycles increasing by 0.5°C at each cycle as described [152]. The data including analysis of amplification process and melting curve, and threshold cycle number were provided automatically by IQ5 Real-time PCR system software version 2.0 (BioRad, US) in the form of Microsoft Excel as described [152]. The comparatively change folds of mRNAs levels in different samples were calculated with this formula as described [152]:

Sample $\Delta C_t = C_{t_{\text{sample}}} - C_{t_{\beta\text{-actin}}}$; $\Delta\Delta C_t = \text{Sample } \Delta C_t - \text{control } \Delta C_t$; the fold of sample vs. control $= 2^{-\Delta\Delta C_t}$

The sequences of target proteins were searched and obtained in NCBI nucleotide database. The target products of real-time PCR were fragments of these proteins located in different exons to make sure DNA residues will not be false detected. The primers were designed according to the purpose. The information of the primers is provided as follows:

Protein Name	Primer sequences	PCR product (bp)
β - actin	Sense: 5'- CTTAGTTGCGTTACACCCTTTC - 3'	151
	Antisense: 5'- ACCTTCACCGTTCCAGTTTT - 3'	
MACF1	Sense: 5'- TTCAATGCACTCATTACCG - 3'	170
	Antisense: 5'- GACTTTTCATCTGGAGATGG - 3'	
Annexin2	Sense: 5'- GACACCTGCTCAGTATGACG - 3'	189
	Antisense: 5'- AAGTCACCAGATGTGTCCGA - 3'	
S100A11	Sense: 5'- GTCTTCCAGAAGTATGCTGG - 3'	206
	Antisense: 5'- GCTAGGCCACCAATCAGATT - 3'	
S100A4	Sense: 5'- CACAAGTACTCGGGCAAAGA - 3'	190
	Antisense: 5'- ACAGGAAGACACAGTACTCT - 3'	

PSMA3	Sense: 5' - GCAGTAGCAGGTTTGTGGC - 3'	176
	Antisense: 5' - CTGCAGCCAAAAGGTCTAAC - 3'	
HMGB1	Sense: 5' - TCAAAGGAGAACATCCTGGC - 3'	199
	Antisense: 5' - TCA GCC TTG ACA ACT CCC TT - 3'	
hnRNP U	Sense: 5' - TGGTCACTAACTACAAGTGG - 3'	173
	Antisense: 5' - GAGAGTTCTACTTCATCACT - 3'	

2.12 Cell motility assay

2.12.1 Transwell assay

Cell motility was assessed by Transwell Migration Assay kit (BD Falcon, USA). The inserts were put into the 24-well plate and both sides of the membrane of the inserts were coated with collagen I (BD, USA) for more than 2 hours. 36 to 48 hours after transfection, the HepG2 cells were trypsinized and cell number was adjusted to 20,000-30,000 cells per 200 µl volume by cell count. The cell suspension was deposited onto the collagen-coated membrane in each insert in the 24-well plate. The plate was put into 37°C, 5% CO₂ incubator and cultured for 8 hours to overnight. This was needed to allow cell migration to take place, from the upper side to the lower side of the collagen-coated membrane. The medium was then removed and the upside cells of the inserts were gently mopped with cotton bars. The cells at the lower side of the membrane were fixed with 100% methanol for 15 min. The cells were stained with 0.04% Giemsa solution (Sigma Aldrich, USA) for 10-15 min. The membrane was washed with saline solution and dried up. Four fields were randomly selected in each insert for counting and microscopic images were taken. Three independent experiments were carried out for cells

transfected with each type of construct.

2.12.2 Surface coating and fluorescent beads based cell motility assay

The 96-well plates used for the assay were coated with collagen I (BD, USA) (0.3 $\mu\text{g}/\mu\text{L}$) and Blue Fluorescent Beads before seeding the cells. The plate was coated with collagen I first and then the Blue Fluorescent Beads (Cellomics, USA). Blue Fluorescent Beads provided in kit were resuspended in the buffer and centrifuged. The supernatant was removed, and the remaining was washed with the Washing buffer by vortexing, centrifuging and resuspending. Blue Fluorescent Bead solution was transferred to a tube containing 7.5 mL 1 $\times$ Washing buffer, vortexed. 75 μL was added to each well of the 96 well coated plates. The plate was then incubated 1 hour at 37°C in the dark. The plate was washed 5 times with washing buffer and buffer of last wash was left in wells.

HepG2 cells transfected with pcDNA 3.1- HBV genotypes A/B/C were seeded into the wells. 16 to 24 hours after transfection, the cells were trypsinized and seeded into the double coated plates. The cells were diluted with serum-free culture medium to a concentration of 1.0×10^4 cells/mL. Pre-warm the prepared plate to 37°C, Remove the buffer and gently add 50 μL (5.0×10^2 cells) of the cell suspension to the well. Add 50 μL serum-free culture medium for negative controls or 50 μL of 20% FBS in culture medium (for a final concentration of 10% serum) for stimulating the cell motility. Incubate for desired time (around 12hours for

transfected HepG2 cells) at 37°C in 5% CO₂. The results were observed using Olympus IX71 inverted fluorescent microscope.

2.13 Microscopy based real-time cell migration tracking assay

Cell migration tracking of HepG2 Cells co-transfected with (1) pEGFP and empty vector pXJ40, (2) pEGFP and pXJ40-HBx-A, (3) pEGFP and pXJ40-HBx-B and (4) pEGFP and pXJ40-HBx-C were trypsinized and seeded into 35mm culture dish in proper concentration 24 hours post-transfection and cultured for 3 hours letting the cells totally adhered to the culture dish. Then the cells were observed for 16 hours with time-lapse microscopy (Olympus Inc.) at 37 °C and 5% CO₂ realized by the on-line CO₂ incubator attached on the microcopy stage. The photo taking, tracking and data recording and analyzing were completed by using the track object application of Meta Morph 7.0 software. The fluorescent images were automatically taken every 20 minutes in 16 hours. The specific regions of HepG2 cells expressing GFP were recognized by Meta Morph 7.0 with their characteristic pixel appearance and distribution as fluorescent tags. Guided by these tags, the cell movement pathway of each HepG2 cell expressing Green Fluorescent Protein (GFP) was followed and tracked. The cell migration distance in every 20 minutes was measured and recorded, from which the average cell migration velocity was calculated. In each sample, at least 10 cells were selected and the migration data of which were analyzed. Based on the automatically generated data, the real-time cell movement distances were shown and average velocity can be calculated.

Chapter 3 : Genotype Specific SH3 Binding Array Assay

In the previous work of our group [153], a variable proline/serine rich region was identified in generally well conserved non-structural viral HBV protein, HBx. The identification of this region has provided us an important clue in understanding the differences in global proteome profiles generated by transfecting genotype specific HBx. Particularly, the region consists of amino acid residues with the characteristic of the PXXP motif, well known as the SH3 domain binding motif. The PXXP motifs have been implicated in intracellular signal transduction. In recent years, such motifs have been reported in viral perturbations of cellular mechanism as well, including hepatitis C virus (HCV) and human immunodeficiency virus (HIV-1). It is believed that NS5A, non-structural protein of HCV, interacts through its C-terminal PXXP motif with Grb2 [31], potentially resulting in a decrease of the mitogenic signaling pathway leading to inhibition of activating protein 1 (AP-1) [154]. Likewise, the relatively conserved proline rich sequences in Nef protein of HIV has been shown to interact specifically with Hck and Lyn from a tested pool of Src family kinases and these interactions are abrogated when the PXXP motif is removed [155]. The Nef viral protein has also been shown to interact with Nef-associated kinases (NAK), a serine / threonine kinase. Interestingly, inoculation of simian immunodeficiency viruses containing specific proline-alanine mutations into *Rhesus Macaques* showed the presence of revertant viruses [156]. In the case of HBV, the activation of c-Src by HBx has been

reported but the mechanism remains elusive. Hence, the functional importance of the HBx proline-rich region was further investigated in this study.

As shown in the Fig. 3.7, HBx protein with 154 amino acids, showed relatively conserved sequences between different genotypes. The only noticeable variation was the region comprising the proline/serine rich domain. Interestingly, HBx B and C showed the similar composition of amino acids in the proline/serine rich region, with more prolines than serine. On the other hand, residues showed higher proportion in HBx genotype A and D. The primary structure of polypeptides determines the secondary, tertiary structure and quaternary structure in all proteins. Proline is a rather unique amino acid because the α -amino group is secondary. The special structure of proline gives it an exceptional conformational rigidity compared to other amino acids. The characteristic of PXXP where x denotes any amino acids was believed to be recognized as core sequence for the SH3 domain ligand.

To further understand the importance of this region of HBx, this region (around 38 aa) was inserted into pGEX-5X-1 in fusion with GST with additional C-terminal 6×histidine residues. His tag was included to be used as detective tag in the binding array assay. The same region for all genotypes A, B, C and D was cloned. Since the full length HBx has been found to be insoluble in *E.coli* cells, we hypothesized that the partial sequence could be used for the interaction with SH3 binding array. Our hypothesis was supported by the fact that 18 amino acids

proline rich sequence from PAK cloned in fusion with GST had similar binding kinetics to full length PAK for binding with Nck via the SH3 domain and ligand domain interaction [157]. To demonstrate the specificity of these proline residues in HBx interaction with cellular SH3 domain containing proteins, 4 prolines in the proline serine rich motifs (GST-PXXP) were mutated to alanine residues by site mutation.

3.1 Cloning of HBx genotypes A/B/C

Our previous investigation indicated that expression of HBx genotype A, B or C was undetectable when cloned in pcDNA3.1 vector (data not shown) by Western blot. Fragments of HBx A/B/C were then amplified by PCR from the original clonings on pcDNA3.1, and ligated into pXJ40. The results shown in Fig. 3.1 indicated that all three genotypes of HBx have been cloned into pXJ40. DNA sequencing was separately carried out. The expression of HBx was confirmed by Western blot using anti-HA antibody (Fig. 3.2), as HA tag was fused in-frame with the genotype-specific HBx. The result showed that HBx of three genotypes was effectively expressed in HepG2 cells.

Figure 3.1: Cloning of HBx - A/B/C in pXJ40.

Electrophoresis of agarose was shown. DNA samples included HBx of individual genotype cloned in pXJ40, and restricted by enzyme to release the HBx inserts. Lane 4 is the 1kb DNA marker (Gene Ruler, #SM0311). Lanes 1, 2, 3 represented HBx –A, HBx –B, HBx –C on pXJ40 respectively.

Figure 3.2: Western-blot analysis of HBx genotype A, B and C in transfected cells.

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

3.2 Genotyping of HBV genotype D

In addition to HBx genotype A, B and C, we also explored the possibility of

cloning and characterizing HBx genotype D. To differentiate HBx genotype D from the other genotypes, genotyping was carried out. The entire HBsAg coding region including PreS1, PreS2 and S was amplified from viral DNA (extracted from clinical samples) by Taq polymerase. The total length of the PreS1, PreS2 and S gene was 1174bp (Figure 3.3). Result of PCR amplification shown in Fig. 3.3 indicated that not all clinical samples can be amplified by the same primer set. For example, no amplification product was detected in sample 2 and 6. The sequencing was performed using an internal primer for the surface antigen starting from position 155bp.

Figure 3.3: PCR result of entire Surface antigen (1174bp) for genotyping.

Lanes 1, 3, 4, 5 and 7 represented PCR amplification from viral DNA extracted while lane 2 and 6 represent the negative control of PCR. The arrow pointed to the amplified products.

Sequences obtained were then analyzed by Vector NTI as described to detect specific restriction sites for genotyping, a modified RFLP method [158]. Specific restriction sites present in one but not other genotypes included *AlwI*, *NlaIV*, *EarI*,

HphI. For example, genotype A was distinguished from others by the restriction enzyme site for *NlaIV*. Another example was Restriction site for *EarI* exists in all the other genotypes except in genotype B. As for genotype C, restriction site for *AlwI* does not exist. And only genotype E can be digested with restriction enzyme *NciI*. There is no restriction site for *HphI* in genotype F. (Fig. 3.4).

Samples with confirmed genotypes were further used for construction of HBx.

Figure 3.4: Restriction Map created in silico by VectorNTI after sequencing of PCR products from HBsAg.

Various restriction enzymes namely *AlwI*, *HphI*, *EarI*, and *NlaI* were used for genotyping.

3.3 Cloning of HBx genotype D

The coding region of HBx of genotype D was amplified by PCR from the confirmed genotype D clinical sample, and then ligated into mammalian expression vector pXJ40 (Fig. 3.5). Similarity to the cloning of HBx genotype A, B and C, the restriction sites *PstI* and *HindIII* were used for the cloning of genotype D. Results shown in Fig. 3.5 indicated that HBx genotype D was successfully cloned into pXJ40. This was further confirmed by sequencing analysis.

Figure 3.5: Cloning of HBx of Genotype D in pXJ40.

Plasmid DNA was extracted from colonies after transformation. The DNA was restricted by PstI and HindIII and separated on agarose gel. Lane 1 to 5 represented DNA from 5 different colonies.

Western Blot analysis was then performed to examine the expression of HBx of genotype D. Results shown in Fig. 3.6 indicated that the HBx protein of genotype D was correctly expressed in the transfected cells.

Figure 3.6: Western blot result of HBx genotype D.

Lane1 is the negative control, empty vector of pXJ40; Lane2 is the HBx-D-pXJ40. The size of the HBx is 17.5KDA as indicated. The Tag antibody used here is HA (Santa Cruz). β- actin is the internal reference.

3.4 Alignments of HBx of distinct genotypes (A/B/C/D)

After sequencing of the clonings, the nucleotide sequences were translated into amino acid sequences *in silico* for sequence alignment.

Results shown in Fig. 3.7, most of the amino acids of difference belong to the same functional group. Noticeable differences can be observed in the region located at amino acids 29-46.

Figure 3.7: Alignments of amino acids of HBx of genotypes A/B/C/D.

HBx from pcDNA3.1 were cloned into pXJ40 and sequenced. HBx-D from clinical sample was cloned into pXJ40 and sequenced. *In silico* translation into amino acids was performed with BioEdit. Most of the HBX sequences were conserved with the only noticeable difference between 29-46aa. Within this region, different genotypes contained different number of proline and serine residue. This differential proportion of proline and serine residues may possibly play different functional role resulting in disease progression variations. Proline regions potentially have restricted rotational freedom and is non-polar compared to serine.

3.5 Generation of mutations within proline/serine rich domain

These constructs were to be used to hybridize with the SH3 panomics array to identify interacting cellular SH3-containing proteins. To strengthen the significance of the findings, point mutations (Proline to Alanine) were generated on each of these constructs. We have previously cloned proline/serine rich regions of HBx genotypes A, B, C and D into pXJ 40 as a fusion protein with HA tag. As the mutants for genotype A, B and C have been generated, P to A mutant of genotype D was generated. This allowed us to probe the SH3 array with all the four genotypes. The site mutagenesis was obtained by using QuickChange Site-Directed Mutagenesis kit, and the mutation was generated based on the whole plasmids with the plus of primers located at the target mutation sites. Plasmids containing mutations were obtained and were augmented by transformation. The plasmid DNA from individual colonies was first examined by endogenous restriction enzymes digestion and positive clones were confirmed by sequencing. The template used for mutagenesis was shown in Fig 3.8. The plasmid consists of 39 amino acids (aa.19~ aa.58) of HBx containing the proline/serine rich region fused with additional 6×histidine at C-terminal.

Figure 3.8: Cloning strategy of GST fusion protein proline rich region of HBx on pGEX-5X-1.

3.6 Protein expression and purification

Prior to the SH3 binding assay, the fusion protein (Proline/serine rich region/GST/HA) was expressed and purified in *E.coli*. To this end, the fusion insert was obtained by PCR amplification and cloned into GST contained expression vector, pGEX-5X-1. To facilitate the protein purification, the amplified fusion insert was cloned in frame with the GST tag. The purification of the fusion protein was achieved by using the commercial GST purification columns packed with GST binding specific beads. On the other hand, the 6×His tag was needed for the detection in SH3 binding assay. For this reason, another additional 6×His tag was added into the fusion insert located downstream of SH3 binding domain of HBx shown in Fig. 3.8.

The fusion protein was successfully expressed in *E.coli* at 37°C for 4 hours with the induction of IPTG. As shown in Fig. 3.9, the fusion proteins (wild type and mutant) were soluble at 37°C. The crude cellular protein extract was incubated with GST 4B sephora columns. Purification results were shown in Fig 3.11. By the specific GST affinity purification, most of the native expressed proteins of *E.coli* were removed and better hybrid efficiency in the SH3 binding array was obtained. In addition, the purified GST fusion proteins were confirmed by Western blot analysis with the anti-His antibody as shown in Fig. 3.12. The standard BSA (2µg/µl) was used as a roughly quantitation method as shown in Fig. 3.10. Quantity of BSA with known concentration was shown by SDS-PAGE, and the quantity of expressed proteins can be obtained by comparing the density between the BSA

bands and targets protein bands. Such a roughly quantitation can be used for the control of hybridization with the SH3 binding array.

Figure 3.9: Expression of Proline/serine rich domain of HBx genotype D on pGEX-5X-1 vector with 6×His.

Lane 1 indicated protein weight marker (RPN756, Biorad). Lane 2: supernatant (soluble part) of the cell lysis with the expression of fusion protein GST-HBx Proline-rich region-6His. Lane 3: insoluble part of cell lysis with the expression of fusion protein GST-HBx Proline-rich region-6His. Lane 4: supernatant (soluble part) of cell lysis with the expression of fusion protein GST-HBx P to A mutant -6His. Lane 5: insoluble part of cell lysis with the expression of fusion protein GST-HBx P to A mutant -6His. The sizes of the bands of marker are shown in the figure. Probe (GST-PXXP) and its alanine mutant (GST-AXXA) were over-expressed and visualized by Coomassie Blue staining.

Figure 3.10: Gradient concentration of standard BSA (2µg/µl) was used for protein quantification.

Lane 1 is protein weight marker (# 0431, Fermentas). The sizes of the bands of marker are shown in the figure. Lane 2 to lane 10 is 2µg/µl BSA of 1µl to 9µl correspondingly.

Figure 3.11: Purification of Proline/serine rich domain of HBx genotype D on pGEX-5X-1 vector with 6×His.

Purification of GST-fusion proteins GST-PXXP; GST-AXXA purification was performed accordingly. Lane 1 is the protein molecular weight marker (#0431, Fermentas). Lanes 2, 3, 4 are fusion protein of PXXP domain in HBx-D in different concentration and lanes 5, 6, 7 are fusion protein of AXXA domain in HBx-D in different concentration. Lanes 8, 9, 10 are BSA (2µg/µl) with different quantity, 1µl, 2µl and 5µl respectively. The BSA can be used for rough quantification of target proteins.

Figure 3.12: Western Blot Analysis of Expression of GST fusion PXXP and AXXA from genotype D in *E.coli*.

Anti-His antibody was used to detect the C-terminal 6 X His tag.

3.7 Identification of interacting SH3 domain-containing cellular proteins

It has previously been used to show the specific binding of Tec family of kinases, Etk, to the proline-rich domain of p53 in response to DNA damage [159]. To identify cellular proteins that interact with HBx through the PXXP SH3 domain binding, SH3 domain arrays (Panomics, US) were incubated with the purified wild type and mutant HBx. In particular, we subjected our test wild probes (GST-PXXP) and mutant probes (GST-AXXA) to Panomics SH3 domain arrays following the manufacturer's instructions. The expressions of our probes were shown above. The protein showed equivalent and strong expression levels in both the HBx genotype D GST-PXXP and GST-AXXA clones and the proteins were subjected to purification using glutathione sepharose 4B. HBx genotype A GST-PXXP and GST-AXXA constructed previously in our lab were also expressed and purified the same as genotype D probes. After purification, 15 µg of purified probes were then subjected to interactions with the Panomics SH3 domain arrays.

SH3 domains of various identified intracellular proteins were immobilized in

duplicates on the Panomics nitrocellulose membranes. A successful assay was indicated by the positive controls located at the right side and bottom with Histidine-tagged ligand. Positive interactions were indicated by the presence of chemiluminescence signal in both spots. Negative interactions were shown by no chemiluminescence while false positive interaction was indicated if one of duplicated spots showed signal. The interaction results of genotypes A/B/C/D wild type GST-PXXP and mutant GST-AXXA probes were shown in Fig. 3.13 (hybridization of genotypes B and C had been carried out by Dr. Tan Tuan Lin). Specific interactions were identified and boxed. Negative interaction (GST alone, spots D13, 14) and weak interactors were eliminated in the screen. The positions of the spots D15 to D20 represented unloaded material to show specificity of interactions. Row E and Columns 21, 22 contain histidine-tagged peptides spotted for position identity. In the case of binding assay with HBx, the interactions were verified by the use of mutant proline-alanine, GST-AXXA on duplicate membrane. The interacting proteins identified through our screen were summarized in Table 3.1. Our results therefore suggested that the proline/serine rich regions of HBx were responsible for its interaction with SH3 domain-containing cellular proteins. And the comparison among different genotypes HBx showed relatively similar results, but not strictly the same. Such binding differences may imply that different HBx has different potential binding affinity with the intracellular SH3 domain containing proteins.

Figure 3.13: Panomics SH3 domains arrays interaction with HBx probes containing SH3 binding domain from HBx genotypes A/B/C/D.

The left panels showed the specific interactions of PXXP probes, genotypes A, B, C and D respectively, binding to various SH3 domains highlighted in red boxes and showed negative interaction to the GST spots (blue box), a negative control. The right panels showed the specific interactions were verified by the use of the probe with proline-alanine mutant (GST-AXXA), genotypes A to D respectively.

Table 3.1: Identified proteins from SH3 binding array.

Column (#, #)	Domain	Full name	Genotype A	Genotype B	Genotype C	Genotype D
A (5, 6)	Abl2	Abelson-related protein, Arg	√	√	-	-
A(19, 20)	OSF	Osteoclast stimulating factor 1	√	√	√	-
B (7, 8)	Tec	Tyrosine-protein kinase, Tec	√	√	√	√
B (11,12)	PIG2	1-phosphatidylinositol-4,5-biphosphate phosphodiesterase gamma 2	√	√	√	√
B (13, 14)	ARH6	Rho guanine nucleotide exchange factor 6	√	√	√	√
B (19, 20)	EFS	Embryonal Fyn-associated substrate	√	√	√	√
C (13, 14)	RHG4	Rho-GTPase-activating protein 4	√	√	√	√
D (5, 6)	VAV3-D2	VAV-3 protein, SH3 domain # 2	√	√	√	-
D (7, 8)	VINE-D1	Vinexin, SH3 domain # 1	-	√	√	-
D (9, 10)	VINE-D3	Vinexin, SH3 domain # 3	√	√	√	√
D (11, 12)	c-Src	Cellular Rous Sarcoma viral oncogene homolog	√	√	√	weak
D (13, 14)	GST	Negative control	-	-	-	-
A(15,16)	NE-DLG	Presynaptic protein SAP102	√	-	-	√
C(9,10)	Nck2-D2	Cytoplasmic protein Nck2, SH3 domain #2	√	-	√	√
D (3,4)	Vav3-D1	Vav-3 protein, SH3 domain #1	weak	-	-	√

Frame in grey: cytoskeletal related proteins

3.8 Activation of HBV replication by SH3 domain of vinexin β

Previously, we have shown that HBx from genotype B can interact with vinexin β resulting in cytoskeletal change and in turn changing the adhesion velocity of HBx containing HepG2 cells [122]. To further understand the function of interaction between HBx and vinexin β, exogenous SH3 binding domain of HBx and SH3 domain of vinexin β were constructed and respectively co-transfected with HBV into HepG2 cells. These two domains were cloned into mammalian expression vector pXJ40, as shown in Fig. 3.14 and Fig. 3.15. For HBx, selected sequence was the same with the fragment in SH3 binding probe. For vinexin β, the first two SH3 domains which were supposed to be pivotal in protein interactions were selected to be cloned into pXJ40, and the sequences were shown as bellow (). PCR was conducted with designed primers including proper ER sites and the plasmids

with full length HBx and vinexin β were used as templates. The fragments were easily got by PCR as shown in the Fig. 3.14. Followed by enzyme digestion, purification, ligation and transformation, positive colonies were selected amplified and enzyme digestion confirmed clonings (Fig. 3.15) were double confirmed by DNA sequencing.

Figure 3. 14: Schematic diagram of domain organization of vinexin β and protein sequences.

A) The boxes in blue show the SH3 domains. B) Protein sequence of vinexin β , the SH3 domains are in blue and underlined. The underlined “KKRK” is supposed to be a potential nuclear localization signal. C) Amino acid sequence cloned (including the first two SH3 domains).

Figure 3.15: PCR products of SH3 binding domain of HBx genotype B and SH3 domain of vinexin β .

As indicated, lane 1 was DNA ladder and the sizes were shown in the figure. Lane 2 was PCR product of proline/serine rich region (SH3 binding domain) of HBx genotype B, 39 amino acids were included and the PCR fragment is 135 bp including ER sites. Lane 3 was the PCR product of two SH3 domains from vinexin β . 132 amino acids were included and the PCR fragment was 414bp including ER sites.

Figure 3.16: Enzyme digestion results of SH3 binding domain of HBx genotype B (BSH3), left panel, and SH3 domain of vinexin β (VSH3), right panel.

The constructed and transformed colonies were confirmed by enzyme digestion (BamHI and XhoI). The sizes are correct. The positive clonings were followed by DNA sequencing for double confirm.

PXJ40-BSH3 and pXJ40-VSH3 were co-transfected with HBV respectively into HepG2 cells. To observe the effects of HBx SH3 binding domain and vinexin SH3 domain to HBV replication, HBsAg specific ELISA (IMX) and Real time RT-PCR were applied.

IMX system has been widely used in clinical diagnosis in hospital for the identification of chronic HBV carriers for its ability to qualitative and semi-quantitate analysis of the HBV replication activity. In this system, Microparticle Enzyme Immunoassay (MEIA) technology was used for the detection of Hepatitis B surface antigen (HBsAg). As shown in Fig. 3.16, HepG2.2.15, stable cell line derived from HepG2, containing two times head-to-tail of HBV genome [136, 160, 161], could release HBsAg significantly. On the other hand, HepG2 alone showed no release of HBsAg. HBV transfected HepG2 cells can also release HBsAg, but not in a extensive extent. When HepG2 cells were co-transfected with HBV and HBx SH3 binding domain or vinexin β SH3 domains, the secreted HBsAg was significantly elevated, especially when vinexin β SH3 domains was introduced. In support of this, HBcAg expression was also increased as shown by real time RT-PCR (Fig 3.17). Containing the first two SH3 domains of native vinexin β , the introduced SH3 domains may play the same role with the native vinexin β in the co- transfected cells. The strong increase of HBV replication caused by SH3 domains of vinexin β could due to the inadequacy of native vinexin β which can be used to interact with HBV. It may implicate the endogenous vinexin β was not sufficient for the interaction with HBV. The

over-expressed SH3 domains of vinexin β was used as the native vinexin β of cells and interacted with HBV and increased the HBV replication in turn.

Figure 3.17: Secreted HBsAg detection by IMX.

Data were presented as the S/CO ratio (Y-axis), referring to the Sample /Cut-Off Rate, which was autogenerated and provided by the system. Ratio>1 was taken to be reactive and ratio<1 was taken for non-reactive. The controls were reagents commercially purchased together with IMX kit.

Figure 3.18: Real time RT-PCR analysis of HBcAg expression in HepG2 cells.

Beta-actin was used as internal control. HBcAg mRNA level of HepG2 cells transfected with HBV was taken as one fold. In comparison, HBcAg mRNA level in HepG2 cells co-transfected with HBV and HBx SH3 binding domain (BSh3) showed slight increase. And HBcAg mRNA level was significantly boosted up by the co-transfection with vinxin SH3 domains.

3.10 Discussion and Conclusions

In this part, HBx genotypes A/B/C/D were cloned into mammalian expression vector pXJ40, and the expressions were confirmed by Western blot analysis. *In silico* alignment of HBx A/B/C/D showed a proline/serine rich region which was relatively different among genotypes. The proline/serine rich region was deemed as characteristic sequence of SH3 and WW binding motif and involved in a wide range of protein-protein interactions. Such binding motifs have recently been identified in virus proteins.

The identified proline/serine rich region was cloned into bacterial vector and successfully expressed in *E.coli*. The purified SH3 binding proteins were used as probe in the hybridization with the commercial SH3 domain arrays (Panomics) to identify cellular proteins. This should lead to identification of HBx interacting protein through the PXXP SH3 domain binding. As a negative control, mutant SH3 binding regions were also constructed, purified and hybridized with SH3 domain arrays.

A number of SH3-containing cellular proteins were identified in this screening, including vinexin β and β PIX. As shown in the Table (3.1) of proteins, various cytoskeletal modifying and signal transducing related proteins were identified by the HBx interacting probe. Specific interaction of HBx proline-rich region was observed with the identification of control c-Src, previously shown to be important in HBV infection [65, 95]. Negative controls of GST only (D13, 14) and unspotted regions (D15 to D20) denoted the specificity of the interaction screen. Other proteins of interests were cortactin (cortical acting-binding protein), β PIX (RhoGEF) and vinexin (cytoskeleton protein). In the previous work of our lab, we have previously characterized the binding of β PIX [150] and vinexin β [122] via this proline rich and SH3 domain interaction. The interactions between HBx and vinexin β as well as cortactin were confirmed by GST pull-down assay [122]. Interestingly, the activated Rac1 by the interaction of HBx proline rich domain with β PIX has been shown to be critical in up-regulating the replication of HBV [150]. Here, the activation of HBV replication was shown enhanced by IMX assay

by detecting the HBsAg expressed and secreted in culture medium.

Our results therefore suggested that the proline residues of HBx were responsible for its interaction with SH3 domain-containing cellular proteins, and the bindings were abrogated by P to A mutant HBx probes. And the comparison among different genotypes HBx showed relatively similar results, but not strictly the same. Such binding differences may imply that different proline/serine regions in HBx have different potential binding affinity with the intracellular SH3 domain containing proteins.

Chapter 4 : Protein Profile in Cells Transfected with Genotype Specific HBx: a Comparative iTRAQ-coupled 2D LC-MS/MS Analysis

HBV can be classified into eight genotypes from A to H based on 8% DNA sequence divergence [162]. Earlier studies have indicated a clinical linkage between serotypes and vaccine escape mutants [163]. It has also been suggested that particular HBV genotypes may correlate with differences in clinical features of viral infection [164, 165]. Genotypes A, B, C and D are the most prevalent genotypes, with genotype B and C more dominant in Asia and Oceania whereas genotype A and D more prevalent in North, Central America, India and Middle East. At the molecular level, it has been found that different genotypes of HBV showed genotype-specific cellular apoptosis potentially related to methylation [166].

Another important aspect of HBV genotype-disease association is the cytoskeletal reorganization, which in turn impacts on tumor metastasis. The invasion and metastasis of cancerous cells is a multi-step process [167]. So far, most of the cell migration and metastasis candidate proteins have been identified in HCC cell lines or liver cancer clinical samples. HCC frequently invades blood vessels and metastasize [168]. In addition, the degradation of the extracellular matrix (ECM) is an initial step as ECM envelopes HCC tissues cancer invasion thus preventing the

movement of cancer cells [169]. For instance, it has been shown that HBx-induced activation of metalloproteases (MMPs) is associated with the enhancement of cell migration [127, 170], resulting in an increase in ECM degradation and an increase in cell migration.

The proteome of a cell is a dynamic entity which changes with different physiological states of the cell. In addition to our recent application to understand drug-cell interactions [139, 153, 171], proteomics analysis has recently been applied to depict a global map of liver disease onset such as liver cirrhosis and HCC. Our preliminary investigation using 2D gel coupled with MS identification has yielded promising indications on the genotype-associated protein profile in HBx producing cells. However, the technique is not as sensitive and fast as compared with LC-MS/MS analysis which has been applied in our investigation on drug-cell interactions [139, 153, 171]. In this work, we applied the iTRAQ-coupled 2-D LC-MS/MS analysis to reveal and quantify the differences in protein expression levels of normal HepG2 cells and those transfected with HBx of three different genotypes (A, B and C).

4.1 HBx genotypes and their expression in transfected cells

HBx of different genotypes from clinical samples were cloned into high mammalian expression vector pXJ40 and confirmed by sequencing. Variability in genetic changes of such relatively highly reserved protein has been observed previously [153], mainly located between amino acid 29 and 46 (Fig. 3.7). In this

region, HBx genotype B is proline-rich, whereas in HBx A and C, several prolines are replaced by serines. This may make a structural and in turn functional difference among the viral HBx proteins of different genotypes as most of the protein-protein interactions are based on the conformation matching. The proline/serine rich domains interact with SH3 domain or WW domain and regulate a series of cellular processes. For example, we have demonstrated that this region can indeed interact with SH3 domain [122]. Significantly, our ongoing investigation indicated that different genotypes of HBx displayed distinct binding affinity towards SH3 domain containing cellular proteins.

To mimic the early HBV infection and characterize early cellular response to HBV infection specifically associated with HBx, HepG2 cells were transfected with HBx of genotypes A/B/C cloned in pXJ40 vector. Electronic transfection with Amaxa kit was applied to obtain a high transfection efficiency around 70~80% (Fig 4.1) The comprehensive protein profile indicating global cellular events in response to genotype specific HBx can then be analyzed using the high throughput, broad-spectrum and high sensitivity proteomics method iTRAQ-coupled 2-D LC-MS/MS.

Figure 4.1: Transfection of HepG2 cells using Amaxa Nucleofector and monitored by transient transfection of pEGFP vector.

A) phase contrast image of cells 48 hours after transfection. B) The corresponding image of the green fluorescent by pEGFP in the same field (Olympus IX71 fluorescent microscopy).

4.2 iTRAQ analysis and differentially expressed proteins

To ensure the feasibility of the approach, BSA was used to as a standard to show the representative spectrum [139]. To establish the protein profile in the HepG2 cells tansfected with HBx compared with empty vector pXJ40, the iTRAQ was used to label as following: empty vector pXJ 40 = iTRAQ 114; HepG2 transfected with pXJ-HBx genotype A = iTRAQ 115; HepG2 transfected with pXJ-HBx genotype B = iTRAQ 116; HepG2 transfected with pXJ-HBx genotype C = iTRAQ 117. In each of the three independent experiments, the total proteins were extracted from the cells after transfection and labeled with iTRAQ reagents with the same assignments of the isobaric ions. ProteinPilot software was used for the identification and quantification of the protein expression levels in differential samples. Fig. 4.2 showed a representative MS spectrum of the peptide

HVGMAVAGLLADAR from the protein of PSMA3. It showed that the same peptide was up-regulated in the samples labeled by iTRAQ 115, 116 and 117 which was consistent with the change of protein expression level.

Totally, 163 unique proteins were identified in three independent experiments as mentioned above and standard errors shown in table 4.1 were generated by the three independent experiments. Identification and quantification proteins with significant expression levels were based on the amount of ProtScore with the condition that Unused protein score was more than 2.0, which gave the confidence value of 99% and one or more peptide hits were found per protein. Other proteins identified with ProtScore less than 2.0 were not included for further analysis. Among the 163 identified proteins, 24 differentially expressed proteins were chosen here and separated into four major groups according to their cellular functions. 5 metabolic proteins, 10 signaling proteins, 6 cytoskeletal proteins and 3 cell growth related proteins were shown here (Table 4.1).

Table 4.1: List of Differentially Expressed Cellular Proteins.

Prot sc	Protein name and Accession no.	avg.A:N	avg.B:N	avg.C:N	Function ([172] and references therein)
metabolic enzymes					
8	(P15531) Nucleoside diphosphate kinase A	0.86±0.10	0.88±0.02	0.95±0.05	Tumor metastatic process associated protein
6.44	(P50454) Serpin H1 precursor	0.97±0.11	1.25±0.08	1.14±0.06	Binds specifically to collagen. Involved as a chaperone in the biosynthetic pathway of collagen
3.93	(Q06830) Peroxiredoxin-1	0.93±0.04	0.90±0.10	0.84±0.04	Involved in redox regulation of the cell
2.59	(P09914) Interferon-induced protein with tetratricopeptide repeats 1 (IFIT-1)	0.94±0.05	0.80±0.11	0.82±0.01	Induced by interferon, contains 10 TPR repeats
2.01	(P04080) Cystatin-B (Stefin-B)	1.10±0.11	1.15±0.05	1.13±0.07	Intracellular thiol proteinase inhibitor. Tightly binding reversible inhibitor of cathepsins L, H and B
Signaling proteins					

6.38	(P07355) Annexin A2	1.33±0.24	1.12±0.01	1.08±0.01	Calcium-binding protein
6.01	(P22626) hnRNP A2 / hnRNP B1	0.83±0.02	0.94±0.05	0.93±0.07	Involved with pre-mRNA processing
3.45	(P09651) hnRNP A1	0.94±0.02	0.93±0.04	1.00±0.05	Involved in the packaging of pre-mRNA into hnRNP particles, transport of poly(A) mRNA from the nucleus to the cytoplasm and may modulate splice site selection
2.11	(P61978) hnRNP K	0.87±0.05	0.85±0.08	0.99±0.03	Binds to pre-mRNA. Binds tenaciously to poly(C) sequences. Likely to play a role in the nuclear metabolism of hnRNAs
2.94	(Q00839) hnRNP U	1.28±0.11	1.30±0.02	1.19±0.05	Binds to pre-mRNA. Has high affinity for scaffold-attached region (SAR) DNA. Bind to double- and single-stranded DNA and RNA
5.22	(P06703) Protein S100-A6	0.90±0.10	0.89±0.04	0.84±0.05	Calcium-binding protein
3.85	(P31949) Protein S100-A11	0.94±0.04	0.86±0.10	0.86±0.02	Calcium-binding protein
2	(P26447) Protein S100-A4	0.97±0.01	0.95±0.04	0.96±0.01	Calcium-binding protein
2.42	(P09429) HMGB1	1.20±0.04	1.12±0.03	1.29±0.09	Facilitate specific proteins DNA binding and linked to cell proliferation and migration
2	(P12004)Proliferating cell nuclear antigen (PCNA)	0.83±0.04	0.74±0.07	0.88±0.01	Involved in the control of eukaryotic DNA replication
Cytoskeleton proteins					
11.23	(P08670) Vimentin	1.15±0.02	1.06±0.03	1.24±0.11	Class-III intermediate filaments found in various non-epithelial cells
12	(P21333) Filamin-A	1.08±0.03	1.07±0.05	1.06±0.04	Promotes orthogonal branching of actin filaments and links actin filaments to membrane glycoproteins
11.95	(P02545) Lamin-A/C	0.84±0.08	0.88±0.10	0.81±0.10	Components of the nuclear lamina
8.47	(O75369) Filamin-B (FLN-B)	1.05±0.03	1.07±0.05	1.12±0.02	Connects cell membrane constituents to the actin cytoskeleton
2.83	(O43707) α -actinin-4	1.25±0.03	1.14±0.06	1.20±0.07	Anchor actin to a variety of intracellular structures
2	(Q96PK2) MACF1	1.61±0.05	1.53±0.20	1.27±0.11	Play a role in cross-linking cytoskeletal proteins by binding intermediate filaments to the N-terminal plectin repeats and microtubules to the C-terminus
Cell growth proteins					
6.01	(P23921)Ribonucleoside-diphosphate reductase large subunit	1.10±0.03	1.11±0.04	1.13±0.20	Provides the precursors necessary for DNA synthesis
2	(P25788) Proteasome subunit alpha type 3	1.30±0.04	1.05±0.05	1.43±0.21	ATP-dependent multicatalytic proteinase
2	(P05386) 60S acidic ribosomal protein P1	1.12±0.04	1.22±0.02	1.17±0.07	Plays an important role in the elongation step of protein synthesis

A: N was the ratio of different protein expression level in the protein extracts from the cells transfected by HBx-A to the negative control (protein extracts from the cells transfected by empty pXJ40); B:N was HBx-B to negative control; C:N was HBx-C to the negative control. The standard errors were generated from three independent experiments, and in each experiment the protein extractions were prepared and labeled with iTRAQ

reagents with the same arrangements of the isobaric ions. The description of protein functions was obtained from Expsy online database [172].

Figure 4.2: A representative MS/MS spectrum showing a peptide, HVGMAVAGLLADAR, from PSMA3.

The ion assignments were as follows: pXJ40: iTRAQ 114; pXJ40-HBx-A: iTRAQ 115; pXJ40-HBx-B: iTRAQ 116; pXJ40-HBx-C: iTRAQ 117.

Figure 4.3: The differential expression level of several selected proteins.

The four columns of each group represented samples from HepG2 cells transfected with empty pXJ40 vector (control), HBx of genotype A, B and C respectively. The normalized ratio between the HBx A, B, C transfected and control was shown.

4.3 Validation of differentially expressed proteins by real-time RT-PCR

To confirm the differentially expressed proteins detected by LC-MS/MS, Real-time RT-PCR was performed to quantify the expression at mRNA level of selected genes (Fig. 4.4). Significant changes ($p<0.05$) were analyzed based on three independent experiments for each transfection. In each one, the total mRNA were extracted and used as templates. Overall, the validation results were consistent with the LC-MS/MS data (Fig. 4.3). MACF1, Annexin A2 and PSMA3 showed similar up-regulation and S100A11, S100A4, and S100A6 showed down-regulation. However, hnRNP showed up-regulation in HBx genotype B while in genotypes A and C, it was down-regulated. Another inconsistency

between real-time RT-PCR and LC-MS/MS data was on S100A11, which was up-regulated in cells transfected with HBx genotype C, but down-regulated in all the three genotypes in LC-MS/MS result. The inconsistency may due to the post-translational modification and regulation.

Figure 4.4: Real-time PCR analysis of mRNA levels of each gene.

Three experiments were carried out. The change folds were compared with the control (HepG2 cells transfected with empty vector pXJ40).

4.4 Validation of differentially expressed proteins by Western blot

In addition to the mRNA level validation by RT-PCR, the changes of proteins were also confirmed by Western blot analysis. HepG2 cells were transfected, two days post transfection, the cell lysate were collected by lysis buffer and centrifugation, and the proteins were quantitated, the same with the part of preparation for LC-MS/MS analysis before labeling. And the same quantity of proteins was analyzed by Western blot with first antibody β -actin. As shown in Fig. 4.5 the last

batch, β -actin showed the same quantity. In the meantime, overexpressed HBx A/B/C were also analyzed by Western blot and the result showed similar expression level. Based on the quantitation of β -actin and HBx, the responding amounts of the samples were used for the analysis of the other proteins. Fig. 4.5 shows the results of S100A11, PSMA3 and Annexin A2. For S100A11, HepG2 cells transfected with pXJ40 shows the strongest expression and it was down-regulated by HBx A/B/C. In the case of PSMA3 and Annexin 2, the expressions of these two proteins were up-regulated by the HBx A/B/C. The trends of differentially changes of these proteins were consistent with the results of LC-MS/MS. Together with the results of Real-time RT PCR, the differentially protein expressions in HepG2 cells were double confirmed by classical molecular biology experiments.

Figure 4.5 Western blot confirmations of the differentially expressed proteins in HepG2 cells transfected with HBx A/B/C.

The expressions of HBx A/B/C were in comparative levels among three genotypes. Beta actin was internal control. PSMA3 and Annexin 2 were up-regulated while S100A11 was

down-regulated by HBx compared with negative control. All the antibodies were from Santa Cruz (US).

4.5 Cell migration assay

The cell movement is a complex and highly organized behavior involving not only actin but also a large group of actin-binding proteins. The proteins identified in our study, MACF1, HMGB1, Annexin A2 were up-regulated and Lamin A/C was down-regulated in HepG2 cells transfected with HBx. Their up or down regulations may impact on the cytoskeleton organization, and cell migration. To assess whether the expression changes of these proteins correlates with the migration of cells transfected with genotype specific HBV genomes, cell migration assays were carried out by using the methods of Cellomics fluorescent beads and microscopy based real time cell movement tracking. In the migration assay by fluorescent beads, the surface was coated with collagen and followed by fluorescent beads, and cells were seeded onto the prepared surface at a rather low concentration. After several hours' incubation, the movement of a single will leave a track on the surface. The cell migration track was indicated by black trace. Result shown in Fig. 4.6 indicated that cells transfected with replicative genotype-specific HBV genomes displayed increased migration. Further more, the cell migrations affected by genotype-specific HBx were also observed by real-time microscopy tracking. In this assay, the HepG2 cells were co-transfected with Green Fluorescent Protein (GFP) and genotype-specific HBx or empty vector which was used as control. The cells expressing GFP have a great contrast with other cells and the

background which facilitated the tracking of the green cells by the software. As shown in Fig 4.7, obviously, during 16 hours observation 24 hours post-transfection, the cells transfected with HBx showed vigorous movement compared with cells transfected with empty vector. Besides, the cells with genotype-specific HBx showed different cell movement abilities. The average velocity of cells with HBx-A was faster than cells with HBx-B and HBx-C. Fig. 4.8 shows the overlay of fluorescence images at time zero and 16 hours after incubation of the migration tracks.

Figure 4.6: Fluorescent beads based cell motility assay.

(A) the normal HepG2 cells; (B) HepG2 cells transfected by HBV genotype A; (C) HepG2 cells transfected by HBV genotype B; (D) HepG2 cells transfected by HBV genotype C. The write arrows in the plates with FBS showed typical tracks of the cell movements. There is no cell movement in the plates without FBS.

(A)

(B)

Figure 4.7: Microscopy based real time cell migration assay.

Cell migration distance (A) and average velocity (B) assay of HepG2 Cells co-transfected with empty pEGFP and empty vector pXJ40, HepG2 cells co-transfected with pEGFP and pXJ40-HBx-A, HepG2 cells co- transfected with pEGFP and pXJ40-HBx-B and HepG2 cells co- transfected with pEGFP and pXJ40-HBx-C. The cells were observed for 16hours

(960min) 24 hours post-transfection. Each error bar represented the standard error of the data calculated from at least 10 cells. A) The real-time cell migration distance at every single time spot. B) Average cell migration velocity of each sample in 16hours.

Figure 4.8: The overlay of cell migration fluorescence images.

The overlay of fluorescence images (at time 0 and 16 hours) of the migration tracks (the redline) of pXJ40, HBx A, HBx B and HBx C transfected HepG2 cells 24 hours post transfection.

4.6 Discussion and Conclusions

The changes in the level of cellular proteins detected by our LC-MS/MS proteomics analysis have been validated by real time RT-PCR. The results were based on three independent experiments. Most of the encoding gene selected in RT-PCR showed consistent changes in their mRNA levels as compared with the changes detected by LC-MS/MS analysis.

Further validation by Western blot provided further support on the changes detected by LC-MS/MS analysis. However, it was found not all seven proteins selected in RT-PCR were suitable for Western blot analysis. In the case of MACF1, with a molecular weight of more than 600kDa, the protein can not be transferred onto the membrane (PVDF or NC). No signal was detected in the Western blot

analysis of hnRNP U, HMGB1 and Lamin A/C. This may be due to the low concentration or non-specific influence of the antibodies.

Many cytoskeleton proteins shown in Table 1 were up-regulated by the transfection of HBx, except Lamin-A/C. The main function of filamin is to connect cell membrane constituents to the actin cytoskeleton, whereas α -actinin anchors actin to a variety of intracellular structures. The up-regulation of these cytoskeletal proteins by HBx may reveal a strong cell movement tendency. On the other hand, vimentin, a major component of intermediate filament, has been found to be over-expressed in prostate cancer [173].

Lamin found to be down-regulated by HBx (Table 1) consists of A and C isoforms which are two alternatively spliced products of the LMNA gene. Three types of lamin exist in human cells: Lamin A, Lamin B, and Lamin C. The amounts of Lamin A and C are present in equal in the lamina of mammals. It is believed that it is lamin A/C but not lamin B1 perform the regulation of nuclear mechanics [174]. The defects of LMNA gene have been associated with various diseases, such as Hutchinson-Gilford progeria syndrome, Werner syndrome, and familial atrial fibrillation. Reduced lamin A/C expression has been found to be a signature of a variety of cancers with cells showing a low degree of differentiation and a high degree of proliferation. Conversely, the increased expression of lamin A/C has been associated with the decrease of cell motility, and increase of focal adhesion [175]. Taken together, the down-regulation of lamin A/C by HBx may contribute to

the decrease in focal adhesion and an increase in cell motility.

In the group of cytoskeletal proteins, microtubule-actin cross-linking factor 1 (MACF1, also referred to as ACF7) was found to be up-regulated in the cells transfected by HBx, especially of genotype A and B. MACF1 is a broadly expressed cytoskeletal crosslinking protein, a member of specktraplakins family proteins which are enormous (>500kDa) proteins able to bind both microtubules (MTs) and actin filaments (FAs) directly. It has been found that MACF1 is an essential integrator of actin-dependent MT dynamics in cytoplasm in specialized directed movements and the MTs are less stable in MACF1 null cells [176]. Recently, it has been suggested that MACF1 regulates focal adhesion complex and cell migration. MACF1 may facilitate the integration of MT and FA dynamics, and in turn enhances cell migration [177]. In the context of HBV infection, the cytoskeleton change of the cells affecting cell adhesion and migration may play an important role in the development of HCC. Our finding on the up-regulated MACF1 in HBx transfected HepG2 cells leading possibly to an increased cell migration was consistent with the reported association between HBV and HCC.

In the group of signaling proteins, several including High Mobility Group Box 1 (HMGB1) and Annexin A2 were up-regulated. The HMGB family of proteins, consisting of three proteins (HMGB1, HMGB2, HMGB3), play an essential roles in transcription regulation, inflammation and metastasis [178]. HMGB1 is ubiquitous in mammalian cells and mostly localize in nucleus. Over-expression of

HMGB1 during carcinogenesis has been correlated with the progression of squamous cell carcinoma in skin [179]. HMGB1 can also promote cell migration. It is shown that vascular smooth muscle cells and fibroblasts respond to extracellular HMGB1 in the chemoattraction assays [180]. The role of HMGB1 in cell migration and metastasis has been supported by the suppression of metastasis formation, associated with the administration of anti-HMGB1 antibodies, from tumor cells implanted under the skin of recipient mice [181]. Additionally, the ability of invasion through the activation of extracellular proteases is crucial in cell migration progress. By binding to several components of the plasminogen activation system [182], enhancement of the activation of t-PA [182], and activation of metalloproteases MMP-2 and MMP-9 [181], HMGB1 implicate itself in the plasmin activation cascade, which is complemented by degradation of HMGB1 caused by activation of proteases via a feedback mechanism.

The up-regulated Annexin A2 from our study, a protein belonging to a widely distributed super protein family, is closely related to calcium and membrane binding. The up-regulation of Annexin A2 has been found in several types of tumor, gastric adenocarcinoma [183], colorectal adenocarcinoma [183, 184], kidney cancer [185], pancreatic adenocarcinoma [186], and HCC [187]. Interestingly, the up-regulation of Annexin A2 has been reported in the HBV replicative cell line HepaRG recently[188]. This is consistent with our results on the up-regulation of Annexin A2. Annexin A2 also regulates plasmin generation [179, 189-193], which has been implicated in many biological processes involving migration, such as

tissue remodeling, angiogenesis, tumor invasion and progression [194, 195]. Taken together, the up-regulation of Annexin A2 by HBx may well be involved in the development of HCC.

Besides the protein highly related to cell migration and cytoskeletal changes, some other proteins were also found differentially changes in the protein profiles such as calcium binding protein S100 family proteins, proteasome subunit alpha type 3 and heterogeneous nuclear ribonucleoproteins.

The S100 proteins are members of the calcium-binding EF-hand motif family including 20 known human members. S100 proteins have a wide range of cellular functions, both within and outside the cells. S100A4 was shown to be related with decreased expression of tumor suppressor genes like p53 [196] and that may explain the possible function of S100A4 in tumorigenesis. It has been found that it is increased in HCC cell lines [197]. S100A11 seems to exhibit different roles depending on the tumor types, with a role in tumor suppression in bladder carcinoma and renal carcinoma while promoting breast cancer. The down-regulation of S100A11, S100A4 and S100A6 found in our study (Table 4.1) was in line with the results published recently by analyzing the HBV replicative cell line HepG2 2.2.15 [198], albeit lesser in magnitude.

Proteasome subunit alpha type 3 (PSMA3) was found to be up-regulated especially in cells transfected with HBx genotype A and C (Table 4.1). The proteasome is a

multi-protein complex that modulates cell proliferation, apoptosis, angiogenesis and motility by regulating the stability of the proteins. It has been found that abnormal activity of proteasome is implicated in colorectal cancer [199]. The proteasome subunit Adrm1 was over-expressed in colorectal cancer compared with the normal mucosa and the knockdown of this subunit caused cell migration and cell proliferation arresting [200]. Similarly, high activity ubiquitin-proteasome pathway was detected in both patient samples and the pancreatic cancer cell line [201]. Moreover, the proteasome inhibition represses NF- κ B and decreases cell migration of lens epithelial cells, and such an inhibition is related with down-regulation of MMP-2 and MMP-9 [202]. The up-regulation of proteasome subunit in our result was consistent with the findings that the ubiquitin-proteasome pathway was stimulated in the liver tumors [203]. It has been discovered that the proteasome functions as a potential cellular target of HBx [204]. HBx has been reported to co-localize with proteasome [205], and interact with PSMA1, PSMA7 and PSMB7 [206]. Proteasome has also been found to be important in regulating the processes of cell growth and formation of cancer [207, 208]. Taken together, the up-regulated PSMA3 by HBx may contribute to the versatile nature of HBx [209] in tumorigenesis.

HnRNP U was found to be up-regulated in the cells transfected with HBx genotype B (Table 1). The down-regulation of hnRNP A2/B1, hnRNP A1 and hnRNP K was in line with recently published result in HepaRG cell line [188]. Our results further proved the down-regulation was associated with genotype specific HBx. The

heterogeneous nuclear ribonucleoproteins (hnRNP) are a big family of proteins having various cellular functions. They are involved in tumor development including cell invasion, angiogenesis [210]. It has been reported that hnRNP K can bind to and modulate the replicative efficiency of HBV. The over-expression of hnRNP K augments HBV replication, while silencing of endogenous hnRNP K results in a significant reduction of HBV viral load [66]. The role of hnRNP U in viral replication by a recent report showing that its N-terminal region specifically targets the 3' long terminal repeat (3'LTR) in the HIV-1 mRNA and block the cytoplasmic accumulation of the viral mRNAs [211]. Further functional analysis should shed new lights on the impact of up-regulation of hnRNP U on HBV replication.

In our result, the expression change folds of most proteins were different in HBx genotypes A, B and C transfected HepG2 cells. Among the 24 proteins listed in Table 1, 11 proteins showed the most remarkable changes by HBx genotype B, 5 proteins by HBx genotype A and 9 proteins by HBx genotype C. Such a difference may be caused by the sequence differences between genotypes. HBx of genotype B and C contained the most proline residues while genotype A contained more serines which may influence the binding affinity to SH3 domain containing proteins. Interestingly, genotype-specific HBV has variant outcome inclines in clinical cases. In our cell migration assay, the HepG2 cells with HBx-A showed stronger movement than genotypes B and C, which may provide some clue in HBV infection disease prognosis related to HCC. This genotype specific difference

has also been reported in the interaction of HBx and Pin1 [98, 212] as well as in our ongoing investigations. These include the specific interaction between HBx genotype B and vimentin β [150], and the repression of tumor suppressor gene by HBx genotype D [166].

In summary, our results suggested that the application of iTRAQ 2D LC-MS/MS based quantitative proteomics is a useful and efficient approach to investigate the protein profile influenced by HBx expression in HepG2 cells. A series of differentially expressed proteins were identified, some of which were highly related with cell migration. The integrated effects of these differentially expression proteins may provide valuable information on the underlying mechanism of the function of HBx in HCC, as well as on HBV-associated HCC.

Chapter 5 : Extracellular Protein Profile of HepG2 Cells Affected by HBV

In recent years, proteomics analysis has been applied to cancer and other disease research with the combination of mass spectrometry and proteinchip [213]. The overall objectives are to gain better understanding the cancer pathogenesis, tumor biomarkers for diagnosis and early detection of human diseases. Biomarkers identified by proteomics analysis can be helpful in monitoring the progression process of cancer and therapeutic efficiency.

It is difficult for early diagnosis of cancers is the absence of specific symptoms. For example, the cancer antigen (CA) 15-3 which is one of the tumor markers in blood for breast cancer is not useful for early diagnosis because of low sensitivity [214]. Therefore, the quantitative detection of HER-2 and carcinoembryonic antigen (CEA) in abnormal nipple discharge have been introduced for early diagnosis of breast cancer in some countries [214]. In the case of HCC, ultrasonography and α -fetoprotein (AFP) have been applied for the screen of high risk patients in early detection. However, AFP has been found to be associated with late stage of HCC, thereby limiting its usefulness as a prognostic biomarker.

It has been well established that chronic HBV infection leads to HCC development. However, the symptoms of these end stage diseases take years to show up.

Therefore, the finding of biomarkers which are correlated with long-term disease outcomes may become one of the powerful tools in early detection of possible severe prognosis.

By using hepatocyte cell based system and 2D coupled LC-MS/MS, we aimed to identify biomarkers for HBV infection/replication associated disease onset. Replicative HBV genome was transfected into HepG2 cell line and the secreted proteins under the replication of HBV were analyzed by LC-MS/MS.

5.1 Secreted protein profile of HBV replicating HepG2 cells

Replicative HBV genomes of genotypes A/B/C were individually constructed in our lab with a 1.1 times of HBV genome cloned in pcDNA3.1 mammalian expression vector upstream of the promoter. Such a strategy allows proper expression of HBV gene by using its own promoters [215]. Replicative HBV genome was transfected into HepG2 cells by electronic transfection (Amaxa), which shows high transfection efficiency (around 70~80%) [150]. The transfected cells were incubated for 2 days at 37°C with 5% CO₂. The culture medium was collected and cell debris was cleared by 5min centrifugation. The supernatant was concentrated by 5kDa ultra-filter concentrators and preceded with protein quantitation. 100ug maximum proteins of each sample was precipitated by acetone and labeled with iTRAQ reagents.

32 unique extracellular proteins in total were identified in three independent experiments. Most of the proteins were down-regulated compared with the negative control, HepG2 cells transfected by the empty vector pXJ40. The identified secreted proteins were shown in Table 5.1. Fig. 5.1 showed the MS spectrum of peptide from EQLGPVTQEFWDNLEK from Apolipoprotein A-I precursor. The change of this protein was consistent with the change fold shown in the Table 5.1.

Table 5.1: List of Differentially Expressed Extracellular Proteins affected by HBV.

Protein name and Accession no.	avg.A:N	avg.B:N	avg.C:N	Function ([172] and refernces therein)
(P01023) Alpha-2-macroglobulin	0.92±0.06	0.96±0.03	1.09±0.01	Inhibitor of all four classes of proteinases
(P02647) Apolipoprotein A-I	0.86±0.01	0.84±0.06	0.83±0.03	Participates in the reverse transport of cholesterol from tissues to the liver
(P02765) Alpha-2-HS-glycoprotein	1.11±0.03	1.20±0.07	1.25±0.02	Promotes endocytosis, possesses opsonic properties and influences the mineral phase of bone. Shows affinity for calcium and barium ions
(P19823)Inter-alpha-trypsin inhibitor heavy chain H2	1.04±0.03	1.07±0.02	1.06±0.04	Acts as a carrier of hyaluronan in serum or as a binding protein between hyaluronan and other matrix protein
(P01024) Complement C3	0.95±0.11	1±0.06	0.90±0.03	Important in the activation of the complement system

A: N was the ratio of different secreted protein expression level from the cells transfected by HBV-A to the negative control; B:N was HBV-B to negative control; C:N was HBV-C to the negative control. The standard errors were calculated from three independent experiments. The protein extractions were prepared and labeled with iTRAQ reagents with the same arrangement of the isobaric ions in each experiment. The description of protein functions was obtained from Expasy online database [172].

Figure 5.1: A representative MS/MS spectrum showing a peptide, EQLGPVTQEFWDNLEK, from Apolipoprotein A-I precursor.

The ion assignments were as follows: pcDNA3.1: iTRAQ 114; pcDNA3.1-HBV-A: iTRAQ 115; pcDNA3.1- HBV-B: iTRAQ 116; pcDNA3.1- HBV-C: iTRAQ 117.

5.2 Discussion and Conclusions

Among the secreted proteins identified by our LC-MS/MS analysis, apolipoprotein A-I and α 2-HS glycoprotein (AHSG) are of high relevance due to their roles in HCC.

AHSG, which is known as fetuin, is a glycoprotein synthesized by hepatocyte. It was found expressed in high levels in serum and mineralized bones [216]. AHSG was believed to be a negative acute-phase protein which showed decrease in serum

after burns, trauma, surgical procedures and severe inflammation. It has been found that AHSR is up-regulated in the serum of HCC patients, and the treatment of Radiofrequency ablation treatment (RFA) reduces its level [216]. On the other hand, decreased level of AHSR has been linked to clinical high mortality rate of liver cirrhosis [217, 218]. Consistent with the published reports, AHSR was found to be up-regulated by HBV replication. It is likely that AHSR is up-regulated at the early stage may have the potential to be a prognostic marker for HCC development.

Another identified protein apolipoprotein A-I is synthesized in small intestine and liver. It is a major component of High density lipoprotein (HDL). It was believed to participate in the reverse transport process of cholesterol from tissues to the liver for excretion. This protein was down-regulated in our study. Other members of the same family, apolipoprotein E and M were found elevated in HCC [159, 219]. The role of apolipoprotein A-I in HCC development was still not clear.

In addition, complement C3 was found down-regulated albeit marginally in our study. As major components of complement system, the down regulation may represent an immune inhibition by the HBV replication. It was observed a fragment of complement C3 was down-regulated in HCC [220].

Compared to intracellular and serum protein profile, the number of secreted proteins in cell culture medium was much less in which rendered the protein

quantitation difficult. The low abundance of protein may also affect the p value. In the future, more samples need to be analyzed to substantiate our findings. Meanwhile, a better protein quantitation may be achieved by adding standard proteins for calibration. Similarly to our analysis on intracellular profile in HBV replicating cells, Western blot analysis should be conducted to confirm the differential expressions identified in LC-MS/MS.

Biomarkers of cancer have so far been identified in clinical samples. However, such efforts have their shortcomings in particular the intrinsic inter-personal variations. It is to be hoped that our approach should provide candidate biomarkers for the HCC and liver cirrhosis diagnostic and therapeutic tools. Our cell line based approach should provide a useful alternative of the common approach, as the background is much more homogenous compared to clinical samples.

Chapter 6 : HBV Activated Rho GTPases Induced Cell Motility

Rho GTPases play important role in cellular signal transduction and are major regulators of actin dynamics, gene expression, cell cycle and so on, the dis-regulation of which may cause tumorigenesis. Recent year's research of tumors and cancer models showed that Rho GTPases are involved in the initiation and progression of many kinds of cancers, including HCC.

Tumor progression is a complex multistage process, making the cells from normal to malignant by a series of alterations including cell adhesion and movements. One of the important characteristics of the tumor cells is that they have the ability to escape from the normal controls during proliferation, differentiation, apoptosis and migration [221]. In metastasis, tumor cells detached from their place of formation, move and form new tumors in distant healthy tissues.

Cell migration is a key activity of many important normal and abnormal biological processes, including tumor cell metastasis [222]. Invasion and metastasis are essential hallmarks of malignant cancer cells. The processes that tumor cells adhere to host cell layers and subsequent migrate transcellularly are important in cancer invasion and metastasis, which involves the extracellular matrix (ECM) degradation [125, 223, 224]. High motility and cell migration might be the

premonition of malignancy of chronic viral infection.

The Rho small GTPases, which contains subfamilies of Rho, Rac, Cdc42, regulates cell migration by reorganizing the actin cytoskeleton, which is considered as a driving force of cell motility [225, 226]. The Rho GTPases act as molecular switches cycling between inactive GDP-bound forms and active GTP-bound forms [227]. Such cycling is regulated by guanosine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs) [228]. The Rho family has been implicated in many steps during the cellular transformation. In addition, Rho GTPases activation has been linked to virus replication such as HIV-1, HSV (Herpes Simplex Virus) and HBV [150, 229, 230]. Rho A, Rho C, and Rac1 has been reported significantly elevated in a variety of tumors, especially in the more metastatic forms [231, 232].

During the tumor progression, the loss of the cell characteristics results in morphological changes, such as cell-cell adhesion, gene expression and motility rate. For example, the adhesion velocity has been shown to be affected by HBV replication [122]. Our laboratory has recently reported that HBV-replicating HepG2 cells display similar morphology with those expressing the constitutively activated Rac1, mediated by HBx [150]. We have also shown that HBV can activate Rac1 GTPases and induce membrane ruffling. HBx interacts with β PIX directly, which is a Rac1 nucleotide exchange factor (GEF), by the our identified SH3 binding region and affects the cell morphology [150]. Our data are in line

with earlier reports on the potential invasion ability of HBx through indirect activation of matrix metalloproteinase (MMPs) [233, 234].

Another member of Rho GTPases, RhoA was shown to be required for cell invasion of MM1 rat hepatoma [235]. The over-expression was detected in clinical researches and over-expression of RhoA was shown correlated with tumor progression and metastasis in liver cancer [236, 237]. Moreover, besides the over-expression, the activity of RhoA was also found to be elevated in tumor tissues [238]. In the study of HCC mouse model, it was found that tumor formation was accelerated by the induction of constitutive active mutant RhoA V14 [239].

In this work, we try to find out if Rho GTPases including Rac1, Cdc42 and RhoA were involved synergistically with HBx in the migration of HBV-replicating cells by using transwell and microscope based cell migration assay.

6.1 Rac1 activated high mobility of HepG2 cells infected by HBV

HepG2 cells were transfected using Nucleofector to achieve high efficiency of transfection, as indicated by the high percentage of cells expressing EGFP (See Fig. 4.1 in chapter 4).

The expression of HBsAg from the replicative genome in transfected HepG2 cells was assessed by RT-PCR (Fig. 6.2), whereas the expression of wild type Rac1,

Rac1^{G12V}, Rac1^{T17N}, Cdc42 wt, Cdc42^{G12V}, Cdc42^{T17N}, RhoA wt, RhoA^{G14V} and RhoA^{T19N} were determined by Western blot analysis with anti-HA antibody. Beta-actin was the internal control (Fig 6.1). It indicated that all nine constructs of Rho GTPase were successfully expressed in HepG2 cells at comparable level. In addition, both HBV surface antigen (HBsAg) and core antigen (HBcAg) were detected in HepG2 cells transfected by replicative HBV genome (lane 2, Panel B, Fig. 6.2) and P-A mutant HBV (lane 3, Panel B, Fig. 6.2), with the empty vector pcDNA3.1 used as control (lane 1, Panel B, Fig. 6.2).

Figure 6.1: Western blot analysis of Rac1, Cdc42 and RhoA expression

A) Expression of Rac1 wt, Rac1^{G12V} and Rac1^{T17N}; B) Cdc42 wt, Cdc42^{G12V} and Cdc42^{T17N}; C) Expression of RhoA wt, RhoA^{G14V} and RhoA^{T19N}; The internal control was β-actin.

Figure 6.2: Expression of HBsAg and HBcAg by RT-PCR.

HBsAg and HBcAg were strongly expressed in wild type HBV replicative genome (Lane 2) and P-A mutant HBV genome (Lane 3). The negative control was HepG2 cells transfected by empty vector pcDNA3.1. β -actin was the internal control (Lane 1).

The motility of HepG2 cells expressing Rac1 alone, and those expressing both Rac1 and replicative HBV genome was then analyzed by transwell assay, as indicated by the number of cells migrated to the lower side of the collagen-coated membrane (see Materials and Methods). Seven sets of constructs were used in this analysis: wild type Rac1 alone; constitutively activated Rac1^{G12V} alone; dominant negative Rac1^{T17N} alone; wild type Rac1 and wild type HBV genome; wild type Rac1 and P-A mutant HBV genome; wild type HBV genome alone; and P-A mutant HBV genome alone.

The cells transfected with various plasmids were seeded into the inserts which could be located in the 24-well plate. Some of the cells would transfer from the seeded side to another side of the insert after several hours' cell culture. The number of transferred cells can be calculated by numbering the cells on the other side of the insert after mopping off the cells on the seeding side. Figure 6.3 showed

the pictures of the cells which had transferred to the other side of the inserts. The cells were fixed and stained before taking pictures and cell numbering. Fig 6.3 showed views from each wells transfected with the plasmids respectively. The calculated and normalized transferred cells were shown in Fig 6.4. Results shown in Fig. 6.3 indicated that HepG2 cells displayed differential motility depending on the transfected constructs, with the highest number of cells (indicated by arrow heads) in cells transfected either with Rac1^{G12V} alone (panel B) or with wild type Rac1/wild type HBV genome (panel D). A graphical representation of the motility assay was summarized in Fig. 6.4 after cell counting and analysis. HepG2 cells transfected with the constitutively activated Rac1^{G12V} showed a 2.5 fold increase in cell motility (lane 2, Fig. 6.4) as compared with those expressing wild type Rac1 alone (lane 1, Fig. 6.4) or those expressing the dominant negative Rac1^{T17N} (lane 3, Fig. 6.4). Interestingly, cells co-transfected with wild type Rac1 and replicative HBV genome showed similar motility as those transfected with constitutively activated Rac1^{G12V}. This suggested that the cell motility was enhanced by HBV replication, likely via activation of wild type Rac1 to its GTP-bound state as recently reported. To validate our findings, the P-A mutant HBV genome (with four proline residues mutated to alanine residues in the proline rich domain) was co-transfected with the wild type Rac1. Results shown in Fig. 6.4 (lane 5) indicated that the P-A mutant HBV genome abrogated the cell motility significantly, to a level comparable with cells expressing. Meanwhile, the expression of replicative HBV genome itself enhanced the cell migration also (Lane 6), which may imply the endogenous Rac1 activation by the HBV

replication. Not surprisingly, the P-A mutant HBV genome abrogated the cell motility (Lane 7). But the cell migration level caused by endogenous Rac1 was less than that caused by the over-expressed Rac1 V12, which suggested that endogenous wild type Rac1 was not sufficient for the activation caused by HBV replication. And such activation was confirmed by the augmentation of cell movements induced by over-expression of Rac1 wt. In other words, Rac1 was activated by HBV replication. Although the amount of endogenous Rac1 was enough for the activation and in turn induce high cell motility. However, over-expressed Rac1 could even augment the activation. [240]

Figure 6.3: Transwell cell motility assay.

Microscopic examination of cells (red dots by staining) migrated from the upper side to the lower side of the collagen-coated membrane (Olympus IX 71 microscope). Representative cells were indicated by arrow heads. A) HepG2 cells transfected by wild type Rac1; B) HepG2 cells transfected with Rac1G12V; C) HepG2 cells transfected by Rac1T17N; D) HepG2 cells transfected with wild type Rac1 and wild type HBV genome; E) HepG2 cells transfected with Rac1 wt and P-A mutant HBV genome; F) HepG2 cells transfected with wild type HBV genome; G) HepG2 cells transfected with P-A mutant

HBV genome.

Figure 6.4: Graphical representation of transwell cell migration assay.

Lane 1, motility of cells expressing wild type Rac1 alone (indicated as Rac1 wt); Lane 2, cells expressing constitutively activated Rac1^{G12V} (indicated as Rac1 V12); Lane 3, cells expressing dominant negative Rac1^{T17N} (indicated as Rac1 N17); Lane 4, cells co-expressing wild type Rac1 and wild type replicative HBV genome (indicated as Rac1 wt&HBV wt); Lane 5, cells expressing wild type Rac1 and P-A mutant replicative HBV genome (indicated as Rac1 wt&HBV mutant); Lane 6, cells expressing wild type replicative HBV genome (indicated as HBV wt); Lane 7, cells expressing P-A mutant replicative HBV genome (indicated as HBV mutant).

To further elucidate the function of HBx in the promotion of cell migration. Microscopy based real time cell migration assay was performed. As shown in Fig. 5.5, 24 hours after transfection, the cells were trypsinized and seeded into 35mm cell culture dish in a concentration of 10^5 /ml. The cells were incubated for 4 hours to allow adhesion and followed by real time tracking for 16hours incubated in the

microscopy integrated cell culture chamber providing 37°C, 5% CO₂. As shown in previous result, Fig. 5.5, the cell migration was enhanced by HBx. Here, as shown in Fig. 6.5, by the co-transduction of Rac1, the motility of the cells was even elevated which indicated that the over-expressed Rac1 can be activated by HBx and in turn caused stronger cell migration. This is in tune with the result discussed above. In our previous CREB binding assay [150], Cdc 42 did not shown to be activated by HBx. However, interestingly, in our recent cell migration assay, when co-transfected with HBx, the over-expressed Cdc42 was also shown to induce the cell migration. But the migration caused by Cdc42 was not as strong as that caused by Rac1 (Fig. 6.6). In total, our results suggested that the cell migration can be caused by HBV replication through with the activation of Rac1 and Cdc42 GTPase.

Figure 6.5: Microscopy based real time cell migration assay of HBx activated Rac1.

Cell migration distance (A) and average velocity (B) assay of HepG2 Cells transfected with various plasmids as shown in the figure. The cells were observed for 16hours (960min) 24 hours post-transfection. Each error bar represented the standard error of the data calculated from at least 10 cells. A) The real-time cell migration distance at every single time spot. B) Average cell migration velocity of each sample in 16hours.

Figure 6.6: Microscopy based real time cell migration assay of HBx activated Cdc42.

Cell migration distance (A) and average velocity (B) assay of HepG2 Cells transfected with various plasmids as shown in the figure. The cells were observed for 16hours (960min) 24 hours post-transfection. Each error bar represented the standard error of the data calculated from at least 10 cells. A) The real-time cell migration distance at every single time spot. B) Average cell migration velocity of each sample in 16hours.

6.2 RhoA activation by genotype specific HBx

Endogenous Rho GTPases are inactive when at their GDP-bound state. At their GTP-bound state, they trigger transduction of signal to various downstream targets.

RhoA is one of the important members of Rho GTPases family. RhoA can activate mDia or ROCK [241]. Through ROCK and mDia, the effectors of RhoA, RhoA mediates variety of phenotypes of transformation by Src and Ras. For RhoA, Rhotekin is a direct effector which contains Rho binding domain specifically recognizing the GTP-bound (activated form) RhoA [242]. The ability of HBx expression in activating RhoA was investigated by a Rho binding specific pull-down assay. HBx genotypes A/B/C/D were used for the assay. The constructs of RhoA wildtype (RhoA wt), constitutive active mutant RhoA V14, dominant negative mutant RhoA N19 and pGEX-RBD (Rho binding domain) were from Dr. Cheng Gee Koh (SBS, NTU). RhoA constructs were put into mammalian vector pXJ40, and RBD from Rhotekin was cloned into bacterial vector pGEX.

The Rho binding assay was performed as follows. GST-RBD with HA tag was expressed and applied for the binding assay. The ability of GST-RBD in recognizing RhoA V14 was demonstrated by the ability to retain the proteins. Genotype specific HBx was co-transfected with RhoA wt, and the RhoA activation ability of HBx was detected by the retaining of GST-RBD. If RhoA wt was activated by HBx, it would transformed into GTP-bound form, and be retained by GST-RBD pre-adsorbed on the GST beads in the column. As detected by Western blot with anti-HA, constitutive activated RhoA V14 bound to GST-RBD as expected and dominant negative mutant RhoA N19 did not show binding to GST-RBD. Genotype specific HBx was then co-transfected into HepG2 with wild type RhoA. Strong binding of RhoA can be detected in the cells transfected by

HBx genotype C (Lane 5, Fig 6.7), and moderate binding by genotypes B (Lane 4, Fig 6.7) and D (Lane 6, Fig 6.7). However, no binding was shown by HBx genotype A (Lane 3, Fig 6.7). Such result indicated that RhoA was activated by HBx genotypes B/C/D in different degree while no activation by HBx genotype A. This result suggested that RhoA was activated by HBx expression, genotype depended. This may caused by the sequences differences among the genotype specific HBx, although highly conserved, mainly located at the proline/serine rich regions as we have defined in previous work [122] and such differences may affect on the activation abilities of Rho GTPases. Proline rich region is famous for the interaction with SH3 binding domain which is contained in many cellular signaling proteins and such interaction was also believed to work in the virus and cell communication.

Microscopy based cell motility assay was also applied to investigate the RhoA activated cell migration in HepG2 cells transfected with HBx. As shown in Fig.6.8, negative control pXJ40 showed weak cell migration velocity and similar situation in the cells transfected by dominant negative RhoA (RhoA N19). The cells transfected with constitutive active mutant RhoA V14 showed strong cell motility and the cells with RhoA wildtype showed moderate cell migration which may be activated by cellular proteins of HepG2 cells. When being co-transfected with HBx and wildtype RhoA, the cells showed much stronger migration than those without HBx (Fig.6.8).

Figure 6.7: Activation status of RhoA by HBx genotypes A/B/C/D.

The binding assay was performed by using GST fusion protein of Rho binding domain of Rhotekin, which can specifically bind to the GTP binding form of RhoA. The up panel summarized binding assay of HBx transfected cells with various constructs related to RhoA, including constitutively activated RhoA (G14V) and dominant negative RhoA (T19N). The antibody used in the Western blot analysis was the mouse monoclonal anti-HA antibodies (Santa Cruz Biotech). Since no internal control can be applied, bottom panel was a SDS-PAGE gel directly stained with Coomassie Blue for the indication of loading consistency of the relevant binding domain, namely the RhoA-binding domain (RBD) of Rhotekin.

Figure 6.8: Microscopy based real time cell migration assay of HBx activated RhoA.

Cell migration distance (A) and average velocity (B) assay of HepG2 Cells transfected with various plasmids as shown in the figure. The cells were observed for 16hours (960min) 24 hours post-transfection. Each error bar represented the standard error of the data calculated from at least 10 cells. A) The real-time cell migration distance at every single time spot. B) Average cell migration velocity of each sample in 16hours.

6.3 Discussion and Conclusions

Rho GTPases comprise a family of proteins which are highly conserved in almost all organisms. Rho GTPases have been implicated in various cellular processes including cytoskeletal reorganization, cell cycle and division, motility,

angiogenesis, and phagocytosis. Previous studies have shown Rho GTPases involved in virus replication. For example, activation of Rac1 and Cdc42 has been associated with HIV replication [229].

Rho GTPases family is involved in cellular responses for the formation and regulation of cell cytoskeletal and morphological changes, such as filopodia and lamellipodia. It has been reported the HBx activates ERKs and PI-3k/akt pathways, and further activates matrix metalloproteinase-9 with a potential function of invasion and metastasis [126]. We have recently demonstrated the same activation of ERK1/2 and AKT by HBx using SH3 binding domain located in HBx. Our recent investigation has also indicated that Rac1 GTPase can be activated by HBV replication, which in turn sustains HBV replication [150]. Interestingly, the Rac1 activation and HBV replication has been found to be significantly decreased in HepG2 cells which were transfected with the replicative P-A mutant HBV genome [150].

Our cell motility analysis has provided an extension on our investigation of the Rac1 activation by HBV replication. By using transwell cell migration assay and microscopy based real time cell movements tracking, the movements of HepG2 cells under the affects of Rho GTPases were analyzed. Our result showed that HBV replication activated Rac1 increase the cell motility while it was abrogated by P-A mutant HBV. Furthermore, the co-transfected HBx and Rac1 also showed an increase in cell movement. Interestingly, the HBx activated Rac1 induced cell

movements were expanded to other two members of Rho GTPases, Cdc42, and RhoA. Although Cdc42 was not shown to be activated by HBV in CRIB binding assay [150], however, in the real time cell migration assay, Cdc42 was found to be activated by HBx and increase the cell migration in a moderate extent in our cell motility assay. This may suggest that the activation of Cdc42 by HBx may not be direct. As for RhoA, it was shown to be activated by HBx in a genotype depending way. The RhoA activation was further shown to lead to an enhancement of cell migration when co-transfected with HBx in the real time cell migration assay.

In summary, by applying the site mutation construction and transwell cell migration assay, our results have shown that the potential relationship between proline serine rich region in HBx (in the context of HBV replication) and cell motility, by the activation of Rho GTPases. Rac1 was activated by HBV replication and the activation increased the cell motility. The cell movement was decreased by the P-A mutant which was located in HBx proline-serine rich region. As a critical SH3 binding domain, proline-serine rich region has the ability to interact with a number of cellular proteins containing SH3 domains or WW domains. Such interactions could induce various cellular responses. As shown in our results, this typical region could be the reason for inducing high cell motility by activating Rho GTPases. As a remarkable characteristic of cancer cell, increased cell motility is an omen of cancer development. Our results should shed new lights in the understanding of mechanism of HCC development.

Chapter 7 : Conclusions

HBx, the smallest non-structural protein of HBV, has been implicated in the process of HCC development, with a possible link to individual genotypes. Nevertheless, the underlying mechanism remains obscure.

In this work, HBx genotypes A/B/C/D were cloned into expression vector for high level of expression in mammalian cells and the successful expressions were confirmed in human hepatocellular liver carcinoma cell line HepG2. Our previous identified proline/serine rich regions located in HBx were selected and used as probe in this study to identify cellular protein which interact with HBx through SH3 binding. P to A mutant probes were also used for binding array assays to show the binding specificity of this region. SH3 domain containing cellular proteins, including cytoskeletal proteins, were identified as HBx-interacting proteins from the commercial SH3 array binding assay. Interestingly, the interaction appeared to be genotype specific. The differences between the genotypes may imply the differential cellular signal targets of genotype specific HBx via proline/serine rich region. The interaction between HBx and cellular protein was made more functionally relevant by the increased HBV replication in cells over-expressing vixen β , one of the identified HBx interacting proteins.

The impact of the interaction between HBx and cellular protein on overall cellular

behavior was then analyzed by 2D coupled LC-MS/MS proteomics analysis. This would provide us with a comprehensive HBx-induced protein profile in HepG2 cells. Our results indicated that proteins were differentially expressed in HepG2 cells transfected by HBx of various genotypes. Proteins associated with cytoskeleton were found to be either up-regulated (MACF1, HMGB1, Annexin A2) or down-regulated (lamin A/C). These may in turn result in the decrease of focal adhesion and increase of cell migration in response to HBx. Levels of other cellular proteins with reported impacts on the function of extracellular matrix (ECM) proteins and cell migration, including Ca^{2+} -binding proteins (S100A11, S100A6, and S100A4) and proteasome protein (PSMA3), were affected by HBx. The differential protein profile identified in this study was also supported by our functional assay which indicated that cell migration was enhanced by HBx. By the application of cell movement detection on fluorescent beads coated surface and microscopy based real time cell tracking, the movements of the cells were recorded and the cell movement velocity was obtained. Taken together, the newly identified proteins should provide useful information on cellular signaling pathways due to early cellular response to HBV genotype-related infection. The comparison of the profiles should give us a platform to analyze the HBx effects to liver cells and the possible pathways to different prognosis caused by different genotypes. Further downstream functional assays, including our reported cell migration assay, should provide new insights on the association between HCC and HBx.

In addition to the intracellular protein profile, secreted protein profile as a result of

HBV replication also analyzed. These proteins may be candidate prognostic biomarkers for HCC, as HBV replication can be considered as a early stage of HCC development. Some proteins including AHSR and apolipoprotein A-I which have been revealed to be differentially expressed in HCC and liver cirrhosis patients' serum were also identified in our study. This is the first report linking these proteins to HBV replication, suggesting their role as potential biomarkers for HCC development.

Rho GTPases actively regulate cellular processes such as actin remodeling, cell adhesion and migration. With the coordination of Rho GTPases and their multiply cytoskeleton related target proteins, the cell adhesion and motility are well controlled. Deregulation of the Rho GTPases may lead to uncontrolled dynamic cytoskeletal reorganization and metastasis, one common feature of cancer developments. Findings from our study strongly implicated the Rho GTPases family of proteins in HBV/HBx induced cell motility. In addition to our recent demonstration on HBV-induced Rac1 activation, we showed that HBV can also enhance cell migration through the PXXP motif in HBx. Significantly, RhoA, another member of Rho GTPases was activated by HBx in the genotype specific manner. In particular, HBx genotype B, C and D activated RhoA, but not genotype A. The activation was supported by the microscopy based real-time migration assay.

Our findings on the influence of HBx on cellular behavior through its interacting

proteins can be summarized in the new diagram (Fig 7.1). For example, the clinically associated role of HBV/HBx in cell metastasis was supported by the activation of Rho GTPases, and changes in level of cytoskeletal protein including lamin A/C, MACF1 and HMGB1. The identification of extracellular proteins, previously linked to HCC, from culture medium of HBV-replicating cells not only supported the role of HBV in HCC development, but also suggested their role as prognostic biomarkers for HCC development. In addition, protein profile analyzed in our study revealed other cellular proteins which have been associated with HBV, such as Annexin 2 and hnRNP K.

Figure 7.1: Hypothetical HBV and HBx regulatory intracellular and secreted proteins.

Arrows indicated the proteins proposed to be affected by HBx directly or indirectly in our results. Cytoskeletal proteins were up or down-regulated by HBx, which potentially result in the increase of cell motility. Rho GTPases, major regulators of cytoskeleton reorganization, were also estimated activated by HBx and resulted in enhancing cell motility. The decreased expressions of calcium binding protein S100 family may benefit the increase of calcium concentration of cells caused by HBx, and result in the activation

of cellular signal pathways and amplification of HBV [52, 63, 243]. The upregulation of proteasome subunit suggested that the ubiquitin-mediated protein degradation pathway was involved in HBV infection, which was in line with the previous research [198, 206]. Secreted protein profiles in HBV replicating cells showed the higher level of proteins such as AHSG, which may be potential biomarker for the prognosis of liver cancer.

Taken together, our work established a useful sequence-to-function platform to analyze HBV-host interactions. In particular, by using iTRAQ coupled 2D LC-MS/MS based quantitative proteomics, the changes in protein profile influenced by HBx expression, allowed the identification of cellular proteins which may well be the downstream effectors of HBV-host interactions via SH3 binding. A series of differentially expressed cellular cytoskeletal proteins identified in this work were highly associated with cell migration, and thus providing new mechanistic insights on our cell observation of changes in cellular adhesion/de-adhesion and migration under the influence of HBV replication/HBx expression. It is to be hoped that our platform can be expanded into further functional analyses for the comprehensive understanding of the underlying mechanism of the involvement of HBx in HCC, as well as on HBV-associated HCC.

Chapter 8 : Future Work

In this work, a proline/serine rich region identified in HBx from our lab was further analyzed functionally. SH3 binding arrays were used for the identification of potential HBx interacting proteins via SH3 binding region (proline/serine region). SH3-containing cellular proteins, interacting with HBx in a genotype-specific manner, were identified in the SH3 screening assay. The significance of the detected interactions would be enhanced by further characterization, including GST pull down assay. In addition to SH3 domain, WW domain is also implicated in many protein interactions. WW domain is even considered to be a competitor of SH3 domain during the complex intracellular protein interactions. The possible WW domain in HBx-cell interaction is based on the following. Firstly, genotype-related differences in the proline/serine rich region in HBx were observed in multi-alignment of HBx (Fig. 3.7). Secondly, Pin1 which contains WW domain but not a SH3 domain interacts with HBx genotype A which contains more serine than proline residue in the proline/serine rich domain [98]. This may imply genotype specific HBx has similar but different binding targets. It would be of interest to functionally differentiate and characterize the correlation of HBx and WW domain containing proteins.

Among the cellular SH3-containing proteins, β PIX (a GEF of Rac1/Cdc42) has been shown to interact with HBx. The Rho GTPases are part of the Ras-related

superfamily of molecular switches that are critical regulators of signal pathways in most cells. They are involved in the regulations of the cytoskeleton, and do so by recruiting a variety of downstream effector proteins. In addition to Rac1 activation, Cdc42 and RhoA were also shown to increase cell motility in the context of HBV replication. And what is more interesting is the activation of RhoA by HBx was genotype specific as genotype A had no effect on it. These findings from the current work led us to ask the question: are there other interactions involved in HBV induced cytoskeletal reorganization besides SH3 binding? How to map out a comprehensive network of such interaction?

The SH3 induced activation signaling pathways were expanded by our proteomics analysis. 2D coupled LC-MS/MS is a powerful technique used for many applications which has very high sensitivity and specificity. A number of signaling proteins and cytoskeletal proteins were identified by the proteomics approach, providing us with a comprehensive protein profile as a result of HBV or HBx interaction with cells via SH3 binding. The significance of the overall involvement of cytoskeletal protein in HBV-HBx –cell interaction was further demonstrated in our cell motility assay. The direct involvement of individual cellular protein can be further analyzed by a combined approach using gene knock-out (siRNA) and cell motility assay in the context of HBV replication. Our proteomics approach also led to the identification of candidate biomarkers for HCC prognosis. These included AHSG and apolipoprotein A-I. Due to the low abundance of secreted proteins and inefficient quantitation, the identification of biomarkers by proteomics analysis can

be strengthened by better sample preparation and improved software for better analysis. These candidate biomarkers, once validated by in vitro analysis (Western blot) and in clinical samples, would hold the promise of providing HCC prognosis in chronic HBV carriers.

It is to be hoped that our sequence-to-function analysis platform can be used for other protein-protein interactions and in particular virus-host interactions. Findings from such characterization should not only provide new insights on better antiviral strategy, but also identify biomarkers for efficient prognosis of related human diseases.

References

- [1] Arbuthnot, P., Kew, M., Hepatitis B virus and hepatocellular carcinoma. *International journal of experimental pathology* 2001, 82, 77-100.
- [2] Lok, A. S., Chronic hepatitis B. *N Engl J Med* 2002, 346, 1682-1683.
- [3] Blumberg, B. S., Larouze, B., London, W. T., Werner, B., *et al.*, The relation of infection with the hepatitis B agent to primary hepatic carcinoma. *The American journal of pathology* 1975, 81, 669-682.
- [4] Andre, F., Hepatitis B epidemiology in Asia, the Middle East and Africa. *Vaccine* 2000, 18 Suppl 1, S20-22.
- [5] Lavanchy, D., Worldwide epidemiology of HBV infection, disease burden, and vaccine prevention. *J Clin Virol* 2005, 34 Suppl 1, S1-3.
- [6] Lavanchy, D., Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J Viral Hepat* 2004, 11, 97-107.
- [7] Nowak, M. A., Bonhoeffer, S., Hill, A. M., Boehme, R., *et al.*, Viral dynamics in hepatitis B virus infection. *Proceedings of the National Academy of Sciences of the United States of America* 1996, 93, 4398-4402.
- [8] Seeger, C., Mason, W. S., Hepatitis B virus biology. *Microbiol Mol Biol Rev* 2000, 64, 51-68.
- [9] Chen, D. S., Hsu, N. H., Sung, J. L., Hsu, T. C., *et al.*, A mass vaccination program in Taiwan against hepatitis B virus infection in infants of hepatitis B surface antigen-carrier mothers. *Jama* 1987, 257, 2597-2603.
- [10] Hsu, H. M., Chen, D. S., Chuang, C. H., Lu, J. C., *et al.*, Efficacy of a mass hepatitis B vaccination program in Taiwan. Studies on 3464 infants of hepatitis B surface antigen-carrier mothers. *Jama* 1988, 260, 2231-2235.
- [11] Chen, H. L., Chang, M. H., Ni, Y. H., Hsu, H. Y., *et al.*, Seroepidemiology of hepatitis B virus infection in children: Ten years of mass vaccination in Taiwan. *Jama* 1996, 276, 906-908.
- [12] Brunetto, M. R., Oliveri, F., Coco, B., Leandro, G., *et al.*, Outcome of anti-HBe positive chronic hepatitis B in alpha-interferon treated and untreated patients: a long term cohort study. *J Hepatol* 2002, 36, 263-270.
- [13] Hoofnagle, J. H., di Bisceglie, A. M., The treatment of chronic viral hepatitis. *N Engl J Med* 1997, 336, 347-356.
- [14] de Man, R. A., Wolters, L. M., Nevens, F., Chua, D., *et al.*, Safety and efficacy of oral entecavir given for 28 days in patients with chronic hepatitis B virus infection. *Hepatology* 2001, 34, 578-582.
- [15] Rizzetto, M., Ciancio, A., Chronic HBV-related liver disease. *Mol Aspects Med* 2008, 29, 72-84.
- [16] Chen, W. N., Oon, C. J., Human hepatitis B virus mutants: significance of molecular changes. *FEBS letters* 1999, 453, 237-242.
- [17] Ganem, D., Pollack, J. R., Tavis, J., Hepatitis B virus reverse transcriptase and its many roles in hepadnaviral genomic replication. *Infect Agents Dis* 1994, 3, 85-93.
- [18] Bartenschlager, R., Junker-Niepmann, M., Schaller, H., The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. *Journal of virology* 1990, 64,

5324-5332.

[19] Hirsch, R. C., Lavine, J. E., Chang, L. J., Varmus, H. E., Ganem, D., Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as well as for reverse transcription. *Nature* 1990, 344, 552-555.

[20] Junker-Niepmann, M., Bartenschlager, R., Schaller, H., A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. *Embo J* 1990, 9, 3389-3396.

[21] Yokosuka, O., Arai, M., Molecular biology of hepatitis B virus: effect of nucleotide substitutions on the clinical features of chronic hepatitis B. *Med Mol Morphol* 2006, 39, 113-120.

[22] Okamoto, H., Tsuda, F., Sakugawa, H., Sastrosoewignjo, R. I., *et al.*, Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. *The Journal of general virology* 1988, 69 (Pt 10), 2575-2583.

[23] Naumann, H., Schaefer, S., Yoshida, C. F., Gaspar, A. M., *et al.*, Identification of a new hepatitis B virus (HBV) genotype from Brazil that expresses HBV surface antigen subtype adw4. *The Journal of general virology* 1993, 74 (Pt 8), 1627-1632.

[24] Norder, H., Courouce, A. M., Magnius, L. O., Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* 1994, 198, 489-503.

[25] Stuyver, L., De Gendt, S., Van Geyt, C., Zoulim, F., *et al.*, A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. *The Journal of general virology* 2000, 81, 67-74.

[26] Arauz-Ruiz, P., Norder, H., Robertson, B. H., Magnius, L. O., Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *The Journal of general virology* 2002, 83, 2059-2073.

[27] Bartholomeusz, A., Schaefer, S., Hepatitis B virus genotypes: comparison of genotyping methods. *Reviews in medical virology* 2004, 14, 3-16.

[28] Lok, A. S., Akarca, U., Greene, S., Mutations in the pre-core region of hepatitis B virus serve to enhance the stability of the secondary structure of the pre-genome encapsidation signal. *Proceedings of the National Academy of Sciences of the United States of America* 1994, 91, 4077-4081.

[29] Bowyer, S. M., van Staden, L., Kew, M. C., Sim, J. G., A unique segment of the hepatitis B virus group A genotype identified in isolates from South Africa. *The Journal of general virology* 1997, 78 (Pt 7), 1719-1729.

[30] Kidd-Ljunggren, K., Oberg, M., Kidd, A. H., The hepatitis B virus X gene: analysis of functional domain variation and gene phylogeny using multiple sequences. *The Journal of general virology* 1995, 76 (Pt 9), 2119-2130.

[31] Theamboonlers, A., Jantaradsamee, P., Kaew-In, N., Tangkijvanich, P., *et al.*, The predominant genotypes of hepatitis B virus in Thailand. *Annals of tropical medicine and parasitology* 1999, 93, 737-743.

[32] Borchani-Chabchoub, I., Gargouri, A., Mokdad-Gargouri, R., Genotyping of Tunisian hepatitis B virus isolates based on the sequencing of preS2 and S regions. *Microbes and infection / Institut Pasteur* 2000, 2, 607-612.

[33] Kidd-Ljunggren, K., Broman, E., Ekvall, H., Gustavsson, O., Nosocomial transmission of hepatitis B virus infection through multiple-dose vials. *The Journal of hospital infection* 1999, 43, 57-62.

[34] Norder, H., Hammas, B., Lee, S. D., Bile, K., *et al.*, Genetic relatedness of hepatitis B viral

strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. *The Journal of general virology* 1993, 74 (Pt 7), 1341-1348.

[35] Arauz-Ruiz, P., Norder, H., Visona, K. A., Magnius, L. O., Genotype F prevails in HBV infected patients of hispanic origin in Central America and may carry the precore stop mutant. *Journal of medical virology* 1997, 51, 305-312.

[36] Blitz, L., Pujol, F. H., Swenson, P. D., Porto, L., *et al.*, Antigenic diversity of hepatitis B virus strains of genotype F in Amerindians and other population groups from Venezuela. *Journal of clinical microbiology* 1998, 36, 648-651.

[37] Kato, H., Orito, E., Sugauchi, F., Ueda, R., *et al.*, Determination of hepatitis B virus genotype G by polymerase chain reaction with hemi-nested primers. *Journal of virological methods* 2001, 98, 153-159.

[38] Kato, H., Orito, E., Gish, R. G., Sugauchi, F., *et al.*, Characteristics of hepatitis B virus isolates of genotype G and their phylogenetic differences from the other six genotypes (A through F). *Journal of virology* 2002, 76, 6131-6137.

[39] Schaefer, S., Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World J Gastroenterol* 2007, 13, 14-21.

[40] Aufiero, B., Schneider, R. J., The hepatitis B virus X-gene product trans-activates both RNA polymerase II and III promoters. *Embo J* 1990, 9, 497-504.

[41] Dandri, M., Petersen, J., Stockert, R. J., Harris, T. M., Rogler, C. E., Metabolic labeling of woodchuck hepatitis B virus X protein in naturally infected hepatocytes reveals a bimodal half-life and association with the nuclear framework. *Journal of virology* 1998, 72, 9359-9364.

[42] Dandri, M., Schirmacher, P., Rogler, C. E., Woodchuck hepatitis virus X protein is present in chronically infected woodchuck liver and woodchuck hepatocellular carcinomas which are permissive for viral replication. *Journal of virology* 1996, 70, 5246-5254.

[43] Doria, M., Klein, N., Lucito, R., Schneider, R. J., The hepatitis B virus HBx protein is a dual specificity cytoplasmic activator of Ras and nuclear activator of transcription factors. *Embo J* 1995, 14, 4747-4757.

[44] Henkler, F., Hoare, J., Waseem, N., Goldin, R. D., *et al.*, Intracellular localization of the hepatitis B virus HBx protein. *The Journal of general virology* 2001, 82, 871-882.

[45] Hoare, J., Henkler, F., Dowling, J. J., Errington, W., *et al.*, Subcellular localisation of the X protein in HBV infected hepatocytes. *Journal of medical virology* 2001, 64, 419-426.

[46] Sirma, H., Weil, R., Rosmorduc, O., Urban, S., *et al.*, Cytosol is the prime compartment of hepatitis B virus X protein where it colocalizes with the proteasome. *Oncogene* 1998, 16, 2051-2063.

[47] Weil, R., Sirma, H., Giannini, C., Kremsdorf, D., *et al.*, Direct association and nuclear import of the hepatitis B virus X protein with the NF-kappaB inhibitor IkappaBalpha. *Molecular and cellular biology* 1999, 19, 6345-6354.

[48] Siddiqui, A., Jameel, S., Mapoles, J., Expression of the hepatitis B virus X gene in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 1987, 84, 2513-2517.

[49] Hohne, M., Schaefer, S., Seifer, M., Feitelson, M. A., *et al.*, Malignant transformation of immortalized transgenic hepatocytes after transfection with hepatitis B virus DNA. *Embo J* 1990, 9, 1137-1145.

[50] Kim, K. H., Seong, B. L., Pro-apoptotic function of HBV X protein is mediated by interaction with c-FLIP and enhancement of death-inducing signal. *Embo J* 2003, 22, 2104-2116.

- [51] Kim, S., Kim, H. Y., Lee, S., Kim, S. W., *et al.*, Hepatitis B virus x protein induces perinuclear mitochondrial clustering in microtubule- and Dynein-dependent manners. *Journal of virology* 2007, 81, 1714-1726.
- [52] McClain, S. L., Clippinger, A. J., Lizzano, R., Bouchard, M. J., Hepatitis B virus replication is associated with an HBx-dependent mitochondrion-regulated increase in cytosolic calcium levels. *Journal of virology* 2007, 81, 12061-12065.
- [53] Clippinger, A. J., Bouchard, M. J., Hepatitis B virus HBx protein localizes to mitochondria in primary rat hepatocytes and modulates mitochondrial membrane potential. *Journal of virology* 2008, 82, 6798-6811.
- [54] Rahmani, Z., Huh, K. W., Lasher, R., Siddiqui, A., Hepatitis B virus X protein colocalizes to mitochondria with a human voltage-dependent anion channel, HVDAC3, and alters its transmembrane potential. *Journal of virology* 2000, 74, 2840-2846.
- [55] Takada, S., Shirakata, Y., Kaneniwa, N., Koike, K., Association of hepatitis B virus X protein with mitochondria causes mitochondrial aggregation at the nuclear periphery, leading to cell death. *Oncogene* 1999, 18, 6965-6973.
- [56] Waris, G., Huh, K. W., Siddiqui, A., Mitochondrially associated hepatitis B virus X protein constitutively activates transcription factors STAT-3 and NF-kappa B via oxidative stress. *Molecular and cellular biology* 2001, 21, 7721-7730.
- [57] Blum, H. E., Zhang, Z. S., Galun, E., von Weizsacker, F., *et al.*, Hepatitis B virus X protein is not central to the viral life cycle in vitro. *Journal of virology* 1992, 66, 1223-1227.
- [58] Xu, Z., Yen, T. S., Wu, L., Madden, C. R., *et al.*, Enhancement of hepatitis B virus replication by its X protein in transgenic mice. *Journal of virology* 2002, 76, 2579-2584.
- [59] Reifenberg, K., Nusser, P., Lohler, J., Spindler, G., *et al.*, Virus replication and virion export in X-deficient hepatitis B virus transgenic mice. *The Journal of general virology* 2002, 83, 991-996.
- [60] Chen, H. S., Kaneko, S., Girones, R., Anderson, R. W., *et al.*, The woodchuck hepatitis virus X gene is important for establishment of virus infection in woodchucks. *Journal of virology* 1993, 67, 1218-1226.
- [61] Zoulim, F., Saputelli, J., Seeger, C., Woodchuck hepatitis virus X protein is required for viral infection in vivo. *Journal of virology* 1994, 68, 2026-2030.
- [62] Melegari, M., Scaglioni, P. P., Wands, J. R., Cloning and characterization of a novel hepatitis B virus x binding protein that inhibits viral replication. *Journal of virology* 1998, 72, 1737-1743.
- [63] Bouchard, M. J., Wang, L. H., Schneider, R. J., Calcium signaling by HBx protein in hepatitis B virus DNA replication. *Science (New York, N.Y)* 2001, 294, 2376-2378.
- [64] Leupin, O., Bontron, S., Schaeffer, C., Strubin, M., Hepatitis B virus X protein stimulates viral genome replication via a DDB1-dependent pathway distinct from that leading to cell death. *Journal of virology* 2005, 79, 4238-4245.
- [65] Klein, N. P., Bouchard, M. J., Wang, L. H., Kobarg, C., Schneider, R. J., Src kinases involved in hepatitis B virus replication. *Embo J* 1999, 18, 5019-5027.
- [66] Tang, H., Delgermaa, L., Huang, F., Oishi, N., *et al.*, The transcriptional transactivation function of HBx protein is important for its augmentation role in hepatitis B virus replication. *Journal of virology* 2005, 79, 5548-5556.
- [67] Cha, M. Y., Ryu, D. K., Jung, H. S., Chang, H. E., Ryu, W. S., 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, 978-986.
- [68] Yen, T. S., Hepadnaviral X Protein: Review of Recent Progress. *J Biomed Sci* 1996, 3, 20-30.

- [69] Tarn, C., Lee, S., Hu, Y., Ashendel, C., Andrisani, O. M., Hepatitis B virus X protein differentially activates RAS-RAF-MAPK and JNK pathways in X-transforming versus non-transforming AML12 hepatocytes. *J Biol Chem* 2001, 276, 34671-34680.
- [70] Bouchard, M. J., Schneider, R. J., The enigmatic X gene of hepatitis B virus. *Journal of virology* 2004, 78, 12725-12734.
- [71] Benn, J., Su, F., Doria, M., Schneider, R. J., Hepatitis B virus HBx protein induces transcription factor AP-1 by activation of extracellular signal-regulated and c-Jun N-terminal mitogen-activated protein kinases. *Journal of virology* 1996, 70, 4978-4985.
- [72] Tarn, C., Bilodeau, M. L., Hullinger, R. L., Andrisani, O. M., Differential immediate early gene expression in conditional hepatitis B virus pX-transforming versus nontransforming hepatocyte cell lines. *J Biol Chem* 1999, 274, 2327-2336.
- [73] Tarn, C., Zou, L., Hullinger, R. L., Andrisani, O. M., Hepatitis B virus X protein activates the p38 mitogen-activated protein kinase pathway in dedifferentiated hepatocytes. *Journal of virology* 2002, 76, 9763-9772.
- [74] Chirillo, P., Falco, M., Puri, P. L., Artini, M., *et al.*, Hepatitis B virus pX activates NF-kappa B-dependent transcription through a Raf-independent pathway. *Journal of virology* 1996, 70, 641-646.
- [75] Menzo, S., Clementi, M., Alfani, E., Bagnarelli, P., *et al.*, Trans-activation of epidermal growth factor receptor gene by the hepatitis B virus X-gene product. *Virology* 1993, 196, 878-882.
- [76] Benn, J., Schneider, R. J., Hepatitis B virus HBx protein deregulates cell cycle checkpoint controls. *Proceedings of the National Academy of Sciences of the United States of America* 1995, 92, 11215-11219.
- [77] Wang, C., Wang, W., Lu, H., [Immunohistochemical study of hepatitis C virus core antigen and HBxAg in liver cirrhosis and hepatocellular carcinoma tissues]. *Zhonghua Zhong Liu Za Zhi* 1997, 19, 85-88.
- [78] Henkler, F., Lopes, A. R., Jones, M., Koshy, R., Erk-independent partial activation of AP-1 sites by the hepatitis B virus HBx protein. *The Journal of general virology* 1998, 79 (Pt 11), 2737-2742.
- [79] Lara-Pezzi, E., Armesilla, A. L., Majano, P. L., Redondo, J. M., Lopez-Cabrera, M., The hepatitis B virus X protein activates nuclear factor of activated T cells (NF-AT) by a cyclosporin A-sensitive pathway. *Embo J* 1998, 17, 7066-7077.
- [80] Lara-Pezzi, E., Roche, S., Andrisani, O. M., Sanchez-Madrid, F., Lopez-Cabrera, M., The hepatitis B virus HBx protein induces adherens junction disruption in a src-dependent manner. *Oncogene* 2001, 20, 3323-3331.
- [81] Thanos, D., Maniatis, T., NF-kappa B: a lesson in family values. *Cell* 1995, 80, 529-532.
- [82] Seto, E., Mitchell, P. J., Yen, T. S., Transactivation by the hepatitis B virus X protein depends on AP-2 and other transcription factors. *Nature* 1990, 344, 72-74.
- [83] Lucito, R., Schneider, R. J., Hepatitis B virus X protein activates transcription factor NF-kappa B without a requirement for protein kinase C. *Journal of virology* 1992, 66, 983-991.
- [84] Su, F., Schneider, R. J., Hepatitis B virus HBx protein activates transcription factor NF-kappaB by acting on multiple cytoplasmic inhibitors of rel-related proteins. *Journal of virology* 1996, 70, 4558-4566.
- [85] Elmore, L. W., Hancock, A. R., Chang, S. F., Wang, X. W., *et al.*, Hepatitis B virus X protein and p53 tumor suppressor interactions in the modulation of apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* 1997, 94, 14707-14712.

- [86] Koike, K., Moriya, K., Iino, S., Yotsuyanagi, H., *et al.*, High-level expression of hepatitis B virus HBx gene and hepatocarcinogenesis in transgenic mice. *Hepatology* 1994, 19, 810-819.
- [87] Koike, K., Moriya, K., Yotsuyanagi, H., Shintani, Y., *et al.*, Compensatory apoptosis in preneoplastic liver of a transgenic mouse model for viral hepatocarcinogenesis. *Cancer Lett* 1998, 134, 181-186.
- [88] Yu, D. Y., Moon, H. B., Son, J. K., Jeong, S., *et al.*, Incidence of hepatocellular carcinoma in transgenic mice expressing the hepatitis B virus X-protein. *J Hepatol* 1999, 31, 123-132.
- [89] Slagle, B. L., Lee, T. H., Medina, D., Finegold, M. J., Butel, J. S., Increased sensitivity to the hepatocarcinogen diethylnitrosamine in transgenic mice carrying the hepatitis B virus X gene. *Mol Carcinog* 1996, 15, 261-269.
- [90] Madden, C. R., Finegold, M. J., Slagle, B. L., Hepatitis B virus X protein acts as a tumor promoter in development of diethylnitrosamine-induced preneoplastic lesions. *Journal of virology* 2001, 75, 3851-3858.
- [91] Pawson, T., Protein modules and signalling networks. *Nature* 1995, 373, 573-580.
- [92] Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., *et al.*, The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* 1992, 70, 431-442.
- [93] Bar-Sagi, D., Rotin, D., Batzer, A., Mandiyan, V., Schlessinger, J., SH3 domains direct cellular localization of signaling molecules. *Cell* 1993, 74, 83-91.
- [94] Pleiman, C. M., Hertz, W. M., Cambier, J. C., Activation of phosphatidylinositol-3' kinase by Src-family kinase SH3 binding to the p85 subunit. *Science (New York, N.Y)* 1994, 263, 1609-1612.
- [95] Klein, N. P., Schneider, R. J., Activation of Src family kinases by hepatitis B virus HBx protein and coupled signaling to Ras. *Molecular and cellular biology* 1997, 17, 6427-6436.
- [96] Rickles, R. J., Botfield, M. C., Zhou, X. M., Henry, P. A., *et al.*, Phage display selection of ligand residues important for Src homology 3 domain binding specificity. *Proceedings of the National Academy of Sciences of the United States of America* 1995, 92, 10909-10913.
- [97] Weng, Z., Rickles, R. J., Feng, S., Richard, S., *et al.*, Structure-function analysis of SH3 domains: SH3 binding specificity altered by single amino acid substitutions. *Molecular and cellular biology* 1995, 15, 5627-5634.
- [98] Datta, S., Banerjee, A., Chandra, P. K., Chakravarty, R., Pin1-HBx interaction: a step toward understanding the significance of hepatitis B virus genotypes in hepatocarcinogenesis. *Gastroenterology* 2007, 133, 727-728; author reply 728-729.
- [99] Kay, B. K., Williamson, M. P., Sudol, M., The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *Faseb J* 2000, 14, 231-241.
- [100] Cheadle, C., Ivashchenko, Y., South, V., Searfoss, G. H., *et al.*, Identification of a Src SH3 domain binding motif by screening a random phage display library. *J Biol Chem* 1994, 269, 24034-24039.
- [101] Rickles, R. J., Botfield, M. C., Weng, Z., Taylor, J. A., *et al.*, Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. *Embo J* 1994, 13, 5598-5604.
- [102] Sparks, A. B., Quilliam, L. A., Thorn, J. M., Der, C. J., Kay, B. K., Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries. *J Biol Chem* 1994, 269, 23853-23856.
- [103] Feng, S., Chen, J. K., Yu, H., Simon, J. A., Schreiber, S. L., Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science (New York, N.Y)* 1994, 266, 1241-1247.

- [104] Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., *et al.*, Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. *Proceedings of the National Academy of Sciences of the United States of America* 1996, 93, 1540-1544.
- [105] Knudsen, B. S., Zheng, J., Feller, S. M., Mayer, J. P., *et al.*, Affinity and specificity requirements for the first Src homology 3 domain of the Crk proteins. *Embo J* 1995, 14, 2191-2198.
- [106] Grabs, D., Slepnev, V. I., Songyang, Z., David, C., *et al.*, The SH3 domain of amphiphysin binds the proline-rich domain of dynamin at a single site that defines a new SH3 binding consensus sequence. *J Biol Chem* 1997, 272, 13419-13425.
- [107] Quilliam, L. A., Lambert, Q. T., Mickelson-Young, L. A., Westwick, J. K., *et al.*, Isolation of a NCK-associated kinase, PRK2, an SH3-binding protein and potential effector of Rho protein signaling. *J Biol Chem* 1996, 271, 28772-28776.
- [108] Kurakin, A., Hoffman, N. G., Kay, B. K., Molecular recognition properties of the C-terminal Sh3 domain of the Cbl associated protein, Cap. *J Pept Res* 1998, 52, 331-337.
- [109] Chen, H. I., Sudol, M., The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proceedings of the National Academy of Sciences of the United States of America* 1995, 92, 7819-7823.
- [110] Schild, L., Lu, Y., Gautschi, I., Schneeberger, E., *et al.*, Identification of a PY motif in the epithelial Na channel subunits as a target sequence for mutations causing channel activation found in Liddle syndrome. *Embo J* 1996, 15, 2381-2387.
- [111] Rentschler, S., Linn, H., Deininger, K., Bedford, M. T., *et al.*, The WW domain of dystrophin requires EF-hands region to interact with beta-dystroglycan. *Biol Chem* 1999, 380, 431-442.
- [112] Bedford, M. T., Chan, D. C., Leder, P., FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *Embo J* 1997, 16, 2376-2383.
- [113] Ernekova, K. S., Zambrano, N., Linn, H., Minopoli, G., *et al.*, The WW domain of neural protein FE65 interacts with proline-rich motifs in Mena, the mammalian homolog of Drosophila enabled. *J Biol Chem* 1997, 272, 32869-32877.
- [114] Bedford, M. T., Reed, R., Leder, P., WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. *Proceedings of the National Academy of Sciences of the United States of America* 1998, 95, 10602-10607.
- [115] Ranganathan, R., Lu, K. P., Hunter, T., Noel, J. P., Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* 1997, 89, 875-886.
- [116] Lu, P. J., Zhou, X. Z., Shen, M., Lu, K. P., Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science (New York, N.Y)* 1999, 283, 1325-1328.
- [117] Schutkowski, M., Bernhardt, A., Zhou, X. Z., Shen, M., *et al.*, Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and Pin1 substrate recognition. *Biochemistry* 1998, 37, 5566-5575.
- [118] Niebuhr, K., Ebel, F., Frank, R., Reinhard, M., *et al.*, A novel proline-rich motif present in ActA of *Listeria monocytogenes* and cytoskeletal proteins is the ligand for the EVH1 domain, a protein module present in the Ena/VASP family. *Embo J* 1997, 16, 5433-5444.
- [119] Tu, J. C., Xiao, B., Yuan, J. P., Lanahan, A. A., *et al.*, Homer binds a novel proline-rich motif and links group I metabotropic glutamate receptors with IP3 receptors. *Neuron* 1998, 21, 717-726.
- [120] Nishizawa, K., Freund, C., Li, J., Wagner, G., Reinherz, E. L., Identification of a

proline-binding motif regulating CD2-triggered T lymphocyte activation. *Proceedings of the National Academy of Sciences of the United States of America* 1998, 95, 14897-14902.

[121] Freund, C., Dotsch, V., Nishizawa, K., Reinherz, E. L., Wagner, G., The GYF domain is a novel structural fold that is involved in lymphoid signaling through proline-rich sequences. *Nat Struct Biol* 1999, 6, 656-660.

[122] Tan, T. L., Feng, Z., Lu, Y. W., Chan, V., Chen, W. N., Adhesion contact kinetics of HepG2 cells during Hepatitis B virus replication: Involvement of SH3-binding motif in HBx. *Biochim Biophys Acta* 2006, 1762, 755-766.

[123] Nagao, T., Inoue, S., Goto, S., Mizuta, T., *et al.*, Hepatic resection for hepatocellular carcinoma. Clinical features and long-term prognosis. *Ann Surg* 1987, 205, 33-40.

[124] Nagao, T., Inoue, S., Yoshimi, F., Sodeyama, M., *et al.*, Postoperative recurrence of hepatocellular carcinoma. *Ann Surg* 1990, 211, 28-33.

[125] Woessner, J. F., Jr., Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *Faseb J* 1991, 5, 2145-2154.

[126] Chung, T. W., Lee, Y. C., Kim, C. H., Hepatitis B viral HBx induces matrix metalloproteinase-9 gene expression through activation of ERK and PI-3K/AKT pathways: involvement of invasive potential. *Faseb J* 2004, 18, 1123-1125.

[127] Yu, F. L., Liu, H. J., Lee, J. W., Liao, M. H., Shih, W. L., Hepatitis B virus X protein promotes cell migration by inducing matrix metalloproteinase-3. *J Hepatol* 2005, 42, 520-527.

[128] Ou, D. P., Tao, Y. M., Tang, F. Q., Yang, L. Y., The hepatitis B virus X protein promotes hepatocellular carcinoma metastasis by upregulation of matrix metalloproteinases. *Int J Cancer* 2007, 120, 1208-1214.

[129] Lara-Pezzi, E., Serrador, J. M., Montoya, M. C., Zamora, D., *et al.*, The hepatitis B virus X protein (HBx) induces a migratory phenotype in a CD44-dependent manner: possible role of HBx in invasion and metastasis. *Hepatology* 2001, 33, 1270-1281.

[130] Grise, F., Bidaud, A., Moreau, V., Rho GTPases in hepatocellular carcinoma. *Biochim Biophys Acta* 2009, 1795, 137-151.

[131] Wong, C. C., Wong, C. M., Tung, E. K., Man, K., Ng, I. O., Rho-kinase 2 is frequently overexpressed in hepatocellular carcinoma and involved in tumor invasion. *Hepatology* 2009, 49, 1583-1594.

[132] Bustelo, X. R., Sauzeau, V., Berenjeno, I. M., GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. *Bioessays* 2007, 29, 356-370.

[133] Etienne-Manneville, S., Hall, A., Rho GTPases in cell biology. *Nature* 2002, 420, 629-635.

[134] Kaibuchi, K., Kuroda, S., Amano, M., Regulation of the cytoskeleton and cell adhesion by the Rho family GTPases in mammalian cells. *Annual review of biochemistry* 1999, 68, 459-486.

[135] Giancotti, F. G., Ruoslahti, E., Integrin signaling. *Science (New York, N.Y)* 1999, 285, 1028-1032.

[136] Hirschman, S. Z., Price, P., Garfinkel, E., Christman, J., Acs, G., Expression of cloned hepatitis B virus DNA in human cell cultures. *Proceedings of the National Academy of Sciences of the United States of America* 1980, 77, 5507-5511.

[137] Manser, E., Huang, H. Y., Loo, T. H., Chen, X. Q., *et al.*, Expression of constitutively active alpha-PAK reveals effects of the kinase on actin and focal complexes. *Molecular and cellular biology* 1997, 17, 1129-1143.

[138] Pontow, S. E., Heyden, N. V., Wei, S., Ratner, L., Actin cytoskeletal reorganizations and coreceptor-mediated activation of rac during human immunodeficiency virus-induced cell fusion.

Journal of virology 2004, 78, 7138-7147.

[139] Sui, J., Tan, T. L., Zhang, J., Ching, C. B., Chen, W. N., 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, 1643-1651.

[140] Liang, R. C., Neo, J. C., Lo, S. L., Tan, G. S., *et al.*, Proteome database of hepatocellular carcinoma. *J Chromatogr B Analyt Technol Biomed Life Sci* 2002, 771, 303-328.

[141] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., *et al.*, Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004, 3, 1154-1169.

[142] Pierce, A., Unwin, R. D., Evans, C. A., Griffiths, S., *et al.*, Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Mol Cell Proteomics* 2008, 7, 853-863.

[143] Griffin, T. J., Xie, H., Bandhakavi, S., Popko, J., *et al.*, iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. *J Proteome Res* 2007, 6, 4200-4209.

[144] Glen, A., Gan, C. S., Hamdy, F. C., Eaton, C. L., *et al.*, iTRAQ-facilitated proteomic analysis of human prostate cancer cells identifies proteins associated with progression. *J Proteome Res* 2008, 7, 897-907.

[145] Tewfik, M. A., Latterich, M., DiFalco, M. R., Samaha, M., Proteomics of nasal mucus in chronic rhinosinusitis. *Am J Rhinol* 2007, 21, 680-685.

[146] Oberg, A. L., Mahoney, D. W., Eckel-Passow, J. E., Malone, C. J., *et al.*, Statistical analysis of relative labeled mass spectrometry data from complex samples using ANOVA. *J Proteome Res* 2008, 7, 225-233.

[147] Griffiths, S. D., Burthem, J., Unwin, R. D., Holyoake, T. L., *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, 81-89.

[148] Aebersold, R., Mann, M., Mass spectrometry-based proteomics. *Nature* 2003, 422, 198-207.

[149] Xiao, J. H., Davidson, I., Matthes, H., Garnier, J. M., Chambon, P., Cloning, expression, and transcriptional properties of the human enhancer factor TEF-1. *Cell* 1991, 65, 551-568.

[150] Tan, T. L., Fang, N., Neo, T. L., Singh, P., *et al.*, Rac1 GTPase is activated by hepatitis B virus replication--involvement of HBX. *Biochim Biophys Acta* 2008, 1783, 360-374.

[151] Shapiro, A. L., Vinuela, E., Maizel, J. V., Jr., Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem Biophys Res Commun* 1967, 28, 815-820.

[152] Zhang, J., Sui, J., Ching, C. B., Chen, W. N., 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, 1595-1607.

[153] Tan, T. L., Chen, W. N., 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, 293-298.

[154] Macdonald, A., Crowder, K., Street, A., McCormick, C., *et al.*, The hepatitis C virus non-structural NS5A protein inhibits activating protein-1 function by perturbing ras-ERK pathway signaling. *The Journal of biological chemistry* 2003, 278, 17775-17784.

[155] Saksela, K., Cheng, G., Baltimore, D., 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. *The EMBO journal* 1995, 14, 484-491.

[156] Khan, I. H., Sawai, E. T., Antonio, E., Weber, C. J., *et al.*, Role of the SH3-ligand domain of simian immunodeficiency virus Nef in interaction with Nef-associated kinase and simian AIDS in rhesus macaques. *Journal of virology* 1998, 72, 5820-5830.

[157] Zhao, Z. S., Manser, E., Lim, L., Interaction between PAK and nck: a template for Nck targets and role of PAK autophosphorylation. *Molecular and cellular biology* 2000, 20, 3906-3917.

[158] Mizokami, M., Nakano, T., Orito, E., Tanaka, Y., *et al.*, Hepatitis B virus genotype assignment using restriction fragment length polymorphism patterns. *FEBS letters* 1999, 450, 66-71.

[159] Jiang, J., Zhang, X., Wu, C., Qin, X., *et al.*, Increased plasma apoM levels in the patients suffered from hepatocellular carcinoma and other chronic liver diseases. *Lipids Health Dis* 2008, 7, 25.

[160] Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P., Charnay, P., Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. *Nature* 1979, 281, 646-650.

[161] Sells, M. A., Chen, M. L., Acs, G., Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA. *Proceedings of the National Academy of Sciences of the United States of America* 1987, 84, 1005-1009.

[162] Mahtab, M. A., Rahman, S., Khan, M., Karim, F., Hepatitis B virus genotypes: an overview. *Hepatobiliary Pancreat Dis Int* 2008, 7, 457-464.

[163] Chen, W. N., Oon, C. J., Lim, G. K., Frequent occurrence of hepatitis B virus surface antigen mutants in subtype adw in vaccinated Singapore infants. *Vaccine* 2001, 20, 639-640.

[164] Kao, J. H., Chen, P. J., Lai, M. Y., Chen, D. S., Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. *Gastroenterology* 2000, 118, 554-559.

[165] Kao, J. H., Hepatitis B viral genotypes: clinical relevance and molecular characteristics. *Journal of gastroenterology and hepatology* 2002, 17, 643-650.

[166] Niu, D. D., Zhang, J. H., Ren, Y. D., Feng, H. X., Chen, W. N., HBx genotype D represses GSTP1 expression and increases the oxidative level and apoptosis in HepG2 cells. *Molecular oncology* 2009, 3, 67-76.

[167] Stetler-Stevenson, W. G., Aznavoorian, S., Liotta, L. A., Tumor cell interactions with the extracellular matrix during invasion and metastasis. *Annual review of cell biology* 1993, 9, 541-573.

[168] Talwalkar, J. A., Gores, G. J., Diagnosis and staging of hepatocellular carcinoma. *Gastroenterology* 2004, 127, S126-132.

[169] Tian, B., Li, Y., Ji, X. N., Chen, J., *et al.*, Basement membrane proteins play an active role in the invasive process of human hepatocellular carcinoma cells with high metastasis potential. *Journal of cancer research and clinical oncology* 2005, 131, 80-86.

[170] Lara-Pezzi, E., Gomez-Gavero, M. V., Galvez, B. G., Mira, E., *et al.*, The hepatitis B virus X protein promotes tumor cell invasion by inducing membrane-type matrix metalloproteinase-1 and cyclooxygenase-2 expression. *J Clin Invest* 2002, 110, 1831-1838.

[171] Sui, J., Zhang, J., Ching, C. B., Chen, W. N., 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, 885-890.

[172]

[173] Wei, J., Xu, G., Wu, M., Zhang, Y., *et al.*, Overexpression of vimentin contributes to prostate cancer invasion and metastasis via src regulation. *Anticancer Res* 2008, 28, 327-334.

[174] Lammerding, J., Fong, L. G., Ji, J. Y., Reue, K., *et al.*, Lamin A and C but not lamin B1

- regulate nuclear mechanics. *J Biol Chem* 2006, 281, 25768-25780.
- [175] Lu, Q. Y., Yang, Y., Jin, Y. S., Zhang, Z. F., *et al.*, Effects of green tea extract on lung cancer A549 cells: proteomic identification of proteins associated with cell migration. *Proteomics* 2009, 9, 757-767.
- [176] Kodama, A., Karakesisoglou, I., Wong, E., Vaezi, A., Fuchs, E., ACF7: an essential integrator of microtubule dynamics. *Cell* 2003, 115, 343-354.
- [177] Wu, X., Kodama, A., Fuchs, E., ACF7 regulates cytoskeletal-focal adhesion dynamics and migration and has ATPase activity. *Cell* 2008, 135, 137-148.
- [178] Bustin, M., Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem Sci* 2001, 26, 152-153.
- [179] Sharma, M. R., Koltowski, L., Ownbey, R. T., Tuszynski, G. P., Sharma, M. C., 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, 146-156.
- [180] Degryse, B., Bonaldi, T., Scaffidi, P., Muller, S., *et al.*, The high mobility group (HMG) boxes of the nuclear protein HMG1 induce chemotaxis and cytoskeleton reorganization in rat smooth muscle cells. *J Cell Biol* 2001, 152, 1197-1206.
- [181] Taguchi, T., Kawai, S., Hashiguchi, T., Reassessment of the diagnostic value of selective lumbosacral radiculography. *J Neuroradiol* 2002, 29, 122-127.
- [182] Parkkinen, J., Rauvala, H., Interactions of plasminogen and tissue plasminogen activator (t-PA) with amphoterin. Enhancement of t-PA-catalyzed plasminogen activation by amphoterin. *J Biol Chem* 1991, 266, 16730-16735.
- [183] Emoto, K., Yamada, Y., Sawada, H., Fujimoto, H., *et al.*, Annexin II overexpression correlates with stromal tenascin-C overexpression: a prognostic marker in colorectal carcinoma. *Cancer* 2001, 92, 1419-1426.
- [184] Duncan, R., Carpenter, B., Main, L. C., Telfer, C., Murray, G. I., Characterisation and protein expression profiling of annexins in colorectal cancer. *Br J Cancer* 2008, 98, 426-433.
- [185] Zimmermann, U., Woenckhaus, C., Pietschmann, S., Junker, H., *et al.*, Expression of annexin II in conventional renal cell carcinoma is correlated with Fuhrman grade and clinical outcome. *Virchows Arch* 2004, 445, 368-374.
- [186] Esposito, I., Penzel, R., Chaib-Harrireche, M., Barcena, U., *et al.*, Tenascin C and annexin II expression in the process of pancreatic carcinogenesis. *J Pathol* 2006, 208, 673-685.
- [187] Mohammad, H. S., Kurokohchi, K., Yoneyama, H., Tokuda, M., *et al.*, Annexin A2 expression and phosphorylation are up-regulated in hepatocellular carcinoma. *Int J Oncol* 2008, 33, 1157-1163.
- [188] Narayan, R., Gangadharan, B., Hantz, O., Antrobus, R., *et al.*, Proteomic analysis of HepaRG cells: a novel cell line that supports hepatitis B virus infection. *J Proteome Res* 2009, 8, 118-122.
- [189] Diaz, V. M., Hurtado, M., Thomson, T. M., Reventos, J., Paciucci, R., Specific interaction of tissue-type plasminogen activator (t-PA) with annexin II on the membrane of pancreatic cancer cells activates plasminogen and promotes invasion in vitro. *Gut* 2004, 53, 993-1000.
- [190] Diaz, V. M., Planaguma, J., Thomson, T. M., Reventos, J., Paciucci, R., Tissue plasminogen activator is required for the growth, invasion, and angiogenesis of pancreatic tumor cells. *Gastroenterology* 2002, 122, 806-819.
- [191] Hajjar, K. A., Jacovina, A. T., Chacko, J., An endothelial cell receptor for plasminogen/tissue plasminogen activator. I. Identity with annexin II. *J Biol Chem* 1994, 269, 21191-21197.
- [192] Mignatti, P., Rifkin, D. B., Biology and biochemistry of proteinases in tumor invasion.

Physiol Rev 1993, 73, 161-195.

[193] Brownstein, C., Deora, A. B., Jacovina, A. T., Weintraub, R., *et al.*, Annexin II mediates plasminogen-dependent matrix invasion by human monocytes: enhanced expression by macrophages. *Blood* 2004, 103, 317-324.

[194] Liotta, L. A., Rao, C. N., Barsky, S. H., Tumor invasion and the extracellular matrix. *Lab Invest* 1983, 49, 636-649.

[195] DeClerck, Y. A., Imren, S., Montgomery, A. M., Mueller, B. M., *et al.*, Proteases and protease inhibitors in tumor progression. *Adv Exp Med Biol* 1997, 425, 89-97.

[196] Grigorian, M., Andresen, S., Tulchinsky, E., Kriajevska, M., *et al.*, Tumor suppressor p53 protein is a new target for the metastasis-associated Mts1/S100A4 protein: functional consequences of their interaction. *J Biol Chem* 2001, 276, 22699-22708.

[197] Cui, J. F., Liu, Y. K., Zhang, L. J., Shen, H. L., *et al.*, Identification of metastasis candidate proteins among HCC cell lines by comparative proteome and biological function analysis of S100A4 in metastasis in vitro. *Proteomics* 2006, 6, 5953-5961.

[198] Tong, A., Wu, L., Lin, Q., Lau, Q. C., *et al.*, Proteomic analysis of cellular protein alterations using a hepatitis B virus-producing cellular model. *Proteomics* 2008, 8, 2012-2023.

[199] Voutsadakis, I. A., The ubiquitin-proteasome system in colorectal cancer. *Biochim Biophys Acta* 2008, 1782, 800-808.

[200] Chen, W., Hu, X. T., Shi, Q. L., Zhang, F. B., He, C., Knockdown of the novel proteasome subunit Adrm1 located on the 20q13 amplicon inhibits colorectal cancer cell migration, survival and tumorigenicity. *Oncol Rep* 2009, 21, 531-537.

[201] Ni, X. G., Zhou, L., Wang, G. Q., Liu, S. M., *et al.*, The ubiquitin-proteasome pathway mediates gelsolin protein downregulation in pancreatic cancer. *Mol Med* 2008, 14, 582-589.

[202] Wang, H. X., Wang, H. M., Lin, H. Y., Yang, Q., *et al.*, Proteasome subunit LMP2 is required for matrix metalloproteinase-2 and -9 expression and activities in human invasive extravillous trophoblast cell line. *J Cell Physiol* 2006, 206, 616-623.

[203] Fang, C. H., Li, B. G., Fischer, D. R., Wang, J. J., *et al.*, Burn injury upregulates the activity and gene expression of the 20 S proteasome in rat skeletal muscle. *Clin Sci (Lond)* 2000, 99, 181-187.

[204] Huang, J., Kwong, J., Sun, E. C., Liang, T. J., Proteasome complex as a potential cellular target of hepatitis B virus X protein. *Journal of virology* 1996, 70, 5582-5591.

[205] Chen, W. N., Oon, C. J., Goo, K. S., Hepatitis B virus X protein in the proteasome of mammalian cells: defining the targeting domain. *Mol Biol Rep* 2001, 28, 31-34.

[206] Hu, Z., Zhang, Z., Doo, E., Coux, O., *et al.*, Hepatitis B virus X protein is both a substrate and a potential inhibitor of the proteasome complex. *Journal of virology* 1999, 73, 7231-7240.

[207] Hochstrasser, M., Ubiquitin-dependent protein degradation. *Annual review of genetics* 1996, 30, 405-439.

[208] Adams, J., The development of proteasome inhibitors as anticancer drugs. *Cancer cell* 2004, 5, 417-421.

[209] Zhang, Z., Protzer, U., Hu, Z., Jacob, J., Liang, T. J., Inhibition of cellular proteasome activities enhances hepadnavirus replication in an HBX-dependent manner. *Journal of virology* 2004, 78, 4566-4572.

[210] Carpenter, B., MacKay, C., Alnabulsi, A., MacKay, M., *et al.*, The roles of heterogeneous nuclear ribonucleoproteins in tumour development and progression. *Biochim Biophys Acta* 2006, 1765, 85-100.

- [211] Valente, S. T., Goff, S. P., Inhibition of HIV-1 gene expression by a fragment of hnRNP U. *Mol Cell* 2006, 23, 597-605.
- [212] Pang, R., Lee, T. K., Poon, R. T., Fan, S. T., *et al.*, Pin1 interacts with a specific serine-proline motif of hepatitis B virus X-protein to enhance hepatocarcinogenesis. *Gastroenterology* 2007, 132, 1088-1103.
- [213] Cho, W. C., Contribution of oncoproteomics to cancer biomarker discovery. *Mol Cancer* 2007, 6, 25.
- [214] Kurebayashi, J., [Biomarkers in breast cancer]. *Gan To Kagaku Ryoho* 2004, 31, 1021-1026.
- [215] Hong, P., Ng, L. F., Ren, E. C., Chen, W. N., A cell-based system for hepatitis B virus replication: significance of clinically enhanced viral replication in relation to deletions in viral core promoter. *Front Biosci* 2005, 10, 2001-2004.
- [216] Kawakami, T., Hoshida, Y., Kanai, F., Tanaka, Y., *et al.*, Proteomic analysis of sera from hepatocellular carcinoma patients after radiofrequency ablation treatment. *Proteomics* 2005, 5, 4287-4295.
- [217] Kalabay, L., Jakab, L., Prohaszka, Z., Fust, G., *et al.*, Human fetuin/alpha2HS-glycoprotein level as a novel indicator of liver cell function and short-term mortality in patients with liver cirrhosis and liver cancer. *Eur J Gastroenterol Hepatol* 2002, 14, 389-394.
- [218] Kalabay, L., Graf, L., Voros, K., Jakab, L., *et al.*, Human serum fetuin A/alpha2HS-glycoprotein level is associated with long-term survival in patients with alcoholic liver cirrhosis, comparison with the Child-Pugh and MELD scores. *BMC Gastroenterol* 2007, 7, 15.
- [219] Yokoyama, Y., Kuramitsu, Y., Takashima, M., Iizuka, N., *et al.*, Protein level of apolipoprotein E increased in human hepatocellular carcinoma. *Int J Oncol* 2006, 28, 625-631.
- [220] Steel, L. F., Shumpert, D., Trotter, M., Seeholzer, S. H., *et al.*, A strategy for the comparative analysis of serum proteomes for the discovery of biomarkers for hepatocellular carcinoma. *Proteomics* 2003, 3, 601-609.
- [221] Lozano, E., Betson, M., Braga, V. M., Tumor progression: Small GTPases and loss of cell-cell adhesion. *Bioessays* 2003, 25, 452-463.
- [222] Nobes, C. D., Hall, A., Rho GTPases control polarity, protrusion, and adhesion during cell movement. *J Cell Biol* 1999, 144, 1235-1244.
- [223] Varner, J. A., Cheresh, D. A., Integrins and cancer. *Curr Opin Cell Biol* 1996, 8, 724-730.
- [224] Gumbiner, B. M., Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell* 1996, 84, 345-357.
- [225] Lauffenburger, D. A., Horwitz, A. F., Cell migration: a physically integrated molecular process. *Cell* 1996, 84, 359-369.
- [226] Mitchison, T. J., Cramer, L. P., Actin-based cell motility and cell locomotion. *Cell* 1996, 84, 371-379.
- [227] Jaffe, A. B., Hall, A., Rho GTPases: biochemistry and biology. *Annu Rev Cell Dev Biol* 2005, 21, 247-269.
- [228] Bos, J. L., Rehmann, H., Wittinghofer, A., GEFs and GAPs: critical elements in the control of small G proteins. *Cell* 2007, 129, 865-877.
- [229] Lu, X., Wu, X., Plemenitas, A., Yu, H., *et al.*, CDC42 and Rac1 are implicated in the activation of the Nef-associated kinase and replication of HIV-1. *Curr Biol* 1996, 6, 1677-1684.
- [230] Hoppe, S., Schelhaas, M., Jaeger, V., Liebig, T., *et al.*, Early herpes simplex virus type 1 infection is dependent on regulated Rac1/Cdc42 signalling in epithelial MDCKII cells. *The Journal of general virology* 2006, 87, 3483-3494.

- [231] Clark, E. A., Golub, T. R., Lander, E. S., Hynes, R. O., Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* 2000, 406, 532-535.
- [232] Hakuma, N., Kinoshita, I., Shimizu, Y., Yamazaki, K., *et al.*, E1AF/PEA3 activates the Rho/Rho-associated kinase pathway to increase the malignancy potential of non-small-cell lung cancer cells. *Cancer Res* 2005, 65, 10776-10782.
- [233] Chung, T. W., Moon, S. K., Lee, Y. C., Kim, J. G., *et al.*, Enhanced expression of matrix metalloproteinase-9 by hepatitis B virus infection in liver cells. *Arch Biochem Biophys* 2002, 408, 147-154.
- [234] Lara-Pezzi, E., Majano, P. L., Yanez-Mo, M., Gomez-Gonzalo, M., *et al.*, Effect of the hepatitis B virus HBx protein on integrin-mediated adhesion to and migration on extracellular matrix. *J Hepatol* 2001, 34, 409-415.
- [235] Yoshioka, K., Imamura, F., Shinkai, K., Miyoshi, J., *et al.*, Participation of rho21 in serum-dependent invasion by rat ascites hepatoma cells. *FEBS letters* 1995, 372, 25-28.
- [236] Li, X. R., Ji, F., Ouyang, J., Wu, W., *et al.*, Overexpression of RhoA is associated with poor prognosis in hepatocellular carcinoma. *Eur J Surg Oncol* 2006, 32, 1130-1134.
- [237] Wang, D., Dou, K., Xiang, H., Song, Z., *et al.*, Involvement of RhoA in progression of human hepatocellular carcinoma. *Journal of gastroenterology and hepatology* 2007, 22, 1916-1920.
- [238] Fukui, K., Tamura, S., Wada, A., Kamada, Y., *et al.*, Expression and prognostic role of RhoA GTPases in hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2006, 132, 627-633.
- [239] Xue, W., Krasnitz, A., Lucito, R., Sordella, R., *et al.*, DLC1 is a chromosome 8p tumor suppressor whose loss promotes hepatocellular carcinoma. *Genes Dev* 2008, 22, 1439-1444.
- [240] Feng, H., Zhang, J., Li, X., Chen, W. N., HBX-mediated migration of HBV-replicating HepG2 cells: insights on development of hepatocellular carcinoma. *J Biomed Biotechnol* 2009, 2009, 930268.
- [241] Narumiya, S., Tanji, M., Ishizaki, T., Rho signaling, ROCK and mDia1, in transformation, metastasis and invasion. *Cancer Metastasis Rev* 2009, 28, 65-76.
- [242] Ito, H., Iwamoto, I., Morishita, R., Nozawa, Y., *et al.*, Possible role of Rho/Rhotekin signaling in mammalian septin organization. *Oncogene* 2005, 24, 7064-7072.
- [243] Bouchard, M. J., Puro, R. J., Wang, L., Schneider, R. J., Activation and inhibition of cellular calcium and tyrosine kinase signaling pathways identify targets of the HBx protein involved in hepatitis B virus replication. *Journal of virology* 2003, 77, 7713-7719.

Appendix

Appendix 1: Primer Sequences for real-time RT-PCR and cloning construction

Protein Name	Primer sequences	PCR product (bp)
β- actin	Sense: 5'- CTTAGTTGCGTTACACCCTTTC - 3'	151
	Antisense: 5'- ACCTTCACCGTTCCAGTTTT - 3'	
MACF1	Sense: 5'- TTCAATGCACTCATTACCG - 3'	170
	Antisense: 5'- GACTTTTCATCTGGAGATGG - 3'	
Annexin2	Sense: 5'- GACACCTGCTCAGTATGACG - 3'	189
	Antisense: 5'- AAGTCACCAGATGTGTCCGA - 3'	
S100A11	Sense: 5'- GTCTTCCAGAAGTATGCTGG - 3'	206
	Antisense: 5'- GCTAGGCCACCAATCAGATT - 3'	
S100A4	Sense: 5'- CACAAGTACTCGGGCAAAGA - 3'	190
	Antisense: 5'- ACAGGAAGACACAGTACTCT - 3'	
PSMA3	Sense: 5'- GCAGTAGCAGGTTTGTGTCG - 3'	176
	Antisense: 5'- CTGCAGCCAAAAGGTCTAAC - 3'	
HMGB1	Sense: 5'- TCAAAGGAGAACATCCTGGC - 3'	199
	Antisense: 5'- TCA GCC TTG ACA ACT CCC TT - 3'	
hnRNP U	Sense: 5'- TGGTCACTAACTACAAGTGG - 3'	173
	Antisense: 5'- GAGAGTTCTACTTCATCACT - 3'	
HBsAg	Forward: 5'-TCACCATATTCTTGGGAACAA-3'	1247
	Reverse: 5'-GTTTTGTTAGGGTTTAAATG-3'	
HBcAg	Forward: 5'-ATCTCCTAGACACCGCCTCA-3'	146
	Reverse: 5'-TTCCAAATTATTACCCACCC-3'	
HBx A	Forward: 5'- CCCAAGCTTGCTGCTAGGCTGTACT-3'	
HBx B	Forward: 5'- CCCAAGCTTGCTGCTAGGCTGTGCT-3'	
HBx C	Forward: 5'- CCCAAGCTTGCTGCTAGGGTGTGCT-3'	
HBx A/B/C/D	Reverse: 5'- AACTGCAGTTAGGCAGAGGTGAAA-3'	
pXJ40-BSH3	Forward: 5'- ACGGGATCCCGTCCCGTCGGCGCTGAA - 3'	135
	Reverse: 5'-CACCTCGAGTTATTATCCGCGTAAAGAGAGGTG-3'	
pXJ40-VSH3	Forward: 5'-ACGGGATCCAAGAAGAGA AAGGCCGCCAGG-3'	414
	Reverse: 5'-CACCTCGAGTTATTACTGCACGTAGCTGGCAGGGAA-3'	

Appendix 2: Primary results of differentially expressed HepG2 intra-cellular proteins affected by HBx A/B/C identified by LC/MS/MS. Run 1

N	Total	%Cov	Name	115:114	116:114	117:114
1	68.07	78.161	Serum albumin precursor - Homo sapiens (Human)	1.074898124	0.986194968	0.974170268
2	30.99	75.467	Actin, cytoplasmic 2 (Gamma-actin) - Homo sapiens (Human)	0.976781845	0.994484127	0.953285575
2	30.99	75.467	Actin, cytoplasmic 1 (Beta-actin) - Homo sapiens (Human)			
3	21.28	68.955	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) - Homo sapiens	1.023411036	1.00870502	1.029028416
4	17.22	67.972	Alpha-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (Non-neural e	0.985749185	1.030470014	1.040233254
5	16.27	39.917	Heat shock protein HSP 90-beta (HSP 84) (HSP 90) - Homo sapiens (Human)	0.972386003	0.956276417	1.047000408
6	15.28	46.537	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1) (eEF1A-1	1.01519084	0.909270346	0.985097229
7	14.19	39.64	Tubulin beta chain (Tubulin beta-5 chain) - Homo sapiens (Human)	0.918619454	0.906580985	1.029726148
7	12.19	36.18	Tubulin beta-2C chain (Tubulin beta-2 chain) - Homo sapiens (Human)			
8	12.9	40.21	Elongation factor 2 (EF-2) - Homo sapiens (Human)	0.997825444	0.96530503	1.014372349
9	12.14	64.783	Pyruvate kinase isozymes M1/M2 (EC 2.7.1.40) (Pyruvate kinase muscle isozyme)	1.068886042	1.01698339	1.065092206
10	12	37.854	Filamin-A (Alpha-filamin) (Filamin-1) (Endothelial actin-binding protein) (Ac	1.060102224	1.085949421	0.96944505
11	11.95	42.922	Lamin-A/C (70 kDa lamin) (Renal carcinoma antigen NY-REN-32) - Homo sapiens (	0.910633445	0.982139587	0.910169899
12	10.94	52.167	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8) - Homo sapien	0.917841136	0.951052427	0.982217491
13	10.79	34.532	Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERp60) (grp60) - Homo sa	0.957021117	1.0294348	1.079187155
14	10.18	43.75	Tubulin alpha-1 chain (Alpha-tubulin 1) (Testis-specific alpha-tubulin) (Tubu	0.955494821	0.99044323	0.956584752
14	10.18	35.698	Tubulin alpha-ubiquitous chain (Alpha-tubulin ubiquitous) (Tubulin K-alpha-1) - Homo sapiens (Human)			
15	13.62	33.06	Heat shock protein HSP 90-alpha (HSP 86) (Renal carcinoma antigen NY-REN-38)	1.001701236	0.956081629	0.935505331
16	9.41	40.61	Transketolase (EC 2.2.1.1) (TK) - Homo sapiens (Human)	0.99288249	1.046499252	1.026983023
17	8.61	83.333	Histone H2B type 1-N (H2B. d) (H2B/d) - Homo sapiens (Human)	0.924045205	0.940537751	1.059020996
17	8.61	71.429	Histone H2B type 1-M (H2B. e) (H2B/e) - Homo sapiens (Human)			
17	8.61	74.603	Histone H2B type 1-H (H2B. j) (H2B/j) - Homo sapiens (Human)			
17	8.61	74.603	Histone H2B type 2-F - Homo sapiens (Human)			
17	8.61	75.2	Histone H2B type 1-C/E/F/G/I (H2B. a/g/h/k/l) (H2B. 1 A) (H2B/a) (H2B/g) (H2B/h) (H2B/k) (H2B/l) - Homo sapiens (Human)			
17	8.61	74.603	Histone H2B type 1-D (H2B. b) (H2B/b) (H2B. 1 B) (HIRA-interacting protein 2) - Homo sapiens (Human)			
17	8.61	74.603	Histone H2B type 1-K (H2B K) (HIRA-interacting protein 1) - Homo sapiens (Human)			
17	8.61	77.778	Histone H2B type 1-L (H2B. c) (H2B/c) - Homo sapiens (Human)			
17	8.61	74.603	Histone H2B type F-S (H2B. s) (H2B/s) - Homo sapiens (Human)			
18	8.47	25.711	Filamin-B (FLN-B) (Beta-filamin) (Actin-binding-like protein) (Thyroid autoan	1.025283098	1.041405678	1.101811886
19	8.16	36.067	Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor (EC 6.3.4.16)	1.067365766	1.035262346	1.075342059

20	8 41.447	Nucleoside diphosphate kinase A (EC 2.7.4.6) (NDK A) (NDP kinase A) (Tumor me	0.83527714	0.885877252	0.967019856
21	7.7 42.169	L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A) (LDH muscle subunit) (L	1.001585364	1.00059092	1.006910682
22	7.7 84.375	Histone H2A type 1-H - Homo sapiens (Human)	0.881202459	1.143456459	0.777017117
22	7.7 83.077	Histone H2A type 1-D (H2A.3) - Homo sapiens (Human)			
22	7.7 83.077	Histone H2A type 1 (H2A.1) - Homo sapiens (Human)			
22	7.7 76.563	Histone H2A type 1-J - Homo sapiens (Human)			
22	7.7 80	Histone H2A type 1-E (H2A.2) (H2A/m) - Homo sapiens (Human)			
22	7.7 80	Histone H2A type 1-B - Homo sapiens (Human)			
22	7.7 66.154	Histone H2A type 1-C - Homo sapiens (Human)			
22	7.7 66.154	Histone H2A type 3 - Homo sapiens (Human)			
22	7.7 62.016	Histone H2A type 2-C (H2A-GL101) (H2A/r) - Homo sapiens (Human)			
22	7.7 63.636	Histone H2A.x (H2a/x) - Homo sapiens (Human)			
22	5.7 33.588	Histone H2A type 1-A - Homo sapiens (Human)			
23	7.59 36.977	Keratin, type I cytoskeletal 18 (Cytokeratin-18) (CK-18) (Keratin-18) (K18) -	0.973689437	0.991268575	0.785387993
24	6.98 42.611	Eukaryotic initiation factor 4A-I (EC 3.6.1.-) (ATP-dependent RNA helicase eI	1.19635582	1.114624858	1.064596653
25	6.81 24.782	60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin	0.959973454	1.022848725	1.027412534
26	6.54 44.242	Interferon-induced 17 kDa protein precursor [Contains: Ubiquitin cross-reacti	1.020855188	0.95482105	0.94512409
27	6.44 31.34	Serpin H1 precursor (Collagen-binding protein) (Colligin) (47 kDa heat shock	1.001760483	1.278514504	1.13058579
28	6.38 39.233	Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I heavy chain) (Chromobind	1.057106376	1.119233012	1.081526995
29	6.02 54.398	Keratin, type I cytoskeletal 17 (Cytokeratin-17) (CK-17) (Keratin-17) (K17) (	0.977983534	0.91740191	0.877768517
30	6.02 42.446	Phosphoglycerate kinase 1 (EC 2.7.2.3) (Primer recognition protein 2) (PRP 2)	0.95634222	0.978106022	0.970825434
31	6.01 34.091	Ribonucleoside-diphosphate reductase large subunit (EC 1.17.4.1) (Ribonucleos	1.070320249	1.129755139	0.94268012
32	6.01 35.411	Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2 / hnRNP B1) - Homo s	0.810896993	0.907110691	0.978776157
33	5.88 35.989	Fructose-bisphosphate aldolase A (EC 4.1.2.13) (Muscle-type aldolase) (Lung c	1.051790357	0.992531776	0.987031817
34	5.77 40.129	Vimentin - Homo sapiens (Human)	1.040378571	1.007321239	1.051701546
35	5.52 39.648	Brain acid soluble protein 1 (BASP1 protein) (Neuronal axonal membrane protei	0.861302316	0.860952497	0.871082544
36	6.04 19.813	Heat shock 70 kDa protein 1 (HSP70.1) (HSP70-1/HSP70-2) - Homo sapiens (Human	0.993784904	0.921440184	0.957621455
37	5.22 26.667	Protein S100-A6 (S100 calcium-binding protein A6) (Calcyclin) (Prolactin rece	0.921905518	0.883231223	0.8504318
38	4.83 41.615	Keratin, type II cytoskeletal 8 (Cytokeratin-8) (CK-8) (Keratin-8) (K8) - Hor	0.844717383	1.095456004	0.846910119
39	4.7 43.519	GTP-binding nuclear protein Ran (GTPase Ran) (Ras-like protein TC4) (Androgen	1.127312183	0.959623873	1.011659265
40	4.7 22.455	L-lactate dehydrogenase B chain (EC 1.1.1.27) (LDH-B) (LDH heart subunit) (LD	1.013187647	1.103864431	1.058676839

41	4.45	30.71	Zinc finger CCH type antiviral protein 1 (Zinc finger CCH domain-containing	1.058091521	1.194561243	0.936047137
42	4.52	27.745	Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein B-48 (Ap	0.997201443	1.035218954	0.925878704
43	4.34	32.026	Eukaryotic translation initiation factor 5A-2 (eIF-5A-2) (eIF-5A2) (Eukaryoti	1.0489434	0.958508849	0.967389822
43	4.34	24.675	Eukaryotic translation initiation factor 5A-1 (eIF-5A-1) (eIF-5A1) (Eukaryotic initiation factor 5A isoform 1) (eIF-4			
44	6.23	47.368	Nucleoside diphosphate kinase B (EC 2.7.4.6) (NDK B) (NDP kinase B) (nm23-H2)	0.953518271	0.880851209	1.041277766
44	4.23	54.745	Putative nucleoside diphosphate kinase (EC 2.7.4.6) (NDK) (NDP kinase) - Homo sapiens (Human)			
45	4.2	28.571	Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin) (Nucleolar prot	0.97574228	0.850580573	1.040689945
46	4.14	56.579	Ubiquitin - Homo sapiens (Human)	1.083095312	1.033843637	1.046540022
47	4.12	18.416	Protein disulfide-isomerase A3 precursor (EC 5.3.4.1) (Disulfide isomerase ER	0.92725414	0.904218435	0.991919339
48	4.11	31.947	4F2 cell-surface antigen heavy chain (4F2hc) (Lymphocyte activation antigen 4	1.06588161	1.366024256	1.042833805
49	4.03	30.688	Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3) - Homo sapiens (Human)	1.136475563	0.9320213	1.090949893
50	4.02	47.899	Beta-2-microglobulin precursor [Contains: Beta-2-microglobulin variant pI 5.3	1.024976134	1.148468852	1.093835235
51	4	30.057	ATP synthase subunit beta, mitochondrial precursor (EC 3.6.3.14) - Homo sapie	0.975193381	0.881460488	1.056760907
52	4	55	Profilin-1 (Profilin I) - Homo sapiens (Human)	0.947568417	0.961343706	1.097435713
53	4	14.879	Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphatase phospho-hydrolase) (PPa	1.118997455	1.045750618	1.027183294
54	3.96	68.932	Histone H4 - Homo sapiens (Human)	0.911446333	0.890382349	1.044110537
55	3.93	65.829	Peroxiorexin-1 (EC 1.11.1.15) (Thioredoxin peroxidase 2) (Thioredoxin-depend	0.953601062	0.917773545	0.859301627
56	3.85	37.143	Protein S100-A11 (S100 calcium-binding protein A11) (Protein S100C) (Calgizza	0.93246758	0.938452363	0.85609293
56	2.4	71.154	Putative S100 calcium-binding protein H_NH0456N16.1 - Homo sapiens (Human)			
57	3.75	37.455	Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose isomerase) (	1.011710644	0.92724365	0.954739809
58	3.7	40.212	Protein DJ-1 (Oncogene DJ1) - Homo sapiens (Human)	1.130449414	1.005277276	1.025914431
59	3.59	25.787	Protein disulfide-isomerase precursor (EC 5.3.4.1) (PDI) (Prolyl 4-hydroxylas	1.112554312	0.969612181	1.056876659
60	3.58	40.157	Phosphoglycerate mutase 1 (EC 5.4.2.1) (EC 5.4.2.4) (EC 3.1.3.13) (Phosphogly	0.918909729	0.938518345	0.961834669
61	3.53	30.551	ATP-dependent RNA helicase A (EC 3.6.1.-) (Nuclear DNA helicase II) (NDH II)	0.975240588	1.042873979	0.909980774
62	3.5	31.2	Galectin-3 (Galactose-specific lectin 3) (Mac-2 antigen) (IgE-binding protein	0.932616353	0.824161172	0.87670368
63	3.45	42.204	Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein) (Sin	0.939664423	0.918186605	1.002915621
64	3.31	28.505	Spliceosome RNA helicase BAT1 (EC 3.6.1.-) (DEAD box protein UAP56) (56 kDa U	1.033096552	0.847647846	1.040819287
64	3.22	26.932	ATP-dependent RNA helicase DDX39 (EC 3.6.1.-) (DEAD box protein 39) (Nuclear RNA helicase URH49) - Homo sapiens (Human)			
65	2.94	13.471	Heterogeneous nuclear ribonucleoprotein U (hnRNP U) (Scaffold attachment fact	1.371914864	1.272784472	1.147521853
66	2.83	15.148	Alpha-actinin-4 (Non-muscle alpha-actinin 4) (F-actin cross-linking protein)	1.271203518	1.103755593	1.216825008
66	1.78	25.336	Alpha-actinin-1 (Alpha-actinin cytoskeletal isoform) (Non-muscle alpha-actinin-1) (F-actin cross-linking protein) - Homo sapie			

67	2.77	21.287	Complement C3 precursor [Contains: Complement C3 beta chain; Complement C3 al	1.70083952	1.919085383	2.26708889
68	2.76	22.297	Calnexin precursor (Major histocompatibility complex class I antigen-binding	1.068298578	1.093580723	1.17816174
69	2.72	31.1	Plastin-3 (T-plastin) - Homo sapiens (Human)	0.899583399	1.105018258	1
70	2.7	24.084	Alpha-2-macroglobulin precursor (Alpha-2-M) - Homo sapiens (Human)	0.928200305	1.016087532	1.107497096
71	2.69	34.596	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit	1.266938329	1.021785855	0.770249307
72	2.67	18.655	Nuclear autoantigenic sperm protein (NASP) - Homo sapiens (Human)	1.084912539	1.458992362	1.244755507
73	2.63	65.306	Transferrin precursor (Prealbumin) (TBPA) (TTR) (ATTR) - Homo sapiens (Huma	1.254021287	1.135172248	1.206001759
74	2.59	26.987	Interferon-induced protein with tetratricopeptide repeats 1 (IFIT-1) (Interfe	1.014157295	0.722166657	0.796146691
75	3.3	38.379	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protei	1.051157355	0.957161069	0.929896593
76	2.5	27.879	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A) (Rotamase A) (C	0.935474694	0.941856802	0.899268866
77	2.42	22.791	High mobility group protein B1 (High mobility group protein 1) (HMG-1) - Homo	1.157254815	1.106149077	1.388212681
77	2.42	19.905	High mobility group protein 1-like 10 (HMG-1L10) - Homo sapiens (Human)			
78	2.38	22.963	40S ribosomal protein S17 - Homo sapiens (Human)	0.880735397	1.137970328	1.199409723
79	2.25	31.56	T-complex protein 1 subunit gamma (TCP-1-gamma) (CCT-gamma) (hTRiC5) - Homo s	0.972139955	0.883970857	0.954242229
80	2.22	35.714	Peroxiorexin-6 (EC 1.11.1.15) (Antioxidant protein 2) (1-Cys peroxiredoxin)	1.032151818	0.939143658	0.907989502
81	2.18	27.987	Polyadenylate-binding protein 1 (Poly(A)-binding protein 1) (PABP 1) - Homo s	3.29100728	6.910613537	1.209091544
81	2.04	32.805	Polyadenylate-binding protein 3 (Poly(A)-binding protein 3) (PABP 3) (Testis-specific poly(A)-binding protein) - Homo sapiens	0.852281809	1.009444118	0.934222758
82	2.18	23.423	Electron transfer flavoprotein subunit alpha, mitochondrial precursor (Alpha-	1.013132453	1.039037585	1.122125268
83	2.18	17.01	40S ribosomal protein S7 - Homo sapiens (Human)	0.957842469	1	1.033931732
84	2.14	15.599	ATP-dependent DNA helicase 2 subunit 1 (ATP-dependent DNA helicase II 70 kDa	0.951824248	1.06091857	1.05076313
85	2.12	55.133	40S ribosomal protein S4, X isoform (Single copy abundant mRNA protein) (SCRI			
85	2.07	38.023	40S ribosomal protein S4, Y isoform 1 - Homo sapiens (Human)	0.89280957	0.836690784	0.995050967
86	2.11	32.829	Heterogeneous nuclear ribonucleoprotein K (hnRNP K) (Transformation up-regula	1.079584479	0.912284791	1.148751736
87	2.1	23.699	Annexin A1 (Annexin I) (Lipocortin I) (Calpactin II) (Chromobindin-9) (p35) (	0.923370659	0.761853039	0.977377534
88	2.1	37.5	Histone H3.1t (H3t) (H3/g) - Homo sapiens (Human)			
88	2.1	50	Histone H3.3 - Homo sapiens (Human)			
88	2.1	35.294	Histone H3.2 (H3/m) (H3/o) - Homo sapiens (Human)			
88	2.1	35.556	Histone H3-like - Homo sapiens (Human)			
88	2.1	35.294	Histone H3.1 (H3/a) (H3/b) (H3/c) (H3/d) (H3/f) (H3/h) (H3/i) (H3/j) (H3/k) (H3/l) - Homo sapiens (Human)			
89	2.09	20.672	Endoplasmic precursor (Heat shock protein 90 kDa beta member 1) (94 kDa gluko	0.850475967	0.759227335	0.808840871
90	2.08	23.964	Malate dehydrogenase, mitochondrial precursor (EC 1.1.1.37) - Homo sapiens (H	1.169510245	1.04492116	1.114498854

91	2.05	26.077	Alpha-1-antitrypsin precursor (Alpha-1 protease inhibitor) (Alpha-1-antiprote	1.248705387	1.432670474	1.167887092
92	2.04	28.616	Ribose-phosphate pyrophosphokinase II (EC 2.7.6.1) (Phosphoribosyl pyrophosph	0.964236736	1.004454613	0.725133121
92	2.03	30.503	Ribose-phosphate pyrophosphokinase I (EC 2.7.6.1) (Phosphoribosyl pyrophosphate		(PRS-I) (PPRibP) - Homo sapiens	
92	2.01	18.553	Ribose-phosphate pyrophosphokinase III (EC 2.7.6.1) (Phosphoribosyl pyrophosphate		III) (PRS-III) (Phosphoribosyl py	
93	2.03	25.875	59 kDa 2'-5'-oligoadenylate synthetase-like protein (p59 OASL) (Thy	0.9918257	1.084786296	1.15449369
94	2.03	39.048	Ig lambda chain C regions - Homo sapiens (Human)	1.063955784	1.025271177	0.961384237
95	2.01	19.692	Probable ATP-dependent RNA helicase DDX17 (EC 3.6.1.-) (DEAD box protein 17)	1.058042288	0.873929203	0.841567159
96	2.01	18.136	DnaJ homolog subfamily A member 1 (Heat shock 40 kDa protein 4) (DnaJ protein	0.985207975	1.21937108	1.03937149
97	2.01	26.712	40S ribosomal protein S16 - Homo sapiens (Human)	1.165787458	1.260217071	0.951746941
98	2.01	22.083	Hepatoma-derived growth factor (HMGF) (High-mobility group protein 1-like 2)	1.012542248	0.89517647	0.972177207
99	2.01	29.412	40S ribosomal protein S20 - Homo sapiens (Human)	1.070285082	2.47487092	0.952915728
100	2.01	28.571	Cystatin-B (Stefin-B) (Liver thiol proteinase inhibitor) (CPI-B) - Homo sapi	1.157113552	1.196008205	1.149711847
101	2.01	10.082	Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Alpha-2-Z-globulin) (Ba-alpha-2	1.232044339	1.12802887	0.897903442
102	7.7	69.231	Histone H2A type 2-A (H2A.2) - Homo sapiens (Human)	0.987627983	1.021046638	0.892147958
102	7.7	62.016	Histone H2A type 2-C (H2A-GL101) (H2A/r) - Homo sapiens (Human)			
103	2	22.176	Cleavage and polyadenylation specificity factor subunit 1 (Cleavage and polya	1.306572437	1.322240233	0.997589648
104	2	34.72	ATP synthase subunit alpha, mitochondrial precursor (EC 3.6.3.14) - Homo sapi	1.178407192	0.950874865	1.134092093
105	2	38.19	Splicing factor, proline- and glutamine-rich (Polypyrimidine tract-binding pr	1.06542623	0.877174616	0.841135859
106	2	33.019	RNA-binding protein 39 (RNA-binding motif protein 39) (RNA-binding region-con	1.133446217	1.043135762	0.9120363
107	2	19.975	Transitional endoplasmic reticulum ATPase (TER ATPase) (15S Mg(2+)-ATPase p97	0.958432674	0.789516151	0.953312635
108	2	30.843	Heterogeneous nuclear ribonucleoprotein F (hnRNP F) (Nucleolin-like protein π	1.354801655	0.928850234	1.058998108
108	2	15.813	Heterogeneous nuclear ribonucleoprotein H' (hnRNP H') (FTP-3) - Homo sapiens (Human)			
108	2	12.027	Heterogeneous nuclear ribonucleoprotein H (hnRNP H) - Homo sapiens (Human)	0.982338011	1.242838144	1.192104459
109	2	58.261	40S ribosomal protein S26 - Homo sapiens (Human)			
109	2	29.565	40S ribosomal protein S26-like 1 - Homo sapiens (Human)			
110	2	64.848	40S ribosomal protein S10 - Homo sapiens (Human)	1.094502091	0.971297801	1.094764709
111	2	40.237	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial precursor (EC 1.9.3.1	1.110797405	0.903015316	1.024292707
112	2	31.532	Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1) - Homo sapi	0.949988008	0.869749725	1.128429294
113	2	18.14	Nascent polypeptide-associated complex subunit alpha (NAC-alpha) (Alpha-NAC)	0.904171407	1.123906851	1.113460779
114	2	16.863	14-3-3 protein epsilon (14-3-3E) - Homo sapiens (Human)	0.935080349	0.862876475	0.913393378
115	2	11.111	Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine hydrolase) (Ad	1.043213129	0.705483258	1.118696378

116	2	20.28	40S ribosomal protein S23 – Homo sapiens (Human)	0.942961156	0.745219886	0.918005049
117	2	21.687	Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Triose-phosphate isomerase) – H	0.647170544	1.125134349	0.82648164
118	2	13.636	Peroxioreoxin-2 (EC 1.11.1.15) (Thioredoxin peroxidase 1) (Thioredoxin-depend	0.818322241	0.936450183	0.895544231
119	2	25.962	Peptidyl-prolyl cis-trans isomerase B precursor (EC 5.2.1.8) (PPIase) (Rotama	0.955629289	0.695971608	1.0578897
120	2	11.479	Fibrinogen gamma chain precursor – Homo sapiens (Human)	0.838864148	0.703943789	1.097124457
121	2	24.752	Protein S100-A4 (S100 calcium-binding protein A4) (Metastasin) (Protein Mts1)	0.975059509	0.917643726	0.96166712
122	2	10.98	Proteasome subunit alpha type 3 (EC 3.4.25.1) (Proteasome component C8) (Macr	1.331171036	1.028310657	1.63886404
123	2	4.9808	Proliferating cell nuclear antigen (PCNA) (Cyclin) – Homo sapiens (Human)	0.795535028	0.758797884	0.878791153
124	2	18.571	60S ribosomal protein L38 – Homo sapiens (Human)	0.904581666	0.706916928	0.813600719
125	2	5.0847	Cellular nucleic acid-binding protein (CNBP) (Zinc finger protein 9) – Homo s	1.241415024	0.788852751	0.935722649
126	2	15.873	Histidine triad nucleotide-binding protein 1 (Adenosine 5'-monophosphoramidas	0.981216431	0.957770169	0.781569421
127	2	7.8261	Macrophage migration inhibitory factor (MIF) (Phenylpyruvate tautomerase) (EC	0.892187655	1.042069674	0.808291316
128	2	14.035	60S acidic ribosomal protein P1 – Homo sapiens (Human)	1.164819598	1.244967818	1.118489027
129	2	12.58	Microtubule-actin cross-linking factor 1, isoform 4 – Homo sapiens (Human)	1.692834973	1.81470716	1.174089432
129	1.99	12.67	Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 (Actin cross-linking family protein 7) (Macrophin-1) (Trabeculin-al	0.981582999	0.916399419	0.916318893
130	1.75	14.784	Plexin-A4 precursor – Homo sapiens (Human)	1.04284811	1.088826776	0.970337152
131	1.73	35.065	T-complex protein 1 subunit delta (TCP-I-delta) (CCT-delta) (Stimulator of TA	1.087208867	1.02410996	1.004477143
132	1.71	18.846	Succinyl-CoA:3-ketoacid-coenzyme A transferase 1, mitochondrial precursor (EC	1.26314044	1.052402139	1.113388777
133	1.7	22.727	40S ribosomal protein S12 – Homo sapiens (Human)	1.014768243	1.115130782	0.991063833
134	1.57	10.562	Titin (EC 2.7.11.1) (Connectin) (Rhabdomyosarcoma antigen MU-RMS-40.14) – Hon	0.944417119	0.974947572	1.019224286
135	1.54	20.548	Transmembrane emp24 domain-containing protein 10 precursor (Transmembrane pro	0.94158566	1.05612123	1.051493645
136	1.53	27.559	Activated RNA polymerase II transcriptional coactivator p15 (SUB1 homolog) (P	0.87408042	0.85365355	0.977053642
137	1.52	15.979	Reticulon-1 (Neuroendocrine-specific protein) – Homo sapiens (Human)	2.891489506	5.671278477	0.912301123
138	1.47	13.052	Regulator of differentiation 1 (Rod1) – Homo sapiens (Human)	0.98057884	0.926824033	0.994871318
139	1.47	24.299	60S ribosomal protein L10 (QM protein) (Tumor suppressor QM) (Laminin recepto	1	1.109677553	1.024170876
139	1.46	19.626	60S ribosomal protein L10-like – Homo sapiens (Human)	1.402849197	1.319385767	1.192116976
140	1.44	32	60S ribosomal protein L31 – Homo sapiens (Human)	1.076479316	1.159925103	0.932855368
141	1.4	21.19	Ankyrin repeat domain-containing protein 52 – Homo sapiens (Human)	0.959130526	1.151205778	0.886159003
142	1.33	42.9	Nuclear migration protein nudC (Nuclear distribution protein C homolog) – Hon			
143	1.32	14.195	Centromeric protein E (CENP-E) – Homo sapiens (Human)			

Appendix 2: Primary results of differentially expressed HepG2 intra-cellular proteins affected by HBx A/B/C identified by LC/MS/MS. Run 2

N	Total	%Cov	Accession,Name	115:114	116:114	117:114
1	20.25	32.184	P02768 Serum albumin precursor - Homo sapiens (Human)	1.416867971	1.570796609	1.482753515
2	16.07	31.733	P63261 Actin, cytoplasmic 2 (Gamma-actin) - Homo sapiens (Human)	1.283501387	1.260685205	1.736766338
2	16.07	31.733	P60709 Actin, cytoplasmic 1 (Beta-actin) - Homo sapiens (Human)			
3	12.14	33.925	P68363 Tubulin alpha-ubiquitous chain (Alpha-tubulin ubiquitous) (Tubulin K-	0.96900326	0.932520509	0.974192679
3	12.11	31.403	Q9BQE3 Tubulin alpha-6 chain (Alpha-tubulin 6) - Homo sapiens (Human)			
3	10.14	33.925	Q71U36 Tubulin alpha-3 chain (Alpha-tubulin 3) (Tubulin B-alpha-1) - Homo sapiens (Human)			
4	11.99	19.475	P08238 Heat shock protein HSP 90-beta (HSP 84) (HSP 90) - Homo sapiens (Huma	0.93168819	1.00589788	0.975973248
5	11.74	18.243	P07437 Tubulin beta chain (Tubulin beta-5 chain) - Homo sapiens (Human)	1.387814522	1.036465883	1.670343161
5	9.74	21.124	Q9BVA1 Tubulin beta-2B chain - Homo sapiens (Human)			
5	9.74	21.124	Q13885 Tubulin beta-2A chain - Homo sapiens (Human)			
6	11.23	21.03	P08670 Vimentin - Homo sapiens (Human)	1.170483351	1.048015714	1.335747242
7	9.25	30.149	P04406 Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) - Homo	0.841161251	0.932492137	0.858973622
8	9.07	22.093	P05783 Keratin, type I cytoskeletal 18 (Cytokeratin-18) (CK-18) (Keratin-18)	0.876859009	0.950780571	0.948316097
9	8.48	14.583	Q04695 Keratin, type I cytoskeletal 17 (Cytokeratin-17) (CK-17) (Keratin-17)	0.832151413	0.861118436	1.029235601
10	8.32	31.45	P14618 Pyruvate kinase isozymes M1/M2 (EC 2.7.1.40) (Pyruvate kinase muscle	0.869397819	0.904057801	0.932250321
11	8.2	23.377	P68104 Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1)	0.883263111	0.909411371	0.976745605
11	6.2	14.039	Q05639 Elongation factor 1-alpha 2 (EF-1-alpha-2) (Elongation factor 1 A-2) (eEFLA-2) (Statin S1) - Homo sapiens (Human)			
12	6.91	22.981	P05787 Keratin, type II cytoskeletal 8 (Cytokeratin-8) (CK-8) (Keratin-8) (K	0.764930964	0.916040123	0.824514985
13	6.49	70.313	Q99878 Histone H2A type 1-J - Homo sapiens (Human)			
13	6.49	70.313	Q96KK5 Histone H2A type 1-H - Homo sapiens (Human)			
13	6.49	69.231	P28001 Histone H2A type 1-E (H2A.2) (H2A/m) - Homo sapiens (Human)			
13	6.49	69.231	P20671 Histone H2A type 1-D (H2A.3) - Homo sapiens (Human)			
13	6.49	69.231	P0C0S8 Histone H2A type 1 (H2A.1) - Homo sapiens (Human)			
13	6.49	69.231	P04908 Histone H2A type 1-B - Homo sapiens (Human)			
13	6.49	63.846	Q93077 Histone H2A type 1-C - Homo sapiens (Human)			
13	6.49	63.846	Q7L7L0 Histone H2A type 3 - Homo sapiens (Human)			
13	6.49	63.566	Q16777 Histone H2A type 2-C (H2A-GL101) (H2A/r) - Homo sapiens (Human)			
13	6.49	47.552	P16104 Histone H2A.x (H2a/x) - Homo sapiens (Human)			
13	6	67.692	Q8IUE6 Histone H2A type 2-B - Homo sapiens (Human)			
13	4.49	45.802	Q96QV6 Histone H2A type 1-A - Homo sapiens (Human)			

14	6. 47	23. 193	Q96PK2	Microtubule-actin cross-linking factor 1, isoform 4 - Homo sapiens (H	1. 535801244	1. 321775517	1. 375199394
15	6	20. 68	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2 / hnRNP B1)	1. 054210067	0. 90153259	0. 927993655
16	5. 7	27. 857	P21333	Filamin-A (Alpha-filamin) (Filamin-1) (Endothelial actin-binding prot	1. 108194051	1. 057943544	1. 104320765
17	5. 22	48. 544	P62805	Histone H4 - Homo sapiens (Human)	0. 860935271	0. 859463692	0. 8919698
18	5. 19	69. 737	P62988	Ubiquitin - Homo sapiens (Human)	1. 070789099	1. 058956981	1. 014884114
19	5. 1	32. 448	P07355	Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I heavy chain) (Ch	1. 573667479	1. 134823802	1. 09330753
20	4. 02	11. 024	P07237	Protein disulfide-isomerase precursor (EC 5. 3. 4. 1) (PDI) (Prolyl 4-hy	1. 333991051	1. 262117863	1. 2219733
21	4. 01	12. 229	P11142	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8) - Homo sapiens (Human)			
21	4	15. 493	P54652	Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2) - Homo sapiens (Human)			
21	4	7. 4883	P34931	Heat shock 70 kDa protein 1L (Heat shock 70 kDa protein 1-like) (Heat shock 70 kDa protein 1-Hom) (HSP70-Hom) - Homo sap			
22	4	15. 531	P02765	Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Alpha-2-Z-globulin) (Ba	1. 401466966	1. 512303948	1
23	4	29. 091	P50454	Serpin H1 precursor (Collagen-binding protein) (Colligin) (47 kDa hea	0. 948886878	1. 303696313	1. 150504739
24	3. 85	7. 2	P31327	Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor (EC 6	1. 992868185	1. 05218327	1. 37652266
25	3. 74	31. 343	075369	Filamin-B (FLN-B) (Beta-filamin) (Actin-binding-like protein) (Thyroi	1. 081614642	1. 116429265	1. 149362617
26	3. 4	18. 072	P02545	Lamin-A/C (70 kDa lamin) (Renal carcinoma antigen NY-REN-32) - Homo s	0. 760641038	0. 779765368	0. 694791615
27	3. 36	18. 788	P62937	Peptidyl-prolyl cis-trans isomerase A (EC 5. 2. 1. 8) (PPIase A) (Rotama	0. 818034053	0. 847833216	0. 942219973
28	3. 23	23. 041	P06733	Alpha-enolase (EC 4. 2. 1. 11) (2-phospho-D-glycerate hydro-lyase) (Non-	0. 881803632	0. 956423283	0. 958121598
29	3. 1	42. 609	P05387	60S acidic ribosomal protein P2 (Renal carcinoma antigen NY-REN-44) -	1. 412437201	0. 891330361	1. 097107649
30	2. 64	18. 293	Q04917	14-3-3 protein eta (Protein AS1) - Homo sapiens (Human)	0. 937207282	0. 947619319	1. 060938835
30	2. 64	18. 219	P61981	14-3-3 protein gamma (Protein kinase C inhibitor protein 1) (KCIP-1) - Homo sapiens (Human)			
30	2. 64	14. 516	P31947	14-3-3 protein sigma (Stratifin) (Epithelial cell marker protein 1) - Homo sapiens (Human)			
30	2. 64	14. 634	P31946	14-3-3 protein beta/alpha (Protein kinase C inhibitor protein 1) (KCIP-1) (Protein 1054) - Homo sapiens (Human)			
30	2. 64	9. 3878	P27348	14-3-3 protein theta (14-3-3 protein tau) (14-3-3 protein T-cell) (HS1 protein) - Homo sapiens (Human)			
30	2	23. 673	P63104	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein 1) (KCIP-1) - Homo sapiens (Human)			
30	2	16. 078	P62258	14-3-3 protein epsilon (14-3-3E) - Homo sapiens (Human)			
31	2. 59	7. 8125	P08758	Annexin A5 (Annexin V) (Lipocortin V) (Endonexin II) (Calphobindin I)	0. 927947521	1. 013292313	0. 998185813
32	2. 55	21. 117	P10809	60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa ch	0. 891090333	0. 854366899	0. 809700668
33	2. 43	10	P01857	Ig gamma-1 chain C region - Homo sapiens (Human)	0. 890417576	1. 124726892	1. 335312605
34	2. 4	38. 889	Q8N257	Histone H2B type 3-B (H2B type 12) - Homo sapiens (Human)	0. 994504273	1. 120775342	0. 918931425
34	2. 12	42. 063	Q99880	Histone H2B type 1-L (H2B. c) (H2B/c) - Homo sapiens (Human)			
34	2. 12	40. 476	P57053	Histone H2B type F-S (H2B. s) (H2B/s) - Homo sapiens (Human)			

34	2.12	33.333	Q99879 Histone H2B type 1-M (H2B.e) (H2B/e)	- Homo sapiens (Human)					
34	2.12	36.508	Q99877 Histone H2B type 1-N (H2B.d) (H2B/d)	- Homo sapiens (Human)					
34	2.12	36.508	Q93079 Histone H2B type 1-H (H2B.j) (H2B/j)	- Homo sapiens (Human)					
34	2.12	36.508	Q5QNW6 Histone H2B type 2-F - Homo sapiens (Human)						
34	2.12	36.508	Q16778 Histone H2B type 2-E (H2B.q) (H2B/q)	(H2B-GL105) - Homo sapiens (Human)					
34	2.12	36.8	P62807 Histone H2B type 1-C/E/F/G/I (H2B.a/g/h/k/l) (H2B.1 A) (H2B/a)	(H2B/g) (H2B/h) (H2B/k) (H2B/l) - Homo sapiens (Human)					
34	2.12	36.508	P58876 Histone H2B type 1-D (H2B.b) (H2B/b)	(H2B.1 B) (HIRA-interacting protein 2) - Homo sapiens (Human)					
34	2.12	36.508	P23527 Histone H2B type 1-O (H2B.n) (H2B/n)	- Homo sapiens (Human)					
34	2.12	36.508	P06899 Histone H2B type 1-J (H2B.r) (H2B/r)	- Homo sapiens (Human)					
34	2.12	36.508	O60814 Histone H2B type 1-K (H2B.k) (HIRA-interacting protein 1) - Homo sapiens (Human)						
34	2.12	30.159	P33778 Histone H2B type 1-B (H2B.f) (H2B/f)	- Homo sapiens (Human)					
35	2.21	13.065	Q06830 Peroxiredoxin-1 (EC 1.11.1.15) (Thioredoxin peroxidase 2) (Thioredoxi	0.816689253 0.931481516 0.778737175					
35	2.28.	283	P32119 Peroxiredoxin-2 (EC 1.11.1.15) (Thioredoxin peroxidase 1) (Thioredoxin-dependent peroxide reductase 1) (Thiol-specific a						
36	2.01	24.51	P61604 10 kDa heat shock protein, mitochondrial (Hsp10) (10 kDa chaperonin)	1 0.934330285 0.833063304					
37	6.49	63.846	Q6F113 Histone H2A type 2-A (H2A.2) - Homo sapiens (Human)						
37	6.49	63.566	Q16777 Histone H2A type 2-C (H2A-GL101) (H2A/r) - Homo sapiens (Human)						
38	4	17.278	P11021 78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kD	1.031747222 0.702177465 0.963746011					
39	2	17.456	P40926 Malate dehydrogenase, mitochondrial precursor (EC 1.1.1.37) - Homo sa	5.378787518 2.057178736 5.158246517					
40	2	11.089	P09914 Interferon-induced protein with tetratricopeptide repeats 1 (IFIT-1)	0.892323054 0.881472933 0.82995975					
41	2	8.4866	P23246 Splicing factor, proline- and glutamine-rich (Polyprimidine tract-bi	1.483855844 1.062240839 1.285309792					
42	2	16.244	Q9UQ80 Proliferation-associated protein 2G4 (Cell cycle protein p38-2G4 homo	1.283645153 1.373889923 0.861585319					
43	2	13.09	O75390 Citrate synthase, mitochondrial precursor (EC 2.3.3.1) - Homo sapiens	2.785855532 2.007375956 3.069790125					
44	2	5.4514	Q96BF6 BTB/P0Z domain-containing protein 14A - Homo sapiens (Human)						
45	2	17.391	P26447 Protein S100-A4 (S100 calcium-binding protein A4) (Metastasin) (Prote	0.976544857 0.996878772 0.957420528					
46	2	9.9338	P05386 60S acidic ribosomal protein P1 - Homo sapiens (Human)	1.083065181 1.196992369 1.231351628					
47	2	14.474	P09429 High mobility group protein B1 (High mobility group protein 1) (HMG-1	1.244345412 1.149754996 1.206974249					
47	2	7.2993	O60361 Putative nucleoside diphosphate kinase (EC 2.7.4.6) (NDK) (NDP kinase) - Homo sapiens (Human)						
48	2	6.929	P18669 Phosphoglycerate mutase 1 (EC 5.4.2.1) (EC 5.4.2.4) (EC 3.1.3.13) (Ph	1.779764295 1.253805757 2.995791912					
48	2	4.3478	P15259 Phosphoglycerate mutase 2 (EC 5.4.2.1) (EC 5.4.2.4) (EC 3.1.3.13) (Phosphoglycerate mutase isozyme M) (PGAM-M) (BPG-depe						
49	2	2.6882	P14866 Heterogeneous nuclear ribonucleoprotein L (hnRNP L) - Homo sapiens (H	1.000715375 0.867060184 0.905409336					
50	2	9.6154	P14649 Myosin light polypeptide 6B (Smooth muscle and nonmuscle myosin light	1.101158619 1.192836642 1.30368948					

51	2	10.169	P08865	40S ribosomal protein SA (p40) (34/67 kDa laminin receptor) (Colon ca	0.88184303	0.793877065	0.839645803
52	2	6.4417	P01859	Ig gamma-2 chain C region - Homo sapiens (Human)	0.844317138	1.034251094	1.688803434
53	2	7.8431	Q9CZV4	Eukaryotic translation initiation factor 5A-2 (eIF-5A-2) (eIF-5A2) (E	0.989974141	1	1.168985248
53	2	7.7922	P63241	Eukaryotic translation initiation factor 5A-1 (eIF-5A-1) (eIF-5A1) (Eukaryotic initiation factor 5A isoform 1) (eIF-5A)	2.981537104	1.974576831	3.486536741
54	2	10.476	P99999	Cytochrome c - Homo sapiens (Human)	0.874081235	0.779082656	0.899945726
55	2	4.9808	P12004	Proliferating cell nuclear antigen (PCNA) (Cyclin) - Homo sapiens (Hu	1.320044756	0.788885415	1.018033504
56	2	12.381	P10599	Thioredoxin (Trx) (ATL-derived factor) (ADF) (Surface-associated sulp	1.309802175	1.266362309	1.553558707
57	11.95	22.131	P07900	Heat shock protein HSP 90-alpha (HSP 86) (Renal carcinoma antigen NY-	0.811953843	0.962877691	0.850382924
58	1.87	24.107	P30041	Peroxioredoxin-6 (EC 1.11.1.15) (Antioxidant protein 2) (1-Cys peroxir	0.971383154	1.000645041	0.777815998
59	3.7	14.509	P08107	Heat shock 70 kDa protein 1 (HSP70.1) (HSP70-1/HSP70-2) - Homo sapien	0.868732333	0.946811855	0.97476691
60	1.7	9.0767	Q8IUG5	Myosin XVIIIb - Homo sapiens (Human)	0.920219522	0.872984686	1.0138307
60	1.7	7.2816	Q00839	Heterogeneous nuclear ribonucleoprotein U (hnRNP U) (Scaffold attachment factor A) (SAF-A) (p120) (ppl20) - Homo sapiens	0.913183212	0.762500083	0.884071813
61	1.7	14.474	P15531	Nucleoside diphosphate kinase A (EC 2.7.4.6) (NDK A) (NDP kinase A) (	1.286548018	1.097204924	1.522073865
62	1.7	10.476	P31949	Protein S100-A11 (S100 calcium-binding protein A11) (Protein S100C) (	0.901139557	0.857904911	0.979795039
63	1.52	42.105	P61956	Small ubiquitin-related modifier 2 precursor (SUMO-2) (Ubiquitin-like	0.941289127	0.948151231	0.965415061
63	1.52	38.835	P55854	Small ubiquitin-related modifier 3 precursor (SUMO-3) (Ubiquitin-like protein SMT3A) (SMT3 homolog 1) (SUMO-2) - Homo sa	1.188141823	1.166906118	1.424201727
63	1.52	20	Q6EEV6	Small ubiquitin-related modifier 4 precursor (SUMO-4) (Small ubiquitin-like protein 4) - Homo sapiens (Human)	1.142655134	1.050152588	1.335917473
64	3.42	38.623	P07195	L-lactate dehydrogenase B chain (EC 1.1.1.27) (LDH-B) (LDH heart subu	4.559123039	4.217110157	8.091453552
65	1.4	50.667	P20674	Cytochrome c oxidase subunit 5A, mitochondrial precursor (EC 1.9.3.1)	1.007849336	0.932519257	0.959953666
66	1.3	3.3708	Q96PG2	Membrane-spanning 4-domains subfamily A member 10 - Homo sapiens (Human)			
67	1.22	21.361	P06576	ATP synthase subunit beta, mitochondrial precursor (EC 3.6.3.14) - Ho			
68	1.1	6.1869	P23921	Ribonucleoside-diphosphate reductase large subunit (EC 1.17.4.1) (Rib			
69	1.1	20	P46783	40S ribosomal protein S10 - Homo sapiens (Human)			
70	1.02	14.569	P13639	Elongation factor 2 (EF-2) - Homo sapiens (Human)			

Appendix 2: Primary results of differentially expressed HepG2 intra-cellular proteins affected by HBx A/B/C identified by LC/MS/MS. Run 3

N	Total	%Cov	Accessi.Name	115:114	116:114	117:114
1	21.91	48.8	P63261 Actin, cytoplasmic 2 (Gamma-actin) - Homo sapiens (Human)	1.067170382	1.107196093	1.248507619
1	21.91	48.8	P60709 Actin, cytoplasmic 1 (Beta-actin) - Homo sapiens (Human)			
2	21.46	36.453	P02768 Serum albumin precursor - Homo sapiens (Human)	1.408526897	1.450983405	1.407720804
3	13.74	31.982	P07437 Tubulin beta chain (Tubulin beta-5 chain) - Homo sapiens (Human)	1.143396735	0.960705876	1.199599028
4	13.14	35.477	P68363 Tubulin alpha-ubiquitous chain (Alpha-tubulin ubiquitous) (Tubulin K-	1.033692122	0.955003738	0.971256196
4	13.11	32.962	Q9BQE3 Tubulin alpha-6 chain (Alpha-tubulin 6) - Homo sapiens (Human)			
4	11.14	35.698	Q71U36 Tubulin alpha-3 chain (Alpha-tubulin 3) (Tubulin B-alpha-1) - Homo sapiens (Human)			
5	12.82	39.827	P68104 Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1)	0.894836366	0.921915829	0.929505169
6	12.61	29.834	P08238 Heat shock protein HSP 90-beta (HSP 84) (HSP 90) - Homo sapiens (Huma	0.996147275	1.024151444	0.994133353
7	12.57	29.399	P08670 Vimentin - Homo sapiens (Human)	1.19591558	1.088092446	1.245443463
8	12.52	37.288	P14618 Pyruvate kinase isozymes M1/M2 (EC 2.7.1.40) (Pyruvate kinase muscle	0.907909989	0.969161689	0.931237221
9	11.37	39.104	P04406 Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) - Homo	0.896468461	0.986387432	0.908559084
10	10.51	43.256	P05783 Keratin, type I cytoskeletal 18 (Cytokeratin-18) (CK-18) (Keratin-18)	0.91579473	0.97583586	0.945452929
11	10.12	37.037	Q04695 Keratin, type I cytoskeletal 17 (Cytokeratin-17) (CK-17) (Keratin-17)	0.887714565	0.910533071	1.019862175
12	9.49	32.712	P05787 Keratin, type II cytoskeletal 8 (Cytokeratin-8) (CK-8) (Keratin-8) (K	0.787111104	0.841487646	0.767083287
13	8.8	79.231	P28001 Histone H2A type 1-E (H2A. 2) (H2A/m) - Homo sapiens (Human)			
13	8.8	79.231	P20671 Histone H2A type 1-D (H2A. 3) - Homo sapiens (Human)			
13	8.8	79.231	P04908 Histone H2A type 1-B - Homo sapiens (Human)			
13	8.8	70.313	Q99878 Histone H2A type 1-J - Homo sapiens (Human)			
13	8.8	70.313	Q96KK5 Histone H2A type 1-H - Homo sapiens (Human)			
13	8.8	69.231	P0C0S8 Histone H2A type 1 (H2A. 1) - Homo sapiens (Human)			
13	8.8	63.846	Q93077 Histone H2A type 1-C - Homo sapiens (Human)			
13	8.8	63.846	Q7L7L0 Histone H2A type 3 - Homo sapiens (Human)			
13	8.8	58.915	Q16777 Histone H2A type 2-C (H2A-GL101) (H2A/r) - Homo sapiens (Human)			
13	7.1	81.119	P16104 Histone H2A. x (H2a/x) - Homo sapiens (Human)			
14	7.4	29.745	P22626 Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2 / hnRNP B1)	1.027161837	0.871396661	0.874514282
15	6.54	23.193	O43707 Alpha-actinin-4 (Non-muscle alpha-actinin 4) (F-actin cross-linking p	1.270166068	1.199554925	1.13345592
16	6.04	23.501	P00558 Phosphoglycerate kinase 1 (EC 2.7.2.3) (Primer recognition protein 2)	1.088245273	1.160919547	1.051242709
17	5.98	35.714	P06733 Alpha-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (Non-	0.856018186	0.950890005	0.890856385
18	5.9	32.078	P02545 Lamin-A/C (70 kDa lamin) (Renal carcinoma antigen NY-REN-32) - Homo s	0.809414864	0.835804582	0.763477087

19	6.06	41.791	P08729	Keratin, type II cytoskeletal 7 (Cytokeratin-7) (CK-7) (Keratin-7) (K	0.954788804	0.966620505	0.881420732
20	5.77	40.663	P10809	60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa ch	0.949640572	0.978381395	0.95041573
21	5.75	20.743	P11142	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8) - Homo sapiens (Human)			
21	4.04	14.085	P54652	Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2) - Homo sapiens (Human)			
21	4.01	13.729	P34931	Heat shock 70 kDa protein 1L (Heat shock 70 kDa protein 1-like) (Heat shock 70 kDa protein 1-Hom) (HSP70-Hom) - Homo sap			
22	5.7	27.857	P07737	Profilin-1 (Profilin I) - Homo sapiens (Human)	1.230686545	1.18844974	1.243348241
23	5.38	85.526	P62988	Ubiquitin - Homo sapiens (Human)	1.079840064	1.055675626	0.983523965
24	5.26	64.078	P62805	Histone H4 - Homo sapiens (Human)	0.891231358	0.854763627	0.867141366
25	5.1	41.298	P07355	Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I heavy chain) (Ch	1.312323833	1.112296033	1.073517718
26	5.1	38.182	P05161	Interferon-induced 17 kDa protein precursor [Contains: Ubiquitin cros	0.905392647	1.069889903	0.828169405
27	4.02	17.717	P07237	Protein disulfide-isomerase precursor (EC 5.3.4.1) (PDI) (Prolyl 4-hy	1.297070861	1.265786886	1.132244706
28	4	27.52	P02765	Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Alpha-2-Z-globulin) (Ba	1.488880396	1.551087856	1.001817465
29	3.95	21.818	P62937	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A) (Rotama	0.932153165	0.89973855	1.031059384
30	3.93	51.111	P06703	Protein S100-A6 (S100 calcium-binding protein A6) (Calcyclin) (Prolac	0.793528676	1.012743711	0.819372237
31	3.87	12.933	P31327	Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor (EC 6	1.978549123	1.04334259	1.313028455
32	5.71	28.746	P11021	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kD	0.848817587	0.79147625	0.845806181
33	3.68	23.96	P30101	Protein disulfide-isomerase A3 precursor (EC 5.3.4.1) (Disulfide isom	0.871230364	0.894391537	0.801694572
34	3.3	27.711	P23528	Cofilin-1 (Cofilin, non-muscle isoform) (18 kDa phosphoprotein) (p18)	0.745539784	0.915733874	0.83015877
35	3.26	66.667	Q93079	Histone H2B type 1-H (H2B.j) (H2B/j) - Homo sapiens (Human)	0.938583434	0.958571613	0.90503794
35	3.26	66.667	Q5QNW6	Histone H2B type 2-F - Homo sapiens (Human)			
35	2.86	53.968	Q8N257	Histone H2B type 3-B (H2B type 12) - Homo sapiens (Human)			
35	3.26	67.2	P62807	Histone H2B type 1-C/E/F/G/I (H2B.a/g/h/k/l) (H2B.1 A) (H2B/a) (H2B/g) (H2B/h) (H2B/k) (H2B/l) - Homo sapiens (Human)			
35	3.26	66.667	O60814	Histone H2B type 1-K (H2B K) (HIRA-interacting protein 1) - Homo sapiens (Human)			
35	3.26	58.73	Q16778	Histone H2B type 2-E (H2B.q) (H2B/q) (H2B-GL105) - Homo sapiens (Human)			
35	3.26	65.873	P57053	Histone H2B type F-S (H2B.s) (H2B/s) - Homo sapiens (Human)			
35	3.26	58.73	P06899	Histone H2B type 1-J (H2B.r) (H2B/r) (H2B.1) - Homo sapiens (Human)			
35	2.82	67.46	Q99880	Histone H2B type 1-L (H2B.c) (H2B/c) - Homo sapiens (Human)			
35	3.26	62.698	Q99877	Histone H2B type 1-N (H2B.d) (H2B/d) - Homo sapiens (Human)			
35	3.26	62.698	P58876	Histone H2B type 1-D (H2B.b) (H2B/b) (H2B.1 B) (HIRA-interacting protein 2) - Homo sapiens (Human)			
35	3.26	58.73	P23527	Histone H2B type 1-O (H2B.n) (H2B/n) (H2B.2) - Homo sapiens (Human)			
35	3.26	51.587	P33778	Histone H2B type 1-B (H2B.f) (H2B/f) (H2B.1) - Homo sapiens (Human)			

35	3.14	57.937	Q99879	Histone H2B type 1-M (H2B.e) (H2B/e) - Homo sapiens (Human)				
36	3.26	33.198	P61981	14-3-3 protein gamma (Protein kinase C inhibitor protein 1) (KCIP-1)	0.968537509	0.843318939	0.807768285	
36	2.71	24.49	P27348	14-3-3 protein theta (14-3-3 protein tau) (14-3-3 protein T-cell) (HS1 protein) - Homo sapiens (Human)				
36	2.01	25.098	P62258	14-3-3 protein epsilon (14-3-3E) - Homo sapiens (Human)				
36	2.64	18.293	Q04917	14-3-3 protein eta (Protein AS1) - Homo sapiens (Human)				
36	2.64	23.79	P31947	14-3-3 protein sigma (Stratifin) (Epithelial cell marker protein 1) - Homo sapiens (Human)				
36	2.64	14.634	P31946	14-3-3 protein beta/alpha (Protein kinase C inhibitor protein 1) (KCIP-1) (Protein 1054) - Homo sapiens (Human)				
36	2	23.673	P63104	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein 1) (KCIP-1) - Homo sapiens (Human)				
37	5.24	19.501	P15531	Nucleoside diphosphate kinase A (EC 2.7.4.6) (NDK A) (NDP kinase A) (	0.774881542	0.897062063	0.774763048	
38	3.1	42.609	P05387	160S acidic ribosomal protein P2 (Renal carcinoma antigen NV-REN-44) -	1.500534773	0.914189041	1.099101543	
39	2.83	34.264	Q9UQ80	Proliferation-associated protein 2G4 (Cell cycle protein p38-2G4 homo	1.204942822	1.175192237	0.900485396	
40	2.62	12.188	P08758	Annexin A5 (Annexin V) (Lipocortin V) (Endonexin II) (Calphobindin I)	0.985826254	1.039278746	1	
41	2.61	27.136	Q06830	Peroxioredoxin-1 (EC 1.11.1.15) (Thioredoxin peroxidase 2) (Thioredoxi	0.867628455	0.852805257	0.980516016	
41	2	22.222	P32119	Peroxioredoxin-2 (EC 1.11.1.15) (Thioredoxin peroxidase 1) (Thioredoxin-dependent peroxide reductase 1) (Thiol-specific a				
42	2.43	16.061	P01857	Ig gamma-1 chain C region - Homo sapiens (Human)	0.945955396	1.15357101	1.337739468	
43	2.38	38.411	P04080	Cystatin-B (Stefin-B) (Liver thiol proteinase inhibitor) (CPI-B) - Ho	1.04874783	1.102047896	1.093335152	
44	2.38	19.048	P31949	Protein S100-A11 (S100 calcium-binding protein A11) (Protein S100C) (	0.970140994	0.885901642	0.845933146	
45	2.27	20.952	P10599	Thioredoxin (Trx) (ATL-derived factor) (ADF) (Surface-associated sulp	1.402379632	0.870662153	1.032657981	
46	2.24	12.987	P63241	Eukaryotic translation initiation factor 5A-1 (eIF-5A-1) (eIF-5A1) (E	1.037939429	0.960866272	1.113572359	
46	2	7.8431	Q9GZV4	Eukaryotic translation initiation factor 5A-2 (eIF-5A-2) (eIF-5A2) (Eukaryotic initiation factor 5A isoform 2) - Homo sa				
47	3.51	20.699	P09651	Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing prote	0.84535789	1.023234177	0.88119507	
48	2.02	26.331	P40926	Malate dehydrogenase, mitochondrial precursor (EC 1.1.1.37) - Homo sa	5.714277744	2.109936237	5.167621613	
49	2.02	12.247	P23921	Ribonucleoside-diphosphate reductase large subunit (EC 1.17.4.1) (Rib	1.076298523	1.307858944	1.141980457	
50	2.02	11.877	P12004	Proliferating cell nuclear antigen (PCNA) (Cyclin) - Homo sapiens (Hu	0.829641938	0.901735735	0.590375185	
51	2.01	20.282	P19338	Nucleolin (Protein C23) - Homo sapiens (Human)		3.200688601		
52	2.01	38.235	P61604	10 kDa heat shock protein, mitochondrial (Hsp10) (10 kDa chaperonin)	1.062372804	0.95829165	0.834577382	
53	2.01	8.6331	P25788	Proteasome subunit alpha type 3 (EC 3.4.25.1) (Proteasome component C	1.251919031	1.100750842	1.236029148	
54	8.8	63.846	Q6F113	Histone H2A type 2-A (H2A.2) - Homo sapiens (Human)				
54	8.8	58.915	Q16777	Histone H2A type 2-C (H2A-GL101) (H2A/r) - Homo sapiens (Human)				
55	2.03	13.559	P08865	40S ribosomal protein SA (p40) (34/67 kDa laminin receptor) (Colon ca	0.977615356	0.91156131	0.958783448	
56	2	19.437	Q04637	Eukaryotic translation initiation factor 4 gamma 1 (eIF-4-gamma 1) (eIF-4G 1) (p220) - Homo sapiens (Human)				

57	2 57. 917 P63162 Small nuclear ribonucleoprotein-associated protein N (snRNP-N) (Sm protein N) (Sm-N) (SmN) (Sm-D) (Tissue-specific-spl	
57	2 39. 583 P14678 Small nuclear ribonucleoprotein-associated proteins B and B' (snRNP-B) (Sm protein B/B') (Sm-B/Sm-B') (SmB/SmB') - Homo	
58	2 21. 358 P23246 Splicing factor, proline- and glutamine-rich (Polypyrimidine tract-bi	1. 576408148 1. 089482546 1. 287645817
59	2 24. 074 P67809 Nuclease sensitive element-binding protein 1 (Y-box-binding protein 1	0. 851873517 1. 0863446 1. 076388121
60	2 26. 824 075390 Citrate synthase, mitochondrial precursor (EC 2. 3. 3. 1) - Homo sapiens	2. 959617138 2. 058856249 3. 075369358
61	2 11. 244 Q96BF6 BTB/POZ domain-containing protein 14A - Homo sapiens (Human)	
62	2 9. 6774 P14866 Heterogeneous nuclear ribonucleoprotein L (hnRNP L) - Homo sapiens (H	1. 063132763 0. 889296412 0. 907054901
63	2 32. 895 P22392 Nucleoside diphosphate kinase B (EC 2. 7. 4. 6) (NDK B) (NDP kinase B) (	1. 205097675 1. 056163549 1. 229204297
63	2 21. 898 060361 Putative nucleoside diphosphate kinase (EC 2. 7. 4. 6) (NDK) (NDP kinase) - Homo sapiens (Human)	
64	2 41. 739 P14174 Macrophage migration inhibitory factor (MIF) (Phenylpyruvate tautomer	0. 922314942 1. 052025437 0. 908373833
65	2 15. 337 P01859 Ig gamma-2 chain C region - Homo sapiens (Human)	0. 89697957 1. 060775042 1. 691872835
66	2 11. 417 P18669 Phosphoglycerate mutase 1 (EC 5. 4. 2. 1) (EC 5. 4. 2. 4) (EC 3. 1. 3. 13) (Ph	1. 890773177 1. 285960317 3. 001236677
66	2 14. 625 P15259 Phosphoglycerate mutase 2 (EC 5. 4. 2. 1) (EC 5. 4. 2. 4) (EC 3. 1. 3. 13) (Phosphoglycerate mutase isozyme M) (PGAM-M) (BPG-depe	
67	2 19. 231 P14649 Myosin light polypeptide 6B (Smooth muscle and nonmuscle myosin light	1. 169841051 1. 223427534 1. 306058884
67	2 18. 543 P60660 Myosin light polypeptide 6 (Smooth muscle and nonmuscle myosin light chain alkali 6) (Myosin light chain alkali 3) (Myos	
68	2 39. 506 Q15843 NEDD8 precursor (Ubiquitin-like protein Nedd8) (Neddylin) - Homo sapiens (Human)	
69	2 16. 535 P06703 Protein S100-A6 (S100 calcium-binding protein A6) (Calcyclin) (Prolac	0. 936372015 0. 926258309 0. 920530431
70	2 14. 286 P99999 Cytochrome c - Homo sapiens (Human)	3. 167504072 2. 025215864 3. 492873669
71	12. 28 29. 235 P07900 Heat shock protein HSP 90-alpha (HSP 86) (Renal carcinoma antigen NY-	1. 391498208 1. 298838854 1. 556382298
72	1. 89 24. 107 P30041 Peroxiredoxin-6 (EC 1. 11. 1. 15) (Antioxidant protein 2) (1-Cys peroxir	0. 843440771 0. 989616811 0. 847334743
73	1. 72 15. 777 Q00839 Heterogeneous nuclear ribonucleoprotein U (hnRNP U) (Scaffold attachm	1. 191424201 1. 337952163 1. 245442805
73	1. 7 19. 439 Q8IUG5 Myosin-18B (Myosin XVIIIb) - Homo sapiens (Human)	
74	1. 7 40. 789 P15531 Nucleoside diphosphate kinase A (EC 2. 7. 4. 6) (NDK A) (NDP kinase A) (	1. 530050039 1. 879992485 1. 476509452
75	1. 65 32. 121 P46783 40S ribosomal protein S10 - Homo sapiens (Human)	1. 763340473 1. 803196549 2. 109605312
76	1. 52 47. 368 P61956 Small ubiquitin-related modifier 2 precursor (SUMO-2) (Ubiquitin-like	1. 366793633 1. 125343323 1. 524840236
76	1. 52 39. 806 P55854 Small ubiquitin-related modifier 3 precursor (SUMO-3) (Ubiquitin-like protein SMT3A) (SMT3 homolog 1) (SUMO-2) - Homo sa	
76	1. 52 28. 421 Q6EEV6 Small ubiquitin-related modifier 4 precursor (SUMO-4) (Small ubiquitin-like protein 4) - Homo sapiens (Human)	
77	3. 41 42. 216 P07195 L-lactate dehydrogenase B chain (EC 1. 1. 1. 27) (LDH-B) (LDH heart subu	0. 760451376 0. 765828848 0. 827682078
78	1. 31 38 P20674 Cytochrome c oxidase subunit 5A, mitochondrial precursor (EC 1. 9. 3. 1)	1 0. 972467065 0. 967169702

Appendix3: Protein profile of differentially expressed secreted proteins in HBV replicating HepG2 cells. **Run 1.**

N	Unuse:Accession Name	115:114	PVal	115:116:114	PVal	116:117:114	PVal	117:118:119
1	89.48 sp P02768 Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	0.9529	0.6575	1.1943	0.0244	0.7317	0.243	
2	8.62 sp P19823 Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapiens GN=ITI2 PE=1 SV=1	1.1961	0.3449	1.1593	0.339	0.7367	0.1925	
3	8.53 sp P02647 Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	1.0552	0.8996	1.0668	0.8803	0.8398	0.489	
4	6.98 sp P01024 Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	0.8686	0.3021	0.9273	0.1366	1.088	0.6221	
5	6.71 sp P01023 Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=1	0.8673	0.2801	0.9015	0.2452	1.0141	0.4492	
6	6.62 sp P02788 Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6							
7	6.53 sp P02771 Alpha-fetoprotein OS=Homo sapiens GN=AFP PE=1 SV=1	0.9472	0.0728	0.6845	0.0207	1.0583	0.4503	
8	6.09 sp P63261 Actin, cytoplasmic 2 OS=Homo sapiens GN=ACTG1 PE=1 SV=1	1.1545	0.1067	1.4806	0.1572	1.2684	0.212	
9	5.15 sp P01857 Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=	0.9231	0.5136	1	1	0.9615	0.2067	
10	4.22 sp P01859 Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=;	1.0471	0.038	0.9715	0.8084	0.9993	0.9585	
11	4.02 sp P69905 Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2	1.1581	0.3893	1.173	0.6737	1.0148	0.9784	
12	4 sp P01834 Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	1.0726	0.3033	1.1298	0.4813	1.2656	0.0049	
13	3.6 sp P02765 Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	1.2751	0.5377	1.5116	0.3888	1.0158	0.9856	
14	3.3 sp P01842 Ig lambda chain C regions OS=Homo sapiens GN=IGLC1 PE=1 SV=							
15	2.49 sp Q99707 Methionine synthase OS=Homo sapiens GN=MTR PE=1 SV=2	1.1684		0.912		0.9608		
16	2.46 sp P01042 Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	0.8377		0.9332		1.058		
17	2.24 sp P01008 Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1	0.7743		0.917		0.8898		
18	2 sp Q68CQ4 Digestive organ expansion factor homolog OS=Homo sapiens GN=	0.9255		1.0004		1.0196		
19	2 sp Q95497 Pantetheinase OS=Homo sapiens GN=VNN1 PE=1 SV=2	0.9705		0.976		1		
20	2 sp P02774 Vitamin D-binding protein OS=Homo sapiens GN=GC PE=1 SV=1	0.7907		0.9087		0.862		
21	2 sp P23142 Fibulin-1 OS=Homo sapiens GN=FBLN1 PE=1 SV=4	1		0.9887		0.8819		
22	2 sp P02679 Fibrinogen gamma chain OS=Homo sapiens GN=FGG PE=1 SV=3							
23	1.75 sp Q81WJ2 GRIP and coiled-coil domain-containing protein 2 OS=Homo sapiens GN=GCC2 PE=1 SV=3	1.0272		1.0562		1.0465		
24	1.73 sp P0C0L5 Complement C4-B OS=Homo sapiens GN=C4B PE=1 SV=1	1.4734		1.1979		0.3422		
25	1.7 sp P01009 Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	1.0977		1.0655		1.1012		
26	1.52 sp P01876 Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=;	0.7458		0.8599		0.924		
27	1.4 sp P02042 Hemoglobin subunit delta OS=Homo sapiens GN=HBD PE=1 SV=2							

Appendix 3: Protein profile of differentially expressed secreted proteins in HBV replicating HepG2 cells. Run 2.

N	UnuseAccession Name	115:114	PVal	115:116:114	PVal	116:117:114	PVal	117:118:119
1	97.44 sp P02768 Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	1.0267	0.8213	1.1735	0.1505	0.8553	0.6437	
2	12.86 sp P01023 Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=1	0.9059	0.3418	0.9887	0.8965	1.0809	0.4465	
3	12.22 sp P19823 Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapiens	1.0697	0.8138	1.0998	0.6242	1.1121	0.7231	
4	12.21 sp P63261 Actin, cytoplasmic 2 OS=Homo sapiens GN=ACTG1 PE=1 SV=1	0.9698	0.4085	0.6801	0.0289	1.0587	0.3328	
5	8.29 sp P01024 Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	0.9345		1		0.9285		
6	6 sp P01834 Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	1.1428	0.2792	1.2628	0.2747	1.164	0.2038	
7	5.96 sp P02647 Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	0.8645	0.4967	0.9035	0.4304	0.8629	0.3533	
8	5.23 sp P02788 Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6	0.9013	0.2221	1.0217	0.6534	1.0722	0.3493	
9	4.66 sp Q14624 Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens	0.7954	0.1241	0.9649	0.8519	0.902	0.6563	
10	4.21 sp P02771 Alpha-fetoprotein OS=Homo sapiens GN=AFP PE=1 SV=1							
11	4 sp P69905 Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2	1.1917		0.8179		0.8929		
12	3.63 sp P02765 Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	1.1123	0.574	1.2583	0.2639	1.2363	0.0392	
13	3.08 sp P01009 Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	1.5205		1.1317		0.4298		
14	2.2 sp P01008 Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1	0.9334		0.9482		1.0047		
15	2.17 sp P02749 Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3	1.0219		1.0565		1.0281		
16	2.03 sp POC0L5 Complement C4-B OS=Homo sapiens GN=C4B PE=1 SV=1	1.0343		1.1183		1.0768		
17	2.03 sp P01857 Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=	1.1709	0.0615	1.4548	0.1241	1.185	0.1095	
18	2.01 sp P01842 Ig lambda chain C regions OS=Homo sapiens GN=IGLC1 PE=1 SV=	1		1.059		0.9701		
19	2 sp Q68CQ4 Digestive organ expansion factor homolog OS=Homo sapiens GN=	0.8813		0.9994		1		
20	2 sp P06396 Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1	1.0544		1.1788		0.6848		
21	2 sp P49747 Cartilage oligomeric matrix protein OS=Homo sapiens GN=COMP	0.8119		0.8866		0.8178		
22	2 sp P01042 Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	1.1817		0.9539		1.0232		
23	2 sp P23142 Fibulin-1 OS=Homo sapiens GN=FBLN1 PE=1 SV=4	0.991		1.0296		1.0276		
24	2 sp P01859 Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=	1.0381		1.024		0.9353		
25	2 sp 095497 Pantetheinase OS=Homo sapiens GN=VNN1 PE=1 SV=2	0.8897		0.916		0.8972		
26	1.86 sp Q6UWW9 UDP-glucuronosyltransferase 2A3 OS=Homo sapiens GN=UGT2A3 PE=2 SV=1							
27	1.84 sp P02790 Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2	1.0625		1.0835		1.1186		
28	1.74 sp P02679 Fibrinogen gamma chain OS=Homo sapiens GN=FGG PE=1 SV=3	0.9349		0.9316		0.9047		

29	1.7	sp	Q5T3I0	G patch domain-containing protein 4	OS=Homo sapiens GN=GPATCH4 PE=1 SV=2			
30	1.7	sp	P02042	Hemoglobin subunit delta	OS=Homo sapiens GN=HBD PE=1 SV=2	0.873	0.9144	0.8664
31	1.52	sp	P00734	Prothrombin	OS=Homo sapiens GN=F2 PE=1 SV=2	0.9095	0.8196	1.0314
32	1.4	sp	Q9BTM1	Histone H2A. J	OS=Homo sapiens GN=H2AFJ PE=2 SV=1	1.179	0.9292	0.8009

Appendix 3: Protein profile of differentially expressed secreted proteins in HBV replicating HepG2 cells. Run 3.

N	Unuse	Accession Name	115:114	PVal	115:116:114	PVal	116:117:114	PVal	117
1	60.74	sp P02768 Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	1.0939	0.6248	1.1144	0.5663	0.9381	0.8981	
2	14.24	sp P19823 Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapien:	1.0118	0.8344	1.0474	0.6993	1.0193	0.871	
3	11.79	sp P01023 Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=1	0.9735	0.7604	0.9655	0.7789	1.11	0.5019	
4	7.71	sp P01024 Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	0.8854		0.9374		0.9345		
5	7.53	sp P02788 Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6	1.12		1.0036		1.1889		
6	5.9	sp P02765 Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	1.144	0.0148	1.2247	0.2279	1.2336	0.2795	
7	4.03	sp P0COL4 Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	1.226		1.0297		1.1459		
8	4	sp P69905 Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2	1.1596		1.0229		0.957		
9	3.15	sp P02042 Hemoglobin subunit delta OS=Homo sapiens GN=HBD PE=1 SV=2	0.8977		0.8576		0.8763		
10	2.89	sp P63261 Actin, cytoplasmic 2 OS=Homo sapiens GN=ACTG1 PE=1 SV=1	0.9091		0.6178		0.7812		
11	2.84	sp Q8WZ42 Titin OS=Homo sapiens GN=TTN PE=1 SV=2							
12	2.77	sp P06396 Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1	0.9362	0.0234	1.157	0.1443	1.2119	0.6315	
13	2.73	sp P01008 Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1	1.0338		1.1101		0.9944		
14	2.69	sp Q14624 Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapien:	0.936		0.8683		1		
15	2.65	sp O95497 Pantetheinase OS=Homo sapiens GN=VNN1 PE=1 SV=2	1.0179		0.9301		0.8452		
16	2.08	sp P02771 Alpha-fetoprotein OS=Homo sapiens GN=AFP PE=1 SV=1							
17	2.03	sp Q86UK5 Limbin OS=Homo sapiens GN=EVC2 PE=1 SV=1							
18	2.03	sp Q5T3I0 G patch domain-containing protein 4 OS=Homo sapiens GN=GPATCH4 PE=1 SV=2							
19	2	sp P01042 Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	1.2377		0.818		1.1019		
20	2	sp P35443 Thrombospondin-4 OS=Homo sapiens GN=THBS4 PE=1 SV=2	0.8438		1		0.9206		
21	2	sp P02647 Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	0.8747		0.7852		0.8087		
22	2	sp P23142 Fibulin-1 OS=Homo sapiens GN=FBLN1 PE=1 SV=4	1.0109		0.8653		1.0167		
23	1.87	sp P02787 Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2							
24	1.7	sp P07195 L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=	0.9509		0.9897		0.9224		
25	1.52	sp Q9H8M2 Bromodomain-containing protein 9 OS=Homo sapiens GN=BRD9 PE=	1		1.0384		1.1146		
26	1.42	sp P00734 Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	0.9895		1.0175		1.0359		
27	1.4	sp Q13535 Serine/threonine-protein kinase ATR OS=Homo sapiens GN=ATR PE=1 SV=3							