Hepatitis B Virus Covalently Closed Circular DNA Interacting HBx and HBc Proteins Global Conservancy Analysis and Related B or T lymphocyte Epitopes

Umar Saeed (✉ umarsaeed15@yahoo.com)

Islamabad Diagnostic Center (IDC)

Zahra Zahid Piracha

Islamabad Diagnostic Center (IDC)

Rizwan Uppal

Islamabad Diagnostic Center (IDC)

Research Article

Keywords: Hepatitis B virus, HBx, HBc, cccDNA, Peptide Vaccines, B or T lymphocyte Epitopes

Posted Date: June 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-455486/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background:

The Hepatitis B Virus HBx and HBc proteins associate with covalently closed circular DNA, which is the main reason for intrahepatic viral persistence and major cause due to which HBV cure has not been achieved yet. The aims of present study were to generate HBV genotype-specific consensus sequences of HBx and HBc, align all ten (A to J) consensus sequences to develop global consensus sequences of HBx and HBc, analyze variable and conserved motifs, and to predict highly conserved B and T cell binding epitopes in HBx and HBc proteins, respectively.

Methods:

237 HBx and 207 HBc sequences, belonging to all HBV genotypes reported globally, were aligned in CLC main workbench to draw global consensus sequences and phylogenetic analysis was performed. The location of possible B cell epitopes were analyzed using immune Epitope Database (IEDB), while possible T cell epitopes were analyzed using ProPred-I and ProPred *in-silico* prediction tools.

Results:

The HBx residues ^{52}H to ^{59}P , which are important for augmentation effect in HBV replication and ^{137}C , which is crucial for transactivation of HBx are conserved in all HBV genotypes. The HBc residues ^{141}S to ^{149}V , which are crucial for pre-genomic RNA packaging, viral DNA synthesis and virion secretion are highly conserved across all the genotypes of HBV. The HBx related epitopes X-B2 and X-B4, and HBc related C-B6 and C-B7 epitopes could be important candidates for B-cell based vaccine. Among HBx related MHC-I epitopes X-M2, X-M5, X-M8, X-M11, X-M12, X-M20, X-M22, X-M25, X-M27 and X-M32, and MHC-II epitopes the X-T4, X-T6 and X-T8 could be important targets for vaccine development. While among HBc related MHC-I epitopes C-M1, C-M2, C-M4, C-M6-11, C-M13, C-M19, C-M24-26, C-M30, C-M34-36, C-M40 and C-M43-45, and MHC-II epitopes the C-T1-3, C-T9, C-T10, C-T12, C-T14 and C-T16-18 epitopes could be ideal epitopes with high conservancy across all HBV genotypes.

Conclusion:

The analysis will aid in screening of novel anti-HBV agents and designing of site-specific inhibitors which may show response against all genotypes. Also the predicted B- or T cell epitopes might be effective for designing antibodies against all majorly occurring genotypes of HBV across the world.

Highlights

HBx and HBc bind to HBV cccDNA, which is the main reason for intrahepatic viral persistence. HBx related epitopes X-B2 and X-B4, and HBc related C-B6 and C-B7 epitopes are crucial for B-cell based vaccine. Among HBx related MHC-I epitopes X-M2, X-M5, X-M8, X-M11, X-M12, X-M20, X-M22, X-M25, X-M27 and X-M32, and MHC-II epitopes the X-T4, X-T6 and X-T8 are crucial for vaccine development. While

HBc related MHC-I epitopes C-M1, C-M2, C-M4, C-M6-11, C-M13, C-M19, C-M24-26, C-M30, C-M34-36, C-M40 and C-M43-45, and MHC-II epitopes the C-T1-3, C-T9, C-T10, C-T12, C-T14 and C-T16-18 epitopes could be critical for peptide vaccine.

Introduction

Hepatitis B virus (HBV) is a partially double stranded DNA virus belonged to the *Hepadnaviridae* family, which has exclusive tropism for hepatocytes (1). HBV infection causes acute hepatitis, which can lead to chronic hepatitis B (CHB), liver fibrosis, fulminant hepatic failure, cirrhosis, and hepatocellular carcinoma (HCC) (1). Globally, 257 million people are living with CHB resulting in 887,000 deaths per year (2).

HBV has 10 major genotypes (A,B,C,D,E,F,G,H,I and J) with diverse geographical distribution (3). HBV genotype A is predominant in Northern Europe, Sub-Saharan and Western Africa, and India; genotype B is widespread in East Asia, China, Japan, Indonesia, Vietnam, Taiwan, Philippines, Greenland, Northern Canada and Alaska; genotype C is primarily observed in Southeast Asia, China, South Korea, Australia, Taiwan, Indonesia, Vietnam and Philippines; genotype D is common in Mediterranean countries, Africa, Europe, India, and Indonesia; genotype E is dominant in Western Africa; genotype F is primarily observed in Central and South America; genotype G is widespread in United States, Germany, and France; genotype H is dominant in Central America; genotype I is prevalent in Laos and Vietnam; and genotype J has been reported dominant in Ryukyu Islands of Japan (3).

Despite of the availability of anti-HBV vaccines, the viral hepatitis has still remained a major public health problem (2). Reverse transcriptase inhibitors such as tenofovir, lamivudine and entecavir are commercially available for treatment of HBV patients, however these have to be taken for life (4). Currently there is no cure of HBV (5). A true cure of HBV requires clearance of intra-hepatic nuclear covalently closed circular DNA (cccDNA) which is the key to virological persistence (5, 6). During HBV life cycle, the cccDNA associate with several cellular histone and non-histone proteins, including HBc and HBx, transcription factors co-activators, and several epigenetic activators and repressors which affect HBV transcription and epigenetic control (7–9).

HBx is a 154-amino acid long protein, which can stimulate viral replication to several folds and essential for initiating and maintaining HBV life cycle, host–virus interactions, and development of HCC (9, 10). In the absence of HBx, the cccDNA rapidly attain silent state (closed conformation) and becomes transcriptionally inactive (6). Short interfering RNA (siRNA) mediated HBx gene silencing inhibited HBV gene expression and replication in HBV genotype B hydrodynamic injection mouse models (11). Recently a new study has shown that cell penetrating antibody targeting HBx in chronic HBV infection mimicking mouse models or cells significantly suppressed hepatitis B virus (12).

The HBc is 183 or 185 amino acids long protein (length depending on the strains), which facilitates formation of nuclear capsid, capable of incorporating pre-genomic (pgRNA) and reverse transcriptase. It facilitate conversion of single-stranded (SS) DNA to the relaxed circular (RC) DNA and generate mature

NC, which could be subsequently enveloped through envelop proteins and released extracellularly as infectious virion (13–15). Both HBx and HBc are crucial for HBV replication. Targeting cccDNA binding viral HBc and HBx proteins could limit viral replication. Thus the aim of the present study was to design global consensus sequences of HBx and HBc proteins and to analyze variable regions and highly conserved motifs which could offer potent target site for development of peptide vaccine, designing site-specific RNA interference and anti-HBV inhibitors. Furthermore we investigated B or T cell epitopes, in HBx and HBc proteins using *in-silico* approaches, which might be capable of generating neutralizing antibodies against all A-J genotypes of HBV.

Material And Methods

Retrieval of HBx and HBc sequences and development of global consensus sequences.

A total of 237 HBx and 207 HBc sequences belonging to all major 10 genotypes of HBV were randomly retrieved from the National Center for Biotechnology Information database (<https://www.ncbi.nlm.nih.gov/nuccore/>). These sequences were reported from all over the world including USA, Canada, Brazil, Venezuela, Argentina, France, Germany, Netherlands, Mexico, Belgium, China, South Korea, Japan, Indonesia, Thailand, Viet Nam, Turkey, Pakistan, India, South Africa, Ghana, Côte d'Ivoire, Nicaragua, Mauritius and Martinique. The HBx and HBc amino acid sequences were fed into CLC Main Workbench v. 8 (Qiagen GmbH, Hilden, Germany). Using multiple sequence analysis feature of the software, we constructed consensus sequences of each genotype (A-J). The consensus sequences of all ten genotypes thus constructed were subsequently aligned to obtain global consensus sequences of HBx and HBc.

Peptide designing and phylogenetic analysis

The global consensus sequences were analyzed to find highly conserved regions and variable residues indifferent domains and motifs of HBx and HBc proteins. Short stretches of amino acids from the highly conserved regions of HBx and HBc proteins were selected from the global consensus sequence alignments; these peptides could be better targets for potential peptide-based vaccines testing and useful for designing inhibitory compounds. To draw phylogenetic trees all 237 HBx and 207 HBc sequences were aligned in CLC Main Workbench v. 8 (Qiagen GmbH, Hilden, Germany) and subjected to unweighted pair group method with arithmetic mean (UPGMA) with boot strap value of 100.

B-and T-lymphocyte epitopes prediction.

The location of B and T cell epitopes were mapped in the global consensus sequences of HBx and HBc. Target epitopes for B-lymphocytes in HBx and HBc were predicted for possible antibody binding through Kolaskar and Tongaonkar antigenicity prediction method using Immune Epitope Database (IEDB) (<https://www.iedb.org/>). Possible T-lymphocytes epitopes in HBx and HBc were analyzed for promiscuous major histocompatibility complex class-I (MHC-I) and II (MHC-II) through ProPred-I (<http://crdd.osdd.net/raghava/propred1/>) and ProPred *in-silico* prediction tools

(<http://crdd.osdd.net/raghava/propred/>), respectively. Predicted B-cell and T cell epitopes in HBx and HBc were further analyzed for epitope conservancy analysis *via* IEDB analysis resource (<http://tools.iedb.org/conservancy/>). To analyze if these peptides could trigger autoimmunity, selected epitopes with 70%-100% conservancy, were compared with human proteome *via* Mimicry Peptide Database(miPepBase) (<http://proteininformatics.org/mkumar/mipepbase/>).

Results

Global consensus sequence of HBx and HBc proteins and analysis of highly conserved regions.

HBV covalently closed circular DNA (cccDNA), associates with cellular and viral DNA binding proteins, HBc (*via* direct binding) and HBx (through indirect binding) to form viral episome that serve as template of viral transcription and protein expression in HBV-infected hepatocyte nuclei (8, 9). HBx promotes transcriptionally active state of viral episome and enhance HBV replication to several folds; while HBc performs crucial roles in structural organization of episome by altering nucleosome numbers and leads to higher episome copy numbers. Attacking cccDNA binding viral proteins (HBc and HBx) could be important antiviral approach to limit HBV replication.

We hypothesized that conserved residues of HBx and HBc might be important in developing novel anti-HBV agents, designing peptide-based vaccines and site-specific inhibitors. Therefore, to analyze highly conserved domains of HBx and HBc we constructed consensus sequence of all ten major genotypes of HBV reported globally, and aligned the genotypic consensus sequences to develop global consensus sequences of HBx and HBc, respectively.

Figure 1 shows alignment of consensus sequences of HBx from all ten HBV (A-J) genotypes; the global consensus sequence is shown at the base. Different motifs and domains of HBx were analyzed for amino acid conservancy and variability (Fig. 1A). The graphs represents percentage conservancy of HBx amino acids. Conserved residues were labeled by their symbols, however variable amino acids were denoted by symbol "X". Stretches of highly conserved amino acids could serve as peptide vaccine (16, 17). Therefore, small peptide fragments (Table 1) were deduced from highly conserved regions of HBx global consensus sequence (Fig. 1A and B). A phylogenetic tree of 237 HBx sequences belonging to A-J genotypes, reported from different countries across the world was constructed (Fig. 2). The sequences from various genotypes were clustered together on the basis of evolutionary relatedness.

Table 1
Position and sequence of highly conserved regions of HBx which might be used for peptide vaccine.

Position of Peptides	Sequence of Peptides
6–20	CCQLDPARDVLCLRP
48–59	DHGAHLSLRGLP
63–77	FSSAGPCALRFTSAR
93–101	LHKRTLGLS
132–143	FVLGGCRHKLVC
145–154	PAPCNFFTSA

Similarly, we aligned HBc consensus sequences from all ten HBV genotypes; and constructed global consensus sequence, shown at the base (Fig. 3A), and analyzed various domains and motifs for conservancy or variability of amino acids. The graphs depict percentage conservancy of HBc amino acids. Highly conserved residues are shown by their symbols, while variable regions are labeled by “X”. Short peptides were selected (Table 2) from highly conserved regions of HBc global consensus sequence (Fig. 2A and B), which might offer potent target site for development of peptide vaccine or designing site-specific anti-HBV inhibitors. The phylogenetic tree of 207 HBc sequences from 10 genotypes, from all over the world was constructed (Fig. 4), which indicate evolutionary relatedness between the sequences.

Table 2
Position and sequence of highly conserved regions of HBc which might be used for peptide vaccine.

Position of Peptides	Sequence of Peptides
3–11	IDPYKEFGA
13–26	VELLSFLPSDFFPS
29–39	DLLDTASALYR
42–66	LESPEHCSPHHTALRQAILCWGELM
99–115	QLLWFHISCLTFGRETV
119–128	LVSFGVWIRT
129–152	PPAYRPPNAPILSTLPETTVVRRR
157–176	RRRTPSPRRRRSQSPRRRRS

B-lymphocyte epitopes of HBx and HBc proteins and conservation analysis

Global consensus sequences of HBx and HBc proteins were subjected to analysis in IEDB for prediction of different B-cell epitopes which might produce neutralizing antibodies against all HBV genotypes. Each HBx related B-cell epitope was given a distinctive name, from X-B1 to X-B7 (Table 3), while HBc related B-cell epitopes were denoted as C-B1 to C-B7 (Table 4). The predicted epitopes were subjected to epitope conservation analysis *via* IEDB conservation analysis resource. The location, length and percentage conservancy of each epitope of HBx and HBc were refracted in Tables 3 and 4, respectively. Among HBx related B-cell epitopes, X-B2 (DVLCLRP) and X-B4 (AGPCALR) were considered to be conserved in all ten genotypes (Table 3). However among HBc related B-cell epitopes, C-B6 (IRQLLWFHISCLTF) and C-B7 (VLEYLVSFVG) were considered highly conserved in all HBV genotypes (Table 4). We utilized miPep Base facility to analyze whether these peptide trigger autoimmunity. And none of the peptides triggered autoimmune response. Aforementioned epitopes, predicted from global consensus sequences, might be capable of generating strong neutralizing antibodies against all (A-J) genotypes of HBV.

Table 3
HBx related B-cell epitopes and their conservancy analysis from HBV genotypes A-J.

Name	B-Cell epitopes	Peptide Length	Epitope Conservancy (%)
X-B1	RLCCQLD	7	70
X-B2*	DVLCLRP	7	100
X-B3	LRGLPVCAFS	10	70
X-B4*	AGPCALR	7	100
X-B5	ILPKVL	6	40
X-B6	LKVFVLGG	8	60
X-B7	HKLVCSPAPC	10	60

Table 4
B-cell epitopes and their conservation in HBV HBc sequences from all ten genotypes of HBV.

Name	B-Cell epitopes	Peptide Length	Epitope Conservancy (%)
C-B1	SVELLSFLPSD	11	60
C-B2	FPSVRDL	7	60
C-B3	TASALYRE	8	70
C-B4	PEHCSPHHTALRQAILCWG	19	50
C-B5	RDLVVSYY	8	30
C-B6	IRQLLWFHISCLTF	14	80
C-B7	VLEYLVSFVG	10	80

T-lymphocyte MHC-I and MHC-II specific epitopes of HBx and HBc proteins and conservation analysis.

The location of possible T-cell MHC-I epitopes was identified in global consensus sequence of HBx and HBc through ProPred-I, for 47 MHC-I alleles. In total, 323 different HBx related MHC-I epitopes were predicted, which were subsequently subjected to conservation analysis *via* IEDB epitope conservation analysis. We selected epitopes with 70–100% conservancy (Table 5). Among these epitopes, X-M2, X-M5, X-M8, X-M11, X-M12, X-M20, X-M22, X-M25, X-M27 and X-M32 were conserved in HBV HBx consensus sequence and among all ten genotypes of HBV. Similarly, 182 different HBc related MHC-I epitopes were predicted and subjected for epitope conservation analysis. Epitopes with 70–100% conservancy were selected and presented in Table 6. The C-M1, C-M2, C-M4, C-M6-11, C-M13, C-M19, C-M24-26, C-M30, C-M34-36, C-M40 and C-M43-45 epitopes remained conserved in HBc consensus sequence and all HBV genotypes.

Table 5
HBx related T-cell class I MHC-specific epitopes and their conservancy analysis from all HBV genotypes.

Name	Class I MHC-specific T-cell epitopes	Peptide Length	Epitope Conservancy (%)
X-M1	QLDPARDVL	9	80
X-M2*	SAGPCALRF	9	100
X-M3	VLHKRTLGL	9	80
X-M4	VLCLRPVGA	9	80
X-M5*	VLGGCRHKL	9	100
X-M6	CQLDPARDV	9	80
X-M7	HLSLRGLPV	9	90
X-M8*	FVLGGCRHK	9	100
X-M9	ALRFTSARR	9	70
X-M10	MAARLCCQL	9	70
X-M11*	CALRFTSAR	9	100
X-M12*	SSAGPCALR	9	100
X-M13	DHGAHLSLR	9	70
X-M14	LRFTSARRM	9	70
X-M15	DPARDVLCL	9	80
X-M16	LRGLPVCAF	9	70
X-M17	RRMETTVNA	9	70
X-M18	ARRMETTVN	9	70
X-M19	ARLCCQLDP	9	70
X-M20*	FSSAGPCAL	9	100
X-M21	LDPARDVLC	9	80
X-M22*	RDVLCLRPV	9	100
X-M23	GAHLSLRGL	9	80
X-M24	SARRMETTV	9	70
X-M25*	LGGCRHKLV	9	100
X-M26	LPVCAFSSA	9	70

Name	Class I MHC-specific T-cell epitopes	Peptide Length	Epitope Conservancy (%)
X-M27*	APCNFFTSA	9	100
X-M28	FTSARRMET	9	70
X-M29	LSLRGLPVC	9	90
X-M30	SLRGLPVCA	9	70
X-M31	CLRPVGAES	9	80
X-M32*	AGPCALRFT	9	100

Table 6

T-cell class I MHC-specific epitopes and their conservation in HBV HBc sequences from all major HBV genotypes.

Name	Class I MHC-specific T-cell epitopes	Peptide Length	Epitope Conservancy (%)
C-M1*	LLDTASALY	9	100
C-M2*	DIDPYKEFG	9	100
C-M3	LPETTVVRR	9	80
C-M4*	YLVSFQVWI	9	100
C-M5	LEYLVSFQV	9	80
C-M6*	LLWFHISCL	9	100
C-M7*	QLLWFHISC	9	100
C-M8*	DLLDTASAL	9	100
C-M9*	LVSFQVWIR	9	100
C-M10*	HISCLTFGR	9	100
C-M11*	AYRPPNAPI	9	100
C-M12	TLPETTVVR	9	90
C-M13*	WIRTPPAYR	9	100
C-M14	CSPHHTALR	9	80
C-M15	TTVRRRRGR	9	70
C-M16	LKIRQLLWF	9	80
C-M17	VRRRGRSPR	9	80
C-M18	LTFGRETVL	9	80
C-M19*	YRPPNAPIL	9	100
C-M20	MGLKIRQLL	9	80
C-M21	IRQLLWFHI	9	80
C-M22	GRETVLEYL	9	80
C-M23	RRRGRSPRR	9	80
C-M24*	RRRTPSPRR	9	100
C-M25*	RRRRSQSPR	9	100
C-M26*	RRRSQSPRR	9	100

Name	Class I MHC-specific T-cell epitopes	Peptide Length	Epitope Conservancy (%)
C-M27	FGRETVLEY	9	80
C-M28	RETVLEYLV	9	80
C-M29	EHCSPHHTA	9	70
C-M30*	VELLSFLPS	9	100
C-M31	MDIDPYKEF	9	90
C-M32	QAILCWGEL	9	80
C-M33	STLPETTVV	9	90
C-M34*	WFHISCLTF	9	100
C-M35*	LSFLPSDFF	9	100
C-M36*	LLSFLPSDF	9	100
C-M37	ALRQAILCW	9	80
C-M38	GLKIRQLLW	9	80
C-M39	HCSPHHTAL	9	80
C-M40*	SPRRRRRSQS	9	100
C-M41	HTALRQAIL	9	80
C-M42	NMGLKIRQL	9	80
C-M43*	IDPYKEFGA	9	100
C-M44*	FLPSDFFPS	9	100
C-M45*	TPPAYRPPN	9	100

To identify the possible location of T-cell MHC-II epitopes in global consensus sequence of HBx and HBc we utilized Propred *in-silico* analysis facility, for 51 HLA-DR alleles. In total 204 HBx related MHC-II epitopes were predicted, which were subjected for IEDB epitope conservancy analysis. We selected epitopes having 70–100% conservancy, as shown in Table 7. The X-T4, X-T6 and X-T8 were conserved in HBx consensus sequence and all HBV genotypes. Similarly, 203 different HBc related MHC-II epitopes were predicted and analyzed for epitope conservancy as described above. Epitopes with 70–100% conservancy are presented in Table 8. Among these, the C-T1-3, C-T9, C-T10, C-T12, C-T14 and C-T16-18 epitopes were identified to be 100% conserved in HBc consensus sequence and among all (A-J) genotypes of HBV. Using miPepBase, we found that none of the peptides trigger autoimmune response.

Table 7

T-cell class II MHC-specific epitopes and their conservation in HBV HBx protein sequences from genotypes A-J.

Name	Class II MHC-specific T-cell Epitopes	Peptide Length	Epitope Conservancy (%)
X-T1	VLCLRPVGA	9	80
X-T2	LRFTSARRM	9	70
X-T3	LRGLPVCAF	9	70
X-T4*	FVLGGCRHK	9	100
X-T5	FTSARRMET	9	70
X-T6*	VLGGCRHKL	9	100
X-T7	VLHKRTLGL	9	80
X-T8*	FSSAGPCAL	9	100
X-T9	LCLRPVGAE	9	80
X-T10	LHKRTLGLS	9	80

Table 8
HBc related T-cell class II MHC-specific epitopes and their conservancy analysis in all ten genotypes of HBV.

Name	Class II MHC-specific T-cell epitopes	Peptide Length	Epitope Conservancy (%)
C-T1*	YLVSFGVWI	9	100
C-T2*	WFHISCLTF	9	100
C-T3*	YRPPNAPIL	9	100
C-T4	LEYLVSFGV	9	80
C-T5	IRQLLWFHI	9	80
C-T6	VRRRGRSPR	9	80
C-T7	LKIRQLLWF	9	80
C-T8	LRQAILCWG	9	80
C-T9*	FHISCLTFG	9	100
C-T10*	FLPSDFFPS	9	100
C-T11	FGRETVLEY	9	80
C-T12*	LVSFGVWIR	9	100
C-T13	VVRRRGRSP	9	80
C-T14*	WIRTPPAYR	9	100
C-T15	MGLKIRQLL	9	80
C-T16*	VWIRTPPAY	9	100
C-T17*	VELLSFLPS	9	100
C-T18*	LLWFHISCL	9	100

Discussion

The HBx protein contains regulatory (or negative regulatory) and transactivation (orcoactivation) domains (Fig. 1B). The regulatory domain (1 to 50 aa) is dispensable for HBx activity and represses HBx transactivation activity (18). Consensus sequence analysis of regulatory domain shows that region ¹M to ²⁰P is highly conserved, while the region ³¹S to ⁴⁰P is variable, except ³²G and ³⁵G residues which are highly conserved among all HBV genotypes.

The transactivation domain (51 to 154 aa) is essential for augmentation effects on HBV transcription and replication (19). The region ⁵²H to ⁶⁵S amino acids is critical for augmentation effect in HBV replication (19). Consensus analysis shows that this *HLSLRGLPVCAFSS* motif is highly conserved among

all HBV genotypes. Deletion of ¹⁴¹L to ¹⁵⁴A (last 14) amino acids of transactivation domain did not affect transactivation property (20). It has been demonstrated that ¹³²F to ¹⁴⁰K and ¹³⁷C residues are crucial for transactivation of HBx (20). The consensus sequence analysis shows that this *FVLGGCRHK* motif and ¹³⁷C residue are completely conserved among all HBV genotypes. It could be inferred that designing anti-HBV siRNA or inhibitor of these region might target HBx from all HBV genotypes. However a natural mutant of HBx (HBxDelta127) without *FVLGGCRHK* motif and ¹³⁷C, has been reported from chronic hepatitis B, liver cirrhosis and HCC patients, which can induce growth and proliferation of hepatoma cells (21, 22). The HBx protein, due to presence of BH3-like motif (¹¹⁰A to ¹³⁵G) in the C-terminal region, directly associates with anti-apoptotic Bcl-2 family proteins and induce elevated cytosolic calcium levels and promote viral DNA replication (23–25). Consensus sequence analysis showed ¹¹¹Y, ¹¹³K to ¹¹⁵C, ¹¹⁷F, ¹²⁰W to ¹²²E, and ¹³²F to ¹³⁵G residues were completely conserved in BH3-like motif of HBx in all HBV genotypes. Among 13 XaaP motifs in HBx protein, ¹⁰D¹¹P, ¹⁹R²⁰P, ²⁸R²⁹P, ⁴⁵V⁴⁶P in regulatory domain and ⁵⁸L⁵⁹P, ⁶⁷G⁶⁸P, ⁸⁹L⁹⁰P motifs in transactivation domain were completely conserved in all HBV genotypes.

The HBc protein is composed of N-terminal (1 to 140 aa), linker (141 to 149 aa), and C-terminal (150 to 183 or 185 aa depending on the strains) domains (26)(Fig. 3B). The N-terminal domain (NTD) is critical and sufficient for capsid assembly (27, 28). Consensus sequence analysis of CTD shows that region ¹M to ¹¹A of all HBV genotypes is completely conserved, except genotype G that contains additional 12 amino acids *RTTLPYGLFGLD* insertion. This insertion has pleiotropic effects on core protein expression, HBV replication, and virion secretion (29). Insertion of *RTTLPYGLFGLD* motif enhanced core protein levels independent of viral genotype, augments replication in genotype G, while impairs replication in genotype A and D (Gutelius et al., 2011). The region ¹³V to ³⁹R (or ²⁵V to ⁵¹R of genotype G) is highly conserved among all HBV genotypes. The NTD carries protease-like sequence ³⁰L to ³⁵S (or ⁴²L to ⁴⁷S of genotype G) which resembles to retroviral proteases (30). Consensus sequence analysis shows that this *LLDTAS* motif is highly conserved among all HBV genotypes. The region between 74 to 101 amino acids is considered as hypervariable which might lead to development of liver injury (31, 32). Consensus sequence analysis also shows that ⁷⁴X and ¹⁷⁹X residues are most variable among all HBV genotypes.

The linker domain *STLPETTVV* can interfere with NTD, pgRNA packaging in sequence-independent manner, viral DNA synthesis in sequence independent manner (during first step of reverse transcription to initiate single strand DNA) and in sequence dependent manner (during second step of reverse transcription that is extensive plus strand DNA synthesis to generate relaxed circular DNA), and virion secretion in sequence dependent manner (26). Presence of only five amino acids *ETTVV* in linker region were sufficient to generate single stranded DNA synthesis (26). The consensus sequence analysis indicates that linker *STLPETTVV* region is completely conserved among all HBV genotypes, except genotype E which contains *STLPENTVV*. The four cysteine residues at position 48, 61, 107 and 183 are not essential for core particle formation, however these residues can further stabilize HBV core particles or HBc dimers (33). The consensus sequence analysis shows that all of these cysteine residues are 100%

conserved among all HBV genotypes. The ¹³²Y, ¹²⁷R, ¹²⁹P, and ¹³⁹I are critical for HBc dimer formation (34). Consensus sequence analysis indicate that these residues are completely conserved among all HBV genotypes. The amino acid regions between ⁹⁸R to ¹¹⁵V and ¹¹⁷E to ¹⁴⁵E are 100% conserved among all HBV genotypes, indicating potential of this region for the HBV life cycle and/or viral pathogenesis.

The C-terminal domain (CTD) contains highly basic residues (arginine rich, protamine-like) that resemble to histone tails, which are critical for non-specific nucleic acid binding (35, 36). The CTD is dispensable for capsid assembly and functionally plays important role in pgRNA packaging and reverse transcription (37–39). The CTD phosphorylation is important for specific viral RNA packaging (40–42). The three major phosphorylation sites in CTD are ¹⁵⁵S, ¹⁶²S, ¹⁷⁰S, while four minor sites are ¹⁶⁰T, ¹⁶⁸S, ¹⁷⁶S and ¹⁷⁸S (42–45). The consensus sequence analysis shows that except genotype H (which contains ¹⁵⁵A instead of ¹⁵⁵S), all of the aforementioned phosphorylation sites in CTD are 100% conserved among all HBV genotypes. The arginine-rich CTD domain contains distinct nuclear localization and cytoplasmic retention signals (46). Major RNA-recognizing activity of CTD is attributed to ¹⁵⁰R to ¹⁵⁷R sequence (47). The *RRRR* following ¹⁴⁹V also provide RNA binding (35). Consensus sequence analysis shows that this *RRRGRSPR* motif is highly conserved among all HBV genotypes. In HBV genotype A between amino acids ¹⁵²R and ¹⁵³G, there exists two additional amino acids *DR*, therefore generating *RRRDRGRSPR* motif. The DNA-recognizing activity of CTD is attributed to three repeated *SPRRR* motifs within 157 to 177 amino acid sequence. Consensus sequence analysis shows that these motifs are completely conserved among all HBV genotypes, except genotype H which contains ¹⁵⁵A. Among 16XaaP motifs in HBc protein, ⁴D⁵P, ¹⁹L²⁰P, ²⁴F²⁵P, ⁷⁸D⁷⁹P, ¹²⁸T¹²⁹P, ¹²⁹P¹³⁰P, ¹³³R¹³⁴P, ¹³⁴P¹³⁵P, and ¹³⁷A¹³⁸P in NTD, ¹⁴³L¹⁴⁴P in linker domain, and ¹⁶⁰T¹⁶¹P, ¹⁶²S¹⁶³P, and ¹⁷⁰S¹⁷¹P motifs in CTD were 100% conserved among all HBV genotypes.

Recent findings have implicated active roles of HBc and HBx in epigenetic regulation of viral-host interplay (9, 48). The multiplicity of HBx and HBc functions and their capacity to influence cccDNA mini-chromosome for enhanced viral replication, elevates these proteins as excellent targets for antiviral therapeutics. (9, 48). The HBx and HBc consensus sequences were used to predict highly conserved B and T cell binding epitopes. Several highly or semi-conserved B cell binding epitopes were predicted, we selected highly conserved epitopes with 70–100% conservation among all ten HBV genotypes. Similarly, several MHC-I or -II related epitopes exhibited maximum allele-binding affinity indicating as possible T-cell related epitopes. Among HBx related B-cell binding epitopes, due to complete (100%) conservancy among all HBV genotypes, the X-B2 and X-B4 epitopes might be consider as better targets for B-cell based vaccine development. Similarly among HBc related B-cell binding epitopes, due to high (80%) conservancy among all HBV genotypes, the C-B6 and C-B7 epitopes might be better targets for B-cell based vaccine development. On the other hand, among HBx related MHC-I specific epitopes the X-M2, X-M5, X-M8, X-M11, X-M12, X-M20, X-M22, X-M25, X-M27 and X-M32; while among MHC-II specific epitopes, the X-T4, X-T6 and X-T8 could be adopted for synthetic vaccine against multi-genotypes of HBV. Similarly for HBc related MHC-I specific epitopes the C-M1, C-M2, C-M4, C-M6-11, C-M13, C-M19, C-M24-26, C-M30, C-M34-36, C-M40 and C-M43-45; while among MHC-II specific epitopes, the C-T1-3, C-T9, C-T10, C-T12, C-T14 and C-T16-18 epitopes could be ideal epitopes with high conservancy across all HBV genotypes. The use of

conserved epitopes predicted against NS3-4A from global consensus sequences could provide broader protection against multi-isotypes of hepatitis C virus (49). Our study suggests conserved epitopes against HBx and HBc global consensus sequences that may be invoked as potential targets for development of effective vaccine candidates and conserved residues could also be attributed for designing novel site specific anti-HBV agents which can target all major genotypes of HBV. Though present study indicates B-cell or T-cell related antigens on the basis of *in-silico* analysis, the antigenic potential of aforementioned peptides should be further characterized in HBV infection animal models.

Conclusion

HBx and HBc bind to HBV cccDNA, which is the main reason for intrahepatic viral persistence. HBx related epitopes X-B2 and X-B4, and HBc related C-B6 and C-B7 epitopes are crucial for B-cell based vaccine. Among HBx related MHC-I epitopes X-M2, X-M5, X-M8, X-M11, X-M12, X-M20, X-M22, X-M25, X-M27 and X-M32, and MHC-II epitopes the X-T4, X-T6 and X-T8 are crucial for vaccine development. While HBc related MHC-I epitopes C-M1, C-M2, C-M4, C-M6-11, C-M13, C-M19, C-M24-26, C-M30, C-M34-36, C-M40 and C-M43-45, and MHC-II epitopes the C-T1-3, C-T9, C-T10, C-T12, C-T14 and C-T16-18 epitopes could be critical for peptide vaccine.

Abbreviations

HBV: Hepatitis B Virus

HBc: Hepatitis B Virus Core protein

HBx: Hepatitis B Virus X protein

cccDNA: Covalently Closed Circular DNA

siRNA: Short interfering RNA

CHB: Chronic Hepatitis B,

HCC: Hepatocellular Carcinoma

Declarations

Ethics approval:

The study has been approved by ethical review board of Islamabad Diagnostic Center Pakistan. Humans/animals were not directly involved in the study.

Consent to publication:

All authors approved the submission of the manuscript for publication

Availability of data and material:

The data is available and can be used for the academic or research purposes.

Competing interests:

The authors have no conflict of interest.

Funding:

No funding was allotted for this study

Authors Contribution:

US conceived the study, wrote manuscript, and analyzed the data and principal investigator of the study; ZZP and RU assisted in manuscript writing and data analysis.

Acknowledgements:

We acknowledge the kind efforts of IDC colleagues.

References

1. Seeger C, Mason WS. Molecular biology of hepatitis B virus infection. *Virology* 2015;479–480:672–686.
2. World Health Organization global hepatitis report 2017. Available at <http://www.who.int/hepatitis/publications/global-hepatitis-report2017/en/>. Accessed August 20, 2017.
3. Lin C-L, Kao J-H. Hepatitis B Virus Genotypes and Variants. *Cold Spring Harbor Perspectives in Medicine* 2015;5:a021436.
4. Zoulim F. Emerging drugs for hepatitis B. *Expert Opinion on Emerging Drugs* 2007;12:199–217.
5. Cornberg M, Manns MP. Hepatitis: No cure for hepatitis B and D without targeting integrated viral DNA? *Nat Rev Gastroenterol Hepatol* 2018;15:195–196.
6. Nassal M. HBV cccDNA: viral persistence reservoir and key obstacle for a cure of chronic hepatitis B. *Gut* 2015;64:1972–1984.
7. Lucifora J, Protzer U. Attacking hepatitis B virus cccDNA-The holy grail to hepatitis B cure. *J Hepatol* 2016;64:S41-S48.
8. Belloni L, Pollicino T, De Nicola F, Guerrieri F, Raffa G, Fanciulli M, et al. Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. *Proc Natl Acad Sci U S A* 2009;106:19975–19979.

9. Hensel KO, Rendon JC, Navas MC, Rots MG, Postberg J. Virus-host interplay in hepatitis B virus infection and epigenetic treatment strategies. *FEBS J* 2017;284:3550–3572.
10. Neuveut C, Wei Y, Buendia MA. Mechanisms of HBV-related hepatocarcinogenesis. *J Hepatol* 2010;52:594–604.
11. Li L, Shen H, Li A, Zhang Z, Wang B, Wang J, et al. Inhibition of hepatitis B virus (HBV) gene expression and replication by HBx gene silencing in a hydrodynamic injection mouse model with a new clone of HBV genotype B. *Viol J* 2013;10:214.
12. Zhang J-F, Xiong H-L, Cao J-L, Wang SJ, Guo XR, Lin BY et al. A cell-penetrating whole molecule antibody targeting intracellular HBx suppresses hepatitis B virus via TRIM21-dependent pathway. *Theranostics* 2018;8:549–562.
13. Hu J, Seeger C. Hepadnavirus Genome Replication and Persistence. *Cold Spring Harb Perspect Med* 2015;5:a021386.
14. Zlotnick A, Venkatakrishnan B, Tan Z, Lewellyn E, Turner W, Francis S. Core protein: A pleiotropic keystone in the HBV lifecycle. *Antiviral Res* 2015;121:82–93.
15. Venkatakrishnan B, Zlotnick A. The Structural Biology of Hepatitis B Virus: Form and Function. *Annu Rev Virol* 2016;3:429–451.
16. Khan H, Murad S. Analysis of different HCV NS4B domains for the development of global consensus sequence. *Acta Virol.* 2015;59:276–83.
17. Waheed Y, Saeed U, Anjum S, Afzal MS, Ashraf M. Development of Global Consensus Sequence and Analysis of Highly Conserved Domains of the HCV NS5B Protein. *Hepat Mon.* 2012;12:e6142.
18. Murakami S, Cheong JH, Kaneko S. Human hepatitis virus X gene encodes a regulatory domain that represses transactivation of X protein. *J Biol Chem* 1994;269:15118–15123.
19. Tang H, Delgermaa L, Huang F, Oishi N, Liu L, He F, et al. The transcriptional transactivation function of HBx protein is important for its augmentation role in hepatitis B virus replication. *J Virol* 2005;79:5548–56.
20. Kumar V, Jayasuryan N, Kumar R. A truncated mutant (residues 58–140) of the hepatitis B virus X protein retains transactivation function. *Proc Natl Acad Sci U S A* 1996;93:5647–52.
21. Zhang H, Shan CL, Li N, Zhang X, Zhang XZ, Xu FQ, et al. Identification of a natural mutant of HBV X protein truncated 27 amino acids at the COOH terminal and its effect on liver cell proliferation. *Acta Pharmacol Sin* 2008;29:473–80.
22. Wang Q, Zhang WY, Ye LH, Zhang XD. A mutant of HBx (HBxDelta127) promotes hepatoma cell growth via sterol regulatory element binding protein 1c involving 5-lipoxygenase. *Acta Pharmacol Sin* 2010;31:367–74.
23. Jiang T, Liu M, Wu J, Shi Y. Structural and biochemical analysis of Bcl-2 interaction with the hepatitis B virus protein HBx. *Proc Natl Acad Sci U S A* 2016;113:2074–9.
24. Geng X, Harry BL, Zhou Q, Skeen-Gaar RR, Ge X, Lee ES, et al. Hepatitis B virus X protein targets the Bcl-2 protein CED-9 to induce intracellular Ca²⁺ increase and cell death in *Caenorhabditis elegans*.

Proc Natl AcadSci USA 2012;109:18465–18470.

25. Geng X, Huang C, Qin Y, McCombs JE, Yuan Q, Harry BL, et al. Hepatitis B virus X protein targets Bcl-2 proteins to increase intracellular calcium, required for virus replication and cell death induction. *Proc Natl AcadSci USA* 2012;109:18471–18476.
26. Kuancheng Liu, Laurie Luckenbaugh, Xiaojun Ning, Ji Xi, Jianming Hu. Multiple roles of core protein linker in hepatitis B virus replication. *PLoSPathog* 2018;14:e1007085.
27. Gallina A, Bonelli F, Zentilin L, Rindi G, Muttini M, Milanesi G. A recombinant hepatitis B core antigen polypeptide with the protamine-like domain deleted self assembles into capsid particles but fails to bind nucleic acids. *J Virol* 1989;63:4645–4652.
28. Birnbaum F, Nassal M. Hepatitis B virus nucleocapsid assembly: Primary structure requirements in the core protein. *J Virol* 1990;64:3319–3330.
29. Gutelius D, Li J, Wands J, Tong S. Characterization of the Pleiotropic Effects of the Genotype G-Specific 36-Nucleotide Insertion in the Context of Other Hepatitis B Virus Genotypes. *Journal of Virology* 2011;85:13278–13289.
30. Miller RH. Proteolytic self-cleavage of hepatitis B virus core protein may generate serum e antigen. *Science* 1987;236:722–5.
31. Ehata T, Omata M, Yokosuka O, Hosoda K, Ohto M. Variations in codons 84–101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. *Journal of Clinical Investigation* 1992;89:332–338.
32. Günther S, Baginski S, Kissel H, Reinke P, Krüger D H, Will H, et al. Accumulation and persistence of hepatitis B virus core gene deletion mutants in renal transplant patients are associated with end-stage liver disease. *Hepatology* 1996;24:751–8.
33. Zhou S, Standring DN. Cys residues of the hepatitis B virus capsid protein are not essential for the assembly of viral core particles but can influence their stability. *Journal of Virology* 1992;66:5393–5398.
34. Wynne SA, Crowther RA, Leslie AG. The crystal structure of the human hepatitis B virus capsid. *Mol Cell* 1999;3:771–80.
35. Hatton T, Zhou S, Standring DN. RNA- and DNA-binding activities in hepatitis B virus capsid protein: a model for their roles in viral replication. *J Virol* 1992;66:5232–41.
36. Lewellyn EB, Loeb DD. The arginine clusters of the carboxy-terminal domain of the core protein of hepatitis B virus make pleiotropic contributions to genome replication. *J Virol* 2011;85:1298–1309.
37. Yu M, Summers J. A domain of the hepadnavirus capsid protein is specifically required for DNA maturation and virus assembly. *J Virol* 1991;65:2511–2517.
38. Nassal M. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *J Virol* 1992;66:4107–4116.

39. Kock J, Nassal M, Deres K, Blum HE, von Weizsacker F. Hepatitis B virus nucleocapsids formed by carboxy-terminally mutated core proteins contain spliced viral genomes but lack full-size DNA. *J Virol* 2004;78:13812–13818.
40. Gazina EV, Fielding JE, Lin B, Anderson DA. Core protein phosphorylation modulates pregenomic RNA encapsidation to different extents in human and duck hepatitis B viruses. *J Virol* 2000;74:4721–4728.
41. Lan YT, Li J, Liao W, Ou J. Roles of the three major phosphorylation sites of hepatitis B virus core protein in viral replication. *Virology* 1999;259:342–348.
42. Jung J, Hwang SG, Chwae YJ, Park S, Shin HJ, Kim K. Phosphoacceptor threonine 162 and serines 170 and 178 within the carboxyl-terminal RRRS/T motif of the hepatitis B virus core protein make multiple contributions to hepatitis B virus replication. *J Virol* 2014;88:8754–8767.
43. Liao W, Ou JH. Phosphorylation and nuclear localization of the hepatitis B virus core protein: significance of serine in the three repeated SPRRR motifs. *J Virol* 1995;69:1025–1029.
44. Ludgate L, Adams C, Hu J. Phosphorylation state-dependent interactions of hepadnavirus core protein with host factors. *PLoS One* 2011;6:e29566.
45. Ludgate L, Ning X, Nguyen DH, Adams C, Mentzer L, Hu J. Cyclin-dependent kinase 2 phosphorylates S/T-P sites in the hepadnavirus core protein C-terminal domain and is incorporated into viral capsids. *J Virol* 2012;86:12237–12250.
46. Li HC, Huang EY, Su PY, Wu SY, Yang CC, Lin YS, et al. Nuclear export and import of human hepatitis B virus capsid protein and particles. *PLoS Pathog* 2010;6:e1001162.
47. Pumpens P, Grens E. Hepatitis B core particles as a universal display model: a structure-function basis for development. *FEBS Lett* 1999;442:1–6.
48. Diab A, Foca A, Zoulim F, Durantel D, Andrisani O. The diverse functions of the hepatitis B core/capsid protein (HBc) in the viral life cycle: Implications for the development of HBc-targeting antivirals. *Antiviral Res* 2018;149:211–220.
49. Naeem A, Waheed Y. Sequence analysis of hepatitis C virus nonstructural protein 3-4A serine protease and prediction of conserved B and T cell epitopes. *Biomed Rep* 2017;7:563–566.

Figures

Figure 1

Sequence alignment of HBV genotype specific consensus sequences of the HBx protein and global consensus sequence is shown.

Figure 2

Phylogenetic tree of 237 HBV HBx sequences from all ten genotypes reported across the world

Figure 3

Sequence analysis of the genotypes A-J of the HBc protein and global consensus sequence is shown.