Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2015, 7(5):1136-1146

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Molecular homology modeling & docking studies to predict the 3D structure and drug determination for HBV core of Hepatitis B

Faryal Khattak*, Muhammad Haseeb and Sahar Fazal

Department of Bioinformatics, Mohammad Ali Jinnah University, Islamabad, Pakistan

---

ABSTRACT

Work on the HBV core protein of Hepatitis B has been performed in order to have deep considerate of the structure of this virulent protein. To understand the underlying mechanism and to see the effect of different drugs on it by using different bioinformatics tools. The proposed work has been divided into two stages, building model of the protein structure through the use of bioinformatics tools and then the designing and docking of the proposed drugs. All steps of homology modeling and refinement were carried out by the program MODELLER (Version 9 (9v8)). The law of PROCHECK is to estimate the overall stereo-chemical value of a given model and Energy of protein folds was finding out by using the tool ProSA (Protein Structure Analysis, Version.4). Docking server and DS viewer were used to study the inter actions of the ligands with the protein. Molecular docking server Interaction of the ligand with the protein in terms hydrophobic interactions, hydrophilic interactions and other interactions. Hex dock server helped in binding the ligand with the protein. The 17 residues are conserved in pair wise alignment between target and template. All the sequences shows highest similarity when these sequences were loaded into CUSTAL X. Tertiary structure of HBV core consist of two main domains A+B structure topology. The Alanine shows same interaction between hex and molecular docking server with Tenofovir, Telbivudine, Lamivudine drugs. The tryptophan and Alanine shows same interaction with Tenofovir, Telbivudine, Lamivudine drugs in molecular docking.

Keywords: Homology modeling, Hepatitis B, HBV, Docking and Drug determination

---

INTRODUCTION

Hepatitis B virus (HBV) infection is one of the most important health problems in the world, about ~350 million peoples affected by Hepatitis B infection globally and is a leading cause of end-stage liver disease, hepatocellular carcinoma, and mortality [1,2]. The Hepatitis B is belongs to the hepadnaviruses family partially double stranded DNA its diameter is 42 nm and it is composed of 27 nm nucleocapsid core, enclosed by an external lipo protein coat having the surface antigen. The diameter of virons is 42 nm and has an isometric nucleocapsid or core 27 nm in diameter, and surrounded by an external coat about 4 nm thick. The word surface antigen is use for the protein of viron coat [3]. The size of HBV viron circular, partially duplex DNA molecule is 3.2kb; the 5'cohessive ends uphold its circularity. The infectious particle is known as the Dane particle and its shape is spherical. During the process of budding virus required the viral membrane and endoplasmic reticulum transported the viral particles through secretory pathway surface proteins having three viral surface proteins are form through the Golgi pathway. The size of these three proteins is dissimilar so it can easily differentiate and know by their size as small (HBsAg), middle (HBmAg), or large (HBlAg) and these proteins are required during budding process in ER [4, 5, 10]. The core protein are encoded by the viral genome, the word e-antigen is also used for the pre- core protein. Immune system recognized the core protein, and it is necessary for the development of nulceocapisd. Most-conserved polypeptide

among the mammalian hepadnaviruses is HBcAg that have 68% between HBV and WHV. Core proteins spontaneously assemble into forms resembling core particles [7]. HBV infected constricted untimely in the life cycle direct to chronic hepatitis, subsequently to cirrhosis, and finally to HCC, typically once a time of 30 to 50 years. Those males whose are infected with HBV are more likely to continue steadily infected than women, who are rapidly infected and to develop anti-HBs. Due to direct viral mechanism it is feasible that man is not carcinogenic [6].

EXPERIMENTAL SECTION

Sequences Extraction
Through the way of present studies primary sequences of HBV core (swissprot AC: Q778I9) was retrieved from the SWISSPROT (http://www.us.expasy.org) and PIR (http://www.georgetown.edu) database [11, 12].

Searching Template
Template searching is carried out through BLAST (Psi-BLAST) algorithm against Protein Data Bank (PDB), the sequence which shows high homology to the target sequence has chosen as template [13].

Multiple Sequence Alignment
Multiple sequence alignment was carried out using the program CLUSTAL X (Version: 1.81) to identify the homologous and functionally important regions, the Sequences which are homologous to the target Sequences were retrieved from SWISS-PROT [14].

Phylogenetic Analysis
Phylogenetic analysis is used to establish the evolutionary relationships among organisms. The results of an analysis can be obtained in form of Draw gram and Draw tree [15].

Secondary Structure Prediction
Sequences were submitted to the Consensus Secondary Structure Prediction Server athttp://pbil.ibcp.fr/NPSA/npsa_npsa.html in order to predict the secondary structures of the target sequences [16].

PDB sum
PDB sum is a database that provides an overview of the contents of each 3D structure deposited in the Protein Data Bank (PDB). It shows the molecules that make up the structure (i-e protein chains, DNA, ligands and metal ions) and schematic diagrams of their interactions. Extensive use is made of the freely available RasMol molecular graphics program to view the molecules and their interactions in 3D.

Model Building and Refinement
Three dimensional comparative model of HBV core was constructed using the crystal coordinates of templates on the basis of alignment between target and template sequences. All steps of homology modeling and refinement were carried out by the program MODELLER (Version 9 (9v8)) [17].

Model Visualization and Evaluation
In order to check out the consistency of the alignment and modeling of variable surface loops, structural investigations on the graphics screen using 3D visualization programs, Ds-Viewer was performed [18]. Effectiveness of the predicted model was carried out by the program PROCHECK [19] (Version: 3.4), the Energy Command of the MODELLER is use to check out the geometry, chemistry, and energy distributions of the model [17]. The ProSA (Protein Structure Analysis) web server is used to determine the energy graphs structural design of protein folds to verify the protein structure quality and the statistics of non-bonded interactions between different atom types through ERRAT [20,21].

Proteins Ligand Interactions
We have used Program Ligand Explorer (http://www.kukool.com/ligand) to study the Protein–Ligand interactions.

Molecular Docking Server
It tells about the interaction of the ligand with the protein in terms hydrophobic interactions, hydrophilic interactions and other interactions. It also tells about the electrostatic energy, intermolecular energy, and frequency of the

interacting ligand [23].

Hex Dock Server

This server also helped in binding the ligand with the protein, it takes both the protein and the ligand to be docked, in the .pdb format. It gives the top 100 good results, and a separate file of the best result. The file contains the interactions in terms of hydrophobic, hydrophilic, metallic, hydrogen bond and bridged Hydrogen bond [22].

RESULTS AND DISCUSSION

Sequence homology searches for the query HBV core was conceded by using BLAST algorithm carried out protein data bank. Crystal structure coordinate of 1YDO were selected as a template on the basis of maximum sequences similarity score and lowest E value for constructing the 3D structure of HBV core. Sequence similarity searches of target (SwissProt AC: Q778I9) showed 82% identity to the template 1YDO. MODELLER 8v2 align2d command was used to carry out the alignment between two sequences by inserting the gaps at the beginning and at the end of target sequences .The starting 33 residue of target sequence is absent. This starts from the MSE1 to trp33, which shows in Figure 1.

```
1YDO      MPYPKKVTIKEVGPRDGLQNEPVWIATEDKITWINQLSRTGLSYIEITSFVHPKWIPALRDAIDVAKG
HBV1      --------------------------------MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASA
                                                *  **      *  **  * *
1YDO      IDREKGVTYAALVPNQRGLENALEGGINEACVFMSASETHNRKNINKSTSESLHILKQVNNDAQKANL
HBV1      LYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFH
            **          *    * *
1YDO      TTRAYLSTVFGCPYEKDVPIEQVIRLSEALFEFGISELSLGDTIGAANPAQVETVLEALLARFPANQI
HBV1      ISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPR
                      *                 *       *                 *
1YDO      ALHFHDTRGTALANMVTALQMGITVFDGSAGGLGGCPYAPGSSGNAATEDIVYMLEQMDIKTNVKLEK
HBV1      RRRSQSRESQC-----------------------------------------------------------

1YDO      LLSAAKWIEEKMGKPLPSRNLQVFKSS
HBV1      ---------------------------
```

Figure 1: The sequence homology search for the query HBV core was carried out by using BLAST algorithm against PDB (Protein Data Bank). For constructing the 3D structure prediction 1YDO was selected as a template on the basis of 3D similarity from RCSB as shown in Figure 1

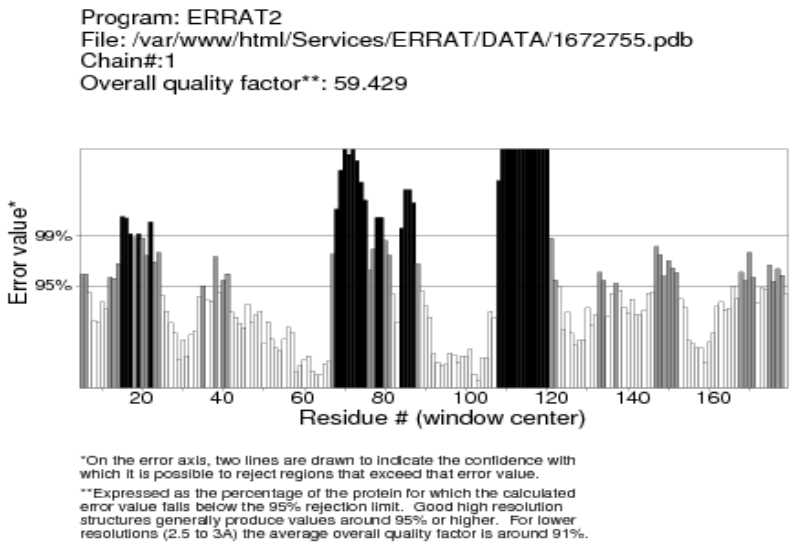

Figure 2: ERRAT result of target sequence showing the overall quality of our model comparing it to the amount of errors allowed in a model

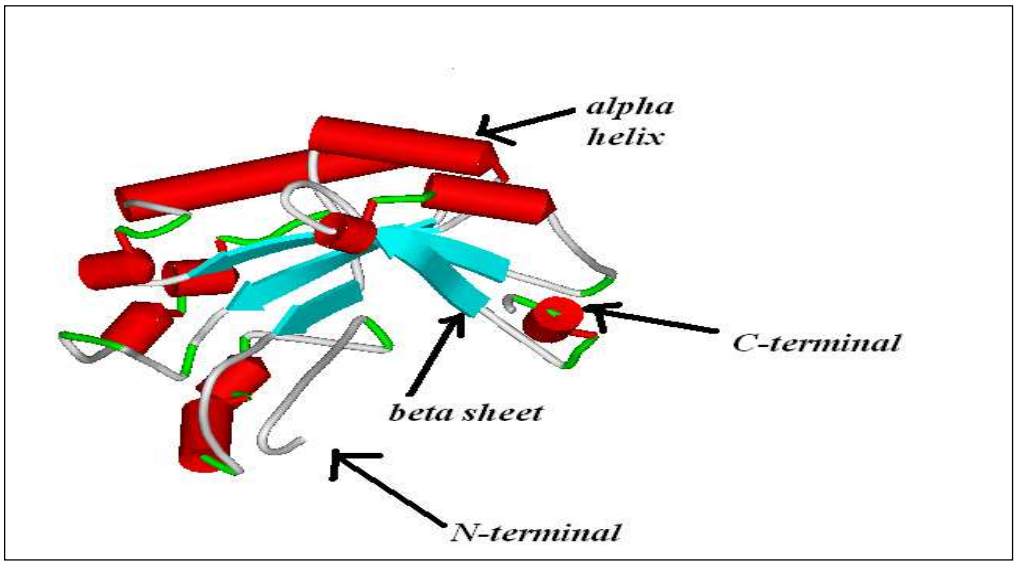

Figure 3: The schematic diagram of the HBV core. One can see the secondary structure like α-helices on the C-terminal, and β-sheets on the N-terminal

As a result of pairwise sequence alignment the different sequences obtained are then run through CLUSTALX, which aligns the evolutionary conserved residues. In order to get best alignment, multiple sequence alignment is used. Most of the residues are conserved throughout the family. Target (HBV) and Template (1YDO) belongs to these clusters which shows that they belong to same ancestor. The phylogenetic tree for HBV as shown in Figure 7. Each branch in the dendrogram represents a point of deviation.

The atomic coordinates of the crystallographic structure 1YDO solved to the resolution of the 2.5 A^o were used as starting mode of the HBV111 HBV core variant structure, Ramachandran plot shows 85.4% core, 13.3% allow, 0.6% generously allowed regions , bad contacts 3, Dihedrals -0.12 , covalent -0.04, overall -0.21. This strongly indicates that the molecular models present good overall stereochemical quality shown in Figure (6a, 6b).

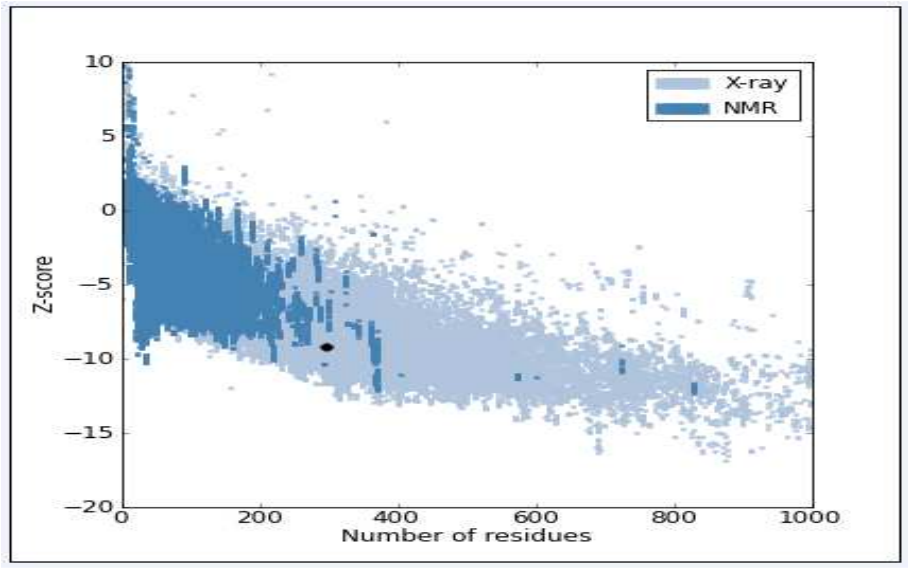

Figure 4a: ProSA plot of HBV core showing the Z value <0

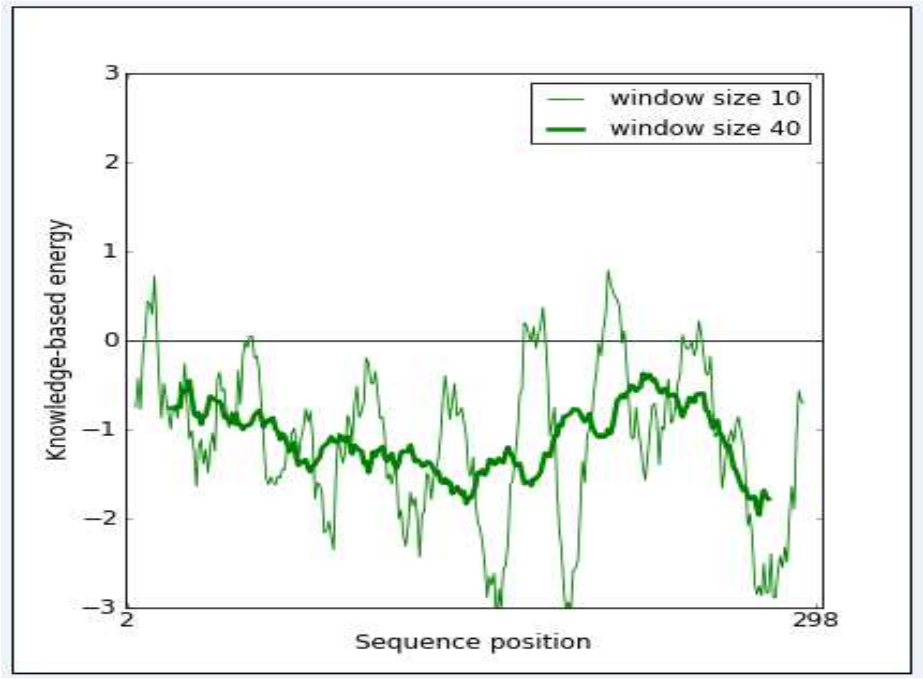

Figure 4b: ProSA plot of HBV core showing the energy graph of residue score of HBV core

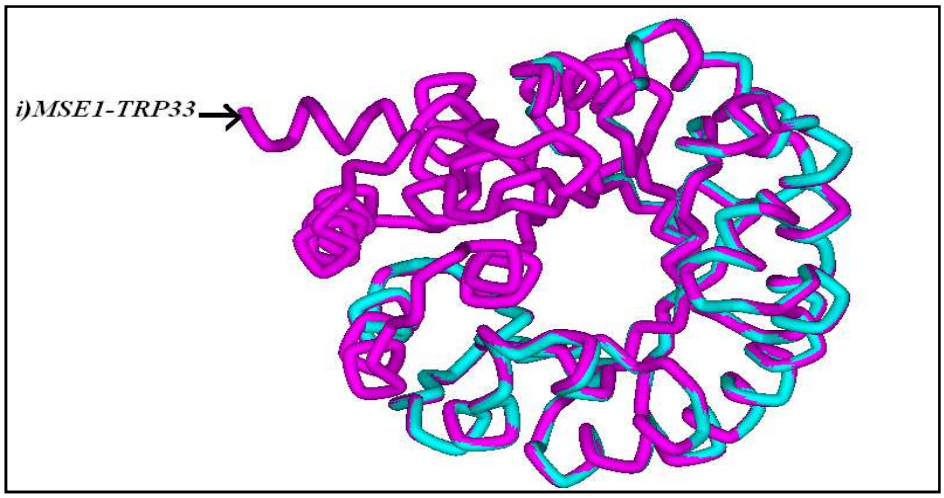

Figure 5: Structural superposition of HBV core (sky blue) onto the crystal structure of 1YDO (purple)

ProSA calculates an overall quality score for a specific input structure. Using the program ProSA the energetic structural design of protein folds is determined. The energy graph of HBV (target) obtained by ProSA whose Z-score is (-1.27) given in Figure (3a, 3b). Which indicate the overall quality of protein structure. ProSA reveals that the predicted model satisfies the criteria for a good quality model. The energy graph for Template 1YDO obtained by ProSA whose Z-score is (-9.21) in Figure (4a, 4b).

ERRAT the overall quality of the model is verified through Errat is 59.42 shows the structure reliability. The structure reliability diagram of ERRAT shows in Figure 2.

The overall description of similarities and differences derived from backbone superposition is given in Figure 6. The RMSD value between template and the predicted model using all main-chain atoms was found to be 1.4091 A^0. Both the sequences showed that the region are well superimposed but there are also some different conformations at

starting regions like, MSE1 to Trp 33 respectively, which is due to insertion of gaps presenting the loop region as shown in Figure 5.

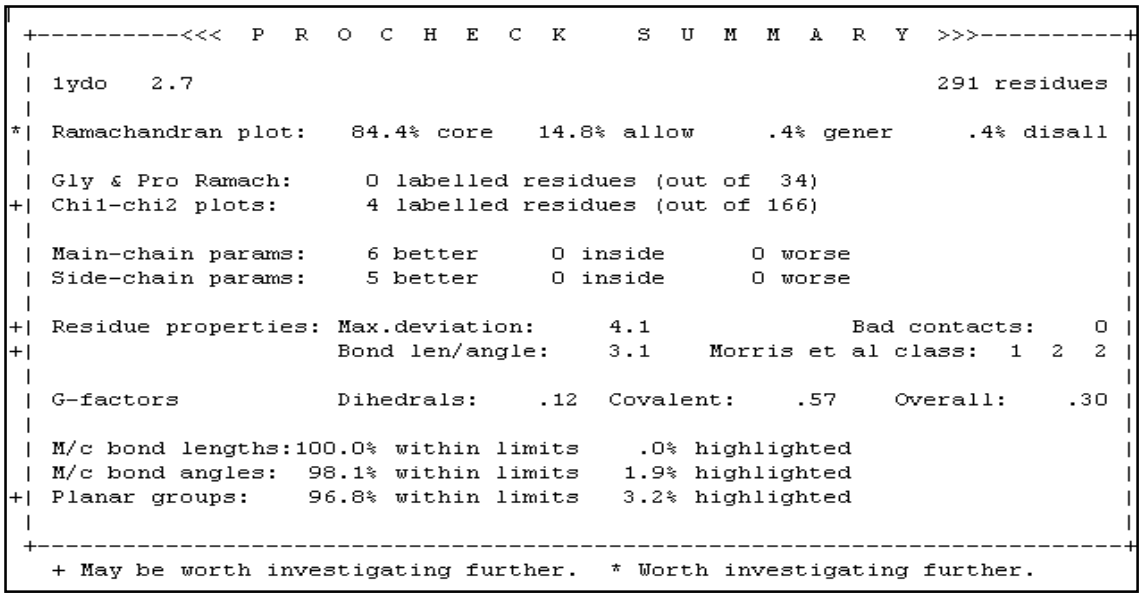

Figure 6a: Graphical representation of Ramachandran plot for HBV core model obtained with PROCHECK

The tertiary structure comprises mainly β-sheets and alpha helices. As shown in figure 3. The N-terminal is mainly consist of antiparellel β-sheets and C-terminal ends with α-helices. N-terminal starts with GLN 7 and antiparellel β-sheets and C- terminal ends with THR 268 and α-sheets.

Interactions as the result of docking with each of drugs are shown in table 9. The predicted structure of HBV core was docked with telbivudine and other three drugs taken randomly from chemspider (http://www. http://www.chemspider.com) on the basis of their molecular weights. Docking of drugs with HBV core was carried out by hex Dock server (http://hexserver.loria.fr/) & molecular dock server. http://www.dockingserver.com/web /.

Drug Compound Docking Analysis

Docking of the drugs was done with the help of online docking server i.e., hex dock server and molecular dock server.

Molecular Docking Server Results

In molecular docking server the drug bind with the target protein its shows different types of interaction with different residues of drug and target protein. The 1st drug telbivudine shows 5 different interactions polar hydrophobic, N bond, Cation-π and others. The drug distance is less than 4A^{o}. It means drugs change the conformation of protein and minimize the effect of the disease. The interaction of the drug and protein is showed in Table 2 and in this table also shows its drug and protein distance and drug atoms and protein residues.

Hex Docking Result

HEX dock bound at a distance less than 4 A^{o}, however each of these drugs has brought conformational changes in the catalytic site residue ALA: 11 and two other residues GLU: 10 and THR: 53, thus I propose that there might be a bit effect on the function of a protein because the drug is targeting the catalytic tirade and some other important residues and our this hypothesis can be confirmed by in-vitro or experimental studies. The result of the hex dock is shows in Table 1.

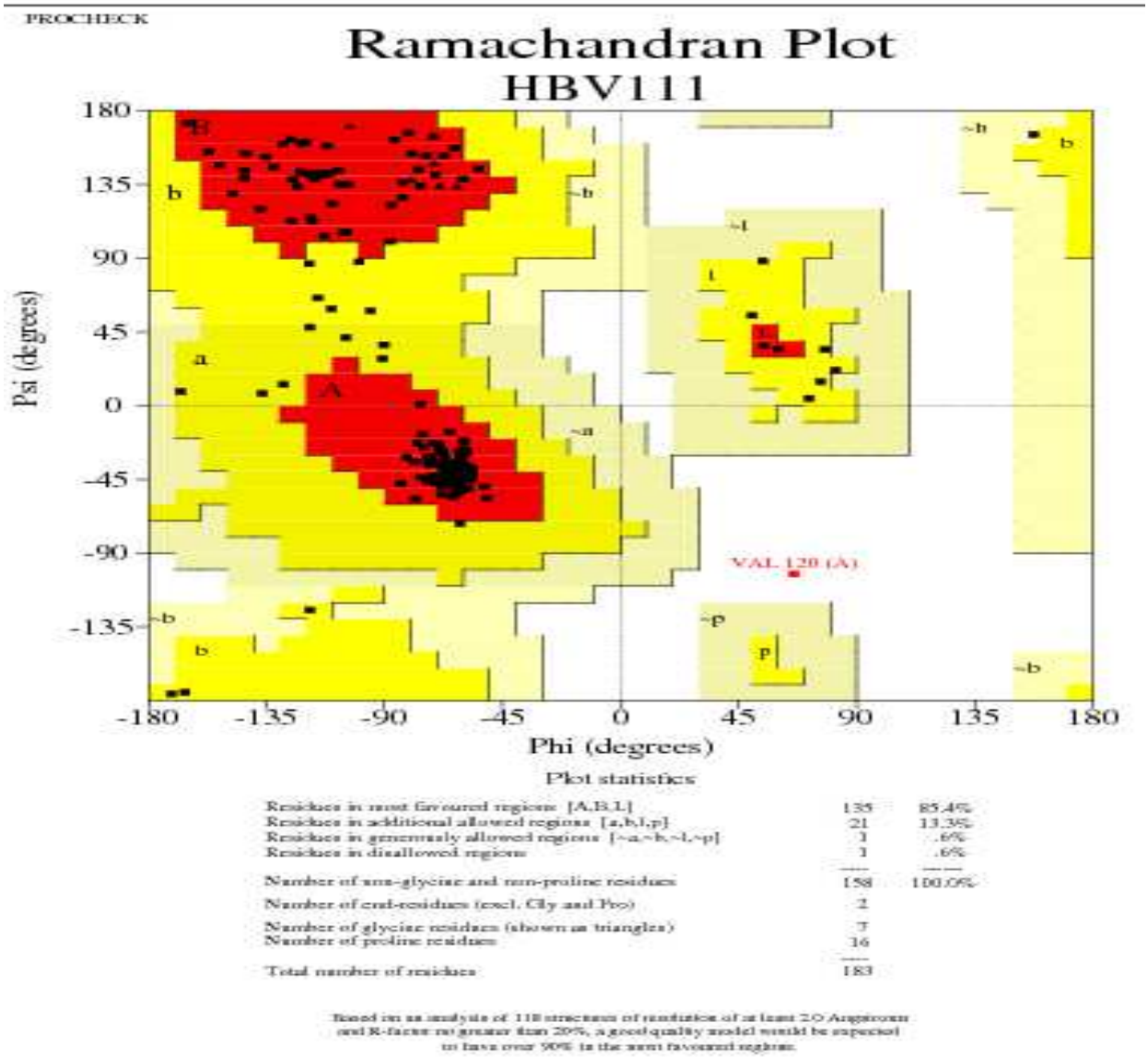

Figure 6b: Ramachandran plot statistics of the HBV core model obtained with PROCHECK

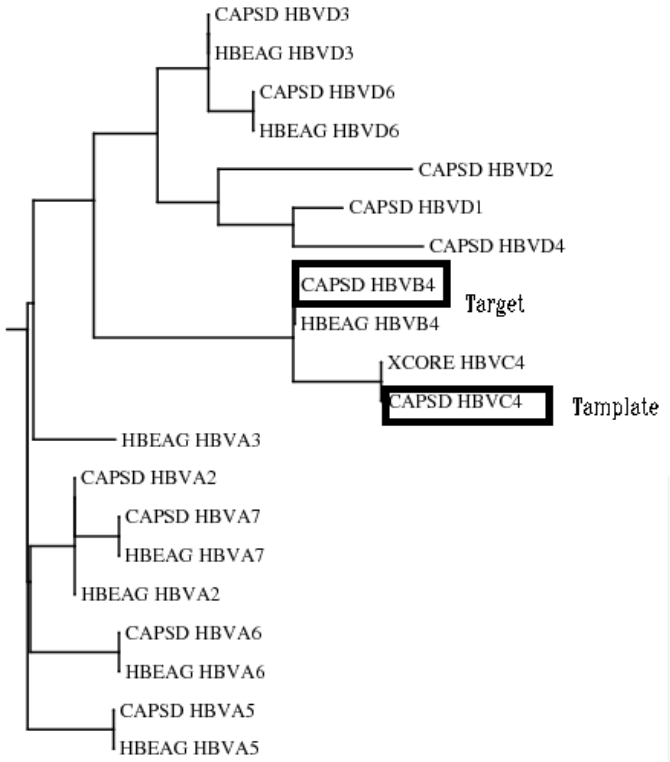

Figure 7: Showing the result of Draw gram

Table 1: Docking results of drugs of hex dock

Drugs	distance	DRUG RESIDUE	Protein residues	INTERACTION
	3.651	N4	SER21	Other
	2.828	O17	THR533	Other
	3.551	O17	ALN54 545454	Hydrophobic
telbivudine	3.674	O17	ALA54	Hydrophobic
	2.877	N1	ALA11	Hydrophobic
	3.365	O5	GLY10	Other
	3.41	N3	GLY10	Other
	3.542	O5	ALA41	Hydrophobic
lamlivudine	3.569	N1	LEU42	Hydrophobic
	3.286	H30	ALA41	Hydrophobic
	3.613	H31	ALA41	Hydrophobic
	3.82	H30	GLU10	Acidic+hydrophlic
	2.744	H30	GLU10	Acidic+hydrophlic
tenofovir	3.714	H30	GLU10	Acidic+hydrophlic

Table 2: Molecular docking results of drugs

Table 2.1: LAMIVUDINE

DRUG RESIDUE	TARGET PROTEIN RESIDUES	INTERACTIONS	DISTANCE
N1	THR109	HYDROGEN BOND	2.81
N1	GLU145	HYDROGEN BOND	2.94
H1	THR109	POLAR	1.95
H2	THR109	POLAR	2.98
H1	THR109	POLAR	3.67
H2	GLU145	POLAR	3.72
C3	ALA69	HYDROPHOBIC	2.1
C8	ALA69	HYDROPHOBIC	3.88
C6	TRP71	HYDROPHOBIC	3.51
C5	TRP71	HYDROPHOBIC	3.63
C2	TRP71	PI	3.82
O2	CYS48	OTHER	3.74
O2	ALA69	OTHER	3.53
H11	ALA69	OTHER	3.5
C5	TRP71	OTHER	3.48
C6	TRP71	OTHER	3.04
H1	THR109	OTHER	2.92
C1	THR109	OTHER	3.73
C4	THR109	OTHER	3.31
H2	GLU145	OTHER	3.81

Table 2.2: TELBIVUDINE

Drug residues	**Target protein residues**	**Interactions**	**Distance**
O3	SER81	N bond	2.97
O4	TRP71	Polar	3.38
O3	TRP71	Polar	3.8
H3	SER81	Polar	2.19
O4	THR 109	Polar	3.67
O4	GLU145	Polar	3.72
C2	ALA69	Hydrophobic	3.36
C2	ALA69	Hydrophobic	3.64
C1	TRP71	Hydrophobic	3.84
C4	ALA80	Hydrophobic	3.34
H3	TRP71	CATION-PI	3.51
c4	SER181	Other	3.75
H3	SER81	Other	2.87
C5M	THR109	Other	3.46
C5	THR109	Other	3.74
N3	THR 109	Other	3.82
C4	THR109	Other	3.4
O4	THR109	Other	3.64
C5M	GLU 145	Other	3.23
N1	ALA69	Other	3.64
O2	ALA69	Other	3.65
N3	ALA69	Other	3.56
H3	ALA69	Other	3.87
O3	TRP71	Other	3.35
C1	TRP71	Other	3.67
O4	TRP71	Other	3.83
H3	GLU77	Other	3.85
O4	ALA80	Other	3.6
H3	ALA80	Other	3.6
C3	SER81	Other	3.42

Table 2.3: TENOFOVIR

DRUG RESIDUES	PROTEIN RESIDUES	INTERACTION	DISTANCE
N5	THR109	Hydrogen bonds	[*2.86*]
N5	GLU145	Hydrogen bonds	[2.58]
O3	TYR88	Polar	[*3.33*]
H1	TYR88	Polar	[*3.67*]
H2	THR109	Polar	[*3.67*]
H3	THR109	Polar	[*2.24*]
H2	GLU145	Polar	[*2.20*]
H3	GLU145	Polar	[*2.27*]
H2	ARG175	Polar	[*3.19*]
N5	ARG175	Polar	[*3.76*]
C3	ALA69	Hydrophobic	[*3.74*]
C1	ALA69	Hydrophobic	[*3.70*]
C2	TRP71	Hydrophobic	[*3.40*]
C4	TRP71	Hydrophobic	[*3.84*]
C6	TRP71	Pi-pi	[*3.68*]
O4	PRO50	Other	[*3.73*]
N3	ALA69	Other	[*3.82*]
O1	ALA69	Other	[3.25]
O3	ALA69	Other	[*3.20*]
H1	ALA69	Other	[*2.56*]
C2	TRP71	Other	[*3.35*]
N1	TRP71	Other	[*3.87*]
C4	SER81	Other	[*3.61*]
O4	TYR88	Other	[*3.74*]
H3	THR109	Other	[*3.32*]
C8	THR109	Other	[3.35]
N4	THR109	Other	[*3.86*]
C5	THR109	Other	[*3.53*]
H2	GLU145	Other	[*3.15*]
H3	GLU145	Other	[*3.41*]
C8	GLU145	Other	[*3.77*]

CONCLUSION

The 17 residues are conserved in pair wise alignment between target and template. All the sequences shows highest similarity when these sequences were loaded into CUSTALX. The tertiary structure of HBV core consist of two main domains A+B structure topology. The ALA shows same interaction between hex and molecular docking server with three different drugs. The TRP and ALA show same interaction with three drugs in molecular docking. The drugs that were docked with HBV core (target) are also docked with template but these drugs do not show any interaction with 1YDO (template).

Acknowledgements
We owe special thanks to Department of Bioinformatics, Mohammad Ali Jinnah University, Islamabad for support in the research work.

REFERENCES

[1] J Poorolajal; R Majdzadeh, *Iranian Journal of Epidemiology*., **2009,** 4(3), 1-8.
[2] JL Dienstag, *New England Journal of Medicine*., **2008,** 359(14), 1486-1500.
[3] S Urban; A Schulze; M Dandri; J Petersen, *Journal of hepatology*., **2010,** 52(2), 282-284.
[4] FB Hollinger, TJ liang. Fields virology, 4th edition, **2001,** 2923-2969.
[5] FJ Mahoney, M Kane, Vaccines, 3rd edition, **1999,** 158-182.
[6] WS Ribinson, London, academic press Ltd, **1994,** 554-569.
[7] WS Robinson; GL mandell; JE Bennett; R dolin, *Churchill living stone*., **1995,** 1406-1539.
[8] M Puoti; C Torti; R Bruno, *Journal Hepatology*., **2006,** 44, S65-S70.
[9] AS Lok; BJ McMahon, *Journal of Hepatology*., **2001,** 34, 1225-1241.
[10] FV Chisari; C Ferrari, *Lippincott-raven*., **1997,** 745-778.
[11] E Gasteiger; E Jung; A Bairoch, *Molecular Biology*., **2001,** 3(3):47-55.
[12] J Tarraga; I Medina; L Arbiza; J H Cepas; T Gabaldon; J Dopazo, *Nucleic Acids Res*., **2007**, 35.
[13] S McGinnis; L Madden, *Nucleic Acids Research*., **2004,** 32.
[14] JD Thompson; TJ Gibson; F Plewniak; F Jeanmougin; DG Higgins, *Nucleic Acids Res*., **1997,** 25(24), 4876-82.
[15] DH Huson; DC Richter; C Rausch; T Dezulian; M Franz; R Rupp, *BMC Bioinformatics*., **2007,** 8:460.
[16] JA Cuff; ME Clamp; AS Siddiqui; M Finlay; GJ Barton, *Bioinformatics*., **1998,** 14(10), 892-893.
[17] N Eswar; B John; N Mirkovic; A Fiser; VA Ilyin; U Pieper; AC Stuart; MA Marti-Renom; MS Madhusudhan; B Yerkovich; A Sali, *Nucleic Acids Res*., **2003**, 31, 3375-3380.
[18] D P Pazel, *IBM Systems J*., **1989,** 28, 307–323.
[19] R A Laskowski; J A Rullmannn; M W MacArthur; R Kaptein; J M Thornton, *J Biomol NMR*., **1996,** 4, 477-486.
[20] M Wiederstein; M J, *Nucleic Acids Res*., **2007**, 35,10–40.
[21] M Y Shen; A Sali, *Protein Science*., **2006**, 15(11), 24–2507.
[22] G Macindoe; L Mavridis; V Venkatraman; M D Devigens; D W Ritchie, *Nucleic Acids Research*., **2010**, 38.