Journal of Computer Science & Systems Biology - Open Access
www.omicsonline.com Research Article JCSB/Vol.2 February 2009

HLA Class I and II Binding Promiscuity of the T-cell Epitopes in Putative Proteins of Hepatitis B Virus

Vijai Singh[1], Indramani [2], Dharmendra Kumar Chaudhary [2], Pallavi Somvanshi[1]*

[1]Bioinformatics Centre, Biotech Park, Sector-G Jankipuram, Lucknow-226021
[2]National Bureau of Fish Genetic Resources, Lucknow-226002

*Corresponding author: Pallavi Somvanshi, Bioinformatics Centre, Biotech Park, Sector –G, Jankipurum, Lucknow (India), Tel +91 522 4012078; Email: psomvanshi@gmail.com

Received January 20, 2009; Accepted February 25, 2009; Published February 27, 2009

Citation: Vijai S, Indramani, Dharmendra KC, Pallavi S (2009) HLA Class I and II Binding Promiscuity of the T-cell Epitopes in Putative Proteins of Hepatitis B Virus. J Comput Sci Syst Biol 2: 069-073.

Abstract

Hepatitis B virus is a human infectious disease universally caused by the hepatitis B virus. Its genome size is 3.215 kb. Immunoinformatics tools have been used to predict the epitopes from seven putative protein viz. polymerase, large-S- and middle –S- Protein, S and X- protein, Precore/Core Protein, Core and E- antigen. Total 50 epitopes were predicted for MHC class I and 55 epitopes for class II MHC molecules. These epitopes showed highest binding score at optimum threshold. Epitopes may use as an antigen for diagnosis and also might be helpful for designing peptide based subunit vaccine against Hepatitis B virus.

Key words: Hepatitis B virus; epitopes; diagnosis; vaccine

Introduction

Hepatitis B virus (HBV) is a dsDNA, enveloped virus that replicates in hepatocytes and belongs to family of the Hepadnavirus. The disease, originally known as "serum hepatitis (Barker *et al* 1996) has caused epidemics in various parts of Asia and Africa, and is endemic in China. (Williams 2006). The primary routes of transmission are through vertical transmission, blood and sexual exposure. Hepatitis B is potentially a life-threatening liver infection caused by hepatitis B virus. It can cause chronic liver disease and lay people at high risk of death from liver cirrhosis and cancer (Chang 2007). It is significantly more transmissible than HIV via blood-borne exposure, and some fluids that do not normally transmit HIV viz. saliva and sweat they contain infectious HBV but at lower levels when compared with blood. In many instances, patient route of infection is not identified. About one- third of world's population has been infected with the hepatitis B virus. This includes 350 million chronic carriers of the virus. The acute illness causes liver inflammation, vomiting, jaundice and death (WHO report).

Highly pathogenic and communicable virus cannot be cultured easily as it always requires high level of biosafety containment facilities. The culture of hepatitis B virus also requires the cell line for growth and all these steps are time consuming and laborious. Therefore, antigen preparation of hepatitis B virus is not easily processed. Several immunoinformatics tools are available for prediction and mapping of antigenic epitopes in protein sequence. It assist in designing subunit vaccines that starts from prediction of antigenic epitope through *in silico* techniques from protein sequence of pathogens independent of their abundance and without need to grow the microorganism *in vitro* (de Groot

and Rappuoli, 2004; Rappuoli, 2001).

Putative proteins of *M. tuberculosis* were predicted for cytoplasmic, integral membrane, secretory protein attached to membrane by Lipid anchor in the subcellular localization. The predictions provide a method to annotate *Mycobacterium* proteomes with subcellular localization information rapidly (Somvanshi a, *et al* 2008). Several proteins are involved in the pathogenesis and regulation of bacterial cell activity. The prediction of protein subcellular localization was used to distinguish the location in cells in *V. cholerae*. These proteins reside in cytoplasm, periplasm, inner membrane, outer membrane and extracellular space (Somvanshi b, *et al* 2008).

Prediction of epitopes in putative protein of hepatitis B virus with Propred and Propred1 are the new strategies to produce antigens for diagnosis. Synthetic peptides can use as vaccines to induce either humoral or cell-mediated immunity requires an understanding of the nature of T-cell and B-cell epitopes has been reported (Singh and Raghava, 2001; Bhasin et al., 2003; Singh and Raghava, 2003). The present study was aimed to predict and map the antigenic epitopes in hepatitis B virus.

Material and Methods

Bioinformatics tools were used for the analysis of genome of hepatitis B virus. The complete sequences of hepatitis B virus (NC_003977) were retrieved from www.ncbi.nlm.nih.gov. The open reading frames were identified from the whole genome using softwares viz. Generunner, ORF finder and DNAstar. The expected molecular weight, and isoelectric point (pI) value were also verified using Generunner and ExPaSy (http://www.expasy.org/). Propred (http://www.imtech.res.in/raghava/propred/) and Propred1 (http://www.imtech.res.in/raghava/propred1/) immunoinformatics tools are available to predict the antigenic epitopes in the complete protein primary sequences. Significant efforts have been made in the last few years as several groups devoted their research toward the development of procedures and algorithms that allow more effective and accurate prediction of MHC binding affinity.

The tools cover maximum number of human leukocyte antigen (HLA) in comparison to other epitope prediction tools. All these epitopes were predicted on the 4% threshold level with the highest binding score with the MHC molecules. T-cells interact through their polymorphic T cell receptor with HLA class I molecules containing endogenously synthesized peptides of 9-11 on the surface of infected cells. The presence of allele-specific amino acid motifs has been demonstrated by sequencing of peptides eluted from MHC molecules. Prediction of T-cell epitopes from protein sequence and have been widely used by experimental researchers without expert knowledge of bioinformatics (Lund *et al*, 2002).

Results and Discussion

In the present study, seven putative proteins of hepatitis B virus were used for the physicochemical analysis such as molecular weight, isoelectric point (pI value) and antigenic nature. The polymerase protein has the highest molecular weight 94.60 kDa and the lowest molecular weight 16.51 kDa was present in X- protein. Isoelectric points of these proteins were ranged between 7.87 to 9.94. The physicochemical propertievs of putative proteins were given (Table 1). The pI value of any protein indicates the stability of protein in that particular isoelectric point.

Putative proteins	Accession number	Expected M.W. (kDa)	pI value
Polymerase	NP647604	94.604	9.83
Large-S- Protein	YP355333	43.673	7.87
Middle –S- Protein	YP355334	31.154	8.02
S- Protein	NP647605	25.381	7.85
X- protein	NP647606	16.510	8.26
Precore/Core Protein	YP355335	24.291	9.29
Core and e- antigen	NP647607	21.097	9.94

Table 1: Physicochemical properties of different putative proteins of Hepatitis B virus.

Protein	T-cell epitopes	Amino acid position	No. of MHC Class II binding alleles	T-cell epitopes	Amino acid position	No. of MHC Class I binding alleles
Polymerase	VRRAFPHCL	536-544	03	FLLSLGIHL	573-581	10
	VVLSRKYTS	750-758	22	GLQPQQGSL	224-232	03
	LVVDFSQFS	387-395	11	FAVPNLQSL	407-415	08
	VVLGAKSVQ	552-560	07	FYPNLTKYL	115-123	04
	FNPDWKTPS	69-77	07	VSPSVPSHL	818-826	04
	YKRETTRSA	157-165	03	MRGTFVAPL	715-723	05
	LQSLTNLLS	411-419	11	ARVTGGVFL	367-375	06
	LVFQTSTRH	183-191	07	FSYMDDVVL	547-555	05
	IGSWGTLPQ	599-607	11	LPIHTAELL	723-731	10
	LNLYPVARQ	681-689	10	LPRLADEGL	25-33	05
	FGRKLHLYS	496-504	06	RPTTGRTSL	807-815	05
	IVQKLKQCF	610-618	07	DATPTGWGL	700-708	03
	FRKLLLLDD	7-15	05	RSRSGAKLI	737-745	03
	YVSLLLLYK	486-494	12	CSRNLYVSL	482-490	03
	YRPLLHLPF	797-805	07			
	YQHFRKLLL	4 -12	09			
Large-S-Protein	FLPLLPIFF	385-393	05	FLFILLLCL	257-265	14
	LRRFIIFLF	250-258	16	SLDSWWTSL	205-213	07
	VNPVPTTAS	150-158	06	SSWAFARFL	328-336	08
	FVQWFVGLS	352-360	08	LYNILSPFL	379-387	04
	FLLTRILTI	193-201	05	MESTTSGFL	175-183	04
	LIFLLVLLD	264-272	10	GFFLLTRIL	192-200	05
				WSPQAQGIL	77-85	05
Middle –S-Protein	LRRFIIFLF	131-139	16	FLFILLLCL	138-146	15
	FVQWFVGLS	233-241	08	SLDSWWTSL	86-94	06
	FLLTRILTI	74 -82	07	SSWAFARFL	209-217	07
	VNPVPTTAS	31- 39	06	LYNILSPFL	260-268	06
	LGPLLVLQA	63-71	06	MESTTSGFL	56-64	07
	LIFLLVLLD	145-153	10	GFLGPLLVL	62-70	03
	FLPLLPIFF	266-274	05			
S- Protein	LRRFJIFLF	76-84	17	FLFILLLCL	83-91	16
	FVQWFVGLS	178-186	08	SLDSWWTSL	31-39	06
	FLLTRILTI	19-27	07	SSWAFARFL	154-162	07
	LGPLLVLQV	8-16	06	LYNILSPFL	205-213	06

	LIFLLVLLD	90-98	10	MESTTSGFL	1-9	04
	FLPLLPIFF	211-219	05	GFFLLTRIL	18-26	05
X- protein	LRGLPVCAF	54- 62	26	TTVNAHQVL	81-89	05
	LRFTSARSM	70-78	11	EIRLKVFVL	126-134	08
	VNAHQVLPK	82-90	07	CLFKDWEEL	115-123	07
	VLHKRTLGL	91-99	11	FSSAGPCAL	63-72	06
	LCLRPVGAE	15-23	03	SPSSSAVPA	39-47	08
	LKVFVLGGC	128-136	03			
Precore/Core Protein	LVVSYVNVN	112-120	10	KEFGASVEL	36-44	07
	VRRRGRSPR	177-185	07	LTFGRETVL	137-145	08
	FGRETVLEY	138-146	04	DLLDTASAL	58-66	05
	LVSFGVWIR	147-155	12	LCWGELMNL	89-97	07
	LMNLATWVG	93-101	04	YRPPNAPIL	161-169	06
	LKIROLLWF	123-131	16	LPSDFFPSI	48-56	09
	VELLSFLPS	41-49	03			
Protein Core and e antigen	LVVSYVNVN	83-91	10	KEFGASVEL	7-15	08
	VRRRGRSPR	148-156	07	LTFGRETVL	108-116	04
	LVSFGVWIR	118-126	12	DLLDTASAL	29-37	05
	LMNLATWVG	64-74	07	LCWGELMNL	60-68	06
	WFHISCLTF	101-109	08	YRPPNAPIL	132-140	06
	LKIRQLLWF	94-102	16	LPSDFFPSI	19-27	08
	VELLSFLPS	12-20	03			

Table 2: The predicted epitopes in the putative proteins of Hepatitis B virus.

The prediction of epitopes in Polymerase, large-s- protein, middle –s- Protein, s- protein, x- protein, precore/core protein, core and e- antigen, proteins of hepatitis B virus were investigated. Total 50 epitopes were predicted for class I MHC and 55 epitopes for class II MHC molecules in these proteins. The predicted epitopes having the MHC alleles for putative proteins from hepatitis B virus were performed (Table 2). Evaluation of synthetic peptides as potential vaccine candidate for flavivirus has been investigated. Using the computational tools for prediction of epitopes and synthetic peptides from E glycoprotein of Murray Valley encephalitis (MVE) and DEN 2 viruses were prepared and their immunogenicity was evaluated in mice (Gao *et al*, 1990). The identification of significant T-cells epitopes from secretory and cell surface proteins virulent proteins of *M. tuberculosis* H37Rv strain was done. The promiscuous nanomer candidate epitopes from HTL and CTL were recognized (Somvanshi a *et al*, 2008). T-cell analyses of synthetic peptides to other viruses have correlated the association between T- and B-cell responses (Hu et al, 1999). A new approach for vaccine design in immunology and the development of bioinformatics tools for T cell epitope prediction from primary protein sequences is essential. Bioinformatics tools have the potential to accelerate research into the design of vaccines and diagnostic tests by exploiting genome sequences. *In silico* analysis could be combined with *in vitro* screening methods to identify the peptides that are immunogenic. Chemically synthesized domains of FMDV (Food-mouth disease virus) VP1 were tested as potential peptide vaccine. The peptide corresponding to amino acid 141-160, 151-160 and 200-213 which are located near the C- terminal end of VP1 and 9-24, 17-32 and 25-41, and the N-terminal end of the VP1, were each bound to a separate insert carrier protein (Bittle et al, 1982).

The prediction of T-cells epitopes in highly virulence sur-

face proteins i.e. hemagglutinin and neuraminidase from Influenza A virus H5N1 was performed. The highly conserved and specific epitopes were predicted the Influenza host and strain specific. These epitopes were used for the immunodiagnostic of Influenza and also as a novel epitope based vaccine candidate (Somvanshi c *et al*, 2008). Immunoinformatics serves as a valuable tool to screen and select antigenic peptide sequences as potential T cell epitopes for binding affinity with HLA alleles. The dengue variants structural protein was studied and promiscuous nanomer candidate epitope for MHC class I and II has been recognized (Somvanshi and Seth, 2008).

Conclusion

New theoretical and immunoinformatic approaches are currently used for prediction of epitopes in the proteins sequence of hepatitis B virus without using their cultures. The prediction of hepatitis B virus nanomer epitope for T-cells is recognized against MHC Class II and MHC class I. It may use as an easy and effective way to diagnose the suspected individuals during a future hepatitis B virus epidemic, thereby reducing and containing the transmission. These epitopes may also use for the vaccination against hepatitis B virus.

References

1. Barker LF, Shulman NR, Murray R, et al. (1996) Transmission of serum hepatitis. 1970 JAMA 276: 841-4.

2. Bhasin M, Singh H, Raghava GPS (2003) MHCBN: a comprehensive database of MHC binding and non-binding peptides. Bioinformatics 19: 666-667.

3. Chang MH (2007) Hepatitis B virus infection. Semin Fetal Neonatal Med 12: 160-7.

4. De Groot AS, Rappuoli R (2004) Genome-derived vaccines. Expert Rev Vaccines 3: 59-76.

5. Hu GJ , Wang RYH, Han DS, Alter HJ, Shih JWK (1999) Characterization of the humoral and cellular immune responses against hepatitis C virus core induced by DNA-based immunization. Vaccine 17: 3160-3170.

6. Lund O, Nielsen M, Kesmir C, Christensen JK, Lundegaard C, et al. (2002) Web-based tools for vaccine design. In HIV *Molecular Immunlogy* (Los Alamos, NM, LA-UR 03-5816: Theoretical Biology and Biophysics Group. Los Alamos National Laboratory pp45-51.

7. Rappuoli R (2001) Reverse vaccinology, a genome-based approach to vaccine development. Vaccine 19: 2688-2691.

8. Singh H, Raghava GPS (2001) ProPred: prediction of HLA-DR binding sites. Bioinformatics 17: 1236-1237.

9. Singh H, Raghava GPS (2003) ProPred1: prediction of promiscuous MHC class-I binding sites. Bioinformatics 19: 1009-1014.

10. Somvanshi P, Seth PK (2009) Prediction of T cell epitopes for the utility of vaccine development from the structural proteins of dengue virus variants using *in silico* methods. Indian J Biotechnol (*In Press*).

11. Somvanshi P, Singh V, Seth PK (2008) *In silico* prediction of epitopes in virulence proteins of *Mycobacterium tuberculosis* H37Rv for diagnostic and subunit vaccine design. J Proteomics & Bioinformatics 1: 143-153.

12. Somvanshi P, Singh V, Seth PK (2008) Prediction of epitopes in hemagglutinin and neuraminidase proteins of influenza A virus H5N1 strain: A clue for diagnostic and vaccine development. J Integrative Biology 12: 61-69.

13. Somvanshi P, Singh V, Seth PK (2008) Assessing the precision of high-throughput computational approaches for the genome-wide subcellular localization of putative proteins from *Vibrio cholerae*. The Internet Journal of Genomics and Proteomics 3(2).

14. Somvanshi P, Singh V, Seth PK (2008) *In silico* analysis of subcellular localization of putative proteins of *Mycobacterium tuberculosis* H37Rv strain. The Internet Journal of Health 7(1).

15. Somvanshi P, Singh V, Seth PK (2008) Phylogenetic and computational proteome analysis of Influenza A virus subtype H5N1. The Internet Journal of Genomics and Proteomics 3(2).

16. Williams R (2006) Global challenges in liver disease. Hepatology 44: 521-6.

17. FAQ about Hepatitis B, Stanford University School of Medicine.