Int J Clin Exp Pathol 2016;9(3):3816-3827
www.ijcep.com /ISSN:1936-2625/IJCEP0020899

Original Article

Amino acids at positions 30β1 and 57β1 of HLA-DR confer susceptibility to or protection from chronic hepatitis B virus infection

Yan Xia[1*], Wenxuan Sun[1*], Xingku Li[1*], Hongyan Wang[2], Xueyuan Yu[1], Xi Jin[1], Bo Du[1], Yuguang Shi[1], Zhen Liu[1], Shuyun Zhang[1]

[1]Research Center, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang Province, P. R. China; [2]Institute of Harbin Hematology & Oncology, Harbin First Hospital, Harbin, Heilongjiang Province, P. R. China. [*]Equal contributors.

Received December 1, 2015; Accepted January 31, 2016; Epub March 1, 2016; Published March 15, 2016

Abstract: This study is to investigate whether structures and electrostatic potentials of human leukocyte antigen (HLA)-DRB1 molecules are relative to the outcomes of hepatitis B virus (HBV) infection. Totally 230 subjects were categorized into the recovered HBV infection group (RH, n = 86) and the chronic HBV infection group (CH, n = 144). Genotypes and amino acid sequences of HLA-DRB1 were obtained. Structure templates suitable for modeling were checked in Protein Data Bank. The atomic coordinates of HLA-DR molecules were determined by MODELLER computer algorithm. The analyses for the polymorphic residues of HLA-DRB1 proteins were performed using the stepwise logistic regressions. CH and RH-associated polymorphic residues 57, 30, and 28 had more significant associations with the outcomes of HBV infection compared with other residues according to "Allele" model. CH and RH-associated polymorphic residues 30 and 57 had more significant associations with the outcomes of HBV infection compared with other residues according to "Genotype" model. Polymorphic residues 14, 16, 25, 28, 30, 37, 38, 57 and 85 were significant residues according to "Allele" model. HLA alleles carried significant residues. Special residues influenced the electrostatic properties of HLA-DRB1 pockets. The residues Lys14, Tyr16, Gln25, Glu28, His/Leu30, Leu37, Leu38, Val57 and Ala85, especially His/Leu30 and Val57, in chronic HBV population were significantly enriched. This suggests that the residues 14, 16, 25, 28, 30, 37, 38, 57 and 85, especially 30 and 57, in the HLA-DR β chain critically influence the structural and electrostatic properties of the peptide-binding groove.

Keywords: Human leukocyte antigen, hepatitis B virus, amino acid polymorphism, three-dimensional structure

Introduction

Hepatitis B virus (HBV) can cause spontaneously recovered HBV infection (RH) and chronic HBV infection (CH) [1, 2]. Some strong genetic factors are found to be involved in the human leukocyte antigen (HLA) complex, especially HLA class II molecules, which influence selection of HBV antigen and the HBV-specific $CD4^+$ T cell response [3-5]. HLA class II alleles mainly include HLA-DR, DQ and DP subregions. Since 1983 when HLA class II typing in HBV infection was investigated [6], a number of HLA-DRB1 alleles associated with the outcomes of HBV infection were reported. For example, DRB1*03, *11, *13 and *14 may be protective factors and DRB1*07, *09 and *12 may be susceptible to CH in the global population [7-10]. Recently, reports by Genome Wide Association Studies proved that HLA-DQ and HLA-DP had independent associations with various HBV-outcomes [11, 12]. However, strong LD and ethnic difference in HLA alleles made it difficult to determine which allele was the most relevant. An analysis of structure characters of HLA molecules provided a unifying mechanism [13, 14].

The HLA class II molecules are heterodimeric glycoproteins, which are able to present peptides to $CD4^+$ T cells. The primary protein sequences (α1 and β1 chains) encoded by the second exon of the HLA class II genes comprise the peptide-binding grooves that accommodate amino acid side chains of the bound peptides.

Table 1. The demographic and serological description of 230 subjects

Characteristics	CH (n = 144)	RH (n = 86)
Age (means ± SD, years)	38.26 ± 12.60	40.71 ± 13.34
Sex (male:female)	111:33	56:30
HBV DNA	$6.26\times10^7 \pm 1.44\times10^8$	-
HBsAg-positive	144 (100%)	0 (0%)
Anti-HBs positive	0 (0%)	27 (31.4%)
HBeAg	110 (76.4%)	19 (22.1%)
Anti-HBe	16 (11.1%)	24 (27.9%)
Anti-HBc positive	132 (91.7%)	78 (90.7%)

Note: CH, chronic hepatitis B virus infection; RH, spontaneous recovery from hepatitis B. Positive scores for HBsAg, anti-HBc, anti-HBs, HBeAg, and anti-HBe were given as the number of patients (%).

The specificity of the peptide-binding grooves is governed by the properties of pockets in the grooves. Typical pockets include pockets P1, P4, P6 and P9. In HLA-DR molecules, HLA-DRα1 chain encoded by HLA-DRA is almost an invariant chain, but HLA-DR β1 chains encoded by HLA-DRB alleles show high polymorphism. Some studies reported that the polymorphic amino acid residues of HLA class II molecules confer susceptibility to or protection from diseases. For example, Asp at residue 57 of HLA-DQβ chain is associated with susceptibility to pulmonary tuberculosis [15], but this position can exert protective effect for type I diabetes [16]. The polymorphic residues can influence the structural and electrostatic properties of peptide-binding groove [17], so the determination of the structural and electrostatic properties of the HLA class II peptide-binding groove associated with diseases may help identify disease mechanism. For example, in primary sclerosing cholangitis, Asn at residue 37 and Gly at residue 86 of HLA-DR β1 chain directly and indirectly confer susceptibility to primary sclerosing cholangitis by making pocket P9 more positively charged [18]. In type I diabetes, Asp transform to Ala at residue 57 of HLA-DQ2 and DQ8 β chain increases the accessible volume of pocket P9 and imparts it strong positive charge. These changes might confer susceptibility to type I diabetes [19, 20]. A systematic characterization of pocket properties (for example, volume and electrostatic charge) highlights the differences in their ability to accommodate specific amino acid side chain (for example, loose or tight binding). For example, Arg at residue 74 of HLA-DR β1 chain confers susceptibility to the autoimmune polyglandular syndrome type 3 by making pocket P4 more positively charged and increasing the strength of interactions with the bound peptide (Tg2098) [21]. In HBV infection, the HLA-DRB1*130101 allele is associated with self-limited HBV infection and plausibly conveys protection against chronicity [22, 23]. To a certain degree, this is due to the fact that HLA-DR13-positive individuals can preferentially recognize HBcAg141-165 (the minimal epitope is ^{147}TVVRRRGRSP^{156}) [22] and induce a vigorous HBV-special $CD4^+$ T cell response [23]. However, HBV-associated HLA class II molecules remain controversial, and no attempts have been made to investigate how specific amino acids affect the structure and electrostatic properties of the peptide-binding groove.

In this study, we aim to investigate how the polymorphic amino acid residues of HLA-DR β1 chains affect the molecular characteristics of HLA-DR peptide-binding groove and various HBV-outcomes. The analyses were performed in the absence of HBV epitopes to enable direct comparison of the structural and electrostatic properties among different HLA-DR peptide-binding grooves.

Materials and methods

Patients

A total of 230 subjects were categorized into two groups: the recovered HBV infection group (RH, n = 86) and the chronic HBV infection group (CH, n = 144). RH group included patients that were recovered from occult HBV infection (n = 32) and acute HBV infection (n = 54). The former was characterized by hepatitis B surface antigen negative [$HBsAg^-$], HBV DNA negative, hepatitis B core antibody positive [anti-HBc^+] and/or hepatitis B surface antibody positive [anti-HBs^+] which came from routine physical examinations in clinic service without any clinical symptoms and vaccination history. The latter was [HBsAg]-positive with high-titer IgM of anti-HBc, HBV DNA ≥ 1.0×10^3 copies/ml and self-limited from liver disease, reversed from $HBsAg^+$, and scored anti-HBs^+ ≥ 24 weeks after their initial presentation by symptomatic treatment. The individuals in CH group had been HBsAg persistence for more than 6 months and HBV DNA ≥ 1.0×10^3 copies/ml. The

Table 2. Association analyses between residues of HLA molecules and outcomes of HBV infection assuming "Allele" model

Residue	Observed Amino Acids	Single-residue LR	Two-residue LR*							
		Basic Model	Basic Model: Residue 14	Basic Model: Residue 16	Basic Model: Residue 25	Basic Model: Residue 28	Basic Model: Residue 30	Basic Model: Residue 37	Basic Model: Residue 38	Basic Model: Residue 85
		P value	*P* value	*P* value	*P* value	*P* value	*P* value	*P* value	*P* value	*P* value
4	R Q	0.538	0.470	0.776	0.470	0.516	0.616	0.601	0.741	0.762
9	E K W	0.246	0.984	**0.009**	0.984	0.260	0.412	0.455	**0.036**	**0.030**
10	E Q Y	0.809	0.709	**0.021**	0.709	0.897	0.764	0.598	0.480	0.107
11	L S V G D P	0.152	0.619	**0.008**	0.619	0.366	0.378	0.613	**0.023**	**0.023**
12	K T	0.851	0.591	**0.014**	0.591	0.822	0.669	0.722	0.240	0.095
13	S H Y G F R	0.448	**0.012**	0.590	**0.012**	0.787	0.497	0.103	0.844	0.794
14	E K	**0.026**	-	**0.006**	0.465	**0.028**	0.120	0.236	**0.010**	**0.009**
16	H Y Q	**0.005**	**0.001**	-	**0.001**	**0.003**	**0.027**	**0.009**	0.088	0.263
25	R Q	**0.026**	0.586	**0.006**	-	**0.028**	0.120	0.236	**0.010**	**0.009**
26	L Y F	0.390	0.152	0.301	0.152	0.942	0.201	0.649	0.501	0.513
28	D E H	0.264	0.284	0.156	0.284	-	**0.024**	0.635	0.290	0.293
30	C Y L G R H	**1.2×10^{-3}**	**0.006**	**0.007**	**0.006**	**2.6×10^{-4}**	-	**0.016**	**0.013**	**0.027**
31	I F V	0.772	0.542	0.366	0.542	0.672	0.277	0.611	0.695	0.461
32	Y H	0.382	0.147	0.508	0.147	0.270	0.464	0.823	0.411	0.322
33	N	0.272	0.400	0.542	0.400	0.383	0.854	0.965	0.411	0.460
37	S N Y F L	**0.015**	0.158	**0.030**	0.158	**0.027**	0.292	-	**0.033**	0.071
38	V A L	**0.017**	**0.007**	0.387	**0.007**	**0.018**	0.185	**0.034**	-	0.740
40	F Y	0.836	0.767	0.705	0.767	0.265	0.814	0.564	0.164	0.740
47	F Y	0.675	0.238	0.790	0.238	0.363	0.276	0.424	0.617	0.332
57	D Y S A	**5.6×10^{-4}**	**0.006**	**0.015**	**0.006**	**3.6×10^{-4}**	0.141	**0.001**	**0.005**	**0.014**
58	A E	0.050	0.084	0.129	0.084	0.077	0.252	0.261	0.096	0.103
60	Y S H	0.557	0.866	0.866	0.866	0.801	0.603	0.083	0.813	0.991
67	L I F	0.774	0.611	0.762	0.611	0.746	0.390	0.520	0.825	0.973
70	D Q R	0.197	0.673	0.881	0.673	0.106	0.355	0.322	0.634	0.975
71	A E K R	0.439	0.774	0.861	0.774	0.735	0.250	0.947	0.882	0.931
73	A G	0.056	0.914	**0.009**	0.914	**0.048**	0.105	0.398	**0.018**	**0.016**
74	A E L Q R	0.217	0.632	0.132	0.632	0.235	0.291	0.854	**0.044**	**0.038**
77	N T	0.946	0.914	0.789	0.914	0.922	0.573	0.844	0.875	0.859
78	V Y	0.111	0.743	**0.008**	0.743	0.246	0. 583	0.602	**0.023**	**0.020**
85	A V	**0.007**	**0.003**	0.360	**0.003**	**0.008**	0.137	**0.027**	0.164	-
86	G V	0.353	0.088	0.962	0.088	0.146	0.110	0.793	0.964	0.684

Note: Stepwise logistic regression was performed using the "Allele" model. LR, likelihood ratio. *, *P* values of likelihood ratio tests of whether residue N improves the model when added to a model with one another residue (14, 16, 25, 28, 30, 37, 38, 57, or 85). For all, $P < 0.05$ is highlighted in bold.

Table 3. Association analyses between residues of HLA molecules and outcomes of HBV infection using a "Genotype" model

Residue	Single-Residue LR	Two-Residue LR*					
	Basic Model	Basic Model: Residue 16	Basic Model: Residue 25	Basic Model: Residue 37	Basic Model: Residue 38	Basic Model: Residue 57	Basic Model: Residue 85
	P value	*P* value	*P* value	*P* value	*P* value	*P* value	*P* value
9	0.176	0.097	0.532	0.242	0.038	0.193	**0.034**
13	0.244	0.393	**0.031**	0.069	0.701	0.520	0.725
14	0.068	**0.032**	0.999	0.450	**0.017**	0.133	**0.017**
16	**0.047**	-	**0.027**	0.066	0.127	0.137	0.176
25	**0.038**	**0.017**	-	0.342	**0.008**	0.100	**0.008**
30	**1.2×10^{-3}**	**0.002**	**0.004**	**0.010**	**0.010**	**0.010**	**0.017**
37	**0.020**	**0.034**	0.188	-	**0.033**	**0.047**	**0.049**
38	**0.019**	0.071	**0.005**	**0.029**	-	0.057	0.692
57	**0.020**	0.083	0.050	**0.043**	0.060	-	0.097
78	0.171	**0.046**	0.816	0.628	**0.032**	0.505	**0.029**
85	**0.010**	0.054	**0.002**	**0.022**	0.238	**0.047**	-

Note: Stepwise logistic regression was performed using the "Genotype" model. LR, likelihood ratio. *, *P* values of likelihood ratio tests of whether residue N improves the model when added to a model with one another residue (16, 25, 37, 38, 57, or 85). For all, $P < 0.05$ is highlighted in bold.

demographics and detailed descriptions of the serology were listed in **Table 1**. All subjects had no familial relationship with each other. Subjects were diagnosed according to the Virus Hepatitis Diagnosis Standard of China (2008/2010) [1, 2]. None of the subjects were positive for hepatitis viruses A, C and D (HAV, HCV and HDV) or human immunodeficiency virus (HIV) by antibody testing. The study protocol conformed to the 1975 Declaration of Helsinki regulations. The study was approved by the ethics committee of Harbin Medical University. Informed consent was obtained from each participant.

HLA-DRB1 data

Six-digit HLA-DRB1 genotypes were available from a previous data [10]. Amino acid sequences encoded by exon two of HLA-DRB1 alleles were obtained from the IMGT/HLA database (http://www.ebi.ac.uk/ipd/imgt/hla/ambig.html) release 3.12.0 (April 2013).

Three-dimensional modeling

Structure templates suitable for modeling were checked in Protein Data Bank (PDB; http://www.rcsb.org/pdb/), from which five structures (PDB entries: 3C5J, 1AQD, 1D5M, 1A6A and 2Q6W) were chosen for further calculation for their highest scores of sequence distance (only β chains) compared to the target amino acid sequence. Multiple sequence alignments were performed with CLUSTAL_W method using DNA-Star (v.7.10, DNASTAR Company, USA) [24]. The atomic coordinates of HLA-DR molecules were determined using comparative protein structure modeling by MODELLER (v.9.13, Andrej Sali, USA) computer algorithm [25]. In order to directly compare the structural and physiochemical characteristics of the peptide-binding groove among different molecules, all calculations were performed in the absence of antigenic peptides. MODELLER generates the new atomic coordinates of the target sequences by satisfying the spatial restraints. Model validation was carried out using Autodock Vina (v.1.1.2, Molecular Graphics Laboratory, USA). Surface electrostatic potentials of each atomic coordinates were calculated and presented by PyMOL (v.1.5.0.3, Delauo Schrodinger LLC, USA) [26] and APBS (v.1.3, APBS Company, USA) [27]. Potentials < -25 kT/e were red, those > 25 kT/e were blue and neutral potentials (0 kT/e) were white.

Statistical analysis

The analyses for the polymorphic residues of HLA-DR β1 chains were performed using the stepwise logistic regressions. Two models were assumed: i) "Allele" model, in which the count of all amino acids at a given residue was

Table 4. Different amino acids at residue positions that showed strong association with the outcomes of HBV infection

Amino acid position	Amino acid variants	Carrier frequency n (%)		*P* value	OR	95% CI
		CH (n = 288)	RH (n = 172)			
14	Glu	248 (86)	160 (93)	**0.023**	0.465	0.237-0.913
	Lys	40 (14)	12 (7)	**0.023**	2.151	1.095-4.224
16	His	209 (73)	145 (84)	**0.004**	0.493	0.303-0.801
	Tyr	73 (25)	25 (15)	**0.006**	1.996	1.211-3.293
	Gln	6 (2)	2 (1)	0.465	1.809	0.361-9.062
25	Arg	248 (86)	160 (93)	**0.023**	0.465	0.237-0.913
	Gln	40 (14)	12 (7)	**0.023**	2.151	1.095-4.224
28	Asp	137 (48)	108 (63)	**0.002**	0.538	0.365-0.791
	Glu	106 (37)	37 (22)	**0.001**	2.125	1.375-3.285
	His	45 (16)	27 (16)	0.983	0.995	0.592-1.672
30	Cys	9 (3)	4 (2)	0.617	1.355	0.411-4.468
	Tyr	138 (48)	112 (65)	**3.4×10^{-4}**	0.493	0.334-0.728
	Leu	40 (14)	12 (7)	**0.023**	2.151	1.095-4.224
	Gly	45 (16)	27 (16)	0.983	0.995	0.592-1.672
	Arg	4 (1)	2 (1)	0.836	1.197	0.217-6.606
	His	52 (18)	15 (9)	**0.006**	2.306	1.255-4.240
37	Ser	53 (18)	35 (20)	0.608	0.883	0.548-1.421
	Asn	66 (23)	44 (26)	0.517	0.865	0.557-1.342
	Tyr	59 (20)	50 (29)	**0.036**	0.629	0.406-0.972
	Phe	58 (20)	28 (16)	0.304	1.297	0.789-2.131
	Leu	52 (18)	15 (9)	**0.006**	2.306	1.255-4.240
38	Val	232 (81)	155 (90)	**0.007**	0.454	0.255-0.811
	Ala	4 (1)	2 (1)	0.836	1.197	0.217-6.606
	Leu	52 (18)	15 (9)	**0.006**	2.306	1.255-4.240
57	Asp	121 (42)	94 (55)	**0.009**	0.601	0.411-0.880
	Val	137 (48)	54 (31)	**0.001**	1.983	1.334-2.947
	Ser	17 (6)	10 (6)	0.969	1.016	0.454-2.273
	Ala	13 (5)	14 (8)	0.109	0.534	0.245-1.164
85	Ala	52 (18)	157 (9)	**0.006**	2.306	1.255-4.240
	Val	236 (82)	157 (91)	**0.006**	0.434	0.236-0.797

Note: The statistical analysis was performed based on χ^2 test. For all, $P < 0.05$ is highlighted in bold. OR, odds ratio. CI, confidence interval.

entered as covariates; ii) "Genotype" model, in which all observed combinations (pairs) of amino acids at a given residue were entered as covariates. "Genotype" model was applied to control the validity of "Allele" model. In both models, the reference was randomly chosen, and thus no assumptions were made on which amino acid or pair of amino acids contributed to high or low risk. The analyses for the frequency of amino acids were performed using χ^2 test. All statistical analyses were performed using the SPSS statistical software package version 17.0 (SPSS, Chicago, IL, USA). $P < 0.05$ was considered statistically significant.

Results

CH and RH-associated polymorphic residues 57, 30, and 28 have more significant associations with the outcomes of HBV infection compared with other residues according to "Allele" model

The amino acid sequences of HLA-DR β1 chains were determined from the genotype of each

Table 5. The frequency and percentage of HLA-DR carrying risky and protective amino acids of 9 polymorphic residue positions

Residue	Amino acid	DR04	DR07	DR08	DR09	DR11	DR12	DR14	DR15	Pocket
14β	**Lys**		52 (11.3)							4
	Glu	42 (9.1)		26 (5.7)	72 (15.7)	34 (7.4)	67 (14.6)	39 (8.5)	67 (14.6)	
16β	**Tyr**			26 (5.7)			67 (14.6)			6
	His	42 (9.1)	52 (11.3)		72 (15.7)	34 (7.4)		39 (8.5)	67 (14.6)	
25β	**Gln**		52 (11.3)							4
	Arg	42 (9.1)		26 (5.7)	72 (15.7)	34 (7.4)	67 (14.6)	39 (8.5)	67 (14.6)	
28β	**Glu**		52 (11.3)		N		67 (14.6)			4,6
	Asp	42 (9.1)		26 (5.7)	N	34 (7.4)		39 (8.5)	67 (14.6)	
30β	**His**				N		67 (14.6)			6,9
	Leu		52 (11.3)		N					
	Tyr	42 (9.1)		26 (5.7)	N	34 (7.4)		39 (8.5)	67 (14.6)	
37β	**Leu**		N		N		67 (14.6)	N	N	9
	Tyr	42 (9.1)	N	26 (5.7)	N	34 (7.4)		N	N	
38β	**Leu**						67 (14.6)			9
	Val	42 (9.1)	52 (11.3)	26 (5.7)	72 (15.7)	34 (7.4)		39 (8.5)	67 (14.6)	
57β	**Val**		52 (11.3)	N	72 (15.7)		67 (14.6)	N		9
	Asp	42 (9.1)		N		34 (7.4)		N	67 (14.6)	
85β	**Ala**						67 (14.6)			1
	Val	42 (9.1)	52 (11.3)	26 (5.7)	72 (15.7)	34 (7.4)		39 (8.5)	67 (14.6)	

Note: The risky amino acids at 9 polymorphic residue positions of HLA-DR β1 chain were shown in bold. The frequency (numbers) and percentage (in brackets) of HLA-DR β1 chain in risky and protective amino acids were listed. N, the HLA-DR β1 chain has an insignificant amino acid at this position.

individual in 230 subjects. Thirty-one polymorphic positions with two or more different amino acids were observed (**Table 2**). First, we numbered different amino acids at the polymorphic residues. Second, a logistic regression was performed for each polymorphic residue and all polymorphic residues were tested with a forward likelihood ratio test. The two lowest *P* values were detected for residues 57 and 30 ($P = 5.6\times10^{-4}$ and $P = 1.2\times10^{-3}$, respectively) which had the strongest associations with the outcomes of HBV infection. Other residues (14, 16, 25, 37, 38 and 85) were also significant ($P < 0.05$). Third, two-residue models were performed for additional effects on top of residues 57, 30, 14, 16, 25, 37, 38 and 85, respectively. No polymorphic residue had statistical significance with residue 57. The residues 16 and 28 ($P = 0.027$ and $P = 0.024$) had statistical significance with residue 30. Significant residues were consistent in the results of overall effect tests with residue 28. The residues 30 and 57 were found to contribute significantly to other residues. These data suggested that residues 57, 30, and 28 had more significant associations with the outcomes of HBV infection compared with other residues.

CH and RH-associated polymorphic residues 30 and 57 have more significant associations with the outcomes of HBV infection compared with other residues according to "Genotype" model

To confirm the validity of the results from the "Allele" model, we performed a stepwise logistic regression using the "Genotype" model in 230 subjects. In the "Genotype" model analyses, residue 30 remained to be CH-associated residues ($P = 0.0012$). Other residues (16, 25, 37, 38, 57 and 85) were also significant ($P < 0.05$) except for residue 14 ($P = 0.068$) and residue 28 ($P = 0.114$). In the analyses of two-residue models, none of the polymorphic residues had statistical significance on top of residue 30 (**Table 3**). These data indicated that residues 30 and 57 consistently had more significant associations with the outcomes of HBV infection irrespective of statistical models.

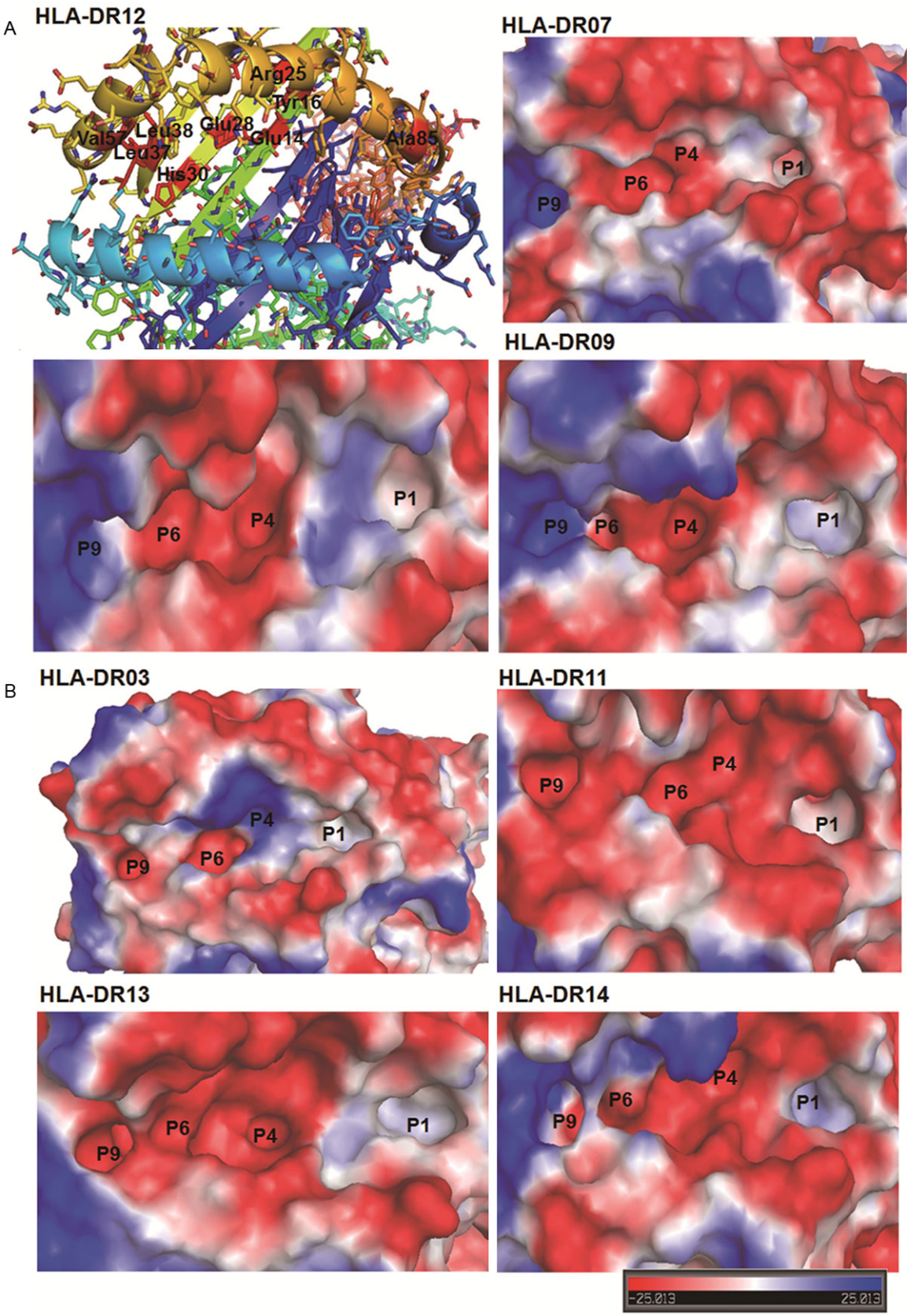
A
HLA-DR12
Arg25
Tyr16
Glu28
Glu14
Leu38
Val57
Leu37
His30
Ala85
HLA-DR07
P4
P1
P6
P9
HLA-DR09
P1
P4
P6
P9
P9
P6
P4
P1
B
HLA-DR03
P4
P1
P9
P6
HLA-DR11
P4
P9
P6
P1
HLA-DR13
P6
P4
P1
P9
HLA-DR14
P4
P6
P1
P9
-25.013
25.013

Figure 1. The electrostatic potential of HLA proteins. The electrostatic potential of (A) risky HLA proteins and (B) protective HLA proteins. The pockets P1, P4, P6 and P9 are labeled at appropriate locations. Potentials < -25 kT/e are shown in red, those > 25 kT/e are shown in blue and neutral potentials (0 kT/e) are shown in white. P1 pockets all present neutral charges. P4 pockets all present negative charges except for HLA-DRB1*030101. All of P6 pockets present negative charges. P9 pockets all present positive charges as shown in (A), and present negative charges as shown in (B). There are significant differences between risky and protective HLA proteins.

CH and RH-associated Lys14βGlu, Tyr16βHis, Gln25βArg, Glu28βAsp, His/Leu30βTyr, Leu37βTyr, Leu38βVal, Val57βAsp, Ala85βVal are related to the outcomes of HBV infection

To define the functions of amino acids at significantly polymorphic residues, we compared the rate of different amino acids at special residues between CH and RH. Residues 14, 25 and 85 were dimorphic. Lys14, Gln25 and Ala85 were significantly higher ($P < 0.05$; OR: 2.151-2.306; 95% CI > 1) and Glu14, Arg25 and Val85 were significantly lower ($P < 0.05$; OR: 0.434-0.465; 95% CI < 1) in CH compared with those in RH, and thus became risky and protective factors, respectively. Residues 16, 28, 30, 37, 38 and 57 were over dimorphic. Tyr16, Glu28, His/Leu30, Leu37, Leu38 and Val57 were significantly higher ($P < 0.05$; OR: 1.983-2.306; 95% CI > 1) and His16, Asp28, Tyr30, Tyr37, Val38 and Asp57 were significantly lower ($P < 0.05$; OR: 0.454~0.629; 95% CI < 1) in CH compared with those in RH, and thus became risky and protective factors, respectively. Other amino acids were unrelated to both CH and RH (**Table 4**). These data suggested that Lys14, Tyr16, Gln25, Glu28, His/Leu30, Leu37, Leu38, Val57 and Ala85 were risky amino acids and Glu14, His16, Arg25, Asp28, Tyr30, Tyr37, Val38, Asp57, Val85 were protective amino acids for chronic HBV infection.

HLA-DR12 and HLA-DR07 carry more risky amino acids than other HLA-DR

To define the HLA-DR molecules carrying risky and protective amino acids, we performed statistical analysis for the frequency and percentage of HLA-DR carrying risky and protective amino acids of 9 polymorphic residue positions (**Table 5**). In 230 subjects, only DR07 at the residue 14 of HLA-DR β1 chain carries the risky amino acid Lys (11.3%); DR08 and DR12 at residue 16 of HLA-DR β1 chain carry the risky amino acid Tyr (5.7% and 14.6%, respectively); only DR07 at residue 25 of HLA-DR β1 chain carries the risky amino acid Gln (11.3%); DR01, DR07, DR10, and DR12 at residue 28 of HLA-DR β1 chain carry the risky amino acid Glu (2.8%, 11.3%, 1.3% and 14.6%, respectively); DR07 and DR12 at residue 30 of HLA-DR β1 chain carry the risky amino acid Leu and His (11.3% and 14.6%, respectively); only DR12 at residues 37 and 38 of HLA-DR β1 chain carries the risky amino acid Leu (14.6%); DR07, DR09 and DR12 at residue 57 of HLA-DR β1 chain carry the risky amino acid Val (11.3%, 15.3% and 14.6%, respectively); and only DR12 at residue 85 of HLA-DR β1 chain carries the risky amino acid Ala (14.6%). At the 9 polymorphic residues, other HLA-DR β1 chains carry protective or insignificant amino acids. Therefore, HLA-DR12 β1 chain totally carries 7 risky amino acids; HLA-DR07 β1 chain totally carries 5 risky amino acids; and β1 chains of HLA-DR01, DR08, DR09 and DR10 have 1 risky amino acid, while other HLA-DR β1 chains have no risky amino acid. The percentage of HLA-DR carrying risky and protective amino acids of the 9 polymorphic residue positions was ≥ 5% (**Table 5**). These data indicated that HLA-DR12 and 07 confer susceptibility to chronic HBV infection.

The risky and protective amino acids at 9 polymorphic residues of HLA-DR β1 chain influence the electrostatic properties of HLA-DR pockets

In order to define the influence of the risk and protective amino acids carried by HLA-DR β1 chain on the electrostatic properties of HLA-DR pockets, we screened out HLA-DR03, DR07, DR09, DR11, DR12, DR13 and DR14 for further studies. Structure templates suitable for modeling were: DR52a_2Q6W_β for HLA-DR07 β1 chain, DR1_1AQD_β for HLA-DR09 β1 chain, DR4_1D5M_β for HLA-DR11 β1 chain, DR52c_3C5J_β for HLA-DR14 β1 chain and DR3_1A6A_β for HLA-DR12, DR03 and DR13 β1 chain. Three-dimensional structures of the analyzed HLA-DR peptide-binding groove were similar (only one structure was shown in **Figure 1**, and others are shown in **Table 5**). However, the special amino acids carried by different HLA-DR β1 chains influenced the electrostatic

properties of HLA-DR pockets (**Figure 1**). Residue 85 was located in pocket P1 that had neutral charge; Residues 14, 25 and 28 were located in pocket P4 that was mainly negatively charged, excepted for DR03 that had positively charged pocket P4; Residues 16, 28, and 30 were located in pocket P6 that was negatively charged; Residues 30, 37, 38 and 57 were located in pocket P9. For HLA-DR07, DR09, and DR12 that carry more risky amino acids (**Table 5**), pocket P9 was mainly positively charged. For HLA-DR03, DR11, DR13 and DR14 that carry protective amino acids (**Table 5**), pocket P9 was negatively charged. These data suggested that His/Leu30, Leu37, Leu38, Val57 in HLA-DR07, DR09, and DR12 β1 chain made pocket P9 positively charged and conferred susceptibility to chronic HBV infection, while Tyr30, Tyr37, Val38, Asp57 in HLA-DR03, DR11, DR13 and DR14 β1 chain made pocket P9 negatively charged and conferred protection against chronic HBV infection.

Discussion

In this study, we found that 9 residues had statistical significance in 31 polymorphic residues. Lys14 Glu, Tyr16 His, Gln25 Arg, Glu28 Asp, His/Leu30 Tyr, Leu37 Tyr, Leu38 Val, Val57 Asp and Ala85 Val had statistical difference between CH and RH group, and distinguished HLA-DR07, DR09, and DR12 from HLA-DR03, DR11, DR13, and DR14 according to their influence on the size and electrostatic properties of HLA-DR pockets.

In pocket P1, most of the residues are conserved, except for residues 85 and 86 [28]. Glyβ86Val can change the size of P1 and determine the selection of anchors [29, 30], but no difference exists between CH and RH. Alaβ85 that was only presented in HLA-DR12 had statistical significance. Although it did not influence the electrostatic potential of P1, Alaβ85, as a small residue, increased the size of P1 [31] and categorized HLA-DR12 into risk factors, suggesting that residue 85, on a certain extent, could influence the outcomes of HBV infection.

The β70Arg/71Arg/74Glu combination belonging to larger amino acids led to smaller pocket P4 of DR14 and DR09 [30, 32]. The attraction of Argβ71 by Gluβ28 left Aspβ70 alone at the rim of the pocket, led to a slightly acidic environment and increased the size of pocket P4 of DR12 and DR07. The pocket P4 of HLA-DR03 had a large positive surface produced by residues β70Gln/71Lys/74Arg [28], which was consistent with our result by surface electrostatic potential calculation. However, residues 70, 71 and 74 in our study had no statistical difference while residues 14, 25 and 28 had. Therefore, residue 28Glu might make DR12 and DR07 different from other alleles and hence, becoming a risk factor for the chronicity of HBV infection. However, the risk amino acids Lys and Gln carried by residues 14 and 25 were only presented in HLA-DR07. In addition, the whole sizes of Lys (135 $Å^3$) and Gln (114 $Å^3$) were similar to those of the protective amino acids Arg (148 $Å^3$) and Glu (109 $Å^3$). Only the electric charges of Gln (neutral charge) and Glu (negative charge) were different, making the negative charges of pocket P4 of DR07 lower than those of other alleles.

The bound peptide stability of pocket P6 was dominated by residues 71 and 30, and its specificity was dominated by β11. Tyrβ30 conferred a far more acidic character, weakened one partial positive charge and allowed less space in the pocket [29]. The acidic residues Gluβ11 and Aspβ66 can form salt bridges with a positively charged anchor, diminishing their mutual repulsion in parallel [29]. HLA-DR03, DR11, DR13 and DR14 were mainly identical at amino acids that comprised pocket P6, so their binding preferences should be the same. The P6 size of HLA-DR12 was increased, because Hisβ30 and Argβ71 were attracted to Gluβ28 but not Gluβ11 [31]. The structure of HLA-DR07 was similar with HLA-DR12, and they might have similar binding ability. Therefore, residues Gluβ28 and Hisβ30 might make DR12 and DR07 different from other alleles and hence, becoming risk factors for the chronicity of HBV infection. However, influence of residue 16 that also had statistical significance, was not described for the structure of pocket P6. At residue 16, Tyr of DR12 is a neutral amino acid rather than a more common positively charged His. Therefore, it was suggested that Tyr might increase negative charges of the pocket P6 of DR12.

The size and preferences of pocket P9 were determined by the salt bridge established between Aspβ57 and Argα76 that generated a small pocket [33]. However, other residues such as β9, β30, β37, β38, and β60 also made

important contributions to the preferences of this pocket. In our dataset, the residues 30, 37, 38 and 57 had statistical significance. HLA-DR03 and HLA-DR13 presented Tyrβ30, Asnβ37, Val38 and Aspβ57, while HLA-DR11 only differed at β37 (Tyrβ37). Although HLA-DR14 presented Alaβ57 and Pheβ37, which were different from other protective molecules, the standard hydrophobic residues Ileα72 and Metα73 made a smaller pocket P9 [30]. At the same time, they all had an electronegative pocket P9. However, HLA-DR07, DR09, and DR12 presented no Asp at residue 57, leading to the absence of an intact salt bridge. In addition, compared with HLA-DR03, His and Leu presented in residues 30, 37 and 38 of HLA-DR12, and Leu presented in residue 30 of HLA-DR07 were smaller. Therefore, they can produce larger and opener P9 [31, 32, 34]. In addition, the lack of Aspβ57 made the P9 pocket produce surplus positive charges around the free Argα76-favoring acidic residues[35], making the P9 pockets of HLA-DR07, DR09 and DR12 present positive charges.

In summary, our data indicated that the residues 14Glu, 16His, 25Arg, 28Asp, 30Tyr, 37Tyr, 38Val, 57Asp and 85Val, especially 30Tyr and 57Asp, in chronic HBV population were significantly enriched. We interpret these data to suggest that the residues 14, 16, 25, 28, 30, 37, 38, 57 and 85, especially 30 and 57, in the HLA-DR β chain critically influence the structural and electrostatic properties of the peptide-binding groove, and thus the range of peptides presented. In the future, we will take advantage of these achievements to develop novel HBV-specific epitopes.

Acknowledgements

This work was supported by the Science and Technology Project of Educational Department of Heilongjiang Province, China (No. 12521186).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Shuyun Zhang, Research Center, The Second Affiliated Hospital of Harbin Medical University, 246 Xuefu Road, Nangang District, Harbin 150086, Heilongjiang Province, P. R. China. Tel: +86-451-86664393; E-mail: Zhang13214501198@163.com

References

[1] Chinese Society of Hepatology and Chinese Society of Infectious Diseases, Chinese Medical Association. The guideline of prevention and treatment for chronic hepatitis B (2010 version). Zhonghua Liu Xing Bing Xue Za Zhi 2011; 32: 405-15.

[2] Ministry Of Health of the People's Republic of China. Diagnostic criteria for viral hepatitis B (ws 299-2008). Health industry standard of the people's republic of China 2008.

[3] Blackwell JM, Jamieson SE and Burgner D. HLA and infectious diseases. Clin Microbiol Rev 2009; 22: 370-85.

[4] Tamori A and Kawada N. HLA class II associated with outcomes of hepatitis B and C infections. World J Gastroenterol 2013; 19: 5395-401.

[5] Yan ZH, Fan Y, Wang XH, Mao Q, Deng GH and Wang YM. Relationship between HLA-DR gene polymorphisms and outcomes of hepatitis B viral infections: a meta-analysis. World J Gastroenterol 2012; 18: 3119-28.

[6] van Hattum J, Schreuder GM and Schalm SW. HLA antigens in patients with various courses after hepatitis B virus infection. Hepatology 1987; 7: 11-4.

[7] Singh R, Kaul R, Kaul A and Khan K. A comparative review of HLA associations with hepatitis B and C viral infections across global populations. World J Gastroenterol 2007; 13: 1770-87.

[8] Ramezani A, Hasanjani Roshan MR, Kalantar E, Eslamifar A, Banifazl M, Taeb J, Aghakhani A, Gachkar L and Velayati AA. Association of human leukocyte antigen polymorphism with outcomes of hepatitis B virus infection. J Gastroenterol Hepatol 2008; 23: 1716-21.

[9] Zhang SY, Gu HX, Li D, Yang SF, Zhong ZH and Li XK. Association of human leukocyte antigen polymorphism with hepatitis B virus infection and genotypes. Jpn J Infect Dis 2006; 59: 353-7.

[10] Li X, Liu W, Wang H, Jin X, Fang S and Shi Y. The influence of HLA alleles and HBV subgenotyes on the outcomes of HBV infections in Northeast China. Virus Res 2012; 163: 328-33.

[11] Nishida N, Sawai H, Matsuura K, Sugiyama M, Ahn SH, Park JY, Hige S, Kang JH, Suzuki K, Kurosaki M, Asahina Y, Mochida S, Watanabe M, Tanaka E, Honda M, Kaneko S, Orito E, Itoh Y, Mita E, Tamori A, Murawaki Y, Hiasa Y, Sakaida I, Korenaga M, Hino K, Ide T, Kawashima M, Mawatari Y, Sageshima M, Ogasawara Y, Koike A, Izumi N, Han KH, Tanaka Y, Tokunaga K and Mizokami M. Genome-wide association study confirming association of HLA-DP with protection against chronic hepatitis B and viral clear-

ance in Japanese and Korean. PLoS One 2012; 7: e39175.
[12] Mbarek H, Ochi H, Urabe Y, Kumar V, Kubo M, Hosono N, Takahashi A, Kamatani Y, Miki D, Abe H, Tsunoda T, Kamatani N, Chayama K, Nakamura Y and Matsuda K. A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population. Hum Mol Genet 2011; 20: 3884-92.
[13] Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG and Strominger JL. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 1993; 364: 33-9.
[14] Painter CA and Stern LJ. Conformational variation in structures of classical and non-classical MHCII proteins and functional implications. Immunol Rev 2012; 250: 144-57.
[15] Delgado JC, Baena A, Thim S and Goldfeld AE. Aspartic acid homozygosity at codon 57 of HLA-DQ beta is associated with susceptibility to pulmonary tuberculosis in Cambodia. J Immunol 2006; 176: 1090-7.
[16] Todd JA, Bell JI and McDevitt HO. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 1987; 329: 599-604.
[17] Nepom BS, Nepom GT, Coleman M and Kwok WW. Critical contribution of beta chain residue 57 in peptide binding ability of both HLA-DR and -DQ molecules. Proc Natl Acad Sci U S A 1996; 93: 7202-6.
[18] Hov JR, Kosmoliaptsis V, Traherne JA, Olsson M, Boberg KM and Bergquist A. Electrostatic modifications of the human leukocyte antigen-DR P9 peptide-binding pocket and susceptibility to primary sclerosing cholangitis. Hepatology 2011; 53: 1967-76.
[19] Jones EY, Fugger L, Strominger JL and Siebold C. MHC class II proteins and disease: a structural perspective. Nat Rev Immunol 2006; 6: 27-82.
[20] Lee KH, Wucherpfennig KW and Wiley DC. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. Nat Immunol 2001; 2: 501-7.
[21] Menconi F, Osman R, Monti MC, Greenberg DA, Concepcion ES and Tomer Y. Shared molecular amino acid signature in the HLA-DR peptide binding pocket predisposes to both autoimmune diabetes and thyroiditis. Proc Natl Acad Sci U S A 2010; 107: 16899-903.
[22] Cao T, Desombere I, Vanlandschoot P, Sällberg M and Leroux-Roels G. Characterization of HLA DR13-restricted CD4(+) T cell epitopes of hepatitis B core antigen associated with self-limited, acute hepatitis B. J Gen Virol 2002; 83: 3023-33.
[23] Diepolder HM, Jung MC, Keller E, Schraut W, Gerlach JT, Grüner N, Zachoval R, Hoffmann RM, Schirren CA, Scholz S and Pape GR. A vigorous virus-specific CD4+ T cell response may contribute to the association of HLA-DR13 with viral clearance in hepatitis B. Clin Exp Immunol 1998; 113: 244-51.
[24] Thompson JD, Higgins DG and Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-80.
[25] Pandit SB, Zhang Y and Skolnick J. TASSER-Lite: an automated tool for protein comparative modeling. Biophys J 2006; 91: 4180-90.
[26] Bramucci E, Paiardini A, Bossa F and Pascarella S. PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. BMC Bioinformatics 2012; 13: S2.
[27] Konecny R, Baker NA and McCammon JA. iAPBS: a programming interface to Adaptive Poisson-Boltzmann Solver (APBS). Comput Sci Discov 2012; 5.
[28] Balbín A, Cárdenas C, Villaveces JL and Patarroyo ME. A theoretical analysis of HLA-DRbeta1*0301-CLIP complex using the first three multipolar moments of the electrostatic field. Biochimie 2006; 88: 1307-11.
[29] Verreck FA, van de Poel A, Drijfhout JW, Amons R, Coligan JE and Konig F. Natural peptides isolated from Gly86/Val86-containing variants of HLA-DR1, -DR11, -DR13, and -DR52. Immunogenetics 1996; 43: 392-7.
[30] James EA, Moustakas AK, Berger D, Huston L, Papadopoulos GK and Kwok WW. Definition of the peptide binding motif within DRB1*1401 restricted epitopes by peptide competition and structural modeling. Mol Immunol 2008; 45: 2651-9.
[31] Chow IT, James EA, Tan V, Moustakas AK, Papadopoulos GK and Kwok WW. DRB1*12:01 presents a unique subset of epitopes by preferring aromatics in pocket 9. Mol Immunol 2012; 50: 26-34.
[32] James EA, Moustakas AK, Bui J, Nouv R, Papadopoulos GK and Kwok WW. The binding of antigenic peptides to HLA-DR is influenced by interactions between pocket 6 and pocket 9. J Immunol 2009; 183: 3249-58.
[33] Agudelo WA, Galindo JF and Patarroyo ME. Electrostatic potential as a tool to understand interactions between malaria vaccine candidate peptides and MHCII molecules. Biochem Biophys Res Commun 2011; 410: 410-5.
[34] Chicz RM, Urban RG, Gorga JC, Vignali DA, Lane WS and Strominger JL. Specificity and

promiscuity among naturally processed peptides bound to HLA-DR alleles. J Exp Med 1993; 178: 27-47.

[35] Nepom BS, Nepom GT, Coleman M and Kwok WW. Critical contribution of beta chain residue 57 in peptide binding ability of both HLA-DR and -DQ molecules. Proc Natl Acad Sci U S A 1996; 93: 7202-6.