Electronic Supplementary Material (ESI) for RSC Advances.

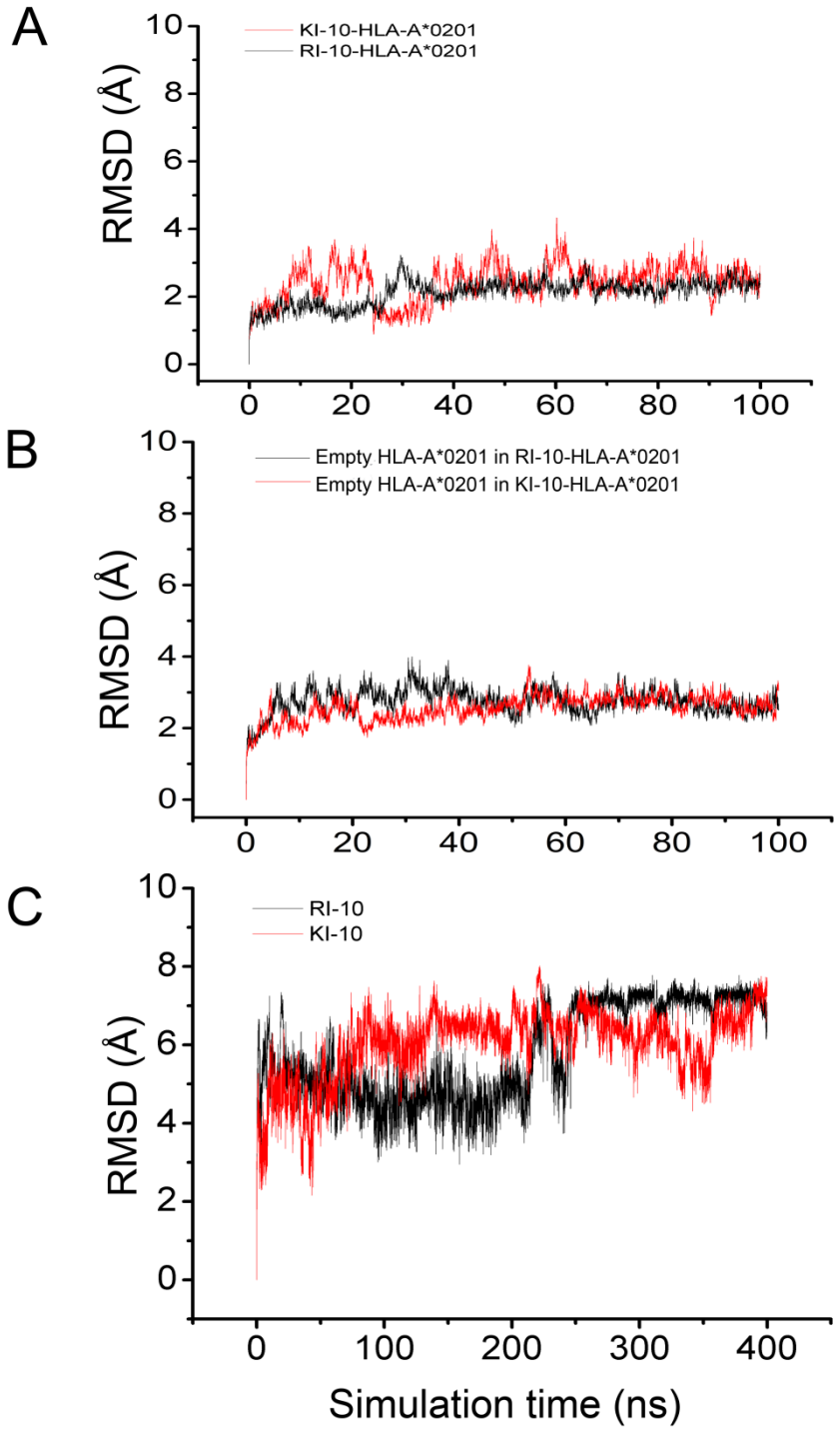

Figure S1. RMSD of HLA-A*0201 in complexes with RI-10/KI-10 peptides, empty HLA-A*0201 molecules and the free peptides in their respective MD simulation systems. (A) RMSD of the protein (including all non-hydrogen atoms) of RI-10-HLA-A*0201 and RI-10-HLA-A*0201 complexes. (B) RMSD of all non-hydrogen atoms of the empty HLA-A*0201 molecules without the peptides. (C) RMSD of all non-hydrogen atoms of the free RI-10 and KI-10 peptides.

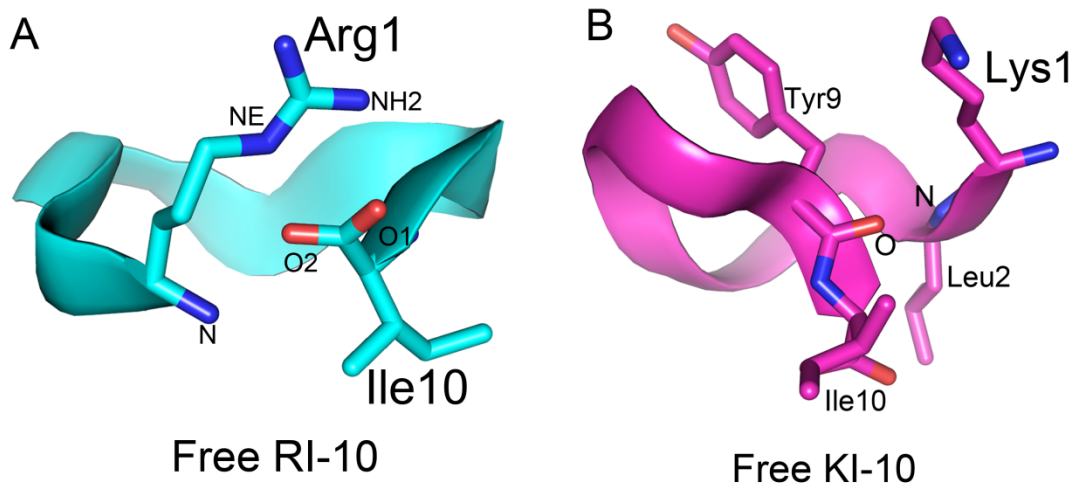

Figure S2. The representative structures of the free peptides in MD simulations. (A) The representative structure (the structures most similar to the average structure) of free RI-10 in the last 150 ns of its MD simulation trajectory. (B) The representative structures of free KI-10 peptide in the last 150 ns of its MD simulation trajectory. The dash lines indicates salt bridges (red) or hydrogen bonds (yellow).

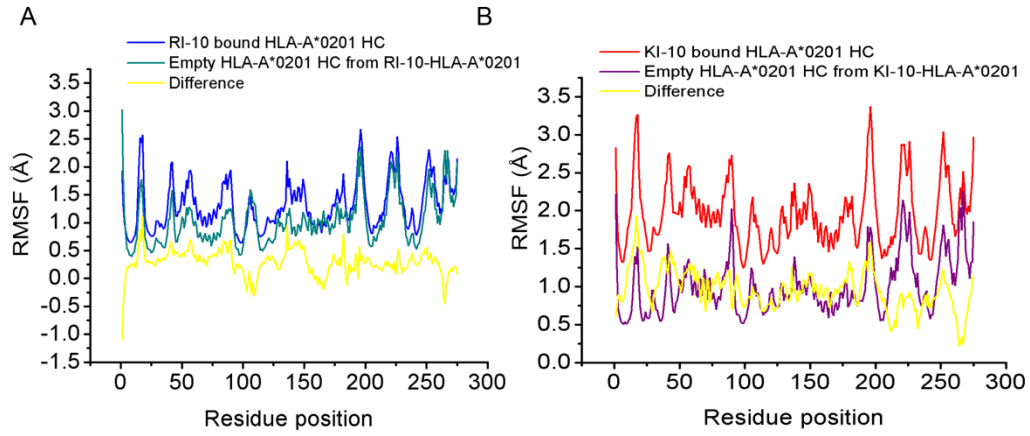

Figure S3. Comparison of the RMSF of individual residues in the MD trajectories of peptide-bound HLA-A*0201 HC and the empty HLA-A*0201 HC extracted from the peptide-HLA-A*0201 complexes. (A) RMSF of the RI-10 bound HLA-A*0201 HC, the empty HLA-A*0201 extracted from RI-10-HLA-A*0201 and their difference. (B) RMSF of KI-10 bound HLA-A*0201 HC, the empty HLA-A*0201 extracted from KI-10-HLA-A*0201, and their difference. The dashed lines in (A) and (B) represent the average difference of all the 275 residues in HLA-A*0201 HC.

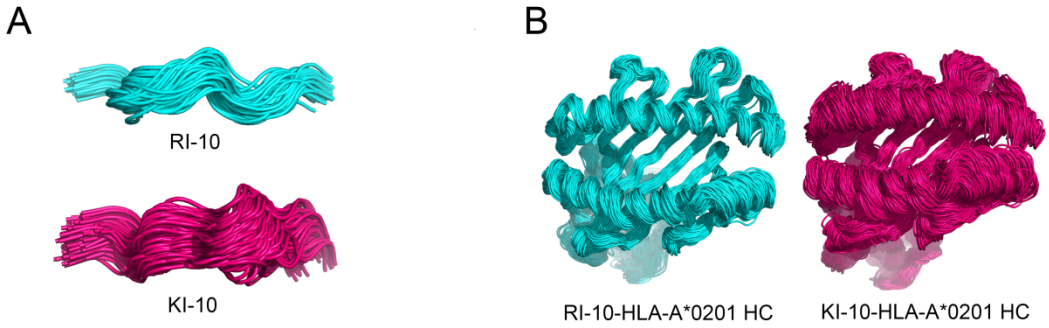

Figure S4. Conformational bundles of the peptides and HLA-A*0201 HC during MD simulations. (A) Conformational bundles of RI-10 and KI-10 peptides representing the 150 ns concatenated MD simulations of the RI-10- and KI-10-HLA-A*0201 complexes. (B) Conformational bundles of HLA-A*0201 HC representing the 150 ns concatenated MD simulations of the RI-10- and KI-10-HLA-A*0201 complexes. For both (A) and (B), all the conformational bundles contain 200 snapshots, sampled every 750 ps from the 150 ns trajectories.

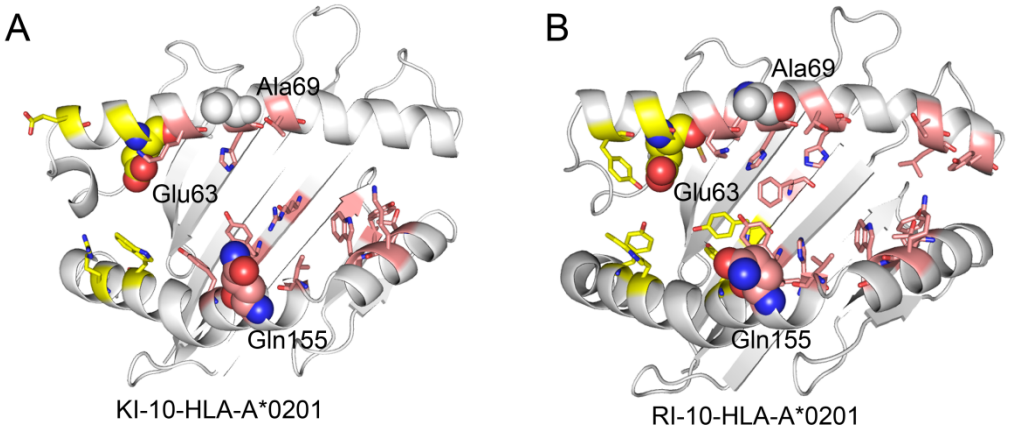

Figure S5. The residue positions of the peptide binding sites and the potential TCR recognition sites in HLA-A*0201 HC. (A) The residue positions of the peptide binding sites and the potential TCR recognition sites in the representative structure (the structures most close to the average structures) of KI-10-HLA-A*0201 for its 150 ns concatenated MD trajectory. (B) The residue positions of the peptide binding sites and the potential TCR recognition sites in the representative structure of RI-10-HLA-A*0201 for its 150 ns concatenated MD trajectory. The peptide binding and potential TCR recognition residues in HLA-A*0201 HC are shown as stick. The residues in HLA-A*0201 HC interacting with the first residue of the peptides are indicated as yellow, while those interacting with the other 9 resides of the peptides (residues 2-10) are indicated as wheat. The potential TCR recognition sites (the “restriction triad”, Glu63, Ala69, and Gln155) are shown as sphere.

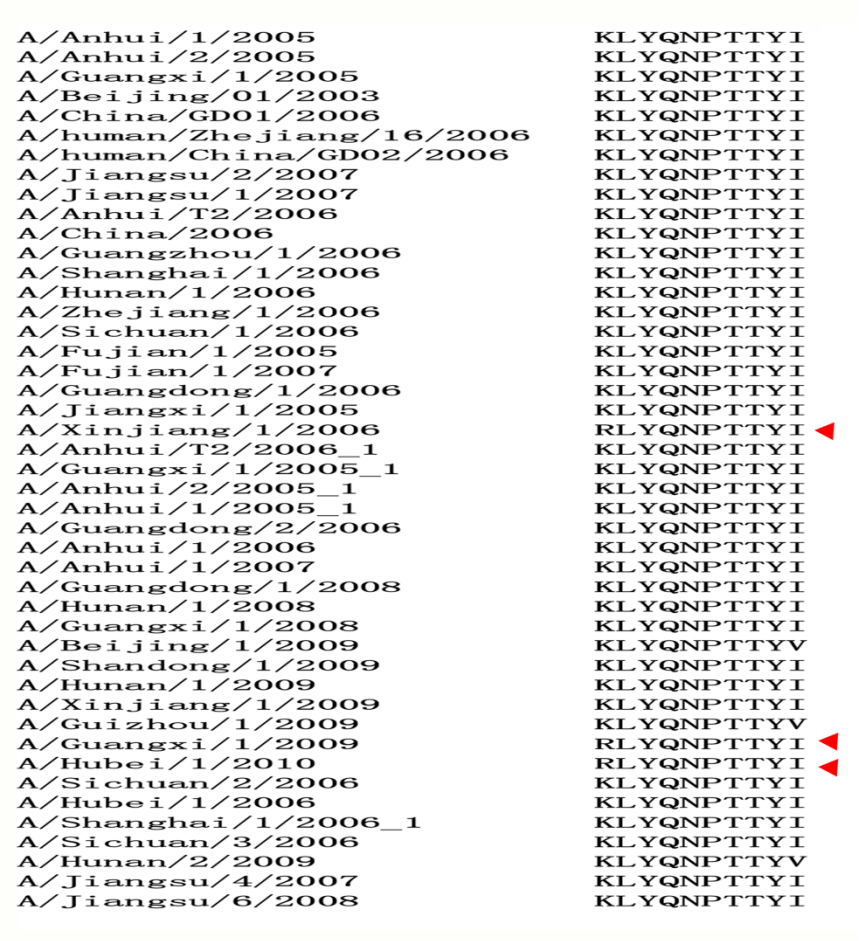

Figure S6. Sequence alignment of hemagglutinins (HA) from all 45 Chinese human H5N1 isolates at the position of RI-10/KI-10. The HA sequences of 45 Chinese human H5N1 isolates containing RI-10 or KI-10 are aligned and RI-10 is indicated with red arrows. The sequence data were extracted from Influenza Research Database (http://www.fludb.org/brc/home.spg?decorator=influenza).

Table S1. MM/PBSA binding free energy (kJ/mol) calculations for the binding of RI-10 and KI-10 to HLA-A*0201 by STM without considering entropy

	Energy terms	Bond[1]	Angle[2]	Dihed[3]	Impr[4]	Elec[5]	Vdw[6]	Ps[7]	Nps[8]	Total[9]
RI-10	$G_{complex}$	5385.95 (126.06)[10]	14293.69 (193.31)	7087.31 (100.35)	872.87 (43.19)	-41066.64 (777.05)	-6088.86 (152.96)	-6915.23 (97.74)	438.79 (3.81)	-25992.14 (588.15)
	$G_{receptor}$	5234.83 (123.75)	13862.21 (190.02)	6791.65 (100.51)	850.50 (42.60)	-39231.03 (695.74)	-5878.67 (141.75)	-7128.23 (125.33)	433.35 (5.12)	-25065.38 (557.92)
	G_{ligand}	156.56 (21.71)	421.07 (33.18)	219.43 (19.63)	21.92 (6.57)	-1196.56 (283.17)	33.11 (2.31)	-413.7 (19.19)	37.84 (1.27)	-720.33 (25.45)
	ΔG_{bind}	-5.46	10.41	76.23	0.45	-639.05	-243.30	626.7	-32.4	-180.55
KI-10	$G_{complex}$	5382.69 (127.07)	14313.97 (189.70)	7073.91 (106.02)	873.94 (43.14)	-40143.01 (758.46)	-6182.84 (143.09)	-6915.54 (84.55)	438.79 (2.87)	-25158.08 (239.24)
	$G_{receptor}$	5218.07 (128.11)	13828.86 (190.28)	6837.48 (95.50)	850.15 (42.72)	-39638.42 (540.04)	-6005.48 (131.66)	-6978.39 (112.87)	436.82 (4.01)	-25450.91 (189.21)
	G_{ligand}	149.70 (20.58)	447.85 (34.27)	232.47 (95.50)	21.40 (6.72)	-165.98 (9.80)	34.19 (20.52)	-440.33 (18.7)	36.43 (1.26)	315.73 (13.66)
	ΔG_{bind}	14.92	37.26	3.96	2.39	-338.61	-211.55	503.18	-34.46	-22.90

[1]bond energy term.
[2]angle energy term.
[3]dihedral angle energy term.
[4]Impr energy term.
[5]coulombic term.
[6]polar solvation term.
[7]non-polar solvation term.
[8]Van der Waals term.
[9]the sum of the energy terms of 1 to 8.
[10]the value in parentheses are standard errors

Table S2. Entropy (kJ/mol K) calculation and the binding free energy (kJ/mol K) of RI-10 and KI-10 to HLA-A*0201

	Complex	ligand	receptor	ΔS	$^{*}\Delta G_{bind}$'
RI-10	69.11	4.32	65.03	-0.24	-106.15
KI-10	69.49	4.58	65.07	-0.11	11.2

*Binding free energy. $^{*}\Delta G_{bind}$'=ΔG_{bind} - TΔS, ΔG_{bind} is given in table S1, T=310 K.

Table S3. Hydrogen bonds formed between RI-10 or KI-10 and HLA-A*0201 during the MD simulations[1]

RI-10-HLA-A*0201					KI-10-HLA-A*0201				
Peptide		**HLA-A*0201 HC**		**Occupancy (%)[2]**	**Peptide**		**HLA-A*0201 HC**		**Occupancy (%)[2]**
Arg1	O	Tyr159	OH	79.4	**Lys1**	O	Tyr159	OH	31.4
	N	Glu63	OE2	20.4		N	Glu63	OE1	32.3
	N	Glu63	OE1	13.3					
	N	Tyr171	OH	16.8					
	N	Tyr7	OH	12.3					
Leu2	N	Glu63	OE1	23.8	**Leu2**	N	Glu63	OE2	17.7
	N	Glu63	OE2	11.5					
Try3	N	Try99	OH	85.9	**Try3**				
Gln4	O	Lys66	NZ	27.3	**Gln4**				
Asn5					**Asn5**	OD1	Arg97	NH1	10.6
Thr7					**Thr7**	OG1	Thr73	OG1	11.6
Thr8	OG1	Trp147	NE1	41.2	**Thr8**				
Tyr9	O	Trp147	NE1	87.2	**Tyr9**	O	Lys146	NZ	12.6
ILE10	OT2	Thr143	OG1	54.4	**ILE10**	OT2	Thr143	OG1	45.9
	OT1	Tyr84	OH	22.6		OT1	Tyr84	OH	32.0
	OT1	Lys146	NZ	24.4		OT1	Lys146	NZ	16.8
	OT2	Lys146	NZ	42.2		OT2	Lys146	NZ	15.9
	N	Asp77	OD2	46.9					
	OT2	Tyr84	OH	11.3					

[1]The concatenated 150 ns trajectories of RI-10-HLA-A*0201 and K I-10-HLA-A*0201 complexes were used for the hydrogen bond analysis.

[2]Occupancy of a hydrogen bond is defined as the percentage of the number of snapshots in which the hydrogen bond formed in the total number of the snapshots extracted from the simulation trajectories. The standard for the selection of a hydrogen bond is that the distance between the donor atom and the receptor atom is < 3.5 Å, and that the angle formed by the donor atom, the hydrogen atom, and the receptor atom is <30 °. Hydrogen bonds with an occupancy > 10 % are shown.