Human Vaccines

ISSN: 1554-8600 (Print) 1554-8619 (Online) Journal homepage: www.tandfonline.com/journals/khvi19

Highly conserved influenza A sequences as T cell epitopes-based vaccine targets to address the viral variability

Paul ThiamJoo Tan, Asif M. Khan & J. Thomas August

To cite this article: Paul ThiamJoo Tan, Asif M. Khan & J. Thomas August (2011) Highly conserved influenza A sequences as T cell epitopes-based vaccine targets to address the viral variability, Human Vaccines, 7:4, 402-409, DOI: 10.4161/hv.7.4.13845

To link to this article: https://doi.org/10.4161/hv.7.4.13845

Published online: 01 Apr 2011.

Submit your article to this journal

Article views: 650

View related articles

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=khvi20

Human Vaccines 7:4, 402-409; April 2011; © 2011 Landes Bioscience

Highly conserved influenza A sequences as T cell epitopes-based vaccine targets to address the viral variability

Paul ThiamJoo Tan,[1,*] Asif M. Khan[1,2] and J. Thomas August[1]

[1]Department of Pharmacology and Molecular Sciences; School of Medicine; Johns Hopkins University; Baltimore, MD USA;
[2]Department of Biochemistry; Yong Loo Lin School of Medicine; National University of Singapore; Singapore

Key words: T cell epitopes-based vaccine, pan-influenza, immunoinformatics, altered peptide ligands, high conservation, immunodominance, pandemic H1N1

Vaccines are the only proven effective method for prevention of human infectious diseases. Almost all traditional vaccines require activating immunological memory B cells to secrete neutralizing antibodies against invading pathogens. The complication with influenza viruses is the high viral mutation rate that results in immune escape through modification of the B cell epitopes. Studies of T-cell immunity to influenza infection provide an alternative vaccine strategy based on highly conserved T-cell epitopes. In this review, we discuss the importance of T cell-mediated immunity in influenza infection and the need for a targeted vaccine approach focused on highly conserved T-cell epitopes to mitigate immune escape. We propose 15 highly conserved pan-influenza sequences as potential T cell epitopes-based vaccine targets for broad protection and lasting immunity against variant influenza strains.

Introduction

The emergence and rapid spread of the 2009 pandemic novel swine-origin influenza H1N1 strain has resulted in a renewed focus on the development of an alternative influenza vaccine that is not subjected to annual update. Current influenza vaccines involve the production of neutralizing antibodies against the viral structural glycoproteins, namely hemagglutinin (HA) and neuraminidase (NA), by immunizing with inactivated or live attenuated influenza viruses.[1] However, frequent mutations in these glycoproteins allow new influenza virus strains to escape the antibody-mediated immunity conferred by the vaccine. Thus, current influenza vaccines need to be updated annually to match the new circulating strains. Further, there is at least a six-month delay after selection of strains to produce, test and distribute the new influenza vaccines, leaving the general population vulnerable to infection and disease. For example, the delay in vaccine production against the novel 2009 H1N1 strain resulted in a worldwide concern at the height of the pandemic. Another drawback of current influenza vaccines is that circulating strains may mutate and differ from the selected vaccine strains during or after the production phase, thus possibly reducing their effectiveness against influenza outbreaks.[2,3]

In contrast to the highly mutable HA and NA surface glycoproteins, the nonstructural internal proteins of influenza, namely matrix protein 1 (M1), nucleoprotein (NP) and polymerase proteins (PB1, PB2 and PA), are more conserved. Furthermore, we have previously reported that some sequences of nine or more amino acids (aa) were found to be evolutionarily highly conserved among avian and human influenza A viruses that had been recorded over the past 30 years.[4] These conserved sequences are a potential source of T-cell epitopes that would be common to almost all influenza A viruses and provide a basis for an alternative influenza vaccine strategy possibly effective against a broad spectrum of heterosubtypic influenza strains. Herein, we discuss our research on identifying and characterizing selected highly conserved sequences as targets for T cell epitopes-based influenza vaccine.

T-Cell Epitopes are Necessary Elements for the Control and Clearance of Influenza Infection

The immune response to influenza virus infection is characterized by viral activation of B cells and two types of T cells: $CD4^+$ helper T lymphocytes (HTLs) and $CD8^+$ cytotoxic T lymphocytes (CTLs). Helper and cytotoxic T lymphocytes mediate cellular immune responses via their T-cell receptors (TCR) that recognize T-cell epitopes presented by human leukocyte antigen (HLA) molecules. HLA class I molecules of most nucleated cells present epitopes of 8–11 aa to CTLs that kill the infected cells.[5,6] The HLA class I epitopes are derived chiefly from endogenous antigens that are processed by the cytoplasmic proteasomal system. HLA class II molecules are expressed in professional antigen presenting cells, such as dendritic cells, B cells and macrophages and present mainly exogenous antigens that are endocytosed by cell surface pattern recognition receptors and processed by endosomal/lysosomal proteases to epitope peptides of 12–25 aa. These HLA class II epitopes are presented to HTLs, which upon activation, produce secondary signals that in turn activate CTLs and B cells.[7-9]

*Correspondence to: Paul ThiamJoo Tan; Email: paultantj@gmail.com
Submitted: 07/02/10; Revised: 09/17/10; Accepted: 10/03/10
DOI: 10/4161/hv.7.4.13845

Table 1. Percentage representations of the 55 highly conserved influenza A sequences among 2009 pandemic H1N1 strains

Protein	Highly conserved sequences[a]	2009 H1N1[b]
PB2	$_{10}$LMS QSR TRE ILT KTT VDH MAI IKK YTS GRQ EKN P$_{43}$ (34)	97.17% [15]
	$_{45}$LRM KWM MAM KYP ITA DKR I$_{63}$ (19)	**0.12%** [11]
	$_{68}$PER NEQ GQT LWS K$_{80}$ (13)	99.13% [7]
	$_{92}$SPL AVT WWN RNG P$_{104}$ (13)	99.08% [10]
	$_{121}$KVE RLK HGT FGP VHF RNQ VKI RRR VD$_{146}$ (26)	98.21% [13]
	$_{228}$YIE VLH LTQ GTC WEQ MYT PGG EV$_{250}$ (23)	99.19% [7]
	$_{252}$NDD VDQ SLI IAA RNI VRR A$_{270}$ (19)	98.85% [12]
	$_{278}$ASL LEM CHS TQI GG$_{291}$ (14)	99.54% [3]
	$_{345}$LTG NLQ TLK$_{353}$ (9)	99.37% [5]
	$_{356}$VHE GYE EFT MVG$_{367}$ (12)	99.65% [7]
	$_{369}$RAT AIL RKA TRR$_{380}$ (12)	99.88% [3]
	$_{400}$VAM VFS QED CM$_{410}$ (11)	98.90% [12]
	$_{412}$KAV RGD LNF VNR ANQ RLN PMH QLL RHF QKD AKV LF$_{446}$ (35)	98.96% [11]
	$_{466}$DMT PST EMS$_{474}$ (9)	97.4% [11]
	$_{479}$RVS KMG VDE YS$_{489}$ (11)	98.79% [5]
	$_{509}$GNV LLS PEE VSE TQG$_{523}$ (15)	98.73% [5]
	$_{527}$LTI TYS SSM MWE ING PES VL$_{546}$ (20)	99.6% [3]
	$_{548}$NTY QWI IRN WE$_{558}$ (11)	99.65% [3]
	$_{570}$MLY NKM EFE PFQ SLV PKA$_{587}$ (18)	99.6% [5]
	$_{614}$QII KLL PFA AAP P$_{626}$ (13)	99.88% [3]
	$_{628}$QSR MQF SSL TVN VRG SGM RIL$_{648}$ (21)	**0%** [6]
	$_{685}$GVE SAV LRG FLI$_{696}$ (12)	99.83% [3]
	$_{703}$RYG PAL SIN$_{711}$ (9)	99.6% [7]
PB1	$_{1}$MDV NPT LLF LKV P$_{13}$ (13)	**0.23%** [5]
	$_{15}$QNA IST TFP YTG DPP YSH GTG TGY TMD TVN RTH QYS E$_{51}$ (37)	99.42% [8]
	$_{114}$VQQ TRV DKL TQG RQT YDW TLN RNQ PAA TAL ANT IE$_{148}$ (35)	99.31% [7]
	$_{196}$TKK MVT QRT IGK KK$_{209}$ (14)	99.6% [8]
	$_{337}$LSI API MFS NKM ARL GKG YMF ESK$_{360}$ (24)	**0%** [7]
PB1		
	$_{362}$MKL RTQ IPA EML A$_{374}$ (13)	**0%** [5]
	$_{474}$GIN MSK KKS YIN$_{485}$ (12)	99.08% [2]
	$_{487}$TGT FEF TSF FYR YGF VAN FSM ELP SFG VSG$_{516}$ (30)	99.65% [7]
	$_{518}$NES ADM SIG VTV IKN NMI NND LGP ATA QMA LQL FIK DYR YTY RCH RGD TQI QTR RSF E$_{575}$ (58)	92.8% [20]
	$_{655}$MEY DAV ATT HSW$_{666}$ (12)	99.31% [7]
	$_{668}$PKR NRS ILN TSQ RGI LED EQM YQ$_{690}$ (23)	99.31% [10]
PA	$_{29}$KIE TNK FAA ICT HLE VCF MYS DFH FI$_{54}$ (26)	99.05% [10]
	$_{130}$YYL EKA NKI KSE$_{141}$ (12)	99.66% [6]
	$_{143}$THI HIF SFT GEE MA$_{156}$ (14)	99.83% [3]
	$_{185}$RGL WDS FRQ SER GEE TIE E$_{203}$ (19)	**0%** [5]
	$_{298}$HEG EGI PLY DAI KC$_{311}$ (14)	99.44% [6]
	$_{412}$EFN KAC ELT DS$_{422}$ (11)	99.83% [3]
	$_{560}$SRP MFL YVR TNG TSK$_{574}$ (15)	99.33% [10]
HA	$_{388}$FGA IAG FIE$_{396}$ (9)	99.73% [5]
NP	$_{1}$MAS QGT KRS YEQ MET$_{15}$ (15)	99.57% [5]
	$_{35}$GIG RFY IQM CTE LKL$_{49}$ (15)	99.95% [2]
	$_{66}$MVL SAF DER RN$_{76}$ (11)	99.89% [2]
	$_{78}$YLE EHP SAG KDP KKT GGP IY$_{97}$ (20)	99.58% [7]
	$_{110}$LYD KEE IRR IWR QAN NG$_{126}$ (17)	**0.05%** [8]
	$_{137}$MIW HSN LND$_{145}$ (9)	99.95% [2]
	$_{241}$DQV RES RNP GNA EIE DL$_{257}$ (17)	99.68% [5]
	$_{410}$QPT FSV QRN LPF$_{421}$ (12)	99.89% [2]
	$_{461}$GRG VFE LSD E$_{470}$ (10)	100% [1]
M1	$_{1}$MSL LTE VET YVL SI$_{14}$ (14)	99.75% [3]
	$_{122}$GAL ASC MGL IYN RMG$_{136}$ (15)	99.9% [3]
	$_{175}$HEN RMV LAS TTA KAM EQM AGS SEQ AAE AME$_{204}$ (30)	99.06% [12]
	$_{208}$QAR QMV QAM R$_{217}$ (10)	**0%** [2]

[a]Highly conserved sequences are defined as sequences of nine or more amino acids with a minimal representation of 80% within database sequences of human H1N1, H1N2, H3N2, H5N1, avian H5N1 and other avian subtypes that were identified by Heiny et al. (2007). The length of each sequence is indicated in parentheses. [b]The percentage representation was calculated as the number of 2009 pandemic sequences that matched the highly conserved sequences divided by the total number of sequences in the same aligned positions. The percentage representations of eight highly conserved sequences that were not prevalent in the 2009 pandemic strains are in boldface. The numbers in square brackets indicate the total number of unique 2009 pandemic peptides at the aligned positions. Variant peptides at an aligned position are defined as sequences having at least one amino acid different from the highly conserved sequence present at the position. The numbers of 2009 pandemic H1N1 sequences extracted from the NCBI Influenza Virus Resource as of June 2010 were: 1734 PB2, 1738 PB1, 1788 PA, 2607 HA, 1896 NP and 2019 M1.

Numerous studies have shown that CTL epitopes derived from the internal proteins M1 and NP provided partial[10,11] or complete[12,13] protection of mice from heterosubtypic influenza viruses. HTLs play several roles in influenza infection; for example, cytokine secretion to recruit and activate other immune cells in response to viral infection,[14] perforin-mediated cytolysis of virus-infected cells[15] and stimulation of antibody production by B cells,[16] including antibody class switching.[17] In the absence of concurrent help from HTLs, clearance of residual virus in mice was delayed and recall of CD8$^+$ T-cell memory was compromised during secondary response.[18] Detailed reviews of interactions between CTLs and HTLs, and the responses of these cells to influenza infection are available.[19-23]

Influenza Virus Immune Escape

Influenza viruses escape the host immune system by mutations to both surface glycoproteins and internal proteins. This immune

Table 2. Low representations of eight highly conserved influenza A sequences among the 2009 pandemic H1N1 strains

Protein	Variability of highly conserved sequences among 2009 H1N1 [a]	No. of peptides	% representation [b]
PB2 45-63	LRMKWMMAM**KYPI**T**ADKRI**	2	0.12%
	.........R.........	1714	98.85%
	.........R.....N...	6	0.35%
	.........R......R..	3	0.17%
	.........RH........	2	0.12%
	.........R.S.......	2	0.12%
	.........R......E..	1	0.06%
	.........R.......K.	1	0.06%
	.........R........V	1	0.06%
	.........R....G....	1	0.06%
	.........R..V......	1	0.06%
PB2 628-648	QS**RM**QFSSL**T**VNVRGSG**M**RI**L**	0	0%
	.................L...	1728	99.71%
	..G..............L...	1	0.06%
	...I.............L...	1	0.06%
	...L.............L...	1	0.06%
	.........A.......L...	1	0.06%
	.................L..M	1	0.06%
PB1 1-13	MDVNPTLLF**L**K**V**P	4	0.23%
	...........I.	1727	99.42%
	...........T.	3	0.17%
	...........L.	2	0.12%
	.........V.I.	1	0.06%
PB1 337-360	LS**I**AP**I**MFSNKM**AR**LG**K**GYMFESK	0	0%
	..M.....................	1620	93.32%
	..M.............R.......	110	6.34%
	..M..........K..........	2	0.12%
	..M.........T...........	1	0.06%
	..M..V..................	1	0.06%
	..M..V..........R.......	1	0.06%
	..T.....................	1	0.06%
PB1 362-374	MK**L**RTQIP**AE**M**L**A	0	0%
	..I..........	1733	99.77%
	..I........I.	1	0.06%
	..I......K...	1	0.06%
	..I.....G....	1	0.06%
	..I.....S....	1	0.06%
PA 185-203	R**G**LW**D**SFRQSERGEET**I**EE	0	0%
	.S.................	1779	99.72%
	.S..............V..	2	0.11%
	.N.................	1	0.06%
	.S..............T..	1	0.06%
	.S..E..............	1	0.06%
NP 110-126	L**Y**DKEE**IRRI**WRQAN**N**G	1	0.05%
	.........V.......	1879	99.42%
	........KV.......	4	0.21%
	.......K.V.......	2	0.11%
	.C.......V.......	1	0.05%
	.........V.....D.	1	0.05%
	......L..V.......	1	0.05%
	......V..V.......	1	0.05%
M1 208-217	Q**AR**QMV**Q**AMR	0	0%
	.T....H...	2012	99.70%
	.TK...H...	6	0.3

[a]Residues in boldface and underlined were prevalent among human H1N1, H1N2, H3N2, H5N1, avian H5N1 and other avian subtypes but not in the 2009 pandemic H1N1 strains (as of June 2010). The amino acids substituted in the pandemic strains are reflected and identical residues are represented by periods. [b]The percentage representation was calculated as the number of identical peptide sequences divided by the total number of peptide sequences analyzed in the same aligned positions.

escape is sometimes the result of a misdirected response by the immune system. For example, antibody response upon infection by a new influenza variant strain is directed to a previous strain. This phenomenon, termed original antigenic sin,[24,25] has also been observed for T cells: memory CTLs trigger a previous T-cell epitope response as opposed to the new variant epitope, leading to impaired clearance of variant viruses infecting the same individual.[26,27] Therefore, in subsequent variant strain infection, the positions and types of amino acid substitutions surrounding or within T-cell epitopes can result in a combination of events. Mutation at flanking residues of T-cell epitopes can influence antigen processing and abolish presentation.[28-31] Within a T-cell epitope, a single amino acid substitution can affect interactions of the other residues to HLA or TCR molecules because of extensive inter-residue T-cell epitope interactions.[32-34] Thus, variations at anchor residue positions can abolish or diminish binding affinity

Table 3. Fifteen selected highly conserved influenza A sequences as T cell epitopes-based influenza vaccine targets

Protein	Selected highly conserved sequences[a]	2009 H1N1 %[b]	Functional domain[d]
PB2	$_{10}$ LMS QSR TRE ILT KTT VDH MAI IKK YTS GRQ EKN P $_{43}$ (34)	97	1-124:PB1[72,73] 1-269:NP[74]
	$_{121}$ KVE RLK HGT FGP VHF RNQ VKI RRR VD $_{146}$ (26)	98	1-269:NP[74]
	$_{228}$ YIE VLH LTQ GTC WEQ MYT PGG EVR NDD VDQ SLI IAA RNI VRR A $_{270}$ (43)	96	1-269:NP[74]
	$_{345}$ LTG NLQ TLK IRV HEG YEE FTM VG $_{367}$ (23)	99	-
	$_{400}$ VAM VFS QED CMI KAV RGD LNF VNR ANQ RLN PMH QLL RHF QKD AKV LF $_{446}$ (47)	98	-
	$_{527}$ LTI TYS SSM MWE ING PES VLV NTY QWI IRN WE $_{558}$ (32)	99	-
	$_{628}$ QSR MQF SSL TVN VRG SG**M** RIL $_{648}$ (21)	0 (M→L 99[c])	577-736:PB1[73]
PB1	$_{1}$ MDV NPT LLF LK**V** PAQ NAI STT FPY TGD PPY SHG TGT GYT MDT VNR THQ YSE $_{51}$ (51)	0.23 (V→I 98)	1-48:PA[75,76]
	$_{114}$ VQQ TRV DKL TQG RQT YDW TLN RNQ PAA TAL ANT IE $_{148}$ (35)	99	-
	$_{337}$ LS**I** API MFS NKM ARL GKG YMF ESK $_{360}$ (24)	0 (I→M 93)	267-493:RNA[77]
	$_{487}$ TGT FEF TSF FYR YGF VAN FSM ELP SFG VSG **I**NE SAD MSI GVT VIK NNM INN DLG PAT AQM ALQ LFI KDY RYT YRC HRG DTQ IQT RRS FE $_{575}$ (89)	0.06 (I→V 92)	506-659:PB2[76]
	$_{655}$ MEY DAV ATT HSW IPK RNR SIL NTS QRG ILE DEQ MYQ $_{690}$ (36)	98	493-757:RNA[77]
PA	$_{29}$ KIE TNK FAA ICT HLE VCF MYS DFH FI $_{54}$ (26)	99	1-209:Endonuclease activity[78,79]
NP	$_{78}$ YLE EHP SAG KDP KKT GGP IY $_{97}$ (20)	99	1-180:RNA[80]
M1	$_{175}$ HEN RMV LAS TTA KAM EQM AGS SEQ AAE AME $_{204}$ (30)	99	165-252:vRNP[81]

[a]Residues in boldface and underlined had very low or no representation in the majority of 2009 pandemic H1N1 strains. $PB1_{1-51, 337-360, 487-575}$ and $PB2_{628-648}$ were selected although they had low or no representation because the remaining length of 9 or more aa sequences was still recognizable as potential T-cell epitopes because the binding core of HLA I and II molecules is typically 9 aa. The length of each sequence is indicated in parentheses. [b]Percentage representation of each highly conserved sequence among the 2009 pandemic H1N1 sequences analyzed as of June 2010. The amino acid substitution that predominated in the 2009 pandemic H1N1 strains and its percentage representation. [d]Functional domain residues followed by the molecule interacted or its function. vRNP, viral ribonucleoprotein.

©2011 Landes Bioscience. Do not distribute.

to HLA molecules.[35,36] At TCR contact positions, variations can lead to reduction of CD8+ cytolytic activity,[37] loss of epitope recognition[38] or formation of altered peptide ligands (APLs). APLs bear structural similarity to the original T-cell epitopes, allowing them to bind to HLA but with amino acid substitutions that alter their interaction with the TCR.[39] Effects of APL include failure to activate T cells[40-42] and induction of T-cell anergy.[43,44] As an approach to minimize these problems of mutational loss of T-cell immunity and the formation of APLs, we focused on identifying and characterizing highly conserved sequences of influenza A viruses that were recorded over the past 30 years, both avian and human, for consideration as potential T cell epitopes-based influenza vaccine targets.

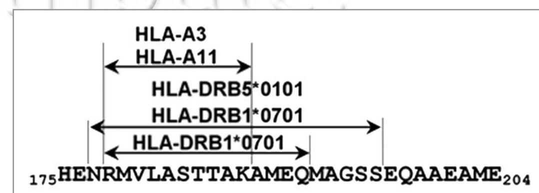

Figure 1. Human HLA class I and II T-cell epitopes mapped in the highly conserved $M1_{175-204}$ sequence. The reported T-cell epitopes were obtained from the IEDB (www.immuneepitope.org) as of June 2010.

Proposal for Development of a T cell Epitopes-Based Influenza Vaccine

Development of T cell epitopes-based vaccines that will induce lasting T-cell immunity and protect against viruses with high mutation rates, including HIV and hepatitis C virus, is currently the goal of many research groups. An ideal T cell epitopes-based vaccine would thus contain evolutionarily conserved T-cell epitopes that are antigenically stable, capable of being restricted by a large number of highly polymorphic HLA molecules, and non-identical or non-similar to human self-proteins.

Immunoinformatics plays several important roles in developing a T cell epitopes-based vaccine. It enables management of large number of influenza sequences retrieved from public and specialized databases, performance of large-scale analysis of conservation of influenza sequences, prediction of T-cell epitopes restricted by different HLA alleles and determination of non-identity or non-similarity to human self-proteins. The NCBI Influenza Virus Resource[45] (www.ncbi.nih.gov/genomes/FLU/), a public specialized repository, provides a convenient platform for bioinformatic analyses of reported influenza virus isolates collected from samples worldwide with tools such as BLAST, multiple sequence alignment and phylogeny programs. We had previously identified highly conserved sequences of influenza A from the analyses of about 36,000 sequences of pathogenic human and avian subtypes that were recorded over the past 30 years.[4] Conservation analyses of this large viral sequence dataset that represents a long temporal period and wide geographical

10 LMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNP 43

121 KVERLKHGTFGPVHFRNQVKIRRRVD 146

228 YIEVLHLTQGTCWEQMYTPGGEVRNDDVDQSLIIAARNIVRRA 270

345 LTGNLQTLKIRVHEGYEEFTMVG 367

400 VAMVFSQEDCMIKAVRGDLNFVNRANQRLNPMHQLLRHFQKDAKVLF 446

527 LTITYSSSMMWEINGPESVLVNTYQWIIRNWE 558

628 QSRMQFSSLTVNVRGSGMRIL 648

PB2

1 MDVNPTLLFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSE 51

114 VQQTRVDKLTQGRQTYDWTLNRNQPAATALANTIE 148

337 LSIAPIMFSNKMARLGKGYMFESK 360

487 TGTFEFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVTVIKN 532

533 NMINNDLGPATAQMALQLFIKDYRYTYRCHRGDTQIQTRRSFE 575

655 MEYDAVATTHSWIPKRNRSILNTSQRGILEDEQMYQ 690

PB1

29 KIETNKFAAICTHLEVCFMYSDFHFI 54

PA

78 YLEEHPSAGKDPKKTGGPIY 97

NP

175 HENRMVLASTTAKAMEQMAGSSEQAAEAME 204

M1

Figure 2. 163 putative HLA supertype CTL epitopes predicted by NetCTL-1.2 (www.cbs.dtu.dk/services/NetCTL). Default settings were used in the prediction.

distribution revealed 55 sequences highly conserved in at least 80%, many in 95 to 100%, of the human and avian influenza virus isolates. Assessment of the conservation of these 55 historically stable sequences against all reported 2009 pandemic H1N1 sequences (as of June 2010, **Table 1**) demonstrated that the majority of the sequences remained evolutionarily robust, with 47 sequences present in an average of 99% of the reported pandemic strains. The remaining eight highly conserved sequences were found in none or less than 0.25% of all the reported pandemic strains (**Table 2**). However, most of the mutations in the predominant 2009 H1N1 pandemic strain corresponding to these eight conserved sequences were conservative substitutions. Furthermore, these mutations were found at the beginning or the end of the conserved sequences, leaving the remaining length of 9 or more consecutive aa recognizable as potential T-cell epitopes where the binding core of HLA I and II molecules is typically 9 aa.

From the highly conserved sequences, we identified 15 sequences as vaccine targets based on several criteria. These 15 sequences had high representation in majority of the pandemic 2009 H1N1 strains and the six pathogenic influenza subtypes (human H1N1, H3N2, H1N2, H5N1, avian H5N1 and other avian subtypes) (**Table 3**). Their high representation in the large dataset of viral sequences spanning three decades increases the probability that future circulating influenza strains will contain these sequences. Moreover, 11 of the 15 selected highly conserved sequences are associated with functional domains that further suggests they are likely to remain conserved as mutations in these domains might disrupt protein function or affect viral fitness, unless compensated functionally by multiple co-mutations.[38,46] The relatively long lengths of the 15 highly conserved sequences (20–89 aa) make them relevant for both HLA class I and II epitopes for recognition by CTL and HTL, respectively. Furthermore, a long peptide sequence containing multiple epitopes is more efficient in stimulating epitope-specific cellular immune responses in contrast to short epitopes alone.[47] Moreover, these sequences may contain linear B-cell epitopes as suggested by studies of mice immunized with NP which were protected against influenza challenge.[48] Multiple highly conserved sequences from different internal influenza proteins are included to increase the breadth of T-cell responses. In addition, vaccination with multiple sequences, as opposed to a single sequence, would increase the repertoire of epitopes available for selection based on HLA restriction.

Within the 15 highly conserved sequences, 23 T-cell epitopes had been mapped by use of HLA transgenic mice expressing Class I A24 and Class II DR2, DR3 and DR4.[49] The Immune Epitope Database (IEDB, www.iedb.org)[50] also listed experimentally characterized human HLA class I and II T-cell epitopes in the highly conserved $M1_{175\text{-}204}$ where, just within the 30 aa, two HLA class I and three class II T-cell epitopes were clustered and overlapped (**Fig. 1**). It is likely that these 15 highly conserved sequences (total of 537 aa) further contain other HLA T-cell epitopes as the number of mapped human T-cell epitopes restricted by different HLA alleles is currently small but is growing as more research groups report data from studies with HLA transgenic mice or HLA-typed human peripheral blood mononucleocytes.[51-54] This is supported by several putative epitopes predicted by use of immunoinformatics tool such as NetCTL-1.2,[55] which integrated predictions of proteasomal C terminal cleavage, TAP transport efficiency and peptide-HLA class I binding. A total of 163 CTL epitopes of 12 HLA supertypes were predicted where 36 were restricted by two or more HLA supertypes (**Fig. 2**). HLA supertypes are classified groups of HLA molecules sharing similar peptide binding specificity.[56-58] Notably, PB1 $_{501}$FVANFSMEL$_{508}$ was predicted as restricted by seven HLA supertypes (A1, A2, A26, B7, B8, B39 and B62). The high number of promiscuous T-cell epitopes binding to multiple HLA supertype alleles within the conserved sequences helps to maximize human population coverage where the frequency at which the supertype alleles are expressed in various ethnicities is remarkably high.[58] For example, the $PA_{29\text{-}54}$ highly conserved sequence containing the nonamer FMYSDFHFI was experimentally shown to bind to at least five major HLA class I A2 supertype alleles (A*0201, A*0202, A*0203, A*0206 and A*6802).[59]

Since HLA molecules are also involved in recognition of self peptides, attention must be given to the possibility of "molecular mimicry" between epitopes within the influenza highly conserved sequences and human sequences, which may contribute towards autoimmunity. Autoimmunity occurs when there is a breakdown of self tolerance, resulting in the adaptive immune system attacking itself. Currently, there is a limited understanding of the onset of autoimmunity but genetic predisposition, environmental factors and chronic viral infections have been attributed.[60] Although influenza infection is rarely implicated as a trigger for autoimmune diseases, cases of adverse autoimmune events have been reported following fulminant influenza infection[61] and current influenza vaccination.[62] None of our 15 selected highly conserved influenza sequences was identical to human peptides in our BLAST analyses as defined by 100% identity to at least seven consecutive amino acids of human proteins. Our finding corroborates that influenza peptide sequences overlap minimally with human self-peptide repertoire.[63,64] However, it should be noted that viral antigens do not need a 100% identity with self antigens to cross-react with self T cells.

In summary, our proposed 15 highly conserved sequences are candidates for development of a broad and lasting T-cell epitopes-based influenza vaccine based on the following criteria:

(1) High representation in major pathogenic human and avian influenza subtypes, including 2009 pandemic H1N1 strains.

(2) Potential relevance to CTL, HTL and linear antibody epitopes due to their relatively long sequence length (at least 20 aa).

(3) The evolutionary conservation and association with functional domains indicate conservation in future circulating strains.

(4) Occurrence of known and putative HLA-supertype restricted T-cell epitopes for broad coverage of human population.

(5) Non-identity to human self peptides to avoid molecular mimicry.

Challenges for T Cell Epitopes-Based Vaccines

The paradigm of developing a T cell epitopes-based vaccine is in its infancy. In a viral infection, the majority of host T-cell

©2011 Landes Bioscience. Do not distribute.

populations are focused on a limited number of viral peptides, termed immunodominant epitopes, at the expense of subdominant epitopes which are recognized by less abundant T-cell populations.[65] Antiviral responses to immunodominant and subdominant epitopes form a hierarchy that is reproducible among individuals of an inbred mouse strain.[66,67] The process where dominant epitopes suppress the immunogenicity of subdominant epitopes is called immunodomination.[68] This immunodominant phenomenon is observed for both CTLs and HTLs.[65,69,70] The mechanisms underlying immunodominance hierarchy are complex and poorly understood. A myriad of factors determining immunodominance are involved, starting from antigen processing, intracellular competitive binding to HLA molecules, recognition of a diverse TCR repertoires, to memory T cells. In humans, binding affinity to HLA molecules alone does not dictate immunodominance.[71] This difference maybe ascribed to the complex history of repeated exposure to different influenza strains in human adults. The effects on the immunodominance hierarchy by repeated exposure to heterosubtypic influenza strains and vaccination followed by natural viral infection in human need further investigation.

In addition, we need to identify and characterize a greater repertoire of T-cell epitopes restricted by different HLA alleles in human studies. Formulations and protocols for immunizing humans with T-cell epitopes also need to be explored with the possible combination of DNA vaccines and peptides. For DNA vaccines, there is also the possible development of more efficient expression plasmid vectors and addition of cytokine sequences for enhanced immune responses. Our laboratory is investigating targeted delivery to HLA class II compartment of a chimeric vaccine composed of the lysosomal associated membrane protein and the highly conserved influenza sequences as well as the efficacy of DNA expression via delivery by in vivo electroporation.

References

1. Fiore AE, Bridges CB, Cox NJ. Seasonal influenza vaccines. Curr Top Microbiol Immunol 2009; 333:43-82.
2. de Jong JC, Beyer WE, Palache AM, Rimmelzwaan GF, Osterhaus AD. Mismatch between the 1997/1998 influenza vaccine and the major epidemic A(H3N2) virus strain as the cause of an inadequate vaccine-induced antibody response to this strain in the elderly. J Med Virol 2000; 61:94-9.
3. Jin H, Zhou H, Liu H, Chan W, Adhikary L, Mahmood K, et al. Two residues in the hemagglutinin of A/Fujian/411/02-like influenza viruses are responsible for antigenic drift from A/Panama/2007/99. Virology 2005; 336:113-9.
4. Heiny AT, Miotto O, Srinivasan KN, Khan AM, Zhang GL, Brusic V, et al. Evolutionarily conserved protein sequences of influenza a viruses, avian and human, as vaccine targets. PLoS One 2007; 2:1190.
5. Shastri N, Schwab S, Serwold T. Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. Annu Rev Immunol 2002; 20:463-93.
6. Bjorkman PJ, Parham P. Structure, function and diversity of class I major histocompatibility complex molecules. Annu Rev Biochem 1990; 59:253-88.
7. Esser MT, Marchese RD, Kierstead LS, Tussey LG, Wang F, Chirmule N, et al. Memory T cells and vaccines. Vaccine 2003; 21:419-30.
8. Zinkernagel RM, Hengartner H. On immunity against infections and vaccines: credo 2004. Scand J Immunol 2004; 60:9-13.
9. Pulendran B, Ahmed R. Translating innate immunity into immunological memory: implications for vaccine development. Cell 2006; 124:849-63.
10. Roy S, Kobinger GP, Lin J, Figueredo J, Calcedo R, Kobasa D, et al. Partial protection against H5N1 influenza in mice with a single dose of a chimpanzee adenovirus vector expressing nucleoprotein. Vaccine 2007; 25:6845-51.
11. Okuda K, Ihata A, Watabe S, Okada E, Yamakawa T, Hamajima K, et al. Protective immunity against influenza A virus induced by immunization with DNA plasmid containing influenza M gene. Vaccine 2001; 19:3681-91.
12. Matsui M, Kohyama S, Suda T, Yokoyama S, Mori M, Kobayashi A, et al. A CTL-based liposomal vaccine capable of inducing protection against heterosubtypic influenza viruses in HLA-A*0201 transgenic mice. Biochem Biophys Res Commun 2010; 391:1494-9.
13. Epstein SL, Kong WP, Misplon JA, Lo CY, Tumpey TM, Xu L, et al. Protection against multiple influenza A subtypes by vaccination with highly conserved nucleoprotein. Vaccine 2005; 23:5404-10.
14. Tripp RA, Sarawar SR, Doherty PC. Characteristics of the influenza virus-specific $CD8^+$ T cell response in mice homozygous for disruption of the H-2lAb gene. J Immunol 1995; 155:2955-9.
15. Brown DM, Dilzer AM, Meents DL, Swain SL. CD4 T cell-mediated protection from lethal influenza: perforin and antibody-mediated mechanisms give a one-two punch. J Immunol 2006; 177:2888-98.
16. Dienz O, Eaton SM, Bond JP, Neveu W, Moquin D, Noubade R, et al. The induction of antibody production by IL-6 is indirectly mediated by IL-21 produced by $CD4^+$ T cells. J Exp Med 2009; 206:69-78.
17. Scherle PA, Gerhard W. Functional analysis of influenza-specific helper T cell clones in vivo. T cells specific for internal viral proteins provide cognate help for B cell responses to hemagglutinin. J Exp Med 1986; 164:1114-28.
18. Belz GT, Wodarz D, Diaz G, Nowak MA, Doherty PC. Compromised influenza virus-specific CD8(+)-T-cell memory in CD4(+)-T-cell-deficient mice. J Virol 2002; 76:12388-93.
19. Castellino F, Germain RN. Cooperation between $CD4^+$ and $CD8^+$ T cells: when, where and how. Annu Rev Immunol 2006; 24:519-40.
20. Schepers K, Arens R, Schumacher TN. Dissection of cytotoxic and helper T cell responses. Cell Mol Life Sci 2005; 62:2695-710.
21. Brown DM, Roman E, Swain SL. CD4 T cell responses to influenza infection. Semin Immunol 2004; 16:171-7.
22. Strutt TM, McKinstry KK, Swain SL. Functionally diverse subsets in CD4 T cell responses against influenza. J Clin Immunol 2009; 29:145-50.
23. Brown LE, Kelso A. Prospects for an influenza vaccine that induces cross-protective cytotoxic T lymphocytes. Immunol Cell Biol 2009; 87:300-8.
24. Francis T. Influenza: The newe acquayantance. Ann Intern Med 1953; 39:203-21.
25. Kim JH, Skountzou I, Compans R, Jacob J. Original antigenic sin responses to influenza viruses. J Immunol 2009; 183:3294-301.
26. Klenerman P, Zinkernagel RM. Original antigenic sin impairs cytotoxic T lymphocyte responses to viruses bearing variant epitopes. Nature 1998; 394:482-5.
27. Kundu SK, Dupuis M, Sette A, Celis E, Dorner F, Eibl M, et al. Role of preimmunization virus sequences in cellular immunity in HIV-infected patients during HIV type 1 MN recombinant gp160 immunization. AIDS Res Hum Retroviruses 1998; 14:1669-78.
28. Yellen-Shaw AJ, Wherry EJ, Dubois GC, Eisenlohr LC. Point mutation flanking a CTL epitope ablates in vitro and in vivo recognition of a full-length viral protein. J Immunol 1997; 158:3227-34.
29. Sant'Angelo DB, Robinson E, Janeway CA Jr, Denzin LK. Recognition of core and flanking amino acids of MHC class II-bound peptides by the T cell receptor. Eur J Immunol 2002; 32:2510-20.
30. Beekman NJ, van Veelen PA, van Hall T, Neisig A, Sijts A, Camps M, et al. Abrogation of CTL epitope processing by single amino acid substitution flanking the C-terminal proteasome cleavage site. J Immunol 2000; 164:1898-905.
31. Maverakis E, Beech JT, Schneider S, Sercarz EE. Presentation of a determinant by MHC class II can be prevented through competitive capture by a flanking determinant on a multideterminant peptide. J Autoimmun 2008; 31:59-65.
32. Theodossis A, Guillonneau C, Welland A, Ely LK, Clements CS, Williamson NA, et al. Constraints within major histocompatibility complex class I restricted peptides: presentation and consequences for T-cell recognition. Proc Natl Acad Sci USA 2010; 107:5534-9.
33. Anderson MW, Gorski J. Cutting edge: TCR contacts as anchors: effects on affinity and HLA-DM stability. J Immunol 2003; 171:5683-7.
34. James EA, Moustakas AK, Bui J, Nouv R, Papadopoulos GK, 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.
35. Voeten JT, Bestebroer TM, Nieuwkoop NJ, Fouchier RA, Osterhaus AD, Rimmelzwaan GF. Antigenic drift in the influenza A virus (H3N2) nucleoprotein and escape from recognition by cytotoxic T lymphocytes. J Virol 2000; 74:6800-7.
36. Rimmelzwaan GF, Boon AC, Voeten JT, Berkhoff EG, Fouchier RA, Osterhaus AD. Sequence variation in the influenza A virus nucleoprotein associated with escape from cytotoxic T lymphocytes. Virus Res 2004; 103:97-100.
37. Berkhoff EG, Geelhoed-Mieras MM, Verschuren EJ, van Baalen CA, Gruters RA, Fouchier RA, et al. The loss of immunodominant epitopes affects interferon-gamma production and lytic activity of the human influenza virus-specific cytotoxic T lymphocyte response in vitro. Clin Exp Immunol 2007; 148:296-306.
38. Rimmelzwaan GF, Kreijtz JH, Bodewes R, Fouchier RA, Osterhaus AD. Influenza virus CTL epitopes, remarkably conserved and remarkably variable. Vaccine 2009; 27:6363-5.
39. Sloan-Lancaster J, Allen PM. Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. Annu Rev Immunol 1996; 14:1-27.
40. Hlavac F, Choppin J, Guillet JG. Self-peptide ligands affect T cell recognition of the homologous influenza A matrix virus peptide M.58-66: modification of the HLA-A2.1/peptide complex structure and T cell antagonism. Hum Immunol 1997; 54:48-53.

©2011 Landes Bioscience. Do not distribute.

41. Li X, Li R, Li Z. Influenza virus haemagglutinin-derived peptides inhibit T-cell activation induced by HLA-DR4/1 specific peptides in rheumatoid arthritis. Clin Exp Rheumatol 2006; 24:148-54.
42. Boon AC, de Mutsert G, Graus YM, Fouchier RA, Sintnicolaas K, Osterhaus AD, et al. Sequence variation in a newly identified HLA-B35-restricted epitope in the influenza A virus nucleoprotein associated with escape from cytotoxic T lymphocytes. J Virol 2002; 76:2567-72.
43. Tsitoura DC, Holter W, Cerwenka A, Gelder CM, Lamb JR. Induction of anergy in human T helper 0 cells by stimulation with altered T cell antigen receptor ligands. J Immunol 1996; 156:2801-8.
44. Mirshahidi S, Ferris LC, Sadegh-Nasseri S. The magnitude of TCR engagement is a critical predictor of T cell anergy or activation. J Immunol 2004; 172:5346-55.
45. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, et al. The influenza virus resource at the National Center for Biotechnology Information. J Virol 2008; 82:596-601.
46. Berkhoff EG, de Wit E, Geelhoed-Mieras MM, Boon AC, Symons J, Fouchier RA, et al. Fitness costs limit escape from cytotoxic T lymphocytes by influenza A viruses. Vaccine 2006; 24:6594-6.
47. Wild J, Bojak A, Deml L, Wagner R. Influence of polypeptide size and intracellular sorting on the induction of epitope-specific CTL responses by DNA vaccines in a mouse model. Vaccine 2004; 22:1732-43.
48. Carragher DM, Kaminski DA, Moquin A, Hartson L, Randall TD. A novel role for non-neutralizing antibodies against nucleoprotein in facilitating resistance to influenza virus. J Immunol 2008; 181:4168-76.
49. Tan PT, Heiny AT, Miotto O, Salmon J, Marques ET, Lemonnier F, et al. Conservation and diversity of influenza A H1N1 HLA-restricted T cell epitope candidates for epitope-based vaccines. PLoS One 2010; 5:8754.
50. Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, et al. The immune epitope database 2.0. Nucleic Acids Res 2010; 38:854-62.
51. Greenbaum JA, Kotturi MF, Kim Y, Oseroff C, Vaughan K, Salimi N, et al. Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population. Proc Natl Acad Sci USA 2009; 106:20365-70.
52. Alexander J, Bilsel P, del Guercio MF, Stewart S, Marinkovic-Petrovic A, Southwood S, et al. Universal influenza DNA vaccine encoding conserved $CD4^+$ T cell epitopes protects against lethal viral challenge in HLA-DR transgenic mice. Vaccine 2010; 28:664-72.
53. Assarsson E, Bui HH, Sidney J, Zhang Q, Glenn J, Oseroff C, et al. Immunomic analysis of the repertoire of T-cell specificities for influenza A virus in humans. J Virol 2008; 82:12241-51.
54. Sun Y, Liu J, Yang M, Gao F, Zhou J, Kitamura Y, et al. Identification and structural definition of H5-specific CTL epitopes restricted by HLA-A*0201 derived from the H5N1 subtype of influenza A viruses. J Gen Virol 2010; 91:919-30.
55. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics 2007; 8:424.
56. Doytchinova IA, Flower DR. In silico identification of supertypes for class II MHCs. J Immunol 2005; 174:7085-95.
57. Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. BMC Immunol 2008; 9:1.
58. Sette A, Sidney J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics 1999; 50:201-12.
59. Alexander J, Bilsel P, del Guercio MF, Marinkovic-Petrovic A, Southwood S, Stewart S, et al. Identification of broad binding class I HLA supertype epitopes to provide universal coverage of influenza A virus. Hum Immunol 2010; 71:468-74.
60. Getts MT, Miller SD. 99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: triggering of autoimmune diseases by infections. Clin Exp Immunol 2010; 160:15-21.
61. Tam CC, O'Brien SJ, Petersen I, Islam A, Hayward A, Rodrigues LC. Guillain-Barre syndrome and preceding infection with campylobacter, influenza and Epstein-Barr virus in the general practice research database. PLoS One 2007; 2:344.
62. Toplak N, Avcin T. Influenza and autoimmunity. Ann NY Acad Sci 2009; 1173:619-26.
63. Louzoun Y, Vider T, Weigert M. T-cell epitope repertoire as predicted from human and viral genomes. Mol Immunol 2006; 43:559-69.
64. Burroughs NJ, de Boer RJ, Kesmir C. Discriminating self from nonself with short peptides from large proteomes. Immunogenetics 2004; 56:311-20.
65. Sercarz EE, Lehmann PV, Ametani A, Benichou G, Miller A, Moudgil K. Dominance and crypticity of T cell antigenic determinants. Annu Rev Immunol 1993; 11:729-66.
66. Chen W, Bennink JR, Morton PA, Yewdell JW. Mice deficient in perforin, $CD4^+$ T cells or CD28-mediated signaling maintain the typical immunodominance hierarchies of $CD8^+$ T-cell responses to influenza virus. J Virol 2002; 76:10332-7.
67. Chen W, Anton LC, Bennink JR, Yewdell JW. Dissecting the multifactorial causes of immunodominance in class I-restricted T cell responses to viruses. Immunity 2000; 12:83-93.
68. Liu J, Ewald BA, Lynch DM, Nanda A, Sumida SM, Barouch DH. Modulation of DNA vaccine-elicited $CD8^+$ T-lymphocyte epitope immunodominance hierarchies. J Virol 2006; 80:11991-7.
69. Yewdell JW. Confronting complexity: real-world immunodominance in antiviral $CD8^+$ T cell responses. Immunity 2006; 25:533-43.
70. Sant AJ, Chaves FA, Jenks SA, Richards KA, Menges P, Weaver JM, et al. The relationship between immunodominance, DM editing and the kinetic stability of MHC class II: peptide complexes. Immunol Rev 2005; 207:261-78.
71. Gelder C, Davenport M, Barnardo M, Bourne T, Lamb J, Askonas B, et al. Six unrelated HLA-DR-matched adults recognize identical $CD4^+$ T cell epitopes from influenza A haemagglutinin that are not simply peptides with high HLA-DR binding affinities. Int Immunol 1998; 10:211-22.
72. Biswas SK, Nayak DP. Influenza virus polymerase basic protein 1 interacts with influenza virus polymerase basic protein 2 at multiple sites. J Virol 1996; 70:6716-22.
73. Poole E, Elton D, Medcalf L, Digard P. Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. Virology 2004; 321:120-33.
74. Biswas SK, Boutz PL, Nayak DP. Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. J Virol 1998; 72:5493-501.
75. Perez DR, Donis RO. A 48-amino-acid region of influenza A virus PB1 protein is sufficient for complex formation with PA. J Virol 1995; 69:6932-9.
76. Gonzalez S, Zurcher T, Ortin J. Identification of two separate domains in the influenza virus PB1 protein involved in the interaction with the PB2 and PA subunits: a model for the viral RNA polymerase structure. Nucleic Acids Res 1996; 24:4456-63.
77. Gonzalez S, Ortin J. Distinct regions of influenza virus PB1 polymerase subunit recognize vRNA and cRNA templates. EMBO J 1999; 18:3767-75.
78. Dias A, Bouvier D, Crepin T, McCarthy AA, Hart DJ, Baudin F, et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature 2009; 458:914-8.
79. Hara K, Schmidt FI, Crow M, Brownlee GG. Amino acid residues in the N-terminal region of the PA subunit of influenza A virus RNA polymerase play a critical role in protein stability, endonuclease activity, cap binding and virion RNA promoter binding. J Virol 2006; 80:7789-98.
80. Albo C, Valencia A, Portela A. Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. J Virol 1995; 69:3799-806.
81. Baudin F, Petit I, Weissenhorn W, Ruigrok RW. In vitro dissection of the membrane and RNP binding activities of influenza virus M1 protein. Virology 2001; 281:102-8.

©2011 Landes Bioscience. Do not distribute.