

RESEARCH ARTICLE

Open Access

# Residue analysis of a CTL epitope of SARS-CoV spike protein by IFN-gamma production and bioinformatics prediction

Jun Huang<sup>1,3</sup>, Yingnan Cao<sup>2,4</sup>, Xianzhang Bu<sup>2</sup> and Changyou Wu<sup>1\*</sup>

## Abstract

**Background:** Severe acute respiratory syndrome (SARS) is an emerging infectious disease caused by the novel coronavirus SARS-CoV. The T cell epitopes of the SARS CoV spike protein are well known, but no systematic evaluation of the functional and structural roles of each residue has been reported for these antigenic epitopes. Analysis of the functional importance of side-chains by mutational study may exaggerate the effect by imposing a structural disturbance or an unusual steric, electrostatic or hydrophobic interaction.

**Results:** We demonstrated that N50 could induce significant IFN-gamma response from SARS-CoV S DNA immunized mice splenocytes by the means of ELISA, ELISPOT and FACS. Moreover, S366-374 was predicted to be an optimal epitope by bioinformatics tools: ANN, SMM, ARB and BIMAS, and confirmed by IFN-gamma response induced by a series of S358-374-derived peptides. Furthermore, each of S366-374 was replaced by alanine (A), lysine (K) or aspartic acid (D), respectively. ANN was used to estimate the binding affinity of single S366-374 mutants to H-2 Kd. Y367 and L374 were predicated to possess the most important role in peptide binding. Additionally, these one residue mutated peptides were synthesized, and IFN-gamma production induced by G368, V369, A371, T372 and K373 mutated S366-374 were decreased obviously.

**Conclusions:** We demonstrated that S366-374 is an optimal H-2 Kd CTL epitope in the SARS CoV S protein. Moreover, Y367, S370, and L374 are anchors in the epitope, while C366, G368, V369, A371, T372, and K373 may directly interact with TCR on the surface of CD8-T cells.

**Keywords:** SARS-CoV, CTL, Epitope, Residue

## Background

Severe acute respiratory syndrome (SARS) is an emerging infectious disease caused by the novel coronavirus SARS-CoV [1,2]. The fatality rate is as high as 15% for patients younger than 60 years old and can be higher than 50% for patients 60 years or older. Nearly 40% of infected patients develop respiratory failure that requires assistant ventilation [3].

Coronaviruses (CoVs) are positive-strand RNA viruses. The virion consists of a nucleocapsid (N) core surrounded by an envelope containing three membrane proteins, spike (S), membrane (M) and envelope (E),

which are common to all members of the genus [4,5]. The M and E proteins are important for viral particle assembly and N is important for viral RNA packaging. The S protein, which provides the virion with a corona-like appearance, binds to host receptors and mediates membrane fusion [6]. The successful development of effective treatments and vaccines against SARS-CoV depends on understanding the roles of various immune effectors in protective immunity and on identifying protective viral antigens recognized by these effector cells. In a preliminary study, the S protein fragment S358-374 was shown to stimulate the production of IFN- $\gamma$  by CD8-T cells from immunized BALB/c mice [7].

The capacity of a diverse array of peptides to bind to an individual class I molecule is due to anchor residues in the peptides [8]. The surface features of the binding cleft of the class I MHC molecule are complementary to

\* Correspondence: changyou\_wu@yahoo.com

<sup>1</sup>Institute of Immunology, Zhongshan School of Medicine; Key Laboratory of Tropical Disease Control Research of Ministry of Education, Sun Yat-sen University, Guangzhou, China

Full list of author information is available at the end of the article

side chains of the anchor residues in the displayed peptide. The amino acid residues lining the binding sites may vary among different class I allelic variants [9,10]. Here, an SARS CoV S protein CTL epitope, S366–374, was identified and the functions of individual residues were evaluated by bioinformatics tool prediction and by IFN- $\gamma$  responses induced by a series of modified S366–374 peptides.

## Methods

### Mice

Female BALB/c mice, 6–8 weeks old, were purchased from Zhongshan University Animal Center (Guangzhou, China) and maintained in our animal care facility under pathogen free conditions. For experimental purposes, six to eight week-old female mice were used. All experiments were performed according to the guidelines in the Institutional Animal Committee of Zhongshan School of Medicine, China.

### SARS-CoV S DNA vaccine

Plasmids encoding SARS-CoV spike (S) protein was constructed as described [11], and kindly provided by Dr. Gary J. Nabel from Vaccine Research Center, NIAID, National Institutes of Health, MD, USA. Plasmid DNA was purified by plasmid-purified kit (QIAGEN, USA). The 260/280 ratios ranged from 1.8 to 2.0. The endotoxin content from purified plasmid DNA was found below 20 U/ml. The endotoxin level within this range had no effect on the immune response.

### Synthesis, purification and analysis of S358-374 analogs

The peptides were synthesized by solid phase using an Fmoc strategy [12]. 2-Chlorotrityl chloride resin loaded at 1.0 mmol/g (GL BIO Company, Shanghai, China) which was used in peptide synthesis and was chosen as the solid support. N-protected Fmoc amino acids were used. For functionalized amino acids, the following derivatives were used: Fmoc-Cys(trt)-OH, Fmoc-Leu-OH, Fmoc-Val-OH, Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH, and Fmoc-Lys(Boc)-OH and were from GL BIO Company (Shanghai, China). The coupling reagent was 2-(1 H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate for all peptides. Syntheses of the different peptides were performed from 0.07 g of 2-Chlorotrityl chloride resin. An excess of 8 eq of each amino acid was used.

Fmoc deprotection was performed with a solution of piperidine in dimethylformamide in a 2/8 (v/v) ratio. Final deprotection of the peptides from the resin was performed in a mixture containing trifluoroacetic acid, phenol, water, triisopropylsilane in a 88/5/5/2 (v/v) ratio for 3.5 h. Peptides were then precipitated by addition of

cold diethyl ether and dissolved in a mixture of 0.1% trifluoroacetic acid in water/acetonitrile and lyophilized. Purity of all peptides was checked by analytical high performance liquid chromatography on a Waters instrument using a C18 column (Novarpack, 5  $\mu$ m, 300  $\text{\AA}$ , 10.0  $\times$  200 mm) and all were at least 70% pure (UV detection at 214 and 254 nm). They were characterized by electrospray mass spectrometry.

### Immunization of mice

Female BALB/c mice were injected (i.m.) with 50  $\mu$ g/mouse of SARS-CoV S plasmid DNA in 100  $\mu$ l of sterile PBS. Mice were boosted twice at 2–3 weeks interval.

### Cell culture and IFN- $\gamma$ ELISA

Mice were sacrificed. Spleen from individual mouse was harvested one to two weeks after the final boost vaccination. Single cell suspensions were prepared and plated in a 96-well micro-titer plate at  $4 \times 10^5$  cells/200  $\mu$ l per well. Pooled SARS CoV S peptides (1  $\mu$ g/ml for each) or single peptide (1  $\mu$ g/ml) with anti-mouse CD28 mAb (1  $\mu$ g/ml) were added to cultures. Supernatants of cell cultures were collected 72 h later, and levels of IFN- $\gamma$  were assessed by specific ELISA kit (BD PharMingen) according to the manufacturer's protocol. The detection limit of the IFN- $\gamma$  assay kit was 3.13 pg/ml.

### IFN- $\gamma$ ELISPOT

Assessment of SARS-CoV S-specific IFN- $\gamma$  producing cells after vaccination was determined by ELISPOT (Diaclone, France) according to the manufacturer's protocol. In brief, single cell suspensions were prepared from spleens of mice after vaccination, and plated in 96-well microplate precoated with anti-IFN- $\gamma$  antibody specific for ELISPOT. Cells were incubated overnight in the presence or absence of peptide (1  $\mu$ g/ml) and anti-CD28 (1  $\mu$ g/ml). The plates were then washed and alkaline phosphatase conjugated anti-mouse IFN- $\gamma$  antibody was added, developed with ready-to-use BCIP/NBT, and read by Champ Spot II ELISPOT reader (Sage Creation, China).

### Cell surface and intracellular cytokine staining

Single-cell suspensions from spleens of mice after vaccination were stimulated with or without SARS CoV S peptides plus anti-CD28 (1  $\mu$ g/ml) for 5 h at 37°C and 5% CO<sub>2</sub>. Brefeldin A (10  $\mu$ g/ml, Sigma) was added in the last 4 h incubation. Cells were washed, fixed with 4% paraformaldehyde and permeabilized in PBS buffer containing 0.1% saponin (Sigma), 0.1% BSA and 0.05% NaN<sub>3</sub> overnight at 4°C. Cells were then stained with conjugated mAbs specific for CD4, CD8 and intracellular cytokine IFN- $\gamma$  for 20–30 min at 4°C in dark. Cells (300,000) were acquired on flow cytometer (BD Calibur)

and data were analyzed with FlowJo program, version 6.0 (Tree Star, Inc., USA). Isotype matched controls for cytokines were included in each staining.

### Bioinformatics analysis

T cell epitope prediction tools, ANN, SMM and ARB provided by IEDB are publicly available in a website ([http://tools.immuneepitope.org/analyze/html/mhc\\_binding.html](http://tools.immuneepitope.org/analyze/html/mhc_binding.html)) Version 2009-09-01B.

Artificial Neural Network (ANN) is a connectionist models that consist of a number of interconnected units that can be activated by transmitting signals [13,14]. ANN can tolerate a degree of erroneous data, and can classify nonlinear data, which makes them highly suitable for processing noisy biological information. ANN applications have been described for predictions of MHC class I binding peptides and for MHC class II peptides. The prediction accuracy of ANN-based methods was reported to be close to 80% sensitivity and 80% specificity [15].

Stabilized Matrix Method (SMM) is a T cell epitope predictive tool based on quantitative matrices. It has been successfully applied to predicting peptide binding to MHC molecules, peptide transport by the transporter associated with antigen presentation (TAP) and proteasomal cleavage of protein sequences [16,17]. The sensitivity of SMM-based methods was reported to be close to 60% [18].

Average Relative Binding (ARB) is a matrix method, which allows combination of searches involving different peptide sizes and alleles into a single global prediction [19]. ARB has achieved a favorable performance in predicting MHC I and MHC II molecules [20,21].

Application of these tools was according to prompt dialog box showed in the website. In brief: a) Choose MHC class I model. b) Enter sequence of various peptides. c) Choose prediction method as ANN, SMM and ARB, respectively. d) Specify what to make binding predictions for. Select MHC source species as mouse, allele as H-2K<sup>d</sup>, length as 9. e) Specify output. These tools predict IC<sub>50</sub> values for peptide binding to specific MHC molecules. IC<sub>50</sub> value means binding ability. Note that binding to MHC is necessary but not sufficient for recognition by T cells.

HLA Peptide Binding Predictions tool provided by Bioinformatics and Molecular Analysis Section (BIMAS) ([http://www.bimas.cit.nih.gov/cgi-bin/molbio/ken\\_parker\\_comboform](http://www.bimas.cit.nih.gov/cgi-bin/molbio/ken_parker_comboform)) was used to compare the binding kinetics of peptides. The analysis is based on coefficient tables deduced from the published literature by Dr. Kenneth Parker [22]. Application of this tool was according to prompt dialog box provided by the website. Higher half lime value means better binding ability.

### Statistics

Statistical evaluation of differences between means of experimental groups was performed by analysis of variance and a non-parametric two-tailed *t* test. P value <0.05 was considered to be significant.

### Results

#### N50 is a MHC-I restricted peptide in SARS-CoV S protein

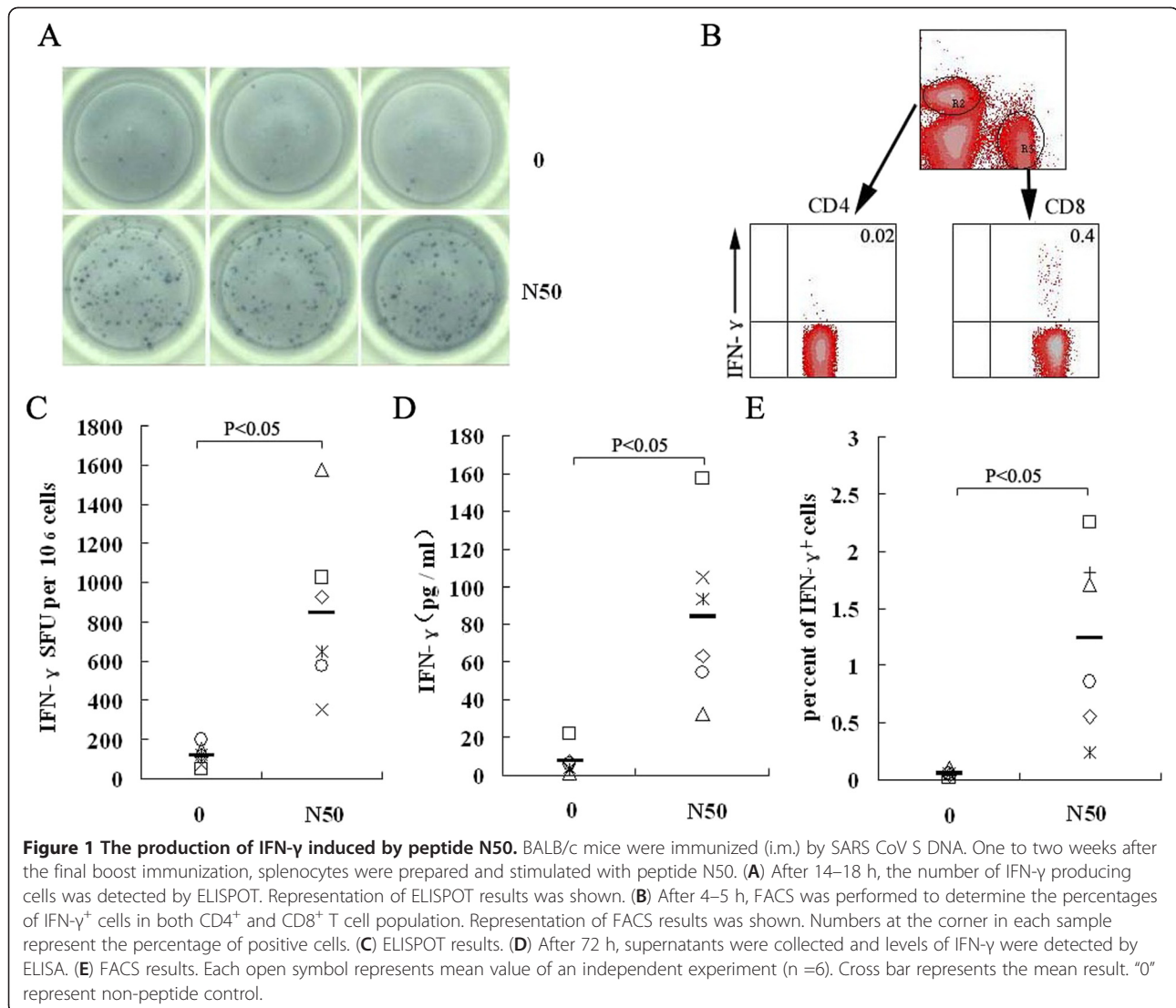
To identify SARS CoV S epitopes, the potential SARS-CoV S epitopes were tested repeatedly by splenocytes from DNA vaccine immunized BALB/c mice. ELISA and ELISPOT results indicated that the adjacent peptides P50 and P51 possessed the same ability to induce IFN- $\gamma$  production [9]. The overlapping sequence between P50 and P51 (N50, KCYGVSATKL) was synthesized. ELISA (Figure 1A), ELISPOT (Figure 1B/D) and FACS results indicated that peptide N50 could induce IFN- $\gamma$  production. The FACS results showed that N50 could only induced CD8<sup>+</sup> T cells to produce IFN- $\gamma$  (Figure 1C/E).

#### Amino acid residue L<sup>374</sup> is essential for stimulation of IFN- $\gamma$ production in response to S365–374

To identify the optimal epitope in S365–374, a series of S358–374-derived peptides were synthesized and used to stimulate splenocytes from SARS-CoV S DNA vaccine immunized BALB/c mice. The fraction of IFN- $\gamma$ -producing T cells was determined by ELISPOT (Figure 2A), and the level of IFN- $\gamma$  in supernatants was measured by ELISA (Figure 2B). Both results indicated that IFN- $\gamma$  was produced only in response to peptides preserving residue L<sup>374</sup>. Thus, S367–374 (YGVSATKL), S365–374 (KCYGVSATKL), and S364–374 (FKCYGVSATKL) could elicit robust IFN- $\gamma$  production. Only S370–374 (SATKL) was inactive, likely due to weak affinity to MHC-I (data not shown). In contrast, L<sup>374</sup> deleted peptides, including S369–373 (VSATK), S366–373 (CYGVSATK), and S363–373 (FKCYGVSATK) could not induce IFN- $\gamma$  production. The IFN- $\gamma$  response induced by S365–374 was much stronger than that induced by S367–374 ( $P < 0.05$ ).

#### S366–374 is the optimal epitope

To identify the optimal epitope, we analyzed the binding affinity of S365–374 peptides to H-2 K<sup>d</sup>, H-2 D<sup>d</sup>, and L<sup>d</sup> by several bioinformatics tools. The MHC-binding scores were determined by three peptide-binding prediction methods: artificial neural network (ANN) [23], stabilized matrix method (SMM) [16], and average relative binding (ARB) [19]. Predicted binding scores were expressed as IC<sub>50</sub> values, which represented the equilibrium dissociation constant (KD) of the peptide in relation to a particular MHC molecule. The binding affinities of all 9 and 10



amino acid peptide stretches in S358–381 were predicted. The data indicated that the binding of 9 aa peptides was stronger than all 10 aa peptides and that these 9 aa peptides bind with higher affinity to H-2 K<sup>d</sup> than to H-2 D<sup>d</sup> or H-2 L<sup>d</sup> (data not shown). Therefore, we concluded that the optimal epitope should be an H-2 K<sup>d</sup> restricted 9 amino acids peptide. In addition, the results demonstrated that S366–374 (CYGVSATKL) was the highest affinity peptide to H-2 K<sup>d</sup> (Table 1).

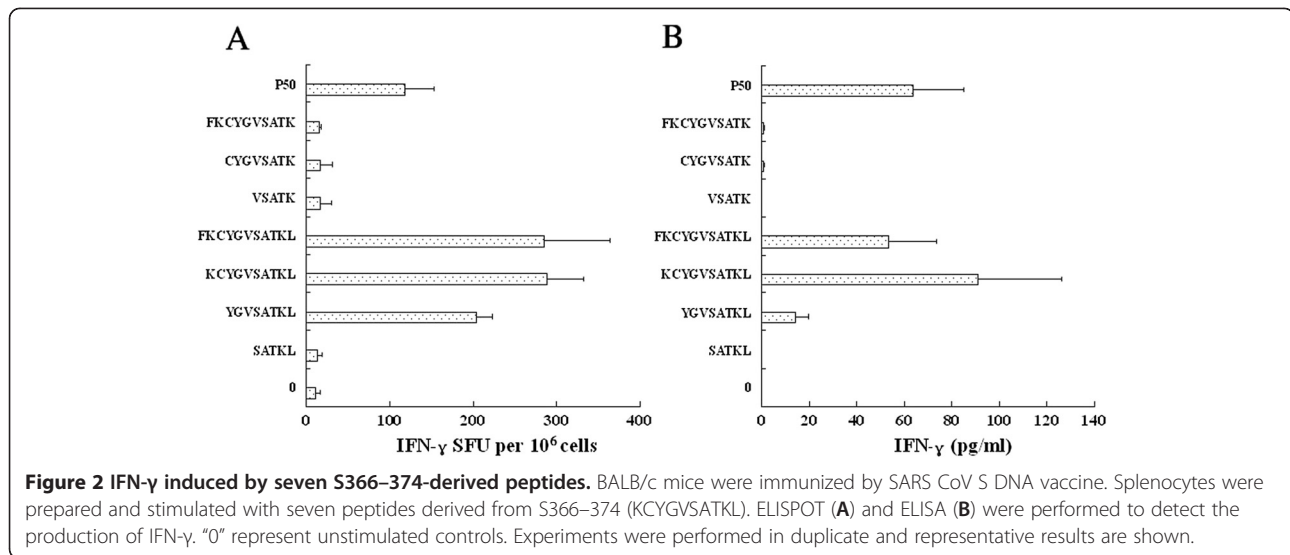
The epitope mapping tool BIMAS [22] (<http://www.bimas.cit.nih.gov>) was used to compare the binding kinetics of three peptides containing L<sup>374</sup> as the N-terminal peptide (S366–374, S365–374, and S367–374). The half time of disassociation from H-2 K<sup>d</sup> of each of these molecules was estimated. The score for S366–374 (CYGVSATKL) was 2880, much higher than for S365–374 and S367–374 (Table 2).

#### Predicted binding affinity of the S366–374 series with one mutated residue

To elicit an effective IFN- $\gamma$  response, an epitope must bind to an MHC molecule first. Artificial neural network (ANN) was used to estimate the binding affinity of single S366–374 mutants to H-2 K<sup>d</sup> (Table 3). The IC<sub>50</sub> value of wild type S366–374 (CYGVSATKL) was 59.2 nM. When Y<sup>367</sup> and L<sup>374</sup> were replaced by alanine (A), lysine (K), or aspartic acid (D), the IC<sub>50</sub> value increased dramatically (to more than 28,000 nM), indicating that Y<sup>367</sup> and L<sup>374</sup> were important for peptide binding and might serve as the main anchors in this epitope.

Although G<sup>368</sup>A (G<sup>368</sup> replaced by A) and G<sup>368</sup>K isoforms of S366–374 possessed higher affinity than wild type S366–374, they could not elicit IFN- $\gamma$  responses (Figure 3), indicating that G<sup>368</sup> might directly contact the T-cell receptor. The residues V<sup>369</sup>, A<sup>371</sup>, and K<sup>373</sup>





had functions similar to  $G^{368}$ , while the role of  $S^{370}$  was distinct. The  $IC_{50}$  value of  $S^{370}A$  was 216.5 nM, about 4 times higher than wild type S366–374. The  $IC_{50}$  values of  $S^{370}K$  and  $S^{370}D$  were nearly 1000 nM, indicating that  $S^{370}$  might be a weak anchor to H-2  $K^d$ . The role of  $T^{372}$  was similar to  $S^{370}$ , as the  $IC_{50}$  values of  $T^{372}A$ ,  $T^{372}K$ , and  $T^{372}D$  ranged from nearly 1000 nM to 3000 nM.

**Table 1 Predicted MHC-peptide binding**

Sequence	ANN	SMM	ARB
STFFSTFKC	38169.9	44479.5	16214.6
TFFSTFKCY	34824.9	468823.1	191908.2
FFSTFKCYG	36550.6	165788.7	202312.4
FSTFKCYGV	37493.9	270563	24611.6
STFKCYGVS	37025.3	36095.8	11060.5
TFKCYGWSA	36950.6	857351.7	1000000
FKCYGVSAT	28781	9145.4	180.1
KCYGVSATK	36702.2	4243.3	476309.3
CYGVSATKL	59.2	84.1	3.7
YGVSATKLN	39136.2	114095.1	1000000
GVSATKLND	39141.7	224004.7	1000000
VSATKLNDL	26880.9	23063.8	59717.3
SATKLNDLC	35465.1	10192.3	206549.5
ATKLNDLCF	39187.4	4083163	1000000
TKLNDLCFS	38468.4	807152.2	357259.1
KLNDLCFSN	37970.1	12958.7	203755.1

The binding affinities of 9 aa peptides in S358–382. (STFFSTFKCYGVSATKLNDLCFSN) to H-2  $K^d$  were predicted by computer algorithms. Numbers in the table are values of  $IC_{50}$  (nM) determined by ANN, SMM, and ARB.  $IC_{50}$  = 50% inhibitory concentration (low  $IC_{50}$  values indicate high affinity binding).

**The production of IFN-γ induced by one residue-mutated S366–374**

To confirm the key residues in S366–374, 27 analogs were synthesized in which each of the nine residues was replaced by A, K, or D. These peptides were used to stimulate splenocytes from SARS CoV S DNA immunized mice. ELISPOT (Figure 3A) and ELISA (Figure 3B) were performed to detect IFN-γ production.

The frequency of IFN-γ producing cells induced by  $K^{365}A$  or  $K^{365}D$  was similar to that induced by wild type S365–374 (KCYGVSATKL), and there was no obvious difference in the IFN-γ levels in the culture cell supernatants following peptide treatment. When  $C^{366}$  was replaced by A, the frequency of IFN-γ producing cells and the IFN-γ levels in the supernatant were also similar to wild type S365–374. When  $C^{366}$  was replaced by K or D, the frequency of IFN-γ producing cells was decreased by nearly 50%, and the level of IFN-γ in the supernatant was much lower than that induced by wild type S365–374. The  $S^{370}$  residue appeared to be more important than  $C^{366}$  for induction of IFN-γ as IFN-γ responses induced by  $S^{370}A$  were similar to that induced by wild type S365–374, while treatment with  $S^{370}K$  and  $S^{370}D$  peptides resulted in a 75% lower frequency of IFN-γ producing cells and a significantly reduced supernatant IFN-γ concentration compared to wild type S365–374.

**Table 2 Estimated half time of disassociation of peptides**

Sequence	Score
YGVSATKL	96
CYGVSATKL	2880
KCYGVSATKL	69.12

Score means Estimate of Half Time of Disassociation of the Subsequence.

**Table 3 Predicted MHC-I binding affinities of A-, K-, and D-substituted S366–374 peptides**

Sequence	IC50 (nM)	Sequence	IC50 (nM)	Sequence	IC50 (nM)			
C <sup>366</sup> A	AYGVSATKL	66	C <sup>366</sup> K	KYGVSAATKL	15.5	C <sup>366</sup> D	DYGVSAATKL	551.5
Y <sup>367</sup> A	CAGVSATKL	25435	Y <sup>367</sup> K	CKGVSATKL	22549	Y <sup>367</sup> D	CDGVSATKL	28232
G <sup>368</sup> A	CYAVSATKL	7.3	G <sup>368</sup> K	CYKVSATKL	18.5	G <sup>368</sup> D	CYDVSATKL	150
V <sup>369</sup> A	CYGASATKL	13.8	V <sup>369</sup> K	CYKGSATKL	69.6	V <sup>369</sup> D	CYDGSATKL	791.1
S <sup>370</sup> A	CYGVAATKL	216.1	S <sup>370</sup> K	CYGVKATKL	1118	S <sup>370</sup> D	CYGVDATKL	959.9
	CYGVSAATKL	59.2	A <sup>371</sup> K	CYGVSKTKL	40.3	A <sup>371</sup> D	CYGVSDTKL	112.9
T <sup>372</sup> A	CYGVSAAKL	1995	T <sup>372</sup> K	CYGVSAKKL	3122	T <sup>372</sup> D	CYGVSAADKL	882.4
K <sup>373</sup> A	CYGVSAATL	13.3		CYGVSAATKL	59.2	K <sup>373</sup> D	CYGVSAATDL	38.5
L <sup>374</sup> A	CYGVSAATKA	9265	L <sup>374</sup> K	CYGVSAATKK	22172	L <sup>374</sup> D	CYGVSAATKD	25522

Numbers in the table are values of IC<sub>50</sub> (nM) predicted by artificial neural network (ANN).

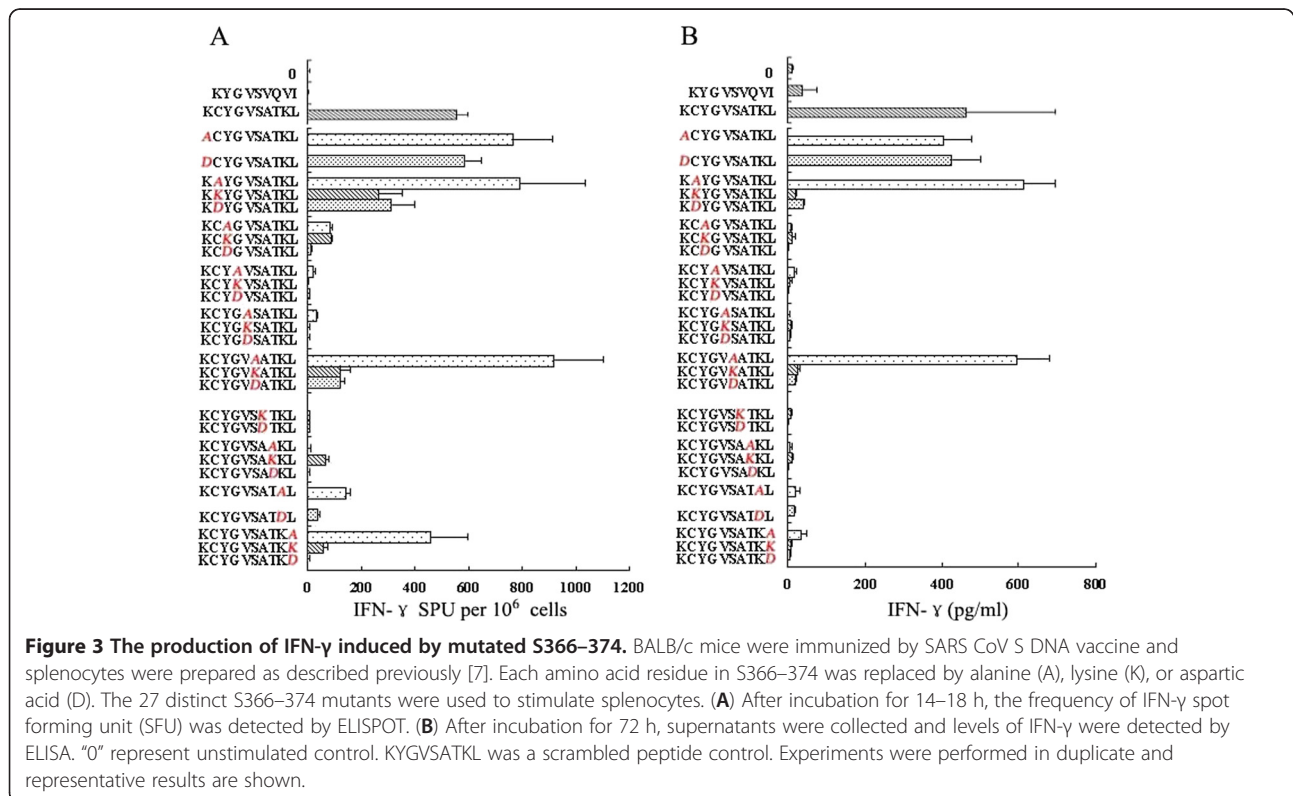
The L<sup>374</sup> was also an essential residue in the epitope. No IFN-γ response was detected in mutated peptides without L<sup>374</sup> (Figure 2). However, when L<sup>374</sup> was replaced by A, IFN-γ responses could still be detected. The frequency of IFN-γ producing cells following L<sup>374</sup>A treatment was about half that induced by wild type S365–374. When L<sup>374</sup> was replaced by K, some IFN-γ spot forming units (SFUs) could also be detected. In addition, Y<sup>367</sup> was important for IFN-γ induction; although about 200 SPU per 10<sup>6</sup> cells could be induced by Y<sup>367</sup>A or Y<sup>367</sup>K, the level of IFN-γ in the supernatant was significantly reduced compared to that induced by wild type S365–374. The K<sup>373</sup> residue had a role similar

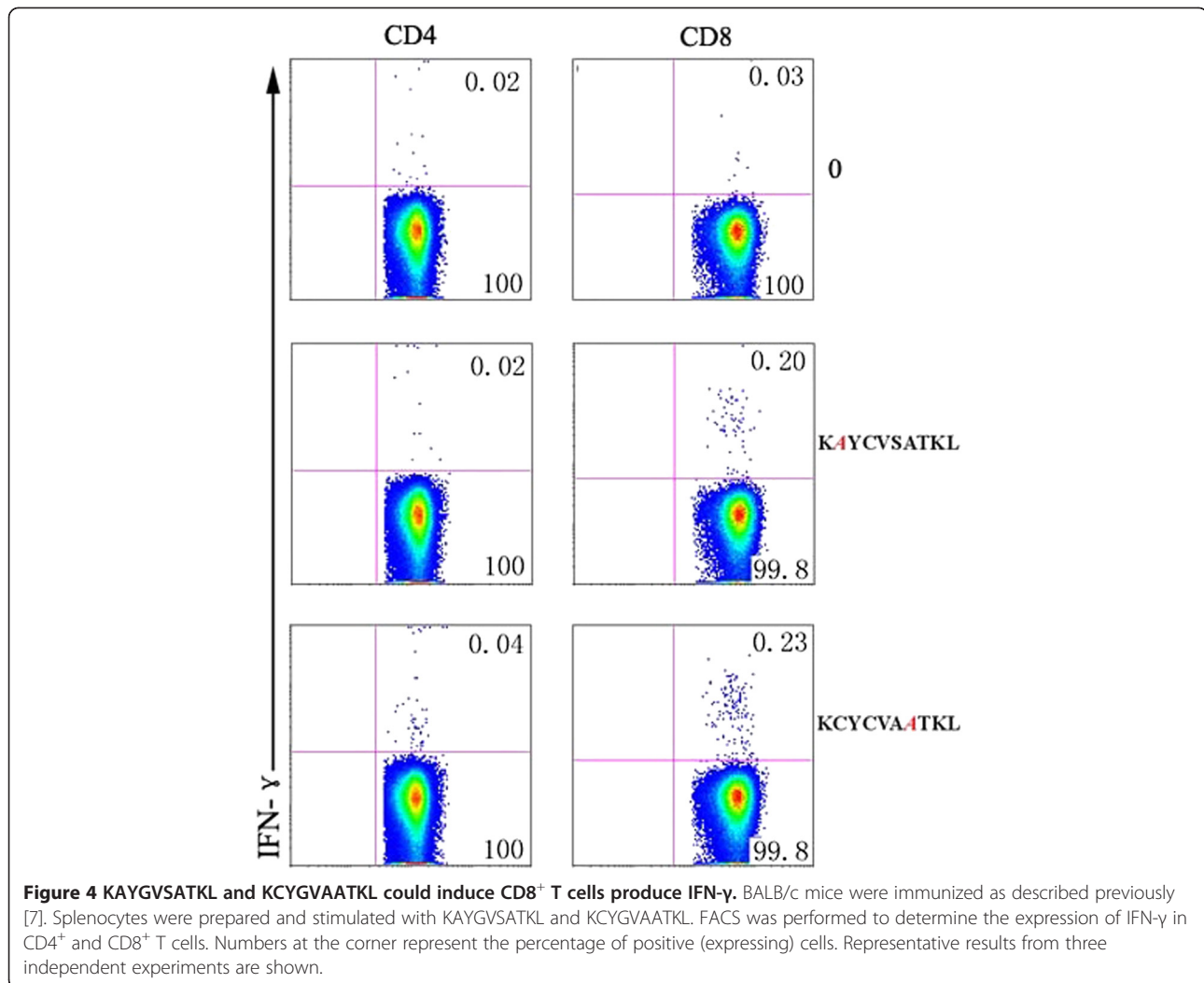
to Y<sup>367</sup>. When K<sup>373</sup> was replaced by A or D, the IFN-γ responses were similar to those induced by Y<sup>367</sup> mutated peptides.

The most important residues in the epitope were G<sup>368</sup>, V<sup>369</sup>, A<sup>371</sup>, and T<sup>372</sup>. Almost no IFN-γ response could be detected when these residues were replaced by A, K, or D.

#### C<sup>366</sup>A and S<sup>370</sup>A are H-2 K<sup>d</sup> restricted epitopes

As shown in Figure 4, IFN-γ responses induced by C<sup>366</sup>A and S<sup>370</sup>A were similar to that elicited by wild type S365–374, indicating that C<sup>366</sup> and S<sup>370</sup> could be replaced by A without affecting peptide function. To





further confirm this result, fluorescence activated cell sorting (FACS) was performed. The result showed that C<sup>366</sup>A and S<sup>370</sup>A could only stimulate CD8<sup>+</sup> T cells to produce IFN- $\gamma$ , indicating that both were H-2 K<sup>d</sup> restricted epitopes.

## Discussion

The T cell epitopes of the SARS CoV spike protein are well known, but systematic evaluation of the functional and structural roles of each residue has not been reported for these antigenic epitopes [24-26]. In the preliminary study [7], by using a synthesized peptide pool of SARS CoV S protein to stimulate the splenocytes from SARS CoV S DNA vaccine immunized mice, we identified that the peptides of P50 and P51 could induce IFN- $\gamma$  responses. P50 and P51 contained a same amino sequence, N50 (S365-374, KCYGVSATKL). In present study, N50 was synthesized and used to stimulate the splenocytes from immunized mice, ELISA, ELISPOT and FACS results all indicated that N50 contained a main

CD8<sup>+</sup> T cell epitope (Figure 1). Moreover, S366-374 (CYGVSATKL) was shown to be an optimal H-2 K<sup>d</sup> restricted epitope of the SARS CoV S protein by both bioinformatics prediction (Tables 1 and 2) and a functional INF- $\gamma$  release assay (Figure 2, Figure 3).

To induce effective T cell responses, the T cell epitope must bind to an MHC molecule. All peptides that bind to class I molecules contain a carboxyl-terminal anchor [27,28]. The anchor residues at both ends of the peptide are buried within the binding cleft, holding the peptide firmly in place. A previous study demonstrated that nonameric peptides bound preferentially and that the main contacts between class I MHC molecules were residue 2 at the amino-terminal end and residue 9 at the carboxyl terminus of the nonameric peptide. These anchors were generally hydrophobic residues (e.g., leucine and isoleucine) [29].

Site-directed mutagenesis is a powerful tool for probing protein or peptide structure and function. Alanine-scanning, lysine-scanning, or aspartic acid-scanning by

systematic replacement of side-chains with alanine, lysine, or aspartic acid have been used widely to study binding sites on proteins [22,30]. Analysis of the functional importance of side-chains by mutational study may exaggerate the effect by imposing a structural disturbance or an unusual steric, electrostatic, or hydrophobic interaction. Alanine-scanning is the least disruptive to the peptide structure because alanine is uncharged and has the smallest amino acid side group next to glycine, and thus has been particularly useful for mapping protein functional domains. In contrast, glycine can change the main chain conformation of the protein [31], so individual alanine mutations are preferred to infer the roles of individual amino acid residues. Charged residues such as lysine and aspartic acid are capable of forming ion pairs and hydrogen bonds, so they often play an important role in protein binding as well as in the recognition of interacting proteins. Thus, lysine and aspartic acid are often used as substitutes to study electrostatic effects between proteins [32].

Computational prediction and modeling of MHC/peptide binding can greatly facilitate peptide screening, with tremendous savings in time and experimental effort. Using these methods, T cell epitopes in many vaccine candidates were identified [33,34]. A number of prediction servers are available for identification of peptides that bind MHC molecules. Lin et al. have reported a comparative evaluation of thirty prediction servers for seven human MHC-I molecules. It showed that the best prediction server across all HLA molecules in this study is NETM\_ANN, closely followed by IEDB ANN and IEDB SMM [35,36].

In this study, both peptide-binding prediction methods and functional experiments were used to evaluate the roles of the different residues within the optimal epitope S365–374. Though K<sup>365</sup> was excluded from this optimal epitope (Tables 1 and 2), the wild type S365–374 containing K<sup>365</sup> induced a strong IFN- $\gamma$  response (Figure 2). Thus, K<sup>365</sup> was included in these synthesized mutant peptides. A comparison of IFN- $\gamma$  responses showed that mutants K<sup>365</sup>A or K<sup>365</sup>D were as potent as wild type S365–374 (Figure 3), demonstrating that K<sup>365</sup> is not in this optimal epitope and that residue in this position do not influence the function of S366–374.

The L<sup>374</sup> at position P9 was predicted to be the carboxyl-terminal anchor in the epitope by ANN (Table 3) and the functional experiment confirmed that analogs without L<sup>374</sup> could not induce IFN- $\gamma$  secretion from immunized mouse splenocytes (Figure 2). However, L<sup>374</sup>A could still induce significant IFN- $\gamma$  secretion, while L<sup>374</sup>K and L<sup>374</sup>D could not, suggesting that electrostatic effects and hydrophobic interaction may play an important role at this position.

Residue 2 at the amino-terminal end is another important anchor in many epitopes. In S366–374, Y<sup>367</sup> at position P2 was predicted to be another important anchor by ANN (Table 3). The functional assay showed that Y<sup>367</sup> mutated peptides (Y<sup>367</sup>A and Y<sup>367</sup>K) could induce IFN- $\gamma$  responses, but that IFN- $\gamma$  production was lower than that induced by S366–374. That might indicate an important role for the Y side chain in determining the binding affinity to H-2 K<sup>d</sup> [10].

The S<sup>370</sup> at P5 is also a functionally significant residue in this epitope. Though not a traditional anchor, bioinformatics tools indicated that it might act as a weak anchor for H-2 K<sup>d</sup> binding. The functional assay indicated that IFN- $\gamma$  responses induced by S<sup>370</sup> mutant epitopes were stronger than L<sup>374</sup> mutant epitopes.

X-ray crystallographic analyses of peptide-class I MHC complexes has revealed how the peptide-binding cleft in a given MHC molecule interacts stably with peptides [37]. Vesselin Mitaksov described the crystal structure of the MHC class I protein H2-K<sup>d</sup> in complex with the antigenic peptide TYQRTRALV (Flu) derived from an influenza nucleoprotein. They found that Flu residues Tyr<sup>P2</sup>, Thr<sup>P5</sup>, and Val<sup>P9</sup> were sequestered into the B, C, and F pockets of the K<sup>d</sup> groove, respectively [32]. In the sequence of S366–374, the positions Tyr<sup>367</sup>, Ser<sup>370</sup> and Leu<sup>374</sup> were at P2, P5, and P9 as well, implying that Tyr<sup>367</sup>, Ser<sup>370</sup> and Leu<sup>374</sup> are anchors in this optimal CTL epitope.

All three bioinformatics tools indicated that C<sup>366</sup> was included in this optimal epitope (Table 1). In its absence, however, S367–374 could still induce an IFN- $\gamma$  response, albeit weaker than that induced by wild type S365–374 (Figure 2). Thus, C<sup>366</sup> could influence the function of this epitope. The ANN tool predicted that C<sup>366</sup> was not an anchor to H-2K<sup>d</sup> and C<sup>366</sup> mutant peptides could still induce strong IFN- $\gamma$  responses, especially C<sup>366</sup>A, so our results indicate that C<sup>366</sup> is not an important residue for the peptide-TCR interaction. Conversely, mutations of G<sup>368</sup>, V<sup>369</sup>, A<sup>371</sup>, T<sup>372</sup>, and K<sup>373</sup> induced only modest IFN- $\gamma$  production, demonstrating that these residues made greater contributions in presenting to TCR in this optimal CTL epitope. As predicted by ANN, T<sup>372</sup> contributed to H-2 K<sup>d</sup> binding (Table 3). Indeed, IFN- $\gamma$  responses induced by T<sup>372</sup> mutant peptides were significantly weaker than peptides containing T<sup>372</sup>, indicating that T<sup>372</sup> might contribute to the TCR interaction.

## Conclusions

In the present study, we demonstrate that S366–374 is an optimal H-2 K<sup>d</sup> CTL epitope in the SARS CoV S protein. Moreover, Y<sup>367</sup>, S<sup>370</sup>, and L<sup>374</sup> are anchors in the epitope, while C<sup>366</sup>, G<sup>368</sup>, V<sup>369</sup>, A<sup>371</sup>, T<sup>372</sup>, and K<sup>373</sup> may directly interact with TCR on the surface of CD8-T cells.



## Abbreviations

SARS: Severe acute respiratory syndrome; CoV: Coronavirus; S: Spike protein; SFU: Spot forming unit; FACS: Fluorescence activated cell sorter; CTL: Cytotoxic lymphocyte; BFA: Brefeldin A; ANN: Artificial neural network; SMM: Stabilized matrix method; ARB: Average relative binding.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

JH carried out the immunoassays and helped to draft the manuscript. YC JH carried out peptide synthesis, purification and analysis. XB participated in its design and coordination and helped to draft the manuscript. CW conceived of the study, and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

We are grateful to Drs. Gary J. Nabel, Zhiyong Yang, Richard A. Koup, and Robert T. Bailer, VRC, NIAID, NIH, USA for kindly providing us with SARS CoV S DNA plasmid and SARS CoV S peptides. This work was supported by grants from the Guangdong Recruitment Program of Creative Research Groups, the National Natural Science Foundation of China (30340012), the Education Bureau of Guangdong Province, and Scientific and Technological Foundation of Guangzhou (2003Z3-E0491).

## Author details

<sup>1</sup>Institute of Immunology, Zhongshan School of Medicine; Key Laboratory of Tropical Disease Control Research of Ministry of Education, Sun Yat-sen University, Guangzhou, China. <sup>2</sup>School of Pharmaceutical Science, Sun Yat-sen University, Guangzhou, China. <sup>3</sup>Department of Pathogenic Biology and Immunology, Guangzhou Medical College, Guangzhou, China. <sup>4</sup>Department of Pharmacy, Xinhua College of Sun Yat-sen University, Guangzhou, China.

Received: 23 April 2012 Accepted: 31 August 2012

Published: 10 September 2012

## References

1. Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA, et al: **Identification of a novel coronavirus in patients with severe acute respiratory syndrome.** *N Engl J Med* 2003, **348**:1967–1976.
2. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, Lim W, Nicholls J, Yee WK, Yan WW, Cheung MT, et al: **Coronavirus as a possible cause of severe acute respiratory syndrome.** *Lancet* 2003, **361**:1319–1325.
3. Tse GM, To KF, Chan PK, Lo AW, Ng KC, Wu A, Lee N, Wong HC, Mak SM, Chan KF, et al: **Pulmonary pathological features in coronavirus associated severe acute respiratory syndrome (SARS).** *J Clin Pathol* 2004, **57**:260–265.
4. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, et al: **Characterization of a novel coronavirus associated with severe acute respiratory syndrome.** *Sci* 2003, **300**:1394–1399.
5. Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Khattri J, Asano JK, Barber SA, Chan SY, et al: **The genome sequence of the SARS-associated coronavirus.** *Sci* 2003, **300**:1399–1404.
6. He Y, Zhou Y, Siddiqui P, Jiang S: **Inactivated SARS-CoV vaccine elicits high titers of spike protein-specific antibodies that block receptor binding and virus entry.** *Biochem Biophys Res Commun* 2004, **325**:445–452.
7. Huang J, Cao Y, Du J, Bu X, Ma R, Wu C: **Priming with SARS CoV S DNA and boosting with SARS CoV S epitopes specific for CD4+ and CD8+ T cells promote cellular immune responses.** *Vaccin* 2007, **25**:6981–6991.
8. Laugel B, van der Berg HA, Gostick E, Cole DK, Wooldridge L, Boulter J, Milicic A, Price DA, Sewell AK: **Different T cell receptor affinity thresholds and CD8 coreceptor dependence govern cytotoxic T lymphocyte activation and tetramer binding properties.** *J Biol Chem* 2007, **282**:23799–23810.
9. Anderson MW, Gorski J: **Cutting edge: TCR contacts as anchors: effects on affinity and HLA-DM stability.** *J Immunol* 2003, **171**:5683–5687.
10. Lee Y, Ferrari G, Lee SC: **Estimating design space available for polypeptides through consideration of major histocompatibility complex binding motifs.** *Biomed Microdevices* 2010, **12**:207–222.

11. Yang ZY, Kong WP, Huang Y, Roberts A, Murphy BR, Subbarao K, Nabel GJ: **A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice.** *Nature* 2004, **428**:561–564.
12. Thompson RE, Jolliffe KA, Payne RJ: **Total synthesis of microcin B17 via a fragment condensation approach.** *Org Lett* 2011, **13**:680–683.
13. Meister GE, Roberts CG, Berzofsky JA, De Groot AS: **Two novel T cell epitope prediction algorithms based on MHC-binding motifs; comparison of predicted and published epitopes from Mycobacterium tuberculosis and HIV protein sequences.** *Vaccin* 1995, **13**:581–591.
14. Buus S, Lauemoller SL, Worning P, Kesmir C, Frimurer T, Corbet S, Fomsgaard A, Hilden J, Holm A, Brunak S: **Sensitive quantitative predictions of peptide-MHC binding by a 'query by committee' artificial neural network approach.** *Tissue Antigens* 2003, **62**:378–384.
15. Brusica V, Bajic VB, Petrovsky N: **Computational methods for prediction of T-cell epitopes—a framework for modelling, testing, and applications.** *Methods* 2004, **34**:436–443.
16. Peters B, Sette A: **Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method.** *BMC Bioinform* 2005, **6**:132.
17. Peters B, Tong W, Sidney J, Sette A, Weng Z: **Examining the independent binding assumption for binding of peptide epitopes to MHC-I molecules.** *Bioinform* 2003, **19**:1765–1772.
18. Doytchinova IA, Guan P, Flower DR: **EpIjen: a server for multistep T cell epitope prediction.** *BMC Bioinform* 2006, **7**:131.
19. Bui HH, Sidney J, Peters B, Sathiamurthy M, Sinichi A, Purton KA, Mothe BR, Chisari FV, Watkins DI, Sette A: **Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications.** *Immunogenet* 2005, **57**:304–314.
20. Peters B, Bui HH, Frankild S, Nielson M, Lundegaard C, Kostem E, Basch D, Lamberth K, Harndahl M, Fleri W, et al: **A community resource benchmarking predictions of peptide binding to MHC-I molecules.** *PLoS Comput Biol* 2006, **2**:e65.
21. Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peters B: **Peptide binding predictions for HLA DR DP and DQ molecules.** *BMC Bioinform* 2010, **11**:568.
22. Parker KC, Bednarek MA, Coligan JE: **Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains.** *J Immunol* 1994, **152**:163–175.
23. Nielsen M, Lund O: **NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction.** *BMC Bioinform* 2009, **10**:296.
24. Wang B, Chen H, Jiang X, Zhang M, Wan T, Li N, Zhou X, Wu Y, Yang F, Yu Y, et al: **Identification of an HLA-A\*0201-restricted CD8+ T-cell epitope SSp-1 of SARS-CoV spike protein.** *Blood* 2004, **104**:200–206.
25. Yang J, James E, Roti M, Huston L, Gebe JA, Kwok WW: **Searching immunodominant epitopes prior to epidemic: HLA class II-restricted SARS-CoV spike protein epitopes in unexposed individuals.** *Int Immunol* 2009, **21**:63–71.
26. Zhi Y, Kobinger GP, Jordan H, Suchma K, Weiss SR, Shen H, Schumer G, Gao G, Boyer JL, Crystal RG, et al: **Identification of murine CD8 T cell epitopes in codon-optimized SARS-associated coronavirus spike protein.** *Virology* 2005, **335**:34–45.
27. Boesteanu A, Brehm M, Mylin LM, Christianson GJ, Tevethia SS, Roopenian DC, Joyce S: **A molecular basis for how a single TCR interfaces multiple ligands.** *J Immunol* 1998, **161**:4719–4727.
28. Robinson RA, Lee DR: **Studies of tum- peptide analogs define an alternative anchor that can be utilized by Ld ligands lacking the consensus P2 anchor.** *J Immunol* 1996, **156**:4266–4273.
29. Malik A, Houghten R, Corradin G, Buus S, Berzofsky JA, Hoffman SL: **Identification of a nonameric H-2Kk-restricted CD8+ cytotoxic T lymphocyte epitope on the Plasmodium falciparum circumsporozoite protein.** *Infect Immun* 1995, **63**:1955–1959.
30. Mitaksov V, Fremont DH: **Structural definition of the H-2Kd peptide-binding motif.** *J Biol Chem* 2006, **281**:10618–10625.
31. Wells JA: **Additivity of mutational effects in proteins.** *Biochem* 1990, **29**:8509–8517.
32. Arabshahi A, Frey PA: **Standard free energy for the hydrolysis of adenylylated T4 DNA ligase and the apparent pKa of lysine 159.** *J Biol Chem* 1999, **274**:8586–8588.

33. Srinivasan KN, Zhang GL, Khan AM, August JT, Brusic V: **Prediction of class I T-cell epitopes: evidence of presence of immunological hot spots inside antigens.** *Bioinform* 2004, **20**(Suppl 1):i297–i302.
34. Soam SS, Khan F, Bhasker B, Mishra BN: **Prediction of MHC class I binding peptides using probability distribution functions.** *Bioinformation* 2009, **3**:403–408.
35. Lin HH, Ray S, Tongchusak S, Reinherz EL, Brusic V: **Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research.** *BMC Immunol* 2008, **9**:8.
36. Zhang GL, Srinivasan KN, Veeramani A, August JT, Brusic V: **PREDBALB/c: a system for the prediction of peptide binding to H2d molecules, a haplotype of the BALB/c mouse.** *Nucleic Acids Res* 2005, **33**:W180–W183.
37. Bolin DR, Swain AL, Sarabu R, Berthel SJ, Gillespie P, Huby NJ, Makofske R, Orzechowski L, Perrotta A, Toth K, *et al*: **Peptide and peptide mimetic inhibitors of antigen presentation by HLA-DR class II MHC molecules. Design, structure-activity relationships, and X-ray crystal structures.** *J Med Chem* 2000, **43**:2135–2148.

doi:10.1186/1471-2172-13-50

**Cite this article as:** Huang *et al.*: Residue analysis of a CTL epitope of SARS-CoV spike protein by IFN-gamma production and bioinformatics prediction. *BMC Immunology* 2012 **13**:50.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

