# ORIGINAL PAPER

# Identification of CTL epitopes in hepatitis C virus by a genome-wide computational scanning and a rational design of peptide vaccine

Toshie Mashiba • Keiko Udaka • Yasuko Hirachi • Yoichi Hiasa • Tomoya Miyakawa • Yoko Satta • Tsutomu Osoda • Sayo Kataoka • Michinori Kohara • Morikazu Onji

Received: 1 October 2006 / Accepted: 22 November 2006 / Published online: 16 January 2007 © Springer-Verlag 2007

**Abstract** Developing a peptide-based vaccine for the highly variable hepatitis C virus (HCV) remains a challenging task. Variant viruses not only escape antigen presentation but also persist in a patient as quasi-species. Such variants are often antagonistic to the responding T cell repertoire. To overcome these problems, we herein propose

**Electronic supplementary material** Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00251-006-0185-3 and is accessible for authorized users.

T. Mashiba · Y. Hiasa · M. Onji Department of Gastroenterology and Metabology, Graduate School of Medicine, Ehime University, Toh-on, Ehime, Japan

K. Udaka (⊠) · Y. Hirachi · S. Kataoka
Department of Immunology, Kochi Medical School, Nankoku,
Kochi 783-8505, Japan
e-mail: udaka@kochi-u.ac.jp

K. Udaka · Y. Hirachi · S. Kataoka
 CREST (Core Research for Evolutional Science and Technology),
 JST (Japan Science and Technology Agency),
 Kawaguchi, Saitama 332-0012, Japan

T. Miyakawa · T. Osoda Department of Bioinformatics Business Promotion, NEC corporation, Tokyo, Japan

Y. Satta Graduate University for Advanced Studies (Sokendai), Hayama, Japan

M. Kohara Department of Microbiology, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan a cocktail vaccine consisting of a few epitope peptides, which make it possible to outpace the emergence of variant viruses. To design such a vaccine, we developed a way to identify HLA-A\*2402-binding peptides efficiently by means of the computational scanning of the whole genome of the pathogen. Most of the predicted peptides exhibited strong binding to the HLA-A\*2402 molecule, while also inducing CD8 T cell responses from the patients' peripheral blood mononuclear cells (PBMCs). Peptide-induced T cells were capable of lysing HCV-expressing HepG2 cells which process antigens endogenously. The amount of HCV core antigen in the patients' livers suggested that the lytic activity of the peptide-induced T cells was clearly in a range suitable for therapeutic use. If T cells were activated under optimal conditions by high density peptides, then they tended to be relatively tolerant of single amino acid variations for cytolysis. Finally, an analysis of the viral population isolated in Japan suggested no obvious changes due to immune evasion in the viral genome even in a host population highly biased toward HLA-A\*2402.

Keywords Cytotoxic T lymphocyte  $\cdot$  Hepatitis C virus  $\cdot$  Peptide  $\cdot$  HLA  $\cdot$  Vaccine

### Introduction

Hepatitis C virus (HCV) remains a serious threat due to its persistence and the fact that it can also cause liver cirrhosis and cancer. Even the most advanced treatment combining pegylated interferon- $\alpha$  and ribavirin only has a sustained viral clearance rate of just more than 50%, and this falls even further with the HCV genotype 1 and for patients with other comorbidities (Feld and Hoofnagle 2005; Tsubota et al. 2004). The induction of cytotoxic T lymphocytes (CTLs) specific for HCV-infected cells has been a promising strategy for viral containment. A number of HLA class I-binding peptides have so far been identified (Battergay et al. 1995; Cerny et al. 1995; Kurokohchi et al. 2001; Nakamoto et al. 2003). However, finding HLAbinding peptides and screening them for T cell responses is a laborious, expensive, and time-consuming process. Therefore, HLA-binding has not yet been examined for a number of known T cell epitopes. However, if the affinity is low, even if a high density of exogenously added peptide may stimulate T cells, the infected hepatocytes may not present a sufficient number of peptides to be recognized by such T cells. Therefore, it is important to develop a method to design a peptide-based vaccine more efficiently and effectively.

HCV is an RNA virus with a high rate of mutation due to the absence of any proof-reading activity in its RNAdependent RNA polymerase. Viral escape from immune attack by epitope changes has been studied by a number of groups and remains a major concern in vaccine development (Ray et al. 2005; Timm et al. 2004; Weiner et al. 1995). The first type of such mutations leads to poor processing of MHC-binding peptides due to an alteration of the amino acid sequences for proteasome cleavage (Kimura et al. 2005; Seifert et al. 2004). The second type leads to poor binding of epitope peptides to the MHC class I molecules (Chang et al. 1997; Cox et al. 2005; Erickson et al. 2001). The third type induces altered responses known as antagonism or anergy in T cells due to a poor recognition by TCR (Chang et al. 1997; Cox et al. 2005; Erickson et al. 2001; Grakoui et al. 2003; Kaneko et al. 1997). The virus not only changes, but it also exists as a quasi-species, i.e., a mixed population of distinct, but closely related, variants. These variants are often antagonistic to the responding T cell repertoire, and thus, annihilate the T cell responses. The quasi-species of viruses are kept in a dynamic yet lasting balance, and thus, effectively suppress the cytolytic activities of a dominant T cell repertoire (Chang et al. 1997). Therefore, the key to a successful vaccine lies in a design that can eradicate the heterogeneous viral population before escaped mutants become prevalent.

## Materials and methods

#### Cells and antibodies

RzM6 is a HepG2 transfectant with a full genome HCV1b isolate AY045702 (Tsukiyama-Kohara et al. 2004). The expression of HCV can be induced by the tamoxifeninduced expression of the doubly transfected Cre that mediates the removal of an intervening sequence in the 5' promoter region of the HCV genome. C1R was a gift from Dr. P. Cresswell (Edwards et al. 1982). C1R-A24 is a HLA-A\*2402-transfected C1R cell line and was kindly provided by Dr. Takiguchi (Karaki et al. 1993). HepG2 was purchased from ATCC. T2-A24 is a TAP-deficient T2 cell line transfected with HLA-A\*2402 and was a gift from Dr. A. Tsuboi (Osaka University).

## Peptides

The peptides were manually synthesized using Fmoc chemistry and then were purified by HPLC to a purity of >95% using a C18 Microbondasphere column (Japan Waters, Tokyo). The peptides were examined by mass spectrometry using Voyager DE-RP (Applied Biosystems Japan, Tokyo). The concentrations were determined by a MicroBCA assay using bovine serum albumin (BSA) as the standard (Pierce, Rockford, IL).

## Peptide-binding assay

The binding of peptides to the HLA-A\*2402 molecule was measured by acid stripping and a reconstitution assay as previously described by Zeh et al. (1994) with minor modifications. Briefly, C1R-A24 cells were exposed to pH 3.3 citrate phosphate buffer and then were reconstituted with graded concentrations of peptide and 0.1 µM human β2-microglobulin (Sigma, M-4890, St. Louis, MS) in DMEM containing 0.25% BSA. An FITC-labeled mAb 17A12 (Tahara et al. 1990) was used to detect the properly folded and peptide-bound HLA-A\*2402 molecules. The fluorescence intensity was measured by FACScan (Becton-Dickinson Japan, Tokyo). Both highand low-binding peptides, HER2-63 TYLPTNASL and Met149 RVWE SATPL, respectively, were always included in the assay, and their binding was used to normalize the variations between experiments. The affinity of a peptide was calculated as previously described (Udaka et al. 2000).

#### Peptide-specific cell lines

The peripheral blood mononuclear cells (PBMCs) from patients or healthy individuals were stimulated weekly with 1  $\mu$ M peptide in 10% fetal calf serum (FCS) containing 10 U/ml recombinant human IL-2. After five time stimulations, cells were tested for killing activity. The patients had been diagnosed to be suffering from chronic hepatitis with HCV genotype 1b. All the patients were positive for both the anti-HCV antibodies and viral RNA in the serum by polymerase chain reaction (PCR). Informed consent was obtained from all patients. The study protocol was approved by the Human Research Committee of Ehime University.

# <sup>51</sup>Cr release assay

A peptide-specific cytolysis assay was conducted against C1R-A24 cells or T2-A24 cells in the presence or absence of 1  $\mu$ M peptide. An HLA-A\*2402-binding peptide, HER2-63 (TYLPTNASL, log Ka; 7.3), was used as a negative control. Target cells were labeled with <sup>51</sup>Cr-sodium chromate at either 37 or 26°C, respectively, for 1 h, and then, they were loaded with 1  $\mu$ M peptide before the addition of effector cells. The *E/T* ratio was 10–20. Percent specific lysis during 3.5 h incubation at 37°C was calculated as (experimental release–spontaneous release)/ (total release–spontaneous release)×100. The cytolysis of cells that naturally present HCV peptides was measured using tamoxifen-treated RzM6 cells as a target for 3 h at an *E/T* ratio of 10.

## Measurement of HCV core protein

Liver tissue specimens from the patients were taken by a biopsy after informed consent was obtained from all the patients. The cell lines and tissue specimens were lysed in Radioimmunoprecipitation assay (RIPA) buffer (1% sodium dodecyl sulfate (SDS), 1% NP40, 10 mM Tris–HCl, pH 8.0 and 0.14 M NaCl) and the supernatant was subjected to the measurement of core protein using an HCV Ag enzyme-linked immunosorbent assay (ELISA) kit (Ortho-Clinical Diagnostics, Tokyo, Japan). The protein concentrations were determined with a DC protein assay (BIO-RAD, Hercules, CA).

#### Results

Generation of a program to predict HLA-A\*2402-binding peptides

We have previously described a computational method to analyze the specificity of MHC class I-binding peptides (Udaka et al. 2002). The method utilizes a data mining technique, a query learning algorithm based on hidden Markov models (HMMs). This algorithm finds peptides whose binding properties are hard to predict by using a prototype prediction program established with the existing binding data (supplementary material 2). We synthesized such peptides and measured their binding to MHC molecules. By feeding the newly obtained binding data back into the data pool, the prediction program could thus be improved. This cyclic learning that combines data analysis and peptide-binding experiments is repeated until a satisfactory prediction can be achieved. This time, we examined the HLA-A\*2402 molecule, the most frequent allele among Asians, e.g., 33% for the Japanese (Tokunaga et al. 1997) and ~10% in the Western countries (Imanishi et al. 1992). The original HMMs employed a cyclic model (Udaka et al. 2002), but this was switched to a more rational parallel model during the learning (Fig. 1a). After 6 rounds of cyclic learning, examining 400 peptides altogether, i.e., 222 before active learning and 178 newly synthesized during learning, a final prediction program was established. However, further learning was still possible when necessary. The binding data for an additional 105 peptides that had been set aside to monitor the progress of the learning were also added to the data pool to generate the final prediction program. This program predicts the binding affinity by approximating the log  $K_a$  values, and thus, gives each peptide a score in real numbers. The performance of the program was assessed by 10-fold crossvalidation using the binding data on 505 peptides described above. The coefficient of the correlation between the predicted scores and actual affinity was 0.80 (Fig. 1b). Most of the known T cell epitope peptides have an affinity of 5.5 or higher in the log  $K_a$  terms. Among the peptides whose predicted scores were 5.5 or higher, 93% (117/126) actually exhibited a log  $K_a$  value of 5.5 or higher. Therefore, the accuracy (sensitivity) of the prediction was 93%. On the other hand, among those peptides whose affinity was 5.5 or higher, 60% (113/187) had scored 5.5 or higher. Therefore, the coverage was 60%. If the threshold had been raised to 6.0, then the accuracy would have been 93% (67/72) and the coverage thus would have fallen to 43% (67/156).

Genome-wide screening of hepatitis C virus for HLA-A\*2402-binding peptides

Using the program developed above, we scanned the entire genome of hepatitis C virus (HCV) genotype 1b for HLA-A\*2402-binding peptides. We chose a prototype Japanese isolate GenBank D20908 and a subclone of D89815 for the analysis. The latter was chosen due to our initial plan to use cells transduced with that clone as a target. Several amino acids differ between D89815 and its subclone pBRT703'X (Dr. Y. Matsuura, personal communication). The highscoring peptides were synthesized and subjected to HLAbinding assays. Some peptides (~7% of the peptides synthesized) were hard to synthesize, and thus, were excluded. Several known epitope peptides also scored high, and their binding was examined. The results are shown in Table 1. High-binder peptides (log  $K_a > 5.5$ ) are in bold. The major anchor amino acids identified by Rammensee et al. for HLA-A\*2402 are Y/F at P2 and I/L/F at P9 (http:// www.syfpeithi.de/). The peptides in italics do not fulfill



b



Fig. 1 The analysis and prediction of HLA-A\*2402-binding peptides. a A hidden Markov model used for analyzing the specificity of HLA-A\*2402-binding peptides. A general description of the model and the data mining algorithm are given in the reference (Udaka et al. 2000) and supplementary material 2. Six independent paths were designed. The first node S0 depicts the initial state and the last node S57 designates the final state. In the states from S1 to S56, the frequency of every amino acid is calculated to give the sum of the probabilities for 20 amino acids to be one in each state. Arrows depict transition probabilities. The states S1 and S56 were introduced to accommodate lateral shifts of the binding motif by one amino acid in both directions. Nine states in each row depict amino acid positions in a peptide of nine amino acids starting from the N-terminal amino acid on the left. The initial state can take any path and S1 in the case of an N-terminal shift of the binding motif. The frequency of transition was tentatively given as 0.1429 (=1 / 7) for each arrow, but it is replaced by an actual frequency once the model is trained using peptide-binding data. The final state can be reached from any path and from S56. The probability of transition from a row to the final state is tentatively given as 0.5 for each arrow in the illustration, but this too is replaced by actual probability. b The correlation of the predicted HLA-A\*2402-binding scores and experimentally determined binding of peptides. The correlation was examined by a tenfold cross validation using the peptide-binding data obtained during cyclic training of the models

two anchor requirements. The high-binder peptides could also be identified among such peptides (underlined peptides).

Responses of the PBMCs from chronic hepatitis patients

Due to the limits of the cell culture, 15 peptides (log  $K_a > 6$ ) were randomly chosen out of the peptides noted in Table 1 and then were tested for the induction of cytotoxic T lymphocytes (CTLs). PBMCs from patients with chronic hepatitis due to HCV 1b and from healthy individuals were stimulated with peptides and tested for cytolysis by <sup>51</sup>Cr release assay. A poor cytolytic activity of CD8 T cells despite a robust production of IFN- $\gamma$  have been observed among both tumor-bearing patients and patients suffering from chronic viral infections (Appay et al. 2000; Huang et al. 2005; Wherry et al. 2003). Therefore, in this study, with the aim of developing a curative vaccine, we used a cytolysis assay. As shown in Fig. 2a and b, most of the peptide-stimulated cell lines exhibited cytotoxicity. Interestingly, there were several peptides to which many patients responded. These peptides can thus be good targets for immunotherapy. Healthy individuals occasionally exhibited some responses but less frequently than the patients (Fig. 2a,b). Most patients exhibited cytolytic activities, often to several peptides. This strongly indicates that in the patients, a HCV-specific T cell repertoire has expanded and the cytolytic activity can therefore be induced if antigenic peptides are provided in a stimulatory environment in vitro.

Peptide-specific cell lines were further examined for restricting HLA molecules and peptide dependency for

 Table 1
 Prediction of HLA-A\*2402-binding peptides in the HCV 1b
 genome

Protein	Peptide	Sequence	Predicted score	Actual binding normalized log	
				K <sub>a</sub>	
Core	C36shpB <sup>a</sup>	LLPRRGPRL	5.30	7.71	
	C77shpB	AQPGYPWPL	5.98	5.36	
	C169shpB	LPGCSFSIF	4.98	4.91	
E1	C234sh	NFSRCWVAL	5.46	5.92	
	C360shpB	AYYSMVGNW	5.73	6.47	
E2	C616shpB	<b>WHYPCTVNF</b>	5.66	6.39	
	C666shpB	<u>LLSTTEWQI</u>	5.35	8.34	
	C674sh	ILPCSFTTL	6.75	7.66	
	C688pB	<u>GLIHLHQNI</u>	6.00	5.86	
	C764pB	GILPFFMFF	6.26	7.70	
	C789sh	ALYGVWPLL	5.80	6.98	
	C789pB	AFYGVWPLL	6.02	7.69	
NS2	C834sh	YYKVFLARL	5.49	7.12	
	C838shpB	FLARLIWWL	5.24	6.18	
	C876pB	<i>LMCAVHPEL</i>	5.52	6.59	
	C975sh	VFSDMETKL	5.54	7.31	
	C992shpB	ACGDIISGL	5.26	3.91	
	C1010shpB	ILLGPADSF	5.43	5.50	
NS3	C1031shpB	AYSQQTRGL <sup>b</sup>	5.76	6.54	
	C1291sh	ITYSTYCKF	5.27	6.98	
	C1291pB	ITYSTYGKF	5.16	6.37	
	C1349pB	ATPPGSVTF	6.53	6.51	
NS4B	C1760shpB	<u>FWAKHMWNF<sup>c</sup></u>	5.93	8.10	
	C1956shpB	LLKRLHQWI	5.20	6.86	
NS5A	C1976pB	<u>WLRDVWDWI</u>	5.27	6.34	
	C1986pB	<i>TVLADFKTW</i>	6.25	6.52	
	C1987pB	VLADFKTWL	5.26	6.63	
	C2132sh	RYAPVCKPL	5.96	6.15	
	C2132pB	RYAPACKPL	5.47	6.76	
	C2139pB	PLLRDEVTF	6.39	5.09	
	C2173shpB	<u>SMLTDPSHI</u>	5.55	6.94	
	C2251sh	VILDSFDPI	6.05	5.32	
	C2251pB	VILDSFEPL	6.26	5.00	
	C2289shpB	ARPDYNPPL	5.47	3.93	
NS5B	C2422shpB	SYTWTGALI	5.51	7.13	
	C2593shpB	ALYDVVSTL	5.99	6.39	
	C2841shpB	<u>RMILMTHFF<sup>c</sup></u>	6.01	7.42	
	C2843shpB	ILMTHFFSI	5.57	7.90	
	C2844shpB	<i>LMTHFFSIL</i>	5.20	5.92	
	C2962shpB	SQLDLSGWF	5.66	< 3	

High-affinity peptides (log  $K_a > 5.5$ ) are shown in bold. The sequences in italics represent the peptides that do not fulfill two major anchor requirements. The high-affinity peptides that would not have been identified by the criteria of two major anchors are underlined.

<sup>b</sup> Epitope reported by Kurokohchi et al. (2001)

<sup>c</sup> Epitopes reported by Nakamoto et al. (2003)

<sup>&</sup>lt;sup>a</sup> The number refers to the position of the first amino acid. The suffix sh stands for a sequence from Genbank D20908, and pB, that from the subclone pBRT703'X of D89815. A substantial number of nucleotide mismatches have been identified between pBRT703'X and D89815 (Matsuura, personal communication).

Fig. 2 The cytolytic activities of the peptide-induced cell lines from patients and healthy individuals. Peptide-specific cell lines were established from PBMCs by weekly stimulations with peptides. Each bar demonstrates lytic activity by a cell line from an individual. The assay used allogeneic C1R-A24 in a or T2-A24 cells in **b** as a target and was performed in the presence of indicated peptide at a concentration of 1 µM. These target cells exhibited substantial background lysis either without any peptide or with a negative control peptide HER2-63. Therefore, specific lysis is shown as the value from which the background lysis with HER2-63 has been subtracted. Nonoverlapping responders were used for a and b. The first (open) and third (dark gray) bars in **a** and the fourth (*closed*) bar are cell lines from healthy individuals



recognition. The responding cell lines were predominantly  $CD8^+$  T cells as is shown in Fig. 3a. When tested against C1R or C1R-A24 target cells, several cell lines exhibited a stronger cytolytic activity against C1R-A24 than C1R in the presence of peptides (Fig. 3b). It was not possible to perform this assay for all the cell lines shown in Fig. 2a,b due to the limited number of cells. For a few cell lines, the cytolysis of C1R was too high for one to see any peptide dependency. This was most likely due to some alloreactivity or NK-like activity because we used a low level of rIL2 during the cell culture.

Cytolytic activity of the peptide induced cell lines against cells naturally presenting the antigen

Although the peptide-induced T cell lines were highly cytolytic against tumor targets heavily loaded with exogenous peptides, it is not clear whether these effecter cells are also capable of killing target cells that present as endogenous antigens through natural antigen processing. To examine this, we tested a HepG2 transfectant RzM6 (Tsukiyama-Kohara et al. 2004) as a target. RzM6 carries a full-length HCV 1b genome (GenBank AY045702). HepG2 naturally expresses HLA-A\*2402. The expression of the HCV gene from the CAG promoter can be conditionally induced by the Cre-mediated removal of the floxed intervening sequence that has been inserted in the 5' terminal region. The RzM6 cells had been doubly transfected with the Cre expression construct whose expression could be induced by tamoxifen. The precise 5' and 3' trimming at the ribozyme sequences eventually produces a full-length HCV. As is shown in Table 2, the HLA-A\*2402binding peptides expressed in RzM6 differ by several amino acids from the HCV isolates shown in Table 1. Tamoxifen-treated RzM6 cells were lysed by most of the peptide-specific cell lines established as shown in Fig. 4. This indicated that most of the peptides identified as HLA-A\*2402-binders from genomic sequences were actually processed endogenously and then presented on the cell surface. Interestingly, most of the peptide-induced CTL lines could lyse RzM6 cells whose epitope sequences carry amino acid substitution(s) compared with the peptides used for stimulation. Although such variant peptides could be antagonistic to part of the CTL repertoire (Kaneko et al. 1997) once the T cells were optimally activated by highdensity peptides, then the responding T cells were also found to be cytolytic for such variant peptides. This is encouraging to those developing a peptide vaccine against highly variable pathogens like HCV. Interestingly, the peptide C616shpB is located in a sequence context which is not ideal for proteasomal cleavage (Nielsen et al. 2005; Nussbaum et al. 2001). However, this peptide also seems to

Fig. 3 Peptide-specific cell lines recognize peptides in the context of HLA-A\*2402. a Cell surface profile of a cell line specific for C2422shpB shown as an example. The enrichment of CD8 cells was obvious for cell lines that exhibited peptidedependent lysis after 3 to 5 weekly stimulations with HLA-A\*2402-binding peptides. b Specific lysis of C1R (open bars) or C1R-A\*2402 (closed bars) target cells in the presence of peptides used for the stimulation of the corresponding cell lines. Lysis without peptide has been subtracted



be presented on RzM6. Although inefficient proteasomal cleavage has been reported to be one of the mechanisms for viral escape (Seifert et al. 2004), not all the peptides that are predicted to demonstrate poor cleavage seem to be spared from antigen presentation. Under normal conditions, peptides are in relatively short supply in comparison to the newly synthesized MHC class I molecules in the endoplasmic reticulum (ER) (Lie et al. 1990). Slowly cleaved peptides may still have a chance to bind to empty MHC class I molecules.

HCV peptide presentation in the infected hepatocytes

Importantly, these peptide-specific T cells have, however, not been sufficient to contain viral infections in the patients. The expression of HCV proteins in tamoxifen-treated RzM6 cells (RzM6-Tx) may have been higher than that in the infected hepatocytes of the patients. We, therefore, examined next the expression of core protein as a representative antigen in the patients' livers. As shown in Fig. 5, the expression of core protein of HCV 1b was high 1b clones analyzed in this study

Peptides D20908 pBRT703'X RzM6 C36 LLPRRGPRL C616 WHYPCTVNF - - - - - A - -- - - - - -**LLSTTEWOI** C666 - - V ILPCSFTTL C674 V - -C789 ALYGVWPLL - F - - - - - - -C834 YYKVFLARL - - - E - - - - -- - - - - - -C975 VFSDMETKL - - - - V - - - V C1031 AYSOOTAGL . . . . . . . . . - - - - -----G--C1291 ITYSTYCKF - - - G - -C1760 FWAKHMWNF - - - - - -- - - - - - - - -C1956 LLKRLHOWI - - H - -C2132 RYAPVCKPL A - -A - - - -C2173 **SMLTDPSHI** C2422 SYTWTGALI - - - - - -C2843 ILMTHFFSI . . . . . . . . . . . . . . . . . .

 Table 2 Amino acid variations in the epitope sequences of the HVC

in RzM6-Tx, whereas liver tissues from patients suffering from chronic hepatitis had lower and variable levels of expression. Therefore, the lytic activity of the peptidespecific CTL lines against RzM6-Tx may not necessarily guarantee the lysis of infected hepatocytes. It is not known

what proportion of hepatocytes in the patients actually express core antigen. In a report of immunostaining of the liver tissues from chronic hepatitis patients, 1-5% of the hepatocytes express core antigen to a level detectable by specific antibodies (Nouri-Aria et al. 1995). Agnello et al. reported that usually 50% or more, but not all, of the hepatocytes are infected by HCV in chronically infected patients (Agnello et al. 1998). If 50% of the hepatocytes expressed core protein, then the results shown in Fig. 5 would indicate that the expression level of the antigen in individual hepatocytes could be comparable to or within a few fold differences from RzM6-Tx in the patients. If so, there is a good chance that infected hepatocytes are lysed by the peptide-induced CTLs. Therefore, the therapeutic potential of peptide-based vaccine seems to be realistic provided that T cells are activated under optimal conditions with high-density peptides and in an immunostimulatory environment including activated APCs and proper helper activities.

# Variability of HCV 1b genomes among Japanese isolates

As HCV is an RNA virus that has a high mutation rate, the sequence variation among HCV isolates is extensive. The



Fig. 4 Lysis of HCV 1b-transfected RzM6 cells by peptide-specific cell lines. The expression of the HCV genome in RzM6 had been induced by tamoxifen-induced Cre expression as described in the Materials and methods section. Peptide-specific cell lines from HCV

1b-infected patients and healthy individuals lysed RzM6 more than the parental HepG2 cells. Individual bars represent the lytic activities of the independent cell lines established from different individuals



Fig. 5 Core protein expression in a HCV-transfected HepG2 cell line and the liver tissue specimens from HCV-infected patients. The amount of core protein was measured by ELISA in reference to the recombinant core protein. The average value for HCV-negative liver tissues was subtracted as background. The expression of HCV 1b in RzM6 had been induced by tamoxifen (RzM6-Tx). Negative control tissue specimens were from a patient with alcoholic liver cirrhosis; LC (Al) and two HBV-infected patients (WH, HS). Biopsy specimens from three HCV 1b RNA-positive patients (IY, NS, HS) with chronic hepatitis were examined. All samples were measured at a total protein concentration of 0.2 mg/ml

emergence of escape variants during the course of HCV infection has been a subject of serious concern in vaccine development (Chang et al. 1997; Cox et al. 2005; Kaneko et al. 1997; Timm et al. 2004). HLA-A\*2402 has a gene frequency of around 30% in the Japanese population (Tanaka et al. 1996; Tokunaga et al. 1997). One-half of all Japanese most likely are at least heterozygous for this allele. Therefore, an HCV virus may thus encounter an HLA-A\*2402-bearing host on every other transmissions on average. Such evolutionary pressure over thousands of years may thus have left some footprints on the viral genome. We therefore examined all Japanese isolates of HCV 1b identified to date for whether there is any evidence of amino acid changes in the HLA-A\*2402-binding peptides that may have helped the virus escape from immune attack. We analyzed the non-synonymous/synonymous (NS/S) substitution rates (Nei and Gojobori 1986) between HLA-A\*2402-binding peptides and the rest of the genome. Hyper-variable regions (HVR) 1 and 2 were analyzed separately because those regions may have been under different evolutionary pressure such as an evasion from antibody responses. The full genome sequence was available for 70 isolates (supplementary material 1).

We first generated an evolutionary tree based on the rapidly changing sequences in HVR 1 and 2 regions (Fig. 6). According to the tree, we calculated and compared the ratio of NS/S substitutions between the HLA-A\*2402binding peptides and then the rest of the genome excluding HVR1 and 2 regions (Table 3). Amino acid substitutions were frequently observed throughout the HCV 1b genome.



**Fig. 6** Phylogenetic tree of the HCV 1b viruses isolated in Japan. An evolutionary relationship among the Japanese isolates of HCV 1b was examined by comparing the sequences in hyper variable regions (HVRs) 1 and 2. The numbers indicate the substitutions per site. The genbank accession numbers of the sequence data are shown

Region <sup>a</sup>	Nonsynonymous substitutions <sup>b</sup>			Synonymous substitutions <sup>c</sup>			Ratio
	K <sub>N</sub>	$L_{\rm N}$	d <sub>N</sub>	Ks	$L_{\rm S}$	$d_{\rm S}$	$d_{ m N}/d_{ m S}$
Whole	201	6543	0.0307	556	2,418	0.230	0.13
HVR	20.1	58.7	0.342	8.0	22.3	0.359	0.95
$\Delta$ HVR	181	6,485	0.0279	548	2,396	0.229	0.12
BS (A24)	12.9	518.4	0.0249	42.2	177.6	0.238	0.11
Non-BS	188	6,025	0.0312	514	2,240	0.229	0.14

 Table 3 Ratio of nonsynonymous versus synonymous substitutions

<sup>a</sup> *HVR* Hyper Variable Region,  $\Delta$ *HVR* sequences excluding HVR, *BS(A24)* HLA-A\*2402-binding sites, *Non-BS* sequences excluding BS & HVR <sup>b</sup> K<sub>N</sub>, L<sub>N</sub>, and d<sub>N</sub> stand for the mean number of nonsynonymous substitutions, the mean number of nonsynonymous sites, and the mean number of nonsynonymous substitutions per site, respectively.

 $^{c}K_{S}$ ,  $L_{S}$ , and  $d_{S}$  stand for the mean number of synonymous substitutions, the mean number of synonymous sites, and the mean number of synonymous substitutions per site, respectively.

HVRs exhibited a marked increase in the NS/S substitution rate, thus, indicating a strong selective pressure for amino acid changes in those regions. In contrast, the NS/S ratio in the HLA-A\*2402-binding peptides was not significantly higher than that in other regions of the genome. In several longitudinal studies that followed the changes in individual patients, frequent amino acid substitutions were accumulated in the epitopes for CTLs and antibodies (Cox et al. 2005; Tester et al. 2005; Timm et al. 2004). An elevated NS substitution rate in the HLA class I-binding epitopes (Cox et al. 2005), therefore, suggests that substitutions positively contributed to the survival of the virus. In contrast to longitudinal studies in individual patients, this analysis indicates that such a bias is not obvious in a viral population that has been circulating among a host population which is heavily biased for HLA-A\*2402. Ray et al. reported that in the absence of selection, amino acid variation tends to converge toward the consensus sequence, the structure of which is likely better adapted to the function of viral proteins. Such convergence may have rapidly occurred when the virus infected the non-HLA-A\*2402-bearing hosts. Judging from the present analysis, it would thus be possible to develop a peptide vaccine for the public where the virus is under continuous selection by one-half of the host population. This is an encouraging result, provided that a strategy is developed to induce immune responses quickly and thoroughly before escape variants emerge.

# Discussion

One of the most troublesome features of HCV for vaccine development is its genetic instability. Patients carry quasispecies of variant viruses which can sometimes be antagonistic to each other for virus-specific T cells. One of the authors and her colleagues have previously shown that the antagonistic inhibition of CTL-mediated cytolysis requires the expression of two closely related peptides on the same antigen-presenting cell that should occur only under limited conditions in a natural infection. More problematically, however, an antagonistic peptide expressed alone on a single antigen-presenting cell can effectively impair antigen-induced proliferation (Kaneko et al. 1997) which, therefore, would normally require a stronger engagement of TCRs than cytolysis (Valitutti et al. 1996). This selectively and effectively limits the expansion of agonist-specific CTLs (Kaneko et al. 1997). When designing a peptide vaccine against a variable target like HCV, it has to be kept in mind that the use of a particular peptide always bears a risk of inducing antagonistic responses in the responding T cell repertoire.

In this study, we observed that the peptide induced CTL lines, if properly stimulated, were cytolytic against the HepG2 transfectant that naturally presents HCV peptides. The transfectant carried amino acid substitutions in some of the epitopes, but it was thereafter effectively killed by CTL lines which were stimulated by wild type peptides. It is not known, however, whether this observation is relevant to natural infections because the transfectant may express more HCV antigens than the liver tissue specimens obtained from chronic hepatitis patients. Under suboptimal conditions, variant peptides may act as antagonists, or they could show a poor presentation in the HLA class I molecules. To overcome these problems, T cells have to be stimulated fully, at least, at the local site of immunization. To provide a stimulatory environment for T cells, the following points should be considered: (1) Epitope peptides should be provided at high density to overcome poor responses of the T cell repertoire that may not have an optimal affinity. An exogenously added synthetic peptide has an advantage over a DNA vaccine in this regard. If a DNA vaccine is used, then, it has to be expressed at a high level and compete for MHC class I presentation with cellular

proteins. For a peptide vaccine, choosing peptides with an optimal affinity for HLA class I molecules is a critical step for achieving a high epitope density. The epitope search program presented herein is thus considered to be a powerful tool to help design peptides. Alternatively, in the future, it may be possible to intentionally avoid a highaffinity dominant epitope against which the T cell repertoire may be anergized. In such a case, it may be better to target subdominant epitopes of intermediate affinity. (2) Antigenpresenting cells at a local site of immunization should be activated fully to induce offensive responses against infected hepatocytes, which may not be an optimal antigen-presenting cell. (3) Helper T cells also need to be activated. In patients chronically exposed to HCV antigens, the immune system may have fallen into a state similar to self-tolerance against HCV (Grakoui et al. 2003; Semmo et al. 2005), or it could simply be exhausted due to the high load of viral replicates circulating in a body, i.e., approximately  $10^{10}$  to  $10^{12}$  newly synthesized virions per day (Grakoui et al. 2003). Helper T cell recruitment and their optimal activation thus remains an important issue to be resolved. Whether or not the helper epitopes need to be of HCV origin or they can be from different antigen sources also remains an important question.

Interestingly, viral changes have been reported to be extensive in the acute phase of an infection, but they appear to subside when the disease enters a chronic phase. In addition, the degree of the immune response against the infected hepatocytes seems to have some correlation with the mutation rate (Chang et al. 1997). In the chronic phase, it takes several months before escaped variants form a visible fraction in patients (Chang et al. 1997; Cox et al. 2005; Erickson et al. 2001; Kaneko et al. 1997). Evolutionary studies on viral changes in HCV also suggest that it takes several months before an amino acid change becomes stabilized in the viral population (Cox et al. 2005; Timm et al. 2004). Considering this time scale, we propose a cocktail vaccine containing a few epitope peptides. One amino acid change in an epitope may occur sooner or later, but the incidence would be much lower for a variant virus which happens to have amino acid changes in two epitopes simultaneously. If three peptides could be used for immunization, then the possibility of a variant virus carrying mutations in all three epitopes of developing would be negligible. If the effector CTLs could be recruited in a timely manner, then a chance to eradicate viruses would thus be obtained. The HLA-A\*2402 prediction program is a powerful tool for designing multiple peptide vaccines. A more efficient viral eradication by a broadranging T cell repertoire, in contrast to the vulnerability of the T cell repertoire directed against a single major epitope to viral escape, has previously been demonstrated in infected patients (Tester et al. 2005).

In this study, we focused on the development of a HLA class I-binding peptide vaccine. However, such peptides alone would usually not be sufficient to induce CTLs. HCV-specific CD8 T cells can be abundantly found in patients especially at the early stage of an infection. Those CTLs become less cytotoxic along the course of chronic transition. A loss of helper T cell activities has been cited by several groups as a cause for the annihilation of the CTL activities (Day et al. 2003; Grakoui et al. 2003; Semmo et al. 2005; Tester et al. 2005; Thimme et al. 2002). In addition, the suppression of the CTL activities by a population of the CD4<sup>+</sup>CD25<sup>+</sup> regulatory T cells has also been demonstrated (Boettler et al. 2005; Rushbrook et al. 2005; Sugimoto et al. 2003). A strategy is thus needed to overcome these problems. Helper peptides can be processed by APCs from exogenously added proteins. Therefore, the activation of APCs and helper T cells along with CD8 T cells in a more stimulatory environment than the chronically infected liver is thus considered to be a crucial point for developing a curative vaccine for HCV.

Acknowledgements This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas (C) Genome Science from the Ministry of Education, Culture, Sports, Science and Technology of Japan and a grant from Formation Program of Research Centers for emerging & reemerging infectious diseases -research & development of life science fields responding to the needs of society. We also would like to express our gratitude to the superb technical assistance of Ms. Satomi Yamanaka. We thank Dr. Yoshiharu Matsuura for providing the HCV 1b clone. We also thank Dr. Masafumi Takiguchi and Dr. Peter Cresswell for providing the cell lines. In addition, the recombinant human IL-2 was a kind gift from Dr. Shinsuke Taki.

## References

- Agnello V, Abel G, Knight G, Muchmore E (1998) Detection of widespread hepatocyte infection in chronic hepatitis C. Hepatology 28:573–584
- Appay V, Nixon D, Donahoe S, Gillespie G, Dong T, King A, Ogg G, Spiegel H, Conlon C, Spina C, Havlir D, Richman D, Waters A, Easterbrook P, McMichael A, Rowland-Jones S (2000) HIVspecific CD8(+) T cells produce antiviral cytokines but are impaired in cytolytic function. J Exp Med 192:63–75
- Battergay M, Fikes J, Di-Bisceglie A, Wentworth P, Sette A, Celis E, Ching W, Grakoui A, Rice C, Kurokohchi K, Berzofsky J, Hoofnagle J, Feinstone S, Akatsuka T (1995) Patients with chronic hapatitis C have circulating cytotoxic T cells which recognize hepatitis C virus-encoded peptides binding to HLA-A2.1 molecules. J Virol 69:2462–2470
- Boettler T, Spangenberg H, Neumann-Haefelin C, Panther E, Urbani S, Ferrari C, Blum H, von Weizsaecker F, Thimme R (2005) T cells with a CD4+CD25+ regulatory phenotype suppress in vitro proliferation of virus-specific CD8+ T cells during chronic hepatitis C virus infection. J Virol 79:7860–7867
- Cerny A, McHutchinson J, Pasquinelli C, Brown M, Brothers M, Grabsheid B, Fowler P, Houghton M, Chisari F (1995) Cytotoxic T lymphocyte response to hepatitis C virus-derived peptides containing the HLA A2.1 binding motif. J Clin Invest 95:521– 530

- Chang K, Rehermann B, McHutchison J, Pasquinelli C, Southwood S, Sette A, Chisari F (1997) Immunological significance of cytotoxic T lymphocyte epitope variants in patients chronically infected by the hepatitis C virus. J Clin Invest 100:2376–2385
- Cox A, Mosbruger T, Mao Q, Liu Z, Wang X, Yang H, Sidney J, Sette A, Pardoll D, Thomas D, Ray S (2005) Cellular immune selection with hepatitis C virus persistence in humans. J Exp Med 201:1741–1752
- Day C, Seth N, Lucas M, Appel H, Gauthier L, Lauer G, Robbins G, Szczepiorkowski Z, Casson D, Chung R, Bell S, Harcourt G, Walker B, Klenerman P, Wucherpfennig K (2003) Ex vivo analysis of human memory CD4 T cells specific for hepatitis C virus using MHC class II tetramers. J Clin Invest 112:831–842
- Edwards P, Smith C, Neville A, O'Hare M (1982) A humanhuman hybridoma system based on a fast-growing mutant of the ARH-77 plasma cell leukemia-derived line. Eur J Immunol 12:641–648
- Erickson A, Kimura Y, Igarashi S, Eichelberger J, Houghton M, Sidney J, McKinney D, Sette A, Hughes, A Walker C (2001) The outcome of hepatitis C virus infection is predicted by escape mutations in epitopes targeted by cytotoxic T lymphocytes. Immunity 15:883–895
- Feld J, Hoofnagle J (2005) Mechanism of action of interferon and ribavirin in treatment of hepatitis C. Nature 436:967–972
- Grakoui A, Shoukry N, Woolland D, Han J, Hanson H, Ghrayeb J, Murthy K, Rice C, Walker C (2003) HCV persistence and immune evasion in the absence of memory T cell help. Science 302:659–662
- Huang Y, BObholzer N, Fayad R, Qiao L (2005) Turning on/off tumor-specific CTL response during progressive tumor growth. J Immunol 175:3110–3116
- Imanishi T, Akaza T, Kimura A, Tokunaga K Gojobori T (1992) Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: Tsuji K, Aizawa M, Sasazuki T (eds) The Eleventh International Histocompatibility Workshop and Conference Yokohama Japan. Oxford University Press
- Kaneko T, Moriyama T, Udaka K, Hiroishi K, Kita H, Okamoto H, Yagita H, Okumura K, Imawari M (1997) Impaired induction of cytotoxic T lymphocytes by antagonism of a weak agonist borne by a variant hepatitis C virus epitope. Eur J Immunol 27:1782–1787
- Karaki S, Kariyone A, Kato N, Kano K, Iwakura Y, Takiguchi M (1993) HLA-B51 transgenic mice as recipients for production of polymorphic HLA-A B-specific antibodies. Immunogenetics 37:139–142
- Kimura Y, Gushima T, Rawale S, Kaumaya P, Walker C (2005) Escape mutations alter proteasome processing of major histocompatibility complex class I-restricted epitopes in persistent hepatitis C virus infection. J Virol 79:4870–4876
- Kurokohchi K, Arima K, Nishioka M (2001) A novel cytotoxic T-cell epitope presented by HLA-A24 molecule in hepatitis C virus infection. J Hepatol 34:930–935
- Lie W, Myers N, Gorka J, Rubocki R, Connolly J, Hansen T (1990) Peptide ligand-induced conformation and surface expression of the Ld class I MHC molecule. Nature 344:439–441
- Nakamoto Y, Kaneko S, Takizawa H, Kikumoto Y, Takano M, Himeda Y (2003) Analysis of the CD8-positive T cell response in Japanese patients with chronic hepatitis C using HLA-A\*2402 peptide tetramers. J Med Virol 70:51–61
- Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous substitutions. Mol Biol Evol 3:418–426
- Nielsen M, Lundegaard C, Lund O, Kesmir C (2005) The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. Immunogenetics 57:33–41

- Nouri-Aria K, ASallie R, Mizokami M, Portmann B, Williams R (1995) Intrahepatic expression of hepatitis C virus antigens in chronic liver disease. J Pathol 175:77–83
- Nussbaum A, Kuttler C, Hadeler K, Rammensee H, Schild H (2001) PAProC: a prediction algorithm for proteasomal cleavages available on the WWW. Immunogenetics 53:87–94
- Ray S, Fanning L, Wang X, Netski D, Kenny-Walsh E, Thomas D (2005) Divergent and convergent evolution after a commonsource outbreak of hepatitis C virus. J Exp Med 201:1753–1759
- Rushbrook S, Ward S, Unitt E, Vowler S, Lucas M, Klenerman P, Alexander G (2005) Regulatory T cells suppress in vitro proliferation of virus-specific CD8+ T cells during persistent hepatitis C virus infection. J Virol 79:7852–7859
- Seifert U, Liermann H, Racanelli V, Halenius A, Wiese M, Wedemeyer H, Ruppert T, Rispeter K, Henklein P, Sijts A, Hengel H, Kloetzel P, Rehermann B (2004) Hepatitis C virus mutation affects proteasomal epitope processing. J Clin Invest 114:250–259
- Semmo N, Day C, Ward S, Lucas M, Harcourt G, Loughry A, Klenerman P (2005) Preferential loss of IL-2-secreting CD4+ T helper cells in chronic HCV infection. Hepatology 41:1019–1028
- Sugimoto K, Ikeda F, Standanlick J, Nunes F, Alter H, Chang K (2003) Suppression of HCV-specific T cells without differential hierarchy demonstrated ex vivo in persistent HCV infection. Hepatology 38:1437–1448
- Tahara T, Yang S, Khan R, Abish S, Haemmerling G, Haemmerling U (1990) HLA antibody responses in HLA class I transgenic mice. Immunogenetics 32:351–360
- Tanaka H, Akaza T, Juji T (1996) Report of the Japanese central bone marrow data center. Clin Transpl 139–144
- Tester I, Smyk-Pearson S, Wang P, Wertheimer A, Yao E, Lewinsohn D, Tavis J, Rosen H (2005) Immune evasion versus recovery after acute hepatitis C virus infection from a shared source. J Exp Med 201:1725–1731
- Thimme R, Bukh J, Spangenberg H, Wieland S, Pemberton J, Steiger C, Govindarajan S, Purcell R, Chisari F (2002) Viral and immunological determinants of hepatitis C virus clearance persistence and disease. Proc Natl Acad Sci USA 99:15661– 15668
- Timm J, Lauer G, Kavanagh D, Sheridan I, Kim A, Lucas M, Pillay T, Ouchi K, Reyor L, Schulze zur Wiesch J, Gandhi R, Chung R, Bhardwaj N, Klenerman P, Walker B, Allen T (2004) CD8 epitope escape and reversion in acute HCV infection. J Exp Med 200:1593–1604
- Tokunaga K, Ishikawa Y, Ogawa A, Wang H, Mitsunaga S, Moriyama S, Lin L, Bannnai M, Watanabe Y, Kashiwase K, Tankaa H, Akaza T, Tadokoro K, Juji T (1997) Sequence-based association analysis of hLA class I and II alleles in Japanese supports conservation of common haplotypes. Immunogenetics 46:199–205
- Tsubota A, Arase Y, Suzuki F, Suzuki Y, Akuta N, Hosaka T, Someya T, Kobayashi M, Saitoh S, Ikeda K, Kumada H (2004) High-dose interferon alpha-2b induction therapy in combination with ribavirin for Japanese patients infected with hepatitis C virus genotype 1b with a high baseline viral load. J Gastroenterol 39:155–161
- Tsukiyama-Kohara K, Tone S, Maruyama I, Inoue K, Katsume A, Nuriya H, Ohmori H, Ohkawa J, Taira K, Hoshikawa Y, Shibasaki F, Reth M, Minatogawa Y, Kohara M (2004) Activation of the CK1-CDK-Rb-E2F pathway in full genome hepatitis C virus-expressing cells. J Biol Chem 15:14531–14541
- Udaka K, Wiesmueller K.-H, Kienle S, Jung G, Tamamura H, Yamagishi H, Okumura K, Walden P, Suto T, Kawasaki T (2000) An automated prediction of MHC class I-binding peptides based on positional scanning with peptide libraries. Immunogenetics 51:816–828

- Udaka K, Mamitsuka H, Nakaseko Y, Abe N (2002) Empirical evaluation of a dynamic experiment design method for prediction of MHC class I-binding peptides. J Immunol 169: 5744–5753
- Valitutti S, Mueller S, Dessing M, Lanzavecchia A (1996) Different responses are elicited in cytotoxic T lymphocytes by different levels of T cell receptor occupancy. J Exp Med 183:1917–1921
- Weiner A, Erickson A, Kansopon J, Crawford K, Muchmore E, Hughes A, Houghton M, Walker C (1995) Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a

cytotoxic T lymphocyte escape variant. Proc Natl Acad Sci USA 92:2755–2759

- Wherry E, Blattman J, Murali-Krishna K, vander Most R, Ahmed R (2003) Viral persistance alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. J Virol 77:4911–4927
- Zeh H III, Leder G, Lotze M, Salter R, Tector M, Stuber G, Modrow S, Storkus W (1994) Flow-cytometric determination of peptideclass I complex formation. Identification of p53 peptides that bind to HLA-A2. Hum Immunol 39:79–86