Prediction of T-cell epitopes of hepatitis C virus genotype 5a

Maemu P Gededzha, M Jeffrey Mphahlele and Selokela G Selabe

Abstract

Background: Hepatitis C virus (HCV) is a public health problem with almost 185 million people estimated to be infected worldwide and is one of the leading causes of hepatocellular carcinoma. Currently, there is no vaccine for HCV infection and the current treatment does not clear the infection in all patients. Because of the high diversity of HCV, protective vaccines will have to overcome significant viral antigenic diversities. The objective of this study was to predict T-cell epitopes from HCV genotype 5a sequences.

Methods: HCV near full-length protein sequences were analyzed to predict T-cell epitopes that bind human leukocyte antigen (HLA) class I and HLA class II in HCV genotype 5a using Propred I and Propred, respectively. The Antigenicity score of all the predicted epitopes were analysed using VaxiJen v2.0. All antigenic predicted epitopes were analysed for conservation using the IEDB database in comparison with 406, 221, 98, 33, 45, 45 randomly selected sequences from each of the HCV genotypes 1a, 1b, 2, 3, 4 and 6 respectively, downloaded from the GenBank. For epitope prediction binding to common HLA alleles found in South Africa, the IEDB epitope analysis tool was used.

Results: A total of 24 and 77 antigenic epitopes that bind HLA class I and HLA class II respectively were predicted. The highest number of HLA class I binding epitopes were predicted within the NS3 (63%), followed by NS5B (21%). For the HLA class II, the highest number of epitopes were predicted in the NS3 (30%) followed by the NS4B (23%) proteins. For conservation analysis, 8 and 31 predicted epitopes were conserved in different genotypes for HLA class I and HLA class II alleles respectively. Several epitopes bind with high affinity for both HLA class I alleles and HLA class II common in South Africa.

Conclusion: The predicted conserved T-cell epitopes analysed in this study will contribute towards the future design of HCV vaccine candidates which will avoid variation in genotypes, which in turn will be capable of inducing broad HCV specific immune responses.

Keywords: Hepatitis C virus, Genotype 5a, In-silico, T-cell epitopes, Vaccines

Introduction

Hepatitis C virus (HCV) is estimated to infect 185 million people worldwide [1]. Chronic HCV infection leads to progressive liver disease, being one of the major causes of hepatocellular carcinoma and one of the most common indications for liver transplantation [2]. The World Health Organization (WHO) strongly recommends combination therapy with pegylated interferon and ribavirin for chronically infected patients who qualify for treatment [1]. Recently, two NS3 protease inhibitors (boceprevir and telaprevir) have been approved by the US Food and Drug Administration, and the WHO conditionally (until more evidence has accumulated) recommends that these drugs should be given in combination with pegylated interferon and ribavirin for the treatment of chronic HCV genotype 1 infections. Also, the WHO strongly recommends that sofosbuvir be given in combination with ribavirin alone in patients who cannot tolerate interferon and are chronically infected with genotypes 1, 2, 3 and 4 [1]. However, these therapies are still not affordable in most developing countries.
As a result, the development of an effective HCV vaccine is undoubtedly the best solution for the ultimate control of HCV infections, and is a public health priority.

Prophylactic vaccines against viral infections are generally aimed at inducing a humoral (B-cell) immune response, while therapeutic vaccines preferably activate both humoral and cellular (T-cell) immune responses [3]. A successful HCV vaccine will need to stimulate both arms of the adaptive immune response, since while both cellular and humoral immune responses occur in a naturally infected host, the current consensus is that a strong cellular response is vital for viral clearance and protection [4].

The development of an effective HCV vaccine requires an understanding of the host's adaptive immune response to natural infection. As with other viral infections, viral antigens are presented to CD4+ and CD8+ T-cells via human leukocyte antigen (HLA) class II and class I molecules, respectively [5]. Different HLA class alleles have been found to be associated with HCV infection. For example, HLA-A*11, HLA-Cw*04 and HLA-B*53 have been associated with HCV persistence [6,7], while HLA-B*27, HLA-A*11:01, HLA-B*57, HLA-Cw*01:02 and HLA-A*03 have been associated with spontaneous HCV clearance [8,9]. Also, HLA-DRB1*11 and HLA-DQB1*03:01 have been associated with decreased disease severity of HCV infection globally, suggesting they may present HCV-derived epitopes more efficiently to CD4+ T-cells than others and thus capable of viral clearance [10]. During acute HCV infection, development and maintenance of strong specific responses by CD8+ and CD4+ T-cells [11,12] and neutralizing antibodies [13] are associated with viral clearance, with HCV-specific CD8+ and CD4+ T-cells usually being transient or absent in patients who develop persistent infections. The rate of chronic liver disease progression has been shown to be determined by the magnitude of HCV-specific CD4+ T-cell responses, since these cells are essential for both the cellular and humoral responses [14]. CD8+ T-cells are essential for long-term protection against chronic HCV [15], while CD4+ T-cells play a role in viral clearance [16].

HCV infection evades the host's immune system by generating immune escape variants through alteration of the virus HLA-restricted epitopes to avoid being recognized by T-cells and neutralizing antibodies [14]. Thus effective HCV vaccines will need to target protective epitopes that display minimal cross-genotype amino acid variability as they will provide broad potency [17]. Peptides corresponding to protective epitopes are desirable vaccine candidates because they are easy to construct and produce, and they do not contain infectious materials [18]. The first step in the process of epitope-based vaccine design and development is the in-silico prediction of peptide binding affinities to HLA proteins [19].

Genotype 5 accounts for over 50% of HCV infections in South Africa [20], and is becoming more prevalent in Europe and North America [21]. It is the most conserved of HCV genotypes, being classified into only one subtype (5a) [22-24]. The growing prevalence of HCV genotype 5a in different parts of the world necessitates its molecular characterization in order to improve the formulations of vaccine candidates that are in development. Thus the aim of this study was to assess immunological determinants by predicting conserved epitopes in near-full length HCV genotype 5a sequences using a suite of online programmes to help in the designing of new vaccine candidates.

Results
Prediction of T-cell epitopes
For HLA class I, a total of 24 antigenic epitopes were predicted in the consensus near full-length of genotype 5a (Table 1). Epitope NS31325–1333 covered 30 of the 47 HLA class I alleles that were analysed assuring high binding affinity to different alleles. For conservation analysis with other genotypes, 8 of 24 epitopes were 100% conserved for specific genotypes. Epitopes NS31332–1340 and NS5B2557–2565 were highly conserved in all genotypes analysed, and epitope E2684–692 was conserved in all genotypes except for genotype 2, while 5 other epitopes were conserved in either 2 or 3 genotypes analysed. In addition epitope NS4B1832–1840 was conserved in genotype 1a, 1b, 2 and 4 epitope variants at anchor residues position 2 and 9 while E2677–685, NS31325–1333 and NS31357–1365 were conserved in at least 3 genotypes epitope variants each (Table 1). For HLA class II, 77 epitopes were predicted (Table 2). Epitope NS4B1879–1887 and NS4B1880–1888 covered 51 of 51 HLA class II alleles analysed. For conservation analysis with other genotypes, 31 of 71 epitopes were conserved for specific genotypes. Some epitopes were highly conserved in all genotypes (E2507–515, E2509–517, NS31253–1261, NS31254–1262, NS31327–1335, NS31392–1400, NS4B1916–1924 and NS4B1919–1927), while other epitopes were conserved in at least 1 to 4 of the genotypes analysed. Epitope NS4B1886–1894 was conserved at anchor residues position 1, 4, 6 and 9 in all 6 genotypes epitope variants while E2692–700, NS2254–972, NS31418–1426 and NS41591–1569 were conserved in at least 5 genotypes epitope variants each (Table 2). Epitopes NS31385–1593, NS5A12265–2293 and NS5B1889–2897 were predicted to cover both HLA class I and HLA class II alleles. The highest number of HLA class I binding epitopes were predicted within the NS3 (63%), followed by NS5B (21%), and for the HLA class II, the highest number of
| Position | Epitope sequence | No of allele | Antigenecity score | Genotype 1a | Genotype 1b | Genotype 2 | Genotype 3 | Genotype 4 | Genotype 6 |
|----------|-----------------|-------------|--------------------|------------|------------|------------|------------|------------|------------|
| E2       | RCDLEDRDR       | 9           | 2.7428             |            |            |            |            |            |            |
|          | CSFTPTPAL       | 27          | 0.8772             | CSFTLPAL (94) | CSFTLPAL (92) | CSFTPMPAL (73) |            |            |            |
|          | ALSTGLIH        | 29          | 0.9005             | ALSTGLIH (70) | ALSTGLIH (99) | ALSTGLIH (97) | ALSTGLIH (97) | ALSTGLIH (84) | ALSTGLIH (91) |
| NS2      | VPPLQVRGG       | 6           | 1.4312             | VPPLVRGG (85) |            |            |            |            |            |
| NS3      | YAQQTRGGL       | 26          | 0.7638             | YAQQTRGGL (99) |            |            |            |            |            |
|          | AQQTTRGLG       | 5           | 0.7751             | AQQTTRGLG (99) |            |            |            |            |            |
|          | ADLYLVRTH       | 5           | 1.0277             | ADLYLVRTH (99) | ADLYLVRTH (99) | ADLYLVRTH (99) | ADLYLVRTH (99) | ADLYLVRTH (99) | ADLYLVRTH (99) |
|          | TILGTVL         | 30          | 0.7555             | TILGTVL (88) | SILGTVL (94) | TILGTVL (82) |            |            |            |
|          | VLDQETAG        | 5           | 0.5004             | VLDQETAG (96) | VLDQETAG (98) | VLDQETAG (96) | VLDQETAG (100) | VLDQETAG (93) | VLDQETAG (71) |
|          | TPHNIEEV        | 24          | 0.8423             | VPHNIEEV (89) | VPHNIEEV (80) | VPHNIEEV (94) |            |            |            |
|          | HPNIEEV         | 24          | 1.1654             | HPNIEEV (88) | HPNIEEV (76) | HSNIEEV (94) |            |            |            |
|          | EGEIPFYGR       | 9           | 1.0560             | EGEIPFYGR (97) |            |            |            |            |            |
|          | PFYGRAPL        | 7           | 0.9479             | PFYGRAPL (99) | PFYGRAPL (85) | PFYGRAPL (82) |            |            | PFYGRAPL (73) |
|          | TKYIMACMS       | 6           | 0.7008             | TKYIMACMS (99) | TKYIMACMS (70) | TKYIMACMS (70) | TKYIMACMS (93) | TKYIMACMS (93) | TKYIMACMS (87) |
| NS4B     | SASLPMDE        | 5           | 0.5488             | SASLPMDE (99) |            |            |            |            |            |
|          | QIAFTPAAAT      | 5           | 0.5293             | QIAFTPAAAT (99) | QIAFTPAAAT (99) | QIAFTPAAAT (99) | QIAFTPAAAT (99) | QIAFTPAAAT (99) | QIAFTPAAAT (89) |
|          | SGMAGAVG        | 5           | 0.8750             | SGMAGAVG (81) | SGMAGAVG (93) | SGMAGAVG (98) | SGMAGAVG (98) | SGMAGAVG (76) |            |
|          | LIDLAGYG        | 5           | 0.8431             | LIDLAGYG (94) | LIDLAGYG (95) | LIDLAGYG (79) | LIDLAGYG (84) |            |            |
|          | GYGAGVAGA       | 9           | 0.5562             | GYGAGVAGA (99) | GYGAGVAGA (99) | GYGAGVAGA (94) | GYGAGVAGA (94) | GYGAGVAGA (91) | GYGAGVAGA (84) |
| NS5B     | GYGAKEVRS       | 7           | 1.3540             | GYGAKEVRS (95) | GYGAKEVRS (94) | GYGAKEVRS (93) | GYSAKEVRS (79) |            |            |
|          | TIMAKNEVF       | 12          | 0.6303             | TIMAKNEVF (99) | TIMAKNEVF (80) | TIMAKNEVF (83) | TIMAKNEVF (100) | TIMAKNEVF (84) | TIMAKNEVF (93) |
|          | EPSKGKKKP       | 6           | 1.4218             | EPSKGKKKP (93) |            |            |            |            |            |
|          | LASCREAKL       | 28          | 0.5264             | LASCREAKL (99) |            |            |            |            |            |
|          | LHGLSASFSL      | 11          | 0.6722             | LHGLSASFSL (99) | LHGLSASFSL (98) | LHGLSASFSL (82) | LHGLSASFSL (82) | LHGLSASFSL (98) | LHGLSASFSL (98) |

* indicates that the epitope has been experimentally proven to be a true positive.
Bold: indicates percentage of epitope that is 100% conserved in more than 70% of the sequences analysed in each genotype.
Italic: indicates amino acid(s) variation in epitope in comparison to the predicted epitope.
Table 2 HLA class II predicted epitopes of HCV genotype 5a and their antigenicity prediction score, number of allele and conservation (in percentages) in different genotypes

| Position | Epitope sequence | No of allele | Antigenecity score | Genotype 1a | Genotype 1b | Genotype 2 | Genotype 3 | Genotype 4 | Genotype 6 |
|----------|------------------|--------------|--------------------|-------------|-------------|------------|------------|------------|------------|
| E1       |                  |              |                    |             |             |            |            |            |            |
| 320      | WDMMMNWP         | 10           | 1.3610             | WDMMMNWP (99) | WDMMMNWP (97) | WDMMMNWP (97) | WDMMMNWP (97) | WDMMMNWP (93) | WDMMMNWP (93) |
| E2       |                  |              |                    |             |             |            |            |            |            |
| 507      | YCFTPSPVV        | 6            | 1.2112             | YCFTPSPVV (93) | YCFTPSPVV (98) | YCFTPSPVV (88) | YCFTPSPVV (100) | YCFTPSPVV (91) | YCFTPSPVV (91) |
| 509      | FTPSPVVVG        | 17           | 1.3553             | FTPSPVVVG (94) | FTPSPVVVG (98) | FTPSPVVVG (88) | FTPSPVVVG (100) | FTPSPVVVG (89) | FTPSPVVVG (73) |
| 665      | LLHTTTQWA        | 5            | 0.6616             | LLHTTTQWA (100) | LLHTTTQWA (100) | LLHTTTQWA (100) | LLHTTTQWA (100) | LLHTTTQWA (100) | LLHTTTQWA (73) |
| 666      | LHTTTQWAI        | 25           | 0.5863             | LHTTTQWAI (100) | LHTTTQWAI (100) | LHTTTQWAI (100) | LHTTTQWAI (100) | LHTTTQWAI (100) | LHTTTQWAI (100) |
| 692      | LHQNIVDVTQ       | 6            | 0.8564             | LHQNIVDVTQ (100) | LHQNIVDVTQ (100) | LHQNIVDVTQ (100) | LHQNIVDVTQ (100) | LHQNIVDVTQ (100) | LHQNIVDVTQ (87) |
| 739      | LLVCQAEAA        | 22           | 0.5506             | LLVCQAEAA (100) | LLVCQAEAA (100) | LLVCQAEAA (100) | LLVCQAEAA (100) | LLVCQAEAA (100) | LLVCQAEAA (73) |
| P7       |                  |              |                    |             |             |            |            |            |            |
| 803      | LPHRALALD        | 7            | 0.6966             | LPHRALALD (100) | LPHRALALD (100) | LPHRALALD (100) | LPHRALALD (100) | LPHRALALD (100) | LPHRALALD (86) |
| NS2      |                  |              |                    |             |             |            |            |            |            |
| 860      | WVPPLQVRG        | 7            | 1.1998             | WVPPLQVRG (100) | WVPPLQVRG (100) | WVPPLQVRG (100) | WVPPLQVRG (100) | WVPPLQVRG (100) | WVPPLQVRG (82) |
| 861      | VPPPLQVRGG       | 25           | 1.4312             | VPPPLQVRGG (100) | VPPPLQVRGG (100) | VPPPLQVRGG (100) | VPPPLQVRGG (100) | VPPPLQVRGG (100) | VPPPLQVRGG (82) |
| 864      | LQVRGGRDA        | 20           | 0.6368             | LQVRGGRDA (100) | LQVRGGRDA (100) | LQVRGGRDA (100) | LQVRGGRDA (100) | LQVRGGRDA (100) | LQVRGGRDA (86) |
| 880      | FHPALGFEI        | 19           | 1.4374             | FHPALGFEI (100) | FHPALGFEI (100) | FHPALGFEI (100) | FHPALGFEI (100) | FHPALGFEI (100) | FHPALGFEI (93) |
| 892      | LLGLILPLY        | 14           | 0.5940             | LLGLILPLY (100) | LLGLILPLY (100) | LLGLILPLY (100) | LLGLILPLY (100) | LLGLILPLY (100) | LLGLILPLY (82) |
| 895      | ILGPLYLLOQ       | 5            | 0.8264             | ILGPLYLLOQ (100) | ILGPLYLLOQ (100) | ILGPLYLLOQ (100) | ILGPLYLLOQ (100) | ILGPLYLLOQ (100) | ILGPLYLLOQ (88) |
| 938      | LLHLGLRGTG       | 25           | 0.5554             | LLHLGLRGTG (100) | LLHLGLRGTG (100) | LLHLGLRGTG (100) | LLHLGLRGTG (100) | LLHLGLRGTG (100) | LLHLGLRGTG (82) |
| 964      | LRDLAVATE        | 9            | 1.0497             | LRDLAVATE (100) | LRDLAVATE (100) | LRDLAVATE (100) | LRDLAVATE (100) | LRDLAVATE (100) | LRDLAVATE (100) |
| 997      | LAGLPVSAR        | 11           | 0.9120             | LAGLPVSAR (100) | LAGLPVSAR (100) | LAGLPVSAR (100) | LAGLPVSAR (100) | LAGLPVSAR (100) | LAGLPVSAR (100) |
| 1025     | LLAPITAYA        | 22           | 0.5253             | LLAPITAYA (100) | LLAPITAYA (100) | LLAPITAYA (100) | LLAPITAYA (100) | LLAPITAYA (100) | LLAPITAYA (89) |
| NS3      |                  |              |                    |             |             |            |            |            |            |
| 1047     | LTGRDKNEA        | 22           | 1.5633             | LTGRDKNEA (100) | LTGRDKNEA (100) | LTGRDKNEA (100) | LTGRDKNEA (100) | LTGRDKNEA (100) | LTGRDKNEA (90) |
| 1131     | YLVTREHADV       | 15           | 0.7569             | YLVTREHADV (99) | YLVTREHADV (99) | YLVTREHADV (99) | YLVTREHADV (99) | YLVTREHADV (99) | YLVTREHADV (97) |
| 1152     | LLSPRPSYSY       | 5            | 2.2999             | LLSPRPSYSY (100) | LLSPRPSYSY (100) | LLSPRPSYSY (100) | LLSPRPSYSY (100) | LLSPRPSYSY (100) | LLSPRPSYSY (89) |
| 1153     | LSPRPISYLSY      | 5            | 1.3109             | LSPRPISYLSY (100) | LSPRPISYLSY (100) | LSPRPISYLSY (100) | LSPRPISYLSY (100) | LSPRPISYLSY (100) | LSPRPISYLSY (89) |
| 1253     | VLNPSVAAT        | 14           | 1.2489             | VLNPSVAAT (100) | VLNPSVAAT (100) | VLNPSVAAT (100) | VLNPSVAAT (100) | VLNPSVAAT (100) | VLNPSVAAT (89) |
| 1254     | LNPSVAATL        | 6            | 0.9169             | LNPSVAATL (100) | LNPSVAATL (100) | LNPSVAATL (100) | LNPSVAATL (100) | LNPSVAATL (100) | LNPSVAATL (89) |
| 1258     | VAATLGFGLA       | 11           | 1.2617             | VAATLGFGLA (100) | VAATLGFGLA (100) | VAATLGFGLA (100) | VAATLGFGLA (100) | VAATLGFGLA (100) | VAATLGFGLA (89) |
| 1262     | LGFGAYMSK        | 14           | 0.8784             | LGFGAYMSK (87) | LGFGAYMSK (82) | LGFGAYMSK (82) | LGFGAYMSK (82) | LGFGAYMSK (82) | LGFGAYMSK (82) |
| No. | Epitope | Prediction Score | Allele | Conservation (%) |
|-----|---------|------------------|--------|------------------|
| 1264 | FGAYMSKAY | 0.6561 | FGAYMSKAY (100) | FGAYMSKAY (82) |
| 1327 | LGIGTVLDQ | 0.6967 | LGIGTVLDQ (97) | LGIGTVLDQ (88) |
| 1373 | IFCHSKKKC | 1.5437 | IFCHSKKKC (93) | IFCHSKKKC (95) |
| 1415 | VFYGRAIPLA | 0.9522 | VFYGRAIPLA (100) | VFYGRAIPLA (81) |
| 1482 | VSRSORRGR | 2.0649 | VSRSORRGR (100) | VSRSORRGR (93) |
| 1561 | VFGLTNID | 1.1689 | VFGLTNID (100) | VFGLTNID (96) |
| 1583 | FPYLVAYQA | 0.6584 | FPYLVAYQA (85) | FPYLVAYQA (85) |
| 1585 | YLVAYQATV | 0.5020 | YLVAYQATV (99) | YLVAYQATV (85) |
| 1586 | LVAYQATVC | 0.6052 | LVAYQATVC (99) | LVAYQATVC (85) |
| 1631 | VQNEIVLTH | 1.1979 | VQNEIVLTH (100) | VQNEIVLTH (80) |
| 1641 | ITKYMCM | 0.6964 | ITKYMCM (100) | ITKYMCM (94) |
| 1645 | IMACMSADL | 0.8406 | IMACMSADL (100) | IMACMSADL (82) |
| 1733 | LGLIGTAGQ | 1.2450 | LGLIGTAGQ (100) | LGLIGTAGQ (100) |
| 1765 | MWNFSQGIQ | 1.2300 | MWNFSQGIQ (100) | MWNFSQGIQ (96) |
| 1791 | MSFTAAVTS | 0.7414 | MSFTAAVTS (100) | MSFTAAVTS (92) |
| 1812 | LGGWVAFSL | 0.608 | LGGWVAFSL (100) | LGGWVAFSL (95) |
| 1849 | IDILAGYG | 0.7525 | IDILAGYG (100) | IDILAGYG (97) |
| 1863 | LVAFKIMC | 1.0160 | LVAFKIMC (100) | LVAFKIMC (95) |
| 1879 | LVNLPSIL | 0.7592 | LVNLPSIL (100) | LVNLPSIL (90) |
| 1880 | VNLPSIL | 0.9874 | VNLPSIL (100) | VNLPSIL (92) |
| 1883 | LPSLCPGA | 0.7104 | LPSLCPGA (100) | LPSLCPGA (99) |
| 1886 | ILCPALGV | 0.9365 | ILCPALGV (100) | ILCPALGV (98) |
| 1892 | LVGVICAAV | 0.7097 | LVGVICAAV (100) | LVGVICAAV (98) |
| 1893 | VGVICAV | 0.9530 | VGVICAV (100) | VGVICAV (98) |
| 1896 | VCAALRRL | 0.9887 | VCAALRRL (100) | VCAALRRL (95) |
| 1897 | ICAALRRH | 1.2316 | ICAALRRH (100) | ICAALRRH (97) |
| 1916 | MNRLIASFS | 0.5291 | MNRLIASFS (100) | MNRLIASFS (99) |

Table 2 HLA class II predicted epitopes of HCV genotype 5a and their antigenicity prediction score, number of allele and conservation (in percentages) in different genotypes (Continued)
|    | Epitope    | Score | Conservation | NS5A     | NS5B     |
|----|------------|-------|--------------|----------|----------|
| 1919 | LIAFASRGN  | 2.0235 | LIAFASRGN (100) | LIAFASRGN (99) | LIAFASRGN (98) | LIAFASRGN (100) | LIAFASRGN (93) | LIAFASRGN (96) |
| 1922 | FASRGNHVS  | 1.2938 | FASRGNHVS (100) | FASRGNHVS (98) | FASRGNHVS (97) | FASRGNHVS (94) | FASRGNHVS (82) | FASRGNHVS (93) |
| 1994 | WLQAKLLPQ  | 1.1603 | WLQAKLLPQ (100) | WLQAKLLPQ (99) | WLQAKLLPQ (98) | WLQAKLLPQ (100) | WLQAKLLPQ (93) | WLQAKLLPQ (96) |
| 2101 | YITGVTQDN  | 1.2321 | YITGVTQDN (100) | YITGVTQDN (98) | YITGVTQDN (97) | YITGVTQDN (94) | YITGVTQDN (82) | YITGVTQDN (93) |
| 2105 | VTQDNLKC  | 1.0178 | VTQDNLKC (100) | VTQDNLKC (98) | VTQDNLKC (97) | VTQDNLKC (94) | VTQDNLKC (82) | VTQDNLKC (93) |
| 2237 | MGGNITRVE  | 1.2656 | MGGNITRVE (92) | MGGNITRVE (96) | MGGNITRVE (91) | MGGNITRVE (94) | MGGNITRVE (82) | MGGNITRVE (91) |
| 2285 | LPWVARPGY  | 0.6143 | LPWVARPGY (100) | LPWVARPGY (98) | LPWVARPGY (97) | LPWVARPGY (82) | LPWVARPGY (93) | LPWVARPGY (78) |
| 2342 | MSYSWTGAL  | 0.5749 | MSYSWTGAL (76) | MSYSWTGAL (97) | MSYSWTGAL (85) | MSYSWTGAL (91) | MSYSWTGAL (79) | MSYSWTGAL (84) |
| 2451 | LRHHNLVYS  | 1.2321 | LRHHNLVYS (89) | LRHHNLVYS (86) | LRHHNLVYS (97) | LRHHNLVYS (83) | LRHHNLVYS (85) | LRHHNLVYS (96) |
| 2523 | YGAKEVRS  | 1.1739 | YGAKEVRS (100) | YGAKEVRS (85) | YGAKEVRS (92) | YGAKEVRS (82) | YGAKEVRS (76) | YGAKEVRS (91) |
| 2559 | MAKNEVFC  | 0.8466 | MAKNEVFC (100) | MAKNEVFC (81) | MAKNEVFC (85) | MAKNEVFC (85) | MAKNEVFC (85) | MAKNEVFC (96) |
| 2564 | VFAVEPSKG  | 0.7925 | VFAVEPSKG (100) | VFAVEPSKG (89) | VFAVEPSKG (86) | VFAVEPSKG (85) | VFAVEPSKG (85) | VFAVEPSKG (78) |
| 2637 | FSYDTRCFD  | 1.1508 | FSYDTRCFD (95) | FSYDTRCFD (100) | FSYDTRCFD (93) | FSYDTRCFD (87) | FSYDTRCFD (86) | FSYDTRCFD (85) |
| 2660 | YQSCDLQPE  | 1.3802 | YQSCDLQPE (100) | YQSCDLQPE (94) | YQSCDLQPE (94) | YQSCDLQPE (94) | YQSCDLQPE (94) | YQSCDLQPE (94) |
| 2696 | YRRCRASGV  | 1.2642 | YRRCRASGV (99) | YRRCRASGV (99) | YRRCRASGV (98) | YRRCRASGV (98) | YRRCRASGV (97) | YRRCRASGV (97) |
| 2778 | YDLELTSC  | 1.0923 | YDLELTSC (100) | YDLELTSC (96) | YDLELTSC (98) | YDLELTSC (91) | YDLELTSC (91) | YDLELTSC (96) |
| 2850 | FSVLQSQEQ  | 0.8451 | FSVLQSQEQ (84) | FSVLQSQEQ (84) | FSVLQSQEQ (84) | FSVLQSQEQ (84) | FSVLQSQEQ (84) | FSVLQSQEQ (84) |
| 2871 | VYSVTPLDL  | 2.0113 | VYSVTPLDL (100) | VYSVTPLDL (75) | VYSVTPLDL (91) | VYSVTPLDL (87) | VYSVTPLDL (87) | VYSVTPLDL (87) |
| 2883 | IQRLHGLSA  | 0.4422 | IQRLHGLSA (98) | IQRLHGLSA (72) | IQRLHGLSA (73) | IQRLHGLSA (79) | IQRLHGLSA (75) | IQRLHGLSA (96) |
| 2889 | LSASFSLHSY | 0.6007 | LSASFSLHSY (99) | LSASFSLHSY (97) | LSASFSLHSY (97) | LSASFSLHSY (97) | LSASFSLHSY (97) | LSASFSLHSY (97) |

*indicates that the epitope has been experimentally proven to be a true positive. Bold- indicates percentage of epitope that is 100% conserved in more than 70% of the sequences analysed in each genotype. Italic- indicates amino acid(s) variation in epitope in comparison to the predicted epitope.
epitopes were predicted in the NS3 (30%) followed by the NS4B (23%) proteins (Table 3).

**Epitope binding affinity to common South African HLA class I and HLA class II alleles**

Genotype 5 epitopes and their genotypic variants were analysed for their binding affinity to HLA class I and HLA class II alleles most common in South Africa, where genotype 5a is predominating. For HLA class I, 11 of the most common South African HLA-A alleles (HLA-A*01:01, HLA-A*02:01, HLA-A*30:01 and HLA-A*30:02), HLA-B (HLA-B*07:02, HLA-B*08:01 and HLA-B*3501), and HLA-C (HLA-C*04:01, HLA-C*06:01, HLA-C*07:01 and HLA-C*07:02) were analysed. The limitation of Propred 1 is that it does not cover most of the main HLA class I alleles: HLA-A (HLA-A*01:01, HLA-A*30:01 and HLA-A*30:02), HLA-B (HLA-B*08:01 and HLA-C (HLA-C*04:01, HLA-C*06:01, HLA-C*07:01 and HLA-C*07:02) that are observed in South Africa. As a results the IEDB epitope analysis tool was used to predict epitopes of all the 11 most common HLA-A, HLA-B, HLA-C covering HLA class I alleles found in South Africa with the ANN prediction server. Thirteen antigenic epitopes with high binding affinity score of <50 IC$_{50}$nM were predicted. Most epitopes bind with high affinity to single HLA class I alleles with exception of epitope NS5B2889–2907 LSAFSLHSY that bind with high affinity to 2 HLA-A alleles (HLA-A*01:01 and HLA-A*30:02). Four of the epitopes binding to HLA-B*35:01 followed by HLA-A*02:01 with 3 epitopes binding to it, however, none of the epitopes bind to HLA-C*04:01, HLA-C*06:01 and HLA-C*07:02 alleles. The NS31359–1367 shows a level of promiscuity to HLA-A and HLA-B alleles. NS31359–1367 HPNIEEVAL bind with intermediate affinity to HLA-B*07:02, while its genotype 2, 3 and 4 variant HSNIIEEVAL bind with poor affinity and genotype 6 variant HPNITETAL bind with high affinity. For HLA-B*35:01, the NS31359–1367 HPNIEEVAL and genotype 6 variant HPNITETAL bind with high affinity while genotype 2, 3 and 4 variant HSNIIEEVAL bind with poor affinity. Three (E2560–515, NS41774–1782, NS4B1919–1927, and NS4B1920–1928 were highly conserved in all genotypes (Table 5).

**Validation of epitopes**

Seven of the predicted epitopes were previously confirmed experimentally by other studies as true positives in comparison with the epitopes analysed in the IEDB resource database. Majority of the epitopes predicted in this study have not been previously tested experimentally. The ‘true epitopes’ are highlighted by (') in Tables 1, 2 and 4.

**Discussion**

Several studies that have published HCV epitopes focused mainly on genotype 1 [41,42], but most of these studies do not take into account the diversity in other genotypes that are common in developing countries like most African countries. In the present study, predicted antigenic epitopes of HCV genotype 5a proteins from South Africa were analysed followed by conservation with randomly selected genotypes 1–6 references from GenBank. Several studies have confirmed the importance of using immunoinformatics as good predictors for selecting HLA ligands, T-cell epitopes and immunogenicity [43]. As a result, several immunoinformatics methods have been developed to assist in the identification of HLA binding peptides [44,45].

For this analysis, near full-length sequences covering all HCV proteins with the exclusion of the 3’ end of the NS5B were included to maximize number of epitopes predicted. The use of the whole viral genome for developing epitope vaccines has a potential control over the immune response and eliminating the side effects [43], and it also increases the chance of detecting a virus at any developmental stage [46]. It has been shown that multiple epitopes from different parts of the HCV genome are important to produce a vaccine that can elicit strong humoral immune responses and multiple specific cellular

![Image](http://www.virologyj.com/content/11/1/187)

---

**Table 3 Distribution of genotype 5a HLA class I and II predicted epitopes in each of the HCV gene**

| HLA class | No of predicted epitopes | Distribution of the predicted epitopes in each of the HCV gene<sup>a</sup> |
|-----------|-------------------------|---------------------------------------------------------------------|
|           |                         | C     | E1    | E2    | P7    | NS2   | NS3<sup>b</sup> | NS4A  | NS4B  | NSSA  | NS5B  |
| I         | 24                      | 0 (0) | 0 (0) | 3 (12)| 0 (0) | 1 (4) | 15 (63)| 0 (0) | 0 (0) | 0 (0) | 5 (21)|
| II        | 77                      | 0 (0) | 1 (1) | 6 (8) | 1 (1) | 10 (13)| 23 (30)| 0 (0) | 18 (23) | 5 (7) | 13 (7)|

<sup>a</sup>Number of epitopes and their percentage in brackets.

<sup>b</sup>HCV gene with highest number of binding epitopes (in bold).
Table 4 Binding affinity scores of predicted epitopes and their variants to common HLA I allele types prevalent in South Africa

| Gene | Epitope sequence | Genotype of epitope | HLA-A | HLA-B | HLA-C |
|------|-----------------|---------------------|-------|-------|-------|
|      |                 |                     | HLA-A*01:01 | HLA-A*02:01 | HLA-A*30:01 | HLA-A*30:02 | HLA-B*07:02 | HLA-B*08:01 | HLA-B*35:01 | HLA-C*04:01 | HLA-C*06:02 | HLA-C*07:01 | HLA-C*07:02 |
| E2   | ALSTGLHHL       | 1a, 1b, 3, 4, 5a, 6 | +     | +++   | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | ALSTGLHL        | 2                  | +     | +++   | +     | +     | +     | +     | +     | +     | +     | +     |
| NS2  | LLAPITAYA       | 1a, 3, 4, 5a       | +     | +++   | +     | ++   | +     | +     | +     | +     | +     | +     |
|      | LLAPITAY        | 2                  | +     | +++   | +     | ++   | +     | +     | +     | +     | +     | +     |
| NS3  | YAQQTRGY       | 5a                 | +     | +     | +     | +     | +++   | +     | +     | +     | +     | +     |
|      | YAQQTRG        | 1a, 2              | +     | +     | +     | +     | +++   | +     | +     | +     | +     | +     |
|      | YSAQQTRG       | 1b                 | +     | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | YAQQTRG        | 6                  | +     | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | AAYSAQQG       | 1a, 1b, 4, 5a      | +     | +     | +++   | +     | +     | +     | +     | +     | +     | +     |
|      | AAYSAQQG       | 2                  | +     | +     | +++   | +     | +     | +     | +     | +     | +     | +     |
|      | FGAHYMSKAY     | 5a                 | +     | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | FGAHYMSKAH     | 1a, 1b, 2          | +     | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | HPNIEEVAL*      | 1a, 1b, 5a         | +     | +     | +     | +     | ++    | +     | +     | ++    | +     | +     |
|      | HSNIEEVAL      | 2, 3, 4            | +     | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | HPNITETAL      | 6                  | +     | +     | +     | +     | +++   | +     | +     | +     | +     | +     |
|      | LPGSEIEIPF     | 5a                 | +     | +     | +     | +     | ++    | +     | +     | +++   | +     | +     |
|      | LGHEGEIPF      | 2                  | +     | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | LGSEGEIPF      | 3                  | +     | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | LPTTGEIPF      | 4, 6               | +     | +     | +     | +     | ++    | +     | +     | +     | +     | +     |
|      | YLVAYQATV*      | 1a, 1b, 4, 5a, 6   | +     | +++   | +     | +     | +     | +     | +     | +     | +     | +     |
|      | YLTAYQATV      | 2, 3               | +     | +++   | +     | +     | +     | +     | +     | +     | +     | +     |
| NS5A | NHVSPHYV       | 1a, 1b, 3, 4, 5a, 6 | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | NHVAPTHYV      | 2                  | +     | +     | +     | +     | +     | +     | +     | +     | +     |
|      | LPWARPQG       | 5a                 | +     | +     | +     | +     | +     | +++   | +     | +     | +     |
|      | LPWARPQD       | 1a, 3, 4, 6        | +     | +     | +     | +     | +     | +++   | +     | +     | +     |
|      | LPWARPQD       | 2                  | +     | +     | +     | +     | +     | +++   | +     | +     | +     |
## Table 4 Binding affinity scores of predicted epitopes and their variants to common HLA I allele types prevalent in South Africa (Continued)

| NSSB | Epitope | HLA Alleles | + | + | + | + | ++ | ++ | + | ++ | +++ | +++ | ++ |
|------|---------|-------------|---|---|---|---|----|----|---|----|-----|-----|----|
| 2696 | YRRCRASGV | 1a, 1b, 2, 3, 5a | + | + | + | + | + | ++ | + | + | ++ | +++ | ++ |
| 2763 | MTRYSAPP | 1a, 1b, 2, 3, 4, 5a, 6 | + | + | +++ | + | + | + | + | + | + | + | + |
| 2889 | LSAFLHGY (#40) | 1a, 1b, 5a | +++ | + | + | +++ | + | + | ++ | + | + | + | + |
|      | LSAFLHGY | 3 | +++ | + | + | +++ | + | + | ++ | + | + | + | + |
|      | LSAFLHGY | 4 | +++ | + | + | +++ | + | + | ++ | + | + | + | + |

*#-indicates that the epitope has been experimentally proven to be a true positive.

*italics indicates amino acid(s) variation in epitope in comparison to the predicted epitope.

++ indicates high binding affinity (<50 IC50nm), ++ indicates medium binding affinity (>50 IC50nm, <500 IC50nm), + indicates poor binding affinity (>500 IC50nm).
Immune responses [47]. A polyepitope-based strategy with multiple components combining core, E1, and E2 proteins; and conserved T-cell epitopes in the NS proteins has been suggested to be a good vaccine candidate for HCV [48].

High number of epitopes was predicted for HLA class II as compared to class I. The findings of this study are consistent with a study by Shehzadi et al. that predicted epitopes in genotype 3 from Pakistan. The study showed that majority of predicted epitopes were found in the NS3 protein for both HLA class I and HLA class II alleles and most of the epitopes were conserved among different genotypes [49]. Although the NS3 region is one of the conserved regions in HCV, variability in the nucleotide and amino acids has been reported by several studies in the same genotype and also in different genotypes [50,51]. A recent study that analysed 1568 NS3-protease sequences from genotypes 1–6 reported that the protease amino acids sequence was moderately conserved and majority of the amino acids clustered in small regions. Of the 181 amino acids analysed 47% showed <1% variability among all HCV genotypes, and 17.1% amino acid positions showing >25.1% variability [51]. The NS3 is considered to be a good cellular target candidate for a therapeutic vaccine [52] since majority of the HCV viral epitopes recognized by CD8+ and CD4+ T-cells were reported in patient responders to interferon therapy [57] and in spontaneous clearance of HCV [58].

Most of the predicted epitopes in the study sequence were found to be conserved across different HCV genotypes with a higher number of epitopes conserved at the anchor residues. The anchor residues are important for epitope high binding affinity to HLA [59]. Conserved epitopes might influence the immunogenic potential

| Position | Predicted epitopes | Genotype of epitope | HLA II-DRB1 allele typesa |
|----------|-------------------|---------------------|--------------------------|
| E2 507   | YCFTPSPVW        | 1a, 1b, 2, 3, 4, 5a, 6 | +++  +  +  +++         |
| 695      | NIVDTQYLY        | 5a                  | +++  +  +++             |
| NS2 938  | LLHLGRLTG        | 5a                  | +    ++  +++             |
| 1025     | LLAPITAYA        | 1a, 3, 4, 5a        | +    ++  +++             |
| NS3 1129 | DLYLVRTA         | 1a, 1b, 5a          | +    +   +++             |
| 1131     | YLVRTAHD        | 1a, 1b, 5a          | +    +   +++             |
| 1391     | LIFCHSKK        | 1b, 2, 3, 4, 5a, 6  | +    +   +++             |
| 1417     | YYRGLHVV        | 5a                  | +    +++  +              |
| 1464     | FSLDPTFTI       | 1a, 1b, 2, 5a       | +    +++  +              |
| 1535     | TTVRLRAYL       | 3, 5a, 6            | +    +    +              |
| 1562     | FTGLTNIDA       | 5a                  | +    +++  +              |
| 1666     | WAALAYC         | 5a                  | +    +    +              |
| NS4B 1774| YLACLSTLP       | 1a, 1b, 2, 3, 4, 5a, 6 | +    +++  +  +     |
| 1919     | LIASRGNN       | 1a, 1b, 2, 3, 4, 5a, 6 | ++   +++  +++  +++     |
| 1920     | IAASRGNN       | 1a, 1b, 2, 3, 4, 5a, 6 | +    +    +++             |
| NS5A 1994| WLAQAKLPLQ     | 5a                  | +    ++  +++             |
| 2099     | YHYITGTO       | 5a                  | +    ++  +++             |
| NS5B 2450| LLRHHNLVY      | 1a, 3, 5a           | +    ++  +++             |
| 2579     | LVVPDGLV       | 2, 3, 4, 5a, 6      | +    +    +              |

*a+++ indicates high binding affinity (<50 IC50nm), ++ indicates medium binding affinity (>50 IC50nm, <500 IC50nm), + indicates poor binding affinity (>500 IC50nm).
since mutations within the epitopes can increase the chance of immune escape [60]. For a vaccine to be effective globally the selected epitopes must cover HLAs of different populations and it must also be conserved among different genotypes. The high mutation rates of viral epitopes and HLA polymorphisms are some of the challenges that are associated with the development of peptide vaccines [61]. Successful epitope vaccine design requires a broad knowledge of HCV genotype diversity. This will help in the proper selection of conserved HCV-specific T-cell epitopes that will help in avoiding HCV immune evasion [62]. This study attempted to ensure maximal coverage of HLA polymorphism and different genotypes by analyzing conserved epitopes considering different HLA alleles.

Majority of the epitopes predicted from HCV proteins isolated from South African genotype 5a were good binders against HLA alleles that are found worldwide. HLA is both polygenic and polymorphic, and the pool of HLA molecules differs for every individual. Different HLA alleles bind peptides with a particular sequence pattern [63]. For an HLA allele to be covered by a set of epitopes, at least one of the epitopes should be capable of inducing an immune response when bound to the corresponding HLA molecule [46]. The epitopes predicted in this study bind to many HLA alleles including the ones common in South Africa and can be used for designing good vaccine candidates that will eventually work in genetically diverse populations. In-vitro and in-silico studies have showed that HLA alleles preferentially bind to conserved regions of viral proteins in human viruses [64].

Very few epitopes were found to be experimentally true positive, however this can be due to the fact that most of the previous studies focused on genotype 1. A limitation of the study was a lack of in-vivo and in-vitro studies to confirm the predicted immunogenic epitopes, which will be the focus of future studies. However in-silico studies still provide the basis for designing good vaccine candidates.

In conclusion, the results of this study demonstrated antigenic T-cell epitopes that are conserved among genotypes and good HLA binders derived from genotype 5a sequences that can be good candidates for vaccine development. Predicted epitopes analysed in this study will contribute to the future design of an efficient vaccine with the use of conserved epitopes to avoid variation in genotypes and as such, it will be able to induce broad HCV specific immune responses. Conserved epitopes among different genotypes will be experimentally tested in the future to determine their involvement in immune response.

Methods
Ethical statement
The study was approved by the Medunsa Research and Ethics Committee of the Faculty of Health Sciences at the University of Limpopo as project no MREC/p/142/2009-PG. The MREC is registered as an Independent Review Board with a reference no (IRB00005122).

Prediction of T-cell epitopes
Genotype 5a full-length sequences available in the GenBank and 6 of the near full length sequences generated from a previous study conducted by our group [24] were aligned and consensus sequences created using BioEdit [65] for the prediction of T-cell epitopes. For HLA class I, prediction for binding alleles was performed using ProPred I (http://www.imtech.res.in/raghava/propred1/) at a 4% default threshold by keeping the proteosome and immunoproteosome filters on at 5% threshold. ProPred I predicts antigenic epitopes for 47 HLA class I alleles [44]. For HLA class II, prediction was performed using ProPred (http://www.imtech.res.in/raghava/propred/) at a 3% default threshold. ProPred predicts antigenic epitopes for 51 HLA class II alleles [45].

Antigenicity of the epitopes
The Antigenicity score of all the predicted epitopes were analysed using VaxiJen v2.0 online antigen prediction (www.ddg-pharmfac.net/vaxijen/). Epitopes having antigenic score >0.5 were selected as antigenic. Vaxijen server performed well with 87% accuracy at a threshold of 0.5 antigenic score for viruses. Vaxijen v2.0 allows antigen classification based on the physicochemical properties of proteins without recourse to sequence alignment.

Epitope conservation analysis
All predicted epitopes were analyzed for conservation using the IEDB database (http://tools.immuneepitope.org/tools/conservancy/iedb_input) at a threshold of 100% conservation in comparison with 406, 221, 98, 33, 45, 45 randomly selected sequences from each of the HCV genotypes 1a, 1b, 2, 3, 4 and 6 respectively. The epitopes were considered conserved in another genotype if it shows 100% identity across the epitope in at least 70% of sequences in that genotype in the randomly selected sequences used in this study, downloaded from the public database. In addition, epitope variants that were conserved in at least 70% of the sequences were analysed for conservancy for anchor residues at positions 2 and 9 for HLA class I and positions 1, 4, 6 and 9 for HLA class II.

Validation of predicted epitopes
All the predicted epitopes were submitted to IEDB database (http://www.immuneepitope.org/) to confirm if they had been tested previously by other studies. The immuneepitope database contains experimentally confirmed data about antibody, T-cell epitopes, HLA binding, HLA restriction and HLA class.
Common South African HLA alleles

For epitope prediction binding to common HLA alleles found in South Africa, the IEDB epitope analysis tool (http://tools.immuneepitope.org/tools/conservancy/iedb_input) was used for HLA class I using the artificial neural network (ANN) algorithm [66] on the IEDB server, while for Class II ProPred (http://www.inttech.res.in/raghava/propred/) at a 3% default threshold was used. ProPred predicts antigenic epitopes for 51 HLA class II alleles [54]. The most common South African alleles were found in published literature [67].

Abbreviations
HCV: Hepatitis C virus; HLA: human leukocyte antigen.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
MPG designed the study, performed the immunoinformatics analysis and drafted the manuscript. SGS and MJM guided the project and critically reviewed the manuscript. All the authors have read and approved the final manuscript.

Acknowledgements
This work was supported by grants from Medical Research Council, National Research Foundation, National Health Laboratory Service Research Trust, and The Stella & Paul Loewestein Charitable & Educational Trust in South Africa.

Received: 17 March 2014 Accepted: 14 October 2014 Published: 8 November 2014

Page 12 of 13
Bouchera R, Kradane-Miked H, Bouzat R, Rasmussen M, Gattad T, Langa-Vives F, Lemercier B, Lim A, Bérad M, Benmahmed L, Buus S, Rooske R, Lemonnier FA: HLA-A*01:03, HLA-A*24:02, HLA-B*08:01, HLA-B*27:05, HLA-B*35:01, HLA-B*44:02, and HLA-C*07:01 monochain transgenic H-2 class I null mice: novel versatile preclinical models of human T cell responses. J Immunol 2013, 191:589–593.

Doi H, Hiroshi K, Shimazaki T, Eguchi J, Baba T, Ito T, Matsumura T, Nozawa H, Monikawa K, Ishi H, Hiraide A, Sakaki M, Imawari M. Magnitude of CD8 T-cell responses against hepatitis C virus and severity of hepatitis do not necessarily determine outcomes in acute hepatitis C virus infection. Hepatol Res 2009, 39:256–265.

Wentworth PA, Sette A, Ellis E, Sidney J, Southwood S, Cottell S, Keogh E, Wong NC, Livingston B, Alazrad D, Vitelio A, Grey HM, Chisari FV. Identification of A2-restricted hepatitis C virus-specific cytotoxic T lymphocyte epitopes from conserved regions of the viral genome. J Immunol 1996, 156:651–659.

Wang S, Bucchi R, Schiller J, Gao J, VanGundy RH, Hildebrand WH, Eckers DD: Natural epitope variants of the hepatitis C virus impair cytotoxic T lymphocyte activity. World J Gastroenterol 2010, 16:5933–1969.

Fitmaurice K, Petrovic D, Ramamurthy N, Simmons R, Merani S, Gaudieri S, Wentworth PA, Sette A, Celis E, Sidney J, Southwood S, Cottell S, Kleneman P: Molecular footprints reveal the impact of the protective HLA-A*03 allele in hepatitis C virus infection. Gut 2011, 60:1563–1571.

Scognamiglio P, Accapezzato D, Casciaro MA, Cacciani A, Artini M, Bruno G, Chicu ML, Sidney J, Southwood S, Abignani S, Sette A, barnabà V: Presence of effector CD8+ T cells in hepatitis C virus-infected healthy non-alcoholic donors. J Immunol 1999, 162:6661–6669.

Chang KM, Gruner NH, Southwood S, Sidney J, Pape GR, Chisari FV, Sette A: Identification of A2-restricted hepatitis C virus epitopes in patients with acute hepatitis C virus infection. Hum Immunol 2011, 16:1156–1164.

Seereenahesh Sam, Jaramillo A, Crippa J, Lisker-Melman M, Chapman WC, Mohanakumar T: Concomitant augmentation of type 1 CD4+ and CD8+ T-cell responses during successful interferon-alpha and ribavirin treatment for chronic hepatitis C virus infection. Hum Immunol 2003, 64:497–504.

Matsui M, Moriya O, Belladonna ML, Kariyana S, Lemmonner FA, Yoshimoto T, Akatsuka T: Adjunctive treatment of novel cytokines, interleukin-23 (IL-23) and IL-27, for induction of hepatitis C virus-specific cytotoxic T lymphocytes in HLA-A*0201 transgenic mice. J Virol 2004, 78:9093–9104.

Mizukoshi E, Nasimaclini M, Blajevic JB, Mihalk K, Rice CM, Liang TJ, Feinlone SM, Rehenbman B: Molecular and immunological significance of chimpanzee major histocompatibility complex haplotypes for hepatitis C virus infection. Immunol Hum Immunol 1999, 162:1156–1164.

Zhang J, Yamada O, Yoshida H, Iwai T, Araki H: Autogenouso transglutamial inhibition of core protein: implication for switch from translation to RNA replication in hepatitis C virus. Virology 2002, 293:141–150.

Matsuda S, Yamada A, Takao Y, Muramatu N, Yutani S, Ide T, Sata M, Itoh K: A new epitope peptide derived from hepatitis C virus 1b possessing A24-restricted T-cell reaction in chronic hepatitis C virus infection. J Hepatol 2002, 36:1237–1247.

Visterme H, Johnson D, Wiedermeier H: Progress in the development and therapeutic vaccines for hepatitis C virus. J Hepatol 2011, 54:1273–1285.

Eckers DD, Sian T, Gill JC, Sanderstrup G: Epitopes of the NS3 protein of hepatitis C virus: recognition in HLA-DR4 transgenic mice. Immune Biol Cell 2002, 89:106–112.

Day CL, Lauer GM, Robbins GK, McGovern B, Wurcel AG, Gandhi RT, Chung RT, Walker BD: Broad specificity of virus-specific CD4+ T-helper cell responses in resolved hepatitis C virus infection. J Virol 2002, 76:12584–12595.

Nakamuta T, Funatoki K, Mota H, Uragawa T, Nakamura A, Ishihara H, Matsuzaki Y, Tanaka N, Inamai M, Yamada H, Imai H, Gaudieri S, Sette A, Chisari FV: CD8 T-cell responses against hepatitis C virus 1b possessing A2-restricted T-cell reaction. J Hepatol 2003, 39:800–807.

Vider-Shalit T, Raffaeli S, Louzoun Y: Lysosomal targeting and attenuation of the replicating hepatitis C virus. Immune Biol Cell 2004, 32:454–463.

Hammer J, Beluncis C, Bolin D, Padopoulos J, Walsky R, Higelin J, Danho W, Singaligla F, Nagy ZA: High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. Proc Natl Acad Sci U S A 1994, 91:4656–4660.

Yusim K, Kemir C, Gachen B, Addo MM, Altfeld M, Brunak S, Chigaev A, Detours V, Korber BT: Clustering patterns of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1 (HIV-1) proteins reveal imprints of immune evasion on HIV-1 global variation. J Virol 2002, 76:8757–8768.

Vidal-Shalit T, Raafael S, Louzouy Z: Virus-epitope vaccine design: Informatic matching the HLA-I polymorphism to the virus genome. Mol Immunol 2007, 44:1253–1261.

Moler-Rubin J, Maufore EM, Flower DR, Reche PA: Selection of conserved epitopes from hepatitis C virus for pan-populational stimulation of T-cell responses. Clin Dev Immunol 2013, 2013:901493.

De Groot AS, Sbai H, Aubin CS, McMurry J, Martin W: Immuno-informatics: Mining genomes for vaccine components. Immune Biol Cell 2002, 80:255–269.

Singh H, Raghava GP: PredPit: prediction of promiscuous MHC Class I binding sites. Bioinformatics 2003, 19:1009–1014.

Singh H, Raghava GP: PredPro: prediction of HLA-DR binding sites. Bioinformatics 2001, 17:1236–1237.

Toussaint NC, Dönes P, Kohlhauser O: A mathematical framework for the selection of an optimal set of peptides for epitope-based vaccines. PLoS Comput Biol 2008, 4:e1000246.

Zeng R, Liu G, Ling S, Zhang H, Yao Z, Xiu B, He F, Huang R, Wei L: A novel combined vaccine candidate containing epitopes of HCV NS3, core and E1 proteins induces multi-specific immune responses in BALB/c mice. Antivir Res 2009, 84:23–30.

Hu CT: Vaccine Development for Hepatitis C: Lessons from the Past Turn into Promise for the Future. Tzu Chi Med J 2005, 17:61–74.

Shehzadi A, Ur Rahman S, Idrees M: Promiscuous prediction and conservancy analysis of CTL binding epitopes of TCV 3a viral proteome from Pakistan: an in silico approach. Tzu Chi J Med 2011, 46:55.