Promiscuous prediction and conservancy analysis of CTL binding epitopes of HCV 3a viral proteome from Punjab Pakistan: an In Silico Approach

Abida Shehzadi1*, Shahid ur Rehman1, Muhammad Idrees2*

Abstract

Background: HCV is a positive sense RNA virus affecting approximately 180 million people worldwide and about 10 million Pakistani populations. HCV genotype 3a is the major cause of infection in Pakistani population. One of the major problems of HCV infection especially in the developing countries that limits the limits the antiviral therapy is the long term treatment, high dosage and side effects. Studies of antigenic epitopes of viral sequences of a specific origin can provide an effective way to overcome the mutation rate and to determine the promiscuous binders to be used for epitope based subunit vaccine design. An in silico approach was applied for the analysis of entire HCV proteome of Pakistani origin, aimed to identify the viral epitopes and their conservancy in HCV genotypes 1, 2 and 3 of diverse origin.

Results: Immunoinformatic tools were applied for the predictive analysis of HCV 3a antigenic epitopes of Pakistani origin. All the predicted epitopes were then subjected for their conservancy analysis in HCV genotypes 1, 2 and 3 of diverse origin (worldwide). Using freely available web servers, 150 MHC II epitopes were predicted as promiscuous binders against 51 subjected alleles. E2 protein represented the 20% of all the predicted MHC II epitopes. 75.33% of the predicted MHC II epitopes were (77-100%) conserve in genotype 3; 47.33% and 40.66% in genotype 1 and 2 respectively. 69 MHC I epitopes were predicted as promiscuous binders against 47 subjected alleles. NS4b represented 26% of all the MHC I predicted epitopes. Significantly higher epitope conservancy was represented by genotype 3 i.e. 78.26% and 21.05% for genotype 1 and 2.

Conclusions: The study revealed comprehensive catalogue of potential HCV derived CTL epitopes from viral proteome of Pakistan origin. A considerable number of predicted epitopes were found to be conserved in different HCV genotype. However, the number of conserved epitopes in HCV genotype 3 was significantly higher in contrast to its conservancy in HCV genotype 1 and 2. Despite of the lower conservancy in genotype 1 and 2, all the predicted epitopes have important implications in diagnostics as well as CTL-based rational vaccine design, effective for most population of the world and especially the Pakistani Population.
HCV mainly spreads via blood supply, reuse of glass syringes and needles, unsterilized medical equipment, use of tooth brushes of HCV patients, etc [7] and causes of acute and chronic infections [8]. Clinical demonstrations of acute Hepatitis C Viral infection include Jaundice, Fever, Myalgia, Fatigue, Lethargy, Increased ALT, Anorexia and Fulminant hepatic failure [7]. About 80% of HCV infected individuals develop chronic infections [9]. Chronic liver infections develop chronic hepatitis, cirrhosis and hepatocellular carcinoma within a period of 10, 20 and 30 years respectively followed by viral infection [10,11]. Out of 70-80% chronically infected individuals, 20% develop cirrhosis and 1-5% individuals suffer from final stage of liver diseases [12]. Hepatic steatosis is the accumulation of lipids in hepatocytes and is reported for the cause of cirrhosis [13] with the more severe cases being reported in patients infected with HCV genotype 3a [14]. The prevalence of steatosis in Pakistani population is about 61.5-65.5% compared with 32.8-81.2% in western countries [15]. The percentage of males infected with HCV chronic liver stage is higher then females with the age of patients between 40-50 years [5].

HCV is classified into six genotypes each having various subtypes [16-18]. These genotypes are distributed differently in various parts of the world with the genetic variance between them is about one third. The genotypes 1, 2 and 3 have world wide distribution. But the significant differences are observed in subtype distribution. Subtype 1a is mostly found in North America and Europe followed by 2b and 3a. Subtype 1b is frequently found in South East Europe and Tunisia and 2c in North Italy. Genotype 4 is mainly restricted to Middle East and Central Africa and genotype 5 in South Africa. Genotype 6 is distributed throughout South East Asia and also being isolated from Hong Kong and Vietnam [17]. The most frequent HCV genotypic distribution in Pakistan is 3a [49.05%] followed by 3b [17.66%] [19]. The knowledge of HCV distribution is crucial for treatment therapy and vaccination because of its predictive value in terms of response to antiviral therapy and vaccination. Effective responses to antiviral therapy are normally associated with genotype 2 and 3 in comparison to any other genotype [20].

HCV replicates at about $10^{12}$ new HCV viruses/day. Replication is carried out by RNA dependent RNA polymerase. RNA polymerase lacks the “proofreading” ability that ensures the high mutation rate of about 8-18 mutations in genomic RNA/year [21,20]. Such a high mutation rate limits the treatment therapy and vaccination. The current treatment therapy for HCV is INF alpha along with ribavirin limited to about 50% population [22]. Although the response rate is not much deterring, but high dosage, long-term treatment and side effects limits the usage [23,21]. There is the possibility that after next few years, new antiviral agents such as inhibitors of the viral protease, helices or polymerase will further improve the response rate of the current therapeutic agents. However, antiviral therapy is not affordable in most developing countries, where the prevalence of HCV is generally the highest. Thus, given the huge reservoir of HCV worldwide, the development of an effective vaccine may be the cheapest way to control disease associated with HCV infection.

Development of an effective HCV vaccine requires understanding of immune response. Viral immune response is associated with Major Histocompatabilty complex protein (MHC) and T lymphocytes/T cell. MHC are classified into 2 broad categories, MHC I and MHC II [24]. MHC initially recognizes the viral antigenic epitopes and presents to T lymphocytes for degradation. MHC I presents the antigenic epitopes to CD8+ T cells and MHC II presents to CD4+ T cells for viral degradation [25,26]. CD8 T cells also referred to as cytotoxic T cells (CTL or Tc), limit viral infections by initial recognizing and their subsequent killing infected cells and secreting cytokines. CD4 T referred to as helper cells or Th cells and provides growth factors and signals for generation and maintenance of CD8 T cells [27]. T cells recognize the antigens only when they are associated with MHC, surface glycoprotein exposed on surface of all vertebrate cells. The selection of T cell epitopes is also important because these are linear and hence easy to synthesize.

A particular vaccine developed against HCV can’t be effective for Pakistani population because of variations in HCV genomic sequences and distribution with regard to geographical area. Since a large number of Pakistani population is infected by HCV3a and number of patients enrolled in public and hospitals is increasing day by day. So there is a current need to develop a vaccine against HCV in particular to HCV3a that will cover approximately maximum Pakistani population. The current vaccines are DNA vaccine, Peptide vaccine and epitopic vaccines. Epitopes are the small antigenic segments of viral proteins and causes infections in the host. Epitopic vaccines provide more potent and controlled immune response and eliminates the potential lethal effects of the use of whole viral proteins [28]. Promiscuous epitopes (epitopes capable of binding maximum number of HLA alleles) may overcome the population coverage. Secondly the conserved epitopes reduces antigen escape associated with the viral mutation [29]. So the present study was designed for the prediction of promiscuous epitopes and to analyze their conservancy in general population. Any mutation in the peptide/epitope will lower the conservancy, so it was hypothesized to analyze the pI value of the mutated amino acid residue, that if remain in the range as was in original epitope provides the likelihood of that particular epitope to be used for epitopic vaccine.
design having an effective control over viral mutation, immune response with minimum side effects.

Methods
Sequence Retrieval and Analysis
The sequence of fully sequenced HCV 3a genome and protein of Pakistani origin was retrieved from NCBI (GU294484). The number of individual bases in the genome i.e. the number of adenine; cytosine, guanine and thymine were calculated from DDBJ database. The molecular weight of proteins, percentage of highly repeated amino acid and the least repeated amino acid in the viral protein was calculated by using sequence and search analysis tool at PIR database (http://pir.georgetown.edu/).

Epitope Prediction
Promiscuous epitopes of HCV 3a viral proteins were predicted for HLA I and HLA II binding alleles using freely available immunoinformatics tools such as ProPred I, and ProPred respectively. In comparison to other epitope prediction tools, Propred 1 and Propred cover maximum number of Human Leukocyte antigens i.e. HLA and being used for epitopic prediction for HBV and tuberculosis. ProPred1 allows the user to predict antigenic epitopes for 47 MHC I alleles and ProPred allows epitopes prediction for 51 MHC II alleles. Predictions through these tools can be carried out at various thresholds from 1 to 10%. The algorithms designed for the working of these tools are based on linear coefficients of matrices. Maximum of the matrices were retrieved from BIMASS where the score of each peptide is calculated in multiplication and/or sum up manner. For example the score of following peptide retrieved from BIMASS where the score of each peptide coefficients of matrices. Maximum of the matrices were retrieved from BIMASS where the score of each peptide is calculated in multiplication and/or sum up manner.

Score = P(1) × A(2) × C(3) × D(4) × P(5) × G(6) × R(7) × A(8) × A(9)

Where P(1) is score of P at position 1.

Epitope Conservancy Analysis
All the predicted epitopes of HCV 3a proteins of Pakistani origin were subjected for worldwide conservancy analysis among HCV genotype 1, 2 and 3. 5 sequences against each HCV protein (used for epitope prediction) were retrieved from NCBI randomly. The predicted epitopes of HCV 3a (Pakistani origin) along with 5 selected sequences of individual genotypes (genotype 1, 2 and 3; one at a time) were submitted to epitope conservancy analysis tool available at IEDB database (http://tools.immuneepitope.org/tools/conservancy/iedb_input). All the epitopes having 77-100% conservancy were selected while rejecting the epitopes having variation at the anchor residues. The anchor residues in the predicted epitopes were highlighted by making it bold. The epitopes that were 100% conserved in the selected proteins of the 3 viral genotypes 1, 2 and 3 were also fully bold. Epitopes with 88/77% conservancy were with single or double amino acid variation respectively and to highlight them bold format was used in the conservancy column against each genotype.

Asteric sign (*) indicates that one out of five selected sequences either does not respond to epitope conservancy or have conservancy lower then 77%. Double asteric sign (**) indicates that only one sequence responds for 77-100% conservancy to the selected epitope.

Validation of varied amino acids using pl value
The Peptides with single or double amino acid variation were analyzed for their hydrophatic characteristics or pl value [32]. The pl gives the information that the varied residue retained the amino acid group or diverted from its normal group in a particular peptide under consideration and thus provides information to be used or their rejection. All the varied amino acid residue with diverted group (with considerable change of pl value) were separated from other using superscript “D” for single variation and “DD” for diverted group for doubly varied residues. The superscript “D” in doubly varied residues of particular peptides represents the partial variation i.e. one of the varied residue retained the amino acid group while other residue shifted the amino acid group by a considerable change of pl value.

Results
HCV 3a genome of Pakistani origin comprises 9474 bp with GC content 2622 and 2700 respectively. The GC contents are 12.35% higher then AT contents. The genome encodes a polyprotein that subsequently get fragmented into structural and non structural protein of obvious molecular weight. The envelope protein E2 comprises highest molecular weight 38755.3 KDa (Table 1).
Leucine (L) a neutral nonpolar amino acid residue has the highest percent of repetition (13.1%) in E2 protein. The least repeated residue of E2 is a basic polar Lysine (K) (1.4%). The shortest segment viral protein is NS4a (5751.69 KDa molecular weight) comprising 54 amino acid residues. Leucine (L) and Valine (V) have highest percentage of repetition (14.8) and Histidine (H), Methionine (M), Thrreonine (T) and Tryptophan (W) are the least repeated amino acid residues (1.9%). The molecular weight of other viral proteins and percent repetition of their amino acid residue for were listed in Table 1. The percentage of amino acid residues gives an out look for their pI value and their probability of incidence in the antigenic epitopes.

F (Phenylalanine), I (Isoleucine), L (Leucine), M (Methionine), V (Valine), W (Tryptophan) and Y (Tyrosine) were mainly the anchor residues for MHC II predicted epitopes and are nonpolar in nature. Total 150 epitopes were predicted against 51 alleles of MHC II (Table 2). The highest number of epitopes was represented by E2 protein comprising 20% of all MHC II predicted epitopes. VFLLNPICGL, FVILVFLLL, WHINSTVLH, FNLLDVPKA, LEILINTHG, VQYLGYVGGS are the promiscuous binders of 45–50 MHC II alleles. E2 is followed by NS2 and NS3 proteins representing 14.66% of the predicted MHC II epitopes. In case of NS2, 14 promiscuous epitopes were predicted with varying binding efficiency. GSRDGVILL, DGVILLTSL, WAAAGLKD, and LQVWVPPLL are the good binders both in term of score and the HLA allele coverage (21, 28, 27 and 28 alleles respectively). E2 predicted epitopes covers 20 to 28 HLA alleles except the PLLHSTTEL epitope that covers only 11 HLA alleles but with highest binding efficiency. NS3 epitopes covers 8 to 25 HLA alleles and were also ranked on the basis of their binding efficiency predicted by the score. The least represented epitopes were by NS5a-1a. It comprises only one epitope (HVKNGSMRL) as predicted promiscuous binders for 16 MHC I binding alleles. The promiscuous binders of MHC I for other proteins were also predicted and summarized in table 3.

Out of total 150 predicted MHC II epitopes, 75.33% were (77–100%) conserve in genotype 3 (Table 1) against the randomly selected viral proteins. Out of 75.33% conserved peptides of genotype 3, 71.68% peptides were 100% conserve while 22.12% peptides were having single residue variation (88% epitope conservancy). Only the

### Table 1

It comprises the data of HCV genome size, Proteins, Molecular weight and percentage of highly repeated and least repeated amino acid residues in individual bases.

| Bases | No. | Proteins | aa Number | Mol. Wt. | Highly repeated aa | % of repetition | Least repeated aa | % of repetition |
|-------|-----|----------|-----------|---------|------------------|----------------|-----------------|----------------|
| Total bp | 9474 | Capsid | 114 | 12985.8 | R | 18.4 | C/F | 0.9 |
| A | 1974 | Core | 75 | 7638.88 | L | 16 | E/K/M/Y | 1.3 |
| C | 2700 | E1 | 190 | 20643.9 | V | 11.1 | E | 1.1 |
| G | 2622 | E2 | 350 | 38755.3 | L | 13.1 | K | 1.4 |
| T | 2178 | NS3 | 149 | 15423.6 | A/G | 11.4 | N | 0.7 |

Leucine (L) a neutral nonpolar amino acid residue has the highest percent of repetition (13.1%) in E2 protein. The least repeated residue of E2 is a basic polar Lysine (K) (1.4%). The shortest segment viral protein is NS4a (5751.69 KDa molecular weight) comprising 54 amino acid residues. Leucine (L) and Valine (V) have highest percentage of repetition (14.8) and Histidine (H), Methionine (M), Thrreonine (T) and Tryptophan (W) are the least repeated amino acid residues (1.9%). The molecular weight of other viral proteins and percent repetition of their amino acid residue for were listed in Table 1. The percentage of amino acid residues gives an out look for their pI value and their probability of incidence in the antigenic epitopes.

F (Phenylalanine), I (Isoleucine), L (Leucine), M (Methionine), V (Valine), W (Tryptophan) and Y (Tyrosine) were mainly the anchor residues for MHC II predicted epitopes and are nonpolar in nature. Total 150 epitopes were predicted against 51 alleles of MHC II (Table 2). The highest number of epitopes was represented by E2 protein comprising 20% of all MHC II predicted epitopes. VFLLNPICGL, FVILVFLLL, WHINSTVLH, FNLLDVPKA, LEILINTHG, VQYLGYVGGS are the promiscuous binders of 45–50 MHC II alleles. E2 is followed by NS2 and NS3 proteins representing 14.66% of the predicted MHC II epitopes. In case of NS2, 14 promiscuous epitopes were predicted with varying binding efficiency. GSRDGVILL, DGVILLTSL, WAAAGLKD, and LQVWVPPLL are the good binders both in term of score and the HLA allele coverage (21, 28, 27 and 28 alleles respectively). E2 predicted epitopes covers 20 to 28 HLA alleles except the PLLHSTTEL epitope that covers only 11 HLA alleles but with highest binding efficiency. NS3 epitopes covers 8 to 25 HLA alleles and were also ranked on the basis of their binding efficiency predicted by the score. The least represented epitopes were by NS5a-1a. It comprises only one epitope (HVKNGSMRL) as predicted promiscuous binders for 16 MHC I binding alleles. The promiscuous binders of MHC I for other proteins were also predicted and summarized in table 3.

Out of total 150 predicted MHC II epitopes, 75.33% were (77–100%) conserve in genotype 3 (Table 1) against the randomly selected viral proteins. Out of 75.33% conserved peptides of genotype 3, 71.68% peptides were 100% conserve while 22.12% peptides were having single residue variation (88% epitope conservancy). Only the
| Epitope start | Predicted T-cell epitopes | HLA alleles | HCV genotype 1 | HCV Genotype 2 | HCV Genotype 3 |
|--------------|--------------------------|-------------|----------------|----------------|----------------|
| Capsid       |                          |             |                |                |                |
| 43           | LGVRATRKA                | 23          | LGVRATRKT      | LGVRATRKT      | LGVRATRKT     |
| 36           | LPRRGPRPILG              | 15          | LPRRGPRPILG    | LPRRGPRPILG    | LPRRGPRPILG   |
| 106          | WGPNDPRPR               | 16          | WGPNDPRPR      | WGPNDPRPR      | WGPNDPRPR     |
| 34           | YLPLRPRGPR              | 24          | YLPLRPRGPR     | YLPLRPRGPR     | YLPLRPRGPR    |
| 21           | VFPPGQQQI                | 8           | VFPPGQQQI      | VFPPGQQQI      | VFPPGQQQI     |
| 35           | VLP RRGPRPIL            | 9           | VLP RRGPRPIL   | VLP RRGPRPIL   | VLP RRGPRPIL  |
| 95           | WLLSPRGSRT              | 28          | WLLSPRGSRT     | WLLSPRGSRT     | WLLSPRGSRT    |
| 29           | KYVGGYVVPR              | 3           | KYVGGYVVPR     | KYVGGYVVPR     | KYVGGYVVPR    |
| 82           | WPLYGNCGLS              | 10          | WPLYGNCGLS     | WPLYGNCGLS     | WPLYGNCGLS    |
| 85           | YGNSECNAV                | 11          | YGNSECNAV      | YGNSECNAV      | YGNSECNAV     |
| 33           | VYLLPRRGPP              | 1           | VYLLPRRGPP     | VYLLPRRGPP     | VYLLPRRGPP    |
| Core         |                          |             |                |                |                |
| 61           | FLALLSCL                 | 50          | FLALLSCL       | FLALLSCL       | FLALLSCL      |
| 64           | LA LSLCIH                | 45          | LA LSLCIH      | LA LSLCIH      | LA LSLCIH     |
| 15           | FA DLMGIPY               | 41          | FA DLMGIPY     | FA DLMGIPY     | FA DLMGIPY    |
| 24           | LVGAPVGGV                | 44          | LVGAPVGGV      | LVGAPVGGV      | LVGAPVGGV     |
| 63           | LLALLSCIL                | 36          | LLALLSCIL      | LLALLSCIL      | LLALLSCIL     |
| 62           | FLALLSCL                 | 24          | FLALLSCL       | FLALLSCL       | FLALLSCL      |
| 32           | VARALAHGV               | 10          | VARALAHGV      | VARALAHGV      | VARALAHGV     |
| 21           | YIPLVGPV                 | 28          | YIPLVGPV       | YIPLVGPV       | YIPLVGPV      |
| 19           | MG YIP LVA                | 26          | MG YIP LVA     | MG YIP LVA     | MG YIP LVA    |
| E1           |                          |             |                |                |                |
| 58           | YV GATTASI               | 41          | YV GATTASI     | YV GATTASI     | YV GATTASI    |
| 140          | MWAHLRL                 | 39          | MWAHLRL       | MWAHLRL       | MWAHLRL      |
| 2            | W NTSGLYV               | 27          | W NTSGLYV     | W NTSGLYV     | W NTSGLYV    |
| 138          | VMW AHIL                 | 28          | VMW AHIL      | VMW AHIL      | VMW AHIL     |
| 56           | BV YVGATTAE             | 21          | BV YVGATTAE    | BV YVGATTAE    | BV YVGATTAE  |
| 9            | VVL NARSN               | 31          | VVL NARSN     | VVL NARSN     | VVL NARSN    |
| 161          | WGVLAGLAY               | 15          | WGVLAGLAY     | WGVLAGLAY     | WGVLAGLAY    |
| 93           | FLVGOAFTF               | 11          | FLVGOAFTF     | FLVGOAFTF     | FLVGOAFTF    |
| 181          | I MMFGSVG                | 91          | I MMFGSVG     | I MMFGSVG     | I MMFGSVG    |
| 130          | MNNWSPAV                 | 35          | MNNWSPAV      | MNNWSPAV      | MNNWSPAV     |
| 134          | WSPAVGMV                | 6           | WSPAVGMV     | WSPAVGMV      | WSPAVGMV    |
| 132          | MNWSPAVCM               | 14          | MNWSPAVCM    | MNWSPAVCM     | MNWSPAVCM    |
| 169          | Y YSMOQGW                 | 18          | Y YSMOQGW    | Y YSMOQGW     | Y YSMOQGW |
| 47           | WTPMPTVPA               | 21          | WTPMPTVPA     | WTPMPTVPA     | WTPMPTVPA    |
| 172          | MQGNWAKV                 | 25          | MQGNWAKV      | MQGNWAKV      | MQGNWAKV     |
| 145          | ILRLPQLTF               | 19          | ILRLPQLTF     | ILRLPQLTF     | ILRLPQLTF    |
| E2           |                          |             |                |                |                |
| 122          | MLEPHHRPVV             | 3           | MLEPHHRPVV   | MLEPHHRPVV    | MLEPHHRPVV |
| 151          | V LLLNPCCG               | 48          | V LLLNPCCG   | V LLLNPCCG    | V LLLNPCCG  |
| 337          | WEFTVLVL                | 4           | WEFTVLVL     | WEFTVLVL      | WEFTVLVL     |
|     | Predicted HLA II epitopes HCV Proteins of Pakistani origin and their conservancy in Genotype 1, 2 and 3 worldwide (Continued) |
|-----|-------------------------------------------------------------------------------------------------------------|
| 339 | FYV/LVFLLL 46                                                                                               | FIV/LVFLLL |
| 35  | WHINSTVLH 41                                                                                                |                                      |
| 342 | LV/LVFLLLADA | LLFLLLADA | LLFLLLADA | LV/LVFLLLADA |
| 100 | V/LAYAPRP 50                                                                                                |                                      |
| 198 | FRPLLPHRL 47                                                                                                |                                      |
| 218 | VRGLALVD 12                                                                                                |                                      |
| 62  | FNLLDVPKA 45                                                                                                |                                      |
| 26  | LELINTHG 46                                                                                                |                                      |
| 57  | FYYH/FNLND 12                                                                                                | FYYH/FNSTDD |
| 83  | VGPLYRQH 26                                                                                                |                                      |
| 58  | YNH/FNLND 24                                                                                                |                                      |
| 286 | LLHSTTEL 17                                                                                                | LLHSTTEL |
| 129 | WAGTTIDPK 14                                                                                                | WAGTTIDKI |
| 320 | VQYLGVGS 46                                                                                                | VQYLGVGS |
| 159 | LLWGGGLG 14                                                                                                |                                      |
| 293 | LAILPCSFT 7                                                                                                | LAILPCSFT |
| 335 | LKWEF/PVLV 4                                                                                                | LKWEF/PVLV |
| 322 | YLYGVSSGM 5                                                                                                | YLYGVSSGM |
| 300 | FTPMPALST 17                                                                                                | FTPMPALST |
| 245 | FYTVCGG 4                                                                                                |                                      |
| 18  | IVRGPEORL 26                                                                                                |                                      |
| 100 | V/LAYAPRP 4                                                                                                |                                      |
| 257 | VWHRFATAAC 19                                                                                              | VEHRLTACDD |
| 206 | LLQETSRGH 8                                                                                                |                                      |
| 1   | YTGTTAAR 8                                                                                                |                                      |
| 267 | WTRGERCDI 10                                                                                                | WTRGERCEI |
| 310 | IHLHQNIVD 11                                                                                                | IHLHQNIVD |
| **NS2** |                                                                                                           |                                      |
| 101 | VRAHVLVRL 51                                                                                                | VRAHVLVRL |
| 62  | VILLTSLLY 50                                                                                                | VILLTSLLY |
| 73  | LV/LVFLLLADA | LLFLLLADA | LLFLLLADA | LV/LVFLLLADA |
| 153 | LD/LAVATE 7                                                                                                | LDKLAVATE |
| 113 | F/RSTVGK 37                                                                                                |                                      |
| 130 | VGRWFNTYLY 11                                                                                                | VGRWFNTYLY |
| 123 | FQMAILSVG 31                                                                                                | FQMAILSVG |
| 137 | YLYDHLAPM 21                                                                                                | YLYDHLAPM |
| 74  | VFDIAKLLA 23                                                                                                | VFDIAKLLA |
| 107 | LV/LVFLLLADA | LLFLLLADA | LLFLLLADA | LV/LVFLLLADA |
| 108 | VRLCMVRS 51                                                                                                | VRLCMVRS |
| 89  | YF/VRAHVLV 33                                                                                                | YF/VRAHVLV |
| 11  | ILV/LVFLLLADA | LLFLLLADA | LLFLLLADA | LV/LVFLLLADA |
| 37  | YACRCESA 18                                                                                                | YACRCESA |
| 33  | W/WNQYACR 8                                                                                                | W/WNQYACR |
| 185 | ILCGLPVSA 10                                                                                                | ILCGLPVSA |
| 145 | MQHWAAGL 18                                                                                                | MQHWAAGL |
| 50  | VFPLLARGS 21                                                                                                | VFPLLARGS |
| 88  | LY/LIQQAAT 35                                                                                                | LY/LIQQAAT |

*Shehzadi et al. Virology Journal 2011, 8:55 http://www.virologyj.com/content/8/1/55 Page 6 of 13*
| 158 | VATEPVIFS | 14 | VATEPVIFS<sup>CD</sup> | VATEPVIFS | VATEPVIFS |
|-----|----------|----|---------------------|----------|-----------|
| 37  | YATCRCSA  | 19 |                     |          | *         |
| 175 | WGDATAACG | 11 | WGDATAACG<sup>CD</sup> | WGDATAACG<sup>CD</sup> | WGDATAACG |

**NS3**

| 4   | VQLVSTATQ | 46 | VQLVSTATQ<sup>CD</sup> | VQLVSTATQ<sup>CD</sup> | VQLVSTATQ |
|-----|-----------|----|----------------------|----------------------|-----------|
| 43  | LQMYTNVDOQ | 42 |                     |                      |           |
| 129 | VCTRGVAKA | 21 | VCTRGVAKA<sup>CD</sup> | VCTRGVAKA<sup>CD</sup> | VCTRGVAKA |
| 24  | WTYYHGAGS | 13 | WTYYHGAGS<sup>CD</sup> | WTYYHGAGS<sup>CD</sup> | WTYYHGAGS |
| 84  | VIPARRRGD | 18 | VIPARRRGD<sup>CD</sup> | VIPARRRGD<sup>CD</sup> | VIPARRRGD |
| 138 | LOFPVETL  | 45 |                     |                      |           |
| 140 | FIPVETLST | 43 | FIPVETLST<sup>CD</sup> | FIPVETLST<sup>CD</sup> | FIPVETLST |
| 6   | VLSATQTQ  | 19 | VLSATQTQ<sup>CD</sup> | VLSATQTQ<sup>CD</sup> | VLSATQTQ |
| 53  | LQGWPAPPQG | 29 | LQGWPAPPQG<sup>CD</sup> | LQGWPAPPQG<sup>CD</sup> | LQGWPAPPQG |
| 27  | YHGAGSRTL | 22 | YHGAGSRTL<sup>CD</sup> | YHGAGSRTL<sup>CD</sup> | YHGAGSRTL |
| 14  | FLGTTLGGV | 10 |                     |                      |           |
| 77  | LVTRHADV1 | 25 | LVTRHADV1<sup>CD</sup> | LVTRHADV1<sup>CD</sup> | LVTRHADV1 |
| 98  | LSPPPLACL | 12 | LSPPPLACL<sup>CD</sup> | LSPPPLACL<sup>CD</sup> | LSPPPLACL |
| 124 | IFRAAVCTR | 44 |                     |                      |           |

**NS4a**

| 23  | VWVGHIEL  | 43 | VWVGHIEL<sup>CD</sup> | VWVGHIEL<sup>CD</sup> | VWVGHIEL |
|-----|-----------|----|----------------------|----------------------|-----------|
| 3   | WLLGGLLAA | 43 | WLLGGLLAA<sup>CD</sup> | WLLGGLLAA<sup>CD</sup> | WLLGGLLAA |
| 4   | VLLGGLLAL | 40 | VLLGGLLAL<sup>CD</sup> | VLLGGLLAL<sup>CD</sup> | VLLGGLLAL |
| 38  | VPDKEVLQY | 11 | VPDKEVLQY<sup>CD</sup> | VPDKEVLQY<sup>CD</sup> | VPDKEVLQY |
| 24  | VWVGHIELG | 8  | VWVGHIELG<sup>CD</sup> | VWVGHIELG<sup>CD</sup> | VWVGHIELG |
| 10  | LAALAYCLS | 8  | LAALAYCLS<sup>CD</sup> | LAALAYCLS<sup>CD</sup> | LAALAYCLS |
| 16  | YCLSLVCGV | 6  | YCLSLVCGV<sup>CD</sup> | YCLSLVCGV<sup>CD</sup> | YCLSLVCGV |
| 26  | VGHIELGGK | 9  | VGHIELGGK<sup>CD</sup> | VGHIELGGK<sup>CD</sup> | VGHIELGGK |
| 25  | VGHIELGG  | 29 | VGHIELGG<sup>CD</sup> | VGHIELGG<sup>CD</sup> | VGHIELGG |
| 20  | VGVWVGH   | 15 | VGVWVGH<sup>CD</sup> | VGVWVGH<sup>CD</sup> | VGVWVGH |
| 9   | VLAALAYC  | 9  | VLAALAYC<sup>CD</sup> | VLAALAYC<sup>CD</sup> | VLAALAYC |
| 29  | IELGGKPAL | 14 | IELGGKPAL<sup>CD</sup> | IELGGKPAL<sup>CD</sup> | IELGGKPAL |

**NS4b**

| 81  | FFNILGGWV | 41 | FFNILGGWV<sup>CD</sup> | FFNILGGWV<sup>CD</sup> | FFNILGGWV |
|-----|-----------|----|----------------------|----------------------|-----------|
| 153 | VNNLPAILS | 51 | VNNLPAILS<sup>CD</sup> | VNNLPAILS<sup>CD</sup> | VNNLPAILS |
| 152 | VNNLPAIL  | 51 | VNNLPAIL<sup>CD</sup> | VNNLPAIL<sup>CD</sup> | VNNLPAIL |
| 39  | WNFSGIQY  | 16 | WNFSGIQY<sup>CD</sup> | WNFSGIQY<sup>CD</sup> | WNFSGIQY |
| 165 | LWGVVCAA  | 35 | LWGVVCAA<sup>CD</sup> | LWGVVCAA<sup>CD</sup> | LWGVVCAA |
| 82  | FNILGGWVA | 32 | FNILGGWVA<sup>CD</sup> | FNILGGWVA<sup>CD</sup> | FNILGGWVA |
| 81  | FFNILGGWV | 5  | FFNILGGWV<sup>CD</sup> | FFNILGGWV<sup>CD</sup> | FFNILGGWV |
| 63  | LMISGASVT | 9  | LMISGASVT<sup>CD</sup> | LMISGASVT<sup>CD</sup> | LMISGASVT |
| 27  | WQKLEFWH  | 35 | WQKLEFWH<sup>CD</sup> | WQKLEFWH<sup>CD</sup> | WQKLEFWH |
| 167 | VGVCACAIIL | 11 | VGVCACAIIL<sup>CD</sup> | VGVCACAIIL<sup>CD</sup> | VGVCACAIIL |
| 45  | IQYLAGLST | 35 | IQYLAGLST<sup>CD</sup> | IQYLAGLST<sup>CD</sup> | IQYLAGLST |
| 64  | MAFASVTS | 23 | MAFASVTS<sup>CD</sup> | MAFASVTS<sup>CD</sup> | MAFASVTS |
| 84  | ILGGWQAVH | 24 | ILGGWQAVH<sup>CD</sup> | ILGGWQAVH<sup>CD</sup> | ILGGWQAVH |
| 103 | VSGLAGAA  | 10 | VSGLAGAA<sup>CD</sup> | VSGLAGAA<sup>CD</sup> | VSGLAGAA |
| 166 | VVGICAAL  | 31 | VVGICAAL<sup>CD</sup> | VVGICAAL<sup>CD</sup> | VVGICAAL |
| Epitope start Position | Predicted T-cell epitopes | HLA alleles | HCV genotype 1 | HCV Genotype 2 | HCV Genotype 3 |
|-------------------------|--------------------------|-------------|----------------|----------------|----------------|
| **Capsid**              |                          |             |                |                |                |
| 38                      | RRGPRLGVR                | 9           | RRGPRLGVR      | RRGPRLGVR      | RRGPRLGVR      |
| 35                      | VLPRRGPRL                | 25          | VLPRRGPRL      | VLPRRGPRL      | VLPRRGPRL      |
| **Core**                |                          |             |                |                |                |
| 55                      | PGCSFSIFL                | 8           | PGCSFSIFL      | PGCSFSIFL      | PGCSFSIFL      |
| 41                      | RALEDGINF                | 20          | RALEDGINF      | RALEDGINF      | RALEDGINF      |
| 7                       | VIDTLTGF                 | 15          | VIDTLTGF       | VIDTLTGF       | VIDTLTGF       |
| 35                      | ALAHGVRAL                | 24          | ALAHGVRAL      | ALAHGVRAL      | ALAHGVRAL      |
| 24                      | LVGAPVGGV                | 18          | LVGAPVGGV      | LVGAPVGGV      | LVGAPVGGV      |
| 26                      | GAPVGGVAR                | 9           | GAPVGGVAR      | GAPVGGVAR      | GAPVGGVAR      |
| **E1**                  |                          |             |                |                |                |
| 135                     | SPAVGMWWA                | 14          | SPAVGMWWA      | SPAVGMWWA      | SPAVGMWWA      |
| 86                      | GDVCGAVF                 | 19          | GDVCGAVF       | GDVCGAVF       | GDVCGAVF       |
| 144                     | HILRLPTQL                | 22          | HILRLPTQL      | HILRLPTQL      | HILRLPTQL      |
| 156                     | IAGAHWGVL                | 27          | IAGAHWGVL      | IAGAHWGVL      | IAGAHWGVL      |
| 64                      | ASIRGHVGD                | 25          | ASIRGHVGD      | ASIRGHVGD      | ASIRGHVGD      |
| 285                     | PLLHSTTEL                | 11          | PLLHSTTEL      | PLLHSTTEL      | PLLHSTTEL      |

**Table 2 Predicted HLA II epitopes HCV Proteins of Pakistani origin and their conservancy in Genotype 1, 2 and 3 worldwide**

**Table 3 Predicted HLA I epitopes HCV Proteins of Pakistani origin and their conservancy in Genotype 1, 2 and 3 worldwide**
| Protein | 305 | 227 | 71 | 157 | 212 | 6 | 157 | NS2 | 119 | 27 | 128 | 57 | 98 | 95 | 130 | 7 | NS3 | 3 | 23 | 10 | 5 | NS4a | 96 | 40 | 81 | 46 | 102 | 54 | 161 | 141 | 164 |
|---------|-----|-----|----|-----|-----|---|-----|-----|-----|----|-----|----|----|---|-----|---|-----|---|-----|----|-----|-----|-----|-----|----|-----|-----|-----|
| ALSTGLIH | 25 | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLIH | ALSTGLI |
40% peptides of singly varied residues diverted their amino acid group and the pl value while 60% singly varied residues retained the amino acid group as was in the predicted epitope of HCV 3a proteins. 6.19% peptides comprised the 77% epitope conservancy because of double residue variation in the peptides of general population in contrast to predicted epitopes of HCV 3a of Pakistani origin. Out of 6.19%, doubly varied amino acid residues 42.85% peptides retained their amino acid group and nearly same pl value as in case of predicted epitope while 28.57% peptides were having partial group diversion and 28.57% (of doubly varied amino acid residues) peptides diverted their amino acid group because of considerable variation in the pl value. Similar data was also obtained for the HCV genotype 1 and 2 consisting 55.07% conservancy. However, in contrast to genotype 3, only 23.94% predicted epitopes were 100% conserve in randomly selected sequences of genotype 1 and 22.22% in genotype 2. Their rate of single/double residue variation was also predicted and expressed as figure 1.

Out of total 69 predicted MHC I epitopes, 78.26% were (77-100%) conserve in genotype 3 (Table 2) against the randomly selected viral proteins. Out of 78.26% conserved peptides of genotype 3, 72.22% peptides were 100% conserve while 22.22% peptides were having single residue variation (88% epitope conservancy). 40.66% peptides of singly varied residues retained the amino acid group as was in the predicted epitope of HCV 3a proteins while 58.33% singly varied residues diverted their amino acid group and the pl value. 5.5% peptides comprised the 77% epitope conservancy because of double residue variation in the peptides of general population in contrast to predicted epitopes of HCV 3a of Pakistani origin. Out of 5.5%, doubly varied amino acid residues 66.66% peptides were having partial group diversion and 33.33% (of doubly varied amino acid residues) peptides diverted their amino acid group because of considerable variation in the pl value. Similar data was also obtained for the HCV genotype 1 and 2 consisting 55.07% conservancy. However, in contrast to genotype 3, only 21.05% predicted epitopes were 100% conserve in randomly selected sequences of genotype 1 and 2. Their rate of single/double residue variation was also predicted and expressed as figure 2.

**Discussion**

The modern technique for control of HCV infection is a vaccine preparation that can specifically induce antibody-mediated immunity. The rapid advancements in the computational methodologies and immunoinformatics/immuno-bioinformatics provide new strategies for the synthesis of antigen specific epitopic vaccine against infectious agents such as viruses and pathogens. Epitopic vaccine against HIV, malaria and tuberculosis provided promising results and supported the defensive
Figure 1 A comparative analysis of HCV 3a Predictive epitopes against MHC II alleles and their conservancy analysis in Genotype 1, 2 and 3 worldwide.

Figure 2 A comparative analysis of HCV 3a Predictive epitopes predicted against MHC I and their conservancy analysis in Genotype 1, 2 and 3 worldwide.
and therapeutic uses of these vaccines [33]. Thus in the present study, a new systematic immunoinformatics approach was applied for the predicted antigenic epitopes of HCV 3a proteins of Pakistani origin followed by diversity and conservancy in other genotypes (1,2 and 3) in randomly selected HCV sequences from NCBI and mainly belong to Thailand, Cuba, UK, USA, China, Japan, France, Italy and Germany. The immunogenic epitopes identified were nanomers and could be used diagnostically to detect HCV specific CTL responses in the patients and after vaccination. A CTL based HCV vaccine might not efficient enough to prevent from infection but it might protect the body from the disease. The analysis showed that the minimal number of epitopes required to represent the complete antigenicity of the whole proteins are significantly smaller then required to represent full length proteins. The majority of the epitopes reported here had intermediate to high HLA binding affinity.

By the use of an efficient CTL based epitope delivery technology; the predicted epitopes could eventually become vaccines in their own or fused as polytopes. The design of the HCV vaccine using conserved epitopes can avoid viral mutation and thus provides more efficient results. The study shows that the predicted epitopes were highly conserved in HCV genotype 3 and also but less conserved in genotype 1 and 2 both for MHC I and MHC II. Moreover, to ensure the viral detection at all stages of its intracellular evolution we have used all the viral proteins. Therefore, the total number of predicted epitopes were also maximized in correspondence to the number of covered proteins used for the analysis.

Abbreviations
HCV: hepatitis C virus; HLA: human leukocyte antigen; MHC: major histocompatibility complex; CTL: cytotoxic T lymphocytes.

Author details
1 Bioinformatics Division, Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan. 2 Division of Molecular Virology, Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan.

Authors’ contributions
AS and SH designed the study. AS performed the immunoinformatics analysis and drafted the manuscript. MI critically reviewed the manuscript. All authors have read and approved the final manuscript.

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

Received: 25 January 2011 Accepted: 8 February 2011 Published: 8 February 2011

References
1. Qi R, Zhang L, Chi C. Biological characteristics of dengue virus and potential targets for drug design. Acta Biochim Biophys Sin 2008, 40:91-101.
2. Bitvich JB, Fernandez-Salas I, Contreras-Cordero FJ, Marleeney LN, Gonzalez-Rojas UJ, Komar N, Gubler JD, Calisher HC, Beaty JB. Serologic Evidence of West Nile Virus Infection in Horses, Coahuila State, Mexico. Emerging Infectious Diseases 2003, 9(7):853-856.
3. Moening V: Introduction to classical swine fever: virus, disease and control policy. Veterinary Microbiology 2000, 73:93-102.
4. Stiffler DJ, Nguyen M, Sohn AJ, Liu C, Kaplan D, Seeger C. Focal Distribution of Hepatitis C Virus RNA in Infected Livers. PLoS ONE 2009, 4(8):1-7.
5. Raja SN, Janjua AK. Epidemiology of hepatitis C virus infection in Pakistan. J Microbiol Immunol Infect 2008, 41:1-8.
6. Raja SN, Singh NN, Janjua AK, Janjua-Ul-Haq R. Hepatitis C virus infection: An enigma continues. Medicine Today 2006, 4(3):93-101.
7. Das RB, Kundu B, Khandaparkar R, Sahni S. Geographical distribution of hepatitis C virus genotype in India. Indian J Pathol Microbiol 2002, 45(3):325-328.
8. Rychliková M, Biedikowska-Szewczyk K. Hepatitis C - new developments in the studies of the viral life cycle. Acta Biochimica Polonica 2007, 54(4):703-715.
9. Jobarteh Modou, Malfroy Marine, Peterson Ingrid, Jeng Adam, Ramu Sar-Nej, Alabi Abraham, Peterson Kevin, Cotten Matt, Hall Andrew, Sarah Rowland-Jones, Whittle Hilary, Tedder Richard, Jaye Assan, Mdmy Maimuna. Seroconversion of hepatitis B and C virus in HIV-1 and HIV-2 infected Gambians. Virology Journal 2010, 7:230-239.
10. Re DV, Caggiari L, Vita DS, Mazzoro C, Lenzi M, Galli M, Monti G, Ferni C, Zignego LA, Gabrielli A, Sansonno D, Dammacco F, Libra M, Sacchi N, Talamini R, Spina M, Cannizzaro R, Guidoboni M, Dolcelli R. Genetic insights into the disease mechanisms of type II mixed cryoglobulinemia induced by hepatitis C virus. Digestive and Liver Disease 2007, 39(1):65-71.
11. Waiw I, Wairz F, Javed M, Saeed M, Najeeb-ul-Haq Khan H. Effect of vitamin “e” supplements in herapy of chronic hepatitis c: a histological study. Gomal Journal of Medical Sciences 2006, 6(2):81-86.
12. Jia Y, Wei L, Jiang D, Wang J, Cong X, Fei R. Antiviral action of interferon-α against hepatitis C virus replicon and its modification by interferon-γ and interleukin-8. J Gastroenterology and Hepatology 2007, 22:1278-1285.
13. Yoon JE, Hu K. Hepatitis C Virus (HCV) Infection and Hepatic Steatosis. Int J Med Sci 2006, 3(2):53-56.
14. Minakari M, Sarreni KF, Shalmani MH, Molaei M, Zali M. Hepatic Steatosis in Iranian Patients with Chronic Hepatitis C. Med Prin Pract 2008, 17:126-130.
15. Zubair A, Jamal S, Mubarak A. Morphometric Analysis of Hepatic Steatosis in Chronic Hepatitis C Infection. The Saudi Journal of Gastroenterology 2009, 15(1):11-4.
16. Ali Armajad, Habib Ahmad, Ali Izz, Sheema Khan, Gulshan Zadi, Muhammad Idrees: Prevalence of active hepatitis c virus infection in district mansehra Pakistan. Virology Journal 2010, 7:334-338.
17. Ramia S, Ess-Fares J. Distribution of hepatitis C virus genotypes in the Middle East. International Journal of Infectious Diseases 2006, 10:272-277.
18. Ali Armajad, Habib Ahmad, Muhammad Idrees: Molecular epidemiology of Hepatitis C virus genotypes in Khyber Pakhtoonkhw of Pakistan. Virology Journal 2010, 7:203-209.
19. Idrees M, Razuddin S. Frequency distribution of hepatitis C virus genotypes in different geographical regions of Pakistan and their possible routes of transmission. BMC Infectious Diseases 2008, 8(69):1-9, 69.
20. Lauer G, Walker B. Hepatitis C Virus Infection. N Engl J Med 2001, 345(1):41-52.
21. Javid A, Khwaja RA. Treatment and Vaccination for Hepatitis C: Present and Future. J Ayub Med Coll 2008, 20(1):129-133, Abbottabad.
22. Pan Q, Henry DS, Metselaar JH, Schoote B, Kevrekidkboon J, Tihanus WH, Janssen ALH, Laan WJL. Combined antiviral activity of interferon-α and RNA interference directed against epatisis without affecting vector delivery and gene silencing. J Mol Med 2009, 87:713-722.
23. Gozlan J, Lacombe K, Gaul E, Raguen G, Girard P. Complete cure of HBV-HDV co-infection after 24 weeks of combination therapy with pegylated interferon and ribavirin in a patient co-infected with HBV/HCV/HDV/ HGV. Journal of Hepatology 2009, 50:432-434.
24. Tan Lei, Lu Hujian, Dan Zhang, Mingyao Tian, Hu Bo, Wang Zhuyue, Jin Ningyi. Protection against H1N1 influenza challenge by a DNA vaccine expressing H3/H1 subtype hemagglutinin combined with MHC class II-restricted epitopes. Virology Journal 2010, 7:363-376.
25. Reche AP, Reinherz LE: PE PVAC: a web server for multi-epitope vaccine development based on the prediction of supertypic MHC ligands. Nucleic Acids Research 2005, 33:138-142.

26. Zhang LG, Khan MA, Srinivasan NK, August TJ, Brusic V: MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides. Nucleic Acids Research 2005, 33:172-179.

27. Robinson LH, Amara RR: T cell vaccines for microbial infections. Nature Medicine Supplement 2005, 11(4):25-32.

28. Groot DSA, Sbai H, Aubin SC, McMurry J, Martin W: Immuno-informatics: Mining genomes for vaccine components. Immunology and Cell Biology 2002, 80:255-269.

29. Lund O, Nielsen M, Kesmir C, Petersen GA, Lundegaard C, Womning P, Sylvester-Hvid C, Lambeth K, Røder G, Justesen S, Buus S, Brunak S: Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics 2004, 55(12):797-810.

30. Singh H, Raghava SPG: ProPred1: prediction of promiscuous MHC Class-I binding sites. Bioinformatics 2003, 19(8):1009-1014.

31. Somvanshi P, Singh V, Seth KP: In Silico Prediction of Epitopes in Virulence Proteins of Mycobacterium Tuberculosis H37Rv for Diagnostic and Subunit Vaccine Design. Journal of Proteomics & Bioinformatics 2008, 1(3):143-153.

32. Kyte J, Doolittle FR: A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology 1982, 157(1):105-132.

33. Khan MA, Miotto O, Heiny TA, Salmon J, Srinivasan NK, Nascimento MJE, Marques JETA, Brusic V, Tan WT, August TJ: A systematic bioinformatics approach for selection of epitope-based vaccine targets. Cellular Immunology 2006, 244:141-147.

doi:10.1186/1743-422X-8-55

Cite this article as: Shehzadi et al: Promiscuous prediction and conservancy analysis of CTL binding epitopes of HCV 3a viral proteome from Punjab Pakistan: an In Silico Approach. Virology Journal 2011 8:55.