A Comprehensive in Silico Analysis for Identification of Immunotherapeutic Epitopes of HPV-18

Bharti Gupta1 · Anoop Kumar2 · Parikipandla Sridevi1

Accepted: 8 September 2021 / Published online: 20 September 2021 © The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Human papillomavirus (HPV) remains the major cause of cervical cancer, globally. High risk HPV (Hr-HPV) 16 and 18 together account for more than 70% of cervical cancer cases, whereas the hr-HPV-18 is the second most prevalent hr-HPV type, causing about 5.2% of all cancers worldwide. Considering the high prevalence and mortality rate, cervical cancer remains a noteworthy health problem among women. As of now, no registered immunotherapies are available after the HPV infection. Thus, developing an immunotherapeutic candidate against hr-HPV would be of major clinical benefit. Nowadays, the T-and B-cell peptide based targeted vaccines have been considered as the best candidate for vaccine development against viral infections. In this study, both prophylactic and therapeutic vaccine candidates against hr-HPV-18 were predicted. To achieve this, the prediction of T-and B-cell epitopes of major histocompatibility complex (MHC) were accomplished, that can be used for HPV immunotherapy. For MHC-I, a maximum number (20) of potent peptides were found, against HLA-B*51:01 (L1 = 9, L2 = 6, E2 = 4, and E4 = 1) having percentile value < 1 and, immunogenicity scores higher than 0.5, followed by HLA-A*11:01 (L1 = 8, E2 = 7 L2 = 2, and E6 = 1, E7 = 1); 19 epitopes. For MHC-II, the highest number of peptides found, against the HLA-DRB1*04:01 (L2 = 10, E5 = 7, and E4 = 4), HLA-DRB1*04:05 (E5 = 7, E2 = 5, E4 = 5, and L1 = 4) HLA-DPA1/DPB1*01:04:04:05 (E7 = 7, E6 = 5, L2 = 5, and E2 = 2), HLA-DRB5*01:01 (E6 = 6, L1 = 6, and L2 = 6); peptides 21, 21, 19 and 18 respectively. For B-cell, total 94, 16 amino acid long B-cell epitopes were predicted. In conclusion, these predicted epitopes can be valuable candidates for in vitro or in vivo therapeutic vaccine studies against hr-HPV-18 associated cancer.

Keywords Hr-HPV-18 · MHC-I · MHC-II · B-cell · Epitope prediction · Immunotherapy

Introduction

Cervical cancer ranks as the 2nd leading cause of female cancer in India. 96,922 new cervical cancer cases with a crude incidence rate 14.9% are being diagnosed annually (Bruni et al. 2019). High Risk HPVs (Hr-HPVs) primarily 16 and 18 are the oncogenic HPV types are responsible for about 70% of all cervical cancer cases worldwide (Bruni 2019). According to Catalan Institute of Oncology and Information Centre on HPV and Cancer (ICO/IARC) statistics, 469.1 million women aged ≥ 15 years are at risk of development of cervical cancer in India (Bruni 2019). HPVs are circular, double strand, non-enveloped DNA viruses of about ~ 8 kb (Doorbar et al. 2012) that can be divided into early, late, and long control regions (LCR) (Zheng and Baker 2006). Early region includes E1, E2, E4, E5, E6, and E7 protein coding part which are regularly in function and plays important role in HPV oncogenicity. E1 participates in HPV DNA replication process (Castro-Muñoz et al. 2019), along with E2; E2 also works as transcription activator and regulator (Oliveira et al. 2006), E4 (Yajid et al. 2017); (Doorbar 2006) and E5 modulates the productive phase of HPV life cycle (Müller et al. 2015). E5, E6 and E7 are considered as the major hallmarks of HPV infection, these three are also essential for the development of HPV positive carcinoma (Estêvão et al. 2019). E6 and E7
modulates cell cycle control and contributes to viral genome maintenance (Schiffman et al. 2016). E6 inhibits the p53 and causes loss in cell cycle regulation, thus overriding the cell cycle process. E7 mediates the inhibition of retinoblastoma protein and results in unrestrained cell proliferation (Pal and Kundu 2020), being the major onco-players in the process of HPV mediated cervical cancer, while E6 and E7 represents the most effective targets for immunotherapeutic (Pal and Kundu 2020).

Late region of HPV viruses consist of capsid coding protein regions; L1 major capsid protein, and L2 minor capsid protein. Through L1 protein virus initiates its interaction with host cells, gets attached to Heparan Sulfate Proteoglycans (HSPG) receptors of the cell (Posner and Peterson 2013). L1 interaction with HSPG, followed by conformational changes and cleavage of L2 by cellular furin (Bronnimann et al. 2016), thus entering the virus into host cell (Richards et al. 2006). LCR of HPV have multiple functions in regulating the viral transcription (Fang et al. 2020), comprises about 10% of complete genome. Some specific LCR mutations results in the development of cervical cancer (Xi et al. 2017).

It has already been established that persistent infection with hr-HPVs is associated with cancerous lesions, invasion and cancer (Radley et al. 2016). In the majority of HPV infected persons, the infection is cleared by the immune system (Gillison et al. 2000); however, the viral infection can continue to persist and subsequently results into cancer at the site of infection (Frazer 2009). Although such persistent is relatively low, the prevalence of HPV associated cancers is significantly high among the general population worldwide. Prophylactic vaccines against some predominant HPVs have also been developed; Cervarix (GlaxoSmithKline) for HPV 16 and 18, and Gardasil (Merck & Co.) for HPV 16,18,6, and 11 (Ribeiro-Muller et al. 2013). Despite this, developed prophylactic vaccines provide no therapeutic benefit and are only generates the antibodies against the L1 capsid protein of HPVs (Schiller et al. 2012). However, these come under preventative measures for cervical cancer and should be administrated before adolescent age (LaVigne and Lei-tao 2019); (Kaarthigeyan 2012). Moreover, high prevalence and mortality rate due to HPV infection indicating a serious concern to develop effective treatment strategies to control HPV infection and cease the development of cervical cancer.

Therapeutic vaccine development is one of potential treatment method that has now been explored to treat and clear the existing HPV infection, therapeutic vaccines differ from prophylactic vaccines (Zur Hausen 2002), as they aimed to stimulates cell mediated immunity rather than neutralising antibodies, thus targeting and killing the infected cells. Peptide based therapeutic vaccines have the advantages of stability, safety and their feasibility of largescale production (Hung et al. 2008). Several studies showed the constant expression of HPV oncoproteins i.e. E6 and E7 proteins in cervical cancer cases, but not in normal tissues. Due to this E6 and E7 protein of HPV makes them ideal target for the therapeutic vaccine development. Peptide based vaccines are categorized into; specific epitope/short peptides and synthetic long peptides (SLPs), in this study we predicted specific epitope/short peptides as these short peptides binds exogenously to the major histocompatibility complexes (MHCs) (van der Burg et al. 2006). This research paper covers prediction of all the potent epitopes against one of the second most prevalent hr-HPV (hr-HPV-18) worldwide. As several attempts by many researchers have been made to identify potent vaccine target against hr-HPV-16 (Kumar, et al. 2015a,b,c); (Bahmani et al. 2020); (Nakagawa et al. 2004), not much efforts have been made to predict epitopes against hr-HPV-18-E2, E4, E5, E6, E7, L1, and L2 proteins altogether.

**Methods**

**Hr-HPV-18 Protein Sequence Retrieval**

The complete hr-HPV-18 genome ID was NC_001357.1 and the complete sequence was retrieved from the NCBI (https://www.ncbi.nlm.nih.gov/nuccore/NC_001357) database.

**Protein Sequence Analysis**

The number of amino acids, molecular weight of proteins, isoelectric points, percentage of strongly basic, acidic, hydrophobic, polar amino acids, in the hr-HPV-18 E2, E4, E5, E6, E7, L1, and L2 proteins was calculated using Protparam (http://web.expasy.org/protparam/) software.

**Secondary Structure Prediction**

An online server, PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) was used to analyze the secondary structure of hr-HPV-18; E2, E4, E5, E6, E7, L1 and L2 proteins, respectively.

**MHC-I T-Cell Epitopes Prediction**

Prediction of hr-HPV-18 proteins Major Histocompatibility Complex I (MHC-I) T cell epitopes was done using Immune Epitope Database Analysis (IEDB) (http://tools.iedb.org/mhci/). The frequently expressed HLAs were selected for the analysis. A method specific IC50 value was selected along with a low Percentile Rank (PR) value, as a lower IC50 or PR means high affinity. In this study predictions
were performed against 9 mer, peptides with PR < 0.5, and Immunogenicity Score (IS) was set at <1 for research.

**MHC-II T-Cell Epitopes Prediction**

The MHC-II T cell epitope binding prediction was done using IEDB MHC-II binding predictor tool (http://tools.iedb.org/mhcii/). The prediction for 15 mer achieved by Consensus (smm/nn/sturniolo)/(comb.lib./smm/nn), and NetMHCIIpan.

**B-cell Epitopes Prediction**

The ABCpred (ABCpred submission page (osdd.net)) software was used to identify B-cell epitopes against hr-HPV-18 proteins, default parameters with an 0.8 threshold value were used for the prediction method.

**Visualization of the 3D Structure of Predicted Epitopes**

The 3D structure of HPV-18 E2, E4, E5, E6, E7, L1, and L2 proteins were retrieved from the RCSB PDB database (E2 PDB ID: IF9F; E4 PDB ID:6ZFG; E5 PDB ID: 2R5I; E6 PDB ID: 2I0I; E7 PDB ID: 6IWD; L1 PDB ID: 5W1X; L2 PDB ID: 1QQH and the predicted epitope were marked on the 3D structure using visualization chimera software.

**Results**

**Structure Analysis**

The amino acid composition of each protein (hr-HPV-18 E2, E4, E5, E6, E7, L1 and L2 proteins) and individual detailed results of them are given in the form of supporting information Excel files (S1 File). Threonine (Thr/T) and Leucine (Leu/L) were the most frequent amino acids found in hr-HPV-18 proteins.

**Epitope Prediction for MHC-I Alleles**

MHCs are highly polymorphic; they have different alleles within the population with a diverse peptide binding specificity. In alleles which are majorly expressed in the humans were selected in this study from the dbMHC database. MHC-I alleles were HLA-A*02:11, HLA-A*03:01, HLA-A*11:01, HLA-A*24:02, HLA-A*26:01; HLA-B*40:06, HLA-B*08:01, HLA-B*35:03, HLA-B*44:03, HLA-B*51:01; HLA-C*04:01, HLA-C*07:01, HLA-C*07:02, 14:02, and 15:07. In the present study a total of 143 epitopes were predicted against the MHC-I; all the predicted epitopes having a length of 9mer, with less than 1 percentile value were chosen. Moreover, all the peptides had an immunogenicity score higher than 0.5 and percentile value <1. A maximum number of potent peptides found against HLA-B*51:01, and HLA-A*11:01 followed by the other HLAs. The list of potent predicted epitopes, against MHC-I, with high immunogenicity score are given in Table 1. The detailed results of all the epitopes against MHC-I are given in supporting information, as excel files (S2 File).

**Epitope Prediction for MHC-II Alleles**

For MHC-II the HLA reference list from the dbMHC database were chosen and alleles were HLA-DRB1*01:01, HLA-DaRB1*03:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*08:02, HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB1*12:01, HLA-DRB1*13:02, HLA-DRB1*15:01, HLA-DRB3*02:02, HLA-DRB4*01:01, HLA-DRB5*01:01, HLA-DQA1*05:01/DBB1*02:01, HLA-DQA1*05:01/DBB1*03:01, HLA-DQA1*03:01/DBB1*03:02, HLA-DQA1*04:01/DBB1*04:02, HLA-DQA1*01:01/DBB1*05:01, HLA-DQA1*01:02/DBB1*06:02, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*01:03/DBB1*02:01, HLA-DPA1*01:03/DPB1*04:02, HLA-DPA1*03:01/DPB1*04:02, HLA-DPA1*02:01/DPB1*05:01, and HLA-DPA1*02:01/DPB1*14:01. A total 222 potent epitopes were found against the MHC-II. The parameter for peptide selection were chosen at default 15 mer length by the IEDB software. Peptides with percentile rank <1 were selected, the highest number of peptides found against the HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DPA1*01:03/DBB1*04:01, HLA-DRB5*01:01; peptides 21, 21, 19 and 18 respectively. The list of potent predicted epitopes, against MHC-II, with high immunogenicity score are given in Table 2. The detailed results of all the epitopes against MHC-II are given in supporting information, as excel files (S3 File).

**B-Cell Epitope Analysis**

The B-cell epitopes for E2, E4, E5, E6, E7, L1, and L2 proteins were predicted using ABCpred with a threshold value 0.80 and other default parameters. Total 94, 16 amino acid long B-cell epitopes were predated against hr-HPV-18 proteins (E2 = 24, E4 = 05, E5 = 02, E6 = 04, E7 = 01, L1 = 30 and L2 = 28). The list of potent predicted epitopes, with high immunogenicity score are given in Table 3. The detailed results of all the epitopes against B-cell are given in supporting information, as excel files (S4 File). The epitopes were marked on the 3D structure of protein (in Fig. 1).
| Sl.no | Protein region | Epitope | Start | End | Length | Immuno- | Percentile rank |
|-------|----------------|---------|-------|-----|--------|---------|---------------|
|       |                |         |       |     |        | genicity | score         |
| 2     | HPV-18 E2      | STSVGTAK| 230   | 238 | 9      | 0.948363| 0.01          |
|       |                | DSVQILVGY| 354   | 362 | 9      | 0.944315| 0.01          |
|       |                | TPSPYSSSTV| 224   | 232 | 9      | 0.942037| 0.01          |
|       |                | SPYSSSTVSV| 226   | 234 | 9      | 0.937738| 0.01          |
|       |                | NTTPIIHLK| 285   | 293 | 9      | 0.908606| 0.02          |
|       |                | GYNTFYIEF| 168   | 176 | 9      | 0.892103| 0.03          |
|       |                | YVAWDVSYY| 135   | 143 | 9      | 0.840453| 0.03          |
|       |                | RYKTEDWTL| 90    | 98  | 9      | 0.867001| 0.04          |
|       |                | HYRDISSTW| 312   | 320 | 9      | 0.84558 | 0.04          |
|       |                | YYMTDAGTW| 142   | 150 | 9      | 0.841327| 0.04          |
| 12    | HPV-18 E4      | SYSTPPHRI| 20    | 28  | 9      | 0.942668| 0.01          |
|       |                | TRYPILSLL| 10    | 18  | 9      | 0.728053| 0.01          |
|       |                | SIVDLSTHF| 58    | 66  | 9      | 0.65812 | 0.06          |
|       |                | CAVPVVTTRY| 4     | 12  | 9      | 0.684205| 0.06          |
|       |                | DGNSSVVVTL| 78    | 86  | 9      | 0.658506| 0.09          |
| 17    | HPV-18 E5      | VPLLPSVCM| 20    | 28  | 9      | 0.78659 | 0.05          |
|       |                | MLLHHHAI| 61    | 69  | 9      | 0.73391 | 0.05          |
| 19    | HPV-18 E6      | SYVGDTELK| 84    | 92  | 9      | 0.98913 | 0.01          |
|       |                | VYGDTLEKL| 85    | 93  | 9      | 0.9184  | 0.02          |
|       |                | VYCKTVLEL| 33    | 41  | 9      | 0.8948  | 0.03          |
|       |                | NPAEKLRH| 113   | 121 | 9      | 0.84636 | 0.03          |
|       |                | DPTRRPYKL| 6     | 14  | 9      | 0.80712 | 0.03          |
|       |                | DEFYSRIREL| 70    | 78  | 9      | 0.75091 | 0.05          |
|       |                | FEDPTRRPY| 4     | 12  | 9      | 0.82963 | 0.06          |
|       |                | AFKDLFVVY| 48    | 56  | 9      | 0.7399  | 0.07          |
| 27    | HPV-18 E7      | ATLQDIVLH| 6     | 14  | 9      | 0.53839 | 0.26          |
| 28    | HPV-18 L1      | IYNPETQRL| 151   | 159 | 9      | 0.974224| 0.01          |
|       |                | DMVSYIHSM| 448   | 456 | 9      | 0.954823| 0.01          |
|       |                | FPIFLQMAL| 56    | 64  | 9      | 0.963233| 0.01          |
|       |                | LPDPNKFGL| 138   | 146 | 9      | 0.953865| 0.01          |
|       |                | EEYDLQFIF| 430   | 438 | 9      | 0.973715| 0.01          |
|       |                | LPPPSVARV| 74    | 82  | 9      | 0.946367| 0.01          |
|       |                | VPLDICOQSI| 281   | 289 | 9      | 0.928381| 0.01          |
|       |                | FYHAGSSRL| 95    | 103 | 9      | 0.978774| 0.01          |
|       |                | IYNPETQRL| 151   | 159 | 9      | 0.95655 | 0.01          |
|       |                | LYPHRPLPL| 21    | 29  | 9      | 0.948142| 0.01          |
| 38    | HPV-18 L2      | STTSFAFFK| 365   | 373 | 9      | 0.962559| 0.01          |
|       |                | APSPEYIEL| 327   | 335 | 9      | 0.951818| 0.01          |
|       |                | TPLPTVRRV| 213   | 221 | 9      | 0.949236| 0.01          |
|       |                | TRPSSSLITY| 243   | 251 | 9      | 0.865615| 0.01          |
|       |                | YYLWPLYYF| 435   | 443 | 9      | 0.937947| 0.02          |
|       |                | EFLTRPSSL| 240   | 248 | 9      | 0.857795| 0.02          |
|       |                | CPDPDVVK| 27    | 35  | 9      | 0.891908| 0.02          |
|       |                | AFEPVDITIL| 255   | 263 | 9      | 0.832221| 0.02          |
|       |                | EPVDTTLTF| 257   | 265 | 9      | 0.886599| 0.03          |
|       |                | SYSNVTVPL| 383   | 391 | 9      | 0.884489| 0.03          |
Discussion

HPV is the most commonly sexually transmitted infection in the men and women worldwide (de Martel et al. 2017). HPV has an intra epithelial infection cycle; particularly infects mucosal and cutaneous layers (Bansal et al. 2016). Persistent infection with HPV leads to cervical cancer development (Tornesello and Buonaguro 2020).

As of now, two vaccines are approved by the FDA for the prevention of HPV associated cervical cancers (Fontecha et al. 2015). GlaxoSmithKline’s Cervarix® contains HPV 16, 18 virus like particles (VLPs), and Merck’s Gardasil® contains HPV 6, 12 VLPs along with HPV 16, 18 VLPs (Wang and Roden 2013), according to some reports these two prophylactic vaccines are highly efficient in preventing HPV infection (Schiller and Lowy 2018). However, none

| Sl.no | Protein region | Epitope | Start | End | Length | Percentile rank |
|-------|----------------|---------|-------|-----|--------|-----------------|
| 48    | HPV-18 E2     | QRTKFLNTVAIPDSV | 342   | 356 | 15     | 0.61            |
| 49    | HPV-18 E4     | TTRYPLLSLLNSYST | 9     | 23  | 15     | 0.16            |
|       |                | TRYPLLSLLNSYSTP | 10    | 24  | 15     | 0.17            |
|       |                | VTRYPLLSLLNSYSD | 8     | 22  | 15     | 0.17            |
|       |                | RYPLLSLLNSYSTTP | 11    | 25  | 15     | 0.21            |
|       |                | YPLLSSLNSYSTPPH | 12    | 26  | 15     | 0.23            |
| 54    | HPV-18 E5     | CMCAWVLVFYIVYV | 27    | 41  | 15     | 0.12            |
|       |                | MCAYAWVLVFYIVVV | 28    | 42  | 15     | 0.12            |
|       |                | YAWVLVFYIVVITTS | 31    | 45  | 15     | 0.23            |
|       |                | AWVLVFYIVVITSP  | 32    | 46  | 15     | 0.23            |
|       |                | WLVFYIVVITSPAT  | 33    | 47  | 15     | 0.23            |
|       |                | LVVFYIVVITSPAT  | 34    | 48  | 15     | 0.23            |
|       |                | LVVFYIVVITSPATA | 35    | 49  | 15     | 0.23            |
|       |                | WLVFYIVVITSPATA | 33    | 47  | 15     | 0.24            |
|       |                | LVVFYIVVITSPATA | 34    | 48  | 15     | 0.15            |
|       |                | LVVFYIVVITSPATA | 35    | 49  | 15     | 0.15            |
| 64    | HPV-18 E6     | NEKRRFHNIAGHYRG | 122   | 136 | 15     | 0.16            |
|       |                | EKRRFHNIAGHYRCH | 123   | 137 | 15     | 0.16            |
|       |                | KRRFHNIAGHYRQCH | 124   | 138 | 15     | 0.16            |
|       |                | RRFHNIAGHYRQCH  | 125   | 139 | 15     | 0.16            |
| 68    | HPV-18 E7     | LRAFAQQLNLNTLSFV | 83    | 97  | 15     | 0.28            |
| 69    | HPV-18 L1     | PTSIFYHAGSSRLT  | 91     | 105 | 15     | 0.07            |
|       |                | TSIFYHAGSSRLTTV | 92     | 106 | 15     | 0.07            |
|       |                | SIFYHAGSSRLTGV  | 93     | 107 | 15     | 0.07            |
|       |                | TPTSIFYHAGSSRL  | 90     | 104 | 15     | 0.09            |
|       |                | IFYHAGSSRLTGVN  | 94     | 108 | 15     | 0.09            |
|       |                | YPLGRKFLVQAGLR  | 524    | 538 | 15     | 0.13            |
|       |                | PLGRKFLVQAGLRK  | 525    | 539 | 15     | 0.13            |
|       |                | LGKFLVQAGLRKRP  | 526    | 540 | 15     | 0.13            |
|       |                | GRKFLVQAGLRKPT  | 527    | 541 | 15     | 0.13            |
|       |                | RKFLVQAGLRKPTI  | 528    | 542 | 15     | 0.13            |
| 79    | HPV-18 L2     | IHGTHYLYWPLYYFI | 430    | 444 | 15     | 0.08            |
|       |                | FAFKYSPTISSASS  | 369    | 383 | 15     | 0.12            |
|       |                | YLWPLYYFIPKRRK  | 436    | 450 | 15     | 0.14            |
|       |                | LWPLYYFIPKRRKR  | 437    | 451 | 15     | 0.13            |
|       |                | HGTHYLYWPLYYFIP | 431    | 445 | 15     | 0.15            |
|       |                | GTHYLYWPLYYFIPK | 432    | 446 | 15     | 0.16            |
|       |                | PLYYFIPKRRKRPY  | 439    | 453 | 15     | 0.19            |
|       |                | WLPYYFIPKRRKRPV | 438    | 452 | 15     | 0.19            |
|       |                | SFAFKYSPTISSAS  | 368    | 382 | 15     | 0.2              |
|       |                | WLPYYFIPKRRKRPV | 438    | 452 | 15     | 0.13            |
of these vaccines were effective against pre-existing HPV infection or cancers associated with HPV infection. Furthermore, HPV prevalence and mortality is still high in several developed and developing countries, given this rationale, so far, several therapeutic vaccines for HPV infection or cancer clearance have been developed (Namvar, Panahi et al. 2020a; b); (Namvar et al. 2020a, b) but, many of them induced an inadequate immune response. Given this limitation there

| Sl. No | Protein region | Epitope | Start position | Immuno-genicity Score |
|--------|----------------|---------|----------------|----------------------|
| 89     | HPV-18 E2      | QDKIIDHYENDSKDID | 16              | 0.93                 |
|        |                | TFIYEFKSECEKYGNT | 171             | 0.92                 |
|        |                | SSTWHWTGAGNEKTGI | 371             | 0.91                 |
|        |                | TGTWVEHFNGNVIDCN | 186             | 0.91                 |
|        |                | KGGQTVQFYFDGNKDN | 116             | 0.91                 |
|        |                | DWTQLDQCEELWNTEP | 95              | 0.90                 |
|        |                | TWDKQATCVSHRGGLY | 149             | 0.88                 |
|        |                | TVTYHSETQRTKFLNT | 334             | 0.87                 |
|        |                | YHSETQRTKFLNTVAIPDSV | 337  | 0.87                 |
|        |                | GQTSAATRPGHCGLA | 241             | 0.86                 |
| 99     | HPV-18 E4      | PWAPQRPTARRRLLHD | 33              | 0.93                 |
|        |                | LNSYSTPPIHRIPACP | 18              | 0.91                 |
|        |                | PPIHRIPAPCPWAPQRP | 24              | 0.90                 |
|        |                | AVPVTRYPPLLSSLNS | 5               | 0.87                 |
|        |                | LHLQATTKDGNSSVVT | 70              | 0.81                 |
| 104    | HPV-18 E5      | CFCVCMYVCCHVPLL | 9               | 0.86                 |
|        |                | PATAFTVYVFCFLPM | 46              | 0.83                 |
| 106    | HPV-18 E6      | HNIAHGYRGQCCHCCN | 128            | 0.89                 |
|        |                | VGGEDGEGI | 480             | 0.86                 |
|        |                | TRRPYKLPLCETELNT | 8               | 0.82                 |
|        |                | HKCIFDSRIELRHY | 66              | 0.82                 |
| 110    | HPV-18 E7      | KATLQDIVLHLLEPQNE | 5            | 0.82                 |
|        |                | SFVCPCWCAS | 95              | 0.94                 |
|        |                | IPTKENNTG | 953             | 0.823                |
| 113    | HPV-18 L1      | KPTIGPRKRSAPSATT | 539            | 0.93                 |
|        |                | CQSICKYPDYLQMSAD | 286            | 0.92                 |
|        |                | KGTASKRSRPLSQGDCPPLE | 171  | 0.91                 |
|        |                | KFLVQAGLRRKPTIGP | 529             | 0.90                 |
|        |                | GTACKSRPLSQGDCPP | 233             | 0.90                 |
|        |                | ASTQSPVPPQYDATKFK | 346            | 0.90                 |
|        |                | SSLEWDNNDPKYDKLKF | 397            | 0.90                 |
|        |                | AVVNTDDYYTRTSIFYHAGS | 20  | 0.87                 |
|        |                | AITQCDKAAPAENKD | 487             | 0.87                 |
| 122    | HPV-18 L2      | GTIYGARVHAYHDISP | 310             | 0.94                 |
|        |                | FEPTVDTTLTDFPRSDV | 256             | 0.92                 |
|        |                | DISPIAPSPEYIELQP | 322             | 0.91                 |
|        |                | SGTCPPDVVPKVEGTT | 24              | 0.91                 |
|        |                | EEPISSTPLPTVRRVA | 207             | 0.91                 |
|        |                | MVSHTAAARRKASVTD | 1               | 0.91                 |
|        |                | GRTGYIPLGRGNSNTVV | 67              | 0.89                 |
|        |                | TGSSTGGTGYIPLGG | 61              | 0.89                 |
|        |                | VWPIVSTAPASTQYI | 413             | 0.89                 |
|        |                | SPTISSASSYNSVTVP | 375             | 0.88                 |
is an urgent need to develop an effective therapeutic vaccine against HPV associated cervical cancers (Frazer and Chandra 2019). Additionally, for cervical cancer, no FDA-approved immunotherapy is present till date. Immunotherapy to treat cervical cancer or HPV infection can be achieved by development of DNA or peptide based vaccines against the viral genome proteins (Panahi et al. 2018). Peptide based vaccines depends on the recognition of highly immunogenic epitopes as they rely on usage of short peptide fragments to engineer the induction of highly targeted immune responses restricted to a specific MHC. There are several studies focused on the immune-therapeutic target for the HPV-16 (Kumar et al. 2015a, b, c; Kumar et al. 2015a, b, c; Kumar et al. 2015a, b, c; Yufeng Yao 2013), however few studies have been done on the 2nd most common HPV type (i.e., HPV-18) along with HPV-16 causing cervical cancer (Basit Jabbar et al. 2018).

The aim of the study was the prediction of T and B cell epitopes from E2, E4, E5, E6, E7, L1, and L2 oncoproteins of hr-HPV-18. E2, E4, E5, E6 and E7 are crucial for transcription activation, cell cycle deregulation/ transformation, tumor pathogenesis, and virus replication. L1 and L2 are responsible for recognition, attachment and entry into host cell, later they help in viral self-assembly and proliferation (Pinidis et al. 2016). Therefore, the early and late HPV regions are considered as the ideal targets for therapeutic vaccine generation.

Hr-HPV-18 is the second most carcinogenic and prevalent type HPV worldwide (Burni et al. 2019). However it does not show symptoms (Geng et al. 1999) like hr-HPV-16. Currently, people with hr-HPV-18 infections show normal cytology or precancerous lesions, hence, the infection remains undetectable and causes the progression of infection into cervical cancer. Hr-HPV-18 is known to be present in a higher proportion of cervical adenocarcinomas (ADC) ~37% than cervical squamous cell carcinomas (SCC) ~12% (Li et al. 2011). As of now the factors that favor the hr-HPV-18 pathogenies are poorly understood. Thus, in this study, hr-HPV-18 is targeted for the epitope prediction. hr-HPV-18 protein sequences were obtained from NCBI GenBank database, and the majorly expressed alleles for MHC-I and II within the Indian population were chosen for the study. There are various studies which showed the importance of amino acid residues for the functions of proteins (Kumar et al. 2015b). So, amino acid composition was also identified for hr-HPV-18, and Threonine (Thr/T) and Leucine (Leu/L) were the most frequent amino acids. The potent epitopes against antigenic proteins can be identified using bioinformatic software and can be used as immune therapeutic targets against HPV infection (De Groot et al. 2010). Adaptive immunity is accelerated by lymphocytes, precisely by B- and T-cells (Sanchez-Trincado et al. 2017).

B- and T-cells particularly recognizes the molecular components known as antigens. A specific region of the antigen known as epitope is recognized by the immune cells to elicit the immunogenic response (Jespersen et al. 2019). Potent peptides against the antigenic epitopes of hr-HPV-16 E5 and E6 have also been identified (Kumar, Singh, et al. 2015a; b, c) (Kumar et al. 2015a), and shown an promising role in terms of vaccine development. In this study,
IEDB software was used to predict the potential immunogenic epitopes for MHC I and II alleles for hr-HPV-18 oncoproteins. We have screened 143 potent epitopes for MHC I, 222 epitopes for MHC II on the basis of percentile rank and immunogenicity score. In future these identified epitopes further can be studied in vitro and in vivo for the development of effective immunotherapeutic targets. The epitopes having a high immunogenicity score (> 0.9); STVSYG-TAK, DSVQIVGTY, TPSYSSVTV, SPYSSTVSV of E2, SYSTPPHR of E4, MLLLIHIHAL of E5, SYVDTLK, SYVGDLEK of E6, ATLQDIVLH of E7, LLPSSVARV, IYNPETQRL, DVMSIYHSM, FPIFQMAL of L1, and CDPDVVPKV, STTSAFFK, APSPEYIEL of L2 against MHC-I; TTRYPLSSLNSYST, TTRYPLSSLNSYSTP, VTTTRYPLSSLNSYS of E4, CMCAAYAWLVFVYIV, MCAYAWLVFVYIVV, YAWLVFVYIVVITS, AWVL-VFVYIVVITS of E5, NEKRRFHIAGHYRG, EKR-RFHIAGHYRGQ of E6, LRAFQQLFLNTLSF of E7, TSIFYHAGSSRLLL, TSIFYHAGSSRLLTV, SIFY-HAGSSRLTVG, TPTSIFYHAGSSRL of L1, and IHGTYYLWPLYFYI, FAFKYSTISSASS of L2 against MHC-II. The potent B cell epitope predicted were YHSETQRSTKFLTNAIPSV of E2, VGGEDEGI of E6, SFVCPWCAS, IPTKENNTG of E7, and KGTASKS-PLSQGDCPLE, ASTSPVPQYQDATKFK, of L1 can be considered as potential candidates for therapeutic vaccines development. In-silico analysis or experimental evidence are someway limited to the identification of epitopes against oncoproteins E6 and E7 only. However, among our MHC-I predicted epitopes MLLLIHIHAL of E5, and ATLQDIVLH of E7 were also reported in a comprehensive in silico study of hr-HPV-18 (Panahi et al. 2018). In another similar study LLPSSVARV epitope of L1, and CDPDVVPKV epitope against L2 were predicted and analyzed for development of cross-subtype prophylactic vaccine development (Namvar et al. 2019). Moreover, predictive analysis of the thorough biological information characteristics of the T-cell and B-cell dominant epitopes will give rise to the further wet laboratory experiments and development of efficient HPV therapeutic vaccines against hr-HPV-18.

**Conclusion**

In the present study, this is first time that in a laborious in silico study epitope prediction of E2, E4, E5, E6, E7, L1, and L2 proteins of 2nd most common type was targeted i.e., hr-HPV-18 have been investigated altogether. For the development of an effective vaccine both prophylactic and therapeutic vaccine candidates need to be identified. Considering the role of these protein in viral replication, maintenance, oncogenicity, and virus assembly, HPV-18 proteins are good candidates for antigenicity and immunogenicity. In the present study, we have targeted the late region of the HPV-18 (L1 & L2) responsible for the viral entry and structure also an ideal target for the prophylactic vaccine, and Early protein E6, E7 main oncoproteins for the therapeutic target. These predicted targets need to be further validated by in-vitro & in vivo study will be warranted.

**Acknowledgements** PS and BG are thankful to all faculty members of the Department of Biotechnology, Indira Gandhi National Tribal University, Amarkantak, M.P., India for their constant support. AK thankful to SERB DST for young scientist project YSS/ YSS/2015/002074.

**Author Contributions** BG contributed to the conceptualization, data curation, investigation, methodology, resources, validation, visualization, writing of manuscript, AK contributed to the conceptualization, resources, supervision, planning, review and editing of the manuscript. PS contributed to the conceptualization, project administration, resources, supervision, review and editing of the original draft.

**Declarations**

Conflict of interest The authors declare no potential conflict of interest.

**References**

Bahmani B, Amini-bayat Z, Ranjbar MM, Bakhtiari N, Zarnani AH (2020) HPV16-E7 protein T cell epitope prediction and global therapeutic peptide vaccine design based on human leukocyte antigen frequency: an in-silico study. Inter J Peptide Res Ther. https://doi.org/10.1007/s10989-020-10089-5

Bansal A, Singh M, Rai B (2016) Human papillomavirus-associated cancers: a growing global problem. Inter J Appl Basic Med Res 6(2):84. https://doi.org/10.4103/2229-516X.179027

Bonnimann MP, Calton CM, Chiquette SF, Li S, Lu M, Chapman JA, Bratton KN, Schlegel AM, Campos SK (2016) Furin cleavage of L2 during papillomavirus infection: minimal dependence on cyclophilins. J Virol 90(14):6224–6234. https://doi.org/10.1128/JVI.00038-16

Bruni L, Albero G, Serrano B, Mena M, Gómez D, Muñoz J, Bosch F, de Sanjose S, & ICO HPV Information Centre. (2019). Human papillomavirus and related diseases report. ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre), 307. https://hpvcentre.net/statistics/reports/XWX.pdf

Bruni L (2019) Human papillomavirus and related diseases report. HPV india

Castro-Muñoz LJ, Manzo-Merino J, Muñoz-Bello JO, Olmedo-Nieva L, Cedro-Tanda A, Alfaro-Ruiz LA, Hidalgo-Miranda A, Madrid-Marina V, Lizano M (2019) The Human Papillomavirus (HPV) E1 protein regulates the expression of cellular genes involved in immune response. Sci Reports 9(1):1–13. https://doi.org/10.1038/s41598-019-49886-4

De Groot AS, Cohen T, Moise L, Ardito M, Martin W (2010) A method for individualizing the prediction of immunogenicity of protein vaccines and biologic therapeutics: individualized T cell epitope measure (iTEm). J Biomed Biotechnol. https://doi.org/10.1155/2010/961752

De Martel C, Plummer M, Vignat J, Franceschi S (2017) Worldwide burden of cancer attributable to HPV by site, country and HPV type. Inter J Cancer 141(4):664–670. https://doi.org/10.1002/ijc.30716
Fang L, Lin X, Yang Y, Song Z, Ding X, Tan L, Gao P (2020) Hallmarks of HPV carcinogenesis: the role of E6, E7 and E5 oncoproteins in cellular malignancy. Biochim Biophys Acta 1862(2):153–162. https://doi.org/10.1016/j.bbamcr.2019.01.001

Fang L, Lin X, Yang Y, Song Z, Ding X, Tan L, Gao P (2020) Genetic variability, phylogeny and functional implication of the long control region in human papillomavirus type 16, 18 and 58 in Chengdu China. Virol J. https://doi.org/10.1186/s12985-020-01349-3

Fontecha N, Basaras M, Arrese E, Hernández S, Andía D, Cisterna R (2015) Human papillomavirus 16 variants may be identified by E6 gene analysis. Intervirology 58(3):143–148. https://doi.org/10.1159/0003901745

Frazer IH (2009) Interaction of human papillomaviruses with the host immune system: a well evolved relationship. Virology 384(2):410–414. https://doi.org/10.1016/j.virol.2008.10.004

Frazer IH, Chandra J (2019) Immunotherapy for HPV associated cancer. Papillomavirus Res 8:100176. https://doi.org/10.1016/j.pvr.2019.100176

Geng L, Connolly DC, Isacson C, Ronnett BM, Cho KR (1999) Atypical immature metaplasia (AIM) of the cervix: it is related to high-grade squamous intraepithelial lesion (HSIL). Human Pathol 30(3):345–351. https://doi.org/10.1016/S0046-8177(99)90015-1

Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, Zahurak ML, Daniel RW, Vignole M, Symer DE, Shah KV, Sidransky D (2000) Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst 92(9):709–720. https://doi.org/10.1093/jnci/92.9.709

Hupf CF, Wu TC, Monie A, Roden R (2008) Antigen-specific immunotherapy of cervical and ovarian cancer. Immunol Rev 222(1):43–69. https://doi.org/10.1111/j.1600-065X.2008.00622.x

Jespersen MC, Mahajan B, Peters B, Nielsen M, Marcatili P (2019) Antibody specific B-cell epitope predictions: Leveraging information from antibody-antigen protein complexes. Front Immunol. https://doi.org/10.3389/fimmu.2019.00298

Kaarttisgyan K (2012) Cervical cancer in India and HPV vaccination. Indian J Med Paediatr Oncol 33(1):7–12. https://doi.org/10.4103/0971-5851.96961

Kumar A, Hussain S, Yadav IS, Gissmann L, Natarajan K, Das BC, Bharadwaj M (2015a) Identification of human papillomavirus-16 E6 variation in cervical cancer and their impact on T and B cell epitopes. J Virol Methods. https://doi.org/10.1016/j.jviromet.2015.03.008

Kumar A, Hussain S, Yadav IS, Gissmann L, Natarajan K, Das BC, Bharadwaj M (2015b) Identification of human papillomavirus-16 E6 variation in cervical cancer and their impact on T and B cell epitopes. J Virol Methods 218:51–58. https://doi.org/10.1016/j.jviromet.2015.03.008

Kumar A, Singh I, Hussain S, Das BC, Bharadwaj M (2015c) Biologicals Identification of immunotherapeutic epitope of E5 protein of human papillomavirus-16: an in silico approach. Biologicals. https://doi.org/10.1016/j.biologicals.2015.07.002

LaVigne K, Leitao MM (2019) Cervical cancer prevention. Fundam Cancer Prev. https://doi.org/10.1007/978-3-030-15935-1_18

Li N, Franceschi S, Howell-Jones R, Snijders PJF, Clifford GM (2011) Human papillomavirus type distribution in 30848 invasive cervical cancers worldwide: variation by geographical region, histological type and year of publication. Inter J Cancer 128(4):927–935. https://doi.org/10.1002/ijc.25396

Li W, Joshi MD, Singhania S, Ramsey KH, Murthy AK (2014) Peptide vaccine: progress and challenges. Vaccines 2(3):515–536. https://doi.org/10.3390/vaccines2030515

Müller M, Prescott EL, WASSON CW, MacDONALD A (2015) Human papillomavirus E5 oncoprotein: function and potential target for antiviral therapeutics. Future Virol 10(1):27–39. https://doi.org/10.2217/fvl.14.99

Nakagawa M, Kim KH, Moscicki AB (2004) Different methods of identifying new antigenic epitopes of human papillomavirus type 16 E6 and E7 proteins. Clin Diag Lab Immunol 11(5):889–896. https://doi.org/10.1128/CDLI.11.5.889-896.2004

Namvar A, Bolhassani A, Javadi G, Noormohammadi Z (2019) In silico/In vivo analysis of high-risk papillomavirus L1 and L2 conserved sequences for development of cross-subtype prophylactic vaccine. Sci Reports 9(1):1–22. https://doi.org/10.1038/s41598-019-51679-8

Namvar A, Bolhassani A, Javadi G, Noormohammadi Z, Baidya S, Das R, Kabir G, Panahi HA, Bolhassani A, Javadi G, Noormohammadi Z, He J, Yang Y, Chen ZZ, Liu Y, Bao S, Zhao Y, Ding X, Dong D, Bharadwaj M (2020a) Bioinformatics analysis of HPV-68 E6 and E7 oncoproteins for designing a therapeutic epitope vaccine against HPV infection. Virol J 13(1):104266. https://doi.org/10.1186/s12985-019-1168-y

Namvar A, Panahi HA, Agi E, Bolhassani A (2020b) Development of HPV16,18,31,45 E5 and E7 peptides-based vaccines predicted by immunoinformatics tools. Biotechnol Lett 42(3):403–418. https://doi.org/10.1007/s10529-020-20792-6

Oliveira JG, Cölf LA, McBride AA (2006) Variations in the association of papillomavirus E2 proteins with mitotic chromosomes. Proc Natl Acad Sci USA 103(4):1047–1052. https://doi.org/10.1073/pnas.0507624103

Pal A, Kundu R (2020) Human papillomavirus E6 and E7: the cervical cancer hallmarks and targets for therapy. Front Microbiol. https://doi.org/10.3389/fmicb.2019.03116

Panahi HA, Bolhassani A, Javadi G, Noormohammadi Z (2018) A comprehensive in silico analysis for identification of therapeutic epitopes in HPV16, 18, 31 and 45 oncoproteins. PLoS ONE 13(10):1–25. https://doi.org/10.1371/journal.pone.0205933

Pindis P, Tsikouras P, Iatrakis G, Zervoudis S, Kroukouli Z, Bothou R, Galazios G, Vladelarena S, Hospital R (2016) Human papillomavirus ‘life cycle and carcinogenesis. Maeillic 11(1):48

Posner J, Peterson BS (2013) Bone 23(1):1–7. https://doi.org/10.1016/j.virol.2015.03.038

Redley D, Saah A, Stanley M (2016) Persistent infection with human papillomavirus 16 or 18 is strongly linked with high-grade cervical disease. Human Vaccines Immunotherap 12(3):768–772. https://doi.org/10.2217/vin.15.1088616

Ribeiro-Muller L, Seitz H, Muller M (2013) Human Papillomavirus Prophylactic Vaccines and Alternative Strategies for Prevention. Human Papillomavirus and Related Diseases from Bench to Bedside A Diagnostic and Preventive Perspective. https://doi.org/10.5772/55852

Richards RM, Lowy DR, Schiller JT, Day PM (2006) Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection. Proc Natl Acad Sci USA 103(5):1522–1527. https://doi.org/10.1073/pnas.0508815103

Sanchez-Trincado JL, Gomez-Peorsanz M, Reche PA (2017) Fundamentals and methods for T- and B-cell epitope prediction. J Immunol Res. https://doi.org/10.1155/2017/2680160

Schiffman M, Doorbar J, Wentzensen N, De Sanjosé S, Fakhry C, Monk BJ, Stanley MA, Franceschi S (2016) Carniogenic human papillomavirus infection. Nature Rev Dis Prim. https://doi.org/10.1038/nrdp.2016.86
Schiller J, Lowy D (2018) Explanations for the high potency of HPV prophylactic vaccines. Vaccine 36(32):4768–4773. https://doi.org/10.1016/j.vaccine.2017.12.079

Schiller JT, Castellsagué X, Garland SM (2012) A review of clinical trials of human papillomavirus prophylactic vaccines. Vaccine 30(5):123–138. https://doi.org/10.1016/j.vaccine.2012.04.108

Tornesello ML, Buonaguro FM (2020) Human papillomavirus and cancers. Cancers 12(12):1–5. https://doi.org/10.3390/cancers12123772

van der Burg SH, Bijker MS, Welte V, Offringa R, Melief CJM (2006) Improved peptide vaccine strategies, creating synthetic artificial infections to maximize immune efficacy. Adv Drug Deliv Rev 58(8):916–930. https://doi.org/10.1016/j.addr.2005.11.003

Wang JW, Roden RBS (2013) L2, the minor capsid protein of papillomavirus. Virology 445(1–2):175–186. https://doi.org/10.1016/j.virol.2013.04.017

Xi J, Chen J, Xu M, Yang H, Luo J, Pan Y, Wang X, Qiu L, Yang J, Sun Q (2017) Genetic variability and functional implication of the long control region in HPV-16 variants in Southwest China. PLoS ONE 12(8):1–11. https://doi.org/10.1371/journal.pone.0182388

Yajid AI, Zakariah MA, Zin AAM, Othman NH (2017) Potential role of E4 protein in human papillomavirus screening: a review. Asian Pacific J Cancer Prev 18(2):315–319

Zheng Z, Baker CC (2006) Papillomavirus genome structure, expression, and post-transcriptional regulation. Front Biosci 11:2286–2302

Zur Hausen H (2002) Papillomaviruses and cancer: from basic studies to clinical application. Nature Rev Cancer 2(5):342–350. https://doi.org/10.1038/nrc798