

### Introduction

#### 1.1 Human Papilloma Virus

Human papillomavirus (HPV) are a group of viruses which are associated with various proliferative diseases in the infected epithelium (Pfister, 1984). There are more than 70 HPV types have been reported. Some of these, among which HPV types 6, 11, 16, and 18 are the most common, are associated with lesions in the anogenital tract (Pfister, 1987). HPV belong to the papovaviridae family. They consist of 72-capsomere capsid containing the viral genome. Capsomers are composed of two structural proteins: the 57 KD late proteins L1, which accounts for 80% of the viral particle, and the 43-53 KD minor capsid protein L2. Several clinical, molecular, and epidemiological investigations have identified human papillomavirus (HPV) as the major cause of cervical cancer and cervical dysplasia (Walboomers et al., 1999).

Human papillomavirus (HPV) causes cervical cancer, the second biggest cause of female cancer mortality worldwide. Estimates of the number of cervical cancer deaths are around 250,000 per year. The prevalence of genital HPV infection in the world is around 440 million. There are over 100 genotypes of HPV, 40 of which infect human mucosal areas of the upper digestive tract and the ano-genital tract. These are grouped into "high-risk" and "low-risk" types according to the degree of risk of development of cancer after infection with each genotype. Genital HPV infection is extremely common and most often causes no symptoms. A proportion of individuals infected with low-risk HPV types such as HPV-6 or HPV-11 will develop genital warts, whereas a subset of women with high-risk HPVs such as HPV-16 or HPV-18 will develop preneoplastic lesions of cervical intraepithelial neoplasia (CIN). Low-grade cervical dysplasias are common and most regress spontaneously. In contrast, the minority of lesions that progress to high-grade dysplasias tend to persist and/or progress to carcinomas in situ before becoming invasive cancers. The majority of adenocarcinomas of the cervix and of squamous cell cancers (SCC) of the vulva, vagina, penis and anus are caused by HPV-16 and HPV-18 (together accounting for about 70% of cases globally), the remaining 30% being due to other high-risk HPV types (such as HPV-31, -33, -35, -39, -45,-51, -66). The relative importance of
different high-risk types varies between countries and regions, but type 16 has the
greatest contribution to cervical cancer in all regions. HPV is also associated with other
cancers of the anus, head and neck, and rarely, recurrent respiratory papillomatosis in
children (Dayyani et al., 2010).

1.2 HPV infection and its detection

HPV infection is limited to the epithelium so when the HPV Virus particles have
entered host cells, infection is dealt with cell-mediated immunity, not by antibody. At
the same time cytotoxic T-lymphocytes (CTLs) recognize foreign peptide antigens
presented on the infected target cell surface by molecules of major histocompatibility
complex (MHC) where, In the presence of co-stimulatory molecules binding to a CTL
may induce an immune response. Infiltrating T-cells are seen in regressing warts
caused by low-risk HPVs. Cervical cancer show tumour infiltrating lymphocytes that
are predominantly CTLs (Konya and Dillner, 2001). HPV infections can be detected
by testing a sample of cells to see if they contain viral DNA or RNA. The most
common test detects DNA from several high-risk HPV types, but it cannot identify the
types that are present. Another test is specific for DNA from HPV types 16 and 18; the
two types that cause most HPV associated cancers. A third test can detect DNA from
several high-risk HPV types and can indicate whether HPV-16 or HPV-18 is present.
A fourth test detects RNA from the most common high-risk HPV types. These tests
can detect HPV infections before cell abnormalities are evident. Theoretically, the
HPV DNA and RNA tests could be used to identify HPV infections in cells taken from
any part of the body. However, the tests are approved by the FDA for only two
indications: for follow-up testing of women who seem to have abnormal Pap test
results and for cervical cancer screening in combination with a pap test among women
over age 30 (Hagerstown, 2012).

1.3 Cervical cancer

Cervical cancer as the most prevalent cancer in women in the coming years is the main
cause of death, especially in young women. Cervical cancer can also be defined as the
cancer originating at the opening of the womb which progressively migrates to the
whole of cervix (Muller-Schiffmanet al., 2006).
Every year in the UK, over 3,000 women will be diagnosed with cervical cancer and nearly 1,000 women will die from the disease. It is the most common cancer in women aged 35 and under. Cervical cancer is not thought to be hereditary (Bosch et al., 2002) but in 99.7% of cases, is caused by persistent infection with a virus called human papillomavirus (HPV). Some available forms of treatment such as surgery, radiation therapy and chemotherapy are all cytoreductive treatment modalities, so, in addition to killing cancerous cells, healthy cells are also destroyed in the process. Indeed, there is a need to decrease the incidence of cervical cancer and develop better forms for its treatment (Burd-Eileen, 2003).

1.4 HPV Vaccines
There are vaccines to protect against HPV and new bioinformatics tools have been developed that can help for designing of more vaccines. The development of several immunoinformatics and computational biology tools are useful for identification of antigenic regions (epitopes) in the protein sequences which can accelerate the wet laboratory practices. These tools have been developed on the basis of existing and validated data with specific algorithms (Petrovsky et al., 2003). So, these are also helpful in increasing the demand for diagnostic for the detection of HPV and also requires the suitable and potent vaccines for control of infection. There are two vaccines such as gardasil and cervarix against both highly pathogenic HPV types 16 and 18 which give protection for 70% of cervical cancers, 80% of anal cancers, 60% of vaginal cancers, and 40% of vulvar cancers (Vuyst and Clifford, 2009).

The FDA reports that HPV vaccines, gardasil (approved in 2006) and cervarix (approved in 2009), are safe for females ages 9 to 26 years. As of 2009, gardasil is also licensed, and considered safe for males ages 9 through 26 years. Boys and young men may choose to get this vaccine to prevent anal cancer and genital warts. Both vaccines were tested in thousands of people around the world before they were approved. These studies showed no serious side effects and no deaths have been linked to either vaccine. Common, mild side effects include pain where the shot was given, fever, dizziness, and nausea. People may faint after getting any vaccine, including HPV vaccines. Fainting after getting a shot is more common in teens than in young children or adults. To keep people from getting hurt from fainting, a 15-minute waiting period for people of all ages is recommended after any vaccination. Both HPV vaccines are
being monitored for side effects, especially rare ones not seen in the study trials. CDC and FDA doctors and scientists still review all reports of serious side effects reported to the Vaccine Adverse Event Reporting System (VAERS) to watch for potential new vaccine safety concerns that may need further study (The VAERS is a national reporting system that looks at reports of side effects after vaccinations.) The American Cancer Society will watch those reviews and report any concerns about the safety of the vaccines (Saslow and Lawson, 2012).

Immune-based therapies for cancer such as therapeutic vaccines may be an attractive alternative over radiation and chemotherapy for several reasons: radiation and chemotherapy result in serious adverse effects and in many cases are not effective against large tumors or disseminated (metastatic) disease. On the other hand, a potential problem with immunotherapy is finding a good antigen. Since expression of the viral E6 and E7 proteins is required to maintain oncogenic phenotype and because normal cells do not express E6/E7, a therapeutic vaccine targeting these proteins has several advantages: (i) the tumor cell cannot lose their expression as an immune evasion mechanism; (ii) there is no immune tolerance generated against them and an effective immuneresponse in not likely to generated autoimmunity that could compromise healthy cells. These benefits contrast with other vaccines that target tumor-associated antigens (TAA) that are also present in normal cells (e.g., melanoma). Despite of this, immunotherapy for virus-induced cancer has not been successful in the clinic (Goldstone and Neefe, 2002; Hung, 2008).

The two proteins which play a significant role in onset of this malignancy are E6 and E7 proteins and their expression has been seen in the cell lines (Herdman et al., 2006). HPV 16 E7 proteins are small proteins consisting of 98 amino acids and zinc binding phosphoproteins are confined in the nucleus. This high risk HPV E7 protein has a higher binding affinity for pRb protein (Darnell and Schroder, 2007). The HPV 16 is a small non enveloped virus containing double stranded DNA as its genetic material. Its genome has seven functional coding regions in which E7 codes for viral protein that bind to the retinoblastoma tumour suppressor proteins thereby permitting the cell to progress through the cell cycle in the absence of normal mitogenic signals (Conway and Christensen, 2009).
The transmembrane protein 50A is a protein that in humans is encoded by the TMEM50A gene. It is also known as cervical cancer oncogene 9 or Small membrane protein 1. This gene is located in the RH gene locus, between the RHD and RHCE genes. The function of its protein product is unknown; however, its sequence has potential transmembrane domains suggesting that it may be an integral membrane protein. Its position between the RH genes suggests that polymorphisms in this gene may be tightly linked to RH haplotypes and may contribute to selective pressure for or against certain RH haplotypes (Wagner and Flegel, 2000). The commonly used physicochemical properties for this protein included isoelectric point, hydrophobicity and mass weight. So, the calculation of these physicochemical properties is very important and could be used for the design of proteomic experiments. Further analyses of physicochemical properties for the identified proteins might be helpful to discover the experimental bias, which would be important for improving the experiments. The analyses could also be used to discover the different distribution caused by some biological factors, for example the different isoelectric point and hydrophobicity in different sub-cellular localization (Zhang et al., 2006).

An epitopes are also known as antigenic determinant in the protein sequences which is recognized by the major histocompatibility complex (MHC) molecules. These epitopes are short peptides between 8 and 11 amino acids in lengths. The advances immunoinformatics tools enable the systematic identification of potential antigen from pathogenic microorganisms that is possible to develop a safe and effective vaccine against any infectious disease (Pizzat et al., 2000; Rappuoli, 2001). These tools assist the designing of subunit vaccines that can start from prediction of antigenic epitope in protein sequences (De Groot and Rappuoli, 2004). The culture of HPV 16 E7 also requires the cell line for growing and all these steps are time consuming, laborious and cost effective (Somvanshi et al., 2008). Therefore, antigen preparation of HPV 16 E7 is not an easy process thus; immunoinformatics tools were helpful for identification of antigenic site from complete protein sequences which are available in databank. The epitope-based vaccines are a new generation of vaccines which are very well tolerated and have fewer side effects than the conventional vaccines.
1.5 Justification of the work

The goal of an HPV vaccine would be to reduce the prevalence of infection and hence the risk of cervical abnormalities. However, questions arise as to how this would interact with screening, which reduces the progress of cervical abnormalities to serious disease. Furthermore, will a vaccine against one genotype influence the other types within a population and will the patterns of infection and disease remain the same if the vaccine alters the timing and type of HPVs experienced within a population? The available forms of treatment for cervical cancer such as surgery, radiation therapy and chemotherapy are all cytoreductive treatment modalities so in additions to killing cancerous cells, healthy cells are also destroyed in the process. Indeed, there is a need to decrease the incidence of cervical cancer and develop better forms for its treatment. So the main aim of this dissertation work was to design a peptide based vaccine against cervical cancer caused by human papilloma virus Type 16 E7 using bioinformatics approaches to generate novel vaccine candidates. The availability of complete genome sequence of HPV 16, has paved a new way for the study of various proteins expressed in humans with cervical cancer. Through the complete genome analysis of the antigen protein, it is possible to compile a list of their proteome analysis.

The three dimensional structure of both proteins HPV 16 E7 and Human TMEM 50A was not yet available in protein data bank; hence the present work of predicting the 3D model of both proteins were undertaken. Subsequently, modeled 3D-structure of both proteins can be exploited for generation of novel vaccine candidates.
The objectives of the present study were as follows:

1- To analyse the proteins expressed due to caused of cervical cancer in humans.

2- To create a modular structure of a protein expressed in humans susceptible to cervical cancer, using bioinformatics modelling tools.

3- To design a vaccine to prevent cervical cancer in humans in whom this protein is expressed.