CHAPTER I

INTRODUCTION

1. Rational and background

Cervical cancer, the second most common cancer in women worldwide is a leading cause of cancer death among women with overall 5-year survival rate of 40%. It is estimated that approximately 80% of incidence and mortality rate of the cervical cancer occurs in developing countries.

Human papillomavirus (HPV) infection is one of the most common sexually transmitted diseases. Infections with the oncogenic high risk (HR) types of HPVs (approximately 40 types) are implicated as a primary aetiological factor for development to cervical cancer because more than 99.7 % of cervical cancers contain HR-HPV DNA [1, 2]. The most significant oncogenic HPV type in cervical cancer is HPV 16, which has highest prevalence rate about 50-70% in women diagnosed with cervical cancer [3, 4]. Cofactors associated with cervical cancer are: i) Environmental or exogenous cofactors such as hormonal contraceptives, tobacco smoking. ii) Viral cofactors such as infections with HR types, HPV variants, viral load and viral intregration. iii) Host cofactors such as endogenous hormones, genetic factors and factors related to immune response [5].

The genome of HPV is a double-stranded DNA molecule of about 8,000 bp. Three genomic regions have been identified: a late region (L), an early region (E), and a long control region (LCR) or Upstream Regulatory Region (URR). During infection, HPV genomes are found as episomes in the nucleus of infected cells of the normal cervix, where infective viral particles can be isolated. However, in some low-grade and in most of the high-grade lesions of the uterine cervix, including cancer, HPV genomes are found to be integrated into the host genome [6]. A disruption of the E1–E2 region is required for HPV genome integration and results in an increased expression and stabilization of the E6 and E7 transcripts [7]. These gene products (E6 and E7) are essential in the process of HPV induced cellular immortalization and transformation [8]. The E6 and E7 oncoproteins encoded by early genes and their

transcriptional level are regulated by the p97 promoter located in the E6 proximal part of the LCR. LCR contains several specific protein binding sites for both viral and host proteins that regulated for positive and negative acting transcription factor. Nucleotide sequence variation within LCR may result in the risk for progression to cervical cancer because it can alter or increase transcriptional activity of E6 and E7 viral oncogenes. Positive acting transcription factors are nuclear factor-1 (NF-1), activating protein-1 (AP-1), octamer-binding factor-1 (Oct-1), transcription enhancing factor-1 (TEF-1), *trans*-acting transcription factor-1 (SP-1) and glucocorticoid response elements. Negative acting transcription factors are transcriptional factor Yin Yang-1 (YY-1) and CCAAT displacement protein-1 (CDP)/Cut that overlaps the binding site of the E1 replication proteins [9-11]. Nucleotide sequence variation within the part of enhancer regions, which contains several specific protein binding sites such as transcriptional regulator proteins YY-1, AP-1 and NF-1 binding sites are found associated to enhance the expression of E6 and E7 viral oncogenes [10-12]. In the studies of HPV 16 intratype sequence variation, HPV 16 variant can be divided into five major groups by geographic difference such as European (E), Asian (As), Asian American (AA), African1 (Af1), African2 (Af2) and a few minor branches such as North American1 (NA1) [13]. These variants differ in their geographical distribution and probably differ in their oncogenic potential. The AA variant, for example, has been associated with more aggressive invasive cancer and a tendency to occur in younger women [14]. Studies of the HPV 16 variant in geographical distribution based on E6 hybridization patterns in 408 invasive cervical specimens by Yamada *et al.* [4] suggested that 26% of Southeast Asia specimens are As variants and it is rare or absent in other continents. Some studies in China have shown the high prevalence of HPV 16 As variant among other variants and its contribution to the high incidence of cervical cancer in China [13, 15].

In Khon Kaen population, As variant was found in 73.9% of HPV 16 variants and associated with cervical carcinoma [16].

There is no data of the transcriptional activity of LCR of the HPV 16 As variant which is the most common variant of HPV 16 in Khon Kaen women, Thailand. Therefore, this research would like to study prevalence of the HPV 16 As variant in Thai women, determine nucleic acid sequence variation in LCR of the HPV 16 As

variant and evaluate the association between LCR transcriptional activity of the HPV 16 As variant and cervical cancer.

The prevalence of HPV 16 As variant in Thai women was investigated in formalin-fixed, paraffin-embedded cervical tissues derived from 4 regions of Thailand and histologically diagnosed as normal, LSIL, HSIL and SCC. HPV DNA was determined and genotyped for HPV 16 by polymerase chain reaction (PCR) using primers GP5+/GP6+ and reversed line blot hybridization (RLBH), respectively. E6 gene was amplified, sequenced and compared to reference for detection of variant. The association between HPV 16 As variant and cervical cell diseases in Thai women was analyzed. Furthermore, to confirm the result of HPV 16 As variant LCR sequence variation from formalin-fixed, paraffin-embedded cervical tissues, LCR sequence variation from fresh cervical tissue samples was also analyzed and determined for LCR variation of HPV 16 As variant. Some of frequently found HPV 16 As LCR variation were selected for analysis of their transcriptional activities by cloning in a promoterless plasmid containing a reporter gene (luciferase). The constructed vector was transfected in C33A cell line and observed luciferase activity. The transcriptional activity of HPV 16 As variant LCR was compared with HPV 16 prototype to determine the risk factors of cervical cancer. This may provide information for HPV vaccine design strategies and cervical cancer control.

2. Objectives

2.1 To study the HPV prevalence and HPV genotype distribution in FFPE samples with cervical grade lesion from women in 4 region of Thailand.

2.2 To study the prevalence of HPV 16 variants and As variant and their association with cervical grade lesions in Thai women.

2.3 To analyze nucleic acid sequence variation in LCR of HPV 16 As variant.

2.4 To evaluate the transcriptional activity of HPV 16 As variant LCR.

3. Scope and limitation of research

3.1 DNA extraction from FFPE samples from women histologically diagnosed as Normal, LSIL, HSIL, and SCC.

3.2 HPV DNA detection by PCR using GP5+/GP6+ consensus primers.

3.3 Determination of HPV 16 in HPV DNA positive samples by 37 type-specific oligonucleotide probes using RLBH assay

3.4 HPV 16 E6 gene and LCR amplification by PCR using specific primer sets for nucleotide sequencing

3.5 Classification of HPV 16 variants using E6 ORF at nt. 83-559.

3.6 Analysis of the nucleotide sequence variation on HPV 16 E6 gene and LCR regions using the bioinformatic genomic tool Multalin sequence.

3.7 Evaluation of the LCR transcriptional activity of HPV 16 As variant and prototype using Luciferase assay.

3.8 Evaluation of the association between HPV 16 As variant LCR transcriptional activity and its oncogenic potential

4. Anticipate outcome

4.1 The prevalence of HPV infection and HPV genotypes in FFPE samples with various cervical grade lesions from women in 4 regions of Thailand.

4.2 The prevalence of HPV 16 variants in Thai women with different grade cervical lesions.

4.3 The nucleotide sequence variation on LCR of HPV 16 As in Thai women with different grade cervical lesions.

4.4 The LCR transcriptional activity of HPV 16 As variant and HPV 16 E prototype.