JKMU

Journal of
Kerman University of Medical Sciences

https://jkmu.kmu.ac.ir
10.34172/jkmu.2024.04
Vol. 31, No. 1, 2024, 23-28

Original Article

Characterization and Structural Analysis of the Human Papilloma Virus L1 Protein in Iran

Zahra Hasanshahi[1], Behzad Dehghani[1], Ava Hashempour[1*], Elnaz Alamdari[1]

[1]HIV/AIDS Research Center, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran

***Corresponding Author:** Ava Hashempour, Email: thashem@sums.ac.ir

Abstract

Background: Human papillomavirus (HPV) is a small, non-enveloped DNA virus related to human cervical cancer. The genome is maintained within the basal epithelium where the primary infection is latent. During the late phase of infection, the capsid proteins (L1 and L2) are expressed to encapsidate the viral genome, generating the infectious virion particles required for HPV propagation. HPV genome encodes six proteins, namely E6 oncoprotein, E7 oncoprotein, E1 replication protein, E2 regulatory protein, L1 major capsid protein, and L2 minor capsid protein. L1 is the principal part of the current vaccines, and any changes in this region can decrease vaccine efficiency. The aim of this research was to conduct a comparative analysis among Iranian L1 protein sequences with reference sequences to determine the possible substation in this region and to find the physicochemical and structural properties of L1 by using bioinformatics tools to provide comprehensive comprehension of the HPV L1 protein.

Methods: Thirteen Iranian PV sequences of the L1 protein and reference sequences were selected and obtained from the NCBI data bank. CLC Sequence Viewer software was used to translate the alignment. PrediSi and Phobius were employed to predict the signal peptide. The secondary and tertiary structures and structure validations of all sequences were analyzed by Qmean, (PS)2-v2, Phyre2server, Discovery Studio, and I-TASSER.

Results: The findings showed that L1 is highly conserved, and only two mutations were found in this region. No signal peptide was described, and this region's main part included a random coil. The tertiary structure was mapped using different software, and five distinct loops were found.

Conclusion: This study is the first report that investigated the changes in the L1 protein of Iranian patients and provided helpful comprehension of the L1 properties vital for cloning and producing the new generation of virus-like particle (VLP) vaccines. Furthermore, the structural analysis showed several loops that had an indispensable role in antibody binding and the prevention of HPV infections.

Keywords: HPV, In silico, L1, Structural analysis, Physicochemical properties, Bioinformatics

Citation: Hasanshahi Z, Dehghani B, Hashempour A, Alamdari E. Characterization and structural analysis of the human papilloma virus L1 protein in Iran. *Journal of Kerman University of Medical Sciences*. 2024;31(1):23–28. doi: 10.34172/jkmu.2024.04

Received: August 23, 2023, **Accepted**: November 1, 2023, **ePublished**: February 29, 2024

Introduction

Currently, sexually transmitted infections (STIs) are becoming more common in societies, and human papillomavirus (HPV) infection is among the most prevalent STIs (1). While most HPV infections remain transient and do not cause disease, few of them can develop high-grade precancerous or invasive cervical lesions (2-5). Much research has established the vital role of HPV in the development of cervical cancer and nearly all precursor lesions related to cervical cancers are infected by high-risk HPV types (6,7). Among all cancers, in women, cervical cancer is ranked fourth and affects many individuals globally (8). In 2012, 528 000 new cases and 266 000 deaths were reported (9). Studies have indicated that around 2.5 per 100 000 Iranian women have cervical cancer, which is a low rate; its mortality rate has also been reported at 1.04 per 100 000 women. The high-risk age for Iranian women is between 55 and 65 years (10).

HPV is a small non-enveloped DNA-tumor virus with a virion size of ~55 nm in diameter; its genome codes 6 nonstructural viral regulatory proteins (E1, E2, E4, E5, E6, and E7) and two structural viral capsid proteins (L1 and L2) (11). The virus capsid, which is responsible for the transmission, spread, and survival of the virus in the environment, differs by at least 10% in different types and 2–10% in different subtypes based on the L1 genome (12).

The potential to cause cancer is an index to classify HPV types as high-risk HPV (HR-HPV) and low-risk HPV (LR-HPV) groups; types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68 are commonly considered as HR-HPV, and HPV-6 and HPV-11, which typically cause warts, are the most common LR-HPV types (13,14).

Based on previous studies among all HPV types, HPV 16 and 18 are responsible for around 70% of universal cases, and 68 and 73 are considered "possibly" cancer-causing (15). Globally, HPV-16 is the most frequent genotype;

Table 1. Ramachandran plot and Qmean results for selected and reference sequences

Tools	Ramachandran plot			Qmean
	Favored	Allowed	Outlier	Qmean score
I-TASSER	79.2%	14.40%	6.40%	-8.79
(PS)2-v2	95.4%	4.2%	0.4%	-3.73
Phyre2	85.6%	10.4%	4.0%	-7.7

(11.6%) demonstrated type 18 (25). In 2014, Yousefzadeh et al. studied 851 Iranian women aged 18–65 years; they concluded that HPV infection among Iranian females was higher than the previous estimates reported in Iran. The prevalence of HPV-16 and 18 was 7.3% and 2.8%, respectively (26). A total of 436 Iranian women with different cervical lesions or malignancies were investigated by Salehi-Vaziri et al from 2011 to 2013; in 45.4% of cases, HPV infection was detected, and HPV-16 (32.8%) was the most common HR-HPV genotype (16). In the present study, all selected sequences and reference sequences belonged to HPV-16.

Yoshiyuki Ishii, in 2003 described 6 positions (C175, C185, C428, C161, C229, and C379) in the L1 HPV-16 protein with a vital role in the assembly and integrity of L1 capsids through intramolecular bonding (27). Our analysis showed that there were no changes in the mentioned positions and they were completely conserved.

Teimoori et al in 2008 and Hajmohammadi et al in 2016 expressed papillomavirus 16 L1 protein in *Escherichia coli* and showed the stability of this protein in this host (28,29). Furthermore, in 2011, Coimbra et al. succeeded in producing this protein in a eukaryotic host, *Pichia pastoris*, as an integrative vector (30). Finally, in 2013, Abdoli et al used *Spodoptera frugiperda* (Sf9) cells to express the L1 protein (31). In agreement with the findings from previous studies, our results confirmed the stability of this protein in prokaryotic and eukaryotic cells as well as the thermostability of this protein; also, this indicates that the expression in various hosts can be used to develop self-assembled VLP vaccines as well as diagnostic tests.

Bioinformatics has provided reliable tools to predict virus proteins (32,33). Structural analysis of the L1 protein showed several different loops between strands and several α-helices. Analysis showed mutations could not affect this structure significantly. Many studies have indicated I-TASSER as the most reliable tool to predict 3D structures; however, the present results determined that (PS)2-v2 constructed the most reliable structure for the L1 protein, which was confirmed by Ramachandran plot and Qmean results.

Suhandono et al in 2014 found five loop regions in the

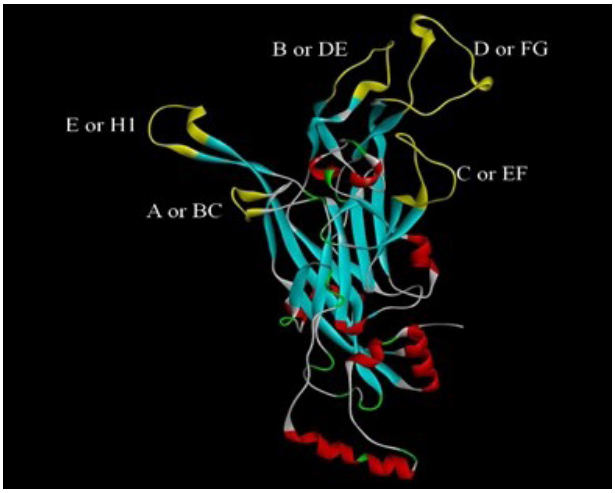

Figure 4. HPV L1 3D model structure. Yellow: 5 identified loops, Red: α-helices

model of the HPV L1 (34). Based on Stanley and colleagues' study in 2006, these loops are the exposed antigens, and any shift at the surface of the loop may contribute to changes in the surface antigen determinants and changes in the antibody needed to identify the virus. Each HPV type has specific loop structures and is functionally active regarding antibody binding (35). In 2006, Carter et al showed DE, FG, and HI are the essential regions for binding by neutralizing the antibodies (36). Moreover, Christensen et al in 2001 showed an immunodominant epitope composed of the FG and HI loops (37). Similar to previous studies, the present study showed five distinct loops that are exposed, and in the selected sequences, we found one substitution that can affect neutralizing antibody binding and the effectiveness of antibodies in preventing HPV infections in the FG or D loop.

Conclusion

To conclude, the results showed L1 was a highly conserved region, and two substitutions, 228 (H to D) and 292 (T to A), did not affect the structure and properties of this protein, which confirmed that the vaccines can still have adequate protection for Iranians. It is suggested that this region should be monitored frequently.

Acknowledgements

The authors would like to thank Shiraz University of Medical Sciences, Shiraz, Iran, for support, and also the Center for Development of Clinical Research of Nemazee Hospital and Dr Nasrin Shokrpour for editorial assistance. Special thanks to Nooshin Zare, Anahid Khanoomi Hasanshahi, and Rozha Rezaei for their generous assistance and Dr Marjan Zare and Dr Behnam Honarvar for statistical and design consultation, respectively.

Authors' Contribution

Conceptualization: Behzad Dehghani.
Data curation: Behzad Dehghani.
Formal analysis: Zahra Hasanshahi.
Funding acquisition: Ava Hashempour.