ISSN: 2321-7782 (Online)
Volume 2, Issue 1, January 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Paper
Available online at: www.ijarcsms.com

In Depth Analysis of the L1 Protection of High Risk Human Papillomavirus – In Silico Approach

Priya Roosvelt[1]
Bioinformatics Infrastructure Facility centre of DBT
Presidency College(Aut)
Chennai - India

Indu Purushothaman[2]
Bioinformatics Infrastructure Facility centre of DBT
Presidency College(Aut)
Chennai - India

Rajarajan S[3]
Bioinformatics Infrastructure Facility centre of DBT
Presidency College(Aut)
Chennai - India

Abstract: ***Human papillomavirus (HPV) causes 80% of cervical cancer in developing countries. The vaccine used for the treatment of cervical cancer is Gardasil and Cervarix. The L1 protein of HPV type is the main element for both the vaccine. But, with regard to the efficacy rate, duration of protection and type of HPV against which protection is provided is found to be very limited. Based on the above restriction prevailing status, the need for an effective and multivalent vaccine for the treatment of the high risk type HPV is in the need. Thus, in the present study, the physiochemical, structural and biological properties of a high risk type of HPV L1 protein were explored. Most of the HPV type shared its structure to the L1 protein of HPV type 18. This study will in turn give us an in depth knowledge about the protein present in the fifteen type of HPV to produce an effective vaccine.***

Keywords: Human Papillomavirus; in silico analysis; L1 protein; vaccine.

I. INTRODUCTION

The second most common cancer caused in women is cervical cancer worldwide and 80% of the cases mainly occurred in developing countries (Ferlay et al.20010). Human papillomavirus (HPV) virus is found to be the main causative agent of cervical cancer (Bosch et al.2002). Every year, an estimate of 490,000 cases and 270,000 deaths caused by cervical cancer is being reported worldwide. In U.S, 11,000 cases and 3900 death were reported in 2008 (Kahn, 2009). Forty out of two hundred types of HPV identified causes genital warts (Bernand, 2010) which is subdivided as high risk types (16,18,31,33,35,39,45,51,52,56,58,59,68,73 and 82), probable high risk types (26,53,66) and low risks (6,11,40,42,43,44,54,61,70,72,81 and CP6108) (Munoz et al. 2006). The major causative agent for the development of cervical carcinoma is HPV type 16 (Kim et al. 2004).

The genome of HPV consists of non-enveloped double stranded DNA viruses of 8Kbp in size (Sinal and Woods. 2005). Three major regions namely early region, a late region and a long control region (LCR) are the main constituent the genome. Six nonstructural regulatory proteins (E1, E2, E4, E5, E6, E7) encode the early region, structural capsid protein (L1- major, L2-minor) encodes late regions which is located downstream of the early regions and finally LCR region is involved in the regulation of viral transcription and replication by providing promoter and multiple transcription factor binding sites (Zheng et al.2006).

The regulation of early transcription and viral DNA replication is implicated by E1 and E2 protein (Doorbar et al. 2012). The viral E2 protein loss its function after the high risk HPV genome gets incorporated into the host cell genome is the main character for the origin of cervical carcinogenesis (Romanczuk et al. 1992). E4 is involved in virion release (Doorbar et al.

1997) and E5, E6, E7 stimulate cell immortalization and transformation and hence called oncogenes (Doorbar, 2006). Throughout the evolution process of HPV, it is found that many nucleotides sequence present in the region of L1 and L2 are conserved and encodes the capsid protein that acts as an envelope and protect the genetic materials present in the HPV system (Zheng et al. 2006). It is accumulated on DNA protein complex (Galloway and McDougall. 1989). A single virus capsid is composed of five molecules of L1 (Baker et al. 1991) and L2 is roughly of 1/10 the abundance of L1 (Doorbar et al. 1987).

The major and minor capsid protein L1 and L2 are about the same size 55kDa but the molecular mass of L2 is found to be migrated to 72kDa as found by expression in in vitro transcription and translation process (Yaegashi et al. 1991). The self-assembly of the L1 major structural proteins will result in the formation of virus particles. When HPV DNA is introduced into virus particles with the help of the L2 protein will result in papillomavirus capsid assembly. Capsid containing L1 and L2 protein complex are found in more number than capsid containing L1 only (Zhou et al. 2004).

The presence of L1 (360 copies) and L2 (72 copies) protein in the viral capsid has been confirmed by immunogold- labeled antibodies (Hagensee et al.1993; Chen et al.2000). The main constituents of the outer capsid shell are the seventy two pentamers of L1 stabilized by intercapsomeric disulfide bonds between highly conserved cysteine residues and linked together by its carboxy-terminal domains (Liu et al.1998). The hydrophobic interactions between a small stretch of amino acids close to the carboxy terminus of L2 and the inner central core of capsomers is required for the association of L1 and L2 inside the capsid.

The vaccine used for the treatment of HPV is Gardasil and Cervarix. The main constituent of Gardasil is major capsid protein L1 of type 6, 11, 16 and 18.While Cervarix is made up of HPV type 16 and 18 L1 protein. Gardasil exhibited an efficacy nearly to 100% against HPV type 16 and 18 and approximately 30% against HPV type 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59 (Schieszer, 2007). After vaccination of Cervarix, it is found to be effective against cervical cancer for 7.3 years (Schwarz, 2007). It is understood that the present vaccine is also found to be against a selective type of HPV and its efficacy differs according to the type. In this paper, we have attempted to study the Physio-chemical, structural and biological properties of a high risk type of HPV L1 proteins. By exploring the Physio-chemical, structural and biological properties of high risk type HPV in the development of effective multivalent vaccine could be attained.

II. Materials and Methods

Retrieval of high risk HPV type protein sequence

The protein sequence of fifteen high risk HPV type L1 (16,18,31,33,35,39,45,51,52,56,58,59,68,73 and 82) was were downloaded from NCBI website.

Sequence information and prediction of antigenicity

The sequence information such as aliphatic index, molecular weight and isoelectric point of each type for both L1 proteins was predicted by CLC bio protein workbench v5.8.1. The antigenicity of each protein was envisaged by Vaxijen v2.0.

Physical properties of proteins

The atomic composition, number of hydrophobic, hydrophilic, negatively and positively charged residues present in each protein was calculated by CLC bio protein workbench v5.8.1.

Secondary structure and domains found in each protein

The secondary structure and domains present in each protein was predicted by CLC bio protein workbench v5.8.1.

Prediction of 3D protein structure

The 3D structure of each protein was predicted by three different structure prediction tools namely Phyre2, Raptor X and (PS) [2]. The quality of predicted protein structure was checked by Pro Q.

Molecular modeling of protein

The structure of predicted protein was completed and bumps if any were removed using What if, a versatile molecular modeling package. The qualities of these proteins were checked by Pro Q.

Structural Bioinformatics studies

The structural Bioinformatics of each protein was carried out by TM-align and cofactor. The comparison and alignment of predicted protein were performed by TM-align. The biological function of each annotated protein was provided by Cofactor server.

III. Results and Discussion

The fifteen high risk HPV L1 protein sequence was retrieved from NCBI and tabulated accordingly with its accession number and the number of amino acids present in it. The antigenicity nature of each L1 protein was calculated and it was found that HPV type 73 (0.61) was the highly antigenic protein among the high risk type studied. While type 31 (0.46), 51 (0.47), 52 (0.48), 56 (0.49), 58 (0.49) and 82 (0.47) was non antigenic. The HPV type 18 L1 proteins possess high molecular weight (79.60) and aliphatic index (63.67). The isoelectric point was high on HPV type 31 L1 proteins (Table 1). The Physio-chemical, structural and biological properties of all high risk HPV L1 proteins was predicted. The Physio-chemical properties of each protein such as molecular weight, aliphatic index and isoelectric point were calculated using CLC bio protein workbench. The pH at which the protein carries no electric charge is termed as isoelectric point and the stability of each protein was known by its aliphatic index. The protective antigen was predicted based on the physicochemical properties of proteins. The physical and chemical properties of each protein were determined by its atomic composition. The pattern of hydrogen bonds between backbone amino and carboxyl groups was defined as the secondary structure of protein.

The atomic composition of each L1 protein namely the number of hydrogen, carbon, nitrogen, oxygen, sulphur, hydrophobic, hydrophilic, negatively charged and positively charged residues was determined. The HPV type 18 showed the highest number of hydrogen (4413), carbon (2862), nitrogen (828), hydrophobic (271) and hydrophilic residue (175). The number oxygen (828), sulphur (29) and negatively charged residues (61) was high in HPV type 56. The negatively charged residues (51) were found to be high in HPV type 33, 52, 58, 68 and 73 (Table2).

The secondary structure and domains present in full and fragment model of each type L1 protein was also studied. The secondary structure of each protein has been mainly made up of alpha helix and beta strand. The number of alpha helix (16) was high in type 33 and the number of the beta strand (32) was found to be high in type 18 HPV. The MHC_II_beta domains were found to be in the HPV type 33, 45,58 and 82 full model only and in type 16, 35, 73 full and fragment model. The fragment of model of HPV type 58 and type 73 were made up of RVT_1 and Hemagglutinin domains respectively. At the same time, NADH_dehy_S2_C domains were found in the fragment model of HPV type 16, 31, 35. Whereas, no domains were found in HPV type 39,51,56,59 and 68 (Table 3).

With the help of the atomic composition and type of chemical bonds present between these atoms, the primary structure of protein can be determined. Thus, in this paper, we have predicted the atomic composition and the number of chemical bonds was predicted for each type of high risk HPV L1 proteins. Also, the number of alpha helix and beta strand which is the main constituents of the secondary structure has been also computed. Thus the patterns of hydrogen bonds among the main- chain peptide groups were defined by the secondary structure of protein. The stable component of protein structure that could fold in parallel is defined as the domains. The domains present in the full and fragment model of each protein was also tabulated (Table 3). The major domains identified are MHC_II_beta, NADH_dehy_S2_C, Trypsin, RVT_1 and Hemagglutinin. The MHC_II_

beta domains involved in antigen processing presentation and immune response (Holling et al. 2004); NADH dehy S2_C domains were involved in mitochondrial electron transport and oxidation-reduction process (Hochstein and Dalton, 1973) and the trypsin domains played a vital role in serine-type endopeptidase activity and proteolysis (Goldman et al. 2006).

The 3D structure of each protein was predicted by Phyre2, Raptor X and (PS) 2 and their qualities were checked by Pro Q. The quality of the protein structure of HPV type 45 (4.1) was the best predicted by Phyre2 when compared with the other predicted proteins. (Table4).

The tertiary structure of the L1 protein of the fifteen type of HPV was predicted using three different protein structure tools namely Phyre2, Raptor X and (PS) 2 tools. The use of HHpred/ HH search package, profile-profile alignment algorithm forms the basis of Phyre 2 tools. For the selection of template and target-template alignment, an automated homology modeling server, (PS) 2 combines PSI-BLAST, IMPALA and T-coffee method. The protein structures that are not similar to the existing model present in protein data bank were predicted by Raptor X server. The WHAT IF, a versatile molecular modeling package was used to model the missing side chains and remove the bumps by rotating the side chains torsion angles. Finally, the quality of the each predicted protein were checked by Pro Q. Based on the –log of P-value, the quality of the protein were determined. The L1 protein of HPV type 45 predicted by Phyre 2 was found to have the best LG score among the other types of HPV than other predicted tools such as Raptor X and (PS)2.

The best TM-score normalized by the length of chain 1 (0.85) and chain 2 (0.96) was secured by HPV type 18 (Table 5). The biological properties of high risk type HPV L1 protein were calculated. The protein structure and enzyme similar to each type along with its PDB hit and the TM - score were tabulated. The structure of the L1 protein of HPV type 16 were similar to the L1 protein of HPV type 18, 31,33,35,39, 51, 52, 56,58, 59, 68 and type 82, while the pentamer structure of the major capsid protein L1 of HPV type 11 were similar to HPV type 16 and 45 L1 protein. The highest TM-score (1.00) were shared between HPV types 31, 51, 56 and type 68 and lowest TM score (0.91) was scored by HPV type 16. The crystal structure of the KatG enzyme at pH6.5 was shared by HPV type 16, 39, 56, 59 and type 68. But, HPV type 18, 31, 33, 39, 56, 59 and 68 L1 protein structures were homologous to the crystal structure of HSV type 1 DNA polymerase enzyme. While the remaining HPV type L1 protein was homologous to specific enzyme (Table6).

Using TM-align, the structural similarity was explored for the L1 protein and its analogous template. From the result, it was found that HPV type 18 showed the best TM-score when compared with the rest of the HPV type. The biological properties such as similar to the other protein structure and enzyme for each type of HPV showed a new dimension in understanding the L1 protein. From this study, it was confirmed that the L1 protein structure of HPV type 16 was highly homologous to most of the L1 protein high risk HPV. The KatG enzyme, Mycobacterium tuberculosis catalase-peroxidase enzyme, E-coli pyruvate dehydrogenase E1 component phosphonolactylthiamin diphosphate, Human milk xanthine oxidoreductase and catalase-peroxidase from Haloarcula marismortui shared similiar structure to HPV type L1 protein and was found to possess oxdioreductase activity; the glucodextranase enzyme from Arthrobacter globiformis 142 and prostate-specific membrane antigen, a tumor marker and peptidase has been described to exhibit hydrolase activity; Herpes simplex virus type 1 DNA polymerase expresses itself as transferase.

The vaccine used for the treatment of cervical cancer is Gardasil and Cervarix. The L1 protein of HPV type is the main element of both the vaccine. But, with regard to the efficacy rate, duration of protection and type of HPV against which protection is provided is found to be very limited. Based on the above restriction prevailing status, the need for an effective and multivalent vaccine for the treatment of the high risk type HPV is in the need.

IV. CONCLUSION

These preliminary studies on the physio-chemical, structural and biological properties sought out the subtle difference between the serotypes of high risk type of HPV and hence the considering the significant differences found could help in better understanding the limitations of the immune response provoked by the current vaccine.

References

1. Ferlay, J., Bray, F., Pisani, P., Parkin, D.M., Globocan, 2000. Cancer incidence, mortality and prevalence worldwide, version 1.0. IARC CancerBase no. 5. Lyons, France, IARC Press, 2001.
2. Bosch, F.X., Lorincz, A., Munoz, N., Meijer, C.J., Shah, K.V. 2002, The causal relation between human papillomavirus and cervical cancer. J Clin Pathol 55,244-65.
3. Kahn, J.A., 2009. HPV vaccination for the prevention of cervical intraepithelial neoplasia. N. Engl. J. Med 361, 271–8.
4. Sinal, S.H., Woods, C.R., 2005. Human papillomavirus infections of the genital and respiratory tracts in young children. Seminars in Pediatric Infectious Diseases, 16, 306-316.
5. Bernard, H.U., Burk, R.D., Chen, Z., van Doorslaer, K., Hausen, H., de Villiers, E.M. 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. Virology 401, 70-79.
6. Zheng, Z.M., Baker, C.C.,2006. Papillomavirus genome structure, expression, and post- transcriptional regulation Front Biosci 11, 2286-2302.
7. Doorbar, J., Quint, W., Banks, L., Bravo, I.G., Stoler, M., Broker, T.R., Stanley, M.A.,2012. The biology and life-cycle of human papillomaviruses Vaccine 30 , F55-70.
8. Romanczuk, H., Howley, P.M.,1992. Disruption of either the E1 or the E2 regulatory gene of human papillomavirus type 16 increases viral immortalization capacity Proc Natl Acad Sci USA 89, 3159-3163.
9. Doorbar, J., Foo, C., Coleman, N., Medcalf, L., Hartley, O., Prospero, T., Napthine, S., Sterling, J., Winter, G., Griffin, H., 1997. Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic Fabs to E4. Virology 238, 40-52.
10. Doorbar, J., 2006 Molecular biology of human papillomavirus infection and cervical cancer Clin Sci (Lond) 110, 525-541.
11. Kim, S.H., Kim, K.S., Lee, E.J., Kim, M.O., Park, J.H., Cho, K.I., Imakawa, K., Hyun, B.H., Chang, K.T,, Lee,H.T., Ryoo, Z.Y ., 2004. Human keratin 14 driven HPV 16 E6/E7 transgenic mice exhibit hyperkeratinosis. Life Sc, 75, 3035-42.
12. Munoz, Nubia., Xavier Bosch, F., Silvia de Sanjose., Herrero, R., Castellsagué, X., Shah, K.V., Peter, J.F., Snijders, Chris, J.L.M. Meijer., 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer New England Journal of Medicine 348, 518-527.
13. Yaegashi, Jenison, N.S.A., Valentine, J.M., Dunn, M., Taichman, L.B., Baker, D.A., Galloway D.A.,1991. Characterization of murine polyclonal antisera and monoclonal antibodies generated against intact and denatured human papillomavirus type 1 virions. J. Virol. 65,1578-1583.
14. Galloway, D. A., McDougall, J.K., 1989 . Human papillomaviruses and carcinomas. Adv. Virus Res. 37,125-171.
15. Baker, T.S., Newcomb, W.W., Olson, N.H., Cowsert, L.M., Olson, C., Brown, J.C., 1991. Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and threedimensional image reconstruction. Biophys. J. 60,1445-1456.
16. Doorbar, J., Gallimore, P.H.,1987. Identification of proteins encoded by the Li and L2 open reading frames of human papillomavirus la. J. Virol. 61,2793-2799.
17. Zhou, J., Sun, X.Y., Louis, K., Frazer, I.H., 1994. Interaction of human papillomavirus (HPV) type 16 capsid proteins with HPV DNA requires an intact L2 N-terminal sequence Journal of virology 68, 619-625.
18. Hagensee, M.E., Yaegashi, N., Galloway, D.A., 1993. Self-assembly of human papillomavirus type 1 capsids by expression of the L1 protein alone or by coexpression of the L1 and L2 capsid proteins. Journal of virology 67, 315-322.
19. Chen, X.S., Garcea, R.L., Goldberg, I., Casini, G., Harrison, S.C., 2000. Structure of small virus-like particles assembled from the L1 protein of human papillomavirus 16. Mol. Cell 5,557–567.
20. Li, M., Beard, P., Estes, P.A., Lyon, M.K., Garcea, R.L., 1998. Intercapsomeric disulfide bonds in papillomavirus assembly and disassembly. J. Virol 72, 2160–2167.
21. Finnen, R.L., Erickson, K.D., Chen, X.S., Garcea, R.L., 2003. Interactions between papillomavirus L1 and L2 capsid proteins. J. Virol 77,4818– 4826.
22. Schieszer, J., 2007. HPV Vaccine Protects Against 10 "Other" Strains Internal Medicine World Report.
23. Schwarz, T.F., 2009. Clinical update of the AS04-Adjuvanted human Papillomavirus-16/18 cervical cancer vaccine, cervarix®. Advances in Therapy 26 , 983–998.
24. Holling TM, Schooten E, Van Den Elsen PJ (2004) Function and regulation of MHC Class II molecules in T-lymphocytes, of mice and men. Hum Immunol 65, 282-290.
25. Hochstein LI and Dalton BP (1973) Studies of a halophilic NADH dehydrogenase I purification and properties of the enzyme. Biochim.Biophys.Acta 302, 216-228.
26. Goldman BS. Nierman WC, Kaiser D, Slater SC, Durkin AS, Eisen J, Ronning CM, Barbazuk WB, Blanchard M, Field C, Halling C, Hinkle G, Iartchuk O, Kim HZ, Mackenzie C, Madupu R, Miller N, Shvartsbeyn A, Kaplan HB (2006) Evolution of sensory complexity recorded in a myxobacterial genome. Proc. Natl. Acad. Sci. USA 103, 15200-15205.

Table I:Physical properties of High risk type L1 protein with accession number

HPV type	Accession number	Number of amino acids	Prediction of antigencity	Molecular weight	Aliphatic index	Isoelectric point
16 isolate QV17722E	AAV91691	531	0.52 antigen	59.03	76.53	8.52
18 isolate BF226	AGU90430	568	0.49 non-antigen	63.67	79.60	8.8
31 ISOLATE QV12357	AEI61093	531	0.46	56.28	66.73	9.04
33 ISOLATE Qv34189	AEI61253	499	0.54 antigen	55.99	69.52	8.75
35 ISOLATE QV31639	AEI61437	502	0.52 antigen	56.15	70.26	8.9
39 ISOLATE BF375	AGU90581	505	0.56 antigen	56.60	69.45	8.75
45 ISOLATE Qv34163	AGU90648	536	0.52 antigen	60.07	75.26	8.8
51 strain Ma (Japan)	ACV88631	504	0.47	56.29	72.76	8.8
52 Isolate Qv07294	AEI61605	529	0.48	59.50	72.74	8.73
56 clone Qv25665	ABO76830	563	0.49	63.45	77.57	8.87
58 strain CNZJ-2	AGQ21871	498	0.49	56.17	67.00	8.82
Type 59 (China)	ACL12341	508	0.53 antigen	56.84	73.05	8.92
68 Isolate Qv33999	AGU90815	505	0.52 antigen	56.80	70.22	8.66
Type 73	CAA63887	503	0.61 antigen	56.42	65.13	8.77
Type 82 subtype IS39/AE2	AAK28456	503	0.47	56.12	74.85	8.97

Table II: Atomic composition, hydrophobic, hydrophilic, negatively and positively charged residues of HPV L1 protein

HPV type	hydrogen	carbon	nitrogen	oxygen	sulphur	Hydro phobic	Hydro philic	-ve charge	+charge
16	4128	2682	702	783	24	252	166	49	53
18	4413	2862	769	826	27	271	175	49	57
31	3875	2522	675	743	23	236	156	47	57
33	3873	2511	665	742	23	231	153	51	57
35	3875	2503	673	751	23	226	162	49	57
39	3883	2517	687	754	24	229	160	50	56
45	4153	2688	729	787	25	252	168	48	55
51	3882	2516	684	750	18	234	159	48	54
52	4122	2683	700	782	25	250	163	51	57
56	4410	2837	768	828	29	267	171	53	61

58	3867	2518	669	745	23	226	157	51	58
59	3939	2542	689	753	20	234	156	50	58
68	3899	2530	689	756	23	228	160	51	56
73	3878	2513	676	754	25	229	160	51	57
82	3886	2506	688	744	18	236	156	47	55

Table III: Prediction of secondary structure and domains present in each HPV type L1 protein

HPV type	Alpha helix	Beta strand	Full model	Fragment model
16	11	30	MHC_II_beta	NADH_dehy_S2_C, MHC_II_beta
18	8	32	-	MHC_II_beta
31	10	30	-	NADH_dehy_S2_C
33	16	28	MHC_ II_ beta	-
35	7	32	MHC_II_beta	MHC_II_beta, NADH_dehy_S2_C
39	4	29	-	-
45	6	33	MHC_ II_ beta	-
51	12	26	-	-
52	10	29	Trypsin	-
56	13	25	-	-
58	13	26	MHC_ II_ beta	RVT_1
59	9	29	-	-
68	6	27	-	-
73	15	20	MHC_ II_ beta	Hemagglutinin, MHC_II_beta
82	10	29	MHC_ II_ beta	-

Table IV:The quality check of 3D structure of L1 protein predicted by Phyre2, Raptor X and $(PS)^2$ performed by Pro Q

HPV type	**Phyre 2**	**Raptor X**	$(PS)^2$
16 isolate QV17722E	2.23	2.39	2.33
18 isolate BF226	3.98	2.55	3.09
31 ISOLATE QV12357	2.35	2.17	1.91
33 ISOLATE Qv34189	2.57	2.21	2.17
35 ISOLATE QV31639	2.36	2.36	1.54

39 ISOLATE BF375	2.30	2.27	3.08
45 ISOLATE Qv34163	4.10	2.58	3.34
51 strain Ma (Japan)	2.31	2.07	2.17
52 Isolate Qv07294	2.46	2.64	2.61
56 clone Qv25665	2.28	2.33	2.41
58 strain CNZJ-2	2.52	2.53	2.17
Type 59 (China)	2.31	2.20	3.18
68 Isolate Qv33999	2.36	2.27	2.98
Type 73	2.25	2.05	1.79
Type 82 subtype IS39/AE2	2.40	2.01	2.04

Table V: TM-score normalized by chain1 and 2 for HPV type L1 protein

HPV type of L1 protein	Length of chain 1	TM-score (normalized by length of chain 1)	Length of chain 2	TM-score (normalized by length of chain 2)	Aligned length
16	481	0.84	455	0.89	422
18	481	0.85	423	0.96	420
31	481	0.84	456	0.89	422
33	481	0.84	454	0.89	420
35	481	0.84	453	0.89	420
39	481	0.84	455	0.89	422
45	481	0.84	426	0.95	420
51	481	0.84	455	0.89	423
52	481	0.84	459	0.88	422
56	481	0.84	453	0.89	421
58	481	0.84	454	0.89	420
59	481	0.84	456	0.89	423
68	481	0.84	456	0.89	423
73	481	0.84	457	0.89	423
82	481	0.84	456	0.88	424

Table VI: The biological properties of HPV type L1 protein

HPV type of L1 protein	Similar structure PDB hit	TM score	Classification	Similar enzymes PDB hit	TM score	Classifcation
16	2r5kE	0.91	Pentamer structure of major capsid protein L1 of HPV type 11	2fxhA	0.29	Crystal structure of KatG at pH 6.5
18	1dzlA	0.97	L1 protein of HPV type 16	2gv9B	0.30	Crystal structure of HSV type 1

						DNA polymerase
31	1dzlA	1.00	L1 protein of HPV type 16	2gv9B	0.29	Crystal structure of HSV type 1 DNA polymerase
33	1dzlA	0.99	L1 protein of HPV type 16	1ug9A	0.29	Crystal structure of glucodextranase from Arthrobacter globiformis 142
35	1dzlA	0.99	L1 protein of HPV type 16	1ug9A	0.29	Crystal structure of glucodextranase from Arthrobacter globiformis 142
39	1dzlA	0.99	L1 protein of HPV type 16	2fxhA	0.29	Crystal structure of KatG at pH 6.5
45	2r5iO	0.99	Pentamer structure of major capsid protein L1 of HPV 18	1sj2A	0.32	Crystal structure of Mycobacterium tuberculosis catalase-peroxidase
51	1dzlA	1.00	L1 protein of HPV type 16	1z8LA	0.29	Crystal structure of prostate-specific membrane antigen, a tumor marker and peptidase
52	1dzlA	0.99	L1 protein of HPV type 16	1itkB	0.30	Crystal structure of catalase-peroxidase from Haloarcula marismortui
56	1dzlA	1.00	L1 protein of HPV type 16	2fxhA	0.29	Crystal structure of KatG at pH 6.5
58	1dzlA	0.99	L1 protein of HPV type 16	1z8LA	0.29	Crystal structure of prostate-specific membrane antigen, a tumor marker and peptidase
59	1dzlA	0.99	L1 protein of HPV type 16	2fxhA	0.29	Crystal structure of KatG at pH 6.5
68	1dzlA	1.00	L1 protein of HPV type 16	2fxhA	0.29	Crystal structure of KatG at pH 6.5
73	1dzlA	0.99	L1 protein of HPV type 16	2ckjA	0.28	Human milk xanthine oxidoreductase
82	1dzlA	0.99	L1 protein of HPV type 16	2qtcB	0.31	E-coli pyruvate dehydrogenase E1 component phosphonolactylthiamin diphosphate