In silico Structural Prediction of E6 and E7 Proteins of Human Papillomavirus Strains by Comparative Modeling

Satish Kumar^{1*}, Lingaraja Jena¹, Vidya W. Bhomale¹, Sangeeta Daf^{2,3}

¹Bioinformatics Centre Mahatma Gandhi Institute of Medical Sciences Sevagram (Wardha) – 442 102, Maharashtra, India E-mails: <u>satishkumar@mgims.ac.in</u>, <u>lingaraj jena@yahoo.co.in</u>, <u>vidyabomale@gmail.com</u>

²Datta Meghe Institute of Medical Sciences (Deemed University) Atrey Layout, Pratap Nagar, Nagpur – 440 022, Maharashtra, India

³Mother Care Hospital 805, Sudampuri, Umrer Road, Nagpur – 440 009, Maharashtra, India E-mail: <u>sangeetadaf@gmail.com</u>

*Corresponding author

Received: June 6, 2011

Accepted: July 12, 2012

Published: July 16, 2012

Abstract: More than 200 different types of Human papillomavirus (HPV) are identified, 40 transmit extensively through sexual contacts affecting the genital tract. HPV strains have been etiologically linked to vaginal, vulvar, penile, anal, oral and cervical cancer (99.7%) as a result of mutations leading to cell transformations due to interference of E6 and E7 oncoproteins with p53 and pRB tumor suppressor genes respectively, besides other cellular proteins. As structures of E6 and E7 proteins are not available, the simultaneous structural analysis of E6 and E7 proteins of 50 different HPV strains was carried out in detail for prediction and validation, using bioinformatics tools. E6 and E7 proteins sequences were retrieved in FASTA format from NCBI and their structures predicted by comparative modeling using modeller9v6 software. Further, most of the HPV strains showed good stereochemistry results in most favored regions when subjected to PROCHECK analysis and subsequently each protein was validated using ProSA-web tool. The work carried out on comparing and exploring the structural variations in these oncogenic proteins might help in genome based drugs and vaccines designing, beyond their limitations.

Keywords: Human papillomavirus, E6 and E7 oncoproteins, Comparative modeling, Modeller.

Introduction

Human papillomavirus (HPV) belongs to a large group of viruses that infect the skin cells and cause them to mutate and grow irregularly. These irregular skin growths commonly called warts, are tiny tumors named as Papillomas. These viruses are epitheliotropic as they induce epithelial hyperproliferation, including cutaneous warts and condylomas in cervical and vaginal epithelia [12]. About 200 different Types of HPV have been identified based on DNA homology; approximately 40 of them affect the genital tract. HPV strains have been etiologically linked to cervical (99.7%) [3], vaginal, vulvar, penile, anal and oral cancers [17, 18].

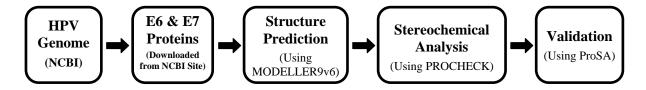
Genome of Human papillomavirus consists of three regions such as Six Early ORFs (Open Reading Frames), Two Late ORFs and Upstream Regulatory Region (URR) [21]. E6 and E7 proteins encoded by E6 and E7 early ORFs, are transforming in nature [10, 11] and have strong binding affinity to p53 and pRB tumor suppressor genes respectively [6, 9, 11]. These viral oncoproteins are major contributors to neoplastic progression by interfering with cell cycle G1-S checkpoint. Among a variety of cellular targets, E6 binds and degrades TP53 protein by forming a complex with the ligase E6AP [14, 20], leading to genetic instability while E7 abrogates pRB protein function through its ubiquitination-mediated degradation, which leads to activation of E2F regulated genes and deregulates the progression through G1 phase of the cell cycle. Integration of viral sequences into the host genome interrupts E2 ORF, leading to the constitutive expression of E6/E7 in the transformed cells [14].

Interestingly, none of the structures of these two oncoproteins of different HPV strains are available; hence, we carried out this study of structure prediction and validation of possible available HPV strains whose genomes are completely sequenced.

Materials and methods

There are 50 strains of HPV whose genomes are completely sequenced and available at genome site of National Centre for Biotechnology Information (NCBI). Hence, all E6 and E7 protein sequences of HPV strains were retrieved in FASTA Format with their respective accession numbers. The Molecular Weight (MW) and Isoelectric Points (pI) of these proteins were calculated using tools available at ExPASy (**Expert Protein Analysis System**) proteomic server of Swiss Institute of Bioinformatics [5]. Then all the results were put in the tabular format separately for E6 and E7 protein sequences.

Each of these protein sequences is aligned based on multiple sequence protein alignment program, BLASTp [1] against Protein Data Bank for characterization at molecular level. Three dimensional structures of E6 and E7 proteins were predicted by comparative modeling using MODELLER9v6 software [15] and visualized in the RasMolv2.5 software [16]. The stereochemistry of each protein was evaluated through PROCHECK analysis [8] and validated using ProSA-web [19].



Results and discussion

Although genome sequence of 50 different strains of HPV is available at NCBI, E6 proteins of HPV 101, 103 & 108 are not available. Accordingly, the comparative results of E6 proteins of 47 different strains of HPV as well as E7 proteins of 50 different strains of HPV were subjected to a tabular format, based on Sequence Length, MW and pI. pI and MW of all the proteins vary with protein sequences (Table 1). The lengths of E6 proteins of all HPV strains under study, range in between 137 to 159 amino acids, having MW in the range of 15.80 kD to 19.18 kD and pI from 5.35 to 9.16 except HPV 96 which has longest E6 protein having length of 225 amino acids, MW 26.04 kD and pI 9.07. However, the number of amino acids in E7 proteins of different strains of HPV varies from 86 to 114 having MW in the range of 9.54 kD to 12.8 kD and pI range of 4.07 to 5.16.

	On the basis of length, WW (kD) & Ir (p1).E6 PROTEINE7 PROTEIN							
HPV TYPE	Accession No	Length	Mol Wt	pI	Accession No		Nol Wt	nI
		0		<u>р</u> 6.80		0		-
HPV1 HPV2	NP 040305.1	140aa	16.317		NP 040307.1	93aa	10.500	4.21
HPV2 HPV4	NP 077116.1 NP 040889.1	159aa	18.301	8.46 7.87	NP 077117.1	92aa		
HPV4 HPV5		140aa	16.487 18.068	5.35	NP 040890.1	100aa 103aa		4.40
-	NP 041365.1	157aa			NP 041366.1			
HPV6	CBY85548.1	150aa	17.290	8.20	CBY85549.1	98aa	10.903	4.43
HPV7	NP 041854.1	154aa	17.881	8.65	NP 041855.1	111aa	12.459	4.80
HPV9	NP 041860.1	148aa	17.278	6.30	NP 041862.1	93aa	10.392	4.68
HPV10	NP 041741.1	148aa	17.563	9.05	NP 041742.1	86aa	9.541	4.68
HPV16	NP 041325.1	158aa	19.187	9.16	NP 041326.1	98aa	11.022	4.20
HPV18	NP 040310.1	158aa	18.871	8.95	NP 040311.1	105aa	11.995	4.70
HPV24	AAA79415.1	140aa	16.319	6.79	AAA79416.1	96aa		4.43
HPV26	NP 041782.1	150aa	17.922	8.95	NP 041783.1	104aa	11.998	4.08
HPV32	NP 041801.1	142aa	16.631	8.65	NP 041802.1	104aa	11.591	4.07
HPV34	NP 041807.1	148aa	17.734	8.79	NP 041808.1	97aa	10.985	4.49
HPV41	NP 040285.1	156aa	17.302	7.45	NP 040286.1	114aa	12.804	4.84
HPV48	NP 043416.1	142aa	16.750	8.28	NP 043417.1	93aa	10.418	4.93
HPV49	NP 041832.1	138aa	16.201	7.02	NP 041833.1	103aa	11.454	4.26
HPV50	NP 043423.1	141aa	16.410	7.42	NP 043424.1	93aa	10.516	5.10
HPV53	NP 041844.1	154aa	18.168	9.03	NP 041845.1	105aa	12.161	4.47
HPV54	NP 043288.1	144aa	17.132	8.74	NP 043289.1	95aa	10.565	4.68
HPV60	NP 043437.1	142aa	16.809	8.37	NP 043438.1	96aa	10.687	4.58
HPV61	NP 043444.1	146aa	16.991	7.00	NP 043445.1	95aa	10.461	4.40
HPV63	NP 040901.1	141aa	16.317	8.04	NP 040902.1	88aa	9.870	4.23
HPV71	AAQ95198.1	157aa	17.875	8.26	AAQ95185.1	94aa	10.571	4.46
HPV88	YP 001672008.1	142aa	16.735	8.14	YP 001672009.1	98aa	10.896	4.54
HPV90	NP 671503.1	148aa	17.173	6.79	NP 671504.1	98aa	10.944	4.58
HPV92	NP 775305.1	138aa	15.808	6.29	NP 775306.1	91aa	10.115	4.35
HPV96	NP 932319.1	225aa	26.048	9.07	NP 932320.1	99aa	11.030	4.38
HPV98	CAW42212.1	153aa	17.725	6.79	CAW42214.1	95aa	10.600	4.54
HPV99	CAW42225.1	155aa	17.647	5.71	CAW42227.1	103aa	11.583	4.31
HPV100	CAW42235.1	152aa	18.182	7.93	CAW42236.1	100aa	11.194	4.38
HPV101	Not available				YP 656499.1	98aa	10.741	4.88
HPV103	Not available				YP 656493.1	100aa	11.543	4.94
HPV104	CAW42247.1	138aa	16.431	6.88	CAW42248.1	104aa	11.673	4.35
HPV105	CAW42259.1	155aa	17.810	5.35	CAW42261.1	101aa	11.268	4.22
HPV107	ABN79867.1	140aa	16.553	7.51	ABN79868.1	102aa	11.582	4.51
HPV108	Not available				YP 002647034.1	99aa	11.165	4.89
HPV109	YP 002756538.1	140aa	16.054	7.35	YP 002756539.1	96aa	10.679	4.75
HPV112	YP 002756545.1		16.393	8.29	YP 002756546.1	97aa	10.833	4.47
HPV113	CAW42270.1	149aa	17.423	6.71	CAW42271.1	92aa	10.508	4.60
HPV116	YP 003084346.1		16.636	8.37	YP 003084347.1	98aa	11.008	4.93
HPV121	YP 003668025.1		16.936	8.64	YP 003668026.1	98aa	11.169	4.34
HPV128	YP 004169263.1		17.000	7.88	YP 004169264.1	96aa	10.919	5.06
HPV129	YP 004169270.1		17.779	8.59	YP 004169271.1	97aa	10.973	4.84
HPV131	YP 004169277.1		16.390	8.73	YP 004169278.1	94aa	10.303	4.81
HPV132	YP 004169284.1		16.721	8.35	YP 004169285.1	92aa	10.445	5.16
HPV134	YP 004169291.1		15.971	8.88	YP 004169292.1	91aa	10.163	5.04
HPV148	YP 004111309.1		16.184	8.37	YP 004111310.1	93aa	10.332	4.71
HPV KI88-03	ACC78256.1	138aa	16.417	6.88	ACC78257.1	104aa	11.687	4.37
HPV RTRX7	AAB61640.1	158aa 157aa	18.166	6.29	ACC78237.1 AAB61641.1	104aa 103aa	11.087	4.37
ΠΓΥΚΙΚΑ/	AAD01040.1	1 <i>J</i> /dd	10.100	0.29	AAD01041.1	10588	11.47/	4.21

Table 1. Comparative analysis of E6 & E7 proteins of 50 different strains of HPV on the basis of length, MW (kD) & IP (pI).

Homology modeling

Three dimensional structure of each protein was predicted by homology modeling. In general, 30% sequence identity is required for generating useful 3D structure models [2, 4, 13]. Kaladhar et al. also predicted structures of E6 & E7 proteins along with other proteins of only HPV Type 92 using Swissmodel server but it seems that the predicted structures are not found validated [7]. In our study, we predicted structures of E6 proteins of 47 and E7 proteins of 50 different strains of HPV by comparative modeling using MODELLER9v6 software and visualized in the RasMolv2.5 software. However, the structure of E7 protein of HPV 41 was not predicted since it had only 27% percentage identity with known structure (PDB ID 1P5Q). Accordingly, total 96 structures of E6 and E7 proteins of 30 different HPV strains were generated in Modeller9v6 software. The E6 proteins of all HPV strains show close similarities with structures (2FK4) available at PDB except HPV 92 (3GE3), HPV 99 (1DQ3) and HPV 116 (1VZ3), while E7 proteins of all HPV strains show close similarities with the structures (2B9D) except HPV 2, 6, 10, 16, 18, 24, 26, 32, 53, 96, 99 and 109 (2EWL).

Evaluation of protein structure quality

The stereo-chemical quality of these predicted structures were then evaluated through PROCHECK analysis. Remarkably the stereochemistry of E6 proteins of HPV 4 & 26 and E7 proteins of HPV 88 & 108 revealed that 94.90% to 95.50% residues were positioned in most favorable region of the Ramachandran plot. Some of E6 (HPV 1, 2, 6, 7, 9, 50, 132) & E7 proteins (HPV 1, 7, 9, 34, 63, 71, 90, 92, 100, 104, 112, 113, 131, RTRX7) show a good quality model range of 90 to 92% and 90 to 94% respectively. Most of E6 (HPV 5, 10, 16, 18, 24, 32, 34, 41, 48, 49, 53, 54, 60, 61, 63, 71, 88, 90, 96, 98, 99, 100, 104, 105, 107, 109, 112, 113, 121, 128, 129, 131, 134, 148, KI88-03 & RTRX7) & E7 proteins (HPV 2, 5, 4, 6, 16, 18, 24, 26, 32, 48, 49, 50, 53, 54, 60, 61, 96, 98, 99, 101, 103, 105, 107, 116, 121, 128, 129, 132, 134, 148, KI88-03) show the range of 85 to 89% and remaining E6 (92, 116) and E7 proteins (10, 109) show the range of 74 to 79% (Table 2).

Validation of 3D structures

3D structures of each protein were subjected to validation using ProSA-web and when analyzed, it revealed a compatible Z-score value of residue energies (Table 2) within the range of native conformations of crystal structures. The ProSA-web based validation analysis showed largely negative Z-score in most of E6 proteins expect HPV 41, 105, 109, 116, 129, 132 and KI88-03, while negative Z-score in all E7 proteins. The residue energies including pair energy, combined energy and surface energy were all negative and had similar surface energy tendency with template.

Figs. 1a and 1d, respectively, show validated 3D structures of E6 & E7 proteins of ONLY HPV 18 amongst all 96 3D structures of 50 HPV strains under study. Figs. 1b and 1e present the respective Z scores and Figs. 1c and 1f – the residue energies of these proteins.

HeV Type E6 PROTEIN E7 PROTEIN HPV 1: Iab.II (%) Z -score [ab.II (%) Z -score [by ProSA-web] Z -score [ab.II (%) Z -score [ab.II (%) HPV-1 90.00 -1.42 92.80 -3.59 HPV-2 91.20 -2.77 84.10 -3.8 HPV-5 85.20 -0.77 80.90 -3.06 HPV-6 90.80 -1.61 81.20 -1.75 HPV-9 91.20 -2.23 91.50 -3.13 HPV-16 81.40 -1.73 84.90 -2.1 HPV-16 81.40 -1.73 84.90 -2.1 HPV-26 94.90 -2.11 83.70 -2.03 HPV-31 88.60 -2.34 90.00 -2.66 HPV-34 88.60 -2.34 90.00 -2.67 HPV-48 86.50 -1.14 82.70 -1.86 HPV-49 87.40 -1.64 82.40 -2.78 HPV-40 89.50 -0.6 87.80 -3.75 </th <th></th> <th colspan="8">of E6 & E7 proteins of 50 different strains of HPV</th>		of E6 & E7 proteins of 50 different strains of HPV							
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$			IEIN	E7 PROTEIN					
tab.l1 (%)Dy ProsA-webtab.l1 (%)Dy ProsA-webHPV-190.00-1.4292.80-3.59HPV-495.50-2.5886.80-3.17HPV-585.20-0.7780.90-3.06HPV-690.80-1.6181.20-1.75HPV-790.10-1.7391.10-2.19HPV-885.50-1.1374.70-1.92HPV-1089.50-1.1374.70-1.92HPV-1681.40-1.7384.40-2.6HPV-2484.40-0.6481.50-2.08HPV-3283.20-2.2185.60-2.44HPV-3283.20-2.2185.60-2.44HPV-3488.60-2.3490.00-2.66HPV-4987.40-1.1288.50-3.44HPV-4987.40-1.1288.50-3.44HPV-5090.00-0.7387.10-2.77HPV-5485.40-1.6482.40-2.78HPV-6189.50-0.687.80-3.31HPV-6388.80-2.0993.40-2.51HPV-6489.50-1.4495.30-3.01HPV-7187.90-1.3394.00-2.72HPV-7285.70-0.1282.60-3.35HPV-7389.90-1.4495.30-3.01HPV-7486.60-1.9983.50-2.55HPV-7589.00-1.4495.30-3.01HPV-7689.50 <td< th=""><th>HPV Type</th><th></th><th>Z-score</th><th></th><th>Z-score</th></td<>	HPV Type		Z-score		Z-score				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $		-	by ProSA-web		by ProSA-web				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-1	_ / / _ / /	-1.42		-3.59				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $		91.20	-2.77	84.10					
$\begin{array}{c c c c c c c c c c c c c c c c c c c $									
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-5	85.20	-0.77		-3.06				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-6	90.80	-1.61						
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-7	90.10	-1.73	91.10	-2.19				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-9	91.20	-2.23	91.50	-3.13				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-10	89.50	-1.13	74.70	-1.92				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-16	81.40	-1.73	84.90	-2.1				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-18	87.70	-1.17	88.40	-2.6				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-24	84.40	-0.64	81.50	-2.03				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-26	94.90	-2.11	83.70	-2.08				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $									
HPV-41 88.30 0.2 Percentage identity is less than 30%. Hence structure cannot be predicted HPV-48 86.50 -1.14 82.70 -1.86 HPV-49 87.40 -1.12 88.50 -3.44 HPV-50 90.00 -0.73 87.10 -2.77 HPV-51 89.40 -1.64 82.40 -2.78 HPV-60 89.50 -0.6 87.80 -3.75 HPV-61 89.20 -2.58 87.10 -2.51 HPV-63 88.80 -2.09 93.40 -2.5 HPV-73 87.90 -1.33 94.00 -2.72 HPV-88 89.20 -1.44 95.30 -3.01 HPV-90 87.90 -1.33 94.00 -2.75 HPV-90 87.90 -0.12 82.60 -2.53 HPV-90 87.90 -0.12 82.60 -2.53 HPV-90 87.90 -0.12 82.60 -3.38 HPV-90 86.40									
Here Here structure cannot be predicted HPV-48 86.50 -1.14 82.70 -1.86 HPV-49 87.40 -1.12 88.50 -3.44 HPV-50 90.00 -0.73 87.10 -2.77 HPV-51 89.40 -1.69 86.30 -2.87 HPV-50 89.50 -0.6 87.80 -3.75 HPV-61 89.20 -2.58 87.10 -2.51 HPV-63 88.80 -2.09 93.40 -2.5 HPV-71 87.90 -1.33 94.00 -2.72 HPV-88 89.20 -1.44 95.30 -3.01 HPV-90 87.90 -1.45 93.00 -3.51 HPV-92 78.00 -1.49 91.00 -2.75 HPV-98 86.40 -1.49 81.20 -3.38 HPV-98 86.40 -1.49 81.20 -3.38 HPV-99 87.90 -0.12 82.60 -2.59 HPV-100 84.10				Percentage identity is least					
$\begin{array}{c c c c c c c c c c c c c c c c c c c $									
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$									
$\begin{array}{c c c c c c c c c c c c c c c c c c c $									
$\begin{array}{c c c c c c c c c c c c c c c c c c c $		90.00			-2.77				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-53	89.40	-1.69	86.30	-2.87				
HPV-61 89.20 -2.58 87.10 -2.51 HPV-63 88.80 -2.09 93.40 -2.5 HPV-71 87.90 -1.33 94.00 -2.72 HPV-88 89.20 -1.44 95.30 -3.01 HPV-90 87.90 -1.45 93.00 -3.51 HPV-92 78.00 -1.49 91.00 -2.75 HPV-96 85.70 -0.12 82.60 -2.53 HPV-98 86.40 -1.49 81.20 -3.38 HPV-99 87.90 -0.93 83.50 -2.59 HPV-100 84.10 -0.8 90.70 -3.76 HPV-101 Not available 89.80 -3.48 HPV-103 Not available 89.80 -3.41 HPV-104 85.90 -0.1 93.00 -3.41 HPV-105 86.40 0.44 89.80 -3.11 HPV-107 89.20 -1.86 88.40 -2.56 HPV-108 N	HPV-54	85.40	-1.64	82.40	-2.78				
HPV-63 88.80 -2.09 93.40 -2.5 HPV-71 87.90 -1.33 94.00 -2.72 HPV-88 89.20 -1.44 95.30 -3.01 HPV-90 87.90 -1.45 93.00 -3.51 HPV-90 87.90 -1.49 91.00 -2.75 HPV-96 85.70 -0.12 82.60 -2.53 HPV-96 85.70 -0.12 82.60 -2.53 HPV-98 86.40 -1.49 81.20 -3.38 HPV-99 87.90 -0.93 83.50 -2.59 HPV-100 84.10 -0.8 90.70 -3.76 HPV-101Not available 86.00 -1.91 HPV-103Not available 89.80 -3.41 HPV-104 85.90 -0.1 93.00 -3.41 HPV-105 86.40 0.44 89.80 -3.11 HPV-107 89.20 -1.86 88.40 -2.56 HPV-108Not available 95.20 -2.79 HPV-109 83.80 0.52 78.00 -1.97 HPV-112 82.70 -1.42 90.70 -2.23 HPV-116 77.70 0.17 89.90 -2.71 HPV-121 89.20 -0.33 89.70 -2.77 HPV-122 83.50 -0.17 92.90 -2.44 HPV-131 85.50 -0.17 92.90 -2.44 HPV-132 90.20 0.13 80.00 -2.27 HPV-134 89.80	HPV-60	89.50	-0.6	87.80	-3.75				
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	HPV-61	89.20	-2.58	87.10	-2.51				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-63	88.80	-2.09	93.40	-2.5				
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	HPV-71	87.90	-1.33	94.00	-2.72				
HPV-9278.00-1.4991.00-2.75HPV-9685.70-0.1282.60-2.53HPV-9886.40-1.4981.20-3.38HPV-9987.90-0.9383.50-2.59HPV-10084.10-0.890.70-3.76HPV-101Not available86.00-1.91HPV-103Not available89.80-3.48HPV-10485.90-0.193.00-3.41HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-12189.20-0.3389.70-2.71HPV-12884.70-1.2288.60-3.83HPV-12983.500.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-14882.20-1.4485.50-2.29HPV-14882.20-1.4485.50-2.29	HPV-88	89.20	-1.44	95.30	-3.01				
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	HPV-90	87.90	-1.45	93.00	-3.51				
HPV-9886.40-1.4981.20-3.38HPV-9987.90-0.9383.50-2.59HPV-10084.10-0.890.70-3.76HPV-101Not available86.00-1.91HPV-103Not available89.80-3.48HPV-10485.90-0.193.00-3.41HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-14882.20-1.4485.50-2.29HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24	HPV-92	78.00	-1.49	91.00	-2.75				
HPV-9987.90-0.9383.50-2.59HPV-10084.10-0.890.70-3.76HPV-101Not available86.00-1.91HPV-103Not available89.80-3.48HPV-10485.90-0.193.00-3.41HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-14882.20-1.4485.50-2.29HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24	HPV-96	85.70	-0.12	82.60	-2.53				
HPV-10084.10-0.890.70-3.76HPV-101Not available86.00-1.91HPV-103Not available89.80-3.48HPV-10485.90-0.193.00-3.41HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24	HPV-98	86.40	-1.49	81.20	-3.38				
HPV-101Not available86.00-1.91HPV-103Not available89.80-3.48HPV-10485.90-0.193.00-3.41HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24	HPV-99	87.90	-0.93	83.50	-2.59				
HPV-103Not available89.80-3.48HPV-10485.90-0.193.00-3.41HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24	HPV-100	84.10	-0.8	90.70	-3.76				
HPV-10485.90-0.193.00-3.41HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24		Not available		86.00	-1.91				
HPV-10586.400.4489.80-3.11HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24					-3.48				
HPV-10789.20-1.8688.40-2.56HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24		85.90	-0.1	93.00	-3.41				
HPV-108Not available95.20-2.79HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24			0.44						
HPV-10983.800.5278.00-1.97HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24			-1.86						
HPV-11282.70-1.4290.70-2.23HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-11386.90-1.0491.20-3.83HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-11677.700.1789.90-2.71HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24			-1.42						
HPV-12189.20-0.3389.70-2.77HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-12884.70-1.2288.60-3.83HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-12983.501.1787.20-2.84HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-13185.50-0.1792.90-2.44HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-13290.200.1380.00-2.27HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-13489.80-0.987.50-3.96HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-14882.20-1.4485.50-2.29HPV-KI88-0386.701.1987.20-3.24									
HPV-KI88-03 86.70 1.19 87.20 -3.24									
		82.20			-2.29				
HPV-RTRX7 89.40 -0.15 92.20 -3.4			1.19		-3.24				
	HPV-RTRX7	89.40	-0.15	92.20	-3.4				

Table 2. PROCHECK analysis and ProSA-web based validation of E6 & E7 proteins of 50 different strains of HPV

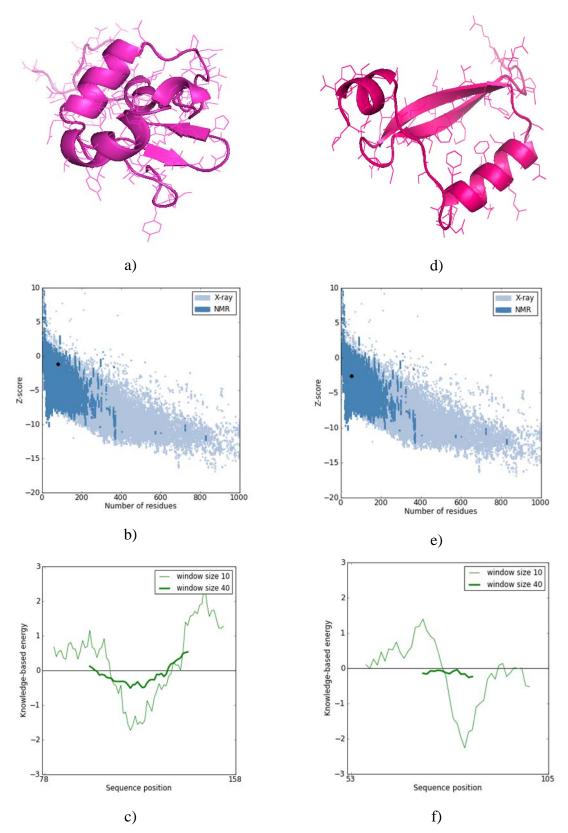


Fig. 1 ProSA-web Z-scores of modeled protein in PDB with respect to their protein length.Z-score is represented in black dot. The energy plots presented with window size 10 & 40.a) 3D structures of E6; b) Z-score plot of E6; c) energy plot of E6;d) 3D structures of E7; e) Z-score plot of E7; f) energy plot of E7.

Conclusion

The variation of causing nongenital cutaneous, nongenital mucosal and anogenital diseases by different HPV types could always remain a challenge to find out the cause behind it. These variations may be in their genomic content leading deviation in their proteomic structures, causing different types of infections as an outcome. As proteins of HPV are directly involved in causing the infection in human, so, it may be of significant interest to explore and analyze their protein structures. In this study, we made an effort to predict the 3D structures of E6 and E7 oncoproteins of 50 different strains of HPV. This study provides simultaneous predicted and validated structures of these HPV proteins. The outcome of this study might provide a platform for simultaneous structural comparative analysis of these proteins and help in finding out variations in their structures so as to explore why different strains of HPV cause different type of cancers. Further, this might also help in exploring for designing specific drugs or vaccines against specific strains of HPV.

Acknowledgement

Authors thank Department of Biotechnology, Ministry of Science & Technology, Government of India for financial support to Bioinformatics Centre wherein this study has been carried out. Authors also thank Shri D. S. Mehta, President, Kasturba Health Society, Dr. B. S. Garg, Dean, Mahatma Gandhi Institute of Medical Sciences, Sevagram and Dr. B. C. Harinath, Director, JBTDRC & Coordinator, BIC for their encouragement.

References

- 1. Altschul S. F., W. Gish, W. Miller, E. W. Myers, D. J. Lipman (1990). Basic Local Alignment Search Tool, J Mol Biol, 215, 403-410.
- 2. Brooks W. H., K. G. Daniel, S. S. Sung, W. C. Guida (2008). Computational Validation of the Importance of Absolute Stereochemistry in Virtual Screening, J Chem Inf Model, 48, 639-645.
- 3. Contreras A., I. Martínez, E. Cruz, M. Lizano (2007). Role of HPV18 E6 in PKB Signal Transduction Pathways, BMC Cancer, 7(1), A7.
- Friesner R. A., J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, J Med Chem, 47(7), 1739-1749.
- Gasteiger E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch (2005). Protein Identification and Analysis Tools on the ExPASy Server. In: John M. Walker (Ed.), The Proteomics Protocols Handbook, Humana Press, 571-607.
- 6. Hiller T., S. Poppelreuther, F. Stubenrauch, T. Iftner (2006). Comparative Analysis of 19 Genital Human Papillomavirus Types with Regard to p53 Degradation, Immortalization, Phylogeny, and Epidemiologic Risk Classification, Cancer Epidem Biomark Prevent, 15, 1262-1267.
- 7. Kaladhar D. S. V. G. K., T. Uma Devi, P. V. N. Rao (2010). An *in silico* Genome Wide Identification, Characterization and Modeling of Human Papilloma Virus strain 92, Int J Eng Sci Techn, 2(9), 4288-4291.
- Laskowski R. A., M. W. MacArthur, D. S. Moss, J. M. Thornton (1993). PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures, J Appl Cryst, 26, 283-291.
- 9. Lowy D. R., D. Solomon, A. Hildesheim, J. T. Schiller, M. Schiffman (2008). Human papillomavirus Infection and the Primary and Secondary prEvention of Cervical Cancer, Cancer, 113(7), 1980-1993.

- 10. Mayers G., E. Androphy (1995). The E6 Protein, Human Papillomaviruses, 1995 Compendium, Part III, 47-57.
- 11. Munger K., A. L. Halpern (1997). HPV16 E7: Primary structure and biological properties. In Human Papillomaviruses 1997 Compendium, Part III, 17-36.
- 12. Nelson P. H., K. H. Vousden, N. L. Hubbert, D. R. Lowy, J. T. Schiller (1989). HPV16 E6 and E7 Proteins Cooperate to Immortalize Human Foreskin Keratinocytes, The EMBO Journal, 8, 3905-3910.
- 13. Plewczynski D., M. Hoffmann, M. von Grotthuss, K. Ginalski, L. Rychewski (2007). *In silico* Prediction of SARS Protease Inhibitors by Virtual High Throughput Screening, Chem Biol Drug Des, 69, 269-279.
- Rosty C., M. Sheffer, D. Tsafrir, N. Stransky, I. Tsafrir, M. Peter, P. Rochefordière, R. Salmon, T. Dorval, J. Thiery, J. Couturier, F. Radvanyi, E. Domany, X. Sastre-Garau (2005). Identification of a Proliferation Gene Cluster Associated with HPV E6/E7 Expression Level and Viral DNA Load in Invasive Cervical Carcinoma, Oncogene, 24(47), 7094-7104.
- 15. Sali A., T. L. Blundell (1993). Comparative Modelling by Statisfaction of Spatial Restraints, J Mol Biol, 234, 779-815.
- 16. Sayle R. A., E. J. Milner-White (1995). RASMOL: Biomolecular Graphics for All, Trends Biochem Sci, 20, 374-376.
- 17. Tungteakkhun S. S., M. Filippova, J. W. Neidigh, N. Fodor, P. J. Duerksen-Hughes (2008). The Interaction between Human Papillomavirus Type 16 and FADD is Mediated by a Novel E6 Binding Domain, J Virol, 82, 9600-9614.
- Watts K. J., C. H. Thompson, T. E. Cossart, B. R. Rose (2001). Variable Oncogene Promoter Activity of Human Papillomavirus Type 16 Cervical Cancer Isolates from Australia, J Clin Microbiol, 35(5), 2009-2014.
- Wiederstein M., M. J. Sippl (2007). ProSA-web: Inter-active Web Service for the Recognition of Errors in Three-dimensional Structures of Proteins, Nucleic Acids Res, 35, 407-410.
- Xuemei J. I., E. M. Sturgis, C. Zhao, C. J. Etzel, Q. Wei, G. Li (2009). Association of p73G4C14-to-A4T14 Polymorphism with Human Papillomavirus Type 16 Status in Squamous Cell Carcinoma of the Head and Neck in Non-Hispanic whites, Cancer, 115(8), 1660-1668.
- 21. Zheng Z. M., C. C. Baker (2006). Papillomavirus Genome Structure, Expression, and Post-transcriptional Regulation, Front Biosci, 11, 2286-2302.



Prof. (Dr) Satish Kumar, M.D. E-mail: satishkumar@mgims.ac.in

MBBS from GMC, Jammu in 1990 and M.D. (Biochemistry) from MGIMS, Sevagram in 1994. Presently continuing as Professor, Biochemistry & Dy Coordinator, Bioinformatics Centre, MGIMS, Sevagram, India. Member, Board of Studies in Biotechnology & Bioinformatics, MUHS, Nashik; Member, Scientific Advisory Committee (SAC), ICMR-NIC National Database on Indian Medical Journals; Member, Faculty of Medicine, IMS, Banaras Hindu University, Varanasi; Editor-in-Chief of Journal of Pathology Research and Journal of BIOINFO Medicine; Member & Referee, National Advisory Board, JK Science, Journal of Medical Education & Research; Principal Investigator on CSIR funded two Projects on Immunodiagnostics in tuberculosis; Patent on M tb ES-31 isolation (184510/410/BOM/99/2001); Coordinator, STP in process Bioinformatics & Biotechnology; Introduced SEVA TB ELISA in 1993 for TB detection in patients attending our hospital; Continuing research work in immunodiagnostics in tuberculosis and in silico cancer work; Co-Investigator on DBT funded project on BTISnet; Associate Editor & Publisher, SEVAMED Quarterly Journal; Involved in developing clinical databases; Establishing Moving Academy on Biomedical Communication; Convener, Online Health Informatics Course (OHIC); Executive Member, Bioinformatics Centre, Raipur; Organized 14 workshops/CMEs on Medical Informatics & Biomedical Communication; Recognized UG, PG and PhD teacher of MUHS, Nashik; On panel of UG, PG & PhD examination under many universities; Published 46 papers (7 International & 39 National); Served on many Institutional committees (HIS, NAAC, College Council etc) with an indomitable commitment.

Lingaraja Jena, M.Sc., Ph.D. E-mail: <u>lingaraj_jena@yahoo.co.in</u>



University Gold Medalist in M.Sc. Bioinformatics from Orissa University of Agriculture & Technology, Bhubaneswar, Odisha, India in 2006. Presently working as Information Officer, Bioinformatics Centre, MGIMS, Sevagram, India. Area of interest are Development of Biological Database & Management, Programming, Bioinformatics Tools application, Computational Biology, Sequence analysis etc. Continuing work in the area of sequence and structure analysis of Mycobacterium tuberculosis & HPV. Published 3 Research Papers (2 National and 1 International).

Vidya W. Bhomale, M.Sc. E-mail: vidyabomale@gmail.com



University Gold Medalist in B.Sc. and College Topper in MSc Biotechnology from RSTM University, Nagpur in 2011. Carried out Project Work on Structural Comparison of E6 and E7 Proteins of different strains of HPV using Bioinformatics tools at Biochemistry, MGIMS, Sevagram. Undergone training in "Recent Agricultural Biotechnology" at Anand Niketan College of Agriculture, Anandwan, Warora. Areas of interest are Molecular biology, Genetic engineering, Pathogens & viruses microbiological studies, Plant biotechnology & breeding. Presently working in the area of protein structure analysis of different strains of HPV.

Prof. (Dr) Sangeeta Daf, M.D.

E-mail: sangeetadaf@gmail.com

MBBS and M.D. (Obstetrics & Gynecology) from MGIMS, Sevagram in 1994 and 2001 respectively. Presently continuing as Professor & HOU, Obstetrics & Gynaecology, Datta Meghe Institute of Medical Sciences (Deemed University), Nagpur - DAGA Memorial Women & Child Care Hospital, Nagpur and Chief OBGY Consultant, Mother & Child Care and Life Care Hospitals, Nagpur, India. Recipient of International Training on Sexual & Reproductive Health & Rights at Uppsala University, SWEDEN & COLOMBO; NTTC for Health Professionals at JIPMER, Pondicherry, India; Master Trainer of Emergency Obstetrics Care (EmOC), International joint venture of JEPHIGO & FOGSI. Able-bodied Skilled in Obstetric Emergencies, Infertility, Gynecological Cancer, Endoscopy, Sexual & Reproductive Health etc (main areas of propel). Held responsibility of Secretary and President, SWOGS (under FOGSI) in successive years. Recognized Teacher by Maharashtra University of Health Sciences, Nashik for Teaching, Supervising and Examining students pursuing M.D. Obstetrics & Gynecology and MBBS Courses. Published 8 Papers in National & International Journals, Delivered talks and also Chaired sessions in more than 20 National & International Conferences/Symposia/CME.

