Original Article

Insights in the interaction of HPV-16 E6 and its variants (E-G350, E-A176/G350, E-C188/G350, AAa and AAc) with MAGI-1

Lilian Esmeralda Araujo-Arcos[1], Sarita Montaño[2],* Ciresthel Bello-Rios[1], Olga Lilia Garibay-Cerdenares[1,3], Marco Antonio Leyva-Vázquez[1] and Berenice Illades-Aguiar[1,*]

[1] Laboratorio de Biomedicina Molecular, Facultad de Ciencias Químico-Biológicas, Universidad Autonóma de Guerrero, Chilpancingo, CP 39090, Mexico; earaujo@uagro.mx (L.E.A.-A.); ciresthelbello@uagro.mx (C.B.-R.); olgaribayce@conacyt.mx (O.L.G.-C.); leyvamarco13@gmail.com (M.A.L.-V.).

[2] Laboratorio de Bioinformática y Simulación Molecular, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Sinaloa, Culiacán Sinaloa, CP 80030, Mexico.

[3] CONACyT-Universidad Autónoma de Guerrero, Chilpancingo, CP 39087, Mexico

* Correspondence: mmontano@uas.edu.mx (S.M.); b.illadesaguiar@gmail.com (B.I.-A.); Tel.: +52-747-499-15-68 (B.I.-A.)

Abstract: Oncogenic protein E6 from Human Papilloma Virus 16 (HPV-16) mediates the degradation of Membrane-associated guanylate kinase with inverted domain structure-1 (MAGI-1), throughout the interaction of its protein binding motif (PBM) with the Discs-large homologous regions 1 (PDZ1) domain of MAG1-1. Generic variation in the E6 gene that translates to changes in the protein's amino acidic sequence modifies the interaction of E6 with the cellular protein MAGI-1. MAGI-1 is a scaffolding protein found at tight junctions of epithelial cells, where it interacts with a variety of proteins regulating signaling pathways. MAGI-1 is a multidomain protein containing two WW (rsp-domain-9), one guanylate kinase-like, and six PDZ domains. PDZ domains played an important role in the function of MAGI-1 and served as targets for several viral proteins including the HPV-16 E6. The aim of this work was to evaluate, with an in silico approach, employing molecular dynamics simulation and protein-protein docking, the interaction of the intragenic variants E-G350 (L83V), E-C188/G350 (E29Q/L83V), E-A176/G350 (D25N/L83V), E6-AAa (Q14H/H78Y/83V) y E6-AAc (Q14H/I27RH78Y/L83V) and E6-reference of HPV-16 with MAGI-1. We found that variants E-G350, E-C188/G350, E-A176/G350, AAa and AAc increase their affinity to our two models of MAGI-1 compared to E6-reference.

Keywords: E6; variants; MAGI-1; molecular dynamic simulation; docking.

1. Introduction

High-risk human papillomaviruses (HR-HPV) are the principal etiological agents of cervical cancer (CC), being the HPV-16 genotype one of the most prevalent worldwide [1]. The encoding proteins E6 and E7 from HPV-16 are the major oncogenic determinants of the disease's progression. These proteins control regulatory functions of the cell cycle, promote proliferation, induce malignant transformation, and facilitate migration and invasion of transformed cells [2].

E6 protein contains 151 amino acids, with a molecular weight of approximately 19 kDa. Structurally, E6 has two zinc fingers, and a LXXLL domain which is vital for its oncogenic potential. Moreover, an important motif of E6 is the PDZ binding motif (PBM), composed of the residues: E148, T149, Q150, and L151 located at the carboxyl terminus of the protein [3]. Any variability in the amino acid sequence in the PBM motif or in neighbouring regions of E6 could modify the

degradation of its target cell proteins [4]. The ability of HPV-16 E6 to retain protein binding with cellular proteins with PDZ domains has been shown to contribute to its transforming activity *in vivo* and *in vitro* systems [5, 6].

Interestingly, the tumorigenic potential of HPV-16 differs among infected women, and it has been proposed that the changes in the amino acidic sequence of E6 are a major risk factor for the development and aggressiveness of the disease [7]. The variants E-G350 (L83V), E-C188/G350 (E29Q/L83V), E-A176/G350 (D25N/L83V), AAa (D25N/L83V), and AAc (Q14H/I27R/H78Y/L83V) are the most prevalent in a population from Guerrero, Mexico, a state with the highest poverty and marginalization rates in the country [8]. World wide these variants were named after their nucleotide changes on the gene E6 and according to the geographical region where they were first isolated [9]. Experimental studies suggested that variants differ in their ability to affect several important cellular processes, including differentiation, apoptosis, immortalization, migration and metastasis [10-12].

Especially, the non-European variants of HPV-16, AAa, AAc, whose sequence shows four mutations (G145T, G188C, C335T and E350G) on the gene sequence resulting in changes of a single amino acid in comparison to E6-reference at positions Q14H, I27R, H78Y, and L83V [13, 14]. It is argued that these differences have mechanistic and functional consequences for the infection promoting viral persistence and increasing the disease progression, aggressiveness and, therefore, the risk of developing cancer [7]. Little is known about the functional impact of variants E-G350, E-C188/G350, E-A176/G350 on the cell´s regulatory systems, but it has been observed that they can be involved in the resistant of apoptosis, migration and invasion by interacting with proteins like p53, Bax and enhancing MAPK signaling, more information is needed to determinate the molecular mechanisms by which the variants alter the virus oncogenic potential [10, 15, 16].

Previously, our group reported that the variants of E6 HPV-16 (E-G350, E-A176/G350, E-C188/G350, AAa and AAc) were the most common and had the most oncogenic potential in a population of southern Mexico [6]. With these results, our group analyzed the effects of the expression of variants AAa, AAc, EA176/G350, E-C188/ G350, E-G350 using an *in vitro* model in C33-A cells transfected with each variant, showing that all variants alter the expression of 431 genes compared to the E6-reference, these genes are involved in cellular processes related to adhesion, angiogenesis, apoptosis, differentiation, cell cycle, proliferation, transcription and protein translation, been MAGI-1´s expression down-regulated on our experimental data [17]. A target protein of HR-HPV E6 is the membrane-associated guanylate kinase with inverted domain structure-1 (MAGI-1), a scaffolding protein found in tight junctions of epithelial cells, where it interacts with a variety of proteins which include β-catenin, Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase (PTEN) and α-actin, these make it a key regulator of signalling pathways at cell−cell junctions [18]. MAGI-1 is a multidomain protein composed of two domains WW (rsp-domain-9) (G300-C333 and L359-L392) a protein-protein binding domain that mediates specific interactions with short proline-rich or proline-containing motifs, one guanylate kinase-like domain, and six PDZ domains located at E17-G105, H472-R55, T643-R721, S813-P895, S970-S1066 and E1124-T1206, that are composed of approximately 80–110 residues, present in the C-terminal portions of signalling proteins [19, 20]. MAGI-1 domains play important roles in protein-protein interaction, especially PDZ domains, which play a role in localizing proteins to the membrane and acting as molecular scaffoldings or adaptors; also these domains serve as targets for several viral proteins [21].

MAGI-1 PDZ1 domain has been identified as a major target of E6 from HPV-16 [22-24]. *In vitro* experiments had proposed that the degradation of MAGI-1 is mediated by the direct interaction of its PDZ-1 (H472-R55) domain with the PBM motif from E6 (E148-L151) [25]. The adjacent amino acids of E6 that may play a key role in the interaction of these proteins are: C103- I104, R135,-C136, C139- S140, S82, G85, L88, S97, N105, R124- F125 and N127- I128 [4, 26]. Since the experiments have been carried out using small portions of the protein (K456-E587), their interaction could be modified

if more protein regions were added. The study of protein-protein interactions remains a challenge these days due to the high cost of production and purification of recombinant proteins and the technical difficulties to crystallize them. The 1491 amino acids of MAGI-1 made it difficult to obtain it´s 3D structure. The inherent technical difficulties to solve the complex structure of proteins, led to the call for integrating complementary computational approaches. We decided to delimit our MAGI-1 models to experimentally proven domains to interact with E6 [27].

Recently, our team has adopted an *in-silico* analysis approach to evaluate the structural changes of E6 and its variants [28, 29]. Since, evidence of the role of the PBM motive of E6 in the progression and aggressiveness of CC has been published [4, 26], and the possible changes in the behaviour of the AA variants and the other variants are not clear. We propose an *in silico* approach to get insights into the interactions of the E6-reference oncoprotein and its variants (E-G350, E-A176/G350, E-C188/G350, AAa and AAc), against MAGI-1 a cellular target protein.

2. Results and Discussion

HPV-16 HPV-16 is accounting for more than 70% of the CC cases [1], it has been well established that the oncoproteins E6 and E7 are responsible for the onset and aggressiveness of the disease. Of these two proteins E6, dysregulates the cell cycle, promotes hipper proliferation, induces malignant transformation, and facilitates migration and invasion of transformed cells in *in vivo* and *in vitro* assays [2]. Also, it has been proposed that the difference in the oncogenic potential of this virus is mediated by the genetic variation that occurs in the E6 gene, which alters the thermodynamics and structural stability on the 3D structure of the protein. However, the molecular insights of the changes remain unknown [7].

In order to get insights into the interaction of the five variants of E6 from HPV-16 with the cellular protein MAGI-1 were adopted an *in silico* approach. Molecular Dinamics (MD) simulations and docking analyses were performed to explore the disruptive effect of mutations in five variants of E6 from HPV-16 and their consequences on the interactions with MAGI-1 255 and 329.

2.1 3D protein structures

Multiple alignments of the sequence of E6 and its variants were performed to evidence amino acidic changes between them. E6 mutations Q14H, D25N, I27R, E29Q and H78Y are found in a non-domain region adjacent to the zinc finger domain 1. While E6 mutations L83V and H78Y are located in an interdomain region between the two zinc finger domains. The 2D structures of the E6-reference and the variants were predicted; the results show that the secondary structure did not change between the variants and E6-reference (Figure 1A).

The crystal structure of E6 present five alfa-helix and four beta-sheets. Also, it contains two zinc molecules forming two finger domains located at C30, C33, C63 C66, and C103, C106, C136, C139 residues (Figure1B). After mutating E6-reference to obtained variants (E-G350, E-A176/G350, E-C188/G350, AAa and AAc) we conducted an structural alignment was using the VMD RMSD tool (Figure 1B). The RMSD values less than two Å represents accurate models. The RMSD values obtained for the variants were 0.51Å for E-G350 (green), 0.66Å for E-C188/G350 (yellow), 0.18 Å for E-A176/G350 (blue), 0.44 Å for AAa (red), and 0.69 Å for AAc (orange) which means that the mutations did not alter the proteins 3D structure, other than the punctual mutations sites (Figure 1B). All models were evaluated by Ramachandran plot showing for all variants that 91.9% of the amino acids fall in the favored region and 8.1% in the allowed region, which means a good stereochemistry for 100% of the residues (Figure S1).
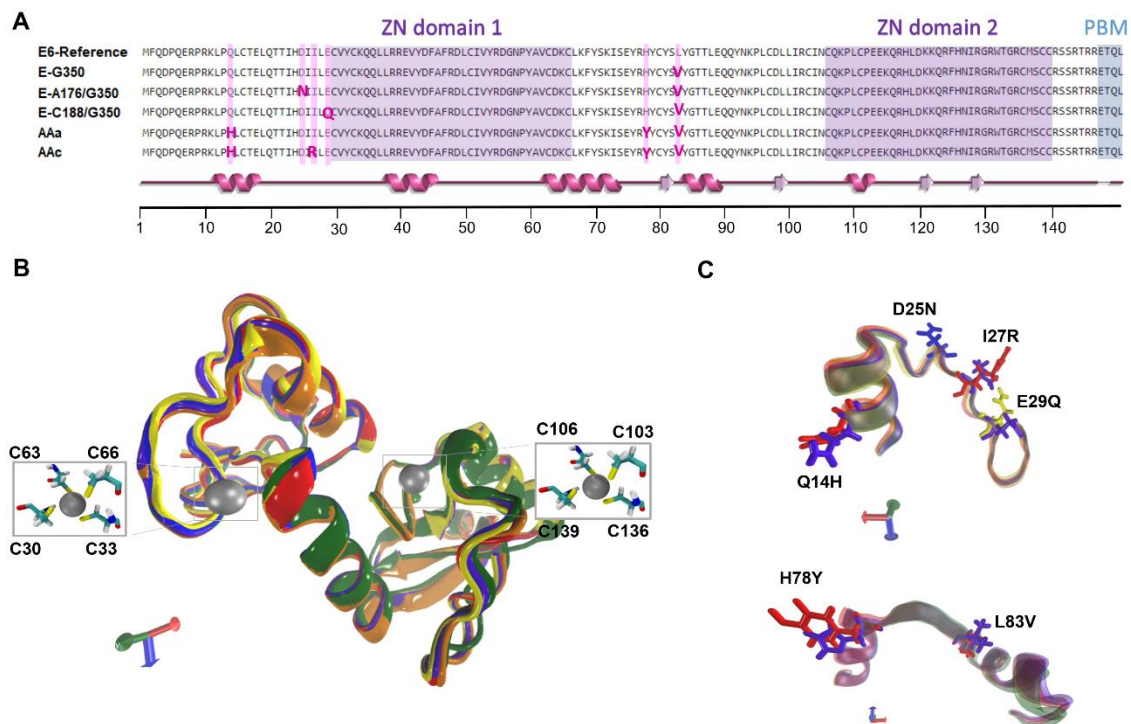
Figure 1. Alingment and super position of the 3D structures of the E6-reference and its variants. (A) Multiple alignment of the sequence of E6 and its variants. Zinc finger domain 1 and 2 are highlighted in purple, PBM in blue and mutations are highlighted in pink. A super position of the secondary structure of all six proteins is shown in pink below the alignment. (B) 3D structure of E6-reference: violet, AAa variant: orange, AAc C variant: red, E-G350, green, E-C188/G350: yellow, E-A176/G350: blue. The silver spheres indicate zinc molecules and the licorice residues correspond to C30, C33, C63, C66, C103, C106, C136 and C139, which make up two zinc finger domains in the proteins 3D structures. The orientation of the proteins is indicated by the axes, X: red, Y: green; Z: blue. (C) Visualization of amino acid changes: Q14H, D25N, I27R, H78Y and L83V are in licorice.

A close up of the mutated residues in the alignment shows that the side chains of mutants H14 and Y78 were exposed to the protein's surface while the amino acids Q14 and H78 side chains of E6-reference were inverted (Figure 1B). According to the program HOPE, which analyses the structural and functional effects of point mutations, H14 is bigger than Q14, bigger residues might lead to bumps on the 3D structure of the protein. H14 is among the observed mutations at this position in other homologous sequences. This sometimes suggests that the mutant is not damaging for the protein's structure and function, on the other hand, the residue is located near a highly conserved position and can gain interactions with target proteins. The Y78 is bigger and more hydrophobic than the H78. This can result in loss of hydrogen bonds and disturbance of correct folding. The accessibility of the residues in the mutants could increase the number of interactions with the MAGI-1 models [30].

Mutations E29Q and D25N remained the same size; therefore, not visible change in 3D structure was observed (Figure 1B). A change in residue charge from negative to neutral can cause loss of interactions with other molecules or residues [30]. Mutant R27 is bigger than I27, this can be observed in the 3D structure comparison (Figure 1B). Also, the change of a neutral to positive residue leads to the possibility of repulsion of ligands or other residues with the same charge. Moreover, the hydrophobicity of the wild type residue is lost which will lead to the loss of hydrophobic interactions [30]. The mutation L83V found in all the variants can cause the proteins to lose interactions with other proteins because, V83 is a smaller than L83 (Figure 1B), loss of

interactions with cellular target proteins could lead to a diminution on the affinity of interactions for the variants [30].

There are many crystal structures in the RCSB PDB server related to the crystalized MAGI-1; however, none of these files corresponds to the complete protein structure. Moreover, there are no reliable software for homology modelling such large proteins [27]. To overcome this drawback, we delimitated our models to domains that have been experimentally shown to interact with E6 [21, 24, 31]. Our first model included the WW1, WW2, PDZ1 domains and was denominated MAGI-1 255 and a second model where we added a highly disordered region of 76 amino acids adjacent to the PDZ1 domain denominated it MAGI-1 329. A sequence alignment of our final models is shown in figure 2A.
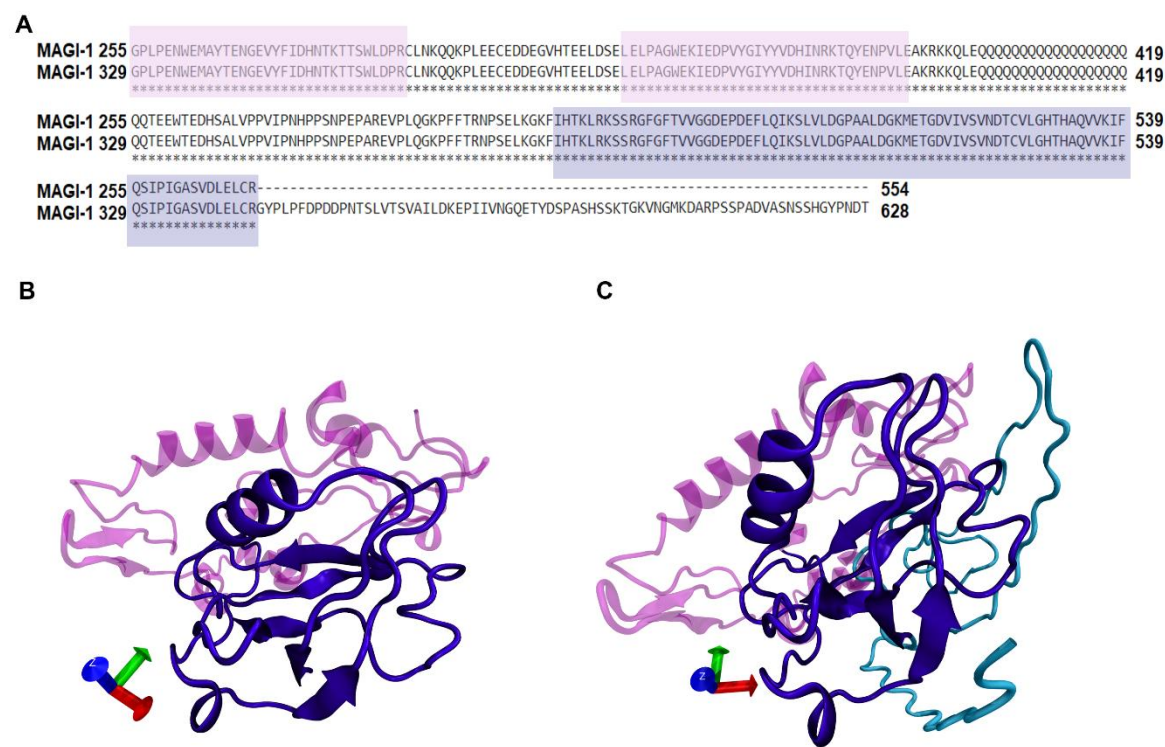

Figure 2. 3D structure of MAGI-1 255 and MAGI-1-329 models. (A) Multiple alignment of the sequence of model MAGI-1 255 and MAGI-1 329 highlighting domains WW1 and WW2 in soft pink and domain PDZ1 in soft purple. (B) 3D model visualization of MAGI-1 255 (C) 3D model visualization of of MAGI-1 329. The WW1 and the WW2 domains are shown in light pink, the PDZ1 domain is shown in violet, and a highly disordered region of 76 amino acids in cyan.

The 3D homology models of MAGI-1 255 (amino acid 300 to 554) and of MAGI-1 329 (amino acid 300 to 628) were obtained using the crystal structure from human MAGI-1 PDZ1 (PDB ID:2KPK) as a template on the I-TASSER server (Figure 2B and 2C). The best models were chosen due to the criteria of good alignment with the template measured by C-Score, TM score, and RMSD values. Model MAGI-1 255 shown in Figure 2B consists of six alfa-helix, one 310 helix, eleven beta-sheets and the rest of the residues appeared lightly twisted in random coils, which are 15.3%, 2.4%, 13.7% and 68.6% respectively of the protein structure. The two WW domains and the interdomain regions from amino acid G300 to amino acid I471 are shown in magenta, and the PDZ1 domain from amino acid H472 to amino acid R554 is shown in purple (Figure 2B). Model MAGI-1 329 consists of seven alfa-helixes, twelve beta-sheets, and the rest of the residues appeared in loops and coils, which are 15%, 14.1%, and 70.9%, respectively the protein structure. For this model, the 3D structure from amino acid G300 to R554 was coloured in purple, and the extra region of 76 (G555 to

T628) amino acids were colored cyan (figure 2C). According to dynamic studies using complementary isothermal titration calorimetry and nuclear magnetic resonance (NMR), the interaction of E6 with MAGI-1 occurs mainly with the PDZ-1 (H472-R554) domain, but different affinity patterns were observed with adjacent regions [4]. To evaluate if the interaction of E6 with our models is modified by adjacent amino acids, we decided to add a highly disordered region of 76 amino acids (G555-T628) to model MAGI-1 329 (Figure 2C)

Ramachandran plot for model MAGI-1 255 and MAGI-1 329 exhibited 92.8% and 91.4% respectively of residues in most favored regions and 7.2% and 8.6% respectively residues are in disallowed regions, which shows a good stereochemistry for more than 90% of the residues, this makes our models acceptable for more refinement with MD simulation (Figure S2).

2.2 Molecular dynamics simulation analysis

To examine the change in the protein dynamics and stability, the 3D models of HPV-16 E6 and its variants, as well as MAGI-1 255 and MAGI-1 329 were refined by MD simulation for 200 ns. Trajectories were analyzed by calculating the root mean square deviation of atomic positions (RMSD), root mean square fluctuation (RMSF), the radius of gyration (Rg), the dPCA analysis and dPCA based clustering (Figure 3).

After 200 ns of the MD simulation and using a snapshot of the most populated cluster of E6 and variants, a structural alignment was done (Figure 3A). The carboxyl terminus of the E-C188/G350, AAa and AAc proteins showed a greater difference compared to E6-reference, while the other variants 3D structures remained very similar compare to E6-reference (Figure 3A).

The RMSD calculation of the E6-reference (purple) during the 200 ns of MD simulation, reached equilibrium at 20 ns of trajectory, while the non-European variants AAa (orange) and AAc (red) were equilibrated at 80 ns, after 150 ns AAa variant loses its equilibrium and recovers it at the end of the DM simulation, probably due to its mutations and its context (Figure 3B). Something similar happens to the variant E-C188/G350 (yellow), which reaches equilibrium at 60 ns, and it´s equilibrium is disturbed from the 150ns to the180ns. This behaviour is attributed to mutations in the E29A and L83V positions that directly affect the structure of the protein, causing disturbs an it´s stability (Figure 3B). Concerning E-A176/G350 (blue), the equilibrium was reached at the first 20ns, but a greater disturbance episode it´s observed from 100ns to the end of the trajectory it is also thought that the nature of mutation D25N on the proteins may contribute to this behaviour. It was also observed that for E6-reference, E-G350 (green), and E6-AAc, the RMSD values during the simulation ranged from 2 Å to 5 Å (Figure 3B). While variants AAa, E-A176/G350 and E-C188/G350 were characterized by higher continuous RMSD fluctuations from the 140 ns to the end of the MD simulation (Figure 3B). The RMSD values of simulated proteins indicated their stability and particular behavior and provided a suitable basis for further analysis.

The Rg presents different grades of compactness during the simulation evidencing a less compactable grade at the end of the trajectory, mainly in the variants E-A176/G350, E-C188/G350 and AAa (Figure 3C in yellow and orange). Meanwhile, E6-reference and variants E-G350 and AAc maintain compactness during the simulation (Figure 3C in green and red). This also confirms that point mutations caused structural destabilizing effects leading to the loss of protein compactness in the E-A176/G350, E-C188/G350 and AAa variants.

Since distance deviations from the starting structure may not necessarily reflect mobility of structural elements, RMSF was used to obtain information on flexibility. According to the data graph in figure 3C, there are six maximum fluctuations peaks areas shared by E6-reference and variants: One at M1 to P5, for E6-reference the fluctuation peak was 16 Å, and for the variants, the highest peak corresponds to E-C188/G350 with 27 Å, the rest of the variants fluctuate from 14 to 24 Å of

distance being E-G350 the lowest peak, this region is composed by coils and turns with non-secondary structure (Figure 3D). The second region at L28 to L50 composed of loops and an alpha helix: reached 15 Å for E6-reference and it wasthe highest fluctuation peak. Fluctuation for variants ranges from 20 to 25 Å; clearly the variants had greater fluctuation in these residues, where E-C188/G350 had the greates fluctuation peaks (Figure 3D). This phenomenon is interesting and it's attributed to mutation E29Q exclusive of the E-C188/G350 variant, while the behavior of the other variants is exclusive of their own structural changes caused by their shared and exclusive mutations. The third peak at C51 to L65, in a coil and two beta-sheets, the variants reached 20 to 23 Å of fluctuation peaks, while E6-reference's fluctuation was only 16 Å (Figure 3D). The fourth region of fluctuation at C80 to L110, was composed of loops, one alpha helix and two beta-sheets and reaches 25 Å for variant E-A176/G350 and E6-reference. For variants AAa and AAc the fluctuation distance was 20 Å. While the peak for variant E-G350 was only 14 Å, evidently less flexible than the other variants and E6-reference (Figure 3D). The fifth fluctuation peak at C111 to C140, in two beta sheets, coils and an one alpha helix, for variants AAa, AAc, E-C188/G350, E-A176/G350 had a fluctuation distance of 20 Å . E6-reference, also, reached a distance of 20 Å. On the other hand, the distance of E-G350 reached 15 Å, and it tends to decrease for the rest of the amino acids at the carboxyl terminus (Figure 3D). Finally the residue-based RMSF of the backbone for the E6-reference displayed less flexible residues than the variants (E-G350, E-C188/G350, E-A176/G350, E6-AAa, and E6-AAc), at the carboxyl terminus (145-151) composed mainly by loops (Figure 3D). Interestingly, this region includes the PBM (ETQV) motif of E6, which is important for this oncoprotein interaction with MAGI-1. Since there is a higher fluctuation in variants compares to E6-reference, it can be deduced that mutations change the structural flexibility of the 3D protein structure.
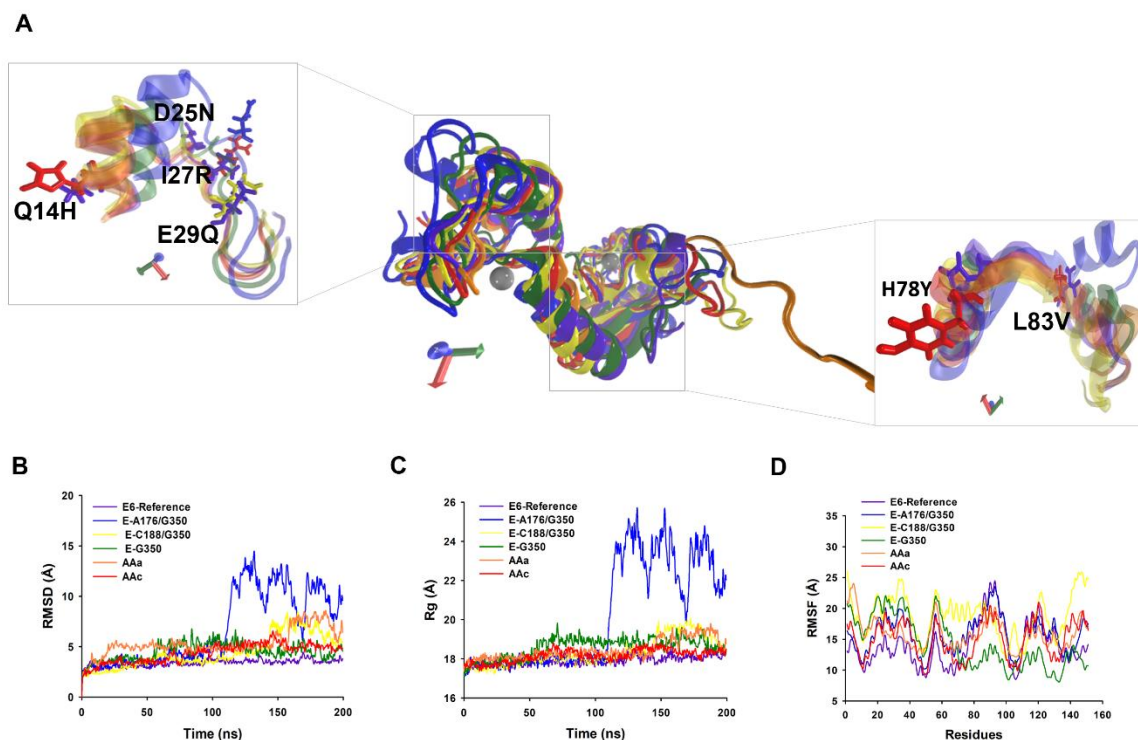

Figure 3. 3D structures of the E6-reference and its variants and conformational stability during 200 ns MD simulation. (A) Super position of average 3D structures of HPV-16 and its variants. Zoom visualization of amino acid changes: Q14H, D25N, I27R, E29Q, H78Y and L83V. E6-reference: violet, AAa variant: orange, AAc variant: red, E-G350, green, E-C188/G350: yellow, E-A176/G350: blue. The silver spheres indicate zinc molecules. The orientation of the proteins is indicated by the axes, X: red, Y: green; Z: blue. (B) RMSD. (C) Radius of gyration. (D) RMSF.

Ramachandran analysis after the MD simulation of these structures shows that more than 98% of the amino acids of the proteins during the simulation remain in the highly favored regions, which means that the protein's conformation are well refined and have native conformations (Figure S3).

For our two MAGI-1 models we showed a snapshot of the most populated cluster from the dPCA clustering analysis in (Figures 4A and B). The PDZ1 domain of our two models have little changes in its 3D structure, but overall it keeps its main secondary structures (Figure 4A and B).

The RMSD of MAGI-1 255 (purple) and MAGI-1 329 (black) models during the 200 ns trajectory showed that both models reached equilibrium before the 100 ns of the simulation and continue stable for the rest of the trajectory. Moreover, MAGI-1 255´s RMSD value after equilibrium ranges from 10 to 13 Å and MAGI-1 329 model´s RMSD values range from 9 to 10 Å, which means our models are reliable for further investigation (Figure 4C). We explored the flexibility of the models by measuring $C\alpha$. The RMSF values of the models through trajectory, mainly 3 regions of MAGI-1 255-model showed more flexible areas, those regions correspond to amino acids G300 to A309 of the WW1 domain, G349 to D379 belong to the WW2 domain and Q399 to H429 that correspond to an interdomain region between the WW2 and PDZ1 domains of MAGI-1 (Figure 4D). MAGI-1 329 model showed three regions with more flexibility which include amino acids E304 to I319 of the WW1 domain, Q399 to V433 corresponding to an inter domain between WW2 and PDZ1 domains and N566 to T628, this region was the most flexible of the two models. Interestingly, this region corresponds to a highly disordered region of the whole protein (Figure 4D) [4]. The Rg show that both models maintained a compacted structure throughout the trajectory of MD simulation simulation (Figure 4E). The Ramachandran analysis shows the refinement of more than 80% of the model´s residues (Figure S4).
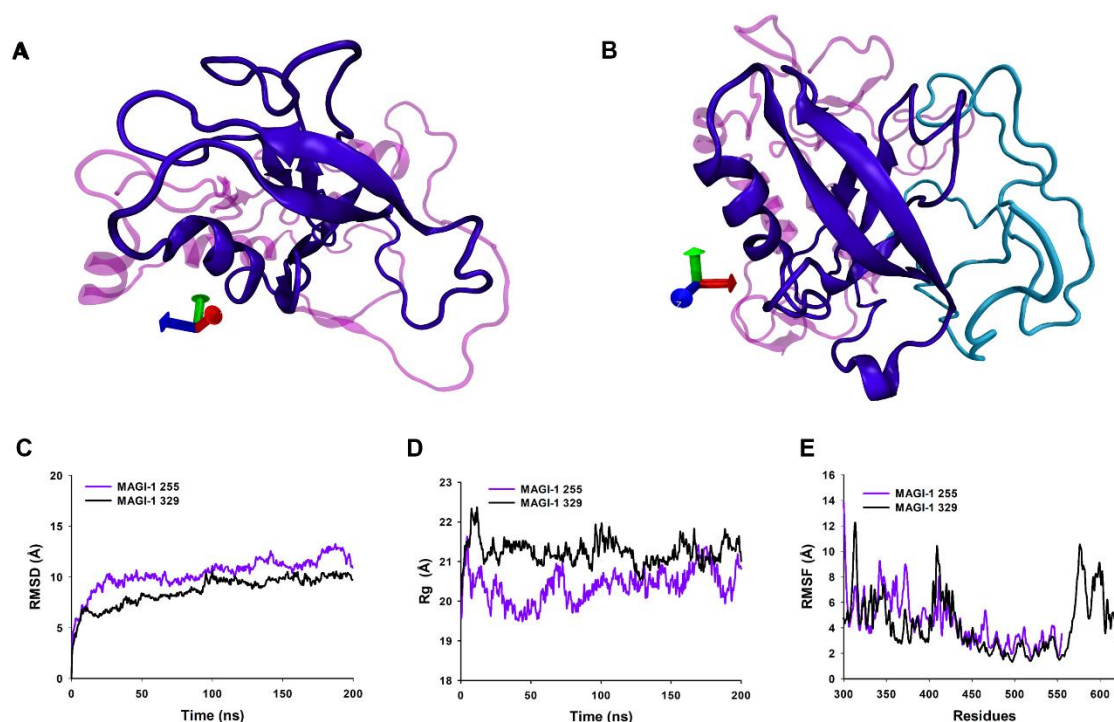

Figure 4. 3D structures of MAGI-1 255 and MAGI-1 329 and conformational stability during 200 ns MDs. (A) Visualization of MAGI-1 255. (B) Visualization of MAGI-1 329. The WW1 and the WW2 domains are shown in light pink, the PDZ1 domain is shown in violet, and a highly disordered region of 76 amino acids in cyan. (C) RMSD. (D) Radius of gyration. (E) RMSF.

2.3 Dihedral Principal component analysis.

dPCA was used to obtain a broader view of dynamic properties with respect to MD simulation simulation results of E6-reference and its variants, MAGI-1 255 and MAGI-1 329. The covariance matrix for the first 20 eigenvectors of E6-reference was 11.40 nm2 and 10.09 nm2, 11.14 nm2, 7.12 nm2, 12.23 nm2 and 10.60 nm2 for the variants E-G350, E-A176/G350, E-C188/G350, AAa and AAc, respectively (Figure 5A). Moreover, the dPCA analysis showed that the first 20 eigenvectors captured 45-57% of the total protein motions (56.7, 45.5, 52.5, 51.1, 55.3 and 53.8%) for E6-reference, E-G350, E-A176/G350, E-C188/G350, AAa and AAc respectively (Figure 5B). Whereas the projections of the first two principal components (PC1 vs PC2) contributed to 15–28% of the collective motions (28.22, 15.11, 24.20, 22.40, 26.66 and 24.59%) for E6-reference E-G350, E-A176/G350, E-C188/G350, AAa and AAc respectively (Figure 5B). These results showed a change in the motion of the atoms of the E-A176/G350, AAa and AAc variants compared to E6-reference and a considerable change in the motion of the atoms of G350 and E-C188/G350 compared to E6-reference, which suggests that the properties of the movements described by the first PCs were different in the six protein systems (Figure 5A and B).
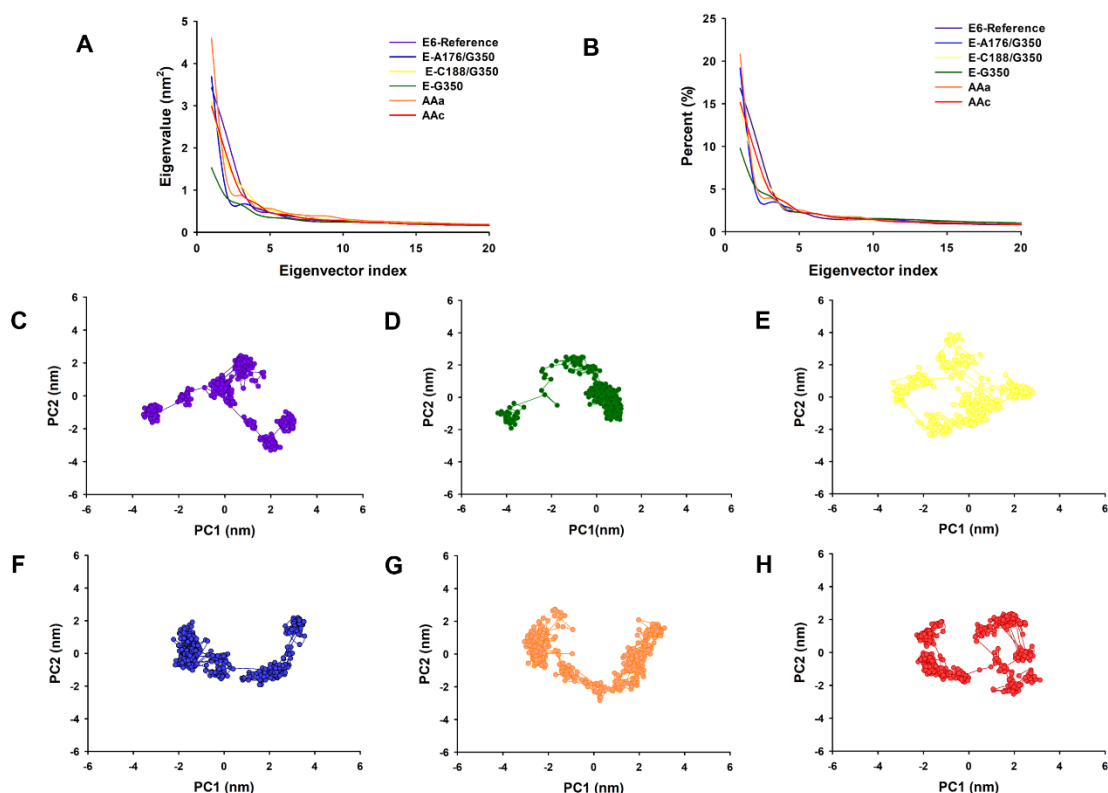

Figure 5. Principal component analysis (PCA) of E6-reference and its variants from HPV-16. (A) The eigenvalues plotted against the corresponding eigenvector indices obtained from the $C\alpha$ covariance matrix constructed from the 200 ns MD trajectory. E6-reference: violet, AAa variant: orange, AAc C variant: red, E-G350, green, E-C188/G350: yellow, E-A176/G350: blue. (B) Percentage of each eigenvector vs. eigenvalues. E6-reference: violet, AAa variant: orange, AAc C variant: red, E-G350, green, E-C188/G350: yellow, E-A176/G350: blue.. Projection of the motion of the structures of the backbone atoms (PC1 vs PC2) (C) E6-reference. (D) E-G350. (E) E-C188/G350 (F) E-A176/G350. (G) AAa and (H) AAc.

The projection of the first two eigenvectors (PC2 vs. PC1) for E6-reference, E-G350, E-A176/G350, E-C188/G350, AAa and AAc (Figure 5C-H), shows that E6-reference system (Fig. 5C) present different mobility behavior compared to the variants systems. The variants E-A176/G350 and AAa were expanded in the conformational space due to their flexibility (Figure 6F and G). The variants E-G350, E-C188/G350 and AAc have more restricted motions, making them the more stable

of the six protein systems (Figure 5 D, E and H). This points out that the punctual mutations of residues affect conformation and motion.

With respect to MAGI-1 255 and MAGI-1 329 the covariant matrix for the first 10 eigenvectors were 31.69 and 35.97 nm2 and the dPCA analysis showed that the first 10 eigenvectors captured 44.8 and 44.6 % of the proteins total motions (Figure 6A). The projections of the first two principal components (PC1 vs PC2) contributed to 29.5 and 39.2% of the collective motions (Figure 6B). These results showed a change in the motion of the atoms of MAGI-1 255 and MAGI-1 329 are

considerable, which means that missing 76 amino acids of model 255 restricts its motions making it more stable. These results support Ramírez et al., 2015 observations about the contribution of the highly disordered 76 amino acid region adjacent to the PDZ-1 domain of MAGI-1 to its behavior.
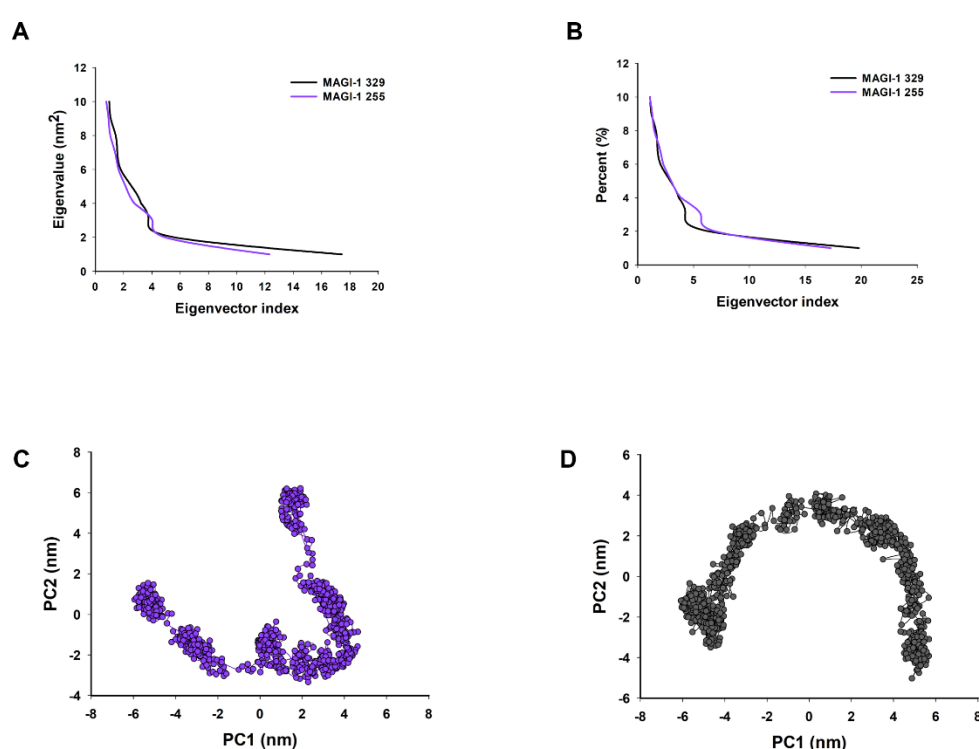

Figure 6. Principal component analysis (PCA) of MAGI-1 255 and MAGI-1 329. (A) First ten igenvalues plotted against the corresponding eigenvector indices obtained from the Cα covariance matrix constructed from the 200 ns MD trajectory. MAGI-1 255 purple and MAGI-1 329 black. (B) Percentage of each eigenvector vs. eigenvalues. 2D Projection of Principal Component Analysis. Projection of the motion of the protein in phase space along the first two principal components. (C) MAGI-1 255 and (D) MAGI-1 329.

2.4 Protein docking

Mutations in proteins can affect protein structure and stability, consequently, these mutations alter the kinetics and thermodynamics of protein-protein interactions (PPI) [32]. Using ClusPro blind base docking method, a representative protein structure of the most populated cluster obtained from the dPCA clustering analysis of the MD simulation refined proteins (E6-reference, E-G350, E-A176/G350, E-C188/G350, AAa and AAc) were docked against the representative structure of MAGI-1 329 and MAGI-1 255, also, obtained from dPCA clustering analysis. Docking resulted in 1000 protein conformations of complexes. The top 10 docked complexes from each problem were analyzed for the lowest energy and residues binding between the two proteins. The best complexes

were selected base on a greater number of cluster members and the lowest energy according to ClusPro guidelines.

The global free binding energy of the E6-reference against MAGI-1 329 and 255 complexes were calculated as -51.90 and -48.14 kcal/mol respectably, using FiberDock [33]. These energies were bigger than the energies obtained from the complexes between the E6 variants and the MAGI-1 models; this means that there is a greater affinity between MAGI-1 models and E6 variants (Table 1). According to our docking results the variants AA, E-A176/G350 and E-G350 showed smaller energy values -148.17, -148.86 and -148.86, we interpreted this as a gain of interaction affinity between these proteins. Moreover, the lower energy docking values were observed with the MAGI-1 329 model (Table 1). In addition, there was an increment in the number of hydrogen bond and salt bridges interactions of E6 and its variants with MAGI-1 329 compared to our MAGI-1 255 model (Table S1).

Table 1. Docking binding affinity of E6 HPV-16 and its variants. Number of cluster members and binding energy.

Complexes	No. Cluster members	Binding energy kcal/mol
E6R/MAGI-1 255	83	-48.14
E-G350/MAGI-1 255	126	-111.14
E-C188/G350/MAGI-1 255	111	-130.89
E-A176/G350/MAGI-1 255	74	-191.34
AAa/MAGI-1 255	111	-138.07
AAc/MAGI-1 255	204	-120.23
Complexes	No. Cluster members	Binding energy kcal/mol
E6R/MAGI-1 329	118	-54.90
E-G350/MAGI-1 329	88	-152.50
E-C188/G350/MAGI-1 329	82	-121.63
E-A176/G350/MAGI-1 329	91	-148.86
AAa/MAGI-1 329	83	-148.17
AAc/MAGI-1 329	88	-166.97

The protein-protein interfaces of the complexes were analyze using PDBsum generate [49]. The top docked complex of each variant against MAGI-1 255 and 329 were subjected to PDBsum to identify the interacting residues. Comparative analysis between the twelve complexes from docking interfaces of E6-reference and its variants identified a list of different amino acids that were shown to be responsible for the interaction with the MAGI-329 and MAGI-1 255.

For the complex E6-reference and MAGI-1 255 (Figure 7A), occurs mainly through amino acids of the WW2 domain (Y74, V79, D80, W66, G65, A64, Y78, I76 and from the PDZ1 domain A234, H231 and G230 with twenty-four amino acids of E6-reference that included: Y81, R77 and Y76 which are adjacent to H78 a highly mutated amino acid in E6 (Table S1). There are nine hydrogen bonds, 3 salt bridges (Table S1).
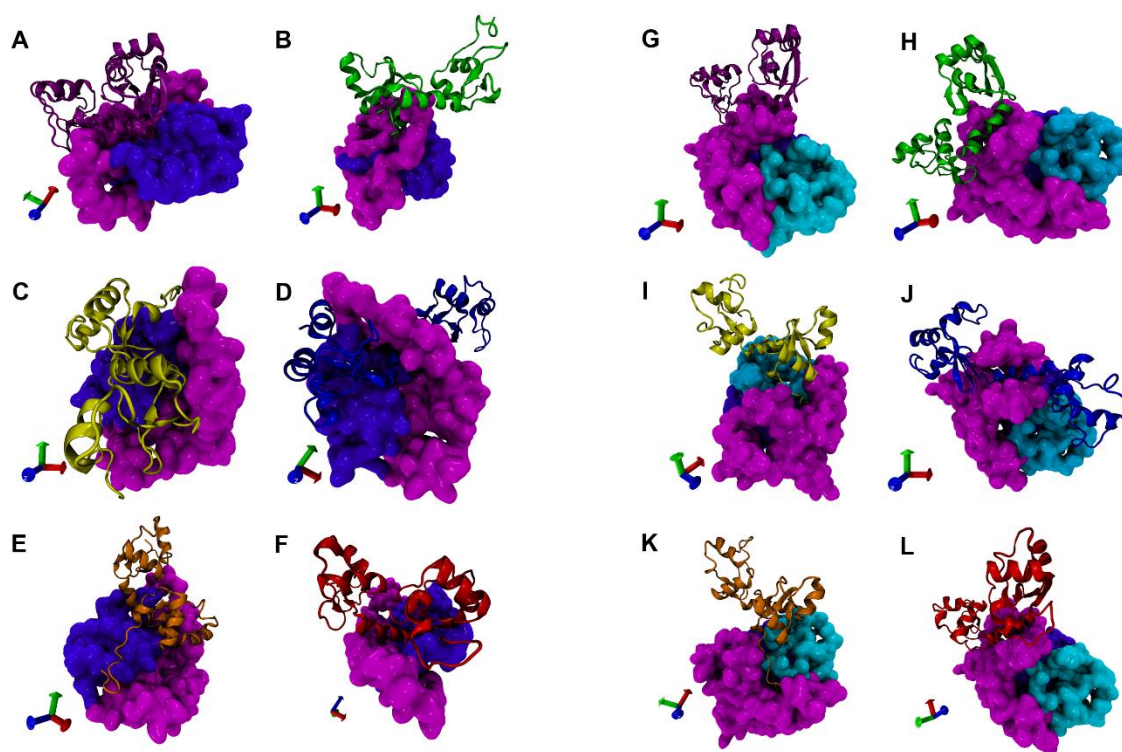
Figure 7. Protein-protein docking of E6, its variants with MAGI-1 255 and MAGI-1 329. Protein-protein docking analysis shows the probable interaction of E6-reference (purple), E-G350 (green), C188/G350 (yellow), E-A176/G350 (blue), AAa (orange) and AAc (red) with MAGI-1 255 (A to F). Docking between MAGI-1 329 and E6-references and its variants (G to L) E6-reference (purple), E-G350 (green), C188/G350 (yellow), E-A176/G350 (blue), AAa (orange) and AAc (red). MAGI-1 255 and MAGI-1 329 are represented in quicksurf in color magenta. The protein-protein docking was performed using the ClusPro 2.0 web server.

On the other hand, the interaction of E6-reference with MAGI-1 329 was through amino acids corresponding to WW2 domain and adjacent non-domain regions. Amino acids C34 to F171 were responsible for most of the interactions with twenty-three amino acids of E6-reference that included: R147, R131, R77, Y76, V31, T32 and H78 from E6-reference (Figure 7G). Some of the E6-reference interacting residues are located closed to the mutation sites, and its interacting residues are different compared to the variants. It is important to point out that H78 is mutated to T78 in the AAa and AAc variants, and the interaction of this residue with the models of MAGI-1 is lost for variant AAa. Also, we observed that R147 is the only residue close to the PBM motif of E6 interacting with our MAGI-1 329 model (Table S2). The differences between the interaction of E6-reference with our two models from MAGI-1 is remarkable, the increase in protein´s coverage results in a gain of interacting residues, an increase of hydrogen bonds, Salt bridges (Table S1). Regarding the missing interaction of the PBM motif with MAGI-1 255 and 329, we believe that the residues could be oriented in a way that avoids the interaction or may be block by adjacent residues. According to the online server HOPE, this could be attributed to changes in size and in charge from the mutations of each variant [50].

Its important to point out that the interacting residues from complex E-C188/G350/MAGI-1 329 included all the amino acids of the PBM motive of E6 (E148, T149, Q150, and 151V/L) interacting mainly with amino acids from the WW1 (G1-R33) domain and amino acids from a highly disordered region of MAGI-1 (G256-T329 in our model). The complex AAa/MAGI-1 329 shows a similar interacting pattern where amino acids from the PBM motif interact mainly with the highly disordered region of MAGI-1 (G256-T329). Different interacting patterns were observed from the rest of the variants and model MAGI-1 255 (Table S2).

For complexes E-G350/MAGI-1 255, E-G350/MAGI-1 329, E-C188/G350/MAGI-1255, E-A176/350/MAGI-1255, E-A176/350/MAGI-1255, AAc/MAGI-1 255 and AAc/MAGI-1 329 detailed information about the amino acids involved in the interaction and the type of bonds can be found in Table S1 and S2.

It is important to understand the changes in PPIs caused by these mutations may alters the affinity and stability of the interaction of E6 with proteins important for tissue homeostasis. The increase in affinity and stability of the interaction of E6 with MAGI-1 result in an increase in the degradation of MAGI-1 and as a consequence in the loss of stability of important cell complexes that maintain cell-cell adherence at the adherents junctions.

3. Materials and Methods

3.1 3D protein structures

Multiple alignments of the sequences were performed using CLUSTAL X 1.81. The secondary structure of the E6 protein and its variants were predicted using PSIPRED server. The crystal structure of VPH-16 E6 protein was obtained from the Protein Data Bank (RCSB PDB), with the identification number: 4XR8, chain H. The E6 structure on this PDB contains 151 residues with four-point mutations in S80C, S97C, S111C, and S140, which were reverted to obtain the E6 reference in the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. After that, all the mutations were carried out in the E6 reference to obtained all the variants of HPV-16. The mutations were done as indicated next, to obtained E-G350: L83V; E-C188/G350: E29Q and L83V; E-A176/G350: D25N and L83V; AAa: Q14H, H78Y, and L83V; for AAc: Q14H, I27R, H78Y, and L83V. The obtained proteins were structurally aligned and visualized using VMD.

To obtained the 3D structure of MAGI-1, a total of 255 and 329 amino acids from the amino acid terminal region of the protein sequence (300-554 and 300-628) were retrieved from the UniProtKB database, (accession number Q96QZ7) and submitted to I-TASSER server as two separate jobs. First, the 3D structure with 225 residues, which comprises WW1, WW2 and the PDZ1 domains of MAGI-1 was obtained by homology modelling using the I-TASSER server, as a template, we selected PDB file: 2KPK, which corresponds to the PDZ-1 domain of the MAGI-1. Furthermore, a 3D structure of 329 amino acids of MAGI-1, which includes the WW1, WW2, PDZ-1 and a 76 amino acidic disordered region of this protein, was obtained by homology modelling on the I-TASSER server using the same template PDB. All the 3D predicted structures were evaluated using the Rampage web server to obtain the Ramachandran plots (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php).

3.2 Molecular dynamics simulation

Parameters for the two $Zn2+$ ions and eight cysteine-ligand coordination of E6 were kindly provided by Justin Lemkul from the Virginia Polytechnic Institute and State University, these parameters included a CYSD patch for the deprotonation of the eight zinc-bound cysteines and a ZN_C patch to covalently link cysteines to the zinc ions. The correct coordination of the deprotonated cysteines and the ion zinc using these patches has been demonstrated in previous studies [34], the CHARMM 36 force field was employed for the application of the patches using CHARMM software [35, 36]. Afterwards, a 200 ns of MD simultion of E6 and its variants were performed using the 2.8 NAMD software package [37] with CHARMM36 and CHARMM22 force fields [36]. For MAGI-1 models we used CHARMM 27 topology and parameter files for proteins. Each system was placed in a cubic box of TIP3P water with a minimum distance of 10 Å between the solute atoms and the edge of the box [38]. To neutralize the systems, we added 7568 water molecules, 21 Na+ and 27 Cl− to the E6-reference. To variant E-G350, we added 7420 water molecules, 21 Na+, and 27 Cl−, to E-C188/G350 variant, 7484 water molecules, 21 Na+, and 28 Cl− were added, to E-A176/G350 variant, 7660 water molecules, 22 Na+ and 29 Cl− were added, to AAa

variant 7566 water molecules, 21 Na+ and 27 Cl− were added and to AAc variant, 7554 water molecules, 21 Na+ and 28 Cl− were added. For MAGI-1 255 we added 10932 water molecules and 18 Na+, and for MAGI-1 329, 11456 water molecules and 23 Na+ were added. Each system was neutralized to 0.15 mol/L of NaCl and submitted to minimization energy for 10,000 steps of steepest descent minimization followed by equilibration for 1 ns under constant temperature 310 K and pressure 1 atm (NPT) ensemble with protein atoms restraints [39, 40]. MD simulation were run for 200 ns, considering all proteins as soluble.

3.3 Trajectory and dPCA analysis

The carma software [44], was used to calculate the root mean square deviation (RMSD) calculates the average deviation in the atomic stability throughout MD simulation, radius of gyration (Rg) measures the compactness and expansion of the molecules, and the root means square fluctuation (RMSF) a parameter to explored the flexibility of the protein through MD simulation, as well as the Principal component analysis (PCA) and dPCA based clustering analysis employing the last 50 ns of the trajectory. dPCA is a standard tool in statistical mechanics used in order to determine the correlated motions of the residues to a set of linearly uncorrelated variables called principal components, and it allows to obtain the large scale collective motions of the atoms on the simulations, which frequently correlates with the proteins biological function and structural properties [41]. Finally, we obtained the PDB files from the most populated cluster analysis and performed a protein-protein docking. Molecular graphics were performed in Sigma plot 12.0. VMD was used to visualize all the 3D proteins [42].

3.4 Protein-protein docking

The protein–protein dockings were carried in ClusPro server [43, 44], the program has been consistently rated among the best global docking methodologies in the CAPRI challenge (Critical Assessment of Predicted Interactions) [43]. For the docking studies, refined models for most populated cluster from E6-reference or its variants were docked within the MAGI-1 (235 and 329) homology models, where MAGI-1 models were the receptors and E6, and its variants were used as a ligand. The conformers with the highest cluster members and the lowest energy calculated in FireDock were taken for analysis on the PDBSum server [33, 45]. All docking complexes were visualized by VMD software [42].

4. Conclusions

We proposed an *in-silico* approach to evidence the differences in the interaction of E6 and five of its natural variants with two models, cellular protein MAGI-1. According to our results variants, AAa and E-C188/G350 showed greater levels of instability, less compactness, a gain of fluctuation regions that are correlated to the increment of active sites. We attribute this behaviour to specific mutations of proteins, and these mutations cause physicochemical changes that affect the behaviour of proteins. Very marked dynamic changes are observed, particularly at the amino and carboxyl termini of proteins, where there is a gain in flexibility in the variants compared to E6-reference. Also, according to the dPCA results a dramatic change of motions behaviour for mutants compared to E6-reference. These differences in structure and mobility incremented the affinity of variants E-C188/G350 and AAa for our models of MAGI-1. E-C188/G350 increases its affinity for our models by three times, increasing the binding bonds by 50%. A similar pattern is observed among all the variants compared to E6-reference. Our results suggest that the physicochemical changes that gave rise to thermodynamic changes of the variants and an increase the affinity for our MAGI-1 models. Here, we were able to represent the possible changes in the physicochemical properties of E6 proteins and the repercussion in the interaction affinity with MAGI-1. An experimental validation will be necessary to evaluate the degradation profile of the MAGI-1 protein mediated by E6-reference and its variants.
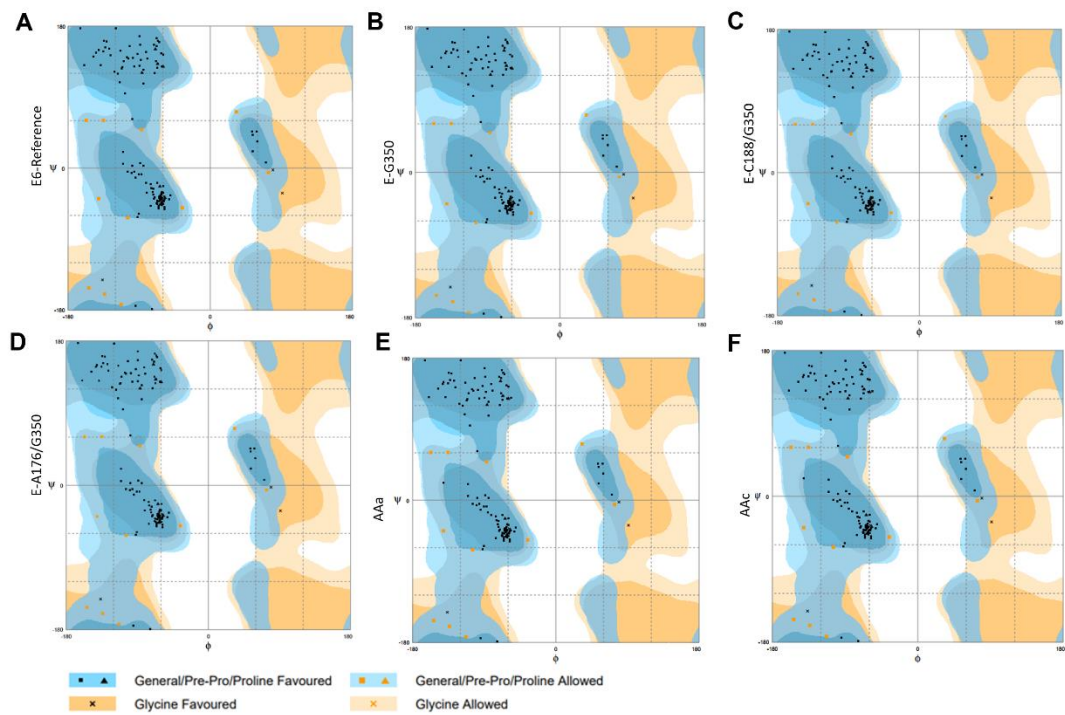
Abbreviations

|	
3D	Three Dimensional
CC	Cervical Cancer
HPV-16	Human Papillomavirus Type 16
PDB	Protein Data Bank
MD	Molecular Dynamic
HPV	Human Papillomavirus
dPCA	Dihedralangles PCA
HR-HPV	High Risk Human Papillomavirus
MAGI-1	Membrane-associated guanylate kinase with inverted domain structure-1
PDZ1	Discs-large homologous regions 1
PBM	protein binding motif
WW	rsp-domain-9
RMSD	Root Mean Square Deviation
Rg	Radius of gyration
RMSF	Root Mean Square Fluctuation
PPI	Protein-protein Interaction
PC	Principal Component

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

S1 Figure. Ramachandran plot analysis of E6 and its variants from HPV-16. (A) E6-reference. (B) E-G350. (C) E-C188/G350. (D) E-A176/G350. (E) E6-AAa. (F) E6-AAc.

Appendix B
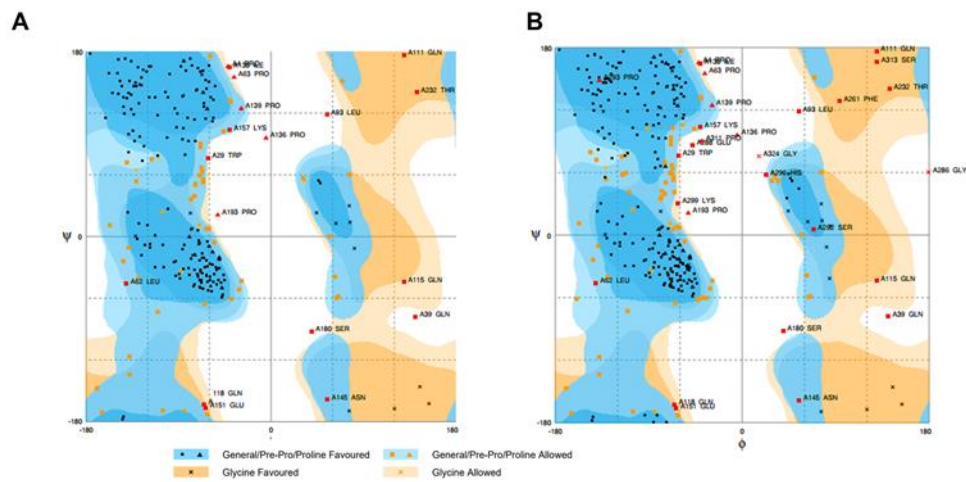
S2 Figure. Ramachandran plot of MAGI-I. (A) Ramachandran plot of MAGI-1 255 model from I-TASSER. (B) Ramchandran plot of I-TASSER model of MAGI-1 329.

Appendix C
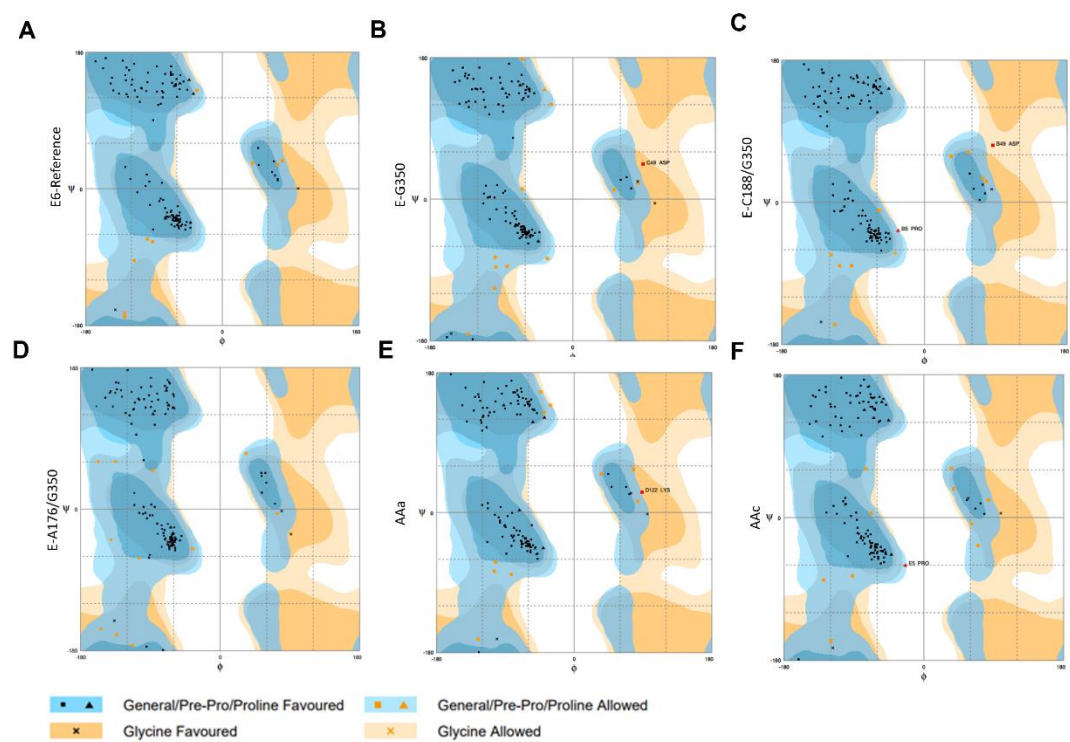
S3 Figure. Ramachandran plot analysis of E6-reference and its variants after MD simulation. (A) E6-reference (D) For the E-G350. (E) variant E-C188/G350. (B) For E6-AAa (C) E6-AAc.

Appendix D
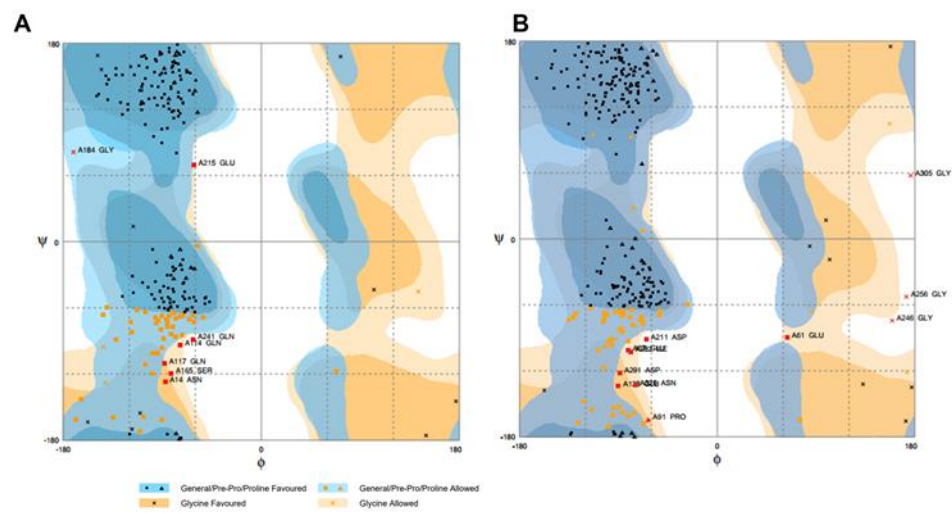
S4 Figure. Ramachandran plot analysis after 200 ns of MD simulation. (A) MAGI-I **255**. (B) **MAGI-1 329**

Table S1. Comparison of types and number of interactions between E6, its variants and our models of MAGI-1. Hydrogen bonds, Salt bridges, Non-bonded contacts.

Complexes	Hydrogen bonds	Salt bridges	Non-bonded contacts
E6R/MAGI-1 255	9	3	156
E-G350/MAGI-1 255	14	9	156
E-C188/G350 /MAGI-1 255	25	7	231
E-A176/G350/ MAGI-1 255	23	10	382
AAa/MAGI-1 255	17	3	239
AAc/MAGI-1 255	13	2	199
Complexes	Hydrogen bonds	Salt bridges	Non-bonded contacts
E6R/MAGI-1 329	11	6	156
E-G350/MAGI-1 329	18	4	220
E-C188/G350 /MAGI-1 329	23	8	298
E-A176/G350/ MAGI-1 329	18	5	212
AAa/MAGI-1 329	24	8	242
AAc/MAGI-1 329	12	6	189

Table S2. Detailed results of amino acidic interaction between E6, its variants and our two MAGI-1 models.

E6 VPH-16	Residues	Residues MAGI1-1 255
E6-referece	Y81, R10, C30, R77, V31, Y32, R55, Y76, C66, A61, F69, K65, V62, V53, L67, N127, C51, C80, R131, G130, Q107, I73, R129 and Y70.	G75, A234, Y74, T161, N163, V79, D80, W66, G65, A64, Y78, I76, H231, G230, Y77, E61, G70, K68 and E67.
E-G350	L72, R146, K68, N105, S143, E75, R135, S142, C139, S138, R144, C106, D44, E114, E113, M137, K121, R117, R141 and C140.	D57, S58, E59, L60, E61, P63, E55, L56, L62, A64, G65, Y78, K68, W66, D80, V79, E67 and H81.
E-C188/G350	R146, T145, K72, R144, R141, E150, S82, R147, E75, R124, R135, Y84, E148, Y79, H78, R77, Y76, F69, C82, R129, Y70, Y81, G130, N127,Y92 and R131.	S58, D57, L62, E67, L60, E59, P63, R162, A64, E61, L56, E54, E55, L167, S165, P164, N163, F160, K230, P193, F196, L229, D194, E70, K68, V221, N224, Q235, S222, C227, F171 and D225.
E-A176/G350	R146, Y84, N105, R124, R147, E148, S82, Y81, Q90, Q91, C33, T145, S143, R131, R129, S142, C80, G130, R77, I73, Y79, Y76, E75, K72, S74, S71, K68, D64, P69, K34, L67, K65, C66, C63, Y70, V62, R55, D56, Y32 and V53.	E55, A64, L167, N163, E54, P164, S165, R162, C227, D225, K85, T226, L229, L60, E61, P63, E70, E59, H231, P193, G230, D194, F159, F196, K68, F160, E67, V79, I69, Y78, H81, D80, R84, Y77, E151, N83, Y88, Y74, I82, E89, Q87, T86, E94, N90, P91 and R150.
AAa	R131, R29, S143, R10, Y32, Y70, F69, S74, R55, I73, Y76, Y78, R77, G130, Y92, K72, W132, N127, Y81, L151, E148, E75, Y79, R144, R124, S82, R147 and H126.	E67, W66, K238, E166, L62, P164, R162, L167, N163, F171, E59, S165, D194, T161, F196, F160, E195, I82, F159, K157, K68, C227, E70, I243, L229, P244, Q241, Q122, S242, G230, H231 and Q119.
AAc	C80, R129, R135, H126, Y81, G130, Y79, Y78,Y76, R124, E75, S82, R77, Y84, V83, Y32, Y70, F69, Q91, K72 and K65.	V221, C227, N163, H231, F171, E252, L229, T161, P164, G230, R162, S165, F160, F196, D194, F159, I76, I82, L167, E70, K168, K68 and E67.

	Residues	Residues MAGI-1 329
E6-referece	D64, K72, K68, K65, Y76, R147, F69, I73, L50, R77, Y70, R129, H78, V62, Q17, L67, R131, C66, V31, D56, Y32, R55 and V53.	D57, E54, S58, K157, A64, L62, E65, T53, G65, W66, V79, Y78, H81, I82, E67, D80, V92, E94, N83, R97, P91, N90 and E89.
E-G350	R77, I73, R129, Y76, D56, H78, I128, Y70, S71, N105, R48, C66, L67, I104, Q107, D74, L50, R131, Y32, F69, K72, R55, V31, V53 and K34.	A64, W66,G65, Q106, Y78, I69, K68, E61, E67, V79, T86, R84, H81, I82, D80, L93, E94, V92, T53, E54, D57, E102, R97, Q105, L101 and Q110.
E-C188/G350	K72, E75, R77, Y76, S142, R144, M137, R117, S143, R141, C140, K68, K65, F69, D64, E148, R147, T145, T149, R146, L151 and Q150.	S269, L270, V284, N285, G286, E55, E59, V51, Q39, F261, D262, L60, V317, D316, T272, T26, N23, T24, R33, E280, P263, N36, D265, P41, H22, D264, L35, L42, N145, D21, E43, E151, T12, L25, Y11, A10, E13, E8, M9 and I20.
E-A176/G350	R147, Y32, R55, G57, Y70, I73, K72, F69, S71, S74, Q91, E75, I128, Y79, H126, Y76, H78, R77, V83, R129, R131, C80, S82, R124 and Y81.	E89, L62, Q87, D316, E61, K157, A64, Y78, T86, W66, G65, V79, E67, D80, D57, Q106, H81, N83, I82, R84, P91, T53, H52, E94, N90, E54, L93, V92, and R97.
AAa	R77, Y76, K72, E75, Y79, R135, S138, S143, R141, S142, L151, K121, R144, T145, Q150, R147, E148, T149 and R146.	L277, E321, D278, V317, A318, D316, A315, N320, L6, L154, S58, E61, E59, L60, P260, K200, L259, P153, V51, V188, D262, F261, D264, L35, P263, Q39, K40, C34, P41, L42, E43, H52, G50 and E55.

| AAc | K65, K68, F69, Y78, R129, S74, R77, L50, V62, Y70, I73, R102, Q107, F45, L67, V53, Y32, Y76, Y60, K72, D56, C33, V31, K34 and R55. | T53, E54, S58, D57, L56, Y78, D80, I69, K68, E67, V79, H81, I82, E94, W66, A64, G65, N83, L62, R84, R97, K157, E61, L93, V92, L101, N90 and E89. |

References

1. Smith, J. S.; Lindsay, L.; Hoots, B.; Keys, J.; Franceschi, S.; Winer, R.; Clifford, G. M., Human papillomavirus type distribution in invasive cervical cancer and high-grade cervical lesions: a meta-analysis update. *Int J Cancer* **2007,** 121, (3), 621-32.

2. Doorbar, J.; Egawa, N.; Griffin, H.; Kranjec, C.; Murakami, I., Human papillomavirus molecular biology and disease association. *Rev Med Virol* **2015,** 25 Suppl 1, 2-23.

3. Ganti, K.; Broniarczyk, J.; Manoubi, W.; Massimi, P.; Mittal, S.; Pim, D.; Szalmas, A.; Thatte, J.; Thomas, M.; Tomaić, V.; Banks, L., The Human Papillomavirus E6 PDZ Binding Motif: From Life Cycle to Malignancy. *Viruses* **2015,** 7, (7), 3530-51.

4. Ramírez, J.; Recht, R.; Charbonnier, S.; Ennifar, E.; Atkinson, R. A.; Travé, G.; Nominé, Y.; Kieffer, B., Disorder-to-order transition of MAGI-1 PDZ1 C-terminal extension upon peptide binding: thermodynamic and dynamic insights. *Biochemistry* **2015,** 54, (6), 1327-37.

5. Shai, A.; Brake, T.; Somoza, C.; Lambert, P. F., The human papillomavirus E6 oncogene dysregulates the cell cycle and contributes to cervical carcinogenesis through two independent activities. *Cancer Res* **2007,** 67, (4), 1626-35.

6. Pim, D.; Thomas, M.; Banks, L., Chimaeric HPV E6 proteins allow dissection of the proteolytic pathways regulating different E6 cellular target proteins. *Oncogene* **2002,** 21, (53), 8140-8.

7. Zehbe, I.; Richard, C.; DeCarlo, C. A.; Shai, A.; Lambert, P. F.; Lichtig, H.; Tommasino, M.; Sherman, L., Human papillomavirus 16 E6 variants differ in their dysregulation of human keratinocyte differentiation and apoptosis. *Virology* **2009,** 383, (1), 69-77.

8. CONEVAL Resultados de pobreza en México 2018 a nivel nacional y por entidades federativas. https://www.coneval.org.mx/Medicion/Paginas/PobrezaInicio.aspx (09/07/2021),

9. Ortiz-Ortiz, J.; Alarcón-Romero, L. e. C.; Jiménez-López, M. A.; Garzón-Barrientos, V. H.; Calleja-Macías, I.; Barrera-Saldaña, H. A.; Leyva-Vázquez, M. A.; Illades-Aguiar, B., Association of human papillomavirus 16 E6 variants with cervical carcinoma and precursor lesions in women from Southern Mexico. *Virol J* **2015,** 12, 29.

10. Asadurian, Y.; Kurilin, H.; Lichtig, H.; Jackman, A.; Gonen, P.; Tommasino, M.; Zehbe, I.; Sherman, L., Activities of human papillomavirus 16 E6 natural variants in human keratinocytes. *J Med Virol* **2007,** 79, (11), 1751-60.

11. Niccoli, S.; Abraham, S.; Richard, C.; Zehbe, I., The Asian-American E6 variant protein of human papillomavirus 16 alone is sufficient to promote immortalization, transformation, and migration of primary human foreskin keratinocytes. *J Virol* **2012,** 86, (22), 12384-96.

12. Richard, C.; Lanner, C.; Naryzhny, S. N.; Sherman, L.; Lee, H.; Lambert, P. F.; Zehbe, I., The immortalizing and transforming ability of two common human papillomavirus 16 E6 variants with different prevalences in cervical cancer. *Oncogene* **2010,** 29, (23), 3435-45.

13. Schiffman, M.; Rodriguez, A. C.; Chen, Z.; Wacholder, S.; Herrero, R.; Hildesheim, A.; Desalle, R.; Befano, B.; Yu, K.; Safaeian, M.; Sherman, M. E.; Morales, J.; Guillen, D.; Alfaro, M.; Hutchinson, M.; Solomon, D.; Castle, P. E.; Burk, R. D., A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia. *Cancer Res* **2010,** 70, (8), 3159-69.

14. Gheit, T.; Cornet, I.; Clifford, G. M.; Iftner, T.; Munk, C.; Tommasino, M.; Kjaer, S. K., Risks for persistence and progression by human papillomavirus type 16 variant lineages among a population-based sample of Danish women. *Cancer Epidemiol Biomarkers Prev* **2011,** 20, (7), 1315-21.

15. Lichtig, H.; Algrisi, M.; Botzer, L. E.; Abadi, T.; Verbitzky, Y.; Jackman, A.; Tommasino, M.; Zehbe, I.; Sherman, L., HPV16 E6 natural variants exhibit different activities in functional assays relevant to the carcinogenic potential of E6. *Virology* **2006,** 350, (1), 216-27.

16. Chakrabarti, O.; Veeraraghavalu, K.; Tergaonkar, V.; Liu, Y.; Androphy, E. J.; Stanley, M. A.; Krishna, S., Human papillomavirus type 16 E6 amino acid 83 variants enhance E6-mediated MAPK signaling and differentially regulate tumorigenesis by notch signaling and oncogenic Ras. *J Virol* **2004,** 78, (11), 5934-45.

17. Zacapala-Gómez, A. E.; Del Moral-Hernández, O.; Villegas-Sepúlveda, N.; Hidalgo-Miranda, A.; Romero-Córdoba, S. L.; Beltrán-Anaya, F. O.; Leyva-Vázquez, M. A.; Alarcón-Romero, L. e. C.; Illades-Aguiar, B., Changes in global gene expression profiles induced by HPV 16 E6 oncoprotein variants in cervical carcinoma C33-A cells. *Virology* **2016,** 488, 187-95.

18. Kotelevets, L.; van Hengel, J.; Bruyneel, E.; Mareel, M.; van Roy, F.; Chastre, E., Implication of the MAGI-1b/PTEN signalosome in stabilization of adherens junctions and suppression of invasiveness. *FASEB J* **2005,** 19, (1), 115-7.

19. Chen, H. I.; Sudol, M., The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci U S A* **1995,** 92, (17), 7819-23.

20. Dobrosotskaya, I.; Guy, R. K.; James, G. L., MAGI-1, a membrane-associated guanylate kinase with a unique arrangement of protein-protein interaction domains. *J Biol Chem* **1997,** 272, (50), 31589-97.

21. Zhang, Y.; Dasgupta, J.; Ma, R. Z.; Banks, L.; Thomas, M.; Chen, X. S., Structures of a human papillomavirus (HPV) E6 polypeptide bound to MAGUK proteins: mechanisms of targeting tumor suppressors by a high-risk HPV oncoprotein. *J Virol* **2007,** 81, (7), 3618-26.

22. Thomas, M.; Banks, L., In vitro assays of substrate degradation induced by high-risk HPV E6 oncoproteins. *Methods Mol Med* **2005,** 119, 411-7.

23. Thomas, M.; Glaunsinger, B.; Pim, D.; Javier, R.; Banks, L., HPV E6 and MAGUK protein interactions: determination of the molecular basis for specific protein recognition and degradation. *Oncogene* **2001,** 20, (39), 5431-9.

24. Thomas, M.; Laura, R.; Hepner, K.; Guccione, E.; Sawyers, C.; Lasky, L.; Banks, L., Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene* **2002,** 21, (33), 5088-96.

25. Charbonnier, S.; Nominé, Y.; Ramírez, J.; Luck, K.; Chapelle, A.; Stote, R. H.; Travé, G.; Kieffer, B.; Atkinson, R. A., The structural and dynamic response of MAGI-1 PDZ1 with noncanonical domain boundaries to the binding of human papillomavirus E6. *J Mol Biol* **2011,** 406, (5), 745-63.

26. Charbonnier, S.; Stier, G.; Orfanoudakis, G.; Kieffer, B.; Atkinson, R. A.; Travé, G., Defining the minimal interacting regions of the tight junction protein MAGI-1 and HPV16 E6 oncoprotein for solution structure studies. *Protein Expr Purif* **2008,** 60, (1), 64-73.

27. Schwede, T., Protein modeling: what happened to the "protein structure gap"? *Structure* **2013,** 21, (9), 1531-40.

28. Rodríguez-Ruiz, H. A.; Garibay-Cerdenares, O. L.; Illades-Aguiar, B.; Montaño, S.; Jiang, X.; Leyva-Vázquez, M. A., In silico prediction of structural changes in human papillomavirus type 16 (HPV16) E6 oncoprotein and its variants. *BMC Mol Cell Biol* **2019,** 20, (1), 35.

29. Bello-Rios, C.; Montaño, S.; Garibay-Cerdenares, O. L.; Araujo-Arcos, L. E.; Leyva-Vázquez, M. A.; Illades-Aguiar, B., Modeling and Molecular Dynamics of the 3D Structure of the HPV16 E7 Protein and Its Variants. *Int J Mol Sci* **2021,** 22, (3).

30. Venselaar, H.; Te Beek, T. A.; Kuipers, R. K.; Hekkelman, M. L.; Vriend, G., Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics* **2010,** 11, 548.

31. Glaunsinger, B. A.; Lee, S. S.; Thomas, M.; Banks, L.; Javier, R., Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene* **2000,** 19, (46), 5270-80.

32. Lorch, M.; Mason, J. M.; Clarke, A. R.; Parker, M. J., Effects of core mutations on the folding of a beta-sheet protein: implications for backbone organization in the I-state. *Biochemistry* **1999,** 38, (4), 1377-85.

33. Mashiach, E.; Schneidman-Duhovny, D.; Andrusier, N.; Nussinov, R.; Wolfson, H. J., FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Res* **2008,** 36, (Web Server issue), W229-32.

34. Godwin, R. C.; Melvin, R. L.; Gmeiner, W. H.; Salsbury, F. R., Binding Site Configurations Probe the Structure and Dynamics of the Zinc Finger of NEMO (NF-κB Essential Modulator). *Biochemistry* **2017,** 56, (4), 623-633.

35. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **2017,** 14, (1), 71-73.

36. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: the biomolecular simulation program. *J Comput Chem* **2009,** 30, (10), 1545-614.

37. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J Comput Chem* **2005,** 26, (16), 1781-802.

38. Jorgensen, W. L., Comparison of simple potential functions for simulating liquid water. In Jayaraman, C., Ed. 1983.

39. Nosé, S., A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984,** 81, (1), 511-519.

40. Hoover, W. G., Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A Gen Phys* **1985,** 31, (3), 1695-1697.

41. Meyer, T.; Ferrer-Costa, C.; Pérez, A.; Rueda, M.; Bidon-Chanal, A.; Luque, F. J.; Laughton, C. A.; Orozco, M., Essential Dynamics: A Tool for Efficient Trajectory Compression and Management. *J Chem Theory Comput* **2006,** 2, (2), 251-8.

42. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J Mol Graph* **1996,** 14, (1), 33-8, 27-8.

43. Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S., The ClusPro web server for protein-protein docking. *Nat Protoc* **2017,** 12, (2), 255-278.

44. Vajda, S.; Yueh, C.; Beglov, D.; Bohnuud, T.; Mottarella, S. E.; Xia, B.; Hall, D. R.; Kozakov, D., New additions to the ClusPro server motivated by CAPRI. *Proteins* **2017,** 85, (3), 435-444.

45. Laskowski, R. A.; Hutchinson, E. G.; Michie, A. D.; Wallace, A. C.; Jones, M. L.; Thornton, J. M., PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem Sci* **1997,** 22, (12), 488-90.