Biophysical Journal, Volume *121*

Supplemental information

Deciphering the mechanisms of HPV E6 mutations in the destabilization of E6/E6AP/p53 complex

Le Li, Xuewei Dong, Yiming Tang, Zenghui Lao, Xuhua Li, Jiangtao Lei, and Guanghong Wei

Supporting Information

Deciphering the mechanisms of HPV E6 mutations in the destabilization of E6/E6AP/p53 complex

Le Li[a], Xuewei Dong[a], Yiming Tang[a], Zenghui Lao[a], Xuhua Li[b], Jiangtao Lei[c] and Guanghong Wei[a]*

a. Department of Physics, State Key Laboratory of Surface Physics, and Key Laboratory for Computational Physical Sciences (Ministry of Education), Fudan University, Shanghai 200438, People's Republic of China.

b. MOE Key Laboratory for Nonequilibrium Synthesis and Modulation of Condensed Matter, School of Physics, Xi'an Jiaotong University, Xi'an 710049, China

c. Institute of Space Science and Technology, Nanchang University, Xuefu Avenue 999, Nanchang City 330031, China.

This supporting material contains analysis methods, six supplemental tables (Tables S1-S6), thirteen supplemental figures (Figures S1-S13) and references.

Analysis methods

Conformational property analysis. We utilized the tools implemented in the GROMACS software package and in-house codes to analyze our trajectory data. The following several parameters are used to examine the influences of the E6 mutations on the structure stabilities of the heterodimer and the heterotrimer: backbone root mean square derivation (RMSD), hydrogen bond (H-bond) number, contact number and distance distributions of residue pairs with the ability to form salt-bridge and cation-π interactions. Here, an H-bond is considered to be formed if the distance between the donor atom D and the acceptor atom A is ≤ 0.35 nm and the D-H···A angle is $\geq 150°$. An atomic contact is considered if the distance between two carbon atoms of nonsequential residues lies within 0.54 nm or the distance between any other two atoms of nonsequential residues comes within 0.46 nm. We first calculated the charge center of the charged group of the four residues: Arg^+, Lys^+, Glu^- and Asp^-, and then calculated the distance of the charge center between two oppositely charged residues. A salt-bridge is formed if the charge center distance is within 0.40 nm(1). A cation-π interaction is considered when the minimum distance between the centroid of aromatic ring and the ε-amino group (NH_3^+) in the side chain of residue Lysine or residue Argine becomes around 0.6 nm(2, 3). We also calculated the binding free energy with the

molecular mechanics/linear Generalized Born surface area (MM/GBSA) method using MMPBSA.py program of Ambertool(4, 5). The binding free energy ($\Delta G_{binding}$) between a ligand and a receptor is calculated as: $\Delta G_{binding} = \Delta E_{vdW} + \Delta E_{elec} + \Delta G_{polar} + \Delta G_{nonpolar}$. Here, E_{vdW} and E_{elec} are, respectively, the van der Waals (vdW) and the electrostatic interaction energies in vacuum. The $G_{polar} + G_{nonpolar}$ is the solvation free energy that is required to transfer a solute from vacuum into the solvent, where $G_{polar}$and $G_{nonpolar}$ are the electrostatic and non-electrostatic contributions to the solvation free energy, respectively. G_{polar} is calculated using the GB implicit solvent model (igb = 1) with a salt concentration of 0.1 M and $G_{nonpolar}$ is estimated using the solvent accessible surface area (SASA).

Community network analysis. Dynamic network analysis was performed using 'gmx covar' tool and our in-house codes. The Cα atom of an amino acid residue is considered as a node of the community network. The covariance value of two nodes was calculated using 'gmx covar' tool implemented in the GROMACS package. Then, the dynamic cross-correlation between two nodes was calculated as:

$$C_{ij} = \frac{Cov(i,\ j)}{\sqrt{Var(i) \cdot Var(j)}} = \frac{< (\vec{r}_i(t) - < \vec{r}_i(t) >) \cdot (\vec{r}_j(t) - < \vec{r}_j(t) >) >}{\sqrt{(< \vec{r}_i(t)^2 > - < \vec{r}_i(t) >^2) \cdot (< \vec{r}_j(t)^2 > - < \vec{r}_j(t) >^2)}}$$

where C_{ij} stands for the dynamic cross-correlation of two nodes (i and j) and Cov(i, j) is the covariance of the two nodes. Var(i) and Var(j) are the variance of node i and j, respectively. The cross-correlation values are zeroed and the corresponding edges are thus removed when the contact probabilities of corresponding residue pairs are less 0.7, in accordance with a number of previous studies(6-9). An atomic contact is taken to be formed using the criteria defined above. The weight of each edge is defined as $-\ln|C_{ij}|$. On the basis of the dynamic network, the shortest path (or the optimal path) between two residues can be obtained using codes developed by *Eargle* and *Sethi*(6). The betweenness of each edge is defined as the number of shortest paths that pass through that edge. The optimal community distribution is calculated using the Girvan-Newman algorithm(10), which iteratively removes the edge with the highest betweenness and recalculates the betweenness of all remaining edges until the modularity of the community network is maximized. The modularity is a measure of the quality of a particular division of a network, and the bigger the modularity, the better the division quality(11). The community analysis was conducted for each replicate trajectory and the results of all independent analysis were averaged. Community network analysis has been used to study the conformational dynamics of proteins(12-14).

As the terminal residues of E6AP (C-terminal residues 382-383), p53C (N-terminal residues 94-95 and C-terminal residues 291-292) and E6 (N-terminal residues 1-3 and C-terminal residues 140-143) have relatively high flexibility, they are excluded in all the analysis. Unless specified, we used the last 200 ns data of E6/E6AP heterodimers and the last 300 ns data of E6/E6AP/p53 heterotrimers for analysis.

Table S1. The binding free energy (kcal mol^{-1}) between E6 and E6AP in WT and F47R^{E6} heterodimers.

System	ΔE_{vdW}	ΔE_{elec}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{binding}$
WT dimer	-50.49 ± 2.06	-666.17 ± 23.58	669.31 ± 24.63	-8.46 ± 0.35	-55.81 ± 1.42
F47R^{E6} dimer	-49.32 ± 0.35	-707.76 ± 14.08	710.17 ± 15.71	-8.64 ± 0.06	-55.55 ± 1.87

Table S2. The binding free energy (kcal mol^{-1}) between E6 and p53 in WT and F47R^{E6} heterotrimers.

System	ΔE_{vdW}	ΔE_{elec}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{binding}$
WT trimer	-55.63 ± 4.35	-92.59 ± 6.60	116.19 ± 3.71	-11.09 ± 0.58	-43.13 ± 1.87
F47R^{E6} trimer	-50.79 ± 8.69	-65.32 ± 20.69	96.52 ± 21.74	-9.72 ± 1.43	-29.31 ± 11.24

Table S3. The binding free energy (kcal mol^{-1}) between E6 and E6AP in WT and R102A^{E6} heterodimers.

System	ΔE_{vdW}	ΔE_{elec}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{binding}$
WT dimer	-50.49 ± 2.06	-666.17 ± 23.58	669.31 ± 24.63	-8.46 ± 0.35	-55.81 ± 1.42
R102A^{E6} dimer	-54.52 ± 3.05	-557.06 ± 48.58	563.61 ± 46.69	-8.99 ± 0.44	-56.95 ± 5.38

Table S4. The binding free energy (kcal mol^{-1}) between E6 and E6AP in the three MD runs of R102A^{E6} heterodimer.

R102A heterodimer					
MD run	ΔE_{vdW}^{WT}	ΔE_{elec}^{WT}	ΔG_{polar}^{WT}	$\Delta G_{nonpolar}^{WT}$	$\Delta G_{binding}^{WT}$
MD-1	-51.66	-521.69	527.91	-8.57	-54.01
MD-2	-54.17	-537.03	546.48	-8.96	-53.69
MD-3	-57.73	-613.45	616.45	-9.442	-61.16

Table S5. The binding free energy (kcal mol^{-1}) between E6 and p53 in WT and R102A^{E6} heterotrimers.

System	ΔE_{vdW}	ΔE_{elec}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{binding}$
WT trimer	-55.63 ± 4.35	-92.59 ± 6.60	116.19 ± 3.71	-11.09 ± 0.58	-43.13 ± 1.87
R102A^{E6} trimer	-57.19 ± 5.88	-117.05 ± 18.12	145.21 ± 13.94	-11.13 ± 0.48	-40.17 ±11.37

Table S6. The binding free energy (kcal mol^{-1}) between E6 and E6AP in WT and L50E^{E6} heterodimers.

System	ΔE_{vdW}	ΔE_{elec}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{binding}$
WT dimer	-50.49 ± 2.06	-666.17 ± 23.58	669.31 ± 24.63	-8.46 ± 0.35	-55.81 ± 1.42
L50E^{E6} dimer	-49.12 ± 6.60	-541.91 ± 39.97	554.01 ± 40.52	-8.62 ± 0.78	-45.63 ± 6.65

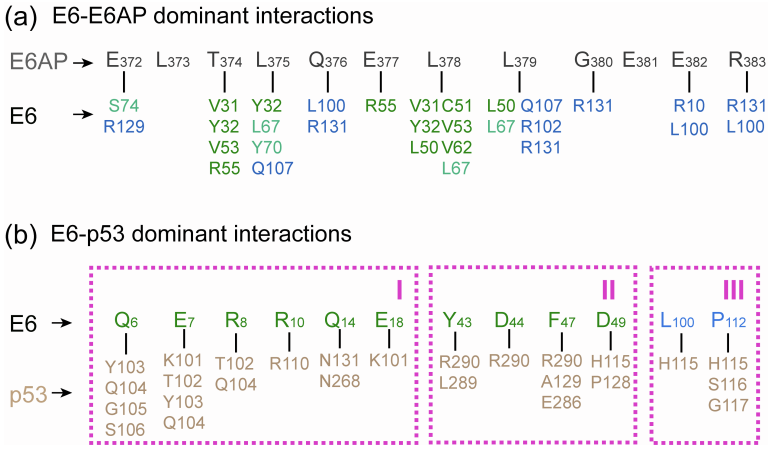

Figure S1. (a) The key interactions between E6 and E6AP in X-ray crystal structure(15). (b) The key interactions between E6 and p53 in X-ray crystal structure(16). Wheat: p53; green: E6N; cyan: E6HL; blue: E6C; gray: LxxLL motif of E6AP.

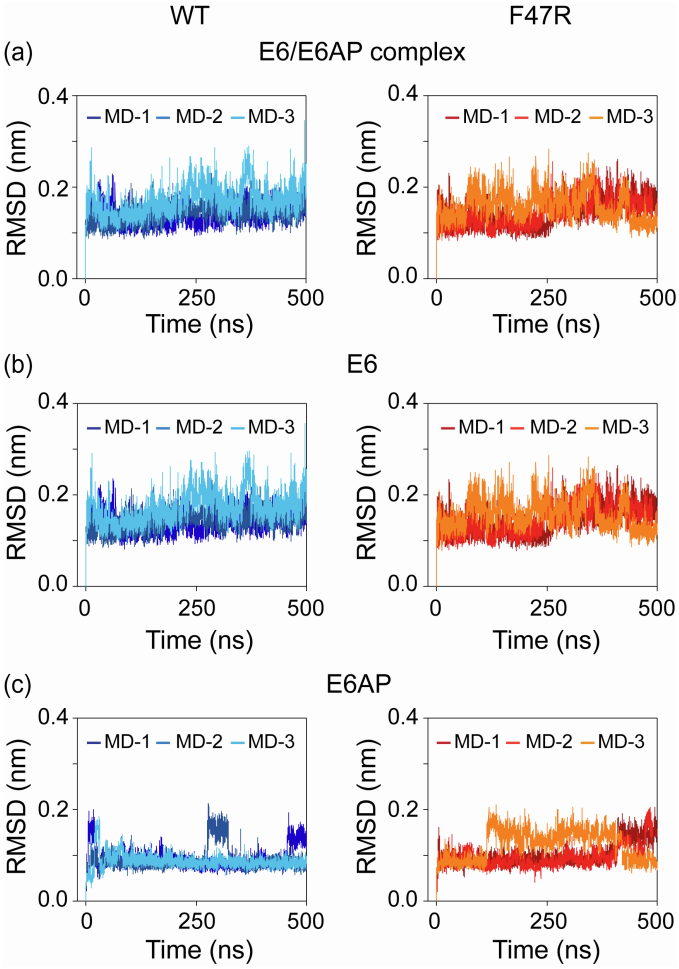

Figure S2. The time evolution of backbone root mean square deviation (RMSD) values of (a) E6/E6AP, (b) E6 and (c) E6AP relative to their initial conformations in WT and F47R mutant E6/E6AP heterodimers.

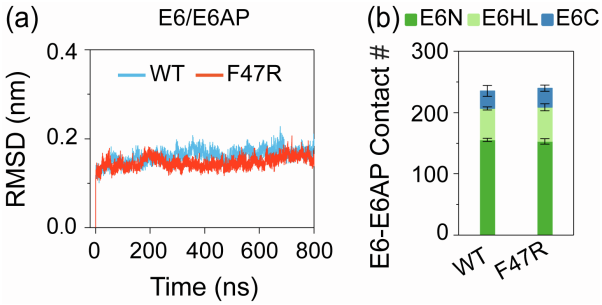

Figure S3. (a) The time evolution of backbone RMSD values of E6/E6AP relative to its initial conformations in WT and F47R mutant E6/E6AP/p53 heterotrimers. (b) The number of native contacts between E6 and E6AP in WT and F47R mutant E6/E6AP/p53 heterotrimers. The data are averaged over three independent MD runs.

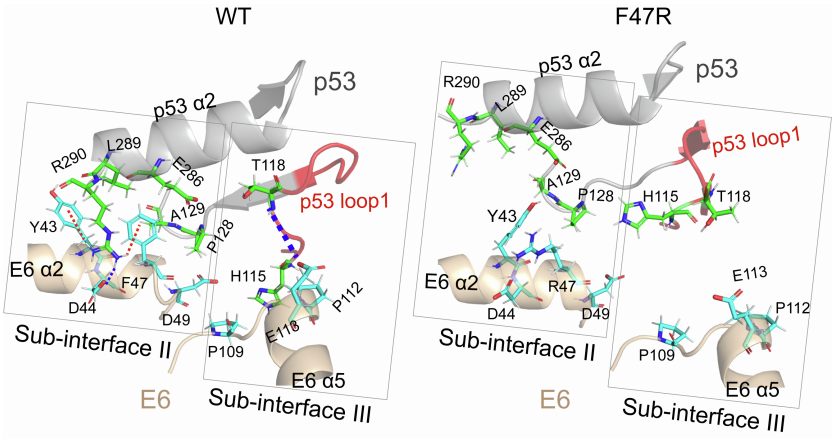

Figure S4. The snapshots (from the center structure of the first cluster) showing the F47R^{E6} mutation-induced alteration of E6-p53 interactions. Thin red dashed line: cation-π interaction; thin blue dashed line: salt-bridge interaction; thick blue dashed line: H-bond.

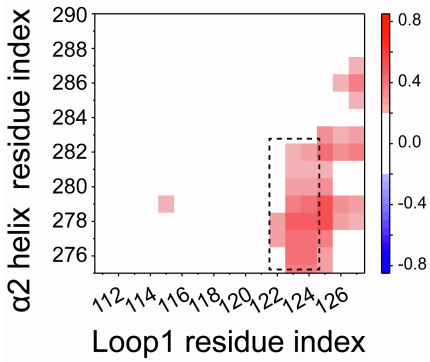

Figure S5. The correlation matrix between α2-helix^{p53} and loop1^{p53} in E6^{F47R}/E6AP/p53 heterotrimer. For clarity, correlations whose absolute value is less than 0.2 are not shown.

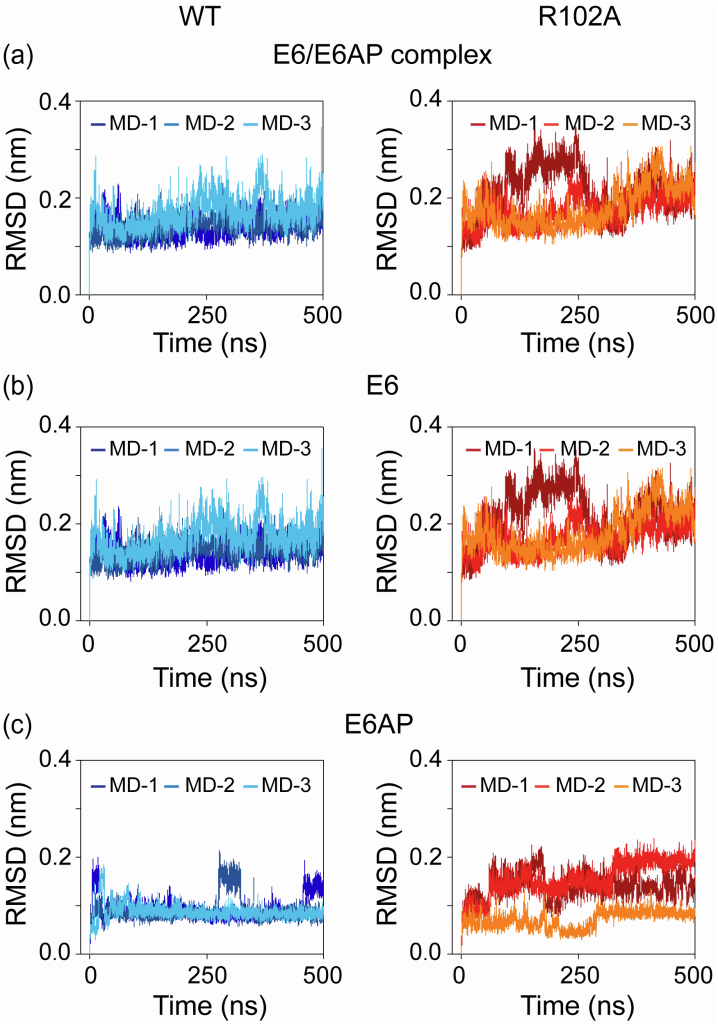

Figure S6. The time evolution of backbone RMSD values of (a) E6/E6AP, (b) E6 and (c) E6AP relative to their initial conformations in WT and R102A mutant E6/E6AP heterodimers.

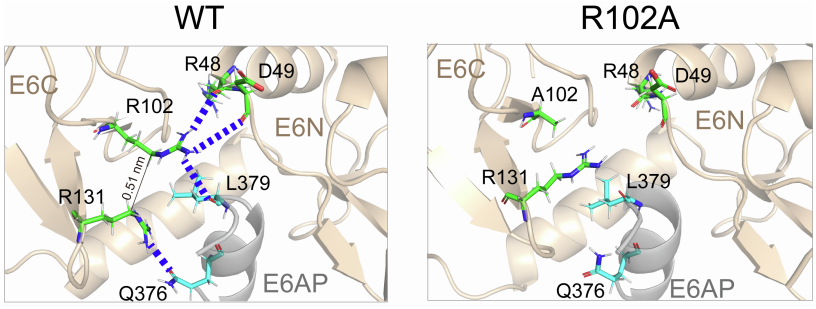

Figure S7. The snapshots (from the center structure of the first cluster of the clustering analysis) showing the R102A^{E6} mutation-induced alteration of E6C-E6N and E6C-E6AP interactions. Thick blue dashed line: H-bond.

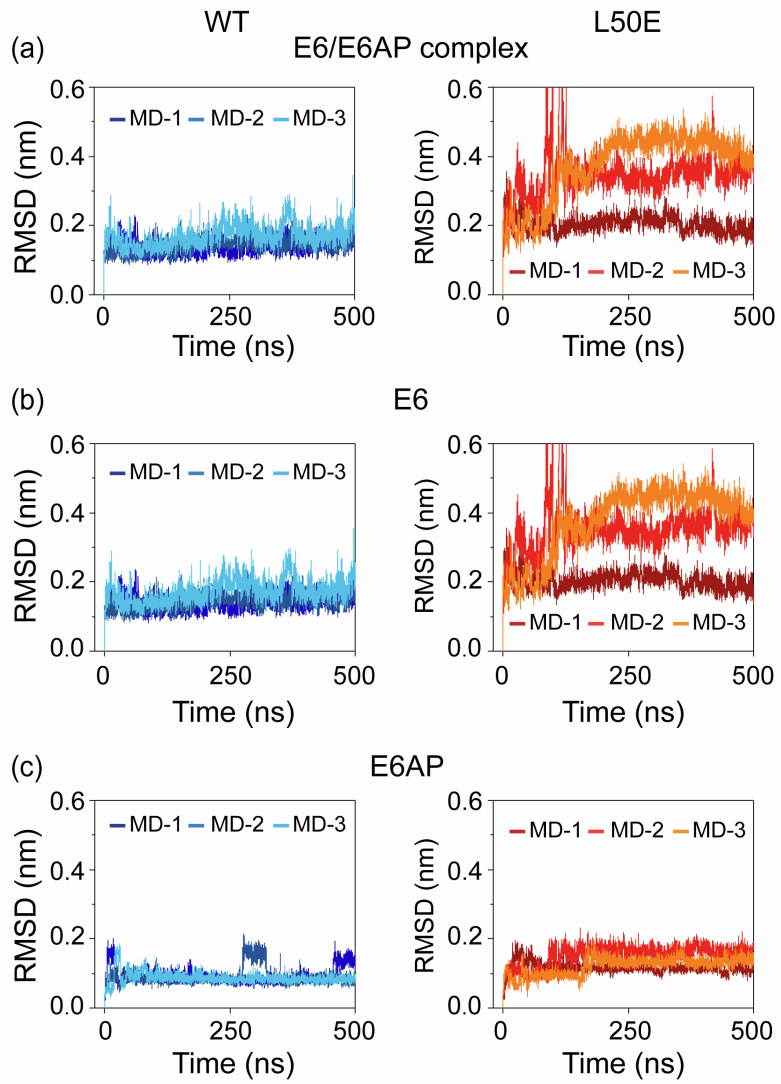

Figure S8. The time evolution of backbone RMSD values of (a) E6/E6AP, (b) E6 and (c) E6AP relative to their initial conformations in WT and L50E mutant E6/E6AP heterodimers.

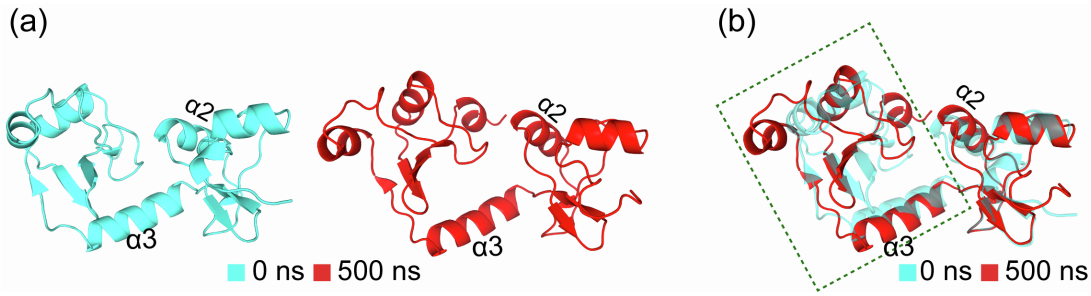

Figure S9. (a) The initial and final conformations of E6 in $E6^{L50E}$/E6AP heterodimer. (b) The final conformation of $E6^{L50E}$ superposed with its initial conformation.

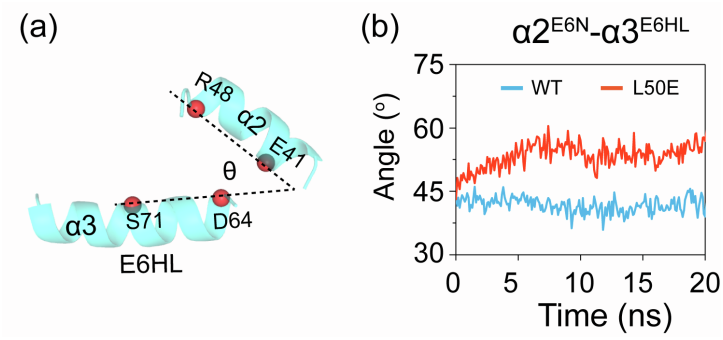

Figure S10. (a) A snapshot showing the angle between α2-helix^{E6N} and α3-helix^{E6HL}. (b) The time evolution of the angle within the first 20 ns simulation time.

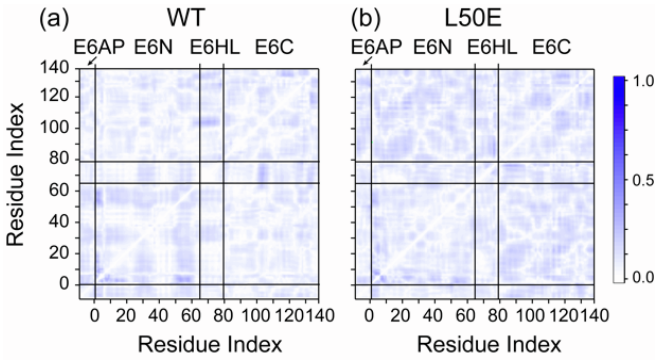

Figure S11. The matrix of Cα-Cα cross-correlation RMSDs of (a) WT and (b) L50E^{E6} heterodimers. Cα-Cα cross-correlation RMSD values were calculated using the data from the three replicate MD simulation of each system.

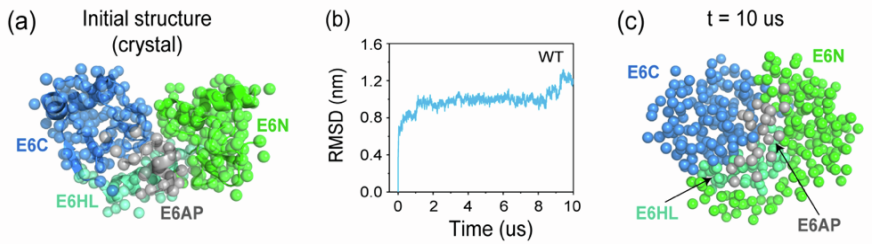

Figure S12. Analysis of simulation results obtained using coarse-grained MARTINI force field for WT E6/E6AP dimer. (a) The snapshot of initial coarse-grained model. (b) Time evolution of RMSD of the heterodimer relative to its initial conformation. (c) The snapshot of heterodimer at t = 10 μs.

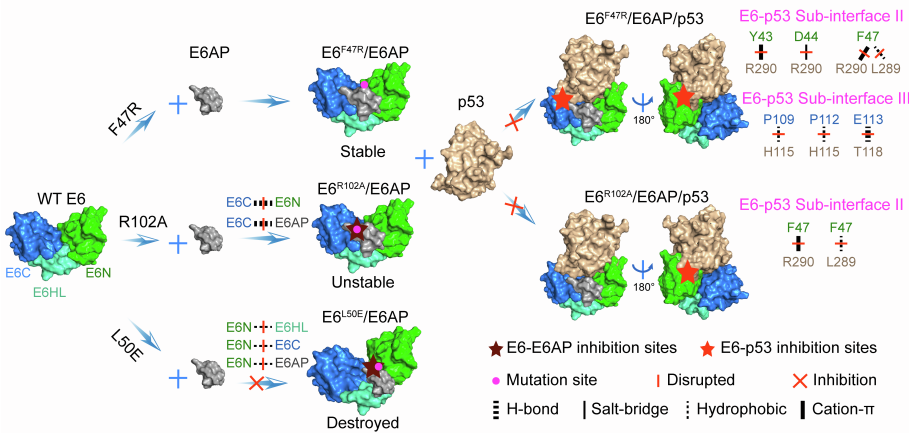

Figure S13. The schematic diagram showing how each p53-degradation-defective HPV16 E6 mutant disrupts the formation of E6/E6AP heterodimer and E6/E6AP/p53 heterotrimer.

References

1. Li, L., X. Li, Y. Tang, Z. Lao, J. Lei, and G. Wei. 2020. Common cancer mutations R175H and R273H drive the p53 DNA-binding domain towards aggregation-prone conformations. *Phys. Chem. Chem. Phys.* 22:9225-9232.
2. Ma, J. C., and D. A. Dougherty. 1997. The cation-π interaction. *Chem. Rev.* 97:1303-1324.
3. Gallivan, J. P., and D. A. Dougherty. 1999. Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci. U. S. A.* 96:9459-9464.
4. Rastelli, G., A. D. Rio, G. Degliesposti, and M. Sgobba. 2010. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J. Comput. Chem.* 31:797-810.
5. Miller III, B. R., T. D. McGee Jr, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg. 2012. MMPBSA. py: an efficient program for end-state free energy calculations. *J. Chem. Theory. Comput.* 8:3314-3321.
6. Sethi, A., J. Eargle, A. A. Black, and Z. Luthey-Schulten. 2009. Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* 106:6620-6625.
7. Zhao, Y., Y. Jian, Z. Liu, H. Liu, Q. Liu, C. Chen, Z. Li, L. Wang, H. H. Huang, and C. Zeng. 2017. Network analysis reveals the recognition mechanism for dimer formation of bulb-type lectins. *Sci. Rep.* 7:1-9.
8. Contreras-Riquelme, S., J.-A. Garate, T. Perez-Acle, and A. J. Martin. 2018. RIP-MD: a tool to study residue interaction networks in protein molecular dynamics. *PeerJ.* 6:e5998.
9. Tang, Y., Y. Yao, and G. Wei. 2021. Unraveling the Allosteric Mechanism of Four Cancer-related Mutations in the Disruption of p53-DNA Interaction. *J. Phys. Chem. B.* 125:10138-10148
10. Girvan, M., and M. E. Newman. 2002. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. U. S. A.* 99:7821-7826.
11. Newman, M. E., and M. Girvan. 2004. Finding and evaluating community structure in networks. *Phys. Rev. E.* 69:026113.
12. Qian, H., Y. Zou, Y. Tang, Y. Gong, Z. Qian, G. Wei, and Q. Zhang. 2018. Proline hydroxylation at different sites in hypoxia-inducible factor 1α modulates its interactions with the von Hippel-Lindau tumor suppressor protein. *Phys. Chem. Chem. Phys.* 20:18756-18765.
13. Lei, J., R. Qi, Y. Tang, W. Wang, G. Wei, R. Nussinov, and B. Ma. 2019. Conformational stability and dynamics of the cancer-associated isoform Δ133p53β are modulated by p53 peptides and p53-specific DNA. *FASEB. J.* 33:4225-4235.
14. Tang, Y., Y. Yao, and G. Wei. 2020. Structural and dynamical mechanisms of a naturally occurring variant of the human prion protein in preventing prion conversion. *Chinese. Phys. B.* 29:108710.
15. Zanier, K., S. Charbonnier, A. O. M. h. O. Sidi, A. G. McEwen, M. G. Ferrario, P. Poussin-Courmontagne, V. Cura, N. Brimer, K. O. Babah, T. Ansari, , I. Muller, R. H. Stote, J. Cavarelli, S. V. Pol, and G. Travé. 2013. Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science.* 339:694-698.
16. Martinez-Zapien, D., F. X. Ruiz, J. Poirson, A. Mitschler, J. Ramirez, A. Forster, A. Cousido-Siah, M. Masson, S. V. Pol, A. Podjarny, G. Travé, and K. Zanier. 2016. Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature.* 529:541-545.