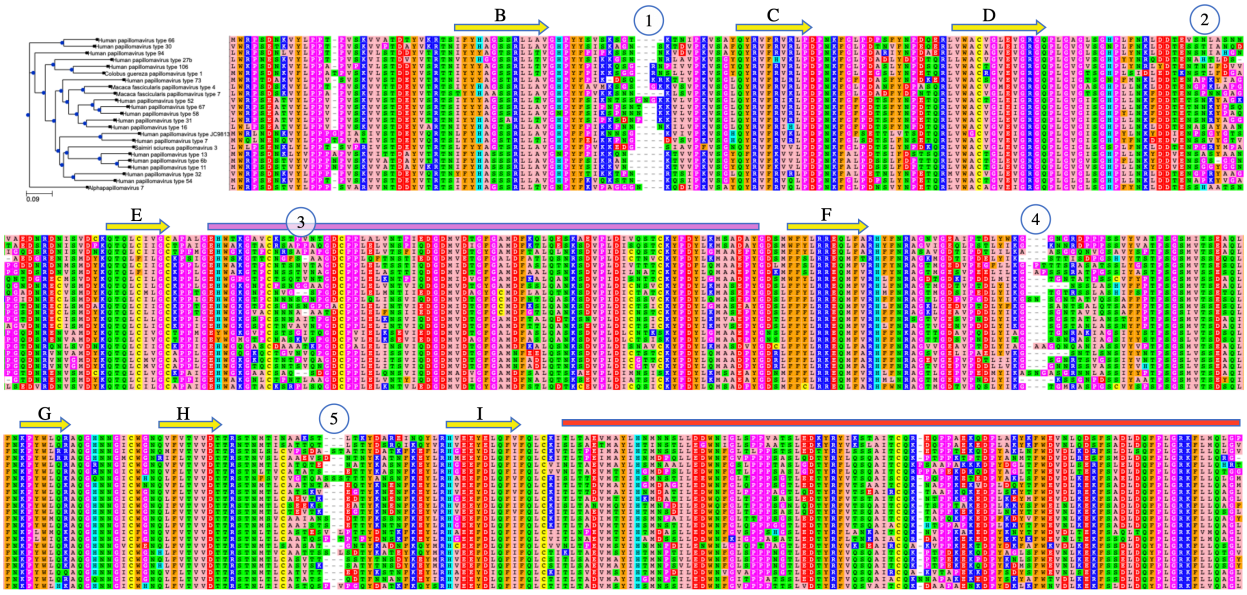
SUPPLEMENTARY MATERIALS

Engineering a Universal Dengue Virus Vaccine using a Virus-Like Particle Scaffold

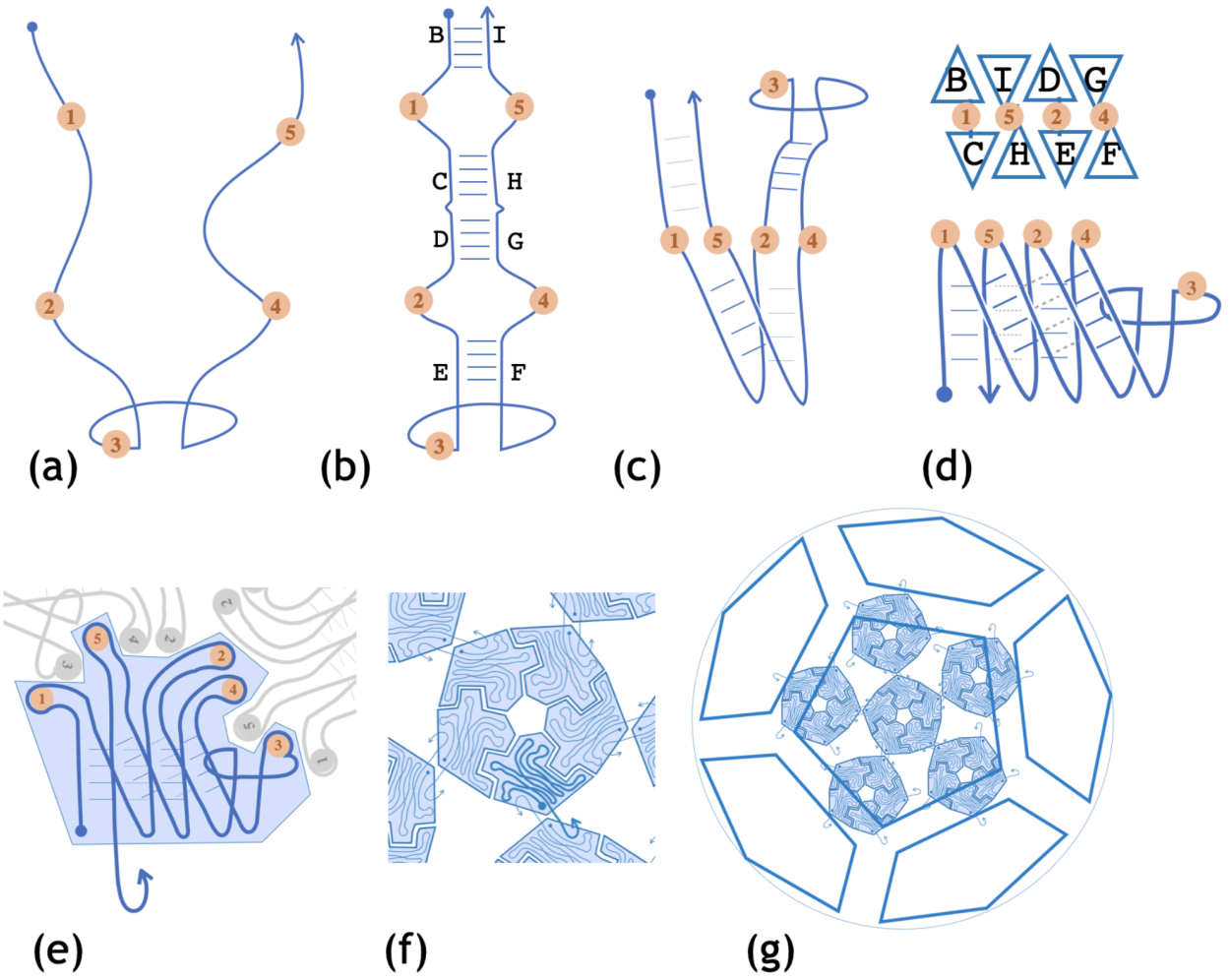
Danielle A. Basore[1,3], Carolyn M. Barcellona[2], Thomas B. Jordan[1, 4], Donna E. Crone[1], Sharon Isern[2], Scott F. Michael[2], Christopher Bystroff[1]

[1]Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, USA, [2]Biological Sciences, Florida Gulf Coast University, Fort Myers, FL, USA, [3]Mercy College, Dobbs Ferry, NY, USA. [4]Charles River Labs, Wilmington MA, USA.

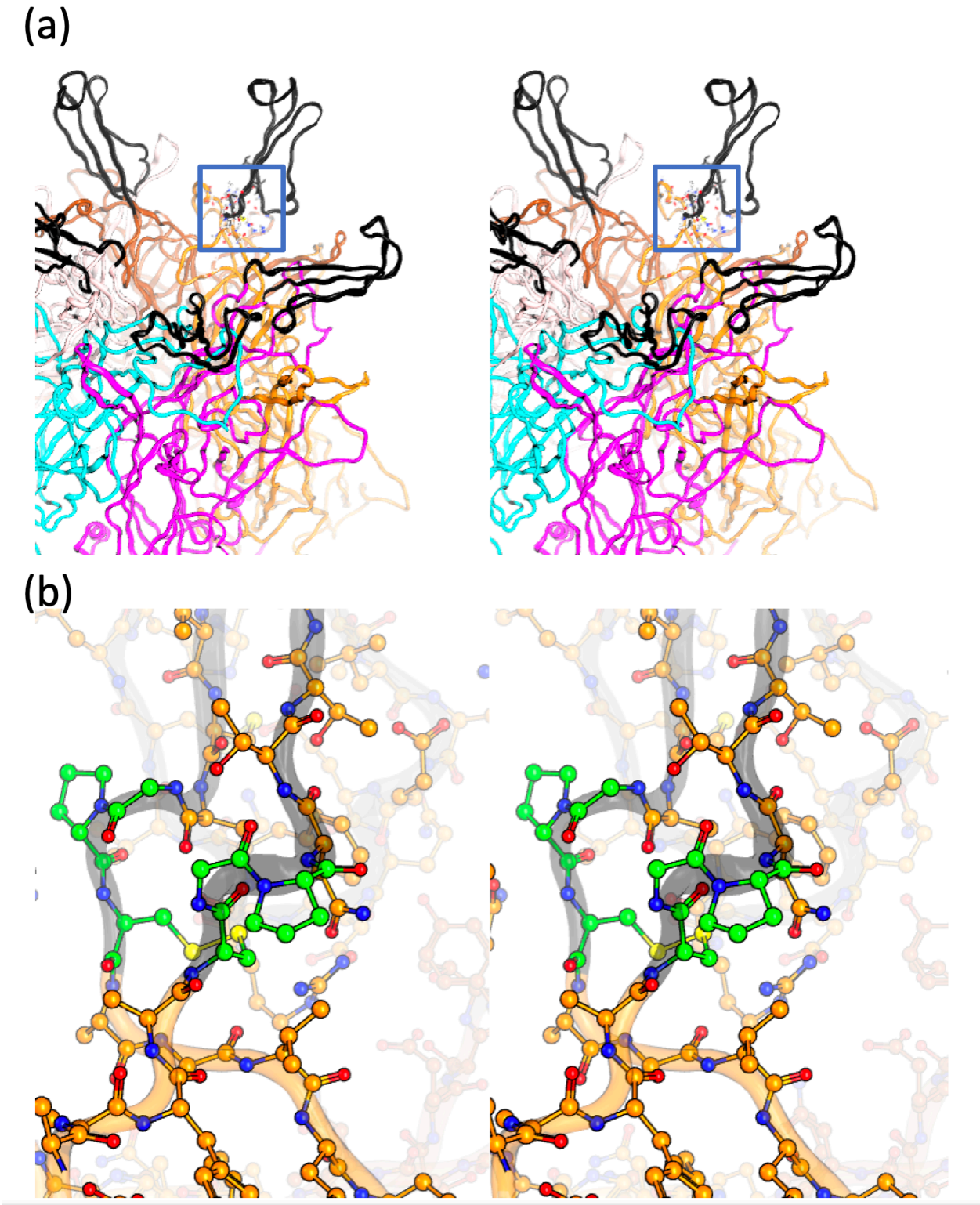
Supplementary Figure 1. Multiple sequence alignment and phylogram for 23 mutually diverse primate papilloma virus L1 sequences, showing sites of high variability. Core secondary structure elements (yellow arrows) are labeled B-I. Purple bar demarks the "irregular" domain. Red bar indicates the "invading arm". Circles are interesting candidate sites for chimeric vaccine design, based on the locations of indels.



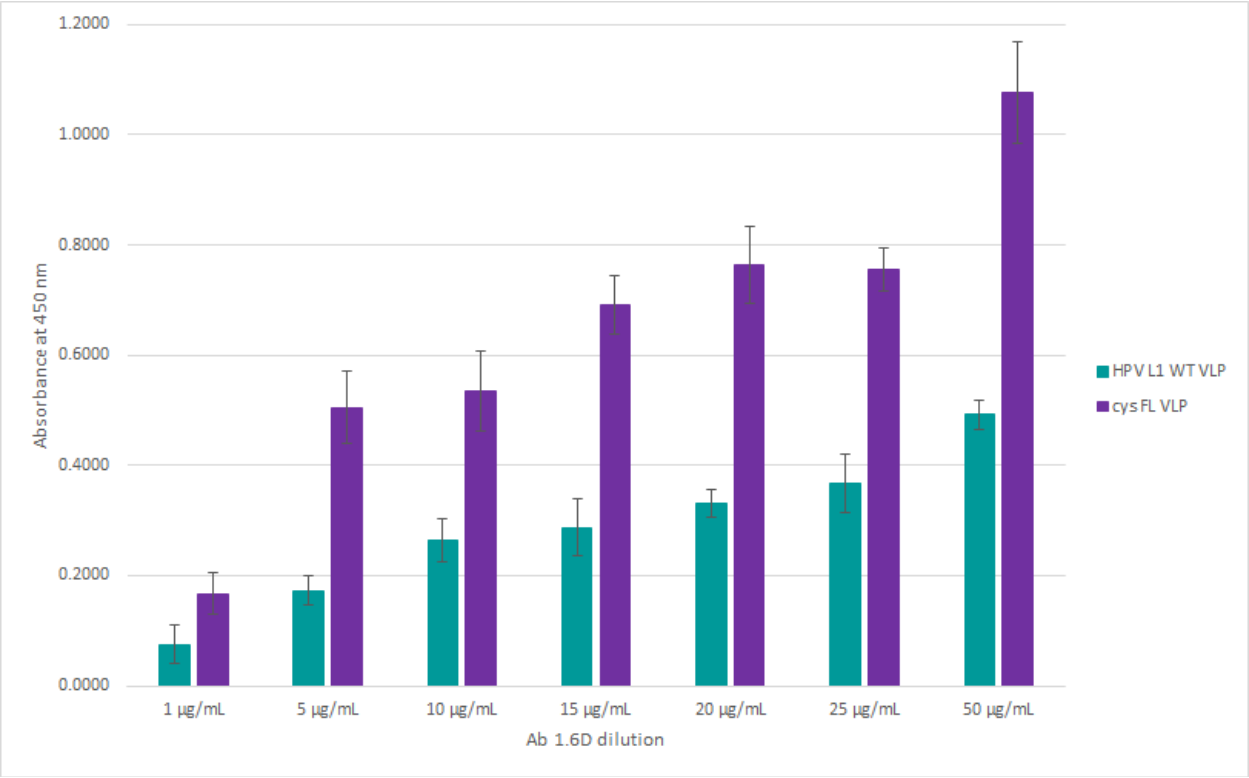
Supplementary Figure 2. An illustration of greek-key, jelly roll protein folding pathway and assembly of the capsomeres and VLPs, showing loop intercalation. (a) Unfolded state. (b) Intermediate 1, a long hairpin. (c) Intermediate 2. Greek key step 1. (d) Finished jelly roll after greek key step 2, monomer. (e) Loop intercalation in pentameric, capsomere state. (f) Invading arm (hook) connects one capsomer to five, creating a group of six. (g) 12 groups of six assemble to form the complete capsid.



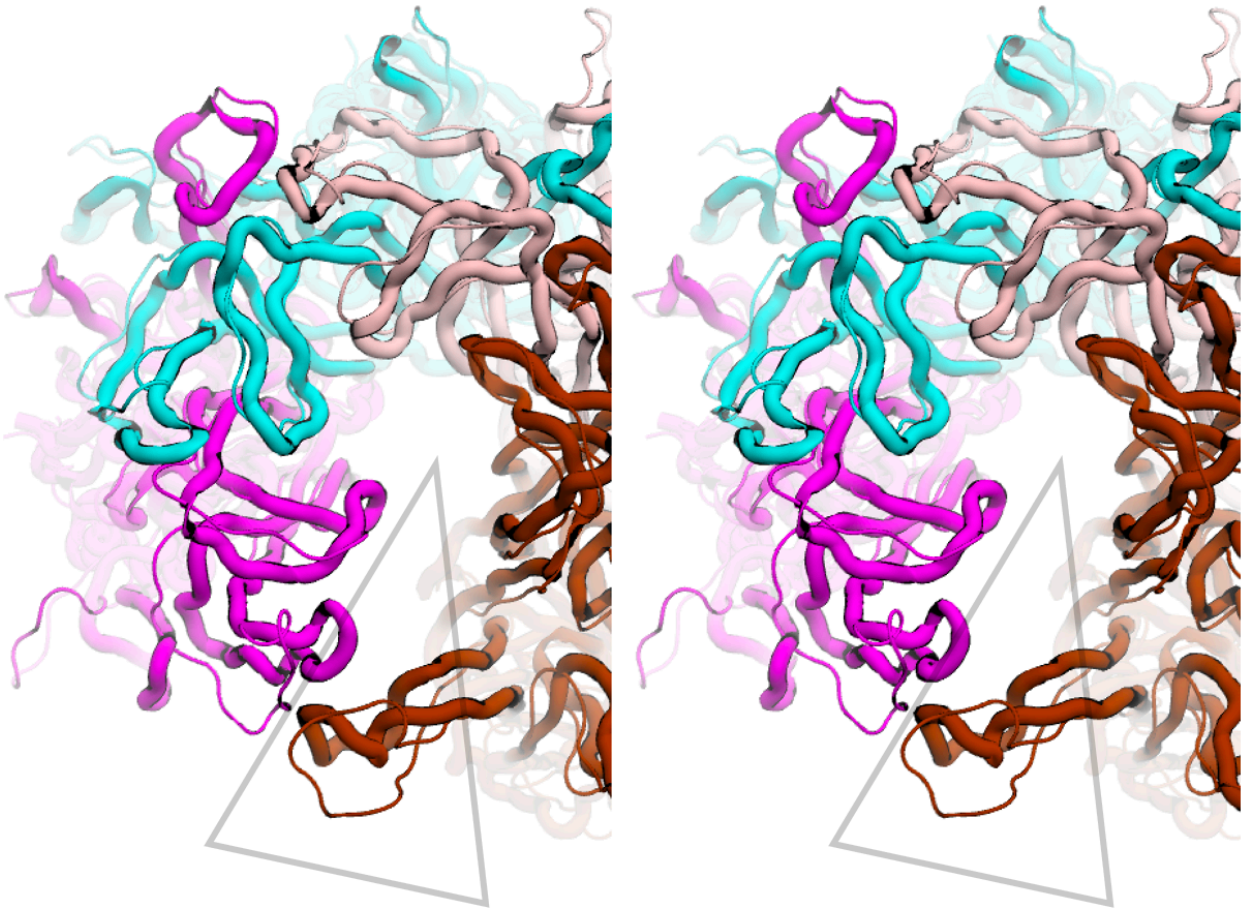
Supplementary Figure 3. (a) Modeled cysFL insertion, near the 5-fold axis of the capsomere. Fusion loop in black ribbons. (b) Blow-up of CGP/GPC motif region (green). Stereo images.



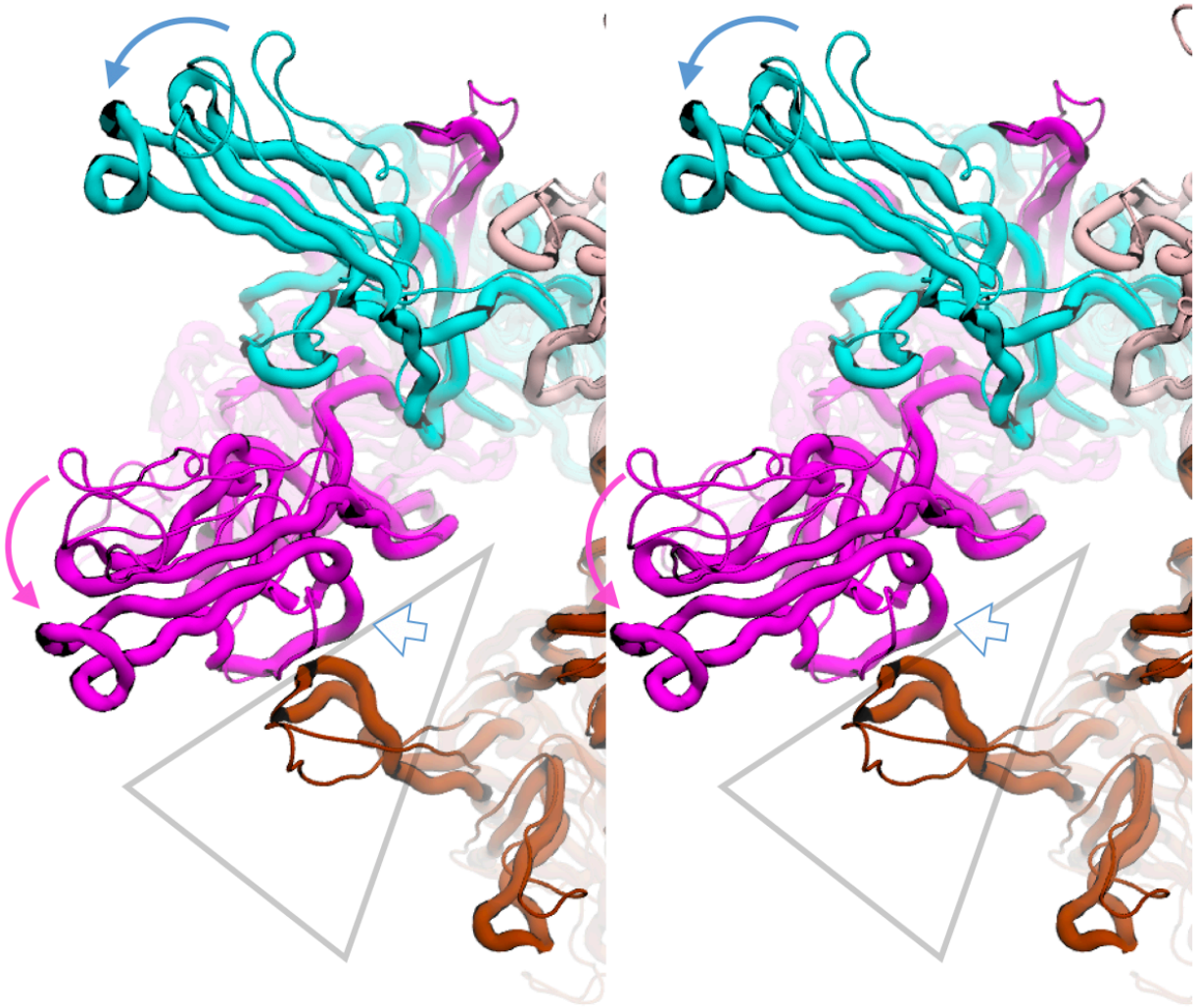
Supplementary Figure 4. Relative staining of HPVL1-WT and HPVL1-FL by conformationally sensitive 1.6D monoclonal antibodies to the FL. Equalized quantities of assembled WT L1 and cysteine locked FL L1 VLPs were probed by ELISA using a conformationally-sensitive human monoclonal Ab against the fusion loop, 1.6D. This Ab specifically detected the FL containing chimeric VLPs, but reacted much less with the WT L1 VLPs.



Supplementary Figure 5. Wild-type HPVL1 tetramer. HPV16_L1wt was simulated as a tetramer for 100ps at 250K in vacuo. Missing monomer indicated by triangle. The tetramer closes slightly during the simulation . See Supplementary Video 1.



Supplementary Figure 6. HPVL1-cysFL tetramer. HPV16_L1_cysFLwas simulated as a tetramer for 100ps at 250K *in vacuo*. Missing fifth monomer indicated by triangle. FL is inserted into the DE loop. The neighboring FL domains (magenta, cyan) shifted slightly (curved arrows) to the outside, as the EF loop shifts into the empty space (open arrow). See Supplementary Video 2.



Supplementary Results 1. **Tetrameric incomplete capsomere molecular dynamics simulations**

In vacuo molecular dynamics (MD) simulations of capsomeres and partial capsomeres of L1-WT (32650 atoms) and L1-FLcys (42835 atoms) were carried out in MOE (Chemical Computing Group, Montreal, Canada) using the Amber10EHT force field, with a timestep of 2 femtoseconds using a Nosé–Poincaré–Anderson thermostat (29). Water molecules were not present. Implicit solvation was done using the reaction field method to scale electrostatics. The extended FL (1 copy, 51 residues), with the CGP/GPC motif was allowed to equilibrate conformationally by fixing all of the atoms up to 3 residues before and after the first inserted loop residue. Any fixed atom more than 9Å away from any unfixed atom was set to be energetically inert to speed the calculations. Models were relaxed incrementally, starting with side chains and ending with all atoms, before running dynamics. A symmetrical pentamer (capsomere) was constructed from the equilibrated monomer and simulations were then carried out for the full capsomere with all atoms unfixed. A homo-tetramer (one monomer omitted) of each L1 construct was simulated for 100 picoseconds at 250 degrees Kelvin. The trajectories were analyzed by visual inspection in MOE. The results of these MD simulations are shown in supplemental videos 1 and 2.

We investigated possible steric effects of the FL in the final stage of capsomere assembly by carrying out equilibrium MD simulations on an incomplete capsomere, a tetramer, starting in its native conformation in the pentamer but minus one monomer. We hypothesized that the addition of the fifth monomer to complete the capsomere would be the slow step because of steric hindrance to monomer docking, whereas adding monomer to a trimer, dimer or monomer would have no steric hindrance. Tentatively, the MD simulation supports this hypothesis. The tetramer did not open up, which would have left an easily accessible empty pocket for the fifth monomer, Instead the tetrameric state closed slightly, with a bridge forming between the HI loop of one L1 monomer to the FG loop of the L1 monomer opposite the empty site, like the shutting of a gate across that gap. If the assembly remained stuck in the tetrameric state because the fifth subunit were sterically blocked from entry, then it would not form VLPs and would be degraded in the cell via the ubiquitin/proteasome pathway, as observed for the L1-FL constructs. But if the smaller wild-type L1 could complete the capsomere, then VLP assembly could proceed and expression levels would be higher in the "spiked" sample. Indeed, that is what we see.

Using short MD simulations such as this, without even the benefit of explicit solvent molecules, one cannot confirm or deny a molecular mechanism, but MD provides a picture that is better than a static structure for testing a hypothesis. Tentatively, we can say the tetrameric intermediate contains a loop bridge between HI and FG and that the tetramer as a whole is not energetically strained in its native conformation. Beyond that, a mechanism for this spiking phenomenon cannot be established with certainty. But inspection of the MD trajectory of the tetramer reveals two insights. First, there is apparently no way to add the fifth subunit without either breaking the FG/HI non-covalent "gate" connection that crosses the empty site, or alternatively by threading the fifth monomer up through the middle of the 4-subunit incomplete capsomere. The bulkier FL would be more difficult to thread through this hole, while the smaller loops of the wild-type L1 would pass easily. As a working hypothesis, the results suggest that the DE and FG loops can tolerate only small insertions unless the expression is spiked with WT-L1.

Second, the FG/HI "gate" may be keeping the assembly from collapsing into a tetrameric ring, thus serving as a "door stop" rather than as a gate, keeping the "door" open for the fifth monomer to join the ring. A tetrameric ring cannot be assembled into an icosahedron, which requires 5-fold symmetry, therefore evolutionary pressure would disfavor L1 proteins that form tetramers. Thus, the intercalating DE, FG and HI loops may serve the purpose of avoiding an unwanted tetrameric structure, promoting the pentamer.

1. Sturgeon, J. B., & Laird, B. B. (2000). Symplectic algorithm for constant-pressure molecular dynamics using a Nosé–Poincaré thermostat. The Journal of Chemical Physics, 112(8), 3474-3482

Supplementary Video 1. **Molecular dynamics simulation of the wild-type HPVL1 tetramer**.

Supplementary Video 2. Molecular dynamics simulation of the HPVL1-cysFL tetrameric state.