Supplemental Methods

An earlier structure of bovine papillomavirus type 1 (BPV1) (1) was appropriately scaled, band-limited to 25Å, then used as a starting model for the analysis of wt HPV16 images. Baker and colleagues have previously demonstrated that at low resolution (circa 25 Å) BPV1 and HPV1 are indistinguishable (2). Using PFT (3) model based search followed by refinement of the HPV images data produced a starting 3D model that was papillomavirus-like. Subsequent search, refinement, and expansion into larger models are described in Figure S1.

Once a stable HPV model was obtained (that is, one that did not change upon further refinement), expanding the current model by 4% produced a larger model. Then all the images were refined against both models. In each round of refinement, images with correlation coefficients below the average correlation coefficient were discarded (e.g. not used in the classification). From the remaining images for each of the (e.g. 2) models, the images were classified as belonging to that model if its correlation coefficient (cc) was larger than for the other model(s). Specifically, we required that the cc must be larger by a certain amount (e.g. 2%) to excluded images that were essentially equal. After this type of supervised classification, new models were calculated and a new set of searching and refinement was carried out for all data with each new model.

Once stable models were obtained (e.g. after a number of cycles of refinement followed by classification) then another larger (e.g. third) model was added. In this manner, we

continued to add larger models until that data did not support a larger model. Each successive model was not restricted in size, but by the results of the supervised classification procedure. Therefore, two events occurred which limited the number of wt models to 5. The first event was that sometimes two models became the same size and had identical features. In this case, the two models were merged into one model. The second event was that multiple sets of refinement caused the largest model (if greater than 633Å) to shrink, so that the largest model was not stable. Specifically, once we had 5 models, adding a 6^{th} larger model resulted in its shrinkage until it became identical to the largest 5^{th} model. When (any) two models became identical they were combined. The 5 final classes appeared stable and did not change upon further refinement and classification. Each class retained a reasonable number of images, and one could discern visible differences between these classes. The results are described in Table S1.

The mutant C175S data contained significantly fewer images, so we simplified the above refinement procedure so that we started with 3 models: the smallest, middle, and largest wt HPV (diameters: 600Å, 619Å and 633Å), and no additional models were added. After a number of cycles of refinement the first and second model had approximate the same diameter (650Å) and the same features, so these two models were combined leaving a total of two classes. These two models were refined until there was no change in size or features, and the results are shown in Table S2.

Micrographs of fully mature capsids assembled within cultured mammalian cells are, for the most part, uniform in size. However, there are occasional ill-formed capsids (see Figure S4).

Supplemental References

1. **Trus BL, Roden RB, Greenstone HL, Vrhel M, Schiller JT, Booy FP.** 1997. Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 A resolution. Nat Struct Biol **4:**413-420.
2. **Baker TS, Newcomb WW, Olson NH, Cowsert LM, Olson C, Brown JC.** 1991. Structures of bovine and human papillomaviruses. Analysis by cryolectron microscopy and three-dimensional image reconstruction. Biophys J **60:**1445-1456.
3. **Baker TS, Cheng RH.** 1996. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. J Struct Biol **116:**120-130.