Supporting Information: Assembly and Stability of Simian Virus 40 Polymorphs

Curt Waltmann,[†] Roi Asor,[‡,¶] Uri Raviv,[*,‡,¶] and Monica Olvera de la Cruz[*,†,§]

[†]*Department of Materials Science and Engineering, Northwestern University, Evanston, IL, 60208*

[‡]*Institute of Chemistry, The Hebrew University of Jerusalem, Edmond J Safra Campus, Givat Ram, Jerusalem, 9190401, Israel*

[¶]*Center for Nanoscale Science and Technology, The Hebrew University of Jerusalem, Edmond J Safra Campus, Givat Ram, Jerusalem, 9190401, Israel*

[§]*Department of Chemistry, Northwestern University, Evanston, IL, 60208*

E-mail: Uri.Raviv@mail.huji.ac.il; m-olvera@northwestern.edu

Here we a describe a few topics which supplement the main text. They include how the connection strength parameter, ϵ, was calibrated, as well as the $N = 11$ configuration, the effect of accurately representing the length of the C-terminal ligand, and changing the fraction of the charge on the N-terminal ligand akin to changing the pH of the solution. We then include a complete description of the force field potentials and parameters. The Euler formula under the constraints that only hexagons and pentagons are allowed is derived and finally we include more details on how the simulations were run.

Calibration of Connection Strength

The strength of VP1 connections, ϵ, is investigated using 6,7,8,9, and 10 k_BT. No assembly of $T = 1$ particles is observed at 6 k_BT, and very little at 7 k_BT. When we use 8 k_BT the maximal yield of $T = 1$ particles are assembled. This is the connection strength that is used in all simulations in the main text. At 9 k_BT the $T = 1$ particles are found alongside completed particles while at 10 k_BT only aggregates are found. In Figure S1, these results are summarized by a combination of simulation snapshots and distributions of bound pentamers to the templates. While, this was intended to calibrate the model experimental results suggest that this strength can be changed in different solution conditions.

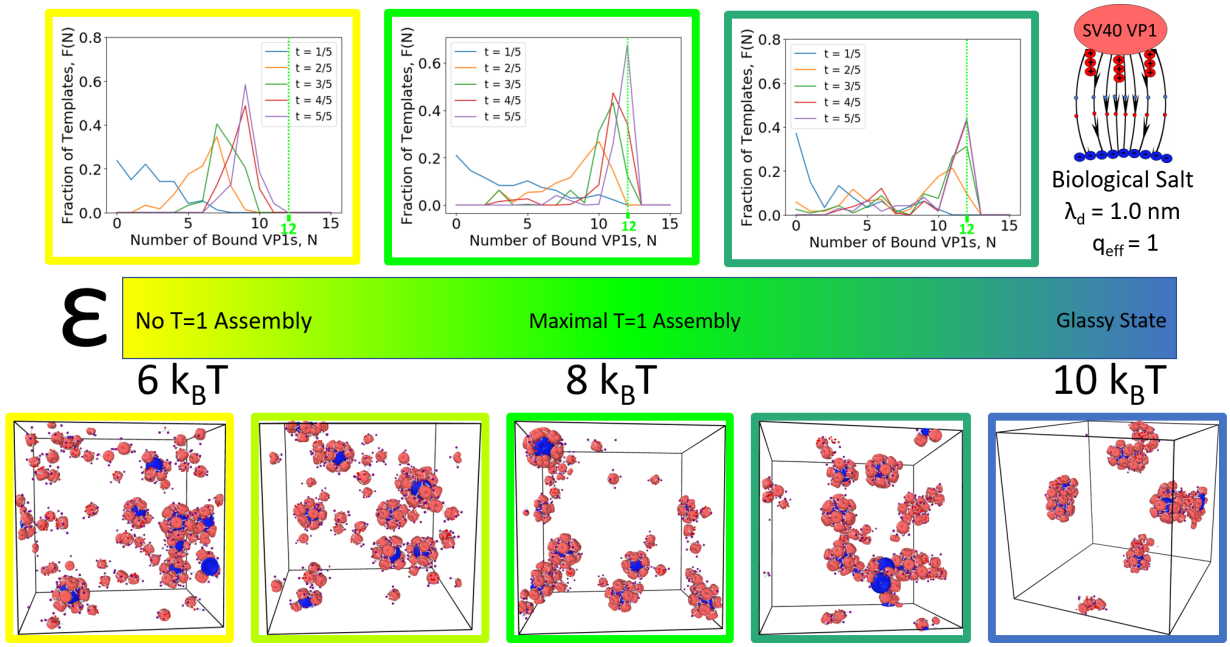


Figure S1: Assembly products at different values of ϵ under biological salt conditions ($\lambda_d = 1$ nm) is shown through distributions of bound pentamers to the templates and snapshots of the simulation box. Maximal assembly is observed at 8 k_BT whereas no assembly is observed at 6 k_BT. At 10 k_BT kinetically trapped, glassy states form. $\epsilon = 7$ and 9 k_BT show intermediate behaviors.

N-terminal Ligand Charge Deprotonation

In order to investigate the affect of pH we slowly reduce the fraction of charge on the N-terminal ligands mimicking the affect of deprotonating the charged amino acids as the pH is raised. At charge fractions near 1 no real difference is observed. Then we see a sharp decrease in the fraction of $T = 1$ particles at an N-terminal ligand charge fraction of .4, where all states $N >= 12$ seem equally probable, as shown in Figure S2. This charge fraction is the only place in the simulations where we see particles with N $\geq$ 6 losing VP1 pentamers during an assembly simulation although it was also observed when the N=13 template was decreased in size. When the charge fraction is further reduced nucleation is blocked entirely.

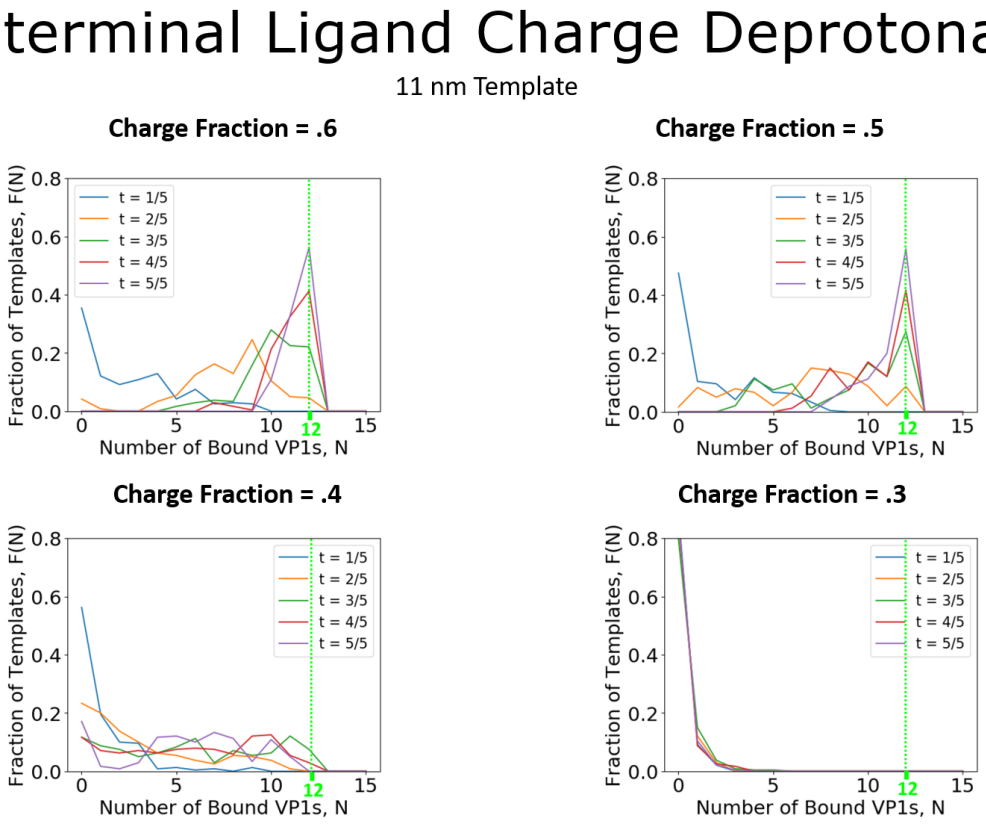


Figure S2: Distributions of the number of bound VP1 pentamers at different N-terminal ligand charge fractions. Simulations are run with $\epsilon = 8\,k_BT$, $\lambda_d = 1.0\,nm$,and an 11 nm template. The charge fraction of the N-terminal ligand is q^2_{eff}. As the fraction of charge is decreased, initially little difference is observed, followed by an increase in the polydispersity of $N \geq 10$ particles increases. Nucleation of the particles then becomes limited and eventually, at N-terminal ligand charge fraction of 0.3, the nucleation of particles is completely blocked.

Force Field

Energies are in units of k_BT and distances are in units of nanometers.

Type Mapping

Colors refer to Figure S3. The globular body is made of 21 pink 'X' beads. Two rings made of 10 beads located in a circle of radius 2 nm and one in the middle. The second ring is 4 nm above the first. All of these beads are part of one rigid body.

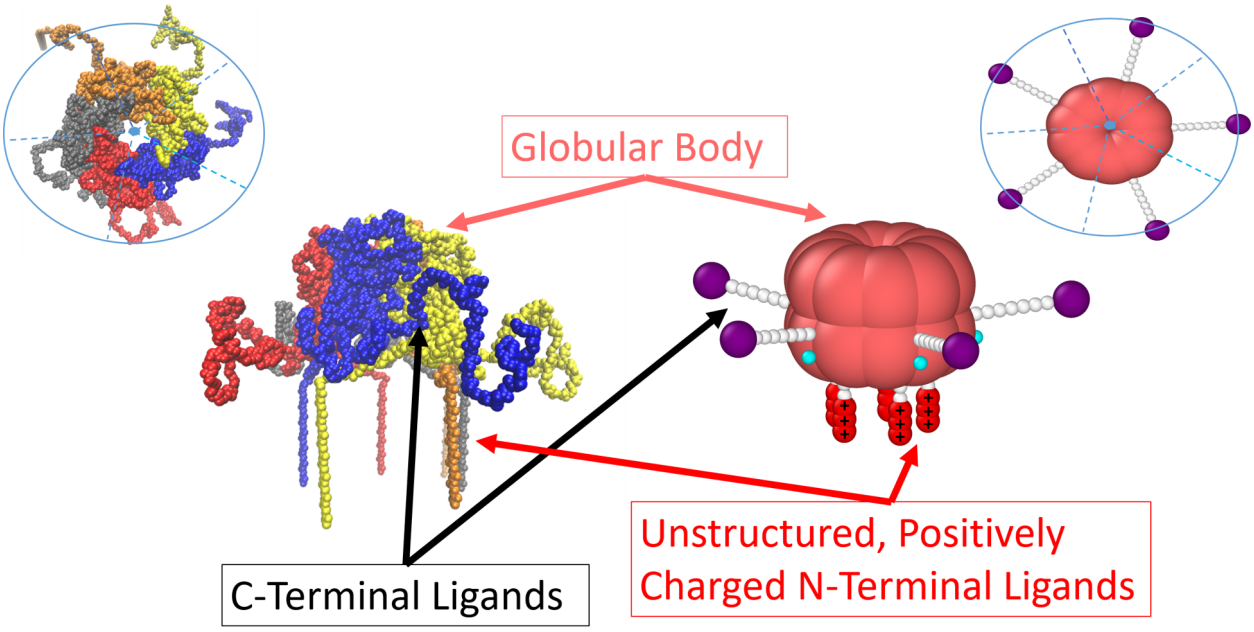
There are also 5 cyan 'D' beads representing the contact site which are located on the surface of the body 0.5 nm below the bottom ring.

The C-terminal ligands are made of 9 beads, 8 white 'P' beads and 1 purple 'C' bead, harmonically bonded to form a chain. The first bead is rigidly attached to the surface of the body 0.5 nm above the base ring. A series of angle potentials centered on beads 3-8 (the rigidly attached bead is number 1) keeps the rest of the beads nearly perfectly aligned as in an $\alpha - helix$, while allowing it to leave the body without angular restriction. When the full length of the chain is included (see Part B of Figure S3), the chain is extended by two light purple 'S' beads which have an equilibrium an equilibrium bond length of 2 nm. These two beads model an extension of the chain by nearly 40 amino acids.

The N-terminal ligands are made of 5 beads, 2 white 'P' beads followed 3 red 'qPp' beads, all harmonically bonded to form a chain. The first bead is rigidly attached to the surface of the VP1 2 nm directly below the bottom ring.

Templates are made of 'qPm' beads, which are uniformly distributed on the radius of a sphere at a packing density which makes them impenetrable to other beads. They also contain a 'center' bead located at the exact center of the rigid template.

A. Original Model of VP1 Coarse Grained Structure



B. Updated Model Including Full C-terminal ligand Length

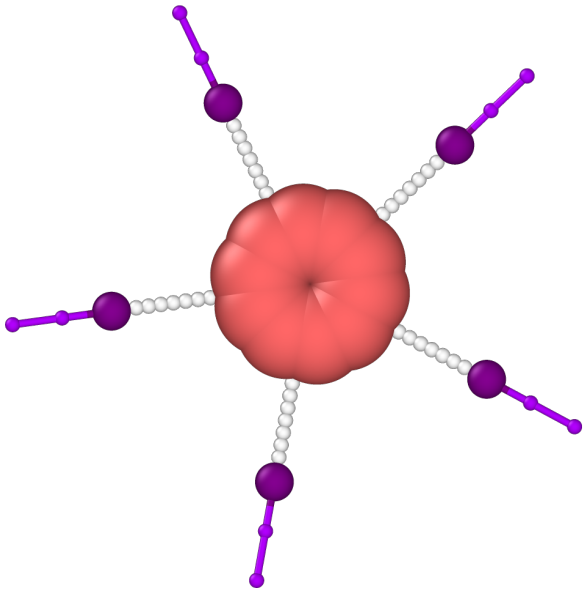


Figure S3: Part A shows the coarse grained structure of the VP1 including the globular body, C-terminal ligands, and N-terminal ligands. A full description of the individual beads that make up these sections and the potentials and parameters, which control their functionality is described in detail in the main text. Part B shows the addition of the full C-terminal ligand length upon updating the model. The parameters and potentials for the 'S' beads that make up this section of the VP1 are also described in detail in the main text.

Volume Excluding Interactions

Volume Excluding interactions are handled through a purely repulsive Lennard-Jones interaction where ϵ=1 k_BT shown in Equation 1. Parameters are shown in Table .

$$U_{Exclusion}(r) = \begin{cases} 0 & r \geq \sigma \\ \frac{\sigma}{r}^{12} - 1 & r \leq \sigma \end{cases} \tag{1}$$

Type	σ(nm)
X	2 no matter the other type (no mixing)
C	0.5
D	0.5
D-D	2.0
P	0.5
qPp	0.5
qPm	1.0

The mixing rule is simply arithmetic average with a few exceptions. 'X' beads where σ is always 2.0 nm to avoid contact site blocking. The σ value for D-D interactions is 2.0 nm in order to saturate the contact site.

Implicit Ions

The effect of the ions is handled through two potentials as described in the main text. Both potentials are in units of KT. The first is the Debye-Huckel 2 between all beads whose names begin with 'q':

$$U_{Debye}(r, q_{eff}, \lambda_d) = \begin{cases} 0 & r \geq 3\lambda_d \\ \frac{-q_{eff}^2 l_b e^{-\frac{\lambda_d}{r}}}{r} & r \leq 3\lambda_d \end{cases} \tag{2}$$

where q_{eff} is in units of the elementary charge e, σ is the Lennard Jones σ, l_b =

$e^2/(4\pi\epsilon_0\epsilon_r k_B T)$ is the Bjerrum length (ϵ_0 and ϵ_r are the permittivity of vacuum and the relative dielectric constant of the medium, respectively)), $\lambda_d = \frac{1}{\sqrt{4\pi l_B c_s}}$ is the Debye length, which in NaCl can be approximated by $\lambda_d(nm) = \frac{0.3}{\sqrt{I(M)}}$, where I is the ionic strength of the solution.

The second potential is implemented as a potential between 'qPp' beads and 'center' beads at the center of the rigid templates. It is essentially a square well with a small linear portion over a range of 0.1 nm such that forces can be calculated in HOOMD. The full form is given in Equation 3, where R is the radius of the template and α is either 1 or 2 k_BT.

$$F_{short}(r, q_{eff}) = \begin{cases} -\alpha & r \leq R + .9 \\ \alpha * {}^{r-(R+1)}/_{.1} & R + .9 < r \leq R + 10 \\ r > R + 1 \end{cases} \tag{3}$$

Contact Site Interaction

The contact is a classic Lennard Jones Potential(Equation 4), where $\sigma = 0.5$ nm and ϵ is a parameter which represents VP1-VP1 contact strength as explained in the main text.

$$U_{LJ}(r, \epsilon) = 4\epsilon(\frac{\sigma}{r}^{12} - \frac{\sigma}{r}^{6}) \tag{4}$$

'S' bead interactions

The two 'S' beads at the end of each C-terminal ligand represent a mostly unstructured 40 amino acid sequence containing positive, negative, hydrophobic, and hydrophylic residues. Thus it can likely find some confirmation where it has a weak positive interaction with any foreign object. To model this the 'S' beads are given a 1 k_BT attractive interaction (i.e., ϵ = 1 k_BT) with all of the other beads in the system (except 'qPm') through a Lennard-Jones potential shown in Equation 4. The σ for 'S' beads is 0.5 nm and all combinations are the same as in the repulsive Lennard Jones see Table . Since we want the 'S' bead interaction

with the entire template object to be $1\,k_BT$ and not the interaction with indvidual 'qPm' beads we borrow from the strategy used in the case of counterion release. Thus there is a Lennard Jones potential between 'S' and 'center' beads, where σ = R+1 and $\epsilon = 1\,k_BT$.

Bonds

Nearly all bonds in the system are harmonic with a bond strength, k = 300 $^{nm}/_{rad^2}$, and equilibrium distance, $r_0 = 0.5\,nm$. The harmonic bond potential is shown in Equation 5.

$$U_{Harmonic_Bond}(k, r) = \frac{k}{2}(r - r_0)^2 \tag{5}$$

The only exception to this is for C-terminal ligand extending bonds containing the 'S' beads, which have use k = 1 $^{nm}/_{rad^2}$ and $r_0 = 2.0\,nm$. As discussed, this is to model the configurations of a much longer section of the chain.

Angles

The only angle potential used is a harmonic angle potential with strength, k = 900 $^{nm}/_{rad^2}$, and equilibrium angle, $\theta_0 = \pi$.

$$U_{Harmonic_Angle}(k, \theta) = \frac{\theta}{2}(\theta - \theta_0)^2 \tag{6}$$

Euler Formula for an Icosahedral Shell of Pentagons and Hexagons

The basic Euler Theorem

$$F = E - V + 2 \tag{7}$$

where F is the number of faces, V is the number of vertices, and E is the number of edges. Under the constraint that the shell is made only of 2 shapes, hexagons and pentagons we also have the following,

$$N = H + P \tag{8}$$

where N is the total number of shapes, H is the number of hexagons and P is the number of pentagons. All edges are shared between two faces.

$$E = \frac{5P + 6H}{2} \tag{9}$$

All vertices in our icosahedral shell are shared by 3 faces

$$V = \frac{5P + 6H}{3} \tag{10}$$

and the number of faces is equal to the number of shapes

$$F = N \tag{11}$$

By substituting N for F into Equation 7, using Equation 11 and then replacing E - V with equations 9 and 10

$$N = \frac{5P + 6H}{6} + 2 \tag{12}$$

And then replacing P with N - H using equation 8 and rearranging it can be shown that

$$H = N - 12 \tag{13}$$

implying that every shape added after 12 must be a hexagon to maintain the given constraints. In other words, all VP1 pentamers added to a structure where N $\geq$ 12 must be

6 coordinated.

Assembly Simulation Methods

Simulations are run by initializing 6 x 6 x 6 simple cubic unit cells with a spacing of 25 nm for a total of 216 VP1 pentamers. This concentration is much higher than what is used experimentally but is useful to accelerate the simulations. 8 templates of a given size are then added to the box in an evenly spaced fashion. Any clashes between VP1s and templates are solved by removing the involved VP1 pentamers. Simulations are run for 200 million timesteps at dt = $.004\,\tau$ using Langevin dynamics.