EUROPEAN SYNCHROTRON RADIATION FACILITY
ESRF User Office
BP 220, F-38043 GRENOBLE CEDEX, France
Delivery address: 6 rue Jules Horowitz, 38043 GRENOBLE, France
Tel: +33 (0)4 7688 2552; fax: +33 (0)4 7688 2020; email: useroff@esrf.fr; web: http://www.esrf.fr

Report for beam time SC-3126 at ESRF

Following in real time the structural changes during the assembly process of SV40 virus

Background

Viruses are sub-microscopic infectious agents that are unable to grow or reproduce outside a host cell. Viruses consist of two or three parts: all viruses have genetic information coded in either DNA or RNA; all have a protein coat (capsid) that protects these genes, and some have a lipid envelope that surrounds them when they are outside a cell. In contrast to bacterial viruses (Bacteriophages) that have an organized condensed DNA structure within the capsid, in animal, human and plant viruses the DNA or RNA organization, and in most cases the packaging structure order, as well as the forces directing assembly are still unknown. We focus in this proposal on the simian vacuolating virus 40 (SV40) that is a member of the polyomavirus family. The SV40 virus particle is small (~45nm), non-enveloped, with a capsid and viral shell structures that have been determined by cryo-transmission tomography and particle averaging, as well as crystallographic methods. The lattice symmetry of the SV40 capsid was found to be that of an irregular T=7d icosahedron of 72 capsomers with an average internal capsid diameter of 29.5nm (32nm upper limit)[1]. SV40 is a DNA virus consisting of a closed circular dsDNA of 5.2kb, stacked with 20-25 nucleosomes originating with the host cell. The nucleosomes can be approximated by cylinderers of 8.36nm in diameter[2] and 6nm in height[3]. This complex of DNA and nucleosomes is called a minichromosome. Experiments show that the minichromosome forms various self-assembled structures in different environments[4]. At high salt concentrations, the nucleusomes tend to fuse into 10 nm granules, while at lower salt concentration the granules open to 10 nm filament and then to nucleusome strings. Recent cryo-electron tomography experiments[5] show what seems like residual specific nucleusome packaging related to the icosahedral symmetry of the virus.

The SV40 virus can be self-assembled in vitro. Therefore it is an excellent model system to study DNA, RNA and protein assembly. SV40 and other viruses may also be used as delivery systems for genetic material, DNA and RNA, for research as well as for gene therapy applications. Recent studies[6] showed that in SV40 the DNA must be highly compacted for packaging. To date, the assembly process of condensed DNA and protein capsid into a virus is poorly understood.

In this project we assembled in-vitro SV40 virus capsid proteins around RNA or DNA molecules. Our main goal was to study the **process of SV40 virus-like particles self-assembly around DNA and RNA nucleotides**.

Using high resolution 3rd generation synchrotron small x-ray scattering (SAXS) source, we investigated the structure of SV40 and the mechanism by which condensed DNA or RNA and capsid protein self assemble into the mature virus-like particles.

Experimental method

This research was conducted in collaboration with the laboratory of Prof. Ariella Openheim, The Hematology department. Prof. A. Oppenheim has developed a unique SV40-based gene delivery system that is constructed *in vitro* from recombinant SV40 capsid proteins, produced in *sf9* insect cells, and plasmid DNA. The lab of Prof. Openheim provided us with DNA and RNA molecules and recombinant SV40 major capsid protein VP1 as well as other reagents and protocols that they have developed.

In-situ high resolution solution SAXS in transmission mode was used to characterize the SV40 virus assembly. The structural information was studied on length scales ranging from a nanometer to some tens of nanometers, under controlled buffer conditions.

The most important aspect of this study was the ability to follow the assembly process as a function of time, using the flow through and stopped-flow setups at ID02.

Results

Our results show that solution x-ray scattering is a very sensitive method to detect small changes in the structure of SV40 and T=7 and T=1 virus like particles . The SAXS signal of the full particle is different from the empty capsid or the virus like particle. This signal is mainly due to the form factor of the particles, related to the scattering particle's shape and size (Figure 1).

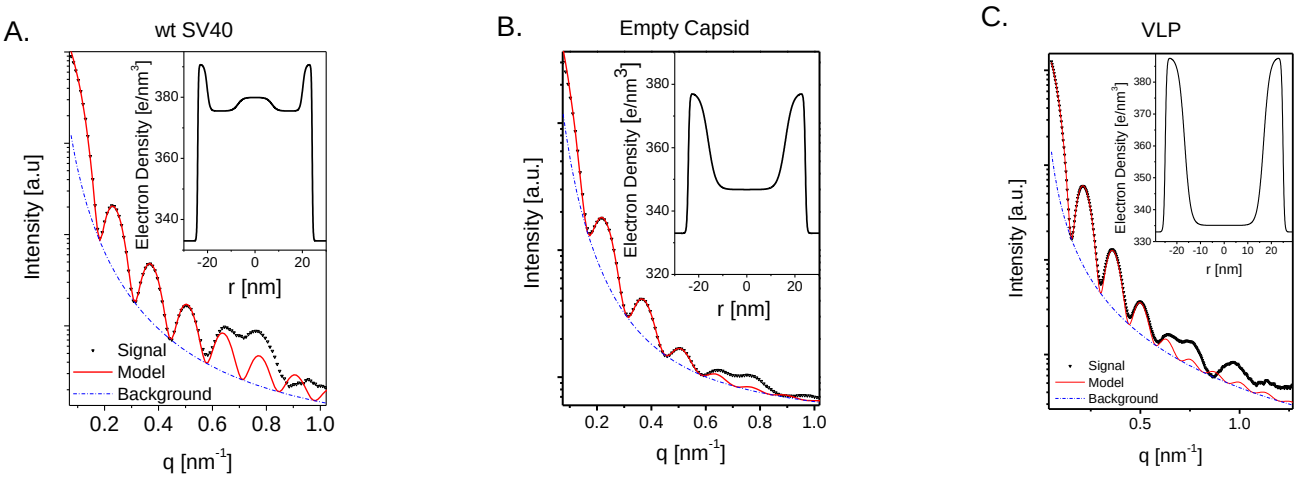

Figure 1. Radially integrated solution small X-ray scattering (SAXS) intensities (open symbols) versus the magnitude of the momentum transfer vector *q* of wt SV40 (a), empty capsid (b) and a virus like particles (VLP) (c). The solid red curves are the best fitted form-factor models of multiple spherical shells with a smoothly varying radial electron density profiles, represented by hyperbolic tangent functions (eq. 1) and shown at the inset of each plot. The broken blue curves show the assumed power-law background functions. The parameters of the three models are presented in Table 1.

Table 1. The structural parameters of the SAXS models (eq. 1).

Model parameter	wt SV40	empty capsid	VLP
R_1 [nm]	7.2	16.0	16.7
$\Delta\rho_1$ (ρ_1) [e/nm^3]	47 (380)	13 (346)	2 (335)
s_1 [nm^{-1}]	0.5	0.3	0.4
R_2 [nm]	20.1	24.4	24.9
$\Delta\rho_2$ (ρ_2) ([e/nm^3]	42 (375)	43 (376)	55 (388)
s_2 [nm^{-1}]	0.9	7.2	1.6
R_3 [nm]	24.1		
$\Delta\rho_3$ (ρ_3) [e/nm^3]	58 (391)		
s_3 [nm^{-1}]	3.0		
background	$0.044q^{-3}+0.1q^{-0.74}$	$0.076q^{-2.84}+0.566q^{-0.36}$	$0.417q^{-1.96}+0.031$

The scattering amplitude is the Furrier transform of the sample's electron density and the measured intensity is the square of the amplitude. In solution the sample is isotropic and therefore all possible orientations contribute to the scattering amplitude and this has to be taken into account. We have been able to simulate the x-ray scattering experiment to model the structure of the viral particles. By fitting the model to the data we were able to determine the size of the full and empty particles at sub-nanometer resolution (Table 1). We find that the radius of the empty particle is larger by ca. 0.3 nm, indicating that the signal is extremely sensitive to the particle size; we note that the resolution of transmission electron microscopy is an order of magnitude lower. Our data shows that we can use SAXS to determine the viron structure. Part of this data is now published in: Szekely P., Ginsburg, A., Ben Nun, T. and Raviv, U. "*Solution X-Ray Scattering Form Factors of Supramolecular Self-Assembled Structures*", Lamgmuir, **26**, 13110-13129, 2010.

To model the data, a form-factor of multiple concentric spherical shells with smoothly varying electron density profiles, represented by hyperbolic tangent functions:

(1) $$\Delta\rho(\mathrm{r}) = 0.5\left\{\Delta\rho_1 + \sum_{\mathrm{i}=1}^{\mathrm{N}-1}\left[\left(\Delta\rho_{i+1} - \Delta\rho_i\right)\right]\cdot \tanh\left[\mathrm{s}_\mathrm{i}\left(r - R_i\right)\right]\right\},$$

was fit to the data. $\Delta\rho(r)$ indicates the radial electron density contrast with respect to the solvent (buffer). The index i represents the i-th layer in the sphere, with an outer radius, R_i, an electron density contrast, $\Delta\rho_i$, and connected to the subsequent layer (i + 1) by a slope s_i. $\Delta\rho_N = 0$ and corresponds to the solvent electron density contrast. To numerically solve arbitrary electron density profiles, we transformed the function describing the radial electron density profile into a series of uniform discrete steps, as explained elsewhere (Ben-Nun, T, et al, submitted).

The next step was to follow the virus assembly process by monitoring the evolution of the SAXS signal with time following introduction of DNA to capsid protein solution (Figure 2).

Our time-dependent low resolution experiments showed that in the early stages of the assembly the particles have a diameter of about 52 nm where is in the end of the assembly process we found that the diameter is about 47 nm.

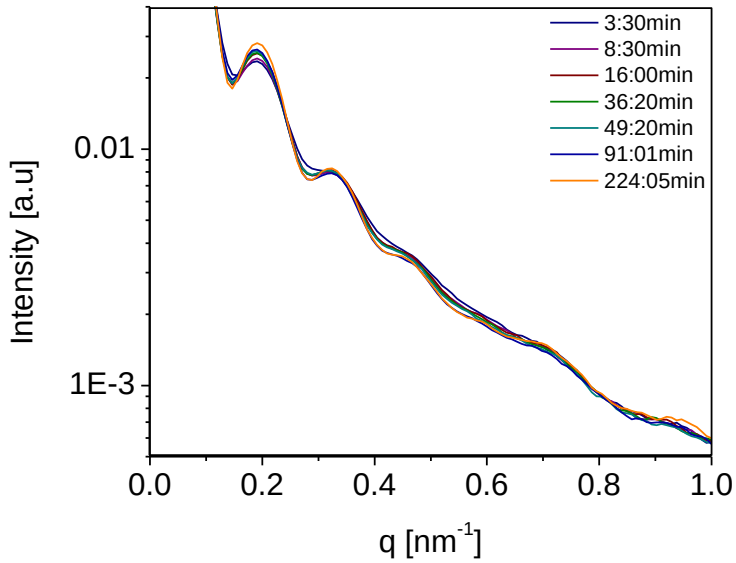

Figure 2. Time dependent radially integrated intensities of assembled DNA + VP1 in assembly buffer. The time indicated is from the mixing moment.

RNA and VP1 reaction:

In the stopped-flow setup we used two syringes for the solution of the VP1 pentamers and for the RNA solution, respectively. For each measurement 100 µl of each solution were rapidly mixed and injected into a quartz capillary. The flow of the mixture was then stopped and the samples were exposed to the X-ray beam for 5 msec, in each scan, as follows. In initial experiments, a waiting time of 30 msec after mixing was followed by measurements every 0.995 sec, to cover the first minute of the reaction. To better follow the initial reaction stage, 4 sequences of 10 measurements were taken at intervals of 200 msec after incubation times of 35, 85, 135 or 185 msec, thereby covering the first two seconds at sampling density of 50 msec. The reported times, however, are slower by up to 30 msec than the actual times as this dead-time is the maximum time it takes for the solution to mix and travel to the point at which the X-ray beam scan the sample.

Data Analysis

The 2D scattering patterns were radially integrated [5]. The scattering intensity, I, was plotted as a function of the magnitude of the momentum transfer vector, q. The data were then analyzed using the software X+ developed in our laboratory.

We measured the steady state structure of a mixture of 5 μM VP1 with 0.3 μM RNA that formed T=1 VLPs. The signals of the pure components were separately measured. Because the stoichiometry of the reaction is 12:1 the mixture contained 1.4 μM VP1 and RNA. The signal of the pure protein was normalized (to 1 μM) and subtracted from the signal of the T=1 VLPs.

TRSAXS curve fitting

When a set of time-resolved scattering curves cross at a certain q value this point is called an isosbestic point. Isosbestic points suggests a two state reaction, i.e. that a mixture of two dominating states exist throughout the reaction and only the relative amounts of the two structures changes rather than the structures themselves. Intermediate structures are likely to form but remain at low concentrations, hence negligibly contribute to the scattering signal.

Under those conditions the TRSAXS intensities can be modelled as a sum of the intensities of the reactants and the final VLP product as follows:

(2) $$I(q,t) = \alpha(t)\times I_{VLP}(q) + \beta(t)\times I_{Pentamers}(q) + \gamma(t)\times I_{RNA}(q)$$

$I_{VLP}(q)$, $I_{Pentamers}(q)$ and $I_{RNA}(q)$ are the intensities of VLPs (Figures 3), pentamers and 524 nt RNA (Figure 4), respectively, normalized to 1μM. The time-dependent coefficients α, β and γ represent the time-dependent concentrations of VLPs, VP1 pentamer and RNA, respectively. In our model we assumed the stoichiometric ratio, dictated by the icosahedra structure (T=1) of the fully assembled particle, in which a single ssRNA molecule is surrounded with 12 VP1 pentamers, enabling a model with a single free parameter:

(3) $$I(q,t) = \alpha(t)\times I_{VLP}(q) + [\beta_0 - 12\alpha(t)]\times I_{Pentamers}(q) + [\gamma_0 - \alpha(t)] \times I_{RNA}(q)$$

β_0 and γ_0 are the initial concentrations of the VP1 pentamers and RNA, respectively (there were no VLPs at the time of mixing. Each TRSAXS intensity measurement was modelled using eq. 3 (Figure 5) from which the best fitted $\alpha(t)$ was obtained.

Figure 3 shows that by mixing the 524 nt RNA with VP1 pentamers ($VP1_5$), we create virus like particles (VLPs) with an outer diameter of 24.5 nm. At low q values, the scattering intensity is well fit by a multi-shell spherical model with high electron density at the center, a lower density layer followed by a high electron density outer layer. This would correspond to a protein shell surrounding an RNA core. The particles appear to be nearly monodisperse as indicated by the adequate fit of the data at the first and third minima. The size is in good agreement with the expected size of T=1 structure. At high q values, as expected, the simple spherical multi-shell model is inadequate to reproduce the data obtained from the T=1 icosahedron.

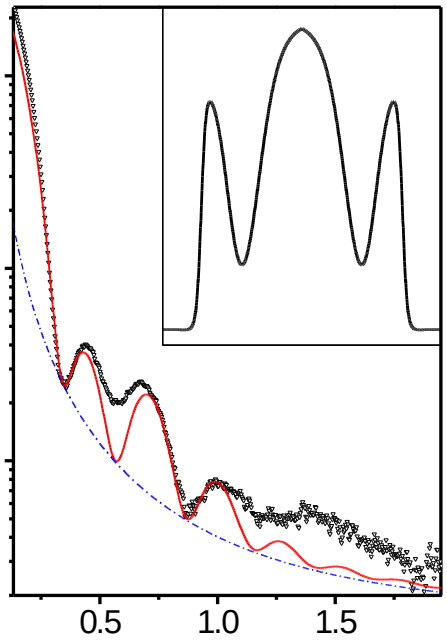

Figure 3. Radially integrated solution small X-ray scattering (SAXS) intensity (open symbols) of steady-state (T=1) RNA virus like particles (VLPs). The solid curve is the best fitted form-factor model of multiple spherical shells with a smoothly varying radial electron density profiles, represented by hyperbolic tangent functions (eq. 1) and shown at the inset. The broken curve shows the assumed power-law background given by $0.261q^{-2.061} + 0.142$. The best fitted parameters for eq. (3) are: $\Delta\rho_1 = 58.9$ e/nm^3, $\Delta\rho_2 = -18.0$ e/nm^3, $\Delta\rho_3 = 48.5$ e/nm^3 (the electron density of the buffer is assume to be similar to that of water: 333 e/nm^3), $R_1 = 5.7$ nm, $R_2 = 8.7$ nm, $R_3 = 12.2$ nm, $S_1 = 0.4$, $S_2 = 0.5$ and $S_3 = 1.7$.

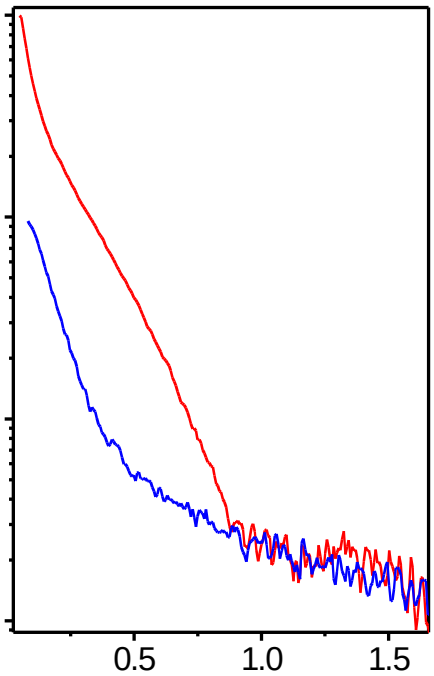

Figure 4. Radially integrated SAXS intensities of a solution of 7.5 μM VP1 pentamers (red curve) and a solution of 1 μM 524 nt RNA (Blue curve), measured using the stopped-flow setup.

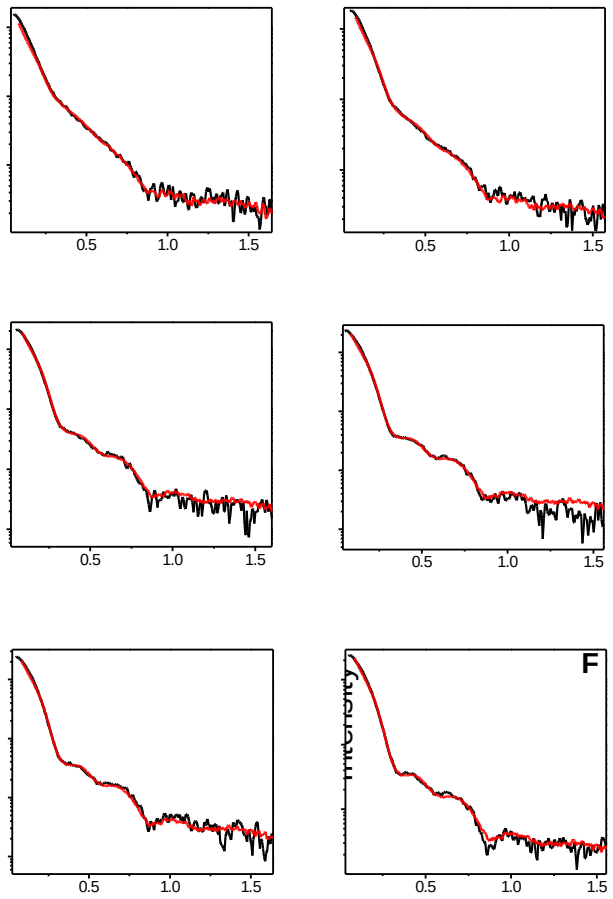

Figure 5. Examples of radially integrated time-resolved SAXS (TRSAXS) intensities (black curves) during the assembly process of the (T=1) RNA VLPs and the models (eq. 3) that best fitted the data (red curves). The initial concentrations of the 524nt ssRNA and VP1 pentamers were 0.5μM and 7.5μM, respectively. The time, *t*, elapsed after mixing the reactants is: (A) 0.035sec (B) 0.085sec (C) 2sec (D) 30sec (E) 40sec (F) 59sec.

A series of time-resolved measurements (Figure 6) of the reaction between 524 nt RNA and VP1 an isosbestic point at $q \sim 0.25$ nm^{-1}, suggesting that the system is composed of two dominating states that contribute to the total scattering intensity throughout the reaction. The scattering data of the pure reaction components (Figure 4) were normalized and the scattering data as in Figure 5 were fit to eq. 3. Based on the fit of the basis spectra, the two states that account for the time-resolved scattering curves are (1) the fully assembled icosahedra T=1 structure and (2) the non-assembled VP1 pentamers and RNA molecules.

From fitting of theses basis spectra the concentration of the RNA-filled VLPs was obtained as a function of time (Figure 7). The first 2-3 seconds of the reaction produces nearly 90% of the steady state VLP concentration. At t = 20 seconds no significant variation in the concentration of the product is detected. Repeating the reaction with slightly lower reactant concentrations, in the flow-through setup, showed that even after 4 min, there is no considerable increase in the concentration of the product (Figure 8), suggesting that the steady-state concentration of VLPs is close to 0.45 μM, under the conditions of our measurements.

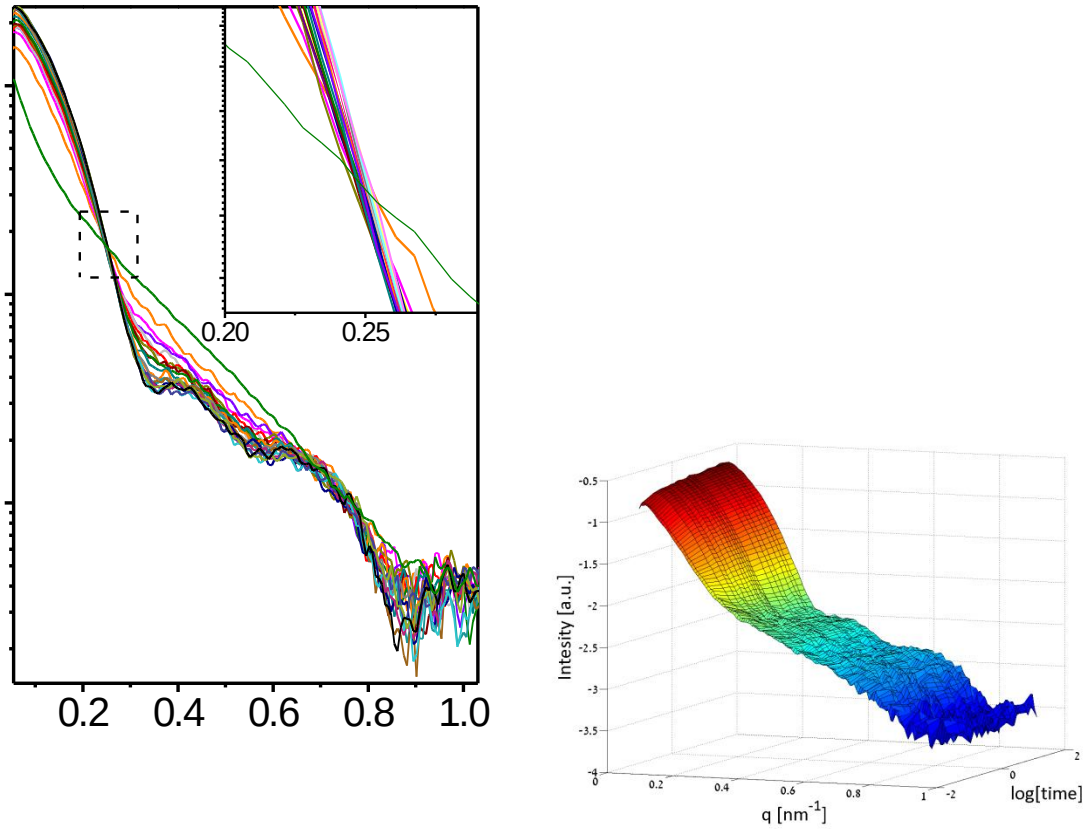

Figure 6. Radially integrated time-resolved SAXS (TRSAXS) measurements during the assembly process of the (T=1) RNA VLPs. The initial concentrations of the 524 nt ssRNA and VP1 pentamers were 0.5µM and 7.5µM, respectively. The time, *t*, elapsed after mixing the reactants is between 0.035sec (Cyan curve) and 59sec (Black curve). The signal that corresponds to $t = 0$ (Green curve) is represented by the sum of the measured scattering intensities of a solution of 7.5µM VP1 pentamers and a solution of 0.5 µM 524 nt ssRNA. The inset shows the isosbestic point at ~q=0.25 nm^{-1} on an expanded scale. A 3D plot of the TRSAXS curves, measured during the first 20 sec of the reaction, are shown at the bottom. The colors correspond to intensity.

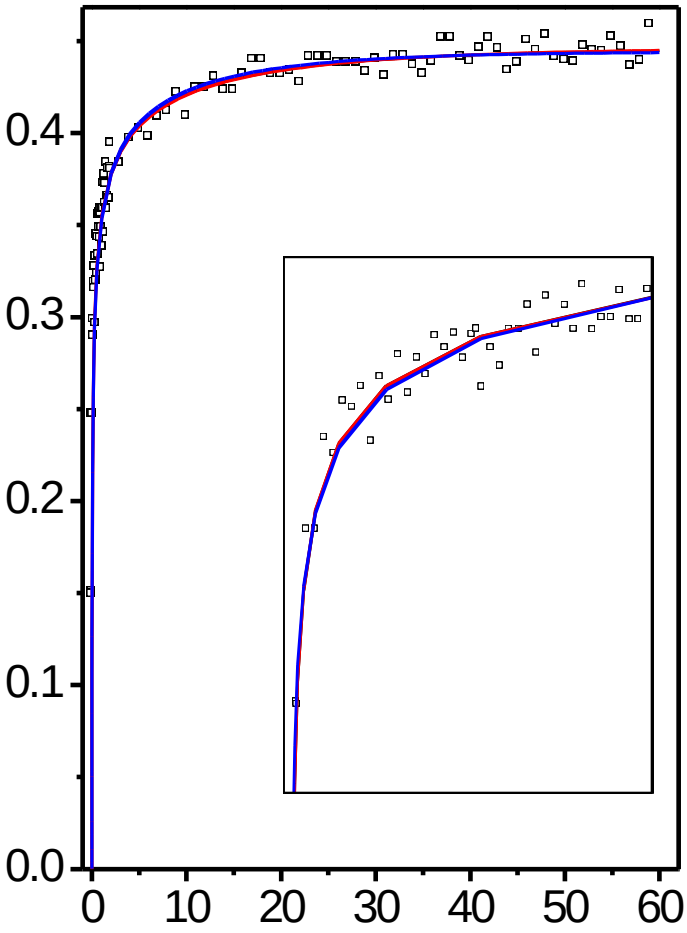

Figure 7. Concentration of formed T=1 RNA VLPs as a function of time, t, elapsed after mixing the 524 nt RNA and VP1 pentamers. The inset shows the first 2 sec on an expanded scale. The concentrations were obtained by fitting the scattering curves shown in Figure 6 to eq. 3 (see text). The reaction was modeled assuming the RNA is initially interacting either with single VP1 pentamers (blue curve), pentamer dimers (red curve), or pentamer trimers (black solid curve; not seen because of overlap with the red curve). This stage is followed by a reaction of the complex with the remaining VP1 pentamers to form T=1 nanoparticles.

We shall estimate the total free energy, ΔG_{VLP}, of the assembly reaction, given by: $12VP1_5$ + RNA → VLP. Using the steady-state concentration of VLPs close to equilibrium we get that the concentration of the $VP1_5$ at steady-state is 2.1 μM and that of the RNA is 0.05 μM. The VLP association constant is: $K_{\mathrm{A}}^{\mathrm{VLP}} = \dfrac{[\mathrm{VLP}]_{eq}}{[RNA]_{eq}[VP1_5]_{eq}^2} = 1.22\times10^{69} M^{-12}$. Therefore $\Delta G_{VLP} = k_B T \ln K_{\mathrm{A}}^{\mathrm{VLP}} = 159 k_B T$, suggesting a mean free energy of about 13 $k_B T$ per VP1 pentamer, where $k_B T$ is thermal energy.

The SAXS data indicate that assembly is very rapid with no measurable concentrations of intermediate, effectively a two state reaction. These data were compared with the SAXS kinetics data and were fitted using the initial values of 0.5 μM RNA and 7.5 μM VP1 and a final value of 0.45 μM in the form of T=1 particle. The suggested assembly mechanism included two phases of bimolecular reactions.
A nucleation step:

(1.) $n \cdot VP1_5$ + RNA $\underset{k_{b_{nuc}}}{\overset{k_{nuc}}{\leftrightarrow}}$ RNA($VP1_5$)$_n$ complex (*nucleation)*

followed by a series of (12 - *n*) elongation steps:

(2.) $VP1_5$ + RNA($VP1_5$)$_{ne}$ complex $\underset{k_D^c}{\overset{k_f}{\leftrightarrow}}$ RNA($VP1_5$)$_{(ne+1)}$complex (elo*ngation*)

where *n* is the number of VP1 pentames that assemble at the initial nucleation stage. *ne* is the number of pentamers in the forming complex during the elongation stage. Nucleation begins with the first subunit binding to RNA with a rate of k_{nuc} and a dissociation rate of $k_{b_{nuc}}$. The dissociation constant is $K_{D_{nuc}} = \frac{k_{b_{nuc}}}{k_{nuc}}$. If the nucleating complex has two or three pentamers, they bind to the RNA complex with the same rate and the same affinity (however, the affinity for the 2nd and 3rd pentamers now includes protein-protein and also protein-RNA interaction). All subsequent reactions occur at a rate of k_f and a dissociation rate of k_b. The dissociation rates are per contact and *c* is the number of contacts. The last pentamer, for example, will make five contacts and have a total dissociation rate of k_b^c. The dissociation constant is $K_D = \frac{k_b}{k_f}$. However the value of *n* is set (1,2 or 3), the model nicely fits the data (Figure 7) and the assembly is faster than the diffusion limit. The VP1 association with RNA (towards assembling a T=1 capsid) is relatively weak (Table 1).

Although the fundamental rate of assembly k_{nuc} was faster than allowable by simple diffusion, the accelerated rate, k_f, is consistent with long range electrostatic "steering". RNA acts as an antenna to attract protein.

The association of protein for construction of RNA-filled T=1 particles was remarkably weak. The first nucleation step, when a VP1 pentamer binds RNA, had a dissociation constant - K_D of about 100 µM. That means that the first step of nucleation is a very rare event. Given the electrostatics, it also suggests that there may be a slow (or rare) conformational change that limits formation of productive nuclei. Given long incubations or removal of peptide segments that inhibit binding, a much stronger K_D may be observed. This separation of contact events from tighter binding has been observed, for example, in antibody-antigen interaction. The estimated nucleation K_D for full-length SV40 pentamers also contrasts with the much higher affinity of truncated pentamer for DNA. This poor assembly of VP1 for RNA may help explain why in vivo VP1 does not package host RNA, although SV40 packaging occurs in the nucleus.

All three of the above nucleation models generate undetectable concentrations of intermediates at early times in the reaction. A maximum of less than 4% of total VP1 is present as an intermediate-sized oligomer. The absence of intermediates results in an apparent two-state assembly reaction.
The fact that the nucleation step binds more tightly than elongation events suggests that subsequent binding may require an energetically expensive induced fit. Also, after the initial dimerization step between RNA and $VP1_5$, all subsequent interactions form at least two contacts. These multiple contacts push the assembly forward.

Table 2. Constants from kinetics curve fits to the concentration of the formed T=1 particles as a function of time.

Calculation	monomeric	Dimeric	trimeric	consensus
[VP1], µM	7.5	7.5	7.5	**7.5**
[RNA], µM	0.5	0.5	0.5	**0.5**
k_f	1.81×10^{10}	7.32×10^{8}	4.3×10^{8}	$\mathbf{1\times10^{9}}$
-log(K_D)	1.92	2.20	1.89	**1.9**
k_{nuc}	9.29×10^{7}	0.22×10^{8}	0.09×10^{8}	$\mathbf{2\times10^{7}}$
-log($K_{D_{nuc}}$)	4.82	3.52	4.48	**4.4**
R_{msd}	0.68×10^{-8}	0.65×10^{-8}	0.62×10^{-8}	

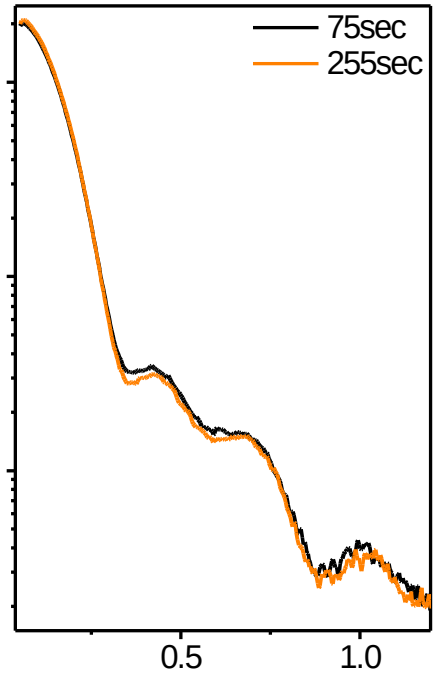

Figure 8. Radially integrated TRSAXS intensities measured after mixing equal volumes of 0.6 µM 524 nt RNA with 10 µM VP1 pentamers to form VLPs. The time, *t*, elapsed after mixing the reactants is indicated in the figure. The reaction was measured at the flow-through setup. The exposure time of each measurement was 0.1 sec.

References

1. Martin, R. G.,*Virology* 1977, 83, 433-437.
2. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J.,*Nature* 1997, 389, 251-260.
3. Dubochet, J.; Adrian, M.; Schultz, P.; Oudet, P.,*Embo Journal* 1986, 5, 519-528.
4. Varshavsky, A. J.; Nedospasov, S. A.; Schmatchenko, V. V.; Bakayev, V. V.; Chumackov, P. M.; Georgiev, G. P.,*Nucleic Acids Research* 1977, 4, 3303-3325.
5. Liddington, R. C.; Yan, Y.; Moulai, J.; Sahli, R.; Benjamin, T. L.; Harrison, S. C.,*Nature* 1991, 354, 278-284.
6. Mukherjee, S.; Abd-El-Latif, M.; Bronstein, M.; Ben-nun-Shaul, O.; Kler, S.; Oppenheim, A.,*PLoS ONE* 2007, 2.

7. Raviv, U.; Needleman, D. J.; Li, Y.; Miller, H. P.; Wilson, L.; Safinya, C. R.,*Proc. Natl. Acad. Sci. USA* 2005, 102, 11167-11172.
8. D.J. Needleman, M.A. Ojeda-Lopez, U. Raviv et al., *Phys. Rev. Lett.* 93 (19), 198104 (2004); D.J. Needleman, M.A. Ojeda-Lopez, U. Raviv et al., *Biophys. J.* 89, 3410 (2005).