Supporting Information: Glass-like Relaxation Dynamics during the Disorder-to-Order Transition of Viral Nucleocapsids

Guillaume Tresset,[*,†] Siyu Li,[‡] Laetitia Gargowitsch,[†] Lauren Matthews,[¶] Javier Pérez,[§] and Roya Zandi[‡]

†*Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France*

‡*Department of Physics and Astronomy, University of California, Riverside, California 92521, United States*

¶*ESRF - The European Synchrotron, 38043 Grenoble, France*

§*SOLEIL Synchrotron, 91192 Gif-sur-Yvette, France*

E-mail: guillaume.tresset@universite-paris-saclay.fr

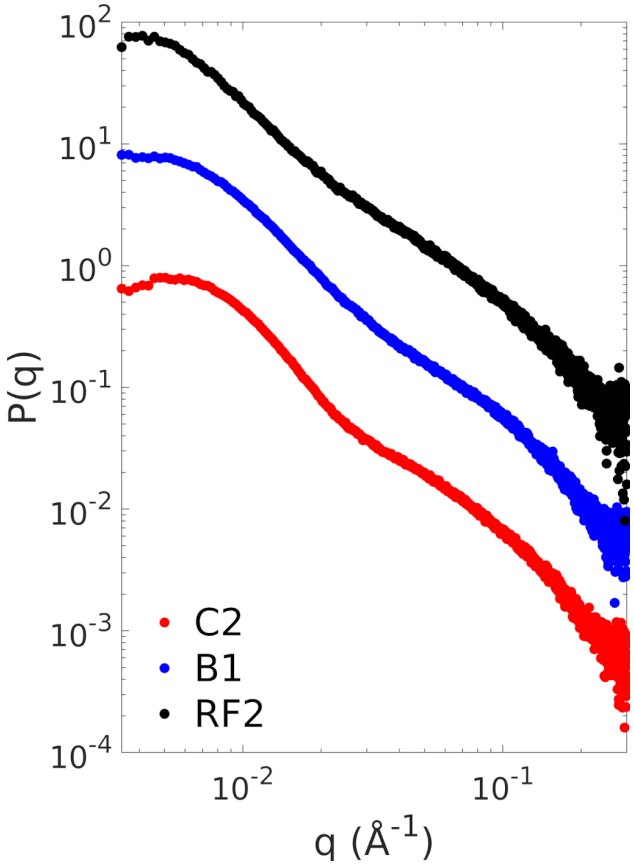

Figure S1: Scattering form factors of various RNAs: RNA C2 (red), B1 (blue) and RF2 (black) at 0.85, 0.69 and 0.60 μM, respectively. The radii of gyration obtained by the Guinier approximation are estimated to be 152, 176 and 214 Å, respectively. The curves are shifted for clarity.

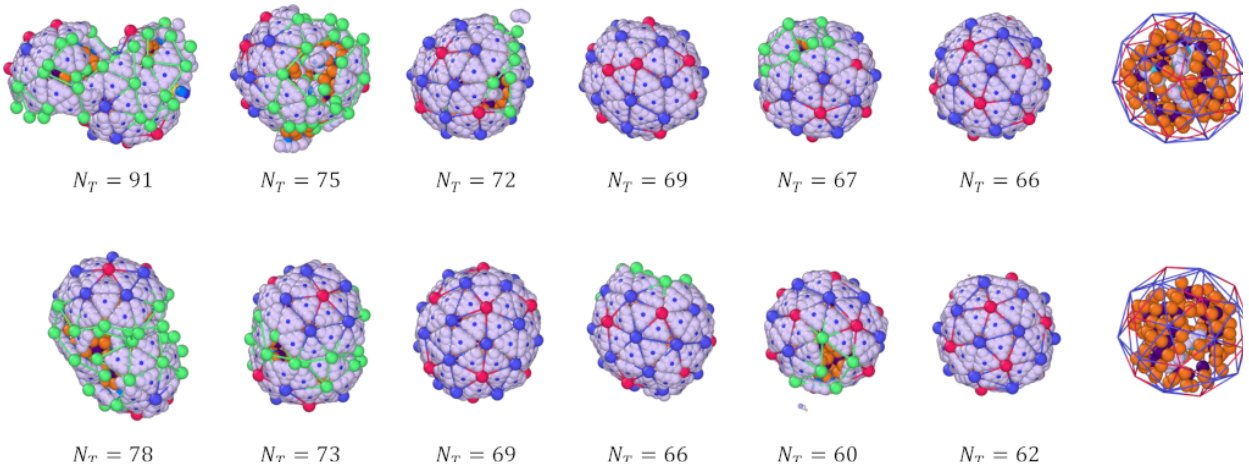

Figure S2: Molecular dynamics simulations starting from initial configuration with 78 subunits. The final structure exhibits a perfect icosahedral symmetry and the excess of subunits ($N_T > 60$) is due to the presence of the trapped subunits inside the capsid (see gray subunits in the rightmost transparent structures).

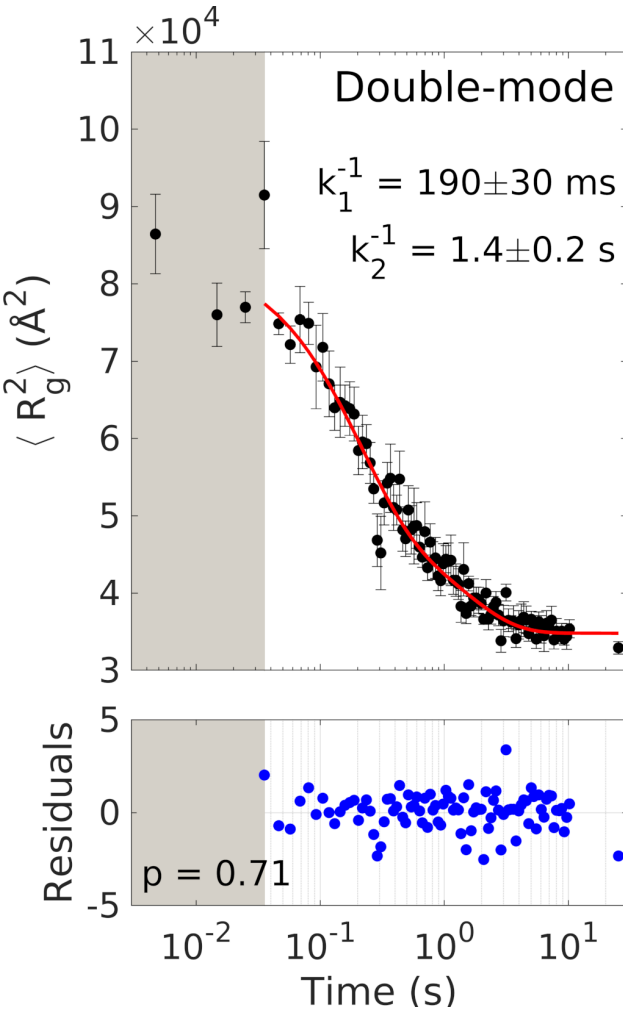

Figure S3: $\langle R_g^2 \rangle$ versus time inferred from TR-SAXS measurements with RNA C2 at a subunit concentration of 25 μM. The red curve is obtained by fitting a double-mode relaxation model, i.e., $Ae^{-k_1 t}+Be^{-k_2 t}+C$, outside the shaded area. The normalized residuals are plotted and the p-value is computed for Pearson's correlation coefficient of the normalized residuals.

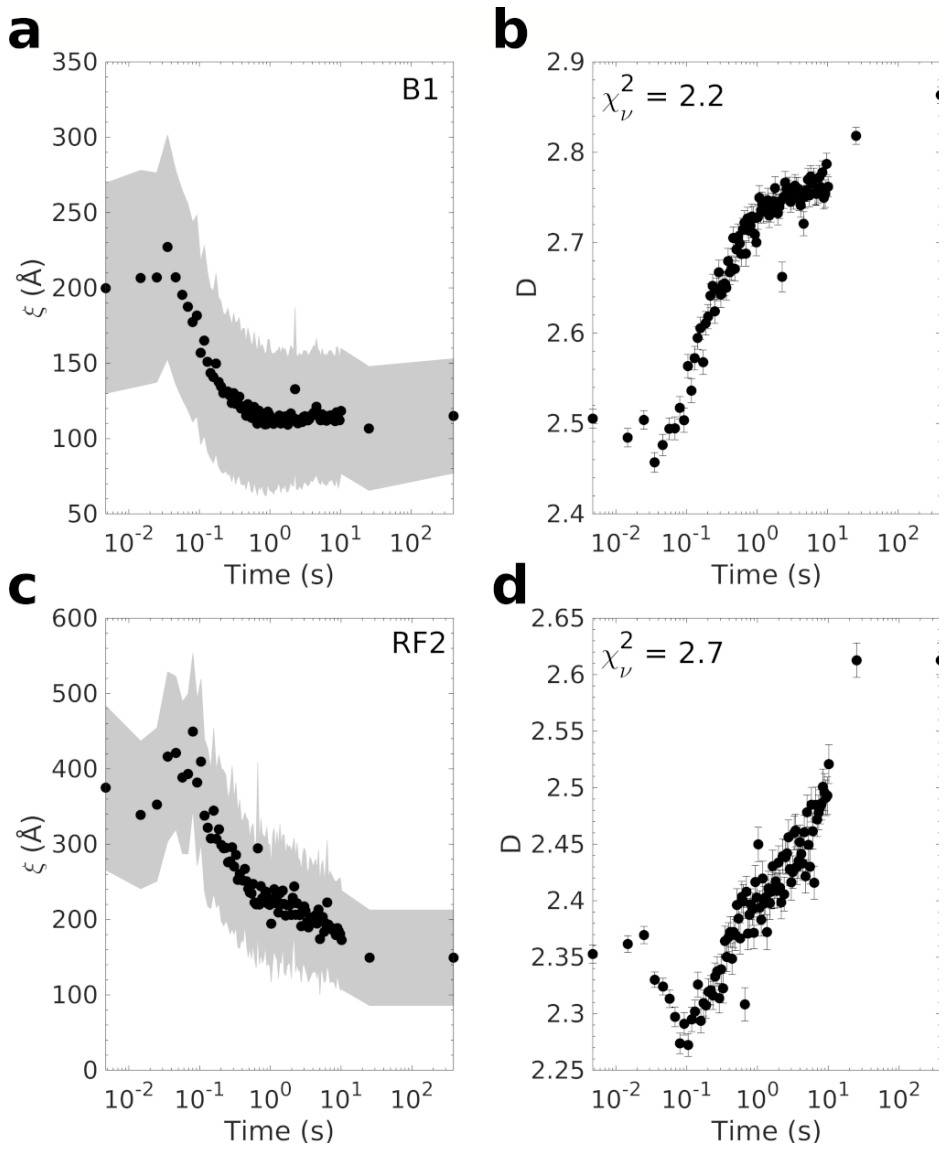

Figure S4: Fitting parameters ξ (a) and D (b) as a function of time inferred from the nanogel model on TR-SAXS patterns for nucleoprotein complexes with RNA B1 and 25 μM of subunits. The shaded area represents the width Δ_ξ of the lognormal distribution of ξ. (c) and (d) are the same as (a) and (b), respectively, but for nucleoprotein complexes with RNA B1 and 25 μM of subunits.

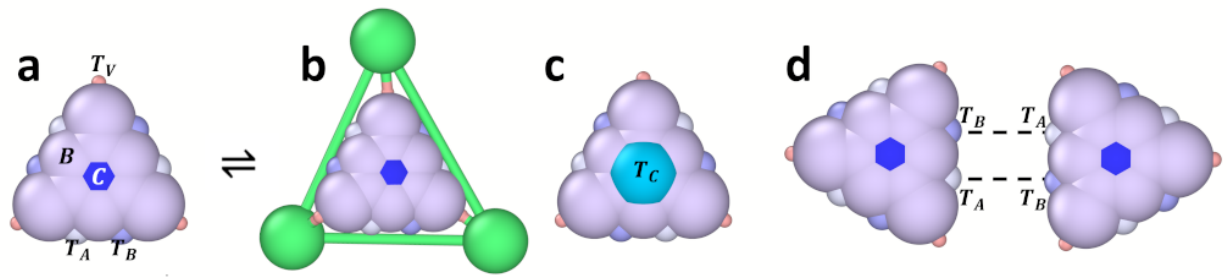

Figure S5: (a) A rigid trimer consists of 20 particles (see supplementary methods below). (b) The rigid trimers can switch to elastic ones, in which case an elastic triangular ring (green triangle) surrounds the rigid body. The process of switching between the rigid and elastic states is reversible. (c) The interior view of the rigid trimer, one particle (cyan) is positively charged. (d) The ligands on two trimers attract each other through LJ potential.

Table S1: Goodness-of-fits of the models used to fit $\langle R_g^2(t) \rangle$ extracted from TR-SAXS measurements with RNA C2 at a subunit concentration of 25 μM. See the supplementary methods for the definition of χ_ν^2 and BIC.

Model	Data n	Parameters k	χ_ν^2	BIC
Glass-like	93	3	1.0	104
Single-mode	93	3	1.9	182
Double-mode	93	5	0.97	108

Supplementary methods

Capsid protein purification

Cowpea chlorotic mottle virus (CCMV) was grown in cowpea leaves (*Vigna unguiculata*) and purified following a protocol originally proposed by Ali *et al.*[S1] Purified virions at a concentration around 5 g.L^{-1} were dialyzed overnight at 4°C against 0.5 M $CaCl_2$, 1 mM EDTA, 1 mM dithiolthreitol, 0.5 mM phenylmethylsulfonyl fluoride, 50 mM Tris-HCl pH 7.5. After centrifugation at 150,000×g for 18 h using an Optima XPN-80 ultracentrifuge (Beckman Coulter Life Sciences, Indianapolis, IN), capsid proteins were aliquoted from the supernatant and their purity was checked with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Only aliquots verifying $A_{260}/A_{280} < 0.66$ were kept, then concentrated using an Amicon Ultra-4 (10 kDa) centrifugal filter unit. Protein concentration was estimated by absorbance measurement at 280 nm. Virions were stored at -80°C in 50 mM sodium acetate pH 4.8 for months, and capsid proteins were kept at 4°C in 0.5 M NaCl, 50 mM Tris-HCl pH 7.5 until use, for no longer than two weeks.

RNA *in vitro* transcription

Plasmids containing a T7 promoter were transformed in 10-β competent *E. coli* cells (New England Biolabs, Ipswich, MA), purified with a NucleoBond Xtra plasmid DNA purification kit (Macherey-Nagel, Düren, Germany), and linearized by appropriate restriction enzyme. Transcription was carried out using a MEGAscript T7 Transcription Kit (Thermo Fisher Scientific) and freshly synthesized RNAs were purified with a MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific) prior to redispersion in ultraPure DNase/RNase-free distilled water (Invitrogen, Carlsbad, CA). RNA concentration was measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). RNA was stored at -80°C until use.

Time-resolved small-angle X-ray scattering (TR-SAXS)

TR-SAXS experiments were carried out at the SWING beamline of the SOLEIL synchrotron (Saint-Aubin, France) and the ID02 beamline of the European Synchrotron Radiation Facilities (ESRF; Grenoble, France). In both cases, the sample-to-detector distance was set to 2 m, which provided q-values ranging from 4.1×10^{-3} to 0.58 Å^{-1} in the former case, and from 3.4×10^{-3} to 0.38 Å^{-1} in the latter case. The experimental setup included a SFM-400 stopped-flow apparatus (BioLogic, France) equipped with four independent syringes with a mixing dead time of a few milliseconds. Scattering patterns were collected with a 5- or 10-ms exposure and the time interval between frames was increased following a geometric progression. The temperature was maintained at 20°C using a thermostated circulating water bath. The two-dimensional scattering images were radially averaged and the uncertainties were calculated using the packages Foxtrot[S2] at the SWING beamline and SAXSutilities[S3] at the ID02 beamline. Intensities were converted into absolute units after subtracting the contribution of the buffer solution.

TR-SAXS data were analyzed with the ATSAS suite.[S4] More specifically, radius of gyration and forward scattering intensity were determined with DATRG. Some analyses and modeling were performed with the SASfit package,[S5] notably to verify the radii of gyration estimated by DATRG and to assess various mass fractal models. The modeling of TR-SAXS data was done with custom-made codes implemented under MATLAB®.

The mean number of subunits $\langle N \rangle$ was inferred from the forward scattering intensity $I_0 \equiv I(q \to 0)$ by[S6]

$$\langle N \rangle = \sqrt{\Gamma^2 + \frac{I_0 - I_0^*}{\Delta b_{\mathrm{S}}^2 c_{\mathrm{RNA}}}} - \Gamma$$

with $\Gamma \equiv (\Delta b_{\mathrm{RNA}}/\Delta b_{\mathrm{S}}) - 1/2$ and $I_0^* \equiv \Delta b_{\mathrm{RNA}}^2 c_{\mathrm{RNA}} + \Delta b_{\mathrm{S}}^2 c_{\mathrm{S}}$. c_{RNA} and c_{S} were the total molar concentrations expressed in μM of RNA and subunits, while $\Delta b_{\mathrm{RNA}} = 0.384$ cm$^{-0.5}$.μM$^{-0.5}$ and $\Delta b_{\mathrm{S}} = 0.0201$ cm$^{-0.5}$.μM$^{-0.5}$ were coefficients estimated experimentally from solutions

of purified RNA C2 and subunits, respectively.

Fitting procedure

Model fitting was performed by least-square minimization under MATLAB®. Let $\{y_i\}$ be data measured as a function of $\{x_i\}$, and $\{\sigma_i\}$ their associated uncertainties. The total number of data is n. A model represented by a function f was evaluated through χ^2_ν defined as

$$\chi^2_\nu = \frac{1}{\nu} \sum_{i=1}^{n} \left[\frac{y_i - f(x_i)}{\sigma_i} \right]^2 \tag{1}$$

where ν is the degrees of freedom, i.e., $\nu = n - k$ with k the number of parameters entering the model f. The closer χ^2_ν is to 1, the better the model. However, this goodness-of-fit does not account well for the fact that if a large number of parameters is used, the model tends to overfit the data. By contrast, the bayesian information criterion (BIC) applies a penalty to nonparsimonious models and can be computed to decide which model gives the best representation of the data in terms of information content.[S7] It is defined by

$$BIC = \nu\chi^2_\nu + k \ln(n). \tag{2}$$

The data being fitted with a model f, the normalized residuals were calculated as follows:

$$R_i = \frac{y_i - f(x_i)}{\sigma_i}. \tag{3}$$

Pearson's correlation coefficient was computed between $\{R_i\}$ and $\{R_{i+1}\}$, and the p-value was estimated from Student's t-distribution under zero-correlation (null) hypothesis with MATLAB®.

Molecular Dynamics Simulation

Molecular dynamics simulations were performed using HOOMD-blue package, where the system was initialized with N_T trimers and one branched polymer, as shown in Fig. S2. The trimer could assume two different conformations, corresponding to the rigid and the elastic states, see Fig. S5a and b. The body of rigid trimers are composed of nine body particles (B), one particle at the center of the mass (C), one charged particle (T_N) interacting with RNA, three vertex ligands (T_V) and six side ligands (T_A and T_B), with the radii of the particles $0.5a$, $0.5a$, $0.5a$, $0.1a$ and $0.2a$, respectively. All particles in the system interact via excluded volume interactions, represented by a soft harmonic repulsive potential,

$$U^{rep}(r) = 0.5\epsilon_r(\sigma - r)^2, \tag{4}$$

where r represents the distance between any two particles and σ is the sum of their radii. The quantity ϵ_r stands for the repulsive strength, which we keep constant, $\epsilon_r = 100k_BT$, in all simulations.

Each rigid trimer edge contains two side ligands (T_A and T_B), which attract each other with a protein-protein interaction strength of $\epsilon = 10k_BT$ (Fig. S5d). The interaction is modeled through Lennard Jones (LJ) potential and can be written as

$$U_{pp}^{LJ}(r) = \begin{cases} 0.5\epsilon_r(\sigma - r)^2 - \epsilon, & \text{if } r < \sigma; \\ \epsilon((\frac{\sigma}{r})^{12} - 2(\frac{\sigma}{r})^6), & \text{if } r \geq \sigma, \end{cases} \tag{5}$$

where the repulsive part of the LJ potential is replaced by the soft harmonic repulsive potential.

An elastic trimer contains a triangular elastic ring, colored green in Fig. S5b, surrounding the rigid triangle. In our simulations as the rigid trimers approach each other, they undergo a conformational change from a rigid state (Fig. S5a) to an elastic state (Fig. S5b). This conformational change is reversible, allowing an elastic trimer to revert to a rigid trimer and

diffuse away from the complex of the genome and capsid protiens. The protein dynamics is simulated through Langevin integrator with a time step of $dt = 0.0005s$. The energy of a trimer in the elastic state involves both stretching and bending energies, which can be written as

$$E_{elasticity} = \frac{1}{2}k_s \sum_i (l_i - l_0)^2 + k_b \sum_i (1 - \cos(\theta_i - \theta_0)), \tag{6}$$

where $l_0 = 3a$ represents the equilibrium size of each trimer side with $a = 3$ nm the fundamental length of the system. The preferred dihedral angle, θ_0, is related to the spontaneous radius of curvature of adjacent subunits. The variables l_i and θ_i denote the length and the dihedral angle of bond i, respectively. The stretching modulus is $k_s = 100k_BT/a^2$ and the bending modulus is $k_b = 500k_BT$. As shown in Fig. S5b, the vertices of elastic triangles are connected to the vertex ligands (T_V) of the rigid triangles with an elastic rod whose energy is

$$E_s = \frac{1}{2}k_s \sum_i (l_i - l_{T0})^2, \tag{7}$$

with $k_s = 100k_BT/a^2$ and l_{T0} equilibrium length of the rod.

The interaction between trimers and the branched polymer is modeled through screened Coulomb interaction, where each trimer carries +9 charges on the inner side, each polymer bead carries −9 charges, and the Debye screening length is $\kappa^{-1} = 0.5a$. The branched polymer is modeled as a Gaussian chain, with a total length of $L = 120a$, and consists of 24 branch points (purple beads on Figs. 3 and S2). Each branch point has four branches, and each branch consists of one monomer making 96 monomers constituting the branches (orange beads on Figs. 3 and S2).

More details of simulations can be found elsewhere.[S8]

References

(S1) Ali, A.; Roossinck, M. J. Rapid and efficient purification of cowpea chlorotic mottle virus by sucrose cushion ultracentrifugation. *J. Virol. Methods* **2007**, *141*, 84–86.

(S2) Thureau, A.; Roblin, P.; Pérez, J. BioSAXS on the SWING beamline at Synchrotron SOLEIL. *J. Appl. Cryst.* **2021**, *54*, 1698–1710.

(S3) Narayanan, T.; Sztucki, M.; Zinn, T.; Kieffer, J.; Homs-Puron, A.; Gorini, J.; Van Vaerenbergh, P.; Boesecke, P. Performance of the time-resolved ultra-small-angle X-ray scattering beamline with the Extremely Brilliant Source. *J. Appl. Cryst.* **2022**, *55*, 98–111.

(S4) Manalastas-Cantos, K.; Konarev, P. V.; Hajizadeh, N. R.; Kikhney, A. G.; Petoukhov, M. V.; Molodenskiy, D. S.; Panjkovich, A.; Mertens, H. D. T.; Gruzinov, A.; Borges, C.; Jeffries, C. M.; Svergun, D. I.; Franke, D. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Cryst.* **2021**, *54*, 343–355.

(S5) Kohlbrecher, J.; Breßler, I. Updates in SASfit for fitting analytical expressions and numerical models to small-angle scattering patterns. *J. Appl. Cryst.* **2022**, *55*, 1677–1688.

(S6) Chevreuil, M.; Law-Hine, D.; Chen, J.; Bressanelli, S.; Combet, S.; Constantin, D.; Degrouard, J.; Möller, J.; Zeghal, M.; Tresset, G. Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte. *Nat. Commun.* **2018**, *9*, 3071.

(S7) Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **1978**, *6*, 461–464.

(S8) Li, S.; Tresset, G.; Zandi, R. Switchable Conformation in Protein Subunits: Unveiling

Assembly Dynamics of Icosahedral Viruses (preprint). https://arxiv.org/abs/2409.00226.