Supporting Information - A Kinetic Zipper Model and the Assembly of Tobacco Mosaic Virus

Daniela Kraft,[1,2] Willem K. Kegel,[1] and Paul van der Schoot[3,4]

[1] *Van't Hoff Laboratory for Physical and Colloid Chemistry, Debye Institute for NanoMaterials Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*
[2] *current address: Center for Soft Matter Research, 4 Washington Place, New York University, NY 10003, USA.*
[3] *Institute for Theoretical Physics, Utrecht University, Leuvenlaan 4, 3584 CE Utrecht, The Netherlands*
[4] *Department of Applied Physics, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

I. LIST OF SYMBOLS

symbol	description
h	energy required for conformational change of the first protein unit
g	energy liberated upon binding of protein to RNA
ϵ	energy liberated upon binding between proteins
N	number of RNA strands
M	number of protein units
q	template length; number of protein units that can adsorb to the RNA
$\lambda = Nq/M$	stoichiometric ratio
ρ_P	dimensionless number density of protein units in solution
$\rho_R(n)$	dimensionless number density of RNA with n adsorbed protein units
ϕ_P	overall dimensionless concentration of proteins present
μ_P	chemical potential of protein units
μ_R	chemical potential of RNA
$Z(n)$	intra-chain partition function
$F(n)$	dimensionless free energy of n protein aggregates adsorbed to one RNA
Ξ	semi-grand canonical partition function
$s = \exp(-\epsilon - g + \mu_P)$	measure for the affinity of proteins for RNA
$S = \phi_P \exp(-\epsilon - g)$	measure for the bare affinity of proteins for RNA
$\sigma \equiv \exp(-h + \epsilon)$	measure for the effect of nucleation and allostery
$\langle\theta\rangle$	average level of coverage of RNA by coat proteins
$P(n)$	probability for having RNA covered with n protein units
$P_{eq}(n)$	equilibrium probability for having RNA covered with n protein units
$n^* = -\ln\sigma/\ln s$	critical number of protein units on RNA above which assembly liberates free energy
$k_+(n), k_-(n)$	assembly and disassembly rates for RNA with n adsorbed protein units
k_+	constant assembly rate for RNA
$\kappa = k + (0)/k_+$	ratio of the on-rate of the first protein aggregate to that of the others
$\tau = k_+ t$	dimensionless time

II. SUPPLEMENTAL FIGURES

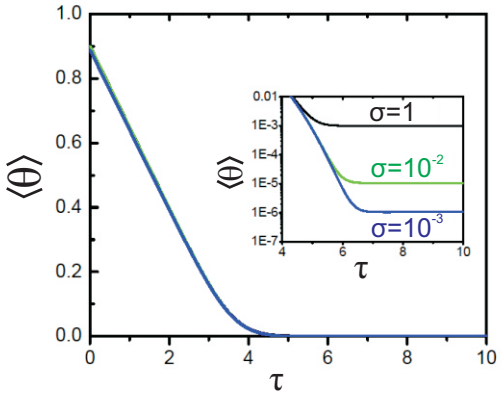

Figure S1: Disassembly kinetics for deep quenches from s=1.1586 to s=0.05923 for different values of σ.

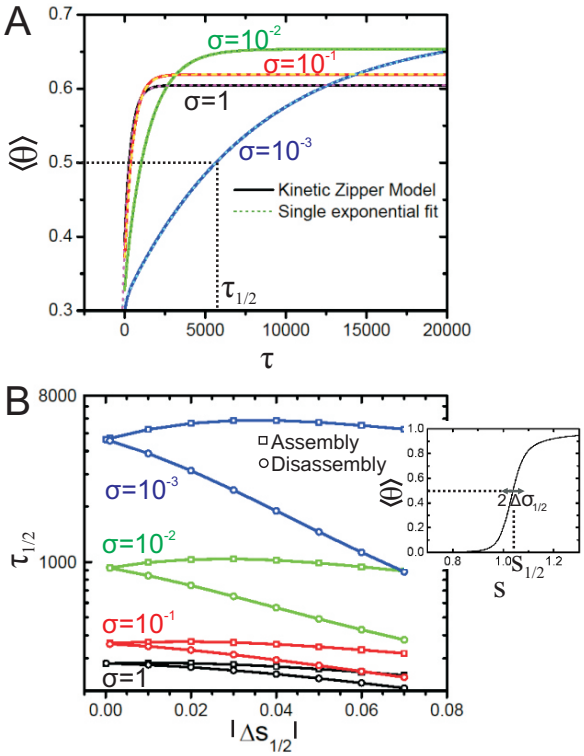

Figure S2: A) Average coverage $\langle\theta\rangle$ as a function of dimensionless time τ for symmetric quenches around $s_{1/2}$ and different degrees of allostery, σ. $\tau_{1/2}$ is the dimensionless time needed to obtain half coverage, $\langle\theta\rangle = 0.5$ ($\Delta s_{1/2} = 0.02$); B) $\tau_{1/2}$ for assembly and disassembly as a function of the symmetric quench depth $\Delta s_{1/2}$ around $s_{1/2} = s(\theta = 0.5)$ for different values of σ. $s_{1/2}$ is the value of s at which $\langle\theta\rangle = 0.5$. (Inset, $\sigma = 0.01$.)

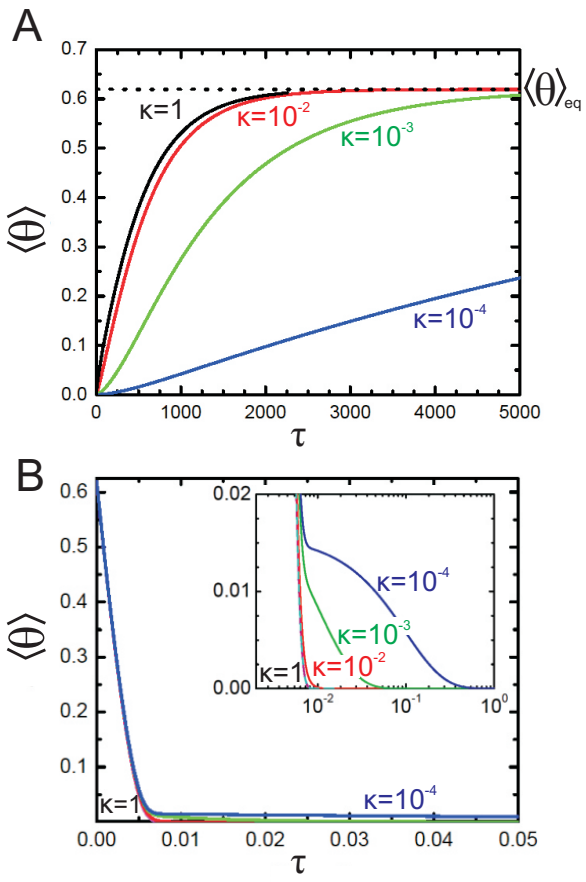

Figure S3: A) Numerical results for the average RNA coverage $\langle\theta\rangle$ as a function of time for different ratios of the assembly rates, $\kappa = k_+(0)/k_+$: assembly from the affinity $s = 10^{-4}$ to $s = 1.03$, and B) disassembly from $s = 1.03$ to $s = 10^{-4}$, and $\sigma = 0.1$.

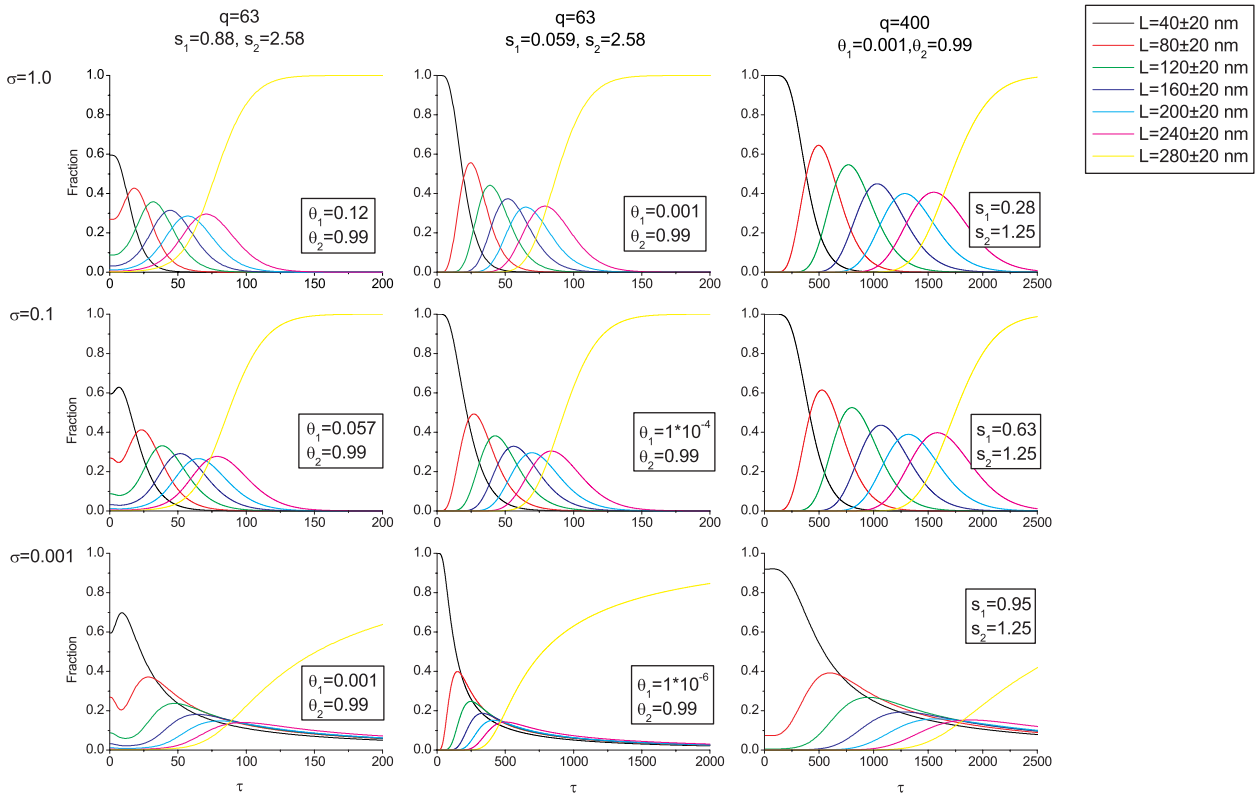

Figure S4: Numerical results from the Kinetic Zipper model for various parameters σ, q and quench starting and ending affinities s as shown in the figure. Following the experiments by Butler and Finch [1], rods shorter than 20 nm were left out, the length distributions were binned in 40nm steps and renormalized such that the probabilities add up to 1 again.

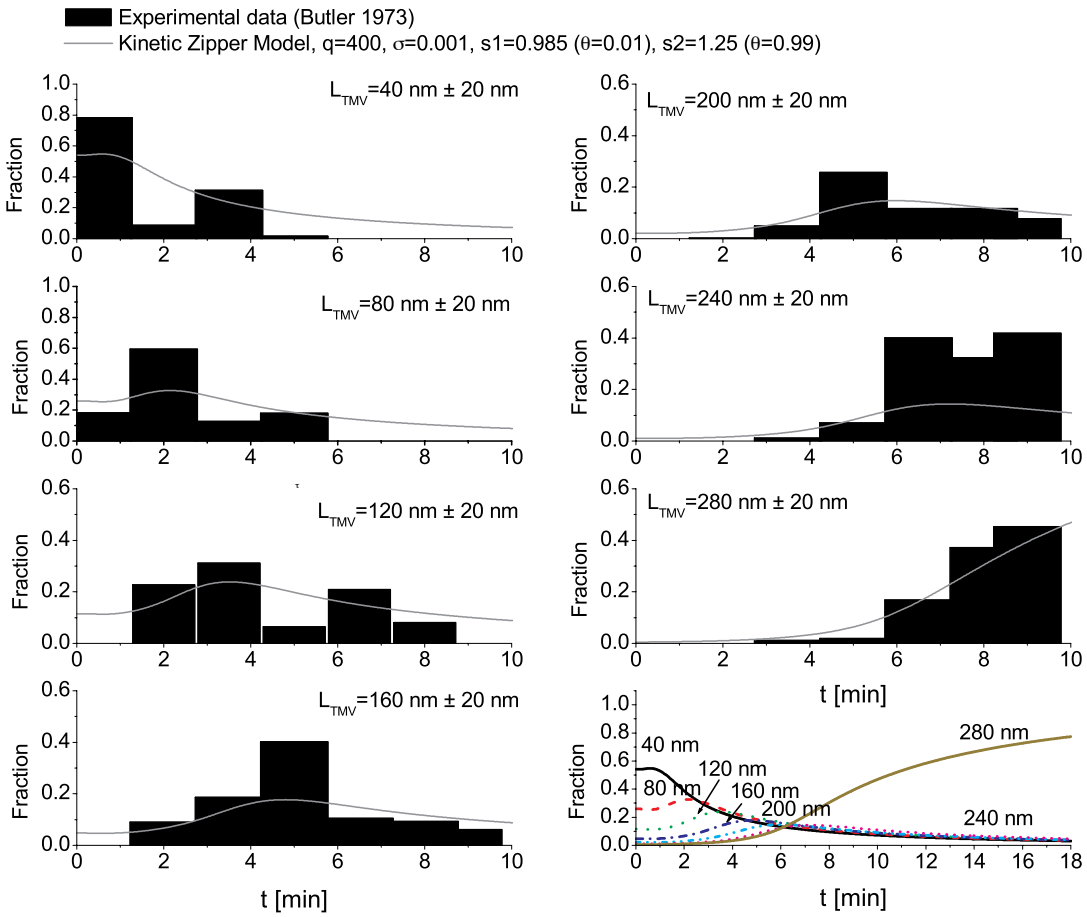

Figure S5: Comparison of experimental data from Butler and Finch[1] with numerical data obtained from our model for $q = 400$, $\sigma = 0.001$ and a quench from $s = 0.985$ to $s = 1.25$.

III. MATRIX NOTATION OF THE KINETIC EQUATIONS

The set of kinetic equations can be simplified further if written in vector notation, $\dot{\bar{f}} \equiv \partial \bar{f}/\partial t = \overline{\overline{M}} \cdot \bar{f}$, where $\bar{f} = (f(0), f(1), \ldots, f(q))$ is a vector of length $q+1$, describing the distribution of partially covered RNA molecules compared to their equilibrium values, and $\overline{\overline{M}}$ is a tri-diagonal $(q+1)\times(q+1)$ matrix describing the model dynamics,

$$\overline{\overline{M}} = \begin{pmatrix} -\kappa & \kappa & 0 & 0 & \cdots & 0 \\ \frac{\kappa}{\sigma s} & -1-\frac{\kappa}{\sigma s} & 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{s} & -1-\frac{1}{s} & 1 & 0 & 0 \\ & & & \vdots & & \\ 0 & \cdots & 0 & \frac{1}{s} & -1-\frac{1}{s} & 1 \end{pmatrix}. \tag{1}$$

IV. INFLUENCE OF THE RATE OF ADSORPTION OF THE FIRST PROTEIN UNIT

In Fig. 3, the assembly rates for adsorption of the first and subsequent protein units were taken to be equal, i.e., $\kappa = 1$ so $k_+(0) = k_+$. The observed nucleation-type assembly kinetics originates solely from the costly conformational switching of the first adsorbed protein aggregate and the melting of the RNA strand. By taking a different assembly rate for the adsorption of the first protein unit into account, implying $\kappa \neq 1$, or, in other words, $k_+(0) \neq k_+$, we find an even stronger delay of the assembly the smaller κ is, and the slower the first step is in comparison with subsequent steps. See Fig. S3A. This is of course to be expected, given our discussion in the manuscript. The disassembly kinetics of the removal of all but the last coat protein unit turns out to be unaffected by the parameter choice of $\kappa \ll 1$. This can be easily understood by considering that κ, just like σ, influences the rate of nucleation of assembly. We have seen that lowering the value of κ or σ introduces a lag time and slows down the rate of assembly. Disassembly, however, is not affected by the choice of the actual values for lower values of κ as well as σ. Only the late-stage kinetics are dominated by desorption rate of the last protein unit as shown in Fig. S3A and B. As the ratio of κ and σ occurs in the kinetic equations, it might be expected that the effect of lowering one value should be equal to increasing the other. Yet, κ also occurs in the equation for the adsorption/desorption of the first/last coat protein and this process seems to dominate over the appearance of κ in the kinetic equation for the probability of one bound protein aggregate, so kinetic and thermodynamic nucleation are not equivalent.

We also note that an increase in the rate of adsorption of the first protein unit and setting $\kappa > 1$ does *not* influence the assembly and disassembly kinetics to any discernible level. This, we believe is caused by the presence of the high-energy intermediates, and hence to be an effect of the allostery. As the assembly kinetics is already well captured by the introduction of the allosteric factor σ, it seems sensible to set $\kappa = 1$ in the remainder of our discussion.

V. QUENCH DEPTH $\Delta s_{1/2}$ AND HALF-COVERAGE TIME $\tau_{1/2}$

It is instructive to probe the assembly dynamics following a shallow quench of the affinity s around the value $s_{1/2}$, defined as that value of s for which $\langle\theta\rangle = 0.5$, see Fig. S1B. Our choice of reference value of $\langle\theta\rangle = 0.5$ is no coincidence of course, as it is the value at the transition point from largely assembled to largely disassembled conformational states. The quench $\Delta s_{1/2}$ is taken symmetrically around $s_{1/2}$, with $\Delta s_{1/2} > 0$ a quench promoting assembly and $\Delta s_{1/2} < 0$ promoting disassembly. Because a significant fraction of RNA molecules has more than a single protein unit adsorbed if the coverage is around 50%, no lag time is observed for small symmetric quenches around $s_{1/2}$. Of interest is the half-coverage time $\tau_{1/2}$, i.e., the time required to achieve a coverage of 50 %, or $\langle\theta\rangle = 0.5$, for different values of the allosteric parameter σ. See Fig. S1 A. On account of the fluctuation dissipation theorem, the relaxation time $\tau_{1/2}$ probes the regression time of spontaneous fluctuations of the coverage of the RNAs, at least in the limit $\Delta s_{1/2} \to 0$. See Fig. S1 B. We find that this half-coverage time $\tau_{1/2}$ depends on both the magnitude and sign of the quench depth $\Delta s_{1/2}$, i.e., whether we take an assembly or a disassembly quench.

For assembly, $\tau_{1/2}$ is only weakly dependent on the quench depth, at least for the (small) values probed. Even though the driving force towards assembly must be stronger with increasing quench depth, which naively should lead to a decrease in assembly times, the nucleation step apparently predominates the overall rate of assembly. See also our discussion in the manuscript. This is much less so for disassembly, where we do see a strong effect of the quench depth, that is, a strong decrease of disassembly times with increasing quench depth. As a consequence, the disassembly rates are always larger than the ones for assembly. Not surprisingly, the asymmetry in assembly and disassembly rates increases with decreasing values of σ, i.e., with increasing allostery.

We also note that the relaxation time $\tau_{1/2}$ increases with decreasing values of σ both for assembly and disassembly, as in fact is to be expected in view of our earlier findings on the influence of κ presented in Fig. S3A. When fitting $\tau_{1/2}(\sigma)$ with a power law for $\sigma < 1$, we find that the exponent of sigma approximates -1 asymptotically.

From this it seems reasonable to assume that in the limit $\sigma \to 0$ the assembly time $\tau_{1/2}$ diverges. This would correspond to the phenomenon of critical slowing down, although in our case a true phase transition occurs only in the limit $q \to \infty$. [5] Similarly, we find that the late stage kinetics of assembly and disassembly can be described quite well by a single exponential decay with a time constant τ_{exp}, as is illustrated in Fig. S1 A for shallow quenches. Just as is the case for the time $\tau_{1/2}$, the relaxation time τ_{exp} for shallow quenches scales with a power of $\sigma < 1$, where the exponent of sigma approximates -1 asymptotically.

Supporting References

[1] Butler, P., and J. Finch, 1973. Structures and roles of polymorphic forms of tobacco mosaic virus protein .7. Lengths of growing rods during assembly into nucleoprotein with viral RNA. *J Mol Biol* 78:637–649.

[2] Zimm, B. H., and J. K. Bragg, 1959. Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains. *The Journal of Chemical Physics* 31:526.

[3] Brooks, C. L., 1996. HelixCoil Kinetics: Folding Time Scales for Helical Peptides from a Sequential Kinetic Model. *The Journal of Physical Chemistry* 100:2546–2549.

[4] Zandi, R., P. van der Schoot, D. Reguera, W. K. Kegel, and H. Reiss, 2006. Classical nucleation theory of virus capsids. *Biophys J* 90:1939–1948.

[5] From classical nucleation and from helix-coil transition theories, we also expect the relevant time scale for crossing the nucleation barrier to scale as $1/s\sigma$.[2–4] This time scale also diverges in the limit $\sigma \to 0$. However, in these models a phase transition cannot occur. Note that we defined σ slightly different, namely to be linearly proportional to the σ employed in the cited models.