Confining annealed branched polymers inside spherical capsids

Alexander Y. Grosberg¹ · Robijn Bruinsma^{2,3}

Received: 28 November 2017 / Accepted: 19 January 2018 / Published online: 14 February 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract The Lifshitz equation for the confinement of a linear polymer in a spherical cavity of radius R has the form of the Schrödinger equation for a quantum particle trapped in a potential well with flat bottom and infinite walls at radius R . We show that the Lifshitz equation of a confined annealed *branched* polymer has the form of the Schrödinger equation for a *quantum harmonic oscillator*. The resulting confinement energy has a $1/R^4$ dependence on the confinement radius R , in contrast to the case of confined linear polymers, which have a $1/R^2$ dependence. We discuss the application of this result to the problem of the confinement of single-stranded RNA molecules inside spherical capsids.

Keywords Branched polymers · Confinement · Viral RNA

1 Introduction

The single-stranded (ss) RNA molecules that constitute the genomes of RNA viruses (“vRNA” molecules) have typical sizes in the range of 10^3 to 10^4 nucleotides with a solution radius of gyration in the range of 20–30 nanometers. Large ss RNA molecules of this type

✉ Alexander Y. Grosberg

¹ Department of Physics and Center for Soft Matter Research, New York University, 726 Broadway, New York, NY 10003, USA

² Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA

³ Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA

have intricate secondary structures where the nucleotides pair with each other according to the Watson-Crick rules in such a way that the molecule as a whole looks like a branched polymer [1]. This opens the possibility of applying the statistical mechanics of branched polymers to describe vRNA molecules [2], somewhat along the lines of the pioneering work by de Gennes [3].

Recently, a number of papers examined the confinement of branched RNA molecules in the context of virus self-assembly [4–9]. These papers concentrated on the role of electrostatic interactions between the negatively charged RNA genome molecules and the positively charged tails of capsid proteins. Indeed, the electrostatic affinity between capsid proteins and RNA is believed to be an important agent driving viral self-assembly and stability [10]. Importantly, the net negative RNA charge considerably exceeds the net positive charge of the capsid protein tails [10] and it was noted in computer simulations [4] that this overcharging effect is significantly enhanced by nucleotide pairing in RNA. In other words, the effect is significantly stronger for a branched polyelectrolyte as compared to a linear one. Subsequent studies confirmed that the efficiency of encapsidation is strongly dependent on the degree of branching [5–8]. This motivates us to consider the role of branching separately from the intricacies of the heavily screened electrostatic interaction, with all its concomitant non-linearity and correlations, and to examine the encapsidation of branched polymers with monomers that interact via a simple short range potential. We will focus here in particular on the effects of *configurational entropy* on the confinement of branched polymers.

Branched polymers are classified as “annealed” or “quenched”. Annealed branched polymers fluctuate over a range of branching configurations, while the branching structure is fixed for quenched branched polymers. Properties of these two classes of branched polymers are significantly different. It is not clear whether vRNA molecules are treated more appropriately as annealed or as quenched branched polymers. Numerical folding algorithms [11] applied to vRNA molecules encounter a large number of secondary structures within $k_B T$ of the groundstate [12]. Also electron microscopy images of vRNA molecules in solution suggest that they do not have a unique secondary structure [13]. This suggests that vRNA molecules are best viewed as annealed branched molecules. On the other hand, electrophoresis experiments indicate that vRNA molecules are not undergoing a thermal average over the different secondary structures on laboratory time scales [14], which would suggest that a quenched description is more appropriate. One might also hypothesize that vRNA should be viewed as a peculiar intermediate case between purely annealed and purely quenched limits. We think that a useful step in that direction would be a more detailed investigation of both annealed and quenched branched polymers placed in a virus-like environments, such as a confining capsid.

In a previous paper [12] we applied Flory theory to estimate the work of confining a *quenched* branched polymers into a spherical container. In the current paper we consider the question of the work of confinement of *annealed* branched polymers. For that purpose, we adopt and present in a simplified pedagogical form the mathematical apparatus developed earlier [15, 16]. We begin focusing on the artificial case where there are no interactions between the monomers unless they are directly connected with each other as part of the polymer. We then include these so-called excluded volume interactions using Flory theory, following the same line of argument as for the case of quenched branched polymers in ref. [12]. Flory theory permits us to look at the enthalpic aspect of confinement under both good and bad solvent conditions. Our conclusion is that conformational entropy is *always* a sub-leading factor for the confinement compared at enthalpy. In other words, the confinement free energy of annealed branched polymers is always enthalpy dominated.

2 Confinement entropy and Lifshitz theory

2.1 Constituent equations

We begin by reminding the reader that the problem of confining a linear polymer in a potential well can be described by a form of the Schrödinger equation. Qualitatively, this is not surprising because a linear polymer chain is mathematically similar to a random walk and the probability distribution of the endpoint of a random walk obeys a form of the diffusion equation. In turn, the diffusion equation is mathematically similar to Schrödinger equation. Our goal here is to develop a similar description for branched polymers. To do so, we first review the case of the use of the Schrödinger equation for a linear polymer following a method that is convenient to generalize to the case of a branched polymer.

We construct the linear polymer by the recursive procedure that is depicted in Fig. 1, panels A and B. Panel A shows the two types of building blocks. There is one homodimer shown as open circles, while there are many heterodimers shown as pairs of open and closed circles. The connection between the two ends of a dimer is an ideal Gaussian polymer with no excluded volume interaction. In other words, the statistical weight of the connection between points $\mathbf{r}$ and $\mathbf{r}'$ is $g(\mathbf{r} - \mathbf{r}') = (D/2\pi a^2)^{D/2} \exp\left[-\frac{D(\mathbf{r}-\mathbf{r}')^2}{2a^2}\right]$, with D space dimension and a “bond length”.

Panel B shows how a linear polymer chain can be constructed by attaching heterodimers to a growing chain. The homodimer acts as a seed or primer that initiates the polymerization reaction. The construction rule is that two dimers can only be connected by conjugating a filled monomer to an empty monomer. Together they become a single filled monomer. The final result is a linear polymer of $N - 2$ filled monomers with 2 empty monomers as endpoints. Apart from the Gaussian springs that link the monomers together, they are also subject to an external field that represent the confining capsid. Empty monomers will be subject to a field $\varphi_1(\mathbf{r})$ and filled monomers to an external field $\varphi_2(\mathbf{r})$. Both are expressed in units of the thermal energy (i.e., we set $k_B T = 1$). We will represent the capsid as strongly confining potential.

The partition sum of a polymer of N monomers with the N th monomer fixed at $\mathbf{r}$ can be conveniently written as $(1/p^*)^N \psi(\mathbf{r})$. Here, $\psi(\mathbf{r})$ is a positive, normalized function that represents the spatial distribution of the end monomer while $(1/p^*)^N$ is the full partition function. The associated free energy is

$$\Omega = N \ln p^*. \quad (1)$$

Fig. 1 Recursive procedure to construct a linear polymer. Panel **a** shows the two types of building blocks. There is one homo-dimer with two identical monomers, shown as open circles, and many heterodimers shown as pairs of open and closed circles. Panel **b** shows a linear polymer chain obtained by the attachment of a heterodimer to the homodimer, which acts as a primer for the polymerization reaction

An important assumption that we will make is that for sufficiently large N the statistical distribution $\psi(\mathbf{r})$ of the N th monomer does not depend on the statistical distribution of the first monomer. The basis for this assumption is that for large N the chain makes many collisions with the wall of the potential well if the radius R or the well is small compared to the radius of gyration $R_0(N)$ of an unconfined polymer. Each collision has the effect of reducing the statistical correlation between the first and last monomer. It should be noted that this assumption is *not* valid for unconfined polymer chains. This assumption is equivalent to that of *ground state dominance* in the Lifshitz theory of polymers [17, Sections 6&7] (see also [15]). The ground state dominance approximation becomes exact in the $N \rightarrow \infty$ limit. For finite N , its applicability is controlled by the spectral gap between ground state and “excited” eigenvalues. In practice, as was nicely demonstrated in a recent paper [9], even rather modest values of N around 50 are usually sufficient for this approximation to provide a reasonable accuracy of a few percent.

We are now ready to carry out the recursion procedure illustrated by Fig. 1. Adding one more heterodimer to a chain of N monomers means generating a polymer of $N + 1$ monomers. The new partition sum is $(1/p^*)^{N+1}\psi(\mathbf{r})$. This same partition sum also can be constructed explicitly as follows. Let the last monomer of the N -monomer chain be located at a point $\mathbf{r}'$ with the statistical weight $(1/p^*)^N\psi(\mathbf{r}')$. In order to add a dimer to the chain we first need to replace the empty monomer at $\mathbf{r}'$ by a filled one and then add an empty monomer at location $\mathbf{r}$. Replacing an empty monomer at $\mathbf{r}'$ by a filled one at $\mathbf{r}'$ and adding an empty monomer at $\mathbf{r}$ has the statistical weight $\exp[-\varphi_1(\mathbf{r}) + \varphi_1(\mathbf{r}') - \varphi_2(\mathbf{r}')]e^{-\varphi_2(\mathbf{r})}$. To complete the construction of the dimer we need to add a connector linking $\mathbf{r}'$ to $\mathbf{r}$ with statistical weight $g(\mathbf{r} - \mathbf{r}')$. Finally, in order to obtain the partition function of the $N + 1$ -mer, we need to integrate over all $\mathbf{r}'$. The result is $\int g(\mathbf{r} - \mathbf{r}')(1/p^*)^N\psi(\mathbf{r}')e^{-\varphi_2(\mathbf{r})}d^d\mathbf{r}'$. Equating the two ways of generating a polymer with $N + 1$ monomers, we arrive at a Schrödinger type equation

$$(1/p^*)\psi(\mathbf{r}) = e^{-\varphi_1(\mathbf{r})}\hat{g}\left[e^{\varphi_1(\mathbf{r})-\varphi_2(\mathbf{r})}\psi(\mathbf{r})\right], \quad (2)$$

where $\hat{g}$ is the operator with kernel $g(\mathbf{r} - \mathbf{r}')$; in other words, $\hat{g} = \exp\left[\frac{a^2}{2D}\nabla^2\right]$, where ∇^2 is the Laplace operator. This equation was first derived in 1968 by Lifshitz [18].

An easy way to recognize that this is indeed a Schrödinger type equation is to assume that the fields φ are smoothly varying fields on length scales of order a , with resulting smooth functions ψ . In that case, we can approximate $\hat{g} \simeq 1 + (a^2/2D)\nabla^2$. This leaves us with a Schrödinger equation (2) for a bound state particle with groundstate energy eigenvalue $-(1/p^*)$ and eigenfunction ψ . Since the ground state does not change sign, the eigenfunction ψ can be chosen positive. Note that we eliminate $\varphi_1(\mathbf{r})$ by re-naming $\psi(\mathbf{r}) \rightarrow e^{-\varphi_1(\mathbf{r})}\psi(\mathbf{r})$; the eigenvalue $-1/p^*$ is thus not affected by the field $\varphi_1(\mathbf{r})$. The physical reason is that there are only two ends, which means that a potential exerted exclusively on the endpoints can only have a negligible effect on the free energy in large N limit.

We can now immediately obtain the free energy of confinement of a linear polymer without self interaction inside a spherical capsid. It is well known that a quantum particle trapped inside a deep, spherical potential well of radius R has a groundstate energy $\epsilon = \varphi_0 + C\hbar^2/(mR^2)$, with φ_0 the depth of the potential, $\hbar$ the Planck constant, m the mass of the particle and $C = \pi^2/2$ a numerical factor of the order of one. The second part is the zero-point energy of the trapped particle. The “translation” between the two versions is that the energy eigenvalue ϵ replaces $1/p^*$ and $\hbar^2/m$ replaces a^2 . If the zero-point energy is small

Fig. 2 In the given branched structure, **a**, we can choose an arbitrary “diameter”—thick line in (**b**), and view it as a decorated linear polymer

compared to the depth of the potential then the confinement free energy of the polymer translates to $N(a/R)^2$ (in units of the thermal energy).¹

Let us now consider in a similar manner the case of an *annealed* branched polymer. The first step is illustrated in Fig. 2: we take a branched structure and choose one “spanning path” through it, connecting two remote endpoints. We will view this path as a linear polymer, albeit decorated with branches along its backbone. For this purpose, we repeat the logic for the case of linear polymer discussed above, with the only additional complication that having a side branch coming from a filled monomer carries some additional statistical weight. We will denote this statistical weight by $\Phi(\mathbf{r})$, which is a measure of the local “branchiness”. Renaming $\psi \rightarrow \Psi$, to emphasize the shift from linear to branched polymer, including a factor of 2 (because the recursion step adds a trimer with *two* equivalent empty monomers, see Fig. 3), and replacing $\varphi_2 \rightarrow \varphi_3$ (because monomers along the backbone are three-functional), we arrive at

$$(1/p^*)\Psi(\mathbf{r}) = 2\hat{g} \left[e^{-\varphi_3(\mathbf{r})} \Phi(\mathbf{r})\Psi(\mathbf{r}) \right]. \quad (3a)$$

A completely new aspect for branched polymer as compared to linear polymers is the necessity to construct another equation for $\Phi(\mathbf{r})$. By construction, this function represents the statistical weight of a branched tree rooted at the point $\mathbf{r}$ of the “main chain”. Since we are dealing with an annealed system, this statistical weight represents the (statistical) sum over all possible branching structures starting from a root at $\mathbf{r}$. The understanding of $\Phi(\mathbf{r})$ as the statistical sum explains the physical meaning of this function which is decisively important

¹To help the reader, we provide here details of this “translation”. Suppose that the external potential $\varphi_2(\mathbf{r})$ is zero inside a sphere of radius $R \gg a$ and very large (infinite) outside. Replacing $\hat{g} \simeq 1 + (a^2/6)\nabla^2$, (2) becomes

$$\left(1 - \frac{1}{p^*}\right)\tilde{\psi} = -\frac{a^2}{6}\nabla^2\tilde{\psi}.$$

The Schrödinger equation inside such a potential well reads

$$\epsilon\psi = -\frac{\hbar^2}{2m}\nabla^2\psi.$$

From the known ground state energy for the quantum case, we can find p^* by replacing $\hbar^2/2m \rightarrow a^2/6$ and $\epsilon \rightarrow 1 - (1/p^*)$. This gives $1/p^* = 1 - Ca^2/3R^2$. Using formula (1), one finds a confinement free energy $\Omega = N \ln p^* \simeq Ca^2/3R^2$.

Fig. 3 Recursive procedure to construct a branched polymer with only ends and three-branches, i.e., with single-functional (shown by empty circles) and three-functional monomers (shown as filled circles). Panel **a** presents building blocks, including one dimer of two ends and many trimers with one branched point and two ends each. Panel **b** shows a branched polymer and attachment of the next trimer to it

for branched polymers. It also gives the prescription for finding it by summing over all possible branch patterns. Focusing on one particular branch rooted at $\mathbf{r}$, first of all, it is possible that this branch represents only a single empty monomer. The statistical weight in this case should be $\hat{g} [e^{-\varphi_1(\mathbf{r})}]$. In all other cases, what is connected to $\mathbf{r}$ is a branch point on which two separate and independent branched structures are rooted. This leads to the statistical weight $\hat{g} [e^{-\varphi_3(\mathbf{r})} \Phi^2(\mathbf{r})]$. Collecting terms together, we arrive at the equation

$$(1/p^*) \Phi(\mathbf{r}) = \hat{g} [e^{-\varphi_1(\mathbf{r})}] + \hat{g} [e^{-\varphi_3(\mathbf{r})} \Phi^2(\mathbf{r})]. \quad (3b)$$

As before, $1/p^*$ plays the role of the groundstate energy eigenvalue. After p^* has been determined, the non-translational part of the free energy of the macromolecule is obtained from formula (1) $\Omega = N \ln p^*$ with N the total number of monomers.²

The arguments above allow us to compute also spatial distribution of densities of end points $\rho_1(\mathbf{r})$ and branched points $\rho_3(\mathbf{r})$. Without going to the details of derivation (see [16]), the answer reads

$$\rho_1(\mathbf{r}) = C e^{-\varphi_1(\mathbf{r})} \Psi(\mathbf{r}), \quad (4a)$$

$$\rho_3(\mathbf{r}) = C e^{-\varphi_3(\mathbf{r})} \Phi^2(\mathbf{r}) \Psi(\mathbf{r}). \quad (4b)$$

Here C is the normalization constant, the same in both equations. The fact that C is the same in both equations reflects a topological property of any tree-like structure with only three-valent branches: the number of ends is always equal to the number of branched points (plus 2, which is negligible for large N). As an annealed branched polymer fluctuates and changes its branchiness, this relation between ends and branches stays fixed, because it is “topologically protected”. See paper [16] for further mathematical details.

²In the original equations derived in the work [16] and also reproduced in [15], (3) include also factors Λ_1 and Λ_3 —the fugacities of the end-points and branch-points, respectively. For the system at hand, this is not necessary, because for any tree the number of ends is always equal to the number of branched points plus two: they are the same in large N limit. In the earlier paper [15], we got rid of these factors by a simple re-naming of variables. In a more general situation, where one may have one, two- and three-valent units in the polymer, i.e., our building set includes both dimers of Fig. 1 and trimers of Fig. 3, then we need Λ_1 , Λ_2 , and Λ_3 to control relative abundance of these building blocks.

2.2 Solution

We now need to solve the pair of coupled Eqs. (3a) and (3b). We will do this following the steps illustrated in Fig. 4. The first step (see Fig. 4a) is a free branched polymer. In that case, both Ψ and Φ do not depend on position so the two equations reduce to

$$(1/p^*) \Phi = 1 + \Phi^2, \quad (5a)$$

$$(1/p^*) = 2\Phi \quad (5b)$$

where we set $\hat{g} \rightarrow 1$. The solution is $\Phi = 1$ and $p^* = 1/2 = p_{\text{free}}^*$ with a free energy $\Omega_{\text{free}} = -N \ln 2$.

The next step (see Fig. 4b) is a branched polymer confined by a deep potential whose width is large compared to the radius of gyration $R_0(N)/a \sim N^{1/4}$ of a free annealed branched polymer. To be definite, we will assume a spherical potential well that acts in the same way on both monomer types so

$$\varphi_1(\mathbf{r}) = \varphi_3(\mathbf{r}) = \varphi(\mathbf{r}) = \begin{cases} 0 & \text{if } |\mathbf{r}| > R \\ \varphi_0 & \text{if } |\mathbf{r}| < R \end{cases}, \quad (6)$$

where $\varphi_0 < 0$ is the depth of the potential well. Inside the well

$$(1/p^*) e^{\varphi_0} \Phi \simeq 1 + \Phi^2, \quad (7a)$$

$$(1/p^*) e^{\varphi_0} \simeq 2\Phi \quad (7b)$$

while Φ and Ψ are both negligible outside the well. The solution is $\Phi \simeq 1$ and $p^* \simeq e^{\varphi_0}/2$ with a free energy $\Omega \simeq \Omega_{\text{free}} + N\varphi_0$ (the approximate signs mean here that the results become exact in the limit of infinite R).

The third step is to reduce the radius R to a value that is significantly less than the unconfined radius of gyration so the polymer collides many times with the wall of the potential. This should increase the free energy by an amount proportional to N :

$$\Omega = \Omega_{\text{free}} + N\varphi_0 + N\omega. \quad (8)$$

The dimensionless quantity ω is positive since confinement must increase the free energy and small compared to one for $R/a \gg 1$. More specifically, it is of the order of the ratio of the number of collisions of the polymer with the wall and the total number of monomers N .

Fig. 4 Confinement of a branched polymer. In the leftmost figure (a), the branched polymer is far from the potential well. It has a free energy Ω_{free} . In the middle figure (b), the entire polymer is placed at the bottom of a deep and very wide potential with a potential energy per monomer $\varphi_0 < 0$. The total free energy is now $\Omega_{\text{free}} + N\varphi_0$. Finally, in the rightmost figure (c), the size of the well is reduced laterally so the polymer is compressed. The compression work is equal to ω per monomer. The total free energy is now $\Omega_{\text{free}} + N\varphi_0 + N\omega$ (see (8))

This confinement free energy per monomer is the quantity that we need to determine (see also Fig. 4).

We can eliminate p^* in favor of ω , using the fact that $\Omega = N \ln p^*$ and re-write (3a) and (3b) in the form³

$$-1 + 2(1 - \omega)\Phi - \Phi^2 \simeq \frac{a^2}{2D} \nabla^2 (\Phi^2) \quad (9a)$$

$$(1 - \omega)\Psi - \Phi\Psi \simeq \frac{a^2}{2D} \nabla^2 (\Phi\Psi) \quad (9b)$$

inside the potential well. This pair of equations must be solved with the boundary condition that both Ψ and Φ vanish at $r = R$.

2.2.1 Solution for $\Phi(x)$ in $D = 1$

The procedure must start with (9a) for the statistical weight $\Phi(\mathbf{r})$ of a tree of branched side chains rooted at $\mathbf{r}$ since this equation does not depend *directly* on the statistical weight $\Psi(\mathbf{r})$ of the main chain at that location. It does however *indirectly* depend on $\Psi(\mathbf{r})$ through the shared eigenvalue ω .

We first consider the one-dimensional ($D = 1$) case, which describes a branched polymer confined between two parallel plates a distance $2R$ apart. The origin $x = 0$ is at the midpoint between the plates. For $D = 1$, (9a) for $\Phi(x)$ has the mathematical form of the equation of motion of a *classical particle* with x/a as the effective (dimensionless) time and $\Phi^2(x) = y(x)$ as the effective position of the particle. More precisely, the equation of motion

$$\begin{aligned} \ddot{y} &= -2[1 - \sqrt{y}]^2 - 4\omega\sqrt{y} \\ &= -\frac{\partial}{\partial y} \left[\underbrace{y^2 + 2y - \frac{8}{3}y^{3/2} + \frac{8\omega}{3}y^{3/2}}_{U(y)} \right]. \end{aligned} \quad (10)$$

is that of a particle moving—without friction—in the “potential energy” landscape $U(y)$ that depends on ω . It is shown in Fig. 5. It is a monotonically increasing function of y , although the slope does become very small for $y \simeq 1$ if ω is small compared to one. This equation of motion must be solved under the boundary conditions that the particle has zero “velocity” $\dot{y} = 0$ at $x = 0$ (because of the $x \leftrightarrow -x$ symmetry), while at $x = R$ the particle “position” $y(R/a) = 0$ must be at the origin since $\Phi \rightarrow 0$ at the two plates.

By substitution, it can be checked that

$$\frac{x}{a} = - \int_{y_0}^{y(x)} \frac{dy'}{\sqrt{2(U(y_0) - U(y'))}}, \quad (11)$$

is a solution of the equation of motion such that $x = 0$ corresponds to midpoint between the plates (it can also be derived from the principle of energy conservation for the particle).

³In order to obtain this result, start from (1) and insert it into the left hand side of (8). Recall that $p_{\text{free}}^* = 1/2$ and $\Omega_{\text{free}} = -N \ln 2$ and insert this into the first term in the right hand side. One finds that $1/p^* = 2e^{-\varphi_0 - \omega} \simeq 2e^{-\varphi_0} (1 - \omega)$. Insert this into (3a) and (3b) with external potentials (6) and replace $\hat{g} \rightarrow 1 + (a^2/2D)\nabla^2$.

Fig. 5 Plot of the effective potential energy $U(y)$ for the equivalent particle equation of motion of (10) for $\omega = 10^{-4}$. The horizontal line is effective mechanical energy $U(y_0)$, which is the first integration constant of the equation of motion. For the sake of visibility, this line is shown for $\delta = y_0 - 1 = 0.4$, which is significantly larger than the actual solution

We introduced the integration constant $y_0 = \Phi(0)^2$ that is the value of $y(x)$ at the midpoint $x = 0$. The value of y_0 is determined by the boundary condition $y(R/a) = 0$ at $x = \pm R$:

$$\frac{R}{a} = \int_0^{y_0} \frac{dy'}{\sqrt{2(U(y_0) - U(y'))}}. \quad (12)$$

The dominant contribution to this integral is from the region $y \simeq y_0$, where the integrand diverges. Expanding $U(y)$ around $y \simeq y_0$ gives

$$U(y_0) - U(y) \simeq \begin{cases} \frac{1}{6}(y_0 - y)^3, & |y_0 - y| \gg \delta^2 + 8\omega, \\ \left(\frac{\delta^2 + 8\omega}{2}\right)(y_0 - y), & |y_0 - y| \ll \delta^2 + 8\omega. \end{cases} \quad (13)$$

The small parameter δ defined by $y_0 = 1 + \delta$ is a measure of the extra branchiness of the polymer caused by the confinement. In principle it has to be determined simultaneously with our other small parameter ω , which is a measure of the amount of collisions between the polymer and the potential wall.

If $y_0 - y$ is large compared to $(\delta^2 + 8\omega)^{1/2}$ then the integral in (11) diverges as $1/(y_0 - y(x))^{1/2}$ for increasing x . This is reasonable since for $\omega = \delta = 0$, the integral should diverge as it corresponds to the limit of infinite R/a . On the other hand, if $y_0 - y$ is small compared to $(\delta^2 + 8\omega)^{1/2}$ then the integral converges. Inserting $(\delta^2 + 8\omega)^{1/2}$ as a cut-off for the divergence of the first case gives

$$R/a \propto 1/(\delta^2 + 8\omega)^{1/4}. \quad (14)$$

We will need an explicit expression for Φ that can be inserted into the equation for Ψ . It is sufficient to confine ourselves to the region near the center $x = 0$ where $U(y_0) - U(y) \simeq \left(\frac{\delta^2 + 8\omega}{2}\right)(y_0 - y)$. Inserting this in (11) gives

$$\frac{x}{a} \simeq -(\delta^2 + 8\omega)^{-1/2} \int_{y_0}^{y(x)} \frac{dy'}{\sqrt{(y_0 - y')}}. \quad (15)$$

This integrates to

$$\Phi(x)|_{x \ll R} \simeq 1 + \delta/2 - C(\delta^2 + 8\omega) \left(\frac{x}{a}\right)^2, \quad (16)$$

after expressing $y(x)$ into $\Phi(x)$. The complete calculation of [15] shows that the numerical constant $C = 1/8$. We also show there, using numerical methods, that the same results apply to $D = 3$, except that the numerical coefficients differ.

2.2.2 Solution for $\Psi(x)$ in $D = 1$

We now turn to (9b) for Ψ , which in $D = 1$ takes the form

$$(1 - \omega) \Psi(x) \simeq \left(1 + \frac{a^2}{2} \frac{d^2}{dx^2}\right) [\Phi(x) \Psi(x)], \quad (17)$$

for $|x| < 2R$. If we introduce $\chi(x) = \Psi(x)\Phi(x)$ then this equation adopts the form of the Schrödinger equation

$$-\frac{a^2}{2} \frac{d^2 \chi}{dx^2} + \frac{1 - \omega}{\Phi(x)} \chi \simeq \chi, \quad (18)$$

with an effective potential energy determined by the branching function Φ . Using the parabolic form (16) and expanding for small x gives the Schrödinger equation for a $D = 1$ harmonic oscillator:

$$\left[-\frac{a^2}{2} \frac{d^2}{dx^2} + C(\delta^2 + 8\omega) \left(\frac{x}{a}\right)^2\right] \chi \simeq \frac{\delta}{2} \chi, \quad (19)$$

where it was assumed that $\delta \gg \omega$ (this will be justified shortly). The “translation” between the two problems, as for the linear polymer before, is that the energy eigenvalue ϵ equals $\delta/2$, the inverse mass $\hbar^2/m = a^2$ and the spring constant $k = 2C(\delta^2 + 8\omega)$. The unusual feature of this version of the Schrödinger equation for a harmonic oscillator is that the spring constant depends on the energy eigenvalue. Imposing the well-known ground state energy eigenvalue of the harmonic oscillator $\epsilon = (1/2)\hbar(k/m)^{1/2}$, it follows that $\omega \sim \delta^2$. Combining this in the earlier scaling relation $R/a \propto 1/(\delta^2 + 8\omega)^{1/4}$ we obtain our central result

$$\omega(R) \propto \frac{1}{(R/a)^4}. \quad (20)$$

The numerical work of [15] shows that proportionality factor is about 94.2. For $D = 3$, we find that $\omega(R) \simeq \frac{A}{(R/a+C)^4}$ with $A = 50.6$ and $C = 2.41$.

3 Monomer-monomer interactions and Flory theory

Let us now return to our main subject—the energetics of the encapsidation of vRNA molecules, viewed here as annealed branched polymers. We have to take into account interaction between monomers that are not adjacent to each other as part of the structure of the polymer. We will do this using Flory theory. Though Flory theory has important limitations [17], it has proven to be a very effective tool to understand the effects of monomer-monomer interactions in polymeric systems. For example, it can be shown—by advanced methods [19]—that in three dimensions the radius of gyration $R_0(N)$ of an annealed branched polymer with repulsive, excluded-volume type interaction depends on the number of monomers as $N^{1/2}$ while the Flory theory of annealed branched polymers finds that $R_0(N)/a \propto N^{7/13}$

[20, 21]. Extensive numerical studies of annealed branched polymers reviewed in [22, 23] have further confirmed the effectiveness of Flory theory.

We will apply Flory theory to the encapsidation of annealed branched polymers, as we did earlier for the case of quenched branched polymers [12]. In Flory theory, a variational free energy $F(R)$ is constructed in terms of the radius R of the polymer in the form of the sum of the conformational entropic free energy term $F_{\text{conf}}(R)$ computed in the absence of excluded volume interaction, but fully accounting for the connectivity of the polymer, plus an interaction free energy $F_{\text{int}}(R)$ of N *disconnected* monomers confined inside a sphere of radius R .

We start with the conformational entropic free energy $F_{\text{conf}}(R)$. In the previous section we showed that $F_{\text{conf}}(R)$ has the form $\frac{A}{(R/a+C)^4}$ in the limit that R is small compared to the radius of gyration $R_0(N)/a \simeq N^{1/4}$ of an ideal annealed branched polymer. In the opposite limit, when R is large compared to $R_0(N)$, $F_{\text{conf}}(R)$ has been shown to scale as $N^{-1/3} R^{4/3}$ [20, 21]. We can (crudely) combine these two limiting cases into a single scaling form $F_{\text{conf}}(R) = G(R/R_0(N))$. The scaling function $G(x)$ is proportional to $x^{4/3}$ for large x and proportional to $1/x^4$ for small x , so it has a minimum around $x = 1$.

For the monomer-monomer interaction free energy $F_{\text{int}}(R)$, we adopt the first two terms $vN^2/R^3 + wN^3/R^6$ of a virial expansion. The first virial term represents two-body interactions while the second term represents three-body interactions. The coefficient w will be assumed positive while the coefficient v changes sign at the transition point between good to bad solvent quality ("theta point").

The Flory variational energy is now

$$F(R) = G\left(\frac{R}{R_0(N)}\right) + \frac{vN^2}{R^3} + \frac{wN^3}{R^6}. \quad (21)$$

The work required to confine the annealed branched polymer inside a spherical capsid of radius R_c is then $\Delta W = F(R_c) - F(R^*)$ where $R^*(N)$ is the equilibrium radius that minimizes $F(R)$. We start with the good solvent case, assuming v/a^3 to be of the order of one. The three-body term can be neglected in this regime, as will be checked shortly. Assume that $R \gg R_0(N)$, in which case we should take the large x limit of the scaling function $G(x)$. Minimizing $F(R)$ with respect to R gives for the equilibrium radius

$$\frac{R^*(N)}{a} \simeq \left(\frac{v}{a^3}\right)^{3/13} N^{7/13}. \quad (22)$$

The condition that $R^*(N) \gg R_0(N)$ is obeyed if $(v/a^3) \gg N^{-5/4}$ and the condition that the three-body term can be neglected is obeyed if $va^3/w \gg N^{-1/2}$. The assumption of the large x limit for the scaling function $G(x)$ is thus consistent for $R \simeq R^*(N)$ and N large compared to one.

If the encapsidation involves only mild compression, so if R_c is close to R^* , then expansion of $F(R)$ to second order in $(R^* - R_c)$ around the minimum gives

$$\Delta W \simeq \frac{vN^2}{R^{*3}} \left(\frac{R^* - R_c}{R^*}\right)^2 \quad (23)$$

(in units of $k_B T$ as before).

In the opposite limit of poor solvent condition, the third virial term must be retained. We now assume that v/a^3 is of the order of minus one while w/a^6 of the order of plus one. The interaction energy F_{int} is then minimized by $R_{\text{int}}^* \propto (w/v)^{1/3} N^{1/3}$ and scales as $(v^2/w)N$. The polymer is condensed into a globule with a density of the order of $-v/w$. Since $R_{\text{int}}^*(N)$ is large compared to $R_0(N)$ in the large N limit, it follows that in the regime

of $R \sim R_{\text{int}}^*$, we are—once again—in the large x limit of the scaling function $G(x)$. In that regime, F_{conf} scales as $N^{-1/3} R^{4/3}$. For $R \sim R_{\text{int}}^*$, $F_{\text{conf}}(R)$ scales $N^{1/9}$ so in the large N limit it can be neglected with respect to F_{int} . The configurational entropy plays only a secondary role for the condensed globule state. For both the good and bad solvent case, the dimensionless compressibility is $\frac{|v|N^2}{R^{*3}}$. This scales with N as $N^{5/13}$ for the good solvent case and as N for the bad solvent case.

We are forced to conclude that the regime of small x of the scaling function, which is the regime for which the results of the previous sections apply, is not directly relevant to the encapsidation of annealed branched polymers in either good or bad solvent conditions.

We can use Flory theory to estimate the work of compressing annealed vRNA molecules into spherical capsids. Using a simple coarse-graining method, we showed in [12] that the estimated number N of monomers for the vRNA molecule of a small $T = 3$ virus with an inner capsid radius of 10 nm is in the range of 300 while the short distance cut-off a is in the range of 1.5 nm. It follows that in the good solvent case with v/a^3 of the order one the limit $(v/a)^3 \gg N^{-5/4}$ is applicable. The equilibrium radius R^*/a of the vRNA molecule is then of the order of $20(v/a^3)$ (in the range of 30 nm) while the compressional modulus is of the order of $9(v/a^3)$. If, for example, the capsid radius is half the equilibrium radius, so if $R_c/R^* = 1/2$, then the work of compression would be in the range of $10k_B T$ for $(v/a^3) = 1$.

4 Conclusion

In summary, we have shown that—within the Lifshitz theory of annealed branched polymers—the $1/R^2$ confinement free energy of linear polymers must be replaced by a $1/R^4$ dependence. This result is consistent with earlier scaling arguments (see ref. [15]). Mathematically, the replacement is due to a “renormalization” of the external potential well produced by the side chains of the polymer. This generates a harmonic potential at the center of the well for the main chain. As a result, the effective Lifshitz equation for the main chain has the form of the Schrödinger equation of a quantum harmonic oscillator, at least for the low-lying levels of the equivalent quantum problem. Higher levels are expected to be dominated by the “bare” potential and are expected to depend on R as $1/R^2$.

An interesting generalization would be to include linear polymer insertions between branch points (by combining recursive procedures of Figs. 1 and 3). This allows a systematic examination of the cross-over between $N(a/R)^4$ confinement entropy for densely branched polymer and $N(a/R)^2$ for non-branched (linear) polymers.

Although mathematically elegant, our analysis is not directly relevant to the energetics of viral RNA encapsidation. The reason is that in both the good solvent and bad solvent regimes the enthalpic free energy contribution to the work of encapsidation overwhelms the entropic conformational free energy contribution. This observation should *not* overshadow the fact that confinement may still lead to changes in the branching structure of annealed branched polymers. Indeed, vRNA molecules in solution are known to form rather elongated “branched backbone” structures while under confinement they resemble dendrimers with the endpoints extending outwards [14]. Flory theory is, however, not sufficient to address this issue. The real challenge, which is beyond the scope of this article, is to examine the conformational and branching entropy of a molecule that is both compact in terms of the enthalpic free energy of interaction between the monomers, yet stretched in the sense that the radius of gyration is large compared to the radius of gyration $R_0(N)$ in the absence of monomer-monomer interactions. Some work in this direction can be found in [24].

Acknowledgments RB would like to acknowledge support from the National Science Foundation under DMR Grant 1309423. The work of AYG was supported partially by the MRSEC Program of the National Science Foundation under Award Number DMR-1420073. RB and AYG thank the Aspen Center for Physics where part of this work was done with the support of the National Science Foundation under Grant No. PHY-1066293.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Yoffe, A.M., Prinsen, P., Gopal, A., Knobler, C.M., Gelbart, W.M., Ben-Shaul, A.: Predicting the sizes of large RNA molecules. *Proc. Natl. Acad. Sci.* **105**, 16153 (2008)
2. Fang, L., Gelbart, W., Ben-Shaul, A.: The size of RNA as an ideal branched polymer. *J. Chem. Phys.* **135**, 155105 (2011)
3. de Gennes, P.-G.: Statistics of branching and hairpin helices for the dat copolymer. *Biopolymers* **6**, 715–729 (1968)
4. Perlmutter, J.D., Qiao, C., Hagan, M.F.: Viral genome structures are optimal for capsid assembly. *eLife* **2**, e00632 (2013)
5. Singaram, S.W., Garmann, R.F., Knobler, C.M., Gelbart, W.M., Ben-Shaul, A.: Role of RNA branchedness in the competition for viral capsid proteins. *J. Phys. Chem. B* **119**, 13991–14002 (2015)
6. Erdemci-Tandogan, G., Wagner, J., van der Schoot, P., Podgornik, R., Zandi, R.: RNA topology remodels electrostatic stabilization of viruses. *Phys. Rev. E* **89**, 032707 (2014)
7. Erdemci-Tandogan, G., Wagner, J., van der Schoot, P., Podgornik, R., Zandi, R.: Effects of RNA branching on the electrostatic stabilization of viruses. *Phys. Rev. E* **94**, 022408 (2016)
8. Li, S., Erdemci-Tandogan, G., Wagner, J., van der Schoot, P., Zandi, R.: Impact of a nonuniform charge distribution on virus assembly. *Phys. Rev. E* **96**, 022401 (2017)
9. Li, S., Orland, H., Zandi, R.: Self consistent field theory of virus assembly. *ArXiv e-prints* (2017)
10. Bruinsma, R.F., Comas-Garcia, M., Garmann, R.F., Grosberg, A.Y.: Equilibrium selfassembly of small RNA viruses. *Phys. Rev. E* **93**, 032405 (2016)
11. Hofacker, I., Fontana, W., Stadler, P., Bonhoeffer, L., Tacker, M., Schuster, P.: Fast folding and comparison of RNA secondary structures. *Monatsh. Chem. (Chemical Monthly)* **125**, 167–188 (1994)
12. Kelly, J., Grosberg, A.Y., Bruinsma, R.: Sequence dependence of viral RNA encapsidation. *J. Phys. Chem. B* **120**, 6038–6050, (2016)
13. Gopal, A., Zhou, Z., Knobler, C.M., Gelbart, W.M.: Visualizing large RNA molecules in solution. *RNA* **18**, 284–299 (2012)
14. Garmann, R., Comas-Garcia, M., Gopal, A., Knobler, C.M., Gelbart, W.M.: The assembly pathway of an icosahedral single-stranded RNA virus depends on the strength of inter-subunit attractions. *J. Mol. Biol.* **426**, 1050–1060 (2013)
15. Grosberg, A.Y., Kelly, J., Bruinsma, R.: The confinement of an annealed branched polymer by a potential well. *Low Temp. Phys.* **43**, 101–109 (2017)
16. Grosberg, A.Y., Gutin, A., Shakhnovich, E.: Conformational entropy of a branched polymer. *Macromolecules* **28**, 3718–3727 (1995)
17. Grosberg, A.Y., Khokhlov, A.R.: *Statistical Physics of Macromolecules*. American Institute of Physics Press, New York (1994)
18. Lifshitz, I.M.: Some problems of the statistical theory of biopolymers. *Zh. Eksp. Teor. Fiz.* **55**, 2408–2422 (1968)
19. Parisi, G., Sourlas, N.: Critical behavior of branched polymers and the lee-yang edge singularity. *Phys. Rev. Lett.* **46**, 871–874 (1981)
20. Gutin, A.M., Grosberg, A.Y., Shakhnovich, E.I.: Polymers with annealed and quenched branchings belong to different universality classes. *Macromolecules* **26**, 1293–1295 (1993)
21. Grosberg, A., Nechaev, S.: From statistics of regular tree-like graphs to distribution function and gyration radius of branched polymers. *J. Phys. A: Math. Theor.* **48**, 345003 (2015)
22. Rosa, A., Everaers, R.: Computer simulations of randomly branching polymers: annealed versus quenched branching structures. *J. Phys. A Math. Theor.* **49**, 345001 (2016)
23. Everaers, R., Grosberg, A.Y., Rubinstein, M., Rosa, A.: Flory theory of randomly branched polymers. *Soft Matter* **13**, 1223–1234 (2017)
24. Kelly, J.: *Physics of Complex Polymeric Molecules*, Ph.D. thesis, UCLA. <https://escholarship.org/uc/item/9hk3m75p> (2017)