Thermodynamics of nanospheres encapsulated in virus capsids

Antonio Šiber*
Institute of Physics, Bijenička cesta 46, 10000 Zagreb, Croatia

Roya Zandi†
Department of Physics and Astronomy, University of California, Riverside, California 92521, USA

Rudolf Podgornik‡
Department of Theoretical Physics, Jožef Stefan Institute, SI-1000 Ljubljana, Slovenia;
Institute of Biophysics, School of Medicine, and Department of Physics, University of Ljubljana, SI-1000 Ljubljana, Slovenia;
and Laboratory of Physical and Structural Biology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, Bethesda, Maryland 20892-0924, USA

(Received 16 November 2009; published 20 May 2010)

We investigate the thermodynamics of complexation of functionalized charged nanospheres with viral proteins. The physics of this problem is governed not only by electrostatic interaction between the proteins and the nanosphere cores (screened by salt ions), but also by configurational degrees of freedom of the charged protein *N* tails. We approach the problem by constructing an appropriate complexation free-energy functional. On the basis of both numerical and analytical studies of this functional we construct the phase diagram for the assembly which contains the information on the assembled structures that appear in the thermodynamical equilibrium, depending on the size and surface charge density of the nanosphere cores. We show that both the nanosphere core charge and its radius determine the size of the capsid that forms around the core.

DOI: 10.1103/PhysRevE.81.051919 PACS number(s): 87.15.nr, 41.20.Cv, 82.35.Rs

I. INTRODUCTION

Viruses have optimized the feat of packaging genome molecules and delivering them to the appropriate cells. In its simplest form, a virus consists of a rigid protein shell (capsid) that surrounds and protects the genetic material (either RNA or DNA) from chemical and physical assaults [1]. A viral capsid contains several copies of either one type of protein or of a few slightly different kinds. A number of *in vitro* self-assembly experiments reveal that the protein subunits of many RNA viruses can assemble spontaneously not only around their own genome but also the genomes derived from other viruses and various nonviral anionic polymers. All these features in addition to their extraordinarily highly symmetric shape and monodisperse size distributions make viruses ideal structures for gene therapy, drug delivery, and various nanotechnology and materials science applications [2–5]. To this end, the number of experiments and theoretical research aimed at understanding the physical basis of assembly of viruses and the factors influencing the structure and size of viral capsids is amazingly soaring [6–21].

The majority of viral capsids have either spherical or elongated structures. Most spherical viruses have structures with icosahedral symmetry and contain $60\mathcal{T}$ protein subunits, where $\mathcal{T}$ is the structural index of viral shells and is determined from the relation $\mathcal{T}=h^2+k^2+kh$ with h and k as nonnegative integers [22]. While the capsid protein of some viruses can form only one structure with a specific size, the capsid protein of many others is more flexible and adopts various structures with different sizes. For example, tobacco mosaic virus (TMV) capsid proteins assemble into tubular structures regardless of the shape and size of their cargo [23]. This indicates that the shape of the virus is solely dictated by the intrinsic property of protein subunits. In the other end of the spectrum, many experiments show that the capsid proteins of cowpea cholorotic mosaic virus (CCMV), a spherical RNA virus, are able to form capsid of various size and shapes [24–26].

Over 35 years ago, Adolph and Butler [25] and more recently Lavelle *et al.* [26] performed a series of *in vitro* experiments with CCMV capsid proteins in the absence of RNA and found that depending on the *p*H and ionic strength several different structures assembled. A notable feature of the constructed shape "phase" diagram based on these experiments is the change from an icosahedral $\mathcal{T}=3$ structure to a cylindrical shape upon a decrease in ionic strength and an increase in *p*H revealing the important role of electrostatic interaction on the size and shape of empty viral shells.

There have been a number of different experiments and theoretical studies [27–34] to investigate the impact on the structure of capsids of the shape and length of genome. To explore the effect of cargo on the morphology of viral capsid, Mukherjee and co-workers examined the assembly of capsid proteins of CCMV around heterogeneous DNA longer than 500 base pairs and found that tubular structures spontaneously formed [9]. Note that CCMV has at $\mathcal{T}=3$ icosahedral structure in its native form.

Quite remarkably, almost half a century ago, Bancroft (see, e.g., Ref. [24]) demonstrated the important role of stoichiometry ratio of capsid proteins and genome in the structure of viral shells. According to their experiments CCMV capsid proteins could encapsidate TMV RNAs which are

*asiber@ifs.hr
†roya.zandi@ucr.edu
‡rudolf.podgornik@fmf.uni-lj.si

 ©2010 The American Physical Society

about 6000 nucleotides in viral particles of various sizes. Depending on the ratio of RNA-protein concentrations, $\mathcal{T}=3$, 4, or 7 structures can form.

More recent experiments with polystyrene sulfonate (PSS, a highly flexible polyelectrolyte chain) and CCMV capsid proteins indicate the significance of length of genome in that the diameter of CCMV viral shells encapsidating PSS varies from 22 to 27 nm when the molecular weight of PSS varies from 400 KDa to 3.4 MDa [11]. The impact of the size of cargo on the diameter of capsid is quite transparent in the experiments of Sun and co-workers in which they found that the capsid proteins of brome mosaic virus (BMV), another spherical plant virus, are able to package the functionalized gold nanospherical particles with diameters of 6, 9, and 12 nm to form viruslike particles with $\mathcal{T}=1$, 2, or 3 structures, respectively [27].

All the aforementioned experiments focus on the important role of size and structure of genome and stoichiometric ratio of genome to protein in determining the structure of viral shells. However, several experimental and theoretical studies reveal that electrostatic interaction is the driving force for the assembly of capsid proteins around anionic cargos, and thus it is crucial to study the impact of cargo charge density on the structure of capsids. In fact, a careful study of several single stranded RNA viruses show that there is a linear relation between the numbers of charges on the capsid inner surface and on their genome [32]. An important question, then, naturally arises: could we change the size of a capsid by changing the net charge of its cargo? More specifically, in the experiments of Sun and co-workers with nanospheres, does a $\mathcal{T}=3$ structure form if one increases the charge density of the 9 nm cores which normally form "pseudo" $\mathcal{T}=2$ structures?

In this paper, we investigate the interplay between the charge density and size of nanocargos in virus assembly. Similar to the experiments of Sun and co-workers, we consider negatively charged nanospheres which interact with positively charged capsid inner surface, under physiological condition, and find that in addition to the diameter of the encapsidated nanospheres, the total net charges on cargos have a significant impact on the size of viral capsids.

An important feature of several RNA viruses, including CCMV and BMV mentioned above, is the presence of cationic polypeptide chains that form the N termini of the capsid protein. Rich in basic amino acids, there is a total of thousands of charges on the N-terminal tails which extend into the capsid interior and are responsible for the absorption of RNA to the inner capsid surface. Very recent *in vitro* studies of Aniagyei *et al.* reveal that a mutant of CCMV coat proteins lacking most of the N-terminal domain, $N\Delta 34$, assembles around negatively charged 12 nm spherical cores to form $\mathcal{T}=2$ structures [34]. Note that native CCMV proteins form a $\mathcal{T}=3$ structure around 12 nm spherical cores. Our calculations also show that the N-terminal arms can have a major impact on the virion structure and (as shown in Figs. 3 and 4) they can significantly modify the free-energy landscape of viral structures. One also has to consider that the deletion of the N-terminal tails might change the preferred angle between the protein subunits, an effect which is not taken into account in the present study. According to our studies, depending on the cargo charge density and the presence of N-terminal tails it might be advantageous for capsid proteins to form relatively smaller or bigger shells compared to their native structures. The effect of N-terminal on the free energy of viral capsids has been investigated previously [35]. Here, we take another approach that enables us to study the energetics of complexation of proteins and core in more details. Our emphasis is also on different aspects of the assembly, in particular the formation of differently sized structures depending on the conditions.

The outline of the paper is as follows. In Sec. II, we present our model to calculate the free energy between capsid inner surface and a rigid sphere including the interaction of positively charged N-terminal tails with the spherical cargo. In Sec. III, we present our numerical results and in Sec. IV we discuss our findings and their implications, and summarize our conclusion.

II. THEORETICAL DESCRIPTION OF ENERGETICS AND THERMODYNAMICS OF THE ASSEMBLY

Here, we consider encapsidation of charged nanoparticles (whose number is n_c) within the virus capsid and the dependence of the formation free energy on the parameters describing the system. We assume that the solution consists of monovalent salt (of bulk concentration c_0), dissolved protein monomers (or, more generally, basic protein subunits, which may be, e.g., protein dimers as is the case for hepatitis B virus—their number is assumed to be n_p), and spherical cores that are perfectly monodisperse with respect to radius R_1 and surface charge density σ_1. All the particles are dissolved in a medium whose dielectric constant is $\epsilon_0\epsilon_r$ (we shall take $\epsilon_r=80$, i.e., water). An assembly problem of this type involves many parameters, including the concentrations of cores (n_c/V) and proteins (n_p/V, where V is the volume of the solution). These two parameters importantly influence the assembly phase diagram, in addition to the energy of the assembled complex, which is a fairly complicated quantity in itself. One can expect formation of variously sized protein capsids (described by different Caspar-Klug $\mathcal{T}$ numbers) around the core, depending both on the properties of the core and on particle concentrations (n_c/V and n_p/V). We shall first, however, examine the energetics of the formed complex (capsid+core).

A. Energetics of the assembled complex

In general, we shall assume that the protein charge is distributed along the flexible N tails. We treat the tails as generic polyelectrolytes with intermonomer bond length a, partial charge p per monomer, and nonelectrostatic excluded volume interactions characterized by the excluded volume v. Our approach is quite similar to that exposed in Ref. [33]. The free energy of the complex in the subspace of fixed total number of polyelectrolyte monomers, N, can be calculated from

$$F=\int f(r)d^3r-\mu\left(\int d^3r|\Psi(\mathbf{r})|^2-N\right), \tag{1}$$

where μ is the Lagrange multiplier enforcing the condition of fixed number of monomers, and

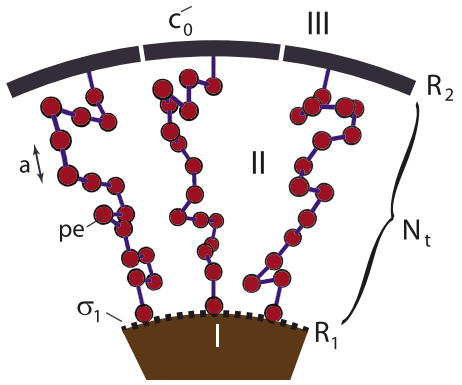

FIG. 1. (Color online) Illustration of the portion of the spherical complex of proteins with flexible tails and a charged core. The parameters discussed in the text are indicated. The configuration of the tails corresponds to the case when the core is poorly charged. For the sake of a clear insight, the illustration shows a discretized model of the polyelectrolyte, although our theory is based on the continuous polyelectrolyte representation (see text).

$$f(r) = k_B T\left[\frac{a^2}{6}|\nabla\Psi(r)|^2 + \frac{v}{2}\Psi(r)^4\right] + ec^+(r)\Phi(r) - ec^-(r)\Phi(r) + pe|\Psi(r)|^2\Phi(r) - \rho_p(r)\Phi(r) - \frac{\epsilon_0\epsilon}{2}|\nabla\Phi(r)|^2 + \sum_{i=\pm}\{k_B T[c^i(r)\ln c^i(r) - c^i(r) - (c_0^i \ln c_0^i - c_0^i)] - \mu^i[c^i(r) - c_0^i]\}. \quad (2)$$

The free energy is a functional of the monomer density field of the polyelectrolyte chains $[\Psi^2(r)]$ and the mean electrostatic potential $[\Phi(r)]$. The complex free energy also depends on the salt concentration fields $[c^{\pm}(r)]$ whose chemical potentials are denoted by $\mu^{\pm}$. One can assume that in addition to the charge located on mobile protein tails, there is also a static density of charge, $\rho_p(r)$, which could in principle be located on the immobile parts of the capsid proteins (outside tails). In the further calculations, we shall neglect the fixed charge on the capsid and assume that the static charge in the system resides exclusively on the core surface, so that

$$\rho_p(r) = \sigma_1\delta(r - R_1), \quad (3)$$

where σ_1 is the charge density at the core surface. The contribution of such localized charges to the electrostatic part of the free energy can be easily separated (if required) from the functional in the form of the boundary term as it was done in Ref. [19]. See Fig. 1 for the illustration of the system that we consider and the relevant parameters characterizing it.

The variation of the free-energy functional with respect to fields Ψ, Φ, and $c^{\pm}$ yields two coupled partial differential equations. The first one is the generalized polyelectrolyte Poisson-Boltzmann (PB) equation,

$$\epsilon_0\epsilon\nabla^2\Phi(r) = 2ec_0\sinh[\beta e\Phi(r)] - ep\Psi(r)^2 + \rho_p(r). \quad (4)$$

The second equation is the Edwards equation,

$$\frac{a^2}{6}\nabla^2\Psi(r) = v\Psi(r)^3 + pe\beta\Phi(r)\Psi(r) - \mu\Psi(r). \quad (5)$$

The Lagrange multiplier (μ) enforces the conservation of the total number of polyelectrolyte monomers,

$$\int d^3r\Psi^2(\mathbf{r}) = N_p N_t = N, \quad (6)$$

where N_t is the number of monomers in a particular tail and N_p is the number of proteins (or, in general, protein subunits) in the complex. The electrostatic boundary conditions are the same as in the standard PB theory. In particular, there is a boundary condition at $r=R_1$, reflecting the finite charge density σ_1 at the surface of the core. There are two additional conditions that need to be specified for the density field $\Psi(r)$. We take the core to be impenetrable and thus presume that

$$\Psi(R_1) = 0. \quad (7)$$

Since we are dealing with the continuous variant of the model, this boundary condition means that *exactly* on the core surface, the density of the monomers is zero. Depending on the conditions, however, the density of the monomers quite close to the core surface may be appreciable. Exactly *how* close to the surface will depend on many parameters, with an important one being the distance between the capsid and the core, i.e., $(R_2-R_1)/a$ ratio.

On the inner part of the capsid one could assume that

$$\Psi(R_2) = \sqrt{\frac{N_p}{4\pi a R_2^2}}, \quad (8)$$

which means that there must be some density of the tails at the capsid wall since they are fixed or grafted. This is the only way that grafting enters our calculation. In our representation, this condition means that the polyelectrolyte anchors to the capsid perpendicularly (see Fig. 1). The boundary conditions should be interpreted carefully due to the continuum nature of the theory that we use, so we once again emphasize the meaning of the boundary conditions in the context of the representation that we have chosen. The boundary conditions specify the values of the polyelectrolyte density field *right at* the core and inner capsid surfaces. However, the density field $[\Psi^2(r)]$ in *immediate* vicinity of these points may be significantly larger, showing the accumulation of monomers close to the surfaces. So, although condition (8) may be interpreted in the discrete variant of the model as if there is a single grafted monomer in a spherical shell of thickness a touching the capsid from the inside, in the continuum variant of the model and depending on the conditions, there may in fact be *many* monomers in the spherical shell of thickness a, although condition (8) still holds. That is why, in the continuum variant of the model, we just say that the condition (8) means that the polyelectrolyte anchors perpendicularly to the inner capsid surface. The problem as we have posed it is well defined for arbitrary separations between the core and the capsid. Note, however, that in order to fit the whole polyelectrolyte in that space, i.e., that the normalization condition in Eq. (6) is satisfied, $\Psi^2(r)$ may have large values of the derivatives due to the

boundary conditions in Eqs. (7) and (8). This also signifies a thermodynamically unfavorable contribution of the entropy of such polyelectrolyte configurations [33], which is given by $k_BTa^2|\nabla\Psi(r)|^2/6$. This term is in fact the only remnant of the discretized representation of the polyelectrolyte, containing explicitly the monomer size. Although there are no ambiguities and problems with our theory in the regime $R_2-R_1\sim a$, there obviously may be some problems in interpretation of the boundary conditions in the discretized representation of the polyelectrolyte. None of the conclusions that we shall present depend on the behavior of our model in this regime. The problems of discretized vs continuum representations of the polyelectrolytes have been previously extensively discussed in the literature, and we point the reader to Refs. [36,37] that discuss the aspects of the continuum polyelectrolyte representation relevant for the situation of interest to us.

The procedure presented thus far is based on the mean-field description of the electrostatics of the problem and ground-state dominance ansatz for the polyelectrolyte field (see Ref. [33] for details). Furthermore, the finite extensibility of the polyelectrolyte is not taken into account. While that was of no essential importance for the problem studied in Ref. [33], it may become of importance in our case since the protein tails are assumed to be grafted to the capsid. To take the geometrical constraints regarding the polyelectrolyte density field into account, we shall now derive an alternative set of equations that we will obtain from a constrained variation of a free-energy functional. Instead of varying $F_{complex}$ over the space of all functions $\Psi(\mathbf{r})$, we shall vary it over the *constrained* set of functions, i.e., polyelectrolyte amplitudes that can be represented as

$$\Psi(r)=\Psi_s(r)+u^2(r), \tag{9}$$

where $u(r)$ is a real function and

$$\Psi_s^2(r)=\frac{N_p}{4\pi ar^2}. \tag{10}$$

The above equation describes the maximally extended polyelectrolyte density, so that no smaller values of $\Psi(r)$ are possible without enlarging the monomer-monomer separation distances a (stretching). Varying $F_{complex}$ over $u(\mathbf{r})$, we obtain the new set of Euler-Lagrange differential equations for $\Phi(r)$ and $u(r)$ as

$$\epsilon_0\epsilon\nabla^2\Phi(r)=2ec_0\sinh[\beta e\Phi(r)]-ep[\Psi_s(r)+u^2(r)]^2+\rho_p(r), \tag{11}$$

$$\frac{a^2}{6}\mathcal{L}_{\Psi_s}[u(r)]=s(r), \tag{12}$$

where $\mathcal{L}_{\Psi_s}(u)$ is a differential operator given as

$$\mathcal{L}_{\Psi_s}(u)=\nabla u\,\nabla\Psi_s+2[u^2\nabla^2u+2u(\nabla u)^2]. \tag{13}$$

We have used here $\nabla^2\Psi_s=0$. The function $s(r)$ is given as

$$s(r)=v(\Psi_s^3u+3\Psi_s^2u^3+3\Psi_su^5+u^7)+\beta e\Phi(\Psi_su+u^3)-2u\mu(\Psi_s+u^2). \tag{14}$$

We still have to specify the boundary conditions. Looking at Eq. (9), one sees that we can no longer put $\Psi(R_1)=0$, so we will be necessarily stuck with finite density of monomers at the core radius. We choose

$$u(R_1)=\delta,\quad u(R_2)=\delta,\quad \delta\to 0 \tag{15}$$

as the appropriate boundary conditions implying that the N tails are *normal* to the surface of the inner core as well as outer capsid wall at R_1 and R_2. In other words we assume that the chains have no overhangs at R_1 and R_2, so that they touch both the core and the capsid only once and perpendicularly. Note that Eq. (8) is automatically satisfied with such a choice.

The complex free energy that we have constructed thus far accounts approximately for electrostatic energy of the system, the entropic and excluded volume effects of the confined polyelectrolyte (protein tails), and for entropic contributions of the salt ions (on the mean-field level). It does not contain, however, the attractive component (nonelectrostatic) of protein-protein interactions. These interactions consist of hydrophobic and van der Waals (vdW) contributions [38] and we denote their value per protein in the complex as $\bar{f}_{p,hydro}$. It is this part of the energy that keeps the dominantly positively charged *empty* capsids together [19,39].

B. Thermodynamics of the assembly

The parameter space for assembly that we are interested in needs to be at least four dimensional (R_1, σ_1, n_c/V, and n_p/V), presuming that the properties of the capsid proteins are kept fixed. Analysis of assembly in such a high-dimensional parameter space would be highly involved [29,40]. It is of interest thus to try to extract the relevant information on the assembly *solely* from the complex free energy, as we have defined it. Such a procedure cannot be expected to be valid for all values of n_c and n_p, but it should be of use when

$$n_c\ll\frac{n_p}{N_p(\mathcal{T}_{max})}, \tag{16}$$

where $\mathcal{T}_{max}$ is the maximal capsid $\mathcal{T}$ number that can be possibly assembled in the experimental circumstances and $N_p(\mathcal{T}_{max})$ is the number of proteins (subunits) in a Caspar-Klug structure of that $\mathcal{T}$ number. Basically, Eq. (16) says that, concerning the "final" (assembled) state, there is no big difference with respect to the $\mathcal{T}$ number of the dominantly assembled structures. The final state will always consist of all cores complexed with certain number of proteins, and the rest will be more or less the same number of proteins that may remain isolated in the solution or perhaps form empty capsids. There will be no free (uncomplexed) cores in the assembled state.

Let us assume that prior to the assembly, the solution contains isolated charged cores and individual viral proteins (or protein dimers or whatever the basic subunit of the as-

sembly might be)—this is the "initial" state. In the initial state, the free energy of the system is

$$F_i = n_p \bar{f}_p^i + n_c F_{core}^i, \tag{17}$$

where $\bar{f}_p^i$ and F_{core}^i are the free energies per isolated protein and core in the initial state, respectively. These may in principle contain also the translational entropy contributions. After the assembly, the free energy is (final state)

$$F_f = n_c F_{complex} + n_c N_p \bar{f}_{p,hydro} + (n_p - n_c N_p)\bar{f}_p^f. \tag{18}$$

The first term on the right-hand side (RHS) of Eq. (18) is what we can calculate—this is the free energy of the complex (note, however, that we do not calculate the entropic term regarding the translational freedom of the assembled structures). The number of complexes is exactly n_c since we assumed that *all* cores are complexed with the proteins. The second term on the RHS of Eq. (18) is the part of the attractive energy of proteins assembled in the complexes that is difficult to calculate. It contains the hydrophobic and van der Waals protein-protein interactions, and per protein the corresponding free energy is $\bar{f}_{p,hydro}$. The third term on the RHS of Eq. (18) is the free energy of the proteins in the final state, which are *not* assembled in the complexes with cores ($n_p - n_c N_p$ of them). They may, however, be assembled in empty capsids. Thus, in general, free energy per such protein in the final state ($\bar{f}_p^f$) is not the same as the free energy per *isolated individual* protein in the initial state ($\bar{f}_p^i$). If the proteins that do not complex with the cores indeed form aggregates (e.g., empty capsids) then one can expect that $\bar{f}_p^f < \bar{f}_p^i$.

The system will proceed from state i to state f if $F_f < F_i$. We want to examine the quantity F_f for complexes that have different $\mathcal{T}$ numbers, i.e., we want to find F_f for several different final states f. For all of these states, the initial assembly state is the same, so for two different final states f_1 and f_2 we can directly compare the corresponding free energies F_{f_1} and F_{f_2} and if $F_{f_1} < F_{f_2}$ we can say that state f_1 is more likely to be realized in the thermodynamical equilibrium. Thus, one should explore the quantity

$$F_f - F_i = n_c(F_{complex} - F_{core}^i) + n_c N_p(\bar{f}_{p,hydro} - \bar{f}_p^f) + n_p(\bar{f}_p^f - \bar{f}_p^i). \tag{19}$$

Note, however, that the last term proportional to n_p does not depend on the structure of the assembled complex, so for our purposes it is of no importance as we want to examine the assembly free energies for different N_p's.

Upon assuming that

$$|F_{complex}| \gg N_p|\bar{f}_{p,hydro} - \bar{f}_p^f|, \tag{20}$$

we can examine $F_{complex}$ for different complexes and compare them mutually (for given initial state, i.e., charge density and radius of the cores) so to judge about the thermodynamically preferred states. In order to be able to construct a phase diagram, i.e., to compare the free energies for *different* initial states, we shall construct the quantity

$$\Delta F \equiv F_{complex} - F_{core}^i. \tag{21}$$

This quantity should contain the biggest part of the assembly free energy, assuming Eq. (20) holds. Note, however, that the interpretation of ΔF as the assembly free energy is approximate and of limited validity. As long as the chemical potential of the free proteins is equal to the nonelectrostatic part of protein-protein interaction, our assumption is valid, but it can be most easily justified by assuming that the assembly of empty shells in solution can take place in the absence of the cores, which is the situation often observed in experiments. In that case, the attractive protein-protein interactions are likely to be similar in the empty capsids (present in the final state of the system) and those that contain the core, so that $\bar{f}_{p,hydro} - \bar{f}_p^f \approx 0$, and condition (20) is automatically satisfied. In the following, we shall term ΔF as the assembly free energy, and we shall refer to $F_{complex}$ (also denoted as F) as the complex free energy or the free energy of the complex.

III. NUMERICAL EVALUATION OF THE MODEL

A. Tail-less protein subunits

As already discussed, a prominent feature of the virus proteins assembled in a capsid is the N tails that protrude into the capsid interior. This feature is typical for many viruses and can influence both the energetics of the protein-genome assembly [33] and its speed. The tails are typically very positively charged, and they are thus expected to play a prominent role in the assembly of proteins with negatively charged cores. The possibility of spatial redistribution of the tails is also expected to influence the assembly, so that one can expect an interplay between the electrostatics of the tails and their configurational entropy. All of these effects are included in our free-energy functional, at least approximately. In this section, we shall emphasize the electrostatic aspect of the capsid proteins and effectively neglect the tail positional degrees of freedom. The problem then reduces to electrostatic interactions only since the entropy of the N tails is quenched. Omitting the hydrophobic, vdW, etc. energies of the capsid proteins, one is left with a purely electrostatic part of the total free energy.

As a first approximation one can smear the charge of the N_p proteins uniformly over the external sphere of radius R_2, so that its surface charge density is σ_2. The problem as described by this model system is still not completely trivial as the PB equation describing it is of course nonlinear [19]. However, some insight can be obtained by linearizing the PB equation and solving it in this approximation [Debye-Hückel (DH)]. The details are elaborated in the Appendix and the final result for the electrostatic free energy is Eq. (A9). One can further simplify it by using $\kappa R_1 \gg 1$, which is a condition typically met for the experiments done on viruses at physiological salt concentrations (~100 mM) [19,39]. In this case,

$$\lim_{\kappa R_1 \gg 1} F_{complex} = \frac{\pi}{\epsilon_0 \epsilon_r \kappa}\{2R_1^2\sigma_1^2 + 4R_1R_2\sigma_1\sigma_2 e^{\kappa(R_1-R_2)} + R_2^2\sigma_2^2 e^{2\kappa(R_1-R_2)} + R_2^2\sigma_2^2\}. \tag{22}$$

One should note here that the salt resides only in compart-

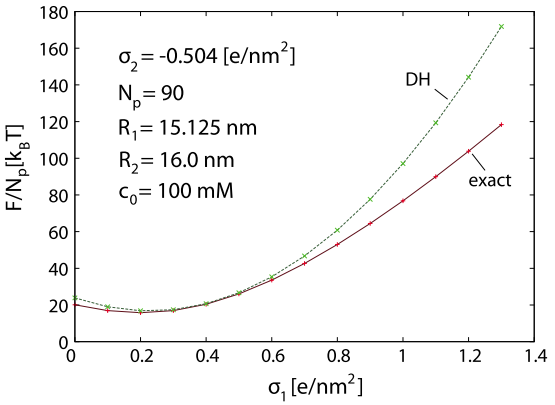

FIG. 2. (Color online) Comparison of the numerically exact results with the analytical DH expression (A9) for the free energy of the capsid with the charged core (complex free energy). The free energy per protein subunit is plotted as a function of the core surface charge density σ_1. The parameters of the calculation are denoted in the figure.

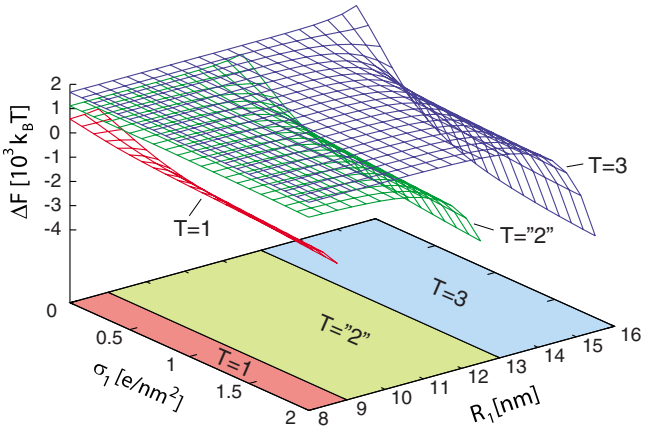

FIG. 3. (Color online) Numerically exact results for the assembly free energy [Eq. (21)] assuming tail-less protein subunits (c_0 =100 mM). The free energies were plotted for capsids with three different $\mathcal{T}$ numbers as functions of the core surface charge density σ_1 and radius R_1. The number of subunits and corresponding capsid radii are N_p=30 ($\mathcal{T}$=1), 60 ($\mathcal{T}$="2"), and 90 ($\mathcal{T}$=3), and R_2=9.24 ($\mathcal{T}$=1), 13.06 ($\mathcal{T}$="2"), and 16 ($\mathcal{T}$=3) nm.

ments III and II (see Fig. 1), so there is no symmetry in the formula regarding σ_1, σ_2 and R_1, R_2. The first term in Eq. (22) is the electrostatic self-energy of the core, and the third term is the self-energy of the protein shell. Note that the self-energy of the core has a prefactor of 2 with respect to the analogous term for the protein shell. This is due to the fact that the shell is screened by the salt both from the inside and the outside, which is not the case for the impenetrable core (note also that the dielectric constant of the core does not figure in the final equations).

The second term in Eq. (22) is the electrostatic interaction free energy between the core and the protein shell. One easily sees that it will be minimized when σ_1 and σ_2 have different signs. We see that it decreases quickly as the distance between the shell and the core increases, i.e., it decreases as $\exp[-\kappa(R_2-R_1)]$.

Although the DH solution enables us to study the energetics of the assembly in more details, at least numerically one needs to check its validity vs the complete nonlinear Poisson-Boltzmann theory (see Ref. [19]). By comparing it to the exact solution of the Poisson-Boltzmann equation one simultaneously checks the numerical results and the analytical formula for the DH solution. This comparison is shown in Fig. 2.

We have chosen the capsid parameters to approximately represent the $\mathcal{T}$=3 capsid of the BMV with capsid radius R_2=16 nm. The surface charge density of the capsid σ_2= $-0.504e/\mathrm{nm}^2$ was obtained by smearing the 18 charges per each of the N_p=90 dimeric tails over the capsid surface. Note that the absolute signs of the charge do not matter; what is important is that the charge on the core is of the opposite sign from the charge on the capsid. The core radius R_1 =15.125 nm was chosen to be in the regime when the attractive interaction is not completely screened by salt—one can see this effect as a minimum in the free energy for some core charge density σ_1. This minimum disappears when the distance between the core and the capsid is larger than $\sim 1/\kappa$. In the chosen case, the minimum is at $\sigma_1 \sim 0.3e/\mathrm{nm}^2$.

As in the case studied in Ref. [19], the DH results always produce larger free energies than the exact Poisson-Boltzmann calculation. They are of course better when the potentials in the solution are small enough, so that the linearization approximation holds—this is the case for not too large surface charge densities, and it can be clearly seen that the DH approximation fails worse as σ_1 (or σ_2) increases, again in complete agreement with the results of Ref. [19].

Having established some confidence in the DH results, we can scrutinize them a bit more closely. If one takes $R_1=R_2$ and $\sigma_1=-\sigma_2$ in Eq. (22), one obtains the *absolute minimum* of the complex free energy in the whole R_1, R_2, σ_1, and σ_2 space. At that point in the parameter space the free energy is exactly *zero*, which is the absolute minimum. This is due to the fact that the core charge exactly neutralizes the protein charge yielding effectively the uncharged shell ($\sigma=\sigma_1+\sigma_2$ =0) of radius $R=R_1=R_2$.

We now insert Eq. (22) in Eq. (21) and find

$$\Delta F = \frac{\pi}{\epsilon_0 \epsilon_r \kappa}\{4R_1R_2\sigma_1\sigma_2 e^{\kappa(R_1-R_2)} + R_2^2\sigma_2^2 e^{2\kappa(R_1-R_2)} + R_2^2\sigma_2^2\}. \tag{23}$$

Examining the assembly free energies [Eq. (21)] of the variously sized (R_1) and charged (σ_1) cores with the proteins assembled in capsids of three different $\mathcal{T}$ numbers, we obtain the exact results for the proteins with no tails shown in Fig. 3. These were obtained by solving full PB equation, without linearizing it. We see that the phase diagram in this case is totally "flat"—the smallest possible $\mathcal{T}$-number structure will always be the one with smallest complex energy, irrespectively of the core charge and its radius (at least in the range of values considered here, and the conditions summarized in Sec. II B). Thus, the phase diagram is governed solely by simple geometry. The same result comes out also in the DH approximation. Note also how the assembly free energy shows practically no dependence on σ_1 and R_1 when R_2 $-R_1 \gg 1/\kappa$. This is due to the fact that in this regime, the

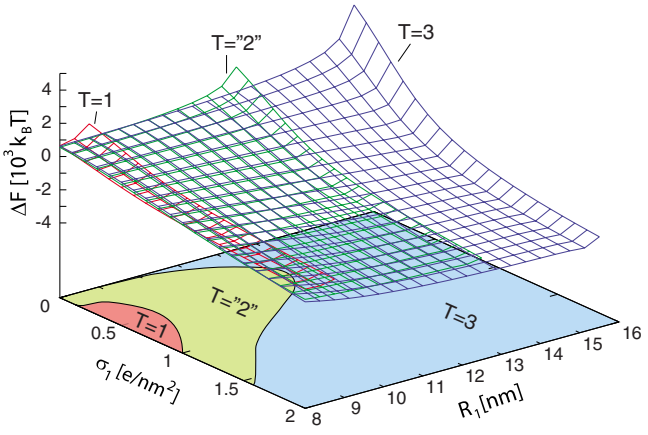

FIG. 4. (Color online) Assembly free energies calculated as functions of the core radius R_1, its charge σ_1, and for three different capsid radii R_2 (and N_p), as quantified by three different $\mathcal{T}$ numbers. The parameters of this calculation are $a=0.5$ nm, $p=1$, $v=0.5$ nm^3, $c_0=100$ mM, and $N_t=18$.

assembly free energy is simply the (positive) electrostatic self-energy of the capsid which depends only on R_2, i.e., the $\mathcal{T}$ number of the capsid [see also Eq. (23) for the Debye-Hückel description of this situation]. When $R_2-R_1\sim 1/\kappa$ (i.e., about 1 nm in our case, $c_0=100$ mM), the electrostatic attraction between the core and the capsid becomes only partially screened by salt ions, and the free energy of the assembly, ΔF, becomes negative suggesting that the assembly is thermodynamically preferable, releasing extra energy in the solution. The assembly is also more efficient for larger core charge densities.

We still need to consider what happens in the case when the protein charges are delocalized on the flexible polyelectrolyte tails. In particular, the decrease in the volume of the space between the core and the capsid when $R_1\rightarrow R_2$ would importantly confine the polyelectrolyte tails. We can thus expect their entropic and self-interaction contributions to become of the largest importance in the region where the assembly free energy of the tail-less monomers shows the most negative values.

B. Protein subunits with N tails

We now assume that the protein charge is distributed along the flexible N tails, i.e., we take all the details of the model developed in Sec. II into account. First, we calculate the free energies of the complex without the account of the finite extensibility of the tails, i.e., we solve Eqs. (4) and (5) with boundary conditions for the polyelectrolyte amplitude as specified in Eqs. (7) and (8). From the thus obtained free energy, we subtract the electrostatic self-energy of the core. The ensuing assembly free energies are shown in Fig. 4.

We see that although the assembly free energies are of the same order of magnitude as in the case of tail-less capsomeres, the shape of the free-energy landscape is quite different and the assembly is now governed both by the core charge and by its radius. This is most easily seen by the complicated shape of the regions denoted by $\mathcal{T}=1$, $\mathcal{T}=$"2", and $\mathcal{T}=3$ in the σ_1-R_1 plane in Fig. 4, which correspond to capsids with the lowest free energy. We also see that when the two radii are close to each other, the free energy steeply rises. This is easy to understand since in this case, the tails are forced to redistribute in a small volume in between the capsid and the core, so the contribution of entropic confinement to the free energy becomes significant. Interestingly, in the case of an infinitely thin capsid studied in the previous section, it was in these regions that the assembly free energy sharply *dropped*, exactly the opposite from the case we have here. It is thus clear at this point that the tails do introduce different physics into the problem. The assembly free energy in general also decreases with σ_1, which is an effect that we saw in the previous section and is due solely to electrostatics. We also observe that the values of the free energy are mostly smaller from the ones obtained in the model of an infinitely thin charged capsid, which essentially means that the tails can adopt such conformations that reduce the electrostatic part of the free energy to a significant extent, especially in the electrostatically unfavorable regime. Note also how the structure with the lowest assembly free energy increases its radius (and total charge) as the core charge increases (e.g., for $R=8.25$ nm one can see the progression from $\mathcal{T}=1$ and $\mathcal{T}=2$ to $\mathcal{T}=3$ structures as σ_1 increases). This can be simply explained by the screening of the core by the capsomer tails—the more charged the core, the more capsomeres (i.e., larger $\mathcal{T}$ numbers) are needed to screen it efficiently. But note here that the tails are assumed to be maximally flexible and can thus easily stretch from practically arbitrary distances (capsid) to the core in order to screen it. For a given core radius R_1, the assembly free energy is positive when $\sigma_1=0$, and it decreases as σ_1 increases, becoming negative for some "critical" core charge density, whose typical values in the range of parameters considered are $\sigma_1\sim e/\mathrm{nm}^2$. Thus, in order for the assembly to proceed spontaneously, the cores need to be sufficiently charged.

To see whether the results are influenced by the maximal extensibility of the tails, in Fig. 5 we plot the assembly free energy using the maximal extensibility ansatz in Eq. (10) together with boundary conditions in Eq. (15). Again, the assembly free energies are of the same order of magnitude, but the borders between the different regions in the σ_1-R_1 plane are quite different with respect to those obtained without the maximal extensibility ansatz. In order to better understand the results obtained thus far, it helps to plot the spatial distribution of the protein monomers, i.e., $\Psi^2(r)$, within the space between the core and the inner capsid radius. A solution for the case when $N_p=90$ (1620 monomers in total), $R_2=16$ nm ($\mathcal{T}=3$), $R_1=9.875$ nm, and $\sigma_1=0.5e/\mathrm{nm}^2$ is shown in Fig. 6.

What Fig. 6 nicely illustrates is that the tails stretch out from the capsid surface toward the core, so that they accumulate around the oppositely charged core. There are still some monomers in the space between the core and the capsid, but the dominant density is situated in a shell around the core. One can also see how the constrained solution stays above the maximal extensibility limit in Eq. (10). One should keep in mind that the boundary conditions for the two solutions are different, so one cannot expect that the constrained variation will always yield larger energy. This is also the case for the displayed calculation where we find $F=927k_BT$, and

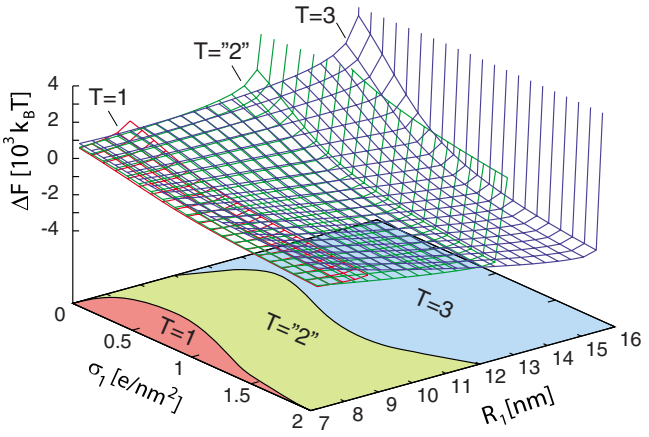

FIG. 5. (Color online) Assembly free energies as functions of the core radius R_1, its charge σ_1, and for three different capsid radii R_2 (and N_p), as quantified by three different $\mathcal{T}$ numbers. The maximal extensibility ansatz was used to obtain these results. The parameters of this calculation are $a=0.5$ nm, $p=1$, $v=0.5$ nm^3, $c_0=100$ mM, and $N_t=18$.

$F=788k_BT$ for the unconstrained and constrained values of the complex free energy, respectively. Note how the constrained polyelectrolyte density in fact approaches closer to the charged core than the unconstrained density due to the different boundary conditions that the two satisfy. The noted effect results in the lowering of the free energy in the constrained case. However, this is not always the case, and it depends on the distance between the core and the capsid (R_2-R_1). When this distance is sufficiently large, the tails cannot accumulate around the core even when they maximally stretch, so that they cannot screen the core efficiently. This effect is not present in the calculation with the unconstrained polyelectrolyte amplitude. This is in fact the most important reason for the different look of boundaries in the σ_1-R_1 plane for the two calculations. Note how in the unconstrained case the complex of core with $\mathcal{T}=3$ capsid has the lowest energy for sufficiently large σ_1 ($\sigma_1>1.5e/\mathrm{nm}^2$) in a huge range of radii R_1 (8–16 nm; see Fig. 4). However, the maximal length of the (fully stretched) tails is $N_ta=9$ nm, so that for $R_1=8$ nm tails are very much extended and only a small part of the polyelectrolyte density can gather around the core. As the unconstrained results do not account for this effect, by breaking the maximal extensibility limit in Eq. (10) the polyelectrolyte density screens the core efficiently and thus lowers the free energy of the $\mathcal{T}=3$ structure with respect to the value obtained in the constrained calculation. This effect is illustrated in Fig. 7 for $\mathcal{T}=3$ ($R_2=16$ nm, $N_p=90$), $R_1=8$ nm, and $\sigma_1=1.5e/\mathrm{nm}^2$. Note how the maximal extensibility limit is severely broken in the unconstrained calculation. The free energies of the complexes are $F=2710k_BT$, and $F=3520k_BT$ in the unconstrained and constrained calculations, respectively, demonstrating the effect discussed.

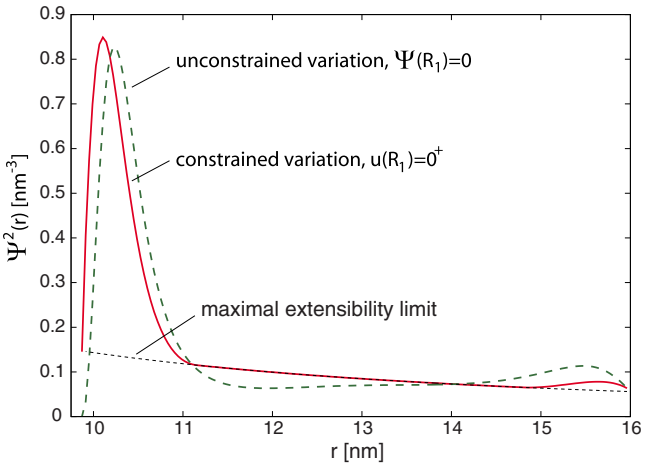

FIG. 6. (Color online) Comparison of the unconstrained monomer density (thick dashed line) and constrained monomer density (full line) approaches. The parameters of the calculation are $\sigma_1=0.5e/\mathrm{nm}^2$, $c_0=100$ mM, $R_1=9.875$ nm, $R_2=16$ nm, $N_p=90$, and $N_t=18$. The maximal extensibility limit [Eq. (10)] is denoted by a thin dashed line.

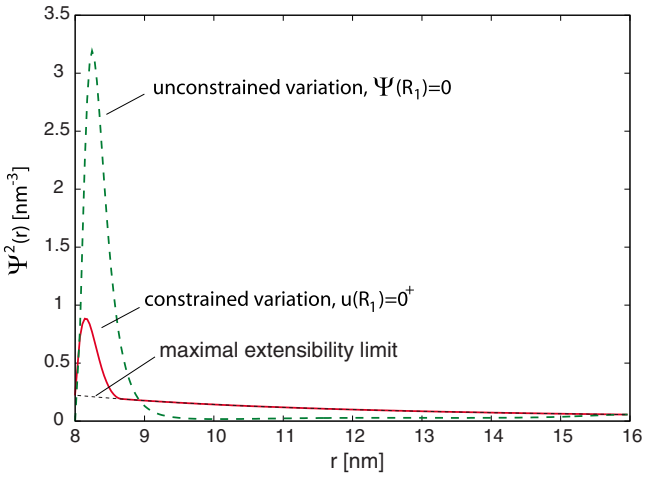

FIG. 7. (Color online) Comparison of the unconstrained monomer density (thick dashed line) and constrained monomer density (full line) approaches. The parameters of the calculation are $\sigma_1=1.5e/\mathrm{nm}^2$, $c_0=100$ mM, $R_1=8$ nm, $R_2=16$ nm, $N_p=90$, and $N_t=18$. The maximal extensibility limit [Eq. (10)] is denoted by a thin dashed line.

It is of interest to examine the efficiency of core screening by the polyelectrolyte tails somewhat closer. We define the ratio

$$\Theta=\frac{\int_{R_1}^{R_1+a}d^3r\Psi^2(r)}{N_tN_p}, \tag{24}$$

which can be interpreted as a percentage of monomers in a shell of thickness a (monomer-monomer separation) around the core. This can also be thought of as the percentage of monomers that cover (or are "in contact with") the core. In Fig. 8 we display the coverage Θ as a function of σ_1 and R_1 for $\mathcal{T}=3$ capsid ($R_2=16$ nm).

What can be easily seen from this plot is the gradual shift of the polyelectrolyte density from the space close to capsid to the shell surrounding the core as the core charge density (σ_1) increases. Due to the maximal extensibility constraint

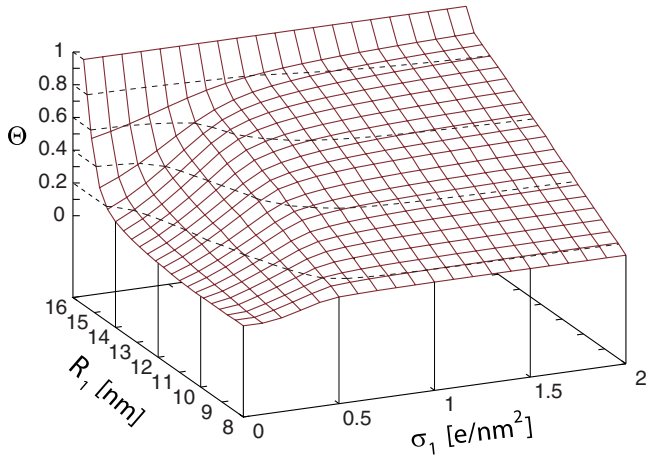

FIG. 8. (Color online) Coverage of the core defined by Eq. (24) as a function of σ_1 and R_1 for a complex with $\mathcal{T}=3$ capsids ($R_2=16$ nm). The calculation was performed with the maximal extensibility constraint on the polyelectrolyte amplitude.

that was included in these calculations, Θ parameter saturates at a value smaller than 1 as σ_1 increases. This simply reflects the fact that the polyelectrolyte must pass through the region between the core and the capsid in order to accumulate around the core, and so much of its density may remain in the space in between the capsid and the core. This effect becomes more important when (R_2-R_1) becomes comparable to the tail length $N_t a$. The charge density of the core at which one observes the significant accumulation of the monomers around the core ($\sigma_1 \sim 0.5e/\mathrm{nm}^2$) is the same as the charge density at which the assembly free energy becomes negative, i.e., the assembly proceeds spontaneously.

IV. CONCLUSIONS

An intriguing feature of our results is that both the core charge and its radius determine the size of the capsid around the core. A particularly interesting case is when the core radius is close (but somewhat smaller) to the $\mathcal{T}=1$ capsid, i.e., $R_1=8$ nm. For sufficiently small core surface charge density ($\sigma_1 < 0.5e/\mathrm{nm}^2$), $\mathcal{T}=1$ structures around the cores shall form. However, if the surface charge density increases over some critical value (around $1.0e/\mathrm{nm}^2$ in the constrained model of the polyelectrolyte), $\mathcal{T}=$"2" capsids shall form, in spite of the fact that the core radius is more than 5 nm smaller from the radius of $\mathcal{T}=$"2" capsid (see Fig. 5). This clearly shows that in addition to core radius, which dominantly influences the assembly process, one needs to have an adequate charge density in order to produce the structures of desired $\mathcal{T}$ number. The same effect is present on the $\mathcal{T}=$"2" and $\mathcal{T}=3$ borders when $R_1 > 11.6$ nm. We show this transition region in another way in Fig. 9, showing the difference in the complex free energy of $\mathcal{T}=2$ and $\mathcal{T}=3$ structures. Note that although the free-energy curves indeed cross for certain values of σ_1 and R_1, their magnitudes remain similar even deeply in the transition region. Thus, one could expect to observe a polydisperse distribution of $\mathcal{T}=$"2" and $\mathcal{T}=3$ structures in a solution with a monodisperse distribution of core size and charge density.

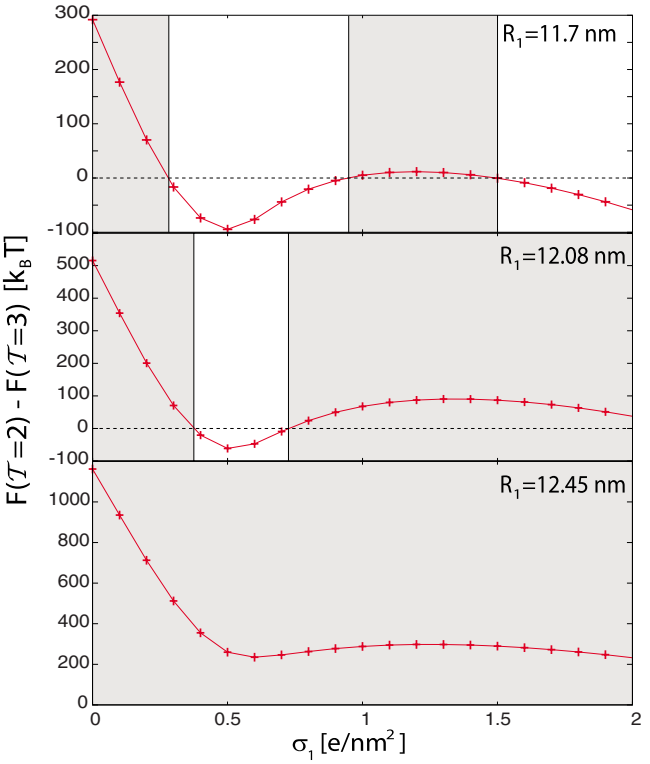

FIG. 9. (Color online) The difference of complex free energy of $\mathcal{T}=2$ and $\mathcal{T}=3$ structures for three different core radii in the transition region as a function of core surface charge density. The parameters of the polyelectrolyte are the same as before and the calculations were performed for the constrained polyelectrolyte model (maximal extensibility limit obeyed). Gray regions denote the surface charge densities σ_1 where the $\mathcal{T}=3$ structure has lower free energy.

An interesting effect is observed for $R_1=12.08$ nm in Fig. 9 where there exist *two* intersections between the free-energy curves for $\mathcal{T}=$"2" and $\mathcal{T}=3$ complexes as σ_1 increases. If the charge density is lower than $0.3e/\mathrm{nm}^2$, $\mathcal{T}=3$ structures have the lowest assembly free energy. Quite surprisingly, upon increasing the charge density, $\mathcal{T}=2$ structures become the dominant ones, i.e., the size of thermodynamically preferred capsids decreases. This is mainly due to the fact that with increasing the charge density, the monomers in the N-terminal tails prefer to sit next to the core, i.e., the electrostatic interaction wins over the chain configurational entropy. However, as we increase the core charge density beyond $0.7e/\mathrm{nm}^2$, it becomes more advantageous to have $\mathcal{T}=3$ structures again as there will be more charges associated with the N tails of $\mathcal{T}=3$ structures.

An even more intriguing situation is illustrated in Fig. 9 for $R_1=11.7$ nm cores. Here, we observe three transition lines instead of two of the previous case. Somewhat counterintuitively, for the charge densities above $1.5e/\mathrm{nm}^2$, $\mathcal{T}=$"2" particles become free-energy minima structures again. This indicates that under physiological conditions, if we increase the charges on the cores significantly, the lowest possible $\mathcal{T}$ structures ($\mathcal{T}=2$ in this case) form, so that more charges on the N tails can sit in the immediate vicinity of the core. This is due to the maximal extensibility constraint, and the effect

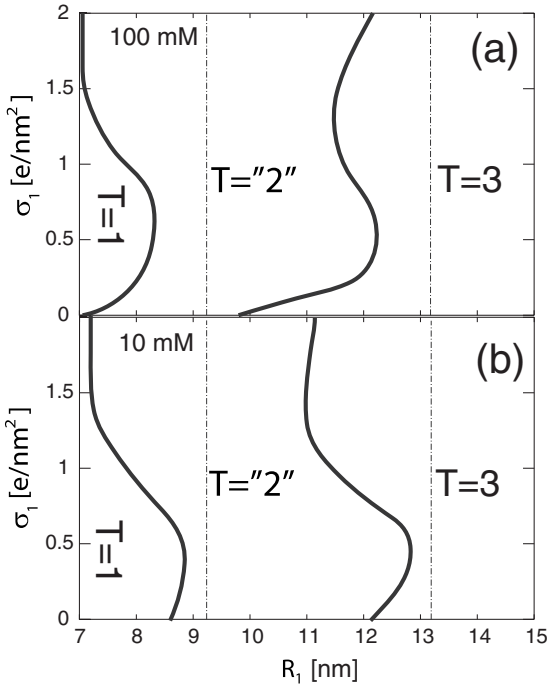

FIG. 10. Regions in the R_1-σ_1 space in which a particular $\mathcal{T}$-number structure has the lowest free energy, as denoted. (a) and (b) show the results for monovalent concentrations of $c_0 = 100$ mM and $c_0 = 10$ mM, respectively. The two vertical dashed-dotted lines denote the radii of $\mathcal{T}=1$ and "2" structures (R_2).

is not present in the calculation with the tails that do not satisfy the constraint (see Fig. 4). It is obvious from the previous discussion that the finite extensibility of the tails is important for determination of the lowest-energy structures. One can see this most easily by comparing Figs. 4 and 5. This effect becomes particularly important for structures in which the core is significantly smaller than the capsid, i.e., when $(R_2-R_1)/(N_t a) \gtrsim 1$.

We have repeated our calculations at lower salt concentrations. The results are presented in Fig. 10. As it is shown in the figure, the overall look of the region boundaries is not dramatically changed.

In summary, we have demonstrated that the thermodynamics of the assembly nontrivially depends on the electrostatic and geometric constraints which include $\kappa(R_1-R_2)$ (electrostatic screening), maximal possible stretching $(R_2-R_1)/(N_t a) \sim 1$, and confinement $(R_2-R_1)/a < 1$ of the protein tails. Our study provides quantitative guidelines for experiments aiming to assemble "hybrid" structures, i.e., protein shells around charged and impenetrable cores, especially in the limit when the number of cores is much smaller than the number of proteins.

ACKNOWLEDGMENTS

The authors would like to acknowledge helpful discussions with B. Dragnea. A.Š. acknowledges support by the Ministry of Science, Education, and Sports of Republic of Croatia (Project No. 035-0352828-2837). R.P. acknowledges the financial support by the Slovenian Research Agency under Contracts No. P1-0055 (Biophysics of Polymers, Membranes, Gels, Colloids and Cells) and No. J1-0908 (Active media nanoactuators with dispersion forces). This study was supported in part by the Intramural Research Program of the NIH, Eunice Kennedy Shriver National Institute of Child Health (R.P.) and NSF Grant No. DMR-06-45668 (R. Z.).

APPENDIX: DERIVATION OF DEBYE-HÜCKEL FORMULAS

We solve the linearized Poisson-Boltzmann equation for the system with tail-less capsomeres. The electrostatic potential can be written in region I (see Fig. 1) as

$$\Phi_I(r) = C, \tag{A1}$$

in region II as

$$\Phi_{II}(r) = A\frac{\exp(-\kappa r)}{r} + B\frac{\exp(\kappa r)}{r}, \tag{A2}$$

and in region III as

$$\Phi_{III}(r) = D\frac{\exp(-\kappa r)}{r}, \tag{A3}$$

where κ is the Debye-Hückel screening length and A, B, C, and D are unknown constants that are to be determined from the two boundary conditions at R_1 and two at R_2. This yields

$$A = \frac{e^{\kappa(R_1-R_2)}[2\kappa R_1^2\sigma_1 e^{\kappa R_2} + (\kappa R_1 - 1)R_2\sigma_2 e^{\kappa R_1}]}{2\epsilon_0\epsilon_r\kappa(1+\kappa R_1)}, \tag{A4}$$

$$B = \frac{R_2\sigma_2 e^{-\kappa R_2}}{2\epsilon_0\epsilon_r\kappa}, \tag{A5}$$

$$C = \frac{R_1\sigma_1 + R_2\sigma_2 e^{\kappa(R_1-R_2)}}{\epsilon_0\epsilon_r(1+\kappa R_1)}, \tag{A6}$$

$$D = \frac{e^{-\kappa R_2}}{2\epsilon_0\epsilon_r\kappa(1+\kappa R_1)}[2\kappa R_1^2\sigma_1 e^{\kappa(R_1+R_2)} + (\kappa R_1 - 1)R_2\sigma_2 e^{2\kappa R_1} + (1+\kappa R_1)R_2\sigma_2 e^{2\kappa R_2}]. \tag{A7}$$

The electrostatic free energy can be written as

$$F = \int d^3r \frac{Q(r)\Phi(r)}{2}, \tag{A8}$$

which yields

$$F = \frac{\pi e^{-2\kappa R_2}}{\epsilon_0\epsilon_r\kappa(1+\kappa R_1)}\{4\kappa R_1^2 R_2\sigma_1\sigma_2 e^{\kappa(R_1+R_2)} + (\kappa R_1 - 1)R_2^2\sigma_2^2 e^{2\kappa R_1} + [2\kappa R_1^3\sigma_1^2 + (1+\kappa R_1)R_2^2\sigma_2^2]e^{2\kappa R_2}\}. \tag{A9}$$

[1] S. J. Flint, L. W. Enquist, V. R. Racaniello, and A. M. Skalka, *Principles of Virology* (ASM Press, Washington, D.C., 2004).
[2] T. Douglas and M. Young, Adv. Mater. **11**, 679 (1999); Nature (London) **393**, 152 (1998); Science **312**, 873 (2006).
[3] S.-W. Lee, C. Mao, C. E. Flynn, and A. M. Belcher, Science **296**, 892 (2002).
[4] J. D. Lewis, G. Destito, A. Zijlstra, M. J. Gonzalez, J. P. Quigley, M. Manchester, and H. Stuhlmann, Nat. Med. **12**, 354 (2006).
[5] M. Comellas-Aragonès, H. Engelkamp, V. I. Claessen, N. A. J. M. Sommerdijk, A. E. Rowan, P. C. M. Christianen, J. C. Maan, B. J. M. Verduin, J. J. L. M. Cornelissen, and R. J. M. Nolte, Nat. Nanotechnol. **2**, 635 (2007).
[6] K. Mayo, D. Huseby, J. McDermott, B. Arvidson, L. Finlay, and E. Barklis, J. Mol. Biol. **325**, 225 (2003).
[7] J. Benjamin, B. K. Ganser-Pornillos, W. F. Tivol, W. I. Sundquist, and G. J. Jensen, J. Mol. Biol. **346**, 577 (2005).
[8] B. K. Ganser, S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist, Science **283**, 80 (1999).
[9] S. Mukherjee, C. M. Pfeifer, J. M. Johnson, J. Liu, and A. Zlotnick, J. Am. Chem. Soc. **128**, 2538 (2006).
[10] J. P. Michel, I. L. Ivanovska, M. M. Gibbons, W. S. Klug, C. M. Knobler, G. J. L. Wuite, and C. F. Schmidt, Proc. Natl. Acad. Sci. U.S.A. **103**, 6184 (2006).
[11] Y. Hu, R. Zandi, A. Anavitarte, C. M. Knobler, and W. M. Gelbart, Biophys. J. **94**, 1428 (2008).
[12] A. Šiber and R. Podgornik, Phys. Rev. E **79**, 011919 (2009).
[13] T. Hu, R. Zhang, and B. I. Shklovskii, Physica A **387**, 3059 (2008).
[14] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, Proc. Natl. Acad. Sci. U.S.A. **101**, 15556 (2004).
[15] D. C. Rapaport, Phys. Rev. Lett. **101**, 186101 (2008).
[16] T. Chen and S. C. Glotzer, Phys. Rev. E **75**, 051504 (2007).
[17] H. D. Nguyen and C. L. Brooks III, Nano Lett. **8**, 4574 (2008).
[18] T. Keef, A. Taormina, and R. Twarock, Phys. Biol. **2**, 175 (2005).
[19] A. Šiber and R. Podgornik, Phys. Rev. E **76**, 061906 (2007).
[20] M. F. Hagan and D. Chandler, Biophys. J. **91**, 42 (2006).
[21] A. Levandovsky and R. Zandi, Phys. Rev. Lett. **102**, 198102 (2009).
[22] D. L. D. Caspar and A. Klug, Cold Spring Harb Symp. Quant Biol. **27**, 1 (1962).
[23] A. Klug, Philos. Trans. R. Soc. London, Ser. B **354**, 531 (1999).
[24] J. B. Bancroft, Adv. Virus Res. **16**, 99 (1970).
[25] K. W. Adolph and P. J. G. Butler, J. Mol. Biol. **88**, 327 (1974).
[26] L. Lavelle, M. Gingery, M. Phillips, W. M. Gelbart, C. M. Knobler, R. D. Cadena-Nava, J. R. Vega-Acosta, L. A. Pinedo-Torres, and J. Ruiz-Garcia, J. Phys. Chem. B **113**, 3813 (2009).
[27] J. Sun, C. DuFord, M.-C. Daniel, A. Murali, C. Chen, K. Gopinath, B. Stein, M. De, V. M. Rotello, A. Holzenburg, C. C. Kao, and G. Dragnea, Proc. Natl. Acad. Sci. U.S.A. **104**, 1354 (2007).
[28] M. F. Hagan, Phys. Rev. E **77**, 051904 (2008).
[29] R. Zandi and P. van der Schoot, Biophys. J. **96**, 9 (2009).
[30] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, Phys. Rev. Lett. **90**, 248101 (2003).
[31] P. van der Schoot and R. Bruinsma, Phys. Rev. E **71**, 061928 (2005).
[32] V. A. Belyi and M. Muthukumar, Proc. Natl. Acad. Sci. U.S.A. **103**, 17174 (2006).
[33] A. Šiber and R. Podgornik, Phys. Rev. E **78**, 051915 (2008).
[34] S. E. Aniagyei, C. J. Kennedy, B. Stein, D. A. Willits, T. Douglas, M. J. Young, M. De, V. M. Rotello, D. Srisathiyanarayanan, C. C. Kao, and B. Dragnea, Nano Lett. **9**, 393 (2009).
[35] M. F. Hagan, J. Chem. Phys. **130**, 114902 (2009).
[36] J. F. Joanny, L. Leibler, and P. G. Degennes, J. Polym. Sci., Polym. Phys. Ed. **17**, 1073 (1979).
[37] R. Podgornik, J. Phys. Chem. **97**, 3927 (1993).
[38] J. Israelachvili, *Intermolecular & Surface Forces* (Academic Press, New York, 1991).
[39] W. K. Kegel and P. van der Schoot, Biophys. J. **86**, 3905 (2004).
[40] A. Šiber and A. Majdandžić, Phys. Rev. E **80**, 021910 (2009).