

Supplementary Information

A theoretical model for the dynamic structure of Hepatitis B nucleocapsid

Dong Meng, Rex P Hjelm, Jianming Hu, and Jianzhong Wu

This supplementary material provides the detailed DFT equations and the simulation protocol for predicting the structure of Hepatitis B virus (HBV) nucleocapsids during various stages of replication.

A. Density functional theory

We use a classical density functional theory (DFT) to calculate the radial distributions of nucleic acids and the C-terminal domain (CTD) of HBV capsids at the beginning and final stages of the replication. The DFT equations and their numerical performance in comparison with experiment and simulation results for free and tethered polyelectrolytes have been reported in a number of previous publications (1-6). It has also been shown that the DFT accurately accounts for the electrostatic interactions and excluded volume effects important for RNA and DNA packaging(7,8). Here we reproduce only the detailed equation used in calculation of the density profiles.

In this work, the HBV nucleocapsid is represented by a polyelectrolyte system with the density profiles of nucleic acids, CTD, and monomeric ions vary only in the radial direction. The DFT predicts that the radial distribution functions of small ions and polyion segments can be calculated from the “modified” Poisson-Boltzmann equations, respectively,

$$\rho_{\alpha}(r) = \rho_{\alpha}^b \exp\left[\beta\mu_{\alpha}^{ex} - \beta\lambda_{\alpha}(r)\right] \quad (1)$$

$$\rho_p(r) = \frac{\exp[-\beta\lambda_p(r)]G^i(r)G^{M+1-i}(r)}{\int 4\pi r^2 dr \exp[-\beta\lambda_p(r)]G^i(r)G^{M+1-i}(r)} \quad (2)$$

$$\rho_{DNA}(r) = \frac{\exp[-\beta\lambda_{DNA}(r)] \int_{|\sigma_{DNA}-r|}^{\sigma_{DNA}+r} dr' r' G_{DNA}(r, r') \exp[-\beta\lambda_{DNA}(r')] G_{DNA}(r', r)}{\int 4\pi r^2 dr \exp[-\beta\lambda_{DNA}(r)] \int_{|\sigma_{DNA}-r|}^{\sigma_{DNA}+r} dr' r' G_{DNA}(r, r') \exp[-\beta\lambda_{DNA}(r')] G_{DNA}(r', r)}$$

where $\beta = 1/(k_B T)$, k_B is the Boltzmann constant, T denotes the absolute temperature; subscript α stands for a monomeric ion (e.g., cations and anions), p for a polymer segment in the coarse-grained representations of RNA, CTD, and ds-DNA; M is the number of segments for the polymer chain, and $\lambda_k(r)$ represents an effective one-body potential for a polyion segment or a monomeric ion. In Eq.(1), the excess chemical potentials of monomeric ions, μ_α^{ex} , are calculated from the mean-spherical approximation (MSA) for the electrolyte at the bulk concentration ρ_α^b (9). In Eq.(2), $G^i(r)$ stands for a chain propagator function (i.e., Green function), which is determined from the recursive relation

$$G^i(r) = \frac{1}{2\sigma_p r} \int_{|\sigma_p-r|}^{r+\sigma_p} dr' \exp[-\beta\lambda_p(r')] G^{i-1}(r') r' \quad (3)$$

for $i = 2, \dots, M$ with $G^1(r) = 1$, and σ_p is the segment diameter. In calculation of the density profiles for the CTD segments, the propagator function for each polymer segment depends on its position along the chain and its identity. In calculating the radial distribution of DNA segments, we adopted the method proposed by Woodward and Forsman for semi-flexible chains (10). Under the “no chain-end effects” assumption, the chain propagator $G_{DNA}(r, r')$ in Eq.(3) is given by

$$G_{DNA}(r, r') = \frac{1}{2\sigma_{DNA} r} \int_{|\sigma_{DNA}-r|}^{r+\sigma_{DNA}} dr'' r'' G_{DNA}(r'', r) e[-\beta\lambda_{DNA}(r'') - \varepsilon_D(1 + \cos\theta \cos\alpha)] I_0(\sin\theta \sin\alpha) \quad (4)$$

where ε_D represents the bending energy, and I_0 is a modified Bessel function

$$I_0(\sin \theta \sin \alpha) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \exp[-\varepsilon_D \sin \theta \sin \alpha \cos \varphi] \quad (5)$$

and the trigonometric functions are given by

$$\cos \alpha = \frac{r^2 + \sigma_{DNA}^2 - r'^2}{2r\sigma_{DNA}}, \quad (6)$$

$$\cos \theta = \frac{r^2 + \sigma_{DNA}^2 - r''^2}{2r\sigma_{DNA}}. \quad (7)$$

The one-body potential accounts for the thermodynamic non-ideality due to intermolecular interactions. It depends on the external potential $\varphi_k(r)$ due to the capsid wall, the mean electrostatic potential $\psi(r)$, and the local excess chemical potential due to excluded volume and correlation effects $\mu_{hsc}^{ex}(r)$:

$$\lambda_k(r) = \varphi_k(r) + Z_k e \psi(r) + \mu_{hsc}^{ex}(r). \quad (8)$$

where Z_k represents the electric valence of particle k , and e is the unit charge. The mean electrostatic potential, in reduced form, is given by

$$\psi^*(r) \equiv \beta e \psi(r) = \psi^*(r=0) - 4\pi l_B \int_0^r dr' \left(r' - \frac{r'^2}{r} \right) \sum_{k=p,+,-} Z_k \rho_k(r'). \quad (9)$$

The Bjerrum length $l_B = 0.78$ nm is related to the distance between two monovalent spheres (small ions or polymer segments) where the electrostatic potential is equal to the thermal energy $k_B T$. In Eq.(9), $\psi^*(r=0)$ is evaluated by applying another boundary condition $\psi(r=\infty)=0$.

The local excess chemical potential is obtained from a functional derivative of the excess Helmholtz energy with respect to the one-body density:

$$\mu_{hsc}^{ex}(r) = \frac{\delta F_{hsc}^{ex}}{\delta \rho_{\alpha}(r)} \quad (10)$$

where F_{hsc}^{ex} is the free energy due to the particle excluded volume and correlation effects. Within the coarse-grained model considered in this work, the excess Helmholtz energy functional can be decomposed into three inter-related contributions, i.e.,

$$F_{hsc}^{ex} = F_{hs}^{ex} + F_{ch}^{ex} + F_{el}^{ex} \quad (11)$$

where F_{hs}^{ex} is the excess Helmholtz energy functional due to hard-sphere repulsions, F_{ch}^{ex} accounts for the effect of chain connectivity on the intramolecular correlations, and F_{el}^{ex} arises from the hard-sphere and electrostatic correlations. Eq.(11) should include additional terms if the solvent-mediated interactions such as hydrophobic or hydration forces are considered.

The Helmholtz energy functional due to hard-sphere interactions is represented by the modified fundamental measure theory (FMT) (11) or the so-called ‘‘White Bear version’’ (12):

$$\begin{aligned} \beta F_{hs}^{ex} = \int d\mathbf{r} \{ & -n_0 \ln(1-n_3) + \frac{n_1 n_2 - \mathbf{n}_{V1} \mathbf{n}_{V2}}{1-n_3} \\ & + \frac{1}{36\pi} [n_3 \ln(1-n_3) + \frac{n_3^2}{(1-n_3)^2}] \frac{(n_2^3 - 3n_2 \mathbf{n}_{V2} \mathbf{n}_{V2})}{n_3^3} \} \end{aligned} \quad (12)$$

where $\{n_{\alpha}, \alpha = 1, 2, 3, V1, V2\}$ are scalar and vector weighted densities that defined in Rosenfeld’s original work(13).

In addition to the excluded-volume effects, the Helmholtz energy should also include contributions due to the intra-chain and electrostatic correlations. The intra-chain correlation accounts for the free-energy penalty in constraining the chain connectivity of the polymeric segments(14). According to an extension of the first-order thermodynamic perturbation theory

(TPT1) for inhomogeneous polymeric systems(15), the excess Helmholtz energy functional due to the intra-chain correlation is given by

$$\beta F_{ch}^{ex} = \frac{1-M}{M} \int n_{0p} \zeta_p \ln y(\sigma_p, n_\omega) d\mathbf{r} \quad (13)$$

where $\zeta_p = 1 - \mathbf{n}_{v2p} \mathbf{n}_{v2p} / n_{2p}^2$ is introduced to account for the effect of local density variation, and $y(\sigma_p, n_\omega)$ is the contact value of the cavity correlation function (CCF) of the polymeric segments. A key assumption of TPT1 is that CCF can be represented by that corresponding to a monomeric reference system(16)

$$y(\sigma_p, n_\omega) = \left[\frac{1}{1-n_3} + \frac{n_2 \sigma_p (1 - \mathbf{n}_{v2} \mathbf{n}_{v2} / n_2^2)}{4(1-n_3)^2} \right] \exp\left(-\frac{\Gamma^2 a_p^2}{4\pi^2 l_B \sigma_p}\right) \exp\left(\frac{l_B Z_p^2}{\sigma_p}\right). \quad (14)$$

where parameters Γ and a_p are calculated from

$$\Gamma = \sqrt{\pi l_B \sum_{k=p,+,-} n_{0k} \left(\frac{1}{1+\Gamma \sigma_k}\right)^2 \left(Z_k - \frac{\pi P_n \sigma_k^2}{2(1-n_3)}\right)^2} \quad (15)$$

$$a_p = \frac{2\pi l_B \left(Z_p - \frac{\pi P_n \sigma_p^2}{2(1-n_3)}\right)}{\Gamma(1+\Gamma \sigma_p)} \quad (16)$$

with

$$P_n = \sum_{k=p,+,-} \frac{2n_{1k} Z_k}{1+\Gamma \sigma_k} \left/ \left(1 + \frac{3}{(1-n_3)} \sum_{k=p,+,-} \frac{n_{3k}}{1+\Gamma \sigma_k}\right)\right. \quad (17)$$

Eq.(14) is derived from a modification of the mean-spherical approximation (MSA) theory for monomeric electrolyte solutions(9).

As for monomeric electrolyte systems, the hypernetted chain approximation (HNC) is often sufficient to describe the Helmholtz energy due to the long-range electrostatic

correlations(17,18). According to HNC, the excess Helmholtz energy can be derived from a quadratic functional expansion with respect to that for a bulk fluid of uniform densities $\{\rho_i^b\}$

$$\begin{aligned} \beta F_{el}^{ex} = & \beta F_{el}^{ex}[\{\rho_i^b\}] - \int d\mathbf{r} \sum_{i=p,+,-} \Delta C_i^{(1)el} (\rho_i(\mathbf{r}) - \rho_i^b) \\ & - \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \sum_{i,j=p,+,-} \Delta C_{ij}^{(2)el} (|\mathbf{r} - \mathbf{r}'|) (\rho_i(\mathbf{r}) - \rho_i^b) (\rho_j(\mathbf{r}') - \rho_j^b) \end{aligned} \quad (18)$$

where $\Delta C_i^{(1)el}$ and $\Delta C_{ij}^{(2)el}$ represent the first-order and the second-order direct correlation functions (DCF) due to the electrostatic interactions in the reference system. In this work, we use analytical expressions of the DCF from the MSA(9).

B. Numerical implementation

Equations (1) and (2) are solved with the Picard iteration method. The numerical procedure starts with an initial guess for the density profiles of polyions and monomeric ions based on their bulk densities. The reduced electrostatic potential $\psi^*(r)$, the effective fields for polyion segments and monomeric ions, $\lambda_k(r)$, and the Green function $G^i(r)$ for polyions is then calculated from Eqs.(3). Subsequently, new density profiles are updated and mixed with the previous results as the new inputs. The iteration procedure repeats until the change in the input and output density profiles of polyions and monomeric ions is smaller than 0.01% at all positions.

References

1. Jiang, T., Z. D. Li and J. Z. Wu. 2007. Structure and swelling of grafted polyelectrolytes: Predictions from a nonlocal density functional theory. *Macromolecules*. 40: 334-343.
2. Li, Z. D. and J. Z. Wu. 2006. Density-functional theory for polyelectrolytes near an oppositely-charged surface. *Physical Review Letters*. 96: 048302.
3. Li, Z. D. and J. Z. Wu. 2006. Density functional theory for planar electric double layers: Closing the gap between simple and polyelectrolytes. *Journal of Physical Chemistry B*. 110: 7473-7484.
4. Wu, J. Z. and T. Jiang. 2008. Ionic effects in collapse of polyelectrolyte brushes. *Journal of Physical Chemistry B*. 112: 7713-7720.
5. Jiang, T. and J. Z. Wu. 2008. Self-organization of multivalent counterions in polyelectrolyte brushes. *J Chem Phys*. 129: 10.1063/1061.2966359.

6. Wang, L., H. J. Liang and J. Z. Wu. 2010. Electrostatic origins of polyelectrolyte adsorption: Theory and Monte Carlo simulations. *J Chem Phys.* 133: 10.1063/1061.3463426.
7. Jiang, T., Z. G. Wang and J. Z. Wu. 2009. Electrostatic Regulation of Genome Packaging in Human Hepatitis B Virus. *Biophys J.* 96: 3065-3073.
8. Li, Z. D., J. Z. Wu and Z. G. Wang. 2008. Osmotic pressure and packaging structure of caged DNA. *Biophys J.* 94: 737-746.
9. Blum, L. 1975. Mean spherical model for asymmetric electrolytes .1. Method of solution. *J. Stat. Phys. (USA).* 30: 1529-1535.
10. Woodward, C. E. and J. Forsman. 2006. Density functional theory for flexible and semiflexible polymers of infinite length. *Physical Review E.* 74: 010801.
11. Yu, Y. X. and J. Z. Wu. 2002. Structures of hard sphere fluids from a modified fundamental measure theory. *Journal of Chemical Physics.* 117: 10156-10164.
12. Roth, R., R. Evans, A. Lang and G. Kahl. 2002. Fundamental measure theory for hard-sphere mixtures revisited: the White Bear version. *Journal of Physics-Condensed Matter.* 14: 12063-12078.
13. Rosenfeld, Y. 1989. Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing. *Physical Review Letters.* 63: 980-983.
14. Chandler, D. and L. R. Pratt. 1976. Statistical mechanics of chemical equilibria and intramolecular structures of nonrigid molecules in condensed phases. *Journal of Chemical Physics.* 65: 2925-2940.
15. Yu, Y. X. and J. Z. Wu. 2002. Density functional theory for inhomogeneous mixtures of polymeric fluids. *Journal of Chemical Physics.* 117: 2368-2376.
16. Jiang, J. W., H. L. Liu, Y. Hu and J. M. Prausnitz. 1998. A molecular-thermodynamic model for polyelectrolyte solutions. *Journal of Chemical Physics.* 108: 780-784.
17. Yu, Y. X., J. Z. Wu and G. H. Gao. 2004. Density-functional theory of spherical electric double layers and zeta potentials of colloidal particles in restricted-primitive-model electrolyte solutions. *Journal of Chemical Physics.* 120: 7223-7233.
18. Li, Z. D. and J. Z. Wu. 2004. Density-functional theory for the structures and thermodynamic properties of highly asymmetric electrolyte and neutral component mixtures. *Physical Review E.* 70: 031109.