

To what extent of ion neutralization can multivalent ion distributions around RNA-like macroions be described by Poisson-Boltzmann theory?

Gui Xiong(熊贵), Kun Xi(席昆), Xi Zhang(张曦), Zhi-Jie Tan(谭志杰)
Citation: Chin. Phys. B . 2018, 27(1): 018203. **doi:** 10.1088/1674-1056/27/1/018203

Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Molecular dynamics simulation of the response of bi-disperse polyelectrolyte brushes to external electric fields

Fen Zhang(张芬), Huan-Da Ding(丁欢达), Chao Duan(段超), Shuang-Liang Zhao(赵双良), Chao-Hui Tong(童朝晖)
Chin. Phys. B . 2017, 26(8): 088204. **doi:** 10.1088/1674-1056/26/8/088204

Theoretical studies on sRNA-mediated regulation in bacteria

Chang Xiao-Xue, Xu Liu-Fang, Shi Hua-Lin
Chin. Phys. B . 2015, 24(12): 128703. **doi:** 10.1088/1674-1056/24/12/128703

RNA structure prediction：Progress and perspective

Shi Ya-Zhou, Wu Yuan-Yan, Wang Feng-Hua, Tan Zhi-Jie
Chin. Phys. B . 2014, 23(7): 078701. **doi:** 10.1088/1674-1056/23/7/078701

Self-consistent field theory of adsorption of flexible polyelectrolytes onto an oppositely charged sphere

Tong Zhao-Yang, Zhu Yue-Jin, Tong Chao-Hui
Chin. Phys. B . 2014, 23(3): 038202. **doi:** 10.1088/1674-1056/23/3/038202

SPECIAL TOPIC — Soft matter and biological physics

To what extent of ion neutralization can multivalent ion distributions around RNA-like macroions be described by Poisson–Boltzmann theory?*

Gui Xiong(熊贵), Kun Xi(席昆), Xi Zhang(张曦)†, and Zhi-Jie Tan(谭志杰)‡

Center for Theoretical Physics and Key Laboratory of Artificial Micro & Nano-structures of Ministry of Education, School of Physics and Technology, Wuhan University, Wuhan 430072, *China*

(Received 31 August 2017; revised manuscript received 23 October 2017; published online 13 December 2017)

Nucleic acids are negatively charged biomolecules, and metal ions in solutions are important to their folding structures and thermodynamics, especially multivalent ions. However, it has been suggested that the binding of multivalent ions to nucleic acids cannot be quantitatively described by the well-established Poisson–Boltzmann (PB) theory. In this work, we made extensive calculations of ion distributions around various RNA-like macroions in divalent and trivalent salt solutions by PB theory and Monte Carlo (MC) simulations. Our calculations show that PB theory appears to underestimate multivalent ion distributions around RNA-like macroions while can reliably predict monovalent ion distributions. Our extensive comparisons between PB theory and MC simulations indicate that when an RNA-like macroion gets ion neutralization beyond a "critical" value, the multivalent ion distribution around that macroion can be approximately described by PB theory. Furthermore, an empirical formula was obtained to approximately quantify the critical ion neutralization for various RNA-like macroions in multivalent salt solutions, and this empirical formula was shown to work well for various real nucleic acids including RNAs and DNAs.

Keywords: nucleic acids, ion binding, Poisson–Boltzmann theory, Monte Carlo simulation

PACS: 82.35.Rs, 87.14.gn, 87.10.Rt, 87.14.G– **DOI:** 10.1088/1674-1056/27/1/018203

1. Introduction

Nucleic acids are of biological significance in genetic storage, expression, and regulation.[1–6] Since nucleic acids are negatively charged molecules, metal ions in solutions could neutralize the negative charges on nucleic acids and favor their folding into native and functional structures.[1,7–15] Specifically, metal ions could enhance the stability of secondary and tertiary structures of nucleic acids[7,16,17] and could significantly influence nucleic acids helix assembly.[18–22] Importantly, multivalent ions such as Mg^{2+} and $Cohex^{3+}$ play very efficient roles in DNA condensation[23–28] and RNA structural folding.[17,29–33]

To understand ion-nucleic acid interactions, several classic polyelectrolyte theories have been developed, including counterion condensation theory,[34–36] Poisson–Boltzmann (PB) theory,[37–42] modified PB theory,[43–48] and tightly bound ion (TBI) theory.[30,49] Among these, the PB theory is a well-defined theory with concise form and efficient solving algorithms, and has been widely applied to charged systems such as colloids, proteins, and nucleic acids.[50–57] However, due to the mean-field nature of the nonlinear PB theory, it ignores the discrete properties of ions, consequently underestimating the distributions of multivalent ions around highly charged polyelectrolytes such as nucleic acids,[27,58–62] and generally underestimates the efficient role of multivalent ions in the folding stability of nucleic acids.[63–68]

All-atom molecular dynamics and Monte Carlo (MC) simulations are widely used as important methods to complement theories and experiments in simulating charged systems and make predictions that agree well with experiments in a wide range of ion conditions.[23,69–76] Since multivalent ions are so important to nucleic acid structures and PB theory is concise and has efficient solving algorithms,[40,41,66] it is interesting to understand how much PB predictions deviate for multivalent ion distributions around nucleic acids and to quantify the extent to which ionic neutralization, multivalent ion distributions around nucleic acids can be approximately described by PB theory.

In the present work, we made extensive calculations on multivalent ion distributions around RNA-like macroions by PB theory and MC simulations, attempting to examine how much PB predictions deviate from the MC simulations. Furthermore, we attempted to obtain an empirical formula for quantifying the extent of ionic neutralization beyond which the multivalent ion distributions around RNA-like macroions can be approximately described by PB theory. Finally, this empirical formula was examined for various real RNAs and

*Project supported by the National Natural Science Foundation of China (Grant Nos. 11374234, 11575128, 11774272, and 11647312).

†Corresponding author. E-mail: xizhang@whu.edu.cn

‡Corresponding author. E-mail: zjtan@whu.edu.cn

© 2018 Chinese Physical Society and IOP Publishing Ltd

DNAs.

2. Model and methods

2.1. Model system for RNA-like macroions and RNAs

In this work, for simplicity, we used spherical macroions (RNA-like macroions) of various sizes and charge densities to model various RNAs structures, as shown in Fig. 1(a). To determine the surface charge densities of the macroions, a statistical analysis was made on 80 RNAs whose sizes (number of nucleotides) N range from 22-nt to 158-nt. The distribution of surface charge density σ for the RNAs is shown in Fig. 1(b), and the PDB codes of the 80 RNAs are listed in the Appendix. According to the size and surface charge density of the RNAs, the radii R of RNA-like macroions were taken as 10 Å, 15 Å, 20 Å, 25 Å, and 30 Å to approximately cover the wide range of RNA size, and the surface charge density values σ on RNA-like spherical macroions were given the typical values around 0.0143 e/Å^2, where $\sigma = N/4\pi R_g^2$ and R_g is the radius of gyration for an RNA. In addition, various real RNAs and DNAs were also used in the work, including 8-bp and 12-bp RNA duplexes, 12-bp and 24-bp DNA duplexes, 28-nt beet western yellow virus (BWYV) pseudoknot (PDB code: 437D), 58-nt rRNA fragment (PDB code: 1HC8), 76-nt tRNA^{Phe} (PDB code: 2K4C), and 158-nt P4-P6 domain of the tetrahymena ribozyme (PDB code: 1GID).

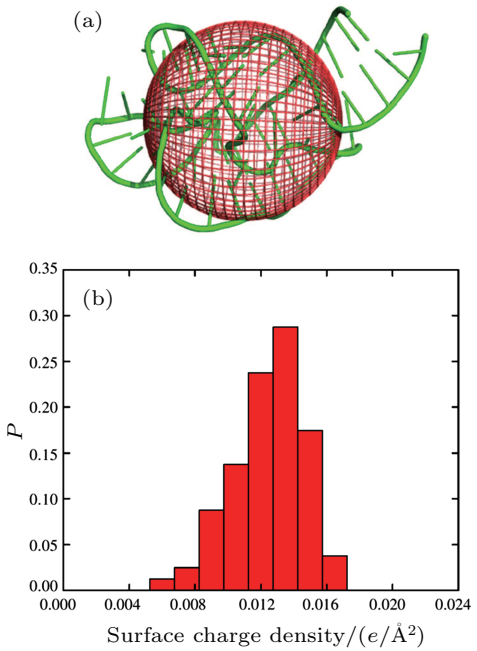


Fig. 1. (color online) (a) Spherical representation of an RNA molecule. (b) Normalized population of surface charge density $\sigma = N/4\pi R_g^2$ from statistical analysis over the 80 RNAs whose lengths N (-nt) are in the range of [22, 158], where N and R_g represent the number of phosphates and the radius of gyration for nucleic acids. The PDB codes of the 80 RNAs are listed in Appendix A.

The RNA-like macroions and real RNAs (and DNAs) were immersed in monovalent, divalent, and trivalent salt solutions, and the solvent was modeled implicitly as a continuous medium of dielectric constant 78. The salts were considered as 1:1, 2:2, and 3:3 salts. The radii of monovalent, divalent, and trivalent cations were taken as 2.7 Å, 3.7 Å, and 4.2 Å,[77] respectively, and the radii of coions were taken as the same as those of the corresponding cations. The salt concentrations covered the typical wide ranges: [1 mM, 100 mM] for monovalent salt, [0.1 mM, 10 mM] for divalent salt, and [0.01 mM, 1 mM] for trivalent salt. For real RNAs and DNAs, the radii of the atoms were taken as their van der Waals radii.[30,51,77] Each phosphorus atom was assumed to carry an electronic charge $-e$, and the other atoms remained neutral.[8,9,17] This has been verified by previous studies as a robust simplified model for the all-atom force field of RNAs and DNAs.[8,9,17] In our systems, the interactions were simplified into Coulombic interaction and hard-core repulsion

$$U_{ij} = \begin{cases} \dfrac{q_i q_j}{4\pi\varepsilon_0\varepsilon r_{ij}}, & r_{ij} \geq a_i + a_j, \\ \infty, & r_{ij} < a_i + a_j, \end{cases} \tag{1}$$

where q_i, a_i and q_j, a_j are the charges and radii of particles i and j (ions or macroions or atoms on RNAs), respectively; r_{ij} is the center-to-center distance between particles i and j; ε_0 is the permittivity of vacuum, and ε is the dielectric constant of solvent.

We employed PB theory and MC simulations to calculate ion distributions around RNA-like macroions or real RNAs and DNAs with different reduced charges N_r due to ion binding, and correspondingly, the reduced surface charge density σ_r was $N_r/4\pi R^2$ for RNA-like macroions. Our calculations, including PB theoretic and MC simulations, were performed for fixed RNA-like macroions, RNAs, and DNAs, which can be less computationally expensive.

2.2. Monte Carlo simulations

Monte Carlo simulations were employed to calculate ion distributions around RNA-like macroions, RNAs, and DNAs.[69,70,78] The simulation system was a canonical ensemble with a collection of ions and a macroion or an RNA (or DNA) in a cubic box at room temperature. In each MC simulation, the macroion or RNA was fixed at the center of the cubic box, and ions were diffusive and mobile in the simulated box. To diminish the boundary effect, the box size was kept at least eight times larger than the Debye–Hückel length of the ionic solution, and periodic boundary condition was employed. According to the Metropolis algorithm, the energy change ΔU due to ion configuration change can be calculated, and the accepted probability for a new configuration is given by $P = \min[\exp(-\Delta U/k_B T), 1]$.[29,71,79] The process was repeated after the system reached equilibrium, and a sufficient number of ion configurations in equilibrium can be used to calculate ion distributions.

2.3. The Poisson–Boltzmann equation and numerical solution

The nonlinear PB theory is a well-established mean-field theory, in which the solvent is treated as a continuous medium with a dielectric constant and ions obey Boltzmann distributions based on mean electrostatic potential ψ[40,44,47]

$$\nabla\cdot[\varepsilon_0\varepsilon(\boldsymbol{r})\nabla\psi(\boldsymbol{r})] = -4\pi\left\{\rho_{\rm f}+\sum_\alpha z_\alpha e c_\alpha^0\,{\rm e}^{-z_\alpha e\psi(\boldsymbol{r})/k_{\rm B}T}\right\}. \quad (2)$$

Here, $\rho_{\rm f}$ is the charge density of fixed charges, and $z_\alpha e$ is the charge of ion species α; $\varepsilon(\boldsymbol{r})$ is the dielectric constant at $\boldsymbol{r}$ and ε_0 is the permittivity of vacuum; and c_α^0 is the bulk concentration of ion species α.

In the present work, the three-dimensional algorithm developed in the TBI theory was used to numerically solve the nonlinear PB equation.[8,9,49] A thin layer of one ion radius was added to a macroion or RNA surface to take the ions' excluded volume layer into account, and the three-step focusing process was employed to compute the detailed electrostatic potential near the macroion or RNA surface.[46,71] The grid size in the first run of the three-step focusing process depends on the salt concentration and was taken to be six times larger than the Debye–Hückel length from the macroion/RNA surface to effectively include the salt effect in solutions. The resolution for the first run varied with the grid size to make the iterative process computationally possible and was generally kept as 1 Å and 0.5 Å per grid for the second and the third runs, respectively. The iteration for a run was continued until the electrostatic potential change $\delta\psi(\boldsymbol{r})$ for an iteration was less than $10^{-4}k_{\rm B}T/e$.[71]

2.4. Ion distributions around RNA-like macroions and RNAs

In this work, ion distributions around RNA-like macroions or RNAs (and DNAs) were characterized by the net ion charge distribution $Q(r)$ within a distance r from RNA-like macroions or RNAs (and DNAs)[29,80,81]

$$Q(r) = \frac{1}{|N|}\int_{\le r}\sum_\alpha z_\alpha c_\alpha(\boldsymbol{r})\,{\rm d}^3\boldsymbol{r}. \quad (3)$$

Here, $c_\alpha(\boldsymbol{r})$ is the ion concentration of ion species α at $\boldsymbol{r}$ and can be calculated from PB theory or MC simulations. For RNA-like spherical macroions, r represents the well-defined radial distance, and for real RNAs and DNAs, r stands for the distance from surface atoms of the molecule. The net ion charge distribution $Q(r)$ is an important quantity generally coupled to ion-mediated nucleic acid stability[29,80] and effective interaction between like-charged particles.[29,78,81]

3. Results and discussion

In this work, by PB theory and MC simulations, we made extensive calculations of ion distributions for RNA-like macroions with typical surface charge density of real RNAs. We mainly focused on the deviations of ion distributions between PB theory and MC simulations, and attempted to derive an empirical formula for a "critical" ion neutralization, beyond which multivalent ion distributions around RNA-like macroions and real RNAs and DNAs can be approximately described by PB theory.

3.1. RNA-like spherical macroions in salt solution

3.1.1. Monovalent ion distributions around RNA-like macroions

We examined monovalent ion distributions around RNA-like spherical macroions at different monovalent salt concentrations. As shown in Figs. 2 and S1 in the supplementary material, the net ion charge distributions of $Q(r)$ increase with the increase of radial distance r, and this increase appears not very sharp, indicating that monovalent ions bind loosely to the macroions. Figure 2 also shows that the net ion charge distribution $Q(r)$ obtained from PB theory is very

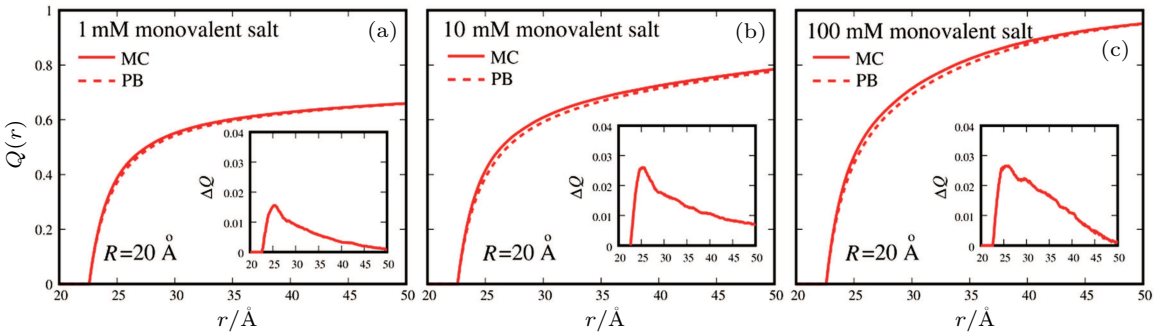


Fig. 2. (color online) Net ion charge distribution $Q(r)$ obtained from PB theory (dashed lines) and MC simulations (solid lines) around an RNA-like spherical macroion in monovalent salt solutions. The radius and surface charge density of the RNA-like spherical macroion are taken to be typical values of 20 Å and 0.0143e/Å^2, respectively. The monovalent salt concentrations are 1 mM (a), 10 mM (b), and 100 mM (c), respectively. Insets show the deviations ΔQ between distributions $Q(r)$ of PB theory and MC simulations. Distributions $Q(r)$ and deviations ΔQ for an RNA-like spherical macroion with a different radius can be found in Fig. S1 in the supplementary material.

close to that from the MC simulations, despite very slight deviation. With an increase of monovalent salt concentration, the deviations between PB theory and MC simulation become slightly larger. Nevertheless, the net ion charge distributions of $Q(r)$ obtained from PB theory are always very close to those from the MC simulations for monovalent salt solution (e.g., deviation $\Delta Q(r) < 4\%$), which is consistent with previous studies.[82–84] This is reasonable. Since Coulomb correlations between monovalent ions are generally weak around polyelectrolytes such as colloids and nucleic acids, the mean field assumption in PB theory can be a good approximation for describing monovalent ion distributions.

For a very high salt concentration, exclusion correlations between ions become strong and can cause the slightly lower predicted ion binding of PB theory since ion correlations can drive ions to self-organize to form low-energy states and consequently favor ion binding.[30,49]

3.1.2. Divalent ion distributions around RNA-like macroions

Divalent ions such as Mg^{2+} are important for RNA structures and functions,[8,9,12] and we examined divalent ion distributions around RNA-like macroions. As shown in Figs. 3 and S2–S3 in the supplementary material, the net ion charge distributions of $Q(r)$ increase with the increase of r in a shaper way and become saturated quickly at smaller r than those for monovalent salt, indicating that divalent ions bind in a tighter manner. Figure 3 also shows that although divalent ion binding becomes stronger for high salt concentration, this salt-concentration dependence of divalent ion binding is less pronounced than monovalent ion binding; see also Fig. S2 in the supplementary material. This is because the higher ionic charge of divalent ions (higher than monovalent ions) causes stronger ion–macroion Coulombic attraction and consequently causes the binding of divalent ions to be more enthalpically dependent than that of monovalent ions.

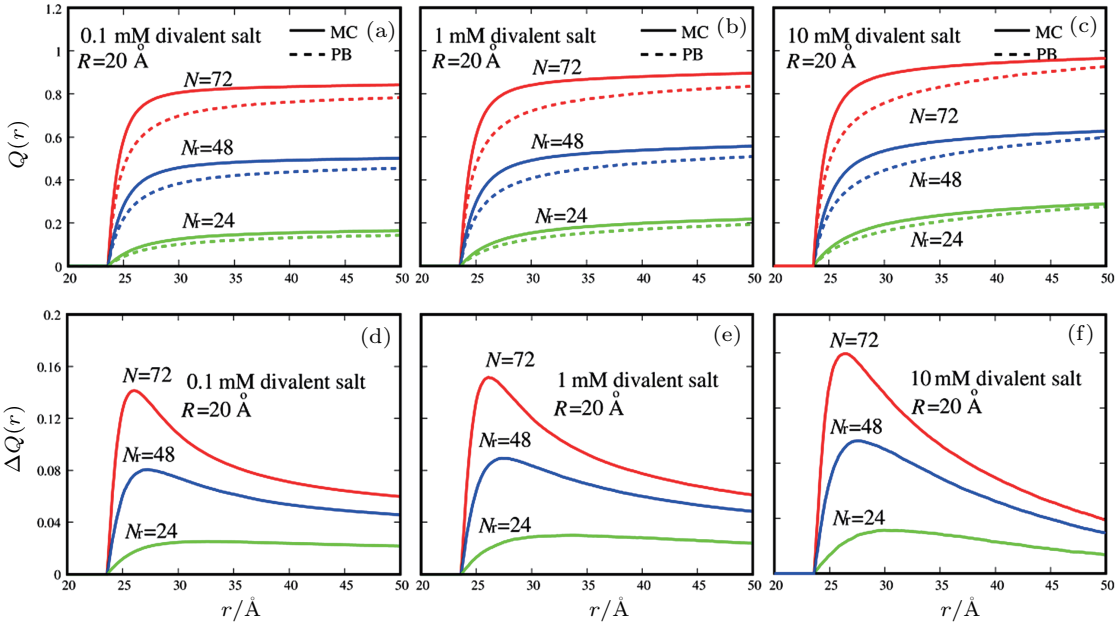


Fig. 3. (color online) (a)–(c) Net ion charge distribution $Q(r)$ obtained from PB theory (dashed lines) and MC simulations (solid lines) around an RNA-like spherical macroion in divalent salt solutions. (d)–(f) Deviations ΔQ between the distributions $Q(r)$ of PB theory and MC simulations. The radius and surface charge density of the RNA-like spherical macroion are taken as the typical values of 20 Å and 0.0143e/Å^2, respectively. Distributions $Q(r)$ and deviations ΔQ for RNA-like spherical macroions with different radii can be found in Figs. S2 and S3 in the supplementary material.

The extensive comparison between PB theory and MC simulations shows that PB theory appears to underestimate the divalent ion distributions, compared to the MC simulations, and these underestimations of ion binding become slightly more apparent for higher salt concentration. This is because PB theory assumes fluid-like ion distribution obeying mean electrostatic field and ignores ion discrete properties such as ion–ion correlations. Such ion–ion correlations can drive ions to self-organize to low-energy micro-states and favor ion binding to polyelectrolytes such as RNAs and DNAs.[30,49] Thus, PB theory apparently underestimates the binding of divalent ions with strong correlations, compared with the MC simulations where ion–ion correlations are explicitly involved. Higher divalent concentration would cause stronger ion–ion correlations around RNA-like macroions and consequently cause more pronounced deviations between the mean-field PB theory and the MC simulations.

3.1.3. Trivalent ion distributions around RNA-like macroions

Trivalent ions can very efficiently play a role in RNA folding[1,2,4,5] and DNA condensation,[25,26] so we examined trivalent ion distributions around RNA-like macroions. As

shown in Figs. 4 and S4–S5 in the supplementary material, the net ion charge distributions of $Q(r)$ increase rapidly with radial distance r and become saturated at small r, indicating that trivalent ions bind to the macroions very tightly. With increasing trivalent salt concentration, $Q(r)$ increases very slightly. As discussed above, the higher ion charge leads to a stronger Coulomb attraction between trivalent ions and RNA-like macroions, which would cause stronger and tighter ion binding. Also, this strong ion binding is more enthalpically dependent than the binding of divalent ions.

Furthermore, figure 4 shows that PB theory underestimates trivalent ion binding around RNA-like macroions more significantly than divalent ions; see Figs. 3, 4 and S2–S5 in the supplementary material. The deviations between PB theory and the MC simulations become slightly more apparent for higher trivalent salt concentration. As discussed above, the correlations between trivalent ions are stronger (than those between divalent ions) and the mean-field approximation adopted in PB theory causes a greater apparent underestimation of trivalent ion binding (e.g., $Q(r)$) than divalent ion binding. The slightly stronger deviation for higher trivalent salt concentration is attributed to the greater number of binding ions and consequently stronger ion–ion correlations around RNA-like macroions.

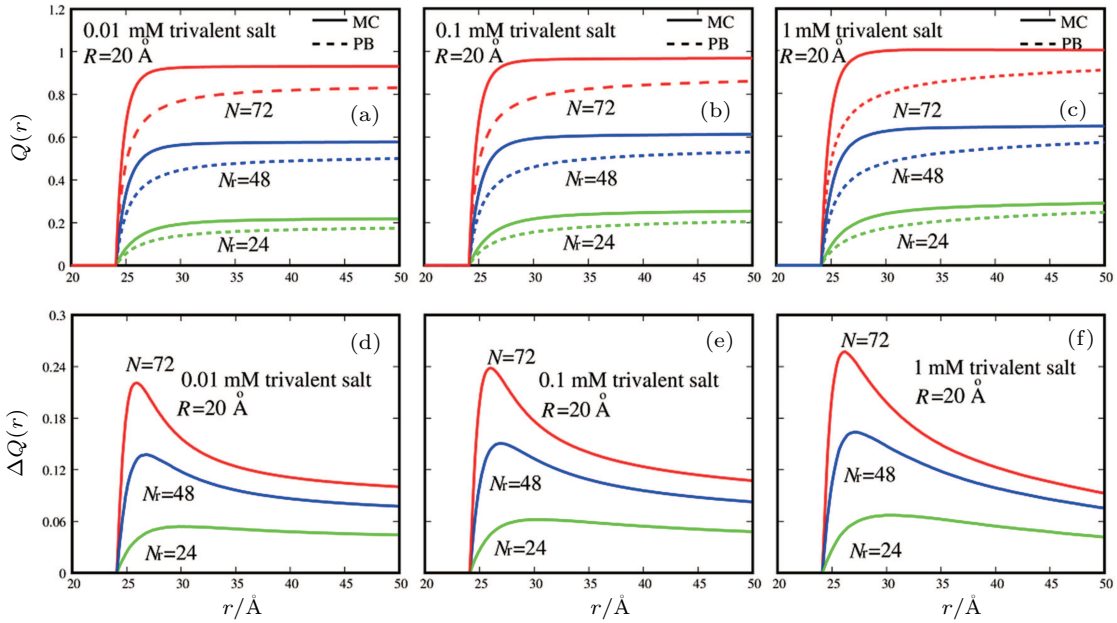


Fig. 4. (color online) (a)–(c) Net ion charge distribution $Q(r)$ obtained from PB theory (dashed lines) and MC simulations (solid lines) around an RNA-like spherical macroion in trivalent salt solutions. (d)–(f) Deviations ΔQ between the distributions $Q(r)$ of PB theory and MC simulations. Radius and surface charge density of the RNA-like spherical macroion are taken to be typical values of 20 Å and 0.0143e/Å^2, respectively. Distributions $Q(r)$ and deviations ΔQ for RNA-like spherical macroions with different radii can be found in Figs. S4 and S5 in the supplementary material.

3.2. To what extent of ion neutralization can multivalent ion distributions around RNA-like macroions be described by PB theory?

As discussed in the above section, PB theory significantly underestimates multivalent ion distribution (e.g., $Q(r)$) around RNA-like macroions because PB theory ignores ion–ion correlations, and correlations between binding ions near macroions are very strong for divalent and trivalent ions. Since such strong correlations between binding ions are induced by the highly charged RNA-like macroions, the correlations between binding ions can become weak when the highly charged macroions get ionic neutralization by binding ions. As shown in Figs. 3, 4 and S2–S5 in the supplementary material for divalent and trivalent salts, when the RNA-like macroions get ionic neutralization, the deviations between the MC simulations and PB theory would become significant less and even become visibly weak. Then an interesting question is: to what extent can ion neutralization and multivalent ion distributions around an RNA-like macroion be approximately described by PB theory?

First, we used $\Delta Q_{\max} = \max[Q_{\mathrm{MC}}(r) - Q_{\mathrm{PB}}(r)]$ to quantify the deviation between the MC simulations and PB theory. As shown in Fig. 5, $\Delta Q_{\max}$ decreases distinctly from a high value to a negligible value when the ionic neutralization of RNA-like macroions gradually becomes strong. For example, for the RNA-like macroion with radius of 20 Å, the deviation $\Delta Q_{\max}$ between the MC simulations and PB theory can become negligible when the "reduced" charge density σ_{r} of the macroion with ion neutralization is less than $\sim 0.005e$/Å^2. Second, to quantify the "critical" extent of ion neutralization, we need to choose a certain value to characterize whether $\Delta Q_{\max}$ is negligible. Our above calculations have shown that $\Delta Q_{\max}$ is always less than $\sim 4\%$ for monovalent salt over a wide range of monovalent salt concentra-

tion, and extensive previous theoretical models combined with experiments have shown that PB theory works quantitatively for describing the effect of monovalent ions on the stability of RNAs and DNAs.[9,10,17,29] Thus, we chose the value of $\sim 4\%$ as a criterion to determine the "critical" extent of ionic neutralization beyond which the multivalent ion distributions around RNA-like macroions can be approximately described by PB theory. Here, interpolation technique was employed since the covered σ_r's cannot be continuous. Based on this criterion, we have obtained a series of critical surface charge densities σ_c by comparing the deviation between MC simulations and PB theory with the criterion for RNA-like macroions with different radii and surface charge densities around the typical value ($\sim 0.0143e/\text{Å}^2$) over wide ranges of divalent and trivalent salt concentrations. Furthermore, as shown in Fig. S6 in the supplementary material, through fitting to the data from the MC simulations and PB theory for divalent and trivalent salt separately, we can derive an empirical formula for describing the critical surface charge densities σ_c

$$z\sigma_c = a\sqrt{\sigma_0} - (b + 0.628/S)\log[c]. \tag{4}$$

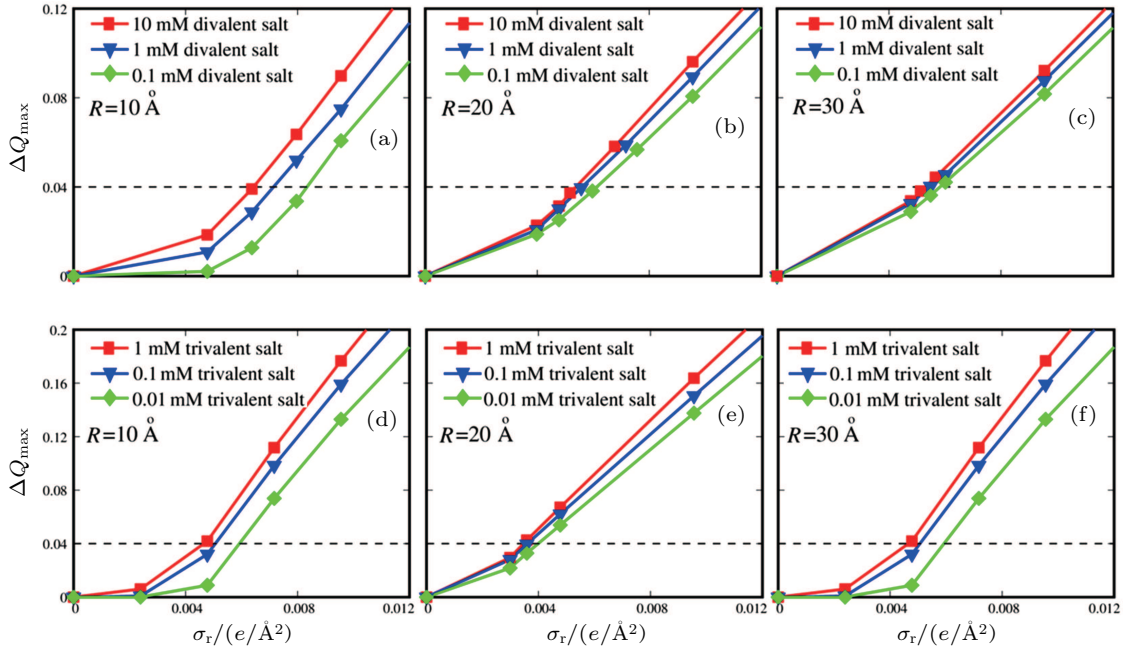


Fig. 5. (color online) Maximum deviation ΔQ_{max} of ΔQ as a function of reduced surface charge density σ_r ($= N_r/4\pi R^2$) due to ion binding to the RNA-like spherical macroions in (a)–(c) divalent and (d)–(f) trivalent salt solutions. The radii of the RNA-like spherical macroions are 10 Å, 20 Å, and 30 Å, respectively. The dashed lines denote the ΔQ_{max} of 4%.

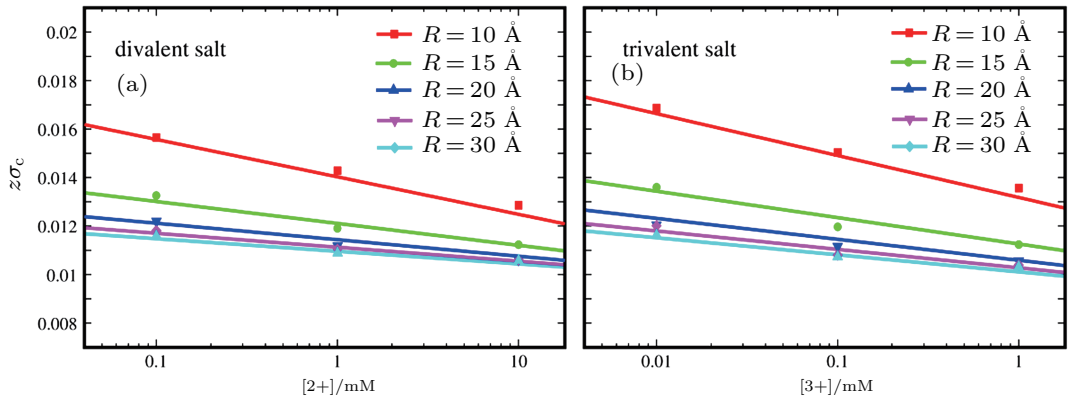


Fig. 6. (color online) "ritical"urface charge density σ_c for RNA-like spherical macroions with surface charge density of $0.0143e/\text{Å}^2$ as a function of divalent salt concentration [2+] (a) and trivalent salt concentration [3+] (b), respectively; z is the valence of divalent ion (a) or trivalent ion (b). The symbols are obtained through extensive comparisons between PB theory and MC simulations, and the lines are from the empirical formula Eq. (4). Different colors denote macroions with different radii: 10 Å (red), 15 Å (green), 20 Å (blue), 25 Å (magenta), and 30 Å (cyan). σ_c for RNA-like spherical macroions with different surface charge densities σ around the typical value of $0.0143e/\text{Å}^2$ can be found in Fig. S6 in the supplementary material.

The parameters a and b are 0.07852 and 0.00017 for divalent salt, and 0.066826 and 0.00025 for trivalent salt, respectively; z is the valence of multivalent ions, $[c]$ is the bulk salt concentration of multivalent ions in mol/L, and S is the macroion surface area in Å^2. The presence of σ_c in Eq. (4) means that when reduced surface charge density σ_r for an RNA-like macroion with ionic neutralization is less than σ_c, the multivalent ion distributions can be approximately described by PB theory. As shown in Figs. 6 and S6 in the supplementary material, the σ_c obtained from the empirical formula is in good agreement with the data from the explicit calculations with MC simulations and PB theory.

3.3. Multivalent ion distributions around real RNAs and DNAs

3.3.1. Ion distributions around RNAs and DNAs

Beyond the above described RNA-like spherical macroions, we examined the multivalent ion distributions around real RNAs and DNAs, including 8-bp and 12-bp RNA duplexes, 12-bp and 24-bp DNA duplexes, 28-nt BWYV pseudoknot, 58-nt rRNA fragment, 76-nt tRNA^{Phe}, and 158-nt P4-P6 domain of the tetrahymena ribozyme. Here, we also used Eq. (3) to calculate the net ion charge distribution while r stands for the distance from surface atoms of the RNA or DNA. As shown in Figs. 7, 8 and S7, S8 in the supplementary material, real RNAs and the above described RNA-like macroions share similar features for net ion charge distributions. (i) The net ion charge distributions of $Q(r)$ increase with increasing distance r from the molecule surface and become saturated at small r, indicating that multivalent ions generally bind to an RNA very tightly. (ii) With an increase of multivalent salt concentration, $Q(r)$ increases only slightly; (iii) PB theory significantly underestimates the multivalent ion binding around RNAs and DNAs. (iv) The deviation between MC simulations and PB theory appears slightly more pronounced for higher-valence multivalent salt.

As discussed above, the higher ion charge leads to stronger Coulomb attraction between ions and RNAs (or DNAs) and causes stronger ion binding, which is consequently more enthalpically dependent and less ion-concentration dependent. The more pronounced deviation between MC simulations and PB theory for ions with higher valence and for higher ion concentration is attributed to the stronger ion–ion correlations. As described above, higher salt concentration and ion valence would cause stronger ion–ion correlations, which can drive ions to self-organize to low energy microstates and favor ion binding.[30,49] Such ion–ion correlations are ignored in PB theory, causing greater apparent deviations of PB theory for higher ion valence and higher concentrations.

In analogy to the above described RNA-like macroions, the value $\Delta Q_{max}(=\max[Q_{MC}(r)-Q_{PB}(r)])$ was used to quantify the deviation between the MC simulations and PB theory, as shown in Fig. S9. The value of ΔQ_{max} for an RNA or DNA decreases distinctively from a large value to a negligible value when the ionic neutralization of the RNA becomes strong gradually. Based on the same criterion as that for RNA-like macroions, we have obtained a series of critical charges N_c for extensive RNAs and DNAs, below which the multivalent ion distributions can be approximately described by PB theory. Here, we also used the interpolation technique for obtaining N_c.

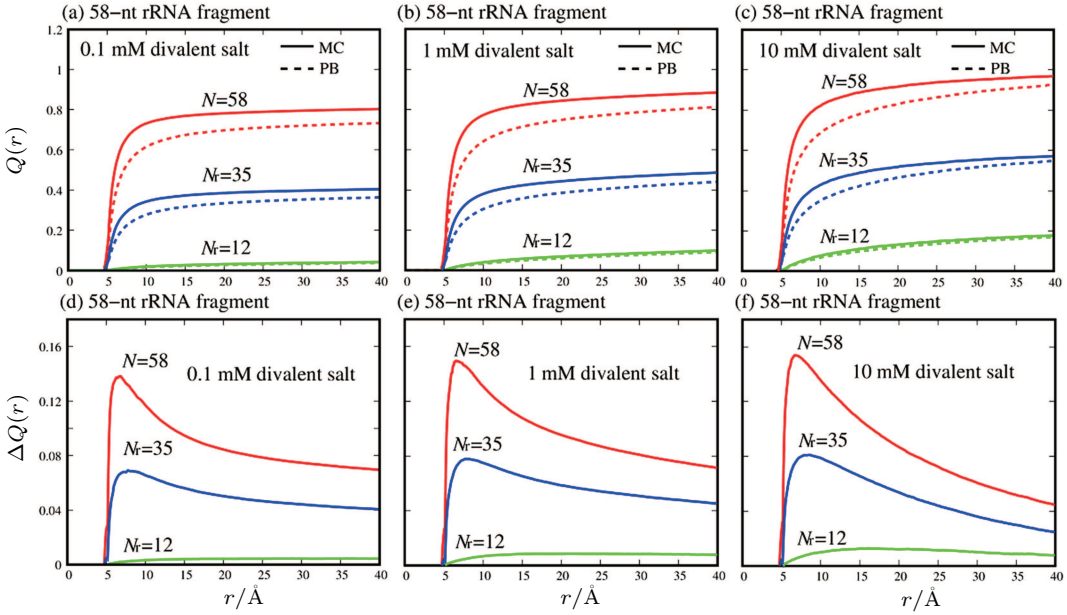


Fig. 7. (color online) (a)–(c) Net ion charge distribution $Q(r)$ obtained from PB theory (dashed lines) and MC simulations (solid lines) around a 58-nt rRNA fragment (PDB code: 1HC8) in divalent salt solutions. (d)–(f) Deviations ΔQ between distributions $Q(r)$ of PB theory and MC simulations. Distributions $Q(r)$ and deviations ΔQ for other RNAs and DNAs in divalent salt solutions can be found in Fig. S7 in the supplementary material.

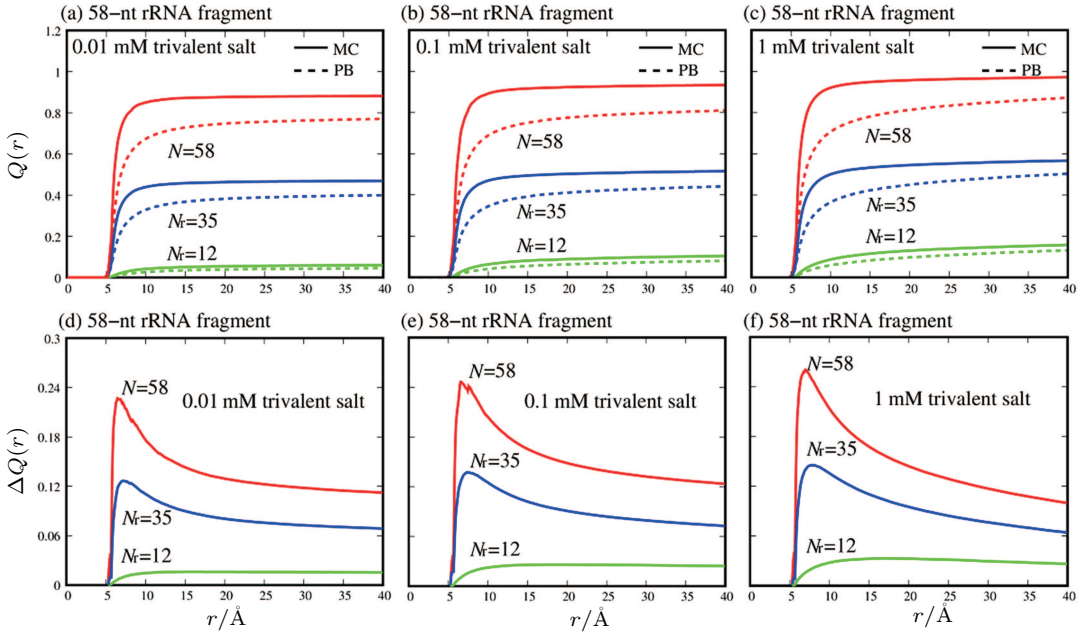


Fig. 8. (color online) (a)–(c) Net ion charge distribution $Q(r)$ obtained from PB theory (dashed lines) and MC simulations (solid lines) around a 58-nt rRNA fragment (PDB code: 1HC8) in trivalent salt solutions. (d)–(f) Deviations ΔQ between distributions $Q(r)$ of PB theory and MC simulations. Distributions $Q(r)$ and deviations ΔQ for other RNAs and DNAs in trivalent salt solutions can be found in Fig. S8 in the supplementary material.

3.3.2. Can an empirical formula work for real RNAs and DNAs?

In the above, we discussed obtaining a series of N_c for extensive RNAs and DNAs. Additionally, we can calculate N_c using empirical formula Eq. (4) by modeling an RNA or DNA as a spherical macroion with surface area $S = 4\pi R_g^2$, where R_g is the radius of gyration; see Fig. S10 in the supplementary material. Since the parameters of RNA-like macroions are from the statistical analyses on 80 RNAs, can the empirical formula for RNA-like macroions be said to describe the critical charge N_c $(= S\sigma_c)$ of real RNAs and DNAs?

To answer this, we first examined the values of N_c obtained from the calculations with MC simulations and PB theory, compared against those calculated from the empirical formula Eq. (4). As shown in Figs. 9 and S11 in the supplementary material, for the RNAs and DNAs of sphere-like shape, the values of N_c from PB theory and MC simulations are very close to those from Eq. (4); while for rod-like RNAs and DNAs, the values of N_c from PB theory and MC simulations deviate apparently from those obtained from Eq. (4). To account for the anisotropic shape for the RNAs and DNAs, we modeled an RNA (or DNA) as an ellipsoid macroion with the same rotation inertia tensor $\boldsymbol{I}$ as the RNA (or DNA). Practically, based on coordinates $\boldsymbol{x}$ of the structures of RNAs and DNAs in the Protein Data Bank, the rotation inertial tensor $\boldsymbol{I}$ respective to the center $\boldsymbol{x}_c$ can be calculated by[85]

$$I_{ij} = \sum_{\alpha} m_{\alpha} \left[\delta_{ij} \sum_{k} (x_{\alpha k} - x_{ck})^2 - (x_{\alpha i} - x_{ci})(x_{\alpha j} - x_{cj}) \right], \quad (5)$$

where the subscripts $i, j = 1, 2, 3$ are the three orthogonal directions. Here, the summation of α is over all atoms, and the values of m_{α} are set to 1 for phosphorus atoms and are set to 0 for other atoms, mainly in order to account for the electrostatic properties of RNAs and DNAs. From the rotation inertia tensor $\boldsymbol{I}$, the principal axis of inertia ω and the moment of inertial λ of an RNA or DNA can be calculated through

$$\boldsymbol{I} \cdot \boldsymbol{\omega} = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = \lambda \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}. \quad (6)$$

Finally, an ellipsoidal shell with the same $\boldsymbol{I}$, $\boldsymbol{\omega}$, and λ is employed to represent the complicated RNA or DNA, and the surface area S of the ellipsoid is used to calculate σ_c in Eq. (4), which is used to determined N_c $(= S\sigma_c)$ further. As shown in Figs. 9 and S10, S11 in the supplementary material, an RNA or DNA with arbitrary structure is represented better by an ellipsoidal macroion than by a spherical one. It is encouraging that the values of N_c from the MC simulations and PB theory are very close to those from Eq. (4) if the RNAs and DNAs are represented by ellipsoidal macroions. This finding is reasonable because spherical representation cannot well describe the shape of RNAs or DNAs with anisotropic structures.

From the above discussions, the empirical formula Eq. (4) can work well for describing N_c of real RNAs and DNAs with ellipsoidal representations, i.e., the empirical formula provides a good estimate of the critical extent of ion neutralization beyond which the multivalent ion distributions around real RNAs and DNAs can be described by PB theory. This quantification of the critical extent of ion neutralization would be helpful in some two-phase polyelectrolyte models for describing multivalent ion–nucleic acid interactions.[86–88] Notably, equation (4) was derived from the ion distributions around RNA-like spherical macroions, which might be somewhat different from the ion distributions around a cylindrical polyelectrolyte such as a long DNA helix.

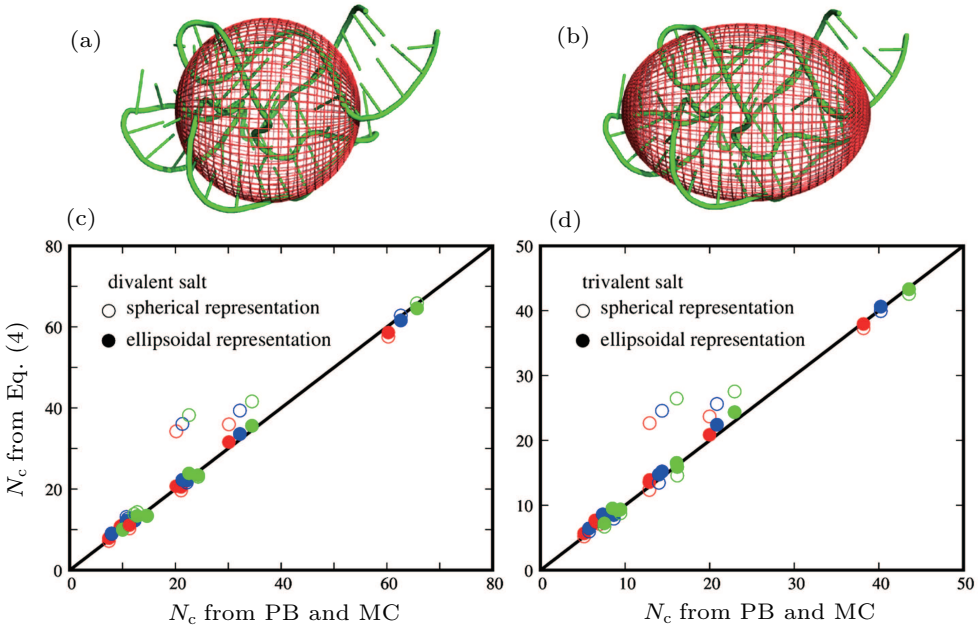


Fig. 9. (color online) Spherical representation (a) and ellipsoidal representation (b) for an RNA or DNA. Here the 58-nt rRNA fragment (PDB code: 1HC8) is shown as a paradigm for spherical and ellipsoidal representations. The spherical and ellipsoidal representations for other nucleic acids used in the work are shown in Fig. S10 in the supplementary material. (c)–(d) Critical charges N_c ($N_c = S\sigma_c$) obtained from PB theory and MC simulations against those from Eq. (4) in divalent (c) or trivalent (d) salt solution. The hollow and solid symbols stand for the data obtained from Eq. (4) with the spherical representation and ellipsoidal representations, respectively. The blue, green and red symbols represent the salt concentrations of 0.1 mM (green), 1 mM (blue), and 10 mM (red) for divalent salt solutions and 0.01 mM (green), 0.1 mM (blue), and 1 mM (red) for trivalent salt solutions, respectively.

4. Conclusions

In this work, we employed PB theory and MC simulations to calculate the ion distributions around RNA-like macroions and various RNAs and DNAs, and our calculations covered wide ranges of ion conditions including monovalent, divalent, and trivalent salts with different bulk concentrations. Through the extensive calculations and detailed comparisons between PB theory and MC simulations, we have obtained the following major conclusions.

1. For an RNA-like macroion in monovalent salt solution, no obvious deviation was found in the representation of net ion charge distribution between PB theory and MC simulations; and PB theory can approximately describe monovalent ion binding to RNA-like macroions.

2. For an RNA-like macroion in divalent or trivalent salt solutions, PB theory appears to underestimate the net ion charge distribution; and the deviation between PB theory and MC simulations becomes more pronounced for higher ion valence and higher ion concentration.

3. When an RNA-like macroion is neutralized to a "critical" extent by ions, the multivalent ion distributions can be approximately described by PB theory, and we obtained an empirical formula to describe this critical extent of ion neutralization for RNA-like macroions.

4. For real RNAs and DNAs in multivalent salt solutions, PB theory obviously underestimates the net ion charge distribution, and when an RNA or DNA gets ion neutralization of a "critical" extent, the multivalent ion distribution can be approximately described by PB theory. This critical ion neutralization extent can be approximately given by the empirical formula derived for RNA-like macroions, and could be useful for future two-phase polyelectrolyte models treating multivalent ion–nucleic acid interactions.

In this work, for simplicity, the radii of coions were kept the same as those of cations. Different radii of coions might bring slightly different ion distributions, which is beyond the scope of the present work and deserves to be discussed elsewhere. The present work has also employed some approximations and simplifications. First, the dielectric discontinuity at the boundary between solvent and RNA-like macroions/RNAs was ignored, which may slightly affect the ion binding at the boundary. Second, RNAs were modeled as spherical macroions, which would ignore some details of structures. Finally, the employment of the empirical formula from the analyses for RNA-like spherical macroions to real RNAs and DNAs also ignored structural details of RNAs and DNAs. Nevertheless, the present work provides extensive calculations and analyses on multivalent ion distributions around various RNA-like macroions and nucleic acids over a wide range of salt conditions, and an empirical formula was derived to measure the "critical" extent of ion neutralization beyond which the multivalent ion distributions around RNAs and DNAs can be described by PB theory. This empirical formula would be very useful for future two-phase polyelectrolyte models for

quantifying multivalent ion-nucleic acid interactions.[79–81]

Appendix A: RNA structures used in the work

The PDB codes of RNAs used to gather statistically normalized populations of surface charge density include 2CJK, 1ARJ, 2AP0, 2EEW, 1NBK, 1YG3, 2XO1, 3SUX, 4EN5, 3Q50, 1L3D, 1Q8N, 2XNW, 2KUW, 2K95, 4QK8, 1HC8, 2EET, 3GOT, 3G4M, 3GX6, 4QLN, 1KAJ, 2APO, 2I1V, 2MIY, 2XO0, 3DS7, 1RNK, 2F87, 2KTZ, 2L8H, 4ENB, 3VRS, 4LVW, 2EEU, 4QK9, 3RKF, 3FO6, 4L81, 4OQU, 2EES, 4LVY, 3D0U, 3GAO, 2G9C, 3F2T, 3IQR, 3LA5, 4PQV, 2K96, 2RPO, 2K4C, 3MJB, 2TPK, 2XNZ, 4OQU, 4OJI, 4ENC, 3Q51, 2XO0, 4LVX, 2EEV, 3GER, 4NYG, 3GX3, 2XNZ, 437D, 2LC8, 2H0Z, 2QWY, 3FO4, 3IWN, 4P5J, 5DUN, 2M8K, 4QLM, 3GX3, 3GES, 12-bp RNA.

Acknowledgements

We are grateful to Profs. Shi-Jie Chen (Univ. Missouri), Xiangyun Qiu (George Washington Univ.), Wenbing Zhang (Wuhan Univ.), and Dr. Yuan-Yan Wu (Yangzhou Univ.) for valuable discussions. Parts of the numerical calculation in this work were performed on the supercomputing system in the Supercomputing Center of Wuhan University.

References

[1] Woodson S A 2005 *Curr. Opin. Chem. Biol.* **9** 104
[2] Draper D E, Grilley D and Soto A M 2005 *Annu. Rev. Biophys. Biomol. Struct.* **34** 221
[3] Chen S J 2008 *Annu. Rev. Biophys.* **37** 197
[4] Wong G C and Pollack L 2010 *Annu. Rev. Phys. Chem.* **61** 171
[5] Lipfert J, Doniach S, Das R and Herschlag D 2014 *Annu. Rev. Biochem.* **83** 813
[6] Bao L, Zhang X, Jin L and Tan Z J 2016 *Chin. Phys. B* **25** 018703
[7] Chen H, Meisburger S P, Pabit S A, Sutton J L, Webb W W and Pollack L 2012 *Proc. Natl. Acad. Sci. USA* **109** 799
[8] Tan Z J and Chen S J 2006 *Biophys. J.* **90** 1175
[9] Tan Z J and Chen S J 2011 *Met Ions Life Sci.* **9** 101
[10] Shi Y Z, Wang F H, Wu Y Y and Tan Z J 2014 *J. Chem. Phys.* **141** 105102
[11] Wang J, Zhao Y J, Wang J and Xiao Y 2015 *Phys. Rev. E* **92** 062705
[12] Shi Y Z, Jin L, Wang F H, Zhu X L and Tan Z J 2015 *Biophys. J.* **109** 2654
[13] Heilman Miller S L, Pan J, Thirumalai D and Woodson S A 2001 *J. Mol. Biol.* **309** 57
[14] Sun L Z, Zhang D and Chen S J 2017 *Annu. Rev. Biophys.* **46** 227
[15] Wang J and Xiao Y 2016 *Phys. Rev. E* **94** 040401
[16] Drozdetski A V, Tolokh I S, Pollack L, Baker N and Onufriev A V 2016 *Phys. Rev. Lett.* **117** 028101
[17] Tan Z J and Chen S J 2011 *Biophys. J.* **101** 176
[18] Tan Z J and Chen S J 2006 *Nucleic Acids Res.* **34** 6629
[19] Tan Z J and Chen S J 2008 *Biophys. J.* **94** 3137
[20] He Z J and Chen S J 2012 *J. Chem. Theory Comput.* **8** 2095
[21] Jacobson D R and Saleh O A 2017 *Nucleic Acids Res.* **45** 1596
[22] Fenley M O, Olson W K, Tobias I and Manning G S 1994 *Biophys. Chem.* **50** 255
[23] Wu Y Y, Zhang Z L, Zhang J S, Zhu X L and Tan Z J 2015 *Nucleic Acids Res.* **43** 6156
[24] Zhang Z L, Wu Y Y, Xi K, Sang J P and Tan Z J 2017 *Biophys. J.* **113** 517
[25] Widom J and Baldwin R L 1980 *J. Mol. Biol.* **144** 431
[26] Bloomfield V A 1997 *Biopolymers* **44** 269
[27] Qiu X Y, Andresen K, Kwok L W, Lamb J S, Park H Y and Pollack L 2007 *Phys. Rev. Lett.* **99** 038104
[28] Li W, Wong W J, Lim C J, Ju H P, Li M, Yan J and Wang P Y 2015 *Chin. Phys. B* **24** 128704
[29] Wang F H, Wu Y Y and Tan Z J 2013 *Biopolymers* **99** 370
[30] Tan Z J and Chen S J 2010 *Biophys. J.* **99** 1565
[31] Andresen K, Das R, Park H Y, Smith H, Kwok L W, Lamb J S, Kirkland E J, Herschlag D, Finkelstein K D and Pollack L 2004 *Phys. Rev. Lett.* **93** 248103
[32] Zhuang X, Bartley L E, Babcock H P, Russell R, Ha T, Herschlag D and Chu S 2000 *Science* **288** 2048
[33] Russell R, Zhuang X, Babcock H P, Millett I S, Doniach S, Chu S and Herschlag D 2002 *Proc. Nalt. Acad. Sci. USA* **99** 155
[34] Manning G S 1978 *Q. Rev. Biophys.* **11** 179
[35] Manning G S 2007 *J. Phys. Chem. B* **111** 8554
[36] Schurr J M and Fujimoto B S 2002 *Biophys. Chem.* **101** 425
[37] Gilson M K, Sharp K A and Honig B H 1988 *J. Comput. Chem.* **9** 327
[38] Sharp K A and Honig B 1990 *J. Phys. Chem.* **94** 7684
[39] Liu H Y and Zou X Q 2006 *J. Phys. Chem. B* **110** 9304
[40] Chen D, Chen Z, Chen C, Geng W and Wei G W 2011 *J. Comput. Chem.* **32** 756
[41] Grant J A, Pickup B T and Nicholls A 2001 *J. Comput. Chem.* **22** 608
[42] Ran S Y and Jia J L 2015 *Chin. Phys. B* **24** 128702
[43] Chu V B, Bai Y, Lipfert J, Herschlag D and Doniach S 2007 *Biophys. J.* **93** 3202
[44] López García J J, Horno J and Grosse C 2011 *Langmuir* **27** 13970
[45] Das T, Bratko D, Bhuiyan L B and Outhwaite C W 1997 *J. Chem. Phys.* **107** 9197
[46] Lu B Z, Chen W Z, Wang C X and Xu X J 2002 *Proteins* **48** 497
[47] Lu B Z, Zhou Y C, Holst M J and McCammon J A 2008 *Commun. Comput. Phys.* **3** 973
[48] Borukhov I, Andelman D and Orland H 1997 *Phys. Rev. Lett.* **79** 435
[49] Tan Z J and Chen S J 2005 *J. Chem. Phys.* **122** 044903
[50] Bacquet R and Rossky P J 1984 *J. Chem. Phys.* **88** 2660
[51] Draper D E 1995 *Annu. Rev. Biochem.* **64** 593
[52] Wette P, Schöpe H J and Palberg T 2002 *J. Chem. Phys.* **116** 10981
[53] Boschitsch A H, Fenley M O and Zhou H X 2002 *J. Phys. Chem. B* **106** 2741
[54] Baker N A 2004 *Method. Enzymol.* **383** 94
[55] Wagoner J and Baker N A 2004 *J. Comput. Chem.* **25** 1623
[56] Wei G W, Zheng Q, Chen Z and Xia K L 2012 *SIAM Rev.* **54** 699
[57] Lyklema J 2015 *J. Colloid Interf. Sci.* **446** 308
[58] Lozada Cassou M 1983 *J. Phys. Chem.* **87** 3729
[59] Vlachy V 1999 *Annu. Rev. Phys. Chem.* **50** 145
[60] Bai Y, Greenfeld M, Travers K J, Chu V B, Lipfert J, Doniach S and Herschlag D 2007 *J. Am. Chem. Soc.* **129** 14981
[61] Qiu X Y, Parsegian V A and Rau D C 2010 *Proc. Natl. Acad. Sci. USA* **107** 21482
[62] Trachman R J and Draper D E 2017 *Nucleic Acids Res.* **45** 4733
[63] Giambaşu G M, Luchko T, Herschlag D, York D M and Case D A 2014 *Biophys. J.* **106** 883
[64] Misra V K and Draper D E 2002 *J. Mol. Biol.* **317** 507
[65] Sushko M L, Thomas D G, Pabit S A, Pollack L, Onufriev A V and Baker N A 2016 *Biophys. J.* **110** 315
[66] Baker N A, Sept D, Joseph S, Holst M J and McCammon J A 2001 *Proc. Natl. Acad. Sci. USA* **98** 10037
[67] Bao L, Zhang X, Shi Y Z, Wu Y Y and Tan Z J 2017 *Biophys. J.* **112** 1094
[68] Ermak D L and McCammon J A 1978 *J. Chem. Phys.* **69** 1352
[69] Deserno M, Holm C and May S 2000 *Macromolecules* **33** 199
[70] Carnal F and Stoll S 2012 *J. Phys. Chem. A* **116** 6600
[71] Zhang X, Zhang J S, Shi Y Z, Zhu X L and Tan Z J 2016 *Sci. Rep.* **6** 23434
[72] Qi W P, Lei X L and Fang H P 2010 *Chem. Phys. Chem.* **11** 2146

[73] Qi W P, Song B, Lei X L, Wang C L and Fang H P 2011 *Biochemistry* **50** 9628
[74] Zhang Y J, Zhang J and Wang W 2011 *J. Am. Chem. Soc.* **133** 6882
[75] Su X H, Lei Q L and Ren C L 2015 *Chin. Phys. B* **24** 113601
[76] Wei J Y, Ran S Y, He L L, Wang X H and Zhang L X 2015 *Chin. Phys. B* **24** 118701
[77] Marcus Y 1985 *Ion Solvation* (New York: Wiley)
[78] Wu Y Y, Wang F H and Tan Z J 2013 *Phys. Lett. A* **377** 1911
[79] Zhang X, Bao L, Wu Y Y, Zhu X L and Tan Z J 2017 *J. Chem. Phys.* **147** 054901
[80] Gutin A M and Shakhnovich E I 1994 *Phys. Rev. E* **50** R3322
[81] Salerno K M, Frischknecht A L and Stevens M J 2016 *J. Phys. Chem. B* **120** 5927
[82] dos Santos A P, Diehl A and Levin Y 2010 *J. Chem. Phys.* **132** 104105
[83] Murthy C S, Bacquet R J and Rossky P J 1985 *J. Phys. Chem.* **89** 701
[84] Jayaram B, Swaminathan S, Beveridge D L, Sharp K and Honig B 1990 *Macromolecules* **23** 3156
[85] Hyeon C, Dima R I and Thirumalai D 2006 *J. Chem. Phys.* **125** 194905
[86] Sun L Z and Chen S J 2016 *J. Chem. Theory Comput.* **12** 3370
[87] Zhu Y H and Chen S J 2014 *J. Chem. Phys.* **141** 055101
[88] Henke P S and Mak C H 2016 *J. Chem. Phys.* **144** 105104