

Polyethylenimine/DNA Ratio Strongly Affects their Nanoparticle Formation: A Large-Scale Coarse-Grained Molecular Dynamics Study

*Subhamoy Mahajan and Tian Tang**

Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada

* Corresponding author, Email: tian.tang@ualberta.ca

Abstract

Polyethylenimine (PEI)-DNA nanoparticles (NPs) has shown a lot of potential in gene delivery. N/P ratio, the ratio between total number of amines in PEIs and total number phosphates in DNAs, is an essential factor determining the efficacy of delivery. In this work, the aggregation of PEIs and DNAs under different N/P ratios is studied using large-scale coarse-grained simulations under the Martini framework. At very low N/P ratio, the aggregation of DNAs is limited, and as the N/P ratio increases the nanoparticles change from a loose linear structure to a compact branched structure. Such a transition in the mode of aggregation is caused by the different alignments of PEIs with DNA backbones prior to aggregation, which dictates their ability to serve as polycation bridges. Except for very large nanoparticles at high N/P ratios, the charge of a nanoparticle is proportional to the number of DNAs in it. Their ratio allows for the definition of an intrinsic property called specific repulsion, which controls the characteristics of the steady-state size distribution of NPs: unimodal for strong specific repulsion, bimodal for moderate specific repulsion, and more or less uniform for weak specific repulsion. Understanding the mechanism behind DNA-PEI NP formation helped us propose a two-step process to generate NPs that are more compact and closer to being spherical.

1. Introduction

Over the last few decades, gene delivery has been gaining attention as a means to treat genetic defects.^{1,2} In this process, genetic materials such as DNAs are “delivered” to the cell to induce a therapeutic response against genetic defects, which can be hereditary or arise from harsh environmental conditions. One challenge in DNA delivery is to find an effective delivery system because DNAs by themselves are ineffective in cell internalization and are prone to degradation.³ Gene delivery was first achieved using viral vectors because viruses are capable of transferring their genetic materials into a cell to reproduce.⁴ Although viral vectors are modified so that they cannot replicate, they may still result in immune response in the host and therefore have significant safety concerns. As a result, non-viral delivery has been investigated due to the low cost, ease of synthesis, and relatively lower safety concerns. The efficacy of non-viral delivery, however, is not yet at par with viral delivery.

There have been numerous studies on non-viral vectors and among them, polyethylenimine (PEI) has shown a lot of potential.^{3,5-9} PEIs are comprised of repeating amine groups, which can form a wide variety of structures from completely linear to highly dendritic. They are highly soluble in water, and due to the presence of many nitrogen atoms, they act as a buffer over a large range of pH.^{1,10} At physiological pH, PEIs are weak bases and positively charged, with ~21% of their nitrogen being protonated.¹¹ Meanwhile, DNAs are strong acids and negatively charged due to the presence of phosphates on their backbone. Adding PEIs to DNAs results in charge neutralization and condensation of DNAs, and aggregation of these condensed DNAs to form nanoparticles (NPs).¹² Properties of the NPs such as hydrodynamic radius, surface charge and shape depend on the structure and molecular weight of PEI being used, molecular modifications made to PEIs such as the grafting of aliphatic chains, as well as the N/P ratio (ratio of the total

number of amines in PEIs to that of phosphates in DNAs). These NP properties play a significant role in the various steps of a gene delivery process.

Take for example the step of NP internalization, i.e., cellular uptake. A successful cellular uptake starts with the adsorption of the NP onto the plasma membrane. The overall negatively charged nature of the plasma membrane promotes the adsorption of NPs with positive surface charge.^{13,14} Adsorption of negatively charged NPs has also been observed due to the presence of cationic sites on the plasma membrane, but the adsorption is limited.¹⁵ Neutral NPs are difficult to be adsorbed¹⁶, but in intravenous delivery they have the advantage of avoiding non-specific interactions with other biomolecules such as proteins before reaching the target cells.^{17,18} Upon adsorption, NPs are internalized and it has been shown that PEI-based gene delivery primarily uses clathrin-mediated endocytosis.¹⁹ This endocytic pathway encapsulates NPs in endosomes with a diameter ranging from 100-150 *nm*^{19,20}, thereby placing a limit on the size of the NPs that can be internalized. It is therefore important to be able to control the size and charge of the NPs, in order to improve their cellular uptake.

Extensive studies have been carried out to modulate the size (in terms of hydrodynamic radius) and charge (in terms of zeta potential) of DNA-PEI NPs by changing the N/P ratio.^{21,22} Zeta potential is calculated to quantify the surface charge of the NP, which is negative for low N/P ratio and positive for high N/P ratio. It is well accepted that for intermediate N/P ratios where the zeta potential is close to zero, the hydrodynamic radius of the NPs is several folds larger than other N/P ratios.²¹ The hydrodynamic radius is not a single value but follows certain distributions.^{10,12,23} Identifying these distributions can be important because a successful gene delivery might be enabled by a range of hydrodynamic radius and not the mean value.^{24,25} Time variation of the distributions has not been addressed, which can be important in understanding the NP formation

process and how we may control it. The NPs are known to form various shapes such as spheres, toroids, rods, and folded loops,^{10,12,23} which might be crucial for cellular uptake^{24,26-28}, but their effect on gene delivery is not well explored.

In recent years, many molecular dynamics (MD) simulations have been performed to generate atomistic images on the aggregation of DNAs and PEIs. Dynamics of DNA-PEI binding and their binding free energy were studied for PEIs with different molecular weight, degrees of branching and protonation ratios using all-atom (AA) MD simulations.²⁹⁻³³ Binding of large polycations to single DNA (or RNA) was also studied with coarse-grained (CG) MD simulations.³⁴⁻³⁶ Aggregation of multiple DNAs was studied using AA MD; these simulations were small in size and the largest system contained only four short DNAs each having 12 base pairs.^{37,38} At a larger length scale, CG simulations were used to study the aggregation of DNAs, modeled as linear polyelectrolytes, by mono- and multi-valent ions.³⁹⁻⁴⁸ Similar CG simulations are not available for DNA aggregation mediated by polycations such as PEI. Due to their size and structure, polymers can aggregate DNAs in ways different from mono- and multi-valent ions. Moreover, modeling DNAs as linear polyelectrolytes³⁹⁻⁴⁸ is highly approximate, which lacks the double helical structure and major/minor grooves. These structural features are important as PEIs have been observed to bind with DNA along major and minor grooves.²⁹ Such details cannot be captured by mean-field theories such as Poisson-Boltzmann. Together with the size limitation in AA simulations, there is a need to study PEI-induced DNA aggregation and NP formation using large-scale CG simulations with models comprised of accurate molecular structures.

In this work, large-scale CG-MD simulations are performed to investigate the aggregation of a large number of DNA and PEI molecules. Since N/P ratio plays a significant role in determining the properties of NPs, we focus on its influence on aggregation dynamics, and NP size, shape and

charge. This is done by simulating 5 systems with different PEI/DNA number ratios (hence different N/P ratios) and at two different salt concentrations. Similar to the experiments of Ogris et al.⁴⁹, the systems are prepared by keeping the number of DNAs fixed while changing the number of PEIs. The rest of the paper is organized as follows. Preparation of the simulated systems and details of the simulations are provided in **Section 2**. Results are presented in **Section 3**. In **Section 3.1-3.3** we focus on systems with only neutralizing salt ions, and address NP shape (**Section 3.1**), charge and size (**Section 3.2**) and the rate of NP growth (**Section 3.3**). Results in the presence of 150 mM salt are presented and compared in **Section 3.4**. Further discussion of the results, their comparison with experiments and practical implications are given in **Section 4**.

2. Methods

The DNA used in the simulations is a Drew-Dickerson dodecamer $d(\text{CGCGAATTCGCG})_2$ in its canonical B form with a charge (Q_{DNA}) of -22 . At the CG level, it is modeled using the stiff elastic network in the Martini force field developed by Uusitalo et al.⁵⁰ The PEI simulated is a 586 Da branched PEI with its structure and protonation sites shown in **Figure 1**.²⁹ Three amines are protonated ($Q_{PEI} = 3$) according to the $\sim 21\%$ protonation ratio at physiological pH.^{11,29} The CG PEI model was developed in our previous work within the Martini framework and was shown to be compatible with Martini DNA.⁵¹ Specifically, each CG bead contains C-C-N atoms with the accompanying hydrogens, and the mapping scheme is displayed as blue enclosures in **Figure 1**. The PEI beads are classified as ‘*t*’, ‘*s*’, ‘*p*’ and ‘*pq*’, which respectively denote beads containing tertiary, secondary, primary and protonated primary nitrogens.⁵¹

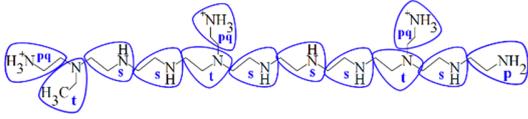


Figure 1: All-atom (black) and coarse-grained (blue) structures of PEI used in this study. The blue enclosures represent the CG mapping scheme. The CG beads were classified based on the type of nitrogen within the bead, i.e. tertiary (*t*), secondary (*s*), primary (*p*), and protonated primary (*pq*).⁵¹

Five systems are simulated each containing the same number of DNAs ($N_{DNA} = 27$) but a different number of PEIs (N_{PEI}), corresponding to the PEI/DNA number ratio (α) of 2, 4, 6, 8 and 10 (**Table 1**). In each system, the principal axes of the 27 DNAs were aligned and arranged in a $3 \times 3 \times 3$ array, which was then placed in a cubic simulation box of side length 25 nm . The distance between the centers of mass of two closest DNAs was 6 nm . The PEIs were randomly placed within the box, and the solutes were solvated with polarizable Martini water⁵². An appropriate amount of Martini K^+ or Cl^- ions⁵³ were introduced to neutralize each system (see **Table 1**). Systems with $\alpha = 2-6$ contained excess DNAs and required cationic (K^+) ions to neutralize the systems, whereas the systems with $\alpha = 8-10$ had excess PEIs and required anionic (Cl^-) ions. For each system in **Table 1**, an additional system was simulated by adding 150 mM KCl and correspondingly removing an appropriate amount of water molecules. Comparison of the two systems at the same α allows us to address the effect of salt.

All simulations were performed using the GROMACS 5 package.⁵⁴ Each system was first energy minimized using steepest-descent, followed by a constrained MD in NPT ensemble for 1 ns , and an unconstrained MD in NPT ensemble for $4 \mu\text{s}$. Since the potential energy surface of the Martini force field is smooth⁵⁵, NVT equilibration after energy minimization was not required. In the constrained MD simulations, all the bonds of PEI and the backbone bonds of DNA were constrained using LINCS⁵⁶. The temperature was maintained at 300 K using velocity rescaling thermostat⁵⁷ with a time constant of 0.1 ps . The pressure was maintained at 1 bar using Berendsen thermostat⁵⁸ with a time constant of 3 ps and compressibility of $3 \times 10^{-4} \text{ bar}^{-1}$. Unconstrained

NPT simulations were conducted at the same temperature and pressure, using the same parameters for the thermostat. However, the pressure was maintained using Parrinello-Rahman barostat⁵⁹ with a time constant of 5 ps and compressibility of $4.5 \times 10^{-4} \text{ bar}^{-1}$. For all simulations, a neighbour-list was maintained within a radius of 1.1 nm using the Verlet scheme⁶⁰, and was updated every 40 time steps. Electrostatic interactions were calculated using a relative dielectric constant of 2.5 and were cut off at 1.1 nm using the reaction field scheme⁶¹. Similarly, van der Waals interactions were cut-off at 1.1 nm using the potential-shift-Verlet scheme. Periodic boundary condition was applied in all directions, and leapfrog integrator was used with a time step of 5 fs. Unless otherwise specified, all the simulation time reported below are scaled by a factor of 4, whereas the simulation parameters (time constants, time step) mentioned above are unscaled.⁵³

Table 1: Summary of simulated systems. N_{DNA} , N_{PEI} , N_{ion} are respectively the total numbers of DNAs, PEIs and ions. The N/P ratio is calculated using $13\alpha/22$, where 13 is the number of nitrogen atoms in each PEI and 22 is the number of phosphorus atoms in each DNA. Corresponding systems with 150 mM KCl had the same values for N_{DNA} and N_{PEI} , but an additional 150mM KCl in each system.

α	N_{DNA}	N_{PEI}	N_{ion}	N/P ratio
2	27	54	432 (K^+)	1.18
4	27	108	270 (K^+)	2.36
6	27	162	108 (K^+)	3.55
8	27	216	54 (Cl^-)	4.73
10	27	270	216 (Cl^-)	5.91

3. Results

3.1. Mode of aggregation and roles of PEI

At the end of the simulation, several DNA-PEI NPs are formed in each system. **Figure 2** shows the largest NP for different α . The size of the NP (s_{NP}) is quantified by the number of DNAs in the NP. Upon visual inspection, in **Figure 2a** where $\alpha = 2$, the aggregation is very limited. In **Figure 2b, c** where $\alpha = 4$ and 6, the DNAs in the NP tend to connect to one another from end-to-

end, leading to a loose linear structure resembling a string. On the contrary, in **Figure 2d, e** where $\alpha = 8$ and 10, the DNAs in the NP have established some side-by-side connections forming more compact branched structures, representing a different mode of aggregation. By calculating the relative shape anisotropy (see Supporting Information (SI) Section S1), the branched structures are closer to being spherical compared with the linear structures. Clearly, the mode of aggregation is dependent on the number of PEIs present in the system and how they bind to the DNA. To facilitate the discussion, we identify different roles of PEI in the dynamics of DNA aggregation. A PEI is defined to be bound to a DNA if the minimum distance between any one bead in the PEI and any one bead in the DNA is less than 0.53 nm , which is the equilibrium distance for the Lennard-Jones interaction potential between DNA and PEI beads (see SI Section S2 for further discussion on DNA-PEI binding). A PEI is categorized as *unbound* if it is not bound to any DNA, *peripheral* if it is bound to only one DNA, and *bridging* if it is simultaneously bound to two or more DNAs to form a polycation bridge.

Figure 3a-c shows the numbers of unbound (N_u), peripheral (N_p) and bridging (N_b) PEIs as functions of simulation time, for different α 's. For all systems, PEIs starts as unbound (**Figure 3a**), but the number quickly reduces as the PEIs become peripheral (**Figure 3b**) or bridging (**Figure 3c**). At any time, N_b is significantly lower than N_p , indicating that the majority of unbound PEIs convert to peripheral PEIs. Furthermore, by examining the number of conversions among free, peripheral and bridging PEIs, we find that direct conversion from unbound to bridging PEI is rare. That is, bridging PEI is formed dominantly from peripheral PEI, and peripheral PEI from unbound PEI (see SI Section S3 for more details).

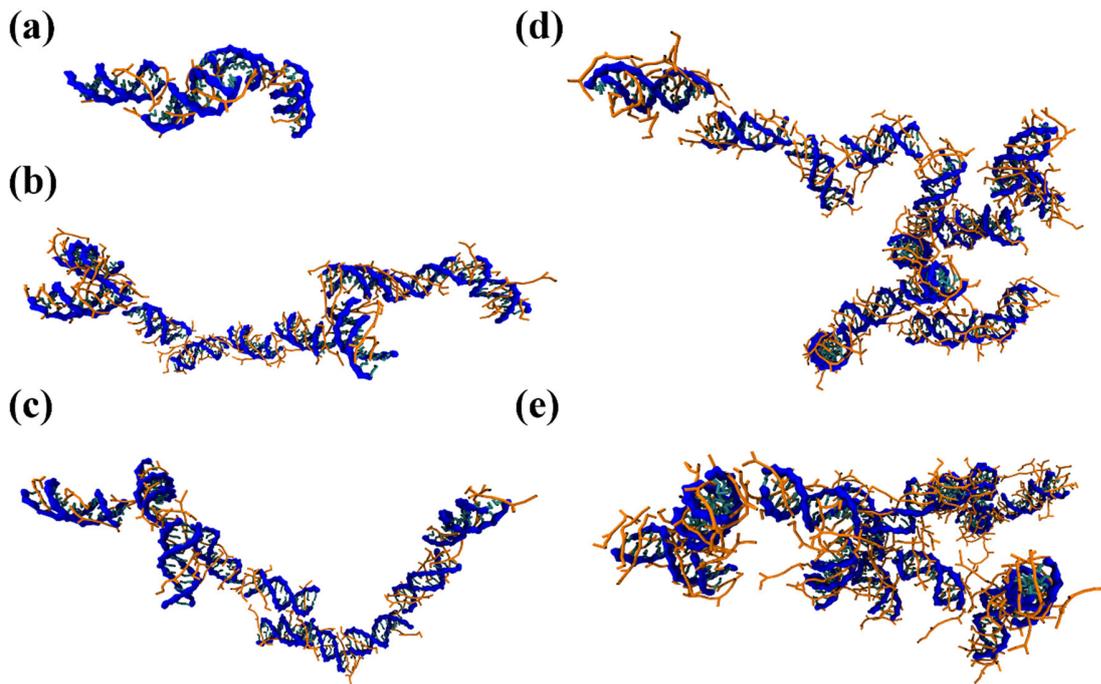


Figure 2: Largest NP formed at the end of 4 μ s simulation. (a) $\alpha = 2$, $s_{NP} = 4$; (b) $\alpha = 4$, $s_{NP} = 10$; (c) $\alpha = 6$, $s_{NP} = 10$; (d) $\alpha = 8$, $s_{NP} = 15$ and (e) $\alpha = 10$, $s_{NP} = 14$. DNAs are represented with blue backbone and cyan base pairs, and PEIs are represented in orange.

In **Figure 3a**, unbound PEIs shows an exponential-like decay, where the rate of decay (negative of the slope) decreases with time for all α . For systems with excess DNA ($\alpha = 2-6$) N_u decreases to zero before 500 ns, whereas for systems with excess PEI ($\alpha = 8-10$) N_u remains non-zero even at the end of the simulations. The non-zero N_u suggests that in excess of PEIs there exists a maximum number of PEIs that can bind to a DNA. To estimate this maximum, we calculate the average number of PEIs bound to a DNA, $(N_p + N_b)/N_{DNA}$, over the last 2 μ s. The average is 7.89 for $\alpha = 8$ and 9.48 for $\alpha = 10$, which implies that the maximum number of PEIs that binds to a DNA increases with α . Given that the number of DNA is the same for all systems, the results suggest that the fashion in which the PEIs bind to the DNA may vary with α .

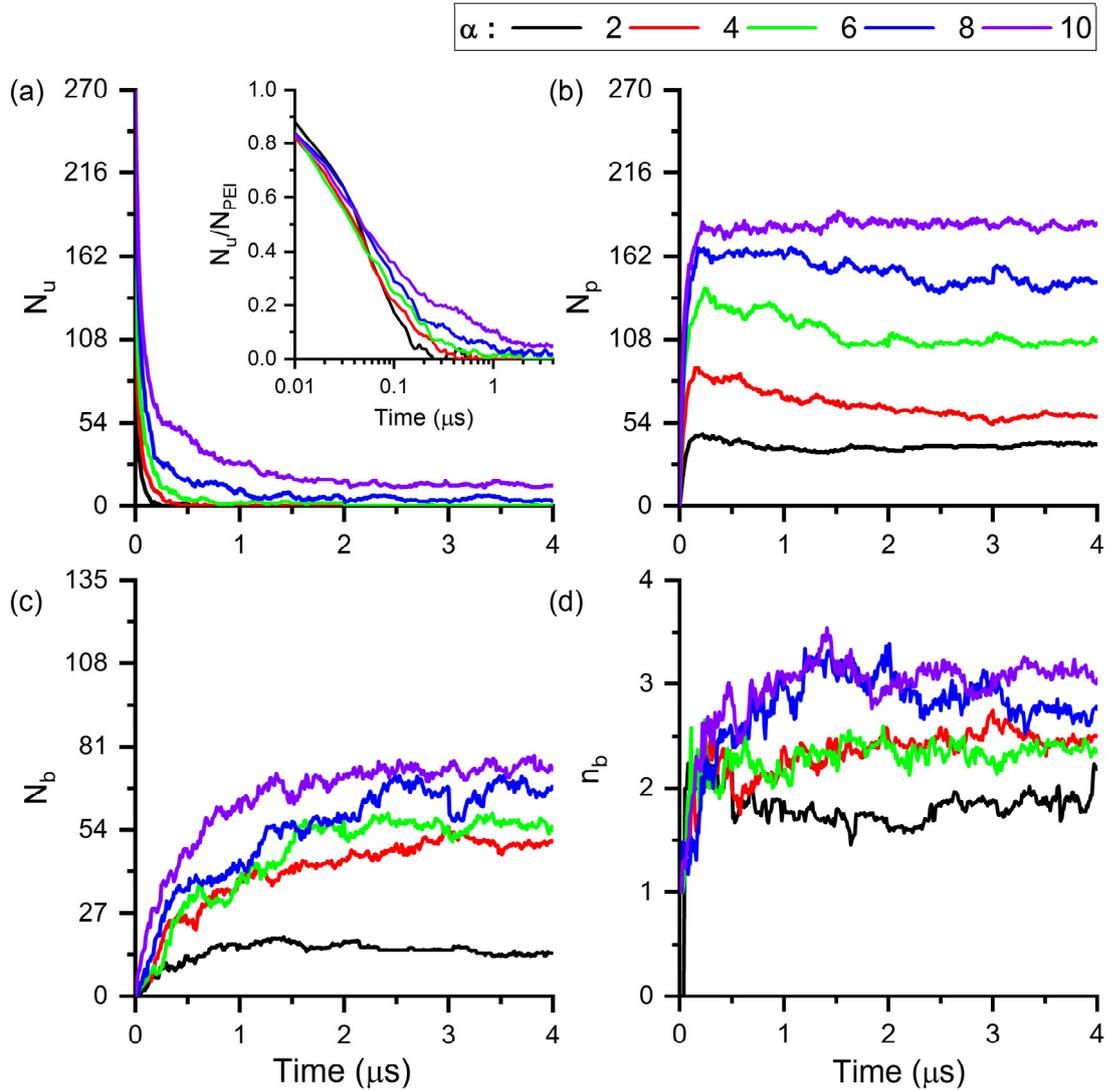


Figure 3: Number of PEIs in different roles, plotted against simulation time: (a) unbound (N_u) (b) peripheral (N_p) and (c) bridging (N_b). (d) Average number of bridging PEI between a pair of bridged DNAs (n_b).

In **Figure 3b**, N_p increases and reaches its maximum value at ~ 250 ns for all α . However, its behavior beyond 250 ns is different. For systems with excess DNA ($\alpha = 2-6$), N_p slowly decreases, which is due to the conversion of peripheral PEIs to bridging ones. At the same time no new peripheral PEI is formed due to the exhaustion of unbound PEIs. For systems with excess PEI ($\alpha = 8-10$), N_p remains almost constant because there is an abundance of unbound PEIs. As peripheral PEIs convert to bridging ones, it is likely that some of the DNA phosphates previously “covered”

by the peripheral PEIs are now exposed, allowing unbound PEIs to bind. Therefore, N_p remains constant due to a dynamic balance between the conversions from unbound to peripheral PEIs and from peripheral to bridging PEIs.

In **Figure 3c**, bridging PEIs increases monotonically with time, and the final equilibrium value of N_b is larger for higher α . A larger number of bridging PEIs could imply a greater number of DNAs being aggregated together; it could also imply a larger number of bridges being formed between a pair of aggregated DNAs. To investigate this, we plot the average number of bridging PEIs between a pair of bridged DNAs (n_b) in **Figure 3d**. Details on the calculation of n_b can be found in SI Section S4. The 5 n_b curves appear to form 3 groups: $\alpha = 2$ corresponds to the lowest value of n_b , $\alpha = 4$ and 6 correspond to an intermediate n_b , and the highest value of n_b occurs for $\alpha = 8$ and 10. Interestingly these 3 groups correlate with the observations made earlier on the mode of aggregation, where the aggregation was limited for $\alpha = 2$, end-to-end for $\alpha = 4$ and 6 and side-by-side for $\alpha = 8$ and 10. More discussions are available in SI Section S4.

Having understood the different roles PEIs play and the conversion between these roles during the aggregation process, we now discuss why different modes of aggregations are found under different PEI/DNA ratio. The aggregation starts by PEIs binding to the DNAs (conversion from unbound to peripheral PEIs), followed by bridging of different DNAs (conversion from peripheral to bridging PEIs). As discussed earlier, different PEI/DNA ratios can lead to different fashions in which PEIs bind to a DNA. To explore this, we first look at single PEI binding to a DNA and report two distinct ways in which a PEI approaches a DNA. **Figure 4a** shows snapshots, taken at different time as indicated, of a single PEI binding to a DNA at one of its terminals. Since the PEI in this study has three protonated beads, it can make three primary points of contact with the DNA. As shown in **Figure 4a**, the PEI first makes a one-point contact (24 ns), and then slowly moves

along the backbone of the DNA to form a two-point contact (29 and 30 *ns*). Some PEIs further align themselves with the backbone or bend over the terminal (the PEI shown at 92 *ns*) to form a three-point contact, although two-point contact is dominant for PEIs binding at the DNA terminals.

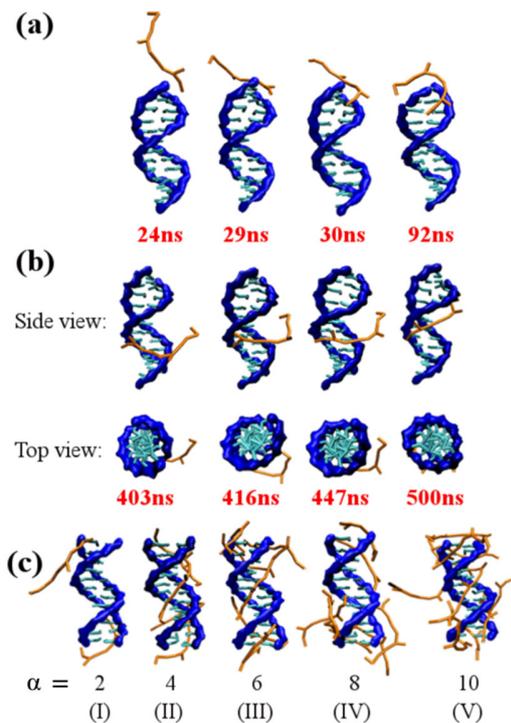


Figure 4: Snapshots of a single PEI binding to a DNA (a) from the terminal and (b) at the center. (c) Snapshots of PEIs bound to a DNA for different PEI/DNA ratios at 500 ns. DNAs are represented with blue backbone and cyan base pairs, and PEIs are represented in orange.

Figure 4b shows snapshots, also taken at different time as indicated, of a single PEI binding to a DNA at its center. Similarly, the PEI first makes a one-point contact (403 *ns*), gradually moves closer to the DNA backbone (416 and 447 *ns*) forming a two-point contact and can eventually form a three-point contact completely aligned with the backbone (500 *ns*). It should be noted that the snapshots in **Figure 4a, b** are selected from our simulation trajectory to clearly show the binding process; they are not an indication that binding at a DNA's center occurs later than binding at one of the terminals.

Figure 4c shows one DNA from each system ($\alpha = 2-10$) at 500 ns which has had PEIs bound to it but has not aggregated with other DNAs yet. At low α values (**Figure 4c-I**), the number of bound PEIs are very few, resulting in limited aggregation as seen in **Figure 2a**. At moderate α values (**Figure 4c-II, III**), there are several peripheral PEIs and they bind to the DNA at different locations. PEIs at the terminals mostly form two-point contact leaving one additional binding site available, whereas PEIs at the center mostly form three-point contact making it difficult to interact with another DNA to form a bridge. If this DNA were to aggregate with another DNA, it would be facilitated by the PEIs at the terminals but not those at the center. This forces the DNAs to undergo end-to-end aggregation forming loose linear NP as seen in **Figure 2b, c**. For systems with excess PEI, i.e., large α (**Figure 4c-IV, V**), the competition among the PEIs to bind with the DNA prevents PEIs binding at the center from complete alignment with the backbone. This misalignment reduces three-point contacts made by center-binding PEIs, allowing them to bind with other DNAs through a lateral side-by-side aggregation, and forming more compact branched NPs as seen in **Figure 2d-e**.

The PEI/DNA ratio α represents how crowded the PEIs are around the DNAs. Since each DNA has 22 deprotonated phosphates and each PEI has 3 protonated amines, on average each DNA can be saturated with 7.33 PEIs. The average number of PEIs bound to a DNA, $(N_p + N_b)/N_{DNA}$, in the last 2 μ s is 2, 4, 6, 7.89 and 9.48 for $\alpha = 2, 4, 6, 8$ and 10 respectively. The first three numbers (< 7.33) represent under-saturation of the DNAs by the PEIs, while the latter two (> 7.33) represent over-saturation. The data shows not only increased PEI crowding with the increase in α , but also the extra accommodation of PEIs when α is large. From **Figure 4a, b** it is clear that bound PEIs takes time (~ 100 ns) to conform to the DNA structure, while unbound PEIs continue to bind with the DNA. In the inset of **Figure 3a**, we plot the ratio between the numbers of unbound PEI and

the total number of PEI (N_u / N_{PEI}) as a function of time using a semi-log plot. For the first 100 ns N_u / N_{PEI} is similar for all α , implying that the initial rate of decay of N_u is proportional to N_{PEI} . In other words, the rate of PEI binding to DNA increases with α and more PEIs can bind to a DNA within this time window for higher α . DNAs for $\alpha = 2, 4, 6$ are under-saturated and the PEIs still have time to adjust their configurations and become aligned to the DNAs at the center. Whereas the extra accommodation of PEIs made by the DNAs at $\alpha = 8$ and 10 are traded with misalignment of center-binding PEIs. The misalignment is expected to be larger for $\alpha = 10$, as on average each DNA in $\alpha = 10$ accommodates 1.6 more PEIs than in $\alpha = 8$. Overall, our results above demonstrate that the PEI/DNA ratio influences the initial rate of PEI binding to individual DNAs and subsequent alignment of PEI to the DNA backbones. The alignment, in turn, affects the mode of aggregation of multiple DNAs, leading to different shapes of the NPs. It should be noted that DNAs can also be over-saturated if a deprotonated phosphate bead simultaneously interacts with multiple protonated amine beads. This is in fact observed in our simulations. While the number of phosphate beads interacting with three amine beads is negligible, a fraction of phosphate beads can interact with two amine beads at the same time, and the number increases with α (see SI Section S5 for further discussions).

3.2. NP charge and size

Figure 5 shows the size distribution of the NPs (black) and the average charge of the NPs at a given size (red). Data used to generate the plots are collected from the last 0.5 μ s of the simulations, which is regarded as steady-state (see SI Section S6 for details). The horizontal axis of each subplot is the size of the NP (s_{NP}) quantified by the number of DNAs in the NP. The time-averaged number of NPs having size s_{NP} are plotted on the left axis representing the size distribution. Meanwhile, the sum of charges from all NPs with size s_{NP} is calculated. Dividing this

total charge by the number of NPs at the same s_{NP} gives the average NP charge (Q_{NP}) and it is plotted on the right axis. If for a given s_{NP} no NPs are found, Q_{NP} is set to zero (shown as the red dashed line in **Figure 5**). When $\alpha = 4$ and for $s_{NP} = 15$, Q_{NP} is non-zero while the average number of NPs appears to be zero, which seems contradictory. We point out that the average number of NPs is in fact not zero but a very small number (0.025), which represents unstable NPs formed for a short period of time during the last $0.5 \mu s$. Similar unstable NPs are observed for $\alpha = 2$ at $s_{NP} = 5$ and $\alpha = 6$ at $s_{NP} = 13$ and 17 . No unstable NPs are observed for systems with excess PEIs.

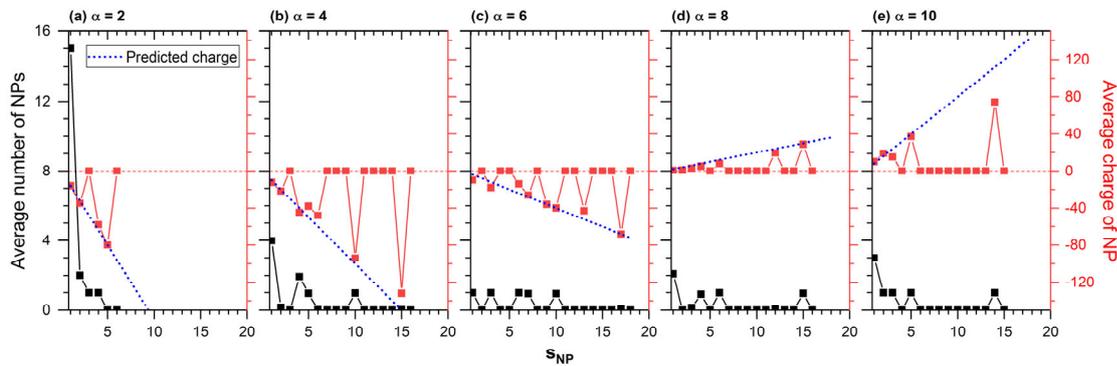


Figure 5: Steady-state NP size distributions (black) and average charge of NPs as a function of size (red) for $\alpha = 2-10$ (a-e). The predicted charge (blue) is calculated using Eq. 1.

Considering only the non-zero Q_{NP} , it is interesting to see that its magnitude increases almost linearly with s_{NP} for all α . Consulting the initial slopes (before $250 ns$) of **Figure 3b, c** it is clear that the rate of formation of peripheral PEIs is much faster than that of bridging PEIs. As a result, for any α the DNAs in the system are likely to have a similar number of peripheral PEIs before they begin to aggregate. Therefore, it is reasonable to hypothesize that if these DNAs aggregate into a NP, its charge would be proportional to the number of DNAs in it. As a very rough approximation, we assume all the PEIs in the system are bound to DNAs before they begin to aggregate. Under this assumption, the NP charge can be predicted (Q_{NP}^{pred}) as a function of s_{NP} and α using Eq. 1, where Q_{DNA} , N_{DNA} , Q_{PEI} and N_{PEI} are the charge and number of a single DNA

and a single PEI respectively. In **Figure 5**, Q_{NP}^{pred} is plotted (blue dots) alongside the average NP charge from the simulations, and are found to be in good agreement with each other. This suggests that the assumption is quite reasonable, which is because most of the PEIs in the systems have become bound to the DNAs before they begin to aggregate. The remaining small number of PEIs that bind to the NPs later in the simulation only results in small deviation from the predicted charge. The deviation of Q_{NP} from Q_{NP}^{pred} is negligible for $\alpha = 2-8$ because the number of unbound PEIs (N_u , see **Figure 3a**) is eventually zero (for $\alpha = 2-6$) or close to zero ($\alpha = 8$). The discrepancy, however, is considerable for $\alpha = 10$ because $N_u (= 13$ at steady-state) is not negligible. The excessive PEIs in $\alpha = 10$ cannot all be accommodated by the DNAs and therefore the charge of NPs is lower than our prediction. The discrepancy is particularly large for large NPs (see $s_{NP} = 14$ in **Figure 5e**) because although some PEIs bind to the NPs later in the simulation, they tend to bind to smaller NPs due to their lower positive charge.

$$Q_{NP}^{pred} = s_{NP}(N_{DNA}Q_{DNA} + N_{PEI}Q_{PEI})/N_{DNA} = s_{NP}(Q_{DNA} + \alpha Q_{PEI}) \quad (1)$$

Since Q_{NP} at the steady-state is proportional to s_{NP} , the quantity Q_{NP}/s_{NP} represents the degree of charge neutralization of DNAs in the NP. While Q_{NP} in **Figure 5** is calculated at the steady-state, it is of interest to study how the degree of neutralization changes during the aggregation process. In **Figure 6a**, the average $\langle Q_{NP}/s_{NP} \rangle$ is shown for different α , where the average is performed over different time windows and additionally for each time window, it is averaged over different s_{NP} . The prediction Q_{NP}^{pred}/s_{NP} from Eq. 1 is also shown for comparison. For $\alpha = 2$ and 4 the prediction (Eq. 1) could accurately estimate $\langle Q_{NP}/s_{NP} \rangle$ for all time windows, whereas for $\alpha = 6-10$ the prediction is accurate only for $t > 1 \mu s$. The lack of agreement for $t < 1 \mu s$ is because not all the PEIs are bound to DNAs during this time window. Although our approximation of NP charge (Eq. 1) is simple, its prediction is accurate for $t > 1 \mu s$ (see SI Section S7 for an extended,

time-dependent model for $Q_{NP}^{pred}(t)$). In the following, to ease the discussion on the aggregation process we will use Q_{NP}^{pred} instead of Q_{NP} which does not change with simulation time.

Q_{NP}^{pred}/s_{NP} takes the values of -16, -10, -4, 2, and 8 for $\alpha = 2, 4, 6, 8$ and 10 respectively. This implies that on average, the NPs in systems with $\alpha = 2-6$ are negative (under-neutralized), and the NPs in systems with $\alpha = 8-10$ are positive (over-neutralized). NPs with the same sign of net charge would have long-range electrostatic repulsion and the strength of repulsion would depend on the product of their net charges. Since the charge of a NP is approximately proportional to its size (Eq. 1), we define an intrinsic property, $(Q_{NP}^{pred}/s_{NP})^2$, as the *specific repulsion*. The specific repulsion directly depends on the degree of neutralization of DNAs in a NP and approximates the strength of repulsion between two NPs with $s_{NP} = 1$. It only depends on α of the system, and we classify it as strong for $\alpha = 2$, moderate for $\alpha = 4$ and 10, and weak for $\alpha = 6$ and 8. Despite the long-range repulsion NPs do aggregate, which can only be achieved through local electrostatic attraction, i.e., the attraction between protonated amines in one NP and deprotonated phosphates in another NP. Therefore, aggregation of DNAs is hindered by long-range electrostatic repulsion but promoted by the diffusion of NPs which brings two NPs close enough so that they may aggregate via local electrostatic attraction.

To understand the effects of diffusion we will borrow some results from the Smoluchowski coagulation equation⁶² (SCE) for spherical Brownian particles. SCE states that the aggregation of two particles is faster if the two particles have a larger ratio between their radii. For example, aggregation is slow for two large particles due to their low diffusion coefficient, and for two small particles due to their small radius of sphere of action which prevents their collision. On the other hand, it is easier for a small particle which has a higher diffusion coefficient to collide and aggregate with a large particle which has a larger radius of sphere of action. It should be noted that

strictly speaking SCE is not directly applicable for our systems because our aggregating particles (DNAs) resemble cylinders and not spheres. However, the qualitative results mentioned above help us better understand the size distribution observed in **Figure 5**.

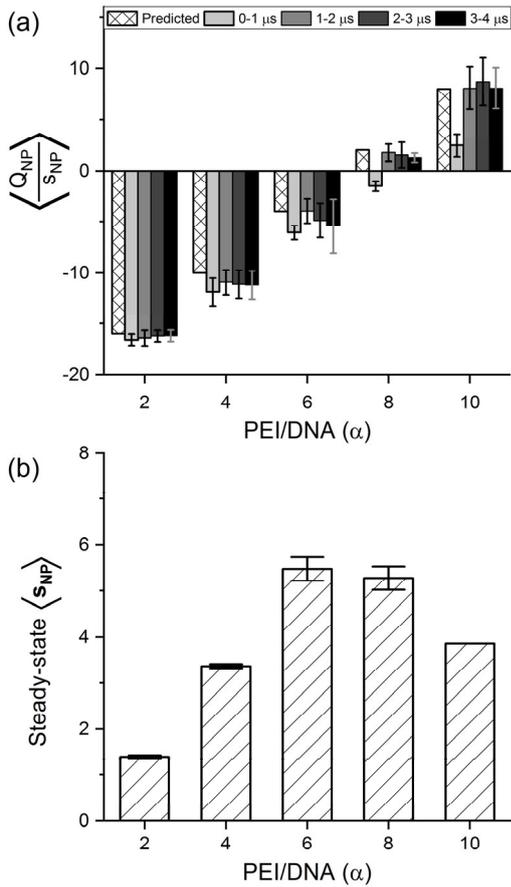


Figure 6: (a) Average charge associated with one DNA in the NPs, representing the degree of neutralization. (b) Steady-state average NP size.

Regardless of the value of α , the NP size distribution in **Figure 5** has a peak located at $s_{NP} = 1$ representing unaggregated DNAs. Excluding $s_{NP} = 1$, three distinct shapes are observed for the size distribution akin to the specific repulsion defined above. For $\alpha = 2$, we hardly notice any aggregation of DNAs (only small-NPs with $s_{NP} \leq 3$ are formed), and the distribution is unimodal with the peak at $s_{NP} = 2$. This is because of the strong specific repulsion, and the lack of PEIs to

form bridges (see **Figure 3c, d**). For $\alpha = 4$ and 10, the distribution appears to be bimodal, with peaks at $s_{NP} = 4$ and 10 for $\alpha = 4$, and at $s_{NP} = 5$ and 14 for $\alpha = 10$. The two peaks represent the formation of moderate ($4 \leq s_{NP} \leq 6$) and large NPs ($s_{NP} \geq 7$), respectively. At these α 's, the specific repulsion is moderate, which at the early stage of the aggregation process allows the aggregation of small NPs to form a few moderate NPs. Afterwards, a small NP would aggregate faster with a moderate NP than another small NP. As a result, a moderate NP would grow in size to form a large NP while new moderate NPs are formed from the aggregation of small NPs. This leads to the bimodal distributions seen in **Figure 5b, e**. The two peaks remain separated because the aggregation of a moderate NP and a large NP is unlikely: they both diffuse slowly and have a large net charge (proportional to s_{NP}), which gives rise to large long-range electrostatic repulsion (Eq. 1). For $\alpha = 6$ and 8 the distribution is multimodal which resembles a uniform distribution. In these cases, the specific repulsion is weak which enhances the aggregation of small NPs at the early stage of the aggregation process. Consequently, many moderately-sized NPs are formed, with a broader size distribution.

$$\langle s_{NP} \rangle = \frac{\sum_{s_{NP}=1}^{s_{NP}=27} s_{NP} N_{NP}(s_{NP}; t)}{\sum_{s_{NP}=1}^{s_{NP}=27} N_{NP}(s_{NP}; t)} = \frac{N_{DNA}}{\sum_{s_{NP}=1}^{s_{NP}=27} N_{NP}(s_{NP}; t)} \quad (2)$$

Figure 6b shows the average NP size ($\langle s_{NP} \rangle$) at the steady-state. $\langle s_{NP} \rangle$ at any time is calculated from Eq. 2, where $N_{NP}(s_{NP}; t)$ is the number of NPs at time t that has size s_{NP} . So $\langle s_{NP} \rangle$ is essentially obtained by dividing N_{DNA} by the total number of NPs. Data in **Figure 6b** are obtained from the steady-state stage of the simulation (last 0.5 μ s), and the error bars arise from the standard deviation in the total number of NPs. The steady-state $\langle s_{NP} \rangle$ is lowest for $\alpha = 2$ (strong specific repulsion), moderate for $\alpha = 4$ and 10 (moderate specific repulsion) and highest for $\alpha = 6$ and 8

(weak specific repulsion). Therefore, the specific repulsion not only controls the characteristic of the size distribution but also the steady-state $\langle s_{NP} \rangle$.

3.3. Rate of NP growth

Figure 7a shows the average NP size ($\langle s_{NP} \rangle$) as a function of time for all α , where $\langle s_{NP} \rangle$ was calculated using Eq. 2. Upon visual inspection, we notice that $\langle s_{NP} \rangle$ exhibits linear trend over certain time windows. For simplicity, three time windows 0-2, 2-2.7 and 2.7-3.5 μs were chosen over which $\langle s_{NP} \rangle$ was approximated by linear functions. Essentially each curve $\langle s_{NP} \rangle$ in **Figure 7a** is fitted with a piecewise linear function which is continuous at 2 and 2.7 μs . Two additional constraints were imposed in the curve fitting. Firstly, $\langle s_{NP} \rangle$ is set to be 1 at $t = 0$, representing initial unaggregated DNAs. Secondly, $\langle s_{NP} \rangle$ at $t = 3.5 \mu\text{s}$ is set to be the steady-state $\langle s_{NP} \rangle$ value shown in **Figure 6b**. The piecewise linear functions are then determined by least squares fitting and shown as dashed lines in **Figure 7a**. The rates of NP growth (r_{NP}), defined as the slopes of the piecewise linear functions, are shown in **Figure 7b** for different α and time windows.

The first observation made from **Figure 7b** is that the initial rate (0-2 μs) increases as α increases from 2 to 6. This is expected because the number of peripheral PEIs increases with α (see **Figure 3b**), which leads to higher probability of converting to bridging PEIs that aggregates DNAs together. The same increase in r_{NP} is observed as α increases from 8 to 10. However, both $\alpha = 8$ and 10 have lower r_{NP} than $\alpha = 6$, and there is a sharp decrease in r_{NP} from $\alpha = 6$ to $\alpha = 8$. At first glance, this is counter-intuitive, especially since the specific repulsion is similar for $\alpha = 6$ and 8. This sharp decrease most likely arises from the difference in the mode of aggregation between $\alpha = 2-6$ and $\alpha = 8-10$. Since $\alpha = 8-10$ leads to branched NPs which are more compact, these NPs are expected to have a smaller radius of sphere of action than $\alpha = 2-6$ for the same s_{NP} .

($s_{NP} > 1$). This reduces the chances of collisions between NPs for $\alpha = 8-10$ (branched) when compared to $\alpha = 6$ (linear) and therefore a sharp decrease in r_{NP} is observed between $\alpha = 6$ and 8.

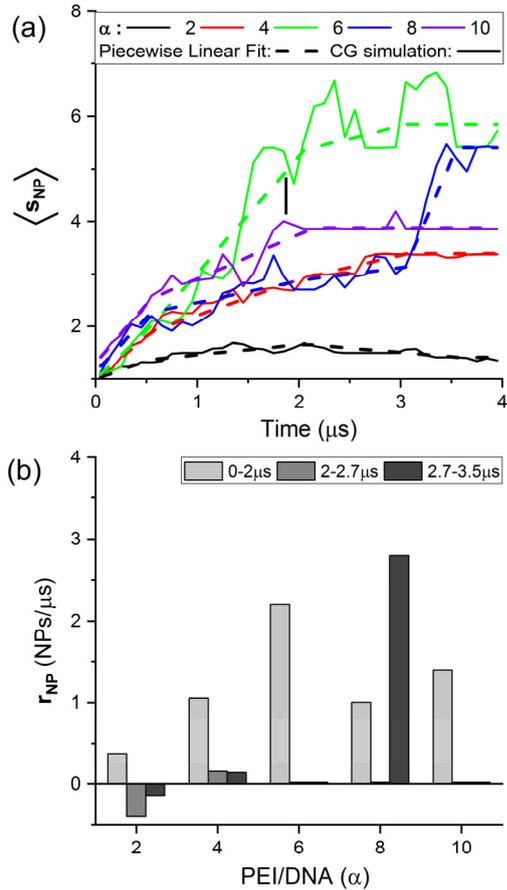


Figure 7: (a) Average NP size as a function of time. (b) Rate of NP growth calculated as the slope of the curves in (a) in the specified time ranges.

Another observation from **Figure 7b** is that except for $\alpha = 8$, r_{NP} decreases with time, i.e. r_{NP} is highest for 0-2 μs followed by 2-2.7 μs , and lowest for 2.7-3.5 μs . This reduction primarily occurs for two reasons. Firstly, as NPs increase in size, their charge increases proportionately (see Eq. 1) and as a result, electrostatic repulsion between them increases which reduces r_{NP} . Secondly, as NPs aggregate the total number of NPs in the system decreases which reduces the probability of collision.⁶² Although according to SCE, as the size of the largest NP increases, the size ratio between the largest and smallest NPs would increase, which could potentially increase r_{NP} , this

effect is secondary in our simulations. Reduction of r_{NP} also occurs due to other reasons which are specific to different systems. For $\alpha = 2$, the average number of bridging PEIs between a pair of bridged DNAs (n_b) is low (see **Figure 3d**), making the aggregate unstable. This instability even makes r_{NP} negative for $t > 2 \mu\text{s}$. For $\alpha = 4$ and 6, the reduction in r_{NP} also occurs due to the instability of loose linear NPs. This can be seen from the relatively low average number of NPs (~ 0) for large NPs in **Figure 5**. Another evidence for the instability is the fluctuations in $\langle s_{NP} \rangle$ for $\alpha = 6$ (**Figure 7a**). For $\alpha = 10$ as the peripheral PEIs accumulate and surround each DNA it becomes more difficult for the phosphates to be accessible to amines in another NP, which reduces local electrostatic attractions and hence reduces r_{NP} to zero.

The system with $\alpha = 8$ did not show a monotonic relation between r_{NP} and time. Instead, r_{NP} is highest for 2.7-3.5 μs , followed by 0-2 μs and is zero for 2-2.7 μs . Since the specific repulsion is weak for both $\alpha = 8$ and 6 while such a behavior was not observed for $\alpha = 6$, it is most likely associated with the different mode of aggregation. One possible explanation for $r_{NP} = 0$ within 2-2.7 μs is that during this period the compact NPs hardly collide with each other. Instead, they diffuse close to each other and begin to aggregate from 2.7 μs . Unlike $\alpha = 8$, we do not observe $r_{NP} > 0$ for $\alpha = 10$ later in the simulation, which can be due to the greater long-range electrostatic repulsion and the lack of exposed phosphates. The PEI/DNA ratio, therefore, influences the rate of NP growth not only by the long-range electrostatic repulsion but also by the mode of aggregation.

3.4. Influence of salt concentration

The aggregation in the presence of 150 mM KCl was qualitatively similar to that in the absence of salts (0 mM KCl), with some minor quantitative differences. To avoid repetition, we will only present key results for $\alpha = 10$, while the rest of the results can be found in SI (Section S8). In

Figure 8a, we plot N_u , N_p and N_b as a function of time for 0 and 150 mM KCl systems. N_u with 150 mM salt is higher for most of the simulation time, implying that the rate of DNA-PEI binding decreases with increase in salt. This can be explained by the competition between PEI and K^+ in binding with the DNA and screening of DNA-PEI attractions. However, in the last 1 μ s of the simulation N_u is similar for both salt concentrations. This is expected because PEIs form more stable binding with DNAs as compared to K^+ . Consistent with the results for N_u , N_p is lower at higher salt concentration. The number of bridging PEIs N_b is practically identical for both salt concentrations, that is, the small changes in N_p do not affect N_b (this is consistent with the discussion in SI Section S4). **Figure 8b** shows the average number of bridging PEIs between a pair of bridged DNAs (n_b) as a function of time, which is also hardly affected by salt concentration.

Figure 8c plots the NP size distribution (black) and average NP charge as a function of s_{NP} (red) for both 0 and 150 mM KCl. When 150 mM KCl is added, the shape of the NP size distribution does not change and remains bimodal. However, the size of the largest NP increases from $s_{NP} = 14$ to $s_{NP} = 19$. The presence of 150 mM KCl has introduced a screening effect on the electrostatic repulsion between smaller NPs, leading to their aggregation into larger NPs.^{7,63} The NP charge in presence of 150 mM KCl also increases linearly with s_{NP} similar to the case of 0 mM KCl. Predicted charge from Eq. 1 is shown as the dotted blue line and is in good agreement with simulation data when $s_{NP} < 6$ (similar to $\alpha = 10$ for 0 mM KCl). In **Figure 8d**, the average NP size $\langle s_{NP} \rangle$ is plotted as a function of time, which shows similar characteristics for both 0 and 150 mM KCl. The initial slope of $\langle s_{NP} \rangle$ is clearly higher at higher salt concentration, representing a higher rate of NP growth. In fact, $\langle s_{NP} \rangle$ is larger for 150 mM KCl during the entire time of the simulations. Both the increased rate of NP growth and size in the presence of 150 mM KCl can be explained by the screening of long-range electrostatic repulsions between NPs by salt.

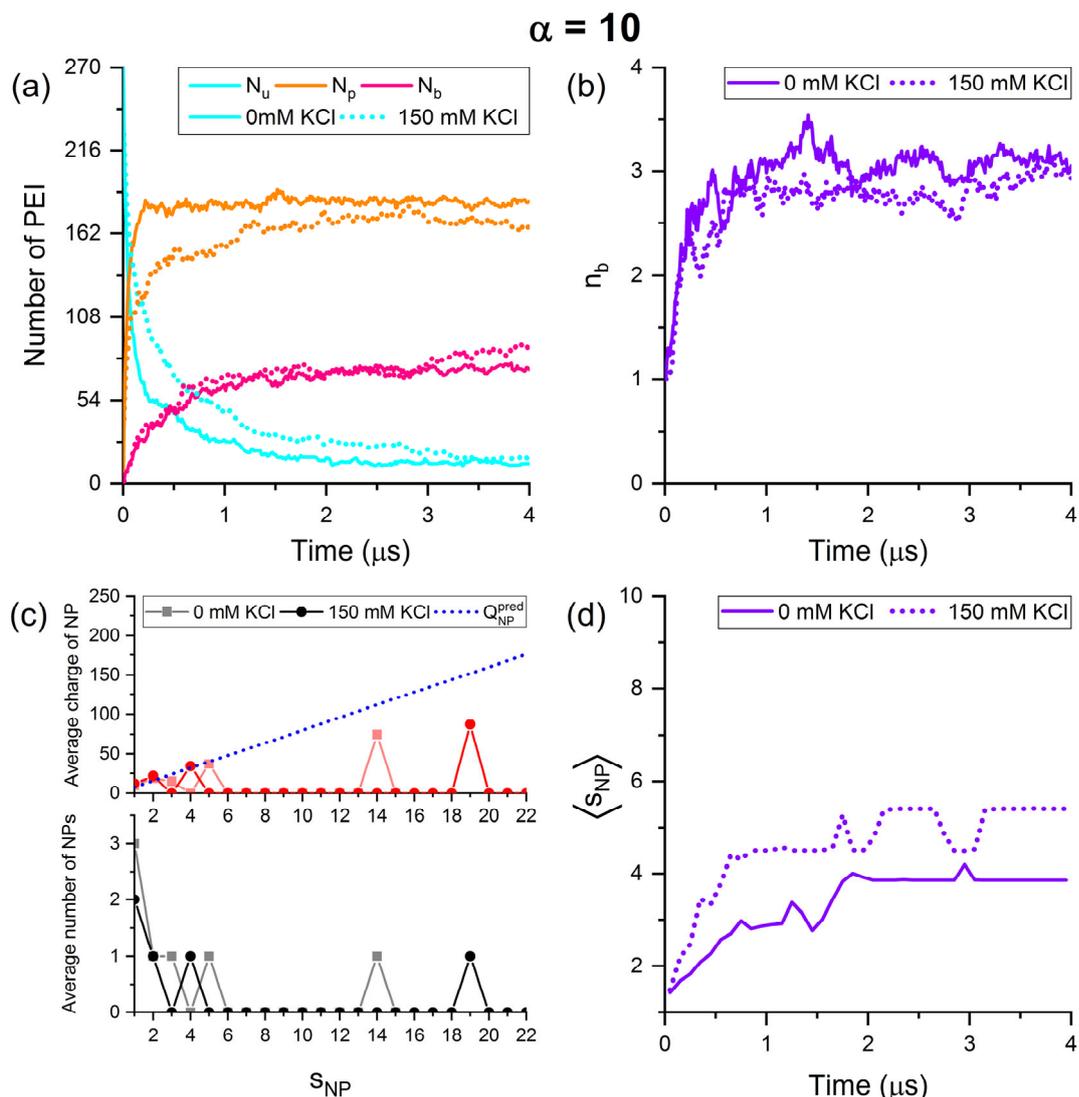


Figure 8: Influence of salt concentration on (a) number of PEIs in different roles; (b) average number of bridging PEIs between a pair of bridged DNAs (n_b); (c) steady-state size distribution (black) and average NP charge vs size (red); (d) average NP size as a function of time. The results shown here are only for $\alpha = 10$, and other results can be found in SI Section S8.

4. Discussion

Charge of DNA-PEI NPs can be quantified in experiments using the zeta potential. The zeta potential is negative for low N/P ratio and positive for sufficiently large N/P ratio, although the exact N/P ratio to cause zero zeta potential varies for different PEIs and DNAs. Previous work has reported zero zeta potential at N/P ratio ~ 5 for 600 Da PEI⁶⁴, ~ 2.3 for 2 kDa PEI²¹, 2.7 to 3.84 for

25kDa PEI^{21,23,64}, ~ 3 for 750 kDa PEI²¹ and ~ 2.2 for 800 kDa¹⁷ PEI. In our simulations of 586 Da PEIs, the NPs are negative for $N/P < 3.55$ and positive for $N/P > 4.73$ (**Table 1**). The location of zero zeta potential can be approximated using linear interpolation, which turns out to be at $N/P \sim 4.4$ (See SI Section S9 for details). These results agree well with the experimental findings.

In experiments, the size of NPs is expressed in terms of hydrodynamic radius measured by dynamic light scattering.^{21,65} Although the NP size distributions we presented earlier in **Figure 5** are in terms of number of DNAs in the NP, calculation of hydrodynamic radius results in distributions that are qualitatively similar (see SI Section S10 for details). For 2k, 25k and 750k Da PEIs, Choosakoonkriang et al.²¹ found the mean hydrodynamic radius to be largest for intermediate N/P ratios at which the zeta potential is nearly zero. This is consistent with our results, where the NP size is largest for $\alpha = 6$ and 8, which has the lowest negative and positive charge respectively. Erbacher et al.²³ studied NP size distribution at different N/P ratios using 25 kDa PEI. They observed bimodal size distribution at $N/P = 2$, consistent with the observation made for $\alpha = 4$ ($N/P = 2.36$) in our simulation. Dunlap et al.¹² reported several peaks in the NP size distribution for a DNA system over-neutralized by 25 kDa PEI at 150mM NaCl, which is similar to the system with $\alpha = 10$ in our simulation. Aggregation of semi-flexible polyelectrolytes by multivalent salts studied using CG simulations showed similar shapes of the size distribution.⁴⁴ Specifically, they observed limited aggregation for highly under-neutralized systems similar to $\alpha = 2$, bimodal distribution for lightly under-neutralized or over-neutralized systems similar to $\alpha = 4$ and 10, and uniform distribution for near-neutral systems similar to $\alpha = 6$ and 8. It should be noted that these distributions were found when the Bjerrum length is about twice the diameter of the beads in the polyelectrolyte, while the results may be different for other Bjerrum lengths.

On the other hand, there are some differences between experimental studies and our results on the NP size distribution. Unimodal size distribution was observed by Erbacher et al.²³ for N/P = 3 and 5, where the zeta potential was nearly zero. Our simulation results predict a more or less uniform distribution at these N/P ratios. Since the PEIs in our work have drastically different molecular weight (586 Da as compared to 25 kDa in Erbacher et al.²³), and different size distributions were observed for 22 kDa and 25 kDa PEIs,¹² we can expect the NP size distribution to be dependent on the molecular weight of PEI.

Choosakoonkriang et al.²¹ found DNAs in NPs to have different circular dichroism (CD) spectra for systems with excess PEIs and excess DNAs. They proposed the changes in CD spectra to be caused by changes in base-base interaction of a DNA. Interestingly, we observed different modes of aggregation due to crowding and misalignment of PEIs to the DNA backbone. The crowding of PEIs in the major and minor grooves might alter the base-base interactions in a DNA and in turn the CD spectra.

The different modes of aggregation and the mechanisms behind can allow us to propose a strategy for controlling the hydrodynamic radius and shape of a NP. While an efficient gene delivery would require compaction and aggregation of multiple NAs, NPs with a very large hydrodynamic radius are difficult for cellular uptake.²⁴ Also, spherical NPs are expected to have higher cellular uptake,²⁴ so there is a need to generate NPs that have an optimal number of NAs, are sphere-like, compact and stable. A numerical experiment was done to achieve this. In particular, for $\alpha = 4$, a large linear NP was formed as shown in **Figure 2b**. Linear NPs larger than this ($S_{NP} > 10$) was unstable, as pointed out in **Section 3.2**. At the end of the 4 μ s simulation, 162 PEIs were added to change α of the system to 10. To make the system electroneutral K^+ ions in the $\alpha = 4$ system were removed and an appropriate amount of Cl^- ions were added. NPT simulation

was run on this new system for another 3.5 μs . Comparing the largest NP before (**Figure 2b**) and after (**Figure S10** in SI Section S11) the simulation, we observe that the NP retains its size (in terms of number of DNAs) but becomes more compact (with reduced hydrodynamic radius). The DNAs in the NP are brought closer by the newly added PEIs through the formation of side-by-side contacts. This implies that compact and stable NPs can be formed using a two-step addition of PEIs. The first step involves the addition of PEIs so that the DNAs are moderately under-neutralized. Aggregation is expected to occur in an end-to-end fashion forming linear NPs, but the number of DNAs in the aggregate will not be too large due to their instability. In the second step, a large number of PEIs are introduced to bypass the N/P ratio associated with weak specific repulsion to prevent further aggregation. The previously formed linear NPs are then condensed into stable NPs that are more compact and closer to being spherical, with the help of excess PEIs (see SI Section S11). Since the largest NP formed at 4 μs for $\alpha = 4$ has smaller s_{NP} than $\alpha = 10$, the NP created using the two-step process would have a lower hydrodynamic radius than the one formed using a single step process.

Finally, we comment on the limitation of this work. As with all MD simulations, the length and time scales of our simulated systems are still small compared with experiments. We have used a low molecular weight (LMW) PEI, 586 Da, while 25 kDa PEI is the most commonly studied PEI in experiments. It should be noted that there are experiments which explored the efficacy of LMW PEIs (600-800 Da).^{5,6,66} As well, there are interests in introducing functional modification (e.g., lipid substitution) to LMW PEIs (0.6-2 kDa) to achieve a balance between efficacy and toxicity.⁹ To be able to observe the binding and aggregation process within the timeframe permitted by MD simulations, we have used DNA and PEI concentrations that are larger than those in experiments. However, the local concentration of DNAs and PEIs near a NP can be much greater than their

concentration in the bulk solution, and we believe it is this local concentration that affects the binding most. Setting a higher concentration in simulation essentially accelerates the diffusion process of the DNAs and PEIs before their binding and aggregation, which would have taken more time to occur in experiments. We also point out that compared with similar works in the literature,³⁵ our systems are considerably larger, and the investigation on multi-NP formation allowed us to obtain size distribution from the simulations which has never been attempted before. The good agreement of our results with experiments has shown great potential of using simulations as a predictive tool to assist in the design of delivery systems. One way of extending the length of the simulations is to use larger time steps, e.g., 10 fs instead of 5 fs used in this work. While simulations of Martini PEI⁵¹ and Martini DNA⁵⁰ separately in polarizable water⁵² were found to be numerically stable for a time step of 10 fs, our initial attempt of using 10 fs for simulating Martini PEI and DNA together in polarizable water failed due to numerical instability. Since large Martini simulations of PEI and DNA together had not been performed before, the reason behind the numerical instability is not yet clear and remains to be explored.

5. Conclusions

Large-scale coarse-grained molecular dynamics simulations are performed to study gene delivery nanoparticles formed from the aggregation of DNAs and PEIs. The effects of PEI to DNA number ratio, or equivalently the N/P ratio, are investigated on the shape, charge and size of the nanoparticles. At very low N/P ratio, the aggregation of DNAs is limited, and as the N/P ratio increases the nanoparticles change from a loose linear structure to a compact branched structure, which is closer to a sphere. Such a transition in the mode of aggregation is caused by the different alignments of PEIs with DNA backbones prior to aggregation, which dictates their ability to serve as bridging PEIs. Except for very large nanoparticles at high N/P ratios, the charge of a

nanoparticle is proportional to the number of DNAs in it. Their ratio allows for the definition of an intrinsic property called specific repulsion, which controls the characteristics of the steady-state size distribution of NPs: unimodal for strong specific repulsion, bimodal for moderate specific repulsion, and more or less uniform for weak specific repulsion. The rate of nanoparticle growth is affected by the N/P ratio through the long-range electrostatic repulsions and mode of aggregation. Presence of salt does not have any qualitative influence on the formation and characteristics of the nanoparticles. Quantitative, adding salt causes an increase in the rate of nanoparticle growth and their steady-state size, due to the screening of electrostatic repulsion between nanoparticles brought by the ions.

6. Associated content

Supporting Information: relative shape anisotropy of nanoparticles, DNA-PEI binding, conversion of PEI roles, average number of bridging PEIs between a pair of bridged DNAs (n_b), additional mechanism of DNA over-saturation, determination of steady-state, time-dependent nanoparticle charge, aggregation in presence of 150 mM KCl, zeta potential, distribution of hydrodynamic radius, formation of nanoparticles using a two-step process.

7. Acknowledgements

We acknowledge the computing resources and technical support from Western Canada Research Grid (WestGrid). Financial support from the Natural Science and Engineering Research Council (NSERC) of Canada and MITCAS is gratefully acknowledged. The authors declare no competing financial interest.

8. References

- [1] Boussif, O.; Lezoualc'h, F.; Zanta, M. A.; Mergny, M. D.; Scherman, D.; Demeneix, B.; Bher, J.-P. A Versatile Vector for Gene and Oligonucleotide Transfer into Cells in Culture and *in Vivo*: Polyethylenimine. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7297–7301.
- [2] Huang, P.; Zhao, J.; Wei, C.; Hou, X.; Chen, P.; Tan, Y.; He, C.-Y.; Wang, Z.; Chen, Z.-Y. Erythrocyte Membrane Based Cationic Polymer-mcDNA Complexes as an Efficient Gene Delivery System. *Biomater. Sci.* **2017**, *5*, 120–127.
- [3] Godbey, W. T.; Barry, M. A.; Saggau, P.; Wu, K. K.; Mikos, A. G. Poly(ethylenimine)-Mediated Transfection: A New Paradigm for Gene Delivery. *J. Biomed. Mater. Res.* **2000**, *51*, 321–328.
- [4] Vasken Aposhian, H.; Barr, S. M.; Keck, K. Experimental Gene Delivery Systems for Mammalian Cells - the Polyoma Pseudovirus System. *Adv. Enzyme Regul.* **1978**, *16*, 275–288.
- [5] Godbey, W. T.; Wu, K. K.; Mikos, A. G. Size Matters: Molecular Weight Affects the Efficiency of Poly(ethylenimine) as a Gene Delivery Vehicle. *J. Biomed. Mater. Res.* **1999**, *45*, 268–275.
- [6] Forrest, M. L.; Koerber, J. T.; Pack, D. W. A Degradable Polyethylenimine Derivative with Low Toxicity for Highly Efficient Gene Delivery. *Bioconjugate Chem.* **2003**, *14*, 934–940.
- [7] Wightman, L.; Kircheis, R.; Rössler, V.; Garotta, S.; Ruzicka, R.; Kursa, M.; Wagner, E. Different Behavior of Branched and Linear Polyethylenimine for Gene Delivery In Vitro and In Vivo. *J. Gene Med.* **2001**, *3*, 362–372.
- [8] Boeckle, S.; von Gersdorff, K.; van der Pijpen, S.; Culmsee, C.; Wagner, E.; Ogris, M. Purification of Polyethylenimine Polyplexes Highlights the Role of Free Polycations in Gene Transfer. *J. Gene Med.* **2004**, *6*, 1102–1111.
- [9] Remant Bahadur, K. C.; Landry, B.; Aliabadi, H. M.; Lavasanifar, A.; Uludağ, H. Lipid Substitution on Low Molecular Weight (0.6-2.0 kDa) Polyethylenimine Leads to a Higher Zeta Potential of Plasmid DNA and Enhances Transgene Expression. *Acta Biomater.* **2011**, *7*, 2209–2217.
- [10] Tang, M. X.; Szoka, F. C. The Influence of Polymer Structure on the Interactions of Cationic Polymers with DNA and Morphology of the Resulting Complexes. *Gene Ther.* **1997**, *4*, 823–832.
- [11] Utsuno, K.; Uludağ, H. Thermodynamics of Polyethylenimine-DNA Binding and DNA Condensation. *Biophys. J.* **2010**, *99*, 201–207.

- [12] Dunlap, D. D.; Maggi, A.; Soria, M. R.; Monaco, L. Nanoscopic Structure of DNA Condensed for Gene Delivery. *Nucleic Acids Res.* **1997**, *25*, 3095–3101.
- [13] Godbey, W. T.; Wu, K. K.; Mikos, A. G. Tracking the Intracellular Path of Poly(ethylenimine)/DNA Complexes for Gene Delivery. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 5177–5181.
- [14] Bieber, T.; Meissner, W.; Kostin, S.; Niemann, A.; Elsasser, H.-P. Intracellular Route and Transcriptional Competence of Polyethylenimine-DNA Complexes. *J. Controlled Release* **2002**, *82*, 441–454.
- [15] Verma, A.; Stellacci, F. Effect of Surface Properties on Nanoparticle–Cell Interactions. *Small* **2010**, *6*, 12–21.
- [16] Villanueva, A.; Cañete, M.; Roca, A. G.; Calero, M.; Veintemillas-Verdaguer, S.; Serna, C. J.; Del Puerto Morales, M.; Miranda, R. The Influence of Surface Functionalization on the Enhanced Internalization of Magnetic Nanoparticles in Cancer Cells. *Nanotechnology* **2009**, *20*, 115103.
- [17] Ogris, M.; Brunner, S.; Schüller, S.; Kircheis, R.; Wagner, E. PEGylated DNA/Transferrin–PEI Complexes: Reduced Interaction with Blood Components, Extended Circulation in Blood and Potential for Systemic Gene Delivery. *Gene Ther.* **1999**, *6*, 595–605.
- [18] Li, S.; Rizzo, M. A.; Bhattacharya, S.; Huang, L. Characterization of Cationic Lipid-Protamine–DNA (LPD) Complexes for Intravenous Gene Delivery. *Gene Ther.* **1998**, *5*, 930–937.
- [19] Khalil, I. A.; Kogure, K.; Akita, H.; Harashima, H. Uptake Pathways and Subsequent Intracellular Trafficking in Nonviral Gene Delivery. *Pharmacol. Rev.* **2006**, *58*, 32–45.
- [20] Hillaireau, H.; Couvreur, P. Nanocarriers’ Entry into the Cell: Relevance to Drug Delivery. *Cell. Mol. Life Sci.* **2009**, *66*, 2873–2896.
- [21] Choosakoonkriang, S.; Lobo, B. A.; Koe, G. S.; Koe, J. G.; Middaugh, C. R. Biophysical Characterization of PEI/DNA Complexes. *J. Pharm. Sci. (Philadelphia, PA, U. S.)* **2003**, *92*, 1710–1722.
- [22] D’Andrea, C.; Pezzoli, D.; Malloggi, C.; Candeo, A.; Capelli, G.; Bassi, A.; Volonterio, A.; Taroni, P.; Candiani, G. The Study of Polyplex Formation and Stability by Time-Resolved Fluorescence Spectroscopy of SYBR Green I-Stained DNA. *Photochem. Photobiol. Sci.* **2014**, *13*, 1680–1689.
- [23] Erbacher, P.; Bettinger, T.; Belguise-Valladier, P.; Zou, S.; Coll, J.-L.; Bher, J.-P.; Remy, J.-S. Transfection and Physical Properties of Various Saccharide, Poly(Ethylene Glycol), and Antibody-Derivatized Polyethylenimines (PEI). *J. Gene Med.* **1999**, *1*, 210–222.

- [24] Chithrani, B. D.; Ghazani, A. A.; Chan, W. C. W. Determining the Size and Shape Dependence of Gold Nanoparticle Uptake into Mammalian Cells. *Nano Lett.* **2006**, *6*, 662–668.
- [25] Chithrani, B. D.; Chan, W. C. W. Elucidating the Mechanism of Cellular Uptake and Removal of Protein-Coated Gold Nanoparticles of Different Sizes and Shapes. *Nano Lett.* **2007**, *7*, 1542–1550.
- [26] Zhang, S.; Gao, H.; Bao, G. Physical Principles of Nanoparticle Cellular Endocytosis. *ACS Nano* **2015**, *9*, 8655–8671.
- [27] Ding, H.-M.; Ma, Y.-Q. Theoretical and Computational Investigations of Nanoparticle-Biomembrane Interactions in Cellular Delivery. *Small* **2015**, *11*, 1055–1071.
- [28] Kettler, K.; Veltman, K.; van de Meent, D.; van Wezel, A.; Hendriks, A. J. Cellular Uptake of Nanoparticles as Determined by Particle Properties, Experimental Conditions, and Cell Type. *Environ. Toxicol. Chem.* **2014**, *33*, 481–492.
- [29] Sun, C.; Tang, T.; Uludağ, H.; Cuervo, J. E. Molecular Dynamics Simulations of DNA/PEI Complexes: Effect of PEI Branching and Protonation State. *Biophys. J.* **2011**, *100*, 2754–2763.
- [30] Sun, C.; Tang, T. Study on the Role of Polyethylenimine as Gene Delivery Carrier Using Molecular Dynamics Simulations. *J. Adhes. Sci. Technol.* **2014**, *28*, 399–416.
- [31] Sun, C.; Tang, T.; Uludağ, H. Molecular Dynamics Simulations for Complexation of DNA with 2 kDa PEI Reveal Profound Effect of PEI Architecture on Complexation. *J. Phys. Chem. B* **2012**, *116*, 2405–2413.
- [32] Kondinskaia, D. A.; Kostritskii, A. Y.; Nesterenko, A. M.; Antipina, A. Y.; Gurtovenko, A. A. Atomic-Scale Molecular Dynamics Simulations of DNA-Polycation Complexes: Two Distinct Binding Patterns. *J. Phys. Chem. B* **2016**, *120*, 6546–6554.
- [33] Ziebarth, J.; Wang, Y. Molecular Dynamics Simulations of DNA-Polycation Complex Formation. *Biophys. J.* **2009**, *97*, 1971–1983.
- [34] Ziebarth, J.; Wang, Y. Coarse-Grained Molecular Dynamics Simulations of DNA Condensation by Block Copolymer and Formation of Core-Corona Structures. *J. Phys. Chem. B* **2010**, *114*, 6225–6232.
- [35] Wei, Z.; Luijten, E. Systematic Coarse-Grained Modeling of Complexation Between Small Interfering RNA and Polycations. *J. Chem. Phys.* **2015**, *143*, 243146.
- [36] Zhan, B.; Shi, K.; Dong, Z.; Lv, W.; Zhao, S.; Han, X.; Wang, H.; Liu, H. Coarse-Grained Simulation of Polycation/DNA-Like Complexes: Role of Neutral Block. *Mol. Pharmaceutics* **2015**, *12*, 2834–2844.

- [37] Sun, C.; Tang, T.; Uludağ, H. Molecular Dynamics Simulations of PEI Mediated DNA Aggregation. *Biomacromolecules* **2011**, *12*, 3698–3707.
- [38] Bagai, S.; Sun, C.; Tang, T. Potential of Mean Force of Polyethylenimine-Mediated DNA Attraction. *J. Phys. Chem. B* **2013**, *117*, 49–56.
- [39] Tom, A. M.; Rajesh, R.; Vemparala, S. Aggregation Dynamics of Rigid Polyelectrolytes. *J. Chem. Phys.* **2016**, *144*, 034904.
- [40] Lee, K.-C.; Borukhov, I.; Gelbart, W. M.; Liu, A. J.; Stevens, M. J. Effect of Mono- and Multivalent Salts on Angle-Dependent Attractions Between Charged Rods. *Phys. Rev. Lett.* **2004**, *93*, 128101.
- [41] Stevens, M. J. Bundle Binding in Polyelectrolyte Solutions. *Phys. Rev. Lett.* **1999**, *82*, 101-104.
- [42] Grønbech-Jensen, N.; Mashl, R. J.; Bruinsma, R. F.; Gelbart, W. M. Counterion-Induced Attraction Between Rigid Polyelectrolytes. *Phys. Rev. Lett.* **1997**, *78*, 2477-2480.
- [43] Fazli, H.; Golestanian, R. Aggregation Kinetics of Stiff Polyelectrolytes in the Presence of Multivalent Salt. *Phys. Rev. E* **2007**, *76*, 041801.
- [44] Sayar, M.; Holm, C. Equilibrium Polyelectrolyte Bundles with Different Multivalent Counterion Concentrations. *Phys. Rev. E* **2010**, *82*, 031901.
- [45] Sayar, M.; Holm, C. Finite-Size Polyelectrolyte Bundles at Thermodynamic Equilibrium. *Europhys. Lett.* **2007**, *77*, 16001.
- [46] Diehl, A.; Carmona, H. A.; Levin, Y. Counterion Correlations and Attraction Between Like-Charged Macromolecules. *Phys. Rev. E* **2001**, *64*, 011804.
- [47] Fazli, H. Mohammadinejad, S.; Golestanian, R. Salt-Induced Aggregation of Stiff Polyelectrolytes. *J. Phys.: Condens. Matter* **2009**, *21*, 424111.
- [48] Tom, A. M.; Rajesh, R.; Vemparala, S. Aggregation of Flexible Polyelectrolytes: Phase Diagram and Dynamics. *J. Chem. Phys.* **2017**, *147*, 144903.
- [49] Ogris, M.; Steinlein, P.; Kursa, M.; Mechtler, K.; Kircheis, R.; Wagner, E. The Size of DNA/Transferrin-PEI Complexes is an Important Factor for Gene Expression in Cultured Cells. *Gene Ther.* **1998**, *5*, 1425–1433.
- [50] Uusitalo, J. J.; Ingólfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.* **2015**, *11*, 3932–3945.

- [51] Mahajan, S.; Tang, T. Martini Coarse-Grained Model for Polyethylenimine. *J. Comput. Chem.* **2019**, *40*, 607–618.
- [52] Yesylevskyy S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLoS Comput. Biol.* **2010**, *6*, 1000810.
- [53] Marrink, S. J.; De Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.
- [54] Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E.; GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- [55] Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. *Chem. Soc. Rev.* **2013**, *42*, 6801–6822.
- [56] Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- [57] Bussi, G.; Donadio, D.; Parinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- [58] Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- [59] Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys. (Melville, NY, U. S.)* **1981**, *52*, 7182–7190.
- [60] Verlet, L. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98–103.
- [61] Tironi, I. G.; Sperb, R.; Smith, P. E.; Van Gunsteren, W. F. A Generalized Reaction Field Method for Molecular Dynamics Simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- [62] von Somluchowski, M. Versuch einer Mathematischen Theorie der Koagulationskinetik Kolloider Lösungen. *Z. Phys. Chem. (Muenchen, Ger.)* **1917**, *92*, 129–168.
- [63] Goula, D.; Remy, J. S.; Erbacher, P.; Wasowicz, M.; Levi, G.; Abdallah, B.; Demeneix, B. A. Size, Diffusibility and Transfection Performance of Linear PEI/DNA Complexes in the Mouse Central Nervous System. *Gene Ther.* **1998**, *5*, 712–717.

- [64] Tang, G. P.; Guo, H. Y.; Alexis, F.; Wang, X.; Zeng, S.; Lim, T. M.; Ding, J.; Yang, Y. Y.; Wang, S. Low Molecular Weight Polyethylenimines Linked by β -Cyclodextrin for Gene Transfer into the Nervous System. *J. Gene Med.* **2006**, *8*, 736–744.
- [65] Putnam, D.; Gentry, C. A.; Pack, D. W.; Langer, R. Polymer-Based Gene Delivery with Low Cytotoxicity by a Unique Balance of Side-Chain Termini. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 1200–1205.
- [66] Godbey, W. T.; Wu, K. K.; Mikos, A. G. Poly(ethylenimine)-Mediated Gene Delivery Affects Endothelial Cell Function and Viability. *Biomaterials* **2001**, *22*, 471–480.