SUPPLEMENTAL MATERIAL

Systematic Coarse-Grained Modeling of Complexation between Small Interfering RNA and Polycations

Zonghui Wei[1] and Erik Luijten[1,2,3,4,a)]

[1]*Graduate Program in Applied Physics, Northwestern University, Evanston, Illinois, 60208, USA*
[2]*Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois, 60208, USA*
[3]*Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, Illinois, 60208, USA*
[4]*Department of Physics and Astronomy, Northwestern University, Evanston, Illinois, 60208, USA*

a) Corresponding author. E-mail: luijten@northwestern.edu

I. PEG-grafted PEI all-atom force field development

The force field for PEG-grafted PEI is developed based upon the CHARMM General Force Field (CGenFF) version 2b8. We employ the "divide-and-conquer" strategy, where the whole polymer is separated into building blocks as indicated by the red lines in the example structure at grafting density 2% in Fig. S1. The PEI backbone has 50 amine groups and the grafted PEG chain has 20 monomers. 82% of the nitrogen sites in PEI backbone are protonated at pH 2. The unprotonated sites are assigned to positions 1, 7, 13, 19, 25, 31, 37, 43, 49.

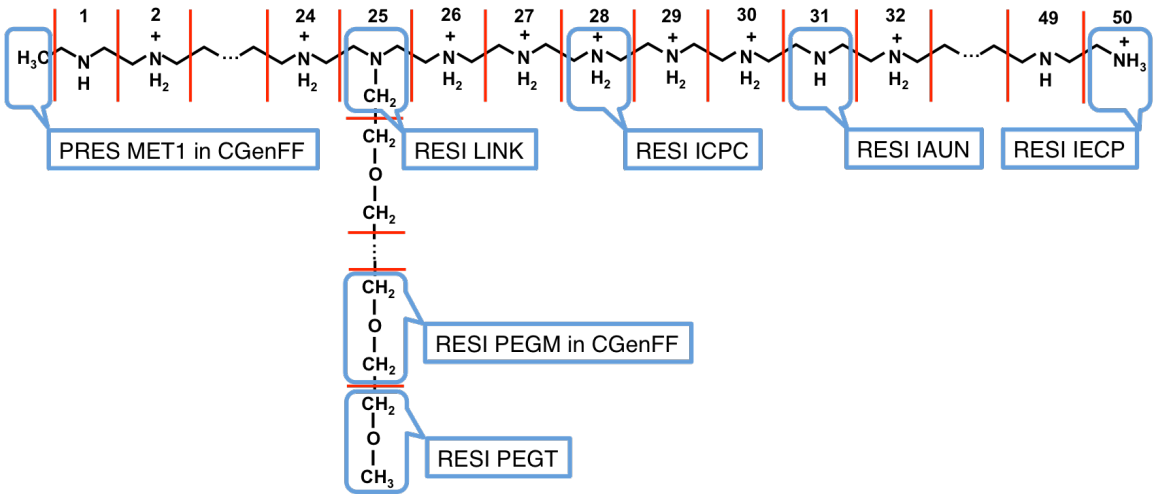


FIG. S1. Separation method and resulting building blocks for the development of the PEG-grafted PEI force field.

The above separation method results in 5 residues (building blocks) to be defined, which are named as RESI LINK, ICPC, IAUN, IECP, and PEGT, respectively. The left end of the PEI backbone exists in CGenFF as a patch MET1. The PEG monomer can also be found in CGenFF, as RESI PEGM.

To cover all the parameters needed after linking the building blocks, we design the compounds containing them based upon the local environment of individual residues within the polymer. For example, for RESI ICPC, which is usually linked to two protonated amine groups in the polymer at pH 2, we place a protonated fragment on either side of RESI ICPC and form a compound called IAPN, $CH_3NH_2^+CH_2CH_2NH_2^+CH_2CH_2NH_2^+CH_3$. Similarly, depending on all the possible local environment of the residues, we need to develop the force field for all the compounds listed in Table S1 to fully cover the parameters needed for the whole polymer. For future reference, the compounds in Table S1 cover the possible linking groups of each residue at all pH conditions, while in this paper we only discuss the case of pH 2.

TABLE S1: Designed compounds.

Compound Name	Chemical Formula
ICPC	$CH_3NHCH_2CH_2NH_2^+CH_2CH_2NHCH_3$
IECM	$CH_3CH_2NHCH_2CH_3$
IECP	$CH_3NHCH_2CH_2NH_3^+$
IAUN	$CH_3NHCH_2CH_2NHCH_2CH_2NHCH_3$
ICUC	$CH_3NH_2^+CH_2CH_2NHCH_2CH_2NH_2^+CH_3$
IAPN	$CH_3NH_2^+CH_2CH_2NH_2^+CH_2CH_2NH_2^+CH_3$
LINK	$CH_3NHCH_2CH_2NCH_2CH_2NHCH_3$ $\vert$ $CH_2CH_2OCH_3$
PLINK	$CH_3NH_2^+CH_2CH_2NCH_2CH_2NH_2^+CH_3$ $\vert$ $CH_2CH_2OCH_3$

We upload all the compounds in Table S1 to the CGenFF program (version 0.9.7.1 beta) and obtain an initial estimate for the force-field parameters. The CGenFF program automatically assigns parameters to a compound by exploiting analogies to small molecules, and accompanies these parameters by an evaluation of their accuracy expressed through a "penalty score." For the compounds listed in Table S1, the resulting partial charges, bond and angle coefficients, as well as some of the dihedral parameters show acceptable penalty scores (below 50). The highest penalty score of the parameters that we accepted is 37. These parameters are adopted directly for simulation. For dihedral parameters with high penalty scores, we choose to further optimize them via quantum-mechanical (QM) relaxed potential-energy surface (PES) scans using Gaussian09.

The Force Field Toolkit (FFTK) implemented in VMD is used to fit the dihedral parameters by reproducing the QM PES using molecular mechanics (MM). The QM relaxed PES is obtained by performing constrained geometry optimization at fixed dihedral angles based on the global minimum structure of the compound at the MP2/6-31G(d) level. For each conformation (horizontal axes in Fig. S2), the dihedral being scanned is fixed at a specific angle while all the other atoms are allowed to relax to find their minimum-energy configurations. Only the dihedrals involving heavy atoms are scanned and due to the symmetry of the global minimum structure, for some compounds, such as ICPC and IECM, only half of the dihedrals involving heavy atoms need to be investigated. Each scanned dihedral angle is varied from –180° to 180° in steps of 10°. All the dihedral scans for each compound are combined in one figure with the lowest energy set to 0. Figure S2 shows comparisons of the PES calculated using Gaussian and our fitted curves obtained using FFTK. Generally, the fitted force field can reproduce the QM PES very well in the low-energy regions (below 3 kcal/mol), indicating that the force field can faithfully describe the intramolecular interactions in PEG-grafted PEI.

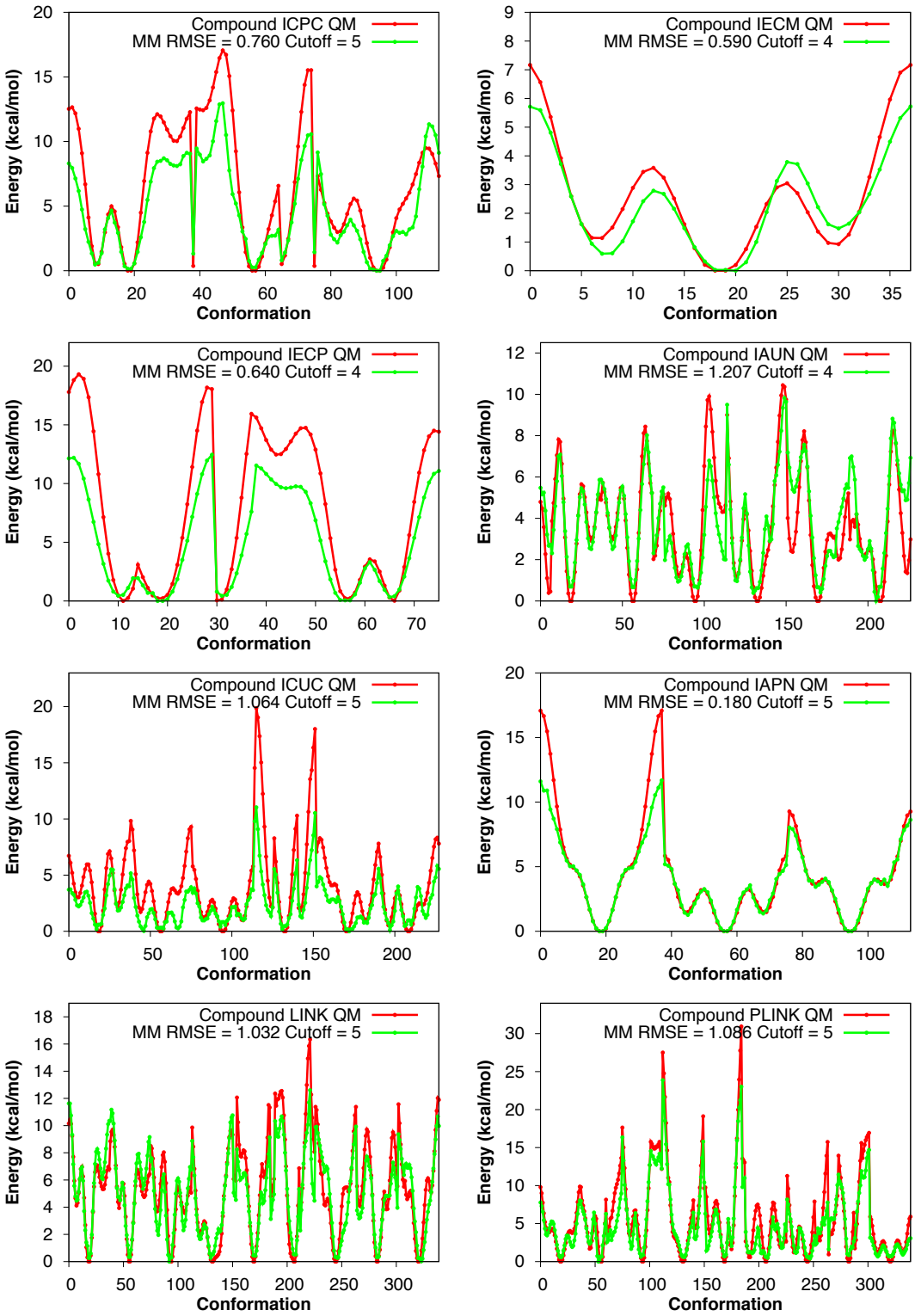


FIG. S2. Comparison of PES in QM calculations and MM calculations using fitted dihedral parameters for each compound. Cut-offs used in optimization and root-mean-square error (RMSE) are shown in the figure for each compound. For a detailed discussion see text.

To validate the developed force field, we compare the all-atom simulation results of non-grafted linear PEI with Ref. S1, where both 20-mer and 50-mer linear PEI chains were simulated using the AMBER force field and the 20-mer simulation results were validated against experimental results. The 50-mer PEI chains simulated using the AMBER force field were either alternately (i.e., 50%) protonated or fully protonated, whereas we employ PEI in which 82% of all nitrogen sites are protonated. Because of this difference in protonation, we only make a qualitative comparison. Figure S3 shows the end-to-end distance and radius of gyration over the last 200 ns of the production run (cf. main text, Sec. II A) and Table S2 presents a numerical comparison to the AMBER data. The average radius of gyration of our CHARMM-based PEI with 82% protonation is fully consistent with the AMBER data, showing a value slightly larger than the 50% protonated case and statistically equal to the fully protonated case. On the other hand, the average end-to-end distance of the 82% protonated PEI is closer to the AMBER data at 50% protonation than at 100% protonation, but this may be ascribed to the large fluctuations in our data.

TABLE S2: Comparison of PEI dimensions with AMBER simulation results of Ref. S1.

50-mer linear PEI	**82% protonated (this work)**	**Alternately protonated (50%)**	**Completely protonated**
End-to-end distance	6.58±2.24	6.77±1.34	8.45±1.23
Radius of gyration (R_g)	2.76±0.34	2.41±0.25	2.76±0.18

We also plot the pair distributions of water oxygen atoms around PEI nitrogen sites and PEI carbon sites in Fig. S3b. Comparing with Fig. 6 in Ref. S1, we note that the first peak position (~0.283 nm in Fig. S3b *vs.* ~0.29 nm in Ref. S1) and second peak position (~0.478 nm in Fig. S3b *vs.* ~0.48 nm in Ref. S1) for nitrogen sites are in excellent agreement. The two dominant peak heights for 82% protonated PEI lie in between those of the alternately protonated PEI and completely protonated PEI. Likewise, for the water distribution around carbon sites, the positions and heights of the first and second peak for 82% protonated PEI lie in between those of the alternately protonated PEI and completely protonated PEI.

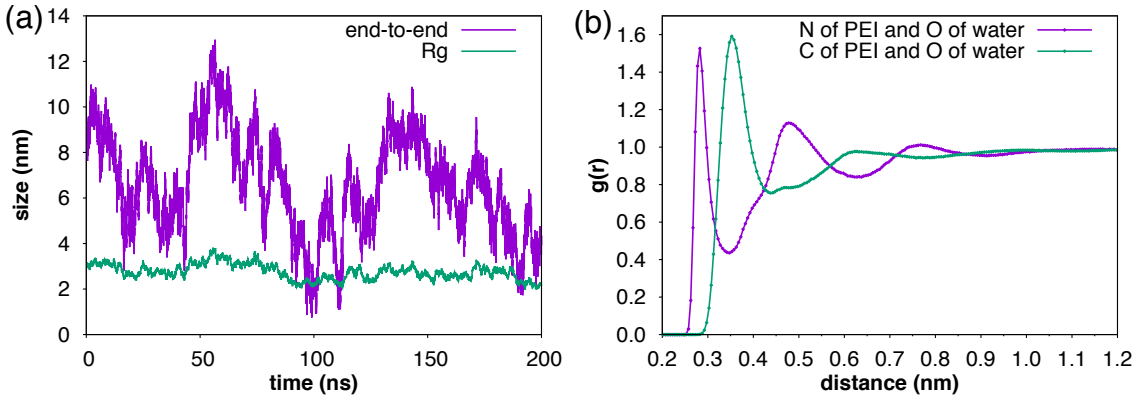


FIG. S3. (a) End-to-end distance and radius of gyration of linear PEI with 82% of the amine groups protonated. (b) Distribution function $g(r)$ of water oxygen atoms around PEI nitrogen sites (purple) and PEI carbon sites (green).

II. Dimensions of the simulation cell in AA simulations of single PEG-*g*-PEI in water and complexation of copolymers and siRNA

TABLE S3: Dimensions of simulation cell for the systems containing a single PEG-*g*-PEI chain.

Grafting density	Size of simulation box (Å^3)	Number of water molecules
0%	61×69×71	9297
2%	76×70×59	9714
4%	84×97×62	15668
8%	77×71×69	11605

TABLE S4: Dimensions of simulation cell for the systems involving complexation of copolymers and siRNA.

Grafting density	Size of simulation box (Å^3)	Number of water molecules
0%	104×104×116	38995
2%	105×97×126	39719
4%	98×100×153	45948
8%	110×122×140	58950

III. Parametrization of the coarse-grained force field for PEG-grafted PEI based on the MARTINI force field

Following the notation of the example structure in Fig. 2 in the main text, we can list all the bond, angle, and dihedral types in the CG model as shown in Table S5.

TABLE S5: Types of bonded interactions in CG model of PEG-grafted PEI. *

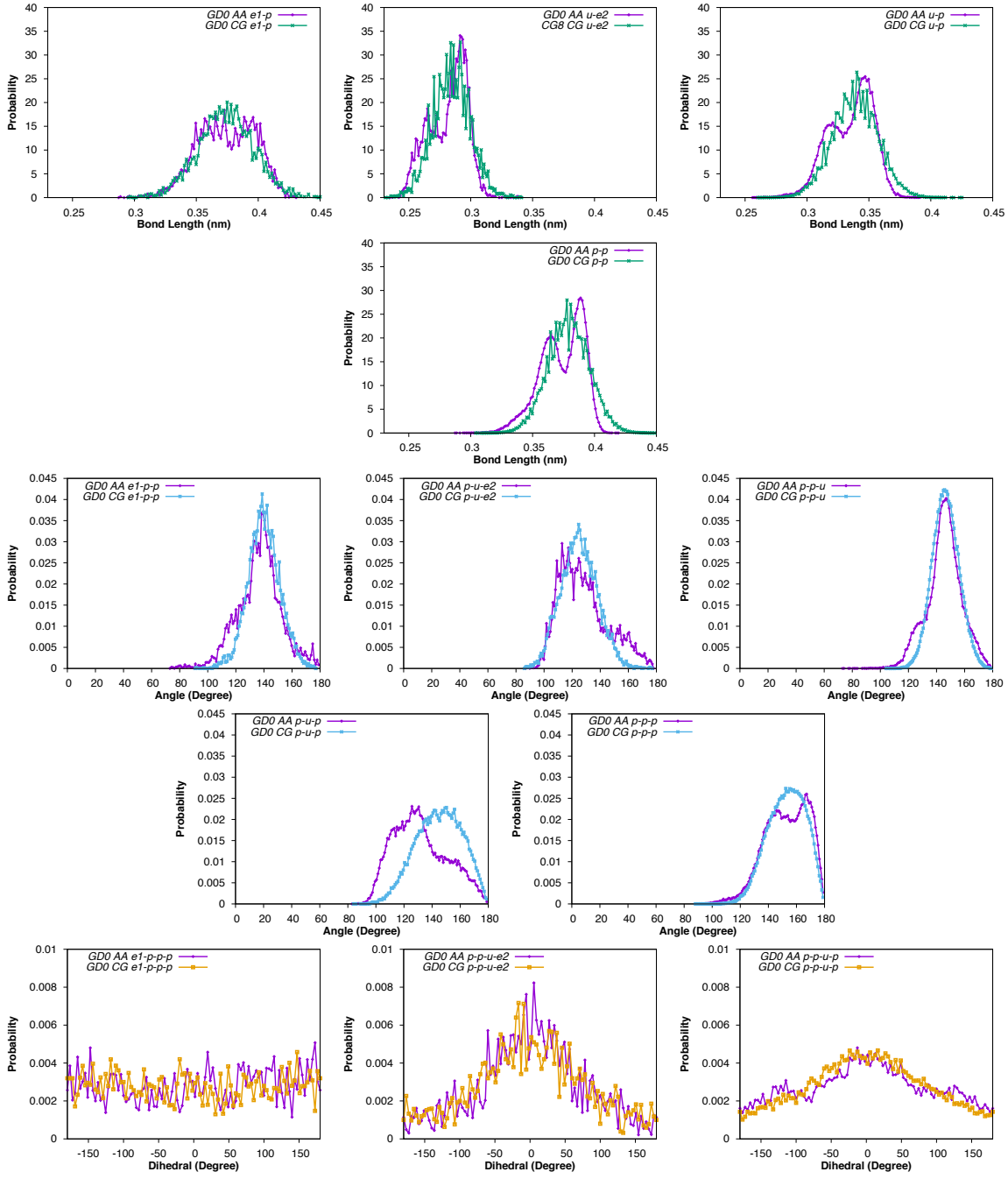
#	Bonds	Angles	Dihedrals
1	e1-p	e1-p-p	e1-p-p-p
2	u-e2	p-u-e2	p-p-u-e2
3	u-p	p-p-u	p-p-u-p
4	p-p	p-u-p	p-p-p-u
5	p-l	p-p-p	p-p-p-p
6	l-g	p-p-l	p-p-p-l
7	g-g	p-l-p	p-p-l-p
8		p-l-g	p-p-l-g
9		l-g-g	p-l-g-g
10		g-g-g	l-g-g-g
11			g-g-g-g

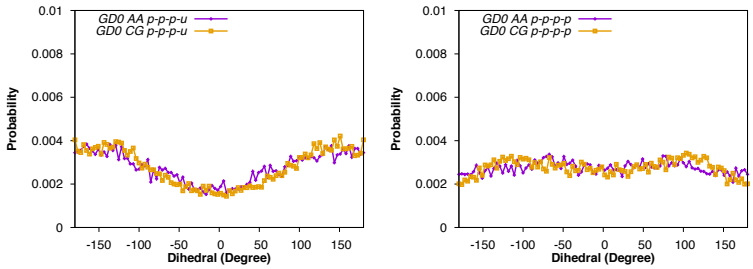
* Note: For grafting density 0% (linear PEI only), there is no interaction type involving bead l and bead g.

The last 200 ns of the AA simulations of isolated PEG-*g*-PEI in water serve as the input for the parametrization of the CG model. We convert the AA trajectory based upon the centers of mass of the atom groups corresponding to the CG beads as shown in Table I in the main text. From the converted trajectory of the CG beads, we plot all the bond, angle, and dihedral distributions. We find that corresponding distributions look similar for different grafting densities, except for some distributions involving the linking bead. Thus, we only need to parameterize based on one grafting density, for which we choose grafting density 2%. The distributions extracted from the AA simulations at different grafting densities are not shown, but all the parametrization results are shown in Fig. S4 for each grafting density.

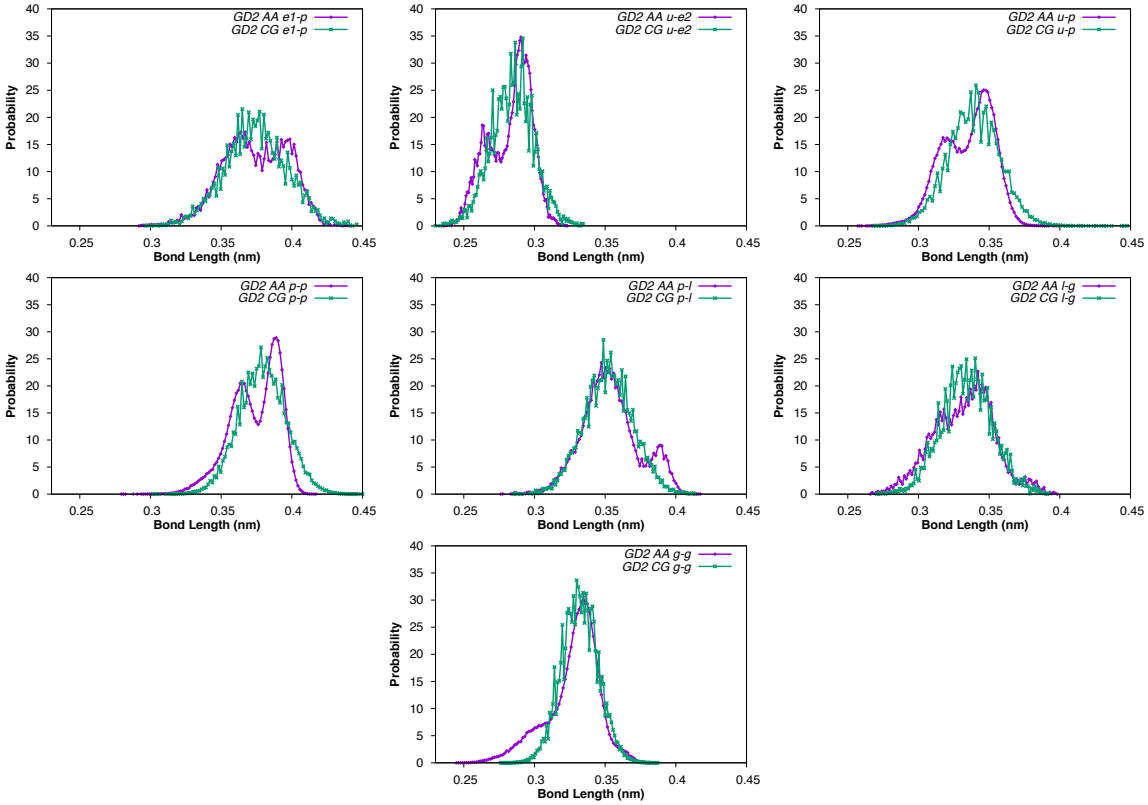
We use type 1 for all bond, angle, and dihedral interactions. Specifically, the stronger type-1 angle potentials are chosen to prevent beads from becoming co-linear, which would cause an instability in the simulations. We create the topology file for the CG simulation of PEG-grafted PEI using an initial estimate from the distribution plots extracted from the AA simulations. Then we perform a CG simulation using this topology file and compare the corresponding distributions between the CG simulation and the AA simulation. Based upon this comparison, specifically of the positions and widths of peaks in the distributions, we adjust the parameters and repeat the CG simulation. After several iterations of parametrization, we are able to achieve a good

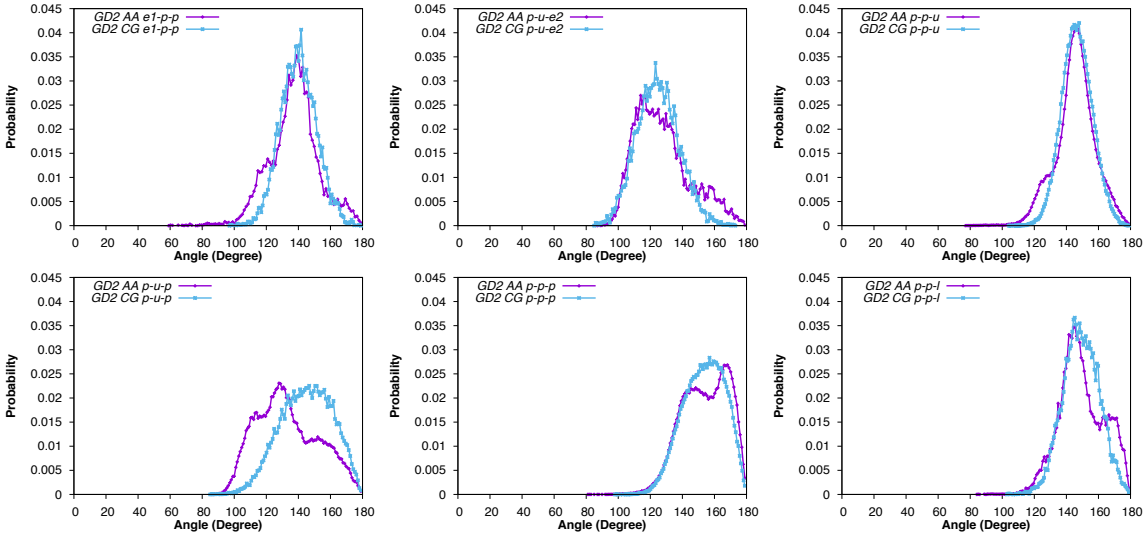
match. The comparison of the AA and CG distributions for all grafting densities is shown in Fig. S4.

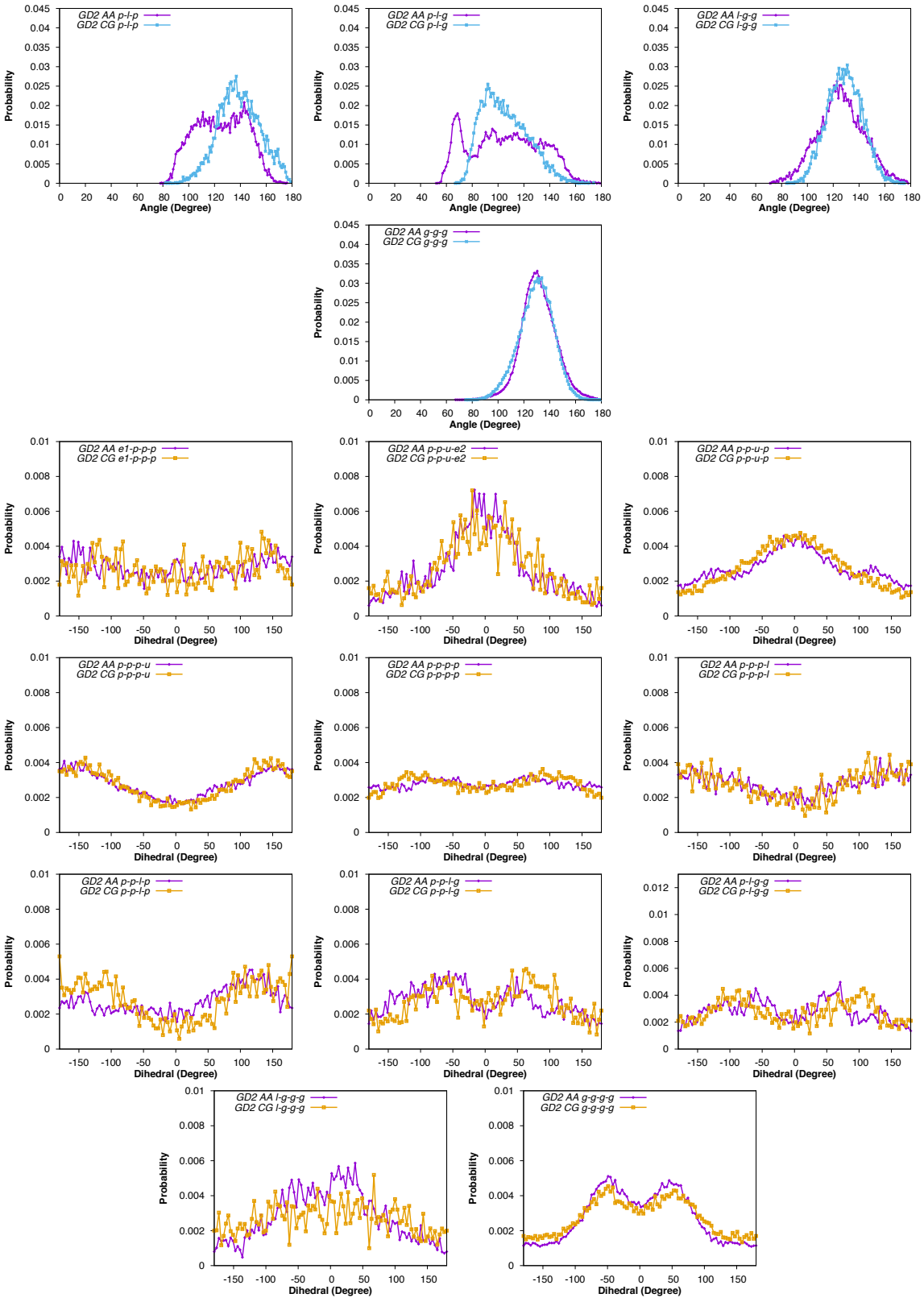




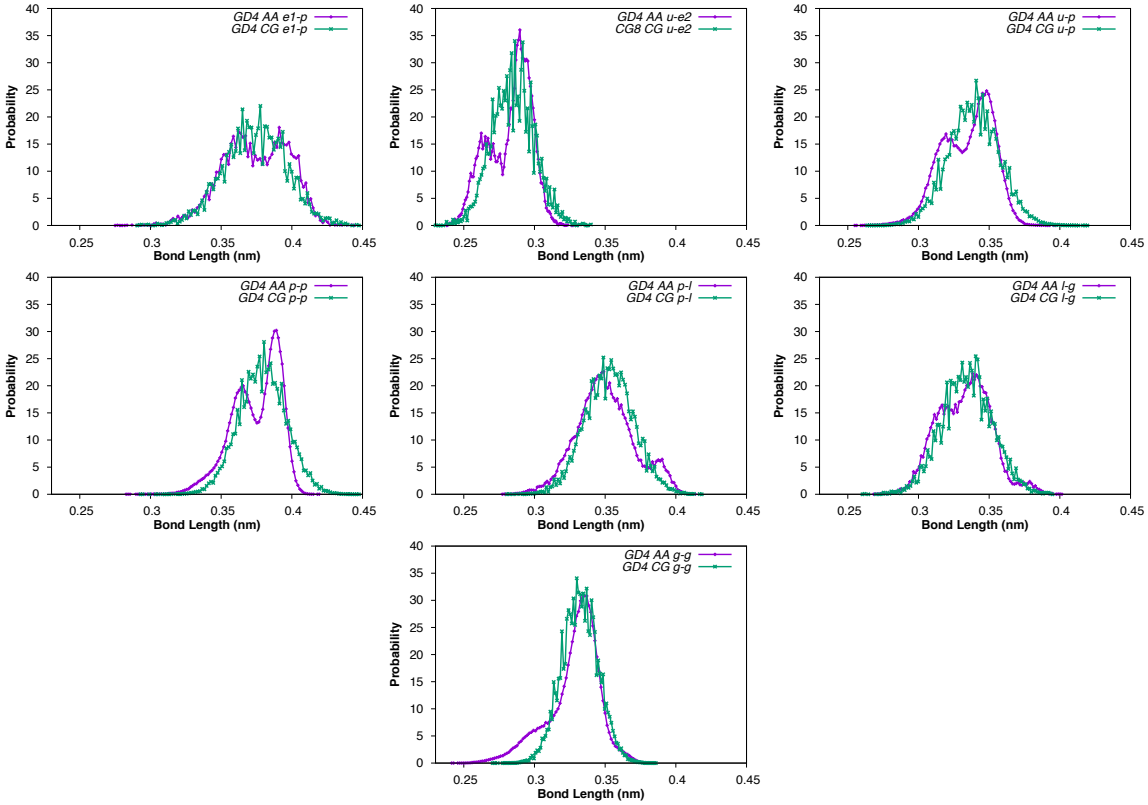
(a) Grafting density (GD) 0%.

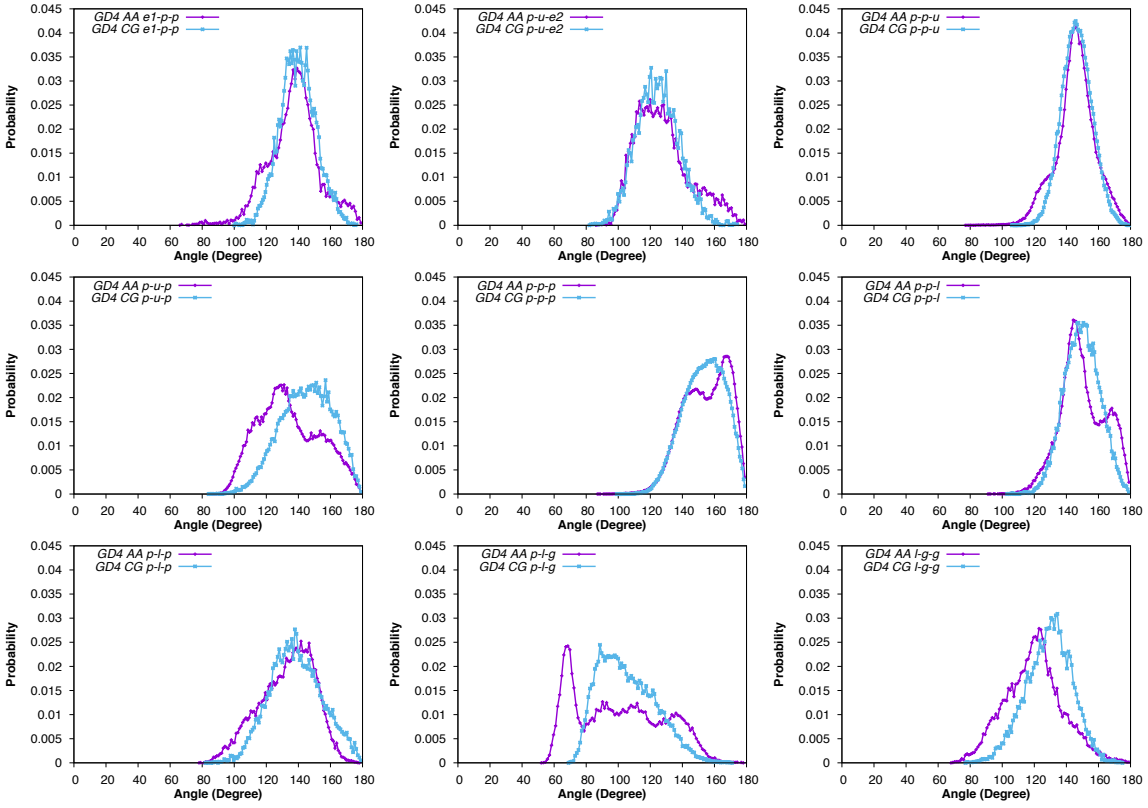


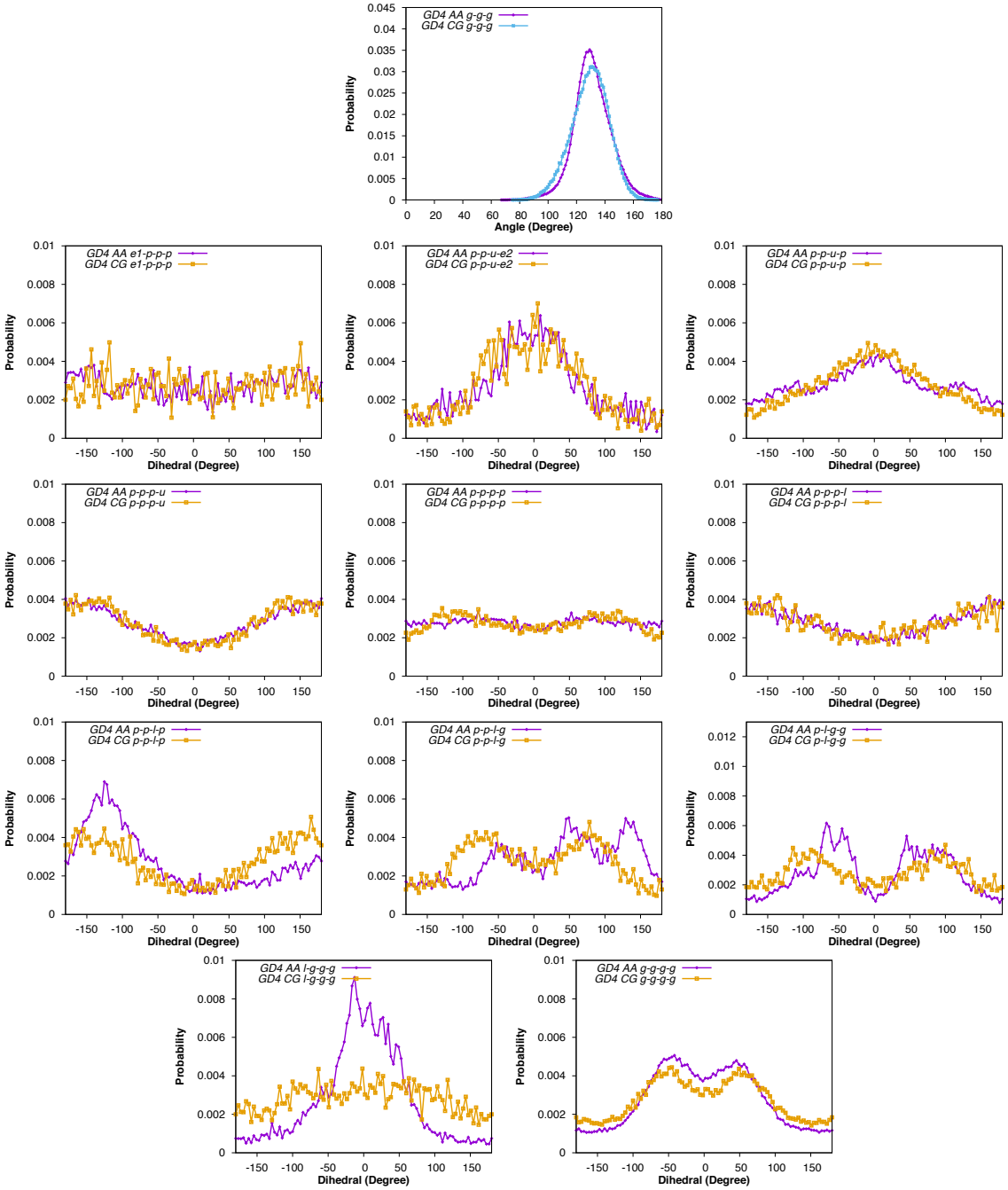




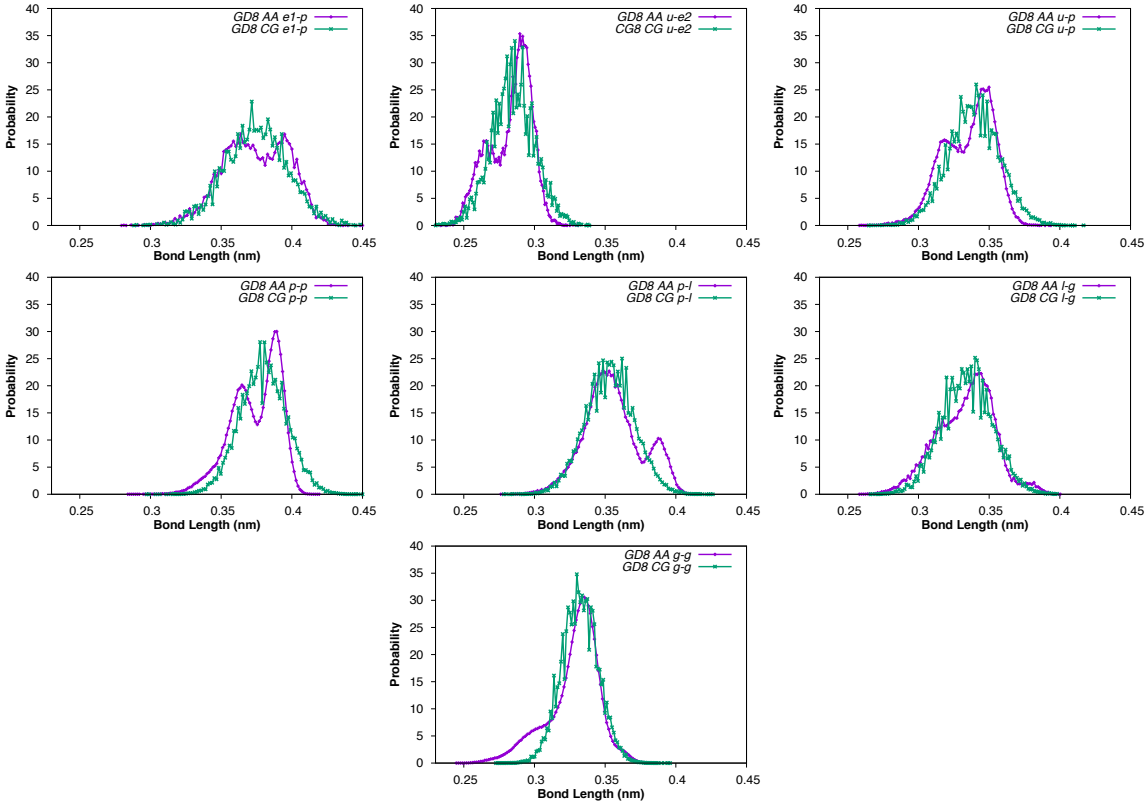
(b) Grafting density (GD) 2%.

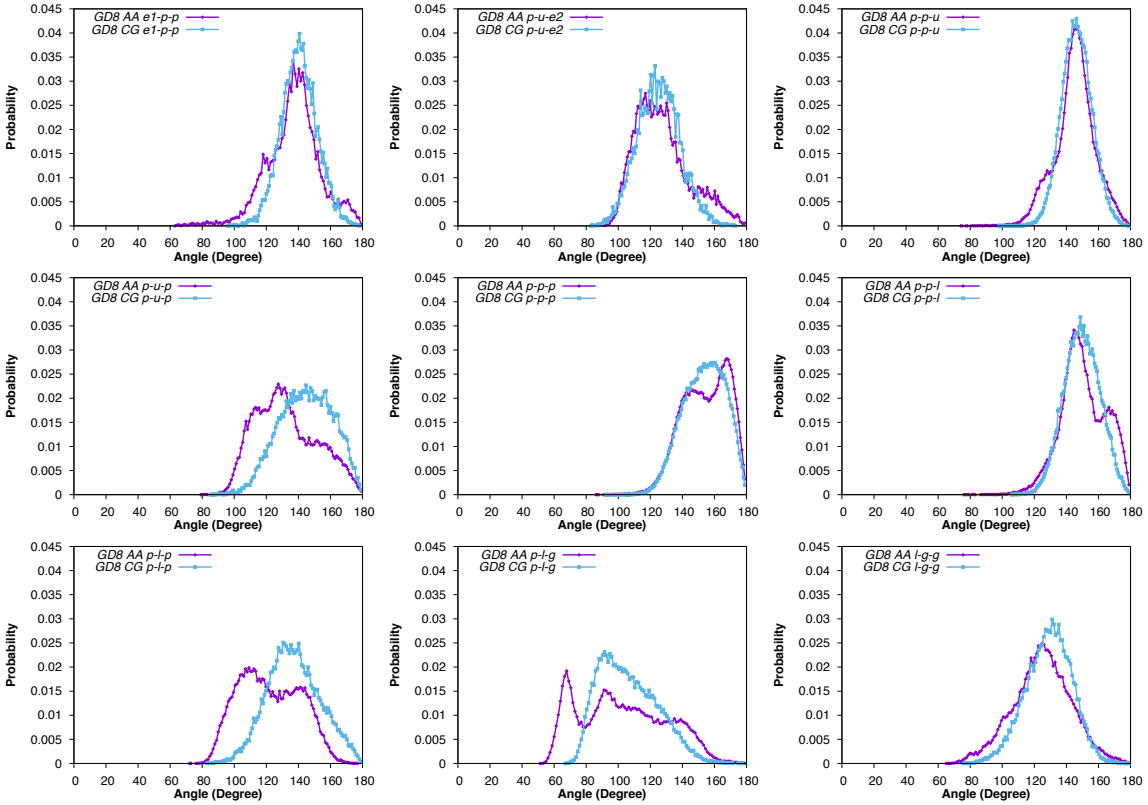

GD4 AA e1-p
GD4 CG e1-p
GD4 AA u-e2
CG8 CG u-e2
GD4 AA u-p
GD4 CG u-p
GD4 AA p-p
GD4 CG p-p
GD4 AA p-l
GD4 CG p-l
GD4 AA l-g
GD4 CG l-g
GD4 AA g-g
GD4 CG g-g
Probability
Bond Length (nm)

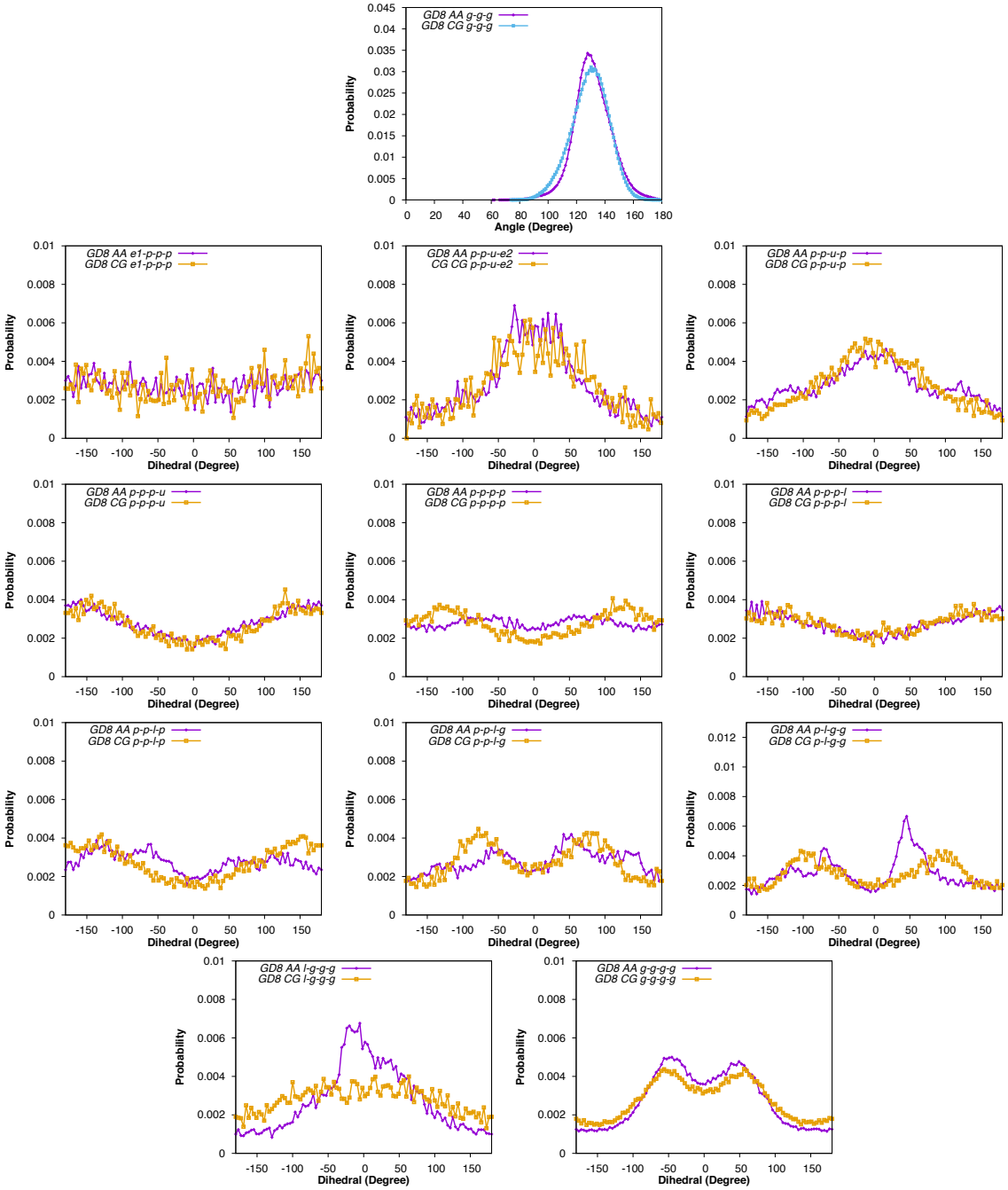

GD4 AA e1-p-p
GD4 CG e1-p-p
GD4 AA p-u-e2
GD4 CG p-u-e2
GD4 AA p-p-u
GD4 CG p-p-u
GD4 AA p-u-p
GD4 CG p-u-p
GD4 AA p-p-p
GD4 CG p-p-p
GD4 AA p-p-l
GD4 CG p-p-l
GD4 AA p-l-p
GD4 CG p-l-p
GD4 AA p-l-g
GD4 CG p-l-g
GD4 AA l-g-g
GD4 CG l-g-g
Probability
Angle (Degree)



(c) Grafting density (GD) 4%.


GD8 AA e1-p
GD8 CG e1-p
GD8 AA u-e2
CG8 CG u-e2
GD8 AA u-p
GD8 CG u-p
GD8 AA p-p
GD8 CG p-p
GD8 AA p-l
GD8 CG p-l
GD8 AA l-g
GD8 CG l-g
GD8 AA g-g
GD8 CG g-g
Probability
Bond Length (nm)


GD8 AA e1-p-p
GD8 CG e1-p-p
GD8 AA p-u-e2
GD8 CG p-u-e2
GD8 AA p-p-u
GD8 CG p-p-u
GD8 AA p-u-p
GD8 CG p-u-p
GD8 AA p-p-p
GD8 CG p-p-p
GD8 AA p-p-l
GD8 CG p-p-l
GD8 AA p-l-p
GD8 CG p-l-p
GD8 AA p-l-g
GD8 CG p-l-g
GD8 AA l-g-g
GD8 CG l-g-g
Probability
Angle (Degree)



(d) Grafting density (GD) 8%.

FIG. S4. Comparison of AA and CG bond, angle, and dihedral distributions for grafting density (a) 0%, (b) 2%, (c) 4%, (d) 8%. Green lines are CG bond distributions. Blue lines are CG angle distributions. Orange lines are CG dihedral distributions.

As illustrated in Fig. S4, most of the distributions match very well. However, in some cases, such as the p-u-p angle distribution at grafting density 0%, further tuning of the parameters cannot improve the matching of the peak positions. In some other cases, the AA distributions are bimodal. To investigate the origin of this bimodal distribution, we

examine the specific case of the u-p bond-length distribution at grafting density 2%. We plot the bond distributions of all bead pairs with u-p type in the AA simulation separately and find that each u-p bond shows this bimodal distribution (data not shown), implying that this bimodality is an intrinsic property of the individual u-p bonds rather than the combination of two unimodal distributions of all the u-p bonds. To clarify this, we specifically focus on the bond distribution of the bead pair 6-7 (p-u) shown in Fig. S5a. Fig. S5b shows the CG bond length as a function of dihedrals 1 and 2. There are eight preferred combinations of dihedral angles. Half of them show the preferred bond length around 0.32 nm and the other half show the preferred bond length of 0.36 nm, corresponding to the two peaks in the bond distribution. Thus, the bond-length distribution, which is a combination of bond, angle, and dihedral motion, calculated from AA simulations can display a bimodal character that reflects the nature of the dihedral potential energy surface. Since such details are difficult to reproduce using CG potentials, we only try to approximate the bimodal distribution using a unimodal distribution.

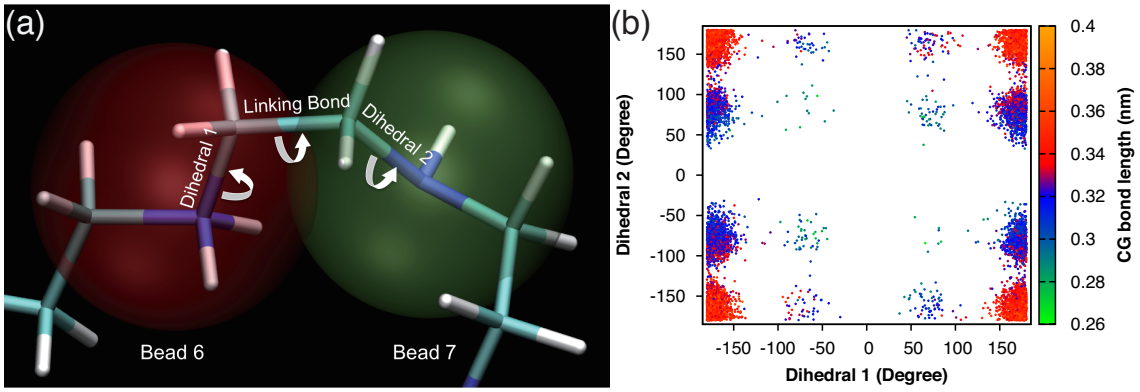


FIG. S5. (a) AA structure and corresponding CG beads of bead 6 (red) and bead 7 (green). (b) CG bond length distribution as a function of dihedrals 1 and 2.

To further validate the CG force field, we plot the radius of gyration of PEG-grafted PEI in the AA simulations and CG simulations at different grafting densities (Fig. S6).

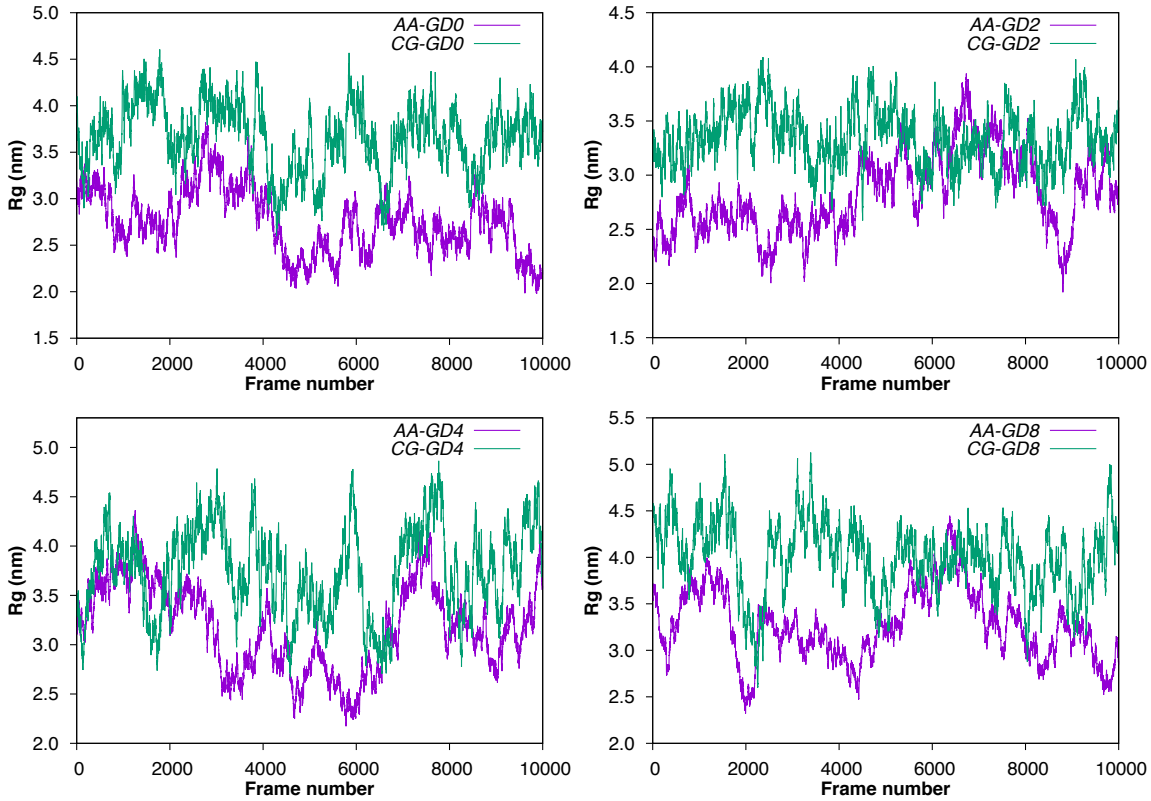


FIG. S6. Comparison of the radius of gyration (R_g) of PEG-grafted PEI in AA and CG simulations. Note: horizontal axis only represents the frame number in the AA or CG simulation; the time step and total simulation time are different.

As shown in Fig. S6, R_g in the CG simulations is usually slightly larger than in the AA simulations. While exploring the role of the water model in this, we found that the size of the polycation in our model with polarizable water matches better to the AA simulations better than the data using the standard CG water model. We tested parametrization of the CG force field for PEG-grafted PEI using the standard water model. Although we can still achieve a good match of bond, angle, and dihedral distributions, the shape of the polymer is not correct. The standard water model overestimates the attraction of the ions to the charged beads. Therefore, after being neutralized by ions, those charged beads do not repel each other any more, which results in the configurations shown in Fig. S7, which have much smaller radii of gyration.

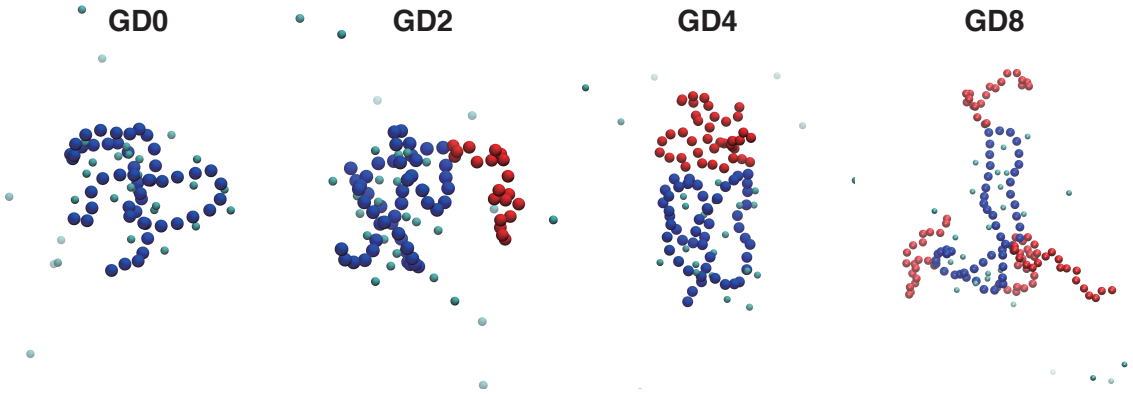


FIG. S7. Configurations of PEG-grafted PEI simulated with the standard MARTINI water model. Color code: blue–PEI, red–PEG, cyan–ions.

III. Network choice of siRNA model

In this work, we use MARTINI DNA in the canonical A form to represent the CG siRNA. The MARTINI model includes an elastic network to be used with the bonded terms in ds-DNA simulations. There are two different elastic network schemes: a "soft" and a "stiff" elastic network. We first tested the soft network in a simulation of isolated siRNA in water. However, we observed that the soft network scheme cannot maintain the intrinsic structure very well. We increased the force constant in the soft elastic network from the original value of 13 kJ mol^{-1} nm^{-2} to 30, 50 and 100 kJ mol^{-1} nm^{-2}. However, as illlustrated Fig. S8, the A-helices deformed significantly during the simulation for all three force constants.

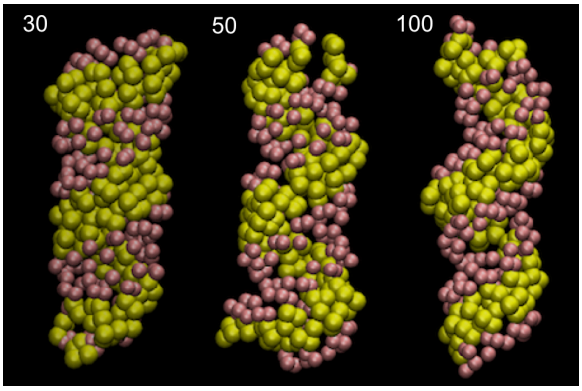


FIG. S8. Configurations of siRNA using the soft elastic network with force constants of 30, 50, and 100 kJ mol^{-1} nm^{-2}.

To avoid such siRNA deformation, we opted to use the stiff elastic network. As shown in Fig. 3 and Fig. 7 of the main text, the shape of the siRNA is stable during the simulation. However, the length reduction along the principal axis observed in the AA simulations cannot be reproduced. In an attempt to recover this length variation, we decreased the force constant in the stiff elastic network model from 500 kJ mol^{-1} nm^{-2} to 300 kJ mol^{-1} nm^{-2} at intervals of 50 kJ mol^{-1} nm^{-2}, and simulate isolated siRNA in water as well as siRNA binding with polymers at different grafting densities. The end-to-end distance along the principal axis as a function of simulation time is shown in Fig. S9 for each case.

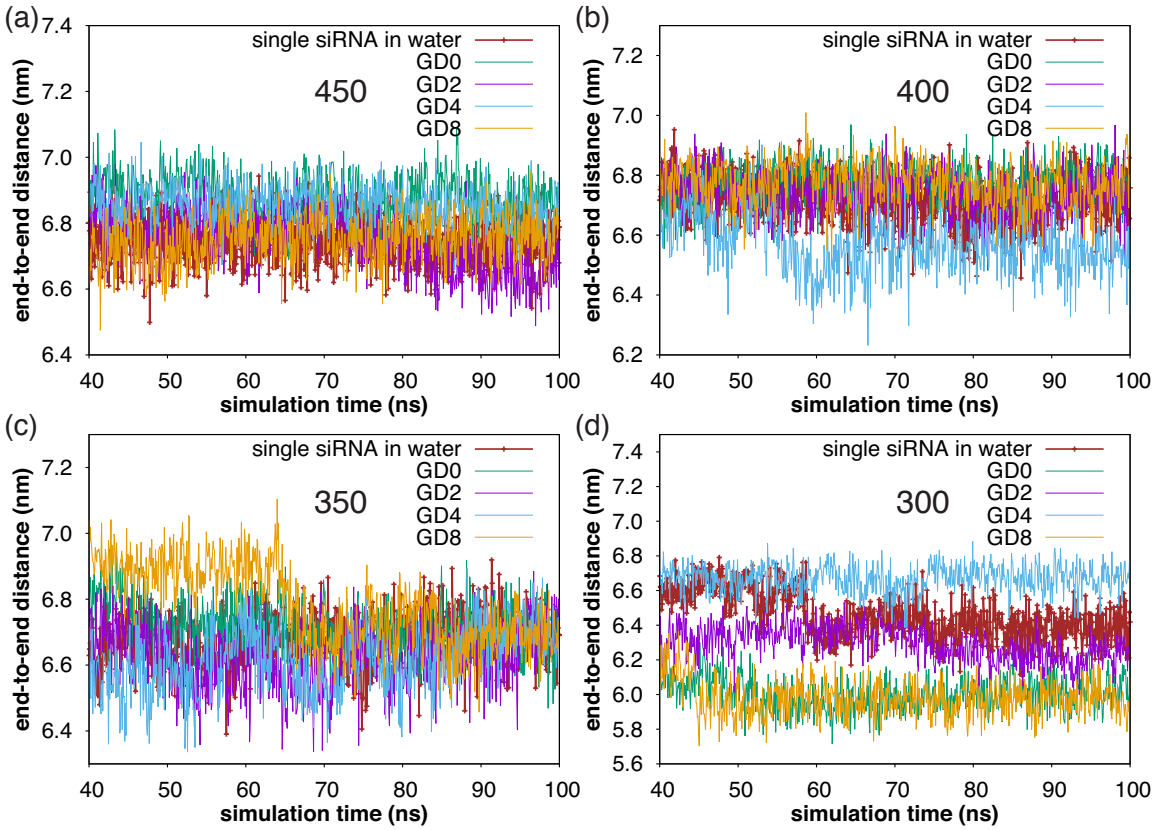


FIG. S9. The end-to-end distance along the principal axis of siRNA in CG simulations with stiff elastic network and force constant of (a) 450 kJ mol^{-1} nm^{-2} (b) 400 kJ mol^{-1} nm^{-2} (c) 350 kJ mol^{-1} nm^{-2} (d) 300 kJ mol^{-1} nm^{-2}.

Clearly, decreasing the force constant of the stiff elastic network cannot reproduce the length reduction along the principal axis upon binding. Instead, the weakened constraints may lead to instability in the siRNA structure, as can be seen, e.g., in Fig. S9b for grafting density 4% (GD4) and Fig. S9c for grafting density 8% (GD8). For the smallest force constant investigated, 300 kJ mol^{-1} nm^{-2} (Fig. S9d), even isolated siRNA in water can be unstable, with a sudden decrease of the minor groove length at around 60 ns.

IV. Radial distribution function (RDF) of water molecules around siRNA and polymers

The RDF plots in Fig. S10 illustrate our choices for the hydration shell limit in AA and CG simulations as 0.3 nm and 0.6 nm, respectively, as these separations permit inclusion of the first peak of the distribution function, which corresponds to direct hydrogen bonding between the macromolecule and water molecules.

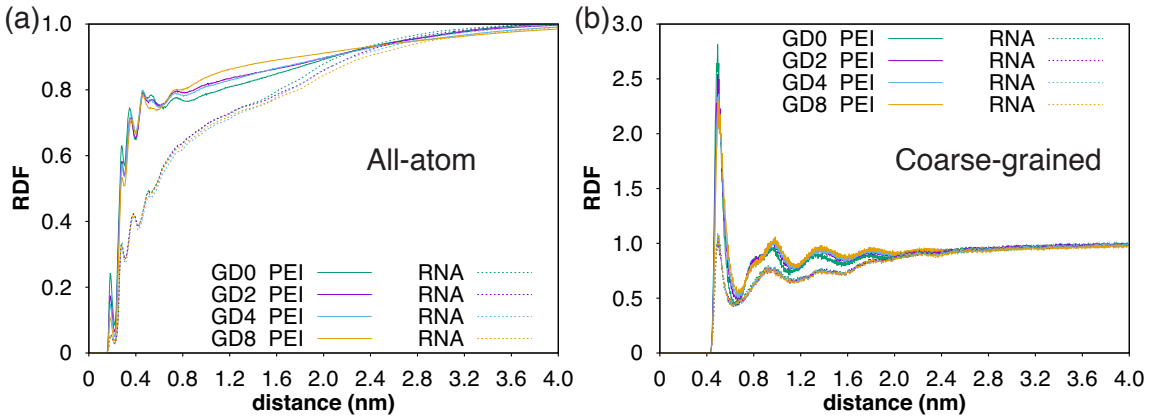


FIG. S10. RDF of water molecules around siRNA and polymers in (a) AA simulations and (b) CG simulations.

V. Bonded coarse-grained force field parameters for PEG-grafted PEI

Note that the following MARTINI force field for PEG-grafted PEI needs to be used in conjunction with the polarizable water model.

TABLE S6: Bond and angle coarse-grained force field parameters for PEG-grafted PEI

Beads	Type	Position of minimum (nm or degrees)	Force constant (kJ mol^{-1} nm^{-2}, kJ mol^{-1})
e1-p	1	0.381	5700
u-e2	1	0.285	12000
u-p	1	0.340	8200
p-p	1	0.395	9700
p-l	1	0.353	9000
l-g	1	0.340	8400
g-g	1	0.335	17000
e1-p-p	1	145.0	45.0
p-u-e2	1	80.0	15.0
p-p-p	1	160.0	20.0
p-p-u	1	148.0	65.0
p-u-p	1	20.0	2.0
p-p-l	1	155.0	35.0
p-l-p	1	100.0	0.5
p-l-g	1	100.0	0.5
l-g-g	1	125.0	35.0
g-g-g	1	129.0	45.0

TABLE S7: Dihedral coarse-grained force field parameters for PEG-grafted PEI

Beads	Type	Multiplicity	Phase ϕ_s (degrees)	Force constant (kJ mol^{-1})
e1-p-p-p	1	1	0	0
p-p-u-e2	1	1	180	3.5
p-p-p-u	1	1	0	0.2
p-p-u-p	1	1	180	3.5
p-p-p-p	1	1	180	0.3
p-p-p-l	1	1	0	0
p-l-g-g	1	1	180	2.0
l-g-g-g	1	1	180	2.2
p-p-l-p	1	1 2	180 0	0.2 0.1
p-p-l-g	1	1 2	180 0	0.3 0.5
g-g-g-g	1	1 2 3 4	180 0 0 0	1.96 0.18 0.33 0.12

Reference

(1) Choudhury, C. K.; Roy, S. *Soft Matter* **2013**, *9*, 2269–2281.