scientific reports

Check for updates

OPEN

Martini compatible coarse-grained model of polyethylenimine for pulmonary gene delivery

Graham Lunt[1,3], Niloofar Hashemi[1,3], Subhamoy Mahajan[1,2] & Tian Tang[1]✉

Pulmonary gene delivery has demonstrated high specificity for respiratory diseases, offering great control on dosage of therapeutics and side effects. On the other hand, intrinsic barriers in pulmonary systems impose new challenges such as crossing the pulmonary surfactant and evading mucus entrapment. Differences in hydrophobicity of plasma membrane and pulmonary surfactant require different chemistries of gene carriers to improve efficacy. Large-scale coarse-grained (CG) molecular dynamics simulations would facilitate the screening of gene carriers and understanding of the molecular mechanisms involved in pulmonary delivery. Among non-viral carriers, polyethyleneimine (PEI) has been a promising candidate that can be synthesized with various molecular weight, degree of branching, and functionalization. In this work, CG models are developed for PEI and its lipid-functionalized form, within the Martini framework, to provide a platform for exploring structure-function relationships of PEI-based pulmonary delivery systems. Special attention is focused on parameterizing the non-bonded interactions associated with CG PEI, to ensure compatibility with Martini proteins, short interfering RNA, and phospholipids that are essential components in pulmonary gene delivery. The non-bonded parameters are validated by comparing all-atom (AA) and CG potential of mean force (PMF) curves, where the root-mean-square deviations between the AA and CG PMF curves are shown to be comparable to or smaller than those reported in Martini literature.

Keywords Coarse-graining, Polyethyleneimine, Lipid substitution, Pulmonary gene delivery, Martini force field, Potential of mean force

Pulmonary diseases pose a significant global health burden, driving the need for advanced therapeutics[1]. Pulmonary gene delivery, which involves introducing genetic material directly into the lungs, could revolutionize treatment for respiratory diseases[2]. Pulmonary systems impose unique challenges such as the complex structure of airways, potential mucociliary clearance[3], the barrier of pulmonary surfactant, and the threat of neutralization by macrophages[4,5]. Mucociliary clearance involves entrapment of particles by negatively charged mucins (glycoproteins) and their removal by upward motion of cilia beneath the mucus layer, complicating drug/gene delivery[3,6]. The physical barrier of pulmonary surfactant is composed of diverse phospholipids, such as 2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) and 1,2-Dipalmitoyl-sn-glycero-3-phosphoglycerol (DPPG). These phospholipids are organized in mono- and multilayer structures at the interface between air and alveolar fluid, and these different structures cyclically transition between one another during respiratory cycles[7]. The presence of surfactant proteins (SP), such as SP-A, SP-B, SP-C and SP-D, are responsible for organizing the surfactants as well as providing protection against foreign materials[3,8]. Finally, macrophages in alveolar fluid can clear foreign pathogens and particles, such as therapeutic agents, using apoptotic pathways[9]. Further details on barriers to pulmonary delivery can be found in the extensive reviews by Whitsett et al.[3]. , de Souza Carvalho et al.[10], Murgia et al.[11], and Baliga et al.[12].

Pulmonary gene delivery specifically targets lungs, offering high efficacy[13] and reduced systemic side effects[14]. Recent advancements in non-viral delivery has improved the precision of transporting nucleic acids to specific pulmonary targets, such as the vascular endothelial cells for treating pulmonary arterial hypertension, or airway epithelial cells affected by cystic fibrosis[15]. Various gene delivery carriers, such as cationic liposomes and polymers, have been explored for pulmonary delivery[16]. Cationic liposomes typically consist of a hydrophobic group for bilayer stabilization and a cationic group for condensing DNA or aggregating RNA[17]. In pulmonary delivery, the interaction of pulmonary surfactant with the liposomes affects their stability, uptake efficiency and biocompatibility[18]. Liposomes can also be modified with ligands such as monoclonal antibodies[19], glycoprotein[20],

[1]Department of Mechanical Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada. [2]Department of Mechanical Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA. [3]Graham Lunt and Niloofar Hashemi: Contributed equally to this work. ✉email: tian.tang@ualberta.ca

and peptides[21] for improving specificity. Cationic polymers, such as poly-l-lysine (PLL) and polyethyleneimine (PEI), can condense DNA or aggregate short interfering RNA (siRNA), protecting them from degradation[22]. These carriers are promising for pulmonary gene therapy due to their high transfection efficiency and versatility[23]. PEI-based gene delivery systems, known for their high transfection rates in the lungs[24], have the potential to treat pulmonary diseases such as cystic fibrosis[25], pulmonary hypertension[26], and lung cancer[27]. However, native PEIs characterized by their wide range of molecular weights presents inherent challenges in gene delivery, including cytotoxicity and limited target specificity[27–31]. Molecular functionalization techniques such as lipid substitution, PEGylation, and ligand conjugation have been demonstrated to be advantageous in addressing the limitations of native PEIs, while simultaneously improving transfection efficiency and biodegradability[32–37]. For instance, ligands can be specifically selected to interact with proteins unique to the target cell[34,38,39]. PEIs with attentive functionalization therefore offer significant advantages for pulmonary delivery, including enhanced transfection efficiency, targeted delivery to specific lung cells, reduced cytotoxicity, improved biodegradability, and enhanced stability[40,41].

Fundamental understanding of mechanisms associated with PEI-based gene delivery can enable systematic and fast design of the carriers including their molecular functionalization. All-atom (AA)[42–44] and coarse-grained (CG)[45–48] molecular dynamics (MD) simulations serve as valuable tools in exploring mechanisms at angstrom to nanometer scales that are typically inaccessible in experiments. These simulations have provided valuable insights into the structural characteristics and dynamics of PEI-based gene delivery systems, shedding light on the interactions between PEI and nucleic acids[49–52], the formation of PEI-nucleic acid nanoparticle (NP)s[46], and the behavior of the NPs at different stages of the in-vitro gene delivery pathway[45,53,54]. To the best of our knowledge, MD simulations exploring pulmonary systems have not been studied for PEI-based NPs. However, pulmonary surfactants have been studied in other contexts. For instance, Duncan et al.[55] studied phase transitions in lung surfactants, and Baoukina and coworkers[56,57] explored their monolayer-bilayer transformations. Investigations concerning the impact of NP shape[58], surface charge density[59], and lipid coating[60] on their translocation through pulmonary surfactant were studied for NPs with simple geometry. A comprehensive understanding of how NPs interact with key components of the surfactant is necessary for the rational design of NP delivery systems in pulmonary applications. Compared with AA simulations, the use of CG MD is advantageous for investigating PEI NPs in pulmonary systems, as it allows for efficient exploration of the interactions between NPs and the complex pulmonary environment, at more realistic length- and timescales.

The goal of this work is to develop a CG model for PEI and its lipid-substituted form, for the simulation of PEI NPs in pulmonary gene delivery. We chose the Martini force field, due to the existence of many parameterized biomolecules in Martini such as nucleic acids and proteins, as well as its wide application in simulating NPs in complex environments[61]. Although Martini CG models have been successfully developed for linear and branched PEIs[62] including our prior work[47], a notable gap remains: there is currently no Martini model available for PEIs with molecular functionalization such as lipid substitution[47]. In addition, the compatibility of existing Martini models with CG PEIs, especially after lipid substitution, is not guaranteed and requires meticulous parameterization and validation. This consideration is essential for ensuring the accuracy of CG simulations in investigating the interactions between PEI NPs and the pulmonary environment. For this purpose, Martini-compatible models are developed for a native PEI and its counterpart with lipid (caprylic acid (CA)) substitution. The interaction of PEI and lipid-substituted PEI with nucleic acid (an siRNA) and key components of pulmonary surfactant (DPPC, DPPG, SP-B) are reparametrized and validated against AA simulations. SP-B is chosen over other SPs as it has been observed to modify properties of pulmonary surfactant that are important to gene delivery such as permeability[63] and folding[56]. CA is chosen as the lipid substitution due to its efficacy and the strong correlation between the degree of substitution and cellular uptake[64].

Methods

AA and CG models

A 2 kDa branched PEI[65] (hereafter referred to as "PEI" for simplicity) with a protonation ratio of 40% to simulate PEI at a pH of 5 to 6[66] was used as the base model (AA structure shown in Fig. 1). The PEI with CA substitution, hereafter referred to as "PEI-CA", was prepared by grafting a CA to one or three of the branches on PEI (substitution sites numbered 1, 2 and 3 in Fig. 1). When grafting CA to only one branch, substitution site 1 was used. Force field parameters for the AA PEI models, developed based on CGenFF[67], were directly adopted from Sun et al.[65].

The CG models for both PEI and PEI-CA were created based on their AA counterparts. The CG PEI model was previously developed[47] by mapping each ethylenimine monomer "CCN" to P1 (for groups with uncharged nitrogen) or Qd (for groups with charged nitrogen) beads, with the bead types determined based on hydration and partitioning free energies of ethylamine[68]. The CG model produced accurate interactions with salts, water, and DNA[47]. The bead mapping of the "CCN" groups is visualized by the red, blue, and black ellipses in Fig. 1. For the CG PEI-CA, atoms on the CA tail were grouped and mapped into the "(O=)CCCC$_{0.5}$" and "C$_{0.5}$CCCC" beads, named formaldehyde and alkane beads respectively (green and orange ellipses in Fig. 1), following a 4.5:1 mapping. Here, C$_{0.5}$ represents a carbon atom that is shared equally between two beads. Consistent with Marrink et al.[68], the bead types were determined based on hydration and partitioning free energies[64]. Specifically, the free energy of "(O=)CCCC$_{0.5}$" was calculated from the average free energy of "(O=)CCC" and "(O=)CCCC", and the free energy of "C$_{0.5}$CCCC" was considered to be the average free energy of "CCCC" and "CCCCC". Based on the free energies, Na and C1 bead types were respectively assigned to the formaldehyde and alkane beads.

The parameters associated with bonds, angles, and dihedrals of CG PEI were determined in a previous work and used without any alterations[47,69]. Additional bonded interactions required for CG PEI-CA were determined using an automated bottom-up scheme[69] (details in "Bonded interactions"). The non-bonded interactions of CG PEI and PEI-CA with Martini polarizable water[70] and salt were kept unaltered, while their interactions with CG

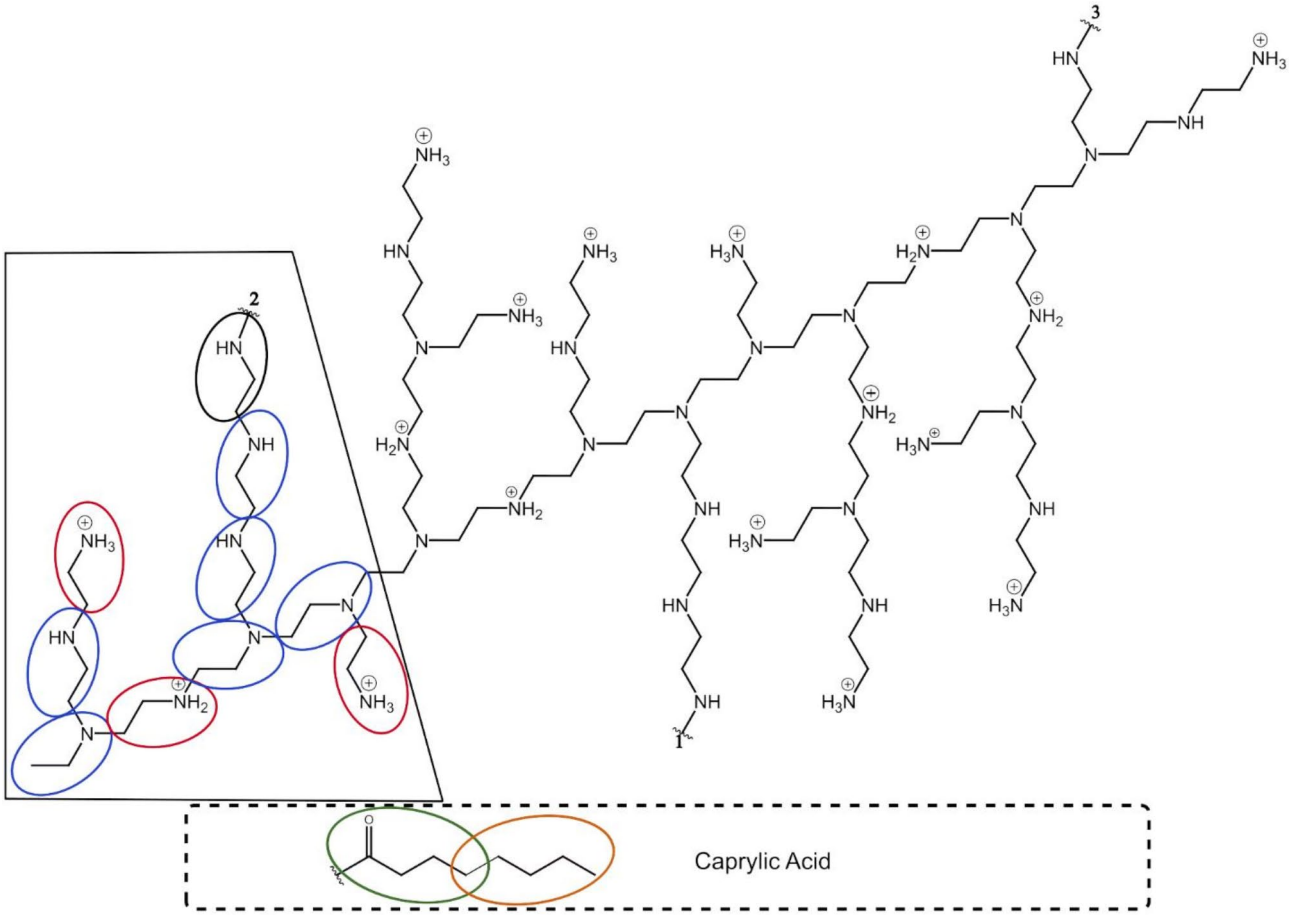

Fig. 1. Molecular structure, protonation sites, and lipid substitution sites of PEI (2 kDa), with the portion enclosed by the black box depicting a smaller PEI (500 Da) used for parameterization of non-bonded interactions. Blue, red, green and orange ellipses show the mapping of atom groups to P1, Qd, Na, and C1 beads, respectively. The black ellipse highlights a substitution site where CA can be grafted. With CA substitution, P1 is assigned to the black bead. In the absence of CA substitution, a proton is added to each substitution site, turning the black bead into Qd. An auxiliary uncharged PEI is also used in parametrization, for which the red (Qd) beads are replaced by blue (P1) beads. This structure was generated using ChemDraw 22.2.0.

phospholipids, SP-B, and siRNA were validated by comparing the CG potential of mean force (PMF) curves to their AA counterparts (details in "Non-bonded interactions"), and non-bonded parameters were reparametrized if necessary. As PMF calculations are costly, the smaller 500 Da subsection highlighted in Fig. 1 was used in these simulations. To facilitate the reparameterization, an uncharged PEI (hereafter referred to as “uncharged PEI”) that contained only P1 beads was introduced as an auxiliary model which allows P1 beads to be reparametrized independently of Qd beads. The uncharged PEI is obtained by removing a hydrogen from each of the three protonated amine groups of the 500 Da subsection in Fig. 1.

The AA and CG DPPC, DPPG, siRNA, and SP-B models were used for validating or reparametrizing their CG non-bonded interaction potentials with CG PEI and PEI-CA. The sources for these models along with the CG bead types are provided in Table 1. The AA models of DPPC and DPPG bilayers were generated using the CHARMM-GUI[71] membrane builder[72,73] and the corresponding CG models were created using the CHARMM-GUI Martini Maker[72,73]. The AA siRNA, sequence given in Table 1, is a smaller portion of a therapeutic siRNA that has been shown to treat lung cancer by inhibiting B-cell lymphoma-2 (BCL-2) expression[74]. The AA structure was generated in the B-form using NAB module of AMBERTOOLS[75]. The CG siRNA was created from the AA model using the Martinize-nucleotide python script[76]. Stiff elastic bonds with a cutoff of 1.0 nm and a force constant of 500 kJ mol^{-1} nm^{-2} were added to preserve the secondary structure of the CG siRNA. The SP-B is a highly hydrophobic protein that has been shown to produce large oligomeric structures[77], but for the purpose of parametrizing PEI only one dimer was used in this work to allow for realistic simulation time for the AA simulations. The AA SP-B created using homology modeling was generously provided by Olmeda et al.[77,78]. The Protonate 3D application within Molecular Operating Environment (MOE)[79] was used to predict the protonation state, charge the terminal protein residues, and determine the rotamer and tautomer states of protein residues in water at 300 K and pH of 6.9 in the absence of salts. A pH of 6.9 was chosen to match the pH of lung surfactants[80]. The CG SP-B was generated from the AA SP-B using the Martinize python script[81]. The CG

Molecule or system	Source of the AA model	Source of the CG model	Martini beads in the CG model
DPPC Bilayer	CHARMM-GUI[71] Membrane Builder [72,73]	CHARMM-GUI[71] Martini Maker[81]	Q0, Qa, Na, C1
DPPG Bilayer	CHARMM-GUI[71] Membrane Builder[72,73]	CHARMM-GUI[71] Martini Maker[81]	P4, Qa, Na, C1
siRNA Sense: 5'-UGUGGAUGA-3' Antisense: 3'-ACACCUACU-5'	Smaller portion of siRNA for silencing BCL-2 gene.[74] Initial structure using AMBERTOOLS NAB.[75]	Built from AA structure using Martenize-nucleotide python script[76]	SN0, SNda, TN0, TT2, TT3, Q0, TG2, TG3, TNa, TA2, TA3, TY2, TY3
SP-B Sequence of a monomer chain: LPYCWLCRALIKRIQAMIPK GALAVAVAQVCRVVPLVAG GICQCLAERYSVILLDTLLGR MLPQLVCRLVLRCS	Homology modeling by Olmeda et al.[77]	Built from AA structure using the Martenize python script[81]	C1, C2, C3, C5, SP1, SC4, SC4, P4, P5, N0, Na, Nd, Nda, Qa, Qd

Table 1. Description of AA and CG models of non-PEI molecules or systems.

topology maintained the cysteine bonds from the AA topology. Additionally, elastic bonds (lower cutoff: 0.5 nm, upper cutoff: 0.9 nm, force constant: 500 kJ mol^{-1} nm^{-2}) maintained the secondary structure of individual monomers. The CHARMM27 and Martini 2.1P force fields were respectively used for the AA and CG models of DPPC, DPPG, siRNA, and SP-B.

Bonded interactions

The bonded interactions in MARTINI 2.1P[81] include bond stretching, bond angle bending, and dihedral angle potentials. The bond stretching potential, V_b, is harmonic,

$$V_{\mathrm{b}}\left(r_{ij}\right)=\tfrac{1}{2}K_{ij}^{\mathrm{b}}\left(r_{ij}-r_{ij}^{\mathrm{eq}}\right)^{2} \tag{1}$$

where r_{ij} is the bond length, K_{ij}^{b} is the force constant, and r_{ij}^{eq} is the equilibrium bond length between beads i and j. The bond angle bending potential, V_{a}, is cosine-harmonic,

$$V_{\mathrm{a}}\left(\theta_{\ ijk}\right)=\tfrac{1}{2}K_{ijk}^{\mathrm{a}}\left[\cos\left(\theta_{\ ijk}\right)-\cos\left(\theta_{\ ijk}^{\ \mathrm{eq}}\right)\right]^{2} \tag{2}$$

where $\theta_{\ ijk}$ is the bond angle, K_{ijk}^{a} is the force constant, and $\theta_{\ ijk}^{\ \mathrm{eq}}$ is the equilibrium bond angle between beads i, j, and k. The dihedral angle potential is given by a sum of $2W$ periodic functions,

$$V_{\mathrm{d}}\left(\phi_{ijkl}\right)=\sum_{w=1}^{2W}K_{ijkl}^{\mathrm{d},\,w}\left(1+\cos\left(n\phi_{ijkl}-\phi_{ijkl}^{\mathrm{eq},\,w}\right)\right) \tag{3}$$

where $\phi_{\ ijkl}$ is the dihedral angle, $K_{ijkl}^{\mathrm{d},\,w}$ and $\phi_{\ ijkl}^{\ \mathrm{eq},\,w}$ are respectively the w^{th} force constant and the equilibrium dihedral angle for beads i, j, k, and l. Parameter n is an integer given by,

$$\left\lfloor n=\frac{w+1}{2}\right\rfloor \tag{4}$$

where $\lfloor\cdot\rfloor$ represent the floor function.

Beads, bonds, bond angles, and dihedrals within CG PEI and PEI-CA are assigned names consistent with our previous work[47]. Bead names p, s, t, pq are used respectively for the primary, secondary, tertiary, and protonated primary beads in PEI and PEI-CA. Additionally, formaldehyde and alkane beads in CA are named 'f' and 'a', respectively. Since each bead has two terminal atoms that bind with other beads, a direction is assigned to it, pointing from the terminal with lower atomic weight to the one with higher atomic weight while neglecting the fractional weights of atoms within a bead. For example, "-CCN-", and "-$C_{0.5}$CCC(=O)-" both have directions from left to right. Note that in the formaldehyde bead, the rightmost carbon is bound to an oxygen atom, making the right terminal heavier than the left. In the alkane bead the terminals are identical, therefore the bead is symmetric and has no associated direction. To avoid ambiguity caused by asymmetry in PEI beads, directions are also assigned to a bond connecting two beads. The directionality of a bond is determined by combining the directions of the two beads it connects, e.g., "CCN-CCN" and "$CCCCC_{0.5}$-$C_{0.5}$CCC(=O)" has a single direction $\rightarrow$ while "CCN-C(=O)$CCC_{0.5}$" has a mixed directionality of $\rightleftarrows$. This allows for a simple nomenclature of all CG bonds, angles, and dihedrals using their bead names and associated directionality of bonds. For example, "CCN-C(=O)$CCC_{0.5}$" is a bond represented by $s \rightleftarrows f$, "CCN-C(=O)$CCC_{0.5}$-$C_{0.5}$CCC" is an angle represented by $s \rightleftarrows f \leftarrow a$, and "-CCN-(CCN)-CCN-C(=O)$CCC_{0.5}$-" is a dihedral represented by $t\left(\rightarrow s\right) \rightarrow s \rightleftarrows f$. All bonded interactions were categorized using this nomenclature and collectively parametrized.

Parameters already determined from previous works[47] were used without change, while any missing parameters (those including f and a beads) were determined using an automated bottom-up technique[69]. Specifically, AA trajectories from the simulation of a single 2 kDa PEI-CA in water were mapped into reference

CG trajectories using the mapping scheme discussed in "AA and CG models", which were used to generate the reference CG distributions for bond lengths, bond angles, and dihedral angles. For simplicity and to avoid confusion with distributions obtained from CG simulation trajectories, these reference CG distributions are hereafter referred to as AA distributions. The initial guesses for the parameters in Eqs. (1)-(3) were determined by direct Boltzmann inversion of the AA distributions. The automated algorithm was then used to iteratively adjust the parameters to minimize the differences between the means and standard deviations of distributions obtained from CG and AA trajectories[69].

The final comparison of the AA and CG bonded distributions was quantified using the root mean squared deviation (RMSD) and normalized-RMSD (NRMSD). Generally, in this work, if a function $f(u)$ evaluated from AA and CG simulations are compared, the RMSD is defined as,

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[f_{AA}(u_i) - f_{CG}(u_i)]^2}, \tag{5}$$

where $f_{AA}(u)$ and $f_{CG}(u_i)$ are the value of the function at u_i evaluated from AA and CG simulations, respectfully, and n is the total number of data points. The NRMSD is defined from the normalization of RMSD by the difference between the maximum and minimum values of $f_{AA}(u)$,

$$NRMSD = \frac{RMSD}{\max(f_{AA}(u)) - \min(f_{AA}(u))} \tag{6}$$

For the AA vs. CG comparison of each bonded distribution, u represented the variable for that bonded interaction (e.g., bond length, bond angle or dihedral angle), while $f(u)$ was the corresponding probability distribution.

Non-bonded interactions

The non-bonded interactions in Martini 2.1P[81] include electrostatic interactions between charged beads described by the Coulombic potential and van der Waals interactions described by the 12 – 6 Lennard-Jones (LJ) potential. The Coulombic potential is given by,

$$V_{\mathrm{el}}(r_{ij}) = \frac{q_i q_j}{4\pi\,\epsilon_0 \epsilon_r r_{ij}} \tag{7}$$

where r_{ij} is the distance between beads i and j with charges q_i and q_j, respectively, ϵ_0 is the permittivity of free space, and ϵ_r is the dielectric constant of water medium. The LJ potential is given by,

$$V_{\mathrm{LJ}}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{8}$$

where ϵ_{ij} and $2^{1/6}\sigma_{ij}$ are the depth and location, respectively, of the minimum in the LJ potential.

The quality of non-bonded interactions was gauged by comparing the AA and CG PMFs between the 500 Da PEI/PEI-CA (Fig. 1, portion enclosed by the black box) and one of the molecules or systems listed in Table 1. Alignment of the PMFs was examined both visually and by comparing the RMSD and NRMSD between AA and CG PMFs to those reported in classic Martini literature. The RMSD and NRMSD were calculated using Eqs. (5) and (6) where $f(u)$ was identified to be the PMF as a function of its reaction coordinate, Z (details in "Simulation details"). If significant differences were observed between AA and CG PMFs, LJ pair interaction parameters between PEI and the molecule were modified, while keeping other pair interactions unchanged. Since electrostatic interactions are determined by the charges of the beads and the water model, they were left unaltered. Fig. 2a shows a flow chart that demonstrates the six-step parameterization and validation process. First, PMF curves for AA and CG PEI were generated and compared using the default Martini 2.1P parameters[68,81]. Second, if a poor matching was found, AA and CG PMFs for the auxiliary uncharged PEI were generated, and the LJ interaction between P1 beads in the uncharged PEI and beads in other molecules were reparametrized. After a reasonable agreement was found, the new parameters were considered "Stage 1" parameters. Third, the CG PMFs for PEI were regenerated with Stage 1 parameters for P1 beads and default parameters for Qd beads, followed by comparison with the corresponding AA PMF. Fourth, for an unsatisfactory matching, the Qd beads in PEI were reparametrized (with P1 bead parameters unaltered), and the parameters that produced acceptable matching were considered "Stage 2" parameters. Fifth, the Stage 2 parameters for PEI and default parameters for CA beads were used to generate the CG PMF for PEI-CA which was compared against the corresponding AA PMF. Finally, similar to previous steps, Na and C1 beads in PEI-CA were reparametrized to generate a good agreement with the AA PMF, and the finalized parameters were referred to as "Stage 3" parameters.

Reparameterization at each stage was conducted by trial-and-error tuning of ϵ_{ij} and σ_{ij} in Eq. (8), guided by the difference between the CG and AA PMFs. Special care was needed to reparametrize non-bonded interactions between a specific pair of molecules without altering their accurate interactions with other molecules. To ensure that only intended bead pairs were subjected to reparameterization, the bead types of a pair of molecules were duplicated and renamed before modifying the cross interaction between them. An example is shown in Fig. 2b where the interaction between Qd beads in PEI and Q0 beads in DPPC required modification. Within the general Martini topology, the LJ parameters for Qd and Q0 bead types were duplicated to generate new bead types "Qdx" and "Q0z". Qdx and Q0z were identical to Qd and Q0, respectively, when interacting with default Martini beads. Additionally, self-interactions were identical to default Martini, e.g., Qdx-Qdx interaction was identical to Qd-Qd interaction. Then, the LJ parameters between Qdx and Q0z beads were altered to fine-tune

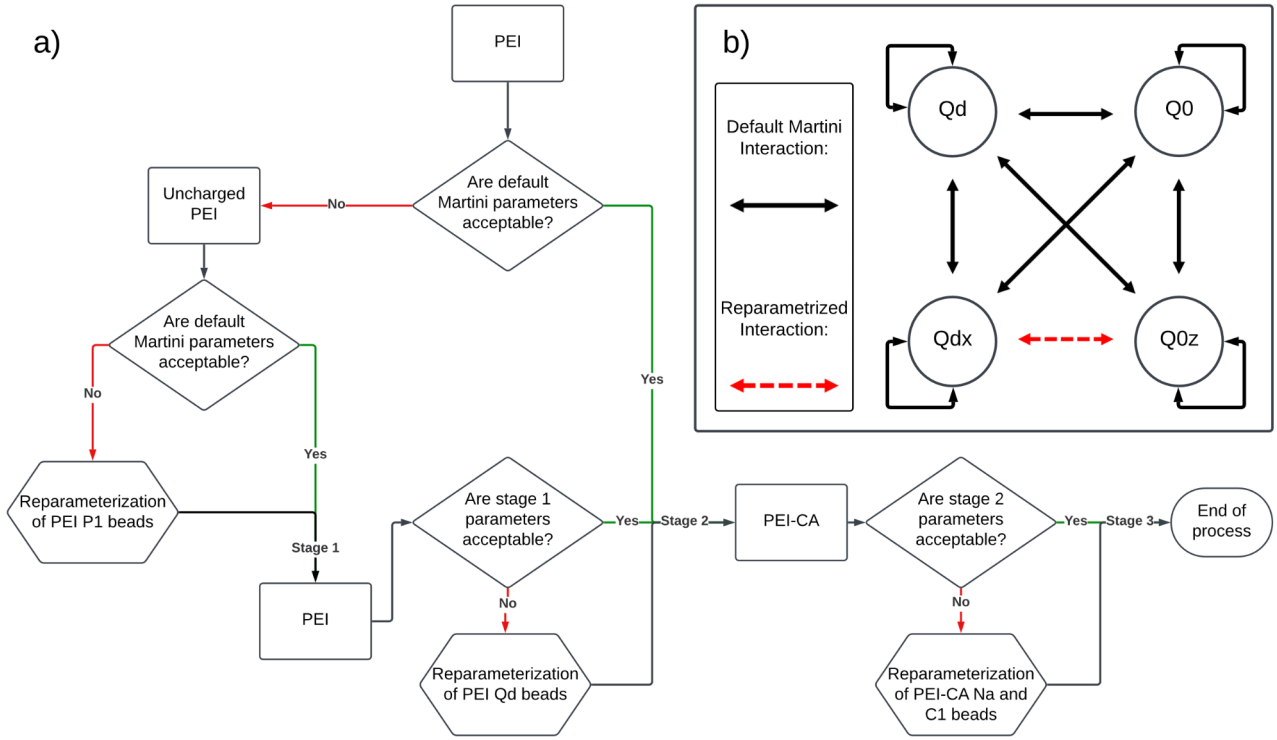

Fig. 2. (**a**) Flowchart of the process for parameterizing non-bonded interactions associated with PEI and PEI-CA. (**b**) An example of reparametrizing the interaction between a Qd bead in PEI and a Q0 bead in DPPC. The red dashed line represents the updated cross interaction while the black lines show the standard Martini interactions that remain unchanged.

PEI-DPPC interactions, without changing how PEI (or DPPC) interacts with itself or other molecules, or how other molecules interact with each other.

Simulation details

To determine the bonded parameters for PEI-CA, AA simulations were run for PEI-CA with 1 CA (substituted at site #1 in Fig. 1) and 3 CA (substituted at all 3 sites in Fig. 1) substitutions. For each simulation, a single PEI-CA was placed in the center of a cubic box sized to allow a 2 nm margin around the PEI-CA that was oriented along its principal axes. The box was filled with TIP3P[82] water, and Cl^- ions were added to neutralize the system. The energy of the system was first minimized using steepest decent followed by NVT equilibration for 100 ps at 300 K and NPT equilibration for 100 ps at 300 K and 1 bar. The temperature was maintained using the stochastic velocity-rescaling thermostat[83] with a time constant of 0.1 ps and the pressure was maintained using the Berendsen barostat[84] with a time constant of 2 ps and compressibility of 4.5×10^{-5} bar $^{-1}$. During equilibration all bonds involving hydrogen atoms were constrained using the LINCS algorithm[85] and position restraints with a force constant of 1000 kJ mol^{-1} nm^2 were applied on the backbone atoms of PEI-CA. Finally, PEI-CA was simulated for 40 ns in the NPT ensemble at 300 K and 1 bar without any position restraints. The temperature and pressure were maintained using the stochastic velocity rescaling thermostat[83] and the Parrinello-Rahman barostat[86] with the same time constants and compressibility used during equilibration. The final 20 ns of the simulations were used for analysis. In all AA simulations, a cutoff of 1.2 nm was used for short-range non-bonded interaction, and long-range electrostatic interactions were handled by the Particle Mesh Ewald method[87]. The equations of motion were solved using the leap-frog integrator with a time step of 2 fs and periodic boundary condition was applied in all directions.

The CG simulations for a single PEI-CA were initialized by converting the final frame of the PEI-CA in the AA simulations to a CG structure using the mapping scheme discussed in "AA and CG models" The CG PEI-CA was solvated with Martini polarizable water using a van der Waals radius of 0.095 nm[70] to ensure the box dimensions remained stable during NPT equilibration. The simulations of PEI-CA in water used default Martini bead types as specified before, and the bonded interactions were parameterized iteratively as described in “Bonded interactions”. During each parameterization step, the CG systems underwent steepest descent energy minimization and NPT equilibration (1 ns), followed by production NPT simulation (5 ns) at the same temperature and pressure as the AA systems. NVT equilibration was excluded as CG simulations eliminate the high-frequency low-amplitude fluctuations in the potential energy surface that require NVT equilibration[88]. The same minimization method, thermostat, and barostats as the AA systems were used except that the pressure time constants for the NPT equilibration and production simulation were changed to 3.0 ps and 5.0 ps, respectively. The short-range electrostatic and LJ potentials were cut off at 1.1 nm and smoothly shifted to zero between 0 and

1.1 nm, and long-range electrostatics were handled using the reaction-field electrostatics[89]. A relative dielectric of 2.5[71] was used for the polarizable water. A periodic boundary condition was applied in all directions and the equations of motion were integrated using the leap-frog algorithm with a time step of 10 fs. The best set of bonded parameters were used to extend the NPT production simulations to 20 ns and the last 10 ns of the simulation was used for further analysis. Both the AA and CG MD simulations for PEI-CA in water were run using Gromacs 2020.4[90].

PMFs were generated by performing umbrella sampling (US) simulations with data analyzed by the weighted histogram analysis method (WHAM)[91]. Each PMF curve required a predefined reaction coordinate, *Z*. Figure 3a shows the reaction coordinate for the interaction of PEI-CA with SP-B, which was the *z*-component of the separation between the centers of mass (COMs) of PEI-CA and SP-B. The reaction coordinate for the interaction between PEI-CA and siRNA was defined in the same way. Figure 3b shows the reaction coordinate for the interaction between PEI-CA and a DPPC (or DPPG) bilayer. A cylinder of 4-nm diameter was first defined with its axis passing through the COM of PEI-CA and perpendicular to the bilayer. The COM location of the lipid molecules within this cylinder was determined and its distance to the COM of PEI-CA was defined as the reaction coordinate. This allowed the simulations to account for localized deformation of the bilayer at small *Z*. Reaction coordinates for the PMF between PEI and siRNA, SP-B, DPPC bilayer, or DPPG bilayer were defined identically to those in the PMF calculations of PEI-CA.

For AA US simulations, the box dimensions varied for each pulmonary gene delivery component (Table 1) that PEI-CA (or PEI; the same approach was applied for PEI and hence not repeated in the description below) was simulated with; these dimensions are found in Table S1 (Section S1 of the Supporting Information). The *z* dimensions of the boxes were sized to be at least twice the largest reaction coordinate. For the systems involving a lipid bilayer, the *x* and *y* dimensions of the box were sized to allow for a 3 nm margin between PEI-CA and the edge of the box. The lipid bilayer was generated using CHARMM-GUI[72,73,92] with the same *x*-*y* dimensions. For the systems involving siRNA or SP-B, the *x* and *y* dimensions of the box were sized to allow for a 1.5 nm margin around the largest molecule. A single PEI-CA molecule was placed at a sufficiently large reaction coordinate such that the non-bonded interactions would be neglibible[82]. The simulation box was filled with TIP3P water[37] and neutralized with Cl^- or Na^+ ions. The system was first energy minimized using the steepest decent algorithm with a force maximum of 100 kJ nm^{-1} mol^{-1}. Then, the system was equilibrated for 100 ps in the NVT ensemble at 300 K, followed by 100 ps NPT equilibration at 300 K and 1 bar. During the equilibration, position restraints were activated on all molecules excluding water, ions, DPPC, and DPPG. The thermostat and barostat, including their time constants and compressibility, were the same as what was used for the single PEI-CA simulations. Next, position restraints were removed except for those on siRNA and SP-B, and a short pulling simulation was run to move the PEI-CA at the speed of 0.01 nm ps^{-1} towards the bilayer, siRNA, or SP-B. The pulling simulation was conducted in the NPT ensemble at 300 K and 1 bar using the Nose-Hoover thermostat[93] with a time constant of 1 ps and the Parrinello-Rahman barostat[86] with a time constant of 3 ps. Frames were taken from the pulling simulation to create initial configurations for the US simulations, where the PEI-CA molecule was positioned in 0.2 nm increments along the reaction coordinate. Each US system was subject to a 100 ps NPT equilibration (same position restraints, thermostat, and barostat as described above for equilibration), followed by US simulation of 20 ns (same thermostat and barostat as described above for pulling simulation) where a biasing potential was applied on the reaction coordinate with a spring constant of 500 kJ mol^{-1} nm^2. In the US simulations involving SP-B, the PEI-CA molecule was observed to drift in the *x*-*y* plane. Because the reaction coordinate was defined along the *z*-axis, drifting in the *x*-*y* plane increased the distance between PEI-CA and SP-B without the increase being reflected in the reaction coordinate. To address this issue, an additional position restraint was placed on the *x* and *y* coordinates of the PEI-CA COM to keep it centered in the *x*-*y* plane during the pulling and US simulations. This effect was observed to a lesser extent with the PEI-CA and siRNA simulations and was irrelevant in the PEI-CA and bilayer simulation due to the planar structure of the bilayer (infinitely large in the *x*-*y* plane under periodic boundary condition applied in all directions). In all AA US simulations, non-bonded interactions were handled identically to the AA simulations with a single PEI-CA. The equations of motion were solved using the leap-frog integrator with a time step of 2 fs. Constraints were applied to bonds involving a hydrogen atom via the LINCS algorithm[85]. The first 5 ns of the US simulations were discarded during analysis.

The CG US simulation box was constructed similar to its AA counterpart and also used the box dimensions in Table S1 (Section S1 of the Supporting Information). For the systems involving a lipid bilayer, care was taken to ensure that the area per lipid was consistent between AA and CG simulations. Each simulation box was filled with Martini polarizable water (van der Waals radius of 0.095 nm) and neutralized with Martini Cl^- or Na^+ ions. The simulation settings of the CG US simulations were identical to the AA US simulations, except for skipping the unnecessary NVT equilibration. Simulation parameters such as thermostat, barostat and integration time step were identical to the single CG PEI-CA simulations. A simulation length of 20 ns for each US window was found to be sufficient because extending the simulation caused minimal change in the PMF and the first 5 ns of the discarded during analysis. Both the AA and CG US simulations were conducted using Gromacs 2020.4[90].

Results and discussion

Accuracy of bonded interactions

The comparison of AA and CG distributions for representative bond length, bond angle, and dihedral angle in the CG PEI-CA is shown in Fig. 4a–f. The remaining bonded distributions that involve the CA beads are presented in Section S2 of the Supporting Information. The CG bonded parameters generated by the bottom-up scheme are reported in Table S1. The bonded interactions unrelated to the formaldehyde or alkane beads are not shown as they have been previously validated for the CG PEI model[47]. Both CG PEI-CA models, with different degrees of lipid substitution, successfully reproduced the position and height of the peaks in the AA

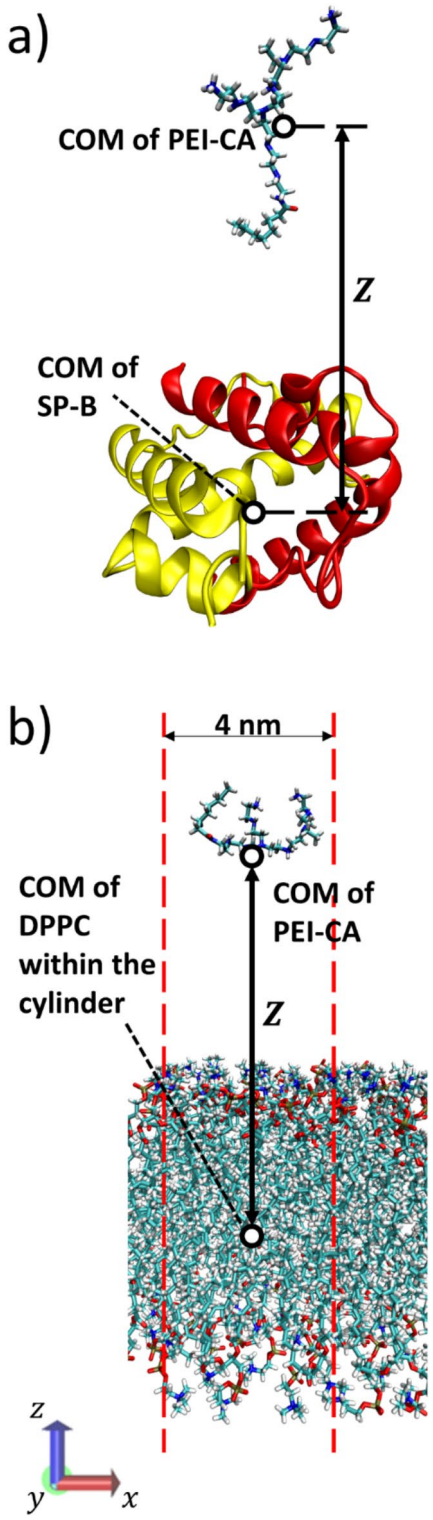

Fig. 3. The reaction coordinates in AA US simulations for the interaction of PEI-CA with (**a**) SP-B and (**b**) a DPPC bilayer. The reaction coordinates in the corresponding CG US simulations were identically defined. Molecule visualizations were generated using VMD 1.9.4.

distributions. This is expected as the bottom-up parametrization scheme[69] adjusts the force field parameters based on both the mean and the standard deviation of the AA distribution. The bond length distribution for PEI-CA with 3 substitutions has the largest discrepancy between the AA and CG models (Fig. 4d) because skewed distributions are not easily reproduced by harmonic potentials. The CG distribution is slightly narrower in this case, but the similarity between AA and CG is comparable to the bonded distributions of CG PEI[47]. The

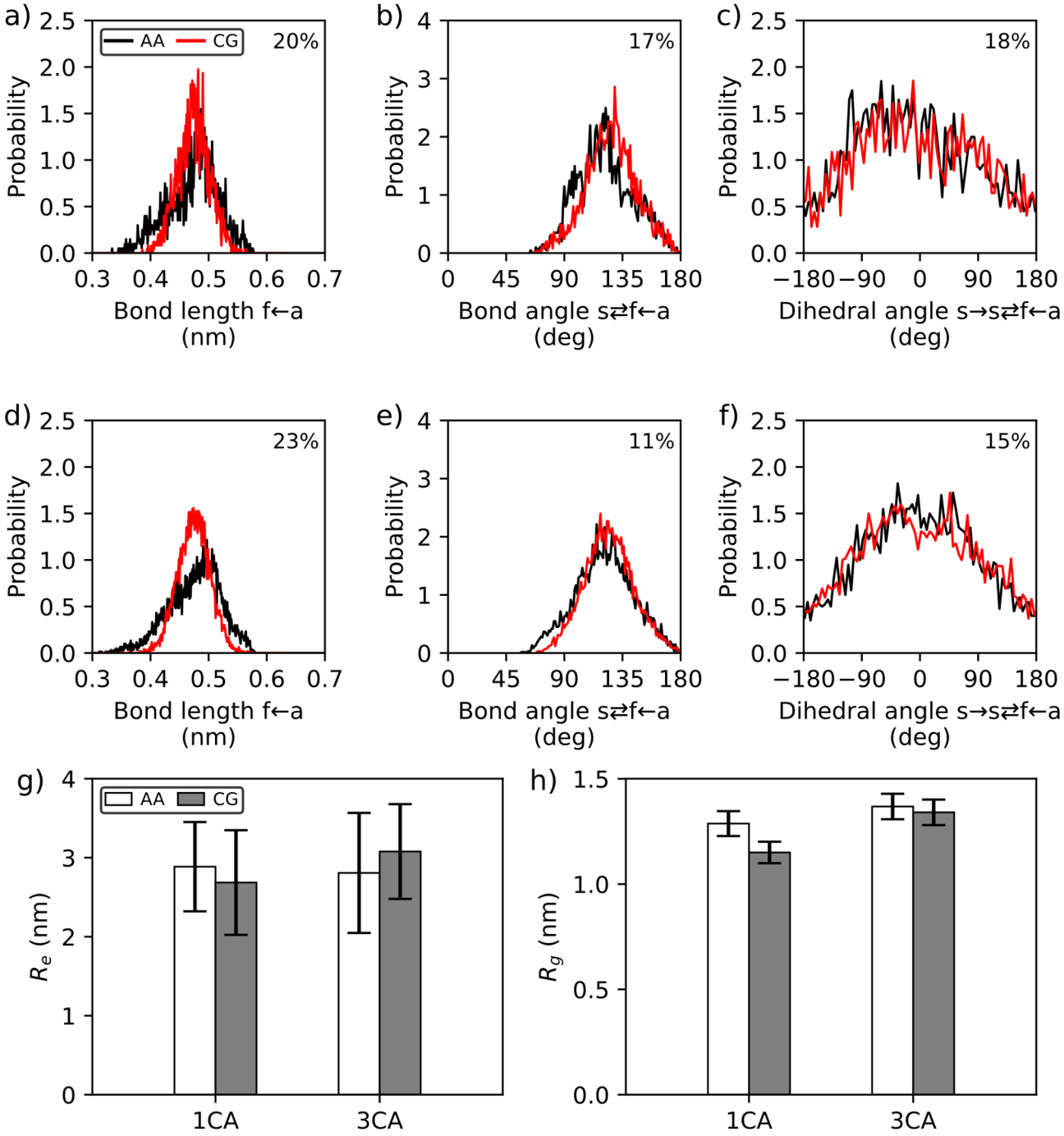

Fig. 4. Comparison of AA and CG distributions of representative bond length, bond angle, and dihedral angle in 2 kDa PEI with (**a**–**c**) 1 and (**d**–**f**) 3 CA substitutions. Percentage reported in each subfigure is the NRMSD calculated according to Eq. (6). Comparison of the end-to-end distance (R_e) (**g**) and radius of gyration (R_g) (**h**) of AA and CG 2 kDa PEI-CA with 1 and 3 CA substitutions. Error bars in (**g**) and (**h**) represent the standard deviation.

NRMSD between the AA and CG distributions calculated according to Eq. (6) are given in the top right corner of the subfigures. The NRMSDs range from 10 to 25% which is within the range of those calculated from classical Martini works[94].

Accuracy of the bonded interactions is further validated by comparing AA and CG end-to-end distance (R_e) and radius of gyration (R_g) of PEI-CA (Fig. 4g and h). For both CG PEI-CA with 1 and 3 lipid substitutions, R_e is defined based on the main backbone chain of PEI (between the tertiary bead in the lower left and the charged primary bead in the upper right of Fig. 1). The mean of AA R_e and R_g is comparable to the CG results, and the largest deviation (11%) is observed for R_g in PEI-CA with 1 substitution. The AA standard

deviations for R_e and R_g are reproduced by the CG model for all cases, which is a feature of the bottom-up parameterization technique[69]. The standard deviations of R_e are larger than those for R_g, which is consistent with previous results for PEI[47]. The AA and CG R_g increases with the number of CA substitutions, due to the increase in molecular volume induced by additional CA tails. Considering the standard deviation, the number of CA substitutions has little influence on R_e regardless of the type of the model (AA or CG).

Reparameterization for interaction with siRNA

The PMFs calculated between PEI/PEI-CA and siRNA are shown in Fig. 5, with the shaded region depicting the estimated error from bootstrap analysis[95]. The CG PMF between uncharged PEI and siRNA with default Martini parameters shows large discrepancy with the AA PMF (black curves in Fig. 5a). The CG uncharged PEI exhibits large attraction with siRNA (deeper PMF well) that is missing in its AA counterpart. To reduce the attraction, the ϵ_{ij} values in the LJ potential (Eq. 8) between the uncharged PEI beads (P1) and siRNA beads are decreased (see "Non-bonded interactions" for details). In particular, ϵ_{ij} values associated with siRNA backbone beads and sidechain beads are modified separately with the goal of matching the long-range slope and short-range energy minimum in the PMF curve, respectively. This strategy is used during every reparameterization step for siRNA. Updated parameters in Table S2 (Section S3 of the Supporting Information) demonstrate a good agreement between CG and AA PMFs for the uncharged PEI, as shown by the black curves in Fig. 5b. In contrast to the uncharged CG PEI, the PMF of CG PEI is more repulsive than the AA PMF (red curves in Fig. 5a). This repulsion is expected to further increase after correcting the unrealistically large attractions between uncharged PEI beads and siRNA. Following the same strategy, ϵ_{ij} between charged PEI beads (Qd) and siRNA beads are increased to enhance their attraction. Table S2 (Section S3 of the Supporting Information) lists the LJ parameters that produce good matching between CG and AA PMFs for PEI, as seen in Fig. 5b. These optimized parameters are applied to PEI-CA, along with the default Martini parameters for the beads in CA substitution (Na and C1). The CG PMF for PEI-CA demonstrates a good match with its AA counterpart (Fig. 5c), as a result, further parametrization for CA substitution is deemed unnecessary. The RMSD and NRMSD before and after each stage of parameterization are given in Table 2.

Reparameterization for Interaction with lipids

Large differences between the CG and AA PMFs for the interaction of PEI with DPPC bilayer are observed with default Martini parameters (red lines in Fig. 6a). The CG and AA PMF curves for uncharged PEI demonstrated a satisfactory match (black lines in Fig. 6a), and modifications to P1 bead parameters are unnecessary. Instead, parameters are adjusted between the charged PEI beads (Qd) and the headgroup beads of DPPC (Qa and Q0) to reduce the large repulsion between CG PEI and DPPC bilayer (Fig. 6a), as well as to capture the local minima at Z = 1–2 nm observed in the AA PMF. The ϵ_{ij} and σ_{ij} parameters between the PEI Qd beads and the DPPC Qa and Q0 beads were systematically adjusted. Among all adjustments, altering σ_{ij} between Qd and Q0 yields the best CG PMF (Fig. 6b) and stage 2 parameters. These optimized parameters (stage 2) are applied to PEI-CA; however, the CG and AA PMF result did not match well (see the dotted line in Fig. 6c). To achieve a better match, ϵ_{ij} parameters between CA beads (Na, C1) and DPPC head group beads (Qa, Q0) are decreased to reduce the large attraction at Z < 1.5 nm observed with default parameters for lipid substitutions. Among all modifications tested, reducing ϵ_{ij} for Na-Q0 interaction results in the best match between CG and AA PMF (dashed line in Fig. 6c). The optimized LJ parameters are provided in Table S2 (Section S3 of the Supporting Information).

The default Martini parameters produce a reasonable agreement between the AA and CG PMFs for interactions between DPPG bilayer and PEI (Fig. 6d), and PEI-CA (Fig. 6e). Therefore, no further adjustments to the CG parameters are made. The RMSD and NRMSD for both DPPC and DPPG bilayers before and after each stage of reparameterization are found in Table 2.

Reparameterization for Interaction with SP-B

The CG PMFs between PEI and SP-B based on the default Martini parameters match well with corresponding AA PMFs (Fig. 7a), consequently, the LJ parameters are not altered. In contrast, the CG PMF between PEI-CA and SP-B (with stage 2 parameters) deviated significantly from the AA PMF (Fig. 7b). The CG PMF demonstrated an uncharacteristic attraction for $Z < 3$ nm and a global free energy minimum at $Z = 1.8$ nm. Similar to the CG PMF, AA PMF demonstrate an attraction for $Z < 3$ nm, but such attractions only result in a local minimum at $Z = 2.3$ nm. The observation that AA and CG PMFs match well in Fig. 7a but not in Fig. 7b indicates an overestimated hydrophobic attraction between the SP-B and CA beads (C1 and Na). This is confirmed by visualizing the AA and CG trajectories in the US simulations, when the reaction coordinates were at their respective PMF local minima (2.3 nm in AA, Fig. 7c, 1.8 nm in CG, Fig. 7d). In both cases the CA atoms/beads act as an anchor, attaching PEI-CA to SP-B. The effect is stronger in the CG system, shown by the consistently closer CA-to-SP-B distance over time (Section S4 of the Supporting Information). To weaken the hydrophobic attractions in the CG system, ϵ_{ij} between the CA beads (Na, C1) and SP-B beads is reduced following the same steps used to reparametrize siRNA-PEI interactions. Specifically, backbone and sidechain beads are parameterized separately to match the long-range slope of the PMF and short-range energy minimum, respectively. The resulting stage 3 parameters (listed in Table S2 in Section S3 of the Supporting Information) render the good agreement between the AA and CG PMFs (Fig. 7b), with RMSD and NRMSD before and after the reparameterization given in Table 2.

Discussion

A CG PEI-CA model is developed with an automated bottom-up coarse-graining technique[69], and it successfully reproduces the mean and variance of the local structural correlations (bonded distributions). Additionally, the CG model accurately predicts global structural properties of PEI-CA such as the end-to-end distance of the linear

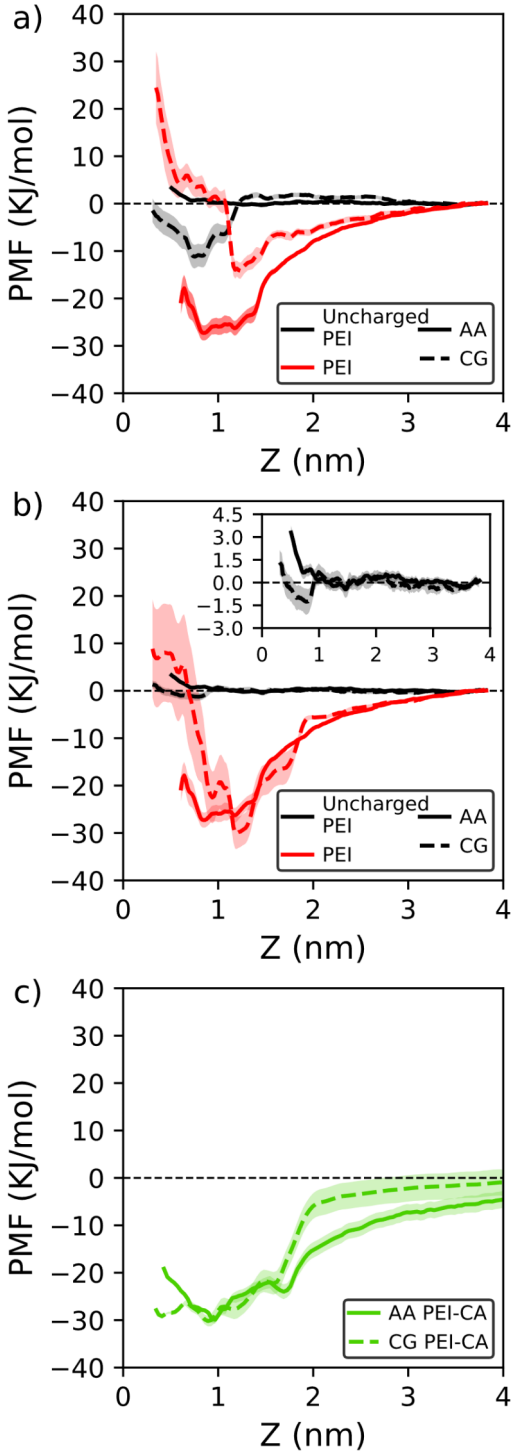

Fig. 5. AA and CG PMF curves for the interaction of siRNA with (**a**) PEI and uncharged PEI using default Martini parameters (before stage 1 in Fig. 2), (**b**) PEI and uncharged PEI using modified Martini parameters (after stage 2 in Fig. 2), (**c**) PEI-CA using modified parameters from (**b**) (after stage 2 in Fig. 2). Inset in (**b**) shows an enlarged image for the PMFs of uncharged PEI.

backbone and the radius of gyration. To improve the CG nonbonded interaction of PEI or PEI-CA with siRNA, SP-B, DPPC bilayer, or DPPG bilayer, a systematic three stage parameterization is introduced that aims to match the AA and CG PMFs between them (Fig. 2). Although trial-and-error is at the core of the reparameterization process, the CG parameters are changed strategically. First, the influences of charged (Qd), uncharged (P1), and

		Uncharged PEI		PEI		PEI-CA	
		Before 1	After 1	Before 2	After 2	Before 3	After
siRNA	RMSD	4.02	0.82	11.41	5.61	4.19	
	NRMSD	110%	22%	41%	20%	14%	
DPPC	RMSD	7.71		43.76	3.46	10.84	4.22
	NRMSD	28%		337%	27%	29%	11%
DPPG	RMSD			10.77		9.98	
	NRMSD			24%		20%	
SPB	RMSD			4.29		1.13	1.56
	NRMSD			11%		17%	6%

PMF comparison from classical Martini literature	RMSD (kJ/mol)	NRMSD
Pulling a DPPC lipid through a DPPC bilayer[68]	18.40	18%
Cytosine and guanine interacting in water[94]	3.90	34%
Lysine pulled out of a dioleoylphosphatidylcholine (DOPC) bilayer[81]	15.21	23%

Table 2. RMSD and NRMSD between AA and CG PMFs before and after each reparameterization stage (1, 2 or 3).
Crossed-out cells represent situations where reparameterization was not needed. Shaded cells represent the values for the finalized parameters. RMSD is given in units of kJ mol^{-1}. RMSD and NRMSD between AA and CG PMFs calculated from data in classical Martini literature are included for comparison.

lipid-substituted (Na, C1) beads in PEI-CA, are decoupled by calculating the PMFs for an auxiliary uncharged PEI (only P1 beads), standard charged PEI (P1 and Qd beads), and PEI-CA (P1, Qd, Na, and C1 beads). The LJ parameters of PEI/PEI-CA with siRNA, SP-B, DPPC, or DPPG are improved systematically by considering the beads that would be responsible for long-range (backbones in SP-B and siRNA, head groups in DPPC and DPPG) and short-range (side chains in SP-B and siRNA, lipid tails in DPPC and DPPG) interactions. The decrease in RMSD and NRMSD between the AA and CG PMFs shown in Table 2 after each reparameterization stage demonstrates the effectiveness of our reparameterization strategy. The PEI-DPPC bilayer and uncharged PEI-DPPG bilayer interactions exhibit the largest improvements, with their NRMSD reduced by an order of magnitude. The final RMSD and NRMSD of the PMFs from our work (shaded cells in Table 2) are comparable and often lower than those obtained in classical Martini works in the literature (Table 2).

Despite the improved accuracy, one should be aware of the discrepancies between the AA and CG PMFs and their influences in replicating physical properties or processes. The discussions on the accuracy of uncharged PEI are avoided due to its lack of relevance in gene delivery. The CG siRNA-PEI pair experiences stronger repulsion than AA at $Z < 1$ nm and the location of minimum in the CG PMF is 0.4 nm larger than that in the AA PMF (Fig. 5b). These results indicate CG siRNA-PEI complexes would be more swollen than their AA counterparts. However, the long-range CG siRNA-PEI interactions and the minimum PMF value exactly reproduce the AA results, indicating accurate modelling of long-range siRNA-PEI attraction and free energy of association. Similar analysis of CG PEI-CA and siRNA interactions indicate accurate modelling of free energy of association but an underestimation of long-range attraction ($Z > 2$ nm; Fig. 5c).

The CG PMF between PEI and DPPC bilayer is lower than its AA counterpart for $Z < 2$ nm (Fig. 6b), suggesting an overestimated attraction between PEI and DPPC. Therefore, when simulating a PEI NP crossing the cell membrane or pulmonary surfactant containing DPPC, a higher association is expected. Similar analysis reveals underestimated attraction between PEI and DPPG (Fig. 6d). Together, the overestimated attraction between PEI and DPPC, and underestimated attraction between PEI and DPPG, might have certain compensatory effect when simulating a cell membrane or pulmonary surfactant containing a mixture of both types of phospholipids. For CG PEI-CA at $Z < 1$ nm, attraction with DPPC (Fig. 6c) is underestimated and attraction with DPPG (Fig. 6e) is overestimated. Similar to PEI, a compensatory affect is expected between PEI-CA and cell membrane or pulmonary surfactant containing a mixture of DPPC and DPPG. Above 1 nm, the CG PMF between DPPG and PEI-CA also predicts underestimated attraction, and no compensatory affect is expected. The interactions CG PEI and PEI-CA have with SP-B accurately models the AA interactions and no significant difference between AA and CG results are expected.

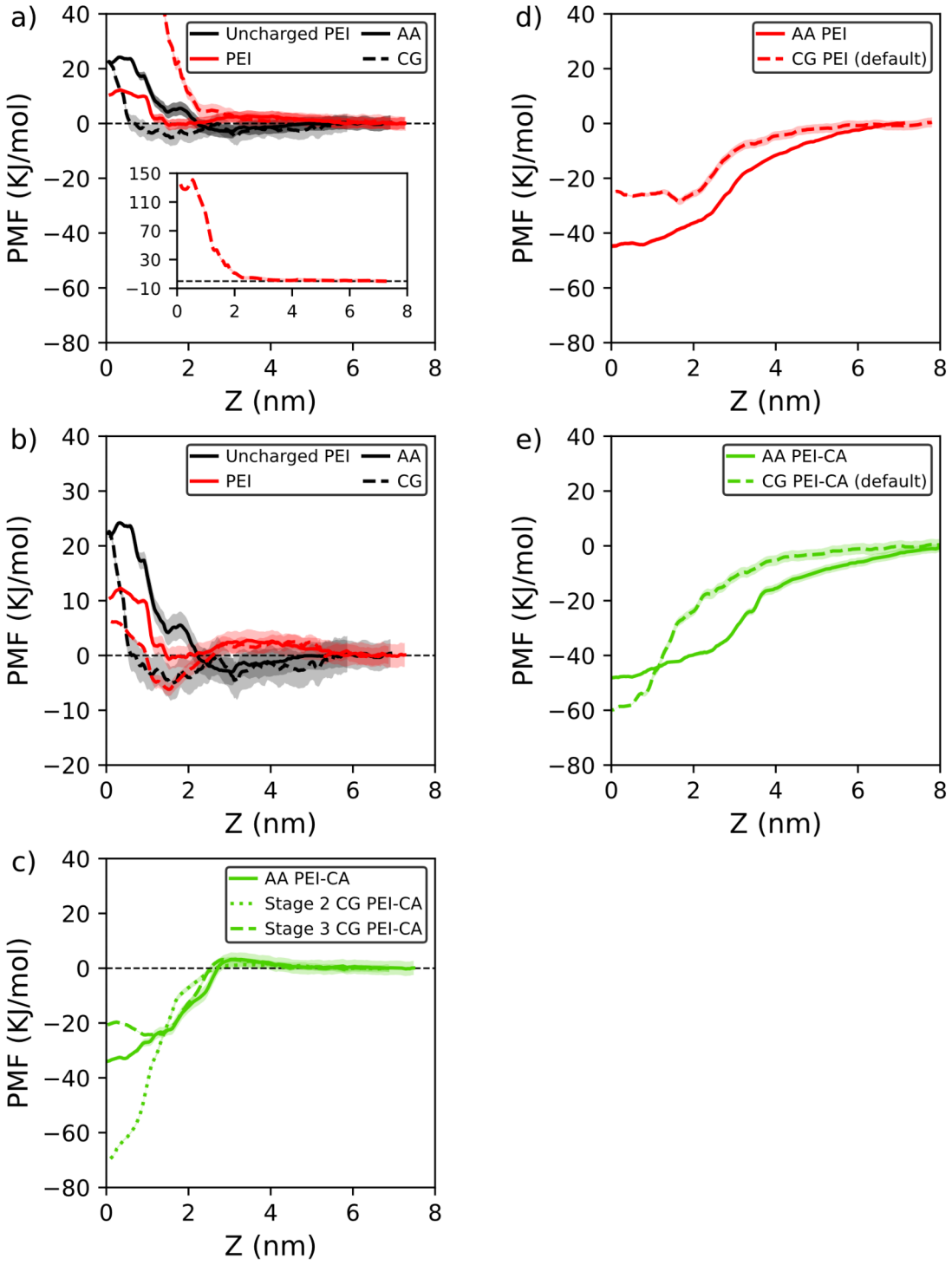

Fig. 6. AA and CG PMF curves for the interaction of DPPC bilayer with (**a**) PEI and uncharged PEI using default Martini parameters (before stage 1 in Fig. 2), (**b**) PEI and uncharged PEI using modified Martini parameters (after stage 2 in Fig. 2) (**c**) PEI-CA using modified parameters from b), as well as parameters after stage 3 in Fig. 2. AA and CG PMF curves for the interaction of DPPG with (**d**) PEI and (**e**) PEI-CA, both using the default Martini parameters. Inset in (**a**) shows the entire range of data for CG PEI. Vertical scale in (**b**) is different from the other subfigures to more clearly show the data.

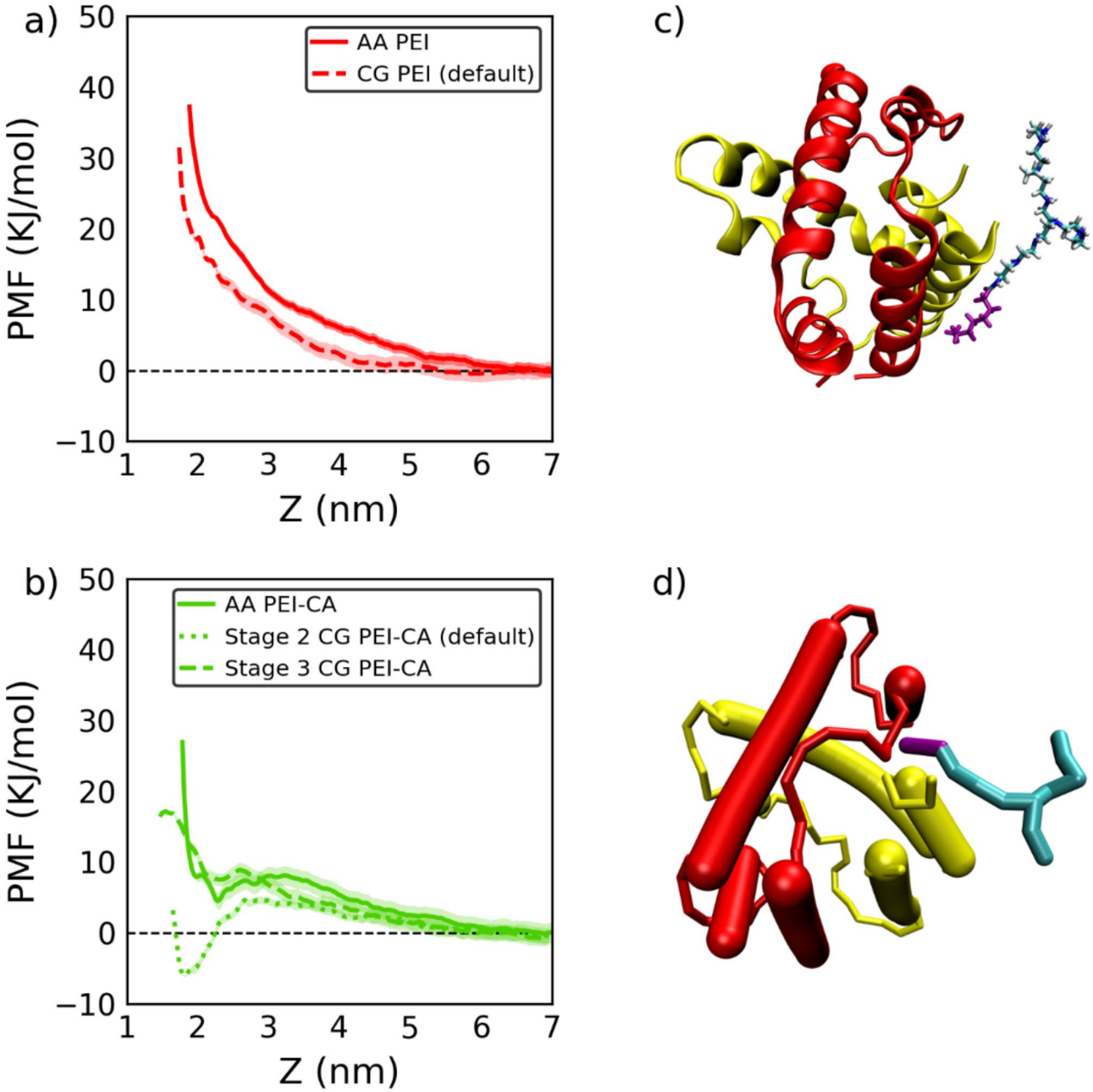

Fig. 7. AA and CG PMFs for the interaction of SP-B with (**a**) PEI using default Martini parameters (before stage 1 in Fig. 2) (**b**) PEI-CA using default and modified Martini parameters (after stage 3 in Fig. 2). Visualizations of hydrophobic interaction anchoring PEI-CA to SP-B in (**c**) AA and (**d**) CG systems. PEI is represented in light blue with a purple CA tail. The SP-B dimer is represented by one red and one yellow monomer, and the helices in (**d**) are represented using Bendix[96]. Molecule visualizations were generated using VMD 1.9.4.

The reparameterization of CG parameters within the Martini framework faces challenges due to fixed functional form of the interaction potentials. Efforts are made to further improve the matching between AA and CG PMFs by exploring the LJ parameter space around the default parameters, but without success. For example, in the case of DPPG, extensive reparameterization was attempted for both PEI and PEI-CA but it only resulted in an increase in RMSD and NRMSD. Given that the PMF comparisons with default Martini parameters produces an NRMSD similar to those from classical Martini works (Table 2) the default parameters are accepted. An automated algorithm that can run US simulations, compare CG and AA PMF curves, adjust LJ parameters based on the comparison, and re-run the US simulations would allow for more effective reparameterization. It is of interest to establish such an algorithm in the future, but it would require (1) careful design of the objective function(s), (2) development of an iterative scheme to systematically optimize the force field parameters, and

(3) consideration on the computational efficiency due to the high computational cost associated with PMF calculations.

Implications and future applications of developed CG models

The CG models developed in this work represent a significant advancement in the computational study of interactions between PEI-based NPs and the pulmonary environment. Parametrized and validated against AA simulations, the Martini compatible native and lipid-substituted PEI models are valuable tools for enabling rational design of pulmonary gene delivery systems with a systematic exploration of NP properties such as size, shape, charge, and surface functional groups. For example, understanding the role of lipid substitution in potentially enhancing the NP's ability to cross the pulmonary surfactant barrier and improving cellular uptake can guide the development of more efficient and targeted gene delivery vectors. Furthermore, these models can be extended to simulate the transport of NPs across the mucociliary barrier and their uptake by epithelial cells. However, such extensions will require further parameterization to accurately represent components like mucins in the mucociliary barrier and epithelial cell membranes' complex lipid and protein composition.

While motivated by pulmonary gene delivery, the developed CG models are not restricted to this particular application. In the literature, Martini-based CG models have been successfully applied to study a wide range of biomolecular processes such as protein-ligand binding[97–99], protein folding[61] and aggregation[100], protein-protein[100,101] and DNA/RNA[76,94] interactions. They have also been used to explore membrane biophysics[102], viral capsid assembly[103], enzyme activity[104], and drug delivery mechanisms[105,106]. The CG model of PEI-CA developed in this work is the first for PEI with a lipid substitution, and the Martini strategy to parametrize it can be extended to PEIs with other molecular substitutions. This work therefore lays the groundwork for parameterizing CG models containing new PEI substitutions that would allow the investigation of substitution-function relationship in various biological contexts including intravenous gene delivery. Since PEI and PEI-CA have been validated against DPPC and DPPG bilayers, one can explore the barriers associated with crossing of not only pulmonary surfactant but also cell membrane. MD simulations that explore protein-PEI interactions are limited in the literature[107] and none are in the context of gene delivery. High accuracy of CG protein-PEI (and PEI-CA) interactions demonstrated in our models would allow one to explore the specificity of the gene carrier based on the protein compositions of a cell membrane or pulmonary surfactant, as well as screening of cytotoxic effects due to protein-PEI interactions. In the context of pulmonary delivery, SP-B has been theorized to facilitate folding of the surfactant[77,108] and permeation of charged particles through the lipid layers[63]. It is of great interest to investigate the effects of SP-B on the mechanisms and efficacy of PEI-based pulmonary gene delivery systems in the future.

Conclusion

A Martini compatible CG model is developed for PEI and its lipid-substituted variant, essential for simulating PEI NPs in pulmonary gene delivery. CG Native ("PEI") and caprylic acid substituted PEI ("PEI-CA") models are developed by determining their bonded and nonbonded interaction parameters. The bonded parameters are obtained from an automated bottom-up parameterization tool, which give rise to accurate predictions of local (probability distributions of bond lengths, angles and dihedrals) and global (end-to-end distance, radius of gyration) structural correlations when compared with the AA counterpart. Parametrization of the nonbonded interactions focuses on PEI and PEI-CA interactions with siRNA and key components of pulmonary surfactant, including DPPC, DPPG, and SP-B. These parameterizations also make PEI and PEI-CA compatible with the key components of cell membrane and can be used to study intravenous gene delivery. A three-stage parameterization strategy is introduced to improve the match between AA and CG PMFs for (Stage 1) an uncharged PEI, (Stage 2) charged PEI without lipid substitution, and (Stage 3) PEI-CA. With the final adopted parameters, the relative differences between the AA and CG PMFs are comparable to or smaller than those reported in the Martini literature. The models presented in this work are expected to facilitate CG simulations that study PEI and PEI-CA based pulmonary (and intravenous) gene delivery systems and contribute to strategy development that enhances their efficacy of the delivery.

Data availability

Final bonded and non-bonded parameters for PEI, siRNA, DPPC, DPPG, and SP-B can be found in the Supporting Information. Other data generated and analyzed in this study are available from the corresponding author on reasonable request.

Received: 1 October 2024; Accepted: 31 January 2025
Published online: 05 February 2025

References

1. Chan, Y. et al. Introduction to chronic respiratory diseases: A pressing need for novel therapeutic approaches. in *Medicinal Plants for Lung Diseases: A Pharmacological and Immunological Perspective* 47–84 (2021). https://doi.org/10.1007/978-981-33-6850-7_2
2. Dua, K. et al. The potential of siRNA based drug delivery in respiratory disorders: Recent advances and progress. *Drug Dev. Res.* **80**, 714–730 (2019).
3. Whitsett, J. A. & Alenghat, T. Respiratory epithelial cells orchestrate pulmonary innate immunity. *Nat. Immunol.* **16**, 27–35 (2015).
4. Sosnowski, T. R. Nanosized and nanostructured particles in pulmonary drug delivery. *J. Nanosci. Nanotechnol.* **15**, 3476–3487 (2015).
5. Liu, D. et al. Nanomedicines targeting respiratory injuries for pulmonary disease management. *Adv. Funct. Mater.* **32**, 2112258 (2022).

6. Huck, B. C. et al. Models using native tracheobronchial mucus in the context of pulmonary drug delivery research: Composition, structure and barrier properties. *Adv. Drug Deliv Rev.* **183**, 114141 (2022).
7. Dziura, M. et al. Simulated breathing: Application of molecular dynamics simulations to pulmonary lung surfactant. *Symmetry (Basel)*. **13**, 1259 (2021).
8. Goerke, J. Pulmonary surfactant: Functions and molecular composition. *Biochim. Biophys. Acta Mol. Basis Dis.* **1408**, 79–89 (1998).
9. Gustafson, H. H., Holt-Casper, D., Grainger, D. W. & Ghandehari, H. Nanoparticle uptake: The phagocyte problem. *Nano Today*. **10**, 487–510 (2015).
10. de Souza Carvalho, C., Daum, N. & Lehr, C. M. Carrier interactions with the biological barriers of the lung: Advanced in vitro models and challenges for pulmonary drug delivery. *Adv. Drug Deliv Rev.* **75**, 129–140 (2014).
11. Murgia, X., de Souza Carvalho, C. & Lehr, C. M. Overcoming the pulmonary barrier: new insights to improve the efficiency of inhaled therapeutics. *Eur. J. Nanomed.* **6**, (2014).
12. Baliga, U. K. & Dean, D. A. Pulmonary gene delivery—realities and possibilities. *Exp. Biol. Med.* **246**, 260–274 (2021).
13. Gao, Y., Liu, X., Chen, N., Yang, X. & Tang, F. Recent advance of Liposome nanoparticles for nucleic acid therapy. *Pharmaceutics* **15**, 178 (2023).
14. Yetisgin, A. A., Cetinel, S., Zuvin, M., Kosar, A. & Kutlu, O. Therapeutic nanoparticles and their targeted delivery applications. *Molecules* **25**, 2193 (2020).
15. Katz, M. G. et al. Targeted gene delivery through the respiratory system: Rationale for Intratracheal Gene transfer. *J. Cardiovasc. Dev. Dis.* **6**, 8 (2019).
16. Gautam, A., Waldrep, C. J. & Densmore, C. L. Delivery systems for pulmonary gene therapy. *Am. J. Respiratory Med.* **1**, 35–46 (2002).
17. Wasungu, L. & Hoekstra, D. Cationic lipids, lipoplexes and intracellular delivery of genes. *J. Controlled Release*. **116**, 255–264 (2006).
18. Ernst, N. et al. Interaction of liposomal and polycationic transfection complexes with pulmonary surfactant. *J. Gene Med.* **1**, 331–340 (1999).
19. Shi, C. et al. Novel drug delivery liposomes targeted with a fully human anti-VEGF165 monoclonal antibody show superior antitumor efficacy in vivo. *Biomed. Pharmacotherapy*. **73**, 48–57 (2015).
20. Gregoriadis, G. Engineering liposomes for drug delivery: Progress and problems. *Trends Biotechnol.* **13**, 527–537 (1995).
21. Wei, X. et al. Liposome-based glioma targeted drug delivery enabled by stable peptide ligands. *J. Controlled Release*. **218**, 13–21 (2015).
22. Cavallaro, G., Sardo, C., Craparo, E. F., Porsio, B. & Giammona, G. Polymeric nanoparticles for siRNA delivery: Production and applications. *Int. J. Pharm.* **525**, 313–333 (2017).
23. Lungwitz, U., Breunig, M., Blunk, T. & Göpferich, A. Polyethylenimine-based non-viral gene delivery systems. *Eur. J. Pharm. Biopharm.* **60**, 247–266 (2005).
24. Günther, M. et al. Polyethylenimines for RNAi-mediated gene targeting in vivo and siRNA delivery to the lung. *Eur. J. Pharm. Biopharm.* **77**, 438–449 (2011).
25. Ferrari, S. et al. Polyethylenimine shows properties of interest for cystic fibrosis gene therapy. *Biochim. et Biophys. Acta (BBA) - Gene Struct. Expr.*. **1447**, 219–225 (1999).
26. Rai, N., Shihan, M., Seeger, W., Schermuly, R. T. & Novoyatleva, T. Genetic delivery and gene therapy in pulmonary hypertension. *Int. J. Mol. Sci.* **22**, 1179 (2021).
27. Gautam, A., Densmore, C. L., Xu, B. & Waldrep, J. C. Enhanced gene expression in mouse lung after PEI–DNA aerosol delivery. *Mol. Ther.* **2**, 63–70 (2000).
28. Pandey, A. P., Sawant, K. K. & Polyethylenimine A versatile, multifunctional non-viral vector for nucleic acid delivery. *Mater. Sci. Eng., C.* **68**, 904–918 (2016).
29. Li, S., Tan, Y., Viroonchatapan, E., Pitt, B. R. & Huang, L. Targeted gene delivery to pulmonary endothelium by anti-PECAM antibody. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **278**, L504–L511 (2000).
30. Goula, D. et al. Rapid crossing of the pulmonary endothelial barrier by polyethylenimine/DNA complexes. *Gene Ther.* **7**, 499–504 (2000).
31. Densmore, C. L. et al. Aerosol delivery of robust polyethyleneimine– DNA complexes for gene therapy and genetic immunization. *Mol. Ther.* **1**, 180–188 (2000).
32. 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.* **7**, 2209–2217 (2011).
33. Neamnark, A. et al. Aliphatic lipid substitution on 2 kDa polyethylenimine improves plasmid delivery and transgene expression. *Mol. Pharm.* **6**, 1798–1815 (2009).
34. Jiang, C. et al. Recent advances in the development of polyethylenimine-based gene vectors for safe and efficient gene delivery. *Expert Opin. Drug Deliv*. **16**, 363–376 (2019).
35. Song, H. et al. Cationic lipid-coated PEI/DNA polyplexes with improved efficiency and reduced cytotoxicity for gene delivery into mesenchymal stem cells. *Int. J. Nanomed.* **7**, 4637–4648 (2012).
36. Bivas-Benita, M., Romeijn, S., Junginger, H. E. & Borchard, G. PLGA-PEI nanoparticles for gene delivery to pulmonary epithelium. *Eur. J. Pharm. Biopharm.* **58**, 1–6 (2004).
37. Kleemann, E. et al. Modified polyethylenimines as non-viral gene delivery systems for aerosol gene therapy: investigations of the complex structure and stability during air-jet and ultrasonic nebulization. *J. Controlled Release*. **100**, 437–450 (2004).
38. De Wolf, F. A. & Brett, G. M. Ligand-binding proteins: their potential for application in systems for controlled delivery and uptake of ligands. *Pharmacol. Rev.* **52**, 207–236 (2000).
39. Frei, A. P. et al. Direct identification of ligand-receptor interactions on living cells and tissues. *Nat. Biotechnol.* **30**, 997–1001 (2012).
40. Ding, L., Tang, S., Wyatt, T. A., Knoell, D. L. & Oupický, D. Pulmonary siRNA delivery for lung disease: Review of recent progress and challenges. *J. Controlled Release*. **330**, 977–991 (2021).
41. Dailey, L. A. et al. Modified polyethylenimines as non viral gene delivery systems for aerosol therapy: Effects of nebulization on cellular uptake and transfection efficiency. *J. Controlled Release*. **100**, 425–436 (2004).
42. Choudhury, C. K., Kumar, A. & Roy, S. Characterization of conformation and interaction of gene delivery vector polyethylenimine with phospholipid bilayer at different protonation state. *Biomacromolecules* **14**, 3759–3768 (2013).
43. Antipina, A. Y. & Gurtovenko, A. A. Molecular-level insight into the interactions of DNA with phospholipid bilayers: Barriers and triggers. *RSC Adv.* **6**, 36425–36432 (2016).
44. Kwolek, U. et al. Interactions of Polyethylenimines with Zwitterionic and anionic lipid membranes. *Langmuir* **32**, 5004–5018 (2016).
45. Mahajan, S. & Tang, T. Polyethylenimine-DNA nanoparticles under endosomal acidification and implication to gene delivery. *Langmuir* **38**, 8382–8397 (2022).
46. Mahajan, S., Tang, T. & Polyethylenimine -DNA ratio strongly affects their nanoparticle formation: A large-scale coarse-grained molecular dynamics study. *J. Phys. Chem. B*. **123**, 9629–9640 (2019).
47. Mahajan, S. & Tang, T. Martini coarse-grained model for polyethylenimine. *J. Comput. Chem.* **40**, 607–618 (2019).

48. Beu, T. A., Ailenei, A. E. & Farcaş, A. Atomistic and coarse-grained modeling of polyethyleneimine. *Chem. Phys. Lett.* **714**, 94–98 (2019).
49. Sun, C., Tang, T. & Uludağ, H. Molecular dynamics simulations of PEI mediated DNA aggregation. *Biomacromolecules* **12**, 3698–3707 (2011).
50. Meneksedag-Erol, D., Tang, T. & Uludağ, H. Mechanistic insights into the role of glycosaminoglycans in delivery of polymeric nucleic acid nanoparticles by molecular dynamics simulations. *Biomaterials* **156**, 107–120 (2018).
51. 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.* **100**, 2754–2763 (2011).
52. Ziebarth, J. D., Kennetz, D. R., Walker, N. J. & Wang, Y. Structural comparisons of PEI/DNA and PEI/siRNA complexes revealed with molecular dynamics simulations. *J. Phys. Chem. B.* **121**, 1941–1952 (2017).
53. Nademi, Y., Tang, T. & Uludağ, H. Steered molecular dynamics simulations reveal a self-protecting configuration of nanoparticles during membrane penetration. *Nanoscale* **10**, 17671–17682 (2018).
54. Nademi, Y., Tang, T. & Uludağ, H. Membrane lipids destabilize short interfering ribonucleic acid (siRNA)/polyethylenimine nanoparticles. *Nanoscale* **12**, 1032–1045 (2020).
55. Duncan, S. L., Dalal, I. S. & Larson, R. G. Molecular dynamics simulation of phase transitions in model lung surfactant monolayers. *Biochim. et Biophys. Acta (BBA) - Biomembr.* **1808**, 2450–2465 (2011).
56. Baoukina, S., Tieleman, D. P., Lung Surfactant Protein, S. P. B. & Promotes Formation of Bilayer reservoirs from monolayer and lipid transfer between the interface and Subphase. *Biophys. J.* **100**, 1678–1687 (2011).
57. Baoukina, S., Monticelli, L., Amrein, M. & Tieleman, D. P. The molecular mechanism of monolayer-bilayer transformations of lung surfactant from molecular dynamics simulations. *Biophys. J.* **93**, 3775–3782 (2007).
58. Luo, Z. et al. The role of nanoparticle shape in translocation across the pulmonary surfactant layer revealed by molecular dynamics simulations. *Environ. Sci. Nano*. **5**, 1921–1932 (2018).
59. Chen, P., Zhang, Z., Gu, N. & Ji, M. Effect of the surface charge density of nanoparticles on their translocation across pulmonary surfactant monolayer: A molecular dynamics simulation. *Mol. Simul.* **44**, 85–93 (2018).
60. Xu, Y. et al. Role of lipid coating in the transport of nanodroplets across the pulmonary surfactant layer revealed by molecular dynamics simulations. *Langmuir* **34**, 9054–9063 (2018).
61. Monticelli, L. et al. The MARTINI coarse-grained force field: Extension to proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).
62. Wei, Z. & Luijten, E. Systematic coarse-grained modeling of complexation between small interfering RNA and polycations. *J. Chem. Phys.* **143**, 243146 (2015).
63. Martínez-Calle, M., Parra-Ortiz, E., Cruz, A. & Olmeda, B. Pérez-Gil, J. towards the molecular mechanism of pulmonary surfactant protein SP-B: At the crossroad of membrane permeability and interfacial lipid transfer. *J. Mol. Biol.* **433**, 166749 (2021).
64. Sun, C., Tang, T. & Uludag, H. A molecular dynamics simulation study on the effect of lipid substitution on polyethylenimine mediated siRNA complexation. *Biomaterials* **34**, 2822–2833 (2013).
65. 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.* **100**, 2754–2763 (2011).
66. Suh, J., Paik, H. J. & Hwang, B. K. Ionization of poly(ethylenimine) and poly(allylamine) at various ph′s. *Bioorg. Chem.* **22**, 318–327 (1994).
67. Vanommeslaeghe, K. et al. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
68. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & De Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B.* **111**, 7812–7824 (2007).
69. Mahajan, S. & Tang, T. Automated parameterization of Coarse-Grained polyethylenimine under a Martini framework. *J. Chem. Inf. Model.* **63**, 4328–4341 (2023).
70. Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D. & Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **6**, e1000810 (2010).
71. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
72. Wu, E. L. et al. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
73. Lee, J. et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 641a (2016).
74. Chen, Z. et al. Mir-342-3p regulates the proliferation and apoptosis of NSCLC cells by targeting BCL-2. *Technol. Cancer Res. Treat.* **20**, (2021).
75. Case, D. A. et al. AmberTools. *J. Chem. Inf. Model.* **63**, 6183–6191 (2023).
76. Uusitalo, J. J., Ingólfsson, H. I., Marrink, S. J. & Faustino, I. Martini coarse-grained force field: Extension to RNA. *Biophys. J.* **113**, 246–256 (2017).
77. Olmeda, B. et al. A model for the structure and mechanism of action of pulmonary surfactant protein B. *FASEB J.* **29**, 4236–4247 (2015).
78. Liekkinen, J. et al. Pulmonary surfactant lipid reorganization Induced by the adsorption of the Oligomeric surfactant protein B complex. *J. Mol. Biol.* **432**, 3251–3268 (2020).
79. Labute, P. & Protonate, D. Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins Struct. Functi. Bioinf.* **75**, 187–205 (2009).
80. Nielson, D. W., Goerke, J. & Clements, J. A. Alveolar subphase pH in the lungs of anesthetized rabbits. *Proc. Natl. Acad. Sci. U S A.* **78**, 7119–7123 (1981).
81. De Jong, D. H. et al. Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687–697 (2013).
82. Jorgensen, W. L. Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *J. Am. Chem. Soc.* **103**, 335–340 (1981).
83. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
84. 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.* **81**, 3684–3690 (1984).
85. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
86. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
87. York, D. M., Darden, T. A. & Pedersen, L. G. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *J. Chem. Phys.* **99**, 8345–8348 (1993).
88. Noid, W. G. & Perspective Coarse-grained models for biomolecular systems. *J. Chem. Phys.* **139**, 090901 (2013).
89. Tironi, I. G., Sperb, R., Smith, P. E. & van Gunsteren W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459 (1995).
90. Abraham, M. J. et al. High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX 1–2*. **GROMACS**, 19–25 (2015).

91. Rosta, E. & Hummer, G. ree energies from dynamic weighted histogram analysis using unbiased Markov state model. *J. Chem. Theory Comput.* **11**, 276–285 (2015).
92. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **97**, 50–58 (2009).
93. Evans, D. J. & Holian, B. L. The nose–Hoover thermostat. *J. Chem. Phys.* **83**, 4069–4074 (1985).
94. 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.* **11**, 3932-3945 (2015).
95. Hub, J. S., De Groot, B. L. & Van Der Spoel D. G-whams-a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010).
96. Dahl, A. C. E., Chavent, M. & Sansom, M. S. P. Bendix: Intuitive helix geometry analysis and abstraction. *Bioinformatics* **28**, 2193–2194 (2012).
97. Negami, T., Shimizu, K. & Terada, T. Coarse-grained molecular dynamics simulations of protein-ligand binding. *J. Comput. Chem.* **35**, 137144 (2014).
98. Delort, B. et al. Coarse-grained prediction of peptide binding to G-Protein coupled receptors. *J. Chem. Inf. Model.* **57**, 562-571 (2017).
99. Dandekar, B. R. & Mondal, J. Capturing protein-ligand recognition pathways in Coarse-Grained simulation. *J. Phys. Chem. Lett.* **11**, 5302-5311 (2020).
100. Javanainen, M., Martinez-Seara, H. & Vattulainen, I. Excessive aggregation of membrane proteins in the Martini model. *PLoS One* **12**, e0187936 (2017).
101. Stark, A. C., Andrews, C. T. & Elcock, A. H. Toward optimized potential functions for protein-protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the MARTINI coarse-grained force field. *J. Chem. Theory Comput.* **9**, 4176–4185 (2013).
102. Srinivasan, S., Zoni, V. & Vanni, S. Estimating the accuracy of the MARTINI model towards the investigation of peripheral protein-membrane interactions. *Faraday Discuss.* **232**, 131-148 (2021).
103. Globisch, C., Krishnamani, V., Deserno, M. & Peter, C. Optimization of an Elastic Network Augmented Coarse Grained Model to Study CCMV Capsid deformation. *PLoS One* **8**, e60582 (2013).
104. Balali-Mood, K., Bond, P. J. & Sansom, M. S. P. Interaction of monotopic membrane enzymes with a lipid bilayer: a coarse-grained MD simulation study. *Biochemistry* **48**, 2135–2145 (2009).
105. Casalini, T. Not only in silico drug discovery: molecular modeling towards in silico drug delivery formulations. *J. Controlled Release* **332** 390-417 (2021).
106. Kjolbye, L. R. et al. Towards design of drugs and delivery systems with the Martini coarse-grained model. *QRB Discov.* **3**(e19), 1–14, (2022).
107. Szefler, B., Diudea, M. V., Putz, M. V. & Grudzinski I. P. Molecular dynamic studies of the complex polyethylenimine and glucose oxidase. *Int. J. Mol. Sci.* **17**, 1796 (2016).
108. Sever, N., Miličić, G., Bodnar, N. O., Wu, X. & Rapoport, T. A. Mechanism of Lamellar Body formation by lung surfactant protein B. *Mol. Cell.* **81**, 49–66 (2021).

Acknowledgements

Digital Research Alliance of Canada is acknowledged for providing the computing resources and technical support. TT acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC; Grant Numbers: RGPIN-2018-04281) and Canada Research Chairs Program (Grant Number: TIER1 2021-00023). The all-atom structure of surfactant protein B was generously provided by Dr. Bárbara Olmeda Lozano and is greatly appreciated.

Author contributions

GL and NH—Investigation, formal analysis, writing of the original draft. SM—investigation, formal analysis, reviewing and editing of the writing. TT—Supervision, funding, resources, and reviewing and editing of the writing. All authors read and approved the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Supporting information

Remaining comparisons between AA and CG bonded interactions for PEI-CA; Additional details on the interactions between PEI-CA and SP-B.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-88848-x.

Correspondence and requests for materials should be addressed to T.T.