

CHARMM Force Field and Molecular Dynamics Simulations of Protonated Polyethylenimine

Titus Adrian Beu * and Alexandra Farcaş

As a gene delivery vector, polyethylenimine (PEI) shows one of the highest transfection efficiencies, while effectively protecting DNA from enzyme degradation. The distinctive charge pattern of protonated PEI is widely considered responsible for fundamental process such as DNA condensation into PEI/DNA polyplexes (which are able to enter cells via endocytosis), proton sponge effect (which triggers the release of polyplexes from endosome), and release of DNA from polyplexes (to be further processed inside the nucleus). Our investigations are largely motivated by the crucial need for a realistic molecular mechanics force field (FF) for PEI, and, accordingly, we focus on two major issues: (1) development of a new atomistic (CHARMM) FF for PEI in different protonation states, rigorously derived from high-quality *ab initio* calculations performed on model polymers, and

(2) molecular dynamics investigations of solvated PEI, providing a detailed picture of the dynamic structuring thereof in dependence on their size and protonation state. The modeled PEI chains are essentially described in terms of gyration radius, end-to-end distance, persistence length, radial distribution functions, coordination numbers, and diffusion coefficients. They turn out to be more rigid than in other computational studies and we find diffusion coefficients in fair agreement with experimental data. The developed atomistic FF proves adequate for the realistic modeling of the size and protonation behavior of linear PEI, either as individual chains or composing polyplexes. © 2017 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24890

Introduction

Gene therapy continues to stir tremendous interest due to its enormous treatment potential for diverse, acquired or inherited diseases.^[1] An effective drug based on a nucleic acid must reach inside the target cell while surviving an entire sequence of biological defense systems, thus critically depending on an efficient delivery vector.^[2] As a general principle, gene carriers are designed to condense DNA/RNA molecules into nm- (μm)-sized complexes by way of specific electrostatic interactions, protecting the payload from enzymatic/non-enzymatic degradation and favoring cellular uptake.

Two main types of gene delivery vehicles are currently used in clinical and research studies: viral and non-viral (synthetic) vectors. Although very effective, the viral vectors pose significant challenges and risks, such as low DNA/RNA loading capacity, immunogenicity, and toxicity. From this perspective, the development of *non-viral* vectors is considered crucial and, given the multiple benefits and recent advances in their fabrication technologies, very promising.^[1–5] They can be designed and synthesized to display essential properties, such as transfection efficiency, specificity, and biocompatibility. While significant progress in engineering non-viral gene vectors has been achieved, clinical results are far from being substantial. The difficulties include among others instability of the complexes in extracellular space, insufficient cellular uptake, and low endosomal escape.^[1,6]

Non-viral gene delivery systems are currently based on polyplexes—DNA complexes with cationic polymers,^[7,8] lipoplexes—cationic lipids, polymeric vesicles,^[9] or liposomes. Owing to their accessible chemistry, cost effectiveness, and controllable

toxicity,^[10] cationic polymers appear to be the most attractive class of non-viral vectors. The most commonly used are polyethylenimine (PEI), poly(L-lysine), cationic dendrimers, glycopolymers, poly(amidoamine), and chitosan.^[11] The electrostatic interactions between the positive amino groups of these polymers and the negative phosphate groups of DNA result in condensed polyionic complexes (polyplexes), which protect DNA from degradation.

Polyethylenimine (PEI: $-(\text{NH} - \text{CH}_2 - \text{CH}_2)_n-$) is composed of repeating NH groups separated by double aliphatic spacers, occurring in linear or branched configurations. PEI was initially used as a gene delivery vector in 1995 by Boussif et al.^[11] and has been recurrently reported to have high transfection efficiency, being extensively used in a variety of drug delivery protocols ever since. The primary, secondary, and tertiary amines of PEI create a considerable buffering capacity, which is suitable for condensing large, negatively charged molecules, such as DNA, in the form of polyplexes. Condensation provides an extremely efficient protection against degradation by cellular nuclease, especially for positively charged polyplexes,^[12] and, due to its particular charge distribution, PEI more effectively protects DNA from enzyme degradation than other cationic polymers. Polyplexes are able to enter cells via

T. A. Beu, A. Farcaş

Faculty of Physics, Department of Biomolecular Physics, University Babeş-Bolyai, Mihail Kogălniceanu Street 1, Cluj-Napoca 400084, Romania
E-mail: titus.beu@phys.ubbcluj.ro

Contract grant sponsor: Executive Unit for Financing Higher Education, Research, Development and Innovation (UEFISCDI), project PN-III-P4-ID-PCE-2016-0474

© 2017 Wiley Periodicals, Inc.

endocytosis, however, the details of the intracellular transport from endosomes to the nucleus remain largely unclear. The large buffering capacity of PEI is at the core of the “proton-sponge” hypothesis, which, although debated, remains the most widely accepted mechanism for PEI escaping from the endosome^[6,13] and releasing the payload.

The transfection efficiency of polyplexes appears to be enhanced by PEI of higher molecular weight,^[12] which, however, increases toxicity, causing necrosis.^[10] This finding obviously advocates for lower molecular weights. Even though branched PEI condenses DNA more efficiently than linear PEI of identical molecular weight,^[10] the latter still proves to be an efficient gene delivery agent under *in vivo* conditions. Optimizing its size and protonation pattern is thus essential for designing effective gene delivery protocols.

Despite the recognized importance of the condensation process and proton sponge effect, the dynamics of PEI is only qualitatively understood. Even the optimal protonation states remain largely disputed and numerical simulations have not been able to sufficiently advance our knowledge in this respect so far. Only few computational studies have dealt with solvated PEI, and even fewer with PEI/DNA complexes, obviously due to the parametrization difficulties of a realistic FF for PEI.

In an elaborate approach to explain the proton sponge effect, Ziebarth et al.^[14] investigated the protonation behavior of PEI in solutions and in the presence of DNA using Monte Carlo simulations. Atomistic parameters for linear PEI 20-mers were determined from molecular dynamics (MD) simulations with the general AMBER GAFF force field (FF)^[15] (not particularly developed for PEI), while charges for entire PEI units were adjusted by the RESP method^[16] from fits to *ab initio* gas phase calculations. The atomistic results were then used to parametrize a coarse-grained bead-spring model. In the ensuing coarse-grained Monte Carlo calculations, the conformational relaxation of linear PEI and the variation in protonation states were studied under different ionic conditions. The computed PEI titration curves in solutions notably showed a too strong dependence on salt concentration as compared to experiments.

The MD simulations of Sun et al.^[17] on DNA/PEI complexes were focused on the effect of PEI branching and protonation state. PEI 13-mers with different branching degrees were simulated in solutions with counterions and then composing DNA/PEI complexes. The used CHARMM27 FF does not provide specific residues for PEI, and they were added by analogy with other residues, while the torsional parameters were somewhat improved by average-quality fits to *ab initio* data. Test simulations yielded results similar to those of Ziebarth et al., suggesting a semi-quantitative realism of the FF.

The MD studies of Choudhury et al.^[18] on the solvation dynamics of linear PEI in different protonation states, were based on the same Amber FF as used by Ziebarth et al., without notable improvements. The simulated structural properties for two different chain lengths (20-mer and 50-mer) revealed that the polymer is pronouncedly coiled at high pH, while under strong acidic conditions the chains are elongated.

In the course of the manuscript preparation, we became aware of the very recent MD study of Wei et al.,^[19] dealing

with the coarse-grained modeling of the complexation between RNA and polyethylene glycol-grafted linear PEI. Therein, as a preparatory step for developing the actual coarse-grained MARTINI FF, a significant step is taken further compared to previous investigations for developing a rigorous atomistic CHARMM FF for PEI, noteworthy, based on methodological elements similar to the ones used in our work. Even so, quite a few important differences persist, among which the most significant are: (1) we define different atom types and residues; (2) we adjust the whole set of CHARMM FF parameters, not only the dihedral contributions; (3) we analyze a broader set of observed quantities for characterizing the atomistic behavior of PEI, which is actually not the main focus of the paper of Wei et al. and is rather described in a Supporting Information.^[20] Nevertheless, several of the results therein provide useful counterparts to our calculations.

In spite of the remarkable recent computational efforts, the structural dynamics of PEI remains an open issue, marked by a stringent need for a realistic FF, desirably based on high quality *ab initio* calculations and the most advanced insights in the field of gene delivery. This was the main incentive for our studies, and, in essence, (1) we parametrized a new additive CHARMM FF for protonated PEI, rigorously derived from high-quality *ab initio* calculations, and (2) we used the developed FF to investigate the structural dynamics of solvated PEI of various sizes and protonation patterns by extensive MD simulations, which were analyzed in terms of various synthetic parameters, such as gyration radius, end-to-end distance, persistence length, radial distribution functions (RDFs), coordination numbers, and diffusion coefficients.

Methodology

CHARMM force field

CHARMM^[21,22] is a widely used additive atomistic FF model for MD. Owing to its comprehensive and versatile functional form—perfectly equivalent to that of other well-established FFs, and to the availability of parameters for a huge, ever growing number of topical biomolecular systems, the CHARMM FF is implemented in several of the state-of-the-art MD simulation packages. The potential energy of the CHARMM FF comprises bonded and non-bonded interactions:

$$U_{\text{bonded}} = \sum_{\text{bonds}} k_b (b - b_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{dihedrals}} k_\psi [1 + \cos(n\psi - \delta)] + \sum_{\text{impropers}} k_\omega (\omega - \omega_0)^2 + \sum_{\text{Urey-Bradley}} k_{\text{UB}} (b^{1,3} - b_0^{1,3})^2 \quad (1)$$

$$U_{\text{non-bonded}} = \sum_{\text{atoms } i,j} \frac{q_i q_j}{\epsilon_0 r_{ij}} + \epsilon_{ij} \left[\left(\frac{r_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right], \quad (2)$$

where

$$r_{ij}^{\min} = (r_i^{\min} + r_j^{\min})/2, \quad \epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}. \quad (3)$$

The bonded terms model bonds, angles, dihedral angles, improper dihedral angles, and Urey-Bradley interactions, whereby k_b , k_θ , k_ψ , k_{cb} , and k_{UB} are the corresponding force constants. b_0 and θ_0 are, respectively, the equilibrium bond lengths and angles, and the dihedral terms feature the multiplicities n and phases δ . It should be noted that, given the particular atomic structure, for realistically modeling bonded interactions within the PEI chains it is sufficient to consider bonds, angles, and dihedrals, ignoring improper angles and Urey-Bradley terms. Non-bonded interactions are represented by electrostatic and van der Waals terms, where q_i are atomic charges, ϵ_{ij} are well depths, and r^{\min} are the corresponding Lennard-Jones radii.

Adjustment of the CHARMM force field

Fitting molecular mechanics (MM) FFs, in general, and CHARMM FFs, in particular, to quantum mechanical (QM) calculations is by no means trivial. Given the considerable number of FF parameters to be adjusted already for molecules with tens of atoms, the uncertainties rapidly increase with system size. Explicitly conceived for "small" molecules and based on a rigorous methodology, fTK (Force Field Tool Kit)^[23,24] is a state-of-the-art software package designed for adjusting CHARMM FF parameters (charges, bonds, angles, and dihedrals) to high quality QM target data generated using the Gaussian package.^[25] fTK can be conveniently operated as a plug-in of the molecular visualization/analysis application VMD.^[26]

An essential feature of CHARMM FFs is their additivity, which enables building chemical compounds of virtually arbitrary sizes from predefined *residues* (essentially, functional groups), composed, in their turn, of specific *atom types*. Technically, developing a CHARMM FF requires: (1) defining residues, atom types, and their connectivity in a *topology file*, and (2) adjusting FF parameters for all atom types and collecting them in a *parameter file*. The realism and versatility of the FF crucially depend on the educated choice of the atom and residue types.

We developed the CHARMM FF based on QM target data generated with Gaussian 09 Rev. A.01^[25] for model polymers, in runs at MP2 level (second-order Møller-Plesset perturbation theory) and using the 6-31G(d) basis set. Following the methodology for developing CHARMM FFs developed by Vanommeslaeghe et al.^[22] and using the fTK v1.1 tool kit^[23,24] to manage the entire process (including the preparation of all the Gaussian input files), we carried out the steps outlined below for each of two model PEI tetramers (one unprotonated and one protonated).

1. The non-bonded Lennard-Jones parameters of each atom type were assigned values from similar atom types defined in the CGenFF component of the latest version of the CHARMM FF (CHARMM36).^[27]

2. The partial atomic charges were optimized for solvated states using QM PEI-water interaction profiles. The latter were generated from separate QM optimizations of PEI-single-H₂O-molecule complexes, each started with an optimized PEI and a single H₂O test molecule opposing one of the PEI atoms accessible as hydrogen-bond donor or acceptor. The resulted QM sets of interaction energies and dipole moments were jointly used as target data for a subsequent iterative fitting procedure of the partial atomic charges.
3. The bond- and angle parameters were adjusted by generating distortions of the internal coordinates (ICs) of each PEI model and monitoring the corresponding total energy changes. Technically, QM and MM Hessian matrices can be generated in terms of ICs by imposing small distortions thereof, δq_i , in opposing directions and evaluating the corresponding second-order energy change. The total QM and MM distortion energies, cumulated over all ICs, can then be defined by

$$\Delta E_i = \sum_j \frac{\partial^2 E}{2 \partial q_i \partial q_j} \delta q_i \delta q_j, \quad (4)$$

and they were matched within fTK in an iterative procedure finally yielding the optimized bond- and angle parameters.

4. The dihedral parameters were optimized using torsion scans. As the functional form of the dihedral FF terms is not harmonic, the Hessian approach cannot be used in their case. Instead, QM Potential Energy Surfaces (PESs) were generated by explicitly scanning the dihedrals of interest. The subsequent optimization of the dihedral parameters was performed via a simulated annealing protocol minimizing the difference between the QM and MM PESs for all scanned dihedrals.

Molecular dynamics

The methodology of choice for realistic numerical simulations of complex biomolecular systems is doubtlessly MD. For our atomistic investigations, we used NAMD,^[28] which, among other state-of-the-art FFs, also implements CHARMM.

NAMD relies on the velocity Verlet integrator to advance the positions and velocities of the atoms, and we used it in all simulations with a time step of 2 fs in conjunction with the SHAKE algorithm [] to constrain all bonds involving H atoms to their nominal length. Throughout, we applied a cutoff distance of 12 Å and a switching function setting in at 10 Å to all short-range non-bonded (van der Waals and real-space electrostatic) interactions.

Periodic boundary conditions were applied to minimize finite size effects and the long-range electrostatic interactions were treated with the Particle Mesh Ewald^[29] method, using a grid spacing of 1 Å. Without significantly affecting the accuracy, the k -space electrostatic contributions were updated only every second step. The temperature was fixed at 310 K using a Langevin thermostat with a damping coefficient of 1 ps⁻¹, while the pressure was kept constant at 1 atm using a Langevin piston.

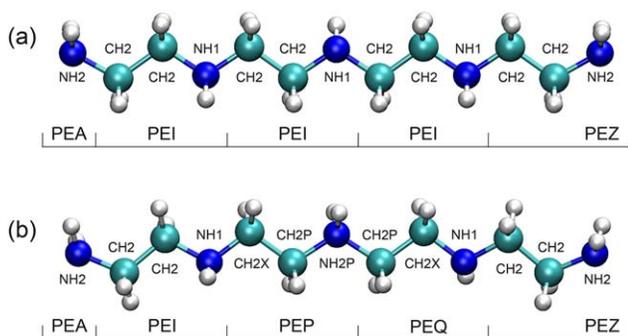


Figure 1. PEI tetramer models used in the parametrization of the CHARMM force field (a) unprotonated and (b) protonated at the central nitrogen), showing the composing residues and atom types. [Color figure can be viewed at wileyonlinelibrary.com]

Results and Discussion

Parametrization of the CHARMM force field

An essential component in any attempt to realistically describe the dynamics of solvated PEI, as such or as part of more complex biomolecular systems, is a reliable MM FF based on QM calculations. The bare quality of the underlying QM data provides, although, no guarantee for the quality of the MM FF, by itself. Besides the already complex polymer topology, difficulties also arise from solvation, the full QM description of which is beyond reach for system sizes of practical interest. Further complications are brought into play by the extensive electrostatic interactions stemming from the charged sites of PEI, water, and counterions. Moreover, directly using site charges provided by static QM models is of little relevance, as they cannot grasp the dynamical behavior of PEI.

Our aim was to construct an additive CHARMM model for linear protonated/unprotonated PEI chains of arbitrary lengths using as few building blocks (residues and atom types) as necessary. Minding the risk of obtaining non-reliable FF parameters for too large chemical compounds, we adjusted our CHARMM FF relative to QM data for two minimal model polymers, which preserve the essential features of protonated/unprotonated PEI chains. Considering that protonation basically replaces an NH group with an NH_2^+ group, we specifically considered two linear tetramer models, PEI4p0—unprotonated (Fig. 1a) and PEI4p1—mono-protonated at the central NH_2^+ group (Fig. 1b). They are composed of 36 and 37 atoms, respectively.

To account for the various vicinities of the involved atoms, yet maintaining a manageable number of FF parameters, we defined the following atom types for modeling PEI chains: NH2—N atom of terminal NH_2 group; HN2—H atom bonded to NH2; NH1—N atom of NH group; HN1—H atom bonded to NH1; NH2P—N atom of protonated NH_2^+ group; HN2P—H atom bonded to NH2P; CH2—C atom adjacent to NH1; CH2P—C atom adjacent to NH2P; CH2X—C atom bridging NH1 and CH2P; HC2—H atom bonded to CH2, CH2P, or CH2X.

Based on the above atom types, we defined five residue types: PEI (backbone $\text{CH}_2\text{-CH}_2\text{-NH1}$)—generic, unprotonated $\text{CH}_2\text{-CH}_2\text{-NH}$ monomer; PEA (backbone NH2)— NH_2 group starting a polymer chain, PEZ (backbone NH2)—patch replacing NH by

NH_2 in the terminal PEI of a chain; PEP (backbone $\text{CH}_2\text{-CH}_2\text{-NH}_2^+$)—charged $\text{CH}_2\text{-CH}_2\text{-NH}_2^+$ monomer starting a protonated segment; PEQ (backbone $\text{CH}_2\text{-CH}_2\text{-NH1}$)—partially charged $\text{CH}_2\text{-CH}_2\text{-NH}$ monomer ending a protonated segment.

It is noteworthy, that the residues PEP and PEQ come in pairs, ensuring a gradual transition to non-protonated segments modeled by PEI. A single protonated NH_2^+ group is actually modeled as a PEP-PEQ sequence, with the unitary protonation charge distributed over both residues and decaying symmetrically about the central N atom. In fact, the local symmetry of the backbone, namely $\text{-NH1-(CH}_2\text{X-CH}_2\text{P-NH}_2\text{)-(CH}_2\text{P-CH}_2\text{X-NH1)}$ underpins the symmetric charge distribution, while the distinction between CH2P and CH2X enables a gradual transition.

Quite in contrast with our residues based on standard functional groups (with C-C-N backbone), Wei et al.^[19] defined residues with symmetric backbone, C-N-C. Symmetric residues are, indeed, appealing for developing coarse-grained residue-based variants of FFs. Nevertheless, they assume by default symmetric protonation environments, making their use for arbitrary protonation patterns difficult. The PEI models used by Wei et al. for parametrizing the FF were trimers, with the central N-groups closer and more sensitive to the terminal groups, as compared to our tetramers. The atom types used by Wei et al. were adopted by analogy from among those defined by the standard CHARMM general FF (CGenFF), by minimizing a “penalty score.” Considering standard (non-specific) values for the charges, Lennard-Jones-, bond-, and angle parameters, Wei et al. solely optimized the dihedral contributions for PEI, using the same fTK application.

Once the residues and atom types are defined, one can proceed with the parametrization procedure implemented in fTK. One crucial insight is that the various classes of parameters (charges, bonds, angles, and dihedrals) are intimately interrelated, and thus adjusting them individually is not guaranteed to realistically reproduce the dynamical behavior of the modeled system. Therefore, we jointly optimized all the parameters by strictly following the work flow outlined by Vanommeslaeghe et al.^[22]

As a general rule concerning all classes of bonded interactions (bonds, angles, and dihedrals), we note that for the force constants involving the same atom types both in the unprotonated PEI4p0 and the protonated PEI4p1 models, we consistently considered in the final set of parameters the values determined from the unprotonated model.

Assignment of Lennard-Jones parameters by similarity

fTK offers no built-in procedure for parameterizing van der Waals interactions, as such. Instead, according to the general methodology, we assigned Lennard-Jones parameters to each atom type by analogy, from similar atom types (aliphatic C and H atoms for CH_2 groups, N and H atoms for methylamine/ NH_2^+ groups) provided by CGenFF,^[27] as listed in Table S1 of the Supporting Information.

Optimization of partial atomic charges from water-interaction profiles

The charge fitting procedure for each model PEI used as reference the geometry-optimized QM configuration, which was

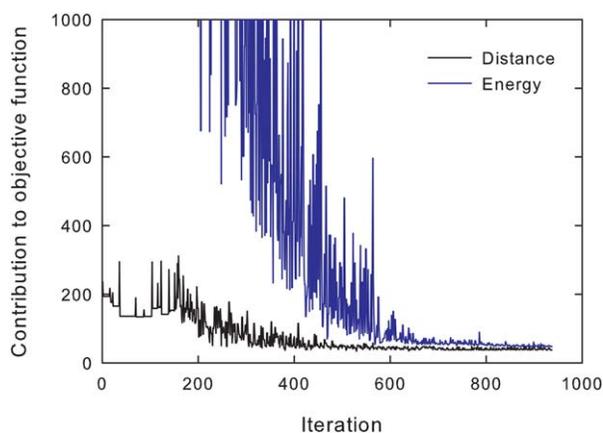


Figure 2. Evolution of the distance and energy contributions to the objective function during the optimization procedure of the partial atomic charges for the unprotonated model tetramer PEI4p0. [Color figure can be viewed at wileyonlinelibrary.com]

used to build PEI-single-H₂O-molecule complexes for each water-accessible PEI atom. For each complex, the H₂O molecule was oriented ideally for hydrogen bonding, minimizing the steric repulsion with the neighboring atoms. The two parameters relevant for hydrogen bonding—the distance between the H₂O molecule and the target atom and the rotation angle of the H₂O molecule about the line connecting them—were then optimized QM, with all other degrees of freedom constrained. To maintain consistency with the remainder of the CHARMM additive FF, the corresponding QM target data was calculated at the HF/6–31G(d) level of theory.

By default, aliphatic hydrogen atoms were assigned the standard CHARMM charge +0.09, and, even although they are involved in calculating the water-PEI interaction energies, they

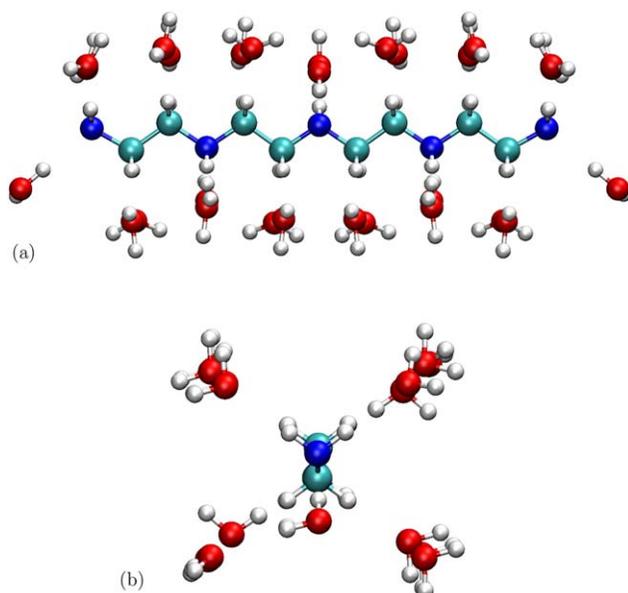


Figure 3. Superimposed optimized configurations of the complexes formed by the unprotonated tetramer PEI4p0 with the test water molecules used for optimizing the atomic charges (side view a) and axial view b)). [Color figure can be viewed at wileyonlinelibrary.com]

Table 1. Residues, atom types, and charges for the unprotonated model PEI4p0.

Residue/atom types	No. of atoms	Q_{ffTK} e	Q_{final} e	ΔQ e
PEA ^[a] , PEZ ^[b]		−0.275	−0.224	0.051
NH2	1	−0.999	−0.948	0.051
HN2	2	0.362	0.362	0.000
PEI		0.003	0.000	−0.003
CH2	2	0.045	0.044	−0.001
HC2	4	0.090	0.090	0.000
NH1	1	−0.795	−0.796	−0.001
HN1	1	0.348	0.348	0.000
PEI terminal		0.221	0.224	0.003
CH2	2	0.068	0.044	−0.024
HC2	4	0.090	0.090	0.000
NH2	1	−0.999	−0.948	0.051
HN2	2	0.362	0.362	0.000

[a] Default patch added before the first PEI residue of the chain. [b] Default patch replacing the NH1–HN1 group in the terminal PEI residue. Q_{ffTK} are the optimized charges yielded by ffTK, Q_{final} are slightly rescaled values which ensure neutrality of the PEI residue and of any unprotonated chain as a whole, and $\Delta Q = Q_{\text{final}} - Q_{\text{ffTK}}$.

were excluded from the optimization. The resulted sets of QM-optimized distances, interaction energies, and dipole moments were jointly used as target data for the subsequent iterative optimization of the partial atomic charges by means of a modified Simplex method, imposing, in addition, neutrality for the unprotonated model PEI4p0, and a total charge +1 for the protonated model PEI4p1. According to the standard procedure, to better approximate the bulk-phase, the QM-optimized distances were shifted by -0.2 \AA , and the interaction energies for neutral molecules were scaled by 1.16.^[23] For the protonated tetramer, we actually performed optimizations with various energy scaling coefficients between 1 and 1.16 and distance shifts between 0 and -0.2 \AA , resulting in an overall limited variability of the optimized charges. The best consistency between the *unprotonated* residues in our two model tetramers was, however, achieved with the same energy scaling and distance shifts. The convergence of the optimization procedure is illustrated in Figure 2 by the evolution of the distance and energy contributions to the objective function. The superimposed configurations of all the optimized PEI₀–H₂O complexes are depicted in Figure 3.

The atomic charges of the unprotonated model PEI4p0 optimized using ffTK (Q_{ffTK} in Table 1), still needed to be slightly rescaled. This was necessary (1) to ensure perfect neutrality of the generic PEI residue, so that it can be used as an independent building block, and (2) to compensate the charges of the starting PEA residue (NH₂ group) and of the terminal PEI (in which the NH unit is replaced by the NH₂ group), so as to preserve neutrality of any unprotonated chain, irrespective of size. For the atom types composing the PEI residue, the final readjusted charges, Q_{final} , differ marginally from Q_{ffTK} , namely by not more than 0.001, while the whole, strictly neutral residue, is affected by a charge difference of only -0.003 . To meet the second constraint, more significant charge changes were necessary for the CH₂ and NH₂ atom types of the terminal PEI

Table 2. Protonated residues and atom types composing the PEI4p1 tetramer model and their optimized charges.

Residue/atom types	No. of atoms	Q_{ffTK}	Q_{final}	ΔQ
		e	e	e
PEP		0.653	0.695	0.042
CH2X	1	-0.064	-0.022	0.042
CH2P	1	0.415	0.415	0.000
HC2	4	0.090	0.090	0.000
NH2P	1	-0.968	-0.968	0.000
HN2P	2	0.455	0.455	0.000
PEQ		0.340	0.305	-0.035
CH2P	1	0.415	0.415	0.000
CH2X	1	-0.064	-0.022	0.042
HC2	4	0.090	0.090	0.000
NH1	1	-0.743	-0.796	-0.053
HN1	1	0.372	0.348	-0.024
Total charge		0.993	1.000	0.007

residue, but these amount to only -0.024 and 0.051 , respectively. While the starting PEA residue is made less negative by $+0.051$, the charge of the terminal PEI as a whole ($+0.221$) is readjusted by only $+0.003$.

The final atomic charges of the unprotonated residues PEA, PEI, and PEZ were used as a basis for establishing the atomic charges for the protonated pair of residues, PEP and PEQ (see Table 2). Along this line, the charges assigned to NH1 and HN1 as part of PEQ were the same as within PEI. The only supplementary change concerned CH2X ($+0.042$) and it ensured that the cumulated charge of PEP and PEQ exactly amounts to $+1$. In view of the nearly unitary charge ($+0.993$) yielded by ffTK, the approach is indeed justified, implying that protonation is largely limited to the PEP-PEQ group, not affecting the neighboring residues. This essentially implied a transfer of charge from PEQ to PEP of about $+0.04$, increasing the charge of the latter by 6%. Finally, PEP and PEQ carry the charges $+0.695$ and $+0.305$.

The very different charges yielded by ffTK for the aliphatic C atoms composing PEP and PEQ, namely $+0.415$ for CH2P and -0.064 for CH2X, fully warrant our choice of considering two distinct carbon types for modeling the gradual charge decay about the protonated NH_2^+ group.

Regarding the possibility exploited by Wei et al.^[19] of defining a protonated residue with a symmetric C-N-C backbone, carrying the entire unitary protonation charge, we note that the sum of the corresponding atomic charges (including the attached H atoms) yielded by ffTK amounts to $+1.132$. This value does not realistically justify limiting the protonation charge to a single C-N-C residue, or, else, suggests the need for supplementary, partially charged, non-symmetric residues on either side, to account for the decaying charge distribution.

Optimization of bond and angle parameters from distortions along internal coordinates

As the Hessian matrix characterizes the local curvature of a PES, it can be used to describe, within the harmonic approximation, energy changes of a system caused by small atomic

displacements about the equilibrium configuration. In essence, comparing MM and QM energy changes enables one to adjust MM FF parameters. Defining, however, the Hessian matrix in terms of individual atomic displacements is not useful for parametrizing a CHARMM FFs, as the latter describes atomic interactions in terms of ICs, defined as bonds, angles, dihedrals, and impropers. The use of independent atomic displacements, instead of simplifying the procedure (as it may appear), couple several ICs, making impossible the unambiguous assignment of FF parameters.

Taking advantage of Gaussian's ability to yield the QM Hessian matrix in the representation of ICs, ffTK evaluates the MM Hessian matrix from distortions along the ICs, as second-order changes of the total energy. The total QM and MM distortion energies, cumulated over all ICs, are defined according to eq. (4) as discrete second-order total energy differentials, and they are matched in an iterative procedure yielding the optimized bond- and angle parameters.

To illustrate the scale of the reduction of the difference between the MM and QM equilibrium bond lengths, we plot in Figure 4 the profiles corresponding to the initial and final optimization steps for PEI4p0. Similarly, Figure S1 (in the Supporting Information) depicts the initial and final profiles of the MM-QM deviations for all the equilibrium angles. While the MM-QM bond deviations reduce by more than one order of magnitude (from a maximum of approximately 0.015 \AA to roughly 0.001 \AA), the deviations in the equilibrium angles decrease by more than 50% to, typically, less than 0.5° .

The optimized values of the equilibrium bond lengths and force constants for the backbone atoms are listed in Table S2 in the Supporting Information. As can be seen, along the backbone, CH2 binds stronger to NH1 than to the neighboring CH2. Methodologically, this would underpin the idea of a symmetric PEI residue, with CH2-NH1-CH2 backbone. Nevertheless, additional shortcomings to those already noted relative to the charge distribution, arising in the case of non-symmetrically protonated chains, will be shown to justify our functional group-based PEI residue, with CH2-CH2-NH1 backbone.

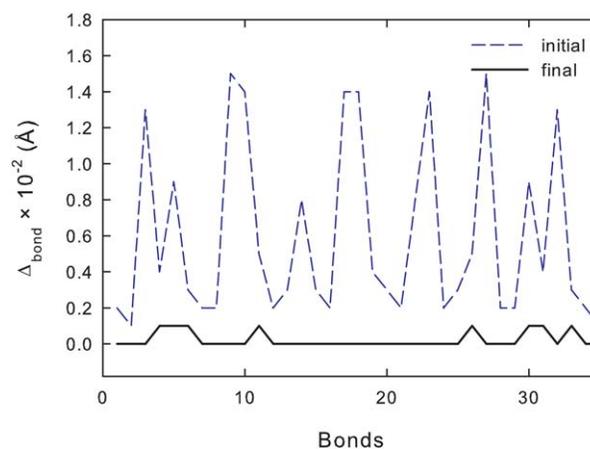


Figure 4. Deviation of the MM equilibrium bond lengths from the QM values for PEI4p0 in the initial and final optimization steps. [Color figure can be viewed at wileyonlinelibrary.com]

Table 3. Data defining the simulated configurations and average box size resulted from simulations.

Size -mer	PEI chain			Solvation		Average box size Å
	Proton. fraction	No. of atoms	Length Å	H ₂ O molecules	Cl ⁻ ions	
14	0	116	51.40	1348	0	17.28
	1/4	119	51.69	1341	3	17.25
	1/3	120	51.78	1340	4	17.25
	1/2	122	51.98	1338	6	17.25
26	0	212	95.34	6682	0	29.31
	1/4	218	95.92	6681	6	29.32
	1/3	220	96.11	6676	8	29.31
	1/2	224	96.50	6669	12	29.30
50	0	404	183.23	23032	0	44.22
	1/4	416	184.38	23016	12	44.21
	1/3	420	184.77	23007	16	44.21
	1/2	428	185.54	22999	24	44.21

Within protonated chains, the CH₂-CH₂ and CH₂X-CH₂P bonds are energetically similar. Conversely, while the equilibrium lengths of the NH₁-CH₂ and NH₁-CH₂X bonds is practically identical, NH₁ appears to bind by about 10% stronger to the intermediate species CH₂X, which mediates the connection with the protonated NH₂P. This obviously calls for asymmetric residues connecting unprotonated and protonated sites. Being asymmetric, both PEP and PEQ meet this requirement, while their sequence PEP-PEQ ensures the local symmetry about the protonated NH₂P.

Comparing the bonding of the unprotonated/protonated N species with their C backbone neighbors, one notices that the protonated bond NH₂P-CH₂P is notably weaker than the unprotonated one, NH₁-CH₂, a slightly increased equilibrium length (1.496 Å vs. 1.446 Å) being accompanied by an almost 20% lower force constant. This makes the definition of a *symmetric* protonated residue centered about N, with a weaker C-N-C backbone less practical. This is even more so as NH₂P-CH₂P is weaker even than the transition bond CH₂X-CH₂P.

The equilibrium parameters for angles resulted from the fTK adjustment procedure are given in Supporting Information Table S3. The C-N-C angle can be seen to progressively increase in sequence for the PEI, PEQ, and PEP residues. Specifically, the equilibrium values for CH₂-NH₁-CH₂, CH₂-NH₁-CH₂X, and CH₂P-NH₂P-CH₂P are 110.9°, 112.3°, and 114.2°, respectively. The C-C-N angle shows a similar increase from 109.7° for CH₂-CH₂-NH₁ in PEI to 111.5° for CH₂X-CH₂P-NH₂P in PEP. As expected, the fully extended length of PEI polymers increases with the protonation state (see Table 3), and this is in part caused by the slightly longer CH₂P-NH₂P bond. Contrary to the weaker CH₂P-NH₂P bond compared to its unprotonated counterpart CH₂-NH₁, the angles centered about N atoms become stiffer with protonation. In fact, the force constant for the CH₂P-NH₂P-CH₂P angle exceeds by about 9% the one for CH₂-NH₁-CH₂.

Unfortunately, the sequence of residues PEP-PEQ-PEP, necessary for modeling close but non-adjacent protonated sites, such as in the case of alternate protonation, is not present in the protonated model tetramer PEI₄p1. A longer model would be required to this end (such as, a doubly protonated PEI pentamer) and this we envision as a future development.

Therefore, the parameters for the CH₂X-NH₁-CH₂X angle are set by similarity equal to those for CH₂-NH₁-CH₂X.

Optimization of dihedral parameters using scans of torsion coordinates

As the functional form of the dihedral energy contributions is not harmonic [see eq. (1)], the Hessian-matrix approach used for bond angles is not applicable. Instead, the optimization procedure implemented in fTK uses scans of torsion coordinates, and the variables to be optimized are the force constants k_{ψ} and phases δ . The latter are fixed to either 0° or 180°, while the corresponding multiplicities n are set based on symmetry considerations regarding the involved functional groups. The used objective function measures the difference between the QM and MM PESs constructed from the explicit scans for the entire set of dihedrals of interest.

Technically, the QM scans imply for each chosen dihedral, a sequence of MP2/6-31G(d) geometry optimizations of the respective PEI model polymer (PEI₄p0 or PEI₄p1) for fixed dihedral angle distortions incrementally covering a symmetric angular range about the equilibrium configuration. Specifically, we distorted each dihedral angle involving the C and N atoms by values between -90° and 90° in steps of 5°. The concrete dihedrals which were subject to optimization are listed in Supporting Information Table S4, while the calculated QM and fitted MM torsion energy profiles for the dihedrals involving only backbone atoms are plotted in Figure 5 for PEI₄p0 and in Figure 6 for PEI₄p1.

The agreement between the MM torsion profiles and the QM references is appreciable. Specifically, while our fit for the CH₂-CH₂-NH₁-CH₂ dihedral of the PEI₄p0 model (Fig. 5b) is affected by a root-mean-square error (RMSE) equal to 0.17, Wei et al.^[19] report a RMSE of 0.59. We ascribe the better quality of our torsional contributions to the fact that we consistently adjusted *all* the bonded parameters, not only the dihedrals. Also, we used longer PEI models, thus reducing the influence of the terminal groups.

For PEI₄p0, the largest deviations between the MM and QM torsion profiles can be observed for the dihedrals involving the NH₂ atom of the ending NH₂ group (see Fig. 5a). This

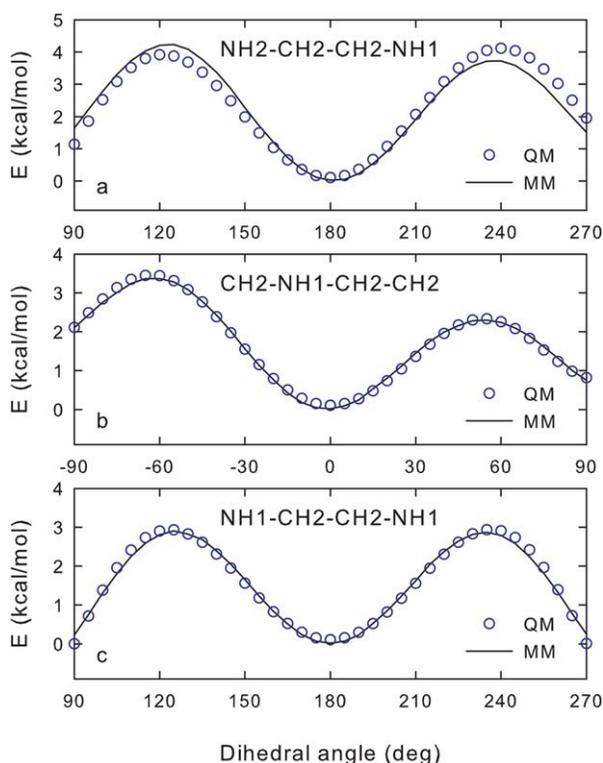


Figure 5. Torsion energy profiles for the dihedrals defining the backbone of the unprotonated model tetramer PEI4p0. [Color figure can be viewed at wileyonlinelibrary.com]

obviously correlates with the just double occurrence in a polymer chain of ending dihedrals like NH₂-CH₂-CH₂-NH₁, which are hence given lower weights in the optimization procedure and finally have higher associated inaccuracies. These larger discrepancies may also stem from the fact that, in adjusting the atomic charges, we slightly lowered the charge on the C atom next to the ending NH₂ group, while assimilating it with the generic CH₂ type and not defining an additional type. This assumed choice was based on the lower weight of the ends within the PEI chains of interest, typically composed of tens of units. The same reasoning would also explain the larger deviations for the CH₂-CH₂-NH₁-CH₂X dihedral of the protonated model PEI4p1 (Fig. 6a), where the left CH₂ is adjacent to a terminal NH₂. In this latter case, the RMSE amounts to 0.25.

As in the optimization of the bond angles, the absence of the sequence of residues PEP-PEQ-PEP within our models is compensated by setting the parameters for the CH₂X-NH₁-CH₂X-CH₂P and CH₂X-NH₁-CH₂X-HC₂ dihedrals by similarity equal to those for CH₂-NH₁-CH₂X-CH₂P and CH₂-NH₁-CH₂X-HC₂, respectively. The adjusted dihedral parameters for the backbone atoms of the PEI4p0 and PEI4p1 models are listed in Supporting Information Table S4.

The backbone dihedrals N-C-C-N become less rigid when the protonated NH₂P atom type is involved, and the force constant for NH₁-CH₂X-CH₂P-NH₂P is by roughly 30% lower than for NH₁-CH₂-CH₂-NH₁. Overall, our FF leads to significantly more rigid PEI polymers than published in the literature. Specifically, as compared to the dihedral force constants reported by Sun et al.,^[17] we find (in kcal/mol) for CH₂-CH₂-NH₁-CH₂ 1.43 instead

of 1.26, for CH₂X-CH₂P-NH₂P-CH₂P 0.33 instead of 0.10, and for NH₁-CH₂-CH₂-NH₁ 2.99 instead of 0.15.

The resulted topology and parameter files containing the complete sets of parameters are provided as Supporting Information.†

Structural dynamics of solvated PEI

We used the developed CHARMM FF in the systematic study of the dynamic structuring of solvated linear PEI chains of different sizes and protonation ratios. To achieve similar uniform protonation states, we considered chains of doubling length (plus 2 unprotonated ending units), specifically composed of 12*n* + 2 monomers (14-mer, 26-mer, and 50-mer). These polymers can accommodate protonation ratios equal to 1/4 (one-in-four), 1/3 (one-in-three), and 1/2 (alternatively protonated). In particular, besides its experimental relevance, alternative protonation is also useful for comparisons with other calculations.^[18]

The main features of the simulated systems, namely single non-protonated or uniformly protonated PEI chains contained along with Cl⁻ counterions in a solvation box, are summarized in Table 3. The number of Cl⁻ counterions matches in each

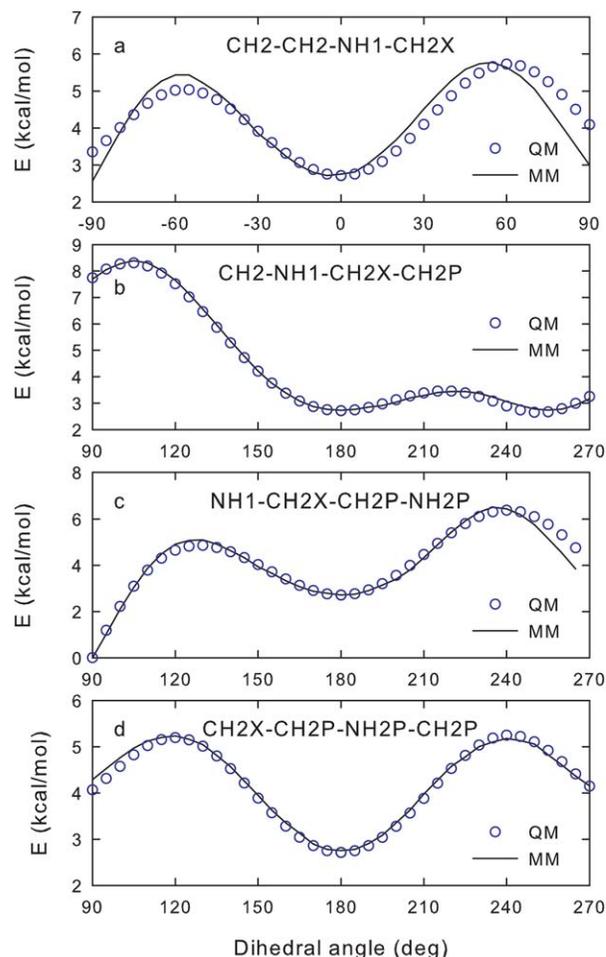


Figure 6. Torsion energy profiles for the dihedrals involving protonated backbone species of the model tetramer PEI4p1. [Color figure can be viewed at wileyonlinelibrary.com]

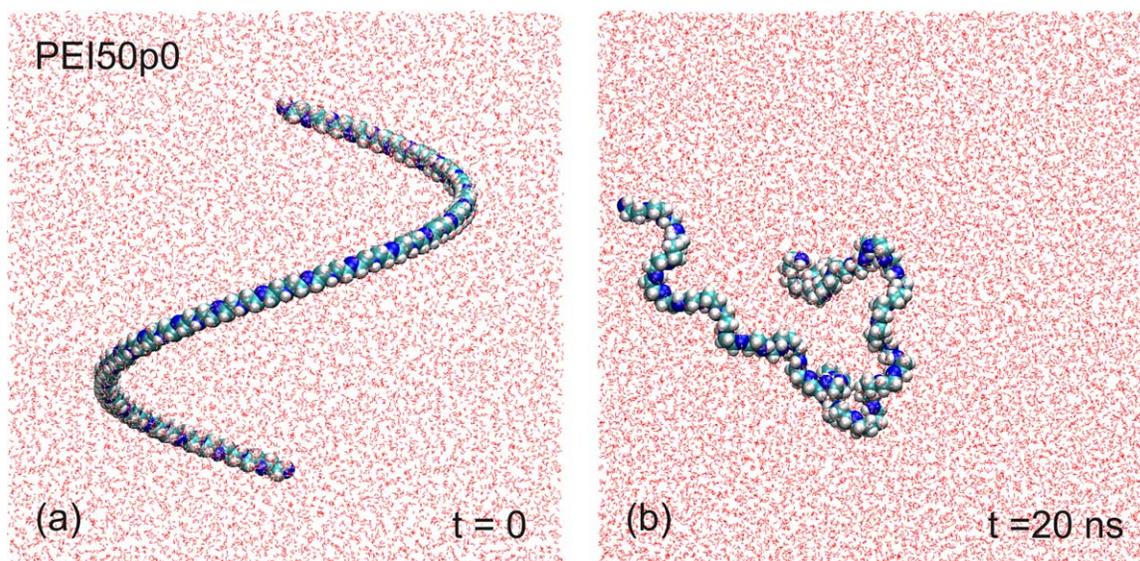


Figure 7. a) Initial configuration with randomly placed water molecules around a helicoidal unprotonated PEI 50-mer and b) final random-coil configuration of a typical trajectory. [Color figure can be viewed at wileyonlinelibrary.com]

case the number of protonated sites. The listed lengths correspond to linear equilibrium configurations of the PEI chains. Protonation can be seen to slightly increase them, and this can be obviously associated with the build-up of Coulomb repulsion along the chain.

As initial configuration for each MD trajectory, we considered a coiled PEI chain, with the extent (radius and axial loop separation) in the range of the expected final gyration radius, embedded in a TIP3P water^[30] cube of sufficient size to screen the chain's interactions with its periodic replicas and also accommodating the Cl^- ions (see Fig. 7a). While the simulation boxes for the 14-mers and 26-mers were dimensioned so as to spaciouly accommodate the fully extended chains along the diagonal, for tractability reasons, the boxes for the 50-mers were in relative terms by about 18% smaller, nevertheless still ample enough to exceed by more than 50% the final gyration radius.

With a view to reduce time correlations and improve statistics, instead of considering a single trajectory for each representative PEI chain, ensembles of 40 trajectories of 20 ns were generated, each started from a different random initial arrangement of the water molecules and Cl^- ions, and including 20 ps of equilibration. The final configuration of a typical trajectory for the unprotonated PEI 50-mer is shown in Figure 7b, while the movie of the entire trajectory is provided as Supporting Information.

We analyzed the dynamic structuring of PEI chains in terms of gyration radius, end-to-end distance, persistence length, RDFs, coordination numbers, and diffusion coefficients.

Gyration radius and end-to-end distance

We characterize the average spatial extent of the PEI chains by means of the gyration radius, R_g , and end-to-end distance, D_{ee} , whereby, in defining the latter, we refer to the N atoms of the ending NH_2 groups. Aiming to give an idea of the tremendous

instantaneous fluctuations affecting the measured quantities, Figure 8 shows the time dependences of R_g for all the 40 trajectories composing the ensemble for the 1/3-protonated PEI 26-mer. On a short time scale, the *ensemble-averaged* time dependence reflects the initially smooth (coiled) structure, which quite quickly condenses into a random coil of fluctuating extent. Already after a couple of nanoseconds, the memory of the initial configuration is lost. To also suggest the distribution of spatial measures, we plotted in Figure 9 the histogram of the gyration radii over the full (time- and trajectory) ensemble of conformations. The plot also provides snapshots of representative conformations—for the maximum (most probable gyration radius) and for the low-probability extremities of the distribution.

Figure 10 presents comparatively the ensemble-averaged time dependences of R_g and D_{ee} for unprotonated and uniformly protonated PEI 14-mers and 50-mers. In plotting protonation-dependent data, we use throughout black for unprotonated chains, while blue, green, and red for 1/4-, 1/3-, and 1/2-protonation, respectively. As can be seen, the average extent of the polymers substantially increases both with the chain length and protonation ratio. While for the 14-mers, the profiles stabilize in less than 1 ns, for the 50-mers, stabilization occurs only after about 3 ns. For this reason, in calculating all the time averages, we consistently discarded the first 3 ns of each trajectory, irrespective of chain size, considering as standard data collection interval the subsequent 17 ns.

The *ensemble-* and *time-averaged* values of the gyration radius, $\langle R_g \rangle$, and end-to-end distance, $\langle D_{ee} \rangle$, can be seen in Figure 11 to depend quasi-linearly on the chain size for each of the four protonation fractions considered. Unprotonated chains, in particular, have the most compact configurations for each chain size, and show a slight saturation tendency. The reduced error bars are, obviously, a beneficial result of the combined ensemble- and time averaging procedure, which amounts to 680 ns of effective data collection for each representative PEI chain.

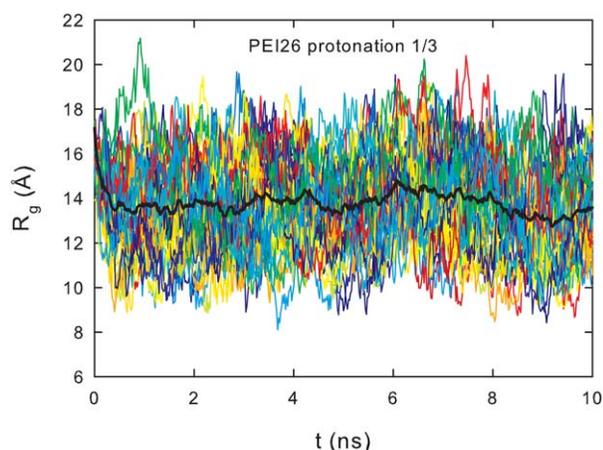


Figure 8. Time evolution of the gyration radius for the ensemble of 40 trajectories run for the 1/3-protonated PEI 26-mer. With black, the corresponding average time dependence. [Color figure can be viewed at wileyonlinelibrary.com]

In Figure 11, the similarity of the ascending patterns of the R_g - and D_{ee} dependences on size and protonation is obvious. Both the chain size and the supplementary electrostatic repulsion between protonated units substantially contribute to the average polymer extent. These manifest dependences will be also shown to be correlated with other dynamic properties, such as the diffusion coefficient.

Our results for the unprotonated and 1/2-protonated PEI 50-mer are compared in Table 4 with those of Choudhury et al.^[18] In accordance with the previous discussion, the increased extent of our PEI chains can be directly traced back to the higher torsional rigidity of our CHARMM FF.

Persistence length

A commonly used physical quantity for characterizing the stiffness of a polymer is the persistence length. Within the continuous worm-like chain model,^[31] it is regarded as the length over which correlations in the direction of the tangent to the

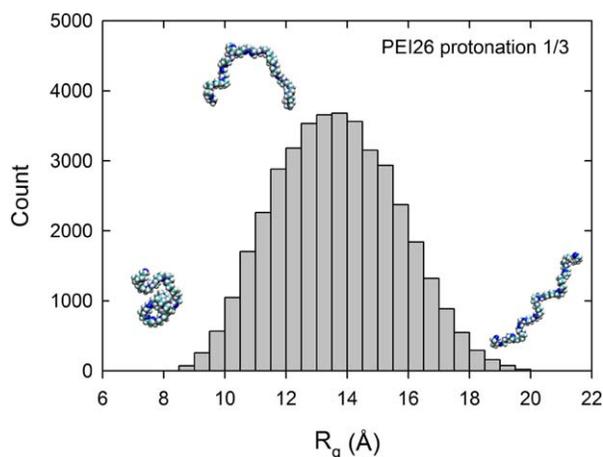


Figure 9. Distribution of the gyration radii for the 1/3-protonated PEI26 over the full time- and trajectory ensemble of conformations. [Color figure can be viewed at wileyonlinelibrary.com]

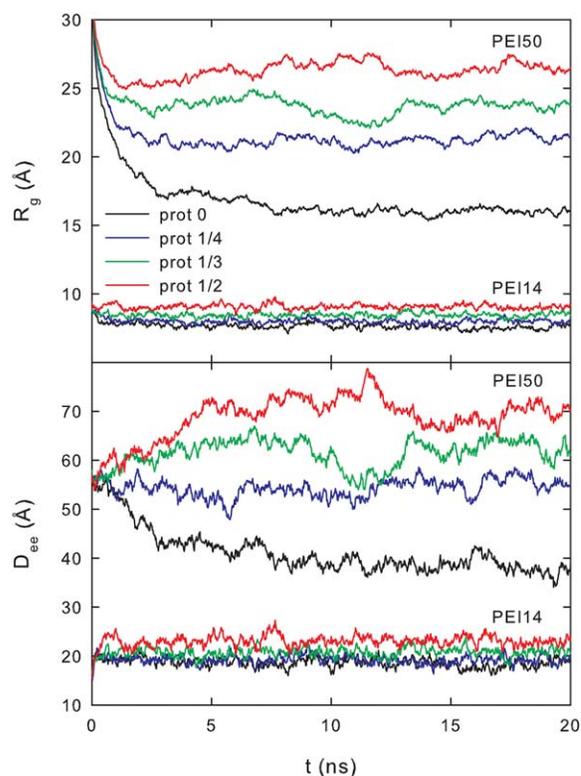


Figure 10. Time evolution of ensemble-averaged gyration radius and end-to-end distance for PEI 14-mers and 50-mers. [Color figure can be viewed at wileyonlinelibrary.com]

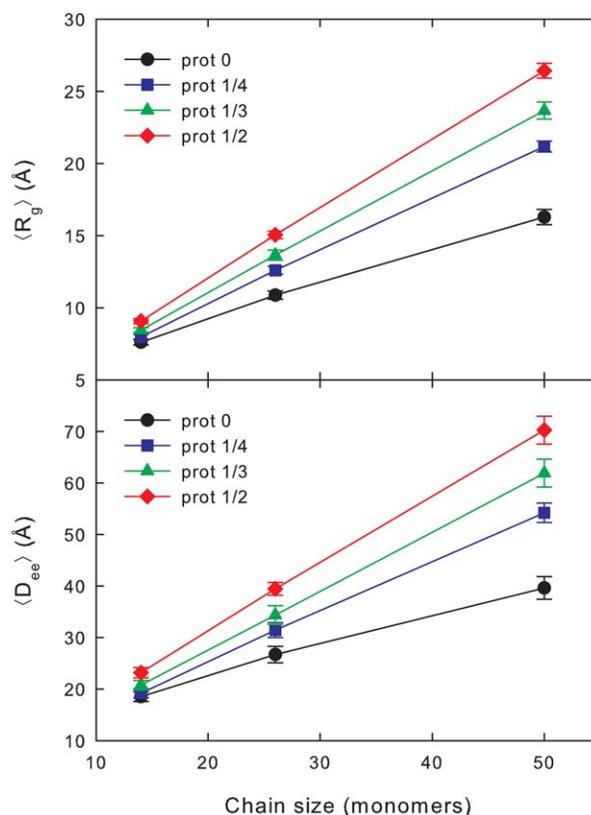


Figure 11. Size dependence of the ensemble- and time-averaged gyration radius and end-to-end distance of the simulated PEI chains. [Color figure can be viewed at wileyonlinelibrary.com]

Table 4. Ensemble- and time-averaged radius of gyration and end-to-end distance for PEI 50-mers.

50-mer	Non-protonated		1/2-protonated	
	Ref. [18]	This work	Ref. [18]	This work
R_g (Å)	12.3 ± 1.2	16.3 ± 0.5	24.1 ± 2.5	26.4 ± 0.5
D_{ee} (Å)	32.5 ± 7.7	39.6 ± 2.1	67.7 ± 13.4	70.2 ± 2.7

chain are (exponentially) lost. The persistence length, λ , can be related to the ensemble- and time-averaged squared end-to-end distance:

$$\langle D_{ee}^2 \rangle = 2\lambda L \left[1 - \frac{\lambda}{L} \left(1 - e^{-\frac{L}{\lambda}} \right) \right], \quad (5)$$

where L is the length of the fully extended polymer. By fitting this model to the $\langle D_{ee}^2 \rangle$ values resulted from our simulations for unprotonated chains (see Fig. 12), we obtained the estimate $\lambda = 4.8 \pm 0.2$ Å, which compares well with the value obtained by Lee^[32] for polyethylene oxide (4.3 Å). While the worm-like chain model fairly describes unprotonated PEI, in the case of 1/3-protonation it significantly departs from the simulated values, especially for the shorter chains. Indeed, taking into consideration by its basic assumptions only *short-range* interactions between *identical* segments, the model proves inadequate for protonated PEI chains, essentially by neither being able to account for different segment types nor for the intra-chain long-range electrostatic interactions involved.

Coordination numbers

The information regarding the internal structure and dynamics of the polymer based on the radius of gyration and end-to-end distance can be illustratively complemented by the coordination analysis. Aimed to synthetically indicate how compact the polymer is, the self-coordination number quantifies the degree of contact between internal atoms by summing pairwise distances:

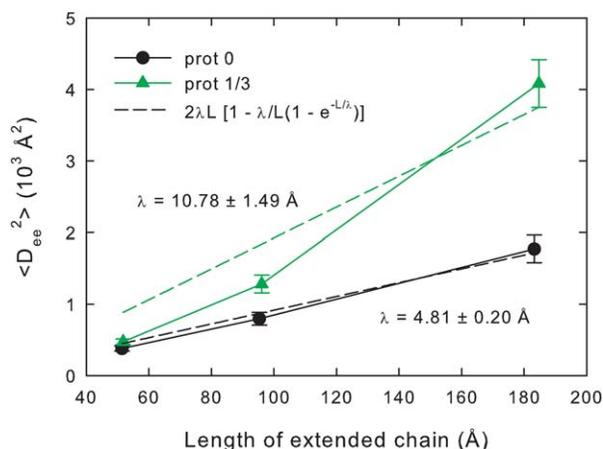


Figure 12. Squared end-to-end distances as functions of fully extended lengths for unprotonated and 1/3-protonated PEI chains, and (dashed) fitting curves yielding the corresponding persistence lengths λ . [Color figure can be viewed at wileyonlinelibrary.com]

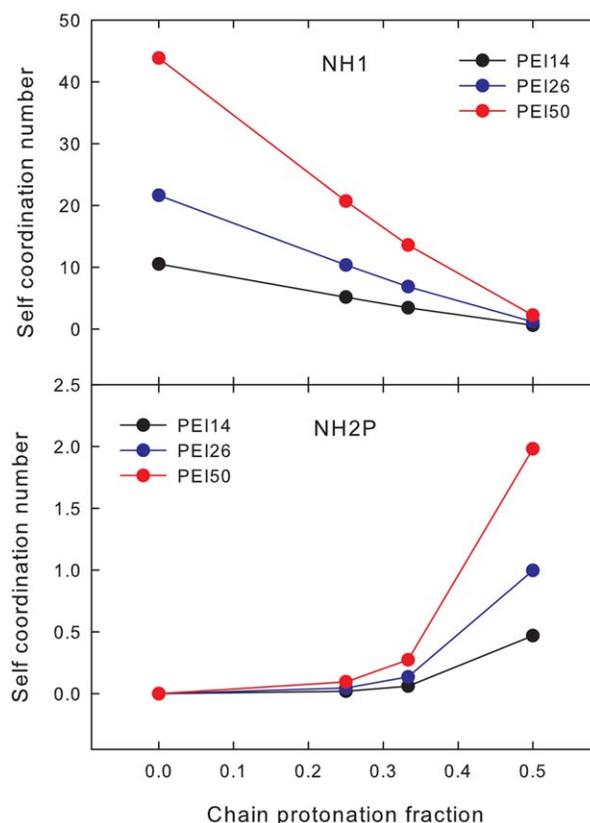


Figure 13. Self-coordination numbers for the NH1 and NH2P nitrogen species as functions of the protonation fraction. [Color figure can be viewed at wileyonlinelibrary.com]

$$C = \sum_{j>i} \frac{1 - (|\mathbf{r}_i - \mathbf{r}_j|/d_0)^n}{1 - (|\mathbf{r}_i - \mathbf{r}_j|/d_0)^m}. \quad (6)$$

In implementing the above formula for calculating time- and ensemble averages for the PEI backbone species, we used the default values recommended in the documentation of the Colvars package,^[33] namely $d_0 = 4$ Å, $m = 12$, and $n = 6$. By definition, for N atoms of a given species, the self-coordination ranges from approximately 0, when all the interatomic distances are much larger than the cutoff d_0 , to $N(N - 1)/2$, when all distances are within the cutoff. Figure 13 illustrates the steady decrease of the self-coordination number for the non-protonated nitrogen NH1 with the chain protonation fraction, and its natural increase with the chain length. By contrast, as expected, the protonated nitrogen species NH2P shows a pronounced increase both with the protonation fraction and the chain size.

Even though the inherent anisotropy in the presence of PEI chains may question the relevance of describing the atomic neighborhood by RDFs, it is yet instructive to investigate their dependence on the main chain features. The average number of neighbors in the first coordination shell of a given atom type (the coordination number), is provided by the integral over the first peak of the corresponding RDF. Given the fluctuations of the simulation box caused by the applied pressure piston, in averaging the RDFs, the partial profiles were normalized relative to the actual box size for each simulation frame.

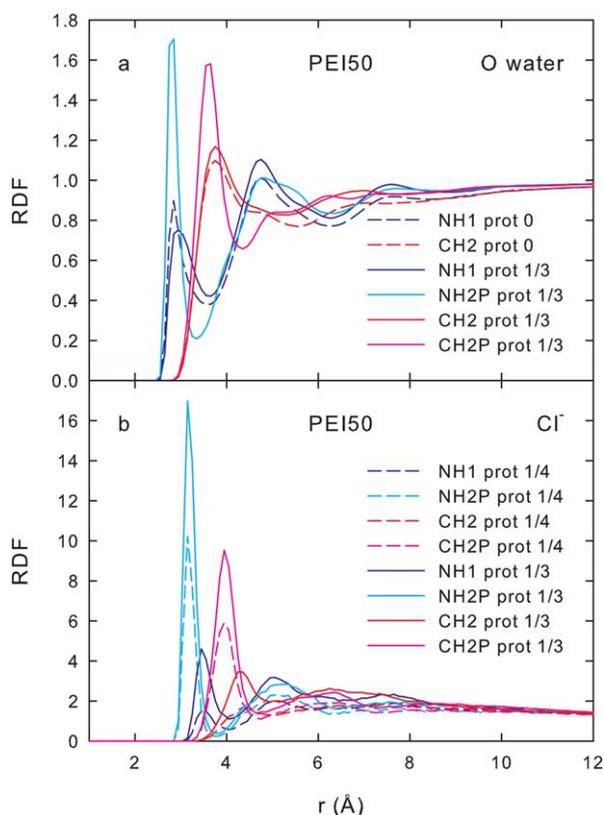


Figure 14. Radial distribution functions for the N and C backbone atoms of the PEI 50-mer with water O atoms (a) and Cl⁻ counterions (b). [Color figure can be viewed at wileyonlinelibrary.com]

In characterizing the structure of the solvent around the PEI chains, we focus on the nitrogen and carbon backbone atoms. The corresponding RDFs with water O and Cl⁻ depicted for the unprotonated and 1/3-protonated 50-mer in Figures 14a and 14b, show a quite different arrangement of the solvent in their vicinities. As compared to the C species, the structuring of oxygen about the N species, whether protonated or not, sets in at lower atomic separations (~ 3 Å vs. ~ 4 Å), extends over a wider range, and has better outlined oscillations. The RDFs for the N species even show distinct, albeit more diffuse, secondary and tertiary coordination shells. In fact, the C-O main maxima fall in the gaps between the N-O peaks.

In the particular case of the CH2-O profiles (both for unprotonated and protonated chains), the maximum is followed by a broad minimum, which makes it difficult to estimate the extent of the major peak. Here, to delimit the major peak, we rather used the discrete inflection point, coinciding in position with the well-defined minimum for the CH2P-O profile. The inflection point is obviously an indication of the overlapping coordination shells of the pairs of consecutive C species along the backbone.

Protonation leads to a clearer defined radial structuring, with narrower and higher major peaks, slightly shifted by the increased electrostatic interactions toward lower inter-atomic distances. The effect of protonation is particularly visible for the protonated nitrogen species. Whereas for NH1-O, both in unprotonated and protonated chains, the second RDF peak

exceeds the first one, in the NH2P-O RDF, the first peak is the major one, dominating all other O peaks.

Figure 14b depicts the RDFs of the Cl⁻ counterions about the N and C species. The pattern of major peaks can be seen to follow the one of the O atoms, however shifted by 0.3–0.4 Å to larger distances. The Cl⁻ coordination shells tightly enclose the O-shells, being enclosed, in their turn, by H-shells, as corroborated by Figure S2b (in the Supporting Information). Specifically, in the 1/3-protonated 50-mer, the O-, Cl⁻, and H-shells around the protonated NH2P species have steadily increasing average radii of 2.85, 3.15, and 3.45 Å, respectively. The coordination shells maintain the same ordering about the CH2P species, with average radii equal to 3.65, 3.95, and 4.15 Å.

As a synthetic measure of the first coordination shells of the N and C backbone atoms, Figure 15 presents the O- and H-coordination numbers as functions of the chain protonation fraction. Somewhat surprisingly, the O-coordination (Fig. 15a) is not primarily affected for protonated residues, but rather at the adjacent non-protonated sites. Indeed, while, in general, the coordination does not pronouncedly depend on the protonation fraction, only the unprotonated atomic backbone species, NH1 and CH2, show a moderate increase of the O-coordination with increasing protonation. Quite in contrast, their protonated counterparts, NH2P and CH2P, feature practically invariable O-shells.

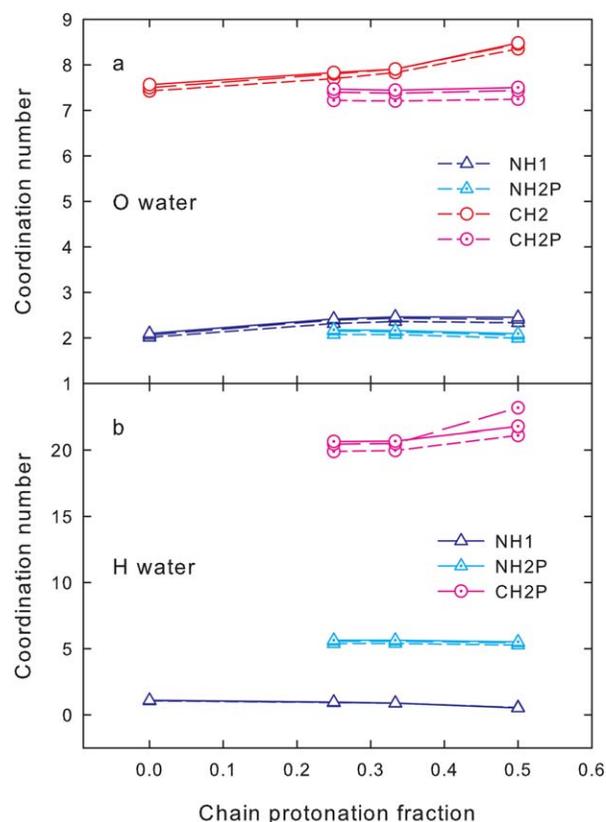


Figure 15. Dependence of the coordination numbers of the N and C backbone atoms on the PEI chain protonation (14-mer with short-dash line, 26-mer with long-dash line, and 50-mer with continuous line). [Color figure can be viewed at wileyonlinelibrary.com]

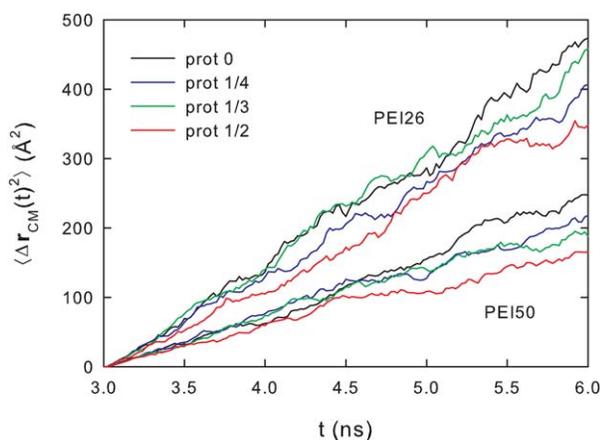


Figure 16. Mean square displacement of the center of mass versus time for PEI 26 and 50-mers with 0, 1/4-, 1/3-, and 1/2-protonation. [Color figure can be viewed at wileyonlinelibrary.com]

In contrast to the O-coordination, the H-coordination (Fig. 15b) varies only for CH2P, increasing significantly in 1/2-protonated PEI chains. Even though practically invariable with the protonation fraction, the hydration numbers for NH1 and NH2P (having partial atomic charges of -0.796 and -0.968 , respectively) are quite different (~ 0.9 and ~ 5.5), and this can be related to the increased electrostatic attraction of the water H atoms (charge $+0.417$) toward protonated nitrogen.

Supplementary considerations on the coordination numbers are to be found in the Supporting Information.

Diffusion coefficients

To characterize the mobility of solvated PEI chains, we studied the dependence of the center-of-mass diffusion coefficient on the chain size and protonation fraction. For the rather intricate evaluation of the diffusion coefficient, we used Einstein's relation:

$$D = \lim_{t \rightarrow \infty} \frac{1}{6t} \langle \Delta r_{CM}(t)^2 \rangle \quad (7)$$

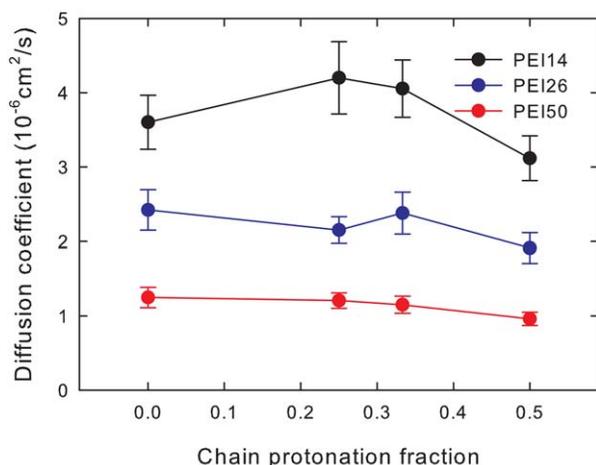


Figure 17. Variation of the diffusion coefficient for PEI 14, 26, and 50-mers with the protonation fraction. [Color figure can be viewed at wileyonlinelibrary.com]

The mean squared displacement (MSD), $\langle \Delta r_{CM}(t)^2 \rangle$, describes the ensemble-averaged spatial spread of the chain's center of mass relative to an (initial) reference position. To produce the correct MSD, real spatial positions need to be used, not the ones repeatedly corrected by virtue of the periodic boundary conditions.

Consistently discarding the initial 3 ns of equilibration from all the trajectories, we present in Figure 16 the subsequent 3 ns of time evolution of the MSD for the PEI 14-mers and 50-mers of 0, 1/4-, 1/3-, and 1/2-protonation. As requested by the general theory, the dependencies are quasi-linear for sufficiently long times. Nevertheless, longer data collection intervals do not automatically guarantee improved results, as, as noticeable from the depicted dependences, particularly for the shorter chains (PEI 14-mers), numerically conditioned decorrelations already set in after about 2 ns.

Overall, the MSD values can be seen to decrease in distinct groups with increasing chain size, and for a given chain size, with increasing protonation fraction. The decrease in mobility is perfectly consistent with the increase in spatial extent that we found, as measured by the gyration radius and end-to-end distance (see Fig. 10). Qualitatively, longer chains with higher protonation fractions, thus less compact and locally more rigid, are naturally expected to show smaller diffusion coefficients.

Figure 17 summarizes the dependence of the diffusion coefficient of PEI, extrapolated from the time dependences of the MSD, on the chain size and protonation fraction. It is readily apparent that, in relative terms, the diffusion coefficient is influenced to a larger extent by the chain size than by the protonation level. While the diffusion coefficient for the 50-mer shows reduced error bars and a steady decrease with the protonation fraction, in the case of the 14-mer the significant error bars obscure a similar tendency.

Our diffusion coefficient for the unprotonated PEI 50-mer (molecular weight 2167), namely $1.25 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$, compares very well with the value of $1.2 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ obtained by Clamme et al.^[34] from two-photon fluorescence correlation spectroscopy measurements for branched PEI chains of molecular weight 2500. We also find remarkable consistency with the experimental value of $1.34 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ obtained by Hostetler et al.^[35] for slightly shorter chains of molecular weight 2140. Moreover, the same steady decrease of the diffusion coefficient with increasing size, seen for the PEI 50-mer in Figure 17, is also reported by Hostetler et al. for PEI chains for molecular weights in the range from 8000 to 20,000.

Conclusions

The central result of our investigations is a new CHARMM FF for linear unprotonated/protonated PEI. Its enhanced realism stems from the fact that, unlike other published studies, in which *only* *dihedrals* were adjusted, we consistently optimized based on high quality *ab initio* calculations *all* the bonded parameters (for bonds, angles, and dihedrals), as well as the partial atomic charges.

One of the distinctive features of the developed FF is the increased stiffness as compared to earlier FFs for PEI. In

particular, it shows increased torsional rigidity (higher dihedral force constants), especially about the C-C bonds. Protonation is found to further increase the PEI chain rigidity and, notably, it does not only affect the protonated nitrogen atoms but to a significant extent also the adjacent residues. The pattern of atomic charges and force constants supports our definition of residues based on functional groups (with C-C-N backbone).

With a view to optimized drug delivery protocols based on PEI, we investigated the dynamic structuring of solvated uniformly protonated PEI chains of various sizes. The gyration radius and end-to-end distance manifestly increase both with chain size and, due to the additional electrostatic repulsion, with the protonation fraction. Direct comparison with other simulations from the literature evidences that our PEI chains are spatially more extended, which can be traced back to our higher force constants.

While protonated and unprotonated nitrogen backbone atoms have rather different O- and H-coordination numbers, these are only marginally affected by the variation of the protonation fraction. Nevertheless, higher protonation increases the hydration numbers for the carbon species. Given the similar, uniform protonation fractions used, the hydration shells of nitrogen and carbon are rather insensitive to the PEI chain size.

The center-of-mass diffusion coefficients of PEI chains vary inversely with the chain length and protonation fraction, consistently with the observed increase in spatial extent, and very well agree with experimental findings.

Keywords: cationic polymers · polyethylenimine · gene delivery · force fields · molecular dynamics

How to cite this article: T. A. Beu, A. Farçaş, *J. Comput. Chem.* **2017**, *38*, 2335–2348. DOI: 10.1002/jcc.24890

 Additional Supporting Information may be found in the online version of this article.

- [1] R. Narain, Ed. *Polymers and Nanomaterials for Gene Therapy*; Elsevier, Woodhead Publishing: Cambridge, **2016**.
- [2] E. Fleige, M. A. Quadir, R. Haag, *Adv. Drug. Deliv. Rev.* **2012**, *64*, 866.
- [3] X. Guo, L. Huang, *Acc. Chem. Res.* **2012**, *45*, 971.
- [4] M. M. Yallapu, M. Jaggi, S. C. Chauhan, *Drug Discov. Today* **2011**, *16*, 457.
- [5] N. Martinho, C. Damge, C. P. Reis, *J. Biomater. Nanobiotechnol.* **2011**, *2*, 510.
- [6] M. Morille, C. Passirani, A. Vonarbourg, A. Clavreul, J. P. Benoit, *Biomaterials* **2008**, *29*, 3477.
- [7] M. Ahmed, R. Narain, *Biomaterials* **2012**, *33*, 3990.
- [8] H. Song, G. Wang, B. He, L. Li, C. Li, Y. Lai, X. Xu, Z. Gu, *Int. J. Nanomed.* **2012**, *7*, 4637.
- [9] R. T. Chacko, J. Ventura, J. Zhuang, S. Thayumanavan, *Adv. Drug. Deliv. Rev.* **2012**, *64*, 836.
- [10] M. A. Mintzer, E. E. Simanek, *Chem. Rev.* **2009**, *109*, 259.
- [11] O. Boussif, F. Lezoualc'h, M. A. Zanta, M. D. Mergny, B. D. D. Scherman, J. P. Behr, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7297.
- [12] W. T. Godbey, A. G. Mikos, *J. Control. Release* **2001**, *72*, 115.
- [13] R. V. Benjaminsen, M. A. Matthebjerg, J. R. Henriksen, S. M. Moghimi, T. L. Andresen, *Mol. Ther.* **2013**, *21*, 149.
- [14] J. D. Ziebarth, Y. Wang, *Biomacromolecules* **2010**, *11*, 29.
- [15] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *25*, 1049.
- [16] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269.
- [17] C. Sun, T. Tang, H. Uludağ, J. E. Cuervo, *Biophys. J.* **2011**, *100*, 2754.
- [18] C. K. Choudhury, S. Roy, *Soft Matter* **2013**, *9*, 2269.
- [19] Z. Wei, E. Luijten, *J. Chem. Phys.* **2015**, *143*, 243146.
- [20] Z. Wei, E. Luijten, *Supplementary material for "Systematic coarse-grained modeling of complexation between small interfering RNA and polycations"*; **2015**. Available at: <https://doi.org/10.1063/1.4937384>
- [21] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evansack, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, C. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586.
- [22] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell Jr., *J. Comput. Chem.* **2010**, *31*, 671.
- [23] C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, J. C. Gumbart, *J. Comput. Chem.* **2013**, *34*, 2757.
- [24] C. G. Mayne, *J. Comp. Chem.* **2013**, *34*, 2757.
- [25] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09, Revision A.01*; Gaussian, Inc.: Wallingford, CT, **2009**.
- [26] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33.
- [27] MacKerell Lab Homepage, *CHARMM force field files*. Available at: http://mackerell.umaryland.edu/charmm_ff.shtml
- [28] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781.
- [29] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577.
- [30] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [31] O. Kratky, G. Porod, *Recl. Trav. Chim. Pays-Bas* **1949**, *68*, 1106.
- [32] H. Lee, R. M. Venable, A. D. M, Jr., R. W. Pastor, *Biophys. J.* **2008**, *95*, 1590.
- [33] G. Fiorin, M. L. Klein, J. Henin, *Mol. Phys.* **2013**, *111*, 3345.
- [34] J. P. Clamme, J. Azoulay, Y. Mely, *Biophys. J.* **2003**, *84*, 1960.
- [35] R. E. Hostetler, J. W. Swanson, *J. Polym. Sci.* **1974**, *12*, 29.

Received: 21 March 2017
Revised: 16 June 2017
Accepted: 30 June 2017
Published online on 27 July 2017