Comment on "Martini Force Field for Protonated Polyethyleneimine"

*Subhamoy Mahajan and Tian Tang**

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

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

Abstract

We comment on the recently published Martini forcefield for linear polyethylenimine (Beu *et al.*, *J. Comput. Chem.*, **2020**, 41, 349-361). With supporting evidence, we demonstrate that the new coarse-graining methodology presented in that work is inconsistent with the Martini methodology and lacks important validation steps which is usually required by Martini forcefields.

Polyethylenimine (PEI) is a polycation that can be synthesized with various molecular weights and degrees of branching, and exhibits different protonation states depending on the pH of the solution. PEI is considered as one of the most potent carriers in gene delivery. In our previous work[1], [2], we developed a Martini coarse-grained (CG) model for linear- and branched-PEIs using four different structures, and two different protonation ratios relevant to gene delivery applications. Recently, Beu *et al.*[3] published an article presenting a new Martini CG model for linear PEIs, where several criticisms were made on our PEI model[1]. Taking the criticisms seriously, we performed investigations into their methodology and claims. We found that the methodology presented by Beu *et al.*[3] is questionable and inconsistent with the Martini framework of coarse-graining[4]–[6], and that the criticisms on our model are unwarranted.

The PEI models studied by Beu *et al.*[3] were linear and comprised of repeating units of -C-N-C- with terminal methyl groups, that is, C-(C-N-C)$_{(12n+3)}$-C. Three different lengths were considered, with 12n+3 = 27, 39 and 51 respectively, along with four "uniform protonation” fractions: 0, 1/4, 1/3 and 1/2. Every -C-N-C- atoms and their accompanying hydrogens were mapped to a CG Martini bead, whereas each of the terminal methyl groups, containing one heavy atom, was also mapped to a CG Martini bead. The bonded interaction parameters for the CG model were derived by applying direct Boltzmann inversion on all-atom (AA) bonded distributions using single and multi-function fits. Afterwards, different CG models were simulated using the same bonded parameters, but with different bead types, namely SNda, SNd, SN0 or P2 for unprotonated bead, and Qd, SQd or SQ0 for protonated bead. Global properties, including radius of gyration (R_g) and end-to-end distance (D_{ee}), of the different CG models were compared with AA PEIs to find the “optimal” bead type combination, which was determined to be SNda-Qd or SNda-SQd in polarizable water, and SN0-SQ0 in non-polarizable water. Using one of the “optimal” bead type combinations, SNda-SQd, potential of mean force (PMF) calculations were performed for the interaction between DNA phosphate groups and protonated PEI beads. In the following, we raise several issues related to the claims made in Beu *et al.*[3]. Most importantly, we explain that the methodology used by Beu *et al.*[3] did not follow the Martini protocol, and their approach to assess the quality of our model was incorrect.

The coarse-graining procedure used by Beu *et al.*[3] deviated significantly from the Martini methodology available in the literature[4]–[6]. Under the Martini framework, a molecule is coarse-grained by first choosing the mapping scheme, then selecting the bead type, and finally determining the bonded parameters through iterative comparison with AA distributions of bond lengths, angles and dihedrals.[4]–[6] The model is then validated by comparing the global properties with AA simulations or experimental results.[4]–[6] On the other hand, Beu *et al.*[3] first obtained the CG bonded parameters by using direct Boltzmann inversion on the AA bonded distributions, and then determined the “optimal” bead type combination by comparing the CG and AA predictions of global properties (R_g and D_{ee}). There are several issues with this approach. First of all, from the method described and results presented in Beu *et al.*[3], the bonded parameters were directly obtained from the Boltzmann inversion without iterative tuning. This method neglects the mutual influences of different bonded interactions and can lead to inaccuracy in the model. In fact, four out of the seven CG dihedral angle distributions in Beu *et al.*[3] could not accurately reproduce the AA distributions (see Figure S8 and S9 in Supporting Information of Beu *et al.*[3]), which is very likely due to their non-iterative parameterization scheme. The distribution for the eighth dihedral angle, PEC-PEI-PEI-PEP, was not reported. While the authors[3] acknowledged that "*torsional coordinates turn out to be very sensitive to the presence of protonated beads*", no iterative improvements were made for their dihedral parameters. In contrast, in our work[1], which followed the Martini methodology, all the dihedral angles associated with the linear PEI accurately reproduced the AA distributions. It should also be noted that Beu *et al.*[3] mischaracterized our work, by stating that our parametrization was carried out by Boltzmann inversion. Although Boltzmann inversion was used in our work to obtain the initial guesses for the parameters, in accordance with the Martini methodology, the bonded parameters were improved over several iterations. Secondly, the choice of bead type in Beu *et al.*[3] was based on global property comparisons, using the pre-determined bonded interactions. The outcome is therefore highly dependent on the quality of the bonded parameters, which significantly influence the values of R_g and D_{ee}. For instance, inaccurate dihedral parameters might negatively impact the choice of bead type combinations. In the Martini literature[4], [5], bead type is determined based on the experimentally measured partition free energy of the molecular analogue between water and different organic solvents, and this step is independent of the bonded parameters. Beu *et al.*[3] proposed a new approach which differed significantly from the widely accepted Martini methodology but did not prove its validity.

Beu *et al.*[3] claimed that when a -C-N-C- residue was mapped to one CG bead, it was rather small compared to to a regular Martini bead ("four-to-one" mapping), and that such a bead needed to be identified with "small" Martini types. This claim is a mischaracterization of Marrink *et al.*[4], [5], where it was stated that four is the average number of heavy atoms that can be represented by a bead using a "four-to-one" mapping; i.e., three, five or more heavy atoms can be mapped into one bead. In fact, the "small" Martini bead was developed to model ring compounds,[5] which can potentially lead to inaccurate modelling of linear PEIs. For example, Lee *et al.*[7] observed that a small subset of linear polyethylene oxides formed ring-like conformations when modelled with "small" Martini bead. In addition, when discussing their Table 3, Beu *et al.*[3] stated *"As an important finding, the best performance in reproducing the AA results with polarizable water (lowest RMSD both for R_g and D_{ee}) is achieved by the SNda–Qd–3–3–P and SNda–SQd–3–1–P FF models...*" This claim is not supported by the data in Table 3 of Beu *et al.*[3]. Among the four models in Table 3 of Beu *et al.*[3], SNda-Qd-3-3-P is clearly the best performer, having the lowest RMSD for both R_g and D_{ee}. The next in line is either SNda-SQd-3-1-P which gives the second lowest RMSD for R_g, or SNda-Qd-3-1-P which gives the second lowest RMSD for D_{ee}. SNda-Qd-3-1-P may actually be seen to match AA results better than SNda-SQd-3-1-P if both R_g and D_{ee} are taken into account. We did not find sufficient evidence for "*the more uniform overall consistency of the CG and AA profiles for the SNda-SQd pair (with both PEI and PEP as 'small' types)*" as claimed by the authors.

Beu *et al.*[3] criticized our choice of bead-type combination, P2-Qd, based on simulations using their bonded parameters. This is incorrect since in a CG model, the mapping scheme, bonded parameters and non-bonded parameters (i.e., bead types) form an integrated "package". One cannot simply combine the bonded parameters from one model and the bead types from another model in order to assess either model. As an experiment, we simulated three linear PEIs using the bonded parameters from our CG model and bead type combination SNda-SQd recommended by Beu *et al.*[3]. These PEIs had a total number of 27, 39 and 51 beads respectively, each containing 1 protonated bead at the terminal. The same lengths were simulated by Beu *et al.*[3], at the protonation ratios of 0, 1/4, 1/3 and 1/2. The three PEIs we simulated did not match these exact protonation ratios because of the following reasons. Firstly, the criticism on our bead type was based on Figure 11 of Beu *et al.*[3], where discrepancy in R_g with AA values was only found for unprotonated PEIs. Therefore, it is more appropriate to conduct our "experiment" for unprotonated or weakly protonated PEIs. On the other hand, our CG model[1], [2] cannot be used for unprotonated PEIs, which are of little relevance to practical gene delivery applications. Consequently, we lowered the protonation ratio as much as possible to create three weakly protonated PEIs. They were named PEI27, PEI39, and PEI51 respectively, based on their total number of beads. For comparison, the same three PEIs were also simulated using the AA model of Sun *et al.*[8] and the CG model with our bonded parameters and the bead type combination of P1-Qd (P1 and P2 are indistinguishable when simulating PEIs in polarizable or non-polarizable water, see Mahajan *et al.*[2]). All simulations were performed with a single PEI in polarizable water[9] using the settings defined in our original work.[1] In **Figure 1**, for both R_g and D_{ee}, the results from AA, CG:P1-Qd and CG:SNda-SQd simulations are comparable. That is, the influence of bead type on the global properties is insignificant, which is very different from the observations of Beu *et al.*[3]. We would like to emphasize that the simulations leading to **Figure 1** were merely an experiment; combining bonded parameters from our model with bead types from another model (or vice versa) is an incorrect approach, and should not have been used to judge the quality of either models.

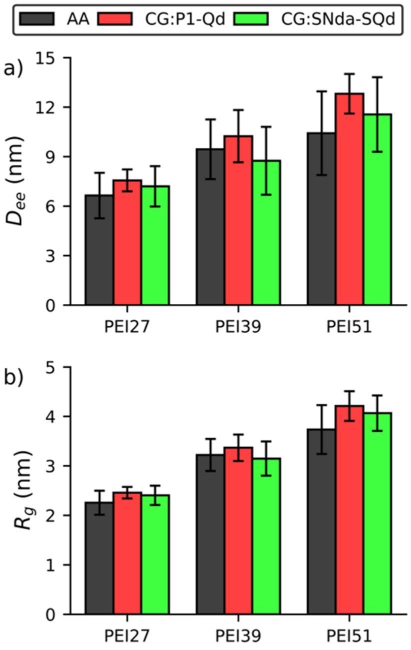

Figure 1: **(a)** End-to-end distance (D_{ee}) and (b) radius of gyration (R_g) of PEIs with 27, 39 and 51 beads of which only one bead was protonated.

Finally, while Beu *et al.*[3] calculated the CG PMF between DNA phosphate groups and protonated PEI beads, no comparison was made with AA simulations to verify the compatibility of their PEI model with Martini DNA[10]. In contrast, our PEI model[1], [2] was demonstrated to be able to reproduce the AA PMF between the centers of mass of a DNA and a PEI. As another experiment, for the 12bp DNA and SL46 PEI in our original work[1], we performed CG PMF calculations using our bonded parameters and the bead type combination of SNda-SQd, and compared the result with the PMFs from AA (force field parameters from [8], [11]), P1-Qd (extended simulations from [1]) and P2-Qd (extended simulations from [2]) models. In **Figure 2**, each PMF curve, including average and standard deviation, was obtained from 1000 bootstrapped trajectories generated from umbrella sampling using GROMACS 5.1.4 [12]. Convergence of all PMFs was tested (not shown) and it took longest (200 ns) for the CG model with SNda-SQd bead combination to converge. For a fair comparison, all CG umbrella sampling simulations were conducted for 200 ns to generate the PMFs. The AA PMF reported in **Figure 2** was based on statistics over 10-20 ns of the AA simulations, while the

CG PMFs were based on statistics over 50-200 ns. Readers may notice that the AA, P1-Qd and P2-Qd PMFs differ slightly from the original publications[1], [2], which is due to averaging over bootstrapped trajectories as well as longer CG simulations. **Figure 2** shows that the PMF associated with SNda-SQd deviates significantly from the AA result, which is in stark contrast to the PMFs for P1-Qd and P2-Qd. Certainly, here our bonded parameters were inappropriately combined with the bead type of Beu *et al.*[3], so the results in **Figure 2** may not be used to directly judge the quality of Beu *et al.*[3]'s model in capturing DNA-PEI interactions. However, this exercise does emphasize the importance of validating the CG DNA-PEI interactions against AA simulations, which was absent in Beu *et al.*[3].

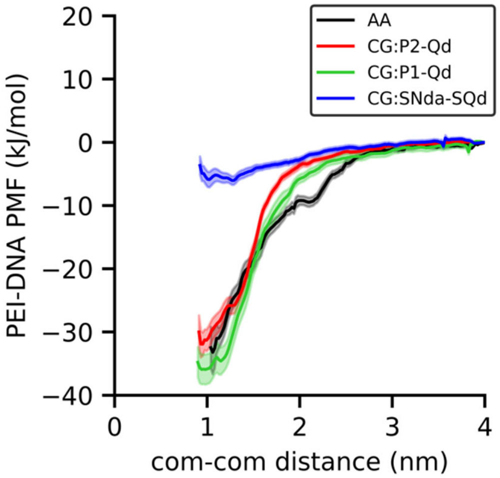

Figure 2: Potential of mean force between centers of mass of DNA[10] and PEI modelled using bead type combinations P1-Qd[2], P2-Qd[1] and SNda-SQd. Each solid line represents the average over 1000 bootstrapped trajectories and the corresponding shaded region represent the standard deviation.

Acknowledgements

S. M. and T. T. acknowledge the computing resources and technical support from Western Canada Research Grid (WestGrid). Funding from the Natural Science and Engineering Research Council (NSERC) of Canada is gratefully acknowledged.

Reference

[1] S. Mahajan, T. Tang, *J. Comput. Chem.*, **2019**, *40*, 607–618.

[2] S. Mahajan, T. Tang, *J. Comput. Chem.*, **2020**, DOI:10.1002/jcc.26214.

[3] T. A. Beu, A. E. Ailenei, R. I. Costinaş, *J. Comput. Chem.*, **2020**, *41*, 349–361.

[4] S. J. Marrink, A. H. de Vries, A. E. Mark, *J. Phys. Chem. B*, **2004**, *108*, 750–760.

[5] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, *J. Phys. Chem. B*, **2007**, *111*, 7812–7824.

[6] S. J. Marrink, D. P. Tieleman, *Chem. Soc. Rev.*, **2013**, *42*, 6801–6822.

[7] H. Lee, A. H. De Vries, S. J. Marrink, R. W. Pastor, *J. Phys. Chem. B*, **2009**, *113*, 13186–13194.

[8] C. Sun, T. Tang, H. Uludağ, J. E. Cuervo, *Biophys. J.*, **2011**, *100*, 2754–2763.

[9] S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta, S. J. Marrink, *PLoS Comput. Biol.*, **2010**, *6*, 1–17.

[10] J. J. Uusitalo, H. I. Ingólfsson, P. Akhshi, D. P. Tieleman, S. J. Marrink, *J. Chem. Theory Comput.*, **2015**, *11*, 3932–3945.

[11] A. D. MacKerell, N. Banavali, N. Foloppe, *Biopolymers*, **2000**, *56*, 257–265.

[12] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah, *SoftwareX*, **2015**, *1*, 19–25.